The Annals of Probability
1974, Vol. 2, No. 5, 964–968

BOOK REVIEWS

M. FRANK NORMAN, *Markov Processes and Learning Models.* Academic Press, New York, 1972, xiii + 274 pp. $ 15.00

Review by RADU THEODORESCU
Universite Laval

This book by Frank Norman is devoted to random processes suggested by certain learning models. It adds in a remarkable way and with a strong personal flavor to the family of existing books dealing with this subject.

As the author states himself (page xi) "no attempt is made to establish the psychological utility of these models." More a probabilist then a psychologist, as the author qualifies himself, he develops a very beautiful mathematical theory motivated by stochastic models for learning. The results, many of which are due to the author, are, at present, enriching the theory of a certain class of stochastic processes; their practical utility is to be tested by the future.

All the examples of learning models (l.m.'s) quoted in the introduction (Chapter 0) have the following structure (page 12). At the beginning of trial n, the subject is characterized by his *state of learning* X_n, which takes on values in a *state space* X. On this trial, an *event* E_n occurs, in accordance with a probability distribution $p(X_n, G) = P(E_n \in G \mid X_n)$ over subsets G of an *event space* E. This, in turn, effects a transformation $X_{n+1} = u(X_n, E_n)$ of state.

Now let $(X, \mathscr{B})$ and $(E, \mathscr{G})$ be measurable spaces, let p be a stochastic kernel on $X \times \mathscr{G}$, and let u be a transformation of $X \times E$ into X, measurable with respect to $\mathscr{B} \times \mathscr{G}$ and $\mathscr{B}$. Following M. Iosifescu (see, e.g., M. Iosifescu and R. Theodorescu, *Random Processes and Learning*, Springer (1969) page 63), we call $((X, \mathscr{B}), (E, \mathscr{G}), p, u)$ a (homogeneous) *random system with complete connections* (r.s.c.c.). Further, we call a sequence $\{X_n, E_n\}_{n \geq 0}$ of random vectors on a probability space $(\Omega, \mathscr{F}, P)$ an *associated stochastic process* if X_n and E_n take on values in $(X, \mathscr{B})$ and $(E, \mathscr{G})$, respectively, $X_{n+1} = u(X_n, E_n)$ and $P(E_n \in G \mid X_n, E_{n-1}, \cdots) = p(X_n, G)$ a.s., for each $G \in \mathscr{G}$.

On page 24 Norman writes: "The concept of r.s.c.c. may be regarded as a generalization and formalization of the notion of a stochastic l.m. Thus we will often call such a system a l.m. or simply a *model*." Consequently, for studying special cases of l.m.'s, first general r.s.c.c.'s or, in Norman's terminology, l.m.'s, are examined. In other words, Norman's main aim is to study r.s.c.c.'s = l.m.'s. It follows that Norman's results will form an interesting and important contribution to the theory of r.s.c.c.'s, which originated in 1935 with a paper by O. Onicescu and G. Mihoc.

Further, consider the following processes: $\{X_n\}_{n \geq 0}$, $\{E_n, X_{n+1}\}_{n \geq 0}$, $\{X_n, E_n\}_{n \geq 0}$, and $\{E_n\}_{n \geq 0}$ generated by an r.s.c.c. It is easily seen that the first three processes

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.* www.jstor.org

are Markovian, whereas the last is not Markovian. For instance the transition operator U of $\{X_n\}_{n\geq 0}$ is defined by $Uf(x) = \int_E p(x, de)f(u(x, e))$. The operator U was introduced for special cases of r.s.c.c.'s in 1937 by W. Doeblin and R. Fortet, and was at that time called by them the operator associated with an r.s.c.c. Norman, in stating his results, is stressing Markovian aspects. Since the origin of these results goes back to r.s.c.c.'s, in our opinion the title of the book is in a certain sense misleading. Of course, *de gustibus et coloribus non disputandum*, but it is a fact that due to the author's own preference for Markov processes the fundamental role of r.s.c.c.'s is glossed over in favor of Markov processes. It should be also remarked that mathematical psychology is only one possible field of application of the theory of r.s.c.c.'s.

