https://doi.org/10.29321/MAJ.10.A01453

Farming will never be a success unless the farmer
had more voice in the disposal of
his produce—P. Morrel.

The
Madras Agricultural Journal

(ORGAN OF THE M. A. S. UNION)

Vol. XVIII] APRIL 1930 [No. 4

ON THE MATHEMATICAL PROBABLE ERROR AS APPLICABLE TO FIELD EXPERIMENTS IN AGRICULTURE [1]

BY

PROF. M. VAIDYANATHAN, M.A., L.T.,
Honorary Reader, University of Madras

i. Apologia

These lectures which the author had the honour of delivering under the auspices of the University are the outcome of an attempt to place before the Agricultural experimentalist the mathematical limitations that govern his '*significance*,' and to explain to him in broad outline the skeleton of the statistical theory that forms the basis of judging inter-varietal differences or the differences between different treatments. To a pure mathematician, the distributions of the errors of statistics and their analytical properties form the crux of the problem, while to an experimentalist the applications of the theory with all its implications bring in not only inferences but suggest new problems and new spheres of work connected with his experimentations. For in fact in no other science are theory and practice so combined, so allied and so interdependent that most of the problems connected with agricultural experimentations have not been correctly diagnosed for want of sufficient progress in statistical ideas and statistical theory. Though indeed a high technical knowledge is necessary to study the peculiarities of the soil, to suggest ways and means for the economic use of the fertilisers, to test the different varieties in respect of the yield and to evolve strains out of them to produce the greatest effect, though these and other investigations are the

[1] Honorary Readership University Lectures for 1929-30 delivered by the author at Madras and Coimbatore.

province and the domain of the agricultural scientist, there is indeed the imperative need for the brains of a statistician in that the scientist's figures require a thorough sorting, scrutiny, and analysis which would lead to a scientific interpretation and inference. The nature of plot experimentation is such that not only at the final stages of the enquiry where the statistician has to interpret the results by a study of the different types of variations that are possible in the yields, but also even at the early stages of the field-trials where there should be a preliminary survey of how best the plots could be arranged to minimise the causes of the errors which hamper the nature of the enquiry, the need for a statistical technique is obviously necessary and even imperative. But mere statistical considerations divorced of the practical needs and simple methods of cultivation become only of academic interest; and here the co-operation of the experimentalist and the scientific statistician is more and more acutely felt.

ii. The basis of the scheme of plot experimentation

(a) General ideas

The Agricultural Research scholar is confronted with the three agencies, the *soil*, the *climate* and the *plant*. Whatever experiments are performed whether with different varieties or with different treatments, the soil and the climate are permanent factors whose influences in the final results cannot be over-estimated. Unless an experiment is repeated on the same soil for a number of years it is not possible to fix up a definite significance between variety and variety or between one manure and another. The difference between one plot and another however close they may be is so strikingly complicated, and the effects of the weather in the matter of cultivation in the same year or different years are so emphatically different that any experiment, if it should be successful, should be performed under conditions which require a preliminary study and a careful scrutiny. Next, the methods of cultivation are all important in the deduction of any inference with regard to the experimentations. High statistical significance depends upon a most careful sowing and above all upon the spacing variations of the different kinds of crops. The special effects of varieties form themselves a separate statistical study. Thus, to judge of differences in any agricultural experiment, a statistician has to study the effects of three factors—climate, soil and plant—with the due recognition of whether the experimentation has been conducted most carefully, consistent with the environment and the conditions of the soil.

(b) Different types of variation

Thus the technique of plot experimentation is a complicated process involving as it does a study or a simultaneous study with respect to the different plots of the differences between manures, between several varieties and between different methods of cultivation. The modern concept of statistics is that it is a science of the study of differences and nowhere either in Biology, Sociology or Economics has this concept been so much exemplified and applied as in this art of field-experimentation. Hence the *method of the arrangement* of the plots to bring out the differences between the several types of factors that produce the variation in them is all important in any scheme of experimentation. For a statistician to draw any conclusions each experiment has to be replicated and the necessity for the repetition of experiments as often as the homogeneity of the soil would allow cannot be over-emphasised. The modern methods of replication of experiments in a limited area constitute a notable advance in plot-technique and have given a statistical import to the very arrangement of the plots on which the external agencies have to be tried. But the number of experiments or replications possible is by the very nature of the experimentation obviously limited and a statistician accustomed to exhaustive figures as in other economic enquiries has to adopt a different theoretical attitude and a different angle of vision to the limited sample of experiments under examination. The statistical necessity therefore demands that with a limited number of 4, 5, or even 10 trials for each variety or manure, the figures must not only be accurate and reliable but must supply data and details concerning all influences at work. What are those influences? Now with small-scale trials where plots of very small sizes are under experimentation such as to bring out the differences of hybridisation, though the soil differences between plot and plot close to each other may not be appreciable, yet in dealing with long strips of several families of plants to judge of their individual characters, the plot variation is a factor that cannot be easily ignored. Even more is the case of the Field-scale trials, where plots of large areas are brought under cultivation under several varieties. Here the difference between one field and another is even more striking and any judgment of the discrepancy of the yields of the several varieties cannot but recognize the innate fertility of the soil which produces sometimes differences out of proportion to the differences in treatment or the efficacies of varieties. Thus an examination of the influences at work in the yields of the different varieties brings out *three* distinct classes of variations occurring in a series of plot-trials with replications of the same variety. Firstly, what might be called *Accidental Errors*[1] which are inevitable in any statistics arises this way, that however much the plots might be homogeneous and alike and however identical might be the treatments of manuring or varieties, the accidental variations such as unequal levelling, unequal channel supply and destruction by insects or spasmodic diseases are bound to exist between plot and plot, and assume some significance particularly in dealing with plots of very small size. Though in field-scale trials such accidental differences may be taken to be *constant* from field to field, in small-sized plots these accidental variations require a study as they would affect the yields in an appreciable way. But with a careful supervision and control of such factors as would bring in these variations the accidental error could be considerably minimised. The second class of errors which we shall call *Systematic Variation* arises in two ways and in two different and emphatic directions. Firstly, the influences of the soil referred to already produce what is known as the *Fertility Gradient* as we pass from plot to plot. If we could analyse the innate fertility of the successive plots as opposed to the effects of the new treatments, it is found that the fertility varies in a systematic manner from point to point. To a mathematician such an analysis would bring out the form of the curve representing the change. So long as this curve is independent of the yields of the plots as the result of the experimentation it would be well if such a curve which we shall call the *Normal Fertility Curve*[2] for the entire field is settled once for all before the other types of variation in the individual plots are studied. If particularly the experimentation could be continued in the same set of plots for different years the study of this curve which could be based upon the observations for years would considerably facilitate further studies regarding the differential effects of the several

[1] In the rest of the paper, *error* and *variation* are used in the identical sense.

[2] This term is borrowed from an article in *The Agricultural Journal of India*, Vol. XX.

varieties or treatments. The other type of the systematic variation is the one due to the treatments themselves. The variations that occur from plot to plot as the effect of different manures in testing different varieties or as the effect of the same manure in producing varietal differences, have thus to be separated from the type of the systematic changes that are due to the fertility gradient. The third class of errors is what may be called *Random Errors* which arise from the inaccuracies due to measurements—either through instruments or the personal eccentricities of the experimentalist. This class of errors is inevitable and however careful one may be, unknown causes operate in producing statistical inaccuracies in the measurements of the successive plots. To a statistician, this type of errors is attributable to 'chance' and the problem is how far chance operates in producing differences between the true value and the observed value of any quantity under measurement. It is here that the statistical theory is showing fresh phases and new developments, and as we shall see later clearer ideas of the statistical limitations of the particular experimentations are necessary to judge of the difference between any two values that may arise by chance. Thus the yield of a plot under a particular variety is a concourse of these three different types of errors affecting the intrinsic yield due to the variety alone. Statistical procedure in an examination of the differential effects thus resolves into a study of these different types of errors and of the significance that has to be attached to the differences between them.

