SB1 Practical 3

Candidate number: 1015294

March 4, 2019

1. Introduction

The following report consists of two main parts each of which analyses the results of one experiment. The first part studies how blood pressure measurements change when are measured at home and by a nurse in a hospital. The aim is to discover whether the two types of measurements follow the same distribution. The first half of this part conducts analysis on the data and decides whether to treat the data as independent pairs or as if all observations are independent. In order to do that the correlation coefficient between home and hospital measurements is estimated using bootstrap. The second half takes into account the acquired results in the first half and applies two different tests in order to discover possible mean and median differences between the two samples.
The second part of this report examines the effect of two different treatments on blood clotting times. The aim is to investigate whether one of the treatments decrease the clotting time. The first half of this part is data analysis and is very similar to the one in the first part. It has the same goal - to determine whether data should be treated in pairs or not. In the second half two tests are presented in order to dicover whether reduction in clotting time exists.

2. Part I

2.1 Data analysis. The data in this experiment consists of 22 observations acquired after recording the diastolic `blood pressure` of 11 individuals who are identified with the `subject` variable. The `blood pressure` of each individual is measured at `home` and by a nurse in a `hospital`. Our data analysis will contain two parts. The first one would examine whether the data should be treated as 11 pairs of `blood pressure` measurements of each individual taken at `home` and in a `hospital` or we could consider having 22 independent observations in which case the trials does not depend on the `subject` variable. The second part inspects the relationship between the variables `home` and `hospital`. The aim of both parts is to help us determine what test approach to apply later in order to determine whether `home` and `hospital` measurements have the same distribution. Firstly, Figure 1 below demonstrates that `blood pressure` might take different interval of values depending on the individual taking part in the trial because it can be seen that `blood pressure` measurements at `home` and in the `hospital` for each subject take a relatively small range of values depending on the individual. The differences in `blood pressure` measurements recorded at `home` and in the `hospital` for the 11 `subjects` are:

```
> home-hospital
 [1]   2  -9  -6  -3 -13   4 -10 -12  -7 -11   1
> summary(home-hospital)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-13.000 -10.500  -7.000  -5.818  -1.000   4.000
```

Considering that `blood pressure` takes values between 70 and 110, then a median of magnitude

7 and a mean of magnitude 5.818 for the difference of `blood pressure` measurements recorded at `home` and in the `hospital` for each `subject` are not substantial. This is an indication that a relation between the variables `blood pressure` and `subject` might exist.

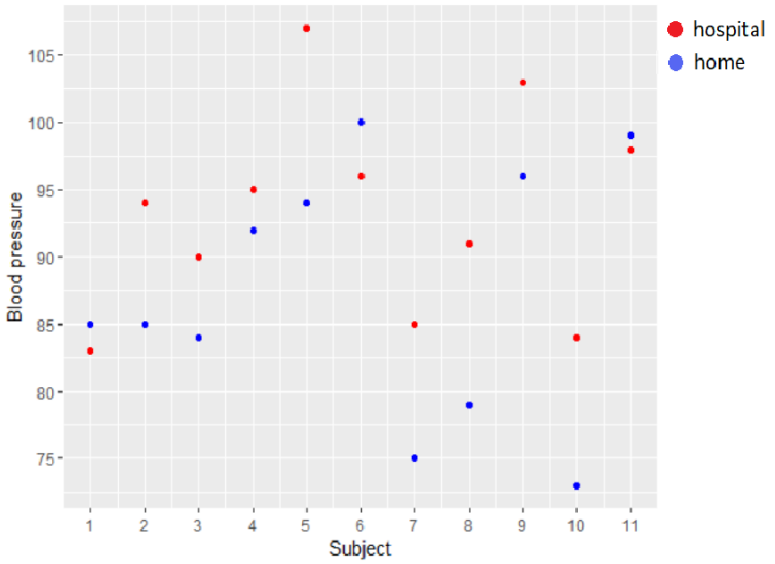


Figure 1: Blood pressure of each subject measured at home and in the hospital.

