Review

Reviewed Work(s): Mathematical Techniques of Applied Probability: by ; Vol. 1-Discrete Time Models: Basic Theory; by ; Vol. 2-Discrete Time Models: Techniques and Applications. by Jeffrey J. Hunter

Review by: Simon Tavaré

Source: *Journal of the American Statistical Association*, Vol. 80, No. 392 (Dec., 1985), pp. 1071-1072

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: https://www.jstor.org/stable/2288587

Accessed: 24-04-2020 13:41 UTC

function; and the difference between it and the limiting characteristic function is computable. It is natural to ask whether the difference between the corresponding distribution functions can be expressed in terms of the difference between the characteristic functions. This problem was considered 40 years ago by Esseen (1944) in his work on the error of approximation in the central limit theorem. His fundamental result was of the following form: Let F and G be two distribution functions and f and g be their characteristic functions. Then for arbitrary $T > 0$,

$$\sup_x |F(x) - G(x)| \leq C \int_{-T}^{T} \left| \frac{f(t) - g(t)}{t} \right| dt + D/T,$$

where C and D are certain positive constants.

The book contains a detailed survey of the most recent advances in this area. Esseen's original result required the existence of a bounded density for at least one of the distributions and the absolute integrability of $F - G$. Recent results place no restrictions on F and G but use a more general metric, the well-known Lévy metric, in the place of Esseen's uniform metric $\sup_x |F(x) - G(x)|$, to describe the distance in the space of distributions. For example, Zolotarev (1971) established a slightly more general version of the following theorem:

Let $L(F, G)$ be the Lévy distance between F and G. Then without any conditions, we have the universal bound

$$L(F, G) \leq \frac{2}{\pi} \int_0^T |f(t) - g(t)| \frac{dt}{t} + 2e \frac{\log T}{T},$$

for $T \geq 1.3$.

4. *Analytic Distribution Functions.* This is another topic of potential interest in mathematical statistics. The useful property of analytic distributions is that their values on an infinite set with a finite limit point uniquely determine the entire distribution. As a consequence, it is often possible to prove the convergence of a sequence of distributions by establishing it only on a domain consisting of an interval.

The analyticity of a distribution function F is related to the tail behavior of the corresponding characteristic function f. The classical result is that if $f(t) = 0[\exp(-c|t|)]$ for $|t| \to \infty$, then $F(x)$ has an extension $F(z)$ to the complex plane that is analytic in the strip $|Im\ z| < c$. A brief chapter describes recent advances.

5. *Infinitely Divisible Distributions.* A distribution function F is said to be infinitely divisible if for each integer $n \geq 1$ there exists a distribution function F_n such that F is the n-fold convolution of F_n. The characteristic function of such a distribution has a well-known canonical form, due to Lévy. During the decade following the publication of Lukacs's earlier book (1970), statisticians investigated the infinite divisibility of specific distributions arising in statistics. The current book has references to these results. Several authors, whose work is cited, have shown that the class of infinitely divisible distributions includes the Student-t distribution, and the lognormal, Pareto, von Mises (for certain parameter values), hyperbolic, and generalized inverse Gaussian distributions. It has also been shown that distributions with densities of the form $C \exp[\beta x - A \exp(\alpha x)]$, with $A > 0$, $\alpha\beta > 0$, are infinitely divisible. The double exponential extreme value distribution is of this type.

Probabilists with a strong interest in characteristic functions will also find much valuable material in the other chapters that have not been described here. The topics include infinite divisibility, unimodality, analytic characteristic functions, and ridge functions. As a probabilist, I am greatly pleased that such a comprehensive and readable book on this important subject is available, and I am sure that those with a primary interest in statistics will have a similar reaction.

SIMEON M. BERMAN
New York University

REFERENCES

Berman, S. M. (1975), "A New Characterization of Characteristic Functions of Absolutely Continuous Distributions," *Pacific Journal of Mathematics*, 58, 323–329.

Esseen, C. G. (1944), "Fourier Analysis of Distribution Functions," *Acta Mathematica*, 77, 1–125.

Heathcote, C. R., and Pitman, J. W. (1972), "An Inequality for Characteristic Functions," *Bulletin of the Australian Mathematical Society*, 16, 1–9.

Lévy, P. (1937), *Théorie de l'Addition des Variables Aléatoires*, Paris: Gauthier Villars.

Lukacs, E. (1970), *Characteristic Functions* (2nd ed.), London: Griffin.

