

REVIEWS

Theory of Probability. By HAROLD JEFFREYS.

[2nd ed. Pp. viii + 411. Clarendon Press, Oxford, 1948. 35s.]

Probability and Induction. By WILLIAM KNEALE.

[Pp. viii + 264. Clarendon Press, Oxford, 1949. 15s.]

Probability Theory for Statistical Methods. By F. N. DAVID.

[Pp. ix + 230. Cambridge University Press, 1949. 15s.]

THE casual reader attracted by the word *probability* in the title of a book must be prepared for at least three possibilities. The work may be a philosophical treatment of the meaning of probability and of its various interpretations; it may be a mathematical treatise developing the theorems of the probability calculus; or (usually on the Continent) it may be a text-book of statistical theory. The current crop of books on probability provides us with an example from each of these categories.

It is a testimony either to the readability of Jeffreys's writing or to the omnivorousness of statisticians' reading that so many of them have read his book though so few of them accept his premises. It may be remembered (*J.I.A.* Vol. LXXI, p. 347) that Jeffreys's work was effectively a treatise on the techniques of statistical inference, a large part of it being concerned with the testing of statistical hypotheses and the estimation of parameters in probability laws. The second edition differs but slightly from the first. The postulational basis of probability theory and its development reappear in the first two chapters with a few verbal changes, though a short appendix has been added on the consistency of the product rule. Likewise chapters VII and VIII, the last in the book, have been reproduced with only minor changes, though the forthright attacks on statistical writers in the former leave room for revision. In the chapters on estimation and significance tests there are a few new sections, some rearrangement of the material, and an improved treatment of (i) the problems of the estimation of intra-class correlation and of the simultaneous testing of two parameters, and (ii) the tests for (a) independence in rare events, (b) a continuous departure from a uniform distribution of chance, and (c) deviation of the mean of a normal universe from an assigned value. Nine new sections were noticed (viz. 3.44, 3.61, 3.9, 4.01, 4.21, 5.01, 5.64, 5.93, 6.3), seven of them between one and two pages in length but one, on 'invariance theory', extending over nine pages of text. This section purports to establish 'the possibility of a consistent rule for stating prior probability over large parts of the subject', and is closely related to some of Perks's arguments in *J.I.A.* Vol. LXXIII, p. 285.

Jeffreys is concerned rather more than most mathematical authors with the logical principles of his subject, but Kneale, on the other hand, is interested only in the philosophical problems of the definition of probability and its use in the interpretation of induction generally. Though he considers various definitions of mathematical probability his book contains little mathematical symbolism and is directed towards philosophers interested in the principles of scientific method.

However, about two-fifths of the contents of *Probability and Induction* is devoted to an excellent discussion of probability theory in which the frequency definition is criticized and the subjectivist standpoint is demolished—the manuscript of the book was largely complete before the war and Jeffreys is not mentioned. The approach Kneale finds most plausible is a modification of von Kries's notion of the 'range' of the characters considered: the probability of an A thing's being B is defined as the proportion of the range of A-ness which belongs also to the range of B-ness. Whatever opinions may be held about this definition there is no gainsaying the fact that Kneale's treatment of his whole subject is skilful and surprisingly easy for the amateur philosopher to follow.

David's book is something else again. It is much more like the type of book we have come to associate with the classic French and German writers on probability theory. After a short, deliberately naïve, chapter on fundamental ideas the author passes to her definitions and preliminary theorems, the binomial law and the Poisson and normal approximations to it. She completes half her book with chapters on Bayes's theorem, genetical applications, and the multi-nomial law. As might be expected from one so well versed in statistical theory and applications and from the stated object of the book, namely, to 'prove in elementary mathematical language those propositions and theorems of the calculus of probability which have been found useful for students of elementary statistics', David writes throughout in a manner calculated to appeal to statisticians. Her account of the methods available for evaluating numerically the sum of a number of binomial terms is unusual, but all the more welcome, in a text on probability theory. And her other innovations, e.g. her introduction of confidence intervals along with Bayes's theorem, are also pleasurable to the reviewer. However, we must except from this category the dubious inclusion of a series of Neyman which is merely a particular example of a generalized Poisson instead of the latter itself, and the rather specialized chapter on 'simple genetical applications' with its somewhat pedantic references to *panmixia* instead of *random mating*.