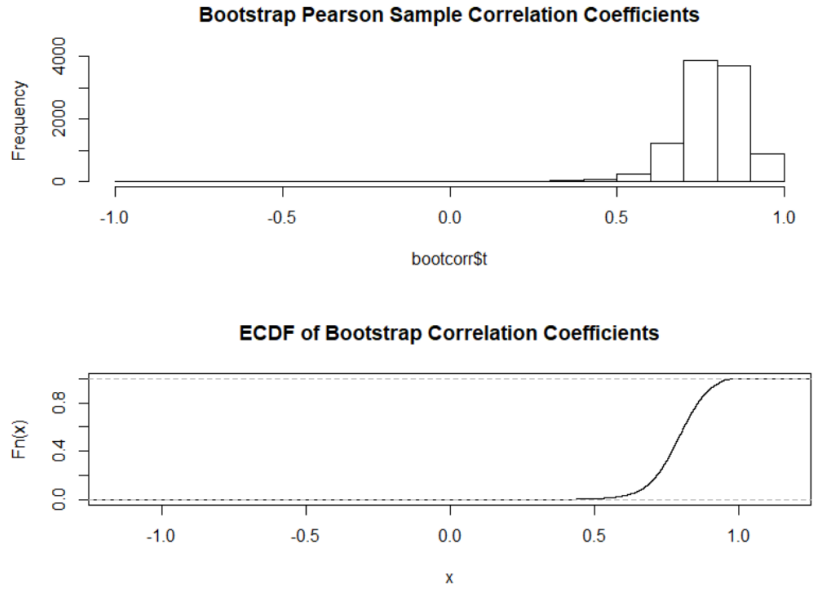
Now we shall calculate the correlation between the values of `blood pressure` taken at `home` and in the `hospital`. If it is significant then we would prefer to treat the data in pairs because `blood pressure` is likely to depend on the tested `subject`. Consider two random variables $\mathbf{X}$ and $\mathbf{Y}$ which represent `blood pressure` values recorded at `home` and in the `hospital`, respectively. The observed values for $\mathbf{X}$ are: 85 85 84 92 94 100 75 79 96 73 99 and the corresponding values for $\mathbf{Y}$ are: 83 94 90 95 107 96 85 91 103 84 98. We want to estimate the Pearson's correlation coefficient ρ between $\mathbf{X}$ and $\mathbf{Y}$. In order to achieve a good approximation we shall estimate ρ in two different ways, first using t-test and then bootstrap. Assume the following setup: the null hypothesis is H_0: $\rho = 0$ and the alternative is H_1: $\rho \neq 0$. If our underlying variables have approximately uncorrelated bivariate normal distribution we could compare our test statistic $t_{obs} = 3.6633$ with a t_9 distribution because our test has $11 - 2 = 9$ degrees of freedom. According to the R code below a sample estimate for the correlation is 0.77 which has a significant p-value equal to 0.005 considering a level of $\alpha = 5\%$. Also t-test gives a 95% confidence interval (0.32, 0.93) which does not contain 0 so indicates a reasonable correlation between $\mathbf{X}$ and $\mathbf{Y}$.

```
> cor.test(home,hospital)

        Pearson's product-moment correlation

data:  home and hospital
t = 3.6633, df = 9, p-value = 0.005208
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3243198 0.9381479
sample estimates:
      cor
0.7736752
```

However, we do not know if normality assumptions are reasonable in our case so we apply bootstrap method as well. We use a non-parametric bootstrap with 11 pairs (x_i, y_i) resampled with replacement from the observed set of 11 pairs, and the correlation coefficient ρ is calculated based on the resampled data. After repeating this operation B=10000 times we obtain:
The bootstrap statistic for the estimated correlation is $\hat{\rho}$ =0.77 as in the t-test. Further, we notice

```
ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = home.hospital, statistic = pearson, R = Brep)


Bootstrap Statistics :
     original     bias    std. error
t1* 0.7736752 0.01002287  0.09342795
> # Bootstrap confidence intervals
> boot.ci(bootcorr, type = c('basic','norm'))
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = bootcorr, type = c("basic", "norm"))

Intervals :
Level      Normal              Basic
95%   ( 0.5805,  0.9468 )   ( 0.6056,  0.9658 )
Calculations and Intervals on Original Scale
```

that bootstrap gives tighter 95% confidence intervals for the estimated correlation than the one we obtained in t-test, but still does not contain 0. Indeed a 95% normal confidence interval for $\hat{\rho}$ is $(0.58, 0.95)$ and a pivotal 95% confidence interval is $(0.60, 0.96)$. Therefore, we can conclude that the correlation coefficient between $\mathbf{X}$ and $\mathbf{Y}$ is significant and data should be treated in pairs. Also, positive correlation means that if `blood pressure` measurements at `home` increase/decrease then those in the `hospital` are likely to increase/decrease too. Bootstrap sample coefficients and their empirical cumulative distribution function can be seen below.



Secondly, we examine the relationship between `home` and `hospital` variables. Figure 3 (b) below shows that when `blood pressure` is measured at `home` it is often smaller than when is measured in the `hospital`. From Figure 3 (a) we can observe that only 3 individuals (1, 6 and 11) have higher `home` measurements than `hospital` ones. Hence, in general we would expect the variable `hospital` give higher `blood pressure` than `home`.

