MATH 180A - INTRODUCTION TO PROBABILITY
MIDTERM #2

FALL 2020

REMEMBER THIS EXAM IS GRADED BY A HUMAN BEING. WRITE YOUR SOLUTIONS NEATLY AND COHERENTLY, OR THEY RISK NOT RECEIVING FULL CREDIT.

IF YOU DO NOT ASSIGN THE PAGES OF YOUR WORK TO THE QUESTIONS OF THE EXAM IN YOUR UPLOAD TO GRADESCOPE, YOU WILL LOSE POINTS (1 POINT FOR EVERY QUESTION THAT YOU FAIL TO ASSIGN THE PAGES TO).

WRITE YOUR NAME AND STUDENT ID ON THE FIRST PAGE OF YOUR UPLOAD

WRITE YOUR SOLUTIONS TO EACH PROBLEM ON SEPARATE PAGES. CLEARLY INDICATE AT THE TOP OF EACH PAGE THE NUMBER OF THE CORRESPONDING PROBLEM. DIFFERENT PARTS OF THE SAME PROBLEM CAN BE WRITTEN ON THE SAME PAGE (FOR EXAMPLE, PART (A) AND PART (B)).

FROM THE MOMENT YOU ACCESS THE MIDTERM ON GRADESCOPE, YOU WILL HAVE A TOTAL OF 70 MINUTES TO COMPLETE AND UPLOAD YOUR SOLUTIONS TO GRADESCOPE. THE EXAM IS WRITTEN TO TAKE 50 MINUTES. IT IS YOUR RESPONSIBILITY TO UPLOAD YOUR SOLUTIONS ON TIME.

THE MIDTERM IS OPEN TEXTBOOK, AND YOU CAN REWATCH LECTURES AND REVIEW THE LECTURE SLIDES. YOU ARE NOT ALLOWED TO USE CALCULATORS, COMPUTER ALGEBRA SYSTEMS, OR SIMILAR SOFTWARE. YOU ARE NOT ALLOWED TO CONSULT OTHER PEOPLE OR RESOURCES ON THE INTERNET. IN PARTICULAR, CONTRACT CHEATING AND OTHER INSTANCES OF ACADEMIC DISHONESTY CAN RESULT IN SUSPENSION OR DISMISSAL.

THIS EXAM IS PROPERTY OF THE REGENTS OF THE UNIVERSITY OF CALIFORNIA AND NOT MEANT FOR OUTSIDE DISTRIBUTION. IF YOU SEE THIS EXAM APPEARING ELSEWHERE, PLEASE NOTIFY THE INSTRUCTOR AT BAU@UCSD.EDU AND THE UCSD OFFICE OF ACADEMIC INTEGRITY AT AIO@UCSD.EDU.

EXCEL WITH INTEGRITY PLEDGE

I pledge to be fair to my classmates and instructors by completing all of my academic work with integrity. This means that I will respect the standards set by the instructor and institution, be responsible for the consequences of my choices, honestly represent my knowledge and abilities, and be a community member that others can trust to do the right thing even when no one is watching. I will always put learning before grades, and integrity before performance. I pledge to excel with integrity.

In addition to the above, I pledge that I did not receive outside assistance with this exam. Outside assistance includes but is not limited to other people, the internet, and resources beyond the textbook, lecture notes/videos, and homework assignments.

To acknowledge that you agree to this pledge, you must copy the sentence:

> I choose to excel with integrity as a member of the University of California, San Diego.

Make sure to sign your name below it and date it as well. Exams without this pledge will *not* be graded.

1. (25 points) Suppose that you are waiting for a bus whose arrival time is distributed as an exponential random variable with mean 1 hour. Once the bus arrives, it takes 1 hour to drive you home. However, if you wait 1 hour for the bus and it still has not arrived, you decide to give up on the bus and walk home, which takes 10 hours. Let Y be the amount of time (in hours) that it takes for you to get home including the time spent waiting for the bus.

(a) (15 points) Calculate the CDF of Y.

Solution. If $X \sim \text{Exp}(1)$, then

$$Y = \begin{cases} X+1 & \text{if } X < 1; \\ 10+1 = 11 & \text{if } X \geq 1. \end{cases}$$

So,

$$F_Y(t) = \begin{cases} \mathbb{P}(X \leq t-1) = 0 & \text{if } t < 1; \\ \mathbb{P}(X \leq t-1) = 1 - \mathbb{P}(X > t-1) = 1 - e^{-(t-1)} & \text{if } 1 \leq t < 2 \\ \mathbb{P}(X < 1) = 1 - e^{-(2-1)} = 1 - e^{-1} & \text{if } 2 \leq t < 11 \\ \mathbb{P}(Y \leq 11) = 1 & \text{if } t \geq 11. \end{cases}$$

Note that Y is neither continuous nor discrete (you did not need to say this to get full credit). $\square$

(b) (10 points) Calculate the expected value $\mathbb{E}[Y]$.

