Homework assignment 3 – Applied Statistics 2018

Hand in your own solutions at the start of the lecture on Friday 11/5 (14.30)

For a project IT-students were asked to consider the interface of an energy monitor for households: is it possible to make a more user-friendly device?

The students designed a new interface with a simpler, more intuitive appearance, in order to make it more user-friendly. One of the aspects they wanted to investigate is whether, for the new interface, it is easier to retrieve last month’s electricity consumption: for the old interface the task completion times of 16 users were observed and for 16 other users the completion times for the new interface (after a trial period of 2 months) were observed.

The results of both samples are shown in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Old</th>
<th>5.5</th>
<th>4.1</th>
<th>4.4</th>
<th>5.8</th>
<th>6.2</th>
<th>2.7</th>
<th>3.8</th>
<th>3.2</th>
<th>6.0</th>
<th>5.0</th>
<th>1.8</th>
<th>6.1</th>
<th>3.9</th>
<th>7.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.625</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1.552</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>New</th>
<th>3.2</th>
<th>5.4</th>
<th>3.0</th>
<th>2.4</th>
<th>4.3</th>
<th>2.2</th>
<th>2.3</th>
<th>4.5</th>
<th>4.0</th>
<th>2.9</th>
<th>3.1</th>
<th>0.9</th>
<th>4.3</th>
<th>3.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.294</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1.168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Should the observed values be interpreted as two independent samples or as paired samples? Motivate your choice,

b. Test the null hypothesis of equal variances cannot be rejected with \( \alpha = 5\% \). (Give 8 steps)

c. Can we state that the new design decreases the mean task completion time?

Conduct the test with \( \alpha = 5\% \).

Grading:

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>total</th>
<th>SPSS below</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>10</td>
<td>+1 Bonus</td>
</tr>
</tbody>
</table>

SPSS-part (directions are for SPSS 14.0, other versions similar):

- Open a new SPSS-file and first go to the “variable view”-tab (left below): name the first variable “TCT”, set the number of decimals to 1 and Label as “Task Completion Time”. Name the second variable “Design” and go to Values: 1 = Old and 2 = New.
- Then return to the Data view and enter all 32 Task Completion Times in the first column and number the Design-variable 1 and 3 respectively.
- Go to Analyze \( \rightarrow \) Compare Means \( \rightarrow \) Independent-samples T-Test and choose “Task Completion Time” as Test variable and “Design” as Grouping variable (Define values 1 and 2). OK/OK
- Check whether the table “Group Statistics” reports the same means and standard deviations as given in the exercise.
- Then consult the table “Independent Samples test” and answer the following questions for the bonus:

1. “Levene’s test on the equality of variances”, an alternative for “our” \( F \)-test on \( H_0: \sigma_1^2 = \sigma_2^2 \) (Ch. 5), What conclusion can you draw from this p-value (which is given as “Sig.” or “observed significance” in the table), at a 5% significance level?

2. The first row of the table “Independent Samples test” shows a “Sig. 2-tailed” (= the 2-sided p-value): explain why this information leads to the same conclusion as in c., and take into account that SPSS only reports the p-value of the 2-sided test.
Solutions:

a. We have two groups of 16 + 16 different: 32 independent observations from 2 populations of task completion times.

b. The following F-test confirms the assumption of equal variances.

1. Probability model: the job completion times $X_1, \ldots, X_{16}$ of the old design and $Y_1, \ldots, Y_{16}$ of the new design are independent with $X_i \sim N(\mu_1, \sigma_X^2)$ and $Y_j \sim N(\mu_2, \sigma_Y^2)$.
2. Test $H_0: \sigma_X^2 = \sigma_Y^2$ (or $\sigma_X = \sigma_Y$) against $H_1: \sigma_X^2 \neq \sigma_Y^2$ with $\alpha = 5\%$.
3. Test statistic $F = \frac{s_X^2}{s_Y^2}$.
4. Distribution under $H_0$: $F \sim F_{16-1}^{16}$
5. Observed value: $F = \frac{s_X^2}{s_Y^2} = \frac{1.552^2}{1.168^2} \approx 1.766$
6. We have a two-sided test: reject $H_0$ if $F \leq c_1$ or $F \geq c_2$.

$P(F_{15}^{15} \geq c_2) = \frac{\alpha}{2} = 0.025$, so according to the $F_{15}^{15}$-distribution: $c_2 = 2.86$

$P(F_{15}^{15} \leq c_1) = P \left( F_{15}^{15} \geq \frac{1}{c_1} \right) = \frac{\alpha}{2} = 0.025$, so $\frac{1}{c_1} = 2.86$, or $c_1 \approx 0.35$

7. Since $F = 1.786$ does not lie in the Rejection Region: we cannot reject $H_0$.
8. At a significance level of 5% we cannot prove that the variances of the job completion times are different.

c. 2 samples t-test:

1. Probability model: the job completion times $X_1, \ldots, X_{16}$ of the old design and $Y_1, \ldots, Y_{16}$ of the new design are independent with $X_i \sim N(\mu_1, \sigma^2)$ and $Y_j \sim N(\mu_2, \sigma^2)$ (Note that we assume equal variances)
2. Test $H_0: \mu_1 = \mu_2$ against $H_1: \mu_1 > \mu_2$ with $\alpha = 0.05$:
3. Test statistic: $T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{S^2 \left( \frac{1}{16} + \frac{1}{16} \right)}}$, where $S^2 = \frac{1}{2} S_X^2 + \frac{1}{2} S_Y^2$ (since $n_1 = n_2 = 16$)
4. $T$ is under $H_0$ t-distributed with $df = n_1 + n_2 - 2 = 16 + 16 - 2 = 30$
5. Observed: $t = \frac{4625 - 3294 - 0}{\sqrt{1.886 \cdot \frac{1}{8}}} \approx 2.74$, where

$s^2 = \frac{n_1 - 1}{n_1 + n_2 - 2} S_X^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_Y^2 = \frac{1}{2} \cdot 1.552^2 + \frac{1}{2} \cdot 1.168^2 \approx 1.886$ ($s \approx 1.375$)
6. Right-sided test with $\alpha = 0.05$. Rejection Region: $t \geq c = 1.697$
    where $c = 1.697$ is taken from the $t_{30}$-table, such that $P(T_{30} \geq c) = \alpha = 5\%$
7. $t = 2.74$ falls in the RR ($2.74 > 1.697$), so we can reject $H_0$.
8. At significance level 5% it is proven that the new design requires on average a shorter task completion time than the old design.
SPSS-output:

### Group Statistics

<table>
<thead>
<tr>
<th>Design</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Completion Time</td>
<td>16</td>
<td>4,625</td>
<td>1,5520</td>
<td>0,3880</td>
</tr>
<tr>
<td>New</td>
<td>16</td>
<td>3,294</td>
<td>1,1682</td>
<td>0,2920</td>
</tr>
</tbody>
</table>

### Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Task Completion Time</td>
<td>Equal variances assumed</td>
<td>2,357</td>
<td>.135</td>
</tr>
<tr>
<td></td>
<td>Equal variances not assumed</td>
<td>2,741</td>
<td>27,867</td>
</tr>
</tbody>
</table>

1. The p-value of Levene’s test on the equality of variances is $13.5\% > 5\% = \alpha$, so do not reject the null hypothesis of equal variances.
2. The 2-tailed p-value of the 2 independent samples t-test $0.010$, so the upper-tailed test has a p-value $\frac{0.010}{2} = 0.005 (= 0.5\%) < \alpha = 5\%$, so reject the null hypothesis in favour of the alternative that the TCT of the new design is structurally smaller.