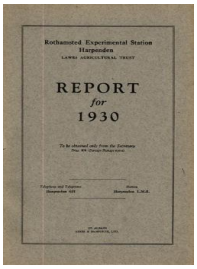
ROTHAMSTED RESEARCH

Report for 1930

Full Table of Content



Statistical Methods and Results

Rothamsted Research

Rothamsted Research (1931) *Statistical Methods and Results ;* Report For 1930, pp 77 - 81 **- DOI: https://doi.org/10.23637/ERADOC-1-63**

centration. There was no correlation between crop growth and phosphate concentration. It is suggested that barley can take up phosphoric acid directly from the solid particles of soil or fertiliser.

XVII. J. CALDWELL. "*A Note on the Dichotomous Branching of the Main Stem of the Tomato (Lycopersicum esculentum).*" Annals of Botany, 1930. Vol. XLIV, pp. 495-498.

Occasionally in the experimental material it was noticed that plants appeared having dichotomously branched stems. One of these is described in this note. It is shown that the arrangement of the leaves indicates that the bifurcation is of the main stem and not axillary in origin. The stelar tissue divides exactly into two—one half going to each of the limits of the fork.

STATISTICAL METHODS AND RESULTS

(Statistical Departments)

(*a*) General

XVIII. F. E. ALLAN. "*The General Form of the Orthogonal Polynomials for Simple Series, with Proofs of their Simple Properties.*" Proceedings of the Royal Society of Edinburgh, 1930. Vol. L, pp. 310-320.

In "Statistical Methods for Research Workers." R. A. Fisher has given a numerical method of polynomial fitting by means of orthogonal functions, developed from their terminal differences. It is shown here that the use of terminal differences may be made to supply direct and simple proofs of the algebraic properties of these polynomials, and a general formula for them.

XIX. F. E. ALLAN. "*A Percentile Table of the Relation between the True and the Observed Correlation Coefficient from a Sample of Four.*" Proceedings of the Cambridge Philosophical Society, 1930. Vol. XXVI, pp. 536-537.

In this paper a table is furnished, for samples of four, of the 95 per cent values of the transformed correlation, z, for different values of the correlation ζ in the population sampled. The table is based on the distribution of the correlation coefficient given by R. A. Fisher in 1915.

XX. R. A. FISHER. "*Moments and Product Moments of Sampling Distributions.*" Proceedings of the London Mathematical Society, 1929. Vol. XXX, Series 2, pp. 199-238.

Much previous work has been expended in studying the distributions of various symmetric functions of the sample values of a variate having a known distribution, and it has been recognised that the moment functions of such statistics must be expressible in terms of the moment functions of the distribution sampled.

Only a few such expressions had, however, been obtained with exactitude, and the great complexity of these gave little promise of a solution of the general problem. It is here shown that, when the Pearsonian moments are replaced by more suitable statistics,

https://doi.org/10.23637/ERADOC-1-63

the expressions are greatly simplified and may be obtained by a direct algebraic method. Further their general form may be derived in terms of combinatorial analysis by the use of two-way partitions, and certain pattern formulae which are the same alike for uni-variate and multi-variate problems. Rules are given, with a general demonstration, by which any particular term of any of these formulae may be obtained independently. The paper contains a table of the pattern formulae of most frequent occurrence, and a table of all the uni-variate formulae required up to the 10th degree, together with a few others of special importance. From these all the corresponding multi-variate formulae may readily be derived.

XXI. R. A. FISHER. "*The Sieve of Eratosthenes.*" The Mathematical Gazette, 1929. Vol. XIV, pp. 564-566.

The note suggests that the celebrated sieve of Eratosthenes has been misunderstood through lack of recognition of the fact that its author probably had in mind not a method of testing whether any particular number is a prime, but a labour-saving device for finding all the primes in a given range of natural numbers.

It is pointed out that a very simple diagrammatic method of doing this has in fact much the appearance of a wicker sieve, and it is suggested that the sieve connected with Eratosthenes' name was in fact a wall diagram of this sort.

XXII. R. A. FISHER. "*Inverse Probability.*" Proceedings of the Cambridge Philosophical Society, 1930. Vol. XXVI, pp. 528-535.

That the principle of inverse probability includes an arbitrary and unsatisfactory element has been recognised by many writers; but their criticism has failed to settle the controversy, since they have supplied no alternative account in mathematical terms of the process of learning by experience.

