MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2011)

Problem Set 9: Solutions

1. The outcomes of successive flips can be viewed as a Markov chain with two states, T and H. The transition probabilities are

$$\begin{aligned} p_{TH} &= 1/3 \\ p_{TT} &= 2/3 \\ p_{HH} &= 3/4 \\ p_{HT} &= 1/4. \end{aligned}$$

Let $X_k, k = 1, \ldots$ denote the outcomes of the flips.

(a) For $k \geq 2$,

$$\begin{aligned} & P(\text{1st tail occurs on kth toss} | X_1 = H) \\ & = P(\text{first k-2 transitions are } H \to H \text{ and the last transition is } H \to T) \\ & = \left(\frac{3}{4}\right)^{k-2} \frac{1}{4}. \end{aligned}$$

(b) Irrespective of the starting state, $P(X_{5000} = H) \approx \pi_H$ where π_H, π_T are steady state probabilities. These probabilities $\pi_H = 4/7$ and $\pi_T = 3/7$ are obtained by solving equations

$$\begin{aligned} \pi_T p_{TH} + \pi_H p_{HH} &= \pi_H \\ \pi_T + \pi_H &= 1 \end{aligned}$$

(c)

$$\begin{aligned} P(X_{5000} = H, X_{5002} = H) &= P(X_{5000} = H) P(X_{5002} = H | X_{5000} = H) \\ &\approx \pi_H P(X_{5002} = H | X_{5000} = H) \\ &= \pi_H (p_{HT} p_{TH} + p_{HH} p_{HH}) \\ &= \frac{4}{7}\left(\frac{1}{4} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{3}{4}\right) \\ &= \frac{124}{336} \end{aligned}$$

(d)

$$\begin{aligned} & P(X_{5001} = \cdots = X_{5000+m} = H | X_{5001} = \cdots = X_{5000+m}) \\ & = \frac{P(X_{5001} = \cdots = X_{5000+m} = H)}{P(X_{5001} = \cdots = X_{5000+m} = H) + P(X_{5001} = \cdots = X_{5000+m} = T)} \\ & = \frac{P(X_{5001} = H) p_{HH}^{m-1}}{P(X_{5001} = H) p_{HH}^{m-1} + P(X_{5001} = T) p_{TT}^{m-1}} \approx \frac{\pi_H p_{HH}^{m-1}}{\pi_H p_{HH}^{m-1} + \pi_T p_{TT}^{m-1}} \\ & \to 1 \text{ as } m \to \infty \end{aligned}$$

(e) Let's examine the expected number of additional tosses until the next heads, given heads currently. This can be found by total expectation, by conditioning on what happens on the next toss. Given that the next toss is tails, the number of additional tosses until we observe the next heads is geometric with parameter $\frac{1}{3}$. Therefore, given tails, the expected number of additional tosses required until we observe the next heads is 3. Hence, the expected number of additional flips required until we observe the next heads, given heads on the current toss is

$$p_{HH} \cdot 1 + p_{HT} \cdot (1+3) = \frac{7}{4}\ .$$

Given that the 375th heads occurs on the 500th toss, the number of additional flips until the 379th heads can be expressed as the sum of four random variables, each with an expectation equal to 7/4. Thus by linearity of expectation, the required answer is $4 \cdot \frac{7}{4} = 7$.

2. (a) The long-term frequency of winning can be found as sum of the long-term frequency of transitions from 1 to 2 and 2 to 2. These can be found from the steady-state probabilities π_1 and π_2, which are known to exist as the chain is aperiodic and recurrent. The local balance and normalization equations are as follows:

$$\begin{aligned} \frac{7}{15}\pi_1 &= \frac{5}{9}\pi_2\ , \\ \pi_1 + \pi_2 &= 1\ . \end{aligned}$$

Solving these we obtain,

$$\pi_1 = \frac{25}{46} \approx 0.54,\ \pi_2 = \frac{21}{46} \approx 0.46\ .$$

The probability of winning, which is the long-term frequency of the transitions from 1 to 2 and 2 to 2, can now be found as

$$\mathbf{P}(\text{winning}) = \pi_1 p_{12} + \pi_2 p_{22} = \frac{25}{46}\frac{7}{15} + \frac{21}{46}\frac{4}{9} = \frac{21}{46} \approx 0.46\ .$$

Note that from the balance equation for state 2,

$$\pi_2 = \pi_1 p_{12} + \pi_2 p_{22}\ ,$$

the long-term probability of winning always equals π_2.

