STAT 251 Written Assignment-1

(40 TOTAL MARKS)

Note 1: Please show steps with proper justifications in your solutions. Partial credits are given to intermediate steps and reasoning. Also define any event and notation that you use in your solutions. Use 3 decimal points in calculation. Don't excessively round off in intermediate steps.

Note 2: Discussion of ideas learned in class is encourage (with other students, TAs or the instructor). This helps the leaning process. But individual work turned in by each student should be your own work. Do not copy or paraphrase solutions from other students or from other sources. DO NOT provide your solutions to another student. Failure to comply with these rules will result in an automatic 0 for your work, and additional academic penalties.

Question 1

For this problem, use dataset `breaking.txt` along with R. An experiment was conducted to select the supplier of raw materials for production of a component. The breaking strength of the component (column `breaking`) was the objective of interest. Four suppliers were considered (column `supplier`). The four operators (column `operator`) can only produce one component each per day

(a) Create a histogram of breaking strength and label the axes properly. Comment on the shape of the distribution. *(3 marks)*

(b) Find the five number summary and the standard deviation for the variable breaking strength. *(2 marks)*

(c) Create a proper graphical display to compare the breaking strength distributions of the four different suppliers and label the axes properly. Which supplier would you recommend? Justify your answer with a brief sentence. *(3 marks)*

(d) Suppose the breaking strength for operator 3 on day 2 was entered incorrectly and should have been 600 instead of 540. How would your answer in part (b) change? *(2 marks)*

Question 2

In the crappies fishing, structures of fishing locations are important, e.g., locations with shelters are preferred. Define the event A as fishing at locations with preferred structures, and define the event B as fishing at locations without preferred structures. Assume all the other conditions are fixed. Assume the probability of catching crappies at preferred locations at a lake is 80% among all kinds of fish, and assume the probability of catching crappies at locations that are not preferred at the same lake is 20%.

a) What is the probability that crappies are caught at both locations with preferred structures l_1 and with not preferred structures l_2? Define the events before computing their respective probabilities. Assume that these events are independent. *(3 mark)*

b) Assume the probability of catching crappies at a preferred location l_1 increases to 85% if crappies are caught at another preferred location l_2. What is the probability that crappies are caught at at least one preferred locations? Define the events before computing their respective probabilities. *(3 mark)*

c) If the probability of catching crappies at locations l_1, l_2 with preferred structures and at locations l_3, l_4 without preferred structures are independent, what is the probability that crappies are successfully caught at all the locations? Define the events before computing their respective probabilities. *(4 mark)*

Question 3

A senior student, Zico, starts to manage his own banking account. His initial deposit is $2,000. Meanwhile, he starts to tutor 10 high school students in Calculus. Every week, each student contacts Zico for tutoring with a probability of $p = 0.8$. If a student contacts Zico, Zico earns $50 and deposits it into his banking account; otherwise, he needs to withdraw $20 for his living expense. Suppose that each student contacts Zico independently every week.

a) Find the probability distribution of the number of tutor sessions each student needs after 5 weeks. *(4 marks)* *(Hint: consider X_i denote the number of tutor sessions the i-th student needs for 5 weeks, and give the pmf of X_i)*

b) Find the expected number of session each student needs after 5 weeks and the corresponding variance by definition. *(3 marks)*

c) Find the expected amount of deposit in Zico's banking account after 5 weeks and the corresponding variance by definition. *(3 marks)*

Question 4

Entry to a certificate program is determined by a national test comprised of two subjects A and B. The scores on subject A are normally distributed with a mean of 80 and a standard deviation of 3, and on subject B are normally distributed with a mean of 70 and a standard deviation of 4. The scores on subject A and B are assumed to be independent.

a) Question a *(3 marks)*

What proportion of students score between 145 and 160 points on the test? (round to the nearest 2 decimal places)

b) Question b *(2 marks)*

If a student needs to score more than 155 points to be admitted to the program. For a random student, what is the probability of getting admitted? (round to the nearest 2 decimal places)

c) Question c *(2 marks)*

If the program admits students that score better than at least 70% of the students who took the test, will a student who scored 153 points be admitted to the program?

d) Question d *(3 marks)*

If two students are randomly selected, calculate the probability that the difference between their subject A scores is less than 10 points. (round to the nearest 2 decimal places)