

Problem 1

(a) Find the $\mathbb{E}[\bar{X}(N)]$ analytically. Does it really estimate the mean of X ?

We have that

$$\bar{X}(N) = \frac{1}{N} \sum_{i=1}^N X_i, \quad X_i \sim \text{Uniform}(0, 1) .$$

We first compute the expectation of X_i . We have

$$\begin{aligned} \mathbb{E}[X_i] &= \int_{-\infty}^{\infty} x \cdot p(X = x) \, dx \\ &= \int_0^1 x \cdot p(X = x) \, dx \\ &= \int_0^1 x \cdot \frac{1}{1-0} \, dx \\ &= \int_0^1 x \, dx \\ &= \left[\frac{1}{2} x^2 \right]_0^1 \\ &= \frac{1}{2} (1^2 - 0^2) \\ &= \frac{1}{2} \end{aligned} \tag{1}$$

Therefore, we can compute the expectation of $\bar{X}(N)$ as follows:

$$\begin{aligned} \mathbb{E}[\bar{X}(N)] &= \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n X_i \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] \\ &\stackrel{(1)}{=} \frac{1}{n} \sum_{i=1}^n \frac{1}{2} \\ &= \frac{1}{n} \cdot n \cdot \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

Therefore, the expected value of $\mathbb{E}[\bar{X}(N)]$ estimates the true mean of X .

(b) Find the $\text{Var}(\bar{X}(N))$ analytically. Hint: find the variance of a single X_i first.

Let us first compute the expectation of $\mathbb{E}[X_i^2]$:

$$\begin{aligned}
 \mathbb{E}[X_i^2] &= \int_0^1 x^2 \cdot p(X=x) \, dx \\
 &= \int_0^1 x^2 \cdot \frac{1}{1-0} \, dx \\
 &= \int_0^1 x^2 \, dx \\
 &= \left[\frac{1}{3} \cdot x^3 \right]_0^1 \\
 &= \frac{1}{3}(1-0) \\
 &= \frac{1}{3}
 \end{aligned} \tag{2}$$

We can now compute the variance of a X_i as follows:

$$\begin{aligned}
 \text{Var}(X_i) &= \mathbb{E}[(X_i - \mathbb{E}[X_i])^2] \\
 &= \mathbb{E}\left[\left(X_i - \frac{1}{2}\right)^2\right] \\
 &= \mathbb{E}\left[X_i^2 + \frac{1^2}{2^2} - 2 \cdot X_i \cdot \frac{1}{2}\right] \\
 &= \mathbb{E}\left[X_i^2 + \frac{1^2}{2^2} - X_i\right] \\
 &= \mathbb{E}[X_i^2] + \mathbb{E}\left[\frac{1}{4}\right] - \mathbb{E}[X_i] \\
 &= \mathbb{E}[X_i^2] + \frac{1}{4} - \frac{1}{2} \\
 &\stackrel{(2)}{=} \frac{1}{3} + \frac{1}{4} - \frac{1}{2} \\
 &= \frac{4}{12} + \frac{3}{12} - \frac{6}{12} \\
 &= \frac{1}{12} .
 \end{aligned}$$

We can now compute the variance of $\bar{X}(N)$ as follows:

$$\begin{aligned}
 \text{Var}(\bar{X}(N)) &= \text{Var}\left(\frac{1}{N} \sum_{i=1}^N X_i\right) \\
 &= \frac{1}{N^2} \cdot \text{Var}\left(\sum_{i=1}^N X_i\right) \\
 &= \frac{1}{N^2} \cdot \sum_{i=1}^N \text{Var}(X_i) \\
 &= \frac{1}{N^2} \cdot \sum_{i=1}^N \frac{1}{12} \\
 &= \frac{1}{N^2} \cdot N \cdot \frac{1}{12} \\
 &= \frac{1}{N} \cdot \frac{1}{12}
 \end{aligned} \tag{3}$$

where Equation 3 holds as each drawing X_i is independent of each other drawing X_j ($\Pr(X_i | X_j) = \Pr(X_i)$), and we can therefore use the property of the variance for i.i.d. variables $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

Therefore, in the limit, we have

$$\lim_{n \rightarrow \infty} \text{Var}(\bar{X}(N)) = \lim_{n \rightarrow \infty} \frac{1}{N} \cdot \frac{1}{12} = 0$$

with a linear convergence rate.

