Stochastic Differential Equations: An Introduction With Applications.

Bernt Øksendal. New York: Springer-Verlag, 1985. xiii + 205 pp. $16.00 (paperback).

For some years now probabilists have been singing the praises of the theory of stochastic differential equations (SDE's) and insisting on their relevance and applicability. Several authors have attempted to give an introduction to the subject for the nonspecialist. Bernt Øksendal has joined the fray with what, I think, is the best attempt at an introductory course yet.

There are many problems with writing such a textbook, not least of which is the need to establish a degree of sophistication in probabilistic techniques in the reader.

There seem to me to be three ways of tackling this problem: assume that the reader knows some measure theory and establish the relevant results in probability theory within the text; assume some knowledge of measure theory and simply state relevant results and concepts in probability; or proceed in an intuitive fashion keeping definitions and technical details to a minimum.

Øksendal has adopted the second approach. My own initial preference was (in the case of a book aimed at getting to the applications as quickly as possible) for the third, but the literature contains so many counterintuitive examples that the second approach, on reflection, seems to be the acceptable compromise. It does, however, have its drawbacks. For example, Chapter 2 of this book moves in two pages from a definition of a random variable to Kolmogorov's extension theorem.

Realistically, I believe that anybody attempting to learn about this subject must accept that a knowledge of a first course in measure-theoretic probability theory is essential, although an ambitious reader may do well to attempt to acquire such background in parallel with the studying of this book by reading, for example, the first half of Chung's (1968) or Breiman's (1968) excellent works.

Chapter 1 presents six problems to motivate the theory contained in the text. Chapter 2, with some mathematical preliminaries, introduces probability theory, stochastic processes, and, in particular, Brownian motion. Chapter 3 is on the concept of the stochastic (Ito) integral and covers this topic with surprising ease and little fuss. The martingale concept is mentioned, but one must refer to Appendix B for a discussion of conditional expectations. Chapter 4 develops the stochastic integral and proves Ito's formula.

Chapter 5 introduces the key concept of SDE's, resisting the temptation to launch into a prolonged discussion of weak and strong (and strict) solutions. Chapter 6 gives the first application—to the filtering problem—and develops the Kalman–Bucy filter. Chapter 7 introduces Ito diffusions and gives a commendably short introduction to hitting distributions, infinitesimal generators, Dynkin's formula, the Feynman–Kac formula, and the Cameron–Martin–Girsanov formula.

Chapter 8 gives an application to the Dirichlet problem and Green's functions. Chapter 9 gives a discussion of optimal stopping for Ito diffusions and characterizes solutions as the least superharmonic majorant of the reward function. Chapter 10 introduces stochastic control, giving the Hamilton–Bellman–Jacobi equation and a converse—to my mind, the shortest and easiest exposition of this subject yet produced. The book concludes with appendixes on normal random variables and conditional expectations, a list of notation, references, and a good index.

Øksendal presents several examples aimed both at increasing the reader's technical skill and at showing applications. There are some infelicities in the English, but these do not impede the reader's understanding, and the few errors in the typing serve mainly as a check on comprehension.

To conclude, the book is to be highly recommended as an introduction to SDE's that should, with a little effort, be accessible both to those in search of applications in their own field and to graduate students who have completed a first course in measure-theoretic probability. For those who wish to go further, a suitable sequel might well be Elliott's (1985) more advanced textbook.

S. D. Jacka
University of Warwick
Coventry, United Kingdom

REFERENCES

Breiman, L. (1968), *Probability,* Menlo Park, CA: Addison-Wesley.
Chung, K. L. (1968), *A Course in Probability Theory,* New York: Harcourt, Brace & World.
Elliott, R. J. (1985), *Stochastic Calculus and Applications,* New York: Springer-Verlag.

Entropy, Large Deviations, and Statistical Mechanics.

Richard S. Ellis. New York: Springer-Verlag, 1985. xiv + 364 pp. $54.00.

