ANOVA etc.

Matthew Allen

07/26/2018

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 1

In the Lesson 3 presentation you saw how to use the Wilcoxon Rank Sum test to compare the difference in median repair times for Macs and PCs. You’ll find the repair dataset in the DS705data package. In this problem we’ll test the hypothesis that the population mean repair times are different for Macs and PCs at the 5% significance level using three different approaches.

### Part 1a

Even though repair times for both computer types are skewed right go ahead and use t.test to test if the population mean times are significantly different. Include your R code below and write a conclusion to the test for practice.

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

# perform t-test to check whether mean repair time for mac and pc are significantly different.  
mac.repair.time <- repair$time[which(repair$type=="Mac")]  
pc.repair.time <- repair$time[which(repair$type=="PC")]  
t.test(mac.repair.time, pc.repair.time, alternative = 'two.sided')$p.value

## [1] 0.0550238

At a 5% level of significance, we fail to reject the hypotheis that the mean repair times for Macs and PCs are different.

### Part 1b

Now use the boot package to construct a 95% BCa confidence interval for the difference in population mean repair times. Use at least 5000 resamples. Use that confidence interval to write a hypothesis test conclusion to this hypothesis test. (Review: you made similar bootstrap CI’s in Lesson 3.)

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

# Construct a 95% BCa confidence interval for the difference in population mean repair times.  
library(boot)  
bootMeanDiff <- function(d,i){  
 means <- tapply(d[i,1],d[,2],mean)  
 means[1]-means[2]  
 }  
boot.object <- boot(repair, bootMeanDiff, R = 5000,   
 strata = repair$type)  
boot.ci(boot.object,conf=.95,type='bca')$bca[4:5]

## [1] -3.4484342 -0.2328956

The 95% interval constructed by bootstrapping is [-3.4218950, -0.1100385]. The interval does not include zero. Thus at a 5% level of significance, we have evidence that the difference in mean repair times is different from 0.

### Part 1c

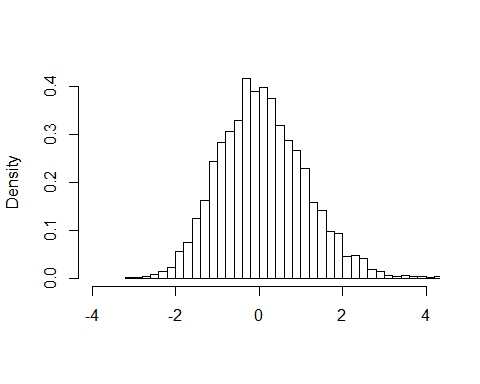
Follow along with with Two Means example in the Bootstrap Hypothesis Testing presentation to bootstrap the two means t test to see if there is a significant difference in population mean repair times. Include a histrogram of the boostrapped t-distribution and write a conclusion to the hypothesis test. (NOTE: in the P value computation slide the last part got cut off, the full code is P <- 2\*min( sum( bootdist < toriginal), sum( bootdist > toriginal ) )/5000.)

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

# two means bootstrap  
xnull <- c(mac.repair.time-mean(mac.repair.time), pc.repair.time-mean(pc.repair.time))  
g <- repair$type  
rs <- rbind(replicate( 5000, sample( xnull[1:62], replace = T) ),  
 replicate( 5000, sample( xnull[63:122], replace = T) ) )  
bootdist <- apply(rs, 2, function(c) t.test(c~g)$statistic )  
  
# compute observed test stat  
toriginal <- t.test( repair$time~g)$statistic  
# for asymmetric distributions the two-tailed P-value is   
# ambiguous. A common solution is to find smaller of the   
# left and right tails and then double it.   
P <- 2\*min( sum( bootdist < toriginal), sum( bootdist > toriginal ) )/5000  
P

## [1] 0.0268

hist( bootdist, probability = TRUE, breaks = 40, main = "",   
 xlab="",xlim=c(-4,4), ylim = c(0,.4))



#curve( dt( x, df = 19), add = TRUE)  
#abline( v = toriginal, col = 'blue')

Replace with your conclusion.

### Part 1d

The bootstrap and theoretical t-distributions give different results here. Which do you trust? Why?

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

Replace with your answer.

## Exercise 2

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. The data set shields is in the DS705data package. The R code to load it is already below.

### Part 2a

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

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

#https://stats.stackexchange.com/questions/11406/boxplot-with-respect-to-two-factors-using-ggplot2-in-r  
  
Machine.A <- data.frame(Diameter=shields$Diameter[which(shields$Machine=="A")])  
Machine.B <- data.frame(Diameter=shields$Diameter[which(shields$Machine=="B")])  
Machine.C <- data.frame(Diameter=shields$Diameter[which(shields$Machine=="C")])  
  
#hist(shields$Diameter[which(shields$Machine=="A")])  
  
machineA\_plot <- ggplot(Machine.A, aes(x=Diameter)) +   
 geom\_histogram(binwidth=1, colour="black", fill="white")  
  
machineB\_plot <- ggplot(Machine.B, aes(x=Diameter)) +   
 geom\_histogram(binwidth=1, colour="black", fill="white")  
  
machineC\_plot <- ggplot(Machine.C, aes(x=Diameter)) +   
 geom\_histogram(binwidth=1, colour="black", fill="white")  
  
  
#https://cran.r-project.org/web/packages/gridExtra/vignettes/arrangeGrob.html  
  
grid.arrange(machineA\_plot, machineB\_plot, machineC\_plot, nrow=1, ncol=3, top="Diameters by Machine", bottom = "Figure 1 Historgrams of Transformed Variables")

## 

### Part 2b

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 2b -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 2c

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

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

Replace this text with your answer here.

### Part 2d

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 2d -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 2e

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 2e -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 2f

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

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

Replace this text with your answer here.

### Part 2g

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 2g -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 2h

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

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

Replace this text with your answer here.

### Part 2i

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 2i -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 2j

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 2j -|-|-|-|-|-|-|-|-|-|-|-

1. Replace this text with your answer here.

# Insert your R code here.

1. Replace this text with your answer here.

### Part 2k

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

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

Replace this text with your answer.

### Part 2l

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 2l -|-|-|-|-|-|-|-|-|-|-|-

Replace this text with your answer here.

### Part 2m

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

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

Replace this text with your answer here.

### Part 2n

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

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

Replace this text with your answer here.

### Part 2o

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 2o -|-|-|-|-|-|-|-|-|-|-|-

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.