Global Journal of Engineering and Technology Review

Journal homepage: www.gjetr.org

Global J. Eng. Tec. Review 5 (2) 30 - 37 (2021)

The Power of T-Test with Large Sample Size under the Different Condition of Sample Size and Significance Level between Real Data, Transformed Data, and Data from Monte Carlo Simulation Technique.

Natcha Mahapoonyanont[1]*, Suwichaya Putuptim[2]

[1]Thaksin University, Songkhla Campus, Songkhla, 90000, Thailand
[2]Phang-nga Community College, Phang-nga, Thailand

ABSTRACT

The power of the test is the probability that the test rejects the null hypothesis (H0) when a specific alternative hypothesis (H1) is true. The probability of occurrence of a type I error is modeled on medical research that tried to avoid the type I error, such as testing of new medicines, etc. The statistical significance level must be set to be as small as possible, and the probability of type II error would be considered later. In behavioral sciences and social sciences research, the researcher wants to avoid a type I error by determining the level of statistical significance. There are arguments of statistical significance that could affect the errors of the findings. Independent variables may have a real influence on the dependent variables but the researcher could not detect them because statistical significance was setting at a low level. Therefore, in some situations, more attention should be paid to the occurrence of the type II error, and less interest in the type I error. This may demonstrate more realistic and valid results. The objectives of this research were to compare the power of test on t – test under the condition of different sample size (n; 30, 60, 90), statistical significance (sig; .001, .01, .05), and type of data (real data, transformed data, simulation data (Monte Carlo Simulation Technique)). The research findings were found that the level of significance, sample size, and type of data had a statistically significant effect on the power of test of the t-test. The interaction between sample size and the level of significance had a statistically significant effect on the power of test of the t-test. The interaction between sample size and the type of data had a statistically significant effect on the power of the test of the t-test. The interaction between the level of significance and the type of data had a statistically significant effect on the power of the test of the t-test. The interaction between the level of significance, sample size, and type of data had a statistically significant effect on the power of the test of the t-test.

Type of Paper: Empirical

Keywords: Power of T-Test; Large Sample Size; Monte Carlo Simulation Technique

1. Introduction

The power of test is the probability that Null hypothesis will be rejected when the null hypothesis is false, which can be written as (1 - β). (Cohen, 1988; Mahapoonyanont et al, 2010) The power of the test has a relationship with the Type I error or level of statistical significance, represented by (α), which is the probability to reject the null hypothesis when the null hypothesis is true. The Type II error, represented by (β), is the probability that the Null hypothesis will be accepted when the null

ISSN 0128-2905.

hypothesis is false. The power of the test is useful for research design s such as creating criteria for determining sizes of a sample or defining the level of statistical significance. (Mahapoonyanont et al, 2010)

*Paper Info: Revised: January 11, 2021
Accepted: March 31, 2021
* Corresponding author: Natcha Mahapoonyanont
E-mail: natcha@tsu.ac.th, natcha.m@hotmail.com
Affiliation: Educational Faculty, Thaksin University, Songkhla Campus, Thailand

The power of test estimation can be divided into two types which are priori and posteriori power analysis. The priori power analysis can be done by studying documents and related researches. This will indicate effect size which is an index that measures the influence of independent variables on dependent variables. The result will be used for determining the size of the sample which is based on the effect size. There are several programs namely SYSTAT, SAS, SPSS, GPOWER, and UCLA to evaluated the posteriori power. (Myer & Well, 2003).

The above-mentioned facts showed that the power of the test is vital to the hypothesis testing of test statistics and research design. As there are few related pieces of research associated with the power of the test, the concept has not been highly regarded or widely known among researchers. Thus, the objectives of this research are to check the influence of the power of the test within the theoretical aspect, so on promote the knowledge of the concept during a wider circle and show that the facility of the test is very important to the experimental research in the behavioral science or scientific discipline. The results of the research would be presented as an alternative for researchers to adapt accordingly. The facility of the test depends on three factors (Naiyapat. 2001; Stevens. 2002) which are level of statistical significance (α — Level), sample sizes, and effect size. There are some ways to reinforce the ability of the test including flexibility in determining the level of statistical significance; increasing the sample size.

