Exercise : Unit 3 (institute)

1. Pick up a data set which is a single factor and completely randomized experiment. Besides, the data set should satisfy the conditions specified by (c.) and (d.).

Analyze the data and answer the following questions:

a. Construct a graphical display to show whether the factor affects the response. What is your conclusion?

b. State the ANOVA model you used in this problem.

c. Use the ANOVA F-test to the original data and the ranks to test the hypothesis that the factor affects the response. Make your conclusion. (the data set you selected should reject the null hypothesis)

d. Check the adequacy of the model. (the data set you selected must satisfy the assumptions: Normality and Constant variance assumptions)
   i. Use the graphical displays to check the assumption.
   ii. Use the hypothesis testing method to check the assumption. (you should mention the name of the testing method)

e. Construct a graphical display (t-distribution with suitable scale factor) to compare the treatment means for the factor. What is your conclusion?

f. Use the four multiple comparison methods to make comparisons between pairs of means. Does you get the same conclusions from these tests and the graphical procedure? If there is any difference, explain it.

g. For any pair of means that are significant different from above multiple comparisons, compute the 95% confidence interval estimates of the mean difference

h. Apply the Krustal-Wallis test to show whether the factor affects the response. Does you get the same conclusions from the usual ANOVA F-test?

i. Fit the orthogonal polynomial for this data. (regression analysis)

j. If we wish to detect a maximum difference in mean response of δ (specified yourself) with a probability of at least 0.9, what sample size should be used? How would you obtain a preliminary of the population variance.

k. If the difference between treatment means is great enough so that the standard deviation of an observation is increased by 25%, what sample size should be used if we wish to detect this with a probability of at least 0.9?

l. If we with to construct a 95% confidence interval of the difference in two treatment means that has an accuracy of δ (specified yourself), how many sample size must be used?
2. Select a data set in which the constant variance assumption or normal assumption is violated.
   a. Use the ANOVA F-test to the original data and the ranks to test the hypothesis that the factor affects the response. Make your conclusion.
   b. Check the adequacy of the model.
   c. Do transformation.
   d. Repeat (a) and (b) using the transformed data.