Report for 1956

Statistics Department

F. Yates

STATISTICS DEPARTMENT

F. YATES

Two members of the scientific staff left during 1956. Averil M. Munns resigned on marriage, and A. W. Whitwell left to take a post with the United Steel Companies Ltd., Rotherham.

New appointments were : Joan Anderson from the Sheep and Wool Branch of the Queensland Department of Agriculture and Stock and O. B. Chedzoy from the English Electric Company Ltd., Kidsgrove, Stoke-on-Trent. D. J. Knight was transferred from the employment of the National Research Development Corporation to Rothamsted staff.

Fifteen temporary workers were accommodated during the year, six of them from various countries of the Commonwealth.

The Agricultural Research Council Statistics Group at Cambridge has now been placed under the technical supervision of the Rothamsted Statistics Department. Members of the unit are working in close collaboration with the staff of the Department, and are making increasing use of the facilities provided by the electronic computer.

THE ELECTRONIC COMPUTER

The electronic computer has continued to give good service, and its versatility is steadily being increased, both by improvements in the ancillary equipment and by additions to our repertoire of programmes.

The high-speed tape reader (Ferranti Mark II) has now been in service for three months. A new fast-input programme enables this reader to take in tape at about 130 characters per second, compared with the 35 characters per second of the old reader. Initial difficulties with the new reader have now been largely overcome. The parity checked code and control system installed last year has been of immense value in detecting reader errors.

A punched-card reader, which enables the machine to read data direct from Hollerith cards, has been linked to the computer. This consists of a summary tabulator chassis modified to give the necessary input information and control signals. The information on cards, being punched in the conventional manner, has to be converted into binary form by the machine. This has set an upper limit of 100 cards per minute on the input speed, and the present input programme is designed to read in and convert a maximum of 27 columns of a card into binary form at this speed. The input programme includes a check on possible reader errors.

During September the disc store was returned to Elliott Bros. for overhaul and respray. The bearings were checked, and no measurable wear was found. The reading heads were all found to be in good condition. After removing the old coating a new, slightly harder and thicker coating was sprayed on the disc. This has resulted in improvement of the head characteristics, and the

performance of the disc since the overhaul has been very satisfactory. It was hoped at the same time to increase the storage capacity of the disc, but this did not prove possible.

Experience has shown that the double triode valve, which is extensively used in many modern computers, especially in the logical circuits, has a much shorter life, 1,500–2,000 hours, when used in this way than that obtained when used in conventional ways (10,000 hours or more). The valve manufacturers are investigating the causes for this premature failure, and one firm has produced a double triode with a special cathode. At the request of this firm we are at present running a life test on these valves, using a standard form of valve as control. After 1,600 hours' running 37 per cent of the control valves and 5 per cent of the test valves have failed.

The advent of the electronic computer has naturally resulted in a large increase in the volume of numerical work that has been undertaken in collaboration with scientists at other research institutes and experimental stations. The more interesting items of this type are mentioned in subsequent paragraphs of this report.

In April, owing to increasing demands on machine time, a more detailed form of log-keeping was introduced. The analysis of these figures for the period April–December is given in the accompanying table.

Record of machine operation over the period 1 April–31 December 1956 *

Details of operation				Hours	Percentage
System modification (M)...	...	...	...	94·3	5·3
Scheduled maintenance (S)	...	...	...	208·3	11·7
Unscheduled maintenance (U)	...	...	...	46·7	2·6
Programme checking (C)	...	...	...	574·5	32·1
Production runs (P)	...	...	...	757·2	42·3
Abortive efforts (A)	...	...	...	47·3	2·7
Idle time (I)	...	...	...	58·4	3·3
Total machine hours (T)	...	...	...	1787·0	100·0
Normal equivalent working hours (N) ...		...		1416·0	

Total time as percentage of normal laboratory time $= \dfrac{T}{N}$ $\qquad = 126\cdot3\%$

% availability $\qquad\qquad = \dfrac{C + P + I}{T}$ $\quad = 77\cdot8\%$

% efficiency $\qquad\qquad = \dfrac{C + P + I}{T - S - M}$ $\quad = 93\cdot6\%$

% usage $\qquad\qquad = \dfrac{C + P}{C + P + I}$ $\quad = 95\cdot8\%$

* This excludes the period 28 August to 9 October, when the disc store was removed for examination and respray. During this period the punched-card reader was linked to the machine, but no details of time spent on this modification were kept.

