Score: ________ / **35**

PSTAT 120A / **MIDTERM EXAM** / Summer 2022 Instructor: **Ethan Marzban**

NAME: ______________________________ **Perm Number:** ________________

SECTION (circle one): 3:30 - 4:20pm (Lucas) 5 - 5:50pm (Moya) 8 - 8:50am (Moya)

Instructions:

- You will have **55 minutes** to complete this exam.
- You are allowed the use of a single **8.5 × 11-inch** sheet, front and back, of notes. You are also permitted the use of **calculators**; the use of any and all other electronic devices (laptops, cell phones, etc.) is prohibited.
- Unless otherwise specified, simplification is not needed; however, all integrals and infinite sums (unless otherwise specified) must be evaluated.
  - One exception is that, whenever applicable, answers may be left in terms of Φ, the standard normal c.d.f..
- **Problem 9(c) is a bonus question; please note that bonus questions will be graded on an all-or-nothing scale.**
- Good Luck!!!

Honor Code: In signing my name below, I certify that all work appearing on this exam is entirely my own and not copied from any external source. I further certify that I have not received any unauthorized aid while taking this exam.

× ______________________________

Multiple Choice Questions:

Question:	1	2	3	4	5	Total
Points:	1	1	1	1	1	5
Score:						

Short-Answer Questions:

Question:	6	7	8	9	Total
Points:	5	8	11	6	30
Score:					

1 Multiple Choice Questions

Please fill in the bubble(s) ***on the exam below*** *corresponding to your answer. You do not need to submit any additional work for these questions.*

1. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and three events $A, B, C \in \mathcal{F}$, which of the following correctly computes $\mathbb{P}(A_1 \cup A_2 \cup A_3)$? [1pts.]

   ◯ $\mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3)$

   √ $1 - \mathbb{P}(A_1^{\complement}) \cdot \mathbb{P}(A_2^{\complement} \mid A_1^{\complement}) \cdot \mathbb{P}(A_3^{\complement} \mid A_1^{\complement} \cap A_2^{\complement})$

   ◯ $\mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3)$

   ◯ $1 - \mathbb{P}(A_1) - \mathbb{P}(A_2) - \mathbb{P}(A_3)$

   ◯ None of the other answer choices

   Solution: This follows from DeMorgan's Law, as well as the multiplication rule for three events:

   $$\begin{aligned}\mathbb{P}(A_1 \cup A_2 \cup A_3) &= 1 - \mathbb{P}(A_1^{\complement} \cap A_2^{\complement} \cap A_3^{\complement}) \\ &= 1 - \mathbb{P}(A_1^{\complement}) \cdot \mathbb{P}(A_2^{\complement} \mid A_1^{\complement}) \cdot \mathbb{P}(A_3^{\complement} \mid A_1^{\complement} \cap A_2^{\complement})\end{aligned}$$

2. Consider a random variable X with p.m.f. given by [1pts.]

k	-2	1
$p_X(k)$	$1/3$	$2/3$

   Which of the following is the correct value of $\mathbb{E}[X]$?

   √ 0

   ◯ 1/3

   ◯ 2/3

   ◯ 1

   ◯ 2

   ◯ None of the above

   Solution: This follows from the definition of expectation:

   $$\mathbb{E}[X] = \sum_k k p_X(k) = (-2)\left(\frac{1}{3}\right) + (1)\left(\frac{2}{3}\right) = 0$$

3. **Fill in the Blanks:** Discrete random variables have state spaces that are __________, whereas continuous random variables have state spaces that are __________. [1pts.]

   ◯ finite; infinite

   ◯ countable; uncountable

   √ **at most countable; uncountable**

   ◯ uncountable; at most countable

   ◯ uncountable; countable

   ◯ None of the above.

Solution: See the definition on Slide 6 of Slide Deck 3.

4. In a bag of 100 marbles, 40 are blue and the remaining 60 are gold. Yaz draws marbles one by one at random, replacing the marble each time. If X denotes the number of marbles (including the final marble) Yaz has to draw before she observes her 3rd blue marble, which of the following accurately describes the distribution of X? [1pts.]

