NOVA SCHOOL OF BUSINESS & ECONOMICS

2494 - COMPUTATIONAL THINKING & DATA SCIENCE

2021-22, Spring Semester

Week 5

Deadline: March, 9

1. A manufacturing company's quality control personnel have recorded the proportion* of defective items for each of 500 monthly shipments of one of the computer components that the company produces. The data are in file **quality.xlsx**. The quality control department manager does not have sufficient time to review all of these data. Rather, she would like to examine the proportion of defective items for a sample of these shipments.

   Write a program that generate a simple random sample of size 25 and construct a 95% confidence interval for the mean proportion of defective items over all monthly shipments.

   Does the 95% CI contain the actual population mean in this case? What proportion of many similarly constructed confidence intervals should include the population mean?

   * number of defective items in the shipment divided by the total number of items in the shipment.

2. A market research consultant hired by a leading soft drink company wants to determine the proportion the proportion of consumers who favor its low-calorie brand over the leading low-calorie competitor in a particular geographic region. A random sample of 250 consumers from the market in investigation is provided in **softdrink.xlsx**.

   Write a program that finds out a 90% confidence interval for the proportion of all consumers in this market who prefer the company's brand. The files contains the gender and age group for each customer in the sample. Find a separate 90% confidence interval for each gender for the proportion who prefer the company's brand. Then do the same for each age group. Compare the CIs obtained.