Varadhan's lemma for large deviations

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto

June 2, 2015

1 Large deviation principles

Let $(\mathscr{X}, d)$ be a Polish space. A function $f : \mathscr{X} \to [-\infty, \infty]$ is called **lower semicontinuous** if $c \in \mathbb{R}$ implies that $f^{-1}(c, \infty]$ is an open set in $\mathscr{X}$; equivalently, $c \in \mathbb{R}$ implies that $f^{-1}[-\infty, c]$ is a closed set in $\mathscr{X}$; equivalently, for any convergent sequence x_n in $\mathscr{X}$,[1]

$$f(\lim_{n\to\infty} x_n) \leq \liminf_n f(x_n).$$

If K is a nonempty compact subset of $\mathscr{X}$ and $f : \mathscr{X} \to [-\infty, \infty]$ is lower semicontinuous, there is some $x_0 \in K$ such that $f(x_0) = \inf_{x\in K} f(x)$.[2]

A **rate function** is a function $I : \mathscr{X} \to [0, \infty]$ such that (i) there is some $x \in \mathscr{X}$ with $I(x) < \infty$ and (ii) for any $c \in [0, \infty]$, $f^{-1}([0, c])$ is compact (namely, the sublevel sets of I are compact). Condition (ii) implies that a rate function is lower semicontinuous.

For $S \in \mathscr{B}_{\mathscr{X}}$, the Borel σ-algebra of $\mathscr{X}$, we define

$$I(S) = \inf_{x\in S} I(x).$$

Denote by $\mathscr{P}(\mathscr{X})$ the collection of Borel probability measures on $\mathscr{X}$ and let I be a rate function. A sequence $P_n \in \mathscr{P}(\mathscr{X})$ is said to **satisfy a large deviation principle with rate n and with rate function I** if for any closed subset C of $\mathscr{X}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(C) \leq -I(C)$$

and for any open subset U of $\mathscr{X}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n(U) \geq -I(U).$$

[1] http://individual.utoronto.ca/jordanbell/notes/semicontinuous.pdf, p. 2, Theorem 3.

[2] http://individual.utoronto.ca/jordanbell/notes/semicontinuous.pdf, p. 5, Theorem 8.

Unless we say otherwise, when we speak about a large deviation principle we shall use the rate n. Because $P_n(\mathcal{X}) = 1$ for each n, it follows that $I(\mathcal{X}) = 0$ and thus that $\inf_{x\in\mathcal{X}} I(x) = 0$.

We first establish that if a sequence of Borel probability measures satisfies a large deviation principle then its rate function is unique.[3]

Lemma 1. *If P_n satisfy a large deviation principle with rate functions I and J, then $I = J$.*

Proof. Let $x \in \mathcal{X}$ and let $B_N = B_{1/N}(x)$, the open ball with center x and radius $1/N$. Then, because $I(B_{N+1}) \leq I(\mathrm{cl}(B_{N+1}))$ and because $\mathrm{cl}(B_{N+1}) \subset B_N$ and so $I(\mathrm{cl}(B_{N+1})) \leq I(B_N)$,

$$\begin{aligned} -I(x) &\leq -I(B_{N+1}) \\ &\leq \liminf_{n\to\infty} \frac{1}{n} \log P_n(B_{N+1}) \\ &\leq -J(\mathrm{cl}(B_{N+1})) \\ &\leq -J(B_N). \end{aligned}$$

Because J is lower semicontinuous, as $N \to \infty$,

$$J(x) \leq \liminf_{N\to\infty} J(B_N),$$

hence

$$-I(x) \leq -J(x),$$

i.e. $J(x) \leq I(x)$. Likewise we obtain $I(x) \leq J(x)$, so $I(x) = J(x)$, for any $x \in \mathcal{X}$. □

2 Varadhan's lemma

For sequences α_n and β_n of positive real numbers, we write

$$\alpha_n \simeq \beta_n$$

if

$$\lim_{n\to\infty} \frac{1}{n}(\log \alpha_n - \log \beta_n) = 0.$$

Lemma 2. $\alpha_n + \beta_n \simeq \alpha \vee \beta_n$.

Lemma 3. *If A is a set and $f, g : A \to [-\infty, \infty]$ are functions, then*

$$\inf_A f - \inf_A g \leq \sup_A (f - g).$$

[3] Frank den Hollander, *Large Deviations*, p. 30,Theorem III.8.

