3
History of Statistics

Fisher to W.R. Buckland: 20 July 1954

. . . You will know that many of these ideas[1] and the definitions of the corresponding technical terms originated in my work which, as these ideas developed, contains sufficiently clear statements as to the meanings of the words introduced, e.g. I never used the word efficiency of estimation in reference to the mean square deviation of estimates from finite samples, still less did I ever introduce this practice which had been indeed latent, though not explicit, in all books on the theory of errors since the time of Gauss. My definition depends on amount of information and is aimed at affording an invariant comparison for transforms of the parameter such as a mean square deviation cannot give.

[1] Fisher had agreed to act as a consultant on terms relating to estimation in the preparation of M.G. Kendall and W.R. Buckland's *Dictionary of statistical terms* and then found himself in difficulties over the definitions presented to him. The *Dictionary* was published in 1957 by Oliver and Boyd, Edinburgh.

Fisher to H.D. Dufrénoy: 12 May 1936

Allow me to congratulate you on your fine paper, 'Méthodes statistiques appliquées à la pathologie végétale', which will, I am sure, be most valuable to all engaged in agronomic and phyto-pathological work.

Perhaps you will permit me to make clear a certain confusion, originating in America, as to the work developed in this country, which has, I see, influenced your mode of expression in some places.

The mathematical problem solved by 'Student' in 1908 constituted a most important refinement of the theory of errors, the possibility of which had been previously overlooked by mathematical writers on this subject. He supplied for the first time a rigorously exact test of significance for the mean of a finite sample drawn from a normally distributed population.

Perceiving the importance of what he had done, and being myself interested in the mathematical problems presented by the exact treatment of finite samples, I later extended his procedure to cover such experimentally important cases as a comparison of the means of two samples, the test of significance of a regression coefficient, and for the difference between the means of two regression coefficients, all of which problems may be reduced to the same distribution as that found by 'Student'.

Further, in a number of other problems, such as the comparison of the means of more than two samples, the significance of an interclass correlation, that of a multiple correlation, and the test of significance of deviations from regression formulae of any assigned form, I found that the principles of 'Student's' approach might be applied, but in these cases gave rise to a more general distribution, the variate of which I have denoted by z.

In respect to notation, 'Student' at first expressed his distribution in terms of the symbol Z, standing for the ratio of the mean value of the sample to the standard deviation as estimated by the followers of Professor Pearson by dividing the sum of the squares of deviations from the mean by the number of observations. Later he appreciated the advantage of using the number of degrees of freedom as a divisor in estimating for variance, and at my suggestion adopted the further modification of comparing the mean (or its deviation from its expected value) with its own standard error, rather than with that of a single observation; the joint effect of these modifications being to use the quantity t connected with Z by the relation $t^2 = nZ^2$ where n is the number of degrees of freedom.

By doing this he overcame the difficulty that for increasing values of n the value at which Z becomes significant is diminished without limit, whereas that at which t becomes significant tends to a finite limit determined by the normal distribution. In 1925 'Student' and I published a series of four short papers in *Metron*, giving the new Tables which 'Student' had calculated, with an explanation of their applications, method of calculation, notation, etc.. I did not decide to use the symbol z for my more general distribution until I was assured that 'Student' was convinced of the desirability of changing his notation and the form of presentation of his Tables. Consequently the use of Z for 'Student's' distribution is an occasion of unnecessary confusion to workers in this subject.

With respect to the 'method of pairing', since in 1908 'Student' had solved only the problem of the significance of the mean of the unique sample, the only practical applications to experimentation which he could exhibit were those in which pairing had been used. He has himself expressly disclaimed the invention of the method of pairing, 'which', he says, 'must be as old as Noah', and to which the improvement he has made in the *theory of errors* is not now confined in this application. His protest was, of course, aroused by the conviction of Professor Love that pairing was an essential part of 'Student's' method.

Pray excuse this long digression on a point which you may think unimportant, and which is certainly principally of academic and historical interest.

Fisher to C. Eisenhart: 28 *October* 1938

. . . I think your idea of a book on 'Readings in Statistical Methods' on the

lines of a source book is really a brilliant one. I think mathematics has suffered more than most subjects from the professional teacher acting as rédacteur and with the best intentions — usually the elimination of difficulties and complications — often presenting the ideas of his originals very inadequately. This is especially so when, as is too often the case, he is himself only familiar with their work at third hand.

Naturally, the selection of such an anthology will be greatly a matter of judgment, and what is worth choosing will depend greatly on the interests of the supposed readers. The field chosen might be very wide, or considerably more narrow, and this is a point which it will be necessary to think about carefully, I should judge, at all stages during the growth of the collection.

As regards my own work, I have tried to look at it from the point of view of such a compiler, and I send you a bibliography in which the various strains of ideas have been followed by classifying the different papers.[1] My early work in mathematical statistics was principally concerned with giving the exact solution of a number of problems of distribution which, until then, had been almost totally neglected. These lead to the exact tests of significance. Initially these problems all appear under different disguises, although later the whole system could be telescoped, and can now be taught quite briefly in a course on the *Analysis of Variance*. I have marked the whole series of which this is the central idea with an α, putting it in brackets where the new solutions were given, incidentally, as problems arose in special lines of work. The α series really culminates in 1928–9 with such problems as the *general sampling distribution of the multiple correlation coefficient*, the *tests of significance in harmonic analysis*, and the general formulae for *moments and product moments of sampling distributions*.

