BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 49, Number 4, October 2012, Pages 597–601
S 0273-0979(2011)01357-X
Article electronically published on October 26, 2011

Large deviations for stochastic processes, by Jin Feng and Thomas G. Kurtz, Mathematical Surveys and Monographs, 131, American Mathematical Society, Providence, R.I., 2006, xii + 410 pp., US $99.00; All AMS Members: US $79.20, ISBN 978-0-8218-4145-7

The theory of large deviations deals with the estimation of small probabilities, particularly those that are exponentially small in some natural parameter. The general goal is to identify the constant in the exponent that dictates the exponential rate of decay. In many situations the constant can be "explicitly" calculated and turns out often to be characterized by a variational formula. It is not surprising because, if $\{P_n(\cdot)\}$ is a sequence of probability measures, for any two sets A, B even if they are not disjoint,

$$\lim_{n\to\infty}\frac{1}{n}\log P_n(A\cup B)=\max\{\lim_{n\to\infty}\frac{1}{n}\log P_n(A),\lim_{n\to\infty}\frac{1}{n}\log P_n(B)\}. \tag{0.1}$$

The following is a mathematically precise formulation. Let X be a complete separable metric space and $\{P_n\}$ a sequence of probability measures defined on the Borel subsets of X. We say that P_n satisfies the large deviation principle with a rate function $I(x) : X \to [0,\infty]$ if the following is valid:

- The function $I(x)$ is lower semi-continuous and the level sets

$$K_\ell = \{x : I(x) \le \ell\} \quad \text{are compact for every } \ell < \infty. \tag{0.2}$$

- For every closed set $C \in X$

$$\limsup_{n\to\infty}\frac{1}{n}\log P_n[C] \le -\inf_{x\in C} I(x). \tag{0.3}$$

- For every open set $G \in X$

$$\liminf_{n\to\infty}\frac{1}{n}\log P_n[G] \ge -\inf_{x\in G} I(x). \tag{0.4}$$

Together they imply that for any set A for which

$$\inf_{x\in A^o} I(x) = \inf_{x\in \bar{A}} I(x)$$

we have

$$\lim_{n\to\infty}\frac{1}{n}\log P_n[A] = -\inf_{x\in A} I(x). \tag{0.5}$$

One can think of $I(x)$ as the local rate of decay of P_n near x, and it can be recovered by

$$I(x) = \lim_{\epsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log P_n[B(x,\epsilon)],$$

where $B(x,\epsilon)$ is the ball around x of radius ϵ. From (0.1) and the covering property of compact sets, the local upper bound leads immediately to (0.3) for compact sets. To go from compact sets to closed sets one needs the following exponential tightness estimate.

2010 *Mathematics Subject Classification.* Primary 60F10; Secondary 60-02, 60J05, 60J25.

©2011 American Mathematical Society
Reverts to public domain 28 years from publication

Given any $\ell < \infty$, there is a compact set K_ℓ such that its complement K_ℓ^c satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[K_\ell^c] \le -\ell. \tag{0.6}$$

Upper bounds in probability are usually obtained through Tchebychef's inequality. If $F(x)$ is a continuous function and we can estimate

$$\limsup_{n\to\infty} \frac{1}{n} \log \int e^{nF(x)} dP_n \le \psi(F), \tag{0.7}$$

then by Tchebychef's inequality,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n[B(x,\epsilon)] \le -F(x) + \psi(F) + o(\epsilon),$$

which leads to

$$I(x) \ge \sup_F [F(x) - \psi(F)].$$

For most problems there is a natural class of function $\mathcal{F}$ for which one can estimate (0.7) and the supremum above is taken over $\mathcal{F}$.

The lower bounds are often obtained by "tilting". It is enough to show that for $x \in X$ with $I(x) < \infty$,

$$\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n[B(x,\epsilon)] \ge -I(x).$$

One can show this by exhibiting a sequence $Q_n \ll P_n$ with the property $Q_n \to \delta_x$ as $n \to \infty$ and

$$\limsup_{n\to\infty} \frac{1}{n} \log H(Q_n|P_n) \le I(x),$$

where

$$H(Q|P) = \int \log \frac{dQ}{dP}\, dQ = \int \frac{dQ}{dP} \log \frac{dQ}{dP}\, dP.$$

The proof is quite elementary. It follows from expressing for any neighborhood G of x, using Jensen's inequality

$$\begin{aligned} P_n(G) &= Q_n(G) \cdot \frac{1}{Q_n(G)} \int_G \exp\,[-\log \frac{dQ_n}{dP_n}] dQ_n \\ &\ge Q_n(G) \cdot \exp\,[-\frac{1}{Q_n(G)} \int_G \log \frac{dQ_n}{dP_n}] dQ_n. \end{aligned}$$

Another way to obtain the lower bound is through the Gärtner-Ellis theorem.