The weak law of large numbers asserts that if S_n is the sum of n iid random variables, $X_1, \ldots, X_n$ say, then as $n \to \infty$, S_n/n converges in probability to $\mu = E(X_1)$. The theory of large deviations is concerned with the rate at which such convergence occurs. In this setting, a large deviation theorem would show that for any set A of possible values for S_n/n, with $\mu \notin A$, $\Pr(S_n/n \in A)$ decreases to 0 exponentially, at a rate that is a function of the set A. (This function, I say, is called the entropy function, and it may be calculated explicitly.) For large n, $\Pr(S_n/n \in A)$ is approximately $\exp(-nI(A))$. Note that (by the Borel–Cantelli lemma) such a result is stronger than the *strong* law of large numbers, and furthermore it gives extremely useful information (in fact, bounds) on probabilities that may well be of interest.

More generally, if an ergodic theorem or similar result guarantees the convergence of a sequence of random quantities to a constant, it is natural to ask about the rate at which the corresponding distributions converge to a point mass. For example, the preceding setting could be generalized by dropping the requirement of independence or by studying the sequence of *empirical measures* (the nth of which is obtained by placing mass n^{-1} at each of the observed values $X_1, \ldots, X_n$) that converges to the distribution of the X's. More generally still, the so-called *empirical process* is a (random) stationary measure on the space of *sequences,* which in the independent case will converge to a product measure; again one can ask about the rate at which this convergence occurs. This gives rise to three successively more general levels on which one can study convergence and prove large deviation results: partial sums (level 1), empirical measures (level 2), and empirical processes (level 3). Results at different levels are related by so-called contraction principles.

Statistical mechanics is concerned with the modeling of physical systems consisting of many constituent "particles." In the two cases covered in this book, the particles are the molecules of an ideal gas and the electrons (each occupying a fixed position) in a ferromagnet. With each particle is associated a random variable (the position and velocity of the gas molecule and the $\{-1, 1\}$-valued quantum mechanical spin of the electron) whose distribution is specified. In the gas, values for different molecules are independent, but a crucial facet of the ferromagnetic model is that there may be a dependence between the spins at different sites, usually a function of the distance between the sites. Individual values of these random variables are unobservable—that is, microscopic; one can only observe macroscopic properties of the system, such as energy, pressure, or overall magnetization. Interest centers on the relationship between these microscopic and macroscopic quantities and, in the ferromagnetic case, on the possible existence of phase transitions, that is, whether, in the limit as the number of particles becomes infinite, a particular microscopic specification gives rise to a unique macroscopic state. In the physical context, entropy is a measure of the number of microstates compatible with a particular macrostate. In spite of the book's title, large deviation theory and statistical mechanics are two quite different fields. There *is* a connection, historically, in that the concept of entropy is crucial in both settings, but more important, as the author so elegantly demonstrates, the theory of large deviations can be a very valuable tool in the analysis of statistical mechanical models.

The book's first chapter provides an introduction to large deviations, studying the case of iid random variables on a finite state space and using elementary combinatorial methods to prove level 1, 2, and 3 large deviation theorems and contraction principles. Chapter 2 extends this to the more general setting, stating large deviation principles for each level and, in particular, what is perhaps the most central result of the book, a large deviation principle first proved by the author for (not necessarily independent) random vectors taking values in $\mathbf{R}^d$. It goes on to give the relevant contraction principles, consequences of exponential convergence, and an extension of Laplace's method for the asymptotics of certain types of integral. The proofs of the main results are deferred to the final chapters of the book. Properties of convex functions and the Legendre–Fenchel transform are crucial; these are developed in Chapter 6. Ellis's level 1 large deviation principle is proved in Chapter 7, and Chapters 8 and 9 use this theorem to prove level 2 and level 3 results, respectively, for iid random vectors with a finite state space. The remaining chapters deal with statistical mechanics. The discrete ideal gas (in one dimension) is the subject of Chapter 3; the Maxwell–Boltzmann distribution, various thermodynamic equations, and the Gibbs variational formula and principle are natural consequences of large deviation results. Chapters 4 and 5 compose more than a third of the book and contain a detailed and complete discussion of ferromagnetic models, first on Z, then on Z^D and the circle, covering the Ising, Curie–Weiss, and more general models; finite and infinite volume Gibbs states; spontaneous magnetization; phase transitions; critical phenomena; and, where appropriate, central limit theorems. Correlation inequalities and large deviation results are the main tools used. The book has four appendixes, collecting some background results from probability and proving and further developing several theorems from the main body of the text. Some of the results and several of the proofs in the book are, in fact, new.

