Exercise 1 (Interpretation of a Confidence Interval)

In a prospective study, researchers compare the mean weight loss (in kg) in adults under diet with additional medication (treatment) and diet only (placebo). They report a mean difference (μ) of 2kg in weight loss and a corresponding 95% CI of [0.53,3.47]. Which of the following results are true, which are wrong?

- ☐ 95% of the study participants have a weight loss between 0.53 kg and 3.47 kg.

```
# Wrong, a CI of 95% doesn't mean that 95% of the
# data lie within that interval.
```

- ☐ The confidence interval covers the true difference in weight loss between treatment and placebo with a probability of 95%.

```
# True
```

- ☐ The treatment effect is statistically significant from 0 at the 5% level.

```
# True, the CI doesn't cover the 0
```

- ☐ We have no evidence against $H_0 : \mu = 0$ at the 5% level.

```
# Wrong, there is evidence against the null.
```

Exercise 2 (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 more varied work conditions would have any impact on arm movement. The accompanying data was obtained from a sample of $n = 16$ subjects (s. below). Each observation is the amount of time (in minutes), expressed as a proportion of total time observed, during which arm elevation was below 30 degrees. The two measurements from each subject were obtained 18 months apart. During this period, working conditions were changed and subjects were allowed to engage in a wider variety of work tasks.

```r
# before
before <- c(81, 87, 86, 82, 90, 86, 96, 73, 74, 75, 72, 80, 66, 72, 56, 82)

# after change
after  <- c(78, 91, 78, 78, 84, 67, 92, 70, 58, 62, 70, 58, 66, 60, 65, 73)

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

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

```r
# 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 = -3.2791, df = 15, p-value = 0.005072
## alternative hypothesis: true difference in means is not equal to 0
## 90 percent confidence interval:
##  -10.358687  -3.141313
## sample estimates:
## mean of the differences
##                   -6.75

# Obviously, the time decreases on average by 6.75 minutes. We reject
# the null hypothesis that te means are equal to 0.
```

Exercise 3 (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 programme, the muscle activation during a flight of 10 pilots was measured before and after the training. This can be done by the aid of electrodes on the skin. The dataset `training.txt` can be downloaded from the webpage:

`https://bsick.github.io/Biostatistics-Fall-2018/`

(**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")
```

(a) 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.
```

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

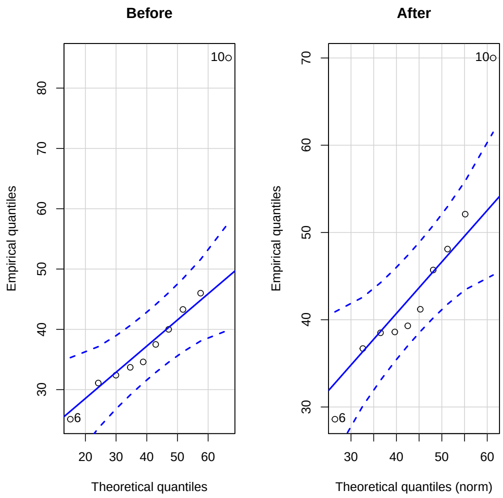
```
# We use a QQ-Plot to check the assumption of normality.
library(car)

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

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

```
# 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
```

```
# 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

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

(d) As seen in the lecture and the previous task, a t-test is not robust against outliers. 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=...)`)

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

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

# 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).
```