This problem set covers material from Week 10, dates 4/21-4/23. Unless otherwise noted, all problems are taken from the textbook. Problems can be found at the end of the corresponding chapter.

Instructions: Write or type complete solutions to the following problems and submit answers to the corresponding Canvas assignment. Your solutions should be neatly-written, show all work and computations, include figures or graphs where appropriate, and include some written explanation of your method or process (enough that I can understand your reasoning without having to guess or make assumptions). A general rubric for homework problems appears on the final page of this assignment.

Monday 4/21

1. A popular model for population growth of microscopic organisms in large environments is exponential growth. At time 0, we introduce $v > 0$ organisms into a large tank. Let X represent the rate of growth, which is random but we can model the rate according to the following PDF:
$$f_X(x) = 3(1 - x^2) \qquad 0 < x < 1$$
We are more interested in the size of the population after $t > 0$ minutes. Letting Y represent the population size, the population growth model defines $Y = ve^{Xt}$.

   Find the PDF of the population size f_Y (assuming v and t fixed) using change-of-variables (if it applies). Otherwise, use the CDF method.

2. Suppose $X \sim f_X$ and $Y \sim f_Y$ are independent random variables, each with support on $(0, \infty)$. Define the random variable $R = \frac{X}{Y}$. Find the PDF of R $f_R(r)$ using change-of-variables in two approaches: one where you define your second transformation as $W = Y$ and another where you define your second transformation as $W = X$. Your answer in both cases will be in the form of an integral.

3. Suppose U and $V \overset{\text{iid}}{\sim} \text{Exp}(\lambda)$. We will find the joint distribution of $X = U + V$ and $Y = \frac{U}{U+V}$, as well as the marginal distribution of Y alone.
   (a) First, obtain the joint PDF $f_{U,V}$. Don't forget support!
   (b) Use the change-of-variables formula to find the joint PDF $f_{X,Y}(x, y)$. Don't forget the support!
   (c) Based on your previous answer, are X and Y independent?
   (d) Find the marginal PDF of Y. What named distribution is this?

Wednesday 4/23

4. Alice walks into a post office with 2 clerks. Both clerks are serving other customers, but Alice is next in line. The clerk on the left takes $\text{Exp}(\lambda_1)$ time to serve a customer, independent of the clerk on the right who takes $\text{Exp}(\lambda_2)$ time. Let T_1 be the time until the clerk on the left is done with their current customer, and T_2 likewise for the clerk on the right. Alice will go to whichever clerk is available next, and leave immediately after.

(a) On a previous homework, you showed $P(T_1 < T_2) = \frac{\lambda_1}{\lambda_1+\lambda_2}$ using integration. Let's try showing this another way that avoids integration altogether, and instead uses some nice theorems that we recently learned. To get you started, let $X_1, X_2 \stackrel{\text{iid}}{\sim} \text{Exp}(1) \equiv \text{Gamma}(1, 1)$. Then at some point, it could be helpful to introduce the term $\lambda_1 X_1$...

(b) Find the expectation of the total time T Alice spends in the post office. *Hint: can you re-write T as a sum of two possibly new random variables? Then eventually consider the following LoTP applied to continuous PDFs:*

$$f_X(x) = \underbrace{f_{X|A}(x|A)}_{\substack{\text{PDF of } X \text{ given} \\ \text{event } A \text{ occurred}}} P(A) + \underbrace{f_{X|A^c}(x|A^c)}_{\substack{\text{PDF of } X \text{ given} \\ \text{event } A^c \text{ occurred}}} P(A^c)$$

5. An important transformation of the n random variables $X_1, \dots, X_n$ is sorting them in increasing order to produce the following new random variables $Y_1, \dots, Y_n$:

$$\begin{aligned}
Y_1 &= X_{(1)} = \min\{X_1, \dots, X_n\} \\
Y_2 &= X_{(2)} = \text{ second smallest of } X_1, \dots, X_n \\
&\vdots \\
Y_{n-1} &= X_{(n-1)} = \text{ second largest of } X_1, \dots, X_n \\
Y_n &= X_{(n)} = \max\{X_1, \dots, X_n\}
\end{aligned}$$

The subscript notation with the parentheses is the common way to denote order statistics. $Y_j = X_{(j)}$ is called the **j-th order statistic**.

Our interest mostly lies in the case where the X_i are IID from a continuous distribution with PDF f_X, CDF F_X, and support S_X. In this case, we can derive the PDF of $X_{(j)}$ by "story". When we do so, the marginal PDF of the j-th order statistic is:

$$f_{X_{(j)}}(x) = n\binom{n-1}{j-1} f_X(x)(F_X(x))^{j-1}(1 - F_X(x))^{n-j}, \quad x \in S_X$$

(a) In words, can you describe why this PDF "makes sense" to be the PDF of the j-th order statistic by explaining what each component of the PDF tells us? That is, build this PDF by story by describing what role the n plays, the $\binom{n-1}{j-1}$ plays, etc.

(b) Now suppose that $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Unif}(0, 1)$. Find and name the PDF of the j-th order statistic.

Friday 4/25

None!

General rubric

Points	Criteria
5	The solution is correct *and* well-written. The author leaves no doubt as to why the solution is valid.
4.5	The solution is well-written, and is correct except for some minor arithmetic or calculation mistake.
4	The solution is technically correct, but author has omitted some key justification for why the solution is valid. Alternatively, the solution is well-written, but is missing a small, but essential component.
3	The solution is well-written, but either overlooks a significant component of the problem or makes a significant mistake. Alternatively, in a multi-part problem, a majority of the solutions are correct and well-written, but one part is missing or is significantly incorrect.
2	The solution is either correct but not adequately written, or it is adequately written but overlooks a significant component of the problem or makes a significant mistake.
1	The solution is rudimentary, but contains some relevant ideas. Alternatively, the solution briefly indicates the correct answer, but provides no further justification.
0	Either the solution is missing entirely, or the author makes no non-trivial progress toward a solution (i.e. just writes the statement of the problem and/or restates given information).
Notes:	For problems with multiple parts, the score represents a holistic review of the entire problem. Additionally, half-points may be used if the solution falls between two point values above.
Notes:	For problems with code, well-written means only having lines of code that are necessary to solving the problem, as well as presenting the solution for the reader to easily see. It might also be worth adding comments to your code.