(*c*) **The randomness of the plot arrangement**

This analysis of variations leads to the question how best we could arrange the plots to get at the greatest benefit out of the experimentation. The accidental errors being supposed constant, the first type of the systematic variation must be brought to a minimum which as already explained is due entirely to the differences between plot and plot owing to the fertility gradient of the field and not owing to any external agencies. Thus any arrangement must be conceived with a view to reduce this class of errors which is entirely due to circumstances not within our control. It is obvious that one method of securing this aim is to arrange the plots *at random* with respect to the several varieties or treatments and their replications so that the differences in fertility between plot and plot may not contribute to the differences between the treatments or varieties. In the case of the small sized plots the best way of securing the randomness is to get cards bearing the names of varieties and shuffle them freely so that the replications may not occur in any preconceived manner. Thus if we are testing 5 varieties in 20 plots, so that each variety recurs four times, we may have twenty cards with the names of varieties written down and shuffle the twenty cards and produce them in any random order, such as :

A D E B B A C D A B A C B E D C D C E E

If it should be possible to conduct experiments for several years in the same set of plots this arrangement could be altered continuously year to year to produce the desired effect. Mathematically this sort of chance randomness may be produced in a number of ways, we could either shuffle cards or take the last digits of a set of values from the logarithmic tables or produce a set of figures where the systematic variations have no place. But[1] it must be recognized that any random arrangement with respect to a number of varieties

[1] *The Principles and Practice of Yield trials* by Engledow and Yule, pp. 67-68.

is not always convenient from the point of view of cultivation and may even lead to mistakes which cannot be easily repaired. The simplicity of the arrangement combined with the randomness are features that must be borne in mind in scheming the plot-trials. But when dealing with plots of large sizes covering fairly huge blocks, even this random arrangement cannot reduce the effects of the fertility gradient and the only course open is to recognise the differences between the blocks but to secure randomness among the varieties in the individual blocks themselves. We shall be facing then two types of the fertility gradient, that due to the differences between the blocks and the one due to the plots themselves in the same block. Later on we shall be discussing methods of analysing the variances due to the several factors that cause the variation; but anyhow our aim must be to reduce these variances due to the soil to make the plot trials really effective and productive of the real significance. The following is an arrangement by which the uniformity between varieties could be secured by this process, where if ranks be given to them in the several blocks the sum total of the ranks of each variety is the same. Thus with five varieties and four blocks an arrangement like the following secures to each variety positions in the blocks whose sum is the same, and secures an even balancing to the several varieties in the blocks.

1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
A B C D E	E D C B A	A B C D E	E D C B A

where A, B, C, D, E represent different varieties and the sum of the ranks of each variety equals 12. By this method we have to a certain extent eliminated the differences between plots in the same block but still the differences between the blocks themselves remain. As each block has its own fertility gradient and as the significance of varietal differences cannot but recognise the differences between the blocks, still the block-variations can be reduced by an extended method of the above scheme of 'balancing.' Thus with five varieties, instead of having four blocks if we could have *five* so that each block has all the five varieties grown on it, then not only could we 'balance' the positions of the varieties but also we could 'equalise' their positions as in the following scheme:

1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
A B C D E	B A E C D	D E A B C	E C D A B	C D B E A

A, for example, occupies all the five positions in the several blocks so also the other varieties; so that by such an arrangement of 'balancing' and 'equalising' the plots and the blocks display more or less the same fertility differences. Thus whatever arrangement could be conceived of to suit the facilities of cultivation and environment, the conditions necessary to meet the soil inequality between plot and plot in any experimentation cannot be overlooked.

iii. Mathematical treatment of the different types of variations

(*a*) Systematic Variation

Whether we test different varieties or whether we test different treatments with respect to the same variety the systematic variation from plot to plot does require a mathematical understanding and study. If indeed we could have a fairly appreciable number of plots under trial, covering a fairly large area, then the study of the systematic variation reduces into an examination of the best curve that could fit in with the given data. But where there are only a few plots under examination the curve that could pass through a few values only cannot be relied upon to indicate the systematic changes. Anyhow the idea of passing the best curve through a given set of values cannot be lightly brushed aside, as its applications in an examination of a large series of trials are indeed of very high statistical importance. The method of curve-fitting is usually known as 'Graduation' or 'Smoothing,' and it consists of two distinct aspects, *one* securing a suitable *smooth* curve which would remove all roughness in the data and thus secure an apparent perfection to the given set of observations influenced by the other types of errors as well; and the *other* being that not only should we obtain any *smooth* curve but *the one* which would produce the greatest agreement to the given data. The curve that would satisfy both these conditions is the one we seek for in graduating a given set of values. Thus the process of Graduation secures the *most probable* values of a series of observations influenced by systematic changes (as we observe in the plot-fertility due to the heterogeneity of the soil). Now with regard to the mathematical methods of graduation, they are indeed too full, and the particular method to be adopted depends upon the nature of the enquiry on hand. An actuary, for example, uses what is known as 'the summation formula'[1] and his literature is full of it. The summation formula is based upon this concept—that of a given set of figures each figure influences the neighbouring ones accidentally as it frequently happens in the age distribution of a population (where, e.g., a man aged 39 gives his age to be 40). Here the best method of Graduation must be the one which combines the figures so that when differences above a certain order are neglected the original function is reproduced. It is obvious that this method is not suited to a series of plot-trials, for any single observation cannot be supposed to affect any other observed result, and each figure is taken to be the most reliable. The following is a useful method[2] which we have employed in graduating 25 plot-figures (Vide Table I). Assuming a curve of the second degree $y = a + bx + cx^2$ to pass through the twenty-five values we seek the best values of a, b and c by assuming that the moments[4] of 0, 1, 2 order of the given set and the graduated set are equal. Such an assumption is tantamount to the geometrical significance that with respect to the given curve[3] and the assumed curve (in this case the second degree curve) their areas, their centres of gravity and their moments of inertia are identical. If we could think of a higher degree curve involving greater number of parameters, we have to take higher moments and compute as many moments as there are parameters. But it is not advisable to increase the degree of the polynomial, though theoretically it is possible—for with $(n+1)$ constants we can fit in a polynomial of the nth degree—, because to that extent the accuracy of the law of variation is sacrificed. Hence we could conveniently stop at the second or the third degree. The first aspect of graduation is now satisfied in virtue of which we have secured a smooth curve of a polynomial.[5] But to produce greater agreement we can make another assumption that a very small quantity ϵ can be found such that the difference between the observed and this intermediary value $f(x) = (a + bx + cx^2)$ is of the form $\epsilon\, f^1(x)$[6] following

[1] C. F. Whittaker and Robinson, *Calculus of Observations*, pp. 288-291.

[2] The method is explained in full in *Calculus of Observations* (Vide Supra) pp. 303-315.

[3] The Area of the curve is the one bounded by the curve, the X axis, and the extreme ordinates.

[4] Moment of the nth order $= \Sigma x^n y \,/\, \Sigma y$.

[5] The form of the curve may suggest other Analytical types which are also necessarily smooth.

[6] $f^1(x)$ is the first differential coefficient of $f(x)$.

the expansion of $f(x+\epsilon)$ by the Taylor's Theorem. It only remains to choose ϵ and we can take a series of small values for it, ·001, ·01 and so on and test which value of ϵ would give the best agreement to the observed set of values. Table I gives the graduated values of the observed yields, and their differences, and the column 6 expresses the ratio of the difference (due to graduation) to the graduated result. Graph I is of the equation $y=400-\cdot 16x^2$ which is the graduated curve of the yields of the twenty-five plots, ϵ being taken to be zero. We shall, according to our theory, then take the differences between the graduated and the ungraduated due to the other types of variation.

GRAPH NO. 1

(Vide Page 197)

Graduation of 25 plots under five different treatments. ($y=400-\cdot 16\ x^2$.)

