Chapter 14

Solved Problems

14.1 Probability review

Problem 14.1. Let X and Y be two $\mathbb{N}_0$-valued random variables such that $X = Y + Z$, where Z is a Bernoulli random variable with parameter $p \in (0,1)$, independent of Y. Only one of the following statements is true. Which one?

(a) $X + Z$ and $Y + Z$ are independent

(b) X has to be $2\mathbb{N}_0 = \{0, 2, 4, 6, \dots\}$-valued

(c) The support of Y is a subset of the support of X

(d) $\mathbb{E}[(X + Y)Z] = \mathbb{E}[(X + Y)]\mathbb{E}[Z]$.

(e) none of the above

Solution: The correct answer is (c).

(a) False. Simply take $Y = 0$, so that $Y + Z = Z$ and $X + Z = 2Z$.

(b) False. Take $Y = 0$.

(c) True. For m in the support of Y (so that $\mathbb{P}[Y = m] > 0$), we have

$$\mathbb{P}[X = m] \geq \mathbb{P}[Y = m, Z = 0] = \mathbb{P}[Y = m]\mathbb{P}[Z = 0] = \mathbb{P}[Y = m](1 - p) > 0.$$

Therefore, m is in the support of X.

(d) False. Take $Y = 0$.

(e) False.

Problem 14.2. A fair die is tossed and its outcome is denoted by X, i.e.,

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

After that, X independent fair coins are tossed and the number of *heads* obtained is denoted by Y.
Compute:

1. $\mathbb{P}[Y = 4]$.

2. $\mathbb{P}[X = 5 | Y = 4]$.

3. $\mathbb{E}[Y]$.

4. $\mathbb{E}[XY]$.

Solution:

1. For $k = 1, \ldots, 6$, conditionally on $X = k$, Y has the binomial distribution with parameters k and $\frac{1}{2}$. Therefore,

$$\mathbb{P}[Y = i | X = k] = \begin{cases} \binom{k}{i} 2^{-k}, & 0 \le i \le k \\ 0, & i > k, \end{cases}$$

   and so, by the law of total probability.

$$\mathbb{P}[Y = 4] = \sum_{k=1}^{6} \mathbb{P}[Y = 4 | X = k] \mathbb{P}[X = k]$$

$$= \tfrac{1}{6} \left(2^{-4} + \binom{5}{4} 2^{-5} + \binom{6}{4} 2^{-6} \right) \left[= \tfrac{29}{384} \right].$$

(14.1)

2. By the (idea behind the) Bayes formula

$$\mathbb{P}[X = 5 | Y = 4] = \frac{\mathbb{P}[X = 5, Y = 4]}{\mathbb{P}[Y = 4]} = \frac{\mathbb{P}[Y = 4 | X = 5] \mathbb{P}[X = 5]}{\mathbb{P}[Y = 4]}$$

$$= \frac{\binom{5}{4} 2^{-5} \times \frac{1}{6}}{\frac{1}{6} \left(2^{-4} + \binom{5}{4} 2^{-5} + \binom{6}{4} 2^{-6} \right)} \left[= \tfrac{10}{29} \right].$$

3. Since $\mathbb{E}[Y | X = k] = \frac{k}{2}$ (the expectation of a binomial with $n = k$ and $p = \frac{1}{2}$), the law of total probability implies that

$$\mathbb{E}[Y] = \sum_{k=1}^{6} \mathbb{E}[Y | X = k] \mathbb{P}[X = k] = \tfrac{1}{6} \sum_{k=1}^{6} \tfrac{k}{2} \left[= \tfrac{7}{4} \right].$$

4. By the same reasoning,

$$\mathbb{E}[XY] = \sum_{k=1}^{6} \mathbb{E}[XY | X = k] \mathbb{P}[X = k] = \sum_{k=1}^{6} \mathbb{E}[kY | X = k] \mathbb{P}[X = k]$$

$$= \sum_{k=1}^{6} k \mathbb{E}[Y | X = k] \mathbb{P}[X = k] = \tfrac{1}{6} \sum_{k=1}^{6} \tfrac{1}{2} k^2 \left[= \tfrac{91}{12} \right].$$

Problem 14.3.

1. An urn contains 1 red ball and 10 blue balls. Other than their color, the balls are indistinguishable, so if one is to draw a ball from the urn without peeking - all the balls will be equally likely to be selected. If we draw 5 balls from the urn at once and without peeking, what is the probability that this collection of 5 balls contains the red ball?

2. We roll two fair dice. What is the probability that the sum of the outcomes equals exactly 7?

3. Assume that A and B are *disjoint* events, i.e., assume that $A \cap B = \emptyset$. Moreover, let $\mathbb{P}[A] = a > 0$ and $\mathbb{P}[B] = b > 0$. Calculate $\mathbb{P}[A \cup B]$ and $\mathbb{P}[A \cap B]$, using the values a and b:

Solution:

1.

$$\mathbb{P}[\text{"the red ball is selected"}] = \frac{\binom{10}{4}}{\binom{11}{5}} = \frac{5}{11}.$$

2. There are 36 possible outcomes (pairs of numbers) of the above roll. Out of those, the following have the sum equal to $7 : (1,6), (2,5), (3,4), (4,3), (5,2), (6,1)$. Since the dice are fair, all outcomes are equally likely. So, the probability is

$$\frac{6}{36} = \frac{1}{6}.$$

3. According to the axioms of probability:

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] = a + b, \ \ \mathbb{P}[A \cap B] = \mathbb{P}[\emptyset] = 0.$$

Problem 14.4.

1. Consider an experiment which consists of 2 independent coin-tosses. Let the random variable X denote the number of heads appearing. Write down the probability mass function of X.

2. There are 10 balls in an urn numbered 1 through 10. You randomly select 3 of those balls. Let the random variable Y denote the maximum of the three numbers on the extracted balls. Find the probability mass function of Y. You should simplify your answer to a fraction that does not involve binomial coefficients. Then calculate: $\mathbb{P}[Y \geq 7]$.

3. A fair die is tossed 7 times. We say that a toss is a **success** if a 5 or 6 appears; otherwise it's a **failure**. What is the distribution of the random variable X representing the number of successes out of the 7 tosses? What is the probability that there are exactly 3 successes? What is the probability that there are no successes?

4. The number of misprints per page of text is commonly modeled by a Poisson distribution. It is given that the parameter of this distribution is $\lambda = 0.6$ for a particular book. Find the probability that there are exactly 2 misprints on a given page in the book. How about the probability that there are 2 *or more* misprints?

Solution:

1.

$$p_0 = \mathbb{P}[\{(T,T)\}] = \frac{1}{4},$$

$$p_1 = \mathbb{P}[\{(T,H),(H,T)\}] = \frac{1}{2},$$

$$p_2 = \mathbb{P}[\{(H,H)\}] = \frac{1}{4},$$

$$p_k = 0, \text{ for all other } k.$$

2. The random variable Y can take the values in the set $\{3,4,\ldots 10\}$. For any i, the triplet resulting in Y attaining the value i must consist of the ball numbered i and a pair of balls with lower numbers. So,

$$p_i = \mathbb{P}[Y = i] = \frac{\binom{i-1}{2}}{\binom{10}{3}} = \frac{\frac{(i-1)(i-2)}{2}}{\frac{10\cdot9\cdot8}{3\cdot2\cdot1}} = \frac{(i-1)(i-2)}{240}.$$

Since the balls are numbered 1 through 10, we have

$$\mathbb{P}[Y \geq 7] = \mathbb{P}[Y = 7] + \mathbb{P}[Y = 8] + \mathbb{P}[Y = 9] + \mathbb{P}[Y = 10].$$

So,

$$\begin{aligned}
\mathbb{P}[Y \geq 7] &= \frac{6\cdot5}{240} + \frac{7\cdot6}{240} + \frac{8\cdot7}{240} + \frac{9\cdot8}{240} \\
&= \frac{1}{240}(30 + 42 + 56 + 72) \\
&= \frac{200}{240} = \frac{5}{6}.
\end{aligned}$$

3. X has a binominal distribution with parameters $n = 7$ and $p = 1/3$, i.e., $X \sim b(7, 1/3)$.

$$\mathbb{P}[X = 3] = \binom{7}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^4 = \frac{560}{2187},$$

$$\mathbb{P}[X = 0] = \left(\frac{2}{3}\right)^7 = \frac{128}{2187}.$$

4. Let X denote the random variable which stands for the number of misprints on a given page. Then

$$\mathbb{P}[X = 2] = \frac{0.6^2}{2!}e^{-0.6} \approx 0.0988,$$

$$\begin{aligned}
\mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X < 2] \\
&= 1 - (\mathbb{P}[X = 0] + \mathbb{P}[X = 1]) \\
&= 1 - \left(\frac{0.6^0}{0!}e^{-0.6} + \frac{0.6^1}{1!}e^{-0.6}\right) \\
&= 1 - \left(e^{-0.6} + 0.6e^{-0.6}\right) \\
&= 1 - 1.6e^{-0.6} \approx 0.122.
\end{aligned}$$

Problem 14.5. Let X and Y be two Bernoulli random variables with the same parameter $p = \frac{1}{2}$. Can the support of their sum be equal to $\{0, 1\}$? How about the case where p is not necesarily equal to $\frac{1}{2}$? Note that no particular dependence structure between X and Y is assumed.

Solution: Let p_{ij}, $i = 0, 1$, $j = 0, 1$ be defined by

$$p_{ij} = \mathbb{P}[X = i, Y = j].$$

These four numbers effectively specify the full dependence structure of X and Y (in other words, they completely determine the distribution of the random vector (X, Y)). Since we are requiring that both X and Y be Bernoulli with parameter p, we must have

$$p = \mathbb{P}[X = 1] = \mathbb{P}[X = 1, Y = 0] + \mathbb{P}[X = 1, Y = 1] = p_{10} + p_{11}. \tag{14.2}$$

Similarly, we must have

$$
\begin{align}
1 - p &= p_{00} + p_{01}, \tag{14.3}\\
p &= p_{01} + p_{11}, \tag{14.4}\\
1 - p &= p_{00} + p_{10} \tag{14.5}
\end{align}
$$

Suppose now that the support of $X + Y$ equals to $\{0, 1\}$. Then $p_{00} > 0$ and $p_{01} + p_{10} > 0$, but $p_{11} = 0$ (why?). Then, the relation (14.2) implies that $p_{10} = p$. Similarly, $p_{01} = p$ by relation (14.4). Relations (14.3) and (14.5) tell us that $p_{00} = 1 - 2p$. When $p = \frac{1}{2}$, this implies that $p_{00} = 0$ - a contradiction with the fact that $0 \in _{X+Y}$.

When $p < \frac{1}{2}$, there is still hope. We construct X and Y as follows: let X be a Bernoulli random variable with parameter p. Then, we define Y depending on the value of X. If $X = 1$, we set $Y = 0$. If $X = 0$, we set $Y = 0$ with probabilty $\frac{1-2p}{1-p}$ and 1 with probability $\frac{p}{1-p}$. How do we know that Y is Bernoulli with probability p? We use the law of total probability:

$$\mathbb{P}[Y = 0] = \mathbb{P}[Y = 0 | X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = 0 | X = 1]\mathbb{P}[X = 1] = \tfrac{1-2p}{1-p}(1-p) + p = 1 - p.$$

Similarly,

$$\mathbb{P}[Y = 1] = \mathbb{P}[Y = 1 | X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = 1 | X = 1]\mathbb{P}[X = 1] = (1 - \tfrac{1-2p}{1-p})(1-p) = p.$$

14.2 Random Walks

Problem 14.6. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a symmetric simple random walk. For $n \in \mathbb{N}$ the **average of the random walk** on the interval $[0, n]$ is defined by

$$A_n = \tfrac{1}{n} \sum_{k=1}^{n} X_k.$$

1. Is $\{A_n\}_{n \in \mathbb{N}_0}$ a simple random walk (not necessarily symmetric)? Explain carefully using the definition.

2. Compute the covariance $\text{Cov}(X_k, X_l) = \mathbb{E}[(X_k - \mathbb{E}[X_k])(X_l - \mathbb{E}[X_l])]$, for $k \leq l \in \mathbb{N}$

3. Compute the variance of A_n, for $n \in \mathbb{N}$.

(*Note:* You may need to use the following identities:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad \text{and} \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \text{for } n \in \mathbb{N}.)$$

Solution:

1. No it is not. The distribution of A_2 is

$$\begin{pmatrix} -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

In particular, its support is $\{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\}$. For $n = 1$, $A_1 = X_1$, which has the support $\{-1, 1\}$. Therefore, the support of the difference $A_2 - A_1$ cannot be $\{-1, 1\}$ (indeed, the difference must be "able" to take fractional values). This is in contradiction with the definition of the random walk which states that all increments must have distributions supported by $\{-1, 1\}$.

2. We write $X_k = \sum_{i=1}^{k} \xi_i$, and $X_l = \sum_{j=1}^{l} \xi_j$, where $\{\xi_n\}_{n \in \mathbb{N}}$ are the (iid) increments of $\{X_n\}_{n \in \mathbb{N}_0}$. Since $\mathbb{E}[X_k] = \mathbb{E}[X_l] = 0$, we have

$$\mathrm{Cov}(X_k, X_l) = \mathbb{E}[X_k X_l] = \mathbb{E}\left[\left(\sum_{i=1}^{k} \xi_i\right)\left(\sum_{j=1}^{l} \xi_j\right)\right]$$

When the sum above is expanded, the terms of the form $\mathbb{E}[\xi_i \xi_j]$, for $i \neq j$ will disappear because ξ_i and ξ_j are independent for $i \neq j$. The only terms left are those of the form $\mathbb{E}[\xi_i \xi_i]$. Their value is 1, and there are k of them (remember $k \leq l$). Therefore,

$$\mathrm{Cov}(X_k, X_l) = k.$$

3. Let $B_n = \sum_{k=1}^{n} X_k = n A_n$. We know that $\mathbb{E}[A_n] = 0$, and that $\mathrm{Var}[A_n] = \frac{1}{n^2} \mathrm{Var}[B_n]$, so it suffices to compute

$$\mathrm{Var}[B_n] = \mathbb{E}[B_n^2] = \mathbb{E}\left[\left(\sum_{k=1}^{n} X_k\right)\left(\sum_{l=1}^{n} X_l\right)\right]$$

We expand the sum on the right and group together equal ($k = l$) and different indices ($k \neq l$) to get

$$\mathrm{Var}[B_n] = \sum_{k=1}^{n} \mathrm{Var}[X_k] + 2 \sum_{1 \leq k < l \leq n} \mathrm{Cov}(X_k, X_l) = \sum_{k=1}^{n} k + 2 \sum_{1 \leq k < l \leq n} k.$$

For each $k = 1, 2, \ldots, n$, there are exactly $n - k$ values of l such that $k < l \leq n$. Therefore, $\sum_{1 \leq k < l \leq n} k = \sum_{k=1}^{n} k(n - k)$, and so

$$\mathrm{Var}[B_n] = \sum_{k=1}^{n} k + 2 \sum_{k=1}^{n} k(n - k) = (2n + 1) \sum_{k=1}^{n} k - 2 \sum_{k=1}^{n} k^2$$

$$= (2n + 1) \frac{n(n + 1)}{2} - 2 \frac{n(n + 1)(2n + 1)}{6} = \frac{n(n + 1)(2n + 1)}{6}.$$

Therefore,

$$\mathrm{Var}[A_n] = \frac{(n + 1)(2n + 1)}{6n}.$$

Another way to approach this computation uses the representation $X_k = \sum_{i=1}^{k} \xi_k$, so

$$\mathrm{Var}[B_n] = \mathbb{E}\left[\left(\sum_{k=1}^{n} X_k \right)^2 \right] = \mathbb{E}\left[\left(\sum_{k=1}^{n} \sum_{i=1}^{k} \xi_i \right)^2 \right]$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{n} \sum_{k=i}^{n} \xi_i \right)^2 \right] = \mathbb{E}\left[\left(\sum_{i=1}^{n} (n - i + 1) \xi_i \right)^2 \right]$$

$$= \sum_{i=1}^{n} (n - i + 1)^2 = \sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6},$$

where we use the fact that $\mathbb{E}[\xi_i \xi_j] = 0$ for $i \neq j$.

Problem 14.7. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a simple random walk. The **area A_n under the random walk** on the interval $[0, n]$ is defined by

$$A_n = \sum_{k=1}^{n-1} X_k + \tfrac{1}{2} X_n.$$

(*Note:* Draw a picture of a path of a random walk and convince yourself that the expression above is really the area under the graph.)

Compute the expectation and variance of A_n.

Problem 14.8. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a simple symmetric random walk ($X_0 = 0$, $p = q = \frac{1}{2}$). What is the probability that X_n will visit all the points in the set $\{1, 2, 3, 4\}$ at least once during the time-interval $n \in \{0, \ldots, 10\}$? (*Note:* write your answer as a sum of binomial coefficients; you don't have to evaluate it any further.)

