Biostatistics: Exercise 04

Beate Sick, Lisa Herzog

06.10.2020

Exercise 01: Neck and shoulder disorders

Musculosceletal neck-and-shoulder disorders are common among office staff who perform repetitive tasks using visual display units. A study was carried out to determine whether varying working conditions have an impact on arm movement. The accompanying data was obtained from a sample of $n = 16$ subjects (s. below). Each observation is the time (in minutes), expressed as a proportion of the total observation time during which arm elevation was below 30 degrees. For each subject, the two measurements were obtained 18 months apart. During this period, working conditions were changed and subjects were allowed to engage in a wider variety of working tasks.

```
# before
before <- c(83, 86, 86, 83, 90, 86, 95, 73, 74, 74, 72, 81, 66, 72, 56, 75)

# after change
after <- c(80, 90, 78, 79, 84, 67, 91, 70, 58, 64, 70, 59, 66, 60, 65, 73)

# pairwise difference
diff <- after - before
```

- Does the data suggest that the true average time during which elevation is below 30 degrees differs before and after changing working conditions? Perform an appropriate test at the 10% significance level (R-Hint: `t.test(..., alternative="...", paired=..., conf.level=...)`).

```
# Since each subject is examined before and after the study,
# we have to perform a paired t-test or a one-sample t-test on the
# differences. since we want to know whether the time during which
# arm elevation is below 30 degrees changes into either direction,
# we have to perform a two-sided test:
t.test(after, before, alternative = "two.sided",
       paired = TRUE, conf.level = 0.9)
##
##  Paired t-test
##
## data:  after and before
## t = -3.001, df = 15, p-value = 0.008954
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
##  -9.702952 -2.547048
## sample estimates:
## mean of the differences
##                  -6.125

# From the result of the t-test, we can see that
# the time decreases on average by 6.75 minutes.
```

```
# We reject # the null hypothesis that the means are equal to 0
# because the 95% CI does not cover the zero
# or, equivalently, bedause the p-value is smaller than 0.05
```

Exercise 02: Muscle activation training

In order to minimize the forces acting on the spine when flying a sports airplane, it is important that pilots activate certain groups of muscles in the belly and the back during the flight. To test the effectiveness of a new training program, the muscle activation of 10 pilots was measured during a flight before and after training. This was done by using electrodes on the skin. The dataset `training.txt` can be downloaded from the webpage. (R-Hint: Since it is a .txt file, which is separted by \t, you have to read it in with `dat <- read.table(..., sep="\t", header=TRUE)`)

```
# Read the data
dat <- read.table(file = paste0(dir,"data/training.txt"), header = TRUE,
                  sep = "\t")
```

- Is the design of the experiment paired or unpaired?

```
# The design of the experiment is paired since we compare
# measurements of the same pilot before and after training.
```

- Use an appropriate plot to check whether muscle activity before and after training is normally distributed. Interpret your plot.

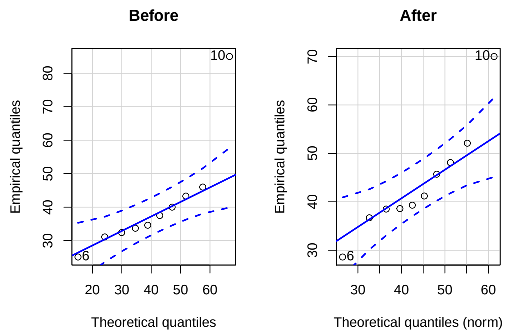
```
# We use a QQ-Plot to check the assumption of normality.
library(car)
```

```
## Loading required package: carData
```

```
par(mfrow = c(1,2))
qqPlot(dat$before, dist = "norm",
    mean = mean(dat$before),
    sd = sd(dat$before),
    xlab = "Theoretical quantiles",
    ylab = "Empirical quantiles",
    main = "Before")
```

```
## [1] 10  6
```

```
qqPlot(dat$after, dist = "norm",
    mean = mean(dat$after),
    sd = sd(dat$after),
    xlab = "Theoretical quantiles (norm)",
    ylab = "Empirical quantiles",
    main = "After")
```




```
## [1] 10  6
```

```r
# Apart from one outlier the data seems to be normally distributed.
```

- Perform a pairwise, two-sided t-test at the 5% significance level to investigate whether muscle activation changes by training. Interpret your results. What happens if you remove the outlier? (**R-Hint**: In order to remove an observation (a row) from a dataset, you can write `dat[-row,]`.)

```r
# We perform a paired t-test to investigate the difference.
# We use a two-sample t-test because we want to investigate
# if there is an effect in either direction.
t.test(dat$after, dat$before, alternative = "two.sided", paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  dat$after and dat$before
## t = 1.4667, df = 9, p-value = 0.1765
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.632348  7.652348
## sample estimates:
## mean of the differences
##                    3.01
```

```r
# There is no evidence for the alternative that the means are
# different from 0. We see that the muscle activity on average
# increases by 3.01 points but the effect is non-significant.

# Remove the outlier (10th observation)
dat2 <- dat[-10,]
t.test(dat2$after, dat2$before, alternative = "two.sided", paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  dat2$after and dat2$before
## t = 9.8516, df = 8, p-value = 9.49e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.838140 6.184082
## sample estimates:
## mean of the differences
##                5.011111
```

```r
# After removing the outlier, we see a significant improvement.
# There is high evidence that the differences within the pairs are
# different from 0.
```

- As seen in the lecture and the previous task, a t-test is not robust against outlier. Additionally, the t-test shouldn't be applied to small datasets (≤ 10) because we can't ensure that the data is normally distributed. Apply a more appropriate test (R-Hint: `wilcox.test(..., alternative="...", paired=...)`)

```r
# We could use a Wilcoxon test instead
wilcox.test(dat$after, dat$before, alternative = "two.sided", paired=TRUE)
```

```
##
##  Wilcoxon signed rank exact test
##
## data:  dat$after and dat$before
## V = 45, p-value = 0.08398
## alternative hypothesis: true location shift is not equal to 0
```

```r
# The test shows no significant difference at the 5% level. However,
# there is weak evidence for the alternative that the means are
# different from 0 (p<0.1).
```

Exercise 03: t-test with simulated data

With the following code we perform a simulation study. Read carefully through the code, execute it line by line and try to understand it.

