Inference for Paired Data

Matt Allen

06/19/2018

## Exercise 1

To reduce ankle injuries, restrictive appliances such as taping and spatting (applying tape over the shoe and sock) have been employed. As part of a study at UWL, subjects also completed a 5-point Likert-type scale survey regarding their perceptions of the movement of each ankle appliance during exercise.

Researchers would like to compare the central values for perceptions of the movement of taped ankles compared to spatted ankles using and to estimate the difference with 90% confidence.

### Part 1a

Load the data set AnkleMovement.rda from the DS705 package.

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

# Load the data AnkleMovement  
require(DS705data)

## Loading required package: DS705data

data("AnkleMovement")

### Part 1b

Create a new variable of the differences, with the perceptions of the spatted ankle (spat) subtracted from the perceptions of the taped ankle (tape).

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

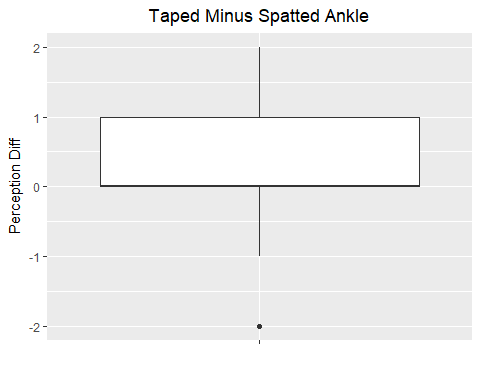
# Create variable of difference in perception.  
perception\_diff <- AnkleMovement$tape - AnkleMovement$spat

### Part 1c

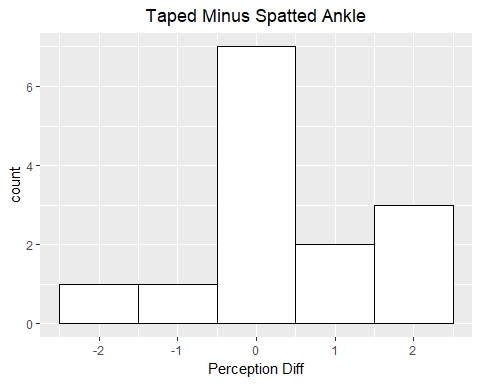
Create a boxplot and histogram for the sample of differences.

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

# Create Boxplots and Histograms for sample differences  
# https://www.r-bloggers.com/make-a-box-plot-with-single-column-data-using-ggplot2-tutorial/  
library(ggplot2)  
perception\_diff\_df <- data.frame(perception\_diff)   
ggplot(data = perception\_diff\_df, aes(x = "", y = perception\_diff)) +   
 geom\_boxplot() + labs(x="", y="Perception Diff") + ggtitle("Taped Minus Spatted Ankle") +  
 theme(plot.title = element\_text(hjust = 0.5))



ggplot(perception\_diff\_df, aes(x=perception\_diff)) +   
 geom\_histogram(binwidth=1, colour="black", fill="white") + labs(x = "Perception Diff") +  
 ggtitle("Taped Minus Spatted Ankle") +  
 theme(plot.title = element\_text(hjust = 0.5))



### Part 1d

Comment on the suitability of this data for the paired t-test, the Wilcoxon signed rank test, and the sign test.

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

It is difficult to judge if the distribution is normal, which would be the best time to use a paired t-test. Also based on the box plot and the histogram, the data may be left skewed, which may be more evidence the distribution of differences not being normal. Also the sample size is small. There is also no explicit statement that the sample is random. We must assume that the sample is random. Although the Wilcoxon signed rank test is less powerful than the t-test, it is a better choice when samples are random, but non-normal. The population of differences should also be approximately symmetric about the median. Based on the boxplot and the histogram, it appears that the difference may not be symmetric about the median. Wilcox signed rank test is more powerful than sign test, but the sign test only requires that the sample data be randomly sampled.