(c) The probability density function of $\hat{X}(N)$ is given as

$$f_{\hat{X}(N)}(x) = \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot x^{\frac{N-1}{2}} \cdot (1-x)^{\frac{N-1}{2}}.$$

Let us now compute the expectation of $\hat{X}(N)$:

$$\begin{aligned}
\mathbb{E}[\hat{X}(N)] &= \int_0^1 x \cdot f_{\hat{X}(N)}(x) \, dx \\
&= \int_0^1 x \cdot \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot x^{\frac{N-1}{2}} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x \cdot x^{\frac{N-1}{2}} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x^{\frac{N-1}{2}+1} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x^{\frac{N+1}{2}} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+1}{2}! \frac{N-1}{2}!}{\left(\frac{N+1}{2} + \frac{N-1}{2} + 1\right)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+1}{2}! \frac{N-1}{2}!}{\left(\frac{N-1}{2} + \frac{N-1}{2} + 1 + 1\right)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+1}{2}! \frac{N-1}{2}!}{\left(2 \cdot \frac{N-1}{2} + 2\right)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+1}{2}! \frac{N-1}{2}!}{(N-1+2)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+1}{2}! \frac{N-1}{2}!}{(N+1)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{1}{2} \cdot \frac{\left(\frac{N+1}{2}\right)! \frac{N-1}{2}!}{N!} \\
&= \frac{N!}{\left(\frac{N-1}{2}\right)! \left(N - \frac{N-1}{2}\right)!} \cdot \frac{1}{2} \cdot \frac{\left(\frac{N+1}{2}\right)! \frac{N-1}{2}!}{N!} \\
&= \frac{1}{\left(\frac{N-1}{2}\right)! \left(N - \frac{N-1}{2}\right)!} \cdot \frac{1}{2} \cdot \frac{\left(\frac{N+1}{2}\right)! \frac{N-1}{2}!}{1} \\
&= \frac{1}{\left(N - \frac{N-1}{2}\right)!} \cdot \frac{1}{2} \cdot \frac{\left(\frac{N+1}{2}\right)!}{1} \\
&= \frac{1}{\left(\frac{N+1}{2}\right)!} \cdot \frac{1}{2} \cdot \frac{\left(\frac{N+1}{2}\right)!}{1} \\
&= \frac{1}{2}
\end{aligned} \tag{4}$$

where we used the given known integral rule to obtain Equation 4, and Equation 5 results from

$$\begin{aligned}
N - \frac{N-1}{2} &= \frac{2N}{2} - \frac{N-1}{2} \\
&= \frac{N + N - (N-1)}{2} \\
&= \frac{N+1}{2} .
\end{aligned}$$

Therefore, the expectation of the of $\hat{X}(N)$ is $\frac{1}{2}$, i.e., the true mean of X .

(d) Find the $\text{Var}(\hat{X}(N))$ analytically. Which distribution has the better variance $\bar{X}(N)$ or $\hat{X}(N)$?

We compute the expectation of $\hat{X}(N)^2$:

