4 Large deviations in spaces of functions

4a Some motivation

Sanov's theorem for a discrete distribution with d atoms was obtained in dimension $(d-1)$; for a nonatomic distribution, infinite dimension is required.

On the other hand, the sum of n random variables is a symmetric (that is, permutation invariant) function of them; the same holds for the empirical distribution (sample frequencies). But we are also interested in more general, non-symmetric functions. For example, given i.i.d. random variables $X_1, X_2, \dots$ with the expectation 1, consider

$$\mathbb{P}\big(\exists n \;\; X_1 + \dots + X_n \le -100\big) \,.$$

(Think about the risk of ruin in gambling, insurance or finance.) Or, like this:

$$\mathbb{P}\big(\exists k, \ell : 1 \le k \le \ell \le n, \, X_k + \dots + X_\ell \le -100\big) \,.$$

Or, for a large 2-dimensional array $n \times n$ of such i.i.d. random variables one may ask about existence of a $k \times k$ sub-array whose sum is ≤ -100; etc. Here, infinite dimension is involved even if each random variable takes only two values.

4b A joint compactification

When dealing with a sequence of models, for $n = 1, 2, \dots$, and interested in the limit as $n \to \infty$, it may help to embed these models into a single compact space.

Recall the Banach space $L_p[0,1]$, for $p \in (1, \infty)$, of all equivalence classes of measurable functions $[0,1] \to \mathbb{R}$, with the norm $\|f\|_p = \big(\int_0^1 |f(x)|^p \,\mathrm{d}x\big)^{1/p}$. Its dual space is $L_q[0,1]$ for $q = \frac{p}{p-1}$ (that is, $\frac1p + \frac1q = 1$); if $f \in L_p$ and $g \in L_q$ then $\langle f, g \rangle = \int_0^1 f(x) g(x) \,\mathrm{d}x$ is well-defined, and $|\langle f, g \rangle| \le \|f\|_p \|g\|_q$

(Hölder). Every linear[1] functional ℓ on L_p is of the form $\langle \cdot, g\rangle$ for some $g \in L_q$.[2] The unit ball $B_p = \{f \in L_p : \|f\|_p \le 1\}$ in the norm topology is separable,[3] but not compact.[4] Here is a weaker, compact topology on B_p.

Given $f, f_1, f_2, \dots \in B_p$, all $g \in L_q$ such that $\langle f_n, g\rangle \to \langle f, g\rangle$ (as $n \to \infty$) are a linear[5] subspace (think, why); if it is the whole L_q, one says that $f_n \to f$ in the weak topology.[6] This topology is metrizable on B_p;[7] in particular, it corresponds to the norm

$$\|f\|_{\mathrm{int}} = \sup_{a\in(0,1)} \left| \int_0^a f(x)\,\mathrm{d}x \right| .$$

Weak compactness of B_p is easy to prove. We take $g_1, g_2, \dots \in B_q$ dense in B_q. Given a sequence $(f_k)_k$ in B_p, we take a subsequence (denoted by $(f_k)_k$ again) such that $\langle f_k, g_1\rangle$ converges; then, such that also $\langle f_k, g_2\rangle$ converges; and so on. The diagonal construction ensures that $\lim_k \langle f_k, g_i\rangle$ exists for all i. We get a linear functional $\ell(g) = \lim_k \langle f_k, g\rangle$ on L_q; it is $\ell(g) = \langle f, g\rangle$ for some f; and $f_k \to f$ weakly.

We turn to probability measures on L_p that are concentrated on finite-dimensional subspaces, but the rate function that describes their large deviations is not. If puzzled, recall Sect. 1a: the binomial distributions are concentrated on rational numbers, but their rate function $\gamma(\cdot)$ is not.

Let μ be a probability measure on $\mathbb{R}$ such that,[8] for a given $q \in (1, \infty)$,

$$\Lambda_\mu(t) = \mathcal{O}(|t|^q) \quad \text{as } t \to \pm\infty ,$$

and $\Lambda'_\mu(0) = 0$ (that is, expectation zero), and $\Lambda''_\mu(0) > 0$ (that is, not a single atom). We introduce a random element S_n of L_p, where $\frac{1}{p} + \frac{1}{q} = 1$,

$$S_n = nX_1 1\!\!1_{(0,\frac{1}{n})} + \dots + nX_n 1\!\!1_{(\frac{n-1}{n},1)}$$

[1] I mean, algebraically linear and continuous (that is, bounded).

[2] Hint: first, $\ell(1\!\!1_A) = \int_A g$ by Radon-Nikodym. Second, take f such that $fg = |g|^q$, that is, $f = |g|^{q/p} \operatorname{sgn} g$; then, for every measurable A such that g is bounded on A we have $\ell(f \cdot 1\!\!1_A) = \int_A fg = \int_A |g|^q$ and $\|f \cdot 1\!\!1_A\|_p = \bigl(\int_A |g|^q\bigr)^{1/p}$, thus $\|\ell\| \ge \bigl(\int_A |g|^q\bigr)^{1/q}$.

[3] Rational step functions are dense; rational piecewise linear functions are also dense.

[4] Try $f_n(x) = \sin nx$, or the Rademacher functions $f_n(x) = \frac{\cos 2^n \pi x}{|\cos 2^n \pi x|}$.

[5] I mean, algebraically linear and closed.

[6] In this case the convergence is uniform on compact subsets of L_q, but (generally) not uniform on B_q.

[7] But not on the whole L_p; never mind.

[8] More generally, one may require $\forall t\ \ \Lambda_\mu(t) < \infty$ and use Orlicz spaces (more general than L_p spaces).

where $X_1, \dots, X_n$ are independent random variables, each distributed μ. Here is the corresponding cumulant generating function:

$$\Lambda_n(g) = \ln \mathbb{E} \exp\langle S_n, g\rangle = \Lambda_\mu\biggl(n\int_0^{1/n} g \biggr) + \dots + \Lambda_\mu\biggl(n\int_{\frac{n-1}{n}}^1 g \biggr) \quad \text{for } g \in L_q \,.$$

It is easy to guess that $\frac1n \Lambda_n(g) \to \int_0^1 \Lambda_\mu\bigl(g(x)\bigr)\,\mathrm{d}x$ as $n\to\infty$.

4b1 Lemma. $\Lambda'_\mu(t) = \mathcal{O}\bigl(|t|^{q-1}\bigr)$ as $t \to \pm\infty$.

