5. General principles

5.1. Contraction principles.

Theorem 5.1 (Contraction principle)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces and $f : \mathcal{X} \to \mathcal{Y}$ a continuous function. Consider a good rate function $I : X \to [0, \infty]$.*

(1) For each $y \in \mathcal{Y}$, define

$$I'(y) := \inf\{I(x) : x \in \mathcal{X}, y = f(x)\}.$$

Then, I' is a good rate function on $\mathcal{Y}$

(2) If I controls the LDP associated with a family of probability measures $\{\mu_\varepsilon\}$ on $\mathcal{X}$, then I' controls the LDP associated with the family of probability measures $\{\mu_\varepsilon \circ f^{-1}\}$ on $\mathcal{Y}$.

Proof. (1) I' is non-negative by definition. To prove the lower semicontinuity, observe that by compactness of $\Psi_I(\alpha)$,

$$\Psi_{I'}(\alpha) = \left\{y : \inf_{x:f(x)=y} I(x) \le \alpha\right\} = \{f(x) : I(x) \le \alpha\} = f(\Psi_I(\alpha))$$

is again compact as desired.

(2) Note that for each $A \subset \mathcal{Y}$, $\inf_{y\in A} I'(y) = \inf_{x\in f^{-1}(A)} I(x)$, where $f^{-1}(A)$ is open (respectively, closed) if A is open (respectively, closed). With this LDP is established naturally. □

Theorem 5.2 (Inverse contraction principle)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff topological spaces. Suppose that $g : \mathcal{Y} \to \mathcal{X}$ is a continuous bijection, and that $\{\nu_\varepsilon\}$ is an exponentially tight family of probability measures on $\mathcal{Y}$. If $\{\nu_\varepsilon \circ g^{-1}\}$ satisfies the LDP with the rate function $I : \mathcal{X} \to [0, \infty]$, then $\{\nu_\varepsilon\}$ satisfies the LDP with the good rate function $I' := I \circ g$.*

Proof. The non-negativity of I' is clear from definition, and, for every $\alpha < \infty$,

$$\Psi_{I'}(\alpha) = \{y : I(g(y)) \le \alpha\} = g^{-1}(\Psi_\alpha).$$

is closed, implying the lower semicontinuity of I'. Therefore, I' is a rate function. Now since $\{\nu_\varepsilon\}$ is exponentially tight, for every α, it remains to show that $\{\nu_\varepsilon\}$ satisfies the weak LDP.

For the upper bound, fix any compact set $K \subset \mathcal{Y}$ and apply the large deviation upper bound for $\nu_\varepsilon \circ g^{-1}$ to obtain

$$\limsup_{\varepsilon\to0} \varepsilon \log \nu_\varepsilon(K) = \limsup_{\varepsilon\to0} \varepsilon \log \nu_\varepsilon(g^{-1} \circ g(K)) \le -\inf_{y\in K} I'(y).$$

To demonstrate the lower bound, fix any $y \in \mathcal{Y}$ with $I(y) = \alpha < \infty$ and any open neighborhood $G \subset \mathcal{Y}$ of y. In addition, choose, thanks to the exponential tightness, a compact set K such that

$$\limsup_{\varepsilon\to0} \varepsilon \log \nu_\varepsilon(K) < -\alpha.$$

Hence, since $g(K^c) = g(K)^c$ is open (a consequence of Hausdorff topology), it follows from the LDP for $\nu_\varepsilon \circ g^{-1}$ that

$$-\alpha > \limsup_{\varepsilon\to0} \varepsilon \log \nu_\varepsilon(g^{-1} \circ g(K^c)) \ge -\inf_{x\in g(K^c)} I(x). \tag{5.1}$$

In addition, observe that $g^{-1}|_{g(K)}$ is continuous, there exists an open neighborhood G' of $g(y)$ such that

$$G' \subset g(G \cap K) \cup g(K^c) = g(G \cup K^c),$$

which in turn gives

$$\max\Big\{\limsup_{\varepsilon\to 0} \varepsilon \log \nu_\varepsilon(G \cap K), \limsup_{\varepsilon\to 0} \varepsilon \log \nu_\varepsilon(K^c)\Big\}$$
$$\geq \liminf_{\varepsilon} \varepsilon \log \nu_\varepsilon(g^{-1}(G')) \geq -I'(y).$$

Inequality (5.1) then rules out $\limsup_{\varepsilon\to 0} \varepsilon \log \nu_\varepsilon(K^c) \geq -I'(y)$ and asserts

$$\limsup_{\varepsilon\to 0} \varepsilon \log \nu_\varepsilon(G) \geq \limsup_{\varepsilon\to 0} \varepsilon \log \nu_\varepsilon(G \cap K) \geq -I'(y)$$

Since the neighborhood G of y is arbitrary, the proof of the lower bound is complete. □

Corollary 5.3. *Let $\{\mu_\varepsilon\}$ be an exponentially tight family of probability measures on X equipped with the topology τ_1. If $\{\mu_\varepsilon\}$ satisfies an LDP with respect to a Hausdorff topology τ_2 on X that is coarser than τ_1, then the same LDP holds with respect to the topology τ_1.*

Proof. It follows straightforwardly from Theorem 5.2. □

5.2. **Varadhan's integral lemma.**

In this subsection we assume $\mathcal{X}$ to be a regular topological space.

