ECON-C4100 - Econometrics I
Statistical tests, Exercise session 2

Antto Jokelainen, Ellen Sahlström

Hypothesis testing

- A hypothesis is a yes/no-question formulated based on a scientific problem.

- We differentiate between one-sided hypothesis testing (*Is the mean for group a smaller than the mean for group b?*) and two-sided hypotheses (*Is the difference in means different from zero?*).

- A null hypothesis must be formulated, together with an alternative hypothesis. **These two cover all possible scenarios.**

Hypothesis testing

- One-sided: $H_1 : \mu_1 - \mu_2 < d_0$
- Two-sided: $H_0 : \mu_1 - \mu_2 = d_0$, vs. $H_1 : \mu_1 - \mu_2 \neq d_0$

- The hypothesis can be tested using a t-test.

T-statistic

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - d_0}{SE(\bar{Y}_1 - \bar{Y}_2)} \tag{1}$$

where

$$SE(\overline{Y}_1 - \overline{Y}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{2}$$

This will give us the value of the t-statistic. One can remember, that a t-statistic larger than 1.96 (absolute value) allows us to reject the null hypothesis at a 5% significance level.

P-value

- Using the t-statistic we can calculate a p-value.
- The p-value is (simply put) the probability to observe the given result when the null hypothesis is true.
- A small p-value allows us to reject the null hypothesis.
- The "level of significance" refers to the p-value
- With a large n, the p-value is calculated as $p = 2\phi(-|t^{act}|)$.

Degrees of freedom

- When calculating certain statistics, the size of the sample is used and denoted n.

- However, as the calculation consists of several parts, the calculations "eat" some of the sample size.

- Therefore, we talk about degrees of freedom, as $n - x$, where x is the correction for degrees of freedom, the number of calculations done before that step.

- Some probability distributions are defined using degrees of freedom, such as Student's t, F and χ^2-distribution.