$$\begin{aligned}
\mathbb{E}[\hat{X}(N)^2] &= \int_0^1 x^2 \cdot f_{\hat{X}(N)}(x) \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x^2 \cdot x^{\frac{N-1}{2}} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x^{\frac{N-1}{2}+2} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \int_0^1 x^{\frac{N+3}{2}} \cdot (1-x)^{\frac{N-1}{2}} \, dx \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(\frac{N+3}{2} + \frac{N-1}{2} + 1)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(\frac{N-1}{2} + \frac{N-1}{2} + 1 + 2)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(2 \cdot \frac{N-1}{2} + 3)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N-1+3)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N+2)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{N+1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N+2) \cdot (N+1)!} \\
&= \binom{N}{\frac{N-1}{2}} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N+2) \cdot N!} \\
&= \frac{N!}{(\frac{N-1}{2})! (N - \frac{N-1}{2})!} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N+2) \cdot N!} \\
&= \frac{1}{(\frac{N-1}{2})! (N - \frac{N-1}{2})!} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2}! \frac{N-1}{2}!}{(N+2)} \\
&= \frac{1}{(N - \frac{N-1}{2})!} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2}!}{(N+2)} \\
&= \frac{1}{\frac{N+1}{2}!} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2}!}{(N+2)} \\
&= \frac{1}{\frac{N+1}{2}!} \cdot \frac{1}{2} \cdot \frac{\frac{N+3}{2} \cdot (\frac{N+1}{2})!}{(N+2)} \\
&= \frac{1}{2} \cdot \frac{\frac{N+3}{2}}{(N+2)} \\
&= \frac{1}{2} \cdot \frac{\frac{N+3}{2} \cdot 2}{(N+2) \cdot 2} \\
&= \frac{1}{2} \cdot \frac{N+3}{(N+2) \cdot 2} \\
&= \frac{1}{2} \cdot \frac{N+2+1}{(N+2) \cdot 2} \\
&= \frac{1}{2} \cdot \left(\frac{N+2}{(N+2) \cdot 2} + \frac{1}{(N+2) \cdot 2} \right)
\end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \cdot \left(\frac{1}{2} + \frac{1}{2N+4} \right) \\
 &= \frac{1}{4} + \frac{1}{4N+8}
 \end{aligned}$$

We can finally compute the variance of $\hat{X}(N)$ as follows:

$$\begin{aligned}
 \text{Var}(\hat{X}(N)) &= \mathbb{E}[\hat{X}(N)^2] - \mathbb{E}[\hat{X}(N)]^2 \\
 &= \mathbb{E}[\hat{X}(N)^2] - \left(\frac{1}{2} \right)^2 \\
 &= \mathbb{E}[\hat{X}(N)^2] - \frac{1}{4} \\
 &= \frac{1}{4} + \frac{1}{4N+8} - \frac{1}{4} \\
 &= \frac{1}{4N+8}
 \end{aligned}$$

which, in the limit, converges to zero:

$$\lim_{N \rightarrow \infty} \text{Var}(\hat{X}(N)) = \lim_{N \rightarrow \infty} \frac{1}{4N+8} = 0$$

As the convergence rate of $\text{Var}(\hat{X}(N))$ is slower than $\text{Var}(\bar{X}(N))$, and that both have the same (correct) expectation, we can conclude that $\bar{X}(N)$ has the better variance, i.e., it is better to estimate the mean of a uniform distribution by using the mean of the samples instead of the median.

(e) See provided code.

Problem 2

(a) Show that if $E < 0$, then the game is losing in a long run and if $E > 0$, then the game is winning in a long run.

Let π be the stationary distribution of the Markov chain, and

$$E = \pi_0 \cdot (B1_+ - B1_-) + (\pi_1 + \pi_2) \cdot (B2_+ - B2_-) .$$

We remark that E is composed of two sums, the first sum (which includes π_0) denotes the expected return when in state 0. As π is the stationary distribution, the state 0 is reached with π_0 probability at each time step in the long run. The $(B1_+ - B1_-)$ part is positive if probability of gaining one dollar is greater than losing one dollar when playing game B while in state 0.¹

Because the second part of the equation follows the same structure (one could multiply out $(\pi_1 + \pi_2)$ to obtain the same argument for the two remaining states of the Markov chain), we can conclude that E denotes the expected *gain* (or *return*) when playing game B in the long run.

Therefore, if the expected gain is negative, the game is losing in the long run. On the other hand, a positive expected gain would imply that the game is winning.

(b) We have three states in our Markov chain: 0, 1, and 2, which denotes the number that results from the current capital modulo 3. We now define the transition matrix P as follows:

$$P = \begin{pmatrix} 0.0 & 0.095 & 0.905 \\ 0.255 & 0.0 & 0.745 \\ 0.745 & 0.255 & 0.0 \end{pmatrix}$$

We remark that each play results in a win or a loss, and therefore a probability of 0.0 to stay in the same state after playing a round, hence, the diagonal elements of P are all 0.0.