Theorem 0.1. *If*

$$\Phi(t) = \lim_{n\to\infty} \frac{1}{n} \log \int e^{tnF(x)} dP_n$$

exists and is differentiable at some $t = t_0$ *with* $\Phi'(t_0) = m$, *then*

$$\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log P_n[|F(x) - m| \le \epsilon] \ge -mt_0 + \Phi(t_0) = -\sup_t [mt - \Phi(t)].$$

The following facts follow easily from the definitions, provided (0.6) holds.

Theorem 0.2. *If* P_n *satisfies* (0.6), *then a large deviation principle with a rate function* $I(x)$ *holds if and only if, for any bounded continuous function* $F(x)$,

$$\lim_{n\to\infty} \frac{1}{n} \log \int e^{nF(x)}\, dP_n = \sup_x [F(x) - I(x)].$$

Theorem 0.3. *If $\{P_n\}$ satisfies* (0.6) *and for any $x \in X$ and any sequence $Q_n \to \delta_x$*

$$\liminf_{n\to\infty} H(Q_n|P_n) \geq I(x),$$

then the large deviation upper bound (0.3) *holds for P_n with the rate function $I(x)$.*

In order to establish the large deviation principle for a sequence $\{P_n\}$ with rate function $I(x)$, it is clearly sufficient to show the existence of the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \int e^{nF(x)}\, dP_n = \psi(F)$$

for a rich enough class of functions. This can sometimes be shown by establishing sub-additivity in n. If P_n is the distribution of the mean of n independent and identically distributed random variables

$$\Phi(t) = \frac{1}{n} \log \int e^{ntx}\, dP_n(x) = \log \int e^{tx} dP_1(x)$$

and as was proved by Cramér, a large deviation principle holds with

$$I(x) = \sup_t [tx - \Phi(t)].$$

The current volume is the latest among several books and monographs [2], [3], [4], [9] and [12] that describe aspects of this theory. It focuses on Markov processes. The approach of this particular monograph is best illustrated through an example. Consider on R^d a vector field

$$X = \sum_{j=1}^{d} b_j(x) \frac{\partial}{\partial x_j},$$

and perturb it by small noise. In other words consider the stochastic differential equation

$$dx_i(t) = b_i(x(t))dt + \sqrt{\epsilon} \sum_j \sigma_{i,j}(x(t)) d\beta_j(t); x(0) = x.$$

We are interested in the large deviation property of the distribution P_x^ϵ of the solution $x_\epsilon(t)$ at time $t = 1$. If we can evaluate

$$\Lambda(x, f) = \lim_{\epsilon\to 0} \epsilon \log E^{P_x^\epsilon}[e^{\frac{f(x(1))}{\epsilon}}]$$

and show that

$$\lambda(x, f) = \sup_y [f(y) - I_x(y)],$$

we have essentially established the large deviation principle for P_x^ϵ with rate function $I_x(y)$. The function

$$u(t, x) = \epsilon \log E^{P_x^\epsilon}[e^{\frac{f(x(1))}{\epsilon}}]$$

solves

$$u_t = \frac{\epsilon}{2} \sum_{i,j} a_{i,j}(x) u_{x_i,x_j j}(x) + \sum_{i,j} a_{i,j}(x) u_{x_i}(x) u_{x_j}(x) + \sum_j b_j(x) u_{x_j}(x)$$

with $u(0, x) = f(x)$. Here $a_{i,j}(x) = \sum_k \sigma_{i,k}(x)\sigma_{j,k}(x)$. Equivalently we need the limit as $\epsilon \to 0$ of $u^\epsilon(0, x)$ where

$$u_t^\epsilon + \frac{\epsilon}{2} \sum_{i,j} a_{i,j}(x) u^\epsilon_{x_i,x_j}(x) + \sum_{i,j} a_{i,j}(x) u^\epsilon_{x_i}(x) u^\epsilon_{x_j}(x) + \sum_j b_j(x) u^\epsilon_{x_j}(x) = 0$$

with $u^\epsilon(1,x) = f(x)$. The limit will be the solution $u(t,x)$ of the Hamilton-Jacobi equation

$$u_t + \sum_{i,j} a_{i,j}(x) u_{x_i}(x) u_{x_j}(x) + \sum_j b_j(x) u_{x_j}(x) = 0; u(1,x) = f(x)$$

evaluated at $t = 0$, given by

$$u(0,x) = \sup_{\substack{x(\cdot) \\ x(0)=x}} [f(x(1)) - \frac{1}{2}\int_0^1 \langle a^{-1}(\dot{x}(t) - b(x(t))), (\dot{x}(t) - b(x(t)))\rangle dt]$$

showing that the rate function is

$$I_x(y) = \frac{1}{2} \inf_{\substack{x(\cdot) \\ x(0)=x \\ x(1)=y}} [\int_0^1 \langle a^{-1}(\dot{x}(t) - b(x(t))), (\dot{x}(t) - b(x(t)))\rangle dt].$$