(b) This question is one of determining the probability of absorption into the recurrent class $\{1A, 2A\}$. This probability of absorption can be found by recognizing that it will be the ratio of probabilities

$$\frac{p_{1,1A}}{p_{1,1A} + p_{1,1B}} = \frac{\frac{2}{15}}{\frac{2}{15} + \frac{1}{15}} = \frac{2}{3}\ .$$

More methodically, if we define a_i as the probability of being absorbed into the class $\{1A, 2A\}$, starting in state i, we can solve for the a_i by solving the system of equations

$$\begin{aligned} a_1 &= p_{1,1A} + p_{11}a_1 + p_{12}a_2 \\ &= \frac{2}{15} + \frac{1}{3}a_1 + \frac{7}{15}a_2 \\ a_2 &= p_{21}a_1 + p_{22}a_2 \\ &= \frac{5}{9}a_1 + \frac{4}{9}a_2\ , \end{aligned}$$

from which we determine that $a_1 = \frac{p_{1,1A}}{p_{1,1A}+p_{1,1B}} = \frac{2}{3}$.

(c) Let A, B be the events that Jack eventually plays with decks $1A$ & $2A$, $1B$ & $2B$, respectively, when starting in state 1. From part (b), we know that $\mathbf{P}(A) = a_1 = \frac{2}{3}$ and $\mathbf{P}(B) = 1 - a_1 = \frac{1}{3}$. The probability of winning can be determined as

$$\mathbf{P}(\text{winning}) = \mathbf{P}(\text{winning}|A)\mathbf{P}(A) + \mathbf{P}(\text{winning}|B)\mathbf{P}(B) \ .$$

By considering the corresponding the appropriate recurrent class and solving a problem similar to part (a), $\mathbf{P}(\text{winning}|A)$ and $\mathbf{P}(\text{winning}|B)$ can be determined; in these cases, the steady-state probabilities of each recurrent class are defined under the assumption of being absorbed into that particular recurrent class. Let's begin with $\mathbf{P}(\text{winning}|A)$. The local balance and normalization equations for the recurrent class $\{1A, 2A\}$ are

$$\begin{aligned} \frac{3}{5}\pi_{1A} &= \frac{1}{5}\pi_{2A} \ , \\ \pi_{1A} + \pi_{2A} &= 1 \ . \end{aligned}$$

Solving these we obtain,

$$\pi_{1A} = \frac{1}{4}, \ \pi_{2A} = \frac{3}{4} \ ,$$

and hence conclude that

$$\mathbf{P}(\text{winning}|A) = p_{1A,2A}\pi_{1A} + p_{2A,2A}\pi_{2A} = \pi_{2A} = \frac{3}{4} \ .$$

Similarly, the local balance and normalization equations for the recurrent class $\{1B, 2B\}$ are

$$\begin{aligned} \frac{3}{4}\pi_{1B} &= \frac{1}{8}\pi_{2B} \ , \\ \pi_{1B} + \pi_{2B} &= 1 \ . \end{aligned}$$

Solving these we obtain,

$$\pi_{1B} = \frac{1}{7}, \ \pi_{2B} = \frac{6}{7} \ ,$$

and hence conclude that

$$\mathbf{P}(\text{winning}|B) = p_{1B,2B}\pi_{1B} + p_{2B,2B}\pi_{2B} = \pi_{2B} = \frac{6}{7} \ .$$

Putting these pieces together, we have that

$$\begin{aligned} \mathbf{P}(\text{winning}) &= \mathbf{P}(\text{winning}|A)\mathbf{P}(A) + \mathbf{P}(\text{winning}|B)\mathbf{P}(B) \\ &= \frac{3}{4} \cdot \frac{2}{3} + \frac{6}{7} \cdot \frac{1}{3} \\ &= \frac{11}{14} \approx 0.79 \ , \end{aligned}$$

meaning that Jack substantially increases the odds to his favor by slipping additional cards into the decks.