From various documents and related researches, the researcher realized that a good experiment design must consider many related factors such as level of statistical significance and sample size for instance. Researchers have to determine the appropriate value of sample size and effect size to obtain a suitable level of power of the test. Hence, to obtain an appropriate point, the researcher aims to use secondary data which is real data so that the researcher can see various situations simultaneously and obtain the best condition. We will determine whether a different type of data, the power of the test of the t - test at different levels of statistical significance, and the sample size would have different values or not.

This research aimed to study the power of test of the t-test under the different condition of sample size (n = 30, 60, 90), significance level (.001, .01, .05), and type of data (real data, transformed data, and data from monte carlo simulation technique).

2. Research objectives

1. To study the power of test of the t-test under the condition of the difference of sample size, level of statistical significance, and type of data.

2. To compare the power of test according to the following cases;
2.1 Between the different significance level
2.2 Between the different sample sizes
2.3 Between the different data types

2.4 Interaction between significance level and sample size
2.5 Interaction between significance level and data type
2.6 Interaction between sample size and data type
2.7 Interaction between significance level, sample size, and data

3. Research hypothesis

Hypothesis 1: the level of significance had a statistically significant effect on the power of test of the t-test.
Hypothesis 2: the sample size had a statistically significant effect on the power of test of the t-test.
Hypothesis 3: the type of data had a statistically significant effect on the power of test of the t-test.
Hypothesis 4: the interaction between sample size and the level of significance had a statistically significant effect on the power of test of the t-test.
Hypothesis 5: the interaction between sample size and the type of data had a statistically significant effect on the power of test of the t-test.
Hypothesis 6: the interaction between the level of significance and the type of data had a statistically significant effect on the power of test of the t-test.
Hypothesis 7: the interaction between level of significance, sample size, and type of data had a statistically significant effect on the power of test of the t-test.

4. Literature Review

4.1 Power of test

In the past, research on the power of test has been conducted on many different issues, addressing some factors influencing the power of the test, such as the study of the power of a test of one-way ANOVA after data were transformed of large samples. The research results were found that the power of the test has increased when the sample size and the statistical significance level were increased. (Mahapoonyanont, 2010)

The Anderson-Darling test for normality was found to be the best in control of type I errors at all significant levels and sample sizes. The Q statistic test can control type I errors better than the D statistic test. In small samples, the D Statistic test and the Anderson-Darling test for normality had similar high power of the test. For the medium and large samples, the Q Statistic test had the highest power of the test. (Charnarong, 2005)

Power of test in the simple linear regression analysis when there is an inconsistent variation of the error, the data with normal and the log-normal distribution of errors had a standard deviation approached 0. The sample sizes of 20, 25, 30, and 40, Breusch–Pagan test had the power of test higher than modified Levene's test. But as the samples were larger (50 and 60), the two test statistics had similarly higher power of the test and tends to be the

And when the error was exponential distribution, the Modified Levene test statistic had higher power of test when sample size and significance level increased. The model of error variance does not affect the power of test of the Breusch–Pagan test and modified Levene's test. (Chaibut. 2004)

4.2 Data transformation

Many statisticians mentioned the definition and explanation of data transformation, Kirk (1968) defined data transformation as a change of the set of scores with definite characteristics. Kirk explained that the three main reasons for data transformation in variance analysis are 1) to obtain homogeneity of variance, 2) so that the data will be a normal distribution, 3) increase the effect size of the experiment. Data transformations consist of 1) Square-root Transformation, 2) Logarithmic Transformation, 3) Reciprocal Transformation, 4) Angular or Inverse sine Transformation.

Moreover, Myers; & Well (2003), and Maxwell; & Delaney (2004) mentioned that data transformation is a process that transforms data set from non-normal distribution to normal distribution and reduces the heterogeneity of variances by applying logarithm and square root in the calculation.