Proof. For $s,t>0$, by convexity, $\Lambda_\mu(t+s) \ge \Lambda_\mu(t) + s\Lambda'_\mu(t)$; if t is large enough, then $s\Lambda'_\mu(t) \le C(t+s)^q$ for all $s>0$; and $\min_{s>0} \frac1s (t+s)^q = \frac{q^q}{(q-1)^{q-1}} t^{q-1}$. The case $t<0$ is similar. $\square$

4b2 Lemma. There exists C such that for all $g_1, g_2 \in L_q$,

$$\biggl| \int_0^1 \Lambda_\mu\bigl(g_1(x)\bigr)\,\mathrm{d}x - \int_0^1 \Lambda_\mu\bigl(g_2(x)\bigr)\,\mathrm{d}x \biggr| \le C \|g_1-g_2\|_q \bigl(1 + \|g_1\|_q + \|g_2\|_q \bigr)^{q-1} \,.$$

Proof. Using 4b1, we take C such that

$$\forall t_1,t_2 \;\; |\Lambda_\mu(t_1) - \Lambda_\mu(t_2)| \le C|t_1-t_2| \bigl(\max(1,|t_1|,|t_2|) \bigr)^{q-1} \,;$$

then

$$\biggl| \int_0^1 \Lambda_\mu\bigl(g_1(x)\bigr)\,\mathrm{d}x - \int_0^1 \Lambda_\mu\bigl(g_2(x)\bigr)\,\mathrm{d}x \biggr| \le \int_0^1 |\Lambda_\mu\bigl(g_1(x)\bigr) - \Lambda_\mu\bigl(g_2(x)\bigr)|\,\mathrm{d}x \le$$
$$\le C \langle |g_1-g_2|, \bigl(\max(1,|g_1|,|g_2|)\bigr)^{q-1} \rangle \le C\|g_1-g_2\|_q \| \bigl(\max(1,|g_1|,|g_2|)\bigr)^{q-1} \|_p \,,$$

and $\|\bigl(\max(\dots)\bigr)^{q-1}\|_p = \|\max(\dots)\|_q^{q-1}$, and finally, $\|\max(\dots)\|_q \le \|1 + |g_1| + |g_2| \|_q \le \|1\|_q + \|g_1\|_q + \|g_2\|_q$. $\square$

4b3 Proposition. For every $g \in L_q[0,1]$,

$$\frac1n \Lambda_n(g) \to \Lambda_\infty(g) = \int_0^1 \Lambda_\mu\bigl(g(x)\bigr)\,\mathrm{d}x \quad \text{as } n\to\infty \,,$$

and $\frac1n \Lambda_n(g) \le \Lambda_\infty(g)$ for all n.

Proof. Introducing linear operators $A_n : L_q \to L_q$ by $A_n g = \bigl(n\int_0^{1/n} g \bigr) \mathbb{1}_{(0,\frac1n)} + \dots + \bigl(n\int_{\frac{n-1}{n}}^1 g \bigr) \mathbb{1}_{(\frac{n-1}{n},1)}$ we have $\|A_n\| \le 1$ and $A_n g \to g$ (in the norm topology) for all $g \in L_q$ (indeed, such g are a subspace containing all continuous functions). We note that $\frac1n \Lambda_n(g) = \Lambda_\infty(A_n g)$; by 4b2, $\Lambda_\infty(A_n g) \to \Lambda_\infty(g)$. Also, $\frac1n \Lambda_n(g) \le \Lambda_\infty(g)$, since $\Lambda_\mu\bigl(n\int_{k/n}^{(k+1)/n} g(x)\,\mathrm{d}x \bigr) \le n\int_{k/n}^{(k+1)/n} \Lambda_\mu\bigl(g(x)\bigr)\,\mathrm{d}x$ by convexity of Λ_μ. $\square$

4b4 Example (Normal distribution, see 2a2)**.** Let $\frac{\mu(\mathrm{d}x)}{\mathrm{d}x} = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$, then $\Lambda_\mu(t) = t^2/2$, and $\Lambda_\infty(g) = \int_0^1 \frac{1}{2} g^2(x)\,\mathrm{d}x = \frac{1}{2}\|g\|_2^2$. Every $p \in (1,2]$ may be used.

4b5 Example (Fair coin, see 2a3)**.** Let $\mu(\{-1\}) = 1/2 = \mu(\{+1\})$, then $\Lambda_\mu(t) = \ln\cosh t$, and $\Lambda_\infty(g) = \int_0^1 \ln\cosh g(x)\,\mathrm{d}x$. Every $p \in (1,\infty)$ may be used.

Note that $\frac{1}{t}\Lambda_\mu(t)$ converges to $\Lambda'_\mu(+\infty) \in (0,+\infty]$ as $t \to +\infty$, and to $\Lambda'_\mu(-\infty) \in [-\infty, 0)$ as $t \to -\infty$; the least closed interval of full measure μ is $[\Lambda'_\mu(-\infty), \Lambda'_\mu(+\infty)] \cap \mathbb{R}$. Also,

$$\lim_{t\to+\infty} \frac{1}{t}\Lambda_\infty(tg) = |\Lambda'_\mu(-\infty)| \int_0^1 g^-(x)\,\mathrm{d}x + |\Lambda'_\mu(+\infty)| \int_0^1 g^+(x)\,\mathrm{d}x \ge \\ \ge \min\bigl(|\Lambda'_\mu(-\infty)|, |\Lambda'_\mu(+\infty)|\bigr)\|g\|_1 \,. \tag{4b6}$$

For the (one-dimensional) distribution ν_n of $\langle S_n, g\rangle$ we have $\Lambda_{\nu_n}(t) = \ln \mathbb{E}\exp\bigl(t\langle S_n, g\rangle\bigr) = \Lambda_n(tg)$, thus,

$$\frac{1}{n}\Lambda_{\nu_n}(t) \to \Lambda_\infty(tg) = \int_0^1 \Lambda_\mu\bigl(tg(x)\bigr)\,\mathrm{d}x \quad \text{as } n \to \infty \,. \tag{4b7}$$

Much more can be said about ν_n, since it corresponds to a sum of n independent (but not identically distributed) random variables. On the other hand, (4b7) itself leads to LDP, according to “finite-dimensional” Sect. 4c below.

4c Gärtner-Ellis theorem

Let probability measures $\nu_1, \nu_2, \dots$ on $\mathbb{R}$ be such that the limit

$$\lim_{n\to\infty} \frac{1}{n}\Lambda_{\nu_n}(t) = \Lambda(t) \in \mathbb{R} \tag{4c1}$$

exists for all $t \in \mathbb{R}$. (In particular, $\nu_n = \nu^{*n}$ satisfy just $\frac{1}{n}\Lambda_{\nu_n}(t) = \Lambda_\nu(t)$.)

Convexity of Λ_{ν_n} implies convexity of Λ, and therefore, existence of one-sided derivatives $\Lambda'(t-) \le \Lambda'(t+)$. However, these can differ (in spite of analyticity of Λ_{ν_n}); recall Example 2c10.

The Legendre transform

$$\Lambda^*(x) = \sup_{t\in\mathbb{R}} \bigl(tx - \Lambda(t)\bigr)$$

is a convex function $\mathbb{R} \to [0,\infty]$ (since $\Lambda(0) = 0$), and $\frac{1}{|x|}\Lambda^*(x) \to \infty$ as $x \to \pm\infty$ (since $\Lambda(\cdot) < \infty$).

