## Statistik un Data Science für die Informatik

Humboldt-Universität zu Berlin, SS 22, Prof. Dr. Alan Akbik

**Übungsblatt 3**

## Task 2a)

**i)**

Test choice: one-tailed test

Elaboration:

one-tailed: the alternative hypothesis: the reaction time is shorter in the control group K than that in the group B.

**ii)**

one-tailed Wilcoxon signed-rank test for independent samples

Justification:

- The sample contains a small number of observations (< 20) which makes it difficult to make an assumption about the normality of the data distribution. In this case a non-parametric alternative of a t-test is preferable.
- The control group K has an outlier with the value 0.91 which can also be taken as a signal of taking a non-parametric test given the size of the sample.
- The individuals are different in each group $\rightarrow$ the observations are independent.
- The sample size is too small to test for variance

**iii)**

The test statistic of 60 is shown not to be statistically significant (p=0.0549), indicating that alcohol consumption does not significantly increase reaction time.

## Task 3b)

**i)**

the visual representation of data for both true and forged bank notes can be found below. Based on the boxplot, the bottom part of the bank notes is on average more narrow than the the top:

```
mean top width = 10.65
mean bottom width = 9.418
```

The choice of the test was based on the following observations:

- the normality assumption is violated for the bottom width: scipy.stats.normaltest yielded p=2.12e-10 $\rightarrow$ the null hypothesis of normal distribution can be rejected).
- Based on the obseravtion above, the samples have different distributions.
- The observations in the groups are paired
- There is no hypothesis on which group has a significantly greater/smaller difference between the bottom and the top width.

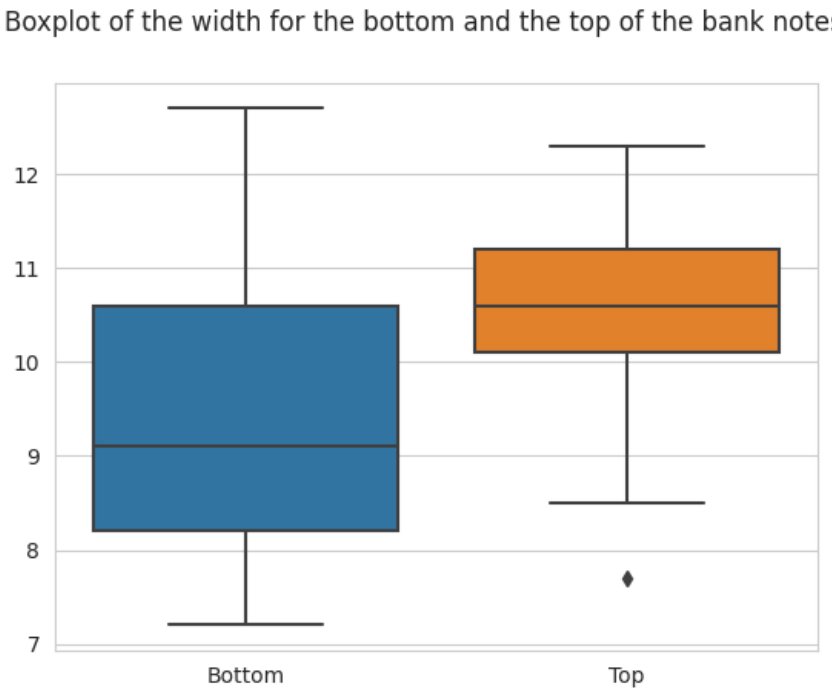


Figure 1: Data spread of bank notes

- The samples have inequal variance: the test for variance equality by scipy.stats.levene() yielded the p-value of 2.84e-21 that enables us to reject the null hypothesis of equal variance.

The test choice: two-tailed Kolmogorov-Smirnov Test
Results:

```
statistic = 0.5
p-value = 4.36e-23
```

The p-value indicates that the result statistically significant. We can reject the null hypothesis of no difference between the top and bottom width.
Conclusion: the top width is on average different from the bottom width of a bank note.

**ii)**

The visual representation of data for both true and forged bank notes can be found below. Based on the boxplot, the bottom part of both true and forged bank noted is on average more narrow than the top. This difference is even more apparent for the true bank notes:
True bank notes:

```
mean top width = 10.168
mean bottom width = 8.305
```

Forged bank notes:

```
mean top width = 11.133
mean bottom width = 10.53

mean top width = 10.65
mean bottom width = 9.418
```

