4 More on the basic notions

4a Every LD-convergent sequence satisfies LDP

Here we prove Prop. 2a11 for every seminorm $\|\cdot\|$ on $C(K)$ that satisfies (2a1), (2a2) and (2a6).

First, let $f \in C(K)$ satisfy $\|f\| \le 1$; we have to prove that $\max(|f|\Pi) \le 1$, that is, $|f(x)|\Pi(x) \le 1$ for all x. However, $1/\Pi(x) \ge |f(x)|$ by (2a7).

Second, let $f \in C(K)$ satisfy $\max(|f|\Pi) \le 1$; we have to prove that $\|f\| \le 1$. By (2a7),

$$|f(x)| \le \frac{1}{\Pi(x)} = \sup\{g(x) : \|g\| \le 1\}$$

for every $x \in K$.

Let $\varepsilon > 0$ be given. For every $x \in K$ there exists $g \in C(K)$ such that $\|g\| \le 1$ and $g(x) > |f(x)| - \varepsilon$. The inequality still holds on some neighborhood of x. By compactness we may cover K by a finite number of such neighborhoods. In other words, we have $g_1, \dots, g_n \in C(K)$ such that $\|g_1\| \le 1, \dots, \|g_n\| \le 1$ and $g_1 \vee \dots \vee g_n > |f| - \varepsilon$ on K. By (2a6), $\|g_1 \vee \dots \vee g_n\| \le \|g_1\| \vee \dots \vee \|g_n\| \le 1$. By (2a1) and (2a2), $\|f\| \le 1 + \varepsilon$ for every $\varepsilon > 0$, which completes the proof.

4b The probability decay rate

We deal with a compact metrizable space K and probability measures μ_n on K satisfying LDP with a rate function I. However, our first lemma does not use μ_n (and its first item does not use the topology of K).

4b1 Lemma. (a) Let $\varphi_n, \varphi : K \to \mathbb{R}$, $\varphi_n \uparrow \varphi$ pointwise; then $(\sup_K \varphi_n) \uparrow (\sup_K \varphi)$ as $n \to \infty$.

(b) Let $\varphi_n, \varphi : K \to \mathbb{R}$ be upper semicontinuous, and $\varphi_n \downarrow \varphi$ pointwise; then $(\max_K \varphi_n) \downarrow (\max_K \varphi)$ as $n \to \infty$.

Proof. (a) For every $\varepsilon > 0$ we take $x \in K$ such that $\varphi(x) > (\sup_K \varphi) - \varepsilon$ and n such that $\varphi_n(x) > (\sup_K \varphi) - \varepsilon$; then $(\sup_K \varphi_n) > (\sup_K \varphi) - \varepsilon$.

(b) We have $(\max_K \varphi_n) \downarrow c$ for some $c \in \mathbb{R}$. For every $\varepsilon > 0$ the sets $\{x \in K : \varphi_n(x) \ge c - \varepsilon\}$ are a decreasing sequence of nonempty closed sets. By compactness, some x belongs to all these sets. Thus, $\varphi(x) = \lim_n \varphi_n(x) \ge c - \varepsilon$ and $\max_K \varphi \ge c - \varepsilon$. □

4b2 Exercise. Without the semicontinuity 4b1(b) need not hold.

Find a counterexample.

4b3 Lemma. Let $f : K \to \mathbb{R}$.

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

$$\liminf_n \|f\|_{L_n(\mu_n)} \ge \sup_K (|f| \mathrm{e}^{-I}) ;$$

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

$$\limsup_n \|f\|_{L_n(\mu_n)} \le \max_K (|f| \mathrm{e}^{-I}) .$$

Proof. (a) We take $f_n \in C(K)$ such that $0 \le f_n \uparrow |f|$. For every j,

$$\liminf_n \|f\|_{L_n(\mu_n)} \ge \liminf_n \|f_j\|_{L_n(\mu_n)} = \max_K \big(f_j \mathrm{e}^{-I} \big) ;$$

however,

$$\max_K \big(f_j \mathrm{e}^{-I} \big) \uparrow \sup_K \big(|f| \mathrm{e}^{-I} \big) \quad \text{as } j \to \infty$$

by 4b1(a).

(b): similar (but using semicontinuity). □

4b4 Corollary.

(a) $$\liminf_n \big(\mu_n(G)\big)^{1/n} \ge \exp(-\inf_G I) \quad \text{for every open } G \subset K ,$$

(b) $$\limsup_n \big(\mu_n(F)\big)^{1/n} \le \exp(-\min_F I) \quad \text{for every closed } F \subset K .$$

4b5 Exercise. Reconsider 2a18 and 2a20 in the light of 4b4.

