Problem 1. A/B testing. (25 points)
Consider two samples with sizes $n_x = n_y = 50$ from:

- two normal distributions with unit variances, means $\mu_x = 0$ and varying μ_y;
- two normal distributions with variances one and two, means $\mu_x = 0$ and varying μ_y;
- any other distributions (not normal) of your choice. Here you should vary one parameter — location of the second distribution.

We would like to test the hypothesis H_0: $\mu_x = \mu_y$ against H_a: $\mu_x \neq \mu_y$ at nominal significance level 5%.

(a) For varying μ_y calculate the real significance level and real power testing H_0 with t-test, Welch test and Mann-Whitney test. Use 10000 simulations for each μ_y.

(b) For each case provide a plot that describes how real significance level and real power do change with μ_y.

(c) Which test is more appropriate in each case?

(d) How will the answers change with $n_x = n_y = 1000$?

Problem 2. (20 points)
Use the dataset `SIC33.csv` with the following variables:

- *output* — Value added.
- *labor* — Labor input.
- *capital* — Capital stock.

Obtain OLS estimates of the model

$$\ln Y_i = \beta_1 + \beta_K \ln K_i + \beta_L \ln L_i + u_i.$$

Use the paired bootstrap with $B = 10000$.

(a) Obtain a bootstrap estimate of the standard error of $\hat{\beta}_K$.

(b) Use this standard error estimate to test H_0: $\beta_K = 0$ against H_a: $\beta_K \neq 0$.

(c) Provide three 95% CI for β_K: classic, heteroskedasticity robust (HC1), naive bootstrap.

(d) Provide 95% bootstrap CI for the product $\beta_K \cdot \beta_L$.

(e) Estimate the same model using median (absolute value) regression. The target function is

$$\min_{\hat{\beta}} \sum_{i=1}^{n} \left| \ln Y_i - \ln \hat{Y}_i \right|.$$

Provide 95% bootstrap CI for each β.