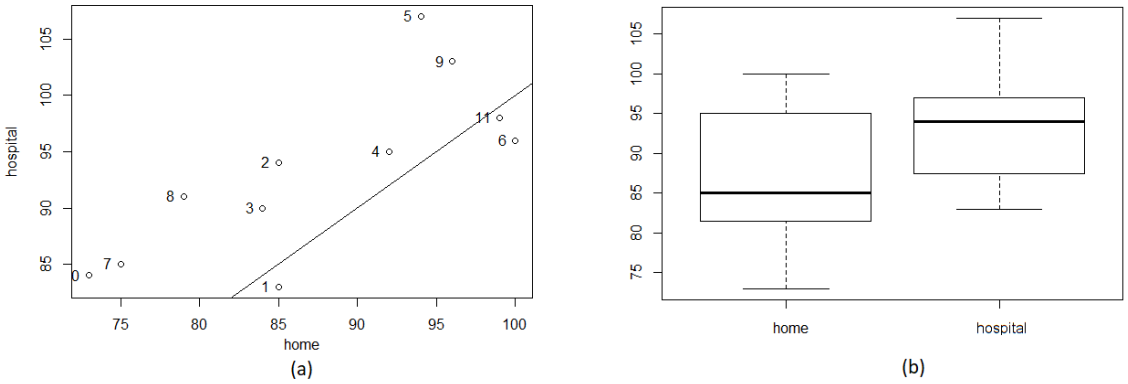


Figure 3: (a) - home against hospital measurements and
(b) - boxplot of home and hopital measurements.

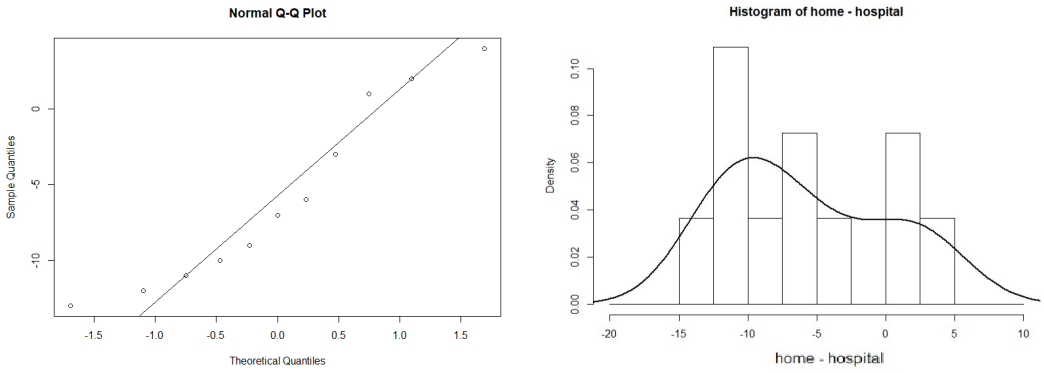
2.2 Tests. In this section we shall execute two different paired tests and analyze the results of each carefully. The first one we consider is t-test which tests whether the true mean difference between the two samples (`home` and `hospital` measurements) is equal or not. The second one is Wilcoxon Signed Rank Test which tests whether the difference between the pairs follows a symmetric distribution around 0. Before we proceed with the testing process, the sample we are interested in is the differences of `blood pressure` measurements taken at `home` and in the `hospital` for each individual i.e the observations are: 2 -9 -6 -3 -13 4 -10 -12 -7 -11 1.

2.2.1 T-test. The assumptions we need to make in order to conduct t-test are:

- Observations must follow a continuous distribution.
- Observations should be independent.
- Observations should not contain outliers.
- Observations should follow approximately normal distribution.

The first two assumption are definitely reasonable to make because our observations do not take discrete values but continuous and independence is a plausible assumption because `blood pressure` measurements on different `subjects` does not depend on each other. Moreover, the observations take values between -13 and 1 which is relatively small range considering that `blood pressure` can take values between 70 and 130 (difference of 60). Hence, it is not dangerous to assume that outliers do not exist in our sample. However, normality assumption might be violated and t-test will not be appropriate. Figure 4 (a) on next page shows that the given sample follows approximately normal distribution except in the tails where in each end there is 1 point which deviates from the normal line. Also Figure 4 (b) indicates the density function on the right does not follow exactly a normally distributed density. Nevertheless, we shall apply t-test but take into account the slight violation of the normality assumption when examining the results. Our null hypothesis is H_0: true mean difference $\mu_d = 0$ and the alternative is H_1: $\mu_d \neq 0$. The R output below indicates that an estimate of the mean of differences is -5.81 with a significant p-value equal to 0.009. This indicates that the distributions of `home` and `hospital` measurements have different means and therefore are not the same. However, we cannot make a definite conclusion about the difference in the means of the two distributions because we are not sure the normality assumption holds.

```
> t.test(home,hospital,paired = TRUE, alt ="two.sided",conf.int = T)

        Paired t-test

data:  home and hospital
t = -3.2267, df = 10, p-value = 0.009071
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -9.835782 -1.800581
sample estimates:
mean of the differences
              -5.818182
```



(a) (b)

Figure 4: (a) -QQ plot for differences in blood measurements taken at home and in the hospital. (b) - Histogram and density function for these differences.

