MAST20004 Probability Semester 1, 2021

Assignment 2: Questions

Due 3 pm, Friday 16 April 2021 April 9, 2021

Name:		
Student ID:		

Important instructions:

- (1) This assignment contains 5 questions, **two** of which will be randomly selected to be marked. Each marked question is worth 10 points and each unmarked question with substantial working is worth 1 point.
- (2) To complete this assignment, you need to write your solutions into the blank answer spaces following each question in this assignment PDF.
 - If you have a printer (or can access one), then you must print out the assignment template and handwrite your solutions into the answer spaces.
 - If you do not have a printer but you can figure out how to annotate a PDF using an iPad/Android tablet/Graphics tablet or using Adobe Acrobat, then annotate your answers directly onto the assignment PDF and save a copy for submission.

Failing both of these methods, you may handwrite your answers as normal on blank paper and then scan for submission (but note that you will thereby miss valuable practice for the exam process). In that case, however, your document should have the same length as the assignment template otherwise Gradescope will reject your submission. So you will need to add as many blank pages as necessary to reach that criterion.

Scan your assignment to a PDF file using your mobile phone (we recommend Cam - Scanner App), then upload by going to the Assignments menu on Canvas and submit the PDF to the **GradeScope tool** by first selecting your PDF file and then clicking on 'Upload PDF'.

Note that here you do not need to submit any Matlab code with your assignment.

- (3) A poor presentation penalty of 10% of the total available marks will apply unless your submitted assignment meets all of the following requirements:
 - it is a single pdf with all pages in correct template order and the correct way up, and with any blank pages with additional working added only at the end of the template pages;
 - has all pages clearly readable;
 - has all pages cropped to the A4 borders of the original page and is imaged from directly above to avoid excessive 'keystoning'.

These requirements are easy to meet if you use a scanning app on your phone and take some care with your submission - please review it before submitting to double check you have satisfied all of the above requirements.

- (4) Late submission within 20 hours after the deadline will be penalised by 5% of the total available marks for every hour or part thereof after the deadline. After that, the Gradescope submission channel will be closed, and your submission will no longer be accepted. You are strongly encouraged to submit the assignment a few days before the deadline just in case of unexpected technical issues. If you are facing a rather exceptional/extreme situation that prevents you from submitting on time, please contact the tutor coordinator **Robert Maillardet** with formal proofs such as medical certificate.
- (5) Working and reasoning must be given to obtain full credit. Clarity, neatness, and style count.

- **Q1.** A random variable N has pmf $p_N(k) = \frac{c}{k(k+1)}$ where $k=2,3,4,\ldots$, and c is a constant.
 - (a) Derive the value of constant c.

(b)	Find a simple expression for the cdf $F_N(x)$ of N (that is, you should simplify any potential sum appearing in the expression).

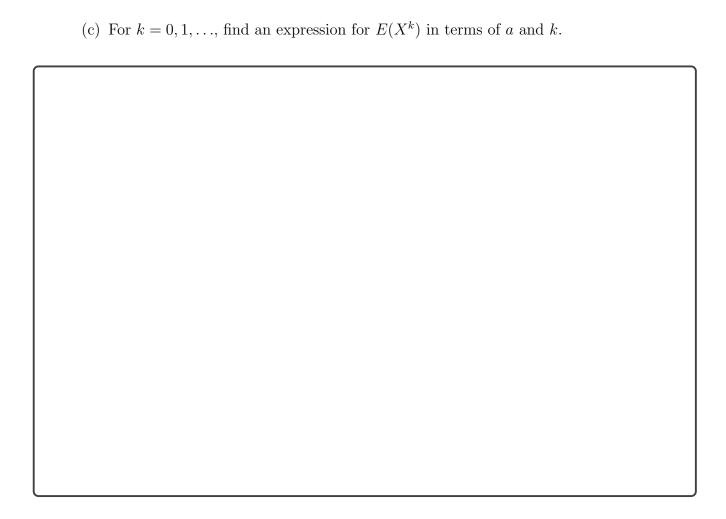
(c)	Find the mean of N , if it exists. If it does not exist, prove that it does not exist.

(d)) Find the standard deviation of N , if it exists. If it does not exist, prove that it does not exist.

(e) Let $Y = (-1)^{-1}$ Justify you	$-1)^N N$. What are the pur answer.	possible values of Y ?	? Does the mean of	Y exist?

- **Q2** A random variable 0 < X < 1 has density $f_X(x) = C_a x^a (1-x)$ for some constant C_a , where a is a real positive parameter.
 - (a) What is the value of the constant C_a ?

(b) What is the CDF of X?



Q3 The lifetime (in weeks) of a certain system component is thought to have an *exponential* distribution, whose density function is given by

$$f(x) = \lambda e^{-\lambda x}, \qquad x \ge 0,$$

for some strictly positive parameter λ .