Closely associated with α was the interest in the more abstract and philosophical problem of estimation. I have marked these papers β. The practical results of these two lines of research have been best made known in *Statistical Methods*. β leads naturally to γ, a series of papers on inductive inference dealing with inverse probability, likelihood and fiducial probability. It is easy to see what totally different emphasis these series deserve, according to the interests of the reader. The group of applications (δ) concerned with heredity, evolution and eugenics commences almost as early, and inevitably overlaps on theoretical points; thus the 1918 paper [*CP* 9] has the first hint of the analysis of variance, which was not fully developed until the 1923 paper on crop variation [*CP* 32]. The 1919 paper [*CP* 11] also contains a pretty solution of the distribution problem. The later papers on human genetics are also very largely methodological. I have marked ϵ papers much concerned with the interpretation of experimental data, and these lead naturally to η, concerned primarily with the design of experiments, always in conjunction with their interpretation.

I am afraid that numerical illustrations of statistical methods not previously used occur practically throughout the series, so that a classification for which

this feature was particularly important would cut across the one I have attempted. I have also marked ζ for the beginning of a series on the interpretation of multiple measurements, which, from the point of view of tests of significance, fits into the analysis of variance, but involves ideas with very different applications. . . .

¹ The following list showing Fisher's classification of his publications was kindly provided by Dr. Eisenhart.

α: *CP* 4, (11), 12, 14, 19, 20, (22), 30, 31, (32), 34, 35, 36, (38), 43, 44, (46), 49, 50, 56, 61, 62, 63, 74, 75, 83, 90, 91, 123, 141; *SMRW*.
β: *CP* 1, 12, 18, 42, 62, 71, 108; *SMRW*.
γ: *CP* 84, 95, 102, 109, 124, 125, 137.
δ: *CP* 9, 11, 15, 24, 25, 47, 55, 59, 65, 68, 69, 70, 71, 72, 81, 86, 87, 93, 96, 105, 106, 113, 114, 115, 116, 118, 119, 130, 131, 132, 133, 135, 141, 142, 143, 144, 153; *GTNS*.
ε: *CP* 16, 21, 22, 32, 37, 39, 51, 58, 64, 77, 104.
η: *CP* 48, 57, 78, 79, 85, 113, 114, 115, 131, 132, 139.
ζ: *CP* 138, 141.

A. *Fisher to Fisher:* 5 *May* 1931

From an announcement by the Agricultural School of Iowa, I notice that you are to give a series of lectures in Iowa City during the coming summer. Someone also told me you expect to arrive in New York at the beginning of June. When you do come, be sure to look me up, so that we can have a chat together. Both of us are — if I so may say — somewhat unorthodox fellows, and although we may not agree on some questions, nevertheless we ought to have a jolly time in talking about other things.

Moreover, another important question should be given some attention. For unless I am totally mistaken, you are not exactly what might be described as a teetotaller. The strain of being continually pure may therefore eventually pall upon you, and you probably will have to unbend as it were so as to bring about the proper reactions in the way of slight dissipations. Permit me therefore to be a guide to you in the exploration of various places of sin such as our American 'speakeasies', which, by the way, offer a most excellent, nay almost perfect, sample of a typical cross-section of American society. At any rate, be sure to look me up.

Now for an entirely different matter. I know perfectly well that you — as what one might call a civilized person — do not take yourself too seriously. But this is not always the case among your adoring disciples, who, in their effort to expand their own egos at the expense of their environment, are apt to overdo a good thing. Take our friend Hotelling, for instance. According to him the adherents of the classical theory of probability must have sunk into a state of ease, almost approaching apathetic luxury, so that the time has become ripe for the Messiah, and lo, according to Hotelling, you are the saviour to lead the statisticians out of the wilderness. As pointed out by Norman Douglas, renunciation and hope, these spiritually correlated Siamese

twins, have always exercised an irresistible charm to polite society. It makes us all feel so comfortable to be told on the one hand that we are going to the devil and on the other hand that the coming Messiah is in the near offing. But what do you think of this new role which has been assigned to you?

Some of your younger disciples, Dr. Shewhart for instance, are like the early Christians imbued with an almost religious fervor. But youth oftentimes speaks from the heart rather than from the brain. . . .

As to the whole theory of small samples, I am myself rather an old hand at the game. In the first decennium of the present century, when I was not yet out of my teens, I became interested in the problem of small samples in connection with the yeast physiology of sparkling wines at a time when some of my near relatives were engaged in the wine industry and in the export of high-grade wines from California. I originally consulted some of my friends . . . in Copenhagen . . . and they told me that the problem, apart from its purely physiological and bacteriological aspects, was mainly one involving mathematical statistics and advised me to consult the late Mr. Gram. I then found that considerable work and that practically a finished theory had been developed both by Gram and by Professor Thiele, some of which, although not nearly all of it, had been described in Thiele's Danish 1889 text on *Almindelig Iagttagelseslaere*.

Bearing in mind these early facts, I am therefore rather surprised, in reading an article by your assistant, Dr. Wishart, in the July 1930 *Biometrika*, of which you were good enough to send me a reprint. Wishart refers to Thiele's work as a paper. This is wrong. The book is a complete and self-sufficient text. Its full title, translated in English, is: *Lectures on the General Doctrine of Observations*. Moreover, Wishart's mention of the symmetric functions by Cayley, and other writers, and that 'later workers were probably unaware that the tables of symmetric functions had already been published in works on pure mathematics', seems to me to be a statement of such radiant improbability that little credence can be placed thereon. For the facts surely must have been known to English readers since 1903, because the tables by Cayley and by the German, Fiedler, as well as Thiele's own tables of 1889 were specifically mentioned in the English text on *Theory of Observations* by Thiele, published in 1903. . . .

There are a lot of other things in your colleague Wishart's paper, which, at least from a historical standpoint, is open to criticism, and we may talk it over when you arrive in this country. Perhaps you may also be able to elucidate me on what your famous 'maximum of likelihood' really implies. . . .

But why go further? The quarrel over inverse probability and over Bayes' Rule is bound to go on between the mathematicians. To me it seems that we take them altogether too seriously. . . .

