

9 Beyond compactness: basic notions

9a	Large deviations principle (LDP)	73
9b	Contraction principle, and ‘tilted LDP’	75
9c	The probability decay rate	76
9d	Exponential tightness	76
9e	Inverse contraction principle	79
9f	Example: endless random walk	80

A simple example of a non-compact space is \mathbb{R} . Here is an instructive example of a seminorm on the Banach space $C_b(\mathbb{R})$ of all bounded continuous functions on \mathbb{R} ;

$$\|f\| = \limsup_{|x| \rightarrow \infty} |f(x)|.$$

It satisfies (2a1), (2a2) and (2a6), however, it is not of the form $\sup(|f| \Pi)$. We feel that it is situated at the points $\pm\infty$ of the extension $[-\infty, \infty]$ of \mathbb{R} , not on \mathbb{R} itself. We exclude such seminorms by requiring

$$f_k \downarrow 0 \text{ pointwise} \implies \|f_k\| \rightarrow 0$$

for all $f_1, f_2, \dots \in C_b(\mathbb{R})$.

9a Large deviations principle (LDP)

Let \mathcal{X} be a Polish space, that is, a separable topological space metrizable by a complete metric.

All bounded continuous functions $\mathcal{X} \rightarrow \mathbb{R}$ are a (generally, nonseparable) Banach space $C_b(\mathcal{X})$.

All (Borel) probability measures on \mathcal{X} are a set $P(\mathcal{X})$. Every $\mu \in P(\mathcal{X})$ gives us a linear functional $C_b(\mathcal{X}) \rightarrow \mathbb{R}$,

$$f \mapsto \int f \, d\mu.$$

The linear functional determines μ uniquely.

Let numbers $p_1, p_2, \dots \in [1, \infty)$ be given such that $p_n \rightarrow \infty$.

9a1 Definition. (a) A sequence $(\mu_n)_n$ of probability measures on a Polish space \mathcal{X} is *LD-convergent* with rate $(p_n)_n$, if the limit

$$\|f\|_{\text{lim}} = \lim_{n \rightarrow \infty} \left(\int |f|^{p_n} d\mu_n \right)^{1/p_n}$$

exists for all $f \in C_b(\mathcal{X})$, and

$$(9a2) \quad f_k \downarrow 0 \text{ pointwise} \implies \|f_k\|_{\text{lim}} \rightarrow 0$$

for all $f_k \in C_b(\mathcal{X})$.

(b) The sequence $(\mu_n)_n$ satisfies *LDP* with rate $(p_n)_n$ and rate function I (a function $\mathcal{X} \rightarrow [0, \infty]$ such that $I^{-1}([0, c])$ is compact for every $c < \infty$), if

$$\lim_{n \rightarrow \infty} \left(\int |f|^{p_n} d\mu_n \right)^{1/p_n} = \max_{x \in \mathcal{X}} (|f(x)|e^{-I(x)})$$

for all $f \in C_b(\mathcal{X})$.

If \mathcal{X} is compact then (9a2) is satisfied automatically (since it holds for the sup-norm). If \mathcal{X} is not compact then (9a2) is violated by the sup-norm (see 9a4 below).

By a rate function (on \mathcal{X}) we mean just a function $I : \mathcal{X} \rightarrow [0, \infty]$ such that $I^{-1}([0, c])$ is compact for every $c < \infty$. A compact set being always closed, a rate function is lower semicontinuous. (See also 9c1.)

On \mathbb{R} (or \mathbb{R}^d), a lower semicontinuous function I is a rate function if and only if $I(x) \rightarrow \infty$ as $x \rightarrow \pm\infty$ (think, why).

9a3 Exercise. Let I be a rate function (on \mathcal{X}). Then

- (a) I reaches its minimum on every closed set;
- (b) the maximum of $|f|e^{-I}$ on \mathcal{X} is reached for every $f \in C_b(\mathcal{X})$;
- (c) the seminorm $\|\cdot\|_I$ on $C_b(\mathcal{X})$ defined by

$$\|f\|_I = \max_{\mathcal{X}} (|f|e^{-I})$$

satisfies (2a1), (2a2), (2a6) and (9a2).

Prove it.