The last eight chapters of David's book are still refreshingly novel material in an elementary text on probability theory. We here find random variables, sampling moments, Lexis theory, estimation, Markoff least squares theory with an application to sampling human populations, characteristic functions, and the Central Limit theorem handled in a straightforward and interesting manner. Actuaries who wish to make a first acquaintance with any of these subjects could not do better than study David's treatment.

In so much that is good it is perhaps ungracious to notice a couple of minor faults of the author. In our opinion she strains after a somewhat spurious mathematical rigour in spite of her disclaimer in the Preface. For example, the proof of Poisson's binomial limit only needs the complicated derivation of David's text if it is desired to include the ordinates at infinity—explicitly excluded by the authoress. Furthermore, the danger of repeating a historical say-so of a previous author without independent verification is illustrated on several occasions. For instance, the relation between the binomial series and the incomplete β -function was not only 'almost certainly known before this century', but is to be found on p. 151 of that great-grandfather of probability text-books, Laplace's *Théorie analytique* (1812); and the claim that the first table of differences of zero appeared in 1925 is easily refuted, since such tables can be traced back to Stirling (1730) and were fairly numerous by the beginning of this century.

The three books we have briefly reviewed are scarcely to be recommended indiscriminately. Each provides an excellent treatment of its subject-matter disguised under a 'probability' title: Jeffreys, a handbook of statistical theory from a logical (though unpopular) probability standpoint; Kneale, a discussion of various probability definitions in their relation to induction; and David, a compendium of probability theorems for reference by the practising statistician.

H. L. S.

Statistics for Economists. By R. G. D. ALLEN.

[Pp. vii + 216. Hutchinson's University Library, 1949. 7s. 6d.]

IN explaining the purpose of this series of books the dust-cover states: 'The general aim of the whole series is to provide popular yet scholarly introductions for the benefit of the general reader, but more especially for the unprofessional student who wishes to pursue his chosen subject systematically up to something like a University standard. The books are written to be intelligible to those who have made no previous study of the respective subject; for this reason as few technical terms as possible are used, and when such terms are necessary they are clearly explained when they first occur.' The first reaction of any self-respecting statistician asked to write a book to comply with these requirements must be to say emphatically that the task is impossible. To give even an elementary introduction to modern statistical method the reader must be assumed to have not only an elementary knowledge of algebra but also a definite liking for this subject.

Prof. Allen has tried hard to exclude algebra in spite of his assertion on p. 9 that 'Statistical method is essentially a branch of mathematics using the processes of reasoning which make up scientific method in general'. Surely it is more correct to say (and Prof. Allen does in effect say it in other places in the book) that a very important procedure in scientific method is the making of comparisons, that statistical method is a branch of scientific method concerned with numerical data relating to the attributes or properties of the individuals of some population or aggregate, and hence the methods of comparison employed range from the use of simple tables and diagrams to the use of very advanced mathematical methods. The mathematical methods bring out relationships and structure which cannot easily be inferred by the simpler comparative methods.

If this be true then a book for the non-mathematical reader must perforce confine itself to a treatment of tables and diagrams and common-sense inference. If the writer goes beyond this the non-mathematical reader will not understand him. On the other hand, if he introduces an inadequate mathematical treatment, such as, for example, Prof. Allen's treatment of the Analysis of Variance (pp. 133-36), neither the non-mathematical reader nor the mathematical reader will feel pleased with the author. Another important branch of the subject which suffers in this book from lack of an adequate mathematical foundation is 'Sampling and Significance', dealt with in the final chapter. An understanding of the ideas contained in this chapter requires some knowledge of probability theory, but all that the reader is told about probability is that 'the frequency distribution is often called a probability distribution'.

It is clear that Prof. Allen himself does not consider the book by itself as adequate reading. There are numerous references to well-known text-books, and

the beginner will have to follow the author's recommendations to the letter if he wishes to acquire a sound knowledge of the subject. A good example of the use of technical terms which will be incomprehensible to the general reader, who is too lazy or has not the time to consult other reading, is the reference to the 'net reproduction rate' on p. 119.