Polya, G. (1949), "Remarks on Characteristic Functions," in *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics, and Probability*, Berkeley: University of California Press, pp. 115–123.

Shimizu, R. (1972), "On the Decomposition of Stable Characteristic Functions," *Annals of the Institute of Statistical Mathematics*, 24, 347–353.

Zolotarev, V. M. (1971), "Estimates of the Difference Between Distributions in the Lévy Metric," *Proceedings of the Steklov Institute of Mathematics*, 112, 232–240.

Mathematical Techniques of Applied Probability (Vol. 1—Discrete Time Models: Basic Theory; Vol. 2—Discrete Time Models: Techniques and Applications).

Jeffrey J. Hunter. New York: Academic Press, 1983. Vol. 1: xiii + 239 pp. $32.00; Vol. 2: xii + 286 pp. $35.00.

These two books are the first two parts of a three-volume treatise on mathematical techniques in applied probability. The series is intended as a text for a one-year first graduate course in stochastic processes and applied probability, the first two books being devoted to discrete time models.

After an introductory chapter on basic probabilistic concepts, Chapter 2 is devoted to a study of generating functions. The 40 or so pages here cover, inter alia, the continuity theorem for probability generating functions, some Abelian and Tauberian theorems, and applications to the solution of difference equations. I suspect that the student may find this rather daunting, since he will probably have little idea as to why such techniques are useful. This chapter seems best approached on a "need-to-know" basis, the instructor pointing out the appropriate parts when required in subsequent chapters.

It is in Chapter 3 that the meat of the subject starts; recurrent events and discrete renewal theory make their debut. Modulo the definition of a recurrent event, this chapter seems pretty standard, covering much the same selection of topics as Chapter 13 of Feller's (1968) classic text. The central result of this subject—the renewal theorem, which asserts that in the aperiodic persistent case, the probability u_n of a renewal at time n converges to the reciprocal of the mean recurrence time—is not proved.

The presence of Chapter 4, some 50 pages on matrix theory, certainly distinguishes Volume 1 from other textbooks in this area. The chapter covers, among other topics, material on generalized inverses and the Perron–Frobenius theorem for nonnegative matrices. The author suggests, and I would agree, that this material be used just for reference.

The early appearance of generating functions, discrete renewal theory, and matrix algebra gives us an accurate prediction of the way in which the author is going to treat the major topic of this work, that of discrete-time, discrete-state-space Markov chains. This approach starts in Chapter 5, the last of Volume 1. Chapter 5 covers classification of states, individual limit theorems for n-step transition probabilities (which are derived in the obvious way from the corresponding renewal theorem cited above), decomposition of the state space, and canonical forms for the transition matrix. The last section draws together in just two pages some connections between limiting and stationary distributions in the regular (finite state space) case. I found this last section inadequate, and a rather strange place to break for the next volume, since the first does not then stand on its own as a textbook.

Having come to the end of Volume 1, this is a good place to compare this author's approach to Markov chains with that of other textbooks in this area. Two features seem to distinguish such books: (a) the examples used to motivate the theory and (b) the way in which the convergence theorem for transition probabilities is approached. With respect to (a), I found the examples here rather flat and uninspiring—nowhere near as much fun as those in Karlin and Taylor (1975) or Ross (1983), for example. From a pedagogical point of view, it is with (b) that I have serious reservations about the present approach. This may be summarized as the classical "analytic" one, which for me removes much of the underlying probabilistic structure from the problem. It seems infinitely preferable to proceed via coupling arguments, in which the Markov chains themselves do the work. This approach, exemplified in Pitman's (1974) paper, and adopted in Grimmett and Stirzaker's (1982) textbook, has the added advantage of being technically simpler and complete. This "probabilistic" route provides as a by-product, almost gratis, proofs of the renewal theorem for recurrent events in both the delayed and ordinary settings.

Volume 2 has four chapters, the first two devoted to further topics in Markov chain theory. Chapter 6 describes analytic methods for calculating transition probabilities, first-passage distributions, and hitting probabilities. Chapter 7, of 100 pages, concentrates on the limiting behavior of transition probabilities. Among topics discussed in the first section are geometric ergodicity in the regular case (better attacked by the coupling argument rather than by the Perron–Frobenius theorem, I suspect), conditions for transience and recurrence in the countably infinite case, and a good discussion of periodicity. Section 2 focuses on methods for obtaining stationary distributions—it is here that generalized inverses make their entrance. There is no mention of computational methods for the infinite case or reversibility. The final two sections cover moments of first-passage time distributions and occupation times, giving an expanded treatment in the style of Kemeny and Snell's (1976) book.

