CS544 Module 4 Assignment

Use the appropriate R functions for the corresponding distributions for the following questions. For the items underlined below (Part1 b, Part2 b, Part3 b), check the answer obtained using the R function with the explicit formula calculation. Show both the solutions.

Part1) Binomial distribution (20 points)

Suppose a student has 40% chance of scoring a perfect score in an exam with randomly selected questions. Each student will be provided 5 attempts.

a) Compute and plot the probability distribution for the number of perfect scores over the 5 attempts (both the PMF and CDF)
b) <u>What is the probability that a student will score a perfect score in exactly 2 out of the 5 attempts?</u>
c) What is the probability that a student will score a perfect score in at least 2 out of the 5 attempts?
d) Simulate the number of perfect scores over 5 attempts for 1000 students. Show the barplot of the frequencies of successes.

Part2) Negative Binomial distribution (20 points)

Suppose a student has 60% chance of scoring a perfect score in an exam with randomly selected questions. The student has to repeatedly take the exam until they achieve three perfect scores.

a) Compute and plot the probability distribution for scoring the three perfect scores (both the PMF and CDF). The student will only go for a maximum of 10 failures before giving up.
b) <u>What is the probability that the student will have the three perfect scores with exactly 4 failures?</u>
c) What is the probability that the student will have the three perfect scores with at most 4 failures?
d) Simulate the number of failures to get the three perfect scores for 1000 students. Show the barplot of the frequencies of the failures.

Part3) Hypergeometric distribution (20 points)

Suppose that your professor created a pool of 60 multiple choice questions and 40 programming questions for the final exam. For each student, a random set of 20 distinct questions from the pool will be presented during the exam. The student has the opinion that the multiple-choice questions are easy to handle than the programming questions.

a) Compute and plot the probability distribution for the number of multiple choice questions out of the 20 questions that the student will be given?
b) What is the probability that the student will have exactly 10 multiple choice questions out of the 20 questions in the exam?
c) What is the probability that the student will have at least 10 multiple choice questions out of the 20 questions in the exam?
d) Simulate the number of multiple choice questions for 1000 students. Show the barplot of the frequencies of the multiple-choice questions.

Part4) Poisson distribution (20 points)

Suppose that, on an average, students email 10 questions per day to the professor.

a) What is the probability that the professor will have to answer exactly 8 questions per day?
b) What is the probability that the professor will have to answer at most 8 questions per day?
c) What is the probability that the professor will have to answer between 6 and 12 questions (inclusive)?
d) Calculate and plot the PMF for the first 20 questions.
e) Suppose the course runs for 50 days. Simulate the number of questions the professor gets per day over the course run. Show a barplot of the frequencies of the number of questions. Show a boxplot of the number of questions. What do you infer from the plots?

Part5) Normal distribution (20 points)

Suppose that visitors at a theme park spend an average of $100 on souvenirs. Assume that the money spent is normally distributed with a standard deviation of $10.

a) Plot the PDF of this distribution covering the three standard deviations on either side of the mean.
b) What are the chances that a randomly selected visitor will spend more than $120?
c) What is the chance that a randomly selected visitor will spend between $80 and $90 (inclusive)?
d) What are the chances of spending within one standard deviation, two standard deviations, and three standard deviations, respectively?
e) Between what two values will the middle 80% of the money spent will fall?
f) If the theme park gives a free T-shirt for the top 2% of the spenders, what will be the minimum amount you have to spend to get the free T-shirt?
g) Show a plot for 10,000 visitors using the above distribution.

Submission:

When the term *lastName* is referenced, please replace it with your last name.

Provide all R code in a single file, **CS544_HW4_LastName.R**. Clearly mark each subpart of each question.

Provide the corresponding outputs from the R console in a single PDF document, **CS544_HW4_LastName.pdf**

Upload the two files to the Assignments section of Blackboard.

Note: Only ONE submission is allowed. Please be sure that what you are submitting is your final submission.