### Part 1e

Because the choice of test is somewhat unclear, as happens often in real life, try all three tests to compare the central values for subject’s perceptions of the movement of taped ankles compared to spatted ankles using .

Do the t-test first:

#### Step 1

Define the parameter in words in the context of the problem.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step1 -|-|-|-|-|-|-|-|-|-|-|-

is defined as the mean of the difference in perception of a spatted ankle subtracted from the perception of a taped ankle.

#### Step 2

State the null and alternative hypotheses for the test using the symbol you defined previously.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step2 -|-|-|-|-|-|-|-|-|-|-|-

#### Step 3

Use R to generate the output for the test you selected.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step3 -|-|-|-|-|-|-|-|-|-|-|-

# paired t test  
t.test(AnkleMovement$tape, AnkleMovement$spat, alternative = "two.sided", paired = TRUE, conf.level = .90)

##   
## Paired t-test  
##   
## data: AnkleMovement$tape and AnkleMovement$spat  
## t = 1.1613, df = 13, p-value = 0.2664  
## alternative hypothesis: true difference in means is not equal to 0  
## 90 percent confidence interval:  
## -0.1874990 0.9017847  
## sample estimates:  
## mean of the differences   
## 0.3571429

#### Step 4

State a statistical conclusion at and interpret it in the context of the problem.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step4 -|-|-|-|-|-|-|-|-|-|-|-

Given a significance level of 10%, we fail to reject the null hypothesis that for taped ankles the mean perceived ankle movement is equal to the mean for spatted ankles for the entire poplulation of athletes.

#### Step 5

Write an interpretation in the context of the problem for the 90% CI for the population mean difference.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step5 -|-|-|-|-|-|-|-|-|-|-|-

Taping and spatting lead to the same perception of movement on a 5 point Likert-type scale. The athletes did not notice a difference in ankle movement between the two methods.

#### Step 6

Perform the Wilcoxon Signed Rank Test.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step6 -|-|-|-|-|-|-|-|-|-|-|-

# Wilcoxon Signed Rank Test  
wilcox.test(AnkleMovement$tape, AnkleMovement$spat, alternative = "two.sided", paired = TRUE, conf.level = .90, conf.int = TRUE)

## Warning in wilcox.test.default(AnkleMovement$tape, AnkleMovement$spat,  
## alternative = "two.sided", : cannot compute exact p-value with ties

## Warning in wilcox.test.default(AnkleMovement$tape, AnkleMovement$spat,  
## alternative = "two.sided", : cannot compute exact confidence interval with  
## ties

## Warning in wilcox.test.default(AnkleMovement$tape, AnkleMovement$spat,  
## alternative = "two.sided", : cannot compute exact p-value with zeroes

## Warning in wilcox.test.default(AnkleMovement$tape, AnkleMovement$spat,  
## alternative = "two.sided", : cannot compute exact confidence interval with  
## zeroes

##   
## Wilcoxon signed rank test with continuity correction  
##   
## data: AnkleMovement$tape and AnkleMovement$spat  
## V = 20.5, p-value = 0.2981  
## alternative hypothesis: true location shift is not equal to 0  
## 90 percent confidence interval:  
## -0.5000113 2.0000000  
## sample estimates:  
## (pseudo)median   
## 0.9999682

As in the paired t-test, we fail to reject the null hypothesis.

#### Step 7

Perform the sign test.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step7 -|-|-|-|-|-|-|-|-|-|-|-

# sign test  
library(signmedian.test)  
signmedian.test(perception\_diff, mu=0, alternative="two.sided", conf.level = .9, conf.int = TRUE)

##   
## Exact sign test  
##   
## data: perception\_diff  
## #(x!=0) = 7, mu = 0, p-value = 0.4531  
## alternative hypothesis: the median of x is not equal to mu  
## 87.5 percent confidence interval:  
## -1 2  
## sample estimates:  
## point estimator   
## 0

As in the paired t-test, we fail to reject the null hypothesis.

