CRAMÉR'S THEOREM FOR POLISH SPACES

YU-LIANG WU

Table of Contents

1. Cramér's theorem for Polish spaces

This note establishes Cramér's theorem for higher dimensional Euclidean spaces and, more generally, for locally convex Hausdorff topological vector spaces. We begin by specifying our general framework.

1.1. Framework and Notation.

Let $\mathcal{X}$ be a topological vector space and $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent and identically distributed (i.i.d.) $\mathcal{X}$-valued random variables with law μ. We are interested in the asymptotic behavior of the empirical means:

$$\hat{S}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \hat{S}_m^n = \frac{1}{n-m}\sum_{i=m+1}^{n} X_i,$$

and their associated laws μ_n. Our primary tool is the following general sufficient condition for a large deviation principle.

Theorem 1.1. *Suppose that $\{\mu_\varepsilon\}$ satisfies a weak LDP with a convex rate function I on a locally convex Hausdorff topological vector space $\mathcal{X}$. Assume that for each $\lambda \in \mathcal{X}^*$, the limit $\Lambda_\lambda(t) = \lim_{\varepsilon\to 0} \varepsilon \Lambda_{\mu_\varepsilon}(t\lambda/\varepsilon)$ exists as extended real numbers and that $\Lambda_{\lambda(t)}$ is lower semicontinuous for $t \in \mathbb{R}$. If for every $\lambda \in \mathcal{X}^*$ and $a \in \mathbb{R}$,*

$$\inf_{x:\langle\lambda,x\rangle>a} I(x) \le \inf_{z>a} \Lambda_\lambda^*(z),$$

then $I = \Lambda^*$, and the family satisfies a weak LDP with rate function Λ^*.

1.2. Cramér's Theorem for Polish Spaces.

Within this framework, we established the LDP for Polish spaces under certain convexity assumptions.

Theorem 1.2 (Cramér's Theorem for Polish Spaces)**.** *Let $\mathcal{X}$ be a locally convex Hausdorff topological real vector space. Assume the following:*

1. *There exists a closed, convex subset $\mathcal{E} \subseteq \mathcal{X}$ such that $\mu(\mathcal{E}) = 1$, and $\mathcal{E}$ is a Polish space under the relative topology.*
2. *The closed convex hull of every compact set $K \subset \mathcal{E}$ is compact.*

Then the sequence $\{\mu_n\}$ satisfies a weak LDP in $\mathcal{X}$ (and $\mathcal{E}$) with rate function Λ^*. Furthermore, for every open convex subset $A \subset \mathcal{X}$,

$$\lim_{n\to\infty}\frac{1}{n}\log\mu_n(A) = -\inf_{x\in A}\Lambda^*(x).$$

Remark 1.3. *The theorem, in particular, applies to any separable Banach space.*

As a corollary of the theorem, we deduce Cramer's theorem for Euclidean spaces.

Corollary 1.4 (Cramér)**.** *The sequence $\{\mu_n\}$ of the laws of empirical means of $\mathbb{R}^d$ -valued i.i.d. random variables satisfies a weak LDP with the convex rate function Λ^*. Moreover, if $0\in\mathring{\mathcal{D}}_\Lambda$, then $\{\mu_n\}$ satisfies the full LDP with the good, convex rate function Λ^*.*

Proof. The weak LDP is a consequence of Theorem 1.2, from which the full LDP follows due to the exponential tightness. □

1.3. **Step 1: Weak LDP.**

We first establish the existence of a weak LDP by constructing a rate function based on the methodology used in the existence and uniqueness of the LDP. Recall that given a topological base $\mathcal{A}$, we define for each $A\in\mathcal{A}$:

$$\overline{\mathcal{L}}_A = \limsup_{n\to\infty}\frac{1}{n}\log\mu_{n(A)} \text{ and } \underline{\mathcal{L}}_A = \liminf_{n\to\infty}\frac{1}{n}\log\mu_{n(A)}.$$

The candidate rate functions are given by:

$$\overline{I}(x) = -\inf_{A\in\mathcal{A}:x\in A}\overline{\mathcal{L}}_A \text{ and } \underline{I}(x) = -\inf_{A\in\mathcal{A}:x\in A}\underline{\mathcal{L}}_A.$$

To prove the weak LDP, it suffices to show $\overline{I}=\underline{I}$, a property which, in this context, follows[1] from the convexity of the rate function.

Lemma 1.5. *Let $\mathcal{A}$ be a base for a Hausdorff topological vector space $\mathcal{X}$, such that (1) $\overline{I}=\underline{I}$ and that (2) for every $A_1,A_2\in\mathcal{A}$, $\overline{\mathcal{L}}_{A_1+A_2}\geq\frac{1}{2}\big(\underline{\mathcal{L}}_{A_1}+\underline{\mathcal{L}}_{A_2}\big)$. Then the rate function I of the weak LDP for $\{\mu_\varepsilon\}$ is convex.*

Proof. It suffices to show that $I\big(\frac{x+y}{2}\big)\leq\frac{I(x)+I(y)}{2}$ for all $x,y\in\mathcal{X}$, as the assertion amounts to that $I(tx+(1-t)y)\leq tI(x)+(1-t)I(y)$ for all $x,y\in\mathcal{X}$ and all $t=k\cdot2^{-n}$ with $k=0,1,...,2^n$ and $n\in\mathbb{N}$. By lower semicontinuity, the inequality further extends to hold for all $t\in[0,1]$.