Proof.

$$\inf_A g = \inf_A (g - f + f) \geq \inf_A (g - f) + \inf_A f,$$

hence

$$\inf_A f - \inf_A g \leq -\inf_A (g - f) = \sup_A (f - g).$$

□

We now prove **Varadhan's lemma**.[4]

Theorem 4 (Varadhan's lemma)**.** *Suppose that P_n satisfies a large deviation principle with rate function I. If $F : \mathcal{X} \to \mathbb{R}$ is continuous and bounded above, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \int_{\mathcal{X}} e^{nF(x)} dP_n(x) = \sup_{x\in\mathcal{X}} (F(x) - I(x)).$$

Proof. Let

$$b = \sup_{x\in\mathcal{X}} F(x), \qquad a = \sup_{x\in\mathcal{X}} (F(x) - I(x)).$$

Because F is bounded above, $-\infty < b < \infty$. Because I is not identically ∞, $-\infty < a \leq b$. For $n \geq 1$ and $S \in \mathscr{B}_{\mathcal{X}}$, define

$$J_n(S) = \int_S e^{nF(x)} dP_n(x).$$

Let $C = F^{-1}([a, b])$. For $N \geq 1$ and $0 \leq j \leq N$, define

$$c_{N,j} = a + \frac{j}{N}(b - a),$$

for which

$$[a, b] = \bigcup_{j=1}^{N} [c_{N,j-1}, c_{N,j}].$$

For $1 \leq j \leq N$, define

$$C_{N,j} = F^{-1}([c_{N,j-1}, c_{N,j}]),$$

for which

$$C = F^{-1}([a, b]) = \bigcup_{j=1}^{N} C_{N,j}.$$

Because F is continuous, each $C_{N,j}$ is closed, and hence applying the large deviation principle,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(C_{N,j}) \leq -I(C_{N,j}).$$

[4]Frank den Hollander, *Large Deviations*, p. 32, Theorem III.13.

For $x \in C_{N,j}$, $F(x) \leq c_{N,j}$, and hence, for each N,

$$J_n(C) = \int_C e^{nF(x)} dP_n(x) \leq \sum_{j=1}^{N} \int_{C_{N,j}} e^{nF(x)} dP_n(x) \leq \sum_{j=1}^{N} e^{nc_{N,j}} P_n(C_{N,j}).$$

Applying Lemma 2,

$$\lim_{n\to\infty} \frac{1}{n} \left(\log \left(\sum_{j=1}^{N} e^{nc_{N,j}} P_n(C_{N,j}) \right) - \log \left(\bigvee_{j=1}^{n} e^{nc_{N,j}} P_n(C_{N,j}) \right) \right) = 0,$$

i.e.

$$\lim_{n\to\infty} \frac{1}{n} \left(\log \left(\sum_{j=1}^{N} e^{nc_{N,j}} P_n(C_{N,j}) \right) - \bigvee_{j=1}^{N} \log(e^{nc_{N,j}} P_n(C_{N,j})) \right) = 0.$$

Then applying the large deviation principle,

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log J_n(C) &\leq \limsup_{n\to\infty} \frac{1}{n} \bigvee_{j=1}^{N} (nc_{N,j} + \log P_n(C_{N,j})) \\
&\leq \bigvee_{j=1}^{N} \left(c_{N,j} + \limsup_{n\to\infty} \frac{1}{n} \log P_n(C_{N,j}) \right) \\
&\leq \bigvee_{j=1}^{N} \left(c_{N,j} - I(C_{N,j}) \right).
\end{aligned}$$

For $x \in C_{N,j}$, $F(x) \geq c_{N,j-1}$, and as $c_{N,j} = c_{N,j-1} + \frac{1}{N}(b-a)$,

$$c_{N,j} \leq \inf_{x \in C_{N,j}} F(x) + \frac{1}{N}(b-a),$$

and using this and Lemma 3 we get

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log J_n(C) &\leq \bigvee_{j=1}^{N} \left(\inf_{x \in C_{N,j}} F(x) + \frac{1}{N}(b-a) - I(C_{N,j}) \right) \\
&= \frac{1}{N}(b-a) + \bigvee_{j=1}^{N} \left(\inf_{x \in C_{N,j}} F(x) - \inf_{x \in C_{N,j}} I(x) \right) \\
&\leq \frac{1}{N}(b-a) + \bigvee_{j=1}^{N} \sup_{x \in C_{N,j}} (F(x) - I(x)) \\
&= \frac{1}{N}(b-a) + \sup_{x \in C} (F(x) - I(x)) \\
&\leq \frac{1}{N}(b-a) + a.
\end{aligned}$$

