Tutorial Business Analytics

Homework 2 - Solution

Exercise 2.4

32 randomly selected men and women participated in a clinical trial. The purpose of the study was to compare vegetarian diets to non-vegetarian diets. The hypothesis to be tested is: “On average, vegetarians eat fewer calories than non-vegetarians”. The sample mean for the 12 vegetarians is $\bar{x}_1 = 1780$ calories per day, while the sample mean for non-vegetarians amounts to $\bar{x}_2 = 1900$ calories per day. Moreover, the sample standard deviations are: $s_1 = 230$ and $s_2 = 250$.

a) Calculate a 95% confidence interval for the average daily intake of each group.

b) How do you assess above hypothesis considering the confidence intervals from question a)?

c) Which test is suitable for testing above hypothesis? Briefly explain your choice and perform the test with significance level $\alpha = 0.05$ and 25 degrees of freedom.

Solution

a) The confidence interval for μ_1 is:

$$\left[\bar{x}_1 \pm t_{\alpha/2}(n_1-1)\cdot\frac{s_1}{\sqrt{n_1}}\right]=[1633.86;1926.14]$$

where $\bar{x}_1=1780$, $n_1=12$ und $s_1=230$

The confidence interval for μ_2 is:

$$\left[\bar{x}_2 \pm t_{\alpha/2}(n_2-1)\cdot\frac{s_2}{\sqrt{n_2}}\right]=[1783.00;2017.00]$$

where $\bar{x}_2=1900$, $n_2=20$ und $s_2=250$

b) The computed confidence intervals are overlapping. Therefore, it cannot be concluded that the daily intake of vegetarians and non-vegetarians differs at a significance level of 5%.

c) Use the “test manual” to solve this exercise.

c) i) 2 samples ii) independent

d) $H_1: \mu_1 < \mu_2$ (Vegetarians eat fewer calories than non-vegetarians)
$H_0: \mu_1 \geq \mu_2$ (Vegetarians eat more or equal calories than non-vegetarians)
$\Leftrightarrow$
$H_1: \mu_D = \mu_2 - \mu_1 > \mu_0 = 0$
$H_0: \mu_D = \mu_2 - \mu_1 \leq \mu_0 = 0$

e) Welch-Test:

$$t_0 = \frac{\bar{x}_2-\bar{x}_1-\mu_0}{s_{\bar{x}-\bar{w}}} \text{ where df=25}$$

$$s^2_{\bar{x}-\bar{w}} = \frac{s_x^2}{n_x}+\frac{s_w^2}{n_w} \qquad s^2_{\bar{x}-\bar{w}} = 7533.3333$$

$$t_0=\frac{120}{86.79}=1.38$$

f) $\alpha=0.05$

g) df=25 which results in $t^c = 1.708$

h) $t_0 = 1.38 \ngtr t^c = 1.708$

$\Rightarrow$ H_0 cannot be rejected. Regarding a significance level of $\alpha=0.05$ it cannot be concluded that vegetarians eat fewer calories than non-vegetarians.

Exercise 2.5

a) Assess whether the following sample could possibly have been taken from a population with mean equal to 0. $(\alpha = 0.05)$

   2 3 2 4 2 4 5 2 1 4 3 0 3 2 4 5 3 3 0 1

   Solve this question manually (pen & paper) and then a second time using R (use the function “t.test()”)

b) Briefly explain the term p-Value.

Solution

a) Use the "test manual" to solve the exercise.

1.) i) 1 sample ii) σ_X unknown

2.)

$H_1 : \mu_x \neq \mu_0 = 0$ (information supplied is not correct)
$H_0 : \mu_x = \mu_0 = 0$ (information supplied is correct)

3.) t-Test

$$t_0 = \frac{\bar{x}-\mu_0}{s_X}\sqrt{n}$$

$\bar{x} = 2.65$ $S_x^2 = 2.134$

$$t_0 = \frac{2.65}{1.461}\sqrt{20} = 8.112$$

4.) $\alpha = 0.05$

5.) $t^c_{1-\frac{\alpha}{2};n-1} = t^c_{0.975;19} = 2.093$

6.) $t_0 > t^c$ true $\Rightarrow$ reject H_0

```
x <- c(2, 3, 2, 4, 2, 4, 5, 2, 1, 4, 3, 0, 3, 2, 4, 5, 3, 3, 0, 1)
t.test(x)
```

$\Rightarrow$ reject H_0

Regarding the significance level $\alpha = 0.05$ it can be concluded that the sample has not been taken from a population with a mean equal to 0.

b) The p-value is the probability to obtain another mean $\bar{x}'$, at least as different from μ_0 than the sample mean $\bar{x}$, given H_0 true. The null hypothesis will be rejected when the p-value turns out to be less than a predetermined significance level (α).