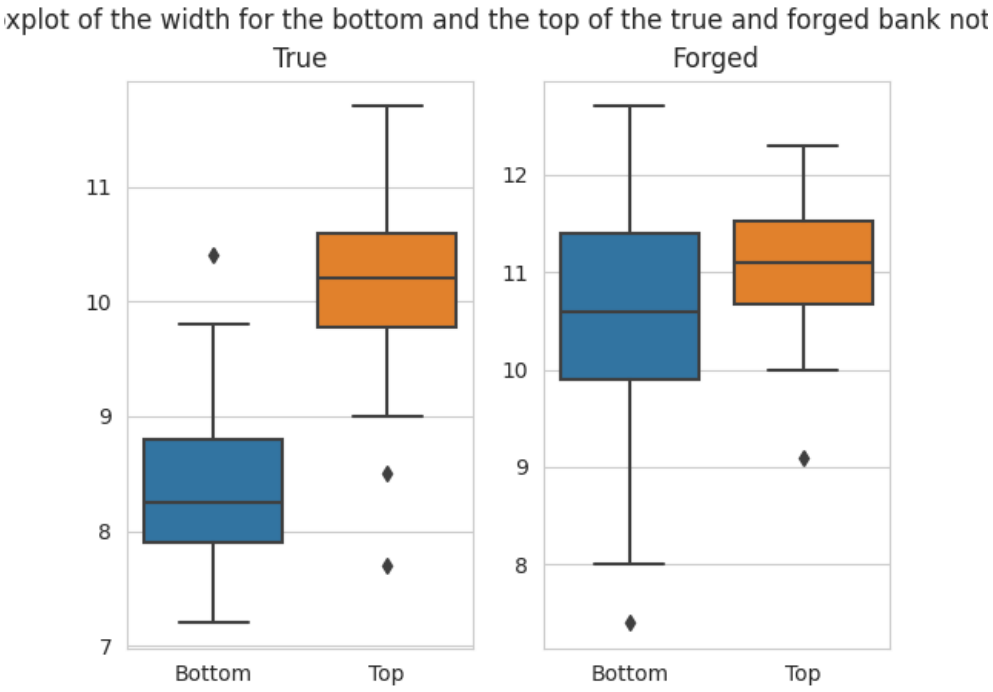


Figure 2: Data spread of true and forged bank notes

The results of the two-tailed Kolmogorov-Smirnov Test for true bank notes:

```
test statistic: 0.88
p-value: 1.349e-40
```

The results of the two-tailed Kolmogorov-Smirnov Test for forged bank notes:

```
test statistic: 0.35
p-value: 7.85e-06
```

The p-values in both groups are considerably below the alpha level of 0.05 and thus statistically significant. We can reject the null hypothesis of no difference between the top and bottom width for both true and forged bank notes.
Conclusion: the top width is on average different from the bottom width of a bank note.

## Task 3c)

**i)**

The visual representation of data for the duration of stay in Clinic 1 and Clinic 2 can be found below. Based on the boxplot, the mean values for both datasets are significantly below 500:

```
Mean duration of stay in Clinic 1: 360.9
Mean duration of stay in Clinic 2: 278.3
```

```
mean top width = 10.65
mean bottom width = 9.418
```

The choice of the test was based on the following observations:

- a test is performed on each sample separately as they are not compared to each other.

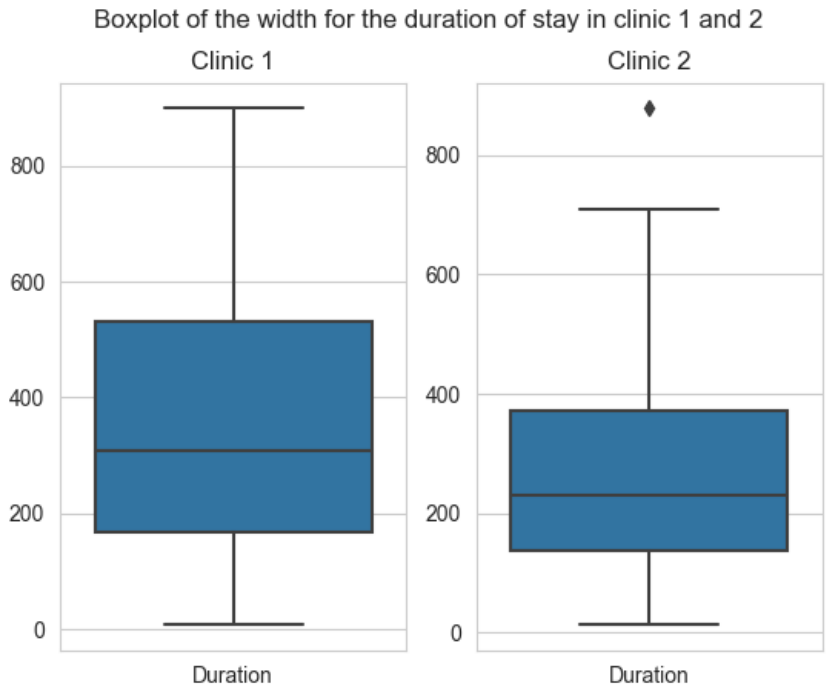


Figure 3: Duration of stay in Clini1 and Clinic 2

- according to the results of the normality test (scipy.stats.normaltest), both samples are non-normally distributed: p-value for the duration of stay in Clinic 1: 0.0007 p-value for the duration of stay in Clinic 2: 0.0194 (the null hypothesis of scipy.stats.normaltest is that data is normally distributed)
- There is an outlier in the sample for Clinic 2 (878 days). Given the size of the sample (28 entries), the outlier is expected to have a certain influence on the data behavior.
- The alternative hypothesis states that the means in each sample is smaller than a certain value (in our case, 500 days).

The test choice: one-sample non-parametric one-tailed Wilcoxon signed-rank test.
Results:

```
Clinic 1
statistic: 1631.0
p-value: 3.015e-08

Clinic 2
statistic: 37.0
p-value: 7.828e-05
```

Both results are statistically significant (both p-values are considerable below the alpha level of 0.05). Thus, the null hypothesis of no difference between the assumed mean of 500 days and the actual means in both samples can be rejected $\rightarrow$ the average stay in each clinic is statistically less than 500 days.

**ii)**

The test choice: one-sample t-test.
Justification:

- normal data distribution (p-value = 0.5387, scipy.stats.normaltest)
- the data are independant

```
statisic = 2.95
p-value=0.0037
```

Interpretation: the result of the one-sample t-test is statistically significant. The mean dose is significantly different from 55 mg/day.