Note that $\sup_{t\ge0}\big(tx-\Lambda(t)\big)=0 \iff x\le\Lambda'(0+)$, and $\sup_{t\le0}\big(tx-\Lambda(t)\big)=0 \iff x\ge\Lambda'(0-)$; thus,

$$\Lambda^*(x)=\begin{cases}\sup_{t\le0}\big(tx-\Lambda(t)\big)>0 & \text{for } x<\Lambda'(0-),\\ 0 & \text{for } \Lambda'(0-)\le x\le\Lambda'(0+),\\ \sup_{t\ge0}\big(tx-\Lambda(t)\big)>0 & \text{for } x>\Lambda'(0+).\end{cases}$$

Consider $\Lambda'(-\infty)\in[-\infty,+\infty)$ and $\Lambda'(+\infty)\in(-\infty,+\infty]$. For $x\in\big(\Lambda'(-\infty),\Lambda'(+\infty)\big)$ the function $t\mapsto tx-\Lambda(t)$ is maximal at some t (maybe, non-unique); then clearly $\Lambda'(t-)\le x\le\Lambda'(t+)$ and $\Lambda^*(x)=tx-\Lambda(t)$. On this open interval Λ^* is finite and convex, therefore continuous. If $x\notin[\Lambda'(-\infty),\Lambda'(+\infty)]$ then $\Lambda^*(x)=+\infty$ (try $t\to-\infty$, $t\to+\infty$). But if $x\in\{\Lambda'(-\infty),\Lambda'(+\infty)\}\cap\mathbb{R}$, two cases are possible: either $\Lambda^*(x)<\infty$ (recall 2c7), or $\Lambda^*(x)=\infty$ (recall 2c8).

4c2 Lemma.

$$\begin{aligned}\nu_n[nx,\infty)&\le\exp\big(-n\Lambda^*(x)+o(n)\big)\quad\text{for } x\ge\Lambda'(0+)\,;\\ \nu_n(-\infty,nx]&\le\exp\big(-n\Lambda^*(x)+o(n)\big)\quad\text{for } x\le\Lambda'(0-)\,.\end{aligned}$$

Proof. Let $x\ge\Lambda'(0+)$ (the other case is similar). For $t\ge0$ we have

$$\begin{aligned}\nu_n[nx,\infty)&\le\frac{\int \mathrm{e}^{tx}\,\nu_n(\mathrm{d}x)}{\mathrm{e}^{tnx}}=\exp\big(\Lambda_{\nu_n}(t)-ntx\big)\,;\\ \frac1n\ln\nu_n[nx,\infty)&\le\frac1n\Lambda_{\nu_n}(t)-tx\to\Lambda(t)-tx\,;\\ \limsup_n\frac1n\ln\nu_n[nx,\infty)&\le-\sup_{t\ge0}\big(tx-\Lambda(t)\big)=-\Lambda^*(x)\,.\end{aligned}$$

□

For tilted measures $\nu_{n,t}$ we have $\Lambda_{\nu_{n,t}}(s)=\Lambda_{\nu_n}(t+s)-\Lambda_{\nu_n}(t)$, thus $\frac1n\Lambda_{\nu_{n,t}}(s)\to\Lambda_t(s)=\Lambda(t+s)-\Lambda(t)$. The corresponding Legendre transform is

$$\Lambda_t^*(x)=\Lambda^*(x)-tx+\Lambda(t)\,,$$

since $\sup_s\big(sx-\Lambda_t(s)\big)=\sup_s\big(sx-\Lambda(t+s)+\Lambda(t)\big)=\sup_s\big((s-t)x-\Lambda(s)+\Lambda(t)\big)=\sup_s\big(sx-\Lambda(s)\big)-tx+\Lambda(t)$. Note that Λ_t^* vanishes on $[\Lambda_t'(0-),\Lambda_t'(0+)]=[\Lambda'(t-),\Lambda'(t+)]$ (only). By 4c2,

$$\begin{aligned}\nu_{n,t}[nx,\infty)&\le\exp\big(-n\Lambda_t^*(x)+o(n)\big)\quad\text{for } x>\Lambda'(t+)\,;\\ \nu_{n,t}(-\infty,nx]&\le\exp\big(-n\Lambda_t^*(x)+o(n)\big)\quad\text{for } x<\Lambda'(t-)\,.\end{aligned}$$

Therefore

$$\nu_{n,t}(na,nb) \to 1 \quad \text{whenever } (a,b) \supset [\Lambda'(t-),\Lambda'(t+)]\,. \tag{4c3}$$

Taking into account that

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\nu_{n,t}}(x) = \exp\bigl(-tx+\Lambda_{\nu_n}(t)\bigr) \ge \exp\bigl(-n\max(ta,tb)+\Lambda_{\nu_n}(t)\bigr) \quad \text{for } x \in (na,nb)$$

we get

$$\nu_n(na,nb) \ge \exp\bigl(-n\max(ta,tb)+n\Lambda(t)+o(n)\bigr) \tag{4c4}$$

whenever $(a,b) \supset [\Lambda'(t-),\Lambda'(t+)]$.

Now we assume that Λ is differentiable (that is, $\Lambda'(t-) = \Lambda'(t+)$ for all $t \in \mathbb{R}$).[1]

4c5 Lemma. If $\Lambda'(0) \le x < \Lambda'(+\infty)$, then for every $\varepsilon > 0$,

$$\nu_n\bigl(nx, n(x+\varepsilon)\bigr) \ge \exp\bigl(-n\Lambda^*(x)+o(n)\bigr)\,.$$

Proof. We take the maximal[2] $t_0 \ge 0$ such that $\Lambda'(t_0) = x$. By (4c4), for every $b > x$ and every t such that $x < \Lambda'(t) < b$ we have $\nu_n(nx,nb) \ge \exp\bigl(-ntb+n\Lambda(t)+o(n)\bigr)$; that is, $\liminf_n \frac1n \ln \nu_n(nx,nb) \ge -tb+\Lambda(t)$ whenever $x < \Lambda'(t) < b$; the latter holds whenever $t > t_0$ is close enough to t_0, therefore it also holds for $t = t_0$: $\liminf_n \frac1n \ln\nu_n(nx,nb) \ge -t_0 b + \Lambda(t_0) = -t_0 x + \Lambda(t_0) - t_0(b-x) = -\Lambda^*(x) - t_0(b-x)$ for all $b > x$. Finally, $\liminf_n \frac1n \ln \nu_n\bigl(nx,n(x+\varepsilon)\bigr) \ge -\Lambda^*(x) - t_0(b-x)$ for all $b \in (x, x+\varepsilon]$, therefore also for $b = x$. $\square$

In combination with (4c2) we get the following.

4c6 Proposition. If ν_n satisfy (4c1) with a differentiable Λ, then

$$\nu_n\bigl(nx,n(x+\varepsilon)\bigr) = \exp\bigl(-n\Lambda^*(x)+o(n)\bigr)$$

for all $x \in \bigl[\Lambda'(0),\Lambda'(+\infty)\bigr)$ and $\varepsilon > 0$.

And, of course,

$$\nu_n\bigl(n(x-\varepsilon),nx\bigr) = \exp\bigl(-n\Lambda^*(x)+o(n)\bigr)$$

for all $x \in (\Lambda'(-\infty),\Lambda'(0)]$ and $\varepsilon > 0$.

[1] Without this assumption Lemma 4c5 still holds for $a \notin \cup_t\bigl[\Lambda'(t-),\Lambda'(t+)\bigr)$.

[2] Recall 2c10…

4c7 Example. Do not think that Λ determines uniquely $\lim_n \frac{1}{n} \ln \nu_n(na, nb)$ in all cases.
Let ν_n be the atom (of probability 1) at $1/n$. Then $\Lambda_n(t) = \frac{1}{n} t$; $\Lambda(t) = 0$ for all t; $\Lambda^*(x) = +\infty$ for all $x \ne 0$; $\Lambda^*(0) = 0$; and $\nu_n(0, \infty) = 1$ for all n.
The atom at $-1/n$ leads to the same Λ (and Λ^*), but here $\nu_n(0, \infty) = 0$ for all n.

