STAT516: Stochastic Modeling of Scientific Data I	Fall 2021

Midterm

Due November 9th by 11:59pm

Instructor: Vincent Roulet *Grader: Alex Jiang*

Upload your answers to the following question on Gradescope in a pdf.
Make a new page for each exercise that will facilitate the grading.

Exercise 1. A factory has a tank for storing waste. Each week the factory produces i cubic meters of waste with probability $q_i = 1/3$ for $i = 0, 1, 2$. The production of waste occurs independently of the amount of waste already in the tank. If the tank contains more than 1 cubic meter by the end of the week, the tank is emptied at the beginning of the following week; otherwise no waste is removed. Let X_n denote the amount of waste in the tank at the end of week n.

1. Explain why $(X_n)_{n\geq 0}$ is a homogeneous Markov chain on $S = \{0, 1, 2, 3\}$.
2. Write down the transition probability matrix $\mathbf{P}$ of $(X_n)_{n\geq 0}$.
3. Show that $(X_n)_{n\geq 0}$ has a unique stationary distribution π.
4. Compute π.
5. Does the sequence $\mathbf{P}^n$ converge for $n \to \infty$? Explain. If it exists, what is the limit.
6. The factory is charged \$75 per cubic meter for the waste removal when the tank holds 2 cubic meters. If it holds 3 cubic meters then the price doubles to \$150 per cubic meter. What is the long-run average weekly cost of waste removal for the factory?

Exercise 2. Trials are performed in sequence. Each trial results into either a failure or a success. If the two most recent previous trials were both successes, the next trial is a success with probability 0.8; otherwise, the chance of success is 0.5. Assume whatever you like about the very first two trials.

1. Define state 1 to be "most recent trial was a failure", state 2 to be "most recent trial was a success, and the preceding trial was a failure" and state 3 to be "last two trials were successes". Let X_n be the state after the n-th trial. Explain why $(X_n)_{n=3}^{\infty}$ is a homogeneous Markov chain, and find its transition matrix $P = (p_{ij})$.
2. Explain why the Markov chain (X_n) is ergodic.
3. Using the chain (X_n), find the long run proportion of trials that are successes.

Exercise 3. Let $(X_n)_{n=1}^{\infty}$ be a sequence of A's and B's formed as follows. The first two letters, X_1 and X_2, are chosen independently and at random, with $\Pr(A) = \Pr(B) = 1/2$. For $n \geq 2$, letter X_{n+1} is selected in a way that depends on previously selected letters.

1. Suppose X_{n+1}, for $n \geq 2$, is selected conditionally on X_n with

$$\Pr(X_{n+1} = A | X_n = A) = 1/2 \quad \text{and} \quad \Pr(X_{n+1} = A | X_n = B) = 1/4.$$

   Find the proportion of A's and B's in a long sequence (X_n).

2. Suppose we instead select X_{n+1}, for $n \geq 2$, conditionally on the preceding two letters with

$$\Pr(X_{n+1} = A | X_n = A, X_{n-1} = A) = \Pr(X_{n+1} = A | X_n = B, X_{n-1} = A) = 1/2,$$
$$\Pr(X_{n+1} = A | X_n = A, X_{n-1} = B) = \Pr(X_{n+1} = A | X_n = B, X_{n-1} = B) = 1/4.$$

   Is $(X_{2n-1})_{n=1}^{\infty}$ a Markov chain? Explain.

3. In the setting from (b), find the proportion of A's and B's in a long sequence.

Exercise 4. Let $(X_n)_{n=0}^{\infty}$ be a homogeneous Markov chain with state space S and a transition matrix $P = (p_{ij})$ that has $p_{ii} < 1$ for all $i \in S$. Let $(Y_n)_{n=0}^{\infty}$ be the sequence of new values of (X_n), and let τ_m be the mth time at which we see a new value of X_n. For instance, if $(X_0, X_1, \ldots, X_{10})$ realizes to $(1, 1, 1, 2, 2, 1, 3, 3, 3, 2, 1)$ then $Y_0 = 1$, $Y_1 = 2$, $Y_2 = 1$, $Y_3 = 3$, $Y_4 = 2$ and $Y_5 = 1$, and $\tau_1 = 3$, $\tau_2 = 5$, $\tau_3 = 6$, $\tau_4 = 9$, $\tau_5 = 10$.

1. Show that all τ_m, $m \geq 1$, are stopping times with respect to (X_n).
2. Name the theorem that implies that (Y_n) is a homogeneous Markov chain. Give the transition matrix Q for (Y_n) in terms of P.
3. Show that if (X_n) is irreducible, then so is (Y_n).
4. Show that if (X_n) is irreducible and recurrent, then so is (Y_n).
5. Suppose (X_n) is irreducible and recurrent with invariant measure $\mathbf{x}$, so $\mathbf{x}^T P = \mathbf{x}^T$. Find an invariant measure $\mathbf{y}$ for the chain (Y_n) in terms of $\mathbf{x}$ and $(p_{ii} : i \in S)$.

Exercise 5 (Optional)**.** Let $(X_n)_{n=0}^{\infty}$ be a homogeneous Markov chain that takes values in state space S with $2 \leq |S| < \infty$. Suppose, throughout all of this problem, that (X_n) is irreducible, and let π be its unique stationary distribution. For a subset $A \subset S$, define $T_1, T_2, \ldots$ to be the times of successive visits to the set A by the chain (X_n), that is,

$$T_1 = \min\{n \geq 1 : X_n \in A\} \quad \text{and} \quad T_{m+1} = \min\{n > T_m : X_n \in A\}.$$

Now define $X_m^A = X_{T_m}$, the state visited on the m-th return to A.

1. Explain why each T_m is a stopping time.
2. State the name of a theorem/result from class that implies that $(X_m^A)_{m=1}^{\infty}$ is a homogeneous Markov chain. It is called the *induced chain* on A.
3. Find a stationary distribution π_A for the induced chain. Is π_A unique?
4. Give an example with $S = \{1, 2, 3\}$ in which the original chain is aperiodic but the induced chain is not.
5. Show that if the original chain (X_n) is reversible and aperiodic, then the induced chain (X_m^A) is aperiodic.