4b6 Corollary. If an open set $G \subset K$ satisfies

(a) $$\inf_G I = \min_{\overline{G}} I$$

then

(b) $$\lim_n \big(\mu_n(G)\big)^{1/n} = \lim_n \big(\mu_n(\overline{G})\big)^{1/n} = \exp\big(-\inf_G I\big) = \exp\big(-\min_{\overline{G}} I\big) ,$$

that is,

(c) $$\lim_n \frac{1}{n} \ln \mu_n(G) = \lim_n \frac{1}{n} \ln \mu_n(\overline{G}) = -\inf_G I = -\min_{\overline{G}} I .$$

4b7 Exercise. 4b6(a) does not imply $\mu_n(\overline{G}) \sim \mu_n(G)$ as $n \to \infty$.

Find a counterexample.

Hint: try $K = [0,1]$, $G = (0,1]$, combine μ_n from Lebesgue measure and an atom at 0, and find appropriate coefficients.

Continuity of I is, of course, sufficient for 4b6(a). Here is a weaker sufficient condition:

$$\limsup_{y \to x, y \in G} I(y) \le I(x) \quad \text{for all } x \in \partial G \,. \tag{4b8}$$

We may also consider $\mu_n(A_n)$ assuming that A_n converge to G in an appropriate sense. To this end we choose a metric on K and, given a set $A \subset K$, we introduce (for any $\varepsilon > 0$)

$$A_{+\varepsilon} = \{ x \in K : \operatorname{dist}(x, A) \le \varepsilon \} \,, \tag{4b9}$$

$$A_{-\varepsilon} = \{ x \in K : \operatorname{dist}(x, \complement A) > \varepsilon \} \,; \tag{4b10}$$

here $\operatorname{dist}(x,A) = \inf_{y \in A} \operatorname{dist}(x,y)$, and $\complement A = \{ x \in K : x \notin A \}$. Note that $A_{+\varepsilon}$ is closed, $A_{-\varepsilon}$ is open, $\cap_\varepsilon A_{+\varepsilon} = \overline{A}$ is the closure of A, and $\cup_\varepsilon A_{-\varepsilon} = A^\circ$ is the interior of A.

4b11 Exercise. Let $A_n \subset K$ be such that A_n is μ_n-measurable.

(a) Let $G \subset K$ be an open set, and

$$A_n \supset G_{-\varepsilon_n} \quad \text{for some } \varepsilon_n \downarrow 0 \,.$$

Then

$$\liminf_n \big(\mu_n(A_n) \big)^{1/n} \ge \exp(- \inf_G I) \,.$$

(b) Let $F \subset K$ be a closed set, and

$$A_n \subset F_{+\varepsilon_n} \quad \text{for some } \varepsilon_n \downarrow 0 \,.$$

Then

$$\limsup_n \big(\mu_n(A_n) \big)^{1/n} \le \exp(- \min_F I) \,.$$

(c) Let $G \subset K$ be an open set such that $\inf_G I = \min_{\overline{G}} I$, and

$$G_{-\varepsilon_n} \subset A_n \subset G_{+\varepsilon_n} \quad \text{for some } \varepsilon_n \downarrow 0 \,.$$

Then

$$\lim_n \big(\mu_n(A_n) \big)^{1/n} = \exp(- \inf_G I) = \exp(- \min_{\overline{G}} I) \,.$$

Prove it.

Hint: (a) the argument of the proof of 4b3(a) works for appropriate f_n, say, $f_n(x) = (1/\varepsilon_n) \operatorname{dist}(x, \complement G) - 1$ if this number lies on $[0,1]$, otherwise 0 (if the number is negative) or 1 (if it exceeds 1); (b) similar (but using semicontinuity), (c) follows from (a), (b).

We can also describe the value $I(x)$ of the rate function at a given point x in terms of probabilities. To this end we choose (once again) a metric on K and use open and closed balls,

$$B(x,r-) = \{y \in K : \operatorname{dist}(x,y) < r\}\,, \quad B(x,r+) = \{y \in K : \operatorname{dist}(x,y \le r\}\,.$$

4b12 Proposition. For every $x \in K$ there exists a function $(0,1) \to \{1,2,\dots\}$, $r \mapsto n_r$, such that

$$\frac{1}{n} \ln \mu_n(B(x,r\pm)) \to -I(x) \quad \text{as } r \to 0+$$

uniformly in $n \ge n_r$. (Here '±' means that the claim holds for closed and open balls.)