We turn to a finite dimension. Let probability measures $\nu_1, \nu_2, \dots$ on $\mathbb{R}^d$ be such that the limit

$$\lim_{n\to\infty} \frac{1}{n} \Lambda_{\nu_n}(t) = \Lambda(t) \in \mathbb{R} \tag{4c8}$$

exists for all $t \in \mathbb{R}^d$. As before, Λ is convex, and

$$\Lambda^*(x) = \sup_{t\in\mathbb{R}^d} \big(\langle t, x \rangle - \Lambda(t) \big)$$

is a convex function $\mathbb{R}^d \to [0, \infty]$, and $\frac{1}{|x|} \Lambda^*(x) \to \infty$ as $|x| \to \infty$ (since Λ is locally bounded).

We consider the interior (possibly, empty) G of the set $\{x : \Lambda^*(x) < \infty\}$.[1]

4c9 Exercise. (a) $x \in G$ if and only if $\liminf_{|t|\to\infty} \frac{\Lambda(t) - \langle t, x\rangle}{|t|} > 0$.

(b) G is convex.

Prove it.

4c10 Theorem. (a) For every nonempty closed set $F \subset \mathbb{R}^d$,

$$\limsup_n \frac{1}{n} \ln \nu_n(nF) \le - \min_{x\in F} \Lambda^*(x) \, .$$

(b) If Λ is differentiable and G is nonempty,[2] then for every open set $U \subset \mathbb{R}^d$,

$$\liminf_n \frac{1}{n} \ln \nu_n(nU) \ge - \inf_{x\in U} \Lambda^*(x) \, .$$

(The infimum over F is reached; think, why.)

4c11 Exercise (upper bound for a half-space)**.** [3]

$$\nu_n\big(\{ nx : \langle t, x \rangle - \Lambda(t) \ge c \} \big) \le \exp\big(-cn + o(n) \big)$$

for all $t \in \mathbb{R}^d$ and $c \ge 0$.

Prove it.

[1] But in Sect. 3 G was a set of t, not x.

[2] The claim still holds when $G = \emptyset$, but the proof is more complicated; see Dembo and Zeitouni, Exer. 2.3.20.

[3] Recall 3a3.

Now we assume that Λ is differentiable.

As before, for every $x \in G$ there exists $t \in \mathbb{R}^d$ (maybe, non-unique) such that $\nabla\Lambda(t) = x$ and $\Lambda^*(x) = \langle t, x\rangle - \Lambda(t)$.

The upper bound 4c11 applies to a half-space not containing the "expectation", since $\langle t, \nabla\Lambda(0)\rangle - \Lambda(t) \le 0$ by convexity.

4c12 Exercise (half-space not containing the "expectation"). [1] If $c > \langle t, \nabla\Lambda(0)\rangle$, then

$$\exists \varepsilon > 0 \;\; \nu_n(\{nx : \langle t, x\rangle \ge c\}) = \mathcal{O}(\mathrm{e}^{-\varepsilon n}) \,.$$

Prove it.

4c13 Exercise (exponential concentration near the "expectation"). [2] If $U \subset \mathbb{R}^d$ is a neighborhood of the point $\nabla\Lambda(0)$, then

$$\exists \varepsilon > 0 \;\; 1 - \nu_n(nU) = \mathcal{O}(\mathrm{e}^{-\varepsilon n}) \,.$$

Prove it.

4c14 Exercise (lower bound). [3] If $G \ne \emptyset$ and $U \subset \mathbb{R}^d$ is open, then

$$\ln \nu_n(nU) \ge -n \inf_{x \in U \cap G} \Lambda^*(x) + o(n) \,.$$

Prove it.

Proof *of Theorem 4c10(b).* We have

$$\inf_{x \in U} \Lambda^*(x) = \inf_{x \in U \cap G} \Lambda^*(x)$$

since, first, $\Lambda^*(x) = \infty$ for $x \notin \overline{G}$, and second, for $x \in \partial G$ and $y \in G$, by convexity, $\Lambda^*\big(x + \varepsilon(y - x)\big) \le \Lambda^*(x) + \varepsilon\big(\Lambda^*(y) - \Lambda^*(x)\big)$ and $x + \varepsilon(y - x) \in G$. It remains to apply 4c14. □

Proof *of Theorem 4c10(a).* If $\Lambda^*(x) > c$ then, by 4c11, x belongs to some open half-space $H = \{y : \langle t, y\rangle - \Lambda(t) > c\}$ such that $\nu_n(nH) \le \exp\big(-cn + o(n)\big)$. If $\Lambda^*(x) > c$ for all x of a compact set F, then $\nu_n(nF) \le \exp\big(-cn + o(n)\big)$, since F is covered by a finite number of such half-spaces (recall 3a6). However, we need it for a closed F, not just compact.

Similarly to the proof of 3a5, we apply 4c11 to $t = \pm e_1, \dots, \pm e_d$ and obtain, for every R,

$$1 - \nu_n\big(n[-R, R]^d\big) \le \exp\big(-n(R - C) + o(n)\big)$$

[1] Recall 3a4.

[2] Recall 3a5.

[3] Recall 3a1.

where $C = \max\big(\Lambda(-e_1), \Lambda(e_1), \dots, \Lambda(-e_d), \Lambda(e_d)\big)$. Thus, for every c there exists a compact $K \subset \mathbb{R}^d$ such that $1 - \nu_n(nK) \le \exp\big(-cn + o(n)\big)$.[1] Finally, $\nu_n(nF) \le \nu_n\big(n(F \cap K)\big) + 1 - \nu_n(nK)$. $\square$

Here is a reason, why dimension $d > 1$ needs more caution than dimension 1. In one dimension, the function $\Lambda^* : \mathbb{R} \to [0, \infty]$ may be discontinuous (in the compact topology of $[0, \infty]$) at $\Lambda'(-\infty)$ or/and $\Lambda'(+\infty)$ (recall (2c7), but its restriction to $[\Lambda'(-\infty), \Lambda'(+\infty)] \cap \mathbb{R}$ is continuous, since Λ^* is convex and lower semicontinuous (as was noted before (3c2)). In higher dimensions, $\Lambda^*|_{\overline{G}}$ need not be continuous.

4c15 Example. There exist $t_n \in \mathbb{R}^2$ and $c_n \in \mathbb{R}$ such that the function $f : x \mapsto \sup_n\big(\langle t_n, x\rangle + c_n\big) \in [0, \infty]$ (evidently convex and lower semicontinuous) is finite on the disk $\{x \in \mathbb{R}^2 : |x| < 1\}$ but has a discontinuous restriction to its boundary.

We choose $\varphi_n \in (0, \pi/2)$, $\varphi_n \downarrow 0$, introduce $\varphi_{n+o.5} = (\varphi_n + \varphi_{n+1})/2$, $x_n = (\cos\varphi_n, \sin\varphi_n) \in \mathbb{R}^2$, $x_{n+0.5} = (\cos\varphi_{n+0.5}, \sin\varphi_{n+0.5}) \in \mathbb{R}^2$, and define t_n, c_n by

$$\langle t_n, x_n\rangle + c_n = 0 = \langle t_n, x_{n+1}\rangle + c_n \, , \quad \langle t_n, x_{n+0.5}\rangle + c_n = 1 \, .$$

In addition, we take $t_0 = 0$, $c_0 = 0$. We get $x_{n+0.5} \to x_\infty = (1, 0)$, $f(x_{n+0.5}) = 1$ for all n, but $f(x_\infty) = 0$.

