STAT201A: Introduction to Probability at an Advanced Level (Fall 2024)
UC Berkeley

Problem Set 6

Due: 10:00pm, Friday, December 6, 2024 (via Gradescope)

1. **(Branching process)** A branching process starts with one individual, i.e. $X(0) = 1$, who reproduces according to the following principle:

# of children	0	1	2
probability	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Individuals reproduce independently of each other and independently of the number of their sisters and brothers. Determine

(a) the probability that the population becomes extinct;

The probability generating function of the number of offspring is

$$\phi(s) = \sum_{k=0}^{2} s^k p_k = \frac{1}{6} + \frac{1}{3}s + \frac{1}{2}s^2.$$

The probability of extinction is the smallest solution s to the equation

$$s = \phi(s).$$

Solving this equation, the probability of extinction is $\frac{1}{3}$.

(b) the probability that the population has become extinct in the second generation, i.e. $\mathbb{P}(X(2) = 0)$;

$$\begin{aligned}\mathbb{P}(X(2) = 0) = \mathbb{P}(X(2) = 0 \mid X(1) = 0) \cdot \mathbb{P}(X(1) = 0) \\ + \mathbb{P}(X(2) = 0 \mid X(1) = 1) \cdot \mathbb{P}(X(1) = 1) \\ + \mathbb{P}(X(2) = 0 \mid X(1) = 2) \cdot \mathbb{P}(X(1) = 2) \\ = 1 \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{3} + \left(\frac{1}{6}\right)^2 \cdot \frac{1}{2} \\ = \frac{17}{72}.\end{aligned}$$

(c) the expected number of children given that there are no grandchildren.

$$\begin{aligned}\mathbb{E}[X(1) \mid X(2) = 0] &= \mathbb{P}(X(1) = 1 \mid X(2) = 0) \cdot 1 + \mathbb{P}(X(1) = 2 \mid X(2) = 0) \cdot 2 \\ &= \frac{\mathbb{P}(X(2) = 0 \mid X(1) = 1) \cdot \mathbb{P}(X(1) = 1)}{\mathbb{P}(X(2) = 0)} \cdot 1 \\ &+ \frac{\mathbb{P}(X(2) = 0 \mid X(1) = 2) \cdot \mathbb{P}(X(1) = 2)}{\mathbb{P}(X(2) = 0)} \cdot 2 \\ &= \frac{\left(\frac{1}{6} \cdot \frac{1}{3}\right)}{\frac{17}{72}} \cdot 1 + \frac{\left(\frac{1}{36} \cdot \frac{1}{2}\right)}{\frac{17}{72}} \cdot 2 \\ &= \frac{6}{17}.\end{aligned}$$

2. **(Random walk)** Random walk on $\{0, 1, 2, 3\}$. Consider the Markov chain (X_n) with transition matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

started with $X_0 = 0$. Define T_j as $\min\{n \geq 1 : X_n = j\}$. Find explicitly the following distributions and expectations.

(a) The distribution of X_2.

$$P^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

$P(X_2 = 0) = \frac{1}{2}, P(X_2 = 1) = P(X_2 = 2) = \frac{1}{4}, P(X_2 = 3) = 0.$

(b) The limit distribution of X_n as $n \to \infty$.

By solving $\pi P = \pi$, we have $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Since the MC has finite state space S and it's irreducible and aperiodic, in limiting distribution theorem, we know for $i, j \in 0, 1, 2, 3$, $\lim_{n\to\infty} [P^n]_{ij} = \frac{1}{4}$. Therefore the limit distribution of X_n is also π.