In fact if we consider the distribution Q_x^ϵ of the solution in the interval $[0,T]$ as a function of t, rather than just the value at time 1, it will converge weakly to the distribution with all its mass at the solution of the deterministic ODE

$$\dot{x} = b(x); x(0) = x.$$

There will be a large deviation result with rate function

$$I(x(\cdot)) = \frac{1}{2}\int_0^T \langle a^{-1}(\dot{x}(t) - b(x(t))), (\dot{x}(t) - b(x(t)))\rangle dt$$

provided $x(0) = x$. Such large deviation principles were first considered in Schilder [8] and Strassen [10] for Browninan motion, and in Varadhan [11] for diffusions with $b = 0$, and by Glass [6] and independently by Ventcel and Freidlin [13] for the general case. This idea started by Hopf [7] and Cole [1] and developed further by Fleming and Souganidis [5], is the general theme of this monograph. While the expectation satisfies a linear partial differential equation, the logarithm will satisfy a nonlinear equation. As $\epsilon \to 0$ we will get, in the limit, a first order nonlinear equation whose solution is given by a variational formula. The identification of the limit as the relevant solution of the limiting nonlinear equation is done through the theory of viscosity solutions.

The book is organized into three parts. The first part discusses the general theory of large deviations along with some of the basic notions. The second part studies the nonlinear semigroups that arise and establishes methods to prove their convergence. The limiting equations in general will fail to have classical solutions and the theory of viscosity solutions is then used as a tool to identify the limit. The last part is devoted to several classes of examples to which the methods developed in the second part are applied.

References

[1] Cole, Julian D. On a quasi-linear parabolic equation occurring in aerodynamics, Quart. Appl. Math, 9, 1951, 225–236. MR0042889 (13:178c)

[2] Dembo, Amir, Zeitouni, Ofer. Large deviations techniques and applications. Second edition. Applications of Mathematics (New York), 38. Springer-Verlag, New York, 1998. xvi+396 pp. ISBN: 0-387-98406-2. MR1619036 (99d:60030)

[3] Deuschel, Jean-Dominique; Stroock, Daniel W. Large deviations. Pure and Applied Mathematics, 137. Academic Press, Inc., Boston, MA, 1989. xiv+307 pp. MR997938 (90h:60026)

[4] Dupuis, Paul, Ellis, Richard S. A weak convergence approach to the theory of large deviations. (English summary) Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1997. xviii+479 pp. ISBN: 0-471-07672-4. MR1431744 (99f:60057)

[5] Fleming, W. H., Souganidis, P. E. A PDE approach to some large deviations problems. Nonlinear systems of partial differential equations in applied mathematics, Part 1 (Santa Fe, N.M., 1984), 441–447, Lectures in Appl. Math., 23, Amer. Math. Soc., Providence, RI, 1986. MR837690 (87i:60032)

[6] Glass, Michael Steven (1970). Perturbation of a first order equation by a small diffusion. Ph.D. dissertation, New York University, New York, NY. MR2619818

[7] Hopf, Eberhard. The partial differential equation $u_t + uu_x = \mu u_{xx}$. Comm. Pure Appl. Math. 3, (1950). 201–230. MR0047234 (13:846c)

[8] Schilder, M. Some asymptotic formulas for Wiener integrals. Trans. Amer. Math. Soc. 125 1966 63–85. MR0201892 (34:1770)

[9] Shwartz, Adam, Weiss, Alan. Large deviations for performance analysis. Queues, communications, and computing. With an appendix by Robert J. Vanderbei. Stochastic Modeling Series. Chapman & Hall, London, 1995. x+556 pp. ISBN: 0-412-06311-5. MR1335456 (96i:60029)

[10] Strassen, V. An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 3 1964 211–226 (1964). MR0175194 (30:5379)

[11] Varadhan, S. R. S. Diffusion processes in a small time interval. Comm. Pure Appl. Math. 20 1967 659–685. MR0217881 (36:970)

[12] Varadhan, S. R. S. Large deviations and applications. CBMS-NSF Regional Conference Series in Applied Mathematics, 46. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984. v+75 pp. ISBN: 0-89871-189-4. MR758258 (86h:60067b)

[13] Ventcel, A. D., Freidlin, M. I. Small random perturbations of dynamical systems. (Russian) Uspehi Mat. Nauk 25 1970 no. 1 (151), 3–55. MR0267221 (42:2123)

S. R. S. Varadhan
Courant Institute, New York University
E-mail address: varadhan@cims.nyu.edu