In his preface the author states that his goal has been "readability rather than completeness" (p. viii), and his style reflects this. As a consequence, someone familiar with the field may find the degree of repetition and the omissions irritating, but the uninitiated owe a considerable debt of gratitude for the relatively gentle way in which the book is developed. Ellis goes to considerable trouble to nurture intuition, often analyzing simple special cases before proceeding to more generality and outlining the general strategy before (and often after) giving detailed arguments. The reader is frequently reminded of definitions and notation, and these are usually restated in full when they appear in theorems. A moderate background in measure-theoretic probability and some basic topology is essential, but the book should be accessible to readers armed with little more than a thorough graduate course in probability. In keeping with his aims, the author also takes care to "remind" the reader of facts from other areas of mathematics (e.g., topology, real analysis, and complex analysis) when these are needed. No background in physics is required, although, particularly when reading the chapter on the ideal gas, I was left wishing that I did, in fact, know more about thermodynamics. (I am not sure whether this should be counted as a virtue or a failing of the book!)

Each chapter ends with a series of notes explaining connections with other work and, where appropriate, pointing to proofs of more general results, and a substantial number of problems. The author suggests that the book could be used to teach a graduate course in large deviations or possibly one in statistical mechanics, although in the latter case, at least with this development, various large deviation results would have to be taken on trust. I would have to agree, although the lack of completeness and the fact that results are only stated with as much generality as is needed are both an advantage and a disadvantage from an instructor's point of view. The book is certainly a very good place for individuals to learn about the three subjects in its title.

It is perhaps worth mentioning that the theory of large deviations encompasses considerably more than the study of the decay rate of noncentral probabilities for sequences of random variables or random measures. This was, however, the historical starting point for the theory, and it remains a sensible place from which to commence any study of the area. In contrast to the present work, the books by Stroock (1984) and Varadhan (1984) provide glimpses of the whole panorama of existing results. Each is written in the style of lecture notes, each requires a much wider degree of background knowledge, and each is set at a level harder than that of Ellis's book.

In conclusion, it should be said that this book is really two books, an introduction to the theory of large deviations and an introduction to equilibrium statistical mechanics—two different areas, each with its own beauty and its own flavor, but, as the author hopes, each enriched by the other. The book is, quite simply, enjoyable, stimulating, and rewarding. For most people it will not be essential reading, but it presents a wonderful opportunity to learn about and gain insight into either field—or both. For those who are at all interested, it is an opportunity that should not be lightly missed.

PETER DONNELLY
University College London

REFERENCES

Stroock, D. W. (1984), *An Introduction to the Theory of Large Deviations,* New York: Springer-Verlag.
Varadhan, S. R. S. (1984), *Large Deviations and Applications,* Philadelphia: Society for Industrial and Applied Mathematics.

Asymptotic Expansions for General Statistical Models.

J. Pfanzagl (with the asistance of W. Wefelmeyer). Berlin: Springer-Verlag, 1985. vii + 505 pp. $36.00 (paperback).

This book is based largely on the author's own work on second-order efficiency of tests and estimators in large samples. It is an important contribution to large-sample theory.

Fisher (1925), in a special context, recognized that there may be more than one efficient estimator: he proposed a measure of second-order efficiency among these. Rao (1961, 1963) carried Fisher's work much further. Modern treatments of the ideas of Fisher and Rao may be found in Efron (1975) and Ghosh and Subramanyam (1974). Consider the truncated squared error less $h(a) = \min\{a^2, d^2\}$ and expand $E_\theta h(\sqrt{n}(T_n - \theta))$ as $1/I(\theta) + n^{-1}\psi(\theta) + o(n^{-1})$. Here T_n is a first-order efficient estimator of a real-valued parameter θ, and $I(\theta)$ is the Fisher information per observation. The estimator T_n with the smallest ψ value is said to be second-order efficient by these authors. Ghosh and Subramanyam (1974) proved that in exponential families a bias-corrected maximum likelihood estimator (MLE) is second-order efficient. This assertion remains true if h is replaced by a bounded, bowl-shaped loss function that is smooth near 0. Appropriate extensions hold for vector-valued parameters.