This paper briefly summarises the author's view that confusion has arisen through assuming that mathematical probability is the only measure of rational belief, and is applicable to all kinds of uncertain knowledge. It is suggested that from knowledge of a population we can express our incomplete knowledge of a sample in terms of probability, whereas knowing a sample we must express our incomplete knowledge of the population in terms of a different mathematical quantity, termed *likelihood*, which does not obey the laws of probability.

There are, however, certain cases in which statements in terms of probability can be made with respect to unknown populations. These are the typical statements of tests of significance, and the logical distinction between these and the statements of inverse probability, to which they bear a superficial resemblance, is examined.

XXIII. J. O. IRWIN. "*On the Frequency Distribution of the Means of Samples from Populations of Certain of Pearson's Types.*" Metron, 1930. Vol. VIII, pp. 1-55.

In a previous paper the author has given a general solution for the frequency distribution of the means of samples of any size, drawn at random from any population whatever, expressed as a definite integral. The present paper applies this solution to the

https://doi.org/10.23637/ERADOC-1-63

particular cases of Pearsonian Type I and Type VII curves for integral values of the exponents p and q, which enter into the equation of the Type I curve, and of the exponent m which enters into the equation of the Type VII curve.

The cases $p = 1, 2, 3, 4$, $q = 1, 2, 3, 4$ are discussed in detail for samples of size 2, 3 and 4 from Type I populations, as are the cases $m = 2, 3, 4$ from samples of size 2, 3 and 4 from Type VII populations. For the Type VII populations the cases $m = 1$ for any size of sample, and $m = 5, 6, 7, 8$ for samples of 2 have also been considered in detail.

XXIV. H. G. SANDERS. *"A Note on the Value of Uniformity Trials for Subsequent Experiments."* Journal of Agricultural Science, 1930. Vol. XX, pp. 63-73.

The question attacked is whether soil variations are sufficiently constant from year to year to give useful corrections to the yields of experimental plots from their yields under previous uniformity trials, and the data investigated were the published results of uniformity trials carried out on two fields at Aarslev (Denmark) between 1906 and 1911. In one case the plots did tend to keep constant in their relative yields, and the precision of an experiment would be increased by nearly 150 per cent, if the regression on the mean yield in the three previous years were used; with the other field, however, the plots showed no constancy in yield (when the variation due to strips was taken out as in modern experimental methods), and consequently previous uniformity trials could give no assistance.

XXV. J. WISHART. *"The Derivation of Certain High Order Sampling Product Moments from a Normal Population."* Biometrika, 1930. Vol. XXII, pp. 224-238.

In a recent paper on the derivation of moments and product moments of sampling distributions, R. A. Fisher dealt among other things with measures of departure from normality, and gave approximate expressions for the semi-invariants of these statistics. If a higher degree of approximation is desired, further high order product moments are required, and these are deduced in this paper, while certain simple relations existing between the formulae, which will be demonstrated elsewhere, are stated, for sampling from a normal population, thus enabling the high order results to be derived from simple expressions already known.

(*b*) Genetics

XXVI. R. A. FISHER. *"The Evolution of Dominance; Reply to Professor Sewall Wright."* The American Naturalist, 1929. Vol. LXIII, pp. 553-556.

The calculations which led Professor Sewall Wright to consider that the selective intensities available for the modification of dominance, are insufficient to have brought about great results are, in a different notation, identical with those that originally led the

https://doi.org/10.23637/ERADOC-1-63

author to attach importance to them. A slight mathematical error has, however, led Professor Wright in the special case chosen for examination, to the conclusion that the selective intensity starting from a low value decreases continuously, whereas in reality, it increases in that case without limit. The conclusion that mutations have had time to become in many cases completely recessive can only be rejected by assuming that small selective intensities do not bring about effects proportional to their magnitude. Although it is inevitably impossible to demonstrate extremely slow processes experimentally, yet there are general reasons for concluding that there is no such restriction upon such small selective intensities as Professor Wright is obliged to postulate.

XXVII. R. A. FISHER. *"The Evolution of Dominance in Certain Polymorphic Species."* The American Naturalist, 1930. Vol. LXIV, pp. 385-406.