Solution. We use the fact that Y is a function of a continuous random variable with a known density. So,

$$\mathbb{E}[Y] = \int_0^1 (x+1)e^{-x}\,dx + \int_1^\infty 11e^{-x}\,dx.$$

The first integral can be computed using integration by parts:

$$\int (x+1)e^{-x}\,dx = -(x+1)e^{-x} + \int e^{-x}\,dx = -(x+1)e^{-x} - e^{-x} + C.$$

So,

$$\int_0^1 (x+1)e^{-x}\,dx = -(x+1)e^{-x} - e^{-x}\Big|_0^1 = -2e^{-1} - e^{-1} + 2 = 2 - 3e^{-1}.$$

The second integral can be computed using the tail probability for the exponential distribution:

$$\int_1^\infty 11e^{-x}\,dx = 11\mathbb{P}(X \geq 1) = 11e^{-1}.$$

So,

$$\mathbb{E}[Y] = 2 - 3e^{-1} + 11e^{-1} = 2 + 8e^{-1}.$$

$\square$

2. (25 points) Let $X \sim \mathrm{Geom}(p)$, where $p \in (0, 1)$. Compute

$$\mathbb{E}\left[\frac{1}{X!}\right],$$

where we recall that $X!$ is the factorial. To receive full credit, your final answer should not contain an infinite series.

Solution. This is a direct computation:

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{X!}\right] &= \sum_{k=1}^{\infty} \frac{1}{k!}(1-p)^{k-1} p \\
&= \frac{p}{1-p} \sum_{k=1}^{\infty} \frac{(1-p)^k}{k!} \\
&= \frac{p}{1-p} \left(\sum_{k=0}^{\infty} \frac{(1-p)^k}{k!} - 1 \right) \\
&= \frac{p}{1-p} \left(e^{1-p} - 1 \right).
\end{aligned}$$

$\square$

3. (25 points) 250000 randomly chosen individuals were interviewed to estimate the unknown fraction $p \in (0,1)$ of the population that like bagels. The resulting estimate is $\hat{p}$. Suppose that we want to construct a 98% confidence interval $(\hat{p}-\varepsilon, \hat{p}+\varepsilon)$. How large must we choose ε? You may leave your answer in terms of the inverse Φ^{-1} of the CDF of the standard normal.

Solution. Recall the equation

$$\mathbb{P}(|\hat{p}-p| < \varepsilon) \geq 2\Phi(2\varepsilon\sqrt{n}) - 1.$$

So, we want

$$2\Phi(2\varepsilon\sqrt{n}) - 1 \geq .98,$$

where $n = 250000$. Solving for ε in the above, we get

$$\varepsilon \geq \frac{\Phi^{-1}(.99)}{2\sqrt{n}} = \frac{\Phi^{-1}(.99)}{2\sqrt{250000}}.$$

This can be simplified to

$$\varepsilon \geq \frac{\Phi^{-1}(.99)}{1000},$$

but this was not necessary for full credit. □

4. (25 points) Let X be the random variable with density

$$f(x) = \begin{cases} 1 & \text{if } x \in (0,1); \\ 0 & \text{if } x \notin (0,1). \end{cases}$$

Let $Y = \ln(\sqrt{X})$.

(a) (15 points) Compute the moment generating function $M_Y(t)$ of Y. Hint: do not try to compute the density of Y.

Solution. By definition, the moment generating function is

$$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t\ln(\sqrt{X})}] = \mathbb{E}[\sqrt{X}^t] = \mathbb{E}[X^{t/2}].$$

We compute the latter expectation as a function of a random variable that we have a density for:

$$E[X^{t/2}] = \int_0^1 x^{t/2}\,dx = \begin{cases} \left.\dfrac{x^{\frac{t}{2}+1}}{\frac{t}{2}+1}\right|_0^1 = \dfrac{1}{\frac{t}{2}+1} & \text{if } t \neq 0 \text{ and } t > -2; \\ 1 & \text{if } t = 0; \\ +\infty & \text{if } t \leq -2. \end{cases}$$

Note that since $\frac{1}{\frac{t}{2}+1}$ at $t = 0$ is 1, we can simply write this as

$$M_Y(t) = \begin{cases} \dfrac{1}{\frac{t}{2}+1} & \text{if } t > -2; \\ +\infty & \text{if } t \leq -2. \end{cases}$$

$\square$

(b) (10 points) Use the moment generating function to compute the nth moment of Y.

Solution. We need to compute the Taylor series of $\frac{1}{\frac{t}{2}+1}$. Rather than computing derivatives, we use the fact that it can written as a geometric series:

$$\frac{1}{\frac{t}{2}+1} = \frac{1}{1-(-\frac{t}{2})} = \sum_{n=0}^{\infty}\left(-\frac{t}{2}\right)^n = \sum_{n=0}^{\infty}\frac{(-1)^n n!}{2^n}\frac{t^n}{n!}$$

for $|t| < 2$. So, the nth moment is

$$\frac{(-1)^n n!}{2^n}.$$

As a fun exercise, you can show that $-\log(X^r) \sim \mathrm{Exp}(1/r)$ for $r > 0$. Try to think of how you might do this using part (a). $\square$