Tutorial 8

1. The number of pages in magazines devoted to advertisements varies widely from magazine- to-magazine and from issue-to-issue within the same magazine. Advertising expenditures, and therefore the number of advertising pages in magazines, tend to be the highest during periods of economic growth. The data in the file `weeklies.txt` gives the number of advertising pages in the current issues of 19 weekly magazines and the number of advertising pages in the same issue of the previous calendar year.

   (a) At the 5% level of significance, is there evidence of a difference in the mean number of advertising pages in the current issues compared to the previous year? Set up a 95% confidence interval estimate of the mean difference in the number of advertising pages in the current issues compared to the previous year. Use SAS.

   (b) What assumptions did you make? Use SAS to check these assumptions.

   (c) Repeat the two questions above using R and Python.

   (d) In R and Python, write the code to form a 99% CI for the mean of difference in the number of advertising pages in the current issues compared to the previous year. Compare this 99% CI vs the 95% CI derived above, which interval is wider? Explain the reason why it is wider.

2. The purchasing director for an industrial parts factory is investigating the possibility of purchasing a new type of milling machine. He determines that the new machine will be bought if there is evidence that the parts produced have a higher average breaking strength than those from the old machine. The data file `machine.txt` represents the breaking strength of samples of 50 parts from the old and the new machines.

   Use SAS to find out if there is evidence that the purchasing director should buy the new machine. What assumptions did you make? Check these assumptions.

   Repeat using R and Python.

3. In many manufacturing processes the term "work-in-process" (often abbreviated WIP) is used. In a book manufacturing plant the WIP represents the time it takes for sheets from a press to be folded, gathered, sewn, tipped on end sheets, and bound. The data set `wip.txt` represents samples of 20 books at each of two production plants and the processing time (defined as the time in days from when the books came off the press to when they were packed in cartons) for these jobs.

   Use a nonparametric method in SAS to determine if there is a difference between the processing times for the two plants.

   Repeat using R and Python.