Week 6 HW Submission: ANOVA etc.

Christian Acosta

Jul 04, 2016

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 -|-|-|-|-|-|-|-|-|-|-|-

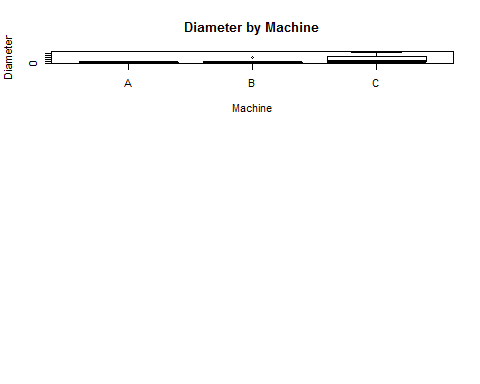
require(DS705data)

## Loading required package: DS705data

data(shields)  
shields

## Machine Diameter  
## 1 A 18.1  
## 2 A 2.4  
## 3 A 2.7  
## 4 A 7.5  
## 5 A 11.0  
## 6 B 8.7  
## 7 B 56.8  
## 8 B 4.4  
## 9 B 8.3  
## 10 B 5.8  
## 11 C 29.7  
## 12 C 18.7  
## 13 C 16.5  
## 14 C 63.7  
## 15 C 18.9  
## 16 C 107.2  
## 17 C 19.7  
## 18 C 93.4  
## 19 C 21.6  
## 20 C 17.8

attach(shields)  
par(mfrow=c(3,1))  
boxplot(Diameter~Machine, main="Diameter by Machine", ylab="Diameter", xlab="Machine")  
  
par(mfrow=c(1,3))



### 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 are relatively close between the three machines. However, machine 'C' has large variability and is overall much larger in diameter than machines A and B. Machine C is also heavily skewed to the right, while A and B are so tightly packed it's hard to tell (a histogram will be able to help).

### Part 3

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

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

Both the normality and distribution are in question here. Machines A & C have right skewness and machine B has a tight cluster of values with one major outlier. In addition, the spread in values of Machine C virtually guarantees that variance is unequal between all three machines.

### 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. Null Hypothesis: The population means are equal between Machines A, B & C. Alternative Hypothesis: The population means are NOT equal between Machines A, B & C.

#machine.model <- lm(Diameter~Machine)  
#summary(machine.model)  
#anova(machine.model)  
oneway.test(Diameter~Machine, data=shields, var.equal=T)

##   
## One-way analysis of means  
##   
## data: Diameter and Machine  
## F = 2.727, num df = 2, denom df = 17, p-value = 0.09393

#mean(Diameter[Machine=='C'])  
#shields

1. Fail to reject null hypothesis (alpha = 0.05).

## There is not enough evidence to suggest that population means (diameter) are different between Machines A, B & C(P=0.094).

### 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. Null Hypothesis: The population means are equal between Machines A, B & C. Alternative Hypothesis: The population means are NOT equal between Machines A, B & C.

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

##   
## One-way analysis of means (not assuming equal variances)  
##   
## data: Diameter and Machine  
## F = 3.974, num df = 2.000, denom df = 8.409, p-value = 0.06096

1. Fail to reject null hypothesis (alpha = 0.05).

There is not enough evidence to suggest that population means (diameter) are different between Machines A, B & C(P=0.061).

### Part 6

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

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

## Welch ANOVA assumes data normality. The previous boxplots suggest that the data is very different between machines A, B & C. A Shapiro test could be used to test for data normality.

### 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. Null Hypothesis: The population medians are equal between Machines A, B & C. Alternative Hypothesis: The population medians are NOT equal between Machines A, B & C.

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 null hypothesis (alpha = 0.05).

## There is enough evidence to suggest that population medians (diameter) are different between Machines A, B & C (P=0.007).

### Part 8

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

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

The Kruskal-Wallis test requires equal variances. Machine C has a much higher standard deviation than machines A & B. This could be tested using a levene test.

### 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. Null Hypothesis: The population means are equal between Machines A, B & C. Alternative Hypothesis: The population means are NOT equal between Machines A, B & C.

F.machines <- oneway.test(Diameter~Machine, data = shields)$statistic  
F.machines

## F   
## 3.974033

#F=3.974  
resA <- Diameter[Machine=='A'] - mean(Diameter[Machine=='A'])  
resB <- Diameter[Machine=='B'] - mean(Diameter[Machine=='B'])  
resC <- Diameter[Machine=='C'] - mean(Diameter[Machine=='C'])  
pop.null <- data.frame(res=c(resA, resB, resC), Machine)  
  
with(pop.null, tapply(res, Machine, mean))

## A B C   
## 5.329071e-16 -3.552714e-16 2.131975e-15

pop.null

## res Machine  
## 1 9.76 A  
## 2 -5.94 A  
## 3 -5.64 A  
## 4 -0.84 A  
## 5 2.66 A  
## 6 -8.10 B  
## 7 40.00 B  
## 8 -12.40 B  
## 9 -8.50 B  
## 10 -11.00 B  
## 11 -11.02 C  
## 12 -22.02 C  
## 13 -24.22 C  
## 14 22.98 C  
## 15 -21.82 C  
## 16 66.48 C  
## 17 -21.02 C  
## 18 52.68 C  
## 19 -19.12 C  
## 20 -22.92 C

B <- 10000; Fstar1 <- numeric(B)  
for (i in 1:B){  
 pop.null <- data.frame(res = sample(c(resA, resB, resC), replace = T), Machine)  
 Fstar1[i] <- oneway.test(res~Machine, data=pop.null, var.equal = T)$statistic  
}  
p.approx1 <- sum(Fstar1 > F.machines)/B  
p.approx1 #.0344.