I am more inclined to quarrel with you over the introduction by you in statistical method of some outlandish and barbarous technical terms. They stand out like quills upon the porcupine, ready to impale the sceptical critic.

Where, for instance, did you get that atrocity, *a statistic*? . . .

Why is it that your disciples always are so serious? Whenever I see them ascend the rostrum to expound — *ex cathedra* as it were — their pet dogmas in a solemn and gravelike manner, I cannot help thinking that they take themselves altogether too seriously and that humour could be used far more advantageously for administering a wholesome dose of truth.

But be this as it may, however. In spite of my scepticism towards some of your work, I cannot help having a sneaking fondness for R.A. Fisher. He is so refreshingly unorthodox, a cooling breeze as it were in the sultry air of our academic circles. Therefore by all means look me up whenever you happen to be in New York. We need a breeze.

Fisher to A. Fisher: 15 May 1931

Many thanks for your long, grave, frivolous, breezy and generally Rabelaisian letter. I must certainly try and hunt you up as I pass through New York, probably on Monday, June 1st; but I shall be in the devil of a hustle and dry as a cinder, for, though I have no doubts about Copenhagen, even your eloquence has not convinced me that New York is the place for a binge.

You ask me a lot of fine enigmatical questions, of which I have found one that I can half answer. I do not know in the least how seriously I take myself, and not being an alienist I never met anyone who did, but I certainly don't take myself seriously enough to feel any responsibility for justifying, explaining, or excusing anything written by anybody else. Stuff from my own laboratory, the publications of which I officially 'approve', is passed out on the comprehensive ground that I approve of anyone who has thought about a subject expressing his own opinion over his own name. This probably needs explaining, as a good many continental institutions and faculties take a very different view of the matter, and even Denmark may have been infected by this characteristically Alpine view, prevalent in Italy, France and Germany, of intellectual organisation. I think my attitude is fairly general in England, though Pearson's laboratory is an extreme exception, and, even if I had not believed it to be a salutary national tradition, I should have adopted it in view of the 'awful warning' which his laboratory provides.

You must have observed for longer than I have, one remarkably interesting characteristic of your adopted country, I mean its susceptibility to the influence of vogue. Pondering upon it, I have often thought that it brought to American Science quite as many advantages as disadvantages. It is a real advantage that new ideas should sometimes be tried out, quickly and eagerly, up to the limits of their usefulness, even if this involves many blunders, inflated reputations, and misapplications of research funds. In my own country a man's reputation is normally at its highest just 20 years after his death, at which date his copyrights expire, but we are sometimes so impatient, in the case of the very long lived, as to applaud their opinions

when, apart from the effects of senility, they are about 50 years out of date. In fact, I have felt that, if an eye were kept on its weaknesses, the reign of Vogue was rather beneficial. It was not until I saw Mr. Grove's review of my book on Statistical Methods in the *American Mathematical Monthly*[1] that the humourous and other possibilities of *rival vogues* occurred to me, and I wondered if you, of whom, among others, Mr. Grove evidently regarded himself as the spokesman, were prepared to stage Danish and English statistics, like rival mannequins, jealous for the privilege of setting the fashion. For myself I found the scene rather disgusting, though admittedly extremely funny.

Personally, I am rather annoyed; partly because five to ten years ago I had quite an uphill job to free English opinion from the authoritarian opinion that all enlightenment was to be found in the immaculate gospel of Pearson, and I see no need or profit in undertaking the same thankless task for Thiele; partly because I have always been a particular admirer of the Danish work; and yet it looks as though, for the next five years or so, I should have to add to every explanation of a modern result some sort of codicil as — 'this is how Thiele blundered over this' — 'This is how Gram handled that' — 'This is how Arne Fisher or Frisch or Steffenson still approach the matter'. Intellectually it will be easy enough. 40 years do not pass without some progress, and progress in the last 10 has been particularly rapid. A result that is worth publishing is an improvement on the methods of *all* previous writers, and the more active they are the more surely will they be found fumbling at it ineffectually. But that is not the way I should prefer to treat either the great writers of the past or the ablest contemporary workers. I should prefer to write for their appreciation, as I should talk over a dinner table.

On the whole, then, I would rather give my beard a Messianic cut, if this would encourage Shewhart and Hotelling to go ahead and do their job in a workmanlike manner. In any case you will not expect me to be too lenient with the doctrine of the infallibility of the Holy Writ of the nineteenth century, whether in Danish or in English.

P.S. I had forgotten your point about vocabulary. I use special words as the best way of expressing special meanings. Thiele and Pearson were quite content to use the same words for what they were estimating and for their estimates of it. Hence the chaos in which they left the problem of estimation. Those of us who wish to distinguish the two ideas prefer to use different words, hence 'parameter' and 'statistic'. No one who does not feel this need is under any obligation to use them. Also, to Hell with pedantry!

I see I have coupled Pearson's name with Thiele's too often. I do *not* mean that his work is rotten with mathematical errors, only that he had no more glimmer than Pearson of some of the ideas we now use.

[1] Grove, C.C. (1930). *Am. Math. Month.* 37, 547-50.

Fisher to R. Frisch: 6 *March* 1931

I am much obliged for the copy you send me of your letter to Prof. Hotelling on the question of the terminology you advocate for the coefficients of the logarithm of the characteristic function. It is perhaps a pity that Laplace did not once and for all give a name to these coefficients, as well as to the function in the expansion of which they occur, which function is of course analytically of much the greater importance. Perhaps we may take it that he left it to the good sense of subsequent mathematicians to say explicitly what they meant, without quarrelling about points of terminology.

The concurrent use of alternative terms for the same objects does not in practice give rise to any inconvenience, and has the great advantage that each person may suit his own convenience; by this process moreover, the more generally useful term, in respect of brevity, of its descriptive quality and of the facility with which it associates itself with a wider terminological system, will in time inevitably come into the more general use.

