Week 6 HW Submission: ANOVA etc.

Paul Isaacson

Feb. 21, 2017

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.

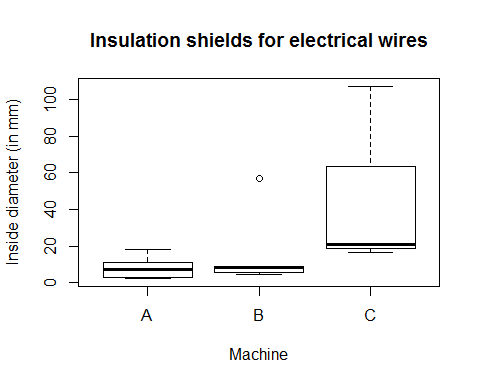
Load the data set shields from the DS705data package (alternately there is shields.rda file in the weekly download folder that can be loaded using the load() command)

### Part 1

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

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

require(DS705data)  
data(shields)  
with(shields, boxplot(Diameter ~ Machine, main='Insulation shields for electrical wires', xlab='Machine', ylab='Inside diameter (in mm)'))



### 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 -|-|-|-|-|-|-|-|-|-|-|-

The medians for the samples from machines A and B are very similar. The median for the sample from machine C is approximately twice a large. The variance for the samples from machine A, machine B, and machine C have different variability. The variability for the sample from machine B has the least variability followed be the sample from machine A. The variability for the sample from machine C is the greatest. As for skewness, all three samples are right-skewed. The sample from machine A is only slightly skewed. The samples from machines B contain outliers leading to a right-skew. The sample from machine C is substantially more skewed.

### Part 3

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

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

The condition of equal variance does not appear to be met. Using the rule of thumb, the max sd (34.52) divided by the min sd (6.52) is over 2 (5.29).

### 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 -|-|-|-|-|-|-|-|-|-|-|-

fit <- aov(Diameter~Machine, data=shields)  
summary(fit)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Machine 2 4141 2070.5 2.727 0.0939 .  
## Residuals 17 12907 759.3   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

1. Do not reject at (P = 0.0939). There is not sufficient evidence to support that not all the means are the same for the populations of shields produced by machines a, machines b, and machines c.

### 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 -|-|-|-|-|-|-|-|-|-|-|-

oneway.test(Diameter~Machine, data=shields, var.equal=FALSE)

##   
## One-way analysis of means (not assuming equal variances)  
##   
## data: Diameter and Machine  
## F = 3.974, num df = 2.0000, denom df = 8.4087, p-value = 0.06096

1. Do not reject at (P = 0.0939). There is not sufficient evidence to support that not all the means are the same for the populations of shields produced by machines a, machines b, and machines c.

### Part 6

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

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

While robust, the ANOVA test does require that the condition of that the population approximates a normal distribution. This breaks down when the distribution has heavy tails or the distribution is very skewed. For the sample from machine c, the distribution is very right-skewed.

### 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 -|-|-|-|-|-|-|-|-|-|-|-

kruskal.test(Diameter ~ Machine, data=shields)

##   
## Kruskal-Wallis rank sum test  
##   
## data: Diameter by Machine  
## Kruskal-Wallis chi-squared = 9.8914, df = 2, p-value = 0.007114

1. Reject at (P = 0.007114). There is sufficient evidence to support that not all the medians are the same for the populations of shields produced by machines a, machines b, and machines c.

### Part 8

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

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

It does not appear that the population distributions of the three machines are the same shape. The condition of equal variance does not appear to be met. Using the rule of thumb, the max sd (34.52) divided by the min sd (6.52) is over 2 (5.29).

### 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 -|-|-|-|-|-|-|-|-|-|-|-

F.obs <- oneway.test(Diameter ~ Machine, data=shields)$statistic  
groups <- shields$Machine  
resA <- with(shields, Diameter[Machine=='A'] - mean(Diameter[Machine=='A']))  
resB <- with(shields, Diameter[Machine=='B'] - mean(Diameter[Machine=='B']))  
resC <- with(shields, Diameter[Machine=='C'] - mean(Diameter[Machine=='C']))  
#pop.null <- data.frame(res=c(resA, resB, resC), groups)  
#with(pop.null, tapply(res, groups, mean))  
  
B <- 10000   
Fstar1 <- numeric(B)  
for (i in 1:B){  
 res <- sample(c(resA, resB, resC), replace=TRUE)   
 pop.null <- data.frame(res, groups)  
 Fstar1[i] <- oneway.test(res ~ groups, data=pop.null,   
 var.equal=FALSE)$statistic  
}  
p.approx1 <- sum(Fstar1 > F.obs) / B  
p.approx1

## [1] 0.0576

1. Do not reject at (P = 0.0576). There is not sufficient evidence to support that not all the means are the same for the populations of shields produced by machines a, machines b, and machines c.

### 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 -|-|-|-|-|-|-|-|-|-|-|-

source('anovaResampleFast.R')  
with(shields, anovaResampleFast(Diameter, Machine))

## [1] "Some bootstrap estimates of the test statistic could not be computed"  
## [1] "Effective number of bootstrap samples was 995"  
## [1] "Assuming unequal variances - using Welch corrected F"  
## [1] "observed F: 3.97403265249524"  
## [1] "observed p-value: 0.0609641818408098"  
## [1] "resampled p-value: 0.346733668341709"

## $f.obs  
## [1] 3.974033  
##   
## $p.obs  
## [1] 0.06096418  
##   
## $p.boot  
## [1] 0.3467337

1. Do not reject at (P = 0.3467337). There is not sufficient evidence to support that not all the means are the same for the populations of shields produced by machines a, machines b, and machines c.

### Part 11

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

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

I am more inclined to use the bootstrap procedure with the pooled residuals. This is because it matches more closely with the other tests we have run. Additionally, the sample sizes are small and we need larger samples to overcome the outliers from the sample from machine B.

### 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 -|-|-|-|-|-|-|-|-|-|-|-

All four are problematic. Standard ANOVA Breaks down because the variances are not equal. Welch corrected ANOVA Is problematic because the distributions are skewed. The Kruskal-Wallis test is problematic because the three samples do not appear to come from populations with the same shape. Finally, the bootstrap procedure is less than ideal due to the small sample sizes.

### Part 13

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

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

I would say we do not have strong evidence to suggest the mean inside diameters are different between the three samples.

### Part 14

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

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

The small sample sizes make the results much less conclusive.

### 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 ) ) )  
kruskal.test(x ~ groups)

##   
## Kruskal-Wallis rank sum test  
##   
## data: x by groups  
## Kruskal-Wallis chi-squared = 0.84306, df = 2, p-value = 0.656

The Kruskal-Wallis test can only detect when the medians are different. In this case, the medians are the same but the variability and shape of each distributions are different.