BOOK REVIEWS

Statistical decision functions. By A. Wald. New York, Wiley, 1950. 9+179 pp. $5.00.

This book, published shortly before Wald's death in an airplane crash in India, is an outgrowth and generalization of the author's papers during the period 1939–1949 on the general theory of statistical decision. It represents a remarkable application to statistical theory of the methods and spirit of modern mathematics.

The general statistical decision problem as formulated by Wald can be described in the following terms. There is given a stochastic process $X = \{X_i\}$ and a class Ω of distributions which is known to contain the true distribution F of X as an element. Any rule δ for sampling from X and arriving at a terminal decision d^t belonging to a given class D^t is called a *decision function*. For any F and d^t there is given a function $W(F, d^t) \geqq 0$ representing the loss involved in taking the decision d^t when F is the true distribution of X. All these preliminaries go to define the *risk function* $r(F, \delta)$, which is the expected value of the loss plus that of the cost of sampling, given the decision function δ and true distribution F. The object of the statistician is to choose δ so as to minimize $r(F, \delta)$. However, the statistician controls only δ while F is, so to speak, controlled by Nature. The problem is therefore analogous to that of the zero-sum two-person game of von Neumann, with the difference that, whereas the statistician wishes to minimize $r(F, \delta)$, we can hardly say that Nature wishes to maximize it.

The problem in choosing δ arises from the fact that no δ will simultaneously minimize $r(F, \delta)$ for all F in Ω. Two alternatives are considered by Wald. (1) If F has a known prior probability distribution ξ in Ω (that is, if Nature's mixed strategy is known to the statistician), then the decision problem is solved by that δ which minimizes the average risk

$$\int_{\Omega} r(F, \delta) d\xi; \tag{1}$$

such a δ is called a *Bayes solution* corresponding to ξ. (2) If a prior distribution ξ of F does not exist or is not known, the statistician who, in the spirit of the theory of games, regards Nature as an opponent out to maximize $r(F, \delta)$, should adopt a *minimax solution* δ for which sup $r(F, \delta)$ with respect to F is as small as possible. Wald himself feels that this attitude is "perhaps not unreasonable."

Alternatives (1) and (2) represent extremes of optimism and pessimim respectively concerning the statistician's attitude toward Nature. Most statisticians will probably find themselves somewhere between the two extremes, and faced with the problem of how to utilize their rather vague feelings about the frequency with which various possible F's can be expected to occur. For this reason, this reviewer suspects that minimax solutions as such are likely to be of little interest in statistics. For example, on p. 142 Wald cites the minimax point estimate of the mean θ of a binomial variate; the corresponding risk function is a constant $r_0=[2(1+N^{1/2})]^{-2}$. The traditional (non-minimax) estimate has risk function $\theta(1-\theta)/N$. For large N the ratio of this to r_0 is near zero except in a small interval about $\theta=1/2$, where it is slightly greater than 1. It is very hard to believe in the superiority of the minimax estimate in this case, which is by no means unusual in its nature.

To those who are indifferent to minimax solutions the principal interest of the book will lie in the main theorem that, under very general conditions, the class $\mathcal{B}$ of all Bayes solutions is *essentially complete* in the following sense: for any decision function δ there exists a δ^* in $\mathcal{B}$ such that $r(F, \delta^*) \leqq r(F, \delta)$ for all F in Ω. (Mathematically, this theorem represents a highly nontrivial extension of the method of Lagrange multipliers in the calculus of variations.) There is obviously no loss involved in restricting the choice of a decision function to any essentially complete class, in particular to $\mathcal{B}$. But even a minimal essentially complete class will usually be so large that further reduction is necessary before the statistician can turn the problem of selecting a decision rule over to the experimenter. One criterion for reduction, the minimax principle, has already been mentioned. Other criteria exist (unbiasedness, invariance, and so on) but are not dealt with in the present volume.

The book makes effective use of the modern theory of measure and integration, and operates at a high level of rigor and abstraction. For this reason few statisticians will be prepared to read it, yet its ultimate liberating effect on statistical theory will be great. It is to be hoped that so rich and stimulating a book as this will reach an audience among mathematicians.

HERBERT ROBBINS

Introduction to the theory of algebraic functions of one variable. By C. Chevalley. (Mathematical Surveys, no. 6.) New York, American Mathematical Society, 1951. 12+188 pp. $4.00.

Here is algebra with a vengeance; algebraic austerity could go no