

ECE440 - Introduction to Random Processes

Midterm Exam

October 23, 2020

Instructions:

- This is an **individual** take-home exam, **collaborations are not allowed**.
- Write clearly and show all your work.
- Your solutions should be submitted via Gradescope as a single pdf file.
- The estimated amount of time required to complete this exam is 2.5 hours.
- **The submission deadline is 10 pm ET, Friday October 23, 2020.**
- Late submissions will not be accepted.
- Perfect score: 100 points.
- This exam has 12 numbered pages.

Name: _____ **SOLUTIONS** _____

Problem	Max. Points	Score	Problem	Max. Points	Score
1.	18		6.	10	
2.	10		7.	10	
3.	10		8.	8	
4.	8		9.	16	
5.	10				
			Total	100	

GOOD LUCK!

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

$$\mathbf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

and initial distribution $\mathbf{P}(X_0 = 1) = 1$ and $\mathbf{P}(X_0 = 2) = 0$. Unless otherwise stated, suppose that $0 < a < 1$ and $0 < b < 1$.

(a) (1 points) $\mathbf{P}(X_5 = 2 \mid X_4 = 1, X_3 = 2, X_1 = 1) = ?$

a

From the Markov property it follows that

$$\mathbf{P}(X_5 = 2 \mid X_4 = 1, X_3 = 2, X_1 = 1) = \mathbf{P}(X_5 = 2 \mid X_4 = 1) = P_{12} = a.$$

(b) (2 points) $\mathbf{P}(X_3 = 2, X_2 = 2 \mid X_1 = 1) = ?$

$a(1-b)$

Using the definition of conditionally probability and the Markov property one finds

$$\begin{aligned} \mathbf{P}(X_3 = 2, X_2 = 2 \mid X_1 = 1) &= \mathbf{P}(X_3 = 2 \mid X_2 = 2, X_1 = 1) \mathbf{P}(X_2 = 2 \mid X_1 = 1) \\ &= \mathbf{P}(X_3 = 2 \mid X_2 = 2) \mathbf{P}(X_2 = 2 \mid X_1 = 1) \\ &= P_{22} \times P_{12} = a(1-b). \end{aligned}$$

(c) (3 points) $\mathbb{E}[X_1] = ?$

$1+a$

To obtain the unconditional pmf of X_1 , we use the law of total probability and find

$$\mathbf{P}(X_1 = 1) = \sum_{i=1}^2 \mathbf{P}(X_1 = 1 \mid X_0 = i) \mathbf{P}(X_0 = i) = P_{11}\mathbf{P}(X_0 = 1) + P_{21}\mathbf{P}(X_0 = 2) = P_{11} \times 1 = 1-a.$$

Hence, $\mathbf{P}(X_1 = 2) = a$. So it follows the expectation is $\mathbb{E}[X_1] = 1 \times (1-a) + 2 \times a = 1+a$.

(d) (8 points) Prove that

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$$

and provide justification for the existence of the limit.

The Markov chain is ergodic for $0 < a < 1$ and $0 < b < 1$. Indeed, both states communicate so it is irreducible. Because the state space S is finite, the single class has to be positive recurrent. State 1 (hence state 2) is aperiodic because $P_{11} = 1 - a > 0$.

For ergodic Markov chains the limiting probabilities P_{ij}^n (i.e., the entries of matrix \mathbf{P}^n) converge as $n \rightarrow \infty$. Both rows of the limiting matrix should be identical because $\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j$, regardless of the initial condition i . The unique stationary distribution $\boldsymbol{\pi} = [\pi_1, \pi_2]^T$ satisfies

$$\begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix} \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix}, \quad \pi_1 + \pi_2 = 1.$$

Solving the linear system yields $\boldsymbol{\pi} = \left[\frac{b}{a+b}, \frac{a}{a+b} \right]^T$.

(e) (2 points) From now on, suppose that $a = b = 1$. $\mathbf{P}(X_{26} = 1 \mid X_1 = 2) = ?$

1

For $a = b = 1$, the Markov chain deterministically cycles around states 1 and 2. Both states have period 2. Given $X_1 = 2$, then $P_{21}^{2n} = 0$ and $P_{21}^{2n-1} = 1$ for all $n \geq 1$. Hence $\mathbf{P}(X_{26} = 1 \mid X_1 = 2) = P_{21}^{25} = 1$.

(f) (2 points) Still with $a = b = 1$, calculate

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \mathbb{I}\{X_i = 2\}$$

and provide justification for the existence of the limit.