To prove the claim, we first note that Hausdorff topological vector space is regular. By virtue of the regularity and lower semicontinuity, we can find for every $x,y\in\mathcal{X}$ neighborhoods $A\in\mathcal{A}$ of $\frac{x+y}{2}$, $A_1\in\mathcal{A}$ of x and $A_2\in\mathcal{A}$ of y such that $A\supset\overline{A_1}+\overline{A_2}$ satisfying

$$-I^\delta\Big(\frac{x+y}{2}\Big)\geq\overline{\mathcal{L}}_A\geq\frac{1}{2}\big(\overline{\mathcal{L}}_{A_1+A_2}\big)\geq\underline{\mathcal{L}}_{A_1}+\underline{\mathcal{L}}_{A_2}\geq\frac{1}{2}I(x)+\frac{1}{2}I(y).$$

This concludes the proof. □

Lemma 1.6. *Assume that Theorem 1.2(1) holds true. Then $\{\mu_n\}$ satisfies the weak LDP in $\mathcal{X}$ with a convex rate function I.*

---

[1] Note that a Hausdorff topological vector space is automatically regular, which ensures our construction yields the desired rate function.

Proof. The lemma follows from Lemma 1.5 with a base $\mathcal{A}$ of topology consisting of convex (open) sets, as argued in the following.

Note that for all convex sets A, $\mu_n(A) = \mu_n(A \cap \mathcal{E})$ and thus $\overline{\mathcal{L}}_A = \underline{\mathcal{L}}_A$ exists by subadditivity. In addition, we know that for any convex open sets A_1 and A_2,

$$\mu_n(A_1)\mu_n(A_2) = P_\mu\Big(\hat{S}_n \in A_1, \hat{S}_n^{2n} \in A_2\Big)$$
$$\leq P_\mu\Big(\hat{S}_{2n} \in \frac{1}{2}(A_1 + A_2)\Big) = \mu_{2n}\Big(\frac{1}{2}(A_1 + A_2)\Big),$$

from which the weak LDP holds for $\{\mu_n\}$ with rate function

$$I(x) = - \inf_{A \in \mathcal{A} : x \in A} \underline{\mathcal{L}}_A,$$

by Lemma 1.5. ☐

1.4. **Step 2: Full LDP.** The second hurdle is extending the weak convergence to the full LDP using the convexity of open sets.

Lemma 1.7. *Under the assumptions of Theorem 1.2, for every open convex subset $A \subset \mathcal{X}$, we have:*

$$\lim_{n\to\infty} \frac{1}{n} \log \mu_n(A) = - \inf_{x \in A} I(x),$$

where I is the convex rate function established in Lemma 1.6.

Proof. By the definition of our candidate rate function, we have $-I(x) \leq \underline{\mathcal{L}}_A$, which implies $-\inf_{x\in A} I(x) \leq \underline{\mathcal{L}}_A$. To obtain the reverse inequality, we employ subadditivity and exponential tightness. Recall that any open set A in a Polish space is itself Polish under a compatible metric, such as:

$$d'(x,y) = d(x,y) + \left| \frac{1}{d(x, A^c)} - \frac{1}{d(y, A^c)} \right|.$$

Consequently, for any $\delta > 0$, there exists $N \in \mathbb{N}$ and a compact set $C \subset A \cap \mathcal{E}$ such that:

$$\underline{\mathcal{L}}_A - 2\delta \leq \frac{1}{N} \log \mu_N(A) - \delta \leq \frac{1}{N} \log \mu_N(\overline{\text{co}}(C)) \leq \overline{\mathcal{L}}_{\overline{\text{co}}(C)} \leq - \inf_{x \in A} I(x),$$

where the third inequality is a result of Fekete's lemma. This completes the proof. ☐

Proof of Theorem 1.2. In view of Theorem 1.1, it suffices to show that

$$\frac{1}{n} \log \mathbb{E}\Big[e^{t\langle \lambda, n\hat{S}_n \rangle}\Big] = \log \mathbb{E}\big[e^{t\langle \lambda, X_1\rangle}\big] =: \Lambda_\lambda(t)$$

is lower semicontinuous and

$$\inf_{x : \langle \lambda, x\rangle > a} I(x) \leq \inf_{z > a} \Lambda_\lambda^*(z) \text{ for all } a \in \mathbb{R} \text{ and } \lambda \in \mathcal{X}^*.$$

The former follows obviously from Fatou's lemma. As for the latter, we consider the following two cases: (1) $\lambda = 0$ and (2) $\lambda \neq 0$.

(1) We assume $\lambda = 0$, so that

$$\Lambda_\lambda^*(x) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{if } x \neq 0, \end{cases}$$

Given that $\mu_n(\mathcal{X}) = 1$, it follows from Lemma 1.7 that $\inf_{x \in \mathcal{X}} I(x) = 0$. Hence, the proposed inequality holds.

(2) If $\lambda \neq 0$, the half-plane $H_a \coloneqq \{x : \langle \lambda, x \rangle > a\}$ is open and convex. By Lemma 1.7:

$$-\inf_{x \in H_a} I(x) = \overline{\mathcal{L}}_{H_a} \geq \sup_{\delta > 0} \overline{\mathcal{L}}_{H_{a+\delta}}.$$

To conclude the proof, define $Y_i = \langle \lambda, X_i \rangle$ and $\hat{Z}_n = \sum_{i=1}^n Y_i$, so the logarithmic moment generating function of Y_1 is $\Lambda_{\lambda(t)}$. By the one-dimensional Cramér's theorem, we have:

$$\inf_{\delta > 0} \overline{\mathcal{L}}_{H_{a+\delta}} = \sup_{\delta > 0} \left[-\inf_{z > a+\delta} \Lambda_\lambda^*(z) \right] = -\inf_{z > a} \Lambda_\lambda^*(z).$$

This proves the theorem. □

References

Research Unit of Mathematical Sciences, University of Oulu, P.O. Box 3000, 90014 University of Oulu, Finland.
Email address: `yu-liang.wu@oulu.fi`
URL: `s92077.github.io`