   ◯ Bern(40)

   ◯ Bern(0.4)

   ◯ Bin(3, 0.4)

   ◯ NegBin(40, 0.4)

   √ **NegBin**(3, 0.4)

   ◯ HyperGeom(40, 100, 3)

   ◯ Poisson(0.4)

   ◯ None of the above.

Solution: Since sampling is done with replacement, we have independence across trials. If we define a success to be "Yaz draws a blue marble," then X tracks the number of trials needed to observe 3 successes, meaning $X \sim \text{NegBin}(3, 0.4)$.

5. Which of the following statements is true in general? [1pts.]

   ◯ Pairwise independence implies mutual independence.

   √ **Two mutually dependent events can be conditionally independent**

   ◯ There are 2^n computations needed to establish the mutual independence of n events

   ◯ Pairwise independence is a stronger condition than mutual independence.

   ◯ All of the above answer choices are **false**.

Solution:

- The first statement is incorrect; we saw several examples (in lecture and also on homework) of events that are pairwise independent but not mutually independent
- The second statement is correct.
- The third statement is incorrect; the number of computations is $2^n - n - 1$ (by an Extra Practice Problem on Worksheet 2)
- The fourth statement is simply a rewording of the first statement, which is false.
- Since the second statement is true, the fifth statement is false.

2 Short Answer Questions

Please mark your final answers in the spaces provided below each question. ***Be sure to show all of your work!***

6. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and suppose A and B are two events. Prove the identity [5pts.]

$$\mathbb{P}(A \setminus B) = \mathbb{P}(A) \cdot \mathbb{P}(B^{\complement} \mid A)$$

Solution: Perhaps the simplest proof was to recall that $A \setminus B = (A \cap B^{\complement})$ meaning, by the multiplication rule,

$$\mathbb{P}(A \setminus B) = \mathbb{P}(A \cap B^{\complement}) = \mathbb{P}(B^{\complement} \mid A)\mathbb{P}(A)$$

thereby completing the proof. ■

....

Alternatively, we could have started with the Set Difference Rule, and then used the multiplication rule:

$$\begin{aligned} \mathbb{P}(A \setminus B) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) && \text{[Set Difference Rule]} \\ &= \mathbb{P}(A) - \mathbb{P}(A) \cdot \mathbb{P}(B \mid A) && \text{[Multiplication Rule]} \\ &= \mathbb{P}(A)\left[1 - \mathbb{P}(B \mid A)\right] && \text{[Factorizing]} \\ &= \mathbb{P}(A) \cdot \mathbb{P}(B^{\complement} \mid A) && \text{[Proved on a Worksheet]} \end{aligned}$$

again completing the proof. ■

7. The management of *GauchoStay* apartments is quite lazy, and has allowed an ant infestation to manifest. There is a 10% chance that a randomly selected unit will have an infestation problem, independently of all other units. The exterminator goes from unit to unit, but is forgetful and could visit the same unit twice. **For this problem, there is no need to simplify your answers.**

   (a) What is the probability that the exterminator observes exactly 3 infested units among the first 7 units they examine? [2pts.]

   Solution: Let X denote the number of infested units observed among the first 7 units; then $X \sim \text{Bin}(7, 0.1)$ and

$$\mathbb{P}(X = 3) = \binom{7}{3}(0.1)^3(0.9)^4$$

   (b) What is the probability that the 12^{th} unit the exterminator examines is the third infested unit they observed? [2pts.]

   Solution: Let Y denote the number of units needed to be observed before encountering the third infested unit; then $Y \sim \text{NegBin}(3, 0.1)$ and

$$\mathbb{P}(Y = 12) = \binom{12-1}{3-1}(0.1)^3(0.9)^{12-3} = \binom{11}{2}(0.1)^3(0.9)^9$$

(c) What is the expected number of units the exterminator must visit before observing their second infested unit? [2pts.]

Solution: Let Z denote the number of units needed to be observed before encountering the second infested unit; then $Z \sim \text{NegBin}(2, 0.1)$ and

$$\mathbb{E}[Z] = \frac{2}{0.1} = \boxed{\text{20 units}}$$

(d) Now, suppose that there are 100 units in *GauchoStay* and 10 of them are infested. Additionally, suppose that the exterminator now takes care to not examine the same apartment twice. What is the probability that the exterminator observes exactly 3 infested units in a sample of 6 units? [2pts.]

