L2. Testing Hypotheses

http://edu.modas.lu/data

1. For **beer** data, calculate confidence intervals for the proportions of men and women who prefer dark beer. How many people should you question to ensure that your margin of error is < 5%?

2. Based on the **cancer** data set, calculate the interval estimation of mean (`mean`) survival time for men and women (Male=1 Female=2) subpopulations. Test the significance of the gender difference in time before the event (death).
    Note: focus on dead patients (censoring status 1=censored, 2=dead).

3. During a new study, you tested a new drug against infection. You found that 70 out of 100 mice survived, whereas the survival rate after the standard therapy was 60% of the infected population. Is this enhancement statistically significant? Use error level α=0.05.

4. In the previous task, try to estimate the minimal number of animals needed to find a significant difference between the standard survival rate of 60% and the assumingly correct 70% efficiency of the new drug.

5. An average survival time for glioblastoma patients (age<50) is 18 months. You have developed a new treatment, which should increase the survival time. You performed a pilot clinical trial (10 patients) in order to determine the positive effect. As a result, you observed the average survival of 20 months. The experimentally observed standard deviation was equal to 5 months.

6. Some COVID-19 cases were diagnosed in a single group of colleagues in a company. An epidemiological investigation was conducted to identify the virus's spread. The table gives the numbers of positive

Contacts	Positive	Tested
Low	22	325
High	10	65

    and totally tested colleagues regarding contacts with the infected group. Please test whether the infected proportions differ for "low" and "high" contacts.

7. Newly hatched **chickens** were randomly allocated into six groups, and each group was given a different feed supplement. Their weights in grams after six weeks are given along with feed types. Compare the mean weights under the following diets: "sunflower" -vs.- "soybean," "sunflower" -vs.- "casein." Do both parametric and non-parametric tests.

8. In the study **insulin**, aimed at the effect of vagus nerve stimulation on insulin secretion mongrel, dogs with varying body weights are used as model organisms. The amount of immunoreactive insulin in pancreatic venous plasma just before stimulation and 5 minutes after stimulation is given for 7 dogs. Is there any significant effect of stimulation? Provide proper hypotheses, test and calculate the p-value. Do both parametric and non-parametric tests.

9. Compare the blood pH for "129S1/SvImJ" and "A/J" mouse strains (data: **mice**). Provide proper hypotheses, p-value of the test, and conclusion. Do both parametric and non-parametric tests.

10. Run hypothesis testing for each variable of the **mice** dataset, comparing male and female populations. Correct the resulting p-values for multiple testing using Benjamini-Hochberg FDR.