This, of course, is what we should all desire, in a philosophical spirit, even if our own special associations have given us a strong preference for one particular usage; and it is perhaps all to the good, that willy-nilly this is the process which slowly perhaps, but inevitably, determines the use of words in the English language.

Apart from its associations, which perhaps ought to weigh with a Scandinavian, I have, personally, always felt a certain awkwardness, both logical and linguistic, about the use of the term half-invariants and semi-invariants, and should be glad to see them replaced by more suitable terms, which should of course include a term for the function in the expansion of which they appear. Mathematically these coefficients are not in any sense invariant; and it is not very clear by what process of calculation the fraction *half* is arrived at! Linguistically, 'half-invariant', the use of which by Thiele gives it the stronger claim to consideration of the two, is what is called a bastard word, and for this reason will not be readily admitted into good usage. The term 'semi-invariant' is an attempt to repair this defect, but is scarcely so euphonious as to make up for its other deficiencies. There is an added inconvenience if one is led to speak of the semi-invariantive function, which would seem to be the proper cognate term.

You do not develop in your letter any objections to the terms cumulant and cumulative function, though one is led to feel that you have some strong objections which you do not mention. I must conclude, therefore, that your objection is based solely upon the undesirability of making a change in a usage with which you are already familiar. I should advise that whenever such a change of usage seems inconvenient, it should not be made, and then the newer terms will be used only by those who feel that they have real advantages. You will not, however, and I am glad to see that your letter does not, question the right of mathematicians to make use of whatever terms they find most suitable for the expression of their meaning.

R. Frisch to Fisher: 7 *April* 1931

Thank you very much for your letter of March 6. Of course, I do not question the right of mathematicians to make use of whatever terms they find most suitable for the expression of their meanings. I have only expressed my personal opinion in the matter of the term 'semi-invariants'. I believe that no pressing need exists for introducing another term than this and that the mere fact that the term 'semi-invariants' has been used for a long time is an asset with this term which weighs very heavily.

As to the term 'cumulants', the term is not self-explanatory and I was not present at the meeting when Hotelling explained it at Cleveland, so I probably do not understand quite what it is meant to cognate. Does it refer to some properties of semi-invariants of a *function* of observations? Or does it refer to the fact that these parameters are defined by integrals, that is, by 'cumulation'? In the latter case it seems that the term cumulation is just as proper in connection with the original moments as with anything else.

Whatever the situation may be in this respect, I must admit that I have still some difficulty in seeing the necessity of a change in the present case.

I have recently received a letter from Dr. Craig who is of the same opinion as me in this matter.

Fisher to R. Frisch: 17 *April* 1931

I wonder if I shall have an opportunity of discussing these questions of terminology or what will be more interesting, the subject matter of them, when later in the Summer I shall be at Minneapolis. I had, of course, fully understood from your previous letter that you would prefer to retain Thiele's term, and I had hoped to make quite clear that I saw no reason why you should not.

The cumulative property has nothing to do with the use of integrals in the definition of the population moments and thence of any parameters functionally related to the moments. It is the property, of which both Laplace and Gauss fully saw the importance, in studying the cumulative effects of independent causes, namely that certain parameters, such as the mean, the variance, etc., of the resultant distribution were simple sums of the corresponding parameters of the component distributions. It is this property of what I call the cumulants that has attracted the attention of mathematicians from the first. In the higher theory, of course, the importance of the property is seen to lie in its application to the cumulative function, the logarithm of the characteristic function in Levy's sense, since this is not always to be expanded in a power series.

For my own part I am perfectly content to leave it, as in the end it must in any case be left, to the convenience of teachers and users of mathematics to decide what terms they shall apply. There is no need, therefore, for me to reiterate the reasons for my own preference.

Fisher to R. Frisch: 19 *March* 1934

I have looked through Arne Fisher's paper and, as you asked me, am sending you my comments. The paper seems to have two objects, which might perhaps with advantage be kept distinct; one is to establish the merit and priority of certain work by Thiele and subsequent work as influenced by him, especially in recent times Bertelsen and Arne Fisher, and the other is to deplore the influence in contemporary America of some recent English work with which the critic is much less familiar. The confusion of the two subjects weakens the paper through the author's anxiety to say at the same time (i) The formulae of which modern writers are so proud are to be found identically in Thiele's writings and (ii) The modern work is erroneous in comparison to Thiele's. The confusion is also, I suggest, generally undesirable in making it difficult to refer with ordinary generosity to a distinguished earlier writer from whom one may occasionally differ, if in such references one is forced to dissociate himself from the kind of well-meaning but indefensible claims which a modern writer like Arne Fisher has put forward on his behalf.

For example, it seems proper to give Thiele credit for first giving the formulae corresponding to the partitions (2^2), (2^3), (3^2) and (4^2) but rather childish to make a song about the formulae for 1^r which must have been familiar at least since Laplace introduced the characteristic function, or about the mean values of $\mu_2, \ldots, \mu_6$, i.e. of the primary series of statistics with which he chose to work. But the set of results given by Thiele can scarcely be claimed as giving the analytical expression in general for partitions of the type r^i, still less for the more general partitions in which the parts are unequal. When I wrote my paper on the solution of the general problem by partitional methods I was surprised in looking through the works of Tchouproff and Church to find how few seemed to have been given by previous writers, and, had I known of it, should have been glad to cite Thiele amongst the others as the earliest writer who had realised that exact formulae of this kind could be derived by direct algebra.