Theorem 5.4 (Varadhan)**.** *Suppose that $\{\mu_\varepsilon\}$ satisfies the LDP with a good rate function $I : \mathcal{X} \to [0,\infty]$, and let $\varphi : \mathcal{X} \to \mathbb{R}$ be any continuous function. Assume further either the tail condition*

$$\lim_{M\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon)\geq M\}}\Big] = -\infty. \tag{5.2}$$

or the following moment condition for some $\gamma > 1$,

$$\limsup_{\rho\to 0} \varepsilon \log \mathbb{E}\Big[e^{\frac{\gamma\varphi(Z_\varepsilon)}{\varepsilon}}\Big] < \infty. \tag{5.3}$$

Then,

$$\lim_{\rho\to 0} \varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}}\Big] = \sup_{x\in\mathcal{X}}[\varphi(x) - I(x)].$$

The theorem is a consequence of the following three lemmas.

Lemma 5.5. *If $\varphi : \mathcal{X} \to \mathbb{R}$ is lower semicontinuous and the large deviations lower bound holds with $I : \mathcal{X} \to [0,\infty]$, then*

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}}\Big) \geq \sup_{x\in\mathcal{X}}[\varphi(x) - I(x)].$$

Lemma 5.6. *If $\varphi : \mathcal{X} \to \mathbb{R}$ is an upper semicontinuous function for which the tail condition (5.2) holds, and the large deviations upper bound holds with the good rate function $I : \mathcal{X} \to [0,\infty]$, then*

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}}\Big) \leq \sup_{x\in\mathcal{X}}[\varphi(x) - I(x)].$$

Lemma 5.7. *Condition (5.3) implies the tail condition (5.2).*

Proof of Lemma 5.5. Fix $x \in \mathcal{X}$. Since φ is lower semicontinuous, for every $\delta > 0$ there exists open neighborhood G of x such that

$$-\varphi(x) \geq -\inf_{y \in G} \varphi(y) - \delta.$$

Hence, by Markov's inequality and the lower bound of LDP,

$$\begin{aligned}
&\liminf_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}}\Big) \geq \liminf_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{Z_\varepsilon \in G}\Big) \\
&\geq \inf_{y \in G} \varphi(y) + \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(G) \\
&\geq \inf_{y \in G} \varphi(y) - \inf_{y \in G} I(y) \geq \varphi(x) - I(x) - \delta,
\end{aligned}$$

proving the inequality since x and δ are arbitrary. □

Proof of Lemma 5.6. To begin, note that by semicontinuity of φ, I and regularity of $\mathcal{X}$, for each $x \in \mathcal{X}$ and each $\delta > 0$ there is an open neighborhood A_x such that

$$\inf_{y \in \overline{A_x}} I(y) \geq I(x) - \delta \quad \text{and} \quad \sup_{y \in A_x} \varphi(y) \leq \varphi(x) + \delta,$$

which gives a local estimate $\mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{Z_\varepsilon \in A_x}\Big) \leq e^{\frac{\varphi(x)+\delta}{\varepsilon}} \mu(\varphi^{-1}(A_x))$. Hence, for each $\alpha < \infty$, we may find a finite cover $\big\{A_{x_i}\big\}_{i=1}^N$ of the compact set $\Psi_I(\alpha)$ such that

$$\begin{aligned}
&\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon) \leq M\}}\Big] \\
&\leq \limsup_{\varepsilon \to 0} \max\Big\{M - \alpha, \max_{1 \leq i \leq N} \Big[\varphi(x_i) + \delta + \varepsilon \log \mu_\varepsilon\Big(A_{x_i}\Big)\Big]\Big\} \\
&\leq \max\Bigg\{M - \alpha, \max_{1 \leq i \leq N} \Bigg[\varphi(x_i) + \delta - \inf_{x \in \overline{A}_{x_i}} I(x)\Bigg]\Bigg\} \\
&\leq \max\Bigg\{M - \alpha, \max_{x \in \Psi_I(\alpha)} [\varphi(x) - I(x) + 2\delta]\Bigg\}.
\end{aligned}$$