9a4 Exercise. Let $I : \mathcal{X} \rightarrow [0, \infty]$ be a lower semicontinuous function. If the seminorm $f \mapsto \sup_{\mathcal{X}} (|f|e^{-I})$ satisfies (9a2) then I is a rate function.

Prove it.

Hint: otherwise, take ε and $x_1, x_2, \dots \in \mathcal{X}$ such that $I(x_k) \leq \varepsilon$ and $\text{dist}(x_k, x_l) \geq 2\varepsilon$ whenever $k \neq l$; consider $f_n(x) = (1 - \frac{1}{\varepsilon} \text{dist}(x, \{x_n, x_{n+1}, \dots\}))^+$.

9a5 Exercise. Let $I_1, I_2 : \mathcal{X} \rightarrow [0, \infty]$ be lower semicontinuous. If $\sup_{\mathcal{X}}(|f|e^{-I_1}) = \sup_{\mathcal{X}}(|f|e^{-I_2})$ for all $f \in C_b(\mathcal{X})$ then $I_1 = I_2$.

Prove it.

Hint: similar to 2a12.

We generalize 2a11 and 2a14 as follows.

9a6 Proposition. Let a seminorm $\|\cdot\|$ on $C_b(\mathcal{X})$ satisfy (2a1), (2a2), (2a6) and (9a2). Then the function $I : \mathcal{X} \rightarrow [0, \infty]$ defined by

$$e^{I(x)} = \sup\{f(x) : \|f\| \leq 1\}$$

is a rate function, and

$$\|f\| = \max_{\mathcal{X}}(|f|e^{-I}) \quad \text{for all } f \in C_b(\mathcal{X}).$$

9a7 Exercise. Prove Proposition 9a6.

Hint: recall 4a. Given f and ε , find g_1, g_2, \dots such that $\|g_k\| \leq 1$ and $g_1 \vee g_2 \vee \dots > |f| - \varepsilon$ on \mathcal{X} . Apply (9a2) to the functions $(|f| - \varepsilon - g_1 \vee \dots \vee g_n)^+$.

9a8 Corollary. If $(\mu_n)_n$ is LD-convergent (with rate $(p_n)_n$) then $(\mu_n)_n$ satisfies LDP (with rate $(p_n)_n$) with one and only one rate function I , namely,

$$e^{I(x)} = \sup\{f(x) : \lim_{n \rightarrow \infty} \|f\|_{L_{p_n}(\mu_n)} \leq 1\}.$$

9a9 Exercise. Prove Corollary 9a8.

Similarly to 2a19,

$$(9a10) \quad \min_{x \in \mathcal{X}} I(x) = 0.$$

9b Contraction principle, and ‘tilted LDP’

Let $\mathcal{X}_1, \mathcal{X}_2$ be Polish spaces, $F : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ a continuous map, $(\mu_n)_n$ a sequence of probability measures on \mathcal{X}_1 , and $(\nu_n)_n$ its image on \mathcal{X}_2 (that is, $\nu_n(B) = \mu_n(F^{-1}(B))$ for Borel sets $B \subset \mathcal{X}_2$).

9b1 Theorem. (a) If $(\mu_n)_n$ is LD-convergent (with rate $(p_n)_n$), then $(\nu_n)_n$ is LD-convergent (with rate $(p_n)_n$).

(b) If $(\mu_n)_n$ satisfies LDP with rate $(p_n)_n$ and rate function I_1 , then $(\nu_n)_n$ satisfies LDP with rate $(p_n)_n$ and rate function I_2 defined by

$$I_2(y) = \min\{I_1(x) : x \in \mathcal{X}_1, F(x) = y\}.$$

If $F^{-1}(\{y\}) = \emptyset$ then the minimum is $+\infty$ by definition. Otherwise, the minimum is reached by 9a3(a).

9b2 Exercise. Prove Theorem 9b1.

Hint: similar to 2b2. And do not forget to prove that I_2 is a rate function.

9b3 Exercise. Generalize Theorem 2c1 to Polish spaces.

9c The probability decay rate

First, the notion of semicontinuity.

9c1 Exercise. Generalize 2a8, 2a9 to Polish spaces.

Hint: when proving (a) \implies (d), enforce $\varphi(\cdot) > 0$ by a transformation (say, $e^{\varphi(\cdot)}$), and then consider $f_n(x) = \max\{c : \forall y (\text{dist}(x, y) < c/n \implies \varphi(y) \geq c)\}$. You get continuous (but generally unbounded) functions.

