

STA290 Homework 2

Due September 25

All solutions should be typed in L^AT_EX.

1. Write a function to calculate (approximate) the permutation distribution of the Welch t.test. Conduct a simulation study to compare the p-values obtained under the permutation distribution to those using Satterthwaite's approximation in the function t.test and the Wilcoxon Rank Sum test. In the simulation study generate data under the null hypothesis (equal means) and with unequal sample sizes and variances i.e. $Y_i \sim N(0, 1)$, for $i = 1, \dots, 5$, and $X_i \sim N(0, 5)$, for $i = 1, \dots, 20$ with the sample sizes not too large. Try for around 100 simulations. Include your program and results from your simulation study. *The verbatim environment in L^AT_EX is useful for typesetting computer code*
2. The data for this problem consist of lengths of time (in seconds) taken by subjects to see a three-dimensional object (a spiral ramp coming out of the page) in a random dot stereogram. 42 out of the 77 subjects received at most verbal prior information only (NV), while the remaining 35 subjects received both visual and verbal prior information (VV). Construct normal probability plots of the NV and VV observations. Find a suitable power transformation that makes the distribution of the transformed data approximately normal. Use and contrast various appropriate statistical procedures to answer whether there is a significant difference between the training procedures. Using your permutation test from above, plot the histogram representing your estimate of the permutation distribution and the approximate t distribution using the Satterthwaite approximation. How well do they agree? Provide a one page write up of your analysis, including hypotheses tested, assumptions and conclusions (including any appropriate figures and tables). *Data are under Assignments link in stereo.dat*
3. Researchers randomly assigned metastatic breast cancer patients to either a control group or a group that received weekly 90-minute sessions of group therapy and self-hypnosis to see whether treatment improved the patients quality of life. The group therapy involved discussion and support for coping with the disease, but the patients were not led to believe that the therapy would affect the prognosis of their disease. Surprisingly, it was noticed in follow-up 10 years later, that the group therapy patients appeared to have lived longer. The data on the number of months of survival after beginning of the study are shown. Three of the women in the treatment group were still alive at the time of follow up, so their survival times are only known to be larger than 122 months. Is there evidence of an effect of the group therapy treatment on survival time, and if so, how much more can a breast cancer patient expect to live if she receives this therapy? Analyze the data as best as possible given the methods covered, and write a brief report on your findings. (limit to 2 pages with any supporting figures or tables)

Survival Times for Patients in Control Group: (n=24)

2, 6, 8, 10, 12, 12, 14, 14, 14, 16, 16, 16, 18, 18, 18, 20, 22, 22, 26, 34, 36, 38, 40, 48

Survival Times for Group Therapy Patients: (n=34)

2, 2, 4, 4, 4, 6, 6, 8, 10, 10, 12, 14, 16, 16, 18, 20, 22, 32, 36, 46, 46, 48, 48, 48, 58, 58, 66, 72, 72, 82, 122, 122*, 122*, 122*

** these patients were still alive at the end of the 122 month study period*