First letting $\delta \to 0$ and then letting $\alpha \to 0$ yields

$$\mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon) \leq M\}}\Big] \leq \max_{x \in \mathcal{X}} [\varphi(x) - I(x)].$$

The proof is concluded, by assumption (5.2) that

$$\begin{aligned}
&\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\Big(e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}}\Big) \\
&= \lim_{M \to \infty} \limsup_{\varepsilon \to 0} \max\Big\{\varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon) \leq M\}}\Big], \varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon) > M\}}\Big]\Big\} \\
&= \lim_{M \to \infty} \limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{E}\Big[e^{\frac{\varphi(Z_\varepsilon)}{\varepsilon}} \mathbb{1}_{\{\varphi(Z_\varepsilon) \leq M\}}\Big].
\end{aligned}$$

□

Proof of Lemma 5.7. The lemma follows from the fact that for any non-negative random variable X and any $\gamma > 1$,

$$\mathbb{E}(X^\gamma \mathbb{1}_{X\geq\alpha}) = \alpha^\gamma \mathbb{P}(X \geq \alpha) + \int_\alpha^\infty x^{\gamma-1}\mathbb{P}(X \geq x)\, dx.$$

Hence, $\mathbb{E}(X^\gamma \mathbb{1}_{X\geq\alpha}) \geq \alpha^{1-\gamma}\mathbb{E}(X\mathbb{1}_{X\geq\alpha})$, which proves the theorem by taking $X = \varphi(Z_\varepsilon)/\varepsilon$ and $\alpha = \varphi(M)/\varepsilon$. □

5.3. **Bryc's inverse Varadhan lemma.**

Recall the following definition.

Definition 5.8. A topological space $\mathcal{X}$ is said to be completely regular if for any closed set F and any point $x \notin F$, there is continuous function $f : \mathcal{X} \to [0, 1]$ such that $f(x) = 1$ and $f(y) = 0$ for all $y \in F$.

For each Borel measurable function $f : \mathcal{X} \to \mathbb{R}$, define

$$\Lambda_f := \lim_{\varepsilon\to 0} \varepsilon \log \int_{\mathcal{X}} e^{\frac{f(x)}{\varepsilon}} d\mu_\varepsilon(x).$$

Theorem 5.9 (Bryc)**.** *Let $\mathcal{X}$ be a completely regular topological space and $\mathcal{B}_{\mathcal{X}} \subset \mathcal{B}$. Suppose that the family $\{\mu_\varepsilon\}$ is exponentially tight and that the limit Λ_f exists for every $f \in C_b(\mathcal{X})$. Then $\{\mu_\varepsilon\}$ satisfies the LDP with the good rate function*

$$I(x) = \sup_{f\in C_b(\mathcal{X})} \big\{f(x) - \Lambda_f\big\}. \tag{5.4}$$

Furthermore, for every $f \in C_b(\mathcal{X})$,

$$\Lambda_f = \sup_{x\in\mathcal{X}}\{f(x) - I(x)\}. \tag{5.5}$$

Proof. Since $\Lambda_0 = 0$, it follows that $I \geq 0$. Moreover, it is lower semicontinuous since it is the supremum of continuous functions. According to Remark 1.6, it suffices to show that $\{\mu_\varepsilon\}$ satisfies the weak LDP with rate I. Finally, by Varadhan's lemma (Theorem 5.4), (5.4) implies (5.5).

The proof of upper bound of the LDP is essentially Theorem 6.2(3) by, formally, replacing $\langle\lambda, x\rangle = e^{\lambda(x)}$ by $\lambda \in C_b(\mathcal{X})$.

For the lower bound, fix $x \in \mathcal{X}$ and an open neighborhood G of x. Since $\mathcal{X}$ is completely regular, choose $f \in C_b(\mathcal{X})$ such that $f(x) = 1$ and $f(y) = 0$ for all $y \in G^c$. For $m > 0$, define $f_m = m(f - 1)$ so that

$$\int e^{\frac{f_m(x)}{\varepsilon}}\, d\mu_\varepsilon(x) \leq e^{-\frac{m}{\varepsilon}} + \mu_\varepsilon(G) \leq e^{-\frac{m}{\varepsilon}} + \mu_\varepsilon(G).$$