Using the "multiscale" approach as in 2c10 one can construct a probability measure μ on $\mathbb{R}^2$ such that Λ_μ behaves like f above.

4d Exponential tightness

We return to the random elements S_n of L_p, introduced in Sect. 4b.

4d1 Proposition. There exists $\varepsilon > 0$ such that for all R large enough,

$$\sup_n \frac{1}{n} \ln \mathbb{P}\big(\|S_n\|_p \ge Rn\big) \le -\varepsilon R^p \, .$$

Using the weak topology of L_p we have, for every C, a compact set $K \subset L_p$ such that $\mathbb{P}\big(S_n \notin nK\big) = \mathcal{O}(\mathrm{e}^{-Cn})$. This is called *exponential tightness.*

Recall that $\Lambda_n(g) = \ln \mathbb{E} \exp\langle S_n, g\rangle$ for $g \in L_q$, $\frac{1}{p} + \frac{1}{q} = 1$; $\frac{1}{n}\Lambda_n(g) \to \Lambda_\infty(g) = \int_0^1 \Lambda_\mu\big(g(x)\big)\,\mathrm{d}x$ by 4b3; and $\frac{1}{n}\Lambda_n(g) \le \Lambda_\infty(g)$. Also, $\Lambda_\mu(t) = \mathcal{O}(|t|^q)$

[1] So-called exponential tightness; see Sect. 4d.

as $t \to \pm\infty$, which implies[1] $\Lambda_\infty(g) = \mathcal{O}\big(\|g\|_q^q\big)$ as $\|g\|_q \to \infty$. We introduce Legendre transforms

$$\Lambda_n^*(f) = \sup_{g \in L_q} \big(\langle f, g \rangle - \Lambda_n(g) \big) \in [0, \infty]\,;$$
$$\Lambda_\infty^*(f) = \sup_{g \in L_q} \big(\langle f, g \rangle - \Lambda_\infty(g) \big) \in [0, \infty]\,.$$

4d2 Lemma.

$$\Lambda_\infty^*(f) = \int_0^1 \Lambda_\mu^*\big(f(x)\big)\,\mathrm{d}x \quad \text{for all } f \in L_p\,.$$

Proof. First, $\Lambda_\infty^*(f) \le \int_0^1 \Lambda_\mu^*\big(f(x)\big)\,\mathrm{d}x$, since $\langle f, g \rangle - \Lambda_\infty(g) = \int_0^1 f(x)g(x)\,\mathrm{d}x - \int_0^1 \Lambda_\mu(g(x))\,\mathrm{d}x = \int \big(f(x)g(x) - \Lambda_\mu(g(x))\big)\,\mathrm{d}x$, and the integrand does not exceed $\Lambda_\mu^*\big(f(x)\big)$ irrespective of g.

In order to prove that $\Lambda_\infty^*(f) \ge \int_0^1 \Lambda_\mu^*\big(f(x)\big)\,\mathrm{d}x$ we use operators A_n introduced in the proof of 4b3: $A_n f = \sum_k f_k \mathbb{1}_{(\frac{k-1}{n}, \frac{k}{n})}$ where $f_k = n \int_{\frac{k-1}{n}}^{\frac{k}{n}} f(x)\,\mathrm{d}x$; and the same for g.

We have

$$\begin{aligned}
\Lambda_\infty^*(f) &\ge \sup_{g = A_n g} \big(\langle f, g \rangle - \Lambda_\infty(g) \big) = \\
&= \sup_{g_1, \dots, g_n} \Big(\frac{1}{n} \sum_k f_k g_k - \frac{1}{n} \sum_k \Lambda_\mu(g_k) \Big) = \frac{1}{n} \sup_{g_1, \dots, g_n} \sum_k \big(f_k g_k - \Lambda_\mu(g_k) \big) = \\
&\qquad = \frac{1}{n} \sum_k \sup_t \big(f_k t - \Lambda_\mu(t) \big) = \frac{1}{n} \sum_k \Lambda_\mu^*(f_k) = \int_0^1 \Lambda_\mu^*\big(A_n f(x)\big)\,\mathrm{d}x\,.
\end{aligned}$$

Also, $A_n f \to f$ in L_p; we take $n_1 < n_2 < \dots$ such that $A_{n_i} f \to f$ almost everywhere. The lower semicontinuity of Λ_μ^* implies $\Lambda_\mu^*(f(\cdot)) \le \liminf_i \Lambda_\mu^*(A_{n_i} f(\cdot))$ a.e.; by Fatou's lemma, $\int_0^1 \Lambda_\mu^*(f(x))\,\mathrm{d}x \le \liminf_i \int_0^1 \Lambda_\mu^*\big(A_{n_i} f(x)\big)\,\mathrm{d}x \le \Lambda_\infty^*(f)$. $\square$

4d3 Exercise. (a) $\Lambda_n^*(f) = +\infty$ whenever $f \ne A_n f$;

(b) $\frac{1}{n}\Lambda_n^*(nf) = \int_0^1 \Lambda_\mu^*\big(f(x)\big)\,\mathrm{d}x = \Lambda_\infty^*(f)$ whenever $f = A_n f$;

(c) $\frac{1}{n}\Lambda_n^*(nA_n f) \to \Lambda_\infty^*(f)$ for all f.

Prove it.

4d4 Example (Normal distribution, see 2c6 and 4b4)**.** $\Lambda_\mu^*(x) = \frac{1}{2}x^2$, and $\Lambda_\infty^*(f) = \int_0^1 \frac{1}{2} f^2(x)\,\mathrm{d}x = \frac{1}{2}\|f\|_2^2$.

[1] Hint: $\Lambda_\mu(t) \le \text{const} \cdot (1 + |t|^q)$ for all t.

4d5 Example (Fair coin, see 2c7 and 4b5)**.** $\Lambda_\mu^*(x) = \gamma(x)$ is just the function of (1a1); $\Lambda_\infty^*(f) = \int_0^1 \gamma(f(x))\,\mathrm{d}x$; note that $\gamma(x) = +\infty$ for $x \notin [-1,1]$.

4d6 Lemma. There exists $\varepsilon > 0$ such that for all x large enough,

$$\Lambda_\mu^*(x) \ge \varepsilon x^p,\ \Lambda_\mu^*(-x) \ge \varepsilon x^p,\ \ \mu[x,\infty) \le \exp\bigl(-\varepsilon x^p\bigr),\ \mu(-\infty,-x] \le \exp\bigl(-\varepsilon x^p\bigr).$$

Proof. We know that $\Lambda_\mu(t) \le C|t|^q$ for $|t| \ge T$. Thus, for every $t \ge T$ we have $\Lambda_\mu^*(x) \ge tx - C|t|^q$ and $\mu[x,\infty) \le \exp\bigl(C|t|^q - tx\bigr)$. Given $x \ge CqT^{q-1}$, we take $t \ge T$ such that $Cqt^{q-1} = x$. Then $tx - C|t|^q = \varepsilon x^p$, where $\varepsilon = \frac{1}{p(Cq)^{p-1}}$. For $(-x)$ the proof is similar. $\square$

It follows from 4d6 and 4d2 that

$$\inf_{\|f\|_p \ge R} \Lambda_\infty^*(f) \ge \varepsilon R^p \quad \text{for large } R\,.$$

(Hint: $\Lambda_\mu^*(x) \ge \varepsilon|x|^p - \text{const}$ for all x.) It may seem that Prop. 4d1 follows, similarly to Theorem 4c10(a). But no, in the infinite dimension we cannot cover $\{f : \|f\| \ge R\}$ by finitely many half-spaces (not containing 0).