2.2.2 Wilcoxon Signed Rank Test. Wilcoxon Signed Rank Test is a non-parametric approach and is a sensible one to use in our case because our data sample is small and we cannot estimate successfully the distribution of the sample. The assumptions we need to make here are:

- Data are paired and come from the same population.
- Pairs are independent and identically distributed.
- Within pair differences follow a continuous distribution.

These three assumptions are valid because: in the first section we showed that the data set should be treated in pairs, which are clearly independent and identically distributed following a continuous distribution function. Also, note that distances within pairs are all non-zero so we do not discard any pair when conducting WSRT. Our null hypothesis is H_0 : the true median is $\Delta = 0$ against $H_1 : \Delta \neq 0$ After implementing the test we obtain a sample estimate of the median equal to 6 and a 95% confidence interval (1.5, 10.5). This estimated median has a significant p-value of 0.018. Therefore, we reject H_0 and conclude that the two distributions have a location difference of magnitude 6. The R output for these results is provided below.
The p-value in this case can be computed manually as well. Recall the sample of within pair differences is: 2 -9 -6 -3 -13 4 -10 -12 -7 -11 1. Then the observed value of the Wilcoxon test statistic is $1 + 2 + 4 = 7$ by definition of the statistic. There are 19 sums of natural numbers less than or equal to 7: 1, 2, 3, 4, 5, 6, 7, 1+2, 1+3, 1+4, 1+5, 1+6, 2+3, 2+4, 2+5, 3+4, 1+2+3, 1+2+4. Then $P(W \leq 7) = \frac{19}{2^{11}} = 0.0092$. Thus the p-value for our two-sided test is equal to 0.018 as expected. Further, the Lehmann-Hodges estimate for the median is the median of the set of Walsh averages $\frac{X_i+X_j}{2}$ for all $i, j = 1...11$ where X_i are within pair differences. Not surprisingly the LH median estimate is again 6 as shown below.

```
> # Wilcoxon Signed Rank Test
> wilcox.test(home,hospital,paired = TRUE, alt ="two.sided",conf.int = T)

        Wilcoxon signed rank test

data:  home and hospital
V = 7, p-value = 0.01855
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -10.5  -1.5
sample estimates:
(pseudo)median
            -6
> w.averages <- walsh(hospital-home)
> median(w.averages)
[1] 6
```

2.3 Conclusion. At the beginning we examined the relationship between the two samples of `blood pressure` measured at `home` and in the `hospital`. By considering the correlation between the two variables we reached the conclusion that we should consider the data in pairs with one observation from each sample. Then two paired tests were performed with the aim to determine whether the two samples have different means and medians. A certain interpretation for the means could not be made but according to Wilcoxon Signed Rank Test the medians for the two distributions are different. Thus, the distributions of `home` and `hospital` measurements are different.

3. Part II

3.1 Data analysis. The data in this experiment consists of 30 observations acquired after investigating the effect of an `old` and a `new` drug treatment on `blood clotting time`. In the trials were involved 15 `subjects` each of which underwent exactly once both treatment procedures. Again our data analysis goal is to determine whether data should be treated in pairs. In order to determine that observe from Figure 5 below that there is not a clear dependence between `clotting times` in each pair for each individual. The red and blue points for each `subject` seem to be randomly scattered and an obvious pattern between them is not present. Moreover, the within pair differences range between -115 and 420 with mean equal to 117 and median equal to 75, which indicates that these differences appear to take supposedly random values.

```
> old-new
 [1]  225 -115   85  390   75   55   30  325  -40  168  -25  -75  420  255   -7
> summary(old-new)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -115.0   -16.0    75.0   117.7   240.0   420.0
```

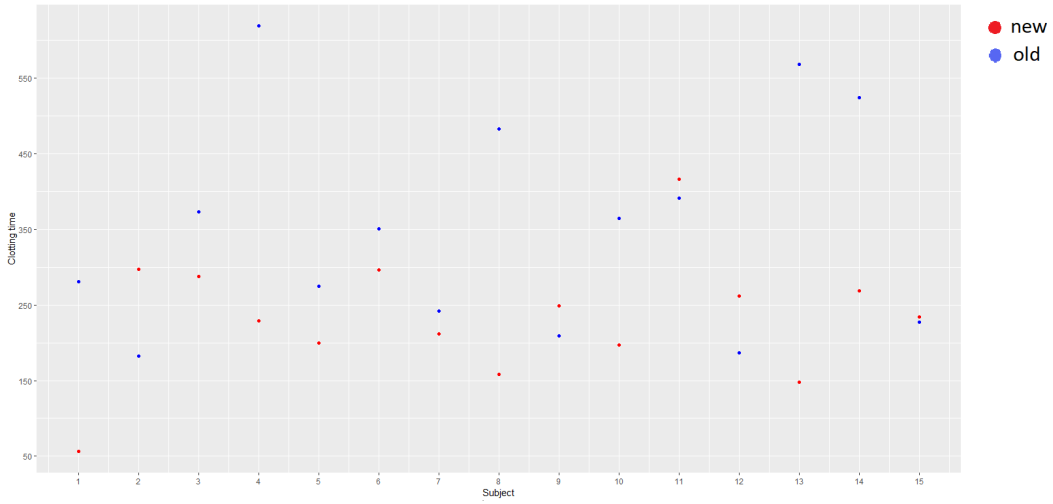


Figure 5: Subject against Clotting time.