(a) The *reliability* of a component is defined as the probability that the component will not have failed by a specified time. If the reliability of the system component at 10.5 weeks is 0.9, find the reliability at 10 weeks.

(b) One hundred components of this type are put in a new system. All components that have failed are replaced at 20 week intervals, and none are replaced at other

Q4 Let X be a continuous non-negative random variable with distribution function F(x) and probability density function f(x) which is such that f(x) = F'(x) for all $x \in [0, \infty)$. Define the failure rate function (also called hazard function) as

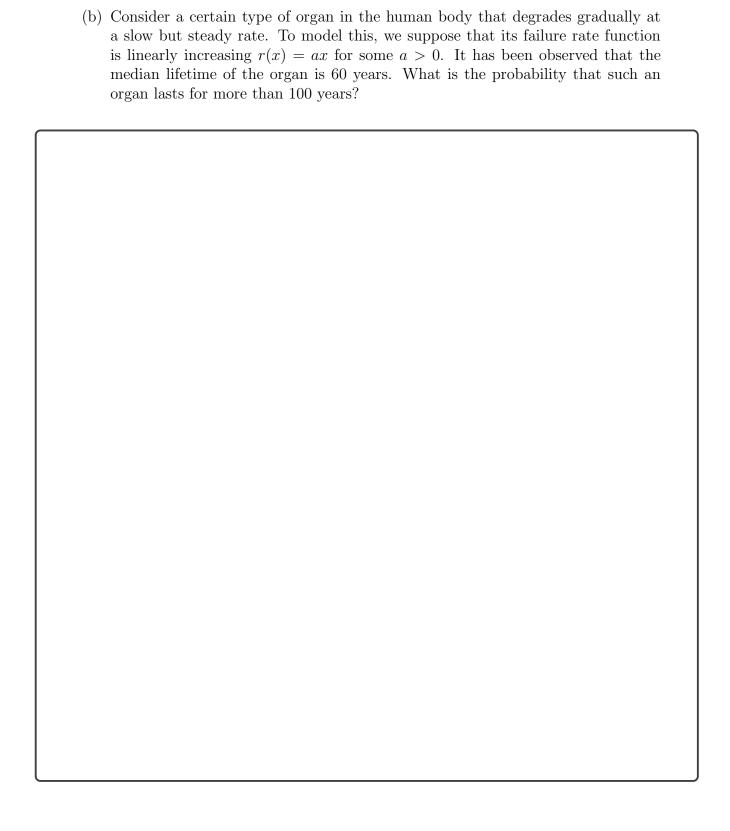
$$r(x) = \frac{f(x)}{1 - F(x)},$$

and the survival function as

$$G(x) = 1 - F(x).$$

(a) Show that for a non negative continuous random variable X with failure rate function r, X has survival function

$$G(x) = \exp\left(-\int_0^x r(u)du\right).$$



- (c) Find the probability density functions of non negative continuous random variables with failure rate functions
 - (i) $r(x) = x^3$,
 - (ii) r(x) = 1/(1+x).

Q5 A bag contains n tokens numbered from 1 to n. A random sample of n tokens is selected from the bag, one at a time (the order matters). A match occurs if the token numbered i is selected on the ith draw.

In this exercise, we are going to investigate if selecting the tokens with replacement or selecting without replacement does affect the probability of at least one match. We will also try to understand to which extent the number of tokens in the bag affects the probability of at least one match.

The Matlab program $Assignment2_Q5_2021.m$ simulates nreps repetitions of the experiment of drawing n tokens, with and without replacement. Read through the program and make sure you understand what each line is computing.

(i) What is the theoretical probability that there is at least one match?

(a)	Ιİ	the	draws	are	done	with	replacement	,	

	matches?
(iii) If n is large, what distribution can be used to approximatches? Justify your answer.	ximate the number of

(iv) Explain what Lines 24, 26 and 38 are computing.
(v) Uncomment Line 42 and replace the "??" with the right expression in terms of pmf_1 to display the empirical probability that there is at least one match. Copy your answer here.

(vi) Uncomment Lines 48 and 49 and replace the "??" in Line 48 by the expression you found in (a). Copy your answer here.

(b) If the draws are done without replacement, it can be shown that

 $P(\text{at least one match}) = 1 - \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!}\right)$ $= 1 - \sum_{k=0}^n \frac{(-1)^k}{k!}.$

It is much harder to describe the full distribution of the number of matches in this case.

Lines 51–60 compute and display the theoretical probability that there is at least one match without replacement using the above formula.

Uncomment Line 43 and replace the "??" with the right expression in terms of pmf_2 to display the empirical probability that there is at least one match. Copy your answer here.

