MA4L3 Large Deviation Theory
Example Sheet 1

2023, term 2
Stefan Adams

Students should hand in solutions to the questions by noon 12 pm Thursday week 3 (26.01.2023) online via the moodle page.

Exercise 1: [**Cramér's theorem**] In the setting of Cramér's theorem from the lecture the logarithmic moment generating function is $\Lambda(\lambda) := \log \mathbb{E}[\mathrm{e}^{\lambda X_1}]$, $\lambda \in \mathbb{R}$, and the rate function is $I(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}$, $x \in \mathbb{R}$.

(a) [5 points] Assume that $\mathbb{P}(X_1 = a) = 1$. Check that $I(a) = 0$ and $I(z) = \infty$ for $z \neq a$.

(b) [10 points] Compute the rate function in Cramér's theorem when $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 2) = \mathbb{P}(X_1 = 3) = \frac{1}{5}$ and $\mathbb{P}(X_1 = 4) = \frac{2}{5}$.

(c) [15 points] Compute the rate function I for the following distributions: Poisson and Exponential.

Exercise 2: [**Logarithmic moment generating function**]

Suppose that $(X_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence of $\mathbb{R}$-valued random variables with law $P \in \mathcal{M}_1(\mathbb{R})$ such that $\Phi(\lambda) := \mathbb{E}[\mathrm{e}^{\lambda X_1}] < \infty$ for all $\lambda \in \mathbb{R}$. The logarithmic moment generating function for the law P is

$$\Lambda(\lambda) := \log \mathbb{E}[\mathrm{e}^{\lambda X_1}] = \int_{\mathbb{R}} \mathrm{e}^{\lambda x}\, P(\mathrm{d}x)\,, \quad \lambda \in \mathbb{R}\,,$$

and its domain is $\mathcal{D}_\Lambda := \{\lambda \in \mathbb{R} \colon \Lambda(\lambda) < \infty\}$. We say Λ is steep at $\partial\mathcal{D}_\Lambda$ (or simply steep) if

$$\lim_{\lambda\to\partial\mathcal{D}_\Lambda,\, \lambda\in\mathcal{D}_\Lambda} |\Lambda'(\lambda)| = \infty\,.$$

(a) [10 points] Use Hölder's inequality and Fatou's lemma to show that $\lambda \mapsto \log \mathbb{E}[\mathrm{e}^{\lambda X_1}]$ is convex and lower-semicontinuous on $\mathbb{R}$.

(b) [10 points] Show that the steepness of Λ implies that $\mathcal{D}_I = \mathbb{R}$, where I is the rate function in Cramér's theorem,

$$I(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}\,.$$

Exercise 3: [**Sanov's theorem and pair empirical measures**]

(a) [15 points] Let Ω be a finite set and $\nu \in \mathcal{M}_1(\Omega)$ be a probability measure. Show that the relative entropy

$$H(\cdot|\nu) \colon \mathcal{M}_1(\Omega) \to [0,\infty], \mu \mapsto H(\mu|\nu) =: I_\nu(\mu)$$

has the following properties:

(i) I_ν is finite, continuous and strictly convex on $\mathcal{M}_1(\Omega)$.

(ii) $I_\nu(\mu) \geq 0$ with equality if and only if $\mu = \nu$.

(b) [15 points] Let Ω be a finite set and $\nu \in \mathcal{M}_1(\Omega)$. Show with the help of the SLLN (Strong law of large numbers) that

$$\mathrm{d}(L_n^{(2)}, \nu \otimes \nu) \to 0 \text{ as } n \to \infty \qquad \mathbb{P} - \text{a.s.},$$

where $L_n^{(2)} = \frac{1}{n} \sum_{i=1}^n \delta_{(X_i, X_{i+1})}$ is the pair empirical measure with periodic boundary condition $X_{n+1} = X_1$ and d is the total variation distance.

Exercise 4: [**General theory**] [20 points] Prove the following statement: A sequence of probability measures $(\mu_n)_{n \in \mathbb{N}}$ on a Polish space can have at most one rate function associated with its large deviation principle (LDP).

Hand-in times (submission on moodle page) – assignment sheets:

Sheet	Week	Date submission	Discussion in Support class/lecture
1	3	26.01.2023 at noon 12 pm	Week 4, Wednesday 01.02.2023
2	5	09.02.2023 at noon 12 pm	Week 6, Wednesday 15.02.2023
3	8	02.03.2023 at noon 12 pm	Week 9, Wednesday 08.03.2023
4	10	16.03.2023 at noon 12 pm	Week 10, Friday 17.03.2023