International Journal Of

Recent Scientific Research

ISSN: 0976-3031

Volume: 7(1) January -2016

CHIKHLA JUN GOGOI., COMPARISON OF POWER BETWEEN T TEST AND WILCOXON TEST IN CASE OF LOCATION PROBLEM

Chikhla Jun Gogoi

THE OFFICIAL PUBLICATION OF

INTERNATIONAL JOURNAL OF RECENT SCIENTIFIC RESEARCH (IJRSR)

http://www.recentscientific.com/ recentscientific@gmail.com

Available Online at http://www.recentscientific.com

International Journal of Recent Scientific Research
Vol. 7, Issue, 1, pp. 8215-8218, January, 2016

ISSN: 0976-3031

International Journal of Recent Scientific Research

RESEARCH ARTICLE

COMPARISON OF POWER BETWEEN T TEST AND WILCOXON TEST IN CASE OF LOCATION PROBLEM

Chikhla Jun Gogoi

Department of Statistics Dibrugarh University

ARTICLE INFO

Article History:

Received 15th September, 2015
Received in revised form 21st November, 2015
Accepted 06th December, 2015
Published online 28st January, 2016

Key words:

power, wilcoxon test, simulation.

ABSTRACT

In case of location problem for we usually use the parametric t test, provided the parent distribution is normal. In this paper the comparison is made to test the power of parametric t test and non parametric wilcoxon test in case of location problem. Here we done the comparison under normal, exponential and logistic distribution. Monte carlo simulation is done for the purpose of comparison at 5% and 10% level of significance.

INTRODUCTION

In general case of location problem most often we use t test. But the fact is that if the parent distribution is non normal then the t test is not considered as the most powerful test. In that situation we may prefer some other tests such as non parametric Wilcoxon test. In this paper power comparison is made between t test and Wilcoxon test for two independent samples under normal and logistic distribution.

Test statistics

T test--The independent samples *t*-test is used when two separate sets of independent and identically distributed samples are obtained, one from each of the two populations being compared. For example, suppose we are evaluating the effect of a medical treatment, and we enroll 100 subjects into our study, then randomly assign 50 subjects to the treatment group and 50 subjects to the control group. In this case, we have two independent samples and would use the unpaired form of the *t*-test. The randomization is not essential here – if we contacted 100 people by phone and obtained each person's age and gender, and then used a two-sample *t*-test to see whether the mean ages differ by gender, this would also be an independent samples *t*-test, even though the data is observational.

Equal sample sizes, equal variance

This test is only used when both

- The two sample sizes (that is, the number, *n*, of participants of each group) are equal;
- It can be assumed that the two distributions have the same variance.

Violations of these assumptions are discussed below.
The *t* statistic to test whether the means are different can be calculated as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2} \cdot \sqrt{\frac{1}{n}}}$$

Where

$$s_{X_1X_2} = \sqrt{s^2_{X_1} + s^2_{X_2}}$$

Here $s_{X_1X_2}$ is the grand standard deviation (or pooled standard deviation), 1 = group one, 2 = group two. $s^2_{X_1}$ and

Corresponding author:* **Chikhla Jun Gogoi
Department of Statistics Dibrugarh University

$s^2_{X_2}$ are the unbiased estimators of the variances of the two samples. The denominator of t is the standard error of the difference between two means.

For significance testing, the degrees of freedom for this test is $2n - 2$ where n is the number of participants in each group.

unequal sample sizes, equal variance

This test is used only when it can be assumed that the two distributions have the same variance. (When this assumption is violated, see below.) The t statistic to test whether the means are different can be calculated as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where

$$s_{X_1X_2} = \sqrt{\frac{(n_1 - 1)s^2_{X_1} + (n_2 - 1)s^2_{X_2}}{n_1 + n_2 - 2}}.$$

Note that the formulae above are generalizations of the case where both samples have equal sizes (substitute n for n_1 and n_2).

$s_{X_1X_2}$ is an estimator of the common standard deviation of the two samples: it is defined in this way so that its square is an unbiased estimator of the common variance whether or not the population means are the same. In these formulae, n = number of participants, 1 = group one, 2 = group two. $n - 1$ is the number of degrees of freedom for either group, and the total sample size minus two (that is, $n_1 + n_2 - 2$) is the total number of degrees of freedom, which is used in significance testing.

Wilcoxon Test

W = $\sum_{i=1}^{N} iZ_i$, N= $n_1 + n_2$

Where Z_i is a indicator variable. It take the value 1, if the ith observation from first sample and zero, otherwise.

The Monte Carlo Study

For the simulation study of the t- test, Wilcoxon test, three families of distributions are selected. These are – the Normal, the Logistic.

