Homework 6

The level of significance α=0.05. You need to (i) write down the null and hypothesis; (ii) the corresponding SAS or R codes; (iii) your conclusion

For Problems 1 and 2, t-test is used.

1. A clinical dietician wants to compare two different diets, A and B, for diabetic patients. She hypothesizes that diet A (Group 1) will be better than diet B (Group 2), in terms of lower blood glucose. She plans to get a random sample of diabetic patients and randomly assign them to one of the two diets. At the end of the experiment, which lasts 6 weeks, a fasting blood glucose test will be conducted on each patient. She also expects that the average difference in blood glucose measure between the two groups will be about 8 mg/dl. Furthermore, she also assumes the standard deviation of blood glucose distribution for both groups to be 16. The dietician wants to know the number of subjects needed in each group assuming equal sized groups. What distribution assumption does she need to make for the blood glucoses for each group? Find the corresponding sample sizes needed for statistical powers being 0.5, 0.6, 0.7, 0.8, and 0.9.
2. An audiologist wanted to study the effect of gender on the response time to a certain sound frequency. He suspected that men were better at detecting this type of sound then were women. He took a random sample of 20 male and 20 female subjects for this experiment. Each subject was be given a button to press when he/she heard the sound. The audiologist then measured the response time - the time between the sound was emitted and the time the button was pressed. Now, he wants to know what the statistical power is based on his total of 40 subjects to detect the gender difference assuming the standard deviations is 0.2 second. What distribution assumption does she need to make for response time for each group? Find the corresponding statistical powers when the mean difference is 0.05, 0.1, 0.15, 0.2, and 0.25 second.

For Problems 3 and 4, Wilcoxon-Mann-Whitney Test is used.

1. With the same set up as problem 1, but the dietician suspects the distribution assumptions in Problem 1 is not valid. She feels the blood glucose in diet A follows uniform distribution [50, 80] and the blood glucose in diet B follows N(75,100). Find the corresponding sample sizes needed for statistical powers being 0.5, 0.6, 0.7, 0.8, and 0.9.
2. With the same set up as problem 2, but the audiologist suspects the distribution assumption in Problem 2 is not valid. He thinks it is better to assume that the times in male group follows gamma distribution with shape parameter α=2 and scale parameter λ=0.05; and the times in female group follows exponential distribution with scale parameter λ=0.2. Find the corresponding statistical power when there are n subjects in each group for n=10, 20, 30, and 40
3. In a study of the percentage of raw material that responds in a reaction, researchers identified the following five factors: the feed rate of the chemicals (FeedRate), ranging from 10 to 15 liters per minute; the percentage of the catalyst (Catalyst), ranging from 1% to 2%; the agitation rate of the reactor (AgitRate), ranging from 100 to 120 revolutions per minute; the temperature (Temperature), ranging from 140 to 180 degrees Celsius; the concentration (Concen), ranging from 3% to 6%. The dataset is as following:

FeedRate Catelyst AgitRate Temp Concen ReactionPercentage

10.0 1.0 100 140 6.0 37.5

10.0 1.0 120 180 3.0 28.5

10.0 2.0 100 180 3.0 40.4

10.0 2.0 120 140 6.0 48.2

15.0 1.0 100 180 6.0 50.7

15.0 1.0 120 140 3.0 28.9

15.0 2.0 100 140 3.0 43.5

15.0 2.0 120 180 6.0 64.5

12.5 1.5 110 160 4.5 39.0

12.5 1.5 110 160 4.5 40.3

12.5 1.5 110 160 4.5 38.7

12.5 1.5 110 160 4.5 39.7

1. The researchers feel the reaction percentage can be fitted by a linear model with the five factors as independent variables. Write the correspondent models and fit the models.
2. Is there any variable not significant? If yes, remove the insignificant variable(s), and refit the model, compare the adjusted r-square.
3. Write down the final model.
4. Give the 95% confidence interval of the parameter
5. Find the 95% confidence interval of the expected mean of the first observation
6. Find the 95% confidence interval of the expected mean and the predicted value when FeedRate=10; Catelyst=2.0; AgitRate=120; Temp=130; Concen=7