Proof. By 4b4,

$$\liminf_n \big(\mu_n(B(x,r-))\big)^{1/n} \ge \exp\big(-\inf_{B(x,r-)} I\big)\,,$$
$$\limsup_n \big(\mu_n(B(x,r+))\big)^{1/n} \le \exp\big(-\min_{B(x,r+)} I\big)\,.$$

We choose n_r such that

$$\big(\mu_n(B(x,r-))\big)^{1/n} \ge \exp\big(-\inf_{B(x,r-)} I\big) - r\,,$$
$$\big(\mu_n(B(x,r+))\big)^{1/n} \le \exp\big(-\min_{B(x,r+)} I\big) + r$$

for all $n \ge n_r$. By lower semicontinuity of I,

$$\inf_{B(x,r\pm)} I \uparrow I(x) \quad \text{as } r \to 0+\ .$$

We have

$$\exp\big(-\inf_{B(x,r-)} I\big) - r \le \big(\mu_n(B(x,r-))\big)^{\frac{1}{n}} \le \big(\mu_n(B(x,r+))\big)^{\frac{1}{n}} \le \exp\big(-\min_{B(x,r+)} I\big) + r$$

$$\searrow \; e^{-I(x)} \; \swarrow$$

therefore

$$\big(\mu_n(B(x,r\pm))\big)^{1/n} \to e^{-I(x)} \quad \text{as } r \to 0+$$

uniformly in $n \ge n_r$. □

In fact, the (necessary) condition

(4b13)
$$\lim_{r\to 0+} \liminf_n \big(\mu_n(B(x,r-))\big)^{1/n} = \lim_{r\to 0+} \limsup_n \big(\mu_n(B(x,r+))\big)^{1/n} \quad \text{for } x \in K$$

is also sufficient for LD-convergence of $(\mu_n)_n$. I give no proof. (See also [3, Sect. 3.1, Remark 3.1(c)] and [1, Th. 4.1.11].)

4c Restriction and conditioning

Any large deviation is done in the least unlikely of all the unlikely ways!

den Hollander [4, p. 10]

Let probability measures μ_n on a compact metrizable space K satisfy LDP with a rate function I. Assume that an open set $G \subset K$ satisfies (4b8) (which always holds if I is continuous).

4c1 Proposition. For every $f \in C(K)$,

$$\lim_n \biggl(\int_G |f|^n \, \mathrm{d}\mu_n \biggr)^{1/n} = \lim_n \biggl(\int_{\overline{G}} |f|^n \, \mathrm{d}\mu_n \biggr)^{1/n} = \sup_G \bigl(|f| \mathrm{e}^{-I} \bigr) = \max_{\overline{G}} \bigl(|f| \mathrm{e}^{-I} \bigr) .$$

Proof. We may assume that $f(\cdot) \ge 0$, since only $|f|$ is relevant. Moreover, we may assume that $f(\cdot) > 0$, since strictly positive functions are dense among weakly positive functions. Thus, we assume that $f = \mathrm{e}^{-h}$, $h \in C(K)$.

We define probability measures ν_n on K by

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\mu_n} = c_n \mathrm{e}^{-nh}$$

and apply Theorem 2c1 (change of measure): $(\nu_n)_n$ satisfies LDP with the rate function $J = I + h - a$, $a = \lim_n \frac{1}{n} \ln c_n$. Condition (4b8) is satisfied also by J, thus 4b6 can be applied to $(\nu_n)_n$ giving

$$\lim_n \bigl(\nu_n(G) \bigr)^{1/n} = \lim_n \bigl(\nu_n(\overline{G}) \bigr)^{1/n} = \exp \bigl(- \inf_G J \bigr) = \exp \bigl(- \min_{\overline{G}} J \bigr) .$$

However,

$$\biggl(\int_G f^n \, \mathrm{d}\mu_n \biggr)^{1/n} = \biggl(\int_G \mathrm{e}^{-nh} \, \mathrm{d}\mu_n \biggr)^{1/n} = \bigl(c_n^{-1} \nu_n(G) \bigr)^{1/n} \to \mathrm{e}^{-a} \lim_n \bigl(\nu_n(G) \bigr)^{1/n}$$

as $n \to \infty$; the same holds for $\overline{G}$. Also,

$$\sup_G \bigl(f \mathrm{e}^{-I} \bigr) = \sup_G \mathrm{e}^{-(h+I)} = \sup_G \mathrm{e}^{-(J+a)} = \mathrm{e}^{-a} \exp \bigl(- \inf_G J \bigr) ;$$

the same holds for $\overline{G}$. □

It may happen that $I(x) = +\infty$ for all $x \in \overline{G}$. Let us exclude this case. Then $\mu_n(\overline{G}) \ne 0$ for all n large enough, and we may introduce *conditional measures,* — probability measures ν_n on $\overline{G}$ such that

$$\int f \, \mathrm{d}\nu_n = \frac{1}{\mu_n(\overline{G})} \int_{\overline{G}} f \, \mathrm{d}\mu_n \tag{4c2}$$

for all bounded Borel functions $f : \overline{G} \to \mathbb{R}$. The set $\overline{G}$ is another compact metrizable space, and we may consider LDP on this space.

