Week 6 HW Submission: ANOVA etc.

Your name here

Oct. ??, 2015

Create a Word docx from this R Markdown file for the following exercise. Submit the R markdown file and resulting Word docx file via D2L Dropbox.

## Exercise

This exercise is based on the data and experimental design from exercises 8.42 & 8.43 in the Ott textbook.

A small corporation makes insulation shields for electrical wires using three different types of machines. The corporation wants to evaluate the variation in the inside diameter dimension of the shields produced by the machines. A quality engineer at the corporation randomly selects shields produced by each of the machines and records the inside diameters of each shield (in millimeters). The goal is to determine whether the location parameters (i.e. mean or median) of the three machines differ.

Load the data set shields from the DS705data package (alternately there is shields.rda file in the weekly download folder that can be loaded using the load() command)

### Part 1

Construct side-by-side boxplots for the inside diameters of the insulation shields for the three machines.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1 -|-|-|-|-|-|-|-|-|-|-|-

require(DS705data)

## Loading required package: DS705data

data(shields)

### Part 2

Comment on what you see in the boxplots. How do the medians compare visually? Do the samples look like they have roughly the same variability? Is there severe skewness or outliers in any of the samples? Be specific.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 2 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 3

Which data conditions for ANOVA appear not to be met, if any? Provide reasoning for your answer.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 4

Conduct an analysis of variance test (the standard one that assumes normality and equal variance). (i) State the null and alternative hypotheses, (ii) use R to compute the test statistic and p-value, and (iii) write a conclusion in context at .

### -|-|-|-|-|-|-|-|-|-|-|- Answer 4 -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 5

Conduct an analysis of variance test with the Welch correction. (i) State the null and alternative hypotheses, (ii) use R to compute the test statistic and p-value, and (iii) write a conclusion in context at .

### -|-|-|-|-|-|-|-|-|-|-|- Answer 5 -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 6

Which data conditions for Welch ANOVA are not met, if any? Provide reasoning for your answer.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 6 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 7

Conduct a Kruskal-Wallis test. (i) State the null and alternative hypotheses, (ii) use R to compute the test statistic and p-value, and (iii) write a conclusion in context using .

### -|-|-|-|-|-|-|-|-|-|-|- Answer 7 -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 8

Which data conditions for the Kruskal-Wallis test are not met, if any? Provide reasoning for your answer.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 8 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 9

Conduct a bootstrapped ANOVA test using pooled residuals and unequal variances as in the notes. (i) State the null and alternative hypotheses, (ii) use R to compute the test statistic and p-value, and (iii) write a conclusion in context . Do not use a helper function, instead mimic the code in the notes using a for loop to construct the boostrapped sampling distribution.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 9 -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 10

Repeat the bootstrapped ANOVA test using unpooled residuals and unequal variances as in the notes. (i) State the null and alternative hypotheses, (ii) use R to compute the test statistic and p-value, and (iii) write a conclusion in context . Go ahead and use the helper function or t1waybt do do this problem.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 10 -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 11

Which seems better and why, the bootstrap procedure with the pooled or unpooled residuals?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 11 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer.

### Part 12

Do any of the four statistical inference procedures used here provide a clear answer to the question of whether or not the three machines produce the same average inside diameter for the insulation shields?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 12 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 13

If you were responsible for conducting the statistical analysis here, what would you report to the engineer?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 13 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 14

What impact do you think samples of sizes 5, 5, and 10 play here?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 14 -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 15

Often the Kruskall Wallis test is presented as a test of

the population distributions are all the same

the population distributions are not all the same,

but this is not what KW tests as this example shows. Take 3 random samples of size 100 from normal distributions having mean 0 and standard deviations 1, 10, and 50. If KW were testing the hypotheses above, then we should reject since these three distributions are clearly different. Run the KW test. You should get a large -value. Why did you get a large -value when the distributions are so different?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 15 -|-|-|-|-|-|-|-|-|-|-|-

set.seed(321)  
x <- c( rnorm(100,0,1), rnorm(100,0,10), rnorm(100,0,50))  
groups <- factor( (rep( c('A','B','C'), each=100 ) ) )  
# complete the code here ...

Replace this text with your explanation.