To address this issue formally we shall find an estimate for the Pearson correlation coefficient between the two variables `old` and `new`. The observed values for these variables are: `old` = (281 182 373 619 275 351 242 483 209 365 391 187 568 524 227) and `new` = (56 297 288 229 200 296 212 158 249 197 416 262 148 269 234). Denote with ρ the true correlation between the two variables. We implement a non-parametric bootstrap with 15 pairs (x_i, y_i) resampled with replacement from the observed set of 15 pairs, and the correlation coefficient ρ is calculated based on the resampled data (R output see below). Repeationg this procedure B = 10000 times we obtain an estimate for the correlation $\hat{\rho} = -0.098$ which is very near to 0 i.e no correlation between variables `old` and `new`. Furthermore, by bootstraping confidence intervals we acquire a 95% normal confidence interval equal to $(-0.5, 0.3)$ and a pivotal 95% confidence interval equal to $(-0.47, 0.37)$ both of which contain 0. Thus, we conclude there is no correlation between the variables `old`:= **X** and `new`:= **Y** and consider the combined sample **Z**= (**X**,**Y**). Bootstrap sample coefficients and their empirical cumulative distribution function can be seen in Figure 6. Note that the sample correlation coefficients follow a fairly normal distribution.

```
ORDINARY NONPARAMETRIC BOOTSTRAP


Call:
boot(data = old.new, statistic = Pearson, R = Brep)


Bootstrap Statistics :
      original      bias    std. error
t1* -0.09800337 -0.01573835   0.2146739
> boot.ci(bootcorr, type = c('basic','norm'))
BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
Based on 10000 bootstrap replicates

CALL :
boot.ci(boot.out = bootcorr, type = c("basic", "norm"))

Intervals :
Level      Normal              Basic
95%   (-0.5030,  0.3385 )   (-0.4743,  0.3703 )
Calculations and Intervals on Original Scale
```

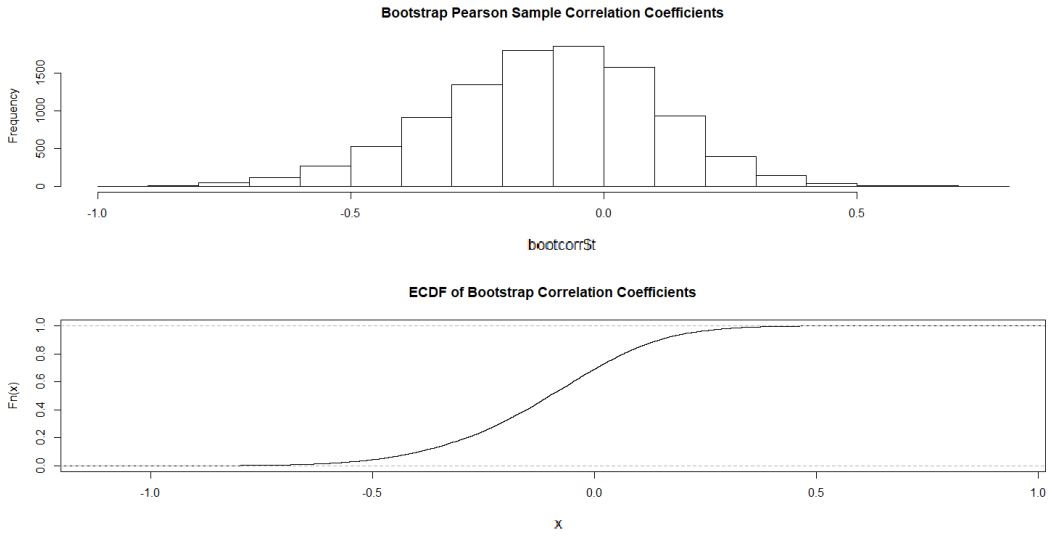


Figure 6

Now we examine the relation between `clotting time` under `old` and `new` drug treatments. Figure 7 demonstrates that the `new` drug reduces `clotting time` and therefore, in the next section we will consider two different tests for mean reduction in `clotting time` and for median reduction in `clotting time`.

