**Problem 1.**

(1-2) By Sanov's theorem for i.i.d. variables, $(\nu_n)_{n\in\mathbb{N}}$ satisfies a LDP with rate function

$$I(\nu) = H(\nu||\pi) = \sum_{k=1}^{N} \nu(k)\,\log\left(\frac{\nu(k)}{\pi(k)}\right).$$

For every $k$, the map $\phi_k : x \mapsto x\,\log(x/\pi(k))$ is well-defined and continuous on $[0,1]$, with $\phi_k(0) = 0$; and it is convex since its second derivative is $1/x > 0$. As a positive linear combination of the functions $\phi_k(\nu(k))$ of the coordinates of $\nu$, $I(\nu)$ is itself continuous and convex. On the open set $O$, the partial derivative of $I_\nu$ with respect to $\nu(k)$ is

$$\frac{\partial I(\nu)}{\partial \nu(k)} = \log\left(\frac{\nu(k)}{\pi(k)}\right) + 1.$$

Since these partial derivatives are continuous on $O$ with respect to all coordinates of $\nu$, $I$ is continuously differentiable on $O$.

(3) Notice that by strict convexity of $I$, the infimum of $I$ under the linear constraints given is attained in the interior of $\mathscr{M}([\![1,N]\!])$, which is $O$. There, one can apply Lagrange's principle: under the constraints $G(\nu) = \sum_{j=1}^{N} \nu(j) = 1$ and $H(\nu) = \nu(k) = \theta$, the minimizer $\nu$ satisfies the equation:

$$\begin{aligned} &dI_\nu = \alpha\,dG_\nu + \beta\,dH_\nu; \\ \iff\quad &\sum_{j=1}^{N}\left(1+\log\left(\frac{\nu(j)}{\pi(j)}\right)\right)d\nu(j) = \alpha\left(\sum_{j=1}^{N} d\nu(j)\right) + \beta\,d\nu(k). \end{aligned}$$

for some constants $(\alpha,\beta)$. This means that if $j \neq k$, then

$$1+\log\left(\frac{\nu(j)}{\pi(j)}\right) = \alpha \qquad ; \qquad \frac{\nu(j)}{\pi(j)} = \mathrm{e}^{\alpha-1} = A,$$

and if $j = k$, then

$$1+\log\left(\frac{\nu(k)}{\pi(k)}\right) = \alpha+\beta \qquad ; \qquad \frac{\nu(k)}{\pi(k)} = \mathrm{e}^{\alpha+\beta-1} = B.$$

The constants $A$ and $B$ are determined by the constraints: $B = \theta/\pi(k)$ and then

$$1-\theta = \sum_{j\neq k} \nu(j) = \sum_{j\neq k} A\,\pi(j) = A\,(1-\pi(k)) \qquad ; \qquad A = \frac{1-\theta}{1-\pi(k)}.$$

One has therefore

$$\begin{aligned} I(\nu) &= \sum_{j=1}^{N} \nu(j)\,\log\left(\frac{\nu(j)}{\pi(j)}\right) = (1-\theta)\log A + \theta\log B \\ &= (1-\theta)\,\log\left(\frac{1-\theta}{1-\pi(k)}\right) + \theta\,\log\left(\frac{\theta}{\pi(k)}\right). \end{aligned}$$

(4) One has $T_{k,n} = H(\nu_n)$ with $H(\nu) = \nu(k)$, which is continuous. We know that $(\nu_n)_{n\in\mathbb{N}}$ satisfies a LDP with rate function $I$, and it is automatically good since

$\mathscr{M}^1(\llbracket 1, N\rrbracket)$, the space of probability measures on a finite set, is compact. Therefore, one can use the contraction principle, and $T_{k,n}$ satisfies a LDP with rate function

$$\inf\{I(\nu) \mid G(\nu) = \theta\} = f(\theta) = (1-\theta)\,\log\left(\frac{1-\theta}{1-\pi(k)}\right) + \theta\,\log\left(\frac{\theta}{\pi(k)}\right).$$