Again I think it is now clear that Thiele's system of presumptive values is not the same as my series of k-statistics and the implication that it is the same does no credit to Thiele, for his intention was clearly different from mine. His principle in making presumptive values was to choose those values of the parameter for which the observed values of his h statistics would be equal to their expectations and this is not the same as the definition of a presumptive value given as Tchouproff's on p. 6. I imagine that Thiele knew what he was about and did what he wanted to do whereas Bertelsen seems to have done as I did and chosen statistics having K as their mean values. Why Bertelsen did so is not explained, and should be of interest, but the implication seems to be that he was aiming at an improved method of estimation, and in consequence fell into the grievous error, of which I alone am accused in the numerical example in the attached sheet. Bertelsen's aim must evidently have been

different from mine since he was a good boy to do in 1927 what was very naughty of me in 1929.

As far as I am concerned anyone who troubled to read my papers would know that I was not concerned with estimation at all, first because, in the paper on the theory of estimation [*CP* 42], I am particular to dismiss the question of bias at an early stage as of no interest to the theory and in the paper on moments of moments [*CP* 74] I go out of my way to point out that moment functions only provide statistical estimates of high efficiency for a special type of distribution (p.200) with a reference to the paper in which I first demonstrated their great inefficiency for all of the Pearsonian types of curve which depart at all widely from the normal [*CP* 18]. My interest in the formulae arose solely from the facts that they can be made exact and general for all distributions, that very few of them had been previously given, and these in very complicated forms, and that no general notions seem to have been developed as to their classification, method of derivation in the more complicated cases, or extension to multivariate distributions. In fact what I had done was just what Craig in a paper published nearly contemporaneously suggested needed to be done if the formulae were to be brought within manageable compass.

The quotation given by Arne Fisher in translation from Thiele's paper 1889, bottom of p. 5, is of interest as showing that Thiele also was only partly concerned with estimation and that he was quite aware that in the current state of the subject he was unable to give any satisfactory account of the principles on which methods of estimation should be based. It seems to me much more to the credit of a man of science that he should be aware of the limitations of his own knowledge and of that of his age than that he should be represented as opposed to the very advances of which, as an active thinker, he must have felt the need. . . .

Fisher to R. Grant: 17 *May* 1950

I am enclosing a letter from Chicago[1] which may interest you, as they are taking the quite extraordinary step of celebrating the 25th anniversary of *Statistical Methods*. Of course, this is partly coincidental, in that its first edition just struck the quarter century, and it will be the half century that they will be particularly concerned to memorialize.

I should not be able without great difficulty to meet their request for a personal account of the origin of the book, because, partly by reason of the book's effect, the whole body of possible readers will be now taking for granted points of view unknown in 1925. However, I will think about it, and may be able to do something. Perhaps you would also consider whether some account of the book as a publication would be worth supplying.

[1] See W.A. Wallis's letter to Fisher (p.325).

R. Grant to Fisher: 19 May 1950

It was a real pleasure to receive your letter of 17th May with the letter from W. Allen Wallis, Committee on Statistics, University of Chicago, U.S.A. which you enclosed. A letter such as this surely goes far to make work worth while!

It all takes my mind back to that day when Frank Crew called relative to your manuscript, how he spoke of its quality, the formative work that it contained, and urged publication if only on the grounds that statistics in future would and must form part of research work in every science.

Some publication particulars have been extracted from the copyright ledger and they are enclosed.[1]

Now I feel there is no one who could write a more acceptable note, such as is required, than Frank Crew. Accordingly, a copy of these publication notes has been sent to him with the request that he gives this suggestion his consideration.

[1] See *J. Am. Statist. Ass.* **46**, 31 (1951).

Fisher to R. Grant: 23 May 1950

Thanks for your letter of 19th May. I shall be very much interested to see what Crew makes of your little assignment. It was [D.W.] Cutler who approached me, probably after consulting Crew, and certainly he came at the right moment, for I did not have to do any mathematical research *ad hoc*, but only had to select and work out in expository detail the examples of the different methods proposed. It was often quite a job to find a good example. Indeed, some were not so good and have been omitted in later editions. . . .

M. Greenwood to Fisher: 10 June 1933

I heard yesterday unofficially but I hope correctly that you are to be the next Galton Professor and should like to be the first to congratulate you and wish you many years of happiness and successful work in the university. There is, in my opinion, no other man alive worthy to sit in old K.P.'s chair; his mantle has descended on your shoulders and a double portion of his spirit is yours. Like all of us, you owe him much. If K.P. had never lived, you would have assuredly been one of our foremost men of science but very likely your field of work would not have been statistics, but perhaps, pure mathematics. You will repay that debt, not by silly adulation of the 'illustrious predecessor'—as so often happens when a great personality leaves the stage—but by carrying the work further. This appointment really makes me happy for—although K.P. has often enough wounded my feelings and insulted my friends—I still love the man and venerate his genius; I should have been sad if his kingdom had fallen into the hands of a second-rater. I hope that when you are my

neighbour[1] we shall often meet and that if there is any way in which I can help you you will not fail to let me do so.

[1] Greenwood was Professor of Epidemiology and Vital Statistics in the School of Hygiene and Tropical Medicine in the University of London.

Fisher to M. Greenwood: 13 June 1933

Thank you immensely for your very kind letter, which expresses better than I could do just how I should like to carry the work forward. I understand that great circumspection and some restraint will be needed to avoid hurting feelings unnecessarily, and this is my chief concern in accepting the post. It will necessarily be some time before the department can be a coherent unit, but I look forward to establishing helpful relations with your School, as well as with the other departments at University College.

Fisher to H.W. Heckstall-Smith: 9 February 1956

Thanks for your letter. I have always regarded the F-test and the z-test as the same, and indeed in 1924 I calculated the variance-ratio as a means of getting their natural logarithms.[1] Of course if everyone had a computing machine at hand on their desk the variance ratios are the quicker, but those without computing machines, and who can use four-figure tables, are about five times more numerous.

