The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 1128–1132

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.

SIDNEY J. YAKOWITZ. *Mathematics of Adaptive Control Processes.* Elsevier, Modern Analytic and Computational Methods in Science and Mathematics, No. 14, 1969. xv + 158 pp. $11.00.

Review by S. ZACKS

University of New Mexico

1. General introduction. The term adaptive control processes had been initiated in the field of automatic self-adjusting control systems early in the 1950's. Later it was adopted by statisticians as a title for the general class of sequential decision problems in which the statistician is faced with an identical decision problem, which repeats at a specified sequence of epochs. At each decision epoch the statistician has to choose a "control" k from a specified set, K, of alternative controls. The choice of a control k_t at epoch t effects the stochastic law according to which the state of the system will be realized at the consecutive epoch, $t+1$; which in turn determines the yield (loss) of the system. The objective can be generally stated as that of choosing a sequence of controls $\{k_1, k_2, \cdots, k_N\}$ so that the total expected yield of the system will be maximized. The time span of the control process, which is reflected by the number N of control epochs, is called also the control or planning horizon. This planning horizon could be finite or infinite. In the infinite horizon case we generally consider the objective of maximizing the total expected discounted yield. The treatment of adaptive control problems of the kind described here involves generally the statistical problem of deciding at each control epoch what is the stochastic (distribution) law associated with each one of the specified controls. As we all know there are Bayes procedures of optimal adaptive control; empirical Bayes procedures, and other procedures of statistical control. Many of the statistical problems of current interest, like problems of optimal sequential design of experiments, quality control, etc. are problems belonging to the general class specified above. The existing literature on these problems is very extensive.

2. Review of the book content. The book under consideration is a very concise monograph on the subject of adaptive control processes. It contains five chapters and nine short appendices. Chapter 1 (only 4 pages) is a very short introduction to the main problem treated in the book. Chapter 2 (20 pages) provides the basic definitions of the concepts involved with control processes (like: state space, decision times, control set functions, policy, trajectory, loss function, control process, a solution, etc.). In Section 2.1 the author treats general loss functions, not necessarily separable (additive) ones. A solution to the control process is defined as a feasible control policy which minimizes the loss (over the whole trajectory). The discussion in Section 2.1 pertains to the deterministic model only. In Section 2.2 the author discusses the fundamental ideas of Dynamic Programming (D.P.) and proves the main theorem (Principle of Optimality of D.P.) for Markovian control processes (a non-stochastic case). Section 2.3 is devoted to some historical background on Dynamic Programming and control processes. Important references are cited (see in particular references [7, 8, 11, 12, 13 and 15] of Chapter 2). Chapter 3 of the book (22 pages) gives a concise development of the theoretical framework for problems of adaptive control processes. In this chapter the author treats the stochastic control problem of the type discussed in the introductory section of this review. The first section of Chapter 3 is devoted to the relevant elements of the adaptive control process, when the model is stochastic. It is shown also how to determine at each stage the appropriate probability distribution law of the trajectory, given the state (or the sample path) of the system at time t. The notion of Markovian adaptive control process is also defined. Section 3.2 presents the D.P. backward induction method of solving a Markovian truncated adaptive control problem (finite horizon). In Section 3.3 we find a discussion of the existence of (optimal) solutions. The main existence theorem for Markovian truncated adaptive control problems is proven (under the common conditions of compactness and continuity of the loss function). In Section 3.4 the author proves that the principle of optimality of D.P. yields (with probability one) an optimal solution to the adaptive control problem. Section 3.5 is devoted to a short historical summary of some important papers on adaptive control processes. The last two chapters (four and five) treat two important statistical problems by the theory of adaptive control processes. In Chapter 4 (22 pages) the celebrated Two Armed Bandit problem (TAB) is discussed. Chapter 5 (20 pages) is devoted to the statistical problem of pattern recognition. The TAB problem can be described in the following general terms. A statistician has to decide which one of two alternative experiments to employ. Experiment E_1 is associated with an observable random variable X having an (unknown) distribution function $F_1(x)$. Experiment E_2 is associated with a random variable Y having an (unknown) distribution function $F_2(y)$. It is assumed that F_1 and F_2 belong to a family $\mathscr{F}$ of distribution functions. If experiment E_1 is performed the gain is $G_1(X)$ and if E_2 is performed the gain is $G_2(Y)$. The problem is how to allocate a finite number, N, of trials among E_1 and E_2 so that the expected total gain will be maximized. We have formulated the TAB problems in more general terms than the classical formulation of the gambling problem associated with the Two Armed Bandit slot machine. The general formulation implies immediately that many experimental design and similar statistical decision problems fall under the category of TAB problems. In case both distribution functions $F_1(x)$ and $F_2(y)$ are known there is no statistical problem. All the N trials are allocated to the experiment with the larger expected gain. If $F_1(x)$ is known but $F_2(y)$ is unknown the problem is generally less difficult than the general TAB problem, and can be reduced to a class of problems called the One Armed Bandit (OAB) problems. Interesting studies on TAB problems have been performed in a Bayesian framework, for certain parametric cases. According to the classical version of a