$\frac{1}{2}$

Even if the Markov chain is not ergodic (recall the states have period 2), the ergodic limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n \mathbb{I}\{X_i = 2\} = \pi_2 = \frac{1}{2} \quad \text{a. s.,}$$

because the process spends exactly half of the time in each state.

2. Consider a probability space $(S, \mathcal{F}, \mathbf{P}(\cdot))$.

(a) (4 points) Let $E \in \mathcal{F}$ be an event. Show that if E is independent of itself the $\mathbf{P}(E)$ is either 0 or 1.

If E is independent of itself, then $\mathbf{P}(E \cap E) = \mathbf{P}(E) \times \mathbf{P}(E)$. Moreover, $E \cap E = E$ so one has

$$\mathbf{P}(E) = \mathbf{P}(E)^2.$$

The solutions to this quadratic equation are $\mathbf{P}(E) = 0$ and $\mathbf{P}(E) = 1$.

(b) (6 points) Suppose that $A, B \in \mathcal{F}$ are independent events. Show that A^c and B^c are independent events.

If A and B are independent events, then $\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B)$. It thus follows that

$$\begin{aligned} \mathbf{P}(A^c \cap B^c) &= 1 - \mathbf{P}(A \cup B) \\ &= 1 - \mathbf{P}(A) - \mathbf{P}(B) + \mathbf{P}(A \cap B) \\ &= 1 - \mathbf{P}(A) - \mathbf{P}(B) + \mathbf{P}(A) \times \mathbf{P}(B) \\ &= (1 - \mathbf{P}(A))(1 - \mathbf{P}(B)) = \mathbf{P}(A^c) \times \mathbf{P}(B^c), \end{aligned}$$

which implies A^c and B^c are independent events as desired. In deriving the first equality we used that $A^c \cap B^c = (A \cup B)^c$, for the second that $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$, and finally the third equality follows from the assumed independence of A and B .

3. (a) (5 points) Let X and Y be random variables. Show that if $\mathbb{E}[X | Y = y] = c$ for some deterministic constant c , then X and Y are uncorrelated.

The covariance between X and Y is given by $\text{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]$. Evaluating the first expectation by conditioning on Y yields

$$\begin{aligned} \mathbb{E}[XY] &= \mathbb{E}_Y [\mathbb{E}_X [XY | Y]] \\ &= \sum_y \mathbb{E}_X [XY | Y = y] \mathbf{P}(Y = y) \\ &= \sum_y y \mathbb{E}_X [X | Y = y] \mathbf{P}(Y = y) \\ &= \sum_y cy \mathbf{P}(Y = y) = c \mathbb{E}[Y]. \end{aligned}$$

The fourth equality follows from the assumption $\mathbb{E}[X | Y = y] = c$. A similar argument yields

$$\mathbb{E}[X] = \mathbb{E}_Y [\mathbb{E}_X [X | Y]] = \sum_y \mathbb{E}_X [X | Y = y] \mathbf{P}(Y = y) = \sum_y c \mathbf{P}(Y = y) = c.$$

Putting all the pieces together, we arrive at

$$\text{cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y] = c \mathbb{E}[Y] - c \mathbb{E}[Y] = 0.$$

The conclusion is that X and Y are uncorrelated. Naturally, the same result holds for continuous X and Y – they choice to argue in the discrete case was arbitrary.

(b) (5 points) Let U and V be random variables. Suppose that $\mathbb{E}[V | U] = U$. Show that $\text{cov}[U, V] = \text{var}[U]$.

We argue as in (a), starting with

$$\mathbb{E}[UV] = \mathbb{E}_U [\mathbb{E}_V [UV | U]] = \mathbb{E}_U [U \mathbb{E}_V [V | U]] = \mathbb{E}[U^2].$$

In deriving the last equality we used the assumed identity $\mathbb{E}[V | U] = U$. The mean of V is

$$\mathbb{E}[V] = \mathbb{E}_U [\mathbb{E}_V [V | U]] = \mathbb{E}[U].$$

All in all, the covariance simplifies to

$$\text{cov}[U, V] = \mathbb{E}[UV] - \mathbb{E}[U] \mathbb{E}[V] = \mathbb{E}[U^2] - \mathbb{E}[U]^2 = \text{var}[U]$$

as we wanted to show.

