Markov Spring 2025 Homework 3

Due Feb. 6 on Gradescope

Plotting and reporting.

- **Document.** Submit a single document that includes all your figures.
- **Figures.** Include clean, labeled plots (axes, units, parameters) with consistent limits across panels. No photos of screens are accepted, although clear screenshots embedded in or appended to your document are Ok. Calculate key results such as empirical means, variances, and probabilities with 3–4 significant digits.
- **Code.** You don't need to include code in your written homework. Instead, create a github account if you don't already have one, create a public repository labeled "Markov2026", and upload your code there, ready to run. Label the code as "HW1_problem4.m", for example. Include the name of your account and repository in the homework. You can use any programming language you want (excel is strongly discouraged, though).
- **Math.** Accompany your calculations with clear and complete sentences. Discuss your results.

Problems

1. **Acceptance–Rejection with Optimized Proposal.** Let X be a random variable with density

$$f(x) = \frac{1}{3}x(1+x)e^{-x}, \qquad x \geq 0.$$

Suppose we want to simulate X using acceptance-rejection using the proposal family

$$g_a(x) = a^2 x e^{-ax}, \qquad x \geq 0,$$

which is a Gamma$(2, a)$ density with rate parameter $a > 0$. The goal is to choose the a that results in the highest acceptance rate.

(a) Let

$$c(a) = \sup_{x \geq 0} \frac{f(x)}{g_a(x)}.$$

Show that $c(a)$ is finite only for $0 < a < 1$.

(b) For $0 < a < 1$, find the value $x^*(a)$ at which

$$h_a(x) = \frac{f(x)}{g_a(x)}$$

is maximized, and derive a closed-form expression for $c(a)$.

(c) Find the value a^* that maximizes the acceptance probability of the algorithm, and compute this acceptance probability.

(d) Plot $f(x)$ and $c(a^*)g_{a^*}(x)$ in the same plot.

(e) Explain how would you efficiently sample from $g_a(x)$.

(f) **[For APPM 5560/STAT 5100 students only.]** Explain how would you efficiently sample from $f(x)$ treating it as a mixture (therefore making this problem unnecessary - I chose it because one can do the algebra nicely...).

2. **Cell cycle.** A cancerous cell can be in the states "Growing" (G), "Synthesizing DNA" (S), and "Dividing" (D). Modeling the cell cycle as a Markov Chain, we assume that every hour the cell can stochastically make the following transitions: if it is on state G, it can switch to state S with probability 1/10 and remain in state G otherwise; if it is in state S, it can switch to state D with probability 1/8 and remain in S otherwise; finally, if it is in state D, it transitions to G with probability 2/5 and remains in D otherwise. (Once the cell divides, we just look at one of the two daughter cells and ignore the other, so the cycle keeps repeating.)

   (a) Write down the probability transition matrix p of the Markov Chain.

   (b) By treating exiting from each state as a geometric random variable, find the expected time that the cell remains in the growing state.

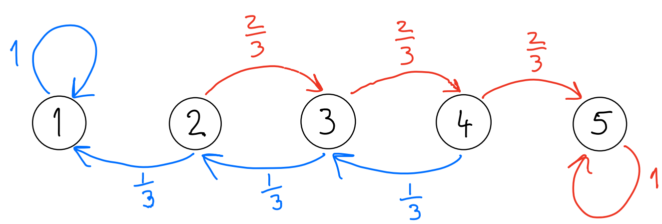
   (c) Find the expected time until a cell completes one cycle, i.e., the time it takes to return to G after starting at G.

   (d) Combining (b) and (c), find the expected fraction of time the cell spends in the G state.

   (e) Calculate p^{50} numerically and explain the relationship between the values you see in this matrix and your answer to (d).

   (f) Simulate the Markov process starting from the state G for 10000 steps. Count the fraction of time that your Chain is in state G and compare with (d).

3. **Markov Chain 1.** Consider a Markov chain with state space $S = 1, 2, 3, 4, 5$. If $X_n = i$ with $i = 2, 3, 4$, then $X_{n+1} = i + 1$ with probability 2/3 and $i - 1$ with probability 1/3. If $X_n = i$ with $i = 1, 5$ then $X_{n+1} = i$ with probability 1 (i.e., the states 1 and 5 are absorbing). This is illustrated in the diagram above.



   (a) Find the probability that $X_5 = 5$ given that $X_0 = 2$ by finding all possible paths for going from state 2 to state 5 in five steps and adding their probabilities.

   (b) Write down the transition probability matrix p of the Markov Chain.

   (c) Compute p^4 and compare your result from (a) with the relevant entry of this matrix.

   (d) Find the probability that state 1 is reached *without visiting state* 4 if $X_0 = 3$.

4. **Markov Chain 2.** Consider the Markov chain with the probability transition matrix between states $\{1, 2, 3, 4, 5, 6\}$ given by

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

   (a) Classify the states as recurrent or transient, and identify the communicating classes of the Markov Chain (no proofs necessary).

(b) Find the probability that $X_5 = 4$ if $X_0 = 1$.

(c) Simulate the Markov Chain starting from $X_0 = 1$ for five time steps 10000 times. Calculate the fraction of these simulations in which $X_5 = 4$ and compare with (b).

(d) Let $\rho_{ii} = P(\text{return to } i \text{ starting from } i)$. Calculate ρ_{33}, ρ_{11}, and ρ_{66}.

(e) Let $\rho_{ij} = P(\text{eventually going to } j \text{ starting from } i)$. Calculate ρ_{14} and ρ_{51}.

5. **Markov Chain 3.** A fair coin is flipped in an ongoing manner, producing outcomes $Y_0, Y_1, Y_2, \ldots$ where each Y_i is either 0 or 1, each with a probability of $\frac{1}{2}$. Consider $X_n = Y_n + Y_{n-1}$ to be the sum of 1's in the n-th and $(n-1)$-the flips for $n \geq 1$. Is the sequence $X_1, X_2, X_3, \ldots$ a Markov chain? Justify your answer mathematically!

6. **Markov Chain 4.** **[Required for APPM 5560/STAT 5100 Students Only.]** Suppose for $n = 2, 3, \ldots$ a fair coin is flipped, obtaining $Y_n \in \{0, 1\}$ with probability 1/2 each. The we define $X_0 = 1$, $X_1 = 1$, and $X_{n+1} = X_{n-1} Y_n$.

   (a) Show that $\{X_n\}_{n=1}^{\infty}$ is not a Markov Chain.

   (b) [Optional +5 points - all or nothing] Let $Z_n = (X_n, X_{n+1})$. Show that $\{Z_n\}_{n=1}^{\infty}$ is a Markov Chain. Specify the state space S and the transition probability matrix.

END