Nonetheless, the disadvantage of data transformation is the difficulty in interpretation. It can be summarized that data transformation can transform data set which has a skewed distribution into normal distribution and reduce the heterogeneity of variances. Besides, data transformation also affects the power of the test by increasing its value. Monte Carlo simulation based researches showed that test statistics, which was obtained from an analysis of transformed data, gave higher power of the test statistics than using raw data and non-parametric statistics. (Rasmussen; &Dunlap. 1991, Freeman; & Modarres. 2003) and Krataithong (1999) conducted studies on the format of data transformation that can transform data into a data set with normal distribution. The best format was considered on the percentage of null hypothesis that was accepted. Their studies were based on Monte Carlo simulation and actual data.

5. Research Methodology

The research design was a 3 x 3 factorial design.

5.2 Data Preparation

The procedure of data preparation in this research as follows:

1) The real data is the score from the measurement of the non-academic variable, in this study used the score of 400 undergraduate students' happiness index from Educational faculty at Thaksin University. And randomly sampled data sets to be a sample size of 30, 60, and 90 for 1,000 pairs per each sample size, total 3,000 pairs.

2) Transformed data that has been transformed by square root transformation method. And randomly sampled data sets to be a sample size of 30, 60, and 90 for 1,000 pairs per each sample size, total 3,000 pairs. (for square root transformation, the formula is $X' = X1/2$)

3) Data from Monte Carlo technique, by using the computer program Microsoft office Excel 2017, repeated 1,000 times in each situation. And randomly sampled data sets to be a sample size of 30, 60, and 90 for 1,000 pairs per each sample size, total 3,000 pairs.

All 9,000 pairs (18,000 data sets) were examined for the normality distribution properties, following the assumption of the t-test and the analysis of variance.

The detail of the data that were used in this study showed in table 1.

5.2 Research tool

Microsoft Office Excel 2017, SPSS statistical package, and Minitab 15 were used in this study.

5.3 Data Analysis

1) Calculate descriptive statistics of data by using SPSS.

2) Calculate the power of test by using Minitab 15.

3) Transformed power of test data set to be normality distribution properties using a two-step approach for transforming continuous variables to normal provided by Templeton (2011).

4) Compare the power of test by using 3 ways ANOVA on SPSS.

Table 1 The detail of 27,000 data sets were used in this study

α	n	Type of data			Total (pairs)	Number of power of test from t - test
		real data (pairs)	transformed data (pairs)	data from monte carlo simulation technique (pairs)		
.001	30	1,000	1,000	1,000	3,000	3,000
	60	1,000	1,000	1,000	3,000	3,000
	90	1,000	1,000	1,000	3,000	3,000
.01	30	1,000	1,000	1,000	3,000	3,000
	60	1,000	1,000	1,000	3,000	3,000
	90	1,000	1,000	1,000	3,000	3,000
.05	30	1,000	1,000	1,000	3,000	3,000
	60	1,000	1,000	1,000	3,000	3,000
	90	1,000	1,000	1,000	3,000	3,000
Total		9,000	9,000	9,000	27,000	27,000

4. Results

4.1 The results of studying the power of test of the t-test under the condition of the difference of sample size, level of statistical significance, and type of data.

Table 2 Statistics of the power of the t-test from real data, transformed data, and data from Monte Carlo simulation technique under the condition of different statistical significance levels for sample size 30, 60, and 90

n	α	Statistics	real data	transformed data	data from Monte Carlo simulation technique
30	.001	M	.012	.014	.019
		SD	.026	.000	.036
	.01	M	.049	.057	.067
		SD	.067	.072	.091
	.05	M	.135	.150	.170
		SD	.120	.127	.150
60	.001	M	.018	.013	.031
		SD	.040	.026	.076
	.01	M	.065	.051	.091
		SD	.087	.066	.012
	.05	M	.162	.140	.196
		SD	.141	.117	.189
90	.001	M	.016	.014	.019
		SD	.043	.027	.000
	.01	M	.056	.054	.046
		SD	.090	.070	.067
	.05	M	.143	.144	.128
		SD	.138	.120	.114

From table 2, it was found that all of the three types of data, real data, transformed data, and data from Monte Carlo simulation technique at the sample size of 30, 60, and 90, the power of test of the t – test at the statistical significance levels of .001, .01 and .05, there is a tendency of higher mean accordingly. Skewness and kurtosis of the power of test at the statistical significance levels of .001, .01 and .05, there was a tendency of a decrease accordingly, that means when a statistical significance levels increase, a series of power of test approach more normal curve.