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

(a) (4 points) For fixed $n > 1$, consider the random variable

$$Y_n = \sum_{i=1}^n \mathbb{I}\{X_i > 0.3\}.$$

Write down an expression for $\mathbb{P}(Y_n = y)$, the probability mass function of Y_n . Make sure you also specify the range of values of y for which $\mathbb{P}(Y_n = y) = 0$.

$$\mathbb{P}(Y_n = y) = \begin{cases} \binom{n}{y} 0.7^y 0.3^{n-y}, & 0 \leq y \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

For $1 \leq i \leq n$, the summands $\mathbb{I}\{X_i > 0.3\}$ are i.i.d. Bernoulli-distributed random variables with parameter $p = \mathbb{P}(X_1 > 0.3) = \mathbb{P}(X_1 \in [0.3, 1]) = 0.7$ (recall that the X_i are uniformly distributed in $[0, 1]$). Because the sum of n i.i.d. Bernoulli random variables with parameter p is Binomial distributed with parameters (n, p) , then it follows that Y_n is Binomial with parameters $(n, 0.7)$. The pmf is

$$\mathbb{P}(Y_n = y) = \begin{cases} \binom{n}{y} 0.7^y 0.3^{n-y}, & 0 \leq y \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

(b) (4 points) Suppose that a, b are deterministic constants such that $a < b$. Using the Central Limit Theorem, write down an approximate expression for $\mathbb{P}(a \leq Y_{1000} \leq b)$.

$$\mathbb{P}(a < Y_{1000} < b) \approx \frac{1}{\sqrt{\pi \times 420}} \int_a^b e^{-\frac{(y-700)^2}{420}} dy$$

We have $\mathbb{E}[Y_{1000}] = 1000 \times 0.7 = 700$ and $\text{var}[Y_{1000}] = 1000 \times 0.7 \times 0.3 = 210$. By virtue of the Central Limit Theorem, we can approximate the distribution of Y_{1000} with that of a Normal, namely

$$Y_{1000} \sim \mathcal{N}(700, 210).$$

The desired probability can thus be approximated as

$$\mathbf{P}(a < Y_{1000} < b) \approx \frac{1}{\sqrt{\pi \times 420}} \int_a^b e^{-\frac{(y-700)^2}{420}} dy.$$

5. (10 points) Consider a model for the evolution of a population and suppose that X_n is the number of individuals in generation n . Suppose the k -th individual in generation n creates $Q_{k,n+1}$ individuals in generation $n+1$, and that the $Q_{k,n}$ are i.i.d. across individuals and generations, and independent of X_0 . Let $\mu = \mathbb{E}[Q_{k,n}]$ and $\sigma^2 = \text{var}[Q_{k,n}]$. Under the preceding assumptions, $X_{\mathbb{N}} = X_0, X_1, \dots, X_n, \dots$ is a Markov chain with state space $S = \{0, 1, 2, \dots\}$ for which

$$X_{n+1} = Q_{1,n+1} + \dots + Q_{X_n,n+1} \quad \text{if } X_n > 0,$$

and $X_{n+1} = 0$ if $X_n = 0$. Let $M_n = \mathbb{E}[X_n]$ and $V_n = \text{var}[X_n]$.

Derive an expression for V_{n+1} in terms of V_n , M_n , μ and σ^2 .

$$V_{n+1} = \sigma^2 M_n + \mu^2 V_n$$

The number of individuals in generation $n+1$ is given by the compound random variable

$$X_{n+1} = \sum_{k=1}^{X_n} Q_{k,n+1}.$$

To compute $V_{n+1} = \text{var}[X_{n+1}]$ we condition on X_n . Because the $Q_{k,n+1}$ are i.i.d. we find that $\mathbb{E}[X_{n+1} | X_n] = \mu X_n$ and $\text{var}[X_{n+1} | X_n] = \sigma^2 X_n$. Using the conditional variance formula

$$\begin{aligned} \text{var}[X_{n+1}] &= \mathbb{E}[\text{var}[X_{n+1} | X_n]] + \text{var}[\mathbb{E}[X_{n+1} | X_n]] \\ &= \mathbb{E}[\sigma^2 X_n] + \text{var}[\mu X_n] \\ &= \sigma^2 \mathbb{E}[X_n] + \mu^2 \text{var}[X_n] = \sigma^2 M_n + \mu^2 V_n. \end{aligned}$$

6. (10 points) Consider a continuous random variable X with probability density function

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

Let $Y = 0$ if $X < 1/2$ and $Y = 2X$ otherwise. Compute $F_Y(y) = \mathbf{P}(Y \leq y)$, the cumulative distribution function of Y .