(*b*) Random fluctuations

Now we shall consider the other type of errors referred to, that due to the random fluctuations and statistical measurements. The Theory of what is known as 'random sampling' is the basis for judging how far the difference between any observed value and the true value could be ascribed to chance. A few fundamental considerations concerning the theory may not be out of place in the discussion of its features. Our object in any statistical enquiry is a study of the characters of a population and the evolution of a mathematical expression involving constants or parameters whose correct values have to be computed from the given data. In cases of the known population, that is, in cases of the distributions whose mathematical expressions are definitely known, the parameters of the expressions have to be deduced from the *statistic* or the *statistics* as computed from the given data. An *Arithmetic average*, for example, is a statistic giving the best value as could be deduced from a given set of observations. What is known as *the Standard deviation* is another statistic describing the given set of observations Similarly other statistics could be computed which would give the best values of the parameters on whose reliable values our description of the population depends. But if the parameters of the population should be correctly deduced there is one fundamental assumption involved in it that the sample under consideration consists of a fairly large number of individuals or theoretically an infinite number. The statistics cannot give reliable values to the parameters unless they are computed from a set of data of an appreciably large number of individuals. But in agricultural experimentation, the number of experiments is obviously limited, and the general problem in dealing with this limited number of a sample of experiments is 'could we attach any significance to the statistics obtained from the limited sample?' The question which requires a statistical answer is this: Given a sample of a population, whether and to what extent could the parameters or the statistics based upon them tally with the parameters or the statistics of the whole population whose features are supposed to be known? A kindred and a more useful question is this:

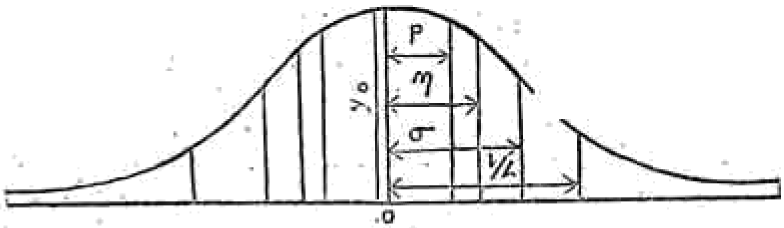
Given two samples of a population, could they be samples from the same population allowing for random errors which could have arisen by chance? The answers to these questions are obviously based on the *theory of probability* based upon the kindred theory, the *theory of errors,* and the correct solutions of these two problems form the key to the innumerable number of problems that arise in the sphere of random sampling such as we have in the limited number of experiments that obtain in the field trials. For example, if one field produces 7 tons and the other 8 tons both being homogeneous, what sort of significance could we attach to this difference, after all would this difference of one ton make an appreciable difference to jump to the conclusion that the second gives a better yield than the first? Or again, if a series of plots are subjected to several varieties whose differences in yield have to be studied, could we attach any significance to the observed differences which might be small or large? Such a comparison obviously depends upon the size of the sample—greater the number of plots subjected to the same variety greater the reliance could we place upon any statistic based upon them. Thus we are led to this consideration:—Given a series of samples each of size n, how could we find out whether they are homogeneous or identical samples or whether any significance could be attached to the differences between them? If the differences between the samples arise by chance only, then the differences can be ignored and the samples can be taken to be identical, but if they should be due to factors other than chance the significance of their differences has to be deduced.

iv. Concept of Probability and Significance

(*a*) The Normal Law and its Implications

If any error should arise by chance only or by circumstances purely accidental, then the law of distribution of such errors, if one such could be theoretically deduced or even assumed, would give us an indication how far and to what extent the amount of an error could be attributed to chance or randomness. Now we shall explain the basis of what is known as 'the *Normal Law of Errors*' which is the classic curve which 'Statistics' has adopted in most of the important enquiries where errors are in question, or where deviations from the best value of a sample have to be judged. First and foremost in the errors that arise by chance only, we include the types of errors such as the inaccuracy of an instrument or the personal equation of the experimentalist, whose causes cannot be diagnosed and whose aggregate sum cannot exceed a reasonably small quantity. But on the other hand, if we are able to discover the conditions and the laws of the actions of those errors however small they may be, the errors are no longer random but transform into a type of the systematic change referred to already. The Normal Law of Errors is thus based on three simple hypotheses, *firstly*, the errors are taken to consist of infinitesimally small errors whose sum total is by itself a very small quantity, *secondly* the chance of a positive error occurring is the same as the chance of a negative error, so that the sum total of the errors of a fairly large number of observations tends to zero, *thirdly* the chance of a small error is greater than the chance of a large error which is obviously a necessary condition in dealing with the errors due to chance. Now these hypotheses are very well illustrated in the resulting graph (Graph No. 2),

GRAPH No. 2



The equation is $y=\frac{1}{\sqrt{2}\,\sigma}e^{-\frac{x^2}{2\sigma^2}}$

y_0 = maximum ordinate (when $x=0$)
or height at the mode

A = Total Area = 1

N = the total number of observations

σ = Standard deviation $=\sqrt{\frac{1}{N}\Sigma(x^2 y)}$

p = Probable error

Area from $x=-p$ to $x=+p$ is half the total area A

giving P from the relation $\left\{\frac{2}{\sqrt{\pi}}\int_0^P e^{-t^2}\,dt=\frac{1}{2}\right\}=0\cdot47696$

where P defines p such that $p=(P\sqrt{2}).\sigma$

$\frac{1}{h}$ = 'Modulus' $=\sigma\times\sqrt{2}$

η = 'Mean absolute error' $=\sigma.\sqrt{\frac{2}{\pi}}$

whose ordinates express the number of times each individual error occurs, as it ranges from zero to infinity. The area bounded by the curve and the x axis thus represents the total number of observations, and the area bounded by ordinates between $x=a$ and $x=b$ represents the proportionate frequency. In the form given $y=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{x^2}{2\sigma^2}}$, the maximum ordinate corresponds to the zero error, and the total area is equal to 1. Thus the frequency of a total of x error ($\pm x$) is the corresponding area bounded by the ordinates for the abscessae $\pm x$. The concept of the probability of any deviation x is thus based on the assumption of the law of errors, and the probability of the occurrence of that deviation is the proportionate area. What are known as Probability integral tables have been constructed giving the probability of a particular error happening. The tables[1] are given in different forms and the best form[2] seems to be the one which would indicate the probability P of exceeding a deviation x positive or negative. From such a table the lesser the value of P, the greater is the significance to be attached to the particular value under consideration that it has arisen by influences other than chance alone.

[1] Vide *Tables for Statisticians and Biometricians*.
[2] T. L. Kelley, *Statistical Method*, pp. 373–385.

(*b*) The best Statistics to describe a population—Meaning of Probable Error

Now in dealing with a population consisting of an infinite number such as a population of random errors governed by the Normal Law, the two statistics that best describe its features and in fact summarise all the information necessary regarding the population are what are known as the *Arithmetic average* and the *Standard deviation*, the first two moments of distribution usually written as $\bar{x}$ and σ. These two statistics are fundamental, one expressing the average of the whole set of values and the other the square root of the average of the squares of the deviations from that average; and the formula for the Normal Law itself indicates that σ is a very important factor as expressing the precision of that average. Hence to test the deviation x it is appropriate to tabulate the values of $\frac{x}{\sigma}$ rather than x itself from which the significance of x could be deduced. If x follows the Normal Law, it is a simple result as could be seen from the form of the equation itself, that $\frac{x}{\sigma}$ also follows the Normal Law. The following is a scheme giving the probability P of exceeding $\pm\frac{x}{\sigma}$ for a deviation x as obtained from the Normal law.

$\frac{x}{\sigma}$	P
·5	·60
·6	·54
1·0	·32
1·5	·14
2·0	·06
2·5	·02
3·0	·003

What is called 'Probable Error' is that value of x which would give a probability of just ·5; and from the above scheme it is seen that where $\frac{x}{\sigma} = ·6$, P takes in a value near to ·5. The Probable Error by itself is not of much value excepting as a mathematical expression roughly equal to 2/3 the standard deviation and where the distribution is *not* normal it is certainly meaningless to attach any significance to it. Now the question is *where* and at what value of P we have to fix up the significance to any observed value. Though no hard and fast rule can be laid, it is convenient to take the point where P = ·05 as the border which corresponds to a value 1·96 for $\frac{x}{\sigma}$; and thus all deviations exceeding 1·96 σ can be regarded significant and those less, insignificant as due to chance.

