BOOK REVIEWS

Correspondence concerning reviews should be addressed to the Book Review Editor, Professor James F. Hannan, Department of Statistics, Michigan State University, East Lansing, Michigan 48823.

DENNIS V. LINDLEY, *Introduction to Probability and Statistics from a Bayesian Viewpoint*, Vols. 1 and 2. Cambridge University Press, New York, 1965. Vol. 1, Probability, xi + 259 pp. $6.00; vol. 2, Inference, xiii + 292 pp. $6.50.

REVIEW BY HERMAN CHERNOFF

Stanford University

The prospective author of a textbook in statistics is bedeviled by a host of problems. The subject matter is delicate and difficult. A complicated historical interplay in theory, textbooks and widespread practice serves to perpetuate a body of knowledge normally justified by dogmas which are controversial and mysterious when examined closely. The technical mathematical background required in a rigorous discussion of relevant theorems involves advanced work in mathematics, but fairly sound intuitive justification can be given with practically no reference to calculus. On the other hand, a body of illustrations requires a sound knowledge of calculus. With or without the use of the formal background, mathematical skill is of fundamental importance in facilitating the penetration to the heart of matters. An attempt to make a text readable and interesting to a large audience must cope with the widely divergent backgrounds of prospective readers.

The jargon and questionable dogmas have encouraged some authors to rewrite statistics in unconventional form. A few such attempts, based on poorly digested readings rather than experience in practice and theory have been unfortunate and rapidly forgotten. In recent years, some well established statisticians have begun to experiment with less conventional treatment.

Professor Lindley is a prominent statistician with many years of experience and a record of important contributions to the fields of Probability and Statistics. He is an outspoken Bayesian with a strong interest in decision theory. In writing this text he has reduced or eliminated a number of the above mentioned difficulties. By limiting the audience to university students of mathematics with a sound knowledge of calculus plus familiarity with matrix algebra, he has assumed students with skill and allowed himself the liberty of heuristic proofs when convenient. A liberal interpretation of his stated objective, of presenting the minimum that any mathematician ought to know about random phenomena, provides considerable choice among topics. A previous more orthodox draft required so much mental juggling to understand the concepts that he chose to present the less conventional and simpler Bayesian approach, where inference is regarded as the study of how data affect belief in hypotheses, belief measured by posterior probability distributions computed by Bayes theorem.

Two other calculated decisions devoted to saving space consisted of confining the discussion of decision theory to a five page description and avoiding data handling and substantial treatment of applications.

This reviewer feels that, as an introductory textbook on Statistics for prospective mathematicians, the books fail to accomplish the stated goals. Pedagogically the author fails to take sufficient advantage of his readers' mathematical sophistication to make the material aesthetic and stimulating. From the point of view of content, it is hard to believe that every mathematician ought to know about analysis of covariance but need not know anything about sequential analysis, design of experiments, or non-parametric techniques such as the Wilcoxon-Mann-Whitney test. Finally, the attempt to rationalize Classical Statistics in terms of Bayesian Statistics still requires substantial mental juggling. We are left with a book which is interesting and valuable as a reference relating Classical Statistics and Bayesian inference but rather tedious to the Mathematician interested in learning what Statistics is about. These bald subjective judgements require some background to be properly interpreted.

A few innocent looking assumptions of consistency in inference, together with the conception of inference as the appropriate way of measuring how evidence changes belief, lead to Bayesian inference and incidentally to the Likelihood Principle. Much of the current controversy on Bayesian inference seems to focus on a conflict between two aims, which this reviewer likes to label *consistency* and *robustness*. (This labeling tends to give the misleading impression that Bayesians are unconcerned with the robustness problem. On the contrary Bayesian techniques provide one useful attack on the problem and Lindley concerns himself with how approximations in the prior affect the posterior.) The Bayesians prove that consistency appropriately axiomatized implies the existence of prior distributions and the relevance of posterior distributions. The "opposition" demands procedures which are universally dependable and insensitive to apparently slight variations in assumptions and beliefs.

Although Bayesians, and Lindley in particular, criticize classical expositions because of their lack of consistency, strict insistence on consistency is too difficult for the Bayesians to cope with. It is effectively impossible to deliver the precise statement of the underlying probability model, or of the exact prior distribution (whose origin may unhappily depend on genetic factors). Precise calculation of posterior probabilities may be difficult to carry out and more difficult to comprehend. Thus the Bayesians adopt more moderate demands on their own methods. As I understand it, the argument proceeds as follows. Consistency implies prior distribution. Hence, understanding must involve the notion of prior distribution. In actual practice we can not be precise and consistent, but if we approximate our prior distributions reasonably well, our results will be reasonably good. On the other hand, formulations which ignore prior distributions can not be reliable although they can conceivably lead to sound results.