Chapter 7 is somewhat misleadingly entitled "Applications of Discrete Time Markov Chains." It is just 22 pages long and describes the elementary properties of the Galton–Watson branching process. This is lifeless material, having neither the breadth of Karlin and Taylor's (1975) chapter or the diversity of examples of Taylor and Karlin's (1984). I suspect that the raison d'être for this part is that it is a convenient place to derive results needed for the final chapter, which is devoted to discrete-time queuing models. The flavor is once more analytic, involving embedded Markov chain methods. There seems to be little or no mention of the intriguing computational and probabilistic problems that arise in the context of network queues and their relatives.

I would not recommend these volumes as potential course books for a course in stochastic processes. The focus of the material is much too narrow, and the pace and style are likely to bury most students without trace. Our course is one semester long, and for many of the students in the class, it will be their one and only exposure to stochastic models at this level. We have to pick our way through the discrete Markovian minefield rather quickly to get on to Poisson processes, continuous-time Markov chains, renewal theory, and so on. Although these volumes may be a useful addendum to an instructor's bookshelves (containing as they do a large amount of detail), I would suggest something like Ross (1983), Taylor and Karlin (1975, 1984), or Grimmett and Stirzaker (1982) to provide a more lively and exhilarating introduction to the subject.

SIMON TAVARÉ
Colorado State University

REFERENCES

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications* (3rd ed.), New York: John Wiley.

Grimmett, G. R., and Stirzaker, D. R. (1982), *Probability and Random Processes*, New York: Oxford University Press.

Karlin, S., and Taylor, H. M. (1975), *A First Course in Stochastic Process* (2nd ed.), New York: Academic Press.

Kemeny, J. G., and Snell, J. L. (1976), *Finite Markov Chains*, New York: Springer-Verlag.

Pitman, J. W. (1974), "Uniform Rates of Convergence for Markov Chain Transition Probabilities," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 29, 193–227.

Ross, S. M. (1983), *Stochastic Processes*, New York: John Wiley.

Taylor, H. M., and Karlin, S. (1984), *An Introduction to Stochastic Modeling*, New York: Academic Press.

Approximation and Weak Convergence Methods for Random Processes, With Applications to Stochastic Systems Theory.

Harold J. Kushner. Cambridge, MA: MIT Press, 1984. xvii + 269 pp. $40.00.

There exist already several fine books on weak convergence, including the original treatises by Billingsley (1968) and Parthasarathy (1967) and the more recent monographs by Billingsley (1971), Kurtz (1981), Bergstrom (1982), and Pollard (1984), as well as the lecture notes by Aldous (1981), which seem to have had some circulation. So why yet another book on approximation and weak convergence methods for random processes, and why by Kushner, who has already published several books on related topics (stochastic approximation, stability and control)?

The answer is provided by the author himself, who offers the following assessment of his philosophy:

> The application of weak convergence ideas usually requires (first) that tightness be proved and (second) that the limit process be characterized. But these activities can often be reduced to a routine verification of conditions on the distributions of the system driving noise and system dynamics, and much of the rest of the book is devoted to both showing how this can be done and doing it on many examples. (p. 28)

For a demonstration of how this can be done, the reader is taken on a guided tour of the recent research of the author and his collaborators at Brown University, notably Huang Hai, which comprises most of Chapters 3–5 (or 107 pages). The so-called *perturbed test function* and *direct averaging* methods are explained, compared, and combined in great detail: "these methods were chosen (and developed) owing to their power and applicability in the modeling and approximation of the types of dynamical systems that arise in communication and control theory, as well as in other areas of applied physical science and engineering" (p. 34). Such systems are excited by random noise and are too complicated to permit an exact analysis, whence the need for approximations. All possible combinations are discussed: the noise can be state-dependent or not, the limiting process can satisfy an ordinary or a stochastic differential equation, the time might be continuous or discrete, and so forth. The reader's endurance may be put to the test, but the point concerning the "routine verification of conditions on the system driving noise and system dynamics" is amply and convincingly demonstrated. In particular, the perturbed test function method uses the kind of infinitesimal operator of the so-called *semigroup of conditional shifts* introduced by Kurtz (1975); coupled with criteria for tightness due to Kurtz (1975) and Aldous (1978), it provides a setting that is very well suited to the approximate study of non-Markovian processes. As an outgrowth of this technique, *perturbed Lyapounov functions* are employed in Chapter 6 to study the ergodic behavior of such processes.