The actuary who is familiar with statistical method and with the handling of numerical data and who wishes to read a comprehensive but relatively brief account of the application of statistical method to economic data can be confidently recommended to read Prof. Allen's book. He will find excellent elementary accounts of the Sources of Published Statistics, Index Numbers and Time Series, with references to further reading. Any actuarial student who has the time will find the book useful in connexion with his study of Finance and Investment. It will give him some of the background for an understanding of National Statistics and the index numbers of various kinds which appear in the financial press.

There is one point of real criticism that the reviewer feels he must make. In any work on sampling and significance the writer should stress the part played by the investigator's general knowledge of the field of inquiry. Is it best for the investigator to be a scientist first and a statistician second? Many students are apt to apply statistical tests rigidly and dogmatically without properly considering whether the assumptions underlying the theoretical model can be justified. Prof. Allen does in the early chapters emphasize the importance of knowing the source of the data and the methods used to collect them, but it is very important that when the more elaborate mathematical techniques are used the investigator should be either an expert himself or closely associated with an expert in the subject under investigation. Prof. Allen may have thought that it was unnecessary to make this point since the book is entitled *Statistics for Economists*. It should be remembered, however, that the publishers hope that the general reader will purchase the book.

H. W. H.

Sampling Methods for Censuses and Surveys. By F. YATES, Sc.D., F.R.S.

[Pp. xiv + 318. Charles Griffin and Co. Ltd., London, 1949. Price 24s.]

REVIEWS of scientific books form, unhappily, a field of literature offering severely restricted possibilities. It is a regrettable, but unavoidable, result that certain words and phrases have been so overworked that one scarcely dares to use them for fear of being accused of laziness, triteness or insincerity. However, when such words or phrases serve best to express one's opinion, it seems an unnecessary concession to 'style' to use ingenious but pointless circumlocutions.

Let it be said, therefore, that the present book *does* fulfil a long-felt want, it *does* break new ground, it *does* contain an exhaustive treatment of its chosen topic, and it *is* likely to prove of value to a wide circle of readers. A further, and most welcome, feature is the essentially statistical approach to the problems considered. A minimum of mathematical symbolism is used in the first five chapters, and no formal mathematical development is attempted. The absence of formal mathematics is not, however, associated with an absence of logic in the arguments used (as all too often happens in 'elementary' text-books). A careful study of this book should result in a clear understanding of the various sampling

methods in current use, the types of problem to which they may most profitably be applied, and the most economical and efficient ways of analysing the data collected. In all parts of the book there is awareness of the effect which practical limitations may have on theoretical considerations, and of the necessity to make the best of the conditions actually prevailing.

This is the first attempt to provide a systematic and comprehensive description of the general theory and practice of modern sampling, and it is of interest to see the way in which Dr Yates has chosen to develop the subject. This can best be described by the following extract from the first chapter of the book.

The method of presentation adopted in this book is to take the various parts of the sampling process in roughly the order they are encountered in the execution of a census or survey, and to discuss the various aspects of each part in turn. Thus Chapters 2 and 3 describe the various types of sample that can be used, and the general principles to be followed in the selection of a sample, Chapter 4 deals with the practical planning of a survey, and Chapter 5 with the problems encountered in its execution and in the abstraction of the results. The remaining chapters are concerned with more strictly statistical problems. Chapter 6 deals with various methods of estimating the population values, Chapter 7 with the estimation of sampling errors and Chapter 8 with the determination of the relative efficiency of the various sampling methods.

As the author himself points out, a knowledge of the last three chapters is essential to a complete understanding of the various methods proposed in the earlier chapters. Nevertheless, a study of the earlier chapters alone will provide a most useful knowledge of the methods of sampling which are available to the investigator.

Sampling theory possesses a rapidly developing vocabulary. The majority of readers of this review will be familiar with the better-known concepts such as sampling unit, bias, random sampling, and random sampling error. Certain further terms, describing methods of drawing a sample which are in current use, may not be so well-known. The terms used in the book include *stratification* (the division of the population into groups or *strata*), *multi-stage sampling* (the division of the population into first-stage sampling units, e.g. villages, and the subdivision of the latter into second-stage sampling units, e.g. households, with, possibly, further subdivision), and *multi-phase sampling* (the collection of fuller information in regard to part or parts of the sample than in regard to the remainder of the sample). These three devices can be used either alone or in combination with each other. It is clear that a great variety of possible sampling schemes can be formed by such combinations. Guidance is given on the suitability of the various possible schemes to different types of problem.