4d7 Lemma. There exists $\varepsilon > 0$ such that

$$\int \exp\bigl(\varepsilon|x|^p\bigr)\,\mu(\mathrm{d}x) < \infty\,.$$

Proof. Using the equality $\exp(\varepsilon|x|^p) = 1 + \varepsilon \int_0^{|x|^p} \mathrm{e}^{\varepsilon u}\,\mathrm{d}u$ we get

$$\int \exp\bigl(\varepsilon|x|^p\bigr)\,\mu(\mathrm{d}x) = 1 + \varepsilon \iint\limits_{0<u<|x|^p} \mathrm{e}^{\varepsilon u}\,\mathrm{d}u\,\mu(\mathrm{d}x) = 1 + \varepsilon \int_0^\infty \mathrm{d}u\,\mathrm{e}^{\varepsilon u} \int\limits_{|x|^p>u} \mu(\mathrm{d}x)\,;$$

Lemma 4d6 gives $\delta > 0$ such that $\int_{|x|^p>u} \mu(\mathrm{d}x) \le 2\mathrm{e}^{-\delta u}$ for large u; it remains to take $\varepsilon < \delta$. $\square$

Proof *of Prop. 4d1.*
Lemma 4d7 gives ε such that $\mathbb{E}\,\exp\bigl(\varepsilon|X_1|^p\bigr) = M < \infty$. We have $\bigl\|\frac{S_n}{n}\bigr\|_p^p = \frac{1}{n}\bigl(|X_1|^p + \dots + |X_n|^p\bigr)$, therefore

$$\begin{aligned}\mathbb{P}\bigl(\|S_n\|_p \ge Rn\bigr) = \mathbb{P}\biggl(\Bigl\|\frac{S_n}{n}\Bigr\|_p^p \ge R^p\biggr) = \mathbb{P}\bigl(|X_1|^p + \dots + |X_n|^p \ge nR^p\bigr) \le \\ \le \frac{\mathbb{E}\,\exp\varepsilon(|X_1|^p + \dots + |X_n|^p)}{\exp\varepsilon nR^p} = M^n \exp(-\varepsilon nR^p)\,,\end{aligned}$$

that is, $\frac{1}{n}\ln\mathbb{P}\bigl(\|S_n\|_p \ge Rn\bigr) \le -\varepsilon R^p + \ln M$; and of course, $\ln M \le \frac{\varepsilon}{2}R^p$ for large R. $\square$

4e Mogulskii's theorem

Recall the weak topology on the closed unit ball B_p of L_p; it is compact. A set $F \subset L_p$ is called sequentially weakly closed,[1] if $F \cap RB_p$ is weakly closed for all $R \in (0,\infty)$. A set $U \subset L_p$ is called sequentially weakly open, if its complement is sequentially weakly closed.

4e1 Theorem. (a) For every nonempty sequentially weakly closed set $F \subset L_p$,

$$\limsup_n \frac{1}{n} \ln \mathbb{P}\big(S_n \in nF \big) \le - \min_{f \in F} \Lambda_\infty^*(f) \,.$$

(b) For every sequentially weakly open set $U \subset L_p$,

$$\liminf_n \frac{1}{n} \ln \mathbb{P}\big(S_n \in nU \big) \ge - \inf_{f \in U} \Lambda_\infty^*(f) \,.$$

4e2 Corollary. Let a nonempty set $A \subset L_p$ satisfy

$$\inf_{f \in A^\circ} \Lambda_\infty^*(f) = \min_{f \in \overline{A}} \Lambda_\infty^*(f) = a$$

where A° and $\overline{A}$ are the interior and closure of A in the sequential weak topology. Then

$$\mathbb{P}\big(S_n \in nA \big) = \exp\big(-an + o(n) \big) \quad \text{as } n \to \infty \,.$$

We choose linearly independent $g_1, g_2, \dots \in B_q$ that span[2] L_q, and note that

$$(f_n \to f \text{ weakly}) \quad \Longleftrightarrow \quad \forall k \;\; \langle f_n, g_k \rangle \xrightarrow[n\to\infty]{} \langle f, g_k \rangle$$

for all $f, f_1, f_2, \dots \in B_p$. We introduce linear operators $T_d : L_p \to \mathbb{R}^d$ by

$$T_d f = \big(\langle f, g_1 \rangle, \dots, \langle f, g_d \rangle \big) \,;$$

they are weakly continuous, and

$$(f_n \to f \text{ weakly}) \quad \Longleftrightarrow \quad \forall d \;\; T_d f_n \xrightarrow[n\to\infty]{} T_d f \,.$$

Denote by $\nu_{d,n}$ the distribution of $T_d S_n$. Similarly to (4b7), by 4b3,

$$\frac{1}{n} \Lambda_{\nu_{d,n}}(t_1, \dots, t_d) \to \Lambda_\infty(t_1 g_1 + \dots + t_d g_d) \quad \text{as } n \to \infty$$

[1] In other words, closed in the bounded weak topology (bw-closed). In fact, every weakly closed set is bw-closed, but the converse fails; never mind.

[2] As a (closed) linear subspace.

for all d and $(t_1,\dots,t_d)\in\mathbb{R}^d$, since $\Lambda_{\nu_{d,n}}(t_1,\dots,t_d)=\ln\mathbb{E}\,\exp\big(t_1\langle S_n,g_1\rangle+\dots+t_d\langle S_n,g_d\rangle\big)=\ln\mathbb{E}\,\exp\langle S_n,t_1g_1+\dots+t_dg_d\rangle=\Lambda_n(t_1g_1+\dots+t_dg_d)$. Theorem 4c10 applies to $\nu_{d,n}$ and Λ_d^*, the Legendre transform of $\Lambda_d:(t_1,\dots,t_d)\mapsto\Lambda_\infty(t_1g_1+\dots+t_dg_d)=\int_0^1\Lambda_\mu\big(t_1g_1(x)+\dots+t_dg_d(x)\big)\,\mathrm{d}x$.

4e3 Exercise. Λ_d is differentiable.

Prove it.[1]

4e4 Exercise.

$$\liminf_{|t|\to\infty}\frac{\Lambda_d(t)}{|t|}>0\,.$$

Prove it.[2]

By 4e4 and 4c9, G contains 0 (and so, $G\ne\emptyset$).

4e5 Exercise. (a) $\Lambda_d^*(T_df)$ is the supremum of $\langle f,g\rangle-\Lambda_\infty(g)$ over all g from the finite-dimensional subspace spanned by $g_1,\dots,g_d$;

(b) $\Lambda_d^*(T_df)\uparrow\Lambda_\infty^*(f)$ as $d\to\infty$.

Prove it.[3]

4e6 Lemma.

$$\min_{f\in F}\Lambda_d^*(T_df)\uparrow\min_{f\in F}\Lambda_\infty^*(f)\quad\text{as }d\to\infty$$

for every weakly closed $F\subset B_p$.