Let $(\mu_n)_n$ satisfy LDP with rate $(p_n)_n$ and rate function I .

9c2 Exercise. Let $f : \mathcal{X} \rightarrow \mathbb{R}$.

(a) If $|f|$ is lower semicontinuous then

$$\liminf_n \|f\|_{L_{p_n}(\mu_n)} \geq \sup_{\mathcal{X}} (|f|e^{-I});$$

(b) if $|f|$ is bounded and upper semicontinuous then

$$\limsup_n \|f\|_{L_{p_n}(\mu_n)} \leq \max_{\mathcal{X}} (|f|e^{-I}).$$

Prove it.

Hints: (a): similar to 4b3(a);

(b): compactness is essential for 4b1(b), but the relation $\max_{\mathcal{X}}(f_je^{-I}) \downarrow \max_{\mathcal{X}}(|f|e^{-I})$ holds provided that f_1 is bounded.

9c3 Exercise. Generalize Corollaries 4b4 and 4b6 to Polish spaces.

9d Exponential tightness

First, the usual tightness.

9d1 Exercise. Let μ be a probability measure on \mathcal{X} . Then for every $\varepsilon > 0$ there exists a finite set $S \subset \mathcal{X}$ such that $\mu(S_{+\varepsilon}) \geq 1 - \varepsilon$. (Recall (4b9).)

Prove it.

Hint: take x_1, x_2, \dots dense in \mathcal{X} and observe that $\mu(\{x_1, \dots, x_n\}_{+\varepsilon}) \rightarrow 1$ as $n \rightarrow \infty$.

9d2 Exercise. Let μ be a probability measure on \mathcal{X} . Then for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that $\mu(K) \geq 1 - \varepsilon$.

Prove it.

Hint: take finite sets S_n such that $\sum_n (1 - \mu((S_n)_{+1/n})) \leq \varepsilon$ and consider $K = \bigcap_n (S_n)_{+1/n}$.

9d3 Exercise. The following three conditions on probability measures μ_1, μ_2, \dots on \mathcal{X} are equivalent:

(a) for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that

$$\sup_n (1 - \mu_n(K)) \leq \varepsilon;$$

(b) for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that

$$\limsup_{n \rightarrow \infty} (1 - \mu_n(K_{+\varepsilon})) \leq \varepsilon;$$

(c) for every $\varepsilon > 0$ there exists a finite set $S \subset \mathcal{X}$ such that

$$\sup_n (1 - \mu_n(S_{+\varepsilon})) \leq \varepsilon.$$

Prove it.

Hint: the implications (c) \implies (b) and (a) \implies (b) are trivial; using 9d1 it is not difficult to prove the implication (b) \implies (c); for proving the implication (c) \implies (a), do in the spirit of 9d2: take finite sets S_k such that $\sum_k \sup_n (1 - \mu_n((S_k)_{+1/k})) \leq \varepsilon$ and consider $K = \bigcap_k (S_k)_{+1/k}$.

9d4 Definition. A sequence $(\mu_n)_n$ of probability measures on \mathcal{X} is *tight*, if it satisfies the equivalent conditions 9d3(a)–(c).

You may add two more conditions to (b), (c) by choosing independently between \limsup and \sup on one hand, and between K and S on the other hand. You may also add one more condition to (a), replacing \sup with \limsup . This way you get $2 + 2 \cdot 2 = 6$ equivalent definitions of tightness!

The weak convergence of probability measures on \mathcal{X} is defined by

$$\mu_n \rightarrow \mu \iff \forall f \in C_b(\mathcal{X}) \int f d\mu_n \rightarrow \int f d\mu$$

for $\mu, \mu_n \in P(\mathcal{X})$.

9d5 Proposition. Every tight sequence contains a (weakly) convergent subsequence.

