DATA SCIENCE **Stochastic Methods**	Name:________________
December 16, 2021 Prof. Marco Ferrante	Student number: SOLUTION

Problem 1. [10] Let $X \sim Exp(\lambda)$ and $Y \sim Geo(p)$ and assume that they are independent.

(i) Compute $\mathrm{E}[XY]$;

(ii) Compute $\mathrm{P}[X > Y]$;

(iii) Is $\mathrm{E}[1/X]$ finite?

Problem 2. [13] Three balls, one red e two black, are distributed between two boxes, labeled A and B. Each period, a box is selected at random, and if it contains the red ball, a ball chosen at random from that box is removed and placed into the other box. Let (X_n, Y_n) denotes the number of red and black balls in urn A after n periods and define the state space $S = \{(0,0),(0,1),(0,2),(1,0),(1,1),(1,2)\}$

(i) Determine the transition probability matrix;

(ii) Classify the states (communication, periodicity, recurrence);

(iii) Does the Markov Chain admits a unique invariant distribution? is this distribution reversible?

(iv) In the long run, what fraction of time is box A empty?

(v) Compute the invariant distribution.

Problem 3. [13] Let $(X_i)_{1\leq i\leq n}$ be a family of i.i.d. random variables with $\mathbb{P}[X_1 = 1] = 1 - \mathbb{P}[X_1 = -1] = p$, with $0 < p < 1$.

(i) Compute $\mu = \mathrm{E}[X_1]$ and $\sigma^2 = Var[X_1]$;

(ii) Prove that $Y_i = (X_i + 1)/2$ is distributed as a Bernoulli random variables $Be(p)$;

(iii) Defined $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, determine an exponential decay for the “upper tail” of $\overline{X}_n - \mu$.

Problem 1 : $X \sim \mathcal{E}xp(\lambda)$, $Y \sim Geo(p)$, $X \perp\!\!\!\perp Y$

1) Since $X \perp\!\!\!\perp Y$, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] = \frac{1}{\lambda} \cdot \frac{1}{p}$

2) $$\mathbb{P}[X > Y] = \sum_{k=1}^{\infty} \mathbb{P}[X > Y \mid Y = k] \cdot \mathbb{P}[Y = k] =$$

$$= \sum_{k=1}^{\infty} \mathbb{P}[X > k \mid Y = k] \cdot \mathbb{P}[Y = k] =$$

by independence $$= \sum_{k=1}^{\infty} \mathbb{P}[X > k] \cdot \mathbb{P}[Y = k] = \sum_{k=1} e^{-\lambda k} \cdot (1-p)^{k-1} p$$

$$= p e^{-\lambda} \sum_{k=1}^{\infty} \left(e^{-\lambda}(1-p)\right)^{k-1}$$

$$= \frac{p e^{-\lambda}}{1 - e^{-\lambda}(1-p)}$$

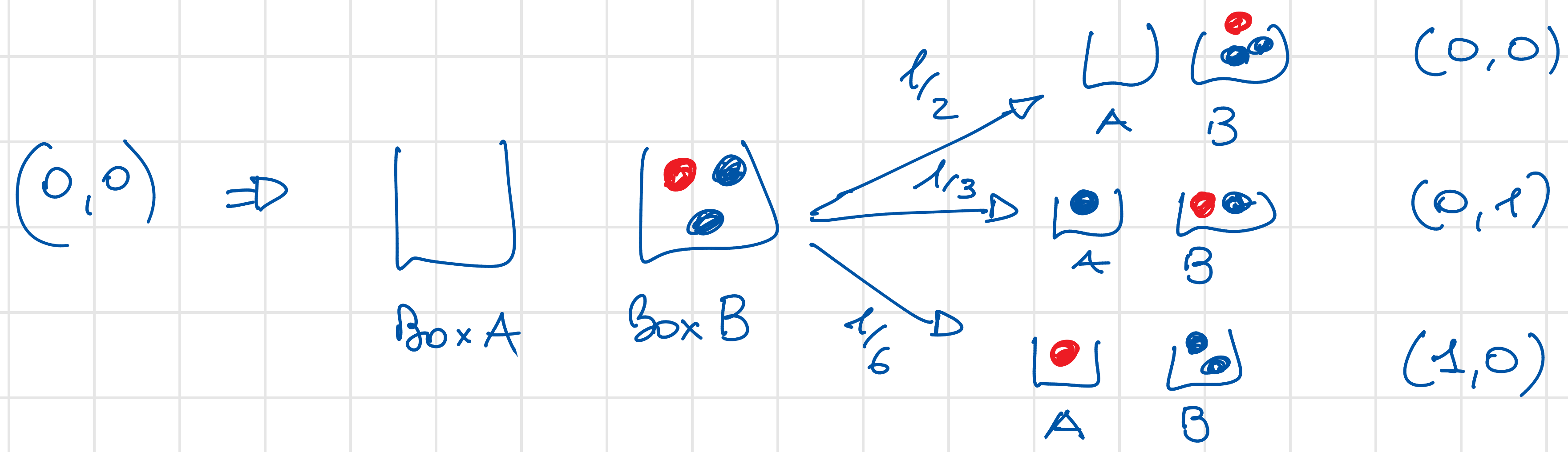
3) $$\mathbb{E}[1/X] = \int_0^{+\infty} 1/x \, \lambda e^{-\lambda x} \, dx \geq \int_0^1 \frac{1}{x} \lambda e^{-\lambda x} \, dx$$

