Biostatistics

In-class exercise week 4 --- Solution
Topic: hypothesis tests, p-value

1) Statistical tests, p-values, confidence intervals

Which of the following statements are correct?

a) That a test gets significant (i.e. $p < \alpha$, with e.g. $\alpha = 0.05$), means that the data shows evidence against the Null hypothesis and one rejects H_0

x True □ False

b) When using a t-test for detecting a truly existing effect, it is more likely that the tests gets significant when using a larger sample.

x True □ False

c) If a test does not get significant, one can conclude that there is no effect.

□ True x False

d) The confidence interval for the expected value which results from a t-test does always contain the mean of the sample.

x True □ False

e) Two 99% CIs for the population mean are determined from two independent samples. If the two confidence intervals do not overlap, then a two-sample t-test will provide a p-value ≤ 2%.

x True □ False

2) Hypothesis tests

a) A randomized clinical study with 500 patients in each of the two treatment groups, A and B, should check if a newly developed drug B leads to a better pain reduction for arthritis patients than the gold standard treatment drug A. We abbreviate the expected value of pain reduction with drug A or B with μ_A or μ_B.
- Is it a paired (dependent) or unpaired (independent) study design?
It is an unpaired study design.

- Formulate the H_0 and H_A for this clinical study
$H_0: \mu_A = \mu_B$, $H_0: \mu_A \neq \mu_B$
Remark: Despite the fact, that we want to show that $\mu_B > \mu_A$ we work in clinical studies always with two-sided hypothesis tests, because then we get a CI that holds all plausible values for the difference of the expected values, which allows to see changes in both direction (this is important for security reasons, since sometimes it turns out that a drug does hurt instead to help, which would not be visible in a one-sided test).

- Name the statistical test, which you plan to use, to test your hypothesis.
Unpaired t-test

- Write down the test statistic T (which can be derived from the data)
$$T = \frac{\bar{X}_1 - \bar{X}_2}{se_{pooled}} \sim t_{(n_1+n_2-2)}$$

- What is the distribution of the test statistic T under H_0?

t-distributed with the parameter, degree of freedom =1000-2=998

b) Reaction times on 500 patients in the morning (t_m) and in the evening (t_e) were measured, to investigate if the reaction time changes with the time of day.
- Is it a paired (dependent) or unpaired (independent) study design?
It is an paired study design.

- Formulate the H_0 and H_A for this study
We compute first for each patient the difference in reaction time $d_i = t_{e_i} - t_{m_i}$ and from these numbers we compute the mean $\bar{d}$
$H_0: \Delta = 0$, $H_0: \Delta \neq 0$,
in words: the Null-Hypothesis is, that the expected value of the differences is zero

- Name the statistical test, which you plan to use, to test your Hypothesis.
Paired t-test

- Write down the test statistic T (which can be derived from the data)
$$T = \frac{\bar{d}}{se(\bar{d})} \sim t_{n-1}$$

- What is the distribution of the test statistic T under H_0?
t-distributed with the parameter, degree of freedom =500-1=499