Solution: The random walk will visit all the points in $\{1, 2, 3, 4\}$ if and only if its maximum M_{10} during $\{0, \ldots, 10\}$ is at least equal to 4. Since we start at $X_0 = 0$, visiting all of $\{1, 2, 3, 4\}$ at least once is the same as visiting 4 at least once. Therefore, the required probability is $\mathbb{P}[M_{10} \geq 4]$. We know (from the lecture notes) that

$$\mathbb{P}[M_n = l] = \mathbb{P}[X_n = l] + \mathbb{P}[X_n = l + 1].$$

Therefore,

$$\mathbb{P}[M_{10} \geq 4] = \sum_{l=4}^{10} \left(\mathbb{P}[X_{10} = l] + \mathbb{P}[X_{10} = l + 1] \right)$$

$$= \mathbb{P}[X_{10} = 4] + 2\mathbb{P}[X_{10} = 6] + 2\mathbb{P}[X_{10} = 8] + 2\mathbb{P}[X_{10} = 10]$$

$$= 2^{-10} \left(\binom{10}{7} + 2\binom{10}{8} + 2\binom{10}{9} + 2 \right).$$

□

Problem 14.9. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a simple symmetric random walk ($X_0 = 0$, $p = q = \frac{1}{2}$). What is the probability that X_n will not visit the point 5 during the time-interval $n \in \{0, \ldots, 10\}$?

Solution: The random walk will not visit the point 5 if only if its maximum M_{10} during $\{0, \ldots, 10\}$ is at most equal to 4. Therefore, the required probability is $\mathbb{P}[M_{10} \leq 4]$. We know (from the lecture notes) that

$$\mathbb{P}[M_n = l] = \mathbb{P}[X_n = l] + \mathbb{P}[X_n = l + 1].$$

Therefore,

$$\mathbb{P}[M_{10} \leq 4] = \sum_{l=0}^{4} \left(\mathbb{P}[X_{10} = l] + \mathbb{P}[X_{10} = l + 1] \right)$$

$$= \mathbb{P}[X_{10} = 0] + 2\mathbb{P}[X_{10} = 2] + 2\mathbb{P}[X_{10} = 4]$$

$$= 2^{-10} \left(\binom{10}{5} + 2\binom{10}{6} + 2\binom{10}{7} \right).$$

Problem 14.10. A fair coin is tossed repeatedly and the record of the outcomes is kept. Tossing stops the moment the total number of heads obtained so far exceeds the total number of tails by 3. For example, a possible sequence of tosses could look like *HHTTTHTHHTHH*. What is the probability that the length of such a sequence is at most 10?

Solution: Let X_n, $n \in \mathbb{N}_0$ be the number of heads *minus* the number of tails obtained so far. Then, $\{X_n\}_{n \in \mathbb{N}_0}$ is a simple symmetric random walk, and we stop tossing the coin when X hits 3 for the first time. This will happen during the first 10 tosses, if and only if $M_{10} \geq 3$, where M_n denotes the (running) maximum of X. According to the reflection principle,

$$\mathbb{P}[M_{10} \geq 3] = \mathbb{P}[X_{10} \geq 3] + \mathbb{P}[X_{10} \geq 4] = 2(\mathbb{P}[X_{10} = 4] + \mathbb{P}[X_{10} = 6] + \mathbb{P}[X_{10} = 8] + \mathbb{P}[X_{10} = 10])$$

$$= 2^{-9} \left[\binom{10}{3} + \binom{10}{2} + \binom{10}{1} + \binom{10}{0} \right] \quad \left[= \tfrac{11}{32} \right].$$

14.3 Generating functions

Problem 14.11. If $P(s)$ is the generating function of the random variable X, then the generating function of $2X + 1$ is

(a) $P(2s + 1)$

(b) $2P(s) + 1$

(c) $P(s^2 + 1)$

(d) $sP(s^2)$

(e) none of the above

Solution: The correct answer is (d): using the representation $P_Y(s) = \mathbb{E}[s^Y]$, we get $P_{2X+1} = \mathbb{E}[s^{2X+1}] = \mathbb{E}[s(s^{2X})] = s\mathbb{E}[(s^2)^X] = sP_X(s^2)$. Alternatively, we can use the fact that

$$\mathbb{P}[2X + 1 = k] = \begin{cases} 0, & k \in \{0, 2, 4, \dots\}, \\ p_{\frac{k-1}{2}}, & k \in \{1, 3, 5, \dots\}, \end{cases}$$

where $p_n = \mathbb{P}[X = n]$, $n \in \mathbb{N}_0$, so that

$$P_{2X+1}(s) = 0 + p_0 s + 0s^2 + p_1 s^2 + 0s^3 + p_2 s^4 + \cdots = sP_X(s^2).$$

Problem 14.12. Let $P(s)$ be the generating function of the sequence $(p_0, p_1, \dots)$ and $Q(s)$ the generating function of the sequence $(q_0, q_1, \dots)$. If the sequence $\{r_n\}_{n \in \mathbb{N}_0}$ is defined by

$$r_n = \begin{cases} 0, & n \leq 1 \\ \sum_{k=1}^{n-1} p_k q_{n-1-k}, & n > 1, \end{cases}$$

then its generating function is given by (*Note:* Careful! $\{r_n\}_{n \in \mathbb{N}_0}$ is not exactly the convolution of $\{p_n\}_{n \in \mathbb{N}_0}$ and $\{q_n\}_{n \in \mathbb{N}_0}$.)

(a) $P(s)Q(s) - p_0 q_0$

(b) $(P(s) - p_0)(Q(s) - q_0)$

(c) $\frac{1}{s}(P(s) - p_0)Q(s)$

(d) $\frac{1}{s}P(s)(Q(s) - q_0)$

(e) $\frac{1}{s^2}(P(s) - p_0)(Q(s) - q_0)$

Solution: The correct answer is (c). Let $R(s)$ be the generating function of the sequence $\{r_n\}_{n \in \mathbb{N}_0}$ and let $\hat{r}$ be the convolution $\hat{r}_n = \sum_{k=0}^{n} p_k q_{n-k}$ of the sequences $\{p_n\}_{n \in \mathbb{N}_0}$ and $\{q_n\}_{n \in \mathbb{N}_0}$, so that the generating function of $\hat{r}$ is $P(s)Q(s)$. For $n > 1$,

$$\hat{r}_{n-1} = \sum_{k=0}^{n-1} p_k q_{n-1-k} = r_n + p_0 q_{n-1}.$$

Therefore, for $s \in (0,1)$,

$$P(s)Q(s) = \sum_{n=0}^{\infty} \hat{r}_n s^n = \hat{r}_0 + \sum_{n=1}^{\infty} \hat{r}_n s^n$$

$$= p_0 q_0 + \sum_{n=1}^{\infty} (r_{n+1} + p_0 q_n) s^n$$

$$= p_0 q_0 + \frac{1}{s}(R(s) - r_1 s - r_0) + p_0 \sum_{n=1}^{\infty} q_n s^n$$

$$= p_0 Q(s) + \frac{1}{s} R(s).$$

Therefore, $R(s) = \frac{P(s) - p_0}{s} Q(s)$.

Problem 14.13.

A fair coin and a fair 6-sided die are thrown repeatedly until the the first time 6 appears on the die. Let X be the number of *heads* obtained (we are including the *heads* that may have occurred together with the first 6) in the count. The generating function of X is

(a) $\frac{1}{2} \frac{s}{6-5s} + \frac{1}{2}$

(b) $\frac{s}{6-5s}$

(c) $\frac{1+s}{7-5s}$

(d) $\frac{1+2s}{7-4s}$

(e) None of the above

   Solution: X can be seen as a random sum of $\frac{1}{2}$-Bernoulli random variables with the number of summands given by G, the number of tries it takes to get the first heads. G is clearly geometrically distributed with parameter $\frac{1}{6}$. The generating function of G is $P_G(s) = \frac{s}{6-5s}$, so the generating function of the whole random sum X is

$$P_X(s) = P_G((\tfrac{1}{2} + \tfrac{1}{2}s)) = \frac{1+s}{7-5s}.$$

Problem 14.14. The generating function of the $\mathbb{N}_0$-valued random variable X is given by

$$P_X(s) = \frac{s}{1 + \sqrt{1-s^2}}.$$

1. Compute $p_0 = \mathbb{P}[X = 0]$.

2. Compute $p_1 = \mathbb{P}[X = 1]$.

3. Does $\mathbb{E}[X]$ exist? If so, find its value; if not, explain why not.

   Solution:

1. $p_0 = P_X(0)$, so $p_0 = 0$,

---

2. $p_1 = P'_X(0)$, and

$$P'_X(s) = \frac{1}{\sqrt{1-s^2}(1+\sqrt{1-s^2})},$$

so $p_1 = \frac{1}{2}$.

3. If $\mathbb{E}[X]$ existed, it would be equal to $P'_X(1)$. However,

$$\lim_{s \nearrow 1} P'_X(s) = +\infty,$$

so $\mathbb{E}[X]$ (and, equivalently, $P'_X(1)$) does not exist.

Problem 14.15. Let N be a random time, independent of $\{\xi_n\}_{n\in\mathbb{N}}$, where $\{\xi_n\}_{n\in\mathbb{N}}$ is a sequence of mutually independent Bernoulli ($\{0,1\}$-valued) random variables with parameter $p_B \in (0,1)$. Suppose that N has a geometric distribution $g(p_g)$ with parameter $p_g \in (0,1)$. Compute the distribution of the random sum

$$Y = \sum_{k=1}^{N} \xi_k.$$

(*Note:* You can think of Y as a binomial random variable with "random n".)

Solution: Independence between N and $\{\xi_n\}_{n\in\mathbb{N}}$ allows us to use the fact that the generating function $P_Y(s)$ of Y is given by

$$P_Y(s) = P_N(P_B(s)),$$

where $P_N(s) = \frac{p_g}{1-q_g s}$ is the generating function of N (geometric distribution) and $P_B(s) = q_B + p_B s$ is the generating function of each ξ_k (Bernoulli distribution). Therefore,

$$P_Y(s) = \frac{p_g}{1 - q_g(q_b + p_B s)} = \frac{\frac{p_g}{1-q_g q_B}}{1 - \frac{q_g p_B}{1-q_g q_B}s} = \frac{p_Y}{1 - q_Y s}, \text{ where } p_Y = \frac{p_g}{1-q_g q_B} \text{ and } q_Y = 1 - p_Y.$$

P_Y can be recognized as the generating function of a geometric random variable with parameter p_Y.

Problem 14.16. Six fair gold coins are tossed, and the total number of *tails* is recorded; let's call this number N. Then, a set of three fair silver coins is tossed N times. Let X be the total number of times at least two *heads* are observed (among the N tosses of the set of silver coins).

(*Note:* A typical outcome of such a procedure would be the following: out of the six gold coins 4 were *tails* and 2 were *heads*. Therefore $N = 4$ and the 4 tosses of the set of three silver coins may look something like $\{HHT, THT, TTT, HTH\}$, so that $X = 2$ in this state of the world.)

Find the generating function *and* the pmf of X. You don't have to evaluate binomial coefficients.

Solution: Let H_k, $k \in \mathbb{N}$, be the number of heads on the kth toss of the set of three silver coins. The distribution of H_k is binomial so $\mathbb{P}[H_k \geq 2] = \mathbb{P}[H_k = 2] + \mathbb{P}[H_k = 3] = 3/8 + 1/8 = 1/2$. Let ξ_k be the indicator

$$\xi_k = \mathbf{1}_{\{H_k \geq 2\}} = \begin{cases} 1, & H_k \geq 2, \\ 0, & H_k < 2. \end{cases}$$

The generating function $P_\xi(s)$ of each ξ_k is a generating function of a Bernoulli random variable with $p = \frac{1}{2}$, i.e.

$$P_\xi(s) = \tfrac{1}{2}(1+s).$$

The random variable N is has a binomial distribution with parameters $n = 6$ and $p = \frac{1}{2}$. Therefore,

$$P_N(s) = (\tfrac{1}{2} + \tfrac{1}{2}s)^6.$$

By a theorem from the notes, the generating function of the value of the random sum X is given by

$$P_X(s) = P_N(P_\xi(s)) = \left(\tfrac{1}{2} + \tfrac{1}{2}(\tfrac{1}{2} + \tfrac{1}{2}s)\right)^6 = (\tfrac{3}{4} + \tfrac{1}{4}s)^6.$$

Therefore, X is a binomial random variable with parameters $n = 6$ and $p = 1/4$, i.e.

$$p_k = \mathbb{P}[X = k] = \binom{6}{k}\frac{3^{6-k}}{4^6}, \ k = 0, \ldots, 6.$$

$\square$

Problem 14.17. Tony Soprano collects his cut from the local garbage management companies. During a typical day he can visit a geometrically distributed number of companies with parameter $p = 0.1$. According to many years' worth of statistics gathered by his consigliere Silvio Dante, the amount he collects from the ith company is random with the following distribution

$$X_i \sim \left(\begin{array}{ccc} \$1000 & \$2000 & \$3000 \\ 0.2 & 0.4 & 0.4 \end{array}\right)$$

The amounts collected from different companies are independent of each other, and of the number of companies visited.

1. Find the (generating function of) the distribution of the amount of money S that Tony will collect on a given day.

2. Compute $\mathbb{E}[S]$ and $\mathbb{P}[S > 0]$.

Solution:

1. The number N of companies visited has the generating function

$$P_N(s) = \frac{0.1}{1 - 0.9s},$$

and the genererting function of the amount X_i collected from each one

$$P_X(s) = 0.2s^{1000} + 0.4s^{2000} + 0.4s^{3000}.$$

The total amount collected is given by the random sum

$$S = \sum_{i=1}^{N} X_i,$$

so the generating function of S is given by

$$P_S(s) = P_N(P_X(s)) = \frac{0.1}{1 - 0.9(0.2s^{1000} + 0.4s^{2000} + 0.4s^{3000})}.$$

2. The expectation of S is given by

$$\mathbb{E}[S] = P'_S(1) = P'_N(P_X(1))P'_X(1) = \mathbb{E}[N]\,\mathbb{E}[X_1] = 9 \times \$2200 = \$19800.$$

To compute $\mathbb{P}[S > 0]$, we note that $S = 0$ if and only if no collections have been made, i.e., if the value of N is equal to 0. Since N is geometric with parameter $p = 0.1$, so

$$\mathbb{P}[S > 0] = 1 - \mathbb{P}[S = 0] = 1 - \mathbb{P}[N = 0] = 1 - p = 0.9.$$

Problem 14.18. Let X be an $\mathbb{N}_0$-valued random variable, and let $P(s)$ be its generating function. Then the generating function of $4X$ is given by

1. $(P(s))^4$,

2. $P(P(P(P(s))))$,

3. $P(s^4)$,

4. $4P(s)$,

5. none of the above.

Solution: (3)

Problem 14.19. Let X and Y be two $\mathbb{N}_0$-valued random variables, let $P_X(s)$ and $P_Y(s)$ be their generating functions and let $Z = X - Y$, $V = X + Y$, $W = XY$. Then

1. $P_X(s) = P_Z(s)P_Y(s)$,

2. $P_X(s)P_Y(s) = P_Z(s)$,

3. $P_W(s) = P_X(P_Y(s))$,

4. $P_Z(s)P_V(s) = P_X(s)P_Y(s)$,

5. none of the above.

Solution: (5)

Problem 14.20. Let X be an $\mathbb{N}_0$-valued random variable, and let $P(s)$ be its generating function, and let $Q(s) = P(s)/(1 - s)$. Then

1. $Q(s)$ is a generating function of a random variable,

2. $Q(s)$ is a generating function of a non-decreasing sequence of non-negative numbers,

3. $Q(s)$ is a concave function on $(0, 1)$,

4. $Q(0) = 1$,

5. none of the above.

Solution: (2)

14.4   Random walks - advanced methods

Problem 14.21. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a simple random walk with $p = \mathbb{P}[X_1 = 1] \in (0, 1)$. Then

(a) X will visit every position $l \in \{\dots, -1, 0, 1, 2, \dots\}$ sooner or later, with probability 1

(b) $\mathbb{E}[X_n^2] < \mathbb{E}[X_{n+1}^2]$, for all $n \in \mathbb{N}_0$.

(c) $\mathbb{E}[X_n] < \mathbb{E}[X_{n+1}]$, for all $n \in \mathbb{N}_0$.

(d) X_{2n} has a binomial distribution.

(e) None of the above.

Solution: The correct answer is (b).

(a) False. When $p < \frac{1}{2}$, the probability of ever hitting the level $l = 1$ is less than 1.

(b) True. $X_{n+1} = X_n + \xi_{n+1}$, so that, by the independence of X_n and ξ_{n+1}, we have

$$\mathbb{E}[X_{n+1}^2] = \mathbb{E}[X_n^2] + 2\mathbb{E}[X_n]\mathbb{E}[\xi_{n+1}] + \mathbb{E}[\xi_{n+1}^2] > \mathbb{E}[X_n^2],$$

because $\mathbb{E}[X_n]\mathbb{E}[\xi] = n(2p - 1) \times (2p - 1) \geq 0$ and $\mathbb{E}[\xi^2] = 1 > 0$.

(c) False. $\mathbb{E}[X_{n+1}] - \mathbb{E}[X_n] = 2p - 1$, which could be negative if $p < \frac{1}{2}$.

(d) False. The support of X_{2n} is $\{-2n, -2n+2, \dots, 2n-2, 2n\}$, while the support of any binomial distribution has the form $\{0, 1, \dots, N\}$, for some $N \in \mathbb{N}$.

(e) False.

Problem 14.22. Throughout the problem, we let τ_k be the first hitting time of the level k, for $k \in \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ for a simple random walk $\{X_n\}_{n \in \mathbb{N}_0}$ with $p = \mathbb{P}[X_1 = 1] \in (0, 1)$. Furthermore, let P_k be the generating function of the distribution of τ_k.