It will be seen that the machine is now working substantial amounts of overtime. Now that the machine is capable of reading

data from punched cards we expect that the demands on it for productive work are likely to increase substantially, and it is probable that systematic shift working will have to be organized.

COMPUTER TECHNIQUES

The Automatic Programming Routine mentioned in the 1955 Report has been further revised and improved and a brief description published (206). It has now proved itself to be an exceedingly useful programme. It was used, for example, in the construction of a programme for the analysis of 2^n factorial experiments. This programme deals with all types of 2^n designs (and also those including factors at 4 and 8 levels) with up to 128 plots. Total and partial confounding, fractional replication and estimation of error from high-order interactions can be simply dealt with. (There is, for example, no need to specify which degrees of freedom are actually confounded.) A simple extension of the programme provides for the analysis of 8×8 quasi-Latin squares. The whole programme contains about 800 orders. A similar programme for 3^n designs is being developed.

Papers on various applications of electronic computers have been prepared. Lipton (198) discussed various practical applications of the computer, summarizing our experience up to the middle of this year. In a further paper (199) he has described two programming techniques of particular use with the one-plus-one address code that is used in several machines employing a magnetic drum store.

The routine analysis of the results of field experiments forms one of our major production tasks on the computer, and Yates, Healy and Lipton have prepared a paper describing the techniques developed for this purpose (205).

EXPERIMENTAL DESIGN AND ANALYSIS

Garner, Patterson and Warren have completed the analysis of the results of the four-course rotation experiment at Rothamsted on organic manures and forms of phosphate. Patterson has also prepared a paper on the statistical aspects of this analysis (201). The examination of the results obtained so far on the Woburn ley–arable experiment has been carried out by Boyd and Dr. H. H. Mann, and a paper has been prepared.

Boyd, Garner and Haines completed the examination of the large series of sugar-beet experiments carried out from 1933 to 1949, and a full report has been prepared (187). Boyd has also prepared a further paper dealing with the economic aspects of the manuring of sugar beet (188).

Lessells has completed a summary on the effect of nitrogen on cereals (191).

We have continued to work in close collaboration with the National Agricultural Advisory Service in the design of both crop and animal experiments. As during last year, the electronic computer has been extensively used in the analysis of experiments of

standard design. The following table shows the numbers dealt with (each variate analysed being reckoned as one unit) :

	Randomized blocks and split plots	Latin squares	3^3	2^n
Rothamsted, etc.	365	271	—	13
National Agricultural Advisory Service	544	53	147	40
National Institute of Agricultural Botany	268	—	—	—
Total	1177	324	147	53

A report on an experiment on diets for laying fowls by Mr. I. W. Rhys, Mr. S. Culpin, Mr. G. E. Forsyth and Westmacott (202), and one on experiments on the feeding of penicillin to pigs (Westmacott and Whitwell 204) have been prepared.

Lessells has analysed the month-to-month variation in the results of soil analyses. A report on this topic has been presented to the Closed Conference of Soil Chemists.

A series of factorial experiments on the inter-relationships of copper, molybdenum and sulphate ions in the metabolism of pregnant and non-pregnant ewes was analysed and interpreted by Leech for the Biochemistry Department of the Ministry of Agriculture Veterinary Laboratory at Weybridge. These results were the basis of a paper presented by Mrs. R. Allcroft at the International Grassland Congress in New Zealand.

SURVEYS

Ten counties (or part counties) have been surveyed during 1956 in the Survey of Fertilizer Practice, and the analysis of the results is proceeding. Because of the importance which the Soil Chemists attach to obtaining soil-analysis data from random samples of fields in different parts of the country, it was decided to resume the practice of taking soil samples in the course of these surveys. Since 1954 information has also been obtained on the use of weed-killers and insecticides. A note based on the results of recent surveys has been prepared on cereal varieties grown in England and Wales (192), also a further note on the use of weed-killers in England and Wales (190). Observations made on the British potato crop were the subject of a paper given by Boyd to the Operational Research Society (189).

