The Annals of Probability
1986, Vol. 14, No. 4, 1428–1431

TWO BOOKS ON LARGE DEVIATIONS

RICHARD S. ELLIS, *Entropy, Large Deviations, and Statistical Mechanics*, Springer, New York, 1985, 368 pages, $54.00.

D. W. STROOCK, *An Introduction to the Theory of Large Deviations*, Springer, New York, 1984, 196 pages, $20.40.

REVIEW BY PETER NEY

University of Wisconsin

Large deviation theory is a part of probability limit theory which forms a natural complement to the laws of large numbers. Where the latter describes the convergence of a sequence of random elements to a mean or equilibrium state, the former focuses on the probabilities that the random quantities are in sets away from the central tendency. Under suitable moment hypotheses these probabilities decay exponentially fast. Interest centers on the exponents or rate functions, which are usually given by variational formulas involving an entropy function. A variety of mathematical techniques come into play, including convexity theory and Legendre transforms, asymptotic Laplace-type arguments, and various transformations or "twisting" of measures. Another source of stimulus to the theory has come from its interaction with other areas such as statistical inference, information theory, statistical physics, and stochastic dynamical systems.

Though there has been a steady history of contributions to the theory since the original paper of H. Cramér in 1938, large deviations has become an especially active research area during the last decade. Much of the impetus for this activity has come from the seminal work of M. Donsker and S. R. S. Varadhan [3] and their students in the United States, and from M. I. Freidlin and A. D. Wentzell in the Soviet Union (see, e.g., [4]).

Recently several monographs and books on the subject have been published. Two of these are the subject of this review.

As the title indicates, Richard Ellis' book is directed to the interface of large deviation theory with statistical mechanics. The relevance of large deviation ideas in statistical mechanics was extensively developed by O. Lanford in his 1971 Batelle Memorial Institute lectures [5]. In his book, Professor Ellis aims to bridge the gap between the physics and the contemporary mathematical developments. The book was this reader's first systematic exposure to the statistical mechanics models, and I feel that it accomplished its goal.

The book can be divided into three parts. The first (Chapters I and II) introduces the basic ideas of LD theory. Rate functions are explicitly calculated for a number of examples by combinatorial methods, which is useful for the beginner. The theory is then carefully developed at three levels. These would

Received March 1986.

deal, e.g., with LD's of sample means, empirical measures, and empirical processes, respectively. (I am not sure who first introduced this terminology.) Associated with these levels are variational formulas for the rate functions at different levels of generality. Sometimes a level n problem can be regarded as a level $n - 1$ problem on a more general state space. For example, the empirical measure of a random sequence taking values in a finite space (level 2) can be regarded as a vector valued random sequence (level 1). Relations between the levels are described via so-called contraction principles, which are also carefully discussed early in the book.

Most of the harder proofs are left to the third part (Chapters VI to IX). Many (though not all) theorems are limited to random sequences taking values in finite state spaces. This makes it possible to use fairly concrete computational arguments, and keeps most of the work at the mathematical level of a good basic probability course. What is gained is that some technical and complicated results are made fairly accessible here. On the other hand, some more sophisticated ideas needed to treat diffusions, Wentzel–Friedlin exit theory, general Markov processes, as well as the ideas requiring infinite dimensional spaces, are not developed.

The centerpiece of the book is the middle part (Chapters III, IV, V), which treats the statistical mechanics models: first the thermodynamics and ideal gas limit laws, then ferromagnetic models on $\mathbb{R}^d$ and higher dimensional lattices, and on the circle. Of course spontaneous magnetization and the existence of critical phenomena is central. Large deviation arguments permeate this treatment. The author's LD theorem for general sequences of random variables (Theorem II.6.1) and Varadhan's theorem on the asymptotics of expectations of random sequences satisfying a LD principle, are used extensively in the statistical mechanics limit theory. There are also some new results on central limit theory for "total spin," and some other limit theorems related to the Curie–Weiss model. As a newcomer to statistical mechanics, I found these chapters to be very enlightening.

This is a book that should be read with pencil and paper at one's side. There are lots of detailed calculations that the reader should be checking and lots of good exercises. Very helpful sections on notes and a complete bibliography are also included. It is clear that a great deal of thought and painstaking care has gone into the writing of this book. It is an important and very useful contribution to the literature on probability limit theory.