1. When $p = \frac{1}{2}$, we have

   (a) $\mathbb{P}[\tau_1 < \infty] = 1$, $\mathbb{E}[\tau_1] < \infty$,
   (b) $\mathbb{P}[\tau_1 < \infty] = 0$, $\mathbb{E}[\tau_1] < \infty$,
   (c) $\mathbb{P}[\tau_1 < \infty] = 1$, $\mathbb{E}[\tau_1] = \infty$,
   (d) $\mathbb{P}[\tau_1 < \infty] \in (0, 1)$, $\mathbb{E}[\tau_1] = \infty$,
   (e) none of the above.

2. Either one of the following 4 random times is <u>not</u> a stopping time for a simple random walk $\{X_n\}_{n \in \mathbb{N}_0}$, or they all are. Choose the one which is not in the first case, or choose (e) if you think they all are.

   (a) the first hitting time of the level 4,

(b) the first time n such that $X_n - X_{n-1} \neq X_1$,

(c) the first time the walk hits the level 2 or the first time the walk sinks below -5, whatever happens <u>first</u>,

(d) the second time the walk crosses the level 5 or the third time the walk crosses the level -2, whatever happens <u>last</u>,

(e) none of the above.

3. We know that P_1 (the generating function of the distribution of the first hitting time of level 1) satisfies the following equation

$$P_1(s) = ps + qsP_1(s)^2.$$

The generating function P_{-1} of the first hitting time τ_{-1} of the level -1 then satisfies

(a) $P_{-1}(s) = ps + qsP_{-1}(s)^2$,

(b) $P_{-1}(s) = qs + psP_{-1}(s)^2$,

(c) $P_{-1}(s) = ps - qsP_{-1}(s)^2$,

(d) $P_{-1}(s) = ps + qsP_{-1}(s)^{-2}$,

(e) none of the above.

4. Let P_2 be the generating function of the first hitting time of the level 2. Then

(a) $P_2(s) = ps + qsP_2(s)^4$,

(b) $P_2(s) = ps^2 + qsP_2(s)$,

(c) $P_2(s) = P_1(s)$,

(d) $P_2(s) = P_1(s)^2$,

(e) none of the above.

Solution:

1. The correct answer is (c); see table at the very end of Lecture 7.

2. The correct answer is (e); first, second, or …hitting times are stopping times, and so are their minima or maxima. Note that for two stopping times T_1 and T_2, the one that happens *first* is $\min(T_1, T_2)$ and the one that happens *last* is $\max(T_1, T_2)$.

3. The correct answer is (b); hitting -1 is the same as hitting 1, if you switch the roles of p and q.

4. The correct answer is (d); to hit 2, you must first hit 1 and then hit 2, starting from 1. The times between these two are independent, and distributed like τ_1.

14.5 Branching processes

Problem 14.23. Let $\{Z_n\}_{n\in\mathbb{N}_0}$ be a simple branching process which starts from one individual. Each individual has exactly three children, each of whom survives until reproductive age with probability $0 < p < 1$, and dies before he/she is able to reproduce with probability $q = 1 - p$, independently of his/her siblings. The children that reach reproductive age reproduce according to the same rule.

1. Write down the generating function for the offspring distribution.

2. For what values of p will the population go extinct with probability 1.
   (*Hint:* You don't need to compute much. Just find the derivative $P'(1)$ and remember the picture from class →)
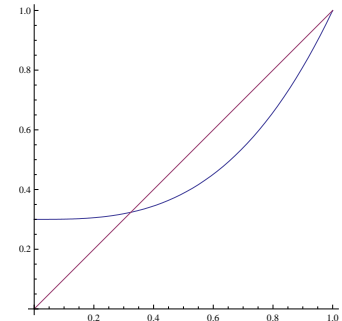
Solution:

1. The number of children who reach reproductive age is binomially distributed with parameters 3 and p. Therefore, $P(s) = (ps + q)^3$.

2. We know that the extinction happens with probability 1 if and only if the graphs of functions s and $P(s)$ meet only at $s = 1$, for $s \in [0,1]$. They will meet once again somewhere on $[0,1]$ if and only if $P'(1) > 1$. Therefore, the extinction happens with certainty if $P'(1) = 3p(p\,1 + q) = 3p \le 1$, i.e., if $p \le 1/3$.

Problem 14.24. Let $\{Z_n\}_{n\in\mathbb{N}_0}$ be a branching process with offspring distribution $(p_0, p_1, \dots)$. We set $P(s) = \sum_{n=0}^{\infty} p_n s^n$. The extinction is inevitable if

(a) $p_0 > 0$

(b) $P(\frac{1}{2}) > \frac{1}{4}$

(c) $P'(1) < 1$

(d) $P'(1) \ge 1$

(e) None of the above

Solution: The correct answer is (c).

(a) False. Take $p_0 = 1/3, p_1 = 0, p_2 = 2/3$, and $p_n = 0$, $n > 2$. Then the equation $P(s) = s$ reads $1/3 + 2/3 s^2 = s$, and $s = 1/2$ is its smallest solution. Therefore, the probability of extinction is $1/2$.

(b) False. Take $p_0 = 0, p_1 = 1/2, p_2 = 1/2$, and $p_n = 0$, $n > 2$. Then the equation $P(s) = s$ reads $1/2s + 1/2s^2 = s$, and $s = 0$ is its smallest solution. Therefore, the probability of extinction is 0, but $P(1/2) = 1/4 + 1/8 > 1/4$.

(c) True. The condition $P'(1) < 1$, together with the convexity of the function P imply that the graph of P lies strictly above the 45-degree line for $s \in (0,1)$. Therefore, the only solution of the equation $P(s) = s$ is $s = 1$.

(d) False. Take the same example as in (a), so that $P'(1) = 4/3 > 1$. On the other hand, the extinction equation $P(s) = s$ reads $1/3 + 2/3s^2 = s$. Its smallest solution if $s = 1/2$.

(e) False.

Problem 14.25. Bacteria reproduce by cell division. In a unit of time, a bacterium will either die (with probability $\frac{1}{4}$), stay the same (with probability $\frac{1}{4}$), or split into 2 parts (with probability $\frac{1}{2}$). The population starts with 100 bacteria at time $n = 0$.

1. Write down the expression for the generating function of the distribution of the size of the population at time $n \in \mathbb{N}_0$. (*Note:* you can use n-fold composition of functions.)

2. Compute the extinction probability for the population.

3. Let m_n be the largest possible number of bacteria at time n. Find m_n and compute the probability that there are exactly m_n bacteria in the population at time n, $n \in \mathbb{N}_0$.

4. Given that there are 1000 bacteria in the population at time 50, what is the expected number of bacteria at time 51?

Solution:

1. The generating function for the offspring distribution is

$$P(s) = \tfrac{1}{4} + \tfrac{1}{4}s + \tfrac{1}{2}s^2.$$

We can think of the population Z_n at time n, as the sum of 100 independent populations, each one produced by one of 100 bacteria in the initial population. The generating function for the number of offspring of each bacterium is

$$P^{(n)}(s) = \underbrace{P(P(\ldots P(s)\ldots))}_{n \text{ Ps}}$$

Therefore, since sums of independent variables correspond to products in the world of generating functions, the generating function of Z_n is given by

$$P_{Z_n}(s) = \left[P^{(n)}(s)\right]^{100}.$$

2. As in the practice problem, the extinction probability of the entire population (which starts from 100) is the 100-th power of the extinction probability p of the population which starts from a single bacterium. To compute p, we write down the extinction equation

$$s = \tfrac{1}{4} + \tfrac{1}{4}s + \tfrac{1}{2}s^2.$$

Its solutions are $s_1 = 1$ and $s_2 = \frac{1}{2}$, so $p = s_2 = \frac{1}{2}$. Therefore, the extinction probability for the entire population is $(\frac{1}{2})^{100}$.

3. The maximum size of the population at time n will occurr if each of the initial 100 bacteria splits into 2 each time, so $m_n = 100 \times 2^n$. For that to happen, there must be 100 splits to produce the first generation, 200 for the second, 400 for the third, ..., and $2^{n-1} \times 100$ for the n-th generation. Each one of those splits happens with probability $\frac{1}{2}$ and is independent of other splits in its generation, given the previous generations. It is clear that for $Z_n = m_n$, we must have $Z_{n-1} = m_{n-1}, \ldots, Z_1 = m_1$, so

$$\begin{aligned}
\mathbb{P}[Z_n = m_n] &= \mathbb{P}[Z_n = m_n, Z_{n-1} = m_{n-1}, \ldots Z_1 = m_1, Z_0 = m_0] \\
&= \mathbb{P}[Z_n = m_n | Z_{n-1} = m_{n-1}, Z_{n-2} = m_{n-2}, \ldots] \mathbb{P}[Z_{n-1} = m_{n-1}, Z_{n-2} = m_{n-2}, \ldots] \\
&= \mathbb{P}[Z_n = m_n | Z_{n-1} = m_{n-1}] \mathbb{P}[Z_{n-1} = m_{n-1} | Z_{n-2} = m_{n-2}, Z_{n-3} = m_{n-3}, \ldots] \mathbb{P}[Z_{n-2} = m_{n-2}, \ldots \\
&= \ldots \\
&= \mathbb{P}[Z_n = m_n | Z_{n-1} = m_{n-1}] \mathbb{P}[Z_{n-1} = m_{n-1} | Z_{n-2} = m_{n-2}] \ldots \mathbb{P}[Z_0 = m_0] \\
&= \left(\tfrac{1}{2}\right)^{m_{n-1}} \left(\tfrac{1}{2}\right)^{m_{n-2}} \ldots \left(\tfrac{1}{2}\right)^{m_0} \\
&= \left(\tfrac{1}{2}\right)^{100(1+2+\cdots+2^{n-1})} = \left(\tfrac{1}{2}\right)^{100 \times (2^n - 1)}
\end{aligned}$$

4. The expected number of offspring of each of the 1000 bacteria is given by $P'(1) = \frac{5}{4}$. Therefore, the total expected number of bacteria at time 51 is $1000 \times \frac{5}{4} = 1250$.

Problem 14.26. A branching process starts from 10 individuals, and each reproduces according to the probability distribution $(p_0, p_1, p_2, \ldots)$, where $p_0 = 1/4$, $p_1 = 1/4$, $p_2 = 1/2$, $p_n = 0$, for $n > 2$. The extinction probability for the whole population is equal to

(a) 1

(b) $\frac{1}{2}$

(c) $\frac{1}{20}$

(d) $\frac{1}{200}$

(e) $\frac{1}{1024}$

Solution: The correct answer is (e). The extinction probability for the population starting from each of the 10 initial individuals is given by the smallest solution of $1/4 + 1/4s + 1/2s^2 = s$, which is $s = 1/2$. Therefore, the extinction probability for the whole population is $(1/2)^{10} = 1/1024$.

Problem 14.27. A (solitaire) game starts with $3 \in \mathbb{N}$ silver dollars in the pot. At each turn the number of silver dollars in the pot is counted (call it K) and the following procedure is repeated K times: a die is thrown, and according to the outcome the following four things can happen

• If the outcome is 1 or 2 the player takes 1 silver dollar from the pot.

• If the outcome is 3 nothing happens.

• If the outcome is 4 the player puts 1 extra silver dollar in the pot (you can assume that the player has an unlimited supply of silver dollars).

• If the outcome is 5 or 6, the player puts 2 extra silver dollars in the pot.

If there are no silver dollars on in the pot, the game stops.

1. Compute the expected number of silver dollars in the pot after turn $n \in \mathbb{N}$.

2. Compute the probability that the game will stop eventually.

3. Let m_n be the maximal possible number of silver dollars in the pot after the n-th turn? What is the probability that the actual number of silver dollars in the pot after n turns is equal to $m_n - 1$?

Solution: The number of silver dollars in the pot follows a branching process $\{Z_n\}_{n \in \mathbb{N}_0}$ with $Z_0 = 3$ and offspring distribution whose generating function $P(s)$ is given by

$$P(s) = \tfrac{1}{3} + \tfrac{1}{6}s + \tfrac{1}{6}s^2 + \tfrac{1}{3}s^3.$$

1. The generating function P_{Z_n} of Z_n is $(P(P(\dots P(s)\dots)))^3$, so

$$\mathbb{E}[Z_n] = P'_{Z_n}(1) = 3(P'(1))^n = 3(\tfrac{1}{6} + 2\tfrac{1}{6} + 3\tfrac{1}{3})^n = \frac{3^{n+1}}{2^n}$$

2. The probability p that the game will stop is the extinction probability of the process. We solve the extinction equation $P(s) = s$ to get the extinction probability corresponding to the case $Z_0 = 1$, where there is 1 silver dollar in the pot:

$$P(s) = s \iff 2 + s + s^2 + 2s^3 = 6s \iff (s-1)(s+2)(2s-1) = 0.$$

The smallest solution in $[0,1]$ of the equation above is $1/2$. To get the true (corresponding to $Z_0 = 3$) extinction probability we raise it to the third power to get $p = \tfrac{1}{8}$.

3. The maximal number m_n of silver dollars in the pot is achieved if we roll 5 or 6 each time. In that case, there will be 3 rolls in the first turn, $9 = 3^2$ in the second, $27 = 3^3$ in the third, etc. That means that after n turns, there will be at most $m_n = 3^{n+1}$ silver dollars in the pot. In order to have exactly $m_n - 1$ silver dollars in the pot after n turns, we must keep getting 5s and 6s throughout the first $n - 2$ turns. The probability of that is

$$(\tfrac{1}{3})^{3^{n-1}+3^{n-2}+\dots+3} = (\tfrac{1}{3})^{\frac{1}{2}3^n - \frac{3}{2}}.$$

After that, we must get a 5 or a 6 in $3^n - 1$ throws and a 4 in a single throw during turn n. There are 3^n possible ways to choose the order of the throw which produces the single 4, so the later probability is

$$3^n(\tfrac{1}{3})^{3^n-1}(\tfrac{1}{6}) = \tfrac{1}{2}(\tfrac{1}{3})^{3^n-n}.$$

We multiply the two probabilities to get

$$\mathbb{P}[Z_n = m_n - 1] = \tfrac{1}{2}(\tfrac{1}{3})^{\frac{1}{2}3^{n+1}-n-\frac{3}{2}}.$$

Problem 14.28. It is a well-known fact(oid) that armadillos always have identical quadruplets (four offspring). Each of the 4 little armadillos has a 1/3 chance of becoming a doctor, a lawyer or a scientist, independently of its 3 siblings. A doctor armadillo will reproduce further with probability 2/3, a lawyer with probability 1/2 and a scientist with probability 1/4, again, independently of everything else. If it reproduces at all, an armadillo reproduces only once in its life, and then leaves the armadillo scene. (For the purposes of this problem assume that armadillos reproduce asexually.) Let us call the armadillos who have offspring *fertile*.

1. What is the distribution of the number of fertile offspring? Write down its generating function.

2. What is the generating function for the number of great-grandchildren an armadillo will have? What is its expectation? (*Note:* do not expand powers of sums)

3. Let the armadillo population be modeled by a branching process, and let's suppose that it starts from exactly one individual at time 0. Is it certain that the population will go extinct sooner or later?

Solution:

1. Each armadillo is fertile with probability p, where

$$p = \mathbb{P}[\text{ Fertile }] = \mathbb{P}[\text{ Fertile } | \text{ Lawyer }]\mathbb{P}[\text{ Lawyer }]$$
$$+ \mathbb{P}[\text{ Fertile } | \text{ Doctor }]\mathbb{P}[\text{ Doctor }]$$
$$+ \mathbb{P}[\text{ Fertile } | \text{ Scientist }]\mathbb{P}[\text{ Scientist }] = \tfrac{1}{2} \times \tfrac{1}{3} + \tfrac{2}{3} \times \tfrac{1}{3} + \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{5}{12}.$$

Therefore, the number of fertile offspring is binomial with $n = 4$ and $p = \tfrac{5}{12}$. The generating function of this distribution is $P(s) = (\tfrac{7}{12} + \tfrac{5}{12}s)^4$.

2. To get the number of great-grandchildren, we first compute the generating function $Q(s)$ of the number of fertile grandchildren. This is simply given by the composition of P with itself, i.e.,

$$P(P(s)) = (\tfrac{7}{12} + \tfrac{5}{12}(\tfrac{7}{12} + \tfrac{5}{12}s)^4)^4.$$

Finally, the number of great-grandchildren is the number of fertile grandchildren multiplied by 4. Therefore, its generating function is given by

$$Q(s) = P(P(s^4)) = (\tfrac{7}{12} + \tfrac{5}{12}(\tfrac{7}{12} + \tfrac{5}{12}s^4)^4)^4.$$

The compute the expectation, we need to evaluate $Q'(1)$:

$$Q'(s) = (P(P(s^4)))' = P'(P(s^4))P'(s^4)4s^3 \text{ and } P'(s) = 4\tfrac{5}{12}(\tfrac{7}{12} + \tfrac{5}{12}s)^3,$$

so that

$$Q'(1) = 4P'(1)P'(1) = \tfrac{100}{9}.$$

3. We need to consider the population of fertile armadillos. Its offspring distribution has generating function $P(s) = (\tfrac{7}{12} + \tfrac{5}{12}s)^4$, so the population will go extinct with certainty if and only if the extinction probability is 1, i.e., if $s = 1$ is the smallest solution of the extinction equation $s = P(s)$. We know, however, that $P'(1) = 5 \times \tfrac{5}{12} = \tfrac{10}{6} > 1$, so there exists a positive solution to $P(s) = s$ which is smaller than 1. Therefore, it is *not* certain that the population will become extinct sooner or later.