A different criterion for efficiency may be based on the measure of concentration of T_n around θ. According to this approach, adopted by Pfanzagl, the estimator T_n of a parametric function $k(\theta)$ is (first-order) efficient if $P_\theta(\sqrt{n}(T_n - k(\theta)) \in I) \geq P_\theta(\sqrt{n}(T_n - k(\theta)) \in I) + o(1)$ for all intervals I that include 0 [or all symmetric measurable convex sets, in case $k(\theta)$ is vector valued]. The comparison is restricted to estimators that are median unbiased up to order $o(1)$. Moreover, some local uniformity conditions are imposed to exorcise the ghosts of superefficiency and so forth. The estimator T_n is second-order efficient if the error terms $o(1)$ are replaced by $o(n^{-1/2})$ throughout. Pfanzagl's theory of first-order and second-order efficient estimators goes far beyond the regular parametric models and covers many interesting nonparametric examples.

A major theme in the book is that first-order efficiency implies second-order efficiency. Pfanzagl (1973) proved this result under a locally uniform version of Cramér's condition. In the present monograph, he also refines a more widely applicable approach of Bickel, Chibishov, and van Zwet (1981). Locally uniform size restrictions up to $o(n^{-1/2})$ for tests and median unbiasedness for estimators already constrain a first-order efficient procedure to be second-order efficient. A clever trick of F. Goetze then allows the author to obtain an asymptotic expansion of the envelop power function, or of maximum concentration probabilities in case of estimation, up to order $o(n^{-1/2})$.

Lehmann (1983) pointed out that in case of iid observations X_j from $N(0, \theta)$, the MLE $\Sigma_1^n X_j^2/n$ is second-order efficient in the sense of Rao (1961, 1963) and Ghosh and Subramanyam (1974) but inadmissible in the second order when the loss is the squared error (p. 426). One may easily check that the estimator $(1-2/(3n))(\Sigma_1^n X_j^2/n)$ is median unbiased up to order $O(n^{-1})$ and second-order efficient by Pfanzagl's criterion, but it is still second-order inadmissible.

Finally, the title of the book is a little misleading. This is a book on second-order efficiency. Although this involves Cramér–Edgeworth expansions up to order $o(n^{-1/2})$, the monograph does not deal with other significant applications of these expansions to statistics.

In spite of its highly technical subject matter, the book is well written. The author provides many interesting examples to illustrate the theory. The present monograph and the author's earlier volume (Pfanzagl 1982) will undoubtedly have a major impact on current research, especially in the large-sample theory of semiparametric and nonparametric models.

R. N. BHATTACHARYA
Indiana University

REFERENCES

Bickel, P. J., Chibishov, D. M., and van Zwet, W. R. (1981), "On Efficiency of First and Second Order," *International Statistical Review,* 49, 169–175.
Efron, B. (1975), "Defining the Curvature of a Statistical Problem," *The Annals of Statistics,* 3, 1189–1242.
Fisher, R. A. (1925), "Theory of Statistical Estimation," *Proceedings of the Cambridge Philosophical Society,* 22, 700–725.
Ghosh, J. K., and Subramanyam, K. (1974), "Second Order Efficiency of Maximum Likelihood Estimators," *Sankhya,* Ser. A, 36, 325–358.
Lehmann, E. L. (1983), *Theory of Point Estimation,* New York: John Wiley.
Pfanzagl, J. (1973), "Asymptotic Expansions Related to Minimum Contrast Estimators," *The Annals of Statistics,* 1, 993–1026.
——— (1982), *Contributions to a General Asymptotic Statistical Theory* (*Lecture Notes in Statistics,* Vol. 13), New York: Springer-Verlag.
Rao, C. R. (1961), "Asymptotic Efficiency and Limiting Information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1), Berkeley, CA: University of California Press, pp. 531–546.
——— (1963), "Criteria of Estimation in Large Samples," *Sankhya,* Ser. A, 25, 189–206.

Modern Concepts and Theorems of Mathematical Statistics.

Edward B. Manoukian. New York: Springer-Verlag, 1986. xvi + 156 pp. $29.80.

This book is a reference book in mathematical statistics. The author writes in the Preface: