Problem Set 9: Solutions
Due December 2, 2011

1. (a) Let t_i be the expected time until the state HT is reached, starting in state i, i.e., the mean first passage time to reach state HT starting in state i. Note that t_S is the expected number of tosses until first observing heads directly followed by tails. We have,

$$\begin{aligned} t_S &= 1 + \frac{1}{2}t_H + \frac{1}{2}t_T \\ t_T &= 1 + \frac{1}{2}t_H + \frac{1}{2}t_T \\ t_H &= 1 + \frac{1}{2}t_H \end{aligned}$$

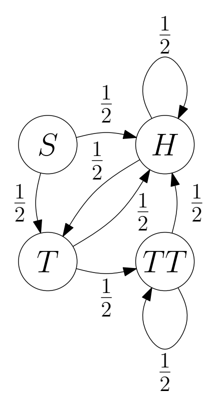
and by solving these equations, we find that the expected number of tosses until first observing heads directly followed by tails is

$$t_S = 4\ .$$

(b) To find the expected number of additional tosses necessary to again observe heads followed by tails, we recognize that this is the mean recurrence time t^*_{HT} of state HT. This can be determined as

$$\begin{aligned} t^*_{HT} &= 1 + p_{HT,H}t_H + p_{HT,T}t_T \\ &= 1 + \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 4 \\ &= 4\ . \end{aligned}$$

(c) Let's consider a Markov chain with states S, H, T, TT, where S is a starting state, H indicates heads on the current toss, T indicates tails on the current toss (without tails on the previous toss), and TT indicates tails over the last two tosses. The transition probabilities for this Markov chain are illustrated below in the state transition diagram:



Let t_i be the expected time until the state TT is reached, starting in state i, i.e., the mean first passage time to reach state TT starting in state i. Note that t_S is the expected number of tosses until first observing tails directly followed by tails. We have,

$$\begin{aligned} t_S &= 1 + \frac{1}{2}t_H + \frac{1}{2}t_T \\ t_T &= 1 + \frac{1}{2}t_H \\ t_H &= 1 + \frac{1}{2}t_H + \frac{1}{2}t_T \end{aligned}$$

and by solving these equations, we find that the expected number of tosses until first observing two consecutive tails is

$$t_S = 6 \; .$$

(d) To find the expected number of additional tosses necessary to again observe heads followed by tails, we recognize that this is the mean recurrence time t^*_{TT} of state TT. This can be determined as

$$\begin{aligned} t^*_{TT} &= 1 + p_{TT,H}t_H + p_{TT,TT}t_TT \\ &= 1 + \frac{1}{2}\cdot 6 + \frac{1}{2}\cdot 0 \\ &= 4 \; . \end{aligned}$$

It may be surprising that the average number of tosses until the first two consecutive tails is greater than the average number of tosses until heads is directly followed by tails, considering that the mean recurrence time between pairs of tosses with heads directly followed by tails equals the mean recurrence time between pairs of tosses that are both tails (or equivalently, the long-term frequency of pairs of tosses with heads followed by tails equals the long-term frequency of pairs of tosses with two consecutive tails[1]). This is a start-up artifact. Note that the distribution of the first passage time to reach state HT (or TT) starting in state S is the same as the conditional distribution of the recurrence time of state HT (or TT), given that it is greater than 1. Although in both cases the *expected values* of the recurrence times are equal (this is what parts (b) and (d) tell us), the conditional expected values of the recurrence time given that it is greater than 1 is not the same in both cases (possible, because the unconditional distributions are not equal).

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,

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

[1] See problem 7.34 on page 399 of the text for a detailed explanation of this correspondence between mean recurrence times and steady-state probabilities.

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.

3. Because of the independence of the per-week profits and the high number of weeks, the desired probability can be well approximated by the Central Limit Theorem. Let X_i be Joe's profit in the ith week. The X_is have a common mean $\mu = 5200$ and variance $\sigma^2 = \frac{1}{12}(6800)^2$.

$$\begin{aligned} P\left(\sum_{i=1}^{52} X_i \le 250000\right) &= P\left(\frac{\sum_{i=1}^{52} X_i - 52\mu}{\sigma\sqrt{52}} \le \frac{250000 - 52\mu}{\sigma\sqrt{52}}\right) \\ &\approx P\left(Z \le \frac{250000 - 52\mu}{\sigma\sqrt{52}}\right) \text{ where Z is a standard normal r.v. (CLT)} \\ &= P\left(Z \le \frac{250000 - 52(5200)}{\frac{1}{\sqrt{12}}6800\sqrt{52}}\right) \\ &= P\left(Z \le \frac{-3\sqrt{3}}{\sqrt{13}}\right) \\ &= 1 - \phi\left(\frac{3\sqrt{3}}{\sqrt{13}}\right) \approx 1 - \phi(1.44) \approx 0.075 \end{aligned}$$

