*The Annals of Mathematical Statistics*
1970, Vol. 41, No. 4, 1381–1383

# 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.*

M. IOSIFESCU AND R. THEODORESCU, *Random Processes and Learning*. Springer–Verlag, New York, 1969. x+304 pp. $17.00.

Review by M. FRANK NORMAN

*University of Pennsylvania*

This monograph consists of three long chapters: 1. A study of random sequences via the dependence coefficient. 2. Random systems with complete connections. 3. Learning.

The dependence coefficient on which Chapter 1 is based is Ibragimov's:

$$\phi(\mathscr{K}_1, \mathscr{K}_2) = \sup_{B \in \mathscr{K}_2}(\operatorname{ess\,sup}_{\omega \in \Omega} |P(B \mid \mathscr{K}_1)(\omega) - P(B)|),$$

where $\mathscr{K}_1$ and $\mathscr{K}_2$ are sub $\sigma$-algebras of a probability space $(\Omega, \mathscr{K}, P)$. This coefficient is used to formulate generalizations of the independence assumptions of a wide variety of classical limit theorems. For example, let $f_1, f_2, \cdots$ be a strictly stationary sequence of real random variables, let $\mathscr{K}_\Lambda$ be the $\sigma$-algebra generated by $\{f_i \colon i \in \Lambda\}$, and let

$$\phi(n) = \sup_{r \geqq 1} \phi(\mathscr{K}_{[1,r]}, \mathscr{K}_{[r+n,\infty)}).$$

Chapter 1 includes extensions, due to Iosifescu, of the Berry–Esséen theorem and the law of the iterated logarithm, in which the classical independence assumption is replaced by $\sum_{n=1}^{\infty} \phi^{\frac{1}{2}}(n) < \infty$ (plus $\phi(1) < 1$ in the latter theorem).

A random system with complete connections (indexed by the nonnegative integers) is a system $\{(W, \mathscr{W}), (X, \mathscr{X}), \{u_n\}_{n \geqq 0}, \{^nP\}_{n \geqq 0}\}$, where $(W, \mathscr{W})$ and $(X, \mathscr{X})$ are measurable spaces, $u_n$ is a measurable transformation from $W \times X$ into $W$, and $^nP$ is a transition probability function on $W \times \mathscr{X}$. Associated with such a system and a $w \in W$ are stochastic processes $\xi_1, \xi_2, \cdots$ and $\zeta_0, \zeta_1, \cdots$ in $X$ and $W$, respectively, such that $\zeta_0 = w$,

$$P(\xi_{n+1} \in A \mid \xi_n, \cdots, \xi_1) = {}^nP(\zeta_n; A),$$

and $\zeta_{n+1} = u_n(\zeta_n; \xi_{n+1})$ with probability 1. The system is *homogeneous* if $u_n$ and $^nP$ do not depend on $n$. The process $\zeta_n$ is Markovian and, in the homogeneous case, has stationary transition probabilities. The term "complete connections" apparently refers to the intricate stochastic interdependence of the variables $\xi_n$.

The learning models considered in Chapter 3 are special random systems with complete connections. In the context of learning theory, the "state" variable $\zeta_{n-1}$

characterizes an experimental subject at the beginning of trial $n$, $n \geqq 1$, and the "event" variable $\xi_n$ reflects the subject's response on this trial and its outcome (or reinforcement or payoff). Thus responses and outcomes are functions $f_n = f(\xi_n)$ of the event sequence, and, under suitable "ergodicity" conditions formulated in Section 2.1, the theory of Chapter 1 yields a number of limit theorems for these variables.

According to its preface, the book "... is intended for mathematicians with a good background in modern probability theory, as well as for applied people working in the field of learning." Unfortunately, most mathematical psychologists interested in learning do not know enough measure theory (for example) to cope with a book at this level. For properly prepared readers, the volume will be a rich source of interesting examples of and results concerning sequences of dependent real random variables and stochastic processes in abstract spaces. A great deal of this material is available only in periodicals, and out-of-the-way ones at that, so the book can hardly help finding a place in the libraries of specialists and institutions.

Granted the appeal of its subject, the monograph has serious deficiencies that greatly diminish its value. The most damaging of these is the fact that, for many of the topics treated, we are presented with a mosaic of almost verbatim cuttings from research papers. I noticed this particularly in the sections drawn from work on learning models. In addition to signaling a lack of integration of the papers considered, this state of affairs leads to some gross confusions. For example, the meaning of $\pi_j$ and $\beta_{ij}$ changes without warning in the middle of page 273. And Subparagraph 3.1.1.1.1 is hopelessly garbled by notational inconsistencies resulting from incomplete adaptation of the numbering of state and event variables in the source paper. In such cases the reader is forced back to the original articles, which are, of course, cited. It goes without saying that errors in these articles are faithfully reproduced, e.g., the incorrect Theorem 3.3.9 (see the second Kanal article mentioned on page 278 for the correction). In the same subparagraph the proof of Theorem 3.3.10 is incomplete, since there are other solutions to $F$'s functional equation than the ones considered. Finally, the estimate $\psi_1(n) = O(n^{-\frac{1}{4}})$ on page 31, which is taken from one of Iosifescu's own papers, appears to be incorrect.

Here are a number of relatively minor infelicities and mistakes. The assumption $\alpha < \infty$ on page 112 is redundant (see Case 1, page 129). In the context of Paragraph 2.1.3.3 it would be more appropriate to let $T_n(w)$ be the support of $Q^n(w; \cdot)$, rather than the set of its atoms. One cannot insure that $A_m$ on page 135 converges to $A$ as $m \to \infty$. The inequality at the bottom of page 140 is invalid if $k > 1$. The proof of Proposition 2.2.21 can be reduced to one line by changing the reference from page 184 of Loève (1963) to $A$. (i) on page 183. The comparison of the estimators $z_{(n)}$ and $\bar{z}_n$ on page 180 is not meaningful, since the relative magnitude of $n$ and $k_n$ is not constrained. All $||\cdot||$'s in Paragraph 3.1.3.2 should be $|\cdot|$'s. In the proof of Theorem 3.1.20 it is incorrectly assumed that finite state models are distance diminishing. "Induction on $n$" is superfluous in deriving (3.2.40) and similar equations. The asymptotic moment formula of Subparagraph 3.2.2.3.3 is valid under the condition $||C|| < 1$ which, for all the reader is told, may not be satisfied in any non-trivial model. (One suspects also that Theorem 2.2.8 may be vacuous.) Finally, there are a substantial number of typographical errors.

On the stylistic front, the presentation in Subsection 2.1.2 of Iosifescu's valuable work on ergodicity in inhomogeneous random systems with complete connections is marred by notational complexity. In many proofs one is overwhelmed with indices. Perhaps details of some of these proofs should have been given for the homogeneous case only. The notation in Paragraphs 2.3.2.2 and 3.1.1.2 is also obtrusive. Though the English is usually clear, there are a great many disconcerting *faux pas.*

It is hoped that the authors will have an opportunity to ameliorate these deficiencies in subsequent printings or editions, and thus move closer to fulfilling the promise of this unique volume.