(*c*) The best statistics to describe a sample

As referred to already, the Normal Law and deductions of significance form it are absolutely of no purpose in dealing with a sample of limited size

n as we have in an agricultural experimentation. The problem to be considered is what are the best statistics to describe a sample of n figures, which would help in judging whether the given sample is a sample of the known population or whether two samples are identical samples. The Arithmetic average and the Standard Deviation are the two statistics, as referred to already, that best describe a population normally distributed, but their best *estimates* as deduced from a sample of such a population could be computed thus:—For the Arithmetic average, the usual procedure of taking $\frac{1}{n}\times$ the sum of the figures of the sample could be taken as the best estimate; while for the standard deviation s, the square root of $\frac{1}{n-1}\times$ the sum of the squares of the deviations (as analogous to $\frac{1}{n}\times$ the sum of the squares of the deviations, for the computation of σ) could be taken to be the most reliable value. (n-1) is substituted for n in the computation of s, for the simple reason that the true Arithmetic average of the population differs from the Arithmetic average of the sample, and that (n-1) in the denominator increases the accuracy of s. The first two moments then with the slight difference in the computation noted already could be taken for all purposes as the two statistics describing the sample, though in cases where the distribution diverges from the normal in an appreciable degree we have to take higher moments and compute other statistics to describe the sample completely. In dealing with the errors that have arisen by chance such as the differences between two samples, their law of distribution may be taken to be *normal*, and the character of the sample may be based upon the first two moments only without loss of precision.

(*d*) Consistent and efficient statistics

Now then to judge of a sample, the law of distribution of the original population is obviously necessary, but when dealing with a sample of errors we have assumed that the original distribution is normal. Now having fixed the characters of the sample by means of its statistics, we have to examine whether and if so how two samples differ with respect to these characters. Statistically we have to examine the distributions of these statistics as obtained from several samples. Taking the Arithmetic average for example, we have to study the law of distribution of the Arithmetic averages obtained from a number of samples each of size n. Such a distribution would indicate the significance to be attached to the difference between the average of one sample and the average of another. Just as we inferred significance from the Normal Law, this new distribution of the average would signify, so far as this particular character is concerned, how far the samples are identical or different. Similarly, we can study the law of the distribution of the Standard Deviation of the sample and deduce the significance. The principle could be extended to any other statistic or to any higher moment whose law of distribution as obtained from several samples all of the same size n could be made the subject of study. Thus theoretically, if two samples are identical they must show *stable* values with regard to *any* moment of their distribution. But two fundamental considerations arise in fixing up a particular statistic for judging the significance of samples. The first is that the standard deviation of the distribution of the statistic must decrease as the size in the sample is increased. This is obviously a necessary condition, otherwise the value of the statistic cannot reach its true value as obtained from the whole population. Such statistics whose variance[1] decreases as n is increased are called *consistent* statistics. The second condition is that the statistic besides being *consistent* must be such that the amount of variance between the several samples must be as low as possible. In other words, of the several statistics that are consistent and possible whose laws of distribution could be inferred, the one whose variance is the least should be preferred for deducing the significance of the sample. Not only should *consistency* be reached but also *efficiency*, if any statistic should be useful to deduce the probability of a particular value of the statistic or its deviation from the true value. As an example, if n is appreciably large it is a well-known result that the average of a normal sample is distributed normally with the standard deviation $\frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation and n the number in the sample. Again, the law of distribution of the standard deviation itself of the sample of appreciable n is also known to be *normal* but with the standard deviation $\frac{\sigma}{\sqrt{2n}}$. Hence for the purpose of efficiency alone the S. D. of the sample is a better statistic than the A. M. We shall examine the distributions of other statistics as deduced from samples where n is very small.

v. Distribution of s, $\frac{x}{s}$, χ^2

(a) General

Just as $\frac{x}{\sigma}$ can be subjected to study in case where n is large, the distribution of $\frac{x}{s}$, where x obtained from a sample has a standard deviation s which differs from σ, could be studied for deducing the significance of x. Thus from a sample of n if $\bar{x}$ be its mean and s its standard deviation, the significance of the sample has to be judged from the value of $\frac{\bar{x}}{s/\sqrt{n}}$ with the help of tables of $\frac{x}{s}$. In general, if any statistical coefficient has its standard deviation known as obtained from the sample, the ratio of its value to the standard deviation could give the significance of its divergence from the assumed true value from the law of distribution of $\frac{x}{s}$. The principle is usefully extended to the case when two samples of size n_1 are compared whose means are $\bar{x}_1$ and $\bar{x}_2$ and whose standard deviations are respectively s_1 and s_2. Here if the differences between the samples could have arisen by chance only the mean of the differences of the corresponding figures in the samples can be taken to be zero, and the significance of $\bar{x}_1 - \bar{x}_2$ when compared with its standard deviation $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_1}}$ * could be judged, and inferences drawn whether the samples are identical or not. Thus the general theorem on which the significance of a sample or differences between two samples could be based is this: If x be the deviation from the true value, whose standard deviation is s,

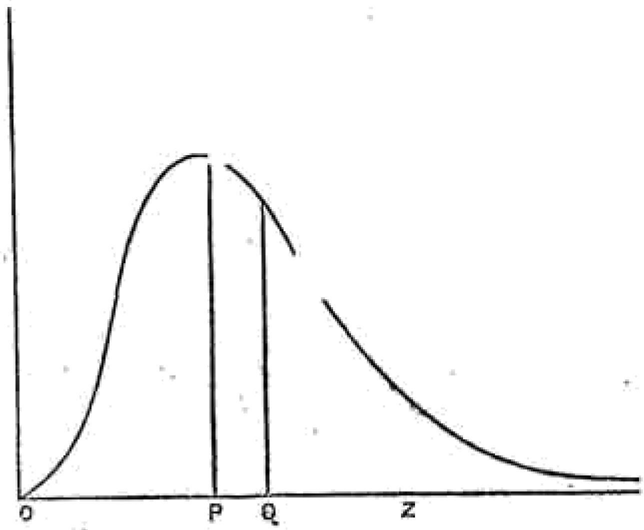
[1] 'Variance' used in the sense 'square of the standard deviation'.

* On the supposition that there is no correlation between the samples.

the law of distribution of $\frac{x}{s}$ would provide the necessary probability integral tables to judge of the significance of this deviation x. As in the case of the Normal Law dealt with already, the probability p of exceeding the deviation $\frac{x}{s}$ could be tabled, and lesser the value of p the greater the significance of x. But it is easily seen from our discussion that while $\frac{x}{\sigma}$ has a normal distribution when n is large, $\frac{x}{s}$ could be distributed only in a different way when n is small as s itself is subject to variation from sample to sample. We have, therefore first to study the frequency of s for samples of size n, and then deduce the frequency of $\frac{x}{s}$, assuming that the original distribution of the population is normal, which, as we have seen, is quite a legitimate assumption in dealing with the errors that arise by accident. The Mathematical deductions of these distributions are explained in Appendix A, and extracts from the Probability Integral Tables for the distribution of $\frac{x}{s}$ for different sizes are given in Table No. 9. The nature of these distributions for the size 10 are shown in Graphs Nos. 3 and 5 which would show how they differ funda-

GRAPH NO. 3

(Vide Appendix A 1)



$$y = y_0 \, s^{(n-2)} \cdot e^{-\frac{1}{2} n \frac{s^2}{\sigma^2}}$$

Now when n = number of samples = 10

let $\frac{s}{\sigma} = z$

then $y = k \; \sigma^8 z^8 e^{-5z^2}$

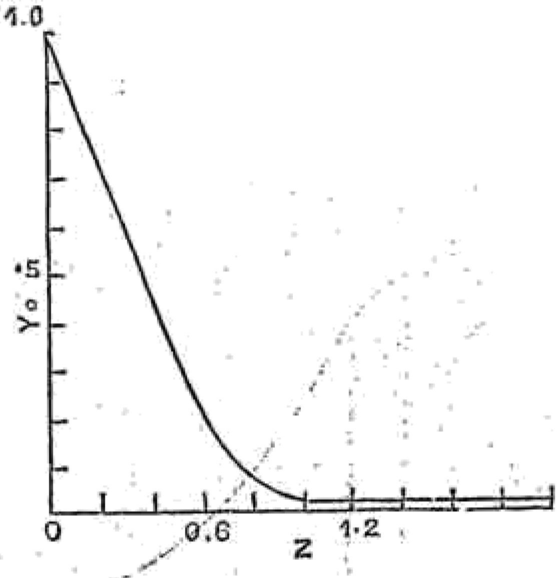
mentally from the Normal Law. It can be further seen that as the number in the sample increases the deviation decreases, and as the deviation increases the Probability decreases, and if, therefore, proper significance has to be attached to any deviation the lesser values of P are reached the greater the number in the sample.