Problem 14.29. *Branching in alternating environments.* Suppose that a branching process $\{Z_n\}_{n \in \mathbb{N}_0}$ is constructed in the following way: it starts with one individual. The individuals in odd generations reproduce according to an offspring distribution with generating function $P_{odd}(s)$ and those in even generations according to an offspring distribution with generating function $P_{even}(s)$. All independence assumptions are the same as in the classical case.

1. Find an expression for the generating function P_{Z_n} of Z_n.

2. Derive the extinction equation.

   Solution:

1. Since $n = 0$ is an "even" generation, the distribution of the number of offspring of the initial individual is given by $P_{even}(s)$. Each of the Z_1 individuals in the generation $n = 1$ reproduce according the generating function P_{odd}, so

$$P_{Z_2}(s) = P_{\sum_{k=1}^{Z_1} Z_{2,k}}(s) = P_{even}(P_{odd}(s)).$$

   Similarly, $P_{Z_3}(s) = P_{even}(P_{odd}(P_{even}(s)))$, and, in general, for $n \in \mathbb{N}_0$,

$$P_{Z_{2n}}(s) = \underbrace{P_{even}(P_{odd}(P_{even}(\ldots P_{odd}(s)\ldots)))}_{2n \text{ Ps}}, \text{ and}$$

$$P_{Z_{2n+1}}(s) = \underbrace{P_{even}(P_{odd}(P_{even}(\ldots P_{even}(s)\ldots)))}_{2n+1 \text{ Ps}}.$$

(14.6)

2. To compute the extinction probability, we must evaluate the limit

$$p = \mathbb{P}[E] = \lim_{n \to \infty} \mathbb{P}[Z_n = 0] = \lim_{n \to \infty} P_{Z_n}(0).$$

   Since the limit exists (see the derivation in the classical case in the notes), we know that any subsequence also converges towards the same limit. In particular, $p = \lim_{n \to \infty} x_{2n}$, where

$$x_n = \mathbb{P}[Z_n = 0] = P_{Z_n}(0).$$

   By the expression for P_{Z_n} above, we know that $x_{2n+2} = P_{even}(P_{odd}(x_{2n}))$, and so

$$p = \lim_{n \to \infty} x_{2n+2} = \lim_n P_{even}(P_{odd}(x_{2n})) = P_{even}(P_{odd}(\lim_n x_{2n})) = P_{even}(P_{odd}(p)),$$

   which identifies

$$p = P_{even}(P_{odd}(p)),$$

(14.7)

   as the extinction equation. I will leave it up to you to figure out whether p is characterized as the smallest positive solution to (14.7).

Problem 14.30. In a branching process, the offspring distribution is given by its characteristic function

$$P(s) = as^2 + bs + c$$

where $a, b, c > 0$.

   (i) Find the extiction probability for this branching process.

(ii) Give a condition for sure extinction.

Solution:

(i) By Proposition 7.5 in the lecture notes, the extinction probability is the smallest non-negative solution to
$$P(x) = x.$$

So, in this case:
$$ax^2 + bx + c = x \implies ax^2 + (b-1)x + c = 0.$$

The possible solutions are
$$x_{1,2} = \frac{1}{2a}\left(-(b-1) \pm \sqrt{(b-1)^2 - 4ac}\right).$$

Recall that it is necessary that $P(1) = a + b + c = 1$. So,
$$\begin{aligned}
x_{1,2} &= \frac{1}{2a}\left(-(b-1) \pm \sqrt{(1-(a+c)-1)^2 - 4ac}\right) \\
&= \frac{1}{2a}\left(1 - b \pm \sqrt{(a+c)^2 - 4ac}\right) \\
&= \frac{1}{2a}\left(1 - b \pm \sqrt{(a-c)^2}\right) \\
&= \frac{1}{2a}\left(a + c \pm |a - c|\right).
\end{aligned}$$

Now, we differentiate two cases:

– If $a > c$, then
$$x_{1,2} = \frac{1}{2a}\left(a + c \pm |a - c|\right) = \frac{1}{2a}\left(a + c \pm (a - c)\right).$$

So,
$$x_1 = \frac{1}{2a}(a + c + a - c) = 1$$

and
$$x_2 = \frac{1}{2a}(a + c - a + c) = \frac{c}{a} < 1.$$

The solution we are looking for is c/a.

– If $a \le c$, then
$$x_{1,2} = \frac{1}{2a}\left(a + c \pm |a - c|\right) = \frac{1}{2a}\left(a + c \pm (c - a)\right).$$

So,
$$x_1 = \frac{1}{2a}(a + c + c - a) = \frac{c}{a} \ge 1$$

and
$$x_2 = \frac{1}{2a}(a + c - c + a) = \frac{a}{a} = 1.$$

The solution we are looking for is 1, i.e., in this case there is sure extinction.

(ii) The discussion above immediately yields the answer to question (ii): the necessary and sufficient condition for certain extinction is $a \leq c$.

Problem 14.31. The purpose of this problem is to describe a class of offspring distributions (pmfs) for which an expression for $P_{Z_n}(s)$ can be obtained.

An $\mathbb{N}_0$-valued distribution is said to be of *fractional-linear type* if its generating function P has the following form

$$P(s) = \frac{as + b}{1 - cs},$$ (14.8)

for some constants $a, b, c \geq 0$. In order for P to be a generating function of a probability distribution we must have $P(1) = 1$, i.e. $a + b + c = 1$, which will be assumed throughout the problem.

1. What (familiar) distributions correspond to the following special cases (in each case identify the distribution and its parameters):

   (a) $c = 0$
   (b) $a = 0$

2. Let A be the following 2×2 matrix

$$A = \begin{bmatrix} a & b \\ -c & 1 \end{bmatrix} \text{ and let } A^n = \begin{bmatrix} a^{(n)} & b^{(n)} \\ c^{(n)} & d^{(n)} \end{bmatrix}$$

   be its nth power (using matrix multiplication, of course). Using mathematical induction prove that

$$\underbrace{P(P(\ldots P(s)\ldots)}_{n \text{ Ps}} = \frac{a^{(n)}s + b^{(n)}}{c^{(n)}s + d^{(n)}}.$$ (14.9)

3. Take $a = 0$ and $b = c = 1/2$. Show inductively that

$$A^n = \frac{1}{2^n} \begin{bmatrix} -(n-1) & n \\ -n & n+1 \end{bmatrix}.$$ (14.10)

   Use that to write down the generating function of Z_n in the linear-fractional form (14.8).

4. Find the extinction probability as a function of a, b and c in the general case (don't forget to use the identity $a + b + c = 1$!).

Solution: *Note:* It seems that not everybody is familiar with the principle of mathematical induction. It is a logical device you can use if you have a conjecture and want to prove that it is true for all $n \in \mathbb{N}$. Your conjecture will be typically look like

For each $n \in \mathbb{N}$, the statement $P(n)$ holds.

where $P(n)$ is some assertion which depends on n (For example, $P(n)$ could be

$$\text{`` } 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \text{ ''} \tag{14.11}$$

The principle of mathematical induction says that you can prove that your statement is true for *all* $n \in \mathbb{N}$ by doing the following two things:

1. Prove the statement for $n = 1$, i.e., prove $P(1)$. (induction basis)

2. Prove that the implication $P(n) \Rightarrow P(n+1)$ always holds, i.e., prove the statement for $n+1$ if you are, additionally, allowed to use the statement for n as a hypothesis. (inductive step)

As an example, let us prove the statement (14.11) above. For $n = 1$, $P(1)$ reads " $1 = 1$ ", which is evidently true. Supposing that $P(n)$ holds, i.e., that

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2},$$

we can add $n + 1$ to both sides to get

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} = \frac{n^2 + 3n + 2}{2} = \frac{(n+1)(n+2)}{2},$$

which is exactly $P(n+1)$. Therefore, we managed to prove $P(n+1)$ using $P(n)$ as a crutch. The principle of mathematical induction says that this is enough to be able to conclude that $P(n)$ holds for each n, i.e., that (14.11) is a ture statement for all $n \in \mathbb{N}$.

Back to the solution of the problem:

1. (a) When $c = 0$, $P(s) = as + b$ - Beronulli distribution with success probability $a = 1 - b$.

   (b) For $a = 0$, $P(s) = \frac{b}{1-cs}$ - Geometric distribution with success probability $b = 1 - c$.

2. For $n = 1$, $a^{(1)} = a$, $b^{(1)} = b$, $c^{(1)} = -c$ and $d^{(1)} = 1$, so clearly

$$P(s) = \frac{a^{(1)} s + b^{(1)}}{c^{(1)} s + d^{(1)}}.$$

Suppose that the equality (14.9) holds for some $n \in \mathbb{N}$. Then

$$\begin{bmatrix} a^{(n+1)} & b^{(n+1)} \\ c^{(n+1)} & d^{(n+1)} \end{bmatrix} = A^{n+1} = A^n A = \begin{bmatrix} a^{(n)} & b^{(n)} \\ c^{(n)} & d^{(n)} \end{bmatrix} \begin{bmatrix} a & b \\ -c & 1 \end{bmatrix} = \begin{bmatrix} a\,a^{(n)} - c\,b^{(n)} & b\,a^{(n)} + b^{(n)} \\ a\,c^{(n)} - c\,d^{(n)} & b\,c^{(n)} + d^{(n)} \end{bmatrix}$$

On the other hand, by the inductive assumption,

$$\underbrace{P(P(\ldots P(s)\ldots))}_{n+1 \text{ Ps}} = \frac{a^{(n)} P(s) + b^{(n)}}{c^{(n)} P(s) + d^{(n)}} = \frac{a^{(n)} \frac{as+b}{1-cs} + b^{(n)}}{c^{(n)} \frac{as+b}{1-cs} + d^{(n)}}$$

$$= \frac{(a\,a^{(n)} - c\,b^{(n)})s + b\,a^{(n)} + b^{(n)}}{(a\,c^{(n)} - c\,d^{(n)})s + b\,c^{(n)} + d^{(n)}}.$$

Therefore, (14.9) also holds for $n + 1$. By induction, (14.9) holds for all $n \in \mathbb{N}$.

3. Here
$$A = \begin{bmatrix} 0 & 1/2 \\ -1/2 & 1 \end{bmatrix} = \tfrac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} = \tfrac{1}{2^1} \begin{bmatrix} (1-1) & 1 \\ -1 & 1+1 \end{bmatrix}$$

so the statement holds for $n = 1$ (induction basis). Suppose that (14.10) holds for some n. Then

$$A^{n+1} = A^n A = \tfrac{1}{2^n} \begin{bmatrix} -(n-1) & n \\ -n & n+1 \end{bmatrix} \tfrac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} = \tfrac{1}{2^{n+1}} \begin{bmatrix} -n & -(n-1)+2n \\ -n-1 & -n+2(n+1) \end{bmatrix}$$
$$= \tfrac{1}{2^{n+1}} \begin{bmatrix} -n & n+1 \\ -(n+1) & n+2 \end{bmatrix}$$

which is exactly (14.10) for $n + 1$ (inductive step). Thus, (14.10) holds for all $n \in \mathbb{N}$. By the previous part, the generating function of Z_n is given by

$$P_{Z_n}(s) = \frac{-\tfrac{1}{2}(n-1)s + \tfrac{1}{2}n}{-\tfrac{1}{2}ns + \tfrac{1}{2}(n+1)}.$$

We divide the numerator and the denominator by $\tfrac{1}{2}(n+1)$ to get the above expression into the form dictated by (14.8):
$$P_{Z_n}(s) = \frac{-\tfrac{n-1}{n+1}s + \tfrac{n}{n+1}}{1 - \tfrac{n}{n+1}s}.$$

Note that $a = -\tfrac{n-1}{n+1}$, $b = \tfrac{n}{n+1}$ and $c = \tfrac{n}{n+1}$ so that $a + b + c = \tfrac{-n+1+n+n}{n+1} = 1$, as required.

4. For the extinction probability we need to find the smallest solution of

$$s = \frac{as + b}{1 - cs}$$

in $[0, 1]$. The equation above transforms into a quadratic equation after we multiply both sides by $1 - cs$

$$s - cs^2 = as + b, \text{ i.e., } cs^2 + (a-1)s + b = 0. \tag{14.12}$$

We know that $s = 1$ is a solution and that $a - 1 = -b - c$, so we can factor (14.12) as

$$cs^2 + (a-1)s + b = (s-1)(cs - b).$$

If $c = 0$, the only solution is $s = 1$. If $c \neq 0$, the solutions are 1 and $\tfrac{b}{c}$. Therefore, the extinction probability $\mathbb{P}[E]$ is given by

$$\mathbb{P}[E] = \begin{cases} 1, & c = 0, \\ \min(1, \tfrac{b}{c}), & \text{otherwise .} \end{cases}$$

Problem 14.32. (Too hard to appear on an exam, but see if you can understand its solution. Don't worry if you can't.) For a branching process $\{Z_n\}_{n \in \mathbb{N}_0}$, denote by S the total number of individuals that ever lived, i.e., set

$$S = \sum_{n=0}^{\infty} Z_n = 1 + \sum_{n=1}^{\infty} Z_n.$$

(i) Assume that the offspring distribution has the generating function given by

$$P(s) = p + qs.$$

Find the generating function P_S in this case.

(ii) Assume that the offspring distribution has the generating function given by

$$P(s) = p/(1 - qs).$$

Find P_S in this case.

(iii) Find the general expression for $\mathbb{E}[S]$ and calculate this expectation in the special cases (i) and (ii).

Solution: In general, the r.v. S satisfies the relationship

$$P_S(s) = \mathbb{E}[s^{1 + \sum_{n=1}^{\infty} Z_n}] = \sum_{k=0}^{\infty} \mathbb{E}[s^{1 + \sum_{n=1}^{\infty} Z_n} | Z_1 = k] \mathbb{P}[Z_1 = k]$$

$$= \sum_{k=0}^{\infty} s \mathbb{E}[s^{Z_1 + \sum_{n=2}^{\infty} Z_n} | Z_1 = k] \mathbb{P}[Z_1 = k]$$

When $Z_1 = k$, the expression $Z_1 + \sum_{n=2}^{\infty} Z_n$ counts the total number of individuals in k separate and independent Branching processes - one for each of k members of the generation at time $k = 1$. Since this random variable is the sum of k independent random variables, each of which has the same distribution as S (why?), we have

$$\mathbb{E}[s^{Z_1 + \sum_{n=2}^{\infty} Z_n} | Z_1 = k] = [P_S(s)]^k.$$

Consequently, P_S is a solution of the following equation

$$P_S(s) = s \sum_{k=0}^{\infty} [P_S(s)]^k \mathbb{P}[Z_1 = k] = sP(P_S(s)).$$

(i) In this case,

$$P_S(s) = s(p + qP_S(s)) \Rightarrow P_S(s) = \frac{sp}{1 - sq}.$$

(iii) Here, $P_S(s)$ must satisfy

$$P_S(s)(1 - qP_S(s)) - sp = 0,$$

i.e.,

$$qP_S(s)^2 - P_S(s) + sp = 0.$$

Solving the quadratic, we get as the only sensible solution (the one that has $P_S(0) = 0$):

$$P_S(s) = \frac{1 - \sqrt{1 - 4spq}}{2q}.$$

Note that $P_S(1) < 1$ if $p < q$. Does that surprise you? Should not: it is possible that $S = +\infty$.

(iii) Using our main "recursion" for $P_S(s)$, we get

$$P_S'(s) = (sP_{Z_1}(P_S(s)))' = P_{Z_1}(P_S(s)) + sP_{Z_1}'(P(s))P'(s).$$

So, for $s = 1$,

$$\mathbb{E}[S] = P_S'(1) = P_{Z_1}(P_S(1)) + P_{Z_1}'(P_S(1))P_s'(1) = 1 + \mu\mathbb{E}[S].$$

If $\mu \geq 1$, $\mathbb{E}[S] = +\infty$. If $\mu < 1$,

$$\mathbb{E}[S] = \frac{1}{1 - \mu}.$$

14.6   Markov chains - classification of states

Problem 14.33. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a simple symmetric random walk, and let $\{Y_n\}_{n \in \mathbb{N}_0}$ be a random process whose value at time $n \in \mathbb{N}_0$ is equal to the amount of time (number of steps, including possibly n) the process $\{X_n\}_{n \in \mathbb{N}_0}$ has spent above 0, i.e., in the set $\{1, 2, \dots\}$. Then

(a) $\{Y_n\}_{n \in \mathbb{N}_0}$ is a Markov process

(b) Y_n is a function of X_n for each $n \in \mathbb{N}$.

(c) X_n is a function of Y_n for each $n \in \mathbb{N}$.

(d) Y_n is a stopping time for each $n \in \mathbb{N}_0$.

(e) None of the above

Solution: The correct answer is (e).

(a) False. The event $\{Y_3 = 1\}$ corresponds to exactly two possible paths of the random walk - $\{X_1 = 1, X_2 = 0, X_3 = -1\}$ and $\{X_1 = -1, X_2 = 0, X_3 = 1\}$, each occurring with probability $\frac{1}{8}$. In the first case, there is no way that $Y_4 = 2$, and in the second one, $Y_4 = 2$ if and only if, additionally, $X_4 = 2$. Therefore,

$$\mathbb{P}[Y_4 = 2|Y_3 = 1] = \frac{1}{\mathbb{P}[Y_3 = 1]}\mathbb{P}[Y_4 = 2 \text{ and } Y_3 = 1]$$
$$= 4\mathbb{P}[X_1 = -1, X_2 = 0, X_3 = 1, X_4 = 2] = \tfrac{1}{4}. \tag{14.13}$$

On the other hand,

$$\mathbb{P}[Y_4 = 2|Y_3 = 1, Y_2 = 1, Y_1 = 1, Y_0 = 0] = \frac{\mathbb{P}[Y_4 = 2 \text{ and } X_1 = 1, X_2 = 0, X_3 = -1]}{\mathbb{P}[X_1 = 1, X_2 = 0, X_3 = -1]} = 0.$$

(b) False. Y_n is a function of the entire past $X_0, X_1, \ldots, X_n$, but not of the individual value X_n.