(d) The expected time until Jack slips cards into the deck is the same as the expected time until the Markov chain enters a recurrent state. Let μ_i be the expected amount of time until a recurrent state is reached from state i. We have the equations

$$\mu_1 = 1 + p_{11}\mu_1 + p_{12}\mu_2 = 1 + \frac{1}{3}\mu_1 + \frac{7}{15}\mu_2$$
$$\mu_2 = 1 + p_{21}\mu_1 + p_{22}\mu_2 = 1 + \frac{5}{9}\mu_1 + \frac{4}{9}\mu_2 \ ,$$

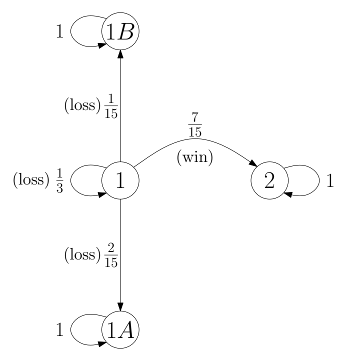
which when solved, yields the expected time until Jack slips cards into the deck,

$$\mu_1 = 9.2 \ .$$

(e) Let S be the number of times that the dealer switches from deck #2 to deck #1, which equals the number of times that he/she switches from deck #1 to deck #2. Let p be the probability that $S = 0$, which is the sum of the probability of all ways for the first change of state to be from state 1 to state 1A or state 1B,

$$p = \left(\frac{2}{15} + \frac{1}{15}\right) + \left(\frac{1}{3}\right)\left(\frac{2}{15} + \frac{1}{15}\right) + \left(\frac{1}{3}\right)^2\left(\frac{2}{15} + \frac{1}{15}\right) + \ldots = \frac{1}{1 - 1/3} \cdot \frac{3}{15} = \frac{3}{10} \ .$$

Alternatively, p is the probability of absorption of the following modified chain into an absorbing state ($1A$ or $1B$), when started in state 1:



As $\mathbf{P}(S > 0) = 1 - p$, and similarly, $\mathbf{P}(S > k + 1 | S > k) = 1 - p$, it should be clear that S will be a shifted geometric, and thus

$$p_S(k) = \left(\frac{7}{10}\right)^k \frac{3}{10} \quad k = 0, 1, 2, \ldots \ .$$

(f) Note that S from part (e) is the total number of cycles from 1 to 2 and back to 1. During the ith cycle, the number of wins, W_i, is a geometric random variable with parameter $q = \frac{5}{9}$. Thus the total number of wins by Jack before he slips extra cards into the deck is

$$W = W_1 + W_2 + \ldots + W_S \ ,$$

which is a random number of random variables, all of which are independent. Conditioned on $S > 0$, W is a geometric (with parameter p) number of geometric (with parameter q) random variables, all conditionally independent, and thus from the theory of splitting Bernoulli processes,

$$p_{W|S>0}(k) = (1 - pq)^{k-1} pq \quad k = 1, 2, \ldots \ ,$$

where $pq = \frac{3}{10} \cdot \frac{5}{9} = \frac{1}{6}$. When $S = 0$, it follows that $W = 0$, and thus by total probability,

$$p_W(k) = \begin{cases} \frac{3}{10} & k = 0 \\ (\frac{7}{10})(\frac{5}{6})^{k-1}\frac{1}{6} & k = 1, 2, \ldots \end{cases} .$$

(g) Let W be the total number of wins before slipping cards into the deck (as in part (f)), and similarly let L be the total number of losses before absorption. We know from part (d) that $\mathbf{E}[W + L] = \mu_1 = 9.2$. From part (f) we can find $\mathbf{E}[W]$ by total expectation,

$$\mathbf{E}[W] = E[W|S = 0]\mathbf{P}(S = 0) + E[W|S > 0]\mathbf{P}(S > 0) = \frac{7/10}{1/6} = \frac{42}{10} = 4.2\ ,$$

because when conditioned on $S > 0$, the number of wins, W, is a geometric random variable with parameter $pq = \frac{1}{6}$. From linearity of expectation, we find

$$\mathbf{E}[L - W] = \mathbf{E}[W + L] - 2\mathbf{E}[W] = 9.2 - 2 \cdot 4.2 = 0.8\ .$$