The book is divided into four parts. The introduction (Chapter 0) is conceived in such a way as to acquaint the reader with models occurring in mathematical learning theory which have motivated the theoretical developments of Parts I and II. Part I (Chapters 1–6) is devoted to distance diminishing models, i.e., to r.s.c.c.'s with (X, d) a metric space, and such that the distance $d(u(x, e), u(y, e))$ tends, in a certain sense, to be less than the distance $d(x, y)$. Part II (Chapters 7–11) is concerned with slow learning. Two types of slow learning are considered. If the probability $P(X_{n+1} \neq X_n \mid X_n = x) = p(x, \{e: u(x, e) \neq x\})$ of any change in x is small, learning occurs *with small probability*. If, on the other hand, $|\Delta X_n|$ is small or, alternatively, $|u(x, e) - x|$ is small, then learning occurs *by small steps*. Chapters 12–16 of Part III deal mainly with applications of the results of Parts I and II to the special l.m.'s, described in the introduction, while Chapters 17 and 18 deal with other special cases of l.m.'s, the Wright model for population genetics and the Ehrenfest model for heat exchange. (The author has pointed out that the square roots in lines 13 and 22 on page 259 should be fourth roots.) The text is so organized that Part I and Part II are independent. Also, the chapters of Part III may be considered independent, except Chapter 13 which depends on Chapter 12.

Judging by its contents, this monograph is intended for mathematicians, mainly probabilists. It may also be used by other people with a good, if not excellent, mathematical background, who are involved in one way or another in mathematical learning theory.

The bibliography contains only works which have influenced the author, and which the author considers to be the most important works in this subject. However, the complete absence should be noted of any historical notes, as well as the absence of the names of O. Onicescu and G. Mihoc who are authors of the concept of chain with complete connections (1935), the concept which led to that of r.s.c.c. Norman should have outlined the part played by this concept, because it is this very concept which led him to such beautiful results!

Clearly written from the mathematical point of view, this book benefits also from Norman's agreeable and simple English *sans faux pas*.

Summing up, Norman's monograph represents an important contribution to the theory of a certain class of stochastic processes motivated by models for learning. It will certainly be well received and appreciated.

UNIVERSITE LAVAL
CITE UNIVERSITE
QUEBEC, CANADA

K. B. ATHREYA AND P. E. NEY, *Branching Processes.* Springer–Verlag, Berlin, 1972, 287 pp.

Review by JOHN A. WILLIAMSON
University of Colorado, Boulder

The authors in their preface state clearly their purpose in writing the book, "to give a unified treatment of the limit theory of branching processes." They hold to their purpose and generally accomplish their end.

The authors concentrate on the decade between the publication of Harris' book (T. E. Harris, *The Theory of Branching Processes* (1963)) and their own. Referring to Harris, the authors write, "only enough material is repeated to make the treatment essentially self-contained. For example, certain foundational questions on the construction of processes to which we have nothing new to add are not developed." The reader can find in this book the classical limit laws, most appearing in their sharpest form, as well as recent new results, and all are presented in a way that can only be called elegant. The emphasis in the book is on single-type processes with the first two chapters devoted to the Galton–Watson process and the next two to the continuous time Markov and Bellman–Harris age-dependent branching processes. Multi-type processes are discussed only in the fifth chapter where, appropriately, the reader is referred to Harris for several proofs.

The basic techniques of functional iteration, of martingales, of convex function bounding, of Taylor's Series remainder estimates, of renewal equation analysis, and of comparison lemma manipulation, together with refinements, are all there for the reader to master. The authors have an interest in, and great facility with, technique. The technical high point, not surprisingly, is their Chapter IV on the age-dependent process.

Having expressed a general enthusiasm for the book, I want next to attempt to identify that group of readers for whom the book will be of most interest, and then finally to discuss what I feel is one serious omission.

This book is definitely required reading for any person intending to attempt research in the branching processes area as well as for anyone who plans only to keep abreast of the literature in this field. Not only do the contents of the various chapters equip the reader for these undertakings, but the problem sets at the end of each chapter provide several open questions and many challenges that test whether the reader has grasped the techniques of the chapter. (Since