ANOVA etc.

Hamid Aboutaher

04/11/2021

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

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

require(DS705data)

## Loading required package: DS705data

data(repair)  
t.test.out = t.test(time~type,data=repair)  
print(t.test.out)

##   
## Welch Two Sample t-test  
##   
## data: time by type  
## t = -1.9442, df = 89.019, p-value = 0.05502  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -3.29705347 0.03583632  
## sample estimates:  
## mean in group Mac mean in group PC   
## 3.150659 4.781267

There is not enough evidence that suggests the mean repair times for Mac(s) and PC(s) are different, therefore we fail to reject the null hypothesis: H\_0: *{} =* {}

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

require(boot)

## Loading required package: 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$type)  
bca.ci = boot.ci(boot.object,conf=.95,type='bca')$bca[4:5]  
print(bca.ci)

## [1] -3.5425190 -0.2331327

We are 95% confident that the population mean repair time for Macs is few hours less than the population mean repair time for PCs.

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

Mac.idx = 1:62;  
PC.idx = 63:122;  
Mac.time = repair$time[Mac.idx]  
PC.time = repair$time[PC.idx]  
Mac.shift = Mac.time - mean( Mac.time )  
PC.shift = PC.time - mean( PC.time )  
set.seed(123) # for reproducibility  
rs = rbind( replicate( 5000, sample(Mac.shift, replace = T) ),  
 replicate( 5000, sample(PC.shift, replace = T) ) )  
bootdist <- apply(rs, 2,   
 function(c) t.test(c[1:62],c[63:122])$statistic )  
toriginal = t.test.out$statistic # from part 1a  
P <- 2\*min( sum( bootdist < toriginal), sum( bootdist > toriginal ) )/5000  
P

## [1] 0.0292

at 95% confident level there is enough evidence to suggest that PC(s) and Mac(s) have different mean repair times. Therefore we reject the null hypothesis in favor of the alternative hypothesis which is Ha: mean(pc) is different than mean(Mac)

### Part 1d

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

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

According the slide 27 the density curve is just a bit skeweed. So the boostrap would not be a good choice. Therefore by elimination process my guess t-distribution to trust.

## Exercise 2

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

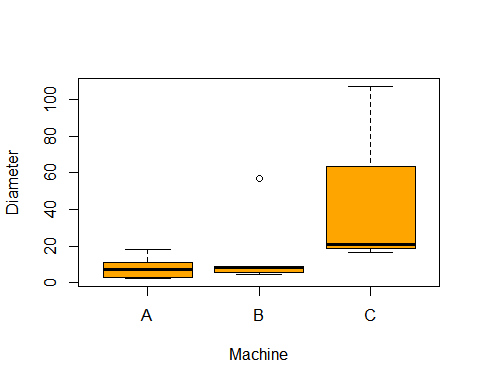
A small corporation makes insulation shields for electrical wires using three different types of machines. The corporation wants to evaluate the variation in the inside diameter dimension of the shields produced by the machines. A quality engineer at the corporation randomly selects shields produced 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 -|-|-|-|-|-|-|-|-|-|-|-

require(DS705data)  
data(shields)  
boxplot(Diameter~Machine,data=shields, col='Orange')



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

It is obvious that C has the largest diameter while A and B have smaller and similar diameter but not quite. Only B has an outlier. C is very skewed with a lot more variation than A and B. Also A looks very symmetric.

### Part 2c

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

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

in 2b we found that C is skeweed and B has an outlier. So there is an unequal variance and population is not normally distributed expect A because of its symmetry. The conditions for ANOVA are not met. A, B, and C seems don’t have same variance.

### 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. null hypothesis: all population have equal mean diameters H0: μ(A)=μ(B)=μ(C) alternative hypothesis: all population mean diameters are NOT equal. Ha: μ(A)≠ μ(B)≠μ(C)

t <- lm(Diameter~Machine, data=shields)  
test <- anova(t)  
test

## Analysis of Variance Table  
##   
## Response: Diameter  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Machine 2 4141 2070.52 2.727 0.09393 .  
## Residuals 17 12907 759.26   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

P <-test$P[1]  
P

## [1] 0.09393099

## at alpha 0.05 we fail to reject the null hypothesis due to lack of evidence that suggest the all populations have the same mean diameters. Therefore we retain the null hpyothesis that the mean diameters of all population are equal.