(h) Using A to again denote the probability of being absorbed into the recurrent class $\{1A, 2A\}$, starting in state 1,

$$\begin{aligned} \mathbf{P}(X_n = 2A|X_{n+1} = 1A) &= \frac{\mathbf{P}(X_{n+1} = 1A|X_n = 2A)\mathbf{P}(X_n = 2A)}{\mathbf{P}(X_{n+1} = 1A)} \\ &= \frac{\mathbf{P}(X_{n+1} = 1A|X_n = 2A)\mathbf{P}(X_n = 2A|A)\mathbf{P}(A)}{\mathbf{P}(X_{n+1} = 1A|A)\mathbf{P}(A)} \\ &\approx \frac{p_{2A,1A}\pi_{2A}}{\pi_{1A}} \\ &= \frac{\frac{1}{5} \cdot \frac{3}{4}}{\frac{1}{4}} \\ &= \frac{3}{5}\ . \end{aligned}$$

Note that the right hand side above equals $p_{1A,2A}$, as clear from the local balance equation $\pi_{1A}p_{1A,2A} = \pi_{2A}p_{2A,1A}$.

3. $\mathbf{P}(\mathrm{D} > \alpha) = \mathbf{P}(|(X - \mu)/\mu| > \alpha) = \mathbf{P}(|X - \mu| > \alpha\mu)$
Using Chebyshev Inequality,

$$\mathbf{P}(|X - \mu| > \alpha\mu) \leq \frac{\sigma^2}{\alpha^2\mu^2} = \frac{1}{r^2\alpha^2}$$

Therefore,

$$\mathbf{P}(D > \alpha) \leq \frac{1}{r^2\alpha^2}$$

$$\mathbf{P}(D \leq \alpha) \geq 1 - \frac{1}{r^2\alpha^2}$$

4. Consider a random variable X with PMF

$$p_X(x) = \begin{cases} p, & \text{if } x = \mu - c; \\ p, & \text{if } x = \mu + c; \\ 1 - 2p, & \text{if } x = \mu. \end{cases}$$

The mean of X is μ, and the variance of X is $2pc^2$. To make the variance equal σ^2, set $p = \dfrac{\sigma^2}{2c^2}$. For this random variable, we have

$$\mathbf{P}(|X - \mu| \geq c) = 2p = \frac{\sigma^2}{c^2},$$

and therefore the Chebyshev inequality is tight.

G1†. Note that Y is $\mathcal{N}(nm, n)$, so

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi n}} e^{-\frac{(y-mn)^2}{2n}} \\ M_Y(s) &= e^{s(nm)+\frac{s^2 n}{2}} \end{aligned}$$

(a)

$$\begin{aligned} P(Y \geq \alpha) &= \int_\alpha^\infty f_Y(y)dy \\ &= \int_\alpha^\infty \frac{1}{\sqrt{2\pi n}} e^{-\frac{(y-mn)^2}{2n}} dy \end{aligned}$$

(b) For $s \geq 0$:

$$\begin{aligned} P(Y \geq \alpha) &\leq e^{-s\alpha} M_Y(s) \\ &= e^{-s\alpha}\left[e^{s(nm)+\frac{s^2 n}{2}}\right] \end{aligned}$$

By differentiating the right-hand side and setting to 0, we obtain

$$\begin{aligned} 0 &= \frac{d}{ds} e^{-s\alpha}\left[e^{snm+\frac{s^2 n}{2}}\right] \\ 0 &= (nm + s^* n - \alpha) e^{-s^*\alpha}\left[e^{s^* nm+\frac{(s^*)^2 n}{2}}\right] \end{aligned}$$

So we have that $s^* = \frac{\alpha - nm}{n}$, and substituting back we get

$$P(Y \geq \alpha) \leq e^{-\frac{(\alpha - nm)^2}{2n}}$$

and is valid when $s^* \geq 0$, which is when $\alpha \geq nm$.

†Required for 6.431; optional for 6.041