4.2 Results of the comparison of the power of test

The results of the comparison of the power of a test of the t-test under the condition of different statistical significance levels, sample size, and type of data have shown in table 3.

Table 3 Comparison of the power of the t-test under the condition of different statistical significance levels, sample size, and type of data

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	91.584	26	3.522	417.171	.000
Intercept	156.107	1	156.107	18488.387	.000
Sig level	86.309	2	43.155	5110.980	.000
Sample size	1.347	2	.674	79.785	.000
Data type	1.019	2	.509	60.333	.000
Sig level * Sample size	.319	4	.080	9.435	.000
Sig level * Data type	.166	4	.042	4.915	.001
Sample size * Data type	1.998	4	.499	59.154	.000
Sig level * Sample size *	.426	8	.053	6.310	.000
Error	227.747	2697	.008		
Total	475.439	2700			
Corrected Total	319.331	2699			

R Squared = .287 (Adjusted R Squared = .286)

From table 3, it was found that the level of significance, sample size, and type of data had a statistically significant effect on the power of test of the t-test. The interaction between sample size and the level of significance had a statistically significant effect on the power of test of the t-test. The interaction between sample size and the type of data had a statistically significant effect on the power of test of the t-test. The interaction between the level of significance and the type of data had a statistically significant effect on the power of test of the t-test. The interaction between the level of significance, sample size, and type of data had a statistically significant effect on the power of test of the t-test.

According to the results of the statistical significance level, sample size, and type of data had a statistically significant effect on the power of test of the t-test. The researchers conducted multiple comparisons using Dunnett's Method to investigate the effects of all independent variables, results as follow;

statistical significance level	.001 (M = .016)	.01 (M = .059)	.05 (M = .152)
.001 (M = .016)		-.043*	-.136*
.01 (M = .059)	.043*		-.092*
.05 (M = .152)	.136*	.092*	

Figure 1 multiple comparison of power of test of the t-test classification by statistical significance level

Sample size	30 (M = .075)	60 (M = .085)	90 (M = .068)
30 (M = .075)		-.010*	.007*
60 (M = .085)	.010*		.017*
90 (M = .068)	-.007*	-.017*	

Figure 2 multiple comparison of power of test of the t-test classification by sample size

Type of data	Real data (M = .073)	Transformed data (M = .071)	Data from Monte Carlo simulation technique (M = .085)
Real data (M = .073)		.002	-.012*
Transformed data (M = .071)	-.002		-.014*
Data from Monte Carlo simulation technique (M = .085)	.012*	.014*	

Figure 3 multiple comparison of power of test of the t-test classification by type of data

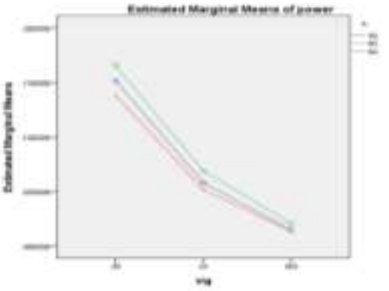


Figure 4 the interaction between sample size and level of significance had a statistically significant effect on the power of test of the t-test

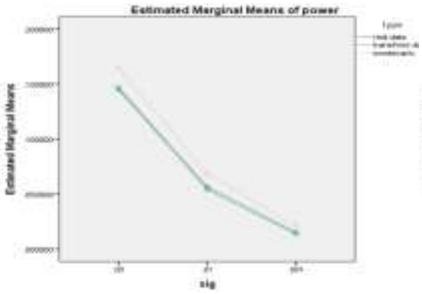


Figure 5 the interaction between type of data And level of significance had a statistically significant effect on the power of test of the t-test