Owing to the misapprehensions that have been spread abroad, especially in the United States, I have added in the 12th edition of *Statistical Methods* (page 226)[2] a short historical note which I hope will prevent expositors from representing the F-test, or as the French have got it 'test de Snedecor', with the z-test, or 'test de Fisher'. I think it was only an afterthought that led Snedecor to say that the capital F he had used was intended as a compliment to myself.[3] In any case you might find it useful to bear in mind that some of your readers are not equipped with calculating machinery, and are used either to slide rules or to logarithms. Of course the differences between logarithms to different bases are quite trivial in practice. . . .

[1] Heckstall-Smith had suggested that in an article he was writing for a medical journal he would need to use the F-test rather than the z-test because natural logarithms would not be tolerated.
[2] *SMRW*, p. 229.
[3] See also Fisher's letter of 16 February 1938 to Snedecor (p. 323).

H. Hotelling to Fisher: 21 October 1930

. . . C.C. Craig, who is here for the year, tells me that the distribution in samples of the standard deviation was published by F.R. Helmert in the *Zeitschrift für Mathematik und Physik* in 1876. His proof is copied in Czuber's *Theorie der Beobachtungsfehler*. . . .

Fisher to H. Hotelling: 14 November 1930

I have just been looking at Helmert, and unless I have misunderstood him he certainly has not anticipated 'Student' in giving the exact distribution of the variance as estimated from squared deviations.

He seems to be only concerned with the distribution of quantities such as

$$(\epsilon_1^m + \ldots \epsilon_n^m)/n$$

when the distribution of ϵ is known; not with the distribution of sums of powers of deviations *from the mean of the sample*.[1]

The last three sections seem to be only a crude kind of discussion of the theory of estimation from such sums of powers as he is discussing, but I cannot find that any account is taken of the fact that the mean from which the deviations are measured is also subject to errors of estimation.

Will you look at it and confirm or correct my views?

[1] Compare with Fisher's letter of 24 July 1934 to E.B. Wilson (p. 326).

H. Hotelling to Fisher: 26 November 1930

. . . Today I read Helmert's article in the *Zeitschrift für Math. u. Physik* for 1876, and agree with you that it is pretty crude. The distribution he reaches on p.203 is the same as that of χ^2, or s^2/σ^2, with a number of degrees of freedom equal to the number in the sample. As you remark, he makes no allowance for the errors of estimate in the mean, though it would have been easy to do so. Of course he missed the vital idea of the distribution of $(\bar{x}-m)/s$. . . .

Fisher to H. Hotelling: 1 May 1931

. . . Is it not a curious fact historically that the first appearance of the normal distribution was as an approximation to the binomial series, whereas if Bernoulli and de Moivre had been more advanced analysts and had got the expression for the true sum of a broken binomial series, they would have found the integral of the z-distribution?

About semi-invariants I find that Thiele actually defines them as the coefficients of $t^r/r!$ in the expansion of

$$S(e^{tx})/n$$

where x is the value observed in the sample. They are therefore in his usage primarily statistics, and only by the current contemporary confusion between statistics and parameters were used to designate the population values. If the point is to be governed by historical accuracy, therefore, now that the distinction has been made and is regarded as essential by statisticians, the term 'semi-invariants' should be confined to the series of statistics

$$m_2, m_3, m_4-3m_2^2, \ldots$$

while 'cumulants' is the first term to be given specifically to the population

parameters, of which Thiele's semi-invariants may be regarded as estimates. As such they are consistent, but as Craig's work has shown, introduce serious algebraic difficulties. Still it is all to the good to have a distinctive name for them.

Fisher to M.G. Kendall: 19 August 1940

Many thanks for your letter and the paper enclosed.[1] I am very pleased to have the latter for the *Annals*, as it makes a worthy addition to the two valuable papers you have written on the subject previously.

I am glad also you find the term 'cumulants' satisfactory for purposes for which I use it. My own reasons, as usual in such cases, were somewhat complex, but I was influenced by the fact that Thiele introduced his term 'half invariants' for a system of statistics calculable from the data, and only later, in the rather casual manner of his time, came to use them for the parametric moment functions to which French and American writers especially later came to apply the term semi-invariants. I have noted Thiele's usage in *Statistical Methods*, pp. 75–78 of the 7th edition.[2] Of course I also wanted a term rather more distinctive than the general term, introduced I believe by Cayley, for the wide class of symmetric functions. I fancy the usage you suggest now corresponds with Cayley's original meaning. Thirdly, I felt some need to avoid such a cumbrous phrase as 'the semi-invariant generating function' for what I call the cumulative function. . . .

[1] Published eventually as Kendall, M.G. (1940). Some properties of k-statistics. *Ann. Eugen.* **10**, 106–11.
[2] *SMRW*, p. 72.

Fisher to F.J. McGuigan: 8 April 1958

Thanks for your interesting letter.[1]

I chose the term 'null hypothesis' without particular regard for its etymological justification but by analogy with a usage, formerly and perhaps still current among physicists, of speaking of a null experiment, or a null method of measurement, to refer to a case in which a proposed value is inserted experimentally in the apparatus and the value is corrected, adjusted, and finally verified, when the correct value has been found; because the set-up is such, as in the Wheatstone Bridge, that a very sensitive galvanometer shows no deflection when exactly the right value has been inserted.

The governing consideration physically is that an instrument made for direct measurement is usually much less sensitive than one which can be made to kick one way or the other according to whether too large or too small a value has been inserted.

Without reference to the history of this usage in physics, I think one can say

that your second interpretation is very nearly right. One might put it by saying that if the hypothesis is exactly true no amount of experimentation will easily give a significant discrepancy, or, that the discrepancy is null apart from errors of random sampling.

Thanks for raising an interesting point.