#### Step 8

Construct a bootstrap confidence interval at a 90% level of confidence for the mean difference in population mean perception of movement for taped and spatted ankles. Use a bootstrap sample size of 5000. Compare this interval with the results of the 90% *t*-interval.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step8 -|-|-|-|-|-|-|-|-|-|-|-

# Bootstrap.  
  
library(boot)  
bootMean <- function(x, i) {  
 # x is a numeric vector  
 # i is a vector of indices for the resampled elements of x  
 mean(x[i])  
}  
set.seed(NULL)  
boot.object <- boot(perception\_diff, bootMean, R=5000)  
boot.ci(boot.object, conf = .9, type = "bca")$bca[4:5]

## [1] -0.2142857 0.7857143

The 90% confidence interval from the paired t-test is [-0.1874990,0.9017847] is similar to the 90% bootstrap confidence interval. They both have the same conclusion that the difference of zero lies in the interval.

#### Step 9

Compare the results of the three hypothesis tests and also whether or not the 90% bootstrap interval agrees with the result of each test. Which procedure should be reported and why?

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e.step9 -|-|-|-|-|-|-|-|-|-|-|-

The paired t-test reports a 90% confidence interval of [-0.1874990, 0.9017847].

Wilcox Signed Rank Test reports a 90% confidence interval of [-0.5000113, 2.0000000].

The sign test reports a 90% confidence interval of [-1,2].

All the tests agree with the bootstrap interval that a difference of 0 is in the 90% confidence interval, so use the t-test because it is the most powerful and most commonly reported.

## Exercise 2

In a nationwide study of insurance claims (in dollars) filed in the previous year, a random sample of 125 claims was selected from all claims for vehicles classified as small, meaning the gross vehicle weight rating (GVWR) is less than 4500 pounds.

For each claim, the insurance company’s estimate for the claim was also provided.

The data frame SmallAuto.rda contains the claims and estimates for each vehicle class.

### Part 2a

Load the data SmallAuto from the DS705 package.

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

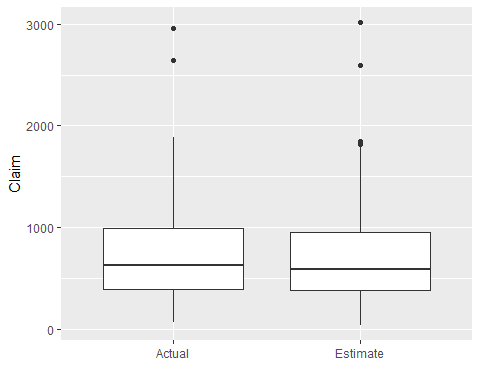
# Load the data.  
data("SmallAuto")

### Part 2b

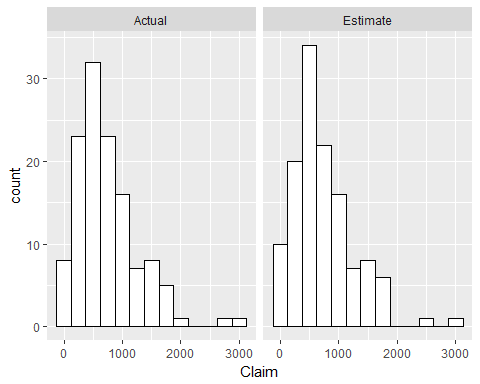
Construct histograms and boxplots for both the estimated claims and actual for the small class of vehicle. Describe the shapes of these distributions.

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

# SmallAuto histogram and boxplot  
ggplot(data = SmallAuto, aes(x = Category, y = Claim)) + geom\_boxplot() + labs(x = "")



ggplot(SmallAuto, aes(x=Claim)) +   
 geom\_histogram(binwidth=250, colour="black", fill="white") +   
 facet\_grid(. ~ Category)



Both the estimate and actual distributions are right skewed.

