Assignment 2

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 7 at 11:59pm (Vancouver Time). You must submit an electronic (preferably PDF) version of your assignment on Canvas.

Problem 1 [15 marks]

Three types of infections in a population, which we call A, B and C for simplicity, are mutually exclusive: they cannot be present together. Infection A afflicts one per cent of the population, infection B afflicts one-half of one per cent of the population, and infection C afflicts two per cent of the population.

Sometimes, one of these infections can be present without the victim having any symptoms. Suppose that 10% of the A-infected people have no symptoms, 5% of B-infected people have no symptoms, and 25% of C-infected people have no symptoms. Symptoms are present only when an infection is present, and the symptoms are identical for all three types of infections.

(a) If a person has symptoms, what are the probabilities that he/she has infections A? B? C? [**5 marks**]

(b) If a person has no symptoms, what are the probabilities that he/she has infections A? B? C? [**5 marks**]

(c) A test with a sensitivity of 95% and a specificity of 90% has been developed for infection A. The test is only applied to people who report symptoms.

  (i) What is the probability that a person testing negative doesn't have infection A? [**2.5 marks**]

  (ii) What is the probability that there will be a screening error based on this test? [**2.5 marks**]

Problem 2 [15 marks]

Consider the following situation:

- There is an urn with 1 red ball and 1 blue ball.
- Every time a ball is drawn (at random) from the urn, it is placed back in the urn along with 2 more balls of the same color, and 1 more ball of the other color.

Let X_i be the number of red balls after the i-th draw, for $i = 1, 2 \ldots.$ Note that X_i is the random variable representing the number of red balls in the urn including the three new balls added after the i-th draw.

(a) Find the probability mass function (*pmf*) of X_3. **[5 marks]**

(b) Compute $E(X_3)$ and $Var(X_3)$. **[5 marks]**

(c) What is the probability that the first ball drawn was red, given that there are exactly 5 red balls after the third ball is drawn. **[5 marks]**

Problem 3 [15 marks]

An investor in the stock market purchased a single share of stock for each of 50 *different* companies. Suppose each stock is worth \$20. Every day, each stock goes up by \$1 with probability p or down by \$1 with probability $(1-p)$. Note that the stocks are independent of each other, and each day is independent of the others as well. Denote the value of the k-th stock after T days (where T is a constant) by S_T^k (for $k = 1, \ldots, 50$). The initial value of the *portfolio* of the investor is

$$\sum_{k=1}^{K} S_0^k = \sum_{k=1}^{50} 20 = 20 \times 50 = 1,000.$$

(a) What is the probability that a share of stock (of any company) was worth \$23 after 5 days if it was worth \$26 after 10 days? **[5 marks]**

(b) What is the expected value and the standard deviation of the value of the investor's portfolio after T days? Your answer should be in terms of T and p. Assume that $T \leq 20$ days. **[5 marks]**

(c) Is it more advantageous for the investor to invest in 50 shares of the *same* company, or keep his current strategy of buying a single share of stock for 50 *different* companies? Explain your answer by comparing the expected value and variance of the portfolio values after T days. Again, you can assume that $T \leq 20$ days. **[5 marks]**

Problem 4 [20 marks]

A sequence of binary bits is transmitted in a certain communication system. Any given bit is decoded erroneously with probability p and decoded correctly with probability $1-p$. Errors occur independently from bit to bit.

(a) Out of a sequence of n bits transmitted, what is the probability that j bits ($j = 1, ..., n$) are erroneously decoded? **[5 marks]**

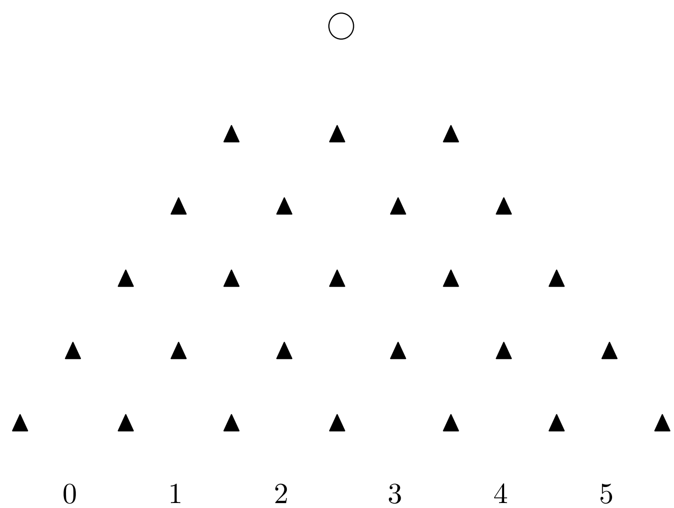
(b) A certain error-correcting code is applied to the communication system described in part (a). The correcting code can correct a single error in any bit, but it cannot correct a run of two or more errors. What is the probability that the code can correct the errors that occur when n-digits are transmitted and j bits ($j = 1, ..., n$) are incorrectly decoded? **[5 marks]**

(c) Write a simulation in R to approximate the probability that a message decoded using the method described in part (b) is correctly decoded for $n = 8, 16, 24, 32$ and $p = 0.01, 0.05, 0.10, 0.15$. That is, you must report $4 \times 4 = 16$ different probabilities (for every possible combination of n and p). **[5 marks]**

(d) Suppose now that the error-correcting code can correct up to 2 consecutive bit errors, but it cannot correct a run of three or more errors. Write a simulation in R to approximate the probability that a message is correctly decoded for $n = 8, 16, 24, 32$ and $p = 0.01, 0.05, 0.10, 0.15$. That is, you must report $4 \times 4 = 16$ different probabilities (for every possible combination of n and p). **[5 marks]**

Problem 5 [20 marks]

In his 1889 publication Natural Inheritance, the renewed British scientist Sir Francis Galton described a pinball-type board that he called a quincunx.



As pictured, the quincunx has five rows of pegs, the pegs in each row being the same distance apart. At the bottom of the board are six cells, numbered 0 through 5. A ball is introduced at the top of the board and will hit the middle peg in the first row and veer either to the right or to the left, then strike a peg in the second row, again veering to the right or to the left, and so on.

(a) If the ball has a 50-50 chance of going either direction each time it hits a peg, what is the probability it ends in cell 0? What about for the other cells? **[5 marks]**

(b) Write a code in R that simulates this experiment and run it 1,000 times, recording the number of times the ball lands on each cell. Compare the frequencies with the theoretical probabilities derived in part (a). **[5 marks]**

(c) Suppose now that the quincunx is enlarged to have 100 rows. Derive the probabilities of landing on cells $0, 1, 2, \ldots, 100$? **[5 marks]**

(d) Update your code to run this experiment and run it 1,000 times. Compare the frequencies with the theoretical probabilities derived in part (c). **[5 marks]**