4. The probability that the airline will have to deny passengers from boarding is the probability that more than 300 passengers show up. Let N be the number of passengers that show up, which is a binomial random variable, with parameters n and $p = 0.9$; thus, $\mathbf{E}[N] = 0.9n$ and $\sigma_N = \sqrt{n(0.1)(0.9)} = \sqrt{0.09n}$. Using the de Moivre - Laplace normal approximation to the binomial,

we have

$$\begin{aligned}\mathbf{P}(N > 300) &= \mathbf{P}(N \geq 300 + 0.5) \\ &= \mathbf{P}\left(\frac{N - 0.9n}{\sqrt{0.09n}} \geq \frac{300.5 - 0.9n}{\sqrt{0.09n}}\right) \\ &\approx 1 - \Phi\left(\frac{300.5 - 0.9n}{\sqrt{0.09n}}\right)\end{aligned}$$

In order for the airline to be approximately 99 percent confident that it will not have to deny boarding to passengers holding tickets,

$$\mathbf{P}(N > 300) \approx 0.01$$

Therefore,

$$\Phi\left(\frac{300.5 - 0.9n}{\sqrt{0.09n}}\right) \approx 0.99$$

and as $\Phi(2.33) \approx 0.99$,

$$\begin{aligned}\frac{300.5 - 0.9n}{\sqrt{0.09n}} &\approx 2.33 \\ 300.5 - 0.9n &\approx 2.33\sqrt{0.09n} \\ (0.9)^2 n^2 - 2(0.9)(300.5)n + (300.5)^2 &\approx (2.33)^2(0.09)n\end{aligned}$$

Solving the quadratic equation above and choosing the lesser of the two solutions for n tell us that $n \approx 320$.

5. (a) Yes, to 0. Applying the weak law of large numbers, we have

$$\mathbf{P}(|U_i - \mu| > \epsilon) \to 0 \text{ as } i \to \infty, \text{ for all } \epsilon > 0$$

Here $\mu = 0$ since $X_i \sim U(-1.0, 1.0)$.

(b) Yes, to 1. Since $W_i \leq 1$, we have for $\epsilon > 0$,

$$\begin{aligned}\lim_{i\to\infty} \mathbf{P}(|W_i - 1| > \epsilon) &= \lim_{i\to\infty} \mathbf{P}(\max\{X_1, \cdots, X_i\} < 1 - \epsilon) \\ &= \lim_{i\to\infty} \mathbf{P}(X_1 < 1 - \epsilon) \cdots \mathbf{P}(X_i < 1 - \epsilon)\} \\ &= \lim_{i\to\infty} (1 - \frac{\epsilon}{2})^i \\ &= 0.\end{aligned}$$

(c) Yes, to 0.

$|V_n| \leq \min\{|X_1|, |X_2|, \cdots, |X_n|\}$
but $\min\{|X_1|, |X_2|, \cdots, |X_n|\}$ converges to 0 in probability. So, since $|V_n| \geq 0$, $|V_n|$ converges to 0 in probability. To see why $\min\{|X_1|, |X_2|, \cdots, |X_n|\}$ converges to 0 in probability, note that:

$$\begin{aligned}\lim_{i\to\infty} \mathbf{P}(|\min\{|X_1|, \cdots, |X_i|\} - 0| > \epsilon) &= \lim_{i\to\infty} \mathbf{P}(\min\{|X_1|, \cdots, |X_i|\} > \epsilon) \\ &= \lim_{i\to\infty} \mathbf{P}(|X_1| > \epsilon) \cdot \mathbf{P}(|X_2| > \epsilon) \cdots \mathbf{P}(|X_i| > \epsilon) \\ &= \lim_{i\to\infty} (1 - \epsilon)^i \text{ since } |X_i| \text{ is uniform between 0 and 1} \\ &= 0.\end{aligned}$$

6. The probability that you will believe the fair coin to be biased is the probability that the fair coin will come up with more than 525 heads out of the 1000 tosses. Let S be the number of times the coin comes up heads, which is a binomial random variable, with parameters $n = 1000$ and $p = 0.5$, so that $\mathbf{E}[S] = (1000)(0.5) = 500$ and $\sigma_S = \sqrt{(1000)(0.5)(0.5)} = 5\sqrt{10}$.

   (a) Using the de Moivre - Laplace normal approximation to the binomial, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 525.5) \\
&= \mathbf{P}\left(\frac{S-500}{5\sqrt{10}} \geq \frac{525.5-500}{5\sqrt{10}}\right) \\
&\approx 1 - \Phi\left(\frac{25.5}{5\sqrt{10}}\right) \\
&= 1 - \Phi(1.6128) \\
&\approx 0.0537.
\end{aligned}$$

   (b) Using the Markov inequality, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 526) \\
&\leq \frac{\mathbf{E}[S]}{526} \\
&= \frac{500}{526} \\
&\approx 0.951.
\end{aligned}$$

   We see that using the Markov inequality gives us a weak upper bound, considering the approximate probability as calculated in part (a).