[1] McGuigan asked what Fisher had in mind when he introduced the term 'null hypothesis', saying that he knew two possible interpretations: (i) that it was an hypothesis asserting that the difference between two groups is null; (ii) that the term came from the word 'nullify' and a null hypothesis is any hypothesis set up for the purpose of being nullifieid.

Fisher to W.L.B. Nixon: 13 January 1956

I was extremely interested in your long letter.[1] . . .

The sources of data given in *Statistical Methods* refer to J.A. Harris (1913) 'On the calculation of intraclass and interclass coefficients of correlation from class moments when the number of possible combinations is large', *Biometrika*, ix, 446–472. This may not be the earliest use of the term intraclass correlation, but it was to Arthur Harris's credit that he distinguished correlations of the interclass and intraclass kinds which, of course, had been confused in the literature of the previous twenty years. Indeed it was largely to give Harris the credit of having made this advance and to show how far it led, i.e. to the analysis of variance applicable to many other problems, that I gave the space to intraclass correlation. . . .

[1] Nixon asked about the origin of the term 'intraclass correlation'.

V.I. Romanovsky to Fisher: 4 December 1935

I would be very obliged to you, if you indicated me how are established two approximative formulae, p.221 of your *Statistical Methods for Research Workers*.

I am also puzzled why you use, in the analysis of variance, $z = \frac{1}{2}\ln(s_1^2/s_2^2)$ instead of s_1^2/s_2^2. . . .

Fisher to V.I. Romanovsky: 20 December 1935

I am glad to have your letter.

So far as I remember I obtained the approximations[1] for the test of the significance of z where both n_1 and n_2 are large by obtaining the moments of the distribution of z, or rather its cumulants, from its characteristic function. I forget the details, but clearly the factor $(1/n_1)-(1/n_2)$ is a simple allowance for the third moment, while the first term is derived from the normal distribution.

I had a good many reasons for using z instead of some function of it in the test of significance in the analysis of variance. One important reason was that in order to make a compact table it is necessary that the test value should be well interpolated by what I call asymptotic interpolation using the reciprocals of the numbers of degrees of freedom, and this is more true of z than of any other simple function. A second point is that half the tabulation is saved by the fact that reversing the sign of z and interchanging n_1 and n_2 we have the 5% and 1% points at the opposite ends of the distribution. Finally, the close analogy between interclass and intraclass correlations is paralleled by that of the values of z obtained from r by the same transformation. The advantages of this transformation I have set out in the book. . . .

[1] *SMRW*, p. 234.

Fisher to G.W. Snedecor: 16 February 1938

. . . I am afraid I have a difficulty which you may not know of in assigning the particular symbol F to the variance ratio,[1] and that is, that a table of this value derived from my table of z was published in India, I think in *Sankhyā*,[2] before yours.[3] The author, if I remember right, used a different symbol — probably x. I feel it would be too complicated to make a sub-heading using the different designations.[4] . . .

[1] Snedecor had suggested to Fisher that 'it would be clarifying to associate your variance ratio table in some manner with F.'
[2] See Mahalanobis, P.C. (1932). Auxiliary tables for Fisher's z test in the analysis of variance. *Ind. J. Agric. Sci.* 2, 679–93.
[3] Snedecor, G.W. (1934). *Analysis of variance and covariance*. Collegiate Press, Ames.
[4] See also Fisher's letter of 9 February 1956 to Heckstall-Smith (p.319).

Fisher to L.H.C. Tippett: 12 November 1931

Many thanks for sending me the copy of your book *The methods of statistics* which arrived here this morning. I had read some of it before and was very glad indeed to see how much use you were making of the methods which from time to time I have put forward. This is a real encouragement to me, and since it has been obvious for some time that most of the books on statistics are much out of date, I hope and think it will help your book, though of course it will expose you to the attack of the more recalcitrant of the Pearson group. . . .

You have evidently found Woo's table[1] useful, which surprises me a little as this involves subtracting $\overline{\eta^2}$ (written $\bar{\eta}^2$), and dividing by σ_η. This makes you use the term correlation ratio in a much wider sense than is usual, for example in place of R the multiple correlation coefficient. Of course, once the analysis

of variance is grasped the equivalence of the test of significance for the whole group of problems including η and R is apparent. But it is the analysis of variance itself which keeps the arithmetic straight and easy, and not the possibility of regarding the ratio of two sums of squares as a sort of η^2. I wonder they did not make Woo work the actual 1 per cent and perhaps 5 per cent points for η^2, as even using interpolation it would have been easy to get the values accurate.

I think your book will do a lot of good and help a good many people. The examples seem to me especially well-chosen.

---

[1] Woo, T.L. (1929). Tables for ascertaining the significance of association measured by the correlation ratio. *Biometrika* 21, 1–66.

L.H.C. Tippett to Fisher: 15 November 1931

Thank you very much for your letter of the 12th and for the comments about the book. . . .

I personally never use η^2 as a test of association, but it *is* used, and therefore seemed desirable that it should be linked up with the other tests. I think that one of the most hopeful ways of getting the Pearsonians to use the Analysis of Variance is by showing its equivalence to η^2. That the former is much neater and easier to calculate should be obvious. . . .

Fisher to L.H.C. Tippett: 17 November 1931

Thanks for your letter. . . .

I am interested in what you say about the Biometric Laboratory and the analysis of variance. In the highest quarters I fear a presentation of the newer methods is acceptable, if at all, only if their origin is disguised; witness Mr. Woo's being introduced to the true distribution of η^2 through Hotelling's paper which Hotelling wrote, as he tells me, in a state of forgetfulness, after having read my 'Goodness of fit of regression formulae' [CP 20], in which the distribution is first given. In this circle my work has now passed through the stages of being regarded as

i) criticism so impudent and ignorant that it may be ignored;

ii) applicable only to such special cases as to be without interest;

iii) the stage of containing some points of value, which are, however, better obtained from other sources.