(c) False. Except in trivial cases, it is impossible to know the value of X_n if you only know how many of the past n values are positive.

(d) False. The fact that, e.g., $Y_{10} = 1$ (meaning that X hits 1 exactly once in its first 10 steps and immediately returns to 0) cannot possibly be known at time 1.

(e) True.

Problem 14.34. Suppose that all classes of a Markov chain are recurrent, and let i, j be two states such that $i \to j$. Then

(a) for each state k, either $i \to k$ or $j \to k$

(b) $j \to i$

(c) $p_{ji} > 0$ or $p_{ij} > 0$

(d) $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$

(e) None of the above

Solution: The correct answer is (b).

(a) False. Take a chain with two states $1, 2$ where $p_{11} = p_{22} = 1$, and set $i = j = 1$, $k = 2$.

(b) True. Recurrent classes are closed, so i and j belong to the same class. Therefore $j \to i$.

(c) False. Take a chain with 4 states $1, 2, 3, 4$ where $p_{12} = p_{23} = p_{34} = p_{41} = 1$, and set $i = 1$, $j = 3$.

(d) False. That would mean that j is transient.

(e) False.

Problem 14.35. Which of the following statements is true? Give a short explanation (or a counterexample where appropriate) for your choice. Here, $\{X_n\}_{n\in\mathbb{N}_0}$ be a Markov chain with state space $\mathbb{Z}$.

1. $\{X_n^2\}_{n\in\mathbb{N}_0}$ must be a Markov chain.

2. $\{X_n^3\}_{n\in\mathbb{N}_0}$ does not have to be a Markov chain.

3. There exists a set $\mathcal{T}$ and a function $f : \mathcal{S} \to \mathcal{T}$ such that $\{f(X_n)\}_{n\in\mathbb{N}_0}$ is not a Markov chain.

4. There exists a set $\mathcal{T}$ and a function $f : \mathcal{S} \to \mathcal{T}$ which is *not* a bijection (one-to-one and onto) such that $\{f(X_n)\}_{n\in\mathbb{N}}$ is a Markov chain,

5. None of the above.

Problem 14.36. Two containers are filled with ping-pong balls. The red container has 100 red balls, and the blue container has 100 blue balls. In each step a container is selected; red with probability 1/2 and blue with probability 1/2. Then, a ball is selected from it - all balls in the container are equally likely to be selected - and placed in the other container. If the selected container is empty, no ball is transfered.

Once there are 100 blue balls in the red container and 100 red balls in the blue container, the game stops (and we stay in that state forever).

We decide to model the situation as a Markov chain.

1. What is the state space $\mathcal{S}$ we can use? How large is it?

2. What is the initial distribution?

3. What are the transition probabilities between states? Don't write the matrix, it is way too large; just write a general expression for p_{ij}, $i, j \in \mathcal{S}$.

4. How many transient states are there?

Solution: There are many ways in which one can solve this problem. Below is just one of them.

1. In order to describe the situation being modeled, we need to keep track of the number of balls of each color in each container. Therefore, one possibility is to take the set of all quadruplets (r, b, R, B), $r, b, R, b \in \{0, 1, 2, \ldots, 100\}$ and this state space would have 101^4 elements. We know, however, that the total number of red balls, and the total number of blue balls is always equal to 100, so the knowledge of the composition of the red (say) container is enough to reconstruct the contents of the blue container. In other words, we can use the number of balls of each color in the red container only as our state, i.e.

$$\mathcal{S} = \{(r, b) : r, b = 0, 1, \ldots, 100\}.$$

This state space has $101 \times 101 = 10201$ elements.

2. The initial distribution of the chain is deterministic: $\mathbb{P}[X_0 = (100, 0)] = 1$ and $\mathbb{P}[X_0 = i] = 0$, for $i \in \mathcal{S} \setminus \{(100, 0)\}$. In the vector notation,

$$a^{(0)} = (0, 0, \dots, 1, 0, 0, \dots, 0),$$

where 1 is at the place corresponding to $(100, 0)$.

3. Let us consider several separate cases, under the understandin that $p_{ij} = 0$, for all i, j not mentioned explicitely below:

   (a) *One of the containers is empty.* In that case, we are either in $(0, 0)$ or in $(100, 100)$. Let us describe the situation for $(0, 0)$ first. If we choose the red container - and that happens with probability $\frac{1}{2}$ - we stay in $(0, 0)$:

   $$p_{(0,0),(0,0)} = \frac{1}{2}.$$

   If the blue container is chosen, a ball of either color will be chosen with probability $\frac{100}{200} = \frac{1}{2}$, so

   $$p_{(0,0),(1,0)} = p_{(0,0),(0,1)} = \frac{1}{4}.$$

   By the same reasoning,

   $$p_{(100,100),(0,0)} = \frac{1}{2} \text{ and } p_{(100,100),(99,100)} = p_{(100,100),(100,99)} = \frac{1}{4}.$$

   (b) *We are in the state $(0, 100)$.* By the description of the model, this is an absorbing state, so

   $$p_{(0,100),(0,100)} = 1.$$

   (c) *All other states.* Suppose we are in the state (r, b) where $(r, b) \notin \{(0, 100), (0, 0), (100, 100)\}$. If the red container is chosen, then the probability of getting a red ball is $\frac{r}{r+b}$, so

   $$p_{(r,b),(r-1,b)} = \frac{1}{2} \frac{r}{r+b}.$$

   Similarly,

   $$p_{(r,b),(r,b-1)} = \frac{1}{2} \frac{b}{r+b}.$$

   In the blue container there are $100 - r$ red and $100 - b$ blue balls. Thus,

   $$p_{(r,b),(r+1,b)} = \frac{1}{2} \frac{100-r}{200-r-b},$$

   and

   $$p_{(r,b),(r,b+1)} = \frac{1}{2} \frac{100-b}{200-r-b}.$$

   (d) The state $(0, 100)$ is clearly recurrent (it is absorbing). Let us show that all other (10200 of them) are transient. It will be enough to show that $i \to (0, 100)$, for each $i \in \mathcal{S}$, because $(0, 100)$ is in a recurrent class by itself. To do that, we just need to find a path from any state (r, b) to $(0, 100)$ through the transition graph:

   $$(r, b) \to (r - 1, b) \to \cdots \to (0, b) \to (0, b + 1) \to \cdots \to (0, 100).$$

Problem 14.37. Let $\{Z_n\}_{n \in \mathbb{N}_0}$ be a Branching process (with state space $\mathcal{S} = \{0, 1, 2, 3, 4, \dots\} = \mathbb{N}_0$) with the offspring probability given by $p_0 = 1/2$, $p_2 = 1/2$. Classify the states (find classes), and describe all closed sets.

Solution: If $i = 0$, the i is an absorbing state (and, therefore, a class in itself). For $i \in \mathbb{N}$, $p_{ij} > 0$ if j is of the form $2k$ for some $k \le i$ (out of i individuals, $i - k$ die and the remaining k have two children). In particular, noone ever communicates with odd $i \in \mathbb{N}$. Therefore, each odd i is a class in itself. For even i, it communicates with all even numbers from 0 to $2i$; in particular, $i \to i + 2$ and $i \to i - 2$. Therefore, all positive even numbers intercommunicate and form a class.

   Clearly, $\{0\}$ is a closed set. Every other closed set will contain a positive even number; indeed, after one step we are in an even state, no matter where be start from, and this even number could be positive, unless the initial state was 0). Therefore, any closed set different from $\{0\}$ will have to contain the set of all even numbers $2\mathbb{N}_0 = \{0, 2, 4, \dots\}$, and, $2\mathbb{N}_0$ is a closed set itself. On the other hand, if we add any collection of odd numbers to $2\mathbb{N}_0$, the resulting set will still be closed (exiting it would mean hitting an odd number outside of it). Therefore, closed sets are exactly the sets of the form $C = \{0\}$ or $C = 2\mathbb{N}_0 \cup D$, where D is any set of odd numbers.

Problem 14.38. Which of the following statements is true? Give a short explanation (or a counterexample where appropriate) for your choice. $\{X_n\}_{n \in \mathbb{N}_0}$ is a Markov chain with state space $\mathcal{S}$.

1. $p_{ij}^{(n)} \ge f_{ij}^{(n)}$ for all $i, j \in \mathcal{S}$ and all $n \in \mathbb{N}$.

2. If states i and j intercommunicate, then there exists $n \in \mathbb{N}$ such that $p_{ij}^{(n)} > 0$ and $p_{ji}^{(n)} > 0$.

3. If all rows of the transition matrix are equal, then all states belong to the same class.

4. If $f_{ij} < 1$ and $f_{ji} < 1$, then i and j do not intercommunicate.

5. If $P^n \to I$, then all states are recurrent. (*Note:* We say that a sequence $\{A_n\}_{n \in \mathbb{N}}$ of matrices converges to the matrix A, and we denote it by $A_n \to A$, if $(A_n)_{ij} \to A_{ij}$, as $n \to \infty$, for all i, j.)

Solution:

1. **TRUE.** $p_{ij}^{(n)}$ is the probability that $X_n = j$ (given that $X_0 = i$). On the other hand $f_{ij}^{(n)}$ is the probability that $X_n = j$ *and* that $X_k \ne j$ for all $k = 1, \dots, n - 1$ (given $X_0 = i$). Clearly, the second event is a subset of the first, so $f_{ij}^{(n)} \le p_{ij}^{(n)}$.

2. **FALSE**. Consider the following Markov chain

.
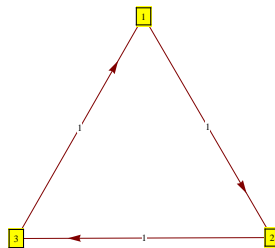
All states intercommunicate, but $p_{12}^{(n)} > 0$ if and only if n is of the form $3k + 1$, for $k \in \mathbb{N}_0$. On the other hand $p_{21}^{(n)} > 0$ if and only if $n = 3k + 2$, for some $k \in \mathbb{N}_0$. Thus, $p_{12}^{(n)}$ and $p_{21}^{(n)}$ are never simultaneously positive.

3. **FALSE**. Consider a Markov chain with the following transition matrix:

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then 1 is an absorbing state and it is in a class of its own, so it is not true that all states belong to the same class.

4. **FALSE**. Consider the following Markov chain:

.
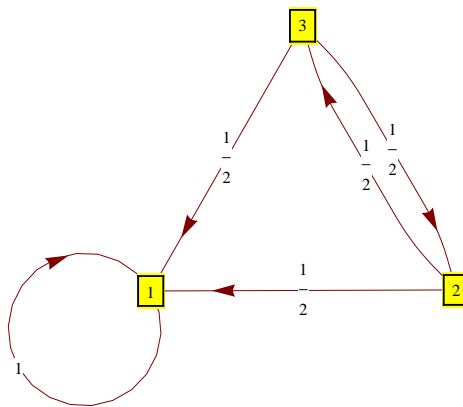
Then the states 2 and 3 clearly intercommunicate. In the event that (starting from 2), the first transition happens to be $2 \to 1$, we are never visiting 3. Otherwise, we visit 3 at time 1. Therefore, $f_{23} = 1/2 < 1$. Similarly, $f_{32} = 1/2 < 1$.

5. **TRUE.** Suppose that there exists a transient state $i \in \mathcal{S}$. Then $\sum_n p_{ii}^{(n)} < \infty$, and, in particular, $p_{ii}^{(n)} \to 0$, as $n \to \infty$. This is a contradiction with the assumption that $p_{ii}^{(n)} \to 1$, for all $i \in \mathcal{S}$.

Problem 14.39. Consider a Markov Chain whose transition graph is given below
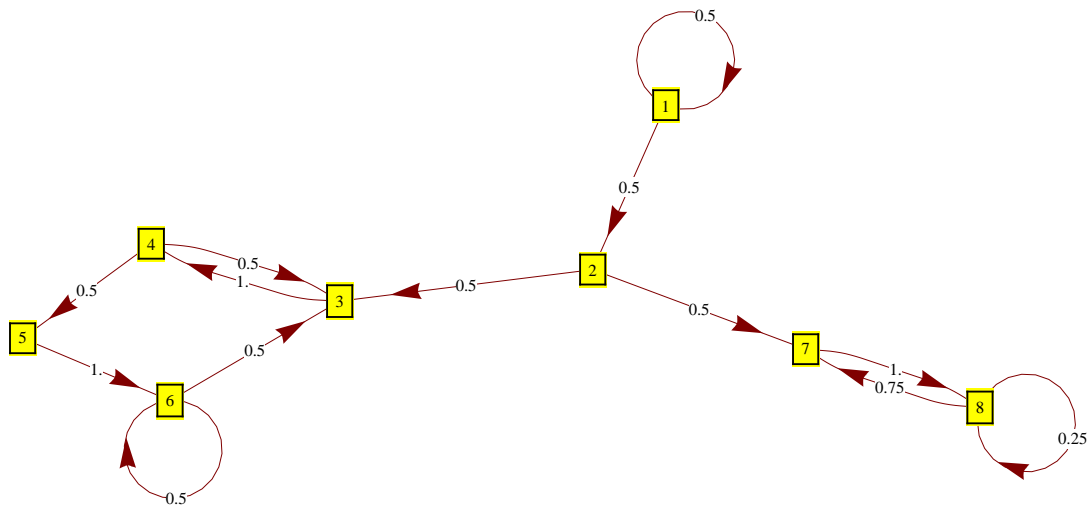
1. Identify the classes.

2. Find transient and recurrent states.

3. Find periods of all states.

4. Compute $f_{13}^{(n)}$, for all $n \in \mathbb{N}$.

5. Using *Mathematica*, we can get that, approximately,

$$
P^{20} = \begin{bmatrix}
0 & 0 & 0.15 & 0.14 & 0.07 & 0.14 & 0.21 & 0.29 \\
0 & 0 & 0.13 & 0.15 & 0.07 & 0.15 & 0.21 & 0.29 \\
0 & 0 & 0.3 & 0.27 & 0.15 & 0.28 & 0 & 0 \\
0 & 0 & 0.27 & 0.3 & 0.13 & 0.29 & 0 & 0 \\
0 & 0 & 0.29 & 0.28 & 0.15 & 0.28 & 0 & 0 \\
0 & 0 & 0.28 & 0.29 & 0.14 & 0.29 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.43 & 0.57 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.43 & 0.57
\end{bmatrix},
$$

where P is the transition matrix of the chain. Compute the probability $\mathbb{P}[X_{20} = 3]$, if the initial distribution (the distribution of X_0) is given by $\mathbb{P}[X_0 = 1] = 1/2$ and $\mathbb{P}[X_0 = 3] = 1/2$.

Solution:

1. The classes are $T_1 = \{1\}$, $T_2 = \{2\}$, $C_1 = \{3, 4, 5, 6\}$ and $C_2 = \{7, 8\}$

2. The states in T_1 and T_2 are transient, and the others are recurrent.

3. The periods are all 1.

4. For $n = 1$, $f_{13}^{(n)} = 0$, since you need at least two steps to go from 1 to 3. For $n = 2$, the chain needs to follow $1 \to 2 \to 3$, for $f_{13}^{(n)} = (1/2)^2$. For n larger than 2, the only possibility is for the chain to stay in the state 1 for $n - 2$ periods, jump to 2 and finish at 3, so $f_{13}^{(n)} = (1/2)^n$ in that case. All together, we have

$$f_{13}^{(n)} = \begin{cases} 0, & n = 1 \\ (1/2)^n, & n \geq 2. \end{cases}$$

5. We know from the notes that the distribution of X_{20}, when represented as a vector $\boldsymbol{a}^{(20)} = (a_1^{(20)}, a_2^{(20)}, \ldots, a_8^{(20)})$ satisfies

$$\boldsymbol{a}^{(20)} = \boldsymbol{a}^{(0)} P^{20}.$$

By the assumption $\boldsymbol{a}^{(0)} = (\frac{1}{2}, 0, \frac{1}{2}, 0, 0, 0, 0, 0)$, so $\mathbb{P}[X_{20} = 3] = a_3^{(20)} = \frac{1}{2}0.15 + \frac{1}{2}0.3 = 0.225$.

Problem 14.40. A country has $m + 1$ cities $(m \in \mathbb{N})$, one of which is the capital. There is a direct railway connection between each city and the capital, but there are no tracks between any two "non-capital" cities. A traveler starts in the capital and takes a train to a randomly chosen non-capital city (all cities are equally likely to be chosen), spends a night there and returns the next morning and immediately boards the train to the next city according to the same rule, spends the night there, ..., etc. We assume that his choice of the city is independent of the cities visited in the past. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be the number of visited non-capital cities up to (and including) day n, so that $X_0 = 1$, but X_1 could be either 1 or 2, etc.