Methods of checking the accuracy of the sampling are described at some length; in particular, the method of *interpenetrating* samples is explained. Interpenetrating samples may be roughly described as independent 'twin' samples each of which is a check on the other. Another point of some importance is the treatment of 'non-response', that is, failure to obtain a measurement from one or more of the selected sample units. It is clear that there may be some loss of 'randomness' in dealing with non-response. The book contains a useful, though not exhaustive, discussion of practical methods of dealing with this situation.

Of special interest is the description (in Chapter 5) of methods of recording collected data. The possibilities of the Cope-Chat card system and of the Hollerith mechanical punched card system are discussed in a lucid and thorough manner.

Such faults as are to be found in this book seem to arise from the conditions of its production rather than from any shortcomings of the author. First, there was some haste necessary in its preparation as the United Nations Sub-Commission on Statistical Sampling desired it to be completed before the 1950 World Censuses of Agriculture and of Population. Secondly, the book is an attempt to present a consolidated account of a subject which is still developing rapidly and which has not, as yet, the benefit of a universally accepted system of terminology and concepts. In his effort to give a definite form to the subject the author does seem to be a little pedantic, not always consistent in his definitions, and perhaps a little dogmatic in some of his opinions. There is, for example, the rather fine distinction between *census* and *survey*, introduced in Chapter 1. So far as can be understood from the book the distinction is based on the extent and complexity of the data collected in regard to each unit. (It may be mentioned that the author points out that both censuses and surveys may be either 'complete' or 'sample'.) The author does not keep strictly to this distinction. Thus on p. 124 there is a reference to 'surveys of the census type'. A further example of an excess degree of detail is the classification of types of defect of sampling frame under five heads (p. 60). These seem to overlap, notably 'incompleteness' (which 'is often confined to units possessing some special characteristic') and 'inadequacy'.

Exception may be taken to the statement that fiducial probability has 'certain logical advantages'; the subsequent description does little to make the concept clear or its logical advantages apparent.

The book is expressly intended for persons with 'little or no previous training in mathematical statistics'. Such persons will probably find Chapters 6-8 rather hard going. This arises from the nature of the subject, and the author cannot be blamed, except in so far as he is, perhaps, rather too sanguine as to the possibility of learning directly from his book, without the necessity of a preliminary general elementary study of statistical theory. In the absence of such previous knowledge these later chapters will not be easily mastered.

The various methods described in the book are amply demonstrated by numerical examples, mostly (but not exclusively) concerned with agricultural problems. A large bibliography provides examples of the application of sampling methods to a wide variety of problems. Actuaries will be particularly interested in an article by G. C. Campbell (*Problems with Sampling Procedures for Reserve Valuations*, *J. Amer. Statist. Ass.* Vol. XLIII, p. 413 (1948)) which appeared too late for inclusion in this bibliography.

The book is completed with a page from a table of random numbers, a concise table of the Normal distribution, and a good index. The size of type used is rather small, but the book is attractively set up and quite easy to read.

N. L. J.

The Elements of Insurance. By W. A. DINSDALE, B.Com., F.I.L.

[Pp. 130. Sir Isaac Pitman and Sons, Ltd., 1949. 10s. 6d.]

Accident Insurance Claims. By J. B. WELSON, LL.M., F.C.I.I., Barrister-at-Law.

[Pp. 169. Sir Isaac Pitman and Sons, Ltd., 1949. 10s. 6d.]

Fire Insurance Law. By H. TAYLOR, B.Com., F.C.I.I., Barrister-at-Law.

[Pp. 115. Sir Isaac Pitman and Sons, Ltd., 1949. 12s. 6d.]

THESE are the first three of a series of handbooks to be issued under the authority of the Chartered Insurance Institute and designed especially for the use of students. Mr Dinsdale's book, which is the one most likely to be of interest to actuaries, deals with the basic principles of insurance, such as insurable interest, good faith, indemnity, etc., with brief subsections describing the application of each to the various branches of insurance. Specimen proposal forms are shown, and policy conditions and acceptance procedure described. There are chapters on reinsurance, claims, and office administration, followed by a bibliography and glossaries of technical terms and Latin phrases.

All the books are excellently produced, and will undoubtedly be of value not only to students but to insurance men generally who wish to refresh their knowledge of the foundations of the business as a whole or to learn something of branches other than their own.