(c) $\mathbb{E}[T_0]$

Define $h_k = \mathbb{E}[\min\{n \geq 0 : X_n = 0\} \mid X_0 = k]$. By one-step analysis, we derive the following system of equations

$$\begin{cases} h_0 = 0, \\ h_1 = \frac{1}{2}h_0 + \frac{1}{2}h_2 + 1, \\ h_2 = \frac{1}{2}h_1 + \frac{1}{2}h_3 + 1, \\ h_3 = \frac{1}{2}h_2 + \frac{1}{2}h_3 + 1. \end{cases}$$

Solving this system, we find

$$\begin{cases} h_0 = 0, \\ h_1 = 6, \\ h_2 = 10, \\ h_3 = 12. \end{cases}$$

Finally, the expected value of T_0 is given by:

$$\mathbb{E}[T_0] = 1 + \sum_{k=0}^{3} h_k P(X_1 = k) = \frac{1}{2} \times 0 + \frac{1}{2} \times 6 + 1 = 4.$$

(d) $\mathbb{E}[T_3]$

Define $g_k = \mathbb{E}\left[\min\{n \geq 0 : X_n = 3\} \mid X_0 = k\right]$. Similar to (c), by one-step analysis, we derive the following system of equations

$$\begin{cases} g_0 = \frac{1}{2}g_0 + \frac{1}{2}g_1 + 1, \\ g_1 = \frac{1}{2}g_0 + \frac{1}{2}g_2 + 1, \\ g_2 = \frac{1}{2}g_1 + \frac{1}{2}g_3 + 1, \\ g_3 = 0. \end{cases}$$

Solving this system, we find

$$\begin{cases} g_0 = 12, \\ g_1 = 10, \\ g_2 = 6, \\ g_3 = 0. \end{cases}$$

Finally, since it definitely takes more than one step from 0 to 3,

$$\mathbb{E}[T_3] = \mathbb{E}[\min\{n \geq 1 : X_n = 3\}|X_0 = 0] = \mathbb{E}[\min\{n \geq 0 : X_n = 3\}|X_0 = 0] = 12.$$

(e) $\mathbb{P}[T_3 < T_0]$

Define $f_k = \mathbb{P}\left(T_3 < T_1 \mid X_1 = k\right)$. We can derive the following system of equations by conditional probabilities

$$\begin{cases} f_0 = 0, \\ f_1 = \frac{1}{2}f_0 + \frac{1}{2}f_2, \\ f_2 = \frac{1}{2}f_1 + \frac{1}{2}f_3, \\ f_3 = 1. \end{cases}$$

Solving this system, we find

$$\begin{cases} f_0 = 0, \\ f_1 = \frac{1}{3}, \\ f_2 = \frac{2}{3}, \\ f_3 = 1. \end{cases}$$

Finally,

$$\mathbb{P}(T_3 < T_1) = \mathbb{P}\left(T_3 < T_1 \mid X_1 = 1\right) P(X_1 = 1) + \mathbb{P}\left(T_3 < T_1 \mid X_1 = 0\right) P(X_1 = 0) = 1/6.$$

3. **(The average number of jobs)** Jennifer is employed for one day at a time. When she is out of work, she visits the job agency in the morning to see if there is work for that day. There is a job for her with probability 1/2. If there is no work, she comes back the next day. When she has a job, she will be called back to the same job for the next day with probability 2/3. When she is not called back, she goes to the job agency again the next morning to look for a new job that she had not had previously. Approximate the average number of jobs Jennifer works in a year.
There are multiple solutions to this problem. A short approximation consist in calculating the average number of days on the same job and the average number of days without a job. Let X be the random variable representing the number of days on a fixed job, X has geometric distribution

with parameter $1/3$. Let Y be a random variable representing the number of days without job, Y also has geometric distribution but starts at 0, with parameter $1/2$. We get $\mathbb{E}[X] = 3$ and $\mathbb{E}[Y] = 1$. Hence $\mathbb{E}[X + Y] = 4$. We obtain $365/4 \approx 91.25$.

4. **(Rain or no rain)** Suppose that at day 0 it is not raining. Then each new day, if it rained yesterday, it will rain with probability 0.7; if it did not rain yesterday, it will rain with probability 0.2.

   (a) Find the stationary distribution.
   This is an irreducible finite state Markov chain. Hence an stationary distribution exists and is unique. It is enough to find π_r and π_n, the stationary distribution for rain and not rain respectively, by solving the system of equations $-0.3\pi_r + 0.2\pi_n = 0$, $0.3\pi_r - 0.2\pi_n = 0$, $\pi_r + \pi_n = 1$. We get $(\pi_r, \pi_n) = (2/5, 3/5)$.

