**Applied Statistics and Design of Experiments**

**Exercise 4**

**February 6, 2018**

*For these exercises you should work in groups of 3-5 students. Part of homework 5 will involve working with your group to refine and extend your answers to some of these questions.*

**Data set: ‘defects.csv’**

The data are from an experiment to compare 4 processing methods for manufacturing steel ball bearings. The 4 process methods were run for one day and a random sample of 1% of the ball bearings from the day was taken from each of the 4 methods. Because the processes produce ball bearings at different rates the sample sizes were not the same for the 4 methods. Each sampled ball bearing had its weight measured to the nearest 0.1 g and the number of surface defects was counted. The variables in the data set are:

Sample: sample number

Method: A, B, C, or D

Defects: number of defects

Weight: weight in g

**Question 1.**

The target weight for the ball bearings is 10 g. For each of the 4 methods test the null hypothesis that the mean weight is equal to 10. Considering the possibility of inflation of type I error rate due to multiple testing, what do you conclude from these results?

Perform all pairwise comparisons of mean weight for the different methods. Report the p-values from all tests. Considering the possibility of inflation of type I error rate due to multiple testing, what conclusions would you draw from these results?

**Question 2**

Compare the mean weights for the 4 methods using ANOVA. What is your conclusion? How does it compare to the conclusions from the pairwise comparisons?

**Question 3**

Assess the assumptions of the ANOVA. Are the assumptions met?

**Question 4**

Perform a simulation study to assess the validity of the ANOVA. Choose distributions for weight that seem to be reasonable based on your analysis of the data. Report the estimated type I error rate of the ANOVA F-test.