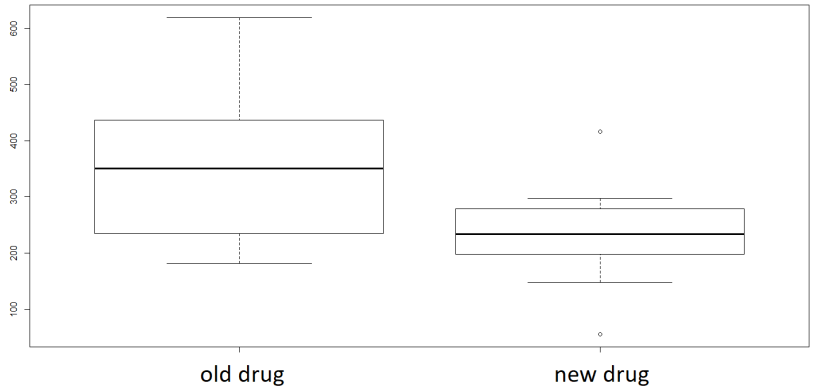


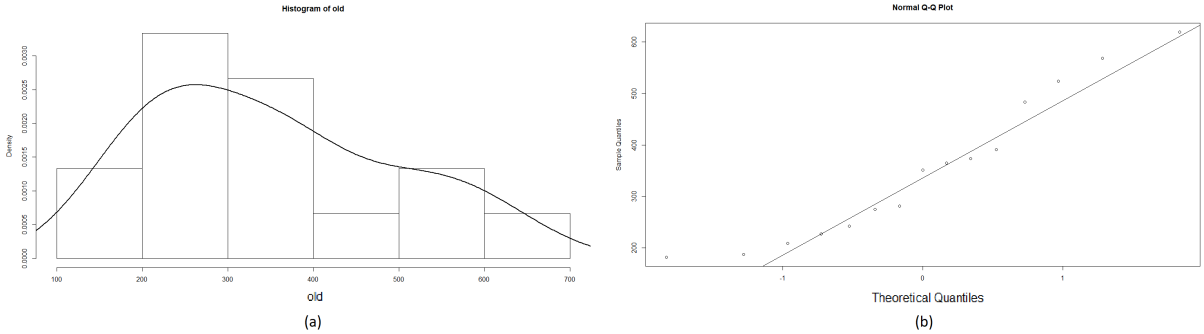
Figure 7: Boxplot of clotting time in seconds after applying old and new drug treatments.

3.2 Tests. The first test we consider is t-test which examines the difference in means of the two variables. The second test is Wilcoxon Rank Sum Test which inspects the location shift in the two samples by considering the median.

3.2.1 T-test. Before proceeding with implementing the test we need to check whether the assumptions for t-test hold. The two assumptions which might be violated are: observations follow an approximately normal distribution and the two samples have equal variance.

In order to inspect whether samples `old`:= $\mathbf{X}$ and `new`:= $\mathbf{Y}$ have the same variance we use the following test statistic:$T = |\text{sd}(\text{rank}(\mathbf{X})) - \text{sd}(\text{rank}(\mathbf{Y}))|$ where sd is the standard deviation. Note we use the standard deviations of the ranks of $\mathbf{X}$ and $\mathbf{Y}$ because we want our test to be robust and not sensible to outliers. Since we do not know the exact distribution of T we apply Monte Carlo method to obtain an estimated p-value. The value for T_{obs} is 1.40 and an approximate p-value is equal to 0.36 as shown below. With this insignificant p-value we do not reject H_0 and can safely conclude that the variances of the two samples $\mathbf{X}$ and $\mathbf{Y}$ are equal.

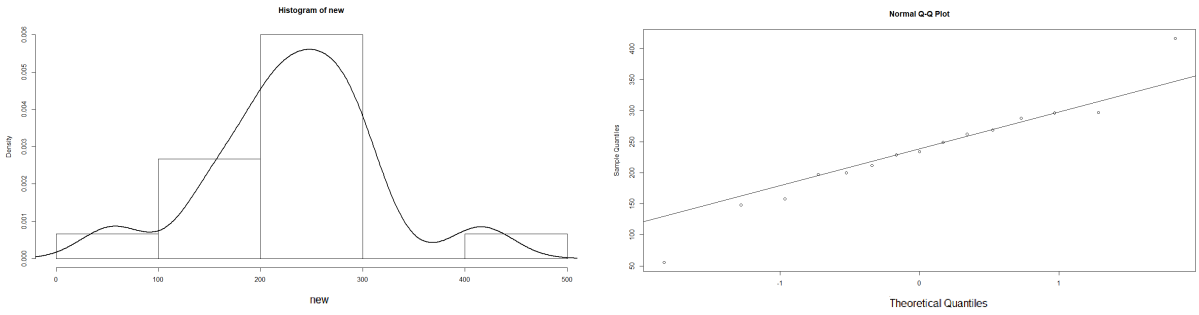
```
> # Test difference in variances, robust method
> z<- c(old,new); rz=rank(z)
> z<- c(old,new); rz=rank(z)
> T=abs(sd(rz[1:15])-sd(rz[16:30]))
> T
[1] 1.407087
> K=1000000; T0=rep(NA,K);
> for (k in 1:K) {i=sample(1:30,30,replace=F) ;rzp=rz[i] ; T0[k]=abs(sd(rzp[1:15])-sd(rzp[16:30]))}
> mean(T<T0)
[1] 0.360246
```



(a) (b)

Figure 8: (a) Histogram and density of sample values of variable old. (b) QQ plot of sample values of variable old.