The stationary distribution of the Markov chain is denoted as

$$\pi := \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 \end{pmatrix} .$$

We now compute the stationary distribution from P by computing the stationary value for each π_i .

$$\begin{aligned} \pi_0 &= \pi_0 \cdot 0.0 + \pi_1 \cdot 0.255 + \pi_2 \cdot 0.745 \\ &= \pi_1 \cdot 0.255 + \pi_2 \cdot 0.745 \end{aligned} \tag{6}$$

¹See the definition of the states in (b).

$$\begin{aligned}
\pi_1 &= \pi_0 \cdot 0.095 + \pi_1 \cdot 0.0 + \pi_2 \cdot 0.255 \\
\pi_1 &= \pi_0 \cdot 0.095 + \pi_2 \cdot 0.255 \\
\pi_1 &\stackrel{(6)}{=} (\pi_1 \cdot 0.255 + \pi_2 \cdot 0.745) \cdot 0.095 + \pi_2 \cdot 0.255 \\
\pi_1 &= \pi_1 \cdot 0.255 \cdot 0.095 + \pi_2 \cdot 0.745 \cdot 0.095 + \pi_2 \cdot 0.255 & | - (\pi_1 \cdot 0.255 \cdot 0.095) \\
\pi_1(1 - 0.024225) &= \pi_2 \cdot 0.745 \cdot 0.095 + \pi_2 \cdot 0.255 & | \div (1 - 0.024225) \\
\pi_1 &= \frac{\pi_2 \cdot (0.255 + 0.070775)}{(1 - 0.024225)} \\
\pi_1 &= \pi_2 \cdot \frac{0.325775}{0.975775} \\
\pi_1 &= \pi_2 \cdot 0.333862827
\end{aligned} \tag{7}$$

Using the fact that $\pi_0 + \pi_1 + \pi_2 = 1$, we can now compute π_2 :

$$\begin{aligned}
1 &= \pi_0 + \pi_1 + \pi_2 \\
1 &\stackrel{(6)}{=} (\pi_1 \cdot 0.255 + \pi_2 \cdot 0.745) + \pi_1 + \pi_2 \\
1 &= \pi_1 \cdot 1.255 + \pi_2 \cdot 1.745 \\
1 &\stackrel{(7)}{=} (\pi_2 \cdot 0.333862827) \cdot 1.255 + \pi_2 \cdot 1.745 \\
1 &= \pi_2 \cdot 0.4189978479 + \pi_2 \cdot 1.745 \\
1 &= \pi_2 \cdot 2.1639978479 \\
\pi_2 &= \frac{1}{2.1639978479} \\
\pi_2 &= 0.4621076684
\end{aligned}$$

from which we can infer

$$\begin{aligned}
\pi_1 &= \pi_2 \cdot 0.333862827 \\
&= 0.4621076684 \cdot 0.333862827 \\
&= 0.1542805726
\end{aligned}$$

and

$$\begin{aligned}
\pi_0 &= \pi_1 \cdot 0.255 + \pi_2 \cdot 0.745 \\
&= 0.1542805726 \cdot 0.255 + 0.4621076684 \cdot 0.745 \\
&= 0.03934154601 + 0.344270213 \\
&= 0.383611759 .
\end{aligned}$$

Therefore, we have the stationary distribution

$$\pi = \begin{pmatrix} 0.383611759 & 0.1542805726 & 0.4621076684 \end{pmatrix}$$

with $\pi = \pi P$ and total sum of one ($0.383611759 + 0.1542805726 + 0.4621076684 = 1$).

Therefore, the expected gain when playing game B in the long run is

$$\begin{aligned} E &= \pi_0 \cdot (B1_+ - B1_-) + (\pi_1 + \pi_2) \cdot (B2_+ - B2_-) \\ &= \pi_0 \cdot (0.095 - 0.905) + (\pi_1 + \pi_2) \cdot (0.745 - 0.255) \\ &= 0.383611759 \cdot (-0.81) + (0.1542805726 + 0.4621076684) \cdot 0.49 \\ &= 0.383611759 \cdot (-0.81) + 0.616388241 \cdot 0.49 \\ &= -0.0086952867 \end{aligned}$$

(c) See provided code.

