Statistical Inference II - PhD ECOSTAT - University of Milano-Bicocca

Homework 1

To submit via e-mail by May 1, 2020, h 14:00.

1 Preliminary test of normality

Many elementary statistical textbooks suggest that assumptions should be checked before conducting statistical tests, and that tests should be chosen depending on a preliminary test about the assumptions. In particular, for comparing two independent samples, the following procedure is widely accepted:

- If for both samples the Shapiro-Wilk test for normality is not significant at $\alpha = 5\%$, use a two-sample Student T test at $\alpha = 5\%$
- If for at least one sample the Shapiro-Wilk test for normality is significant at $\alpha = 5\%$, use a Wilcoxon-Mann-Whitney (WMW) test at $\alpha = 5\%$

Perform a simulation study with

$$\text{(Normality)} \quad X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(0,1),\ Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} N(0,1) \text{ with } n = 10$$

$$\text{(Non-normality)} \quad X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exp}(1),\ Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} \text{Exp}(1) \text{ with } n = 20$$

to compare the Type I error probabilities in testing $H_0 : \mathbb{E}(X) = \mathbb{E}(Y)$ vs $H_1 : \mathbb{E}(X) \neq \mathbb{E}(Y)$ by considering

- the procedure with a preliminary test for normality described above
- always T test at $\alpha = 5\%$
- always WMW test at $\alpha = 5\%$

For the procedure with a preliminary test for normality, evaluate both

- the conditional Type I error probabilities: i) proportion of rejections of T test conditional to Shapiro-Wilk not significant in both samples and ii) proportion of rejections of WMW test conditional to Shapiro-Wilk significant in at least one sample
- the unconditional Type I error probability (proportion of rejections)

Comment on your results.
(To get accurate results, you may need a very large number of replications)

2 Confidence interval for the standardized mean

Assume $Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$.

- Describe how to construct a $1 - \alpha$ confidence interval for the standardized mean $\theta = \mu/\sigma$.
- Suppose we have observed

  $$1.96, 10.12, 8.25, 11.23, 7.24$$

  What is the 95% confidence interval for θ?

3 Testing that the mean is not zero

Assume $Y \sim N(\mu, 1)$.

- Find a rejection region of size $\alpha = 5\%$ for testing $H_0 : \mu \neq 0$ against $H_1 : \mu = 0$
- What is the power of the test?

4 Neyman-Pearson lemma

Write a simple proof of the Lemma.