Solution: Let W denote the number of infected units observed in the sample of 6; then $W \sim \text{HyperGeom}(100, 10, 6)$ and

$$\mathbb{P}(W = 3) = \boxed{\frac{\binom{10}{3}\binom{90}{3}}{\binom{100}{6}}}$$

8. Let X be a continuous random variable with probability density function (p.d.f.) given by

$$f_X(x) = \begin{cases} \frac{2}{9} \cdot x & \text{if } 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Verify that $f_X(x)$ is a valid probability density function. [3pts.]

Solution: We check the two conditions: nonnegativity, and integrating to unity.

(1) If $x \in [0, 3]$ then clearly $x \geq 0$, meaning $f_X(x) \geq 0$ for all $x \in \mathbb{R}$.

(2) We compute

$$\int_0^3 \frac{2}{9} x \, \mathrm{d}x = \frac{2}{9} \cdot \frac{1}{2} \left[x^2 \right]_{x=0}^{x=3} = \frac{1}{9}(9 - 0) = 1 \checkmark$$

Therefore, $f_X(x)$ is a valid probability density function.

(b) Compute [4pts.]

$$\mathbb{E}\left[\frac{1}{(1 + X^2)} \right]$$

Show all of your steps, including any integration you perform!

Solution: By the LOTUS,

$$\mathbb{E}\left[\frac{1}{(1 + X^2)} \right] = \int_0^3 \frac{1}{1 + x^2} \cdot \frac{2}{9} x \, \mathrm{d}x = \frac{1}{9} \int_0^3 \frac{2x}{1 + x^2} \, \mathrm{d}x$$

To evaluate this integral, we make a $u-$substitution: $u = 1 + x^2$, so $\mathrm{d}u = 2x \, \mathrm{d}x$ and

$$\mathbb{E}\left[\frac{1}{(1 + X^2)} \right] = \frac{1}{9} \int_0^3 \frac{2x}{1 + x^2} \, \mathrm{d}x = \frac{1}{9} \int_1^{10} \frac{\mathrm{d}u}{u} = \frac{\ln(10)}{9}$$

(c) Find $F_X(x)$, the cumulative distribution function (c.d.f.) of X. Be sure to consider all cases! [4pts.]

Solution: There are three cases to consider:

- If $x < 0$ then $F_X(x) = 0$
- If $x \geq 3$ then $F_X(x) = 1$
- If $x \in [0, 3]$ then
$$F_X(x) = \int_0^x \frac{2}{9} t \, \mathrm{d}t = \frac{x^2}{9}$$

So, putting everything together:

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x^2}{9} & \text{if } 0 \leq x \leq 3 \\ 1 & \text{if } x > 3 \end{cases}$$

9. In a drawer, you have 2 red socks, 2 white socks, and 2 green socks. You randomly draw a sample of 4 socks, without replacement; let X denote the number of matching pairs in your sample (by matching, we mean in color).

(a) What is the state space S_X of X? [2pts.]

Solution: Though it may be tempting to initially say that $S_X = \{0, 1, 2\}$, let us be a bit more careful. Specifically, is it possible to have zero matching colors? The answer is "no!" This is because we are sampling 4 socks without replacement; we will always have at least one color that is repeated twice in our sample (for those familiar with MATH/PSTAT 8 arguments, the reason for this is the Pigeonhole Principle, but such an argument is not needed for the purposes of this class). Thus,

$$S_X = \{1, 2\}$$

(b) Find the probability mass function of X. [4pts.]

Solution: I admit that there was actually a small ambiguity in the problem which leads to two different (yet valid!) answers. The question really boils down to: "can we tell the two red socks apart?" [This relates to something called **distinguishability** which I did not explcitly mention in lecture.]