$$\geq \lambda e^{-\lambda} \int_0^1 1/x \, dx = +\infty$$

Problem 2

1)

$$P = \begin{array}{c|cccccc} & (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) \\ \hline (0,0) & 1/2 & 1/3 & 0 & 1/6 & 0 & 0 \\ (0,1) & 0 & 1/2 & 1/4 & 0 & 1/4 & 0 \\ (0,2) & 0 & 0 & 1/2 & 0 & 0 & 1/2 \\ (1,0) & 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ (1,1) & 0 & 1/4 & 0 & 1/4 & 1/2 & 0 \\ (1,2) & 0 & 0 & 1/6 & 0 & 1/3 & 1/2 \end{array}$$



and similarly for the other cases.

2) The chain is irreducible

$$(0,0) \longrightarrow (0,1) \longrightarrow (0,2) \longrightarrow (1,2) \longrightarrow (1,1) \longrightarrow (1,0)$$

All the states are recurrent (the unique class is closed)

Moreover $(0,0) \longrightarrow (0,1) \longrightarrow (0,0)$ $\quad$ $(0,0) \longrightarrow (0,1) \longrightarrow (0,1) \longrightarrow (0,0)$

$\Rightarrow (0,0)$ is aperiodic $\Rightarrow$ the chain is aperiodic

3) The Markov Chain admits a unique invariant distribution (by the Ergodic theorem) π.

This distribution is not reversible, since

$$P_{(0,1),(0,0)} = 0 \quad \text{while} \quad P_{(0,0),(0,1)} = 1/3 .$$

4) The long run fraction of time the box A is empty is equal to $\pi_{(0,0)}$.

5) $\pi = \pi P$

$$\begin{cases} \pi_1 = 1/2\,\pi_1 + 1/2\,\pi_4 \\ \pi_2 = 1/3\,\pi_1 + 1/2\,\pi_2 + 1/4\,\pi_5 \\ \pi_3 = 1/4\,\pi_2 + 1/2\,\pi_3 + 1/6\,\pi_6 \\ \pi_4 = 1/6\,\pi_1 + 1/2\,\pi_4 + 1/4\,\pi_5 \\ \pi_5 = 1/4\,\pi_2 + 1/2\,\pi_5 + 1/3\,\pi_6 \\ \pi_6 = 1/2\,\pi_3 + 1/2\,\pi_6 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1 \end{cases} \Rightarrow \begin{cases} \pi_1 = \pi_4 \\ \pi_5 = 4/3\,\pi_1 \\ \pi_2 = 4/3\,\pi_3 \\ \pi_3 = \pi_1 \\ \pi_3 = \pi_6 \\ \pi_1 + \dots + \pi_6 = 1 \end{cases}$$

$$\pi = \left(3/20,\ 4/20,\ 3/20,\ 3/20,\ 4/20,\ 3/20\right)$$

Problem 3 $\quad X_i = \begin{cases} 1 & p \\ -1 & 1-p \end{cases} \qquad 0<p<1$

(1)

$$\mu = \mathbb{E}[X_1] = p + (-1)(1-p) = 2p-1$$

$$\sigma^2 = \mathrm{Var}[X_1] = p + (1-p) - (2p-1)^2 = 1 - 4p^2 - 1 + 4p$$
$$= 4p(1-p)$$

(2) $[Y_i = 1] = [X_i = 1] = p$

$[Y_i = 0] = [X_i = -1] = 1-p \quad \Rightarrow Y_i \sim \mathrm{Bin}(1,p)$

(3) $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad$ Since $X_i = 2Y_i - 1$

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} (2Y_i - 1) = 2 \frac{1}{n} \sum_{i=1}^{n} Y_i - 1$$
$$= 2\overline{Y}_n - 1$$

$$\mathbb{P}[\overline{X}_n - (2p-1) > \varepsilon] = \mathbb{P}[2\overline{Y}_n - 1 - 2p + 1 > \varepsilon]$$
$$= \mathbb{P}[2\overline{Y}_n - 2p > \varepsilon] = \mathbb{P}[\overline{Y}_n - p > \varepsilon/2]$$
$$= \mathbb{P}[\overline{Y}_n > p + \varepsilon/2] \qquad \text{taking } \varepsilon = 2\delta p$$
$$= \mathbb{P}[\overline{Y}_n > p(1+\delta)] \leq e^{-np\frac{\delta^2}{2+\delta}}$$

by Chernoff's Bound for $\mathrm{Bin}(1,p)$ RV's