

# DATA SCIENCE Stochastic Methods

January 22, 2025

**Problem 1.** [12] Let  $\{X_n, n \geq 1\}$  and  $\{Y_n, n \geq 1\}$  be two independent sequences of independent random variables, such that  $X_n \sim \text{Bin}(1, p)$  and  $Y_n \sim \mathcal{U}(0, 1)$  for any  $n$ .

- (i) For any  $n$  define  $Z_n = X_n + Y_n$ . Compute the distribution and expectation of  $Z_n$ ;
- (ii) Compute the limit in probability of the sequence  $\frac{Z_1 + Z_2 + \dots + Z_n}{n}$ , as  $n \rightarrow +\infty$ ;
- (iii) For any  $n$  define  $V_n = X_n \cdot Y_n$ . Compute the distribution, expectation and variance of  $V_n$ ;
- (iv) Compute the limit in probability of the sequence  $\frac{V_1 + V_2 + \dots + V_n}{n}$ , as  $n \rightarrow +\infty$ .

$$\begin{aligned}
 \text{(i)} \quad F_{Z_n}(z) &= P[X_n + Y_n \leq z] = P[X_n + Y_n \leq z | X_n = 0] \cdot P[X_n = 0] \\
 &\quad + P[X_n + Y_n \leq z] \cdot P[X_n = 1] = P F_{Y_n}(z) + (1-p) F_{Y_n}(z-1) \\
 &= \begin{cases} 0 & z < 0 \\ pz & 0 \leq z < 1 \\ p + (1-p)(z-1) & 1 \leq z < 2 \\ 1 & z \geq 2 \end{cases} \quad ; \quad \begin{aligned} E[Z_n] &= E[X_n] + E[Y_n] \\ &= p + \frac{1}{2} \end{aligned}
 \end{aligned}$$

$$\text{(ii)} \quad \text{By the WLLN, } \frac{Z_1 + \dots + Z_n}{n} \xrightarrow[n \rightarrow \infty]{P} E[Z_1] = p + \frac{1}{2}$$

$$\begin{aligned}
 \text{(iii)} \quad F_{V_n}(v) &= P[X_n \cdot Y_n \leq v] = P[X_n Y_n \leq v | X_n = 0] \cdot P[X_n = 0] + \\
 &\quad + P[X_n Y_n \leq v | X_n = 1] \cdot P[X_n = 1] = \\
 &= \begin{cases} 0 & v < 0 \\ (1-p) + p v & 0 \leq v < 1 \\ 1 & v \geq 1 \end{cases} \quad ; \quad \begin{aligned} E[V_n] &= p \cdot \frac{1}{2} \\ \text{Var}[V_n] &= E[X_n^2 \cdot Y_n^2] - (E[X_n Y_n])^2 \\ &= E[X_n^2] \cdot E[Y_n^2] - p^2 \cdot \frac{1}{4} \end{aligned}
 \end{aligned}$$

$$\text{(iv)} \quad \underbrace{V_1 + \dots + V_n}_{n} \xrightarrow[n \rightarrow \infty]{P} E[V_1] = \frac{p}{2}$$



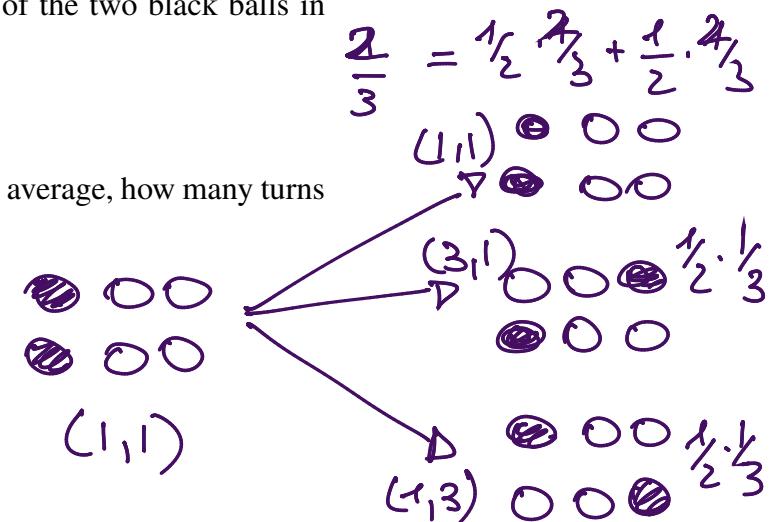
**Problem 2.** [13] We have two rows of 3 balls: in both rows, there are 2 white balls and 1 black ball. We move the balls according to the following scheme: we randomly choose one of the two rows, then we randomly select a ball from that row and move it to the last position in the same row. We repeat this process in the same manner.

- Define a Markov chain that describes the positions of the two black balls in their respective rows;
- Classify the states of the Markov chain;
- Compute the invariant distribution;
- If initially both black balls are in the first position, on average, how many turns will it take for them to return in that position?

(i)

$$\mathcal{S} = \{(i,j) : i, j \in \{1, 2, 3\}\}$$

$$P[X_i = (i,j) | X_0 = (1,1)] =$$



and similarly for the other cases

$$P = \left[ \begin{array}{ccccccccc} & (1,1) & (1,2) & (1,3) & (2,1) & (2,2) & (2,3) & (3,1) & (3,2) & (3,3) \\ (1,1) & \left[ \begin{array}{ccccccccc} \frac{2}{3} & 0 & \frac{1}{6} & 0 & 0 & 0 & \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} & 0 & 0 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{6} \\ \frac{1}{6} & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & 0 & \frac{1}{6} & 0 \\ 0 & 0 & \frac{1}{6} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{6} \\ 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{array} \right] \end{array} \right]$$

(ii) the MC is irreducible and aperiodic

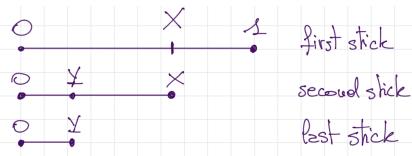
(iii) Since the matrix  $P$  is doubly stochastic, the

invazirized distribution is

$$\left( \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9} \right)$$

(iv)  $m_{(1,1)} = \frac{\ell}{\pi_{(1,1)}} = \frac{1}{1/9} = \boxed{9}$

**Problem 3.** [9] A stick of length one is broken at a random point  $X$ , uniformly distributed over the stick. This remaining stick is broken once more at a random point  $Y$ .



- (i) Determine the expected length of the first broken stick (the second stick in the figure above);
- (ii) Compute the conditional expected length of the last stick, given that  $X = x$ .
- (iii) Compute the expected length of the last stick.

$$\begin{aligned}
 \text{(i)} \quad & \text{The (random) length is } X ; \quad \mathbb{E}[X] = \boxed{\frac{1}{2}} \\
 \text{(ii)} \quad & \text{The (random) length is } Y : \quad \mathbb{E}[Y | X=x] = \boxed{\frac{x}{2}} \\
 \text{(iii)} \quad & \mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y | X]] = \int_0^1 x \cdot \frac{1}{x} dx \\
 & = \left[ \frac{x^2}{2} \right]_0^1 = \boxed{\frac{1}{4}}
 \end{aligned}$$