Proof. We denote $M=\min_{f\in F}\Lambda_\infty^*(f)$ and take $f_d\in F$ such that $\Lambda_d^*(T_df_d)=\min_{f\in F}\Lambda_d^*(T_df)$; clearly, this minimum does not exceed M. Assume the contrary (to the claim of the lemma); $\liminf_{d\to\infty}\Lambda_d^*(T_df_d)=M-4\varepsilon<M$. We take $d_i\to\infty$ such that $\forall i\;\Lambda_{d_i}^*(T_{d_i}f_{d_i})\le M-3\varepsilon$. WLOG, $f_{d_i}\to f_\infty$ weakly (otherwise, choose a subsequence); and $\Lambda_\infty^*(f_\infty)\ge M$, since $f_\infty\in F$. Using 4e5(b) we take d such that $\Lambda_d^*(T_df_\infty)\ge\Lambda_\infty^*(f_\infty)-\varepsilon\ge M-\varepsilon$. For all i large enough we have $\Lambda_d^*(T_df_{d_i})\ge\Lambda_d^*(T_df_\infty)-\varepsilon$ by weak lower semicontinuity of $f\mapsto\Lambda_d^*(T_df)$. Also, $d_i\ge d$. Hence, $\Lambda_{d_i}^*(T_{d_i}f_{d_i})\ge\Lambda_d^*(T_df_{d_i})\ge\Lambda_d^*(T_df_\infty)-\varepsilon\ge M-2\varepsilon$; a contradiction.

□

***Proof** of Theorem 4e1(a).* We denote $M=\min_{f\in F}\Lambda_\infty^*(f)$. WLOG, F is bounded (otherwise we turn to $F\cap RB_p$ with R such that $\sup_n\frac1n\ln\mathbb{P}\big(\|S_n\|_p\ge Rn\big)\le -M$; such R exists by Prop. 4d1); $F\subset RB_p$. By Theorem 4c10(a),

$$\limsup_n\frac1n\ln\mathbb{P}\big(S_n\in nF\big)\le-\min_{x\in T_d(F)}\Lambda_d^*(x)\,,$$

[1] Hint: recall the proof of 4b2.

[2] Hint: use (4b6); all norms on $\mathbb{R}^d$ are equivalent.

[3] Hint: Λ_∞ is continuous.

since $\nu_{d,n}\big(nT_d(F)\big) = \mathbb{P}\big(T_d S_n \in nT_d(F)\big) \ge \mathbb{P}\big(S_n \in nF\big)$. Finally, $\min_{x\in T_d(F)} \Lambda_d^*(x) = \min_{f\in F} \Lambda_d^*(T_d f) \to M$ by 4e6. $\square$

4e7 Lemma. Let $U \subset L_p$ be sequentially weakly open, and $f_0 \in U \cap B_p$. Then there exist d and $\varepsilon > 0$ such that

$$\forall f \in B_p \ \big(\|T_d f - T_d f_0\| \le \varepsilon \implies f \in U \big) .$$

Proof. Assume the contrary: $f_d \in B_p \setminus U$, $\|T_d f_d - T_d f_0\| \le \frac1d$. Taking into account that $\|T_d f - T_d f_0\|$ is increasing in d we have $\|T_d f_n - T_d f_0\| \le \frac1n$ whenever $n \ge d$; thus $T_d f_n \to T_d f_0$ for all d, that is, $f_n \to f_0$ weakly; a contradiction. $\square$

***Proof** of Theorem 4e1(b).* Let $f_0 \in U$; we'll prove that $\liminf_n \frac1n \ln \mathbb{P}\big(S_n \in nU\big) \ge -\Lambda_\infty^*(f_0)$. We take R such that $f_0 \in RB_p$ and $\sup_n \frac1n \ln \mathbb{P}\big(\|S_n\|_p \ge Rn\big) \le -\Lambda_\infty^*(f_0)$; such R exists by Prop. 4d1. Lemma 4e7 gives d and $\varepsilon > 0$ such that $\forall f \in RB_p \ \big(\|T_d f - T_d f_0\| \le \varepsilon \implies f \in U \big)$. It is sufficient to prove that

$$\liminf_n \frac1n \ln \mathbb{P}\bigg(\Big\| T_d \frac{S_n}{n} - T_d f_0 \Big\| < \varepsilon \bigg) \ge - \inf_{x:\|x - T_d f_0\| < \varepsilon} \Lambda_d^*(x) ,$$

since $\inf_{x:\|x-T_d f_0\|<\varepsilon} \Lambda_d^*(x) \le \Lambda_d^*(T_d f_0) \le \Lambda_\infty^*(f_0)$ by 4e5. Theorem 4c10(b) gives the needed inequality, since $\nu_{d,n}\big(\{nx : \|x - T_d f_0\| < \varepsilon\}\big) = \mathbb{P}\big(\|T_d \frac{S_n}{n} - T_d f_0\| < \varepsilon\big)$. $\square$

4e8 Example. Let $X_1, X_2, \dots$ be independent standard normal random variables, and $a, b > 0$. Consider events

$$E_n = \Big\{ \max_{m=0,\dots,n} \sum_{k=1}^m (X_k - a) \ge bn \Big\} .$$

We'll see that

$$\frac1n \ln \mathbb{P}\big(E_n\big) \to \begin{cases} -2ab & \text{for } b \le a, \\ -\frac12 (a+b)^2 & \text{for } b \ge a \end{cases}$$

as $n \to \infty$.

In terms of the random elements S_n of L_p,

$$\frac1n \max_{m=0,\dots,n} \sum_{k=1}^m (X_k - a) = \max_{0\le x\le 1} \int_0^x \Big(\frac1n S_n(u) - a \Big) \, \mathrm{d}u .$$

We introduce the set

$$A = \Big\{ f \in L_p : \max_{0\le x\le 1} \int_0^x \big(f(u) - a \big) \, \mathrm{d}u \ge b \Big\} ,$$

then $E_n = \{ S_n \in nA \}$. According to 4d4, $\Lambda_\infty^*(f) = \frac12 \|f\|_2^2$.

4e9 Exercise. Prove that A satisfies the condition of Corollary 4e2, and find a there.

4e10 Exercise. Formulate and prove a counterpart of 4e9 for

$$\max_{0\le i\le j\le n} \sum_{k=i}^{j} (X_k - a) \ge bn \,.$$

Multidimensional arrays of i.i.d. random variables may be treated similarly. Various geometric bodies may be used instead of the intervals $[i,j]$.

4e11 Exercise. In the situation of 4e8, formulate and prove a statement about the conditional distribution (in the spirit of Prop. 3d2).

As was mentioned in Sect. 4b, the weak topology on B_p is metrizable and, in particular, corresponds to the norm

$$\|f\|_{\mathrm{int}} = \sup_{a\in(0,1)} \biggl| \int_0^a f(x)\,\mathrm{d}x \biggr| \,.$$

On the whole L_p the situation is more complicated; a linear functional $\langle\cdot,g\rangle$ is bounded w.r.t. $\|\cdot\|_{\mathrm{int}}$ if and only if g is (equivalent to) a function of bounded[1] variation. Nevertheless, we have the following fact.