Proof. (sketch) If \mathcal{X} is compact then $C_b(\mathcal{X})$ is separable, and the diagonal argument works. In general, we take compact sets $K_i \subset \mathcal{X}$ such that $\mu_n(K_i) \geq 1/i$ for all n and i , and apply the said above to each K_i . Using the diagonal argument again we get a subsequence $(\mu_{n_k})_k$ such that the limit

$$\lim_{k \rightarrow \infty} \int_{K_i} f d\mu_{n_k}$$

exists for every $f \in C_b(\mathcal{X})$ and every i . However,

$$\int_{K_i} f d\mu_{n_k} \rightarrow \int_{\mathcal{X}} f d\mu_{n_k} \quad \text{as } i \rightarrow \infty$$

uniformly in k . □

In fact, a subset of $P(\mathcal{X})$ is tight if and only if its closure is (weakly) compact (Prohorov's theorem), but we do not need it.

Now we turn to *exponential* tightness.

9d6 Exercise. The following three conditions on probability measures μ_1, μ_2, \dots on \mathcal{X} are equivalent:

(a) for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that

$$\sup_n (1 - \mu_n(K))^{1/p_n} \leq \varepsilon;$$

(b) for every $\varepsilon > 0$ there exists a compact set $K \subset \mathcal{X}$ such that

$$\limsup_{n \rightarrow \infty} (1 - \mu_n(K_{+\varepsilon}))^{1/p_n} \leq \varepsilon;$$

(c) for every $\varepsilon > 0$ there exists a finite set $S \subset \mathcal{X}$ such that

$$\sup_n (1 - \mu_n(S_{+\varepsilon}))^{1/p_n} \leq \varepsilon.$$

Prove it.

Hint: similar to 9d3; (c) \implies (a): $(1 - \mu_n(K))^{1/p_n} \leq (\sum_k (1 - \mu_n((S_k)_{+1/k}))^{1/p_n} \leq \sum_k (1 - \mu_n((S_k)_{+1/k}))^{1/p_n}$.

9d7 Definition. A sequence $(\mu_n)_n$ of probability measures on \mathcal{X} is *exponentially tight* with rate $(p_n)_n$, if it satisfies the equivalent conditions 9d6(a)–(c).

Once again, you may get 6 equivalent definitions...

9d8 Exercise. Every LD-convergent (with rate $(p_n)_n$) sequence is exponentially tight (with rate $(p_n)_n$).

Prove it.

Hint: 9d6(b), and 4b4(b) via 9c3.

9d9 Exercise. Let $(\mu_n)_n$ be exponentially tight (with rate $(p_n)_n$), and the limit $\|f\|_{\text{lim}} = \lim_{n \rightarrow \infty} \|f\|_{L^{p_n}(\mu_n)}$ exists for all $f \in C_b(\mathcal{X})$. Then $\|f\|_{\text{lim}}$ satisfies (9a2), and therefore $(\mu_n)_n$ is LD-convergent (with rate $(p_n)_n$).

Prove it.

Hint: $\int_{\mathcal{X}} |f|^{p_n} d\mu_n = \int_K |f|^{p_n} d\mu_n + \int_{\mathcal{X} \setminus K} |f|^{p_n} d\mu_n \leq (\max_K |f|)^{p_n} + (\varepsilon \max_{\mathcal{X}} |f|)^{p_n}$.

9d10 Proposition. Let a sequence $(\mu_n)_n$ be exponentially tight with rate $(p_n)_n$. Then there exist $n_1 < n_2 < \dots$ such that the sequence $(\mu_{n_k})_k$ is LD-convergent with rate $(p_{n_k})_k$.

9d11 Exercise. Prove Proposition 9d10.

Hint: similar to (the proof of) Proposition 9d5, but consider $\|f\|_{L_{p_n}(\mu_n)}$ rather than $\int f d\mu_n$. And use 9d9.

9e Inverse contraction principle

Let $\mathcal{X}_1, \mathcal{X}_2$ be Polish spaces, $F : \mathcal{X}_1 \rightarrow \mathcal{X}_2$ a continuous map, $(\mu_n)_n$ a sequence of probability measures on \mathcal{X}_1 , and $(\nu_n)_n$ its image on \mathcal{X}_2 (that is, $\nu_n(B) = \mu_n(F^{-1}(B))$ for Borel sets $B \subset \mathcal{X}_2$).