Now we check whether normality assumptions are appropriate. Figure 8 above shows that `old` variable might not follow a normally distributed function. The density of the observations seems not to be normally distributed as shown in the histogram in Figure 8(a). However in the QQ-plot in Figure 8(b) there is only one observation in the left tail which deviates from the normal QQ-line. On the other hand, variable `new` might have a normal distribution. According to Figure 9 it is



(a) (b)

Figure 9: (a) Histogram and density of sample values of variable new (b) QQ plot of sample values of variable new.

symmetric and most of the observations in the QQ-plot are near the normal QQ-line. In conclusion, we will assume normality of the two variables but be very cautious when we make interpretation of the results. Consider null hypothesis is H_0: $\mu_{old} = \mu_{new}$ against the alternative H_1: $\mu_{old} > \mu_{new}$. The R output below indicates that estimates of the means of the two samples are $\hat{\mu}_{old} = 351.80$ and $\hat{\mu}_{new} = 234.07$ with a significant p-value equal to $\frac{0.01}{2}$. Thus, if the two variables are normally distributed there is a reduction in means of `clotting time` when applying the `new` drug treatment compared with the `old` one.

```
> t.test(old,new) # assumes old and new are normal

        Welch Two Sample t-test

data:  old and new
t = 2.7821, df = 22.396, p-value = 0.01076
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  30.0615 205.4052
sample estimates:
mean of x mean of y
 351.8000  234.0667
```

3.2.2 Wilcoxon Rank Sum Test. The assumptions for Wilcoxon Rank Sum Test are met because we already know that the observations are independent and when applying a non-parametric approach such as Wilcoxon Test we do not need the distribution of the variables. The sample we are interested in is `Z` = (`X`,`Y`) = (281 182 373 619 275 351 242 483 209 365 391 187 568 524 227 56 297 288 229 200 296 212 158 249 197 416 262 148 269 234), i.e all 30 `clotting times`. Wilcoxon observed test statistic is the sum of the last 15 entries of Z, which is 182. We want to test whether `X` $-\Delta \sim$ `Y` is true. Then consider null hypothesis H_0: $\Delta = 0$ against H_1: $\Delta > 0$.

```
> # Calculating W_obs
> z<- c(old,new); rz<-rank(z); sum(rz[16:30])
[1] 182
> # W Rank Sum Test
> wilcox.test(old, new, paired = F, alt = "greater", conf.int = T)

        Wilcoxon rank sum test

data:  old and new
W = 163, p-value = 0.01836
alternative hypothesis: true location shift is greater than 0
95 percent confidence interval:
  26 Inf
sample estimates:
difference in location
                    96
```

Note that the observed value of the test statistic in `wilcox.text` function is the Mann-Witney one and this is why it equals $15^2 + \frac{15*16}{2} - W_{obs} = 245 - 182 = 163$. The test gives an estimated difference in location equal to $\hat{\Delta} = 96$ with a p-value equal to 0.02 significant at a 5% level. Therefore, we conclude that the median of the `clotting time` when an `old` treatment is used is bigger than the median of the `clotting time` when an `new` treatment is used. We also obtain a one sided confidence interval CI = $(26, \infty)$ because we consider one sided test. Finally, we note that the

estimated value for Δ from `wilcox.text` is actually the Hodges-Lehmann estimate which is just the least significant estimator.

```
> HodgesLehmann(old,new)
[1] 96
```

3.3 Conclusion. In the analysis we applied two tests in order to discover whether the `new` drug decreases the `clotting time` or not. Firstly, we analyzed the data so that we choose an appropriate test to implement. After considering the correlation betwwen the two variables we concluded that `blood clotting time` does not depend on the tested `subject`. Further, we saw that T-test shows that the mean `clotting time` decreases when the `new` drug is applied, however, normality assumptions in t-test might be violated so we cannot make a definite conclusion about the mean `clotting time`. Then we used Wilcoxon Rank Sum Test and found that a location shift exists between the variables for `old` and `new` treatments. Indeed, when a `new` treatment is used the `clotting time` is smaller than when a `old` treatment is used.

