Available online at www.sciencedirect.com
ScienceDirect
Procedia Social and Behavioral Sciences 9 (2010) 933–937
Procedia
Social and Behavioral Sciences

WCLTA-2010

Power of the test of One-Way Anova after transforming with large sample size data

Natcha Mahapoonyanont[a]*, Tharadeth Mahapoonyanont[b], Nussara Pengkaew[a], Rojarek Kamhangkit[a]

[a]*Faculty of Education, Thaksin University, Songkhla Campus*
[b]*Lantae School, Surin Educational Service Area Office 3*

Abstract

The purposes of this study were 1) to investigate the power of the one-way ANOVA test after transforming it with large sample size data by using Real Data and 5 sample sizes (30, 60, 90, 120 and 150 students) to see if any differences exist between the tests and 2) to test which method yields the most suitable result at which sample sizes. The samples consisted of 3,600 Mathayomsuksa students in Thailand and used multi-stage random sampling. The instrument for the study was a questionnaire about democratic child-rearing with a score reliability of 0.719. The findings explain the effectiveness of the one-way ANOVA test.

Keywords: power of the test; One-way Anova; data transformation; large sample size

1. Introduction

The probability of a Type I error begins in medical research, such as the testing of a newly developed treatment, where researchers try to avoid an error. In order to achieve this, the researchers must define the level of statistical significance as low as possible. The probability of a Type II error would then be considered next. In addition, it has been found that in the behavioral and social sciences, researchers tend to avoid Type I errors by defining low levels of statistical significance too. Nonetheless, there are people who argue that the restrictiveness on the level of statistical significance may contribute to errors because independent variables may have actual influence on the dependent variables. However, the researcher may not be aware of the influence due to a low level of statistical significance. In some situations, researchers should pay more attention to Type II errors but less to Type I errors so as to obtain a suitable level of power for the test (Hays. 1994: 283-285). Howell (1992: 204) stated that a design of experimental research that restricts the level of statistical significance or defines a low level of statistical significance will increase the probability of Type II errors and reduce power of the test. Ongard Naiyapat (2001: 4-5) says that at a low level of statistical significance, researchers would find difficulty in obtaining results in an

* Natcha Mahapoonyanont. Tel.: +6-674-599-651; fax: +6-674-322-521
E-mail address: Natcha.m@hotmail.com, Natcha@tsu.ac.th

doi:10.1016/j.sbspro.2010.12.262

experiment. Researchers should design their research with a test that uses suitable power and is consistent with the conditions of the research rather than placing emphasis on the level of statistical significance.

Cohen (1988) was the first statistician who gave attention and emphasis to the power of tests in behavioral and social sciences. Cohen indicated that test power statistics show that a null hypothesis will be rejected when the null hypothesis is false. Moreover, it is also the most important part of the information on test statistics that can be applied to research data. The power of the test refers to the probability that a null hypothesis will be rejected, which can be written as (1 - β). 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 that a null hypothesis will be rejected when the null hypothesis is true. The Type II error, represented by (β), is the probability that a null hypothesis will be accepted when the null hypothesis is false. The power of the test is useful when making research plans such as criteria for determining sample size or defining level of statistical significance. Power 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 data. 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. The posteriori power analysis can be evaluated by using several programs, namely SYSTAT, SAS, SPSS, GPOWER and UCLA (Myer & Well. 2003: 212 – 215).

The above facts show that the power of the test is important to hypothesis testing, test statistics and research design. As there are very few documents and studies related to test power in the literature, the concept has not been highly regarded or widely recognized among researchers. Thus, the objectives of this research are to study the influence of the power of the test in theoretical aspects and to promote the knowledge of the concept to a wider circle, showing that the power of the test is important to experimental research in the behavioral and social sciences. The result of the research will be presented as an alternative for other researchers to adopt accordingly.

The power of the test depends on three factors which are levels of statistical significance (α - Level), sample size and effect size. There are many ways to enhance the power of the test including flexibility in determining level of statistical significance and increasing the sample size. The basic assumptions of one-way analysis of variance are 1) data must be obtained from the population with normal distribution, 2) data must be obtained using a sampling method, 3) the experiment must adopt a sampling method and 4) variance of data in each level of experimental process must be equal and independent. These basic assumptions of the statistical parameters are very important because the estimation of the statistical value is related to the sample, sample distribution and population. Sir Ronald Fisher, the inventor of the "analysis of variance" method, explained the relationship between average, variance and normal distribution in that normal distribution has two important characteristics which are average and variance. Average value would measure the bias of the estimator. Variance would measure accuracy of the data. When the data distribution is skewed or non-normal, the average value will not reflect the actual value. It is concluded that if these basic assumptions of statistics are violated, the obtained statistics would be a poor representative (Ho Yu. 2005: Online). In order to prevent this problem, it is important to use normal distribution. The method that can transform data into data sets with normal distribution is data transformation. It also allows researchers to analyze variance of small sample size without violating the basic assumptions.

