Assignment 3

STAT 321: Stochastic Systems and Signals
Winter Session 2021-2022 (Term 2)

Assignment Instructions

- Assignments are to be completed individually.
- When answering the questions, writing down the final answer will not be sufficient to receive full marks. Please show all calculations unless otherwise specified. Also define any events and notation that you use in your solutions.
- You may use the language of your choice for the numerical parts, but the free statistical software `R` is recommended. Include any code used in your assignment as an Appendix **at the very end of your assignment**.
- *Due date:* February 20 at 11:59pm (Vancouver Time). You must submit an electronic (preferably PDF) version of your assignment on Canvas.

Problem 1 [10 marks]

A game consists of first rolling a fair 10-faced die once and then tossing a coin 10 times. The number rolled on the die will determine the probability of landing heads for the coin. Let R be the random variable for the value rolled from the die, and N be the random variable for the number of heads from the 10 coin tosses. If $R = r$ is rolled, the coin has probability $r/10$ of landing heads up.

(a) What is the conditional expectation of N given $R = r$? [**2.5 marks**]

(b) What is the conditional variance of N given $R = r$? [**2.5 marks**]

(c) Use your results from parts (a) and (b) to compute the expected value of the number of heads obtained in playing the game once. [**2.5 marks**]

(d) Use your results from parts (a) and (b) to compute the variance of the number of heads obtained in playing the game once. [**2.5 marks**]

Problem 2 [25 marks]

An urn contains 5 red balls, 6 green balls and 3 blue balls. A total of 3 balls are selected without replacement from the urn. Denote X_1 the number of red balls selected and X_2 the number of green balls selected.

(a) Find the joint probability mass function of X_1 and X_2, $p(x_1, x_2)$. **[5 marks]**

(b) Find the marginal probability mass functions of X_1 and X_2, $p_1(x_1)$ and $p_2(x_2)$. **[5 marks]**

(c) Compute the correlation between X_1 and X_2. **[15 marks]**

Problem 3 [25 marks]

Let T be a continuous random variable denoting the amount of time a UBC student waits for the 99 B-Line in the morning to get to class. Suppose T is defined for waiting times between 0 and 30 minutes, with the following probability density function:

$$f(t) = \begin{cases} c(6-t)^2, & 0 \le t \le 30, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the normalizing constant c. **[5 marks]**

(b) Compute the cumulative distribution function $F(t)$. **[5 marks]**

(c) What is the probability that a student will wait more than 10 minutes for the bus? **[5 marks]**

(d) Suppose you have been waiting for 10 minutes already, what is the probability that you will be waiting for at least 20 minutes? **[5 marks]**

(e) Assume that if you wait longer than 20 minutes, you will be late for class. Out of 10 randomly selected students, find the probability that 3 of them will be late to class. **[5 marks]**

Problem 4 [10 marks]

Let $Y \sim \exp(\lambda)$. Define the random variable $X = \lfloor Y \rfloor + 1$, where $\lfloor \cdot \rfloor$ is the floor function. That is, X is a discrete random variable that takes the value k whenever $k - 1 \le Y < k$, for $k = 1, 2, \ldots$. What is the distribution of X? Note that you have to give the name and parameter(s) of the distribution.

Problem 5 [15 marks]

A random walk, also known as a stochastic or random process, can be described by a path that consists of a succession of random steps on some mathematical space such as the integers. A Gaussian random walk is a random walk where the random steps are normally distributed with some mean μ and some variance σ^2. For this problem, we will consider Gaussian random walks, i.e. random walks with steps that come from the standard normal distribution. We can define the state S_n of the random walk at step n as follows:

$$S_n = \sum_{i=1}^{n} X_i, \;\; X_i \sim \mathrm{N}(0,1), \;\; i = 1, \ldots, n$$

(a) Write a simulation in `R` to approximate the expected distance from the starting point, $E\big(|S_n|\big)$, for $n = 10, 20, \ldots, 100$. Use at least N =10,000 replications for each value of n. **[5 marks]**

(b) Identify the distribution of S_n, and give a 95% confidence interval for $|S_n|$, for $n = 10, 20, \ldots, 100$. **[5 marks]**

(c) Write a simulation in `R` to verify your theoretical confidence intervals in part (b). Use at least N =10,000 replications for each value of n. **[5 marks]**