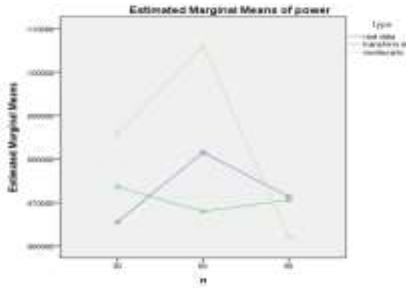
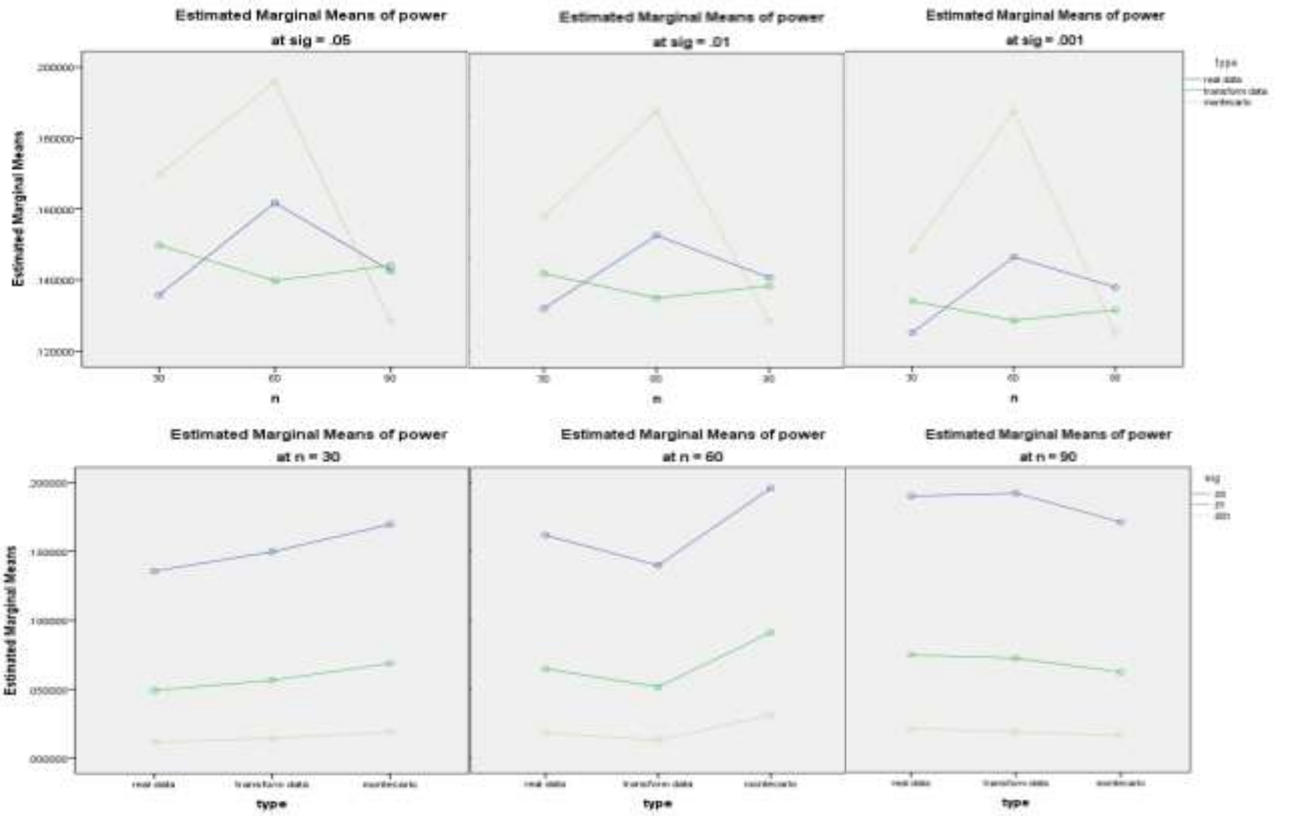


Figure 6 the interaction between type of data and sample size had a statistically significant effect on the power of test of the t-test

From the result, there was an interaction between the level of statistical significance with the sample size and the type of data had a statistically significant effect on the testing power of the T-test, shown in figure 7



