A Weak Convergence Approach to the Theory of Large Deviations

Review Author[s]:
James Lynch

Journal of the American Statistical Association, Vol. 93, No. 443 (Sep., 1998), 1244.

Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28199809%2993%3A443%3C1244%3AAWCATT%3E2.0.CO%3B2-L

Journal of the American Statistical Association is currently published by American Statistical Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/astata.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Sep 6 06:44:16 2004

Great emphasis is given to putting results in their historical perspective. Consequently, the book contains a fair number of short biographies on important probabilists from the early days (Fermat, Pascal, the Bernoullis, DeMoivre, Bayes, Laplace, Poisson, Markov, and so on). Further, an extensive set of exercises (together with solutions to many of them) help the reader in testing his or her newly acquired knowledge. Most of the exercises are fairly standard computationally, and some starred ones challenge the reader to go beyond the immediate application of certain rules.

I found no big surprises in the content; indeed, a fair number of books already contain similar material, either as part of a larger course on probability or as a full course. In the latter category, it is hard to beat the book by Feller (1968). The material covered in *Discrete Probability* comes close to that in the book by Stirzaker (1994). A textbook carrying the treatment of discrete stochastic models much further is the (not so elementary) one by Jacobs (1992).

Discrete Probability will definitely serve its main purpose as a postcalculus-level textbook for a first course in probability. The writing is rather casual but at the same time precise. Students who want to go on to a more advanced course on probability will already have seen the most important notation and models (although it is not clear to me why elements of the sample space Ω are denoted by u rather than the usual ω); those ending with this course will have a fair view of the very basics of probability. I can definitely recommend the book for its intended audience.

Paul EMBRECHTS
ETH, Zurich

REFERENCES

Feller, (1968), *An Introduction to Probability Theory and Its Applications* (Vol. 1, 3rd ed.), New York: Wiley.
Jacobs, K. (1992), *Discrete Stochastics*, Basel: Birkhäuser.
Stirzaker, D. (1994), *Elementary Probability*, Cambridge, U.K.: Cambridge University Press.

A Weak Convergence Approach to the Theory of Large Deviations.

Paul DUPUIS and Richard S. ELLIS. New York: Wiley, 1997. ISBN 0-471-07672-4. xvii + 479 pp. $69.95.

This is a very thoughtfully written book intended for theoretical statisticians who want to learn more about the theory as well as for experts of the theory. Weak convergence plays two fundamentally different roles in the authors' approach. First, analogous to the continuity theorem for convergence in distribution, they establish the equivalence of *Laplace's principle,*

$$\lim \frac{1}{n} \log E(\exp\{-nh(X_n)\}) = \Lambda(h) = \inf\{h(x) + I(x)\},$$

and the *large deviation principle,*

$$\limsup \frac{1}{n} \log P(X_n \in F) \leq -I(F), \quad F \text{ closed},$$

and

$$\liminf \frac{1}{n} \log P(X_n \in G) \geq -I(G), \quad G \text{ open},$$

where $I(A) = -\inf\{I(x)\colon x \in A\}$ is referred to as the *large deviation rate function.* The authors establish the large deviation principle in a variety of applications by determining that Laplace's principle holds. Formulating the theory in this fashion is pedagogically appealing, as there are many analogies between large deviation theory and the more-familiar weak convergence theory.

Second, to calculate the rate function $I(x)$ in these applications, the authors relate $W_n = (1/n)\log E(\exp\{-nh(X_n)\}$ to an optimal control problem. This novel relationship is used to obtain a representation of W_n that is amenable to traditional weak convergence results for determining the limit $\Lambda(h)$ and the rate function $I(x)$.

The general theory is presented in Chapter 1 and used to establish the large deviation principle and to determine the rate function for the empirical distribution function (Sanov's theorem) and the random-walk functional (Mogulskii's theorem) for iid random variables in Chapters 2 and 3. Extensions of Chapters 2 and 3 are presented for dependent random variables (Markov chains and Markov processes) in the remainder of the book.

I found *A Weak Convergence Approach to the Theory of Large Deviations* to be an excellent book with a unique perspective on one of the most active areas in probability theory today. Its self-contained development facilitates reading (and makes it appropriate for more mathematically oriented graduate students), and its control theory perspective is appealing from a research standpoint. This book is well worth reading if one has any interest in this theory.

James LYNCH
University of South Carolina

Modeling and Analysis of Stochastic Systems.

Vidyadhar G. KULKARNI. New York: Chapman and Hall, 1995. ISBN 0-412-04991-0. xiv + 619 pp. $119.95.

This text presents a thorough introduction to stochastic models and gives a comprehensive review of methods for analyzing such models. The author targets graduate students in the areas of operations research, computer science, engineering, and business and public policy. I agree with the author's intentions related to the first three areas, but find the book a bit too difficult for most students in graduate business programs.

The book emphasizes two aspects: modeling a "real-life" system with stochastic elements and analyzing the resulting model. It starts with (discrete-time) Markov chains, follows with Poisson processes and continuous-time Markov chains (hereinafter referred to as Markov processes) and their applications in queueing theory, and ends with renewal processes and their generalizations. This approach allows students to immediately build and analyze probabilistic models from diverse areas such as manufacturing, communications, and inventory systems. Furthermore, the analysis of Markov chain models is often based on matrix algebra and requires the use of numerical packages.

Once the students have digested Markov chain modeling and Poisson processes, they can proceed with the study of renewal processes. By now, they know how to use a variety of probabilistic tools, have seen the use of renewal results in Markov chain studies, and have studied a special case in Poisson process.

I have found that the alternative (and more theoretically appealing) approach that starts with renewal theory often confuses the students. The doctoral stochastic sequence in the School of Industrial and Systems Engineering at Georgia Tech is taken by students from a variety of disciplines, including operations research, industrial engineering, mathematics, and computer science.

The author chose to leave out some advanced topics, such as martingales and the Brownian motion process. Such topics need not be covered in an introductory course sequence but can be taught in follow-up courses from specialized texts (e.g., Harrison 1990). Overall, I agree with the author's choice of contents and presentation plan.

1. OVERVIEW

The text contains nine chapters and eight short appendixes. The appendixes list (without proof) the probabilistic results that are deemed necessary for reading the text.

Chapter 1 presents the fundamental issues involved in analyzing stochastic systems by means of a few easy-to-understand examples, including a single-server queue, a computer performance model, and a manufacturing example. The examples motivate the definition of the main performance measures for stochastic systems such as first passage times and limiting distributions combined with costs/rewards.

Chapter 2 presents Markov chains and focuses on their transient behavior. The first two sections contain the basic definitions and a battery of examples from diverse areas such as operations research, genetics, sociology, manpower planning, neurology, and telecommunications. The remainder of this chapter focuses on marginal distributions and methods for computing them. An impressive variety of modeling exercises ends this chapter.

Chapter 3 studies the limiting behavior of Markov chains. An important contribution of this chapter is the inclusion of Foster's criterion and Pake's lemma (for verifying positive recurrence); the use of these results is illustrated with a slotted ALOHA model. The author uses a simple brand-switching model to illustrate the limiting behavior of finite, irreducible Markov chains. This example demonstrates the nonexistence of a limiting distribution for a periodic chain and the role of the solution to the balance equations. Limiting results for reducible chains are given in a separate section. An entire section (10 pages) is allocated to computational issues. I would have appreciated the inclusion of MATLAB code—I often advise my students to create and use such routines. This chapter ends with a section presenting limiting results for discounted and average costs, a section focused on reversibility, and a section containing stochastic ordering results. The proofs of the limiting results for irreducible Markov chains are