### Part 2c

Create a new variable of the differences for small vehicles, with the difference being the estimated claim amount minus the actual claim amount. The estimated claim amounts in the first half of the vector are paired with the actual claim amounts in the second half of the vector so that row 1 and row 126 form a pair, rows 2 and 127, etc.

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

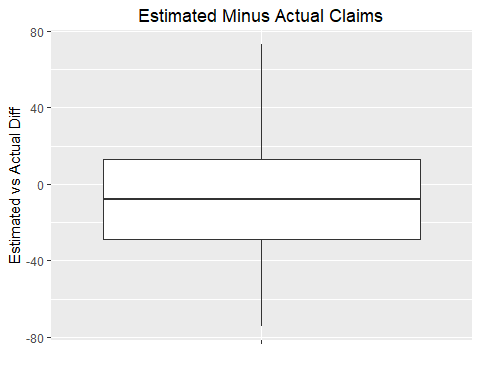
# create vector of differences between estimated and actual  
  
estimated\_claims <- SmallAuto[1:125,]$Claim  
actual\_claims <- SmallAuto[126:250,]$Claim  
  
small\_auto\_est\_act\_diff <- estimated\_claims - actual\_claims

### Part 2c

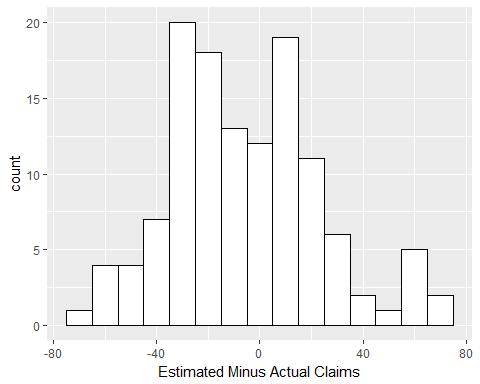
Create a boxplot, histogram, and normal probability plot for the sample of differences. Also, obtain the P-value for a Shapiro-Wilk normality test.

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

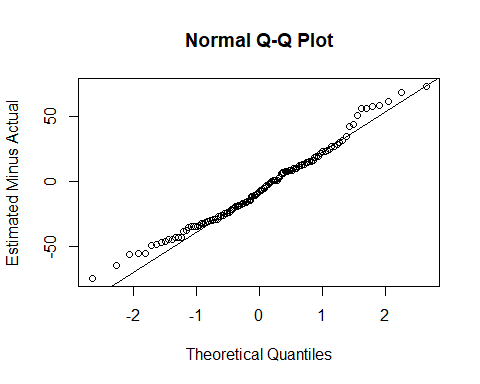
# boxplot, histogram and normal probability plots for sample of differences  
  
small\_auto\_est\_act\_diff\_df <- data.frame(small\_auto\_est\_act\_diff)   
  
ggplot(data = small\_auto\_est\_act\_diff\_df, aes(x = "", y = small\_auto\_est\_act\_diff)) +   
 geom\_boxplot() + labs(x="", y="Estimated vs Actual Diff") + ggtitle("Estimated Minus Actual Claims") +  
 theme(plot.title = element\_text(hjust = 0.5))



ggplot(small\_auto\_est\_act\_diff\_df, aes(x=small\_auto\_est\_act\_diff)) +   
 geom\_histogram(binwidth=10, colour="black", fill="white") + labs(x = "Estimated Minus Actual Claims") #+



# Make normal probability plot of Estimated minus Actual and add fit line.  
{  
 qqnorm(small\_auto\_est\_act\_diff, ylab="Estimated Minus Actual")   
 qqline(small\_auto\_est\_act\_diff)  
}



shapiro.test(small\_auto\_est\_act\_diff)

##   
## Shapiro-Wilk normality test  
##   
## data: small\_auto\_est\_act\_diff  
## W = 0.98264, p-value = 0.1093

### Part 2d

Comment on the shape of the distribution of differences and the suitability of this data for the paired *t*-test, the Wilcoxon signed rank test, and the sign test. Which test will you use?