The second part of the book seems to be devoted mainly to justifying classical procedures by relating them to Bayesian inference. Since Lindley chooses to separate decision making from his conception of inference, the statistician becomes a passive observer whose function is the computation of posterior distributions. Statistical inference is now basically a dull subject. One multiplies a prior by a likelihood and computes a posterior, and that is that. Design of experiments and sequential analysis, which require active participation on the part of the statistician and some economic context for justification, are eliminated.

Without decision making, the justification of classical procedures requires some questionable contortions to be discussed later. Then these procedures are regarded as reasonable if the posterior distributions have appropriate properties. Considerable space is devoted to the presentation of prior distributions and likelihoods which yield appropriate posterior distributions. Except for chi-square goodness-of-fit tests, the methods of non-parametric statistics, which are apparently not easily related to Bayesian Statistics, are quietly ignored. Furthermore, one loses sight of the fact that some of the classical procedures that Lindley justifies were developed in a non-Bayesian framework and survived because they met the test of robustness. It is interesting that Lindley gives fundamentally the same justification for the Analysis of Variance, which is sound under assumptions wildly different from those presented, as for the normal theory method for comparing two variances, a method which is well known to lack robustness.

Confidence Intervals and Significance Tests are introduced as incomplete expressions of posterior beliefs which are useful when the complete description of the posterior is too complicated, as for example in the analysis of variance where there are too many parameters involved. A (Bayesian) confidence set for θ of coefficient β, based on data x, is a set $I_\beta(x)$ of θ with posterior probability β. A significance test of $\theta = \theta_0$ at level $\alpha = 1 - \beta$ is described as follows. The data is significant at level α if θ_0 is not in $I_\beta(x)$.

As an illustrative example, suppose that for given (but unknown θ) X_1, $X_2, \cdots, X_n$ are independent and normal with mean θ and variance 1. The natural prior, representing vague knowledge, is the uniform distribution on $(-\infty, \infty)$ which leads to the posterior distribution for θ which is normal with mean $\bar{X}$ and variance $1/n$. Then $\bar{X} \pm 1.96/n^{\frac{1}{2}}$ is a 95% confidence interval for θ and the test of $\theta = \theta_0$ leads to significance at level .05 if θ_0 is not in $\bar{X} \pm 1.96/n^{\frac{1}{2}}$.

The author indicates that, just as there is some arbitrariness in the choice of a confidence interval, so there must be arbitrariness about a significance test. Indeed, the specific example he gives simply begs for a one sided test, as he points out in more restrained language.

But there is more arbitrariness than he indicates. To be extreme one could define $I_\beta(x)$ as the complement of an interval (a_1, a_2) which contains θ_0 and has posterior probability .05. Then any set of data is significant at the .05 level for testing $\theta = \theta_0$.

Has Lindley succeeded in justifying or understanding the classical test by showing that it is one of a class of tests, some of which lead to nothing intelligible and none of which are complete descriptions of the posterior distribution? I don't think so. To be sure, he indicates a preference for a particular confidence set, namely the one for which elements in the set have greater posterior density than the elements outside the set. This leads to the symmetric two-sided test illustrated above and can be criticized on two grounds. First, the one-sided test is abolished and second, this apparently unique procedure is affected by transformations of the parameter.

Though this reviewer feels that the book fails to attain some of its principle goals, the sum of its positive accomplishments is extremely substantial. The book appears in two parts, entitled *Probability* and *Inference*. Part I on Probability is literally crammed with basic notions and results which are carefully explained. Among the high points is the presentation of the Rényi axiom system from both the frequency point of view and from the degree of belief point of view using betting odds. The genetics applications, and the imaginative pedagogic use of the Poisson distribution of Chapter 4 on stochastic processes, are examples of masterful presentations.

While a novice may find the approach of Part II on Inference tedious, someone better acquainted with the field can appreciate the breadth of interests and the importance of the topics covered. Of particular interest are (i) the theorem describing the effect on the posterior of approximating the prior, (ii) the Behrens distribution, (iii) the special prior distribution developed for the analysis of variance problem when differences among the unknown means have finite variance but the variance of any one of the differences tends to infinity, (iv) the treatment of maximum likelihood, and (v) the large sample Bayesian discussion of goodness of fit tests.

As a sort of postscript, I wish to add a few relatively minor remarks. There seems to be an error in the statement of independence for sampling without replacement from infinite populations. The justification of the axioms in terms of betting odds seems to carry an implicit assumption of linear utility which I think could and should be avoided by a more delicate argument. The book is remarkably free of errors. The collection of problems, many of which were selected from examinations at British Universities, fills an important need for teachers of Statistics.