Daniel Stroock's monograph (based on lecture notes for a course given by the author), requires a more sophisticated background. The author uses all available machinery (of which there is obviously a considerable amount in his arsenal) to develop the ideas leading to current work on the large deviations of functionals of Markov processes.

As an introduction to the kind of results one is looking for in this subject he starts with the LD theorem for Brownian motion (Schilder's theorem), which is a limit theorem for Wiener measure parametrized so that the mass collapses to the unit measure concentrated on the zero function. This is a very pretty result, and the famous rate function $I(f) = \int_0^1 |\dot{f}(t)|^2 \, dt$ makes its appearance here. This chapter is rather independent of the sequel. One meets Schilder's theorem again later from another direction.

After a proof of Varadhan's theorem (on functionals of sequences satisfying the LDP, already mentioned above in the discussion of Ellis' book), the monograph goes on (in Chapter 3) to a very thorough treatment of LD theory for sums of i.i.d. random variables, first $\mathbb{R}^d$ valued, then Banach space valued. Many important techniques are introduced here: use of subadditivity to show existence of limit probabilities; approximation by half spaces in reducing certain problems to the one-dimensional cases; the identification of the rate function as the convex conjugate of the logarithmic generating function of the underlying random variable and the role of convexity theory; various other ideas that appear repeatedly throughout LD theory.

In the case of Gaussian random variables (on a vector space) the rate function can be more explicitly calculated in terms of the covariances. Specializing down to Brownian motion, we find that the i.i.d. theory provides another path to Schilder's theorem. By a general mapping argument, regarding a diffusion as a functional of Brownian motion, the LD theorem for the latter is translated into a corresponding theorem for diffusions (after disposing of some nontrivial technical difficulties). This opens the way to a study of the Wentzel–Friedlin theory which describes the exit behavior of a randomly perturbed dynamical system from a region.

These developments contain many elegant ideas and techniques of interest to probabilists. Chapters 2, 3, and 4 of the book (perhaps together with selected parts of the also excellent notes of R. Azencott [1], with which, as Professor Stroock points out, it has a number of ideas in common), would make a fine basis for a "second" graduate course or seminar in probability theory.

The remainder of the book treats more general large deviation theory related to ergodic phenomena. Chapter 5 gives a preliminary motivating discussion of the kinds of theorems and rate functions one should expect to find. It provides some illuminating heuristic discussion, and is an excellent motivating introduction to the Donsker–Varadhan theory of LD's of Markov processes.

The author then proceeds to approach this theory from a somewhat new point of view; namely via the subadditivity arguments used so successfully in the i.i.d. case in the Bahadur–Zabell [2] approach. This program works very nicely under a strong (uniform recurrence type) condition on the Markov process, and establishes the existence of a rate function, which is then identified as the convex conjugate of the spectral radius of an operator related to the transition mechanism of the Markov chain. This in turn is shown to be the same as the D–V rate function. In the absence of uniform ergodicity the subadditivity method does not seem to work. The LD theorem is proved under other hypotheses which are satisfied by a large class of processes, e.g., Ornstein–Uhlenbeck. The book concludes with a discussion of some new ideas on relations between the above large deviation theory and L. Gross's theory of logarithmic Sobolev inequalities. This approach offers the hope of interesting developments for infinite dimensional processes.

In summary, this monograph is an excellent (advanced level) introduction to recent developments in LD theory, particularly to the circle of ideas introduced by Bahadur and Zabell and Donsker and Varadhan. There are helpful heuristics and motivations, and new approaches to some aspects of the theory are explored.

All mathematics and statistics libraries, and most working probabilists, as well as researchers in related applications, will want to add both of the books under review to their collections.

REFERENCES

[1] AZENCOTT, R. (1980). Grandes deviations et applications. In *Ecole d'Eté de Prob. de St. Flour VIII–1978. Lecture Notes in Math.* **774** 1–176. Springer, Berlin.

[2] BAHADUR, R. R. and ZABELL, S. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probab.* **7** 587–621.

[3] DONSKER, M. and VARADHAN, S. R. S. (1975, 1976, 1983). Asymptotic evaluation of certain Markov process expectations for large time, I, II, III, IV. *Comm. Pure Appl. Math.* **28** 1–47, **28** 279–301, **29** 389–461, **36** 183–212.

[4] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems.* Springer, Berlin.

[5] LANFORD, O. E. (1973). Entropy and equilibrium states in classical statistical mechanics. *Lecture Notes in Phys.* **20** 1–113. Springer Berlin.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706