DATA SCIENCE **Stochastic Methods**

December 16, 2022

Problem 1. [8] Let $X \sim U(0,2)$ and $Y \sim U(0,2)$ be two INDEPENDENT uniform random variables and let $Z = \max\{X, Y\}$.

(i) Compute $P[Z \leq z]$ for $z \in \mathbb{R}$;

(ii) Compute $E[Z]$;

(iii) Compute $E[1/Z]$.

(i) $$\mathbb{P}[Z \leq z] = \mathbb{P}[\max\{X,Y\} \leq z] = \mathbb{P}[X \leq z,\ Y \leq z]$$

$$\overset{(\text{IND.})}{=} \mathbb{P}[X \leq z] \cdot \mathbb{P}[Y \leq z] = \begin{cases} 0 & z < 0 \\ z^2/4 & 0 \leq z < 2 \\ 1 & z \geq 2 \end{cases}$$

Since $Z \geq 0$,

(ii) $$\mathbb{E}[Z] = \int_0^{+\infty} \mathbb{P}[Z > z]\,dz = \int_0^2 \left(1 - z^2/4\right) dz =$$

$$= \left[z - \frac{z^3}{12} \right]_0^2 = 2 - \frac{8}{12} = 2 - \frac{2}{3} = \frac{6-2}{3} = \frac{4}{3}$$

(iii) $$f_Z(z) = z/2 \ \mathbb{1}_{(0,2)}(z)$$

$$\mathbb{E}[1/Z] = \int_0^2 1/z \cdot z/2 \, dz = \int_0^2 1/2 \, dz = \left[\frac{z}{2}\right]_0^2 = 1$$

Problem 2. [8] Let $X \sim Bin(1,p)$ and Y be a Poisson random variable with (random) parameter $X+1$, i.e. $\mathrm{P}[Y=k|X=x] = e^{-(x+1)}\,\frac{(x+1)^k}{k!}$ for $x=0,1$ and $k \in \mathbb{N}$.

(i) Compute $\mathrm{P}[Y=k]$ for any $k \in \mathbb{N}$;

(ii) Compute $\mathrm{E}[Y]$.

(i) $$\mathbb{P}[Y=k] = \mathbb{P}[Y=k \,|\, X=0]\cdot\mathbb{P}[X_0=0] + \mathbb{P}[Y=k \,|\, X=1]\cdot\mathbb{P}[X=1]$$

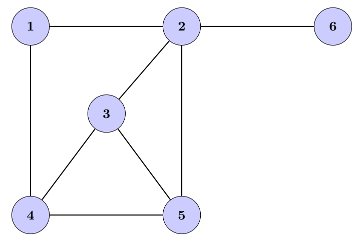
$$= e^{-1}\frac{1}{k!}(1-p) + e^{-2}\frac{2^k}{k!}\cdot p$$

(ii) $$\mathbb{E}[Y] = \sum_{k=0}^{+\infty} k\left(e^{-1}\frac{1}{k!}(1-p) + e^{-2}\frac{2^k}{k!}p\right)$$

$$= (1-p)\sum_{k=0}^{+\infty} k\, e^{-1}\frac{1}{k!} + p\sum_{k=0}^{+\infty} k\cdot e^{-2}\frac{2^k}{k!}$$

$$= (1-p)\,1 + p\cdot 2 = 1+p$$

Problem 3. [8] Define a simple Random Walk $\{X_n, n \geq 0\}$ on the undirected graph:



(i) Compute the probability to go from 6 to 3 in three steps.
(ii) Is the chain irreducible? Is the chain aperiodic?
(iii) Find the invariant distribution.
(iv) Starting from state 6, how many steps are needed on average to go back to 6?

(i) $\mathbb{P}[X_3 = 3 \mid X_0 = 6] = P_{62} \cdot P_{25} \cdot P_{53} = 1 \cdot \frac{1}{3} \cdot \frac{1}{4} = 1/12$

(ii) irreducible (the graph is connected), aperiodic

$(P^2_{22} > 0$ and $P^3_{22} > 0)$.

(iii) $P_1 = 2, P_2 = 4, P_3 = 3, P_4 = 3, P_5 = 3, P_6 = 1, \sum_i P_i = 16$

$(\pi_1, \ldots, \pi_6) = (2/16, 4/16, 3/16, 3/16, 3/16, 1/16)$

(iv) $m_6 = \mathbb{E}[T_6 \mid X_0 = 6] = \frac{1}{\pi_6} = 16$

Problem 4. [10] Let $(X_i)_{1\le i\le n}$ be a family of i.i.d. *Poisson*(λ) random variables.

(i) Compute the moment generating function of X_1;

(ii) Prove that if $\lambda = 1$, then $X_1 + \ldots + X_n$ is a *Poisson*(n) random variable;

(iii) Defined $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and taking again $\lambda = 1$, determine an exponential decay for the "lower tail" of $\overline{X}_n - 1$.

(i) $m_{X_1}(t) = \mathbb{E}[e^{tX_1}] = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} =$

$= e^{-\lambda} \cdot \sum_{k=0}^{+\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t - 1)}$

(ii) if $\lambda = 1$, $m_{X_1}(t) = \ldots = m_{X_n}(t) = e^{(e^t - 1)}$

$m_{X_1+\ldots+X_n}(t) = \prod_{i=1}^n m_{X_i}(t) = e^{n(e^t-1)} \sim$ mgf - Poisson(n)

(iii) Let $\overline{X}_n = 1/n \sum_{i=1}^n X_i$, $\mathbb{E}[\overline{X}_n] = 1$. The lower tail estimate

for $\overline{X}_n - 1 = \overline{X}_n - \mathbb{E}[\overline{X}_n]$ is

$\mathbb{P}[\overline{X}_n - 1 \le -\varepsilon] = \mathbb{P}[\overline{X}_n \le 1-\varepsilon] \le e^{-n \cdot \varepsilon^2/2}$ $\forall\, 0<\varepsilon<1$

Indeed,

$\mathbb{P}[\overline{X}_n \le 1 - \varepsilon] \le e^{-n h(t)}$

where $h(t) = -t(1-\varepsilon) - \lg m_{X_1}(-t) = -t(1-\varepsilon) - (e^{-t} - 1)$

$h'(t) = 0 \iff t = t^* = -\lg(1-\varepsilon)$. Since $\varepsilon + (1-\varepsilon)\lg(1-\varepsilon) \ge \varepsilon^2/2$

for $0<\varepsilon<1$, we get $h(t^*) \ge \varepsilon^2/2$ and

$\mathbb{P}[\overline{X}_n \le 1-\varepsilon] \le e^{-n\varepsilon^2/2}$ $\forall\, 0<\varepsilon<1$