STAT5101: Foundations of Data Science

Assignment 3

Academic year 18/19, First term

Deadline: During Class, Nov 21 (WED), 2018.

1. A population has four members (called A, B, C, and D). You would like to select a random sample of $n = 2$, which you decide to do in the following way: Flip a coin; if it is heads, the sample will be items A and B; if it is tails, the sample will be items C and D. although this is a random sample, it is not a simple random sample. Explain why.

2. The following data represent the number of days absent per year in a population of six employees of a small company:

   1 5 6 8 8 15

   Assuming that you sample with replacement, select all possible samples of $n = 2$ and construct the sampling distribution of the mean. Compute the mean of all sample means and also compute the population mean. Are they equal? What is this property called?

3. The amount of time a bank teller spends with each customer has a population mean $\mu = 3.10$ minutes and standard deviation $\sigma = 0.40$ minutes. Assume the population is symmetrically distributed, if a random sample of 16 customers is selected,

   a. What is the probability that the average time spent per customer will be at least 3 minutes?

   b. There is an 85% chance that the sample mean will be below how many minutes?

   c. If a random sample of 64 customers is selected, there is an 85% chance that the sample mean will be below how many minutes?

4. A study of women in corporate leadership was conducted by Catalyst, a New York research organization. The study concluded that slightly more than 15% of corporate officers at Fortune 500 companies are women. Suppose that you select a random sample of 200 corporate officers, and the true proportion held by women is 0.15.

   a. What is the probability that in the sample, less than 15% of the corporate officers will be women?

   b. What is the probability that in the sample, between 13% and 17% of the corporate officers will be women?

5. Do ringing cell phones disturb business presentations? In a poll of 326 business men and women, 303 answered this question "yes" and only 23 answered "no".

   a. Construct a 95% confidence interval for the population proportion of business men and women who have their presentations disturbed by cell phones.

   b. Interpret the interval constructed in (a).

   c. If you were to conduct a follow-up study that would provide 95% confidence that the point estimate is correct to within ± 0.04 of the population proportion, how large a sample size would be required?

6. The manager of a paint supply store wants to estimate the actual amount of paint contained in 1-gallon cans purchased from a nationally known manufacturer. It is known from the manufacturer's specifications that the standard deviation of the amount of paint is equal to 0.02 gallon. A random sample of 50 cans is selected, and the sample mean amount of paint per 1-gallon can is 0.995 gallon.

   a. Set up a 99% confidence inetrval estimate of the true population mean amount of paint included in a 1-gallon can.

   b. On the basis of your result in (a), do you think that the manager has a right to complain to the manufacturer? Why?

   c. Does the population amount of paint per can have to be normally distributed here? Explain.

7. A consumer group wants to estimate the mean electric bill for the month of July for single-family homes in a large city. Based on studies conducted in other cities, the standard deviation is assumed to be \$25. The group wants to estimate the mean bill for July to within $\pm$ \$4 of the true average with 99% confidence. What sample size is needed?