STAT7203: Applied Probability and Statistics
Week 3 Exercises

1. Consider the water cooling system schematically depicted in Figure 1. The system has four unreliable components: two identical pumps (P1 and P2) and two valves (V1 and V2). The system works if, in the diagram, there is a path from left to right traversing only working components.
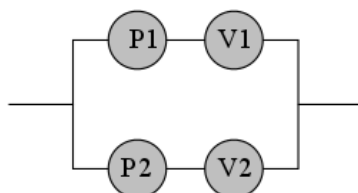

Figure 1: The system works if there is a path from left to right with working components.

   We assume that the pumps and valves fail independently of each other and that the probability that each pump works is 0.9 and the probability that each valve works is 0.8.

   (a) Calculate the probability that the system works.

   (b) Simulate 100,000 outcomes from the system in R and estimate the probability that the system is working as the proportion of times out of the 100,000 outcomes for which the system is working.

   Some useful R commands for this problem:

   - ?: Typing ? followed by the name of the function of interest will display the help text in the command window, for example ?rbinom.
   - rbinom: Simulates random variables from a binomial distribution. Typing rbinom(n,size,prob) returns n random variables from a binomial distribution with size trials and success probability prob.
   - Logical operators & and |: These are the AND (&) and OR (|) operators. These can be applied to vectors of 0's (logical value FALSE) and 1's (logical value TRUE). Try the following:
     ```
     > x = c(1, 1, 0, 0)
     > y = c(1, 0, 1, 0)
     > x | y
     > x & y
     ```
   - The function c is used to create a vector. To access an element of a vector we use square brackets [ ]. What does x[2] output?
   - mean: This function computes the average of the elements of the input.

2. A certain AIDS test has a 0.98 probability of giving a Positive result when the blood is infected, and a 0.07 probability of giving a Positive result when the blood is not infected (a so-called false positive). Suppose 1% of the population carries the HIV virus.

(a) Using Bayes' Rule, determine the probability that a person selected uniformly at random from the population is indeed infected, given that the test yields a Positive result.

(b) Simulate a population of size 100,000 with each individuals AIDS status and the result of the test. Estimate the proportion of the population that returns a positive test result and the proportion of those with a positive test result that have AIDS.

The commands used for the previous question will be useful here as well. You might also want to use

- * Recall in R there are two different multiplications. To multiple matrices and vectors elementwise we use * For a proper matrix multiplication you need to use %*%.
- sum: This function computes the sum of the elements of the input.

3. A packet passes through n routers from the source to the destination. The packet has space in its header for one router label. On passing through the first router, the router writes its label in the header of the packet. On passing through the k-th router, the router overwrites the label in the header with its own label with probability $1/k$, independently of the action of other routers.

What is the probability that a packet arriving at the destination has the label of the j-th router in its header?

4. (Challenge) Let $(a_1, \ldots, a_n)$ be a random permutation of the integers $\{1, 2, \ldots, n\}$ with all permuations equally likely. Let X be the number of values of the permutation that are in the correct position. Use R to simulate random permuations and approximate $\mathbb{P}(X = k)$ for $n = 50$ and $k = 0, 1, 2, \ldots, 5$.

Can you conjecture a general formula for $\mathbb{P}(X = k)$?

Some useful R commands for this problem:

- : (colon) Typing 1:5 produces the vector of integers [1 2 3 4 5].
- rep Typing rep(0,n) creates a vector of zeros of length n. This is useful if you need to create a matrix/vector to store values calculated later.
- sample: Typing sample(n,n) returns a vector containing a random permutation of the integers 1:n.
- ==: Typing A == B does element by element comparisons between A and B and returns an array with elements set to logical 1 (TRUE) where the relation is true and elements set to logical 0 (FALSE) where it is not.
- table: If x is a vector of integers, then table(x) produces a frequency table for x.
- for: You may need to use a FOR loop for this question. See the help for the correct syntax.