(d) We first observe that we can keep the same three states of the Markov chain as introduced in (b). However, we need to update the transition matrix P . We now construct P to model the transition between the states given that the player *plays the random game with equal probabilities every time*, i.e., plays game A with probability 0.5 and game B with probability 0.5. Let us denote the transition matrix of game B as introduced in (b) as P_B . We now construct the transition matrix P_A for game A as follows:

$$P_A = \begin{pmatrix} 0.0 & 0.495 & 0.505 \\ 0.505 & 0.0 & 0.495 \\ 0.495 & 0.505 & 0.0 \end{pmatrix}$$

where in each state, the probability of winning one dollar is 0.495 and the probability of losing one dollar is 0.505. We can now construct the transition matrix for the combined Markov chain, i.e., when playing both games with equal probabilities, as follows:

$$\begin{aligned} P &= 0.5 \cdot P_A + 0.5 \cdot P_B \\ &= \begin{pmatrix} 0.0 & 0.295 & 0.705 \\ 0.38 & 0.0 & 0.62 \\ 0.62 & 0.38 & 0.0 \end{pmatrix}. \end{aligned}$$

We can now compute the stationary distribution of the new Markov chain as done in (b). In the following derivation, we may skip some steps for brevity, as the exact same derivation process as in (b) is followed.

We start by computing π_0 as

$$\pi_0 = \pi_1 \cdot 0.38 + \pi_2 \cdot 0.62 \tag{8}$$

and

$$\begin{aligned}\pi_1 &= \pi_0 \cdot 0.295 + \pi_2 \cdot 0.38 \\ \pi_1 &= (\pi_1 \cdot 0.38 + \pi_2 \cdot 0.62) \cdot 0.295 + \pi_2 \cdot 0.38 \\ \pi_1 &= \pi_1 \cdot 0.1121 + \pi_2 \cdot 0.1829 + \pi_2 \cdot 0.38 \\ \pi_1 &= \pi_2 \cdot 0.6339677892.\end{aligned}$$

Thus, we can compute π_2 :

$$\begin{aligned}1 &= \pi_0 + \pi_1 + \pi_2 \\ 1 &= (\pi_1 \cdot 0.38 + \pi_2 \cdot 0.62) + \pi_1 + \pi_2 \\ 1 &= \pi_1 \cdot 1.38 + \pi_2 \cdot 1.62 \\ 1 &= \pi_2 \cdot 0.6339677892 \cdot 1.38 + \pi_2 \cdot 1.62 \\ 1 &= \pi_2 \cdot 2.4948755491 \\ \pi_2 &= \frac{1}{2.4948755491} \\ \pi_2 &= 0.4008215962\end{aligned}$$

and infer π_0 and π_1 :

$$\begin{aligned}\pi_1 &= \pi_2 \cdot 0.6339677892 \\ &= 0.2541079812\end{aligned}$$

and

$$\begin{aligned}\pi_0 &= \pi_1 \cdot 0.38 + \pi_2 \cdot 0.62 \\ &= 0.2541079812 \cdot 0.38 + 0.4008215962 \cdot 0.62 \\ &= 0.3450704225\end{aligned}$$

Therefore, we can now compute the expected gain when playing the new game by first extracting the probabilities of winning and losing in each state. Let us denote the state $0, 1, 2$ as A, B, C to ease the notation. We have

$$\begin{aligned}A_+ &= 0.295 \\ A_- &= 0.705 \\ B_+ &= 0.62 \\ B_- &= 0.38 \\ C_+ &= 0.62 \\ C_- &= 0.38\end{aligned}$$

and can finally compute the expected gain of the new Markov chain as follows:²

$$\begin{aligned}
 E &= \pi_0 \cdot (A_+ - A_-) + \pi_1 \cdot (B_+ - B_-) + \pi_2 \cdot (C_+ - C_-) \\
 &= \pi_0 \cdot (0.295 - 0.705) + \pi_1 \cdot (0.62 - 0.38) + \pi_2 \cdot (0.62 - 0.38) \\
 &= 0.3450704225 \cdot (-0.41) + 0.2541079812 \cdot 0.24 + 0.4008215962 \cdot 0.24 \\
 &= 0.01570422535
 \end{aligned} \tag{9}$$

Therefore, the expected gain when playing the new game is positive, indicating that, on average, the player will win in the long run.

(e) See provided code.