But apart from this very small knot of incorrigibles I should advise the use of the analysis of variance, with which, as it is only a convenient arithmetical arrangement, no author need be associated and to forget the fancy symbols which from time to time and in various particular cases have been attached to the ratios which may be deduced from the analysis. Of course η^2 is the limiting value of R^2 when the number of adjustable constants is made equal to the number of arrays and it may be useful occasionally to know that R^2 can be

made no higher. But even in the case of R^2 the arithmetical presentation gives a much clearer idea of what it means than that it is the square of a generalised sort of correlation coefficient.

I am very glad to hear about the lectures. I am sure they will do no end of good. I find myself that students knowing little mathematics become extremely keen on these 'advanced' methods simply because, I think, on working through the arithmetic (and certainly not without) they find them readily intelligible, and they take no harm from knowing that they are far more exact mathematically than what is taught to astronomers and biometricians.

W.A. Wallis to Fisher: 10 May 1950

This year marks the twenty-fifth anniversary of the publication of *Statistical Methods for Research Workers*. The *Journal of the American Statistical Association*, of which I have recently become Editor, hopes to mark this event by one or two articles on the character and consequences of that volume. In that connection, I would appreciate your help on two points.

First, can you suggest two or three persons whom I might invite to prepare articles? I naturally want people who have a thorough comprehension of, and sympathy with, your work, so that we will get a treatment that is both deep and broad, and in more than a formal sense. The two names that come first to my mind are Hotelling and Cochran. . . .

Second, would it be possible for you to prepare a paper for us on the history of *SMRW*? I have in mind something on how you came to write it, on any special problems you encountered, on the revisions, deletions, and additions of the various editions, etc. If the publishers have no commercial objections, I believe that statistics on the sales, by year, country, and edition, might be of considerable interest. In general my hope is that you could without embarrassment give a sort of biography of the volume. But anything you see fit to write will be more than welcome. Maybe you would find the next twenty-five years a more stimulating topic. . . .

Fisher to W.A. Wallis: 17 May 1950

Thank you for your two letters of May 10th. I am sure that the choice of Hotelling and Cochran is an excellent one, and hope that they will both have something to say. They are, however, both professed mathematicians, and the book was, of course, written not for mathematicians but for practitioners, who, in so far as they understand their fields of application, are good judges of the kind of statistics which aids them in their work. They constitute, I believe, the real and ultimate judges of such a book as mine, though, of course, the majority of them are rather inarticulate in mathematical circles.

I should suggest, therefore, and have already suggested to the programme committee, that others from one or more fields of application should be invited to show what the book or the ideas which it was intended to express

have done for their own subjects. . . .

As to your suggestion of my preparing a paper on the history of the book, I am not quite sure what can be done, but I will at once consult my publishers and see whether they are interested in putting together an account from their point of view.[1] In preparing recently notes on Messrs. Wiley's republication of some of my old papers I have been greatly impressed by the extent to which thought has changed and by the consequent difficulty of reconstructing for modern readers the state of statistical science in 1925. So at the moment I am not sure whether I could do anything at all adequate. However, I will think about it.[2]

[1] See p.317 for Fisher's correspondence on this with Robert Grant, a director of the publishers, Oliver and Boyd.

[2] The American Statistical Association devoted a session at its meeting in December 1950 to 'The influence of *SMRW* on the development of the science of statistics' and papers under this heading by F. Yates, H. Hotelling, W.J. Youden, and K. Mather were published in *J. Am. Statist. Ass.* **46**, 19-54 (1951).

On his copy of Hotelling's paper 'The impact of R.A. Fisher on statistics', Fisher has written several comments.

On p.35 alongside Hotelling's statement that 'Karl Pearson began as a mathematical physicist, specializing in the theory of elasticity, but with the enthusiastic rise of biology that followed Darwin was drawn into his series of "Mathematical Contributions to Evolution"', Fisher has written 'Response to Galton's endowment'.

On p.36 where Hotelling says '. . . in considering the development of statistics in the first quarter of the twentieth century it is well to bear in mind the British tradition of mathematical physics, which sought primarily for useful formulae and could ignore some incompleteness in the details of their deductions, some over-condensation and abridgement of formal rigor . . .', Fisher has underlined 'useful formulae' and has written in the margin 'New quantitative operations. Heaviside. Matrix algebra'.

On p.37 where Hotelling says Gosset received his B.A. degree in astronomy from Cambridge, Fisher has written 'mathematics'.

On p.42 where Hotelling says Sir John Russell 'persuaded Fisher to leave the classroom for a temporary stay at Rothamsted in 1919', Fisher has written that 'he was unemployed, but had recently been invited by Pearson to come to U.C.' [University College].

On p.46 alongside Hotelling's statement that 'The Continental mathematical tradition of explicit and complete argument has now taken a firm hold both in this country and in England, and should help to minimize errors, misunderstandings, and controversies.', Fisher has scrawled 'Hasn't it!'

Fisher to E.B. Wilson: 24 *July* 1934

. . . A great many statistical discussions in astronomy could be made very much clearer and easier to read, and, I suspect, also to write, if the arithmetical arrangement, which I call the analysis of variance, . . . were familiar to the authors.

Historically, it is interesting that the distribution of s^2 was discovered by an astronomer, Helmert, as early as 1875,[1] but its importance was obviously overlooked, owing to the widespread use of the mean error (Peters' formula) in astronomical work. Perhaps the more so, as Peters was the editor of the journal in which Helmert published his results, which appeared as insignificant excrescences of papers relating principally to the sampling errors of Peters' formula. . . .

[1] Compare with Fisher's letter of 14 November 1930 to Hotelling (p. 320) and also *CP* 165, p. 3.