4e12 Lemma. Λ_∞^* is lower semicontinuous w.r.t. $\|\cdot\|_{\mathrm{int}}$.

Proof. It was seen (recall 4e5) that Λ_∞^* is the supremum of $\langle\cdot,g\rangle - \Lambda_\infty(g)$ when g runs over (finite) linear combinations of $g_1, g_2, \dots$;[2] and the only requirement on these $g_1, g_2, \dots$ was that they span L_q (and are linearly independent). Thus, we may take $g_k = \mathbb{1}_{(0,x_k)}$ for a dense set $\{x_1, x_2, \dots\} \subset [0,1]$. Then each $\langle\cdot,g_k\rangle$ is continuous w.r.t. $\|\cdot\|_{\mathrm{int}}$. $\square$

4e13 Proposition. For every $f \in L_p$ such that $\Lambda_\infty^*(f) < \infty$,

$$\limsup_n \biggl| \frac1n \ln \mathbb{P}\Bigl(\Bigl\| \frac{S_n}{n} - f \Bigr\|_{\mathrm{int}} \le \varepsilon \Bigr) + \Lambda_\infty^*(f) \biggr| \xrightarrow[\varepsilon\to0]{} 0 \,.$$

Proof. We denote $F_\varepsilon = \{ f_1 : \|f_1 - f\|_{\mathrm{int}} \le \varepsilon \}$ and $U_\varepsilon = \{ f_1 : \|f_1 - f\|_{\mathrm{int}} < \varepsilon \}$; F_ε is sequentially weakly closed, and U_ε is sequentially weakly open. In order to prove that

$$\limsup_n \biggl| \frac1n \ln \mathbb{P}\Bigl(\frac{S_n}{n} \in F_\varepsilon \Bigr) + \Lambda_\infty^*(f) \biggr| \xrightarrow[\varepsilon\to0]{} 0 \,,$$

[1] In other words: finite.

[2] Continuity of Λ_∞ was used.

it is sufficient to prove that

$$0 \le \liminf_{\varepsilon} \liminf_{n} \left(\frac{1}{n} \ln \mathbb{P} \Big(\frac{S_n}{n} \in U_\varepsilon \Big) + \Lambda^*_\infty(f) \right) \\ \le \limsup_{\varepsilon} \limsup_{n} \left(\frac{1}{n} \ln \mathbb{P} \Big(\frac{S_n}{n} \in F_\varepsilon \Big) + \Lambda^*_\infty(f) \right) \le 0 \, .$$

The second (middle) inequality is trivial. The first inequality follows from Th. 4e1(b), since $\inf_{f_1 \in U_\varepsilon} \Lambda^*_\infty(f_1) \to \Lambda^*_\infty(f)$ by 4e12. Similarly, the third inequality follows from Th. 4e1(a). □

It means that

$$\frac{1}{n} \ln \mathbb{P} \Big(\Big\| \frac{S_n}{n} - f \Big\|_{\text{int}} \le \varepsilon \Big) \to -\Lambda^*_\infty(f)$$

when $\varepsilon \to 0$ and $n \ge N_\varepsilon$, that is, n grows fast enough when ε tends to 0. Otherwise, if n grows with ε but not fast enough, the situation may differ.

4e14 Exercise. (a) It may happen that

$$\min_{f_1 \in F_\varepsilon} \Lambda^*_\infty(f_1) < \inf_{f_1 \in U_\varepsilon} \Lambda^*_\infty(f_1) = +\infty \, .$$

Find an example.[1]

(b) If $\Lambda^*_\infty(f) < \infty$, then

$$\min_{f_1 \in F_\varepsilon} \Lambda^*_\infty(f_1) = \inf_{f_1 \in U_\varepsilon} \Lambda^*_\infty(f_1)$$

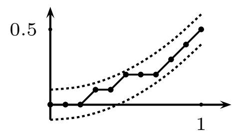
and therefore Corollary 4e2 applies, giving

$$\lim_{n \to \infty} \frac{1}{n} \ln \mathbb{P} \Big(\Big\| \frac{S_n}{n} - f \Big\|_{\text{int}} \le \varepsilon \Big) \xrightarrow[\varepsilon \to 0]{} -\Lambda^*_\infty(f) \, .$$

Prove it.[2]

4e15 Exercise. A fair coin is tossed n times, giving $(\beta_1, \dots, \beta_n) \in \{0,1\}^n$. Consider

$$p_{n,\varepsilon} = \mathbb{P} \bigg(\forall k = 1, \dots, n \quad \Big| \frac{\beta_1 + \dots + \beta_k}{n} - \frac{1}{2} \Big(\frac{k}{n} \Big)^2 \Big| \le \varepsilon \bigg) \, .$$



Prove that

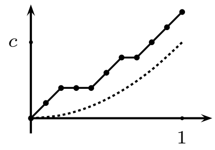
$$\limsup_{n \to \infty} \bigg| \sqrt[n]{p_{n,\varepsilon}} - \frac{\sqrt{\mathrm{e}}}{2} \bigg| \to 0 \quad \text{as } \varepsilon \to 0 \, .$$

[1] Hint: 4d5.

[2] Hint: recall the proof of Th. 4c10(b).

4e16 Exercise. A fair coin is tossed n times, giving $(\beta_1, \dots, \beta_n) \in \{0,1\}^n$. Given $c \in [0,1]$, we consider

$$p_n = \mathbb{P}\Big(\forall k = 1, \dots, n \quad \frac{\beta_1 + \dots + \beta_k}{n} \ge c\Big(\frac{k}{n}\Big)^2 \Big) .$$



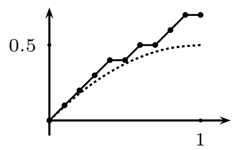
Prove that

$$\begin{aligned} &\sqrt[n]{p_n} \to 1 && \text{for } 0 \le c \le 0.5 \, , \\ &\sqrt[n]{p_n} \to \frac{1}{2c^c (1-c)^{1-c}} && \text{for } 0.5 \le c \le 1 \end{aligned}$$

($0^0 = 1$, as before).[1]

Another example:

$$p_n = \mathbb{P}\Big(\forall k = 1, \dots, n \quad \frac{\beta_1 + \dots + \beta_k}{n} \ge \frac{k}{n} - \frac12 \Big(\frac{k}{n}\Big)^2 \Big) .$$



It appears that

$$\sqrt[n]{p_n} \to \frac{\mathrm{e}^{1/4}}{\sqrt{2}} \quad \text{as } n \to \infty \, .$$

The extremal function is

$$w(x) = \begin{cases} x - 0.5x^2 & \text{for } 0 \le x \le 0.5, \\ 0.5x + 0.125 & \text{for } 0.5 \le x \le 1. \end{cases}$$

In order to prove its extremality, the following lemma helps: $\Lambda^*_\infty\big((w \wedge v)'\big) \le \Lambda^*_\infty(w')$ for every *linear* function $v : [0,1] \to \mathbb{R}$ such that $v(0) \ge 0$ and $v'(\cdot) \ge 0.5$; here $w \wedge v$ is the pointwise minimum.

[1] Hint: guess the extremal function; prove your guess, taking into account that $\int_0^1 \Lambda^*_\mu(f(x))\,\mathrm{d}x \ge \Lambda^*_\mu\big(\int_0^1 f(x)\,\mathrm{d}x\big)$.

Index