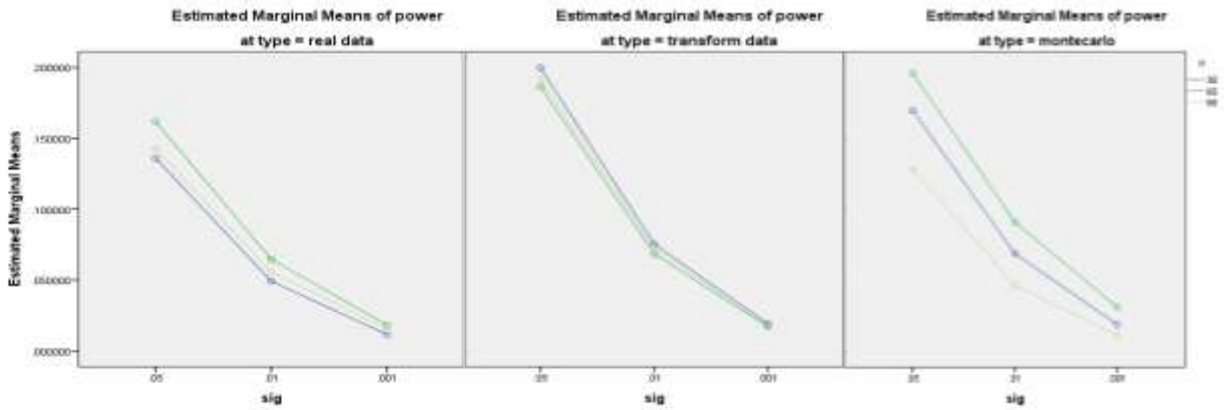


Figure 7 the interaction between statistical significance level with sample size and type of data had a statistically significant effect on the power of test of the t-test

5. Conclusion

According to the results of this study, can conclude that, at the sample size of 30, the researcher should transform data before doing a t-test to get higher power of test. But at the sample size of 60 and 90, the researcher could use real data for the t-test and the sample size at 60 is the most effective and powerful case for research design with a large sample size. And there was significant evidence to prove that the most power of test of the t-test for all types of data and all sample sizes is at the statistical significance level of .05.

Acknowledgements

This work was supported by Thaksin University Research Fund.

References

Chaibut, Sathaporn. (2004). Testing power of a test statistic in the simple linear regression analysis when there is an unstable variance in the discrepancy. Master of Science thesis. Chiang Mai: Chiang Mai University.
Cohen, J., 1992. Statistical power analysis. Current directions in psychological science, 1(3), pp.98-101. https://doi.org/10.1111/1467-8721.ep10768783
Freeman, J. and Modarres, R., 2005. Efficiency of test for independence after Box–Cox transformation. Journal of Multivariate Analysis, 95(1), pp.107-118. https://doi.org/10.1016/j.jmva.2004.08.005
Kirk, R.E., 1968. Experimental design: Procedures for the behavioral. Sciences, 93, pp.237-244.
Kirk, R.E., 1995. Hierarchical designs. Experimental design: procedures for the behavioral sciences. 3rd ed. Pacific Grove, Calif: Brooks/Cole Publishing Co, pp.476-511. DOI: https://dx.doi.org/10.4135/9781483384733.n11
Maxwell, S.E., Delaney, H.D. and Kelley, K., 2017. Designing experiments and analyzing data: A model comparison perspective. Routledge.
Myer, J.L. and Well, A.D., 2003. Research Design and Statistic Analysis. https://doi.org/10.4324/9780203726631
Mahapoonyanont, N., Mahapoonyanont, T., Pengkaew, N. and Kamhangkit, R., 2010. Power of the test of one-way Anova after transforming with large sample size data. Procedia-Social and Behavioral Sciences, 9, pp.933-937. https://doi.org/10.1016/j.sbspro.2010.12.262
Rasmussen, J.L. and Dunlap, W.P., 1991. Dealing with nonnormal data: Parametric analysis of transformed data vs nonparametric analysis. Educational and psychological measurement, 51(4), pp.809-820. https://doi.org/10.1177/001316449105100402
Stevens, J.P., 2012. Applied multivariate statistics for the social sciences. Routledge.
Templeton, G.F., 2011. A two-step approach for transforming continuous variables to normal: implications and recommendations for IS research. Communications of the Association for Information Systems, 28(1), p.4. https://doi.org/10.17705/1CAIS.02804