(5) The variable $T_{k,n}$ is the mean of the i.i.d. Bernoulli random variables $\mathbf{1}_{X_n=k}$, that have parameter $\pi(k)$. By Cramér's theorem, $T_{k,n}$ satisfies a LDP with rate function the Legendre-Fenchel transform of the log-Laplace transform of these Bernoulli variables, which is

$$\Lambda(t) = \log\big(\pi(k)\,\mathrm{e}^t + (1-\pi(k))\big).$$

One computes the LF transform as follows:

$$\begin{aligned}
\Lambda^*(\theta) &= \sup_{t\in\mathbb{R}}\,(\theta t - \Lambda(t)) \\
&= \theta t_0 - \Lambda(t_0) \quad \text{with } \Lambda'(t_0) = \theta; \\
\theta &= \frac{\pi(k)\,\mathrm{e}^{t_0}}{(1-\pi(k)) + \pi(k)\,\mathrm{e}^{t_0}}; \\
t_0 &= \log\left(\frac{\theta\,(1-\pi(k))}{\pi(k)\,(1-\theta)}\right); \\
\Lambda^*(\theta) &= (1-\theta)\,\log\left(\frac{1-\theta}{1-\pi(k)}\right) + \theta\,\log\left(\frac{\theta}{\pi(k)}\right).
\end{aligned}$$

Thus one has recovered the previous result.

(6) One computes

$$\begin{aligned}
\mathbb{E}[\mathrm{e}^{n\,T_{k,n} t}] &= \sum_{x_1,\ldots,x_n} \mathbb{P}[X_1 = x_1,\ldots,X_n = x_n]\,\mathrm{e}^{t\,\mathrm{card}\,\{i \mid x_i = k\}} \\
&= \sum_{x_0,x_1,\ldots,x_n} \pi_0(x_0)\,p(x_0,x_1)\cdots p(x_{n-1},x_n)\,\mathrm{e}^{t\,\mathrm{card}\,\{i \mid x_i = k\}} \\
&= \sum_{x_0,x_1,\ldots,x_n} \pi_0(x_0)\,p^{k,t}(x_0,x_1)\cdots p^{k,t}(x_{n-1},x_n) \\
&= \sum_x \left(\pi_0 (p_{k,t})^n\right)(x).
\end{aligned}$$

As in the lecture, if $r(p_{k,t})$ is the Perron-Frobenius eigenvalue of the positive matrix $p_{k,t}$ and $\pi_{k,t}$ is the corresponding (normalized) positive eigenvector, then

$$\pi_0 = \alpha\,\pi_{k,t} + \text{remainder corresponding to lesser eigenvalues},$$

and

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[\mathrm{e}^{n\,T_{k,n} t}] = \lim_{n\to\infty}\frac{1}{n}\,\log(\alpha\,(r(p_{k,t}))^n) = \log r(p_{k,t}) = \Lambda(t).$$

Notice that the random variables $T_{k,n}$ take their values in the compact set $[0,1]$, so their laws are automatically exponentially tight. One can therefore apply Ellis-Gärtner theorem, which almost gives a LDP except for the problem of exposed points. However, the function $\Lambda$ is finite and defined over $\mathbb{R}$, and as the logarithm of the largest root of a polynomial which depends smoothly on $t$, it is differentiable in $t$, except maybe at a finite number of values $t$, say $(t_1,\ldots,t_M)$. This still implies the density of exposed points and therefore a full LDP for $(T_{n,k})_{n\in\mathbb{N}}$, with good rate function

$$\Lambda^*(\theta) = \sup_{t\in\mathbb{R}} \left(\theta t - \log r(p_{t,k})\right).$$

**Problem 2.**

(1) Fix $\omega$ in the probability space. If $s$ is not an integer multiple of $\frac{1}{n}$, then the path $X_n(\omega)$ is constant around $s$, equal to $\sum_{k=1}^{\lfloor ns\rfloor} \mathbf{1}_{(U_k(\omega)\leq 1/n)}$ (and therefore continuous at $s$). Indeed, the number of terms $\lfloor ns\rfloor$ of the sum stays constant on the interval

$$\frac{\lfloor ns\rfloor}{n} < s' < \frac{\lfloor ns\rfloor + 1}{n}.$$

On the other hand, if $s$ is a multiple of $\frac{1}{n}$, then $X_n(\omega)$ is still constant equal to $\sum_{k=1}^{ns} \mathbf{1}_{(U_k(\omega)\leq 1/n)}$ on an interval to the right of $s$, namely, the interval

$$\frac{ns}{n} \leq s' < \frac{ns+1}{n}.$$

This implies the continuity on the right; and $X_n$ has also a limit to the left of $s$, given by $\sum_{k=1}^{ns-1} \mathbf{1}_{(U_k(\omega)\leq 1/n)}$. So for every $\omega$, $X_n(\omega)$ is indeed in $\mathscr{D}$.