## [1] 0.0335

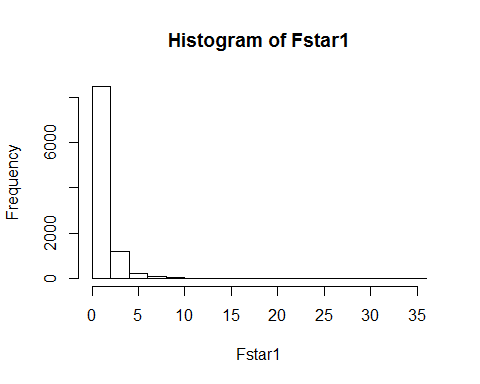
resA; resB; resC

## [1] 9.76 -5.94 -5.64 -0.84 2.66

## [1] -8.1 40.0 -12.4 -8.5 -11.0

## [1] -11.02 -22.02 -24.22 22.98 -21.82 66.48 -21.02 52.68 -19.12 -22.92

hist(Fstar1)



1. Reject null hypothesis (alpha = 0.05).

## There is enough evidence to suggest that population means (diameter), using the boostrap ANOVA method, are different between Machines A, B & C (P=0.0344).

### 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. Null Hypothesis: The population means are equal between Machines A, B & C. Alternative Hypothesis: The population means are NOT equal between Machines A, B & C.

source('anovaResampleFast.R')  
out1 <- anovaResampleFast(Diameter, Machine, B=10000, method=1, var.equal=F)

## [1] "Assuming unequal variances - using Welch corrected F"  
## [1] "observed F: 3.97403265249524"  
## [1] "observed p-value: 0.0609641818408098"  
## [1] "resampled p-value: 0.0549"

1. Fail to reject null hypothesis (alpha = 0.05).

## There is not enough evidence to suggest that population means (diameter), using the boostrap ANOVA method, are different between Machines A, B & C (P=0.0568).

### Part 11

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

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

## I trust the unpooled residuals more because the original machine data was very diverse (different shapes/scales). In addition, the sample sizes were relatively small - how could we be confident that the original sample, which was used to bootstrap data, was representative of the overall population?

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

## No. Given the data appears to be non-normal with unequal variance, the ANOVA, ANOVA Welch correction and Kruskal-Wallis tests are not appropriate. The bootstrapping method is the most appropriate test for the data, however the findings were inconclusive using the unpooled bootstrapping method (p=.057). Given the heavy skewness of the data, and how diverse it is, the trimmed mean method could also be used to evaluate the data.

### Part 13

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

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

## The results are inconclusive at the moment. I would suggest obtaining a larger sampling of data and re-running test methods based on the new data.

### Part 14

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

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

## The small sample sizes play a large role in relaying conclusive results. Evidence suggests that there is a different in diameter between machines, but a larger sample size is needed to produce more definitive results.

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

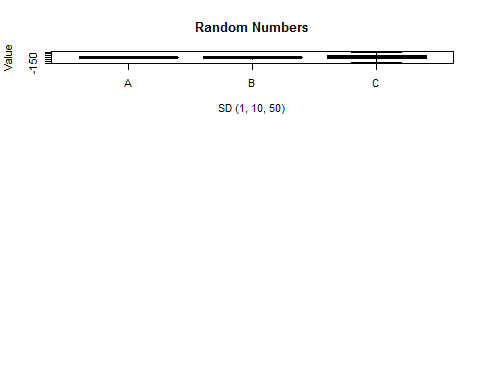
set.seed(321)  
x <- c( rnorm(100,0,1), rnorm(100,0,10), rnorm(100,0,50))  
groups <- factor( (rep( c('A','B','C'), each=100 ) ) )  
groups

## [1] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A  
## [36] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A  
## [71] A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A B B B B B  
## [106] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B  
## [141] B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B B  
## [176] B B B B B B B B B B B B B B B B B B B B B B B B B C C C C C C C C C C  
## [211] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C  
## [246] C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C C  
## [281] C C C C C C C C C C C C C C C C C C C C  
## Levels: A B C

kruskal.test(x~groups)

##   
## Kruskal-Wallis rank sum test  
##   
## data: x by groups  
## Kruskal-Wallis chi-squared = 0.8431, df = 2, p-value = 0.656

par(mfrow=c(3,1))  
boxplot(x~groups, main="Random Numbers", ylab="Value", xlab="SD (1, 10, 50)")



## The Kruskal-Wallis test aims to answer "are the medians different?" Ultimately, the test is ranking pooled data and computing the average rank for the samples. When comparing the mean ranks, because scale is not factored in, a sampling of random numbers will usually have similar distributions around the median.