   (b) How many days should we expect to wait to have rain for the first time?
   Suppose we start from a non-rainy day. Then by considering non-rainy days as failures and rainy days as successes. Then for X the random variable giving the first day with rain, X is geometric with parameter $0.2 = 1/5$. Hence $\mathbb{E}[x] = 5$.

5. **(The game of roulette)** A gambler plays the game of roulette, betting X dollars on red or black. The gambler wins X dollars with probability $p = 18/38$ or loses the bet with probability $q = 20/38$. Suppose that the gambler starts the game with \$500 in his pocket and upper limit on winnings is \$1000.

   (a) Compute the probability of the gambler's ruin for $X = \$10$.
   This problem can be solved by considering a Markov chain on $\{0, 1, \ldots, N\}$, where N is a positive integer; 0 and N are absorbing boundaries; and for $j = 1, \ldots, N - 1$,

$$\mathbb{P}[X_{t+1} = j - 1 \mid X_t = j] = q = 1 - p,$$
$$\mathbb{P}[X_{t+1} = j + 1 \mid X_t = j] = p.$$

   Let R denote the event that you hit the boundary 0 before hitting the boundary N. Define $u_j := \mathbb{P}[R \mid X_0 = j]$. Then, $u_0 = 1$ and $u_N = 0$, while for $j = 1, \ldots, N - 1$,

$$\begin{aligned} u_j &= \mathbb{P}[R \mid X_0 = j] \\ &= \mathbb{P}[R \mid X_0 = j, X_1 = j - 1]\mathbb{P}[X_1 = j - 1 \mid X_0 = j] \\ &\qquad + \mathbb{P}[R \mid X_0 = j, X_1 = j + 1]\mathbb{P}[X_1 = j + 1 \mid X_0 = j] \\ &= (1 - p)\, u_{j-1} + p\, u_{j+1}. \end{aligned}$$

   Now, rewrite the left hand side as $p\mu_j + (1 - p)\mu_j$ and rearrange terms to get

$$p[\mu_{j+1} - \mu_j] = (1 - p)[\mu_j - \mu_{j-1}].$$

Define $r := \frac{1-p}{p}$ and $\Delta_j := \mu_j - \mu_{j-1}$. Then, we obtain

$$\begin{aligned}
\Delta_2 &= r\Delta_1 \\
\Delta_3 &= r\Delta_2 = r^2\Delta_1 \\
\Delta_4 &= r\Delta_3 = r^2\Delta_2 = r^3\Delta_1 \\
&\vdots \\
\Delta_N &= r^{N-1}\Delta_1.
\end{aligned}$$

Further, for all $j = 1, \ldots, N$, we have $\Delta_1 + \Delta_2 + \cdots + \Delta_j = \mu_j - \mu_0 = \mu_j - 1$, where we have used the boundary condition in the last equality. Hence,

$$\mu_j = 1 + \Delta_1 + \Delta_2 + \cdots + \Delta_j = 1 + \Delta_1[1 + r + \cdots + r^{j-1}].$$

Since $\mu_N = 0$, we obtain $\Delta_1 = -1/[1 + r + \cdots + r^{N-1}]$, so

$$\mu_j = 1 - \frac{1 + r + \cdots + r^{j-1}}{1 + r + \cdots + r^{N-1}} = \begin{cases} 1 - \dfrac{j}{N}, & \text{if } r = 1, \\ \dfrac{r^j - r^N}{1 - r^N}, & \text{if } r \neq 1. \end{cases} \tag{1}$$

Using $N = 100$ and $X_0 = 50$ in (1) gives $\mathbb{P}[R \mid X_0 = 50] \approx 0.995$.

(b) Compute the probability of the gambler's ruin for $X = \$100$.

Similarly, this is equivalent to having $N = 10$ and $X_0 = 5$ in the ruin problem, and we immediately obtain $\mathbb{P}[R \mid X_0 = 5] \approx 0.629$.

(c) Compare the above results with the probability of ruin in the case the gambler bets everything on a single turn of the wheel.

If the gambler bets everything on a single turn of the wheel, the probability of ruin is $q = 1 - p = 20/38 \approx 0.526$. This probability is lower than either one of the cases above.