8. Appendix

```
library(Rfit)
library(boot)
library(ggpubr)
library(DescTools)

########## PART 1 ##########
subject <- c(1,2,3,4,5,6,7,8,9,10,11)
home <- c(85,85,84,92,94,100,75,79,96,73,99)
hospital <- c(83,94,90,95,107,96,85,91,103,84,98)
pressure <- data.frame(subject,home,hospital)
pressure

# DATA ANALYSIS
# Figure 1 Blood pressure against Subject
ggplot(pressure, aes(x = subject)) +
  xlab("Subject") + ylab("Blood pressure") +
  geom_point(aes(y = home), color = "blue") +
  geom_point(aes(y = hospital), color = "red") +
  scale_x_continuous(breaks=seq(0,12,1)) +
  scale_y_continuous(breaks=seq(50,120,5))

# Figure 2
boxplot(home-hospital, ylab="Blood pressure")

# Medians and Means of home-hospital
home-hospital
summary(home-hospital)

# Correlation test for Pearson correlation coefficient
cor.test(home,hospital)

# Bootstrap the correlation coefficient
n = length(home)
Brep = 10000

home.hospital <- data.frame(cbind(home,hospital))

pearson <- function(d,i=c(1:n)){
  d2 <- d[i,]
  return(cor(d2$home,d2$hospital))
}
```

```r
bootcorr <- boot(data=home.hospital,statistic=pearson,R=Brep)
bootcorr

# Bootstrap confidence intervals
boot.ci(bootcorr, type = c('basic','norm'))

par(mfrow=c(2,1))
hist(bootcorr$t,main="Bootstrap Pearson Sample Correlation Coefficients")
plot(ecdf(bootcorr$t),main="ECDF of Bootstrap Correlation Coefficients")

# Figure 3
boxplot(home,hospital)
plot(home,hospital)
text(home,hospital, labels=subject, pos = 2)
abline(a=0,b=1)


# Histograms and QQ-plots to see for normality. Figure 4.
hist(home-hospital, freq = F,breaks=seq(-20,10,2.5))
lines(density(home-hospital), lwd=2)
qqnorm(home-hospital)
qqline(home-hospital)

#TESTS For PART 1

# t-test
t.test(home,hospital,paired = TRUE, alt ="two.sided",conf.int = T)

# Wilcoxon Signed Rank Test
wilcox.test(home,hospital,paired = TRUE, alt ="two.sided",conf.int = T)

# Lehmann-Hodges estimator
w.averages <- walsh(hospital-home)
median(w.averages)

################## PART 2 ######################
Subject <- c(1:15)
old <- c(281, 182, 373, 619, 275, 351, 242, 483, 209, 365, 391, 187, 568, 524, 227)
new <- c(56,  297, 288, 229, 200, 296, 212, 158, 249, 197, 416, 262, 148, 269, 234)
blood <- data.frame(Subject,old, new)
```

```r
# DATA ANALYSIS
# Figure 5 Subject against Clotting time
ggplot(blood, aes(x = Subject)) +
  xlab("Subject") + ylab("Clotting time") +
  geom_point(aes(y = old), color = "blue") +
  geom_point(aes(y = new), color = "red") +
  scale_x_continuous(breaks=seq(0,16,1)) +
  scale_y_continuous(breaks=seq(50,700,100))

# Data for within pair differences
old-new
summary(old-new)

# Correlation test for the Pearson correlation coefficient.
cor.test(new,old)

# Bootstrap the Pearson correlation coefficient
m = length(old)
Brep = 10000

old.new <- data.frame(cbind(old,new))

Pearson <- function(d,i=c(1:m)){
  d2 <- d[i,]
  return(cor(d2$old,d2$new))
}
bootcorr <- boot(data=old.new,statistic=Pearson,R=Brep)
bootcorr
boot.ci(bootcorr, type = c('basic','norm'))
hist(bootcorr$t,main="Bootstrap Pearson Sample Correlation Coefficients")
plot(ecdf(bootcorr$t),main="ECDF of Bootstrap Correlation Coefficients")


# Figure 7.
boxplot(old,new)

# Histograms and QQ-plots to see for normality. Figure 8 and 9.
hist(old, freq=F)
lines(density(old), lwd=2)
```

```r
qqnorm(old)
qqline(old)

hist(new, freq=F)
lines(density(new), lwd=2)
qqnorm(new)
qqline(new)

# TESTS for PART 2
#t test
t.test(old,new) # assumes old and new are normal

# Test difference in variances, robust method
z<- c(old,new); rz=rank(z)
T=abs(sd(rz[1:15])-sd(rz[16:30]))
T
K=1000; T0=rep(NA,K);
for (k in 1:K) {i=sample(1:30,30,replace=F) ;rzp=rz[i] ; T0[k]=abs(sd(rzp[1:15])-sd(rzp[16:30]))}
mean(T<T0)

# W Rank Sum Test; note the test statistic in wilcox.test is Mann-Witney's one
wilcox.test(old, new, paired = F, alt = "greater", conf.int = T)

# Calculating W_obs
z<- c(old,new); rz<-rank(z); sum(rz[16:30])

# Hodges-Lehmann estimator
HodgesLehmann(old,new)
```