1. Explain why $\{X_n\}_{n \in \mathbb{N}_0}$ is a Markov chain on the appropriate state space $\mathcal{S}$ and the find the transition probabilities of $\{X_n\}_{n \in \mathbb{N}_0}$, i.e., write an expression for

$$\mathbb{P}[X_{n+1} = j | X_n = i], \text{ for } i, j \in \mathcal{S}.$$

2. Find recurrent and transient classes and periods of all states. Sketch the transition graph for $m = 3$.

3. Let τ_m be the first time the traveler has visited all m non-capital cities, i.e.

$$\tau_m = \min\{n \in \mathbb{N}_0 : X_n = m\}.$$

What is the distribution of τ_m, for $m = 1$ and $m = 2$.

4. Compute $\mathbb{E}[\tau_m]$ for general $m \in \mathbb{N}$. (*Note:* you don't need to use any heavy machinery. In particular, no knowledge of the "absorption and reward" techniques are necessary.)

Solution:

1. The natural state space for $\{X_n\}_{n \in \mathbb{N}_0}$ is $\mathcal{S} = \{1, 2, \ldots, m\}$. It is clear that $\mathbb{P}[X_{n+1} = j | X_n = i] = 0$, unless, $i = j$ or $i = j + 1$. If we start from the state i, the process will remain in i if the traveler visits one of the already-visited cities, and move to $i + 1$ is the visited city has never been visited before. Thanks to the uniform distribution in the choice of the next city, the probability that a never-visited city will be selected is $\frac{m-i}{m}$, and it does not depend on the (names of the) cities already visited, or on the times of their first visits; it only depends on

their number. Consequently, the extra information about $X_1, X_2, \ldots, X_{n-1}$ will not change the probability of visiting j in any way, which is exactly what the Markov property is all about. Therefore, $\{X_n\}_{n \in \mathbb{N}}$ is Markov and its transition probabilities are given by

$$p_{ij} = \mathbb{P}[X_{n+1} = j | X_n = i] = \begin{cases} 0, & j \notin \{i, i+1\} \\ \frac{m-i}{m}, & j = i+1 \\ \frac{i}{m}, & j = i. \end{cases}$$

(*Note:* the situation would not be nearly as nice if the distribution of the choice of the next city were non-uniform. In that case, the list of the (names of the) already-visited cities would matter, and it is not clear that the described process has the Markov property (does it?).)

2. Once the chain moves to the next state, it can never go back. Therefore, it cannot happen that $i \leftrightarrow j$ unless $i = j$, and so, every state is a class of its own. The class $\{m\}$ is clearly absorbing and, therefore, recurrent. All the other classes are transient since it is possible (and, in fact, certain) that the chain will move on to the next state and never come back. As for periods, all of them are 1, since 1 is in the return set of each state.
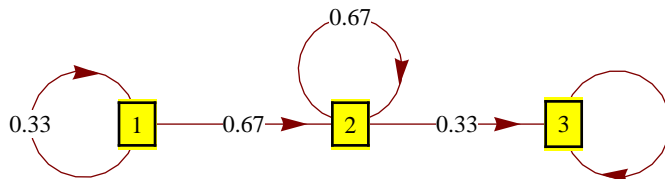
3. For $m = 1$, $\tau_m = 0$, so its distribution is deterministic and concentrated on 0. The case $m = 2$ is only slightly more complicated. After having visited his first city, the visitor has a probability of $\frac{1}{2}$ of visiting it again, on each consecutive day. After a geometrically distributed number of days, he will visit another city and τ_2 will be realized. Therefore the distribution $\{p_n\}_{n \in \mathbb{N}_0}$ of τ_2 is given by

$$p_0 = 0, p_1 = \tfrac{1}{2}, p_2 = (\tfrac{1}{2})^2, p_3 = (\tfrac{1}{2})^3, \ldots$$

4. For $m > 1$, we can write τ_m as

$$\tau_m = \tau_1 + (\tau_2 - \tau_1) + \cdots + (\tau_m - \tau_{m-1}),$$

so that

$$\mathbb{E}[\tau_m] = \mathbb{E}[\tau_1] + \mathbb{E}[\tau_2 - \tau_1] + \cdots + \mathbb{E}[\tau_m - \tau_{m-1}].$$

We know that $\tau_1 = 0$ and for $k = 1, 2, \ldots, m-1$, the difference $\tau_{k+1} - \tau_k$ denotes the waiting time before a never-before-visited city is visited, given that the number of already-visited cities is k. This random variable is geometric with sucess probability given by $\frac{m-k}{m}$, so its expectation is given by

$$\mathbb{E}[\tau_{k+1} - \tau_k] = \frac{1}{\frac{m-k}{m}} = \frac{m}{m-k}.$$

Therefore,

$$\mathbb{E}[\tau_m] = \sum_{k=1}^{m-1} \frac{m}{m-k} = m(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{m-1}).$$

(*Note:* When m is large, the partial harmonic sum $H_m = 1 + \tfrac{1}{2} + \cdots + \tfrac{1}{m-1}$ behaves like $\log m$, so that, asymptotically, $\mathbb{E}[\tau_m]$ behaves like $m \log m$.)

14.7 Markov chains - absorption and reward

Problem 14.41. In a Markov chain with a finite number of states, the fundamental matrix is given by $F = \begin{bmatrix} 3 & 4 \\ \frac{3}{2} & 4 \end{bmatrix}$. The initial distribution of the chain is uniform on all transient states. The expected value of

$$\tau_C = \min\{n \in \mathbb{N}_0 : X_n \in C\},$$

where C denotes the set of all recurrent states is

  (a) 7

  (b) 8

  (c) $\frac{25}{4}$

  (d) $\frac{9}{2}$

  (e) None of the above.

Solution: The correct answer is (c). Using the reward $g \equiv 1$, the vector v of expected values of τ_C, where each entry corresponds to a different transient initial state is

$$v = Fg = \begin{bmatrix} 3 & 4 \\ \frac{3}{2} & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ \frac{11}{2} \end{bmatrix}$$

Finally, the initial distribution puts equal probabilities on the two transient states, so, by the law of total probability, $\mathbb{E}[\tau_C] = \tfrac{1}{2} \times 7 + \tfrac{1}{2} \times \tfrac{11}{2} = \tfrac{25}{4}$.

Problem 14.42. A fair 6-sided die is rolled repeatedly, and for $n \in \mathbb{N}$, the outcome of the n-th roll is denoted by Y_n (it is assumed that $\{Y_n\}_{n\in\mathbb{N}}$ are independent of each other). For $n \in \mathbb{N}_0$, let X_n be the remainder (taken in the set $\{0,1,2,3,4\}$) left after the sum $\sum_{k=1}^n Y_k$ is divided by 5, i.e. $X_0 = 0$, and

$$X_n = \sum_{k=1}^n Y_k \ (\mathrm{mod}\,5), \ \text{for } n \in \mathbb{N},$$

making $\{X_n\}_{n\in\mathbb{N}_0}$ a Markov chain on the state space $\{0,1,2,3,4\}$ (no need to prove this fact).

  1. Write down the transition matrix of the chain.

  2. Classify the states, separate recurrent from transient ones, and compute the period of each state.

3. Compute the expected number of rolls before the first time $\{X_n\}_{n\in\mathbb{N}_0}$ visits the state 2, i.e., compute $\mathbb{E}[\tau_2]$, where

$$\tau_2 = \min\{n \in \mathbb{N}_0 \ : \ X_n = 2\}.$$

4. Compute $\mathbb{E}[\sum_{k=0}^{\tau_2} X_k]$.

(*Note:* For parts (3) and (4), you can phrase your answer as $(B^{-1}C)_{ij}$, where the matrices B and C, as well as the row i and the column j have to be given explicitly. You don't need to evaluate B^{-1}.)

Solution:

1. The outcomes $1, 2, 3, 4, 5, 6$ leave remainders $1, 2, 3, 4, 0, 1$, when divided by 5, so the transition matrix P of the chain is given by

$$P = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

2. Since $p_{ij} > 0$ for all $i, j \in \mathcal{S}$, all the states belong to the same class, and, because there is at least one recurrent state in a finite-state-space Markov chain and because recurrence is a class property, all states are recurrent. Finally, 1 is in the return set of every state, so the period of each state is 1.

3. In order to be able to use "absorption-and-reward" computations, we turn the state 2 into an absorbing state, i.e., modify the transition matrix as follows

$$P' = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 1 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Now, the first hitting time of the state 2 corresponds exactly to the absorption time, i.e., the time τ_C the chain hits its first recurrent state (note that for the new transition probabilities the state 2 is recurrent and all the other states are transient). In the canonical form, the new transition matrix looks like

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Therefore,

$$Q = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}, \text{ and } F = (I-Q)^{-1} = \begin{bmatrix} \frac{5}{6} & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{bmatrix}^{-1}$$

The expected time until absorption is given by an "absorption-and-reward" computation with the reward matrix

$$g_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

so that, since we are starting from the state 0, we get

$$\mathbb{E}[\tau_2] = (Fg_1)_1.$$

4. If the sum went from 0 to $\tau_2 - 1$ (up to, but not including the state where we get absorbed), it would correspond exactly to an "absorption-and-reward" computation with reward g_2 given by

$$g_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}.$$

Therefore, $\mathbb{E}[\sum_{k=0}^{\tau_2-1} X_k] = (Fg_2)_1$. To get the final answer, we need to add the value X_{τ_2} (which is clearly equal to 2) to the sum:

$$\mathbb{E}[\sum_{k=0}^{\tau_2} X_k] = (Fg_2)_1 + 2.$$

Problem 14.43. In a "Gambler's ruin" situation with $a = 3$ ($\mathcal{S} = \{0,1,2,3\}$) and $p = \frac{1}{2}$ compute the probability that the gambler will go bankrupt before her wealth reaches a, if her initial wealth is $X_0 = 2$. You must use the matrix technique (the one in the "Absorption and Reward" lecture with U, Q, etc.) Remember that for $a, b, c, d \in \mathbb{R}$ with $ad - bc \neq 0$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Solution: The states 0 and 3 are absorbing, and all the others are transient. Therefore $C_1 = \{0\}$, $C_2 = \{3\}$ and $T = T_1 = \{1,2\}$. The transition matrix P in the canonical form (the rows and columns represent the states in the order $0, 3, 1, 2$)

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \text{ so that } Q = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \ R = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

The matrix $I - Q$ is a 2×2 matrix so it is easy to invert:

$$(I - Q)^{-1} = \frac{4}{3} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}.$$

So

$$U = (I - Q)^{-1} R = \frac{4}{3} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \frac{4}{3} \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}.$$

Therefore, for the initial "wealth" 2, the probability of going bankrupt before getting rich is $\frac{1}{3}$.

Problem 14.44. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a Markov chain with the following transition matrix

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

Suppose that the chains starts from the state 1.

1. What is expected time that will pass before the chain first hits 3?

2. What is the expected number of visits to state 2 before 3 is hit?

3. Would your answers to (1) and (2) change if we replaced values in the first row of P by any other values (as long as P remains a stochastic matrix)? Would 1 and 2 still be transient states?

4. Use the idea of part (3) to answer the following question. What is the expected number of visits to the state 2 before a Markov chain with transition matrix

$$P = \begin{bmatrix} 17/20 & 1/20 & 1/10 \\ 1/15 & 13/15 & 1/15 \\ 2/5 & 4/15 & 1/3 \end{bmatrix}$$

hits the state 3 for the first time (the initial state is still 1)? Remember this trick for the final.

Solution: The states 1 and 2 are transient and 3 is recurrent, so the canonical decomposition is $\{3\} \cup \{1, 2\}$ and the canonical form of the transition matrix is

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

The matrices Q and R are given by

$$Q = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 1/3 \end{bmatrix}, R = \begin{bmatrix} 0 \\ 1/3 \end{bmatrix},$$

and the fundamental matrix $F = (I - Q)^{-1}$ is

$$F = \begin{bmatrix} 4 & 3 \\ 2 & 3 \end{bmatrix}.$$

1. The reward function $g(1) = g(2) = 1$, i.e. $g = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ will give us the expected time until absorption:

$$v = Fg = \begin{bmatrix} 4 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}.$$

Since the initial state is $i = 1$, the expected time before we first hit 3 is $v_1 = 7$.

2. Here we use the reward function

$$g(i) = \begin{cases} 0, & i = 1, \\ 1, & i = 2. \end{cases}$$

to get

$$v = Fg = \begin{bmatrix} 4 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$

so the answer is $v_1 = 3$.

3. The way the question is posed, the answer is "Maybe". What I intended to ask, though, is the same question about the *third* row. In that case, the answer is "No, they would not". Indeed, these numbers only affect what the chain does after it hits 3 for the first time, and that is irrelevant for calculations about the events which happen prior to that.

   No, all states would be recurrent. The moral of the story that the absorption calculations can be used even in the settings where all states are recurrent. You simply need to adjust the probabilities, as shown in the following part of the problem.

4. We make the state 3 absorbing (and states 1 and 2 transient) by replacing the transition matrix by the following one:

$$P = \begin{bmatrix} 17/20 & 1/20 & 1/10 \\ 1/15 & 13/15 & 1/15 \\ 0 & 0 & 1 \end{bmatrix}$$

The new chain behaves exactly like the old one until it hits 3 for the first time. Now we find the canonical decomposition and compute the matrix F as above and get

$$F = \begin{bmatrix} 8 & 3 \\ 4 & 9 \end{bmatrix},$$

so that the expected number of visits to 2 (with initial state 1) is $F_{12} = 3$.

Problem 14.45. A math professor has 4 umbrellas. He keeps some of them at home and some in the office. Every morning, when he leaves home, he checks the weather and takes an umbrella with him if it rains. In case all the umbrellas are in the office, he gets wet. The same procedure is repeated in the afternoon when he leaves the office to go home. The professor lives in a tropical region, so the chance of rain in the afternoon is higher than in the morning; it is 1/5 in the afternoon and 1/20 in the morning. Whether it rains of not is independent of whether it rained the last time he checked.

On day 0, there are 2 umbrellas at home, and 2 in the office. What is the expected number of days that will pass before the professor gets wet (remember, there are two trips each day)? What is the probability that the first time he gets wet it is on his way home from the office?

Solution: We model the situation by a Markov chain whose state space $\mathcal{S}$ is given by

$$\mathcal{S} = \{(p, u) \; : \; p \in \{H, O\}, u \in \{0, 1, 2, 3, 4, w\}\},$$

where the first coordinate denoted the current position of the professor and the second the number of umbrellas at home (then we automatically know how many umbrellas there are at the office). The second coordinate w stands for "wet" and the state (H, w) means that the professor left home without an umbrella during a rain (got wet). The transitions between the states are simple to figure out. For example, from the state $(H, 2)$ we either move to $(O, 2)$ (with probability 19/20) or to $(O, 1)$ (with probability 1/20), and from $(O, 4)$ we move to (O, w) with probability 1/5 and to $(H, 4)$ with probability 4/5. States (H, w) and (O, w) are made absorbing, and so, all the other states are transient. The first question can be reformulated as a reward problem with reward $g \equiv 1$, and the second one is about absorption. We use *Mathematica* to solve it, and the code is on the right. The answers are: the probability of getting wet for the first time on the way home from the office is about 99%. The expected number of days before getting wet is about 20.

```
ph = 1 / 20; po = 1 / 5;
Transitions = {
    {{"W", "O"}, {"W", "O"}, 1}, {{"W", "H"}, {"W", "H"}, 1},
    {{4, "H"}, {4, "O"}, 1 - ph}, {{4, "H"}, {3, "O"}, ph},
    {{3, "H"}, {3, "O"}, 1 - ph}, {{3, "H"}, {2, "O"}, ph},
    {{2, "H"}, {2, "O"}, 1 - ph}, {{2, "H"}, {1, "O"}, ph},
    {{1, "H"}, {1, "O"}, 1 - ph}, {{1, "H"}, {0, "O"}, ph},
    {{0, "H"}, {0, "O"}, 1 - ph}, {{0, "H"}, {"W", "H"}, ph},
    {{0, "O"}, {0, "H"}, 1 - po}, {{0, "O"}, {1, "H"}, po},
    {{1, "O"}, {1, "H"}, 1 - po}, {{1, "O"}, {2, "H"}, po},
    {{2, "O"}, {2, "H"}, 1 - po}, {{2, "O"}, {3, "H"}, po},
    {{3, "O"}, {3, "H"}, 1 - po}, {{3, "O"}, {4, "H"}, po},
    {{4, "O"}, {4, "H"}, 1 - po}, {{4, "O"}, {"W", "O"}, po}};
Initial = {{{2, "H"}, 1}};
Rain = BuildChain[Transitions, Initial];
P = TransitionMatrix[Rain];

States[Rain]

{{0, H}, {0, O}, {1, H}, {1, O}, {2, H},
 {2, O}, {3, H}, {3, O}, {4, H}, {4, O}, {W, H}, {W, O}}

Q = P[[1 ;; 10, 1 ;; 10]];
F = Inverse[IdentityMatrix[10] - Q];
R = P[[1 ;; 10, 11 ;; 12]];
G = Table[{1}, {i, 1, 10}];
InitialState = StateToPosition[{2, "H"}, Rain];

AbsorptionProbabilities = N[(F.R)[[InitialState]]]

{0.0109781, 0.989022}

ExpectedTime = N[(F.G)[[InitialState]]] / 2

{19.57}
```

Problem 14.46. A zoologist, Dr. Gurkensaft, claims to have trained Basil the Rat so that it can aviod being shocked and find food, even in highly confusing situations. Another scientist, Dr. Hasenpfeffer does not agree. She says that Basil is stupid and cannot tell the difference between food and an electrical shocker until it gets very close to either of them.