Polymorphism in wild populations must usually imply a balance of selective agencies, of which the simplest type is a selective advantage of the heterozygote over both homozygotes. Such a condition should not be confused with the maintenance of a rare mutant type against counter-selection by means of repeated mutations. While such mutations should, on the theory of the selective modification of dominance, tend to become recessive, heterozygotes in polymorphic species will tend to resemble in external appearance whichever homozygote it is most advantageous to resemble. The selective balance must then be maintained by some constitutional disadvantage of the homozygous dominant.

Nabours' experiments with the grouse locust *Apotettix* do, in fact, show such a deficiency of homozygous dominants as is required by this thoery. The average amount of the deficiency is about 7 per cent. In six individual cases the deficiency is statistically significant, and six more show a non-significant deficiency, against two showing a non-significant excess.

The incidence of dominance and linkage in the fish *Lebistes reticulatus* strongly suggests that the colour genes found by Winge are advantageous in the male, but disadvantageous in the female.

The association of the three peculiarities of polymorphism, close linkage and the universal recessive type of dominance is found in molluscs, arthropods and vertebrates. It is tentatively suggested that, at least in the grouse locusts and the snails, the primary cause of the two other phenomena may be found in the closeness of linkage within or between chromosomes. This condition presents an obstacle to normal evolutionary development by gene substitution, and so makes it possible for abnormalities such as duplications to possess occasional advantages, so setting up the stability if the gene-ratio necessary for polymorphism; if the advantage lies in the external appearance, the polymorphism will be manifest, and the variant form will tend to become dominant.

XXVIII. R. A. FISHER. *"Note on a Tri-Colour (Mosaic) Mouse."* Journal of Genetics, 1930. Vol. XXIII, pp. 77-81.

The occurrence is reported of a female mouse showing both black and chocolate markings. Only one such case has occurred out of about 1,500 blacks heterozygous for chocolate, bred in the

https://doi.org/10.23637/ERADOC-1-63

same colony. Mating with a chocolate son gave 30 chocolates, 16 blacks and no tri-colours. The case resembles that of a male guinea-pig reported by Wright and Eaton, which also shows a deficiency of heterozygous offspring. The most probable explanation of both cases is that we have a mosaic, both somatically and in the germinal tissue, originating in non-disjunction. Some apparently analogous cases in mice and rabbits point, however, to a different interpretation for these cases.

XXIX. R. A. FISHER. *"The Distribution of Gene Ratios for Rare Mutations."* Proceedings of the Royal Society of Edinburgh, 1930. Vol. L, pp. 204-220.

The discussion of the distribution of the gene ratio of the author's paper of 1922 is amended by the use of a more exact form of the differential equation to be satisfied. A method of functional equations is developed for dealing with the termini, and is shown to lead to the same solutions as the amended differential equations, in the central portion of the range, for which the latter are valid, and further to give the terminal distribution of rare allelomorphs. The method requires the investigation of a continuous function u_v, of argument v, satisfying the recurrence formula

$$u_{v+1} = e u_v - 1$$

From the asymptotic form of this function its expansion in the neighbourhood of $u=o$ is derived, giving the frequencies of the required distributions.

Exceedingly minute values for the selective advantage or disadvantage make a great difference to (i) the chance of success of a mutation, and (ii) the contribution of such mutations to the specific variance. The order of magnitude to be considered is the inverse of the population of the species. The neutral zone of selective advantage in the neighbourhood of zero is thus so narrow that changes in the environment, and in the genetic constitution of species, must cause this zone to be crossed and perhaps recrossed relatively rapidly in the course of evolutionary change, so that many possible gene substitutions may have a fluctuating history of advance and regression before the final balance of selective advantage is determined.

XXX. J. B. HUTCHINSON. *"The Application of the 'Method of Maximum Likelihood' to the Estimation of Linkage."* Genetics, 1929. Vol. XIV, pp. 519-537.

The "Method of Maximum Likelihood," developed by Dr. R. A. Fisher, is applied to the problem of estimating linkage in cases involving complementary and duplicate factors.

Variances are calculated for existing formulae, and their efficiencies are determined to show that the "Method of Maximum Likelihood" is in all cases superior to any other method of estimation.

The amount of information supplied per plant by Maximum Likelihood formulae for F_2's and backcrosses, and by other formulae for F_2's is calculated and compared with the amount of information supplied per plant by a simple—that is, completely classified—backcross. From figures 2, 3 and 4 of this paper it is possible to estimate the size of family necessary to give any required degree of accuracy.

F

https://doi.org/10.23637/ERADOC-1-63