(2) The jumps of $X_n$ occur at multiples of $\frac{1}{n}$, and more precisely, $X_{n,s} \neq X_{n,s^-}$ if and only if $s = \frac{k}{n}$ and $U_k \leq \frac{1}{n}$. In this case $X_{n,s} - X_{n,s^-} = \mathbf{1}_{(U_k\leq 1/n)}$, so the number of discontinuities $\kappa(n)$ of $X_n$ can be encoded by

$$\begin{aligned}\kappa(n) &= \sum_{k=1}^{n} \mathbf{1}_{(X_n \text{ makes a jump at } s = k/n)} \\ &= \sum_{k=1}^{n} X_{n,k/n} - X_{n,(k/n)^-} = \sum_{k=1}^{n} \mathbf{1}_{(U_k\leq 1/n)} = X_{n,1}.\end{aligned}$$

So we have to compute the limiting law of $X_{n,1}$, which is a sum of $n$ independent Bernoulli variables of parameter $1/n$. The characteristic function of $X_{n,1}$ is

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\zeta X_{n,1}}] = (\mathbb{E}[\mathrm{e}^{\mathrm{i}\zeta\,\mathcal{B}(n^{-1})}])^n = \left(1 + \frac{\mathrm{e}^{\mathrm{i}\zeta} - 1}{n}\right)^n \to_{n\to\infty} \mathrm{e}^{\mathrm{e}^{\mathrm{i}\zeta}-1}.$$

This is well-known to be the characteristic function of a Poisson random variable, as is seen from the calculation

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\zeta\mathcal{P}}] = \sum_{k=0}^{\infty} \frac{1}{\mathrm{e}\,k!}\,\mathrm{e}^{\mathrm{i}k\zeta} = \frac{1}{\mathrm{e}}\,\mathrm{e}^{\mathrm{e}^{\mathrm{i}\zeta}} = \mathrm{e}^{\mathrm{e}^{\mathrm{i}\zeta}-1}.$$

By Lévy criterion of convergence in law, $X_{n,1}$ converges to a Poisson random variable, so one has indeed

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

(3) One looks at $n$ random i.i.d. Bernoulli variables $B_1,\ldots,B_n$, and conditionally to the event that $k$ of them are equal to 1 (and the $n-k$ other equal to 0), one wants to know the law of the $k$-tuples of ordered indices $(n_1 < n_2 < \cdots < n_k)$ such that $B_{n_1} = B_{n_2} = \cdots = B_{n_k} = 1$. However, the law of $(B_1,\ldots,B_n)$ is clearly invariant by permutation of the variables. Fix two $k$-tuples $(n_1 < n_2 < \cdots < n_k)$

and $(m_1 < m_2 < \cdots < m_k)$ and a permutation $\sigma : [\![1,n]\!] \to [\![1,n]\!]$ such that $\sigma(m_i) = n_i$. One has:

$$\begin{aligned}
&\mathbb{P}\left[B_{m_1} = \cdots = B_{m_k} = 1 \mid \sum_{j=1}^n B_j = k\right] \\
&= \frac{\mathbb{P}[B_{m_1} = \cdots = B_{m_k} = 1 \text{ and } B_{j\neq m_i} = 0]}{\mathbb{P}[\sum_{i=1}^n B_j = k]} \\
&= \frac{\mathbb{P}[B_{n_1} = \cdots = B_{n_k} = 1 \text{ and } B_{j\neq n_i} = 0]}{\mathbb{P}[\sum_{i=1}^n B_j = k]} \quad \text{by } \sigma\text{-invariance of } \mathbb{P}; \\
&= \mathbb{P}\left[B_{n_1} = \cdots = B_{n_k} = 1 \mid \sum_{j=1}^n B_j = k\right].
\end{aligned}$$

It follows that the probability of each $k$-tuple is the same, and is therefore the uniform probability $\frac{1}{\binom{n}{k}}$ on subsets of size $k$ in $[\![1,n]\!]$.

Now, if some jump-times of $X_n$ have distance less than $\delta$, then the corresponding set of integers $(n_1 < n_2 < \ldots < n_k)$ contains a pair $n_i < n_{i+1}$ with $i \in [\![1, k-1]\!]$ and $n_{i+1} \in [\![n_i + 1, n_i + n\delta]\!]$. However, to construct such a $k$-tuple, one can

(a) choose an index $i$: $(k-1)$ possibilities,

(b) the set $n_1 < \cdots < n_i < n_{i+2} < \cdots < n_k$: $\binom{n}{k-1}$ possibilities,

(c) and then $n_{i+1}$ in an interval of size $\lfloor n\delta \rfloor$ to the right of $n_i$: $\lfloor n\delta \rfloor$ possibilities.

