NATIONAL UNIVERSITY OF SINGAPORE
Department of Statistics and Applied Probability

2018/19 Semester 2 **ST2137 Computer Aided Data Analysis** **Tutorial 6**

1. 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.
   (a) Use SAS to find out if there is evidence that the purchasing director should buy the new machine. Use a 5% significance level.
   (b) What assumption on the variances did you make for the test in part (a)? Check the assumption.
   (c) Repeat (a) and (b) using R
   (d) Repeat (a) and (b) using SPSS.

2. (Refer to Question 2 in Tutorial 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 cartoons) for these jobs.
   (a) Using a nonparametric method in SAS to determine if there is a difference between the median processing times for the two plants. Use a 5% significance level.
   (b) Repeat (a) using R.
   (c) Repeat (a) using SPSS.

3. 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" (free format with comma as delimiter) 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 precious 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) Repeat (a) using R
   (c) Repeat (a) using SPSS.

4. A flexible working hour program permits employees to design their own 42-hour work week to meet their personal needs. The management of a large manufacturing firm may adopt a flextime program for its administrators and professional employees, depending on the success or failure of a pilot program. Ten employees were randomly selected and given a questionnaire designed to measure their attitude toward their jobs. Each was then permitted to design and follow a flextime workday. After six months, attitudes toward their jobs were again measured. The resulting attitude scores are given in the data file "flextime.txt" (free format with space as delimiter). The higher the score, the more favorable the employee's attitude toward his or her work.
   (a) Use a nonparametric test procedure in SAS to evaluate the success of the pilot flextime program. Use a 5% significance level.

(b) Repeat (a) using R
(c) Repeat (a) using SPSS.

Answers to selected questions

1. (a) H_0: $\mu_N - \mu_O = 0$ against H_1: $\mu_N - \mu_O > 0$. (Assume equal variances). $t_{obs} = 3.93 > t_{0.05}(98) = 1.66$ or p-value = $7.9(10)^{-5}$. Reject H_0
   (b) Equal variances. H_0: $\sigma_N^2 = \sigma_O^2$, against H_1: $\sigma_N^2 \neq \sigma_O^2$. $F_{obs} = 1.26 < F_{0.05}(49, 49) = 1.26$ or p-value = 0.4268. Do not reject H_0.
   SAS code for (a) and (b) (Refer to p8.17-8.19)
   R code for (a) and (b) (Refer to p8.20-8.22)
   SPSS code for (a) and (b) (Refer to p8.23-8.24)
2. H_0: $\mu_1 - \mu_2 = 0$ against H_1: $\mu_1 - \mu_2 \neq 0$. Wilcoxon rank sum test, p-value = 0.0822. Do not reject H_0.
   SAS code (Refer to p8.28-8.30)
   R code (Refer to p8.31)
   SPSS (Refer to p8.32-8.35)
3. H_0: $\mu_{Diff} = 0$ against H_1: $\mu_{Diff} \neq 0$. $|t_{obs}| = 1.85 < t_{0.025}(18) = 2.10$ or p-value = 0.0803. Do not reject H_0. 95% CI for μ_{Diff}: (−21.419, 1.340).
   SAS code (Refer to p8.37-8.38)
   R code (Refer to p8.39)
   SPSS (Refer to p8.40-8.41)
4. Test H_0: $\mu_{Diff} = 0$ against H_1: $\mu_{Diff} > 0$. Signed rank test: p-value = 0.0068. Reject H_0.
   SAS code (Refer to p8.45-8.46)
   R code (Refer to p8.47-8.49)
   SPSS (Refer to p8.50-8.52)