### 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. H0: μ(A)=μ(B)=μ(C) Ha: μ(A)≠ μ(B)≠μ(C) test <- oneway.test(Diameter ~ Machine, data = shields, var.equal=FALSE) P <- test$p.value

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

##   
## 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

P <-test$p.value  
P

## NULL

|  |
| --- |
| (iii) p-value= 0.06 at alpha 0.05 fail to reject null hypothesis in favor of the alternative hypothesis which states that all the population have not equal mean diameters due not enough evidence. Therefore we keep the null hypothesis which states that μ(A)=μ(B)=μ(C) |

### Part 2f

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

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

Poplulation C is skeweed therefore is not normally distributed. Population B has an outlier that makes it impossible to be normally distributed.

### 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. H0: populations have equal medians Ha: populations have not equal medians

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

test$p.value

## NULL

|  |
| --- |
| (iii) at alpha 0.05 p-value=0.007 we reject null hypothesis which states that the population medians are the same in favor of the alternative hypothesis which states that the populations medians are not the same. There is a strong evidence to reject the null hypothesis. |

### Part 2h

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

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

A, B, and C are different population from size and shape perspective.

### 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. H0: μ(A)=μ(B)=μ(C) populations means are equal Ha: μ(A)≠ μ(B)≠μ(C) populations means are different

Fstar1[ is.na( Fstar1 ) ] = 10\*F.obs p.approx1 <- sum( Fstar1 > F.obs )/B p.approx1

F.obs <-oneway.test(Diameter~Machine, data = shields,var.equal = FALSE)$statistic  
resA <-shields$Diameter[shields$Machine=='A']-mean(shields$Diameter[shields$Machine=='A'])  
resB <-shields$Diameter[shields$Machine=='B']-mean(shields$Diameter[shields$Machine=='B'])  
resC <-shields$Diameter[shields$Machine=='C']-mean(shields$Diameter[shields$Machine=='C'])  
B <- 10000;   
Fstar <-numeric(B)  
for (i in 1:B){  
 pop.null <- data.frame(res=c(resA,resB,resC),Machine=shields$Machine)  
 with(pop.null, tapply(res,Machine,mean))  
 Fstar[i]<-oneway.test(res~Machine,data = pop.null,var.equal = FALSE)$statistic  
}  
Fstar[is.na(Fstar)] = 10 \*F.obs  
p <- sum(Fstar > F.obs) / B  
F.obs

## F   
## 3.974033

p

## [1] 0

|  |
| --- |
| (iii) at aplha 0.05 there is not enough evidence to suggest that the means diameter are not equal. failed to reject the null hypothesis which suggests that the means diameter are the same. |

### 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. H0: μ(A)=μ(B)=μ(C) populations means are equal Ha: μ(A)≠ μ(B)≠μ(C) populations means are different

anovaResampleFast(shieldsMachine,B=10000,method=2,var.equal=F)

source('anovaResampleFast.R')  
anovaResampleFast(shields$Diameter,shields$Machine,B=10000,method=2, 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: 0.3392"

## $f.obs  
## [1] 3.974033  
##   
## $p.obs  
## [1] 0.06096418  
##   
## $p.boot  
## [1] 0.3392

1. at alpha 0.05 there is not enough evidence to suggest that the population mean diameters are different. Therefore we fail to reject the null hypothesis.

### Part 2k

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

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

Since the samples are small we eliminate the unpooled. as far as pooled A,B,and C are different in term of size and shape as I explained above.

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

No none of the four provide a clear answer for the reason we have demonstrated which is C is totally different (diameter) than A and B. They have different size and shape. No test would provide an answer.

### Part 2m

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

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

As I just stated in 2l No statistical analysis provides evidence that on average wire diameters created by machines are different.

### Part 2n

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

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

size of 5,5,and 10 are below the required size to make any sort of implication or conclusion.

### 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(445)  
x <- c( rnorm(100,0,1), rnorm(100,0,10), rnorm(100,0,50))  
groups <- factor( (rep( c('A','B','C'), each=100 ) ) )  
data1 <-data.frame(x, groups)  
kruskal.test(x~groups, data = data1)

##   
## Kruskal-Wallis rank sum test  
##   
## data: x by groups  
## Kruskal-Wallis chi-squared = 6.3152, df = 2, p-value = 0.04253

The distributions are different and so do the shape. Because they have different shape.