The study was conducted on computer at the Department of Statistics, Dibrugarh University. To generate the standard normal deviate, the method described in Monte Carlo Method by Hammersly and Handscomb(1964) were used and deviate from the other distributions were generated by using the inverse distribution function on uniform deviates.

In studying the significant levels, we first considered distributions with location parameter equal to zero and with equal scale parameters.

Specifically, we considered the distribution functions F(x -μ_i), where μ_i were the location parameters. For each set of sample N = $\sum_i n_i$, i =1,2 , the experiment was repeated 10,000 times and proportion of rejection of the true null hypothesis was recorded.

For the power study of the tests, random deviates were generated as above for each group and added to « $_i$. Proportion of rejections based on 10,000 replications at the levels .05 for different combinations of « $_i$ were recorded.

Table 1 Empirical level and power of tests under Normal distribution for equal sample sizes

Sample sizes (n_1,n_2)	Location parameters (μ_1, μ_2)	T test	Wilcoxon test
(10,10)	(0,0)	.0540	.0559
	(0,0.05)	.1951	.1947
	(0,1.0)	.5675	.5523
	(0,1.5)	.8874	.8758
	(0,2.0)	.9874	.9828
(30,30)	(0,0)	.0495	.0508
	(0,0.05)	.4741	.4545
	(0,1.0)	.9662	.9583
	(0,1.5)	1.000	1.000
	(0,2.0)	1.000	1.000

Table 2 Empirical level and power of the tests under Normal distribution with unequal sample sizes

Sample sizes (n_1,n_2)	Location parameters (μ_1, μ_2)	T test	Wilcoxon test
(20,25)	(0,0)	.0519	.0507
	(0,0.05)	.3966	.3570
	(0,1.0)	.9064	.8898
	(0,1.5)	.9981	.9971
	(0,2.0)	1.000	1.000

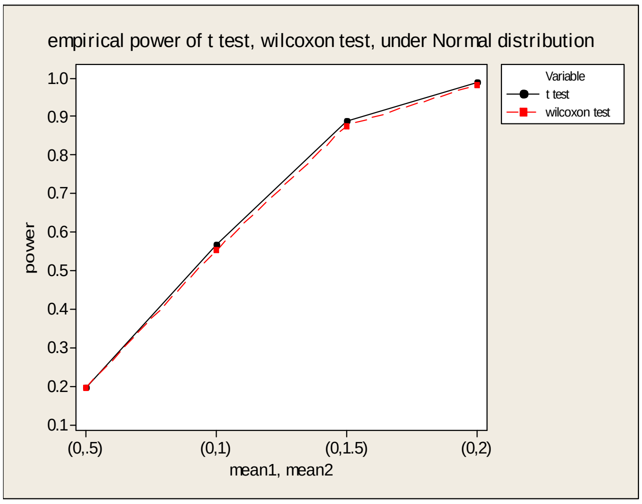

Figure 1 Empirical power of tests under Normal distribution for equal sample sizes (10,10)

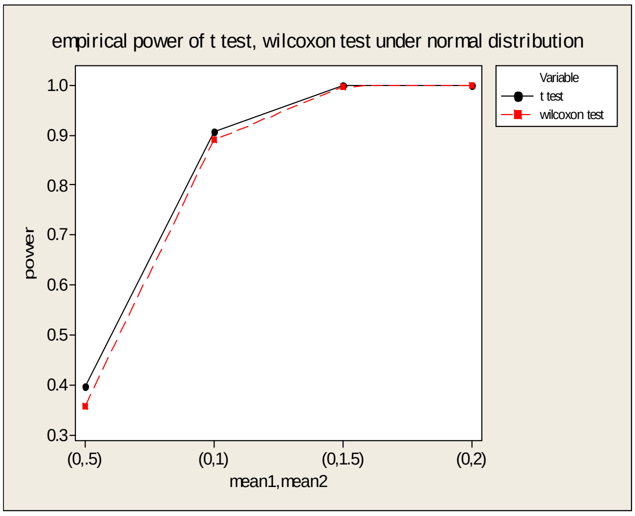

Figure 2 Empirical power of tests under normal distribution for unequal sample sizes (20,25)

Table 3 Empirical level and power of the tests under logistic distribution with equal sample sizes

Sample sizes (n_1,n_2)	Location parameter (μ_1, μ_2)	T test	Wilcoxon test
	(0,0)	.0465	.0533
	(0,0.05)	.0908	.0995
	(0,1.0)	.2302	.2604
(10,10)	(0,1.5)	.4352	.4909
	(0,2.0)	.6529	.6878
	(0,0)	.0499	.0498
	(0,0.05)	.1873	.1934
(30,30)	(0,1.0)	.5622	.5980
	(0,1.5)	.8577	.8978
	(0,2.0)	.9667	.9912

Table 4 Empirical level and power of tests under logistic distribution for unequal sample sizes

