BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences
Prof. Tim Downie

Workshop 11 — Confidence Intervals and Hypothesis tests 1

One-Sample t-Tests

Exercise 1 t-Test by hand: Bottle filling machine
In Lecture 10 (confidence intervals) slides 19 and 20 we had the following example. A machine which fills 1 litre bottles with apple juice. The machine does not fill each bottle with exactly 1000 ml but a random amount that is approximately 1 litre. When operating correctly, the machine fills each bottle with an expected volume of 1000 ml.

The manager now says that if there is good evidence that the machine does not "on average" fill the bottles with 1000 ml then the machine should be repaired. The manager takes a sample of 20 bottles and the volume in each bottle is measured. He then asks you
"To do one of those statistical tests, to see if I need to repair the machine."
He provides you with the sample mean $\overline{x} = 998.6$ and the sample standard deviation $s_x = 2.251$

(a) Write down the null and alternative hypothesis for this t-test.

(b) Calculate the test statistic.

(c) Find the critical value of the test using `qt(1-aplha/2,n-1)` for a significance level of 5%.

(d) Complete the hypothesis test and give your interpretation.

Exercise 2 Significance level
In the same scenario as Exercise 1, the manager argues,
"It is costly to repair the machine. We must reduce the chance that the machine will be repaired, if it is actually functioning properly."

You reply,
"Then you should use a test with a 1% significance level."

(a) Express in words what a 5% and a 1% significance level means in this case.

(b) What scenario fits the type two Error in this example?

(c) The test statistic is the same as in Exercise 1, but now the critical value is different. Find the critical value for this test.

(d) Complete the hypothesis test interpret the result and compare with the decision in Exercise 1.

Exercise 3 Connection between confidence intervals and hypothesis tests: Body temperature
A small sample of body temperatures for 8 healthy people was obtained. The temperatures in Celsius are:

36.8 37.2 37.5 37.0 36.9 37.4 37.9 38.0

The conjecture is that the mean body temperature in the healthy population is $37.0C$

(a) Create a vector in R called `Temp` contining these values.

(b) Calculate the mean and standard deviation for the sample.

(c) Write down the null and alternative hypotheses for the above conjecture.

(d) We use the same R function `t.test()` to find a confidence interval for a population mean and to carry out a t-test. You need to provide the *null value*, `mu=`. The default confidence level is 95%, but can be changed with the argument `conf.level=`

```
> t.test(Temp,mu = 37,conf.level = 0.95)
```

(e) Find the following text in the R-output.

```
95 percent confidence interval:
xxxxxx xxxxxx
```

(f) The hidden values correspond to the limits of the confidence interval. How do you (informally) interpret this confidence interval

(g) Now find the following text in the R-output.

```
        One Sample t-test

data:  Temp
t = xxx, df = xx, p-value = xxx
alternative hypothesis: true mean is not equal to 37
```

(h) What is the test statistic?

(i) Use the p-Value to determine whether the null hypothesis willbe rejected at the 5% level.

(j) What is the interpretation of the hypothesis test?

(k) Make a comparison with the confidence interval and the hypothesis test. Does this correspond to the proposition in lecture slide 22, *Connection between hypothesis tests and confidence intervals*?

Two-Sample t-Tests

Exercise 4 Fuel consumption in cars

The data set `fuel.Rda` contains fuel consumption in miles per gallon for U.S. cars and Japanese cars. You will carry out a t-Test to see if the expected fuel consumption is the same for the two groups.

(a) Load the data into R.

```
>load("Fuel.Rda")
```

(b) Find out how many cars are from each country.

(c) Obtain the sample means and standard deviations split by country. Hint `tapply`

(d) Obtain a box plot for the fuel consumption split by country.

```
> ???(mpg~???,data=???)
```

(e) Before you do the t-Test, what result would you expect just from looking at the box plot?

(f) Notice that the two sample standard deviations are fairly similar for Japanese and for American cars, we can assume that the two groups have the same population standard deviation which we estimate using the pooled standard deviation (see homework reading).

For a two sample t-test, the R specification is

```
> t.test(x~group,data=dataframe)
```

Modify this for the fuel consumption data. Use the argument `var.equal=TRUE` to specify a test where the variances in both groups are the same.

(g) What do you conclude from the output?

1 Homework Reading

The two-sample t-test is used to determine if two population means are equal. A common application of the two-sample t-test is to see if a new manufacturing process or medical treatment is superior to the one.

There are several variations on this test:

- The data may either be two paired or two unpaired samples.

  - By paired samples, we mean that there is a one-to-one correspondence between the values in the two samples. That is, if $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, ..., Y_n$ are the two samples, then X_i corresponds to Y_i. An example would be patient blood pressure before and after an operation: there are n patients and the before blood pressure for patient i is paired with the after blood pressure for patient i. For paired samples, the approach is to calculate difference for each pair $D_i = X_i - Y_i$. A one sample t-Test is then carried out on the differences D_i and null hypothesis $\mu_D = 0$.
  - For unpaired samples, there is no one-to-one correspondence. The sample sizes for the two samples do not have to be equal. The fuel consumption example in Exercise 4 is an unpaired test, there is no one to one connection between the US and the Japanese cars

- The variances of the two samples may be assumed to be equal or unequal. Equal variances yields somewhat simpler formulae.

- In some applications, you may want to adopt a new process or treatment only if it exceeds the current treatment by some threshold. In this case, we can state the null hypothesis in the form that the difference between the two populations means is equal to some constant $H_0 : \mu_1 - \mu_2 = d_0$ where the constant is the desired threshold.

Two sided test for unpaired data

This is the they behind the fuel consumption example in Exercise 4 that you carried out using R. The two sided test for unpaired data has the hypotheses:

$$H_0 : \mu_x = \mu_y \qquad \text{vs} \qquad H_1 : \mu_x \neq \mu_y$$

Test statistic is:

$$t_{\text{stat}} = \frac{\overline{x} - \overline{y}}{\sqrt{s_x^2/n_x + s_y^2/n_y}}$$

where n_x and n_y are the sample sizes, $\overline{x}$ and $\overline{y}$ the sample means, and s_x^2 and s_y^2 are the sample variances of the two groups.

If equal variances are assumed, then the formula reduces to:

$$t_{\text{stat}} = \frac{\overline{x} - \overline{y}}{s_p\sqrt{1/n_x + 2/n_y}},$$

where

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}.$$

S_p is called the *pooled standard deviation* and is a weighted mean of the sample standard deviation of the two groups.

The critical value is the same as in the one sample t-test, but the degrees of freedom is $n_x + n_y - 2$.

The decision rule is the same as in the one sample t-test:
If $|t_{stat}| \leqslant t_{cr}$, we „accept" the null hypothesis $\mu_x = \mu_y$
If $|t_{stat}| > t_{cr}$, the null hypothesis is rejected, we do not believe that $\mu_x = \mu_y$.