4c3 Proposition. Let $\min_{\overline{G}} I \ne +\infty$, then the sequence $(\nu_n)_n$ of conditional measures on $\overline{G}$ satisfies LDP with the rate function $J : \overline{G} \to [0, \infty]$,

$$J(x) = I(x) - \min_{\overline{G}} I \quad \text{for } x \in \overline{G}\,.$$

Proof. Let $f \in C(\overline{G})$; we have to prove that $\bigl(\int_{\overline{G}} |f|^n \, \mathrm{d}\nu_n \bigr)^{1/n} \to \max_{\overline{G}} \bigl(|f| \mathrm{e}^{-J} \bigr)$. By 4c1 (applied to any continuous extension of f), $\bigl(\int_{\overline{G}} |f|^n \, \mathrm{d}\mu_n \bigr)^{1/n} \to \max_{\overline{G}} \bigl(|f| \mathrm{e}^{-I} \bigr)$. Therefore

$$\biggl(\int_{\overline{G}} |f|^n \, \mathrm{d}\nu_n \biggr)^{1/n} = \biggl(\frac{\int_{\overline{G}} |f|^n \, \mathrm{d}\mu_n}{\mu_n(\overline{G})} \biggr)^{1/n} \to \frac{\max_{\overline{G}} \bigl(|f| \mathrm{e}^{-I} \bigr)}{\max_{\overline{G}} \bigl(\mathrm{e}^{-I} \bigr)} = \max_{\overline{G}} \bigl(|f| \mathrm{e}^{-J} \bigr)\,.$$

□

4c4 Exercise. Let I be continuous and $\min_{\overline{G}} I \ne +\infty$, then

$$\frac{\mu_n\bigl(\{ x \in \overline{G} : I(x) \le \min_{\overline{G}} I + \varepsilon \} \bigr)}{\mu_n(\overline{G})} \to 1 \quad \text{as } n \to \infty$$

for all $\varepsilon > 0$.

Prove it.

Hint: use 4c3, and apply 2a20(a) to ν_n.

4c5 Exercise. A fair coin is tossed n times, giving S_n 'heads'. Prove that

$$\mathbb{P}\bigl(S_n \le 0.71n \,\big|\, S_n \ge 0.7n \bigr) \to 1 \quad \text{as } n \to \infty\,.$$

Hint: 3a4 and 4c4.

4c6 Exercise. A fair die is throwed n times, giving the outcomes $1, \dots, 6$ respectively $S_n^{(1)}, \dots, S_n^{(6)}$ times. Prove that

$$\mathbb{P}\bigl(0.15n \le S_n^{(2)}, \dots, S_n^{(6)} \le 0.17n \,\big|\, S_n^{(1)} \ge 0.2n \bigr) \to 1 \quad \text{as } n \to \infty\,.$$

Hint: 3b3 and 4c4.

4c7 Exercise. Let $X_1, \dots, X_n$ be independent, identically distributed random variables, each taking on the three values $-1, 0, 1$ with equal probabilities (1/3). Prove that

$$\mathbb{P}\biggl(\frac{5}{7} - \varepsilon \le \frac{X_1^2 + \dots + X_n^2}{n} \le \frac{5}{7} + \varepsilon \,\bigg|\, \frac{X_1 + \dots + X_n}{n} \ge \frac{3}{7} \biggr) \to 1 \quad \text{as } n \to \infty$$

for every $\varepsilon > 0$.

Hint: 3c, and 4c4.

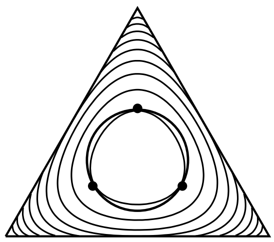
You may also think about the conditional distribution of the frequencies $\frac{k_-}{n}, \frac{k_0}{n}, \frac{k_+}{n}$ (and the mean $\frac{1}{n}(X_1 + \cdots + X_n) = \frac{k_+}{n} - \frac{k_-}{n}$), where

$$k_- = \#\{j : X_j = -1\}\,, \quad k_0 = \#\{j : X_j = 0\}\,, \quad k_+ = \#\{j : X_j = 1\}$$

(recall 3c), the condition being a large deviation of the frequencies from the probabilities in the sense that

$$\Big| \frac{k_-}{n} - \frac{1}{3} \Big|^2 + \Big| \frac{k_0}{n} - \frac{1}{3} \Big|^2 + \Big| \frac{k_+}{n} - \frac{1}{3} \Big|^2 \ge c\,.$$

In terms of the so-called χ^2 statistics, $\chi^2 = \frac{3}{n}\big((k_- - \frac{n}{3})^2 + (k_0 - \frac{n}{3})^2 + (k_+ - \frac{n}{3})^2 \big)$, it means $\chi^2 \ge 3cn$.