The real merriment, though, starts when the methods are applied "on many examples," as promised (Chaps. 8–10, or 50 pages). Here we are treated to a wealth of applications: adaptive antenna arrays, adaptive equalizers and quantizers, stochastic approximation schemes, approximations of solutions to stochastic differential equations (the question of Wong and Zakai), hard limiters, and synchronization systems (phase-locked loops). Some of these, in particular the digital phase-locked loop, are communication devices very hard to analyze by any currently available alternative technique and had been heretofore treated only by heuristics. In fact, it is this part of the book that makes it unique and potentially appealing to such diverse audiences as communication and control engineers, physicists, and mathematical statisticians and probabilists.

In terms of mathematical sophistication, the requirements for the reader are rather severe: they include a working knowledge of stochastic integration, stochastic differential equations (including the martingale formulation of Stroock and Varadhan), and the theory of weak convergence for probability measures. A synopsis of this background material is given in Chapters 1 and 2 (33 pages), in a rather telegraphic style and primarily for purposes of handy reference. The exposition is illustrated by a large number of examples, but there are no exercises or problems. Finally, there are two interesting chapters on singular perturbations (7) and large deviations (11).

The author's writing style leaves some room for improvement and the notation can get quite heavy at times, but on the whole the exposition is readable and rather well motivated. There are a few typographical errors, which although rather obvious, never become annoying. The setting and production of the book by the MIT Press is very good. The only thing that one can cavil about concerns an unfortunate omission: Displayed on the top of each page are the titles of the corresponding chapter and section but not their numbers. However, all cross-references to theorems, lemmas, formulas, assumptions, and so forth, in a different chapter *do* require the number of the chapter (and sometimes of the section); the result is not always pleasant! The price is on the expensive side for the length of the text, but not unreasonable; it is definitely worth paying if you are interested in learning this material.

IOANNIS KARATZAS
Columbia University

REFERENCES

Aldous, D. J. (1978), "Stopping Times and Tightness," *Annals of Probability*, 6, 335–340.

——— (1981), "Weak Convergence and the General Theory of Processes," incomplete draft of monograph, University of California, Berkeley, Dept. of Statistics.

Bergstrom, H. (1982), *Weak Convergence of Measures*, New York: Academic Press.

Billingsley, P. (1968), *Convergence of Probability Measures*, New York: John Wiley.

——— (1971), *Weak Convergence of Measures: Applications in Probability*, Philadelphia: Society for Industrial and Applied Mathematics.

Kurtz, T. G. (1975), "Semigroups of Conditional Shifts and Approximations of Markov Processes," *Annals of Probability*, 4, 618–642.

——— (1981), *Approximation of Population Processes* (CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 36), Philadelphia: Society for Industrial and Applied Mathematics.

Parthasarathy, K. R. (1967), *Probability Measures on Metric Spaces*, New York: Academic Press.

Pollard, D. (1984), *Convergence of Stochastic Processes*, New York: Springer-Verlag.

Mathematics: People, Problems, Results (Vols. 1, 2, and 3).

Douglas M. Campbell and John C. Higgins (eds.). Belmont, CA: Wadsworth, 1984. xvi + 304 pp., iv + 275 pp., iv + 292 pp. $37.95 for the set (paperback).

This is an anthology of 90 essays with the avowed purpose of providing "an introduction to the spirit of mathematics . . . to give the non-mathematician some insight into the nature of mathematics and those who create it" (Vol. 1, p. iii). One can read such a collection as a set of individual parts or as a coherent whole. In this case the collection is so uneven and so chaotically organized ("deliberately untidy" is the editors' own description) that the entire value of the collection lies in making available under one cover a quorum of fine and insightful essays on various aspects of mathematics.

To make it clear at the outset, this anthology, regrettably, is not a successor to J. R. Newman's (1956) four-volume collection *The World of Mathematics*. Newman's collection of 133 essays has a coherence, a taste, and a timelessness born out of the 15 years of its compilation that is entirely lacking in the Campbell–Higgins collection. For example, whereas Newman began his anthology with P. E. B. Jourdain's 70-page essay on "The Nature of Mathematics," this collection begins with an invitation by the editors to contrast two brief extracts on the nature of Egyptian arithmetic; and only after 200