The two decide to see who is right by performing the following experiment. Basil is put in the compartment 3 of a maze that looks like this:

.
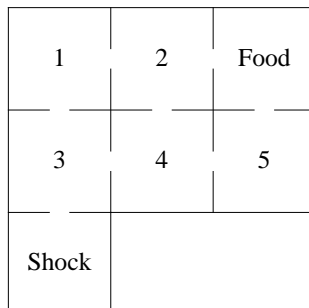
Dr. Gurkensaft's hypothsis is that, once in a compartment with k exits, Basil will prefer the exits

that lead him closer to the food. Dr. Hasenpfeffer's claim is that every time there are k exits from a compartment, Basil chooses each one with probability $1/k$.

After repeating the experiment 100 times, Basil got shocked before getting to food 52 times and he reached food before being shocked 48 times. Compute the theoretical probabilities of being shocked before getting to food, under the assumption that Basil is stupid (all exits are equally likely). Compare those to the observed data. Whose side is the evidence on?

Solution: The task here is to calculate the absorption probability for a Markov chain which represents the maze. Using *Mathematica*, we can do it in the following way:

```
In[30]:= Transitions = {
         {1, 2, 1 / 2}, {1, 3, 1 / 2},
         {2, 1, 1 / 3}, {2, "Food", 1 / 3}, {2, 4, 1 / 3},
         {3, 1, 1 / 3}, {3, 4, 1 / 3}, {3, "Shock", 1 / 3},
         {4, 3, 1 / 3}, {4, 2, 1 / 3}, {4, 5, 1 / 3},
         {5, 4, 1 / 2}, {5, "Food", 1 / 2}};
      Initial = {{3, 1}};
      Maze = BuildChain[Transitions, Initial];


In[33]:= States[Maze]

Out[33]= {1, 2, 3, 4, 5, Food, Shock}

In[34]:= P = TransitionMatrix[Maze];
      Q = P[[1 ;; 5, 1 ;; 5]];
      R = P[[1 ;; 5, 6 ;; 7]];
      F = Inverse[IdentityMatrix[5] - Q];

In[38]:= N[(F.R)[[3]]]

Out[38]= {0.416667, 0.583333}
```

Therefore, the probability of getting to food before being shocked is about 42%. This is somewhat lower than the observed 48% and even though there may be some truth to Dr. Gurkensaft's claims, these two numbers are not very different. *Note:* For those of you who know a bit of statistics, you can easily show that we cannot reject the null hypothesis that Basil is stupid (in the precise sense described in the problem), even at the 90% significance level. In fact, the (one-sided) p-value (using a binomial test) is a bir larger than 0.1. That means that a truly stupid rat would appear smarter than Basil 10% of the time by pure chance.

14.8 Markov chains - stationary and limiting distributions

Problem 14.47. Let $\{X_n\}_{n\in\mathbb{N}_0}$ be a Markov chain with the transition matrix

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

1. Find all stationary distributions.

2. The chain starts from the state $i = 1$. What is the expected number of steps before it returns to 1?

3. How many times, on average, does the chain visit 2 between two consecutive visits to 1?

4. Each time the chain visits the state 1, \$1 is added to an account, \$2 for the state 2, and nothing in the state 3. Estimate the amount of money on the account after 10000 transitions? You may assume that the law of large numbers for the Markov chains provides an adequate approximation.

Solution:

1. Stationary distributions $\pi = (\pi_1, \pi_2, \pi_3)$ satisfy $\pi P = \pi$, i.e.,

$$\begin{aligned}
\tfrac{1}{4}\pi_1 \qquad\quad + \tfrac{1}{3}\pi_3 &= \pi_1 \\
\tfrac{1}{4}\pi_1 + \tfrac{1}{3}\pi_2 + \tfrac{1}{3}\pi_3 &= \pi_2 \\
\tfrac{1}{2}\pi_1 + \tfrac{2}{3}\pi_2 + \tfrac{1}{3}\pi_3 &= \pi_3.
\end{aligned}$$

We also know that $\pi_1 + \pi_2 + \pi_3 = 1$, and that the matrix P is stochastic. Therefore, the third equation below is a linear combination of the first two, and can be exculded from consideration (this is always the case in problems with stationary distributions).

The first equation yields that $\pi_3 = \tfrac{9}{4}\pi_1$, and the second one that $\pi_2 = \tfrac{3}{2}(\tfrac{1}{4}\pi_1 + \tfrac{1}{3}\pi_3) = \tfrac{3}{2}\pi_1$. It remains to find π_1 such that $\pi_1 + \pi_2 + \pi_3 = 1$, i.e $\pi_1 + \tfrac{3}{2}\pi_1 + \tfrac{9}{4}\pi_1 = 1$, i.e., $\pi_1 = (1 + \tfrac{3}{2} + \tfrac{9}{4})^{-1} = \tfrac{4}{19}$. Therefore,

$$\pi = (\tfrac{4}{19}, \tfrac{6}{19}, \tfrac{9}{19})$$

is the only stationary distribution.

2. The number of steps between two returns to a state i (in an irreducible finite chain) is given by

$$\mathbb{E}_i[\tau_i(1)] = \frac{1}{\pi_i}.$$

Therefore, $\mathbb{E}_1[\tau_1(1)] = \tfrac{19}{4}$.

3. The number of visits to the state j between two consecutive visits to the state i (in an irreducible finite chain) is given by

$$\mathbb{E}_i\Big[\sum_{n=0}^{\tau_1(1)} \mathbf{1}_{\{X_n = j\}}\Big] = \tfrac{\pi_j}{\pi_i}.$$

Therefore, our chain will visit the state 2 on average 1.5 times between two visits to the state 1.

4. The chain in question is irreducible and finite, so the law of large numbers applies:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(X_n) = \sum_{i \in \S} f(i)\pi_i.$$

In our case $f(1) = 1$, $f(2) = 2$ and $f(3) = 0$, so the amount of money $M = \sum_{n=0}^{10000} f(X_n)$ can be approximated as

$$M = 10001 \times \tfrac{1}{10001} \sum_{n=0}^{10000} f(X_n) \approx 10001 \times (\$1\tfrac{4}{19} + \$2\tfrac{6}{19}) = 10001 \times \$\tfrac{16}{19} \approx \$8422.$$

Problem 14.48. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a Markov chain whose transition graph is given below:
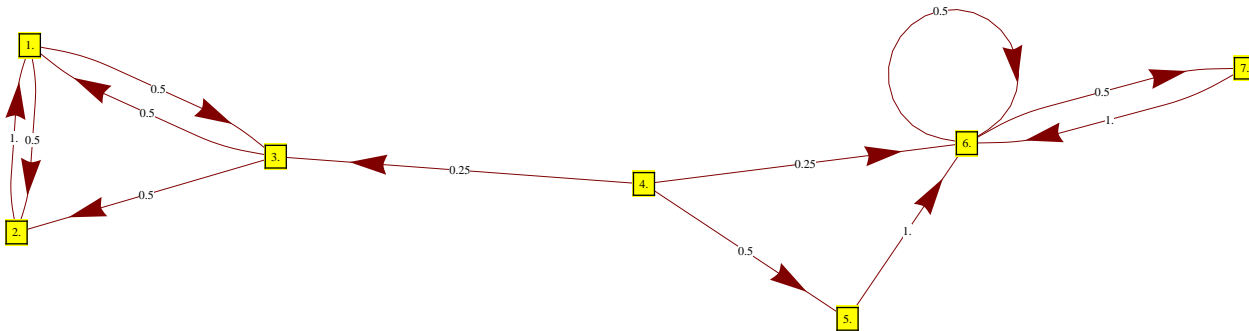
1. Find all stationary distributions.

2. Is there a limiting distribution?

3. If the chain starts from the state 5, what is the asymptotic proportion of time it will spend at state 7? How about the state 5?

Solution:

1. The classes are $C_1 = \{1, 2, 3\}$, $C_2 = 6, 7$, $T_1 = \{4\}$ and $T_2 = \{5\}$. Any stationary distribution π will have $\pi_4 = \pi_5 = 0$ (since 4 and 5 are transient states). The next step is to find (unique) stationary distributions for C_1 and C_2 (viewed as Markov chains in themselves). The transition matrices P_{C_1} and P_{C_2} of C_1 and C_2 are given by

$$P_{C_1} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1. & 0 & 0 \\ 0.5 & 0.5 & 0 \end{bmatrix}, \quad P_{C_2} = \begin{bmatrix} 0.5 & 0.5 \\ 1. & 0 \end{bmatrix}$$

The unique distributions π^{C_1}, π^{C_2} which solve the equations $\pi^{C_1} = \pi^{C_1} P_{C_1}$ and $\pi^{C_2} = \pi^{C_2} P_{C_2}$ respectively are

$$\pi^{C_1} = (\tfrac{4}{9}, \tfrac{1}{3}, \tfrac{2}{9}), \quad \pi^{C_2} = (\tfrac{2}{3}, \tfrac{1}{3}).$$

Therefore, the stationary distributions for P are exactly those of the form

$$\pi = (\alpha_1 \tfrac{4}{9}, \alpha_1 \tfrac{1}{3}, \alpha_1 \tfrac{2}{9}, 0, 0, \alpha_2 \tfrac{2}{3}, \alpha_2 \tfrac{1}{3}),$$

where $\alpha_1, \alpha_2 \geq 0$ satisfy $\alpha_1 + \alpha_2 = 1$.

2. No limiting distribution exists because of the dependence on the initial conditions. To make the argument a bit more rigorous we suppose, to the contrary, that a limiting distribution $\pi = (\pi_1, \ldots, \pi_7)$ exists, i.e.,

$$\pi_j = \lim_{n \to \infty} p_{ij}^{(n)}, \text{ for all } i, j = 1, 2, \ldots, 7.$$

When we start in C_1, the chain will stay there, so $p_{11}^{(n)} + p_{12}^{(n)} + p_{13}^{(n)} = 1$, for all n, and, consequently,

$$\pi_1 + \pi_2 + \pi_3 = \lim_{n \to \infty} (p_{11}^{(n)} + p_{12}^{(n)} + p_{13}^{(n)}) = 1.$$

Similarly, if we start in C_2, we stay in C_2, so $\pi_6 + \pi_7 = 1$, by the same argument. This is a contradiction beacause $\pi_1 + \pi_2 + \cdots + \pi_7$ must be equal to 1, but we have proven that

$$\pi_1 + \pi_2 + \cdots + \pi_7 \geq \pi_1 + \pi_2 + \pi_3 + \pi_6 + \pi_7 = 2.$$

3. When the chain starts from the state 5, it will be in 6 in the next step, and will stay in C_2 forever. The fact that the chain spent one unit of time in the state 5 will be negligible as $n \to \infty$. Therefore, we need to compute the asymptotic proportion of time that the chain with transition matrix P_{C_2} will spend in 7. This is given by $\pi_7^{C_2} = 1/3$. As for the state 5 itself, the chain will spend only one unit of time there, so, the proportion of time spent there will converge to 0.

Problem 14.49. An airline reservation system has two computers, only one of which is in operation at any given time. A computer may break down on any given day with probability $p \in (0,1)$. There is a single repair facility which takes two days to restore a computer to normal. The facilities are such that only one computer at a time can be dealt with. Form a Markov chain that models the situation (make sure to keep track of the number of machines in operation as well as the status of the machine - if there is one - at the repair facility).

1. Find the stationary distribution.

2. What percentage of time (on average) are no machines operable?

3. What percentage of time (on average) is exactly one machine operable?

Solution:

1. Any of the computers can be in the following 4 conditions: *in operation, in repair facility - 2nd day, in repair facility - 1st day, waiting to enter the repair facility*. Since there are two computers, each state of the Markov chain will be a quadruple of numbers denoting the number of computers in each conditions. For example, $(1,0,1,0)$ means that there is one computer in operation and one which is spending its first day in the repair facility. If there are no computers in operation, the chain moves deterministically, but if 1 or 2 computers are in operation, they break down with probability p each, independently of the other. For example, if there are two computers in operation (the state of the system being $(2,0,0,0)$), there are 3 possible scenarios: both computers remain in operation (that happens with probability $(1-p)^2$), exactly one computer breaks down (the probability of that is $2p(1-p)$) and both computers break down (with probability p^2). In the first case, the chain stays in

the state $(2,0,0,0)$. In the second case, the chain moves to the state $(1,0,1,0)$ and in the third one one of the computers enters the repair facility, while the other spends a day waiting, which corresponds to the state $(0,0,1,1)$. All in all, here is the transition graph of the chain:
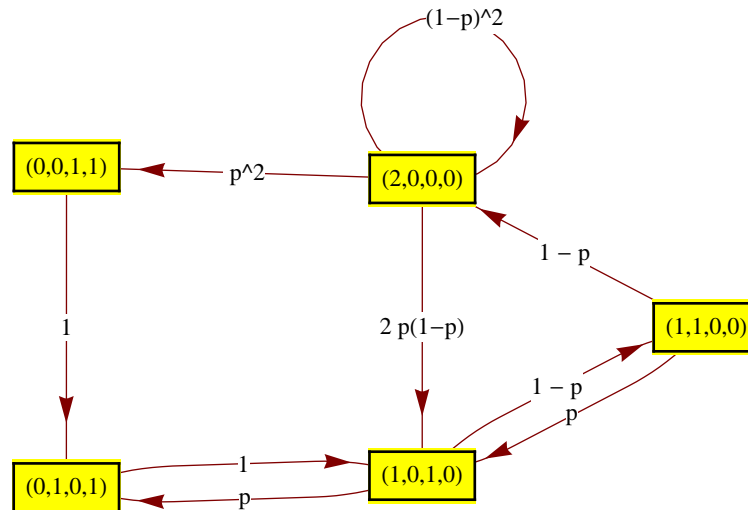
One can use Mathematica (or work by hand) to obtain the unique stationary distribution

$$\pi = \left(\frac{(1-p)^2 p^2}{2p^4-4p^3+p^2+2p+1}, \frac{p^4-3p^3+3p^2}{2p^4-4p^3+p^2+2p+1}, \frac{2p-p^2}{2p^4-4p^3+p^2+2p+1}, \frac{(1-p)(2p-p^2)}{2p^4-4p^3+p^2+2p+1}, \frac{(1-p)^2}{2p^4-4p^3+p^2+2p+1} \right)$$

where the states are ordered as follows $\mathcal{S} = \{(0,0,1,1), (0,1,0,1), (1,0,1,0), (1,1,0,0), (2,0,0,0)\}$.

2. No machines are operable if the system is in $(0,0,1,1)$ or in $(0,1,0,1)$. Therefore, the asymptotic precentage of time no machines are operable is

$$\pi_1 + \pi_2 = \frac{(p-1)^2 p^2 + p^4 - 2p^3 + 3p^2}{2p^4-4p^3+p^2+2p+1}.$$

3. Exactly one machine is operable in states $(1,0,1,0)$ and $(1,1,0,0)$, so the required asymptotic precentage is

$$\pi_3 + \pi_4 = \frac{(p-2)^2 p}{2p^4-4p^3+p^2+2p+1}.$$

Here is the *Mathematica* code I used:

```
In[51]:= $Assumptions = (p > 0) && (1 - p > 0);
         (** just to simplify some internal computations. Ignore **)

In[52]:= Transitions = {
           {"(2,0,0,0)", "(2,0,0,0)", 1 + p^2 - 2 p},
           {"(2,0,0,0)", "(1,0,1,0)", 2 p - 2 p^2},
           {"(2,0,0,0)", "(0,0,1,1)", p^2},
           {"(1,0,1,0)", "(0,1,0,1)", p},
           {"(1,0,1,0)", "(1,1,0,0)", 1 - p},
           {"(1,1,0,0)", "(1,0,1,0)", p},
           {"(1,1,0,0)", "(2,0,0,0)", 1 - p},
           {"(0,1,0,1)", "(1,0,1,0)", 1},
           {"(0,0,1,1)", "(0,1,0,1)", 1}
         };
         Initial = {{"(2,0,0,0)", 1}};
         G = BuildChain[Transitions, Initial];
         P = TransitionMatrix[G];

In[56]:= v = NullSpace[Transpose[P - IdentityMatrix[5]]][[1]];
         pi = v / Simplify[Total[v]]
```

$$Out[57]= \left\{ \frac{(-1+p)^2 \, p^2}{1 + 2 p + p^2 - 4 p^3 + 2 p^4} , \ -\frac{-3 p^2 + 3 p^3 - p^4}{1 + 2 p + p^2 - 4 p^3 + 2 p^4} , \right.$$

$$\left. -\frac{-2 p + p^2}{1 + 2 p + p^2 - 4 p^3 + 2 p^4} , \ -\frac{(-1+p)\left(2 p - p^2\right)}{1 + 2 p + p^2 - 4 p^3 + 2 p^4} , \ \frac{(-1+p)^2}{1 + 2 p + p^2 - 4 p^3 + 2 p^4} \right\}$$

Problem 14.50. Let $\{Y_n\}_{n\in\mathbb{N}_0}$ be a sequence of die-rolls, i.e., a sequence of independent random variables with distribution

$$Y_n \sim \left(\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{array} \right).$$

Let $\{X_n\}_{n\in\mathbb{N}_0}$ be a stochastic process defined by $X_n = \max(Y_0, Y_1, \ldots, Y_n)$. In words, X_n is the maximal value rolled so far.