Therefore,

$$\begin{aligned}
\mathbb{P}[X_n \text{ has jumps separated by less than } \delta \mid \kappa(n) = k] &\leq \frac{(k-1)\lfloor n\delta \rfloor \binom{n}{k-1}}{\binom{n}{k}} \\
&\leq \frac{(k-1)\,k\,n}{n-k+1}\,\delta \leq k^3\,\delta
\end{aligned}$$

by taking the supremum in $n$, obtained for $n = k$ (one always has $n \geq \kappa(n) = k$).

(4) Fix $\varepsilon > 0$. For $K$ big enough, $\sum_{k>K} \frac{1}{\mathrm{e}\,k!} \leq \varepsilon$, and by Question (2), for $n \geq N$, one has therefore

$$\mathbb{P}[\kappa(n) > K] \leq 2\varepsilon,$$

so

$$\begin{aligned}
&\mathbb{P}[X_n \text{ has jumps separated by less than } \delta] \\
&\leq \mathbb{P}[\kappa(n) > K] + \mathbb{P}[X_n \text{ has jumps separated by less than } \delta \text{ and } \kappa(n) \leq K] \\
&\leq 2\varepsilon + \sum_{k\leq K} \mathbb{P}[X_n \text{ has jumps separated by less than } \delta \mid \kappa(n) = k]\,\mathbb{P}[\kappa(n) = k] \\
&\leq 2\varepsilon + \delta \sum_{k\leq K} k^3\,\mathbb{P}[\kappa(n) = k] \leq 2\varepsilon + K^4\delta.
\end{aligned}$$

This is true for $n \leq N$. On the other hand, for $\delta < \frac{1}{N}$, $X_n$ with $n \leq N$ cannot have two jumps separated by less than $\delta$ (they occur at multiple integers of $\frac{1}{n} > \delta$), so the same probability is 0. Hence, for any $\delta$ small enough,

$$\sup_{n\in\mathbb{N}} \mathbb{P}[X_n \text{ has jumps separated by less than } \delta] \leq 2\varepsilon + K^4\delta$$

$$\limsup_{\delta\to 0}\left(\sup_{n\in\mathbb{N}} \mathbb{P}[X_n \text{ has jumps separated by less than } \delta]\right) \leq 2\varepsilon.$$

This is true for every $\varepsilon > 0$, whence the result.

(5) Fix $\varepsilon > 0$: we have to exhibit a relatively compact set $\mathscr{F} \subset \mathscr{D}$ such that $\mu_n(\mathscr{F}) \geq 1-\varepsilon$ for every $n$. Notice that if $X_n$ has no jumps separated by less than $\delta$, then one can find a subdivision of $[0,1]$ that is $\delta$-sparse and such that

$$\max_{i\in[\![1,r]\!]} \sup_{x\in[t_{i-1},t_i)} |X_{n,t_{i-1}} - X_{n,x}| = 0,$$

namely, the subdivision given by the positions of its jumps (plus the endpoints 0 and 1). So,

$$\mathbb{P}[X_n \text{ has no jumps separated by less than } \delta] \leq \mathbb{P}[\omega(X_n,\delta) = 0] = \mu_n[\omega(\cdot,\delta) = 0].$$

By the previous question, for $\delta$ small enough, the left-hand side is always larger than $1-\varepsilon$, and the right hand side is the $\mu_n$-probability of a relatively compact part of $\mathscr{D}$ (by the assumption (i) on the topology of this space). The tightness is therefore shown.

(6) For any $t_1 < t_2 < \cdots < t_r$, the same computations as in Question (2) show that $(X_{n,t_1},\ldots,X_{n,t_r})$ converge towards a vector of independent Poisson variables of parameter $t_1, t_2 - t_1, \ldots, t_r - t_{r-1}$:

$$\lim_{n\to\infty} \mathbb{P}[X_{n,t_1} = k_1, X_{n,t_2} = k_2, \ldots, X_{n,t_r} = k_r] = \frac{(t_1)^{k_1}\,(t_2-t_1)^{k_2}\cdots(t_r-t_{r-1})^{k_r}}{\mathrm{e}^{t_r}\,k_1!\,k_2!\,\cdots\,k_r!}.$$

Since the finite-dimensional laws identify probability measures on $\mathscr{D}$ (by the assumption (ii) on the topology of this space), the limit of a convergent subsequence of the laws $(\mu_n)_{n\in\mathbb{N}}$ is uniquely determined by the previous identity, so by tightness $(\mu_n)_{n\in\mathbb{N}}$ converges indeed. The limiting random process is the standard Poisson process on the interval $[0,1]$.