Now for purposes of referring to the Tables, we need n, and x/s from which P is deduced. But new statistical concepts have been introduced with regard to n. It is easily seen that with a sample of n figures, whose aggregate is to remain constant, $(n-1)$ of these figures could be altered arbitrarily but the nth is automatically fixed up if the sum total should remain the same. Thus a sample of n figures is supposed to have $(n-1)$ degrees of freedom on the supposition that the sum total or their mean should be fixed. Hence instead of referring to the tables for n, the size of the sample, we could enter for $(n-1)$, the number of degrees of freedom, thus ensuring greater accuracy for the corresponding value of P. The same principle could be employed for securing a suitable value for n^1* in testing the significance of any statistical coefficient. If, for example, we are fitting a curve to the given data of size n, say a straight line which involves two constants, to this extent the number of

* n^1 = the number of degrees of freedom.

GRAPH No. 5

(Vide Appendix A III)



Now let $z = \frac{x}{s}$

then $y = \frac{1}{\pi} \left\{ \frac{(n-2)(n-4) \ldots\ldots 4 \cdot 2}{(n-3)(n-5) \ldots\ldots 3 \cdot 1} \right\} (1+z^2)^{-\frac{n}{2}}$

when n is even

$= \frac{1}{\pi} \left\{ \frac{(n-2)(n-4) \ldots\ldots 3 \cdot 1}{(n-3)(n-5) \ldots\ldots 4 \cdot 2} \right\} (1+z^2)^{-\frac{n}{2}}$

when n is odd

$= y_0 (1+z^2)^{-\frac{n}{2}}$.

Here in the graph,

$n = 10$, z is taken to be positive

degrees of freedom is reduced by 2, as any two points in the straight line fix its position, so that the number of degrees of freedom is $(n-2)$. If again we are fitting a curve of the second degree $n^1 = n-3$ and so on. Again, what is known as the correlation co-efficient of a given number of n pairs is obtained by fitting a straight line of closest fit to pass through the n means and n^1 in such a case to test a particular value of the correlation co-efficient $=(n-2)$.

(*b*) Significance of the Means: Illustrations from Agricultural experimentation

We are thus in a position to attach significance to a sample of experiments or to different samples of experiments. In Agricultural experimentation, for example, we have to compare the yields of two varieties grown in two sets of plots and find out whether there is any significant difference between them. Here the problem is to find P for the difference of their means in terms of its standard deviation.

Examples are given in the tables illustrating the application of the theory; Table 3, for example, compares two varieties grown in 10 plots so that $n^1=9$. The mean of the difference between the corresponding figures is seen to be 67·5 whose standard deviation from the column of the differences alone is 64·57 from which P is deduced to be 0·01 shewing that there is significant difference between the varieties. It is easily seen that different conclusions could be reached if we had taken $n^1=10$. Again, instead of computing the standard deviation from the column of differences we could compute it from the standard deviations of the samples themselves,[1] but such a procedure is possible only if there is no correlation between plot and plot, i.e. if there is no systematic variation between the plots themselves. However, we have seen that there are systematic changes from plot to plot and therefore unless their correlation is known we are not warranted to take that the square of the standard error of the differences is equal to the sum of the squares of the standard errors of the samples. Other illustrations are given in Tables 2, 4, 5, and 6 illustrating the significance between varieties, with regard to their productivity on light and heavy soil, to bunch planting, cropping values of seeds, and manuring. The conclusions reached regarding the superiority of one variety to another are all based upon the mean of the differences being taken to be the deviation from the zero value assuming the latter to be the true value of the differences. As another illustration[2] of the theory governing $\frac{x}{s}$, we shall examine the significance of the correlation co-efficient obtained between autumn rainfall and wheat crop for twenty years. Here, $r = -\cdot629$ and its $\text{S.D.} = \frac{\sqrt{1-r^2}}{\sqrt{n-2}} = 1\cdot738$ and $\frac{r}{S.D.} = -3\cdot433$ giving $P < \cdot01$, shewing that the observed value of r is definitely significant. Here it is necessary to point out that it is always risky to rely upon the significance of the correlation co-efficient, or as a matter of that any regression curve, as computed from a limited number of data, for the simple reason that there is no method for knowing anything of the true values of the regression constants and that therefore to adopt statistical methods to attach any significance to the observed co-efficient is only putting the cart before the horse. Exhaustive enquiries alone could secure stable values for the co-efficients in such cases.

[1] $s^2 = s_1^2 + s_2^2$ could be used.

[2] This illustration is given in Page 158 '*Statistical Methods for Research Workers*' by R. A. Fisher.

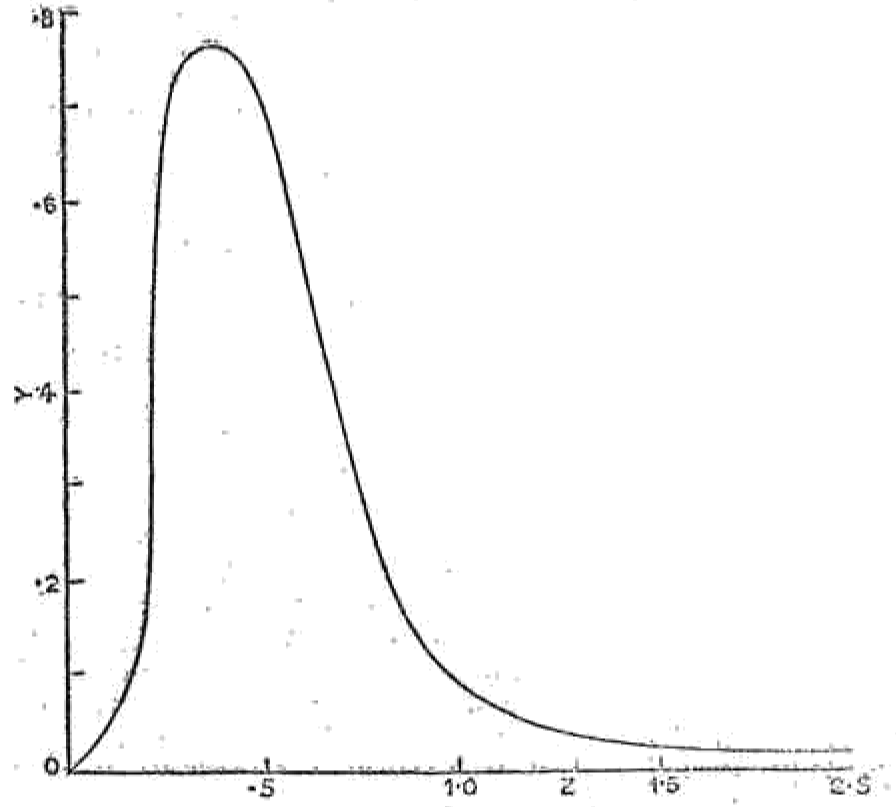
(c) Distribution of χ^2

(1) General

Till now we have been answering the only question how far we could identify two samples of experiments. The other question—the prior question put—, whether a sample is a sample of a known distribution still remains unanswered. The same method of the significance of $\bar{x}$ could be employed, but in all cases there is no guarantee that $\bar{x}$ can be very much relied upon, nor can we be sure that each difference between one sample and another measures to a reliable extent the error from the true value as due to chance. Hence we have to think of other statistics whose distributions could give the significance sought for. The problem in general terms could be stated thus: If $m_1, m_2, m_3 \ldots$ be a set of known values and if $m_1+x_1, m_2+x_2, m_3+x_3, \ldots$ be the corresponding set of observed values, what would be the bets statistic to measure the significance of these errors $x_1, x_2, \ldots$? The statistic must be the one based upon these errors themselves, and so long as we are sure of the known or theoretical set the distribution of $\Sigma\left(\frac{x}{m}\right)^2=\chi^2$ requires a study, where we take the sum of the squares of the proportionate deviations of the individual errors. χ^2 is obviously analogous to s^2, and if n be the number in the sample $\frac{\chi^2}{n}=\frac{s^2}{\sigma^2}$. As the distribution of s^2 is known (vide Graph No. 4), the graph of $\frac{\chi^2}{n}$ or χ^2 itself is deduced. (For the distribution of χ^2 vide