Sample sizes (n_1,n_2)	Location parameter (μ_1, μ_2)	T test	Wilcoxon test
	(0,0)	.0503	.0496
	(0,0.05)	.1547	.1604
(20,25)	(0,1.0)	.4402	.4597
	(0,1.5)	.7672	.7909
	(0,2.0)	.9451	.9560

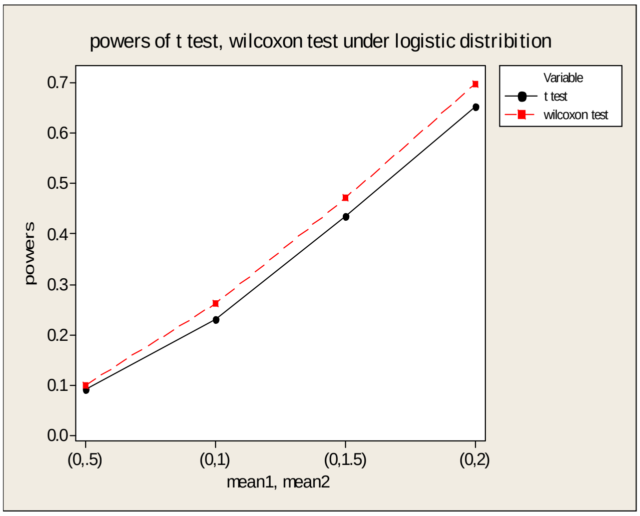

Figure 3 Empirical power of tests under Logistic distribution for equal sample sizes (10,10)

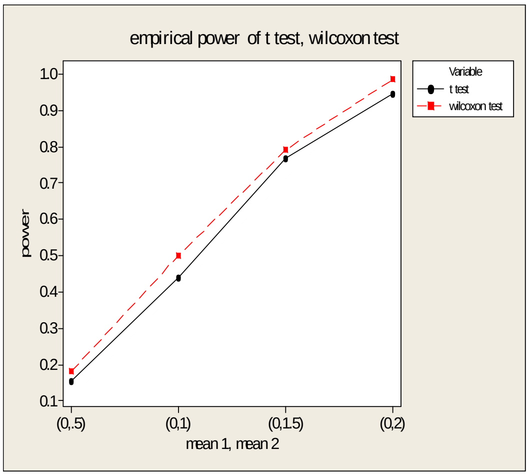

Figure 4 Empirical power of tests under logistic distribution for unequal sample sizes (20,25)

RESULTS

It is seen that traditional test i.e. the t test is more powerful than nonparametric Wilcoxon test in case of normal distribution in both equal and unequal sample sizes. On the other hand if we observe the logistic distribution then Wilcoxon test is more powerful than t test in case of equal and unequal sample sizes. Hence if the parent distribution is normal it is better to prefer t test and when the parent distribution is non normal Wilcoxon test is more preferable.

References

Bovik AC, Munson Jr DC (1986) Edge detection using median comparisons. Comput Vision, Graphics and Image Processing 33: 377–389

Edgington, E.S. (1995) Randomization Tests. 3rd Edition. Marcel Dekker, inc., New York Fried R (2007) On robust shift detection in time series. Computational Statistics and Data Analysis 52: 1063-1074

Hodges JL, Lehmann EL (1963) Estimates of location based on rank tests. Ann Math Statist 34: 598–611

Hoyland A (1965) Robustness of the Hodges-Lehmann estimates for shift. Ann Math Statist 36: 174–197

Keselman HJ, Wilcox RR, Kowalchuk RK, Olejnik S (2002) Comparing trimmed or least squares means of two independent skewed populations. Biometrical Journal 44: 478–489

Lehmann EL (1963a) Robust estimation in analysis of variance. Ann Math Statist 34: 957–966

Lehmann EL (1963b) Asymptotically nonparametric inference: An alternative approach to linear models. Ann Math Statist 34: 1494–1506 16

Lehmann EL (1963c) Nonparametric confidence intervals for a shift parameter. Ann Math Statist 34: 1507–1512

Reed III. JF, Stark (2004) Robust two-sample statistics for equality of means: a simulation study. Journal of Applied Statistics 31: 831–854

R Development Core Team (2009) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Serfling, R.J. (1980) Approximation Theorems of Mathematical Statistics. Wiley, New York

Simonoff JS (1996) Smoothing methods in Statistics. Springer, New York

Wilcox RR, Keselman HJ (2003) Modern robust data analysis methods: measures of central tendency. Psychological Methods 8: 254–274 17

****

How to cite this article:

Chikhla Jun Gogoi.2016, Comparison of Power Between T Test And Wilcoxon Test In Case of Location Problem. *Int J Recent Sci Res.* 7(1), pp. 8215-8218.

ISSN 0976-3031

9 770976 303009 >