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{4}, & 0 \leq y < 1, \\ \frac{y^2}{4}, & 1 \leq y < 2, \\ 1, & y \geq 2. \end{cases}$$

For $y < 0$, then $F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(2X \leq y) = 0$ because $f_X(x) = 0$ for $x \leq 0$.

For $0 \leq y < 1$, then

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(Y = 0) = \mathbf{P}(X < 1/2) = \int_0^{1/2} 2x dx = \frac{1}{4}.$$

For $1 \leq y < 2$ we find

$$\begin{aligned} F_Y(y) &= \mathbf{P}(Y \leq y) = \mathbf{P}(Y = 0) + \mathbf{P}(1 \leq Y \leq y) \\ &= \frac{1}{4} + \mathbf{P}(1 \leq 2X \leq y) \\ &= \frac{1}{4} + \mathbf{P}(1/2 \leq X \leq y/2) \\ &= \frac{1}{4} + \int_{1/2}^{y/2} 2x dx = \frac{y^2}{4}. \end{aligned}$$

Finally, $F_Y(y) = 1$ for $y \geq 2$.

All in all, putting these pieces together yields the desired cdf

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{4}, & 0 \leq y < 1, \\ \frac{y^2}{4}, & 1 \leq y < 2, \\ 1, & y \geq 2. \end{cases}$$

7. (10 points) Suppose that $X_{\mathbb{N}} = X_1, X_2, \dots, X_n, \dots$ is an i.i.d. sequence of random variables such that

$$\begin{aligned} \mathbf{P}(X_1 = 1) &= 4/10, & \mathbf{P}(X_1 = 2) &= 1/10, \\ \mathbf{P}(X_1 = 3) &= 3/10, & \mathbf{P}(X_1 = 4) &= 2/10. \end{aligned}$$

Define

$$T = \min \{n \geq 1 : X_n \in \{1, 2\}\}.$$

$$\mathbb{E} [X_1^2 + X_2^2 + X_5^2 \mid T = 5] = ?$$

$$\frac{126}{5}$$

From the definition of T and the linearity of the expectation operator, then

$$\mathbb{E} [X_1^2 + X_2^2 + X_5^2 \mid T = 5] = 2 \times \mathbb{E} [X_i^2 \mid X_i \in \{3, 4\}] + \mathbb{E} [X_i^2 \mid X_i \in \{1, 2\}].$$

To compute $\mathbb{E}[X_i^2 \mid X_i \in \{3, 4\}]$ the relevant conditional pmf $\mathbb{P}(X_i = x \mid X_i \in \{3, 4\})$ is

$$\mathbb{P}(X_i = 1 \mid X_i \in \{3, 4\}) = 0,$$

$$\mathbb{P}(X_i = 2 \mid X_i \in \{3, 4\}) = 0,$$

$$\mathbb{P}(X_i = 3 \mid X_i \in \{3, 4\}) = \frac{\mathbb{P}(X_i = 3, X_i \in \{3, 4\})}{\mathbb{P}(X_i \in \{3, 4\})} = \frac{\mathbb{P}(X_i = 3)}{\mathbb{P}(X_i = 3) + \mathbb{P}(X_i = 4)} = \frac{3}{5},$$

$$\mathbb{P}(X_i = 4 \mid X_i \in \{3, 4\}) = \frac{\mathbb{P}(X_i = 4, X_i \in \{3, 4\})}{\mathbb{P}(X_i \in \{3, 4\})} = \frac{\mathbb{P}(X_i = 4)}{\mathbb{P}(X_i = 3) + \mathbb{P}(X_i = 4)} = \frac{2}{5}.$$

Hence, $\mathbb{E}[X_i^2 \mid X_i \in \{3, 4\}] = 9 \times \frac{3}{5} + 16 \times \frac{2}{5} = \frac{59}{5}$.