(f) *Show analytically that the provided periodic strategy is winning in a long run.*

We wish to investigate whether playing both games in a given sequence (ABBABBABB...) is also winning.

We observe that the sequence is a repeating pattern of playing game A followed by playing game B two times. We can therefore compute the transition matrix for one of such patterns as follows:

$$P_{ABB} = P_A \cdot P_B \cdot P_B = P_A \cdot P_B^2.$$

Note that the complete transition matrix of the Markov chain $P = (P_{ABB})^\infty$, which is equivalent to playing the game in a repeating pattern of ABBABBABB... in the long run. To this end, we will compute the stationary distribution of P_{ABB} which will indicate whether playing the game in this pattern is winning.

We therefore use P_A and P_B as previously introduced in (b) and (d) and compute P_B^2 :

$$\begin{aligned}
 P_B^2 &= P_B \cdot P_B \\
 &= \begin{pmatrix} 0.69845 & 0.230775 & 0.070775 \\ 0.555025 & 0.2142 & 0.230775 \\ 0.065025 & 0.070775 & 0.8642 \end{pmatrix}.
 \end{aligned}$$

Finally, we can compute P_{ABB} :

$$\begin{aligned}
 P_{ABB} &= P_A \cdot P_B^2 \\
 &= \begin{pmatrix} 0.307575 & 0.14177038 & 0.55065462 \\ 0.38490463 & 0.151575 & 0.46352037 \\ 0.62602038 & 0.22240463 & 0.151575 \end{pmatrix}.
 \end{aligned}$$

We now compute the stationary distribution π of P_{ABB} . We, again, reuse the same derivation process as in (b) and may therefore skip some steps for brevity.

²We could also summarize Equation 9 by clusting π_1 and π_2 into a single term as done in the assignment sheet. However, the formulation in Equation 9 makes the derivation more transparent.

$$\begin{aligned}
 \pi_0 &= \pi_0 \cdot 0.307575 + \pi_1 \cdot 0.38490463 + \pi_2 \cdot 0.62602038 \\
 \pi_0 \cdot (1 - 0.307575) &= \pi_1 \cdot 0.38490463 + \pi_2 \cdot 0.62602038 \\
 \pi_0 &= \pi_1 \cdot \frac{0.38490463}{0.692425} + \pi_2 \cdot \frac{0.62602038}{0.692425} \\
 \pi_0 &= \pi_1 \cdot 0.5558791638 + \pi_2 \cdot 0.9040984655
 \end{aligned} \tag{10}$$

and

$$\begin{aligned}
 \pi_1 &= \pi_0 \cdot 0.14177038 + \pi_1 \cdot 0.151575 + \pi_2 \cdot 0.22240463 \\
 \pi_1 \cdot (1 - 0.151575) &= \pi_0 \cdot 0.14177038 + \pi_2 \cdot 0.22240463 \\
 \pi_1 &= \pi_0 \cdot 0.1670983057 + \pi_2 \cdot 0.2621382326 \\
 \pi_1 &\stackrel{(10)}{=} (\pi_1 \cdot 0.5558791638 + \pi_2 \cdot 0.9040984655) \cdot 0.1670983057 + \pi_2 \cdot 0.2621382326 \\
 \pi_1 &= \pi_2 \cdot 0.455523525 .
 \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 1 &= \pi_0 + \pi_1 + \pi_2 \\
 1 &= (\pi_1 \cdot 0.5558791638 + \pi_2 \cdot 0.9040984655) + \pi_1 + \pi_2 \\
 1 &= \pi_1 \cdot 1.5558791638 + \pi_2 \cdot 1.9040984655 \\
 1 &= \pi_2 \cdot 0.455523525 \cdot 1.5558791638 + \pi_2 \cdot 1.9040984655 \\
 1 &= \pi_2 \cdot 2.6128380267 \\
 \pi_2 &= \frac{1}{2.6128380267} \\
 \pi_2 &= 0.3827255994 .
 \end{aligned}$$

Furthermore,

$$\begin{aligned}
 \pi_1 &= \pi_2 \cdot 0.455523525 \\
 &= 0.3827255994 \cdot 0.455523525 \\
 &= 0.1743405141
 \end{aligned}$$

and

$$\begin{aligned}
 \pi_0 &= \pi_1 \cdot 0.5558791638 + \pi_2 \cdot 0.9040984655 \\
 &= 0.1743405141 \cdot 0.5558791638 + 0.3827255994 \cdot 0.9040984655 \\
 &= 0.4429338863 .
 \end{aligned}$$

Lastly, we compute the expected gain when playing this repeating pattern in the long run. As P now represents three rounds instead of one, we need to compute the values for A_-, A_+, \dots .