9e1 Theorem. Assume that F is one-to-one and $(\mu_n)_n$ is exponentially tight (with rate $(p_n)_n$), then

(a) if $(\nu_n)_n$ is LD-convergent (with rate $(p_n)_n$), then $(\mu_n)_n$ is LD-convergent (with rate $(p_n)_n$);

(b) if $(\nu_n)_n$ satisfies LDP with rate $(p_n)_n$ and rate function I_2 , then $(\mu_n)_n$ satisfies LDP with rate $(p_n)_n$ and rate function I_1 defined by

$$I_1(x) = I_2(F(x)) \quad \text{for } x \in \mathcal{X}_1.$$

Proof. (a) Assume the contrary: $(\mu_n)_n$ is not LD-convergent. Using 9d9 we find $f \in C_b(\mathcal{X}_1)$ such that $\|f\|_{L_{p_n}(\mu_n)}$ does not converge (as $n \rightarrow \infty$). We choose $n_1 < n_2 < \dots$ and $n'_1 < n'_2 < \dots$ such that

$$\lim_k \|f\|_{L_{p_{n_k}}(\mu_{n_k})} \neq \lim_k \|f\|_{L_{p_{n'_k}}(\mu_{n'_k})}$$

(both limits exist, but differ). Using 9d10 we may assume that $(\mu_{n_k})_k$ is LD-convergent with rate $(p_{n_k})_k$, and $(\mu_{n'_k})_k$ is LD-convergent with rate $(p_{n'_k})_k$. The corresponding rate functions I_1, I'_1 on \mathcal{X}_1 differ, since $\max_{\mathcal{X}}(|f|e^{-I_1}) \neq \max_{\mathcal{X}}(|f|e^{-I'_1})$. By Theorem 9b1, I_1 satisfies $I_2(y) = \min\{I_1(x) : F(x) = y\}$, thus, $I_2(F(x)) = I_1(x)$ for all x . Similarly, $I_2(F(x)) = I'_1(x)$ for all x , therefore $I_1 = I_2$; a contradiction.

(b) The relation $I_1(\cdot) = I_2(F(\cdot))$ was verified when proving (a). \square

See also [1, Th. 4.2.4, p. 111]; [2, Lemma 3.12, p. 48].

9f Example: endless random walk

We return to the situation of 7a,

$$X_n\left(\frac{k}{n}\right) = \frac{s_1 + \cdots + s_k}{n} + v\frac{k}{n}, \quad \mathbb{P}(s_k = -1) = \mathbb{P}(s_k = +1) = \frac{1}{2},$$

$v \in (0, 1)$. The events

$$A_n : \quad \exists k \quad X_n\left(\frac{k}{n}\right) \leq -1$$

cannot be treated via 4b6, since the open set

$$\{w : \inf_t w(t) < -1\}$$

is dense in the corresponding compact space (recall 7a, after (7a1)). Indeed, any change of w after a large t is a small change. This is the product topology (recall (5b6)).

In terms of the process

$$Y_n\left(\frac{k}{n}\right) = \frac{s_1 + \cdots + s_k}{n},$$

related to X_n by $X_n(t) = Y_n(t) + vt$, we deal with the dense open set

$$\{w : \inf_t (w(t) + vt) < -1\} = \{w : \exists t \quad w(t) < -vt - 1\}$$

in the compact space denoted in 7e by $\text{Lip}(1)$.

Given a continuous function $h : [0, \infty) \rightarrow (0, \infty)$ such that $1 \ll h(t) \ll t$ for large t (that is, $h(t) \rightarrow \infty$ and $h(t)/t \rightarrow 0$ as $t \rightarrow \infty$), we introduce the set

$$\mathcal{X}_h = \{w \in \text{Lip}(1) : w(\cdot) = o(h(\cdot))\}$$

(that is, $w(t)/h(t) \rightarrow 0$ as $t \rightarrow \infty$) and equip it with the metric

$$\text{dist}(w, w') = \max_t \frac{|w(t) - w'(t)|}{h(t)}.$$

9f1 Exercise. \mathcal{X}_h is a Polish space.

Prove it.

Hint: \mathcal{X}_h is isometric to a closed subset of $C[0, \infty]$ (not just $C[0, \infty)$).

9f2 Exercise. The closure of the open set

$$G = \{w \in \mathcal{X}_h : \min_t (w(t) + vt) < -1\}$$

in \mathcal{X}_h is

$$\bar{G} = \{w \in \mathcal{X}_h : \min_t (w(t) + vt) \leq -1\}.$$

Prove it.