Graph No. 4

(Vide Appendix A II)



$$y=y_0\, s^{\,n-3}\, e^{-\frac{1}{2} n \frac{s^2}{\sigma^2}}$$

Now when n = number of samples = 10

$$\text{let } \frac{s^2}{\sigma^2} = z = \frac{x^2}{10}$$

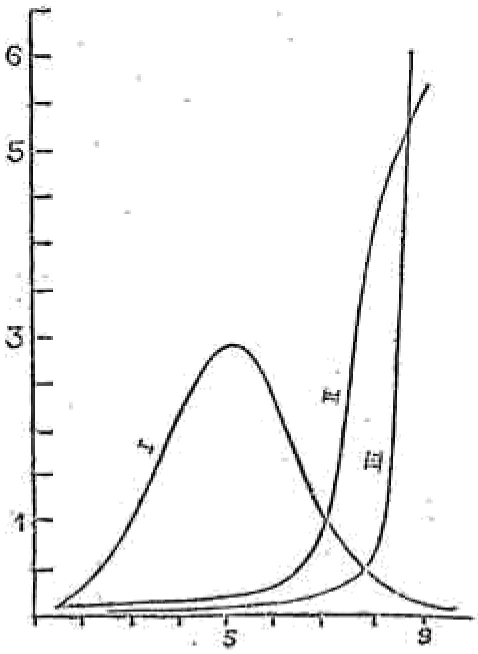
$$y = k \div 7 \; z^{\frac{7}{2}} e^{-5z}.$$

Appendix A). It is easily seen that as χ^2 changes from zero to infinity, P changes from 1 to 0, and greater the value of P greater the correspondence between the theoretical and the observed set of observations. The limit of significance can be placed at P = ·05 as in the case of the other distributions considered already. Extracts are given in Table No. 10 giving the Probability of χ^2 for n^1 the number of degrees of freedom. As an illustration, the numbers with black and red eyes are observed in 33 families of Gammarus (Huxley's Data) and it is for us to test whether the Mendelian ratio 3 : 1 due to linkage is followed in this case. It is found from the 33 frequencies, $\chi^2 = 35{\cdot}620$. From $n = 32$ it is seen that P is not small, and hence all families agree fairly closely in this particular trait.

(2) Binomial and Poisson distributions

The theoretical set referred to from which the deviations of the observed set were measured, might be one which has been decided from the past experience from similar enquiries or one deduced purely from theoretical and statistical considerations. We shall deal with a distribution that is of common use in the Agricultural experimentation, besides the one dealt with already the Normal Law of errors. If chance should be the only governing factor, it is a simple statistical result that the terms of the expansion of $(p+q)^n$ *

Graph No. 6



Binomial Curves
I $(p+q)^8$ and $p = 0{\cdot}5$
II $(p+q)^8$ and $p = 0{\cdot}08$
III $(p+q)^8$ and $p = 0{\cdot}01$

* If p is very small and n increased indefinitely so that np is finite, the limit of the binomial is the Poisson's exponential series. (Vide Biometrica, Vol. 10, pp. 38–).

express the probabilities of happening of 0, 1, 2 . . . . n events of the sample of n experiments. The Arithmetic average of all the terms is np and the standard deviation npq. The Graph No. 6 indicates how with the changes in the value of p, the form changes continuously from one of symmetry to one of increasing asymmetry. The relation between np and npq the first two moments decides the nature of the distribution. Tables have been constructed giving the probabilities for different sizes of the sample as deduced from the binomial distribution and the observed set of figures can thus be made to correspond to the theoretical set. χ^2 could thus be computed, and from the table of χ^2 the significance of a particular value of χ^2 is deduced. Here are two illustrations exemplifying the application of χ^2. (Vide tables 7 and 8) In the first, we assume a binomial law and χ^2 is found to be 7·740. With $n^1=10$, P= ·654 showing that the theoretical and the observed closely agree. In the other there is perfect agreement between the observed and the Poisson exponential.

vi. Concept of Variance—Correlation Co-efficient *

(*a*) General

The ideas of significance which we have dealt with so far lead us to a very important topic connected with the Agricultural experimentation. As we have referred to already the different types of variations and the different factors that produce the variations in the eventual yields of the plots under experimentation have to be analysed and subjected to statistical scrutiny. Though a single sample or different samples of experiments have been tested with regard to homogeneity, the problem still remains how to measure the interaction or correlation, if any, between the members of the same sample. If the whole population could be split up into n samples, each sample containing k observations or experiments, then if all the n samples are homogeneous, the statistic such as the Arithmetic average would shew consistent values in all the samples, and the standard deviation of the statistic will be indeed a very small quantity very close to zero, and thus the identity of the samples is established and there is no correlation between the individuals of the sample. But what ordinarily happens is that the standard deviation of the sample is significantly different from the standard deviation of the average of the sample which is nowhere near zero, in which case the two standard deviations have to be separated and their mutual significance has to be studied. Thus when we deal with the heights of n pairs of father and son of the same family, while the father and son of any family would give a standard deviation of its own, the averages of the father and son as obtained from n samples would give a different standard deviation. The two causes of deviation being different, the values of the two standard deviations and their significant difference based upon their individual degrees of freedom decide whether there is any correlation between any pair of observations. The result could be put in this form: The total variance † of all the n samples containing kn observations, could be split up into two different variances, one due to the deviation of the individuals of each sample from its average, and the other due to the deviations, between the several samples, of the average of each sample. If the latter variance is

* The ideas in this section are taken from several papers dealing with the Intra class correlation. (Vide R. A. Fisher, Pages 176–232).

† 'Variance' is used in the sense of 'the square of the standard deviation.'

significantly greater than the former the correlation, if there should be any, must be positive, but if less it is negative. The following simple equation explains how the total variance is split up:

$$\overset{kn}{\underset{1}{\lessgtr}} (x-\bar{x})^2 = \overset{kn}{\underset{1}{\lessgtr}} (x-\bar{x}_p)^2 + k \overset{n}{\underset{1}{\lessgtr}} (\bar{x}_p-\bar{x})^2$$

where x is any individual observation, $\bar{x}$ the Arithmetic average of the whole, $\bar{x}_p$ the average of a sample. This simple principle of analysing the total variance due to two causes only could be easily extended when the total could be split up into variances due to any number of causes. If, for example, the yield of a particular plot be due firstly to the plot itself, next to manuring and lastly to the variety grown, the variances of these causes could be separately computed, and the sum of these plus the variances due to the sub-classes should be equal to the total variance if there should be no other cause of variation. Again, in the study of rainfall from hour to hour, from day to day, and from month to month each particular item of hourly rainfall belongs to a particular day and to a particular month, the total variance of all the yearly items being thus equal to the monthly and the daily variances plus the variance due to the interaction of causes and deviations in each class. One more illustration from the Agricultural experimentation of plot-trials subjected to different varieties. In the plot-technique, we insisted upon a random arrangement of the varieties in each block, and thus several blocks were dealt with which are themselves subject to variations in fertility. Now to bring out the differences between the varieties markedly, we have to compute separately the variance due to the blocks, the variance due to the varieties and the one due to the random errors of measurement, and we have to examine statistically the significant difference between one variance and another. This analysis of variance is thus a key to the solution of problems where we have to examine the differences between the several causes producing an aggregate result.