The survey of diseases of dairy cattle has been continued in collaboration with Dr. F. W. Withers and the Veterinary Laboratory at Weybridge. A paper is being prepared presenting the results of two years' work in Wiltshire and Devonshire. The Ministry of Agriculture is planning to extend this work to a nation-wide survey, and the proposed method is about to be tried out on a pilot scale in six counties of England. The analysis of the survey of mortality and wastage in pregnant ewes is now completed, and a paper is being prepared in collaboration with Dr. K. S. Sellers. Leech has also collaborated with Mr. J. W. Egdell and Mr. S. B. Thomas of the

N.A.A.S. in the analysis of the bacteriological data from the Survey of Methods of Milk Production (197).

A start has been made on programming the electronic computer for the analysis of survey data.

METHODOLOGY

Healy has devised a direct method for the computation of canonical correlations suitable for an electronic computer (195). This is analogous to the " rotation " technique for the evaluation of latent roots and vectors.

Sir Ronald Fisher has recently developed explicit formulae for the probability integral of Behrens' distribution used in testing the significance of the difference between two means with possibly unequal variance (193). Using these, Healy has computed certain significance points of this distribution for cases when the numbers of observations in each sample are small. These tables will be included in the fifth edition of *Statistical Tables for Biological, Agricultural and Medical Research*.

ENTOMOLOGY

Marjory G. Morris has collaborated with members of the Entomology Department in the analysis of a large experiment on the relative efficiencies of different suction traps.

H. R. Simpson has collaborated with Mr. W. H. Potts of the Colonial Office Tsetse Committee in a study of the possibility of tsetse control by the release of large numbers of male flies sterilized by gamma irradiation (211). Since the tsetse female is believed to mate only once, this should lead to a high proportion of sterile matings. The method has been successfully applied to the eradication of screw-worms in Curaçao. It was required to investigate theoretically the effects of releasing different numbers of sterile males over varying periods so that the practicability of the method could be assessed. A mathematical model of the tsetse population was set up and followed through a number of generations on the computer. By varying the parameters of the model, the importance of the underlying assumptions could be tested and the effects of different release patterns determined.

GENETICS

An increasing amount of numerical work with a genetical background is being undertaken on the computer.

Dr. E. C. R. Reeve (Institute of Animal Genetics, Edinburgh) is investigating the effect of linkage and selection in inbreeding. He considers a population inbred by sib-mating in which the homozygotes at a particular locus are selected against, and estimates the approach to homozygosity at a second locus linked to the first. Mathematically, this involves the repeated multiplication of a vector by a 19×19 non-symmetric matrix, the elements of which depend on the closeness of the linkage and on the degree of selection operating. A programme has been worked out which calculates the appropriate genotype frequencies in successive generations and which

also gives the speed of the ultimate approach to homozygosity. This same programme has been modified to deal with the similar problem of assessing the effect of equal linkage of a gene to two pairs of lethals in a sib-mating system. The matrix in this case is only 10×10 (Gower).

Dr. J. W. B. King (Animal Breeding Research Organization, Edinburgh) has obtained a large body of data on pigs from the Danish progeny testing stations. This comprises a number of different measurements taken on litters of known parentage, and it is required to break down the observed variances and covariances into genetically meaningful components. A start has been made on standard punched-card equipment, but this is extremely laborious, and the job is being transferred to the computer now that the card-reader is available. (Westmacott and Cohen.)

Dr. J. H. Renwick (Galton Laboratory) is estimating linkage from a variety of human pedigree data. Even when all phenotypes in a pedigree are known, the genotypes are usually not exactly ascertainable, and the method of analysis consists in determining the probability of occurrence of the pedigree in terms of the gene frequencies and the recombination fraction, and in choosing the value of the latter which maximises the probability. For large pedigrees this is probably beyond the capabilities of our computer, but interesting results have already been obtained from a set of pedigrees involving only two generations. (H. R. Simpson.)