My assumption was that we absolutely can tell them apart; though we have two red socks, they are clearly two different socks. Therefore, letting R_1 denote the first red sock, R_2 the other red sock, W_1 the first white sock, W_2 the other white sock etc., our outcome space might look something like:

$$\begin{aligned}\Omega = \{&(R_1, R_2, W_1, W_2),\ (R_1, R_2, W_1, G_1),\ (R_1, R_2, W_1, G_2),\\ &(R_1, R_2, W_2, G_1),\ (R_1, R_2, W_2, G_2),\ (R_1, R_2, G_1, G_2),\\ &(R_1, W_1, W_2, G_1),\ (R_1, W_1, W_2, G_2),\ (R_1, W_1, G_1, G_2),\\ &(R_1, W_2, G_1, G_2),\ (R_2, W_1, W_2, G_1),\ (R_2, W_1, W_2, G_2),\\ &(R_2, W_2, G_1, G_2),\ (R_2, W_1, G_1, G_2),\ (W_1, W_2, G_1, G_2)\}\end{aligned}$$

In this way

$$|\Omega| = \binom{6}{4} = 15$$

and our counting would look like:

- The event $\{X = 1\}$ corresponds to "two socks of the same color, and two socks of different colors." Here is how we compute the number of 4-sock samples with this arrangement:
  - First we pick the color of our matched pair; there are $\binom{3}{1}$ ways to do this.

- Then, from that color, we pick both socks: $\binom{2}{2}$.
- Finally, from the remaining 2 colors we pick both of them to constitute our non-matched colors: $\binom{2}{2}$
- Lastly, we pick one sock from each of these non-matched colors: $\binom{2}{1}^2$

Thus,

$$|\{X=1\}| = \binom{3}{1}\binom{2}{2}\binom{2}{2}\binom{2}{1}^2$$

and, since outcomes are equally likely, we simply divide by the number of 4-sock samples that can be drawn from a total of 6 (which is $\binom{6}{4}$) to see

$$\mathbb{P}(X=1) = \frac{\binom{3}{1}\binom{2}{2}\binom{2}{2}\binom{2}{1}^2}{\binom{6}{4}} = \frac{12}{15}$$

- To find $\mathbb{P}(X=2)$ we could simply to 1 minus our probability above; I shall demonstrate the counting, though:
  - First we pick the two colors in which we have matches: $\binom{3}{2}$
  - Then, from each we pick 2 socks: $\binom{2}{2}^2$

  Thus

$$\mathbb{P}(X=2) = \frac{\binom{3}{2}\binom{2}{2}^2}{\binom{6}{4}} = \frac{3}{15}$$

We can see that, thankfully, $3/15 + 12/15 = 1$, so our probabilities do in fact add up to 1 (as they should!)

However, I saw several students made the (implicit) assumption that: "well, we can't really tell the two red socks apart, all that's important is that there are red socks!" In this way, we only need to track the number of each sock of each color in our 4-sock sample, and then compute what value of X this configuration corresponds to:

# Red	**# White**	**# Green**	X
2	1	1	1
2	2	0	2
2	0	2	2
1	2	1	1
1	1	2	1
0	2	2	2

From here, the argument goes: there are 6 outcomes, each equally likely, and so

$$\mathbb{P}(X=1) = \frac{3}{6} = \frac{1}{2} = \mathbb{P}(X=2)$$

Technically speaking, the issue with this approach is that these outcomes are *not* equally likely: take the first row, in which we have 2 red, 1 white, and 1 green. The question becomes *which* white and *which* green were included in our sample? There are really two of each, which leads to many more outcomes.

But, again, this was a much more nuanced distinction that I felt was unfair to assume of you so if you used either of the above arguments you were awarded full points!

(c) Compute $\mathbb{E}[X]$. [1 (bonus)]

Solution: We simply use the definition of expectation:

$$\begin{aligned}\mathbb{E}[X] &= \sum_k k\mathbb{P}(X=k) = 1\cdot\mathbb{P}(X=1) + 2\cdot\mathbb{P}(X=2) \\ &= 1\cdot\frac{12}{15} + 2\cdot\frac{3}{15} = \boxed{\frac{18}{15} = \frac{6}{5} = 1.2}\end{aligned}$$

Or, if you had used $\mathbb{P}(X=1) = 1/2 = \mathbb{P}(X=2)$, then

$$\mathbb{E}[X] = (1)\left(\frac{1}{2}\right) + (2)\left(\frac{1}{2}\right) = \boxed{\frac{3}{2} = 1.5}$$

As was mentioned in the instructions, because this was a bonus question it was graded on an all-or-nothing scale; no partial credit was awarded, and credit was only awarded to the above two values.