1. Explain why X is a Markov chain, and find its transition matrix and the initial distribution.

2. Supposing that the first roll of the die was 3, i.e., $X_0 = 3$, what is the expected time until a 6 is reached?

3. Under the same assumption as above ($X_0 = 3$), what is the probability that a 5 will not be rolled before a 6 is rolled for the first time?

4. Starting with the first value $X_0 = 3$, each time a die is rolled, the current record (the value of X_n) is written down. When a 6 is rolled for the first time all the numbers are added up and the result is called S (the final 6 is not counted). What is the expected value of S?

(*Hints:*

1. You can eliminate 1 and 2 from the state space for parts (2) and (3).)

2. Use the fact that

$$A^{-1} = \begin{bmatrix} \frac{1}{a} & \frac{r}{ab} & \frac{rb+r^2}{abc} \\ 0 & \frac{1}{b} & \frac{r}{bc} \\ 0 & 0 & \frac{1}{c} \end{bmatrix} \text{ when } A = \begin{bmatrix} a & -r & -r \\ 0 & b & -r \\ 0 & 0 & c \end{bmatrix}, \text{ for } a, b, c, r > 0.$$

Solution:

1. The value of X_{n+1} is either going to be equal to X_n if Y_{n+1} happens to be less than or equal to it, or it moves up to Y_{n+1}, otherwise, i.e., $X_{n+1} = \max(X_n, Y_{n+1})$. Therefore, the distribution of X_{n+1} depends on the previous values $X_0, X_1, \ldots, X_n$ only through X_n, and, so, $\{X_n\}_{n\in\mathbb{N}_0}$ is a Markov chain on the state space $\mathcal{S} = \{1,2,3,4,5,6\}$. The transition Matrix is given by

$$P = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 1/3 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 1/2 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 2/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 1/6 \end{bmatrix}$$

2. As stated in the *Hint*, we can restrict our attention to the Markov chain on the state space $\mathcal{S} = \{3,4,5,6\}$ with the transition matrix

$$P = \begin{bmatrix} 1/2 & 1/6 & 1/6 & 1/6 \\ 0 & 2/3 & 1/6 & 1/6 \\ 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 1/6 \end{bmatrix}$$

The state 6 is absorbing (and therefore recurrent), while all the others are transient, so the canonical form of the matrix (with the canonical decomposition $\mathcal{S} = \{6\} \cup \{3,4,5\}$ is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/6 & 1/2 & 1/6 & 1/6 \\ 1/6 & 0 & 2/3 & 1/6 \\ 1/6 & 0 & 0 & 5/6 \end{bmatrix}$$

The fundamental matrix $F = (I - Q)^{-1}$ is now given by

$$F = \begin{bmatrix} 1/2 & -1/6 & -1/6 \\ 0 & 1/3 & -1/6 \\ 0 & 0 & 1/6 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 6 \end{bmatrix}$$

The expected time to absorption can be computed by using the reward 1 for each state:

$$F \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$

Therefore it will take on average 6 rolls until the first 6.

Note: You don't need Markov chains to solve this part; the expected time to absorption is nothing but the waiting time until the first 6 is rolled - a shifted geometrically distributed random variable with parameter 1/6. Therefore, its expectation is $1 + \frac{5/6}{1/6} = 6$.

3. We make the state 5 absorbing and the canonical decomposition changes to $\mathcal{S} = \{5\} \cup \{6\} \cup \{3,4\}$. The canonical form of the transition matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/6 & 1/6 & 1/2 & 1/6 \\ 1/6 & 1/6 & 0 & 2/3 \end{bmatrix}$$

and so

$$Q = \begin{bmatrix} 1/2 & 1/6 \\ 0 & 2/3 \end{bmatrix} \text{ and } R = \begin{bmatrix} 1/6 & 1/6 \\ 1/6 & 1/6 \end{bmatrix}.$$

Therefore, the absorption probabilities are the entries of the matrix

$$(I - Q)^{-1}R = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix},$$

and the answer is $1/2$.

Note: You don't need Markov chains to solve this part either; the probability that 6 will be seen before 5 is the same as the probability that 5 will appear before 6. The situation is entirely symmetric. Therefore, the answer must be $1/2$.

4. This problem is about expected reward, where the reward $g(i)$ of the state $i \in \{3,4,5\}$ is $g(i) = i$. The answer is v_3, where

$$v = F \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 25 \\ 27 \\ 30 \end{bmatrix},$$

i.e., the expected value of the sum is 25.

Note: Can you do this without the use of the Markov-chain theory?

Problem 14.51. A car-insurance company classifies drivers in three categories: *bad*, *neutral* and *good*. The reclassification is done in January of each year and the probabilities for transitions between different categories is given by

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/5 & 2/5 & 2/5 \\ 1/5 & 1/5 & 3/5 \end{bmatrix},$$

where the first row/column corresponds to the *bad* category, the second to *neutral* and the third to *good*.

1. The company started in January 1990 with 1400 drivers in each category. Estimate the number of drivers in each category in 2090. Assume that the total number of drivers does not change in time.

2. A yearly premium charged to a driver depends on his/her category; it is $1000 for *bad* drivers, $500 for *neutral* drivers and $200 for *good* drivers. A typical driver has a total of 56 years of driving experience. Estimate the total amount he/she will pay in insurance premiums in those 56 years. You may assume that 56 years is long enough for the limiting theorems to provide an accurate approximation.

Solution:

1. Equal numbers of drivers in each category corresponds to the uniform initial distribution, $\boldsymbol{a}^{(0)} = (1/3, 1/3, 1/3)$. The distribution of drivers in 2090 is given by the distribution of $\boldsymbol{a}^{(100)}$ of X_{100} given by

$$\boldsymbol{a}^{(100)} = \boldsymbol{a}^{(0)} P^{100},$$

which can be approximated by the limiting distribution of the underlying Markov chain. The limiting distribution is always stationary (and it exists in the present chase thanks to irreducibility and aperiodicity). The stationary distribution solves the linear system $\pi = \pi P$; we easily get

$$\pi = (\frac{4}{14}, \frac{5}{14}, \frac{5}{14}).$$

Therefore, the total of 4200 drivers will be distributed as $(\frac{4}{14} \times 4200, \frac{5}{14} \times 4200, \frac{5}{14} \times 4200) = (1200, 1500, 1500)$.

2. Let the function $f : \mathcal{S} \to \mathbb{R}$ be given by $f(1) = 1000$, $f(2) = 500$ and $f(3) = 200$. The total amount paid in insurance premiums is $\sum_{k=0}^{54} f(X_k)$. By the Law of large numbers for Markov chains, we have

$$\sum_{k=0}^{55} f(X_k) = 56 \times \frac{\sum_{k=0}^{55} f(X_k)}{56} \approx 56 \times (\pi_1 f(1) + \pi_2 f(2) + \pi_3 f(3)) = 30000.$$

14.9 Markov chains - various multiple-choice problems

In each of the following multiple-choice problems choose the answer that *necessarily follows from the information given*. There will be one and only one correct choice.

Problem 14.52. Let C be a class in a Markov chain. Then

(a) C is closed,

(b) C^c is closed,

(c) at least one state in C is recurrent,

(d) for all states $i, j \in C$, $p_{ij} > 0$,

(e) none of the above.

Solution: The answer is (**e**).

(a) False. Take $C = \{(0,0)\}$ in the "Tennis example".

(b) False. Take $C = \{\text{Amélie wins}\}$ in the "Tennis example".

(c) False. This is only true in *finite* chains. Take the Deterministically Monotone Markov Chain as an example. All of its states are transient.

(d) False. This would be true if it read "for each pair of states $i, j \in C$, *there exists $n \in \mathbb{N}$ such that $p_{ij}^{(n)} > 0$*". Otherwise, we can use the "Tennis chain" and the states $i = (40, Adv)$ and $j = (Adv, 40)$. They belong to the same class, but $p_{ij} = 0$ (you need to pass through $(40, 40)$ to go from one to another).

(e) True.

Problem 14.53. In a Markov chain whose state space has n elements ($n \in \mathbb{N}$),

(a) all classes are closed

(b) at least one state is transient,

(c) not more than half of all states are transient,

(d) there are at most n classes,

(e) none of the above.

Solution: The answer is (**d**).

(a) False. In the "Tennis" example, there are classes that are not closed.

(b) False. Just take the Regime Switching with $0 < p_{01}, p_{10} < 1$. Both of the states are recurrent there. Or, simply take a Markov chain with only one state ($n = 1$).

(c) False. In the "Tennis" example, 18 states are transient, but $n = 20$.

(d) True. Classes form a partition of the state space, and each class has at least one element. Therefore, there are at most n classes.

(e) False.

Problem 14.54. Let C_1 and C_2 be two (different) classes. Then

(a) $i \to j$ or $j \to i$, for all $i \in C_1$, and $j \in C_2$,

(b) $C_1 \cup C_2$ is a class,

(c) if $i \to j$ for some $i \in C_1$ and $j \in C_2$, then $k \not\to l$ for all $k \in C_2$ and $l \in C_1$,

(d) if $i \to j$ for some $i \in C_1$ and $j \in C_2$, then $k \to l$ for some $k \in C_2$ and $l \in C_1$,

(e) none of the above.

Solution: The answer is (**c**).

(a) False. Take $C_1 = \{$Amélie wins$\}$ and $C_2 = \{$Björn wins$\}$.

(b) False. Take the same example as above.

(c) True. Suppose, to the contrary, that $k \to l$ for some $k \in C_2$ and $l \in C_1$. Then, since j and k are in the same class, we must have $j \to k$. Similarly, $l \to i$ (l and i are in the same class). Using the transitivity property of the communication relation, we get $j \to k \to l \to i$, and so $j \to i$. By the assumption $i \to j$, and so, $i \leftrightarrow j$. This is a contradiction, however, with the assumption that i and j are in different classes.

(d) False. In the "Tennis" example, take $i = (0,0)$, $j = (30,0)$, $C_1 = \{i\}$ and $C_2 = \{j\}$.

(e) False.

Problem 14.55. Let i be a recurrent state with period 5, and let j be another state. Then

(a) if $j \to i$, then j is recurrent,

(b) if $j \to i$, then j has period 5,

(c) if $i \to j$, then j has period 5,

(d) if $j \not\to i$ then j is transient,

(e) none of the above.

Solution: The answer is (c).
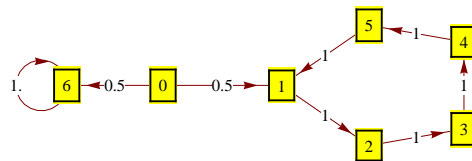
(a) False. Take $j = 0$ and $i = 1$ in the chain in the picture.

(b) False. Take the same counterexample as above.

(c) True. We know that i is recurrent, and since all recurrent classes are closed, and $i \to j$, h must belong to the same class as i. Period is a class property, so the period of j is also 5.

(d) False. Take $j = 6$, $i = 0$ in the chain in the picture.

(e) False.

Problem 14.56. Let i and j be two states such that i is transient and $i \leftrightarrow j$. Then

(a) $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \sum_{n=1}^{\infty} p_{ii}^{(n)}$,

(b) if $i \to k$, then k is transient,

(c) if $k \to i$, then k is transient,

(d) period of i must be 1,

(e) none of the above.

Solution: The answer is (c).
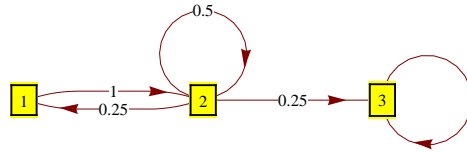
(a) False. In the chain above, take $i = 1$ and $j = 2$. Clearly, $i \leftrightarrow j$, and both of them are transient. A theorem from the notes states that $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$ and $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$, but these two sums do not need to be equal. Indeed, the only way for $i = 1$ to come back to itself in n steps is to move to 2, come back to 2 in $n-2$ steps and then jump to 1. Therefore, $p_{ii}^{(n)} = 1 \times p_{jj}^{(n-2)} \times 0.25$ for $n \geq 3$, $p_{ii}^{(2)} = 1 \times 0.25$ and $p_{ii}^{(1)} = p_{ii} = 0$. Therefore,

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = p_{ii}^{(1)} + p_{ii}^{(2)} + \sum_{n=3}^{\infty} p_{ii}^{(n)} = 0 + 0.25 + 0.25 \sum_{n=1}^{\infty} p_{jj}^{(n)}.$$

This equality implies that the two sums are equal if and only if

$$0.75 \sum_{n=1}^{\infty} p_{ii}^{(n)} = 0.25, \text{ i.e., } \sum_{n=1}^{\infty} p_{jj}^{(n)} = \sum_{n=1}^{\infty} p_{ii}^{(n)} = \frac{1}{3}.$$

We know however, that one of the possible ways to go from 2 to 2 in n steps is to just stay in 2 all the time. The probability of this trajectory is $(1/2)^n$, and so, $p_{jj}^{(n)} \geq (1/2)^n$ for all n. Hence,

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} \geq \sum_{n=1}^{\infty} (1/2)^n = 1.$$

Therefore, the two sums cannot be equal.

(b) False. Take $i = 2$ and $j = 3$ in the chain above.

(c) True. Suppose that $k \to i$, but k is recurrent. Since recurrent classes are closed, i must be in the same class as k. That would mean, however, that i is also recurrent. This is a contradiction with the assumption that i is transient.

(d) False. Take $i = 2$ in the modification of the example above in which $p_{22} = 0$ and $p_{21} = p_{23} = 1/2$. The state 2 is still transient, but its period is 2.

(e) none of the above.

Problem 14.57. Suppose there exists $n \in \mathbb{N}$ such that $P^n = I$, where I is the identity matrix and P is the transition matrix of a finite-state-space Markov chain. Then

(a) $P = I$,

(b) all states belong to the same class,

(c) all states are recurrent

(d) the period of each state is n,

(e) none of the above.

Solution: The answer is **(c)**.

(a) False. Take the Regime-switching chain with

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

Then $P^2 = I$, but $P \neq I$.

(b) False. If $P = I$, all states are absorbing, and, therefore, each is in a class of its own.

(c) True. By the assumption $P^{kn} = (P^n)^k = I^k = I$, for all $k \in \mathbb{N}$. Therefore, $p_{ii}^{(kn)} = 1$ for all $k \in \mathbb{N}$, and so $\lim_{m\to\infty} p_{ii}^{(m)} \neq 0$ (maybe it even doesn't exist). In any case, the series $\sum_{m=1}^{\infty} p_{ii}^{(m)}$ cannot be convergent, and so, i is recurrent, for all $i \in \mathcal{S}$. Alternatively, the condition $P^n = I$ means that the chain will be coming back to where it started - with certainty - every n steps, and so, all states must be recurrent.

(d) False. Any chain satisfying $P^n = I$, but with the property that the n above is not unique is a counterexample. For example, if $P = I$, then $P^n = I$ for any $n \in \mathbb{N}$.

(e) False.

Problem 14.58. Let $\{X_n\}_{n\in\mathbb{N}_0}$ be a Markov chain with state space $\mathbb{Z}$. Then

1. $\{|X_n|\}_{n\in\mathbb{N}_0}$ is a Markov chain,

2. $\{|X_n|\}_{n\in\mathbb{N}_0}$ is not a Markov chain,

3. $\{-3X_n\}_{n\in\mathbb{N}_0}$ is a Markov chain,

4. $\{-3X_n\}_{n\in\mathbb{N}_0}$ is not a Markov chain,

5. none of the above.

Problem 14.59. Let X be a gambler's ruin chain with the upper barier $a > 1$ and $p > \frac{1}{2}$. Then

1. the probability of reaching a before reaching 0 is greater than $\frac{1}{2}$.

2. the probability of (eventual) bankruptcy does not depend on a,

3. the number of recurrent classes depends on a,

4. the limiting distribution exists,

5. none of the above.

Problem 14.60. Let i be a recurrent state with period 3, and let j be another state. Then

1. if $j \to i$, then j is recurrent,

2. if $j \to i$, then j has period 3,

3. if $i \to j$, then j has period 3,

4. if $j \not\to i$ then j is transient,

5. none of the above.

Problem 14.61. Let C_1 and C_2 be two (different) communication classes of a Markov chain. Then

1. $i \to j$ or $j \to i$, for all $i \in C_1$, and $j \in C_2$,

2. $C_1 \cup C_2$ is a class,

3. if $i \to j$ for some $i \in C_1$ and $j \in C_2$, then $k \not\to l$ for all $k \in C_2$ and $l \in C_1$,

4. if $i \to j$ for some $i \in C_1$ and $j \in C_2$, then $k \to l$ for some $k \in C_2$ and $l \in C_1$,

5. none of the above.

Problem 14.62. Two containers are filled with 50 balls each. A container is selected at random (with equal probabilities) and a ball is taken from it and transfered to the other container. If the selected container is empty, nothing happens. The same procedure is then repeated indefinitely. A natural state space on which we model the situation is $\mathcal{S} = \{0, 1, 2, \ldots, 100\}$, where the state $i \in \mathcal{S}$ corresponds to i balls in the first container and $100 - i$ in the second.

1. all states have period 2,

2. $(\frac{1}{2}, \frac{1}{2})$ is a stationary distribution,

3. all states are positive recurrent

4. the number of transient states would change if we started from the state $i = 49$,

5. none of the above.

Problem 14.63. Let $\{X_n\}_{n \in \mathbb{N}_0}$ be a finite-state Markov chain with one transient and two recurrent classes. Then

1. stationary distributions may not exist,

2. the limiting distribution does not exist,

3. the limiting distribution may or may not exist, depending on the particular chain,

4. the number of states is at least 4,

5. none of the above