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

Use the t-test, the distribution is borderline normal, but the sample size is large enough and the t-test is robust enough to work.

### Part 2e

Conduct an appropriate test to see if the population central values for the estimated claim amount is less than for the actual claim amounts for vehicles in the small class using .

#### Step 1

Define the parameter in words in the context of the problem.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 2e.step1 -|-|-|-|-|-|-|-|-|-|-|-

is the mean of the difference between paired estimated claims minus its actual.

#### Step 2

State the null and alternative hypotheses for the test using the symbol you defined previously.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 2e.step2 -|-|-|-|-|-|-|-|-|-|-|-

#### Step 3

Use R to generate the output for the test you selected. Pay close attention to the order of subtraction done behind the scenes in R.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 2e.step3 -|-|-|-|-|-|-|-|-|-|-|-

#this subtracts actual\_claims - estimated\_claims, because alphabetical order of subtraction, so reverse arguments.  
t.test(actual\_claims, estimated\_claims, alternative = "less", paired = TRUE, conf.level = .95)

##   
## Paired t-test  
##   
## data: actual\_claims and estimated\_claims  
## t = 2.1537, df = 124, p-value = 0.9834  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
## -Inf 10.11882  
## sample estimates:  
## mean of the differences   
## 5.718576

#### Step 4

State a statistical conclusion at and interpret it in the context of the problem.

#### -|-|-|-|-|-|-|-|-|-|-|- Answer 2e.step4 -|-|-|-|-|-|-|-|-|-|-|-

At a 5% level of significance, we fail to reject the null hypothesis that the mean difference in the population of estimated claims minus actual claims is greater that or equal to zero.

### Part 2f

Write an interpretation in the context of the problem for a 95% two-sided confidence interval.

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

# calculate confidence interval  
t.test(actual\_claims, estimated\_claims, alternative = "two.sided", paired = TRUE, conf.level = .95)$conf.int

## [1] 0.4632425 10.9739093  
## attr(,"conf.level")  
## [1] 0.95

The 95% confidence interval of the mean difference in the population of estimated claims minus actual claims is [0.4632425, 10.9739093]. This means that the estimate of future claims is greater than the actual. This is important in insurance. The insurance company can reserve conservatively for future claims.

## Exercise 3

The data frame AutoIns is very similar to Small Auto.

In a nationwide study of insurance claims filed in the previous year, a random sample of 125 claims was selected from all claims for vehicles classified as small, meaning the gross vehicle weight rating (GVWR) is less than 4500 pounds A separate sample of 125 claims for vehicles classified as standard, meaning the GVWR is between 4500 and 8500 pounds.

For each claim, the insurance company’s estimate for the claim was also provided.

The data frame AutoIns.rda contains the claims and estimates for each vehicle class. The variables in the data frame are defined as follows:

claim.small = the actual claim amount in dollars for a vehicle in the small class

est.small = the estimated claim amount in dollars for a vehicle in the small class

claim.standard = the actual claim amount in dollars for a vehicle in the standard class

est.standard = the estimated claim amount in dollars for a vehicle in the standard class

### Part 3a

Load the data AutoIns from the DS705 package and look at the structure of the data in the file.

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

# Load the data.  
data("AutoIns")

### Part 3b

Is the data “stacked” or “side-by-side” (“tall” or “wide”)?

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

The data is side by side.

### Part 3c

Which pairs of variables in the data frame are independent of each other? You can use the variable names to identify them.

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

claim.small and claim.standard should be independent of each other. est.small and est.standard should also be independent of each other.

### Part 3d

Which pairs of variables in the data frame are paired (matched pairs)? You can use the variable names to identify them.

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

claim.small and est.small are paired with each other, and claim.standard and est.standard are paired with each other.