From this, we can compute the expected gain for a playing a round of game A in the long run. Note that we can read A_+, A_-, \dots directly from P_A .

$$\begin{aligned} E^{(A)} &= \pi_0 \cdot (A_+ - A_-) + \pi_1 \cdot (B_+ - B_-) + \pi_2 \cdot (C_+ - C_-) \\ &= \pi_0 \cdot (0.495 - 0.505) + \pi_1 \cdot (0.495 - 0.505) + \pi_2 \cdot (0.495 - 0.505) \\ &= (0.495 - 0.505) \cdot (\pi_0 + \pi_1 + \pi_2) \\ &= (-0.01) \cdot 1 \\ &= -0.01 . \end{aligned}$$

For computing the average reward for playing the first round of game B after game A, we compute the probability of being in each state after having played game A:

$$\begin{aligned} \pi^{(A)} &= \pi \cdot P_A \\ &= \begin{pmatrix} 0.4429338863 & 0.1743405141 & 0.3827255994 \end{pmatrix} \cdot \begin{pmatrix} 0.0 & 0.495 & 0.505 \\ 0.505 & 0.0 & 0.495 \\ 0.495 & 0.505 & 0.0 \end{pmatrix} \\ &= \begin{pmatrix} 0.27749113 & 0.4125287 & 0.30998017 \end{pmatrix} . \end{aligned}$$

The expected gain is

$$\begin{aligned} E^{(AB)} &= \pi_0^{(A)} \cdot (A_+ - A_-) + \pi_1^{(A)} \cdot (B_+ - B_-) + \pi_2^{(A)} \cdot (C_+ - C_-) \\ &= \pi_0^{(A)} \cdot (0.095 - 0.905) + \pi_1^{(A)} \cdot (0.745 - 0.255) + \pi_2^{(A)} \cdot (0.745 - 0.255) \\ &= 0.27749113 \cdot (-0.81) + 0.4125287 \cdot 0.49 + 0.30998017 \cdot 0.49 \\ &= 0.129261531 . \end{aligned}$$

Similarly, we can compute the expected gain for playing a round of game B after playing a round of game A.

$$\begin{aligned} \pi^{(AB)} &= \pi^{(A)} \cdot P_B \\ &= \begin{pmatrix} 0.33613004 & 0.1054066 & 0.55846336 \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} E^{(ABB)} &= \pi_0^{(AB)} \cdot (A_+ - A_-) + \pi_1^{(AB)} \cdot (B_+ - B_-) + \pi_2^{(AB)} \cdot (C_+ - C_-) \\ &= 0.33613004 \cdot (-0.81) + 0.1054066 \cdot 0.49 + 0.55846336 \cdot 0.49 \\ &= 0.053030948 . \end{aligned}$$

To total expected reward for playing all three rounds can be computed as follows:

$$\begin{aligned} E &= E^{(A)} + E^{(AB)} + E^{(ABB)} \\ &= (-0.01) + 0.129261531 + 0.053030948 \\ &= 0.172292479 . \end{aligned}$$

this shows that playing all three rounds in this specific order yields a positive expected reward, and therefore is winning in the long run.

(Optional)

We can verify that $\pi^{(AB)}$ is valid by computing the probability of each state after the second round of game B:

$$\begin{aligned} \pi^{(ABB)} &= \pi^{(AB)} \cdot P_B \\ &= \begin{pmatrix} 0.44293388 & 0.17434051 & 0.38272561 \end{pmatrix} . \end{aligned}$$

We remark that $\pi^{(ABB)} = \pi$, which is expected, as we just computed $\pi \cdot P$ step by step.

Statement on the Useage of Gen. AI

I used ChatGPT to help find literature on the properties of integration, differentiation. Otherwise, all code and text has been written entirely by me.