DATA SCIENCE **Stochastic Methods**

February 7, 2023

Problem 1. [10] Let X_1, X_2 and X_3 be three independent Binomial random variables with parameters $(1,p)$, and define $Y = X_1 + X_2 - X_3$.

(i) Compute $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$;

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

(iii) Compute the characteristic function of Y.

(i) $\mathbb{E}[Y] = \mathbb{E}[X_1] + \mathbb{E}[X_2] - \mathbb{E}[X_3] = p + p - p = p$

$\mathrm{Var}[Y] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \mathrm{Var}[-X_3] = p(1-p) + p(1-p) + p(1-p) = 3p(1-p)$

(ii) $$Y = \begin{cases} -1 & p(1-p)^2 \\ 0 & (1-p)^3 + 2p^2(1-p) \\ 1 & p^3 + 2p(1-p)^2 \\ 2 & p^2(1-p) \end{cases}$$

(iii) $$\varphi_Y(\vartheta) = \mathbb{E}[e^{i\vartheta Y}] = p(1-p)^2 e^{-i\vartheta} + (1-p)^3 + 2p^2(1-p) + (p^3 + 2p(1-p)^2)e^{i\vartheta} + p^2(1-p)e^{i2\vartheta}$$

or

$$\varphi_Y(\vartheta) = \mathbb{E}[e^{i\vartheta X_1}] \cdot \mathbb{E}[e^{i\vartheta X_2}] \cdot \mathbb{E}[e^{i\vartheta(-X_3)}]$$

$$= (1-p+pe^{i\vartheta})^2 (1-p+pe^{-i\vartheta})$$

Problem 2. [10] Let $(X_i)_{1\le i\le n}$ be a family of i.i.d. Uniform $(0,3)$ random variables and define $Z_i = \min\{2, X_i\}$.

(i) Compute $\mathrm{P}[Z_1 > z]$ for any $z \in \mathbb{R}$;

(ii) Compute $\mathrm{E}[Z_1]$;

(iii) Prove a Chernoff Bound Upper tail estimate for $\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$.

Hint: use the the Heoffding's inequality.

(i) $$\mathbb{P}[Z_1 \le z] = \begin{cases} 0 & z < 0 \\ z/3 & 0 \le z < 2 \\ 1 & z \ge 2 \end{cases}$$

$$\mathbb{P}[Z_1 > z] = \begin{cases} 1 & z < 0 \\ 1 - z/3 & 0 \le z < 2 \\ 0 & z \ge 2 \end{cases}$$

(ii) $$\mathbb{E}[Z_1] = \int_0^{+\infty} \mathbb{P}[Z_1 > z]\,dz = \int_0^2 (1 - z/3)\,dz$$

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

(iii) $a = 0 \le Z_i \le 2 = b$ $\forall i$. The Hoeffding's inequality says

$$\mathbb{P}\left[n\overline{Z}_n - n\cdot\frac{4}{3} \ge t\right] \le e^{-2t^2/n4} = e^{-t^2/2n}$$

$$\Rightarrow \mathbb{P}\left[\overline{Z}_n \ge 4/3 + \varepsilon\right] \le e^{-\frac{n^2\varepsilon^2}{2n}} = e^{-n\varepsilon^2/2}$$

$\varepsilon = t/n \Rightarrow t = n\cdot\varepsilon$

Problem 3. [12] Let $(X_n)_{n\geq 0}$ be a Markov chain on $\mathbb{N}\setminus\{0\} = \{1,2,3,\dots\}$ with transition probabilities given by

$$p_{i,1} = \frac{1+2i}{(i+1)^2}\ , \quad p_{i,i+1} = \frac{i^2}{(i+1)^2}\ , \quad i \geq 1$$

(i) Is the Markov chain irreducible?

(ii) Is the Markov chain aperiodic?

(iii) Compute $E[X_2|X_0=k]$ for any $k\in\mathbb{N}\setminus\{0\}$;

(iv) Determine the invariant distribution.

(i) YES $\quad P_{i1} > 0 \ \forall i \geq 1, \quad P_{1i}^{(i-1)} > 0 \quad \forall i \geq 1$

(ii) YES $\quad P_{11}^{(2)} > 0, \quad P_{11}^{(3)} > 0$

(iii) $X_0 = 1 \qquad \mathbb{E}[X_2|X_0=1] = P_{11} + P_{12}P_{21} + 2P_{11}P_{12} + 3P_{12}P_{23}$

$X_0 = 2 \qquad \mathbb{E}[X_2|X_0=2] = P_{21}P_{11} + P_{23}P_{31} + 2P_{21}P_{12} + 4P_{23}P_{34}$

In general $X_0 = k,\ k>2 \quad \mathbb{E}[X_2|X_0=k] = P_{k1}P_{11} + P_{k\,k+1}P_{k+1\,1} + 2P_{k1}P_{12} + (k+2)P_{k\,k+1}P_{k+2\,k+2}$

(iv) $\pi P = \pi \Rightarrow \pi_2 = \frac{1}{4}\pi_1,\ \pi_3 = \frac{4}{9}\pi_2,\ \pi_3 = \frac{9}{16}\pi_2, \dots, \pi_{k+1} = \frac{k^2}{(k+1)^2}\pi_k$

$\Rightarrow \pi_1 + \frac{1}{2^2}\pi_1 + \frac{1}{3^2}\pi_1 + \dots + \frac{1}{k^2}\pi_1 + \dots =$

$= \pi_1 \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \right) = 1$

Since $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$

$\pi_1 = 6/\pi^2$

$\pi = \frac{6}{\pi^2}\left(1, \frac{1}{2^2}, \frac{1}{3^2}, \dots, \frac{1}{k^2}, \dots\right)$

$\pi \cong 3.14\dots$