A practice of (robust) statistical testing

November 25, 2024

This exercise aims to demonstrate behavior of a hypothesis tests, when its asssumptions are violated. We will compare a TTest, which assumes a normal distribution, and Mann-Whitney-U test, which does not assume any particular statistical distribution.

1. Implement generators of samples from Normal and Cauchy distributions. The generators should be functions parametrized by position (mean) and number of samples. You can keep the variance fixed.

2. Using library functions, compute histograms of test statistics for T-Test and Mann-Whitney-U tests for the following three cases:

   (a) Under hypothesis H_0, the two compared sets of samples are sampled from the same distribution. We recommend to sample them from Normal(0,1).

   (b) Under hypothesis H_{1_a}, the two compared sets of samples are sampled from two different distributions, but both distributions are normal. We recommend to sample them from Normal(0,1) and Normal(3,1).

   (c) Under hypothesis H_{1_b}, the two compared sets of samples are sampled from two different distributions and the distributions are different. We recommend to sample them from Normal(0,1) and Cauchy(3,1).

   Note: Each repetition of sampling both sets and computing the test statistics provides a single *observation of the test statistic*. Therefore to draw histograms, you need to repeat both sampling and computation of test statistics.

3. Add to the plot distribution of the test statistic under null hypothesis H_0 for both tests. For the TTest, the test statistic follows Student-t distribution with $2n-2$ degrees of freedom (where n is the sample size). The test statistics for Mann-Whitney U Test for large sample sizes follows Normal($\frac{n_1 n_2}{2}$, $\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}$,) where n_1 and n_2 are sample sizes.

   You should observe that distributions fitting the histogram do not change despite changing parameters of the distribution under H_0 hypothesis.

4. Empirically and analytically compute thresholds on test statistics, such that the probability of rejection hypothesis H_0 when it is true (Type I error) is $\alpha = 5\%$. Empirically compute Type II error, falsely accepting hypothesis H_0 while H_1 is true. You should observe how Type II error with respect to the choice of "other" distribution.