STAT5101: Foundations of Data Science

Assignment 2

Academic year 18/19, First term

Deadline: During Class, Oct 10 (WED), 2018.

1. The dean of a business school wishes to form an executive committee of 5 from among the 40 tenured faculty members at the school. The selection is to be random, and at the school there are 8 tenured faculty members in accounting.

   a. What is the probability that the committee will contain at least 1 of the accounting faculty members?

   b. Prof. Chan is a tenured faculty member at the business school, what is the probability that Prof. Chan will be selected as a committee member.

   c. If Prof. Chan is a tenured faculty member in accounting at the business school, what is the probability that Prof. Chan will be selected as a committee member.

   d. According to a survey research, the probability a professor catches a cold during winter is 0.2. Assume 10 professors are randomly selected. Consider the random variable defined by the number of professors, among the 10 professors, that catch a cold during winter. Propose an appropriate distribution for the random variable. Based on the distribution, calculate the probability that at least 2 professors catch a cold.

2. Assume that the flaws along a magnetic tape follow a Poisson distribution with a mean of 0.1 flaw per meter. Let Y denote the distance between two successive flaws.

   a. What is the mean of Y?

   b. What is the probability that there are no flaws in 5 consecutive meters of tape?

3. Cinema advertising is increasing. According to a survey research, the probability a viewer will remember a cinema advertisement is 0.74.

   a. Suppose that 10 viewers of a cinema advertisement are randomly sampled. Consider the random variable defined by the number of viewers who recall the advertisement. What assumptions must be made in order to assume that this random variable is distributed as a binomial random variable?

   Assuming that the number of viewers who recall the cinema advertisement is a binomial random variable,

   b. what is the probability of at least two viewers who recall the advertisement?

   c. what are the mean and standard deviation of this distribution?

4. The time between arrivals of customers at a bank during the noon to 1 P.M. hour has a uniform distribution over an interval from 0 to 120 seconds. What is the probability that the time between the arrival of two customers will be

   a. between 10 and 30 seconds?

   b. What is the expected value and the standard deviation of the time between arrivals?

5. The data **UTILITY.xls** represent the electricity cost in dollars during the month of July 2002 for a random sample of 50 two-bedroom apartments in a large city. Decide whether the data appear to be approximately normally distributed by constructing a normal probability plot.

6. You are trying to set up a portfolio that consists of a corporate bond fund and a common stock fund. The following information about the annual return (per $1,000) of each of these investments under different economic conditions is avaliable along with the probability that each of these economic conditions will occur.

Probability	State of the Economy	Corporate Bonds	Common Stocks
0.10	Recession	-$40	-$120
0.20	Stagnation	60	-30
0.30	Slow growth	80	115
0.25	Moderate growth	105	170
0.15	High growth	100	230

Suppose you wanted to create a portfolio that consists of corporate bonds and common stocks. Compute the portfolio expected return and portfolio risk for proportions 0.4 invested in corporate bonds.

7. A set of final examination grades in an introductory statistics course was found to be normally distributed with a mean of 73 and a standard deviation of 8.
   (a) What percentage of students scored between 65 and 89?
   (b) Only 5% of the students taking the test scored higher that what grade?
   (c) If the professor grades on a curve (gives A's to the top 10% of the class regardless of the score), are you better off with a grade of 81 on this exam or a grade of 68 on a different exam where the mean is 62 and the standard deviation is 3? Show your answer statistically and explain.