   (c) Using the Chebyshev inequality, we have

$$\begin{aligned}
\mathbf{P}(S > 525) &= \mathbf{P}(S \geq 526) \\
&= \frac{1}{2}(\mathbf{P}(S \geq 526) + \mathbf{P}(S \leq 474)) \text{ (by symmetry, since p = 0.5)} \\
&= \frac{1}{2}P(|S-500| \geq 26) \\
&\leq \frac{1}{2}\frac{\sigma_S^2}{26^2} \\
&= \frac{1}{2}\frac{25 \cdot 10}{26^2} \\
&\approx 0.185.
\end{aligned}$$

   We see that the Chesyshev inequality provides a substantial improvement upon the upper bound calculated by the Markov inequality in part (b).

G1†. (a) i. To recover the loss he sustains from a single tails, the player requires the profit from two heads. Thus to break even or win, the player requires at least twice as many heads as tails. For 100 tosses, he needs at least 67 heads.

ii. Let S_{100} be the number of heads in 100 tosses of a fair coin. Therefore, S_{100} is a binomial random variable with $n = 100$ and $p = 1/2$. Thus,

$$\begin{aligned}
p_{S_{100}}(k) &= \begin{cases} \binom{100}{k}\left(\frac{1}{2}\right)^{100}, & k = 0, 1, \ldots, 100 \\ 0, & \text{otherwise,} \end{cases} \\
\mathbf{E}[S_{100}] &= np = 50, \\
\sigma_{S_{100}} &= \sqrt{np(1-p)} = 5.
\end{aligned}$$

(b) $P_w = \mathbf{P}(\text{at least 67 heads in 100 tosses}) = \sum_{k=67}^{100} \binom{100}{k}\left(\frac{1}{2}\right)^{100}$.

(c) Since $S_{100} \geq 0$ and $\mathbf{E}[S_{100}] = 50$, we can use the Markov bound:

$$\mathbf{P}(S_{100} \geq 67) \leq \frac{\mathbf{E}[S_{100}]}{67} \approx \frac{\frac{1}{2}\cdot 100}{\frac{2}{3}\cdot 100} \approx 0.75.$$

(d) Let $X_n = 1$ if the nth toss is heads and $X_n = 0$ if the nth toss is tails. Then $\text{var}(X_n) = \frac{1}{4}$, the X_n's are independent, and

$$S_{100} = \sum_{k=1}^{100} X_k$$

Therefore $\text{var}(S_{100}) = 100 \cdot \frac{1}{4} = 25$, $\sigma_{S_{100}} = 5$, so from the Chebyshev bound,

$$\mathbf{P}(|S_{100} - \mathbf{E}[S_{100}]| \geq 17) \leq \frac{\text{var}(S_{100})}{17^2} = \left(\frac{5}{17}\right)^2 \approx 0.087.$$

From the symmetry of the problem we see that

$$\mathbf{P}(S_{100} \geq 67) = \mathbf{P}(S_{100} - \mathbf{E}[S_{100}] \geq 17) = \mathbf{P}(S_{100} - \mathbf{E}[S_{100}] \leq -17) = \frac{1}{2}\mathbf{P}(|S_{100} - \mathbf{E}[S_{100}]| \geq 17).$$

So we can halve the above bound:

$$\mathbf{P}(S_{100} \geq 67) \leq \frac{\text{var}(S_{100})}{2 \cdot (17)^2} \approx 0.043.$$

(e) We find $\mathbf{P}(S_{100} \geq 67)$ using the CLT as such

$$\begin{aligned}
\mathbf{P}(S_{100} \geq 67) &= \mathbf{P}(S_{100} - \mathbf{E}[S_{100}] \geq 67 - \mathbf{E}[S_{100}]) \\
&= \mathbf{P}\left(\frac{S_{100} - \mathbf{E}[S_{100}]}{\sigma_{S_{100}}} \geq \frac{67 - \mathbf{E}[S_{100}]}{\sigma_{S_{100}}}\right) \\
&\approx 1 - \Phi\left(\frac{67 - \mathbf{E}[S_{100}]}{\sigma_{S_{100}}}\right) \\
&\approx 1 - \Phi\left(\frac{67 - 50}{5}\right).
\end{aligned}$$

Without the 1/2 correction, the probability becomes $1-\Phi\left(\frac{67-50}{5}\right) = 1-\Phi(3.4) \approx 1-.9997 = 0.0003$.

With the 1/2 correction, the probability becomes $1-\Phi\left(\frac{66.5-50}{5}\right) = 1-\Phi(3.3) \approx 1-.9995 = 0.0005$.

So the Double or Quarter game (at least when played this way) looks like a poor way to attempt to make one's fortune. The game is deceptive in that your expected wealth (i.e, $(9/8)^n$) becomes very large due to an ever smaller probability of ever larger wins as $n \to \infty$.

(f) For this example, the Markov and Chebyshev bounds provide very loose upper bounds. The central limit theorem provides a closer approximation.