Hint: first, explain why the minimum is reached.

As before, we endow $\text{Lip}(1)$ with the topology of locally uniform convergence.

9f3 Exercise. The embedding $\mathcal{X}_h \rightarrow \text{Lip}(1)$ is continuous.

Prove it.

The distribution μ_n of the process Y_n is a probability measure on $\text{Lip}(1)$. We want to choose h such that $\mu_n(\mathcal{X}_h) = 1$ and moreover, $(\mu_n)_n$ is exponentially tight in \mathcal{X}_h (not just in $\text{Lip}(1)$).

9f4 Exercise. For every n and $c > 0$,

$$\mathbb{P}\left(\frac{s_1 + \cdots + s_n}{\sqrt{n}} \geq c\right) \leq \exp(-c^2/2).$$

Prove it.

Hint:

$$\mathbb{P}(s_1 + \cdots + s_n \geq c\sqrt{n}) \leq \frac{\mathbb{E} \exp(\lambda(s_1 + \cdots + s_n))}{\exp(\lambda c\sqrt{n})} \leq \exp\left(\frac{n\lambda^2}{2} - \lambda c\sqrt{n}\right)$$

for $\lambda > 0$; choose the optimal λ .

9f5 Exercise. Prove that¹

$$\limsup_{n \rightarrow \infty} \frac{s_1 + \cdots + s_n}{\sqrt{2n \ln n}} \leq 1 \quad \text{a.s.}$$

Hint: $\sum_n \mathbb{P}(s_1 + \cdots + s_n \geq c\sqrt{2n \ln n}) < \infty$ for $c > 1$.

¹In fact, by the law of the iterated logarithm,

$$\limsup_{n \rightarrow \infty} \frac{s_1 + \cdots + s_n}{\sqrt{2n \ln \ln n}} = 1 \quad \text{a.s.},$$

but we do not need it.

We get $\mu_n(\mathcal{X}_h) = 1$ provided that $\sqrt{t \ln t} = o(h(t))$, that is, $\frac{h(t)}{\sqrt{t \ln t}} \rightarrow \infty$ as $t \rightarrow \infty$.

9f6 Exercise. For every n and $c > 0$,

$$\mathbb{P}(\exists t \ Y_n(t) \geq c\sqrt{(t+1)\ln(t+2)}) \leq \frac{2^{-nc^2/2}}{1 - 2^{-c^2/2}}.$$

Prove it.

Hint:

$$\sum_{k=0}^{\infty} \exp\left(-\frac{1}{2}c^2(n+k)\ln\left(\frac{k}{n}+2\right)\right) \leq 2^{-nc^2/2} \sum_{k=0}^{\infty} 2^{-kc^2/2}.$$

9f7 Exercise. If $\frac{h(t)}{\sqrt{t \ln t}} \rightarrow \infty$ as $t \rightarrow \infty$ then $(\mu_n)_n$ is exponentially tight in \mathcal{X}_h .

Prove it.

Hint: $\{w \in \text{Lip}(1) : \forall t \ |w(t)| \leq c\sqrt{(t+1)\ln(t+2)}\}$ is a compact set in \mathcal{X}_h .

Combining 9f7, 9f3 and Theorem 9e1 we conclude that $(\mu_n)_n$ is LD-convergent in \mathcal{X}_h provided that $\frac{h(t)}{\sqrt{t \ln t}} \rightarrow \infty$ as $t \rightarrow \infty$. It satisfies LDP in \mathcal{X}_h with the rate function

$$J(w) = \int_0^{\infty} J_0(w'(t)) dt.$$

Finally, combining 9f2, 4b6 and 7a we get

$$(\mathbb{P}(A_n))^{1/n} \rightarrow \exp\left(-\min_{t>0} tJ_0\left(\frac{1}{t} + v\right)\right) \quad \text{as } n \rightarrow \infty.$$

References

- [1] A. Dembo, O. Zeitouni, *Large deviations techniques and applications*, Jones and Bartlett publ., 1993.
- [2] J. Feng, T.G. Kurtz, *Large deviations for stochastic processes*, 2005, <http://www.math.wisc.edu/~kurtz/feng/ldp.htm>