(*b*) Significance of the difference between two different types of variance—Illustrations

Thus we are led to this statistical problem—what is the best statistic to signify the difference between the variances due respectively to n_1 and n_2 degrees of freedom? In the ideal case when n_1 and n_2 are infinite and when their distributions are normal, the chance that the two variances are equal is just ·5. The best statistic z that would bring out the difference of one variance and another is half the difference of the logarithms between the two which is found to be distributed normally with the standard deviation $\sqrt{\frac{1}{2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$* and thus the Normal Law Probability Integral Tables could be used to measure the significance of the difference between the two variances. Tables are constructed giving for any value of half the difference of the logarithms of the two variances the probability P corresponding to n_1 and n_2.

Here are two examples illustrating this principle of analysis of variance:

[1] (1) A plot of land is divided into 36 patches on which 12 varieties are grown and each patch is divided into three lines at which three different manures are tried. Thus the yield depends upon the particular patch, the

* This is strictly true only if n_1 and n_2 are appreciably large or when they are equal. In other cases the significance is obvious from the variances themselves.

[1] This example and the explanations are taken from Fisher (Vide Supra), page 203.

particular variety and the particular manure. The following is the analysis of the variances due to several causes:

The cause of variance	Number of Degrees of Freedom	Amount of variance
Between varieties ..	11	1·967
Between patches of the same variety ..	24	·727
Between manures ...	2	·175
Differential response of varieties ..	22	·010
Differential response in patches with same variety ..	48	·168
Total ..	107	·670

From the above analysis the following inferences are easily drawn: Firstly while 24 degrees of freedom for the patches give a variance ·727, 48 between the patches growing the same variety give only ·168 proving the significant difference between the plots when compared with the effects of the varieties on the plots. Again, among the varieties themselves for the 22 degrees of freedom the variance is only ·01, which is enough to indicate that there is not much difference between one variety and another. The difference between the manures gives only ·175 which is not very high when compared with the differential response of varieties, shewing that the manures do not demonstrate any significant difference.

[1] (2) Five varieties are grown each in 6 different plots. The total variance is split up into variances due to several causes. Thus we have this scheme:

The Cause of variance	No. of Degrees of Freedom	Amount of variance
Between varieties ..	4	41·36
Between plots growing the same variety ..	25	1715·72
Between plots of the same number serially arranged for the same variety ..	5	424·30
Between varieties taken two and two grown in plots of the same number ..	50	3228·40

[1] This example is taken from '*The Principles and Practice of Yield Trials*' by Engledow and Yule, page 18.

From the above analysis it is clear that while varieties exercise a variance of 41·36 the plots themselves produce an abnormally large variance, showing the decided effect of the plots when compared with the varieties. A further examination shows that in the experiment a fairly high randomness has been secured in the replication of varieties as the variance between the serial numbers of plots is comparatively low. Lastly there is not much significant difference between the adjoining plots and between plots growing the same variety. In every one of these cases the actual probability could be deduced from the tables, and the significance deduced.

APPENDIX A

I

Distribution of s*

Let a sample consist of n data $x_1, x_2, x_3\ x_n$; let the original distribution follow the normal law; and let the averages of the sample and of the original population be respectively $\bar{x}$ and m. Then obviously the probability of the n values occurring=Constant $\times e^{-\frac{1}{2}\frac{\sum(x_s-m)^2}{\sigma^2}} \times \delta x_1 . \delta x_2 \ldots . \delta x_n.$ (1)

Now if $s^2 = \frac{1}{n}\sum (x_s - \bar{x})^2$, after transformation (1) reduces to=Constant $\times e^{-\frac{1}{2}\left(\frac{n s^2}{\sigma^2} + \frac{n(\bar{x}-m)^2}{\sigma^2}\right)} \times s^{n-2}\ \delta \bar{x}\ \delta s.$

From this the law of distribution of s is seen to be

$$y = \text{Constant} \times s^{n-2} e^{-\frac{1}{2}\frac{ns^2}{\sigma^2}} \text{ (Vide Graph 3 for } n=10). \qquad (2)$$

II

Distribution of s^2 and χ^2

From (2) the frequency of s^2 is easily deduced. Since $Q(s^2)\ d(s^2) = Q(s^2) \times 2s\ ds$, it is seen that the law of s^2 is obtained by dividing the frequency of s by $2s$. Thus the frequency of s^2=Constant $\times (s^2)^{\frac{n-3}{2}} e^{-\frac{1}{2}\frac{s^2}{\sigma^2}}$ (3)

The law of χ^2 is again inferred from (3), as $\frac{\chi^2}{n} = \frac{s^2}{\sigma^2}$. This gives that the frequency of χ^2=Constant $\times (\chi^2)^{\frac{n-3}{2}} \times e^{-\frac{1}{2}\frac{\chi^2}{n}}$. (4)

(Vide graph 4 for $n = 10$).

* Vide Biometrica, Vol. 10, pp. 522-523

III

Distribution of $\frac{\bar{x}}{s}$

(Biometrica, Vol. 6, Pages 1–..)

Now it is a well-known result that $\bar{x}$ obtained from samples of n is distributed according to the normal law with a standard deviation $\frac{\sigma}{\sqrt{n}}$ so that the frequency of $\bar{x}$

$$= \text{Constant} \times e^{-\frac{n\,\bar{x}^2}{2\,\sigma^2}} \tag{1}$$

To deduce the frequency of $\frac{\bar{x}}{s} = z$, first we shall take s to be constant from sample to sample. The distribution of $\frac{\bar{x}}{s}$ is then :

$$= \text{Constant} \times s \times e^{\frac{-ns^2 z^2}{2\sigma^2}} \tag{2}$$

But as s varies the probability of s occurring

$$= \frac{\int_{s}^{s+ds} s^{(n-2)} e^{-\frac{ns^2}{2\,\sigma^2}}\, ds}{\int_{o}^{inf.} s^{(n-2)} e^{-\frac{ns^2}{2\,\sigma^2}}\, ds} \tag{3}$$

(from the distribution of s.)

Hence the frequency of $\frac{\bar{x}}{s}$ is obtained by multiplying (2) and (3) and summing for all values of s which reduces to

$$y = \frac{\text{Constant} \times \int_{o}^{inf.} s^{n-1} e^{\frac{-ns^2(1+z^2)}{2\sigma^2}}\, ds}{\int_{o}^{inf.} s^{(n-2)} e^{-\frac{ns^2}{2\sigma^2}}\, ds}$$

By a process of reduction, this integral reduces to

$$y = \tfrac{1}{2}\,\frac{(n\text{-}2)\,(n\text{-}4)\cdot\;\cdot\;\;5\cdot 3}{(n\text{-}3)\,(n\text{-}5)\cdot\;\cdot\;\;4\cdot 2}\,(1+z^2)^{-\frac{n}{2}} \quad \text{(if } n \text{ be odd)}$$

or

$$= \frac{1}{\pi}\,\frac{(n\text{-}2)\,(n\text{-}4)\cdot\;\cdot\;\;4\cdot 2}{(n\text{-}3)\,(n\text{-}5)\cdot\;\cdot\;\;3\cdot 1}\,(1+z^2)^{-\frac{n}{2}} \quad \text{(if } n \text{ be even)} \tag{4}$$

Tables are constructed giving the probability integrals of (4), by the substitution of $z = \tan\theta$ which gives for $(1+z^2)^{-\frac{n}{2}}\,dz = \cos^{n-2}\theta\, d\theta$, which could be easily integrated.

