

Simulation Study For The Comparison of T-Test and the Wilcoxon-Test

Vacide Avsar Guangming Liu

December 8, 2008

Abstract

This paper tries to figure out the differences between the results of the t-test and the Wilcoxon test in a simulation study.

1 Introduction

In this project, we are trying to compare the two statistical hypothesis tests, Student's t-test and the Wilcoxon test by a simulation study.

Student's t-Test is any statistical hypothesis test in which the test statistic has a Student's t distribution if the null hypothesis is true.

Different hypothesis tests make different assumptions about the distribution of the random sample in the data. One of the assumptions for the t-test is that the data are independently sampled from a normally distributed population. This assumption about the population distribution makes the t-test be a parametric statistical test. In some cases, the data within two correlated samples may fail to meet this assumption. When this happens, an appropriate non-parametric alternative test can be found. One of these non-parametric alternative tests is called the Wilcoxon Signed-Rank Test.

Like the t-test, Wilcoxon test involves comparison of the differences between measurements. On the other hand, it does not require assumptions about the form of the distribution of the measurements. It should therefore be used whenever the distributional assumptions that underlie t-test cannot be satisfied.

In our simulation study, we'll compare the size and the power of these two tests in two different cases. For the first case, we'll take the samples from a normally distributed population and for the other case, we'll have samples from uniformly distributed population. Uniform distribution is the one which is not normal but symmetric.

To see the effects clearly, we'll keep sample size, n small and the number of simulated datasets, S large enough:

$$n = 10$$

$$S = 10000$$

For each case, we'll test whether the means of the two samples are equal or not. We expect that the results for the size and power of tests will alter for these two different cases. Comparison of the results will show us which test is more relevant and safer.

Let's start with the first case.

2 Tests for Normally Distributed Population

We generate two sets of S datasets both from normal distribution. The significance level is 0.05 .

2.1 Means of the samples are equal (Case 1)

First, we evaluate the size of the two tests. For this, we take means equal. So,

$$\mu_1 = \mu_2 = 1$$

The variances of two samples are different:

$$(\sigma_1)^2 = \frac{1}{3}$$

$$(\sigma_2)^2 = \frac{1}{12}$$

Our null hypothesis is (Ho: $\mu_1 = \mu_2$)

We use R to generate two samples. Then, under our null hypothesis, we calculate the proportion of rejection of the test by determining the proportion of the p-value that are < 0.05 since we use $\alpha = 0.05$. If the proportion does achieve the given significance level, we get a relevant size of the test. After calculation, we get the size of the t-test as **0.0521** and the size of the Wilcox-test as **0.0507**. (For the results and codes, please see Appendix part). Since they're very close to significance level, they can be said as relevant tests. Then, we can calculate the power of the tests.

2.2 Means are not equal (Case 2)

Now, we will approximate the power of the two tests under an alternative hypothesis which says that μ_1 and μ_3 are not equal.

We reuse the sample 1 with $\mu_1 = 1$ from Case 1 and then generate another data set sample 3 with $\mu_2 = 2$. The variances are not equal:

$$(\sigma_1)^2 = \frac{1}{3}$$

$$(\sigma_2)^2 = \frac{3}{4}$$