TAB problem, we consider two slot machines. One machine yields a dollar gain with probability r and another machine yields a dollar gain with probability s. In a simplified version of this TAB problem the value of s is known but that of r is unknown. It is assumed that r has a prior beta distribution $\beta(p, q)$ on (0, 1). This is in a sense a simple version of an OAB problem. The statistician starts to play on the machine with the unknown probability of success, r. After each trial he decides whether to continue playing on this machine or switch to the other machine. Once the statistician switches to the machine with the known probability of success he is not allowed to switch back. In Section 4.1 of the book we find a complete D.P. formulation of the solution to the above simplified version of the classical TAB problem. In principle the D.P. method yields a solution to the general TAB problem or even to a multi-armed bandit problem (more than two possible experiments). There is no formulation in the book of the general D.P. algorithm for a TAB problem when both distributions (probabilities of success) are unknown. Sections 4.2 and 4.3 are devoted to the multi-armed bandit problem with an infinite number of possible trials, with discounted future gains. It is shown that a solution to the infinite horizon case can be approximated by a sequence of truncated (finite) cases, provided solutions are available to the truncated problems. Several theorems are given to assist in the search for the solution. In Section 4.5 we are provided with some historical outline of the studies on the TAB problem. In Chapter 5 we find a formulation of a pattern recognition machine problem (P.R.M.) in terms of adaptive control processes. This chapter contains five sections. In Section 5.1 the model of P.R.M. is discussed; and preliminary results from probability theory are presented. Section 5.2 treats the subject of supervised learning in the framework of adaptive control. The treatment is, as in previous problems, a Bayesian one. The results of Section 5.2 are generalized in Section 5.3. The problem of selecting a transducer for the P.R.M. is discussed in Section 5.4. Finally, Section 5.5 provides an historical background. In the Appendices the author proves several lemmas and presents a Computer Program and readout for a D.P. solution of a TAB problem, when one machine has known characteristics, and, for a supervised learning problem in a P.R.M. model.

3. Criticism. It is conspicuous that the present monograph is not written by a mathematical statistician or a probabilist. The treatment of the subject is somewhat crude. This is reflected especially in Chapter 4. Furthermore, most mathematical statisticians are familiar with the basic theory of sequential design and analysis of experiments, including the basic results on the TAB problem. The recent book of DeGroot [2] and the paper of Quisel [4] contain material on the TAB problem which is better presented and of a higher interest. Moreover, it is generally a simple matter to write the functional equations for the infinite or truncated Markovian ACP of the type discussed in the present book. In principle the D.P. method can yield, in most cases of interest, a numerical solution to the problem. The execution, however, may be very tedious and sometimes even impractical. One has to reflect upon the simple examples provided in the book (see Example 4.1) to realize how difficult it may be to provide a solution to a nontrivial ACP problem. Recent attempts to simplify the solution are directed, in cases of Markovian ACP problems, towards a solution of an analogous continuous version of the problem, which involves a generalized Stephan free boundary problem of partial differential equations (see Chernoff [1], Grigelionis and Shiryaev [3]). There is no evidence in the book on this current orientation in research.

I have found the notation in several places very cumbersome and uncommon. For example see (3.2) or page 69 line 2 f.b. There are several printing errors and an integral sign is missing in the expression of $dF(\bar{x}_{t'} \mid x, s, t)$ above (3.8) on page 33. The bibliography is generally good, and the computer programs available for the TAB and the P.R.M. problems may be very useful. In summary, it is recommended that the monograph of Prof. Yakowitz be read by students and researchers who study ACP problems; however, one should read the book carefully.

REFERENCES

[1] CHERNOFF, H. (1967). Optimal stochastic control. Tech. Rep. No. 29, Dept. of Statistics, Stanford Univ.

[2] DEGROOT, M. H. (1970). *Optimal Statistical Decisions.* McGraw Hill, New York.

[3] GRIGELIONIS, B. I. and SHIRYAEV, A. N. (1966). On Stephan's problem and optimal stopping rules for Markov processes. *Theor. Probability Appl.* **9** 541–558.

[4] QUISEL, K. (1965). Extensions of the two-armed bandit and related processes with on-line experimentation. Tech. Rep. No. 137, Inst. for Math. Studies in the Social Sciences, Stanford Univ.

M. T. WASAN, *Stochastic Approximation.* Cambridge Univ. Press, 1969. x+202 pp. $9.50.

Review by VÁCLAV DUPAČ

Charles University

This is much more a collection of material than a monograph in the proper sense. The author reproduces smaller or greater parts of a number of different papers, often word for word, without any critical comparison of different results or methods (although some of the results reproduced are more general than others and some of the methods more powerful than others).

Moreover, the book contains some serious oversights and a large number of misprints. In Section 1.3, e.g., Wolfowitz's proof of the Dvoretzky Theorem is reproduced; however, while Wolfowitz states the theorem and its extension together and proves it in three steps (preliminaries, convergence w.p.1, convergence i.m.s.), the author states the theorem and its extension separately, presenting the preliminary step of Wolfowitz's proof as a proof of the first theorem and the remaining two steps as a proof of the second one. Moreover, the extended theorem is formulated incorrectly, as it is the uniformity of the relation $\sum_{n=1}^{\infty} \gamma_n(\cdot) = \infty$ that should be required, not the uniformity of the inequality (4). Section 6.4, Confidence Intervals, is completely false, as is its starting formula (1). (This is