Midterm Exam

November 1, 2021

Instructions:

- This is an open book, open notes exam.
- Calculators are not needed; laptops, tablets and cell-phones are not allowed.
- Perfect score: 100 points.
- Duration: 90 minutes.
- This exam has 12 numbered pages, check now that all pages are present.
- Make sure you write your name in the space provided below.
- Show all your work, and write your final answers in the boxes when provided.

Name:_____

| Problem | Max. Points | Score | Problem | Max. Points | Score |
|---------|-------------|-------|---------|-------------|-------|
| 1. | 24 | | 5. | 12 | |
| 2. | 10 | | 6. | 10 | |
| 3. | 12 | | 7. | 8 | |
| 4. | 12 | | 8. | 12 | |
| | | | Total | 100 | |

GOOD LUCK!

1. Consider a Markov chain $X_{\mathbb{N}} = X_0, X_1, \dots, X_n, \dots$ with state space $S = \{1, 2, 3, 4, 5, 6, 7\}$, state transition diagram



and initial distribution $P(X_0 = 1) = 1$ and $P(X_0 = i) = 0$ for $2 \le i \le 7$.

(a) (2 points) $P(X_5 = 4, X_4 = 3, X_3 = 2, X_2 = 3, X_1 = 2, X_0 = 1) = ?$

(b) (6 points) Specify the communication classes and determine whether they are transient or recurrent.

(c) (6 points) What is the period of state 6?

(d) (4 points) $\lim_{n\to\infty} P_{11}^n = ?$

(e) (6 points) Calculate

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=2}^{n} \mathbb{I}\left\{X_i = 7 \mid X_1 = 5\right\}$$

and provide justification for the existence of the limit.

2. (10 points) Consider i.i.d. continuous random variables X_1, \ldots, X_{10} with probability density function

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Find the approximate probability that $Y = \sum_{i=1}^{10} X_i$ will exceed 7. Write your result in terms of the complementary cumulative distribution function Φ of a standard Normal random variable, that is for $Z \sim \mathcal{N}(0, 1)$ then

$$\Phi(z) = \mathbf{P}\left(Z \ge z\right) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-u^2/2} du.$$



3. Draw a county at random from the United States. Then draw n people at random from that county. Let $0 \le X \le n$ be the number of those people who are infected with COVID-19. If Q denotes the proportion of people in the county with the virus, then Q is also a random variable since it varies from county to county. Given Q = q, we have that $X \sim \text{Binomial}(n, q)$. Also, suppose that the random variable Q is uniformly distributed in the interval [0, 1]. The distribution of X is thus constructed in two steps, leading to a so-termed hierarchical model that we write

 $Q \sim \text{Uniform}[0, 1]$

$$X \mid Q = q \sim \operatorname{Binomial}(n, q).$$

(a) (2 points) $\mathbb{E} [X | Q = q] = ?$



(b) (4 points) $\mathbb{E}[X] = ?$

(c) (6 points) var [X] = ?

4. Suppose that $X_{\mathbb{N}} = X_1, X_2, \ldots, X_n, \ldots$ is an i.i.d. sequence of random variables, which are uniformly distributed in the interval [0, 1].

(a) (8 points) Define the random variable

$$Y = \min\{X_1, X_2\}.$$

Write down an expression for $f_Y(y)$, the probability density function of Y. [Hint: it might be easier to first compute P(Y > y).]

(b) (4 points) Let $Y_{\mathbb{N}} = Y_1, Y_2, \dots, Y_n, \dots$ be the sequence of random variables given by $Y_n = \min\{X_1, \dots, X_n\}, \quad n \ge 1.$

Show that Y_n converges in probability to 0 as $n \to \infty$.

5. Suppose that $X_{\mathbb{N}} = X_0, X_1, \ldots, X_n, \ldots$ is a Markov chain with state space $S = \{1, 2\}$, transition probability matrix

$$\mathbf{P} = \left(\begin{array}{cc} 1-a & a \\ b & 1-b \end{array}\right),$$

where 0 < a < 1 and 0 < b < 1. We define the recurrence time of state $i \in S$ as

$$T_i = \min\{n > 0 : X_n = i\}$$
 given that $X_0 = i$.

Accordingly, T_i is a discrete random variable taking values on the integers $\{1, 2, 3, \ldots\}$.

(a) (6 points) Compute $p_{T_1}(n) = \mathbf{P}(T_1 = n | X_0 = 1)$, the probability mass function of T_1 .



(b) (6 points) $\mathbb{E}\left[T_1 \mid X_0 = 1\right] =$? [Reminder: for your calculations, it may be useful to recall the sum of the geometric series $\sum_{r=1}^{\infty} \alpha^{r-1} = 1/(1-\alpha)$, for $0 < \alpha < 1$.]

6. Suppose that we want to evaluate the integral

$$I = \int_{a}^{b} f(x) dx$$

for some integrable function f. Unlike polynomial, rational or trigonometric functions, if f is complicated then there may be no known closed form expression for I. In these cases, numerical integration methods are appropriate to approximate the value of I.

Here we will explore the simplest version of Monte Carlo integration. Start by writing

$$I = \int_{a}^{b} f(x)dx = \int_{a}^{b} w(x)g(x)dx,$$

where w(x) = f(x)(b - a) and g(x) = 1/(b - a).

(a) (4 points) Show that $I = \mathbb{E}[w(X)]$, where X is a random variable. Specify the distribution of X.

(b) (6 points) Suppose that you can generate N i.i.d. samples from the distribution of X. Describe a method to estimate the value of I, and state any result you use to justify your approximation.

7. (8 points) Consider a random variable X with cumulative distribution function $F_X(x) = P(X \le x)$ given in the following figure.



Sketch $F_Y(y) = \mathbf{P}(Y \le y)$, the cumulative distribution function of $Y = \max\{0, X\}$.



8. (12 points) Suppose that $X_{\mathbb{N}} = X_0, X_1, \ldots, X_n, \ldots$ is a Markov chain with state space $S = \{1, 2, 3, 4, 5\}$ and transition probability matrix

$$\mathbf{P} = \begin{pmatrix} q & p & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ q & 0 & 0 & p & 0 \\ q & 0 & 0 & 0 & p \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Let 0 and <math>q = 1 - p.

Determine the stationary distribution of $X_{\mathbb{N}}$. [Reminder: for your calculations, it might useful to recall the partial geometric sum $\sum_{r=0}^{k} \alpha^r = \frac{1-\alpha^{k+1}}{1-\alpha}$, for $\alpha \neq 1$.]