Since $f_m \in C_b(\mathcal{X})$ and $f_m(x) = 0$,

$$\begin{aligned}
&\max\Big\{\liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(G), -m\Big\}\\
&\geq \{\liminf_{\varepsilon\to 0} \varepsilon \log \int e^{\frac{f_m(x)}{\varepsilon}}\, d\mu_\varepsilon(x) = \Lambda_{f_m}\\
&= -\big[f_m(x) - \Lambda_{f_m}\big] \geq -\sup_{f\in C_b(\mathcal{X})} \big\{f(x) - \Lambda_f\big\} = -I(x),
\end{aligned}$$

from which the lower bound follows by letting $m \to \infty$. □

5.4. **Large deviation on projective spaces.**

Let $(J, \leq)$ be a partially ordered, right-filtering set (for every $i, j \in J$, there exists $k \in J$ such that $i, j \leq k$) and $\{\mathcal{Y}_j\}_{j\in J}$. A *projective system* $(\mathcal{Y}_j, p_{ij})_{i\leq j}$ consists of topological vector spaces $\mathcal{Y}_j$ and continuous maps $p_{ij} : \mathcal{Y}_j \to \mathcal{Y}_i$ such that $p_{ij} \circ p_{jk} = p_{ik}$ and that p_{jj} is the identity map. The *projective limit* $\mathcal{X} = \lim_j \mathcal{Y}_j$ is defined as the (topological) subspace $\mathcal{X}$ of the product topological space $\mathcal{Y} = \prod_{j\in J} \mathcal{Y}_j$, defined as

$$\mathcal{X} = \big\{(y_j)_{j\in J} \in \mathcal{Y} : p_{ij}\big(y_j\big) = y_i, \forall i \leq j\big\},$$

which is naturally associated continuous map $p_j : \mathcal{X} \to \mathcal{Y}_j$.

Theorem 5.10 (Dawson-Gärtner)**.** *Let $\{\mu_\varepsilon\}$ be a family of probability measures on $\mathcal{X}$, such that for any $j \in J$ the Borel probability measures $\{\mu_\varepsilon \circ p^{-1}\}$ on Y_j satisfy the LDP with the good rate function I_j. Then $\{\mu_j\}$ satisfies the LDP with the good rate function*

$$I(\mathbf{x}) = \sup\big\{I_j\big(p_j(\mathbf{x})\big) : j \in J\big\} \text{ for } \mathbf{x} \in \mathcal{X}.$$

Proof. We first show that I is a good rate function. Obviously, I is nonnegative. Note that $\Psi_{I_j}(\alpha)$ is compact by goodness assumption of I_j. By the contraction principle, $I_i(y_i) = \inf_{y_j\in p_{ij}^{-1}(y_i)} I_j\big(y_j\big)$, which, under the goodness assumption, implies $\Psi_{I_i}(\alpha) = p_{ij} \circ \Psi_{I_j}(\alpha)$. Hence, $\Big(\Psi_{I_j}(\alpha), p_{ij}\Big)_{ij}$ is a projective system and

$$\Psi_I(\alpha) = \mathcal{X} \cap \prod_{j\in J} \Psi_{I_j}(\alpha)\Big(= \lim_j \Psi_{I_j}(\alpha)\Big).$$

The lower semicontinuity and goodness of $\Psi_I(\alpha)$ then follows by Tychonoff's theorem.

For lower bound, given any measurable set $A \subset \mathcal{X}$ and every $\mathbf{x} \in A$, there exists an open set $U_j \subset \mathcal{Y}_j$ such that $\mathbf{x} \in p_j^{-1}\big(U_j\big) \cap \mathring{A}$

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(A) \geq \liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon\big(p_j^{-1}\big(U_j\big)\big) \geq -I_j\big(p_j(\mathbf{x})\big) \geq -I(\mathbf{x}).$$

For the upper bound, we first claim that given any $A \subset \mathcal{X}$, we have that, with $A_j := p_j\big(\overline{A}\big)$, $\big(\overline{A_j}, p_{ij}\big)$ is a projective system and $\overline{A} = \lim_j \overline{A_j}$. Observe that $p_{ij}\Big(\overline{A_j}\Big) \subset \overline{A_i}$. Hence, $\overline{A} \subset \lim_j \overline{A_j}$. To prove the other inclusion, observe that for every $\mathbf{x} \in \overline{A}^c$, there exists some open set $U_j \subset \mathcal{Y}_j$ such that $\mathbf{x} \in p_j^{-1}\big(U_j\big) \subset \overline{A}^c$. This implies $p_j(\mathbf{x}) \in U_j \subset A_j^c$ and thus $p_j(\mathbf{x}) \notin \overline{A_j}$, proving $\overline{A} \supset \lim_j \overline{A_j}$. From this claim, we infer that

$$\overline{A} \cap \Psi_I(\alpha) = \lim_j \overline{A_j} \cap \Psi_{I_j}(\alpha) \quad \text{for all} \quad \alpha < \infty.$$

By taking $\alpha < \inf_{x\in\overline{A}} I(x)$, we obtained, by the finite intersection property, there exists j such that $\overline{A_j} \cap \Psi_{I_j}(\alpha)$, and consequently,

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(A) \leq \limsup_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon \circ p_j^{-1}\Big(\overline{A_j}\Big) \leq -\alpha.$$

This finishes the proof as α is arbitrary. $\square$