Because this is true for all N,

$$\limsup_{n\to\infty} \frac{1}{n} \log J_n(C) \le a.$$

On the other hand, for $x \in \mathcal{X} \setminus C$ we have $F(x) < a$ and hence

$$J_n(\mathcal{X} \setminus C) = \int_{\mathcal{X}\setminus C} e^{nF(x)} dP_n(x) \le e^{na},$$

so

$$\limsup_{n\to\infty} \frac{1}{n} \log J_n(\mathcal{X} \setminus C) \le a.$$

Lemma 2 tells us that $J_n(C) + J_n(\mathcal{X} \setminus C) \simeq J_n(C) \vee J_n(\mathcal{X} \setminus C)$:

$$\lim_{n\to\infty} \frac{1}{n}(\log(J_n(C) + J_n(\mathcal{X} \setminus C)) - \log(J_n(C) \vee J_n(\mathcal{X} \setminus C))),$$

whence

$$\limsup_{n\to\infty} \frac{1}{n} \log J_n(\mathcal{X}) = \limsup_{n\to\infty} \frac{1}{n} \log(J_n(C) \vee J_n(\mathcal{X} \setminus C)) \le a,$$

i.e.

$$\limsup_{n\to\infty} \frac{1}{n} \log \int_{\mathcal{X}} e^{nF(x)} dP_n(x) \le \sup_{x\in\mathcal{X}} (F(x) - I(x)).$$

Let $x \in \mathcal{X}$ and let $\epsilon > 0$, and define

$$U_{x,\epsilon} = \{y \in \mathcal{X} : F(y) > F(x) - \epsilon\} = F^{-1}(F(x) - \epsilon, \infty),$$

which is an open subset of $\mathcal{X}$ because F is continuous. Applying the large deviation principle,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n(U_{x,\epsilon}) \ge -I(U_{x,\epsilon}).$$

Because $x \in U_{x,\epsilon}$, $I(U_{x,\epsilon}) \le I(x)$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P_n(U_{x,\epsilon}) \ge -I(x).$$

Together with

$$J_n(U_{x,\epsilon}) = \int_{U_{x,\epsilon}} e^{nF(y)} dP_n(y) \ge \int_{U_{x,\epsilon}} e^{n(F(x)-\epsilon)} dP_n(y) = e^{n(F(x)-\epsilon)} P_n(U_{x,\epsilon}),$$

this yields

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n} \log J_n(\mathcal{X}) &\ge \liminf_{n\to\infty} \frac{1}{n} \log J_n(U_{x,\epsilon}) \\
&\ge \liminf_{n\to\infty} \frac{1}{n} \log \left(e^{n(F(x)-\epsilon)} P_n(U_{x,\epsilon}) \right) \\
&= F(x) - \epsilon + \liminf_{n\to\infty} \frac{1}{n} \log P_n(U_{x,\epsilon}) \\
&\ge F(x) - \epsilon - I(x).
\end{aligned}$$

Because this is true for all $\epsilon > 0$,

$$\liminf_{n\to\infty} \frac{1}{n} \log J_n(\mathcal{X}) \geq F(x) - I(x).$$

Because this is true for all $x \in \mathcal{X}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log J_n(\mathcal{X}) \geq \sup_{x\in\mathcal{X}} (F(x) - I(x)),$$

i.e.,

$$\liminf_{n\to\infty} \frac{1}{n} \log \int_{\mathcal{X}} e^{nF(x)} dP_n(x) \geq \sup_{x\in\mathcal{X}} (F(x) - I(x)),$$

which completes the proof. $\square$

The following theorem states that if a sequence of Borel probability measures satisfies a large deviation principle then the **tilted** sequence of Borel probability measures satisfies a large deviation principle with a **tilted** rate function.[5]

Theorem 5. *Suppose that $P_n \in \mathscr{P}(\mathcal{X})$ satisfy a large deviation principle with rate function I, and let $F : \mathcal{X} \to \mathbb{R}$ be continuous and bounded above. Define for $n \geq 1$ and $S \in \mathscr{B}_{\mathcal{X}}$,*

$$J_n(S) = \int_S e^{nF(x)} dP_n(x).$$

Then the sequence $P_n^F \in \mathscr{P}(\mathcal{X})$ defined by

$$P_n^F(S) = \frac{J_n(S)}{J_n(\mathcal{X})}, \qquad S \in \mathscr{B}_{\mathcal{X}},$$

satisfies a large deviation principle with rate function

$$I^F(x) = \sup_{y\in\mathcal{X}} (F(y) - I(y)) - (F(x) - I(x)).$$

[5]Frank den Hollander, *Large Deviations*, p. 34,Theorem III.17.