Similarly, to find $\mathbb{E}[X_i^2 \mid X_i \in \{1, 2\}]$ the relevant conditional pmf $\mathbb{P}(X_i = x \mid X_i \in \{1, 2\})$ is

$$\mathbb{P}(X_i = 1 \mid X_i \in \{1, 2\}) = \frac{\mathbb{P}(X_i = 1, X_i \in \{1, 2\})}{\mathbb{P}(X_i \in \{1, 2\})} = \frac{\mathbb{P}(X_i = 1)}{\mathbb{P}(X_i = 1) + \mathbb{P}(X_i = 2)} = \frac{4}{5},$$

$$\mathbb{P}(X_i = 2 \mid X_i \in \{1, 2\}) = \frac{\mathbb{P}(X_i = 2, X_i \in \{1, 2\})}{\mathbb{P}(X_i \in \{1, 2\})} = \frac{\mathbb{P}(X_i = 2)}{\mathbb{P}(X_i = 1) + \mathbb{P}(X_i = 2)} = \frac{1}{5},$$

$$\mathbb{P}(X_i = 3 \mid X_i \in \{1, 2\}) = 0,$$

$$\mathbb{P}(X_i = 4 \mid X_i \in \{1, 2\}) = 0.$$

Hence, $\mathbb{E}[X_i^2 \mid X_i \in \{1, 2\}] = 1 \times \frac{4}{5} + 4 \times \frac{1}{5} = \frac{8}{5}$. The final result is

$$\mathbb{E}[X_1^2 + X_2^2 + X_5^2 \mid T = 5] = 2 \times \frac{59}{5} + \frac{8}{5} = \frac{126}{5}.$$

8. (8 points) Consider flipping a coin for which the probability of heads is $p = 1/2$. Let $X_i \in \{0, 1\}$ denote the outcome of a single toss, and let $X_i = 1$ if said outcome is heads. The fraction of heads after n independent tosses is

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

According to the Weak Law of Large Numbers \bar{X}_n converges to p in probability as $n \rightarrow \infty$.

How large should n be so that $\mathbb{P}(0.4 \leq \bar{X}_n \leq 0.6) \geq 0.7$? [Hint: Use Chebyshev's inequality]

84

The X_i are i.i.d. Bernoulli random variables with parameter $p = 1/2$. Hence, $\mathbb{E}[X_i] = p = 1/2$ and $\text{var}[X_i] = p \times (1 - p) = 1/4$. For the sample mean \bar{X}_n , $\mathbb{E}[\bar{X}_n] = \mathbb{E}[X_i] = 1/2$ and $\text{var}[\bar{X}_n] = \text{var}[X_i]/n = 1/(4n)$.

From Chebyshev's inequality,

$$\begin{aligned} \mathbb{P}(0.4 \leq \bar{X}_n \leq 0.6) &= \mathbb{P}(|\bar{X}_n - \mathbb{E}[\bar{X}_n]| \leq 0.1) \\ &= 1 - \mathbb{P}(|\bar{X}_n - \mathbb{E}[\bar{X}_n]| > 0.1) \\ &\geq 1 - \frac{1}{4n(0.1)^2} = 1 - \frac{25}{n}. \end{aligned}$$

The derived lower bound on the probability of interest will exceed 0.7 if $n = 84$.

9. A fair die is tossed many times in succession. The tosses are independent of each other. Initialize $X_0 = 6$ and for each $n \geq 1$, let X_n denote the minimum among the first n tosses.

(a) (12 points) Show that $X_{\mathbb{N}} = X_0, X_1, \dots, X_n, \dots$ is a Markov chain, specify its state space and determine the transition probability matrix.

Let $Y_{\mathbb{N}} = Y_1, \dots, Y_n, \dots$ denote the i.i.d. outcomes of the die tosses. The distribution of Y_n is uniform over the set $\{1, 2, 3, 4, 5, 6\}$. Since $X_0 = 6$, notice we can write for all $n \geq 1$

$$X_n = \min\{Y_1, \dots, Y_n\} = \min\{X_{n-1}, Y_n\}.$$

So $X_{\mathbb{N}}$ is a Markov chain because we have expressed the state as $X_n = f(X_{n-1}, Y_n)$, where $Y_{\mathbb{N}}$ is an i.i.d. process independent of the initial condition. Moreover, since $Y_n \in \{1, 2, 3, 4, 5, 6\}$ then the state space of $X_{\mathbb{N}}$ is $S = \{1, 2, 3, 4, 5, 6\}$.

The transition probability matrix is

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 5/6 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 2/3 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/2 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/3 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

(b) (4 points) Specify the communication classes and determine whether they are transient or recurrent.

Interestingly, in this Markov chain none of the states communicates with each other. So each state belongs to its own communicating class. There is one recurrent class $\mathcal{R} = \{1\}$ with the absorbing state 1. All other states are transient, yielding five classes $\mathcal{T}_i = \{i+1\}$, $i = 1, \dots, 5$.