STAT7203: Applied Probability and Statistics
Assignment 1

Due by 14:00 on Tuesday the 6^{th} of October, 2020. Submission via Blackboard.

The marks for each question are indicate by the number in square brackets. There are a total of 15 marks for this assignment.

1. A continuous random variable X has probability density function

$$f_X(x) = \begin{cases} \frac{1}{x \log 10}, & x \in [0.1, 1] \\ 0, & \text{else.} \end{cases} \tag{1}$$

   (a) Compute the mean and variance of X. [1 mark]

   (b) Determine the probability density function of $Y = (10X)^{-1}$. [1 mark]

   (c) Determine the quantile function of X. [2 marks]

   (d) Hence, write a small function in `R` or `MATLAB` to simulate random variables from this distribution. Supply your code. [1 mark]

   (e) Let X and Y be two independent random variables where X has the probability density function (1) and Y has a continuous distribution of your choice whose support is either $(0, \infty)$ or a subset of $(0, \infty)$. [The list of built-in distributions in `R` can be found by typing `?distributions`. In `MATLAB` this can be found by typing `help stats`.] We are interested in the distribution of the mantissa of the product $Z = XY$. The mantissa of Z can be obtained as

$$W = 10^{(\log_{10} Z - \lceil \log_{10} Z \rceil)},$$

   where $\lceil x \rceil$ denotes rounding up x to the nearest integer. Use this procedure to generate 10 000 independent realisations of W and form a histogram of the result. Compare this with the probability density function (1). How do the two distributions compare? Provide a histogram of the outcomes with the probability density function (1) overlaid. Supply your code. [2 marks]

   (f) Let X_1 and X_2 be two independent random variables with probability density function (1). Determine the probability that $X_1 X_2 \leq 0.1$. [2 marks]

2. Suppose you have four origami books which are labeled $B_1, \ldots, B_4$. The books are kept stacked on your desk. Each day you select one book from the stack and build a model. Books are selected each day at random with probabilities $\mathbb{P}(B_i \text{ is selected}) = (5 - i)/10$ and independently of past selections. At the end of the day you place the selected book on top of the stack.

   Suppose on day 0 the books are stacked in order of the labels. That is, book B_1 is in first position (top), book B_2 is second from the top, book B_3 is third from the top, and book B_4 is on the bottom.

(a) What is the probability that book B_1 is on top of the stack at the end of day n? Does this probability depend on n? [1 mark]

(b) Give an expression for the probability that book B_1 is second from the top of the stack at the end of day n. What is the limiting value as $n \to \infty$? [2 marks]

(c) Simulate this process 10 000 times and provide an estimate of the probability that book B_1 is on the bottom of the stack at the end of day 100. Supply your code. [2 marks]

(d) Let $X_i = 1$ if book B_1 is on the bottom of the stack in the i-th simulation of the process in part (c) and $X_i = 0$ otherwise. Assuming $X_1, X_2, \ldots, X_m$ are independent, show

$$\operatorname{Var}\left(n^{-1}\sum_{i=1}^{m} X_i\right) \leq 1/(4m).$$

[1 mark]

Total [15 marks]