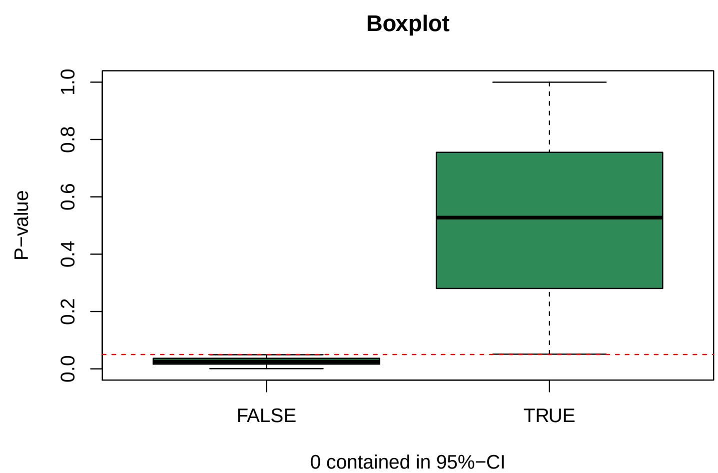
```r
set.seed(3004)
p_val <- c()
ci_lower <- c()
ci_upper <- c()
mean_diff <- c()

for(i in 1:1000){
  mean_sim  <- 1
  sd_sim <- 0.5
  n_sim <- 500
  groupa <- rnorm(n_sim, mean = mean_sim, sd = sd_sim)
  groupb <- rnorm(n_sim, mean = mean_sim, sd = sd_sim)
  test <- t.test(groupa, groupb, mu = 0, paired = FALSE)

  p_val <- c(p_val, test$p.value)
  ci_lower <- c(ci_lower, test$conf.int[1])
  ci_upper <- c(ci_upper, test$conf.int[2])
  mean_diff <- c(mean_diff, abs(diff(test$estimate)))
}

dat <- data.frame("experiment" =  1:1000,
                  "estimate" = mean_diff,
                  "ci_lower" = ci_lower,
                  "ci_upper" = ci_upper,
                  "p_val" = p_val)

dat$ci_zero <- ifelse(dat$ci_lower< 0 & dat$ci_upper > 0, TRUE, FALSE)
```

- Explain, what the code is doing

```
# We sample 500 values from a normal distribution with mean mu = 1 and standard
# deviation sigma = 0.5 and allocate the values to group A and group B.
# Then we perform an unpaired t-test at the 5% significance level with the
# default Null-Hypothesis that the mean difference is equal to zero
# (which is here obviously the case; we would expect not to reject the Null)
# This experiment was repeated 1000 times and the resulting p-values
# and confidence intervals were computed and summarized in a dataframe.
```

- Explain what you can see in the boxplots generated using the following code

```
boxplot(dat$p_val~dat$ci_zero, col = "seagreen", main = "Boxplot",
        xlab = "0 contained in 95%-CI", ylab = "P-value")
abline(h=0.05, col="red", lty=2)
```




```
# For all tests where the CI covers the zero, we get a p-value>0
# This is known as duality of p-value and CI in tests
```

- What is the number of significant tests you would expect in this experiment?

```
# Since we simulate the data, we know the Null hypothesis is true
# The tests were performed at the 5% significance level. Therefore, we have a risk
# of 5% that H0 is falsely rejected (error type I).
# We would expect about 1000*0.05 = 50 tests to show a significant difference.
```

- If you would change the mean to $\mu = 0$ and keep the standard deviation, would you expect more, less or the same number of significant test results?

```
# We would expect the same number of significant tests.
```

- Adapt the code from above so that the true mean difference between `groupa` and `groupb` is 0.1 times the standard deviation. How many significant tests do you get now? Play around with the true difference between the groups - what do you learn? In addition, change the number of observations. What do you observe?

```
set.seed(3004)
p_val <- c()
ci_lower <- c()
ci_upper <- c()
```

```r
mean_diff <- c()

for(i in 1:1000){
  mean_sim <- 1
  sd_sim <- 0.5
  n_sim <- 500
  groupa <- rnorm(n_sim, mean = mean_sim, sd = sd_sim)
  groupb <- rnorm(n_sim, mean = mean_sim + 0.1* sd_sim, sd = sd_sim)
  test <- t.test(groupa, groupb, mu = 0, paired = FALSE)

  p_val <- c(p_val, test$p.value)
  ci_lower <- c(ci_lower, test$conf.int[1])
  ci_upper <- c(ci_upper, test$conf.int[2])
  mean_diff <- c(mean_diff, abs(diff(test$estimate)))
}

dat <- data.frame("experiment" =  1:1000,
                  "estimate" = mean_diff,
                  "ci_lower" = ci_lower,
                  "ci_upper" = ci_upper,
                  "p_val" = p_val)

dat$ci_zero <- ifelse(dat$ci_lower< 0 & dat$ci_upper > 0, TRUE, FALSE)


# The larger the true difference, the higher is the probability that the test gets
# significant.
# The larger the sample size, the more likely it is to observe a significant
# result even with smaller effects that represent a group difference unequal 0.
```