TABLE 1

Graduation of 25 plots under five different varieties

Variety	Plot	Observed Yields	Graduated Values	Difference of (3) and (4)	The Ratio of (5) to (4)
(1)	(2)	(3)	(4)	(5)	(6)
A	1	450·00	399·84	50·16	·13
	2	340·00	392·16	–52·16	·13
	3	430·00	372·96	57·04	·15
	4	370·00	342·14	27·86	·02
	5	380·00	300·00	80·00	·27
B	1	410·00	399·36	10·64	·03
	2	320·00	399·04	–79·04	–·19
	3	400·00	368·64	31·36	·09
	4	310 00	356·00	–46·00	–·13
	5	290·00	315·36	–25·36	–·08
C	1	370·00	398·56	–28·56	–·08
	2	370·00	387·04	–17·04	–·05
	3	360·00	364·00	– 4·00	–·01
	4	400·00	353·76	46·24	·13
	5	400·00	329·44	70·56	·21
D	1	430·00	398·44	31·56	·08
	2	410·00	384·00	26·00	·07
	3	360·00	380·64	–20·64	–·05
	4	330·00	343·16	–13·16	–·04
	5	330·00	322·56	7·44	·02
E	1	370·00	396·00	–26·00	–·07
	2	480·00	389·76	90·24	·23
	3	400·00	376·96	23·04	·06
	4	400·00	359·04	40·96	·11
	5	360·00	307·84	52·16	·17

To illustrate the significance of means

TABLE 2

(*Extracted from Biometrica, Vol. 6*)

Soft and Hard seeds on light and heavy soils

	1899		1900		1901	
	Light	Heavy	Light	Heavy	Light	Heavy
Soft seed ...	7·85	8·89	14·81	13·55	7·48	15·39
Hard seed ...	7·27	8·32	13·81	13·36	7·97	13·13
Difference ...	·58	·57	1·00	·19	-0·49	2·26

	Average
Soft seed ...	11·328
Hard seed ...	10·643
Difference ...	·685

$n^1 = 5$; S. D. $= \cdot 778$; $z = \frac{\bar{x}}{s/\sqrt{n}} = 1 \cdot 93$; P $= \cdot 18$ (nearly).

The difference is not significant, when compared with the soil.

TABLE 3

(*Extracted from Agricultural Journal of India, Vol. 20*)

To illustrate the difference between varieties

Plot Number	Yield of variety A	Yield of variety B	Difference
1	703	670	33
2	705	630	75
3	653	560	93
4	640	615	25
5	700	542	158
6	715	667	48
7	647	702	–55
8	848	750	98
9	918	758	160
10	870	830	40

(1) Average of Difference = 67·5

(2) S. D. of Difference = 64·57

$$z = \frac{\bar{x}}{s/\sqrt{n}} = 3 \cdot 00$$

$$n^1 = 9$$

$$P = \cdot 01 \text{ (nearly.)}$$

The difference is clearly significant.

TABLE 4

To illustrate the problem of Bunch Planting

Variety	Singles		Doubles		Difference	
	Grain	Straw	Grain	Straw	Grain	Straw
A. ...	300	500	305	520	5	20
B. ...	250	320	280	315	30	–5
C. ...	380	600	360	616	–20	16
D. ...	400	620	420	630	20	10

	Grain	Straw
Average Difference	8·75	10·25
S. D. ...	22·07	9·75
$z = \frac{\bar{x}}{s/\sqrt{n}} =$ ...	·67	1·83
$n^1 = 3$ ...	...	...
P = ...	·56 (nearly)	·17 (nearly)

Difference not significant in grain but comparatively significant in straw.

TABLE 5

To illustrate the cropping value of seeds

	Main		Thalus		Nursery		Difference between 2 & 3	
	1		2		3			
	Grain	Straw	Grain	Straw	Grain	Straw	Grain	Straw
Variety A. ...	400	606	410	620	405	630	5	−10
„ B. ...	425	635	420	625	415	600	5	25
„ C. ...	500	700	515	725	525	730	−10	−5
„ D. ...	480	650	490	680	500	620	−10	60
Average ...	451	648	459	662	462	645	−3	17
S. D. of the difference ...	...	...	...	...	...	...	8·67	32·28
$z = \frac{\bar{x}}{\frac{s}{\sqrt{n}}} =$ ...	...	...	...	...	...	...	·7	1·05

$n' = 3$.

Since P is large, there is no significant difference between (2) and (3).

TABLE 6

To illustrate the yields due to different treatments

Plot No.	Treatment A.		Treatment B.		Treatment C.		Treatment D.		Treatment E.		Treatment F.		Treatment G.	
	Gr.	St.	Gr.	St.	Gr.	St.	Gr.	St.	Gr.	St.	Gr.	St.	Gr.	St.
I.	40	100	45	120	48	120	50	115	52	125	55	110	47	94
II.	35	80	38	90	40	120	42	108	42	95	40	100	39	92
III.	38	90	40	102	42	95	39	98	42	100	45	120	38	90
IV.	48	120	50	110	48	120	47	115	50	110	52	115	54	120
V.	50	110	52	120	55	112	60	90	58	120	60	125	62	130
Average ...	42	100	45	109	47	110	48	105	49	110	50	114	48	105
	Gr.	St.												
Average of the average ...	47	108												
S. D.	1·02	1·74												

$n' = 6$

P is very small and there is no significant difference between one treatment and another taken all together, but a similar method could be employed to test the difference between any two treatments.

SIGNIFICANCE OF x^2

TABLE 7

(Extracted from Biometrica, Vol. 10)

Suicides of Women in Eight German States

Number of Suicides.	0	1	2	3	4	5	6	7	8	9	10 and over	Total
Observed frequencies ...	9	19	17	20	15	11	8	2	3	5	3	112
Theoretical or Binomial Frequencies.	12·6	18·4	18·8	16·4	13·2	9·9	7·2	5·1	3·5	2·4	4·5	112

$x^2 = 7 \cdot 740 \quad n' = 10 \quad P = \cdot 654$

This shows that the observed and theoretical closely agree.

TABLE 8

(Taken from Biometrica, Vol. 10)

Number of deaths	0	1	2	3	4 and over
Observed frequencies	109	65	22	3	1
Poisson's Exponential frequencies...	108·7	66·3	20·2	4·1	0·7

Here $x^2 = \cdot 441 \quad n' = 3 \quad P = \cdot 92$

This shows perfect agreement with the theory.

TABLE 9

(Extracted from R. A. Fisher, p. 137)

Probability for $\frac{x}{s}$

n'	P=·9	·8	·7	·6	·5	·4	·3	·2	·1	·05	·02	·01
3	·137	·277	·424	·584	·765	·978	1·250	1·638	2·353	3·182	4·541	5·841
4	·134	·271	·414	·569	·741	·941	1·190	1·533	2·132	2·776	3·747	4·604
5	·132	·267	·408	·559	·727	·920	1·156	1·476	2·015	2·571	3·365	4·032
6	·131	·265	·404	·553	·718	·906	1·134	1·440	1·943	2·447	3·143	3·707
7	·130	·263	·402	·549	·711	·896	1·119	1·415	1·895	2·365	2·998	3·499
8	·130	·262	·399	·546	·706	·889	1·108	1·397	1·860	2·306	2·896	3·355
9	·129	·261	·398	·543	·703	·883	1·100	1·383	1·833	2·262	2·821	3·250
10	·129	·260	·397	·542	·700	·879	1·093	1·372	1·812	2·228	2·764	3.169

n' = the No. of degrees of Freedom.
P = the Probability.

The figures in the body represent the deviation $\frac{x}{s}$

TABLE 10

(Extracted from R. A. Fisher, p. 98)

Probability for χ^2

n'	P=·99	P ·90	·70	·50	·30
3	·115	·584	1·424	2·366	3·665
6	·872	2·204	3·828	5·348	7·231
9	2·088	4·168	6·393	8·343	10·656
10	2·588	4·865	7·267	9·342	11·781
32	14·953	20·599	25·508	29·336	33·530

n' = the No. of degrees of Freedom.
P = the Probability.
The figures in the body represent χ^2.