4c8 Exercise. Generalize 4c1, 4c3 and 4c4, replacing the single set $\overline{G}$ with a sequence of sets A_n such that A_n is μ_n-measurable, and

$$G_{-\varepsilon_n} \subset A_n \subset G_{+\varepsilon_n} \quad \text{for some } \varepsilon_n \downarrow 0\,.$$

Hint: use 4b11(c).

See also [2, Sect. 4] and [1, Sect. 3.3].

4d LDP for product measures

Let K_1, K_2 be compact metrizable spaces, then their product $K = K_1 \times K_2$ is also a compact metrizable space.

Let $\mu_n^{(1)}$ be probability measures on K_1, and $\mu_n^{(2)}$ — on K_2, then their products $\mu_n = \mu_n^{(1)} \times \mu_n^{(2)}$ are probability measures on K.

4d1 Proposition. If $(\mu_n^{(1)})_n$ satisfies LDP with a rate function I_1 and $(\mu_n^{(2)})_n$ — with I_2, then $(\mu_n)_n$ satisfies LDP with the rate function I defined by

$$I(x,y) = I_1(x) + I_2(y) \quad \text{for } x \in K_1,\ y \in K_2\,.$$

Proof. Given $f \in C(K)$, we define $g, g_1, g_2, \dots : K_1 \to \mathbb{R}$ by

$$g_n(x) = \|f(x,\cdot)\|_{L_n(\mu_n^{(2)})} = \bigg(\int |f(x,y)|^n \, \mu_n^{(2)}(\mathrm{d}y) \bigg)^{1/n}\,,$$

$$g(x) = \sup_{K_2} \big(|f(x,\cdot)| \mathrm{e}^{-I_2} \big) = \sup_{y \in K_2} \big(|f(x,y)| \mathrm{e}^{-I_2(y)} \big) \, .$$

Clearly, $g_n \to g$ pointwise. But moreover, $g_n \to g$ uniformly, due to uniform continuity:

$$|g_n(x_1) - g_n(x_2)| \le \sup_{K_2} |f(x_1,\cdot) - f(x_2,\cdot)| \, ,$$
$$|g(x_1) - g(x_2)| \le \sup_{K_2} |f(x_1,\cdot) - f(x_2,\cdot)| \, .$$

We note that $\|g_n\|_{L_n(\mu_n^{(1)})} = \|f\|_{L_n(\mu_n)}$, $\|g_n - g\|_{L_n(\mu_n^{(1)})} \le \sup_{K_1} |g_n - g| \to 0$ and $\|g\|_{L_n(\mu_n^{(1)})} \to \sup_{K_1} \big(|g| \mathrm{e}^{-I_1} \big)$, therefore

$$\|f\|_{L_n(\mu_n)} \to \sup_{K_1} \big(|g| \mathrm{e}^{-I_1} \big) =$$
$$= \sup_{x \in K_1} \Big(\mathrm{e}^{-I_1(x)} \sup_{y \in K_2} \big(\mathrm{e}^{-I_2(y)} |f(x,y)| \big) \Big) = \sup_K (|f| \mathrm{e}^{-I}) \, .$$

□

4d2 Exercise. $(\mu_n)_n$ is LD-convergent if and only if both $(\mu_n^{(1)})_n$ and $(\mu_n^{(2)})_n$ are LD-convergent.

Prove it.

Hint: use the contraction principle for 'only if'.

References

[1] A. Dembo, O. Zeitouni, *Large deviations techniques and applications,* Jones and Bartlett publ., 1993.

[2] R.S. Ellis, *The theory of large deviations and applications to statistical mechanics,* 2006, http://www.math.umass.edu/∼rsellis/pdf-files/Dresden-lectures.pdf

[3] J. Feng, T.G. Kurtz, *Large deviations for stochastic processes,* 2005, http://www.math.wisc.edu/∼kurtz/feng/ldp.htm

[4] F. den Hollander, *Large deviations,* AMS, 2000.

Index