# EXAM: CONVERGENCE OF RANDOM VARIABLES AND LARGE DEVIATIONS

**Problem 1.** We consider a sequence of i.i.d. random variables $(X_n)_{n\in\mathbb{N}}$ with values in a finite set $[\![1,N]\!]$, and with common distribution $\pi \in \mathscr{M}^1([\![1,N]\!])$, such that $\pi(i) > 0$ for every $i \in [\![1,N]\!]$.

(1) Write the large deviation principle for the sequence of empirical measures

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}.$$

Check that the rate function $I(\nu)$ of this LDP is a continuous convex function on $\mathscr{M}^1([\![1,N]\!])$, and that it is differentiable on the dense open set

$$O = \{\nu \in \mathscr{M}^1([\![1,N]\!]) \mid \forall i \in [\![1,N]\!]\,,\ \nu(i) > 0\}.$$

(2) Compute the derivative $dI_\nu = \left(\frac{\partial I}{\partial \nu(1)}(\nu), \ldots, \frac{\partial I}{\partial \nu(N)}(\nu)\right)$.

(3) One fixes a state $k \in [\![1,N]\!]$ and a real number $\theta \in (0,1)$. By applying Lagrange's principle, compute

$$\inf\left\{I(\nu) \mid \nu \in \mathscr{M}^1([\![1,N]\!]) \text{ and } \nu(k) = \theta\right\}.$$

Hint: there are two constraints $G(\nu) = 1$ and $H(\nu) = \theta$, so at the minimizer $\nu$ one should have $dI_\nu = \alpha\, dG_\nu + \beta\, dH_\nu$ for some constants $\alpha, \beta$.

(4) Write the scaled occupation time $T_{k,n} = \frac{\operatorname{card}\{i\in[\![1,n]\!] \mid X_i = k\}}{n}$ of the state $k$ as a continuous function of $\nu_n$. By using the contraction principle, show that $T_{k,n}$ satisfies a large deviation principle, and give the corresponding rate function.

(5) Recover this result by using Cramér's theorem.

(6) More generally, consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ with irreducible transition matrix $p$ on the space $[\![1,N]\!]$, and initial distribution $\pi_0$. Write the Laplace transform $\mathbb{E}[\mathrm{e}^{n\,T_{k,n} t}]$ in terms of the positive matrices

$$p_{k,t}(x,y) = \begin{cases} p(x,y) & \text{if } y \neq k \\ p(x,y)\,\mathrm{e}^t & \text{if } y = k. \end{cases}$$

Use Ellis-Gärtner theory to state a LDP for $(T_{n,k})_{n\in\mathbb{N}}$ in this general case.

**Problem 2.** One denotes $\mathscr{D}$ the vector space of real-valued functions $f : [0,1] \to \mathbb{R}$ such that $f(0) = 0$ and for every $t \in [0,1]$,

$$\lim_{\substack{s\to t\\ s<t}} f(s) \quad \text{and} \quad \lim_{\substack{s\to t\\ s>t}} f(s)$$

exist, the second limit being equal to $f(t)$. If the first limit is not equal to $f(t)$, one says that $f$ has a discontinuity, or a jump at $t$. For $f$ in $\mathscr{D}$ and $\delta > 0$, one sets

$$\omega(f,\delta) = \inf\left\{ \sup_{i\in[\![1,r]\!],\ t_{i-1}\leq x< t_i} |f(x) - f(t_{i-1})| \right\},$$

*Date*: May 14, 2013.

where the infimum is taken over finite subdivisions $0 = t_0 < t_1 < t_2 < \cdots < t_r = 1$ of the interval $[0, 1]$ that are $\delta$-sparse, that is to say that $t_i - t_{i-1} \geq \delta$ for all $i$. One admits that there is a topology of $\mathscr{D}$ that makes it a polish space (separable complete metric space), and such that :

(i) The relatively compact subsets $\mathscr{F} \subset \mathscr{D}$ are those such that

$$\lim_{\delta \to 0} \left( \sup_{f \in \mathscr{F}} \omega(f, \delta) \right) = 0.$$

(ii) A probability measure on $\mathscr{D}$ is entirely determined by its images by the measurable maps

$$\pi_{t_1 < t_2 < \cdots < t_r}(f) = (f(t_1), \ldots, f(t_r)) \in \mathbb{R}^r.$$

Let $(U_k)_{k \in \mathbb{N}}$ be a sequence of i.i.d. random variables that are uniformly distributed on $[0, 1]$: for every $x \in [0, 1]$, $\mathbb{P}[U_k \leq x] = x$.

(1) Let

$$X_{n,t} = \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{1}_{(U_k \leq 1/n)},$$

where $\lfloor nt \rfloor$ denotes the entire part of $nt$. Show that the path $X_n : t \mapsto X_{n,t}$ falls almost surely in $\mathscr{D}$.

(2) Show that the number $\kappa(n)$ of discontinuities of $X_n$ satisfies:

$$\lim_{n \to \infty} \mathbb{P}[\kappa(n) = k] = \frac{1}{\mathrm{e}\, k!}.$$

(3) Show that conditionnally to the event $\kappa(n) = k$, the positions of the jumps $t_1 < t_2 < \cdots < t_k$ of $X_n$ satisfy

$$\begin{aligned}&\text{law of } (nt_1, nt_2, \ldots, nt_k)\\ &= \text{uniform law on the set } \mathfrak{P}_k([\![1, n]\!]) \text{ of subsets } (n_1, \ldots, n_k) \text{ of size } k \text{ in } [\![1, n]\!].\end{aligned}$$

Conditionnally to the same event $\kappa(n) = k$, show that the probability that the random path $X_n$ has two consecutive jumps $t_{i-1}$ and $t_i$ with $t_i - t_{i-1} \leq \delta$ is smaller than

$$\frac{(k-1)\, \lfloor n\delta \rfloor \, \binom{n}{k-1}}{\binom{n}{k}} \leq C(k)\, \delta$$

for some constant $C(k)$.

(4) Show that for any $\delta > 0$,

$$\lim_{\delta \to 0} \left( \sup_{n \in \mathbb{N}} \mathbb{P}[X_n \text{ has consecutive jumps separated by less than } \delta] \right) = 0.$$

(5) Show that the random paths $(X_n)_{n \in \mathbb{N}}$ have laws $(\mu_n)_{n \in \mathbb{N}}$ that form a tight sequence.

(6) Describe the limiting law of $\pi_{t_1 < \cdots < t_r}(X_n)$, and show that $(X_n)_{n \in \mathbb{N}}$ has a limit in law in $\mathscr{D}$. What is this limiting random process?