Dr. A. E. Mourant (Blood Group Reference Laboratory) has asked us to fit blood-group gene frequencies by maximum likelihood to a large number of observations from all over the world. The appropriate programme is likely to be of fairly general application. (Healy.)

COLONIAL WORK

The amount of advisory work for Commonwealth workers has decreased since the departure of Hodnett at the end of last year. Six Commonwealth workers have stayed in the Department for varying periods and a number of queries have been dealt with by correspondence by Church and Westmacott, who have taken over the responsibility for this work.

OTHER WORK

Healy and Lipton, with Dr. E. H. Ashton of Birmingham University, have completed a large-scale study of measurements on the teeth of humans and the great apes with a view to clarifying the relations between these groups and the fossil anthropoids recently discovered in South Africa and elsewhere (196). The extremely heavy computations were carried out on the electronic computer.

Gower has collaborated with J. H. Rayner of the Pedology Department in the development of computer programmes for the analysis of crystallographic data. One of these is a general structure factors programme, and another is designed to calculate two-dimensional Fourier syntheses. These programmes have been used to refine a projection of sodium formaldehyde bisulphite.

173

The programme developed by Rees for the solution of differential equations has been used in collaboration with Dr. R. K. Schofield for the calculation of membrane pressure for a sol of spherical particles (203). Rees has also collaborated with Dr. G. Foster of the London School of Economics in the tabulation of percentage points of a generalized beta distribution which forms the basis for certain tests of significance in multivariate analysis. A paper on this work has been prepared (194), and has been accepted for publication.

Dr. K. L. Blaxter, of the Hannah Dairy Research Institute, and Dr. H. Ruben, Manchester University, have developed a mathematical model of the metabolism of dairy cattle which is intended to form a basis for improved patterns of feeding. The process of fitting the model to observed data is extremely complex, and Lipton is developing a computer programme to cope with this problem so that the adequacy of the model may be assessed.

Leech, with Dr. A. B. Paterson and Dr. P. L. Stuart of the Weybridge Veterinary Laboratory, has investigated the efficacy of routine tuberculin testing of dairy cattle. He has developed a discriminant function which makes the best use of the measured reactions of the cow to different tuberculins (200). He discussed this problem at a meeting of the Royal Society of Medicine and pointed out the importance of the differences in sensitivity of animals of different ages and sexes.

A new edition of *Statistical Tables for Biological, Agricultural and Medical Research* has been prepared and is now in the hands of the printers. Several new tables have been included, notably tables connected with the logit and log–log transformations and a table giving the " product ratio " corresponding to different recombination fractions for the analysis of genetical linkage tests (207).

Yates and Healy prepared a revised statistical section for the third edition of the *Van Nostrand Scientific Encyclopaedia*.

Overseas Assignments

Yates spent two months in the early part of the year in Calcutta and New Delhi on a Technical Assistance Mission of the United Nations. Healy spent seven weeks in the United States and Canada on a grant from the Rockefeller Foundation visiting statistics departments and electronic computing centres. During the trip he took over the secretaryship of the Biometric Society from Dr. C. I. Bliss. Church has been seconded for a year to the Food and Agriculture Organization of the United Nations to work in Ethiopia on a Technical Assistance Mission. His main task is the initiation of agricultural surveys.

Conferences, Lectures, etc.

Healy gave a course in statistics for post-graduate workers in biological and agricultural sciences at the University of Nottingham School of Agriculture. Boyd attended the 19th Congress of the International Institute for Sugarbeet Research held in Brussels in February and also the International Potash Congress in London in

August, where he gave a paper on the effect of potash on crop yield (186). Rees gave colloquia at Imperial College and the London School of Economics on electronic computing in a statistical department. Healy attended the meeting of the British Association in Sheffield and gave a paper with Dr. J. M. Tanner on the genetics of human morphological characters (185). He also contributed to a symposium on the philosophy of experimentation at the Royal Statistical Society Study Group. Lessells gave a lecture on the use of fertilizers on grassland at a Refresher Course for officers of the Yorkshire and Lancashire Province, National Agricultural Advisory Service.