There are many statisticians who give a detailed definition and explanation of data transformation; namely Roger E. Kirk (1968: 63), who defines data transformation as a change in a set of scores with definite characteristics. Kirk explains that the three main reasons for data transformation in variance analysis are 1) to obtain homogeneity of variance, 2) to ensure that the data has normal distribution and 3) to increase the effect size of the experiment. Data transformations consist of 1) Square-Root Transformation (this method is used when the average and variance in the sample are proportional), 2) Logarithmic Transformation (this method is used when the average and standard deviation are proportional), 3) Reciprocal Transformation (this method is used when the square average and standard deviation of the sample are proportional) and 4) Angular or Inverse Sine Transformation (this method is used when average and variance is proportional and the data set has binomial distribution).

Moreover, the new generation of statisticians such as Myers & Well (2003: 223-224) and Maxwell & Delaney (2004: 117) gives their concept of data transformation as a process that transforms data sets from non-normal distribution to normal distribution and reduces heterogeneity of variances by applying logarithms and square-roots in the calculation. Nonetheless, the disadvantage of this definition is the difficulty in interpretation. It can be summarized that data transformation can transform data sets which have a skewed distribution into normal distribution and reduce heterogeneity of variances. In addition, data transformation also has an effect on power of the test by increasing its value. Monte Carlo simulation-based research has shown that test statistics obtained from analysis of transformed data gave higher test power statistics than those using raw data and non-parametric statistics (Rasmussen & Dunlap. 1991: 809-820, Freeman & Modarres. 2003: online). Krataithong (1999) studied the method of data transformation that was calibrated into normal distribution. The best format was determined based on the percentage of null hypotheses that were accepted. Their studies were based on the Monte Carlo simulation and actual data.

From various documents and related studies, one realizes that a good experiment design must consider many related factors such as level of statistical significance and sample size. Researchers have to determine the appropriate value of sample size and effect size in order to obtain a suitable level of test power. In order to obtain this appropriate point, the researcher aims to use secondary data which is actual data so that the researcher can see various situations simultaneously and obtain the best condition. With the SPSS program, we will determine whether different data transformation methods and powers of the one-way ANOVA test at different levels of statistical significance and sample size would have different values or not. The objective of this research is to study the power of one-way analysis variance tests by using actual data under the following conditions: 1) when the sample sizes are different but the data transformation method and level of statistical significance remains constant; 2) when different data transformation methods are used but the sample size and level of statistical significance remains constant; and 3) when there are different level of statistical significance but the sample size and data transformation method remains constant.

2. Research Methodology

This research studies the actual data which is secondary data taken from the report "The Causal Model of Some Factors Affecting Critical Thinking Abilities" (Natcha Mahapunyanont. 2010). The scope of information in this research are:

2.1. Variable in this research are student level variables:

2.1.1. Independent variable

Independent variable is prior achievements in term of grade point average. In this regard, the researcher has divided prior achievements into three groups which are high-score group (GPA from 3.50 and above), medium-score group (GPA between 2.41–3.49), and low-score group (GPA between 0 – 2.40).

2.1.2. Dependent variable

Dependent variable is non-academic learning outcome: the democratic child-rearing as perceived by students.

3. Population and Sample

The population used in "The Causal Model of Some Factors Affecting Critical Thinking Abilities" was a group of students in the 7th through 12th grades in the basic education institutions of the Office of the Basic Education Commission under the Ministry of Education. The total number of students was 1,184,048 (230,704 from the 3rd grade level and 953,344 students from the 4th grade level).

The sample was obtained from multi-stage sampling in 90 schools under the Office of the Basic Education Commission for the year 2008. The sample consisted of 3,600 students.

4. Research Procedures

1. The population is a fake population taken from the secondary data collected from the actual source. The specified sample sizes are 30, 60, 90, 120 and 150.

2. There are three types of data transformation methods: logarithm, square-root, and reciprocal transformation, in addition to the use of raw data.

3. Using specific one-way analysis of variance to calculate test statistics, there are 3 implementation levels. There is an equal experimental unit in each level.

4. The test was conducted on assumptions at .05 and .10 levels of statistical significance and using the SPSS program for calculations.

5. Experimental Implementation

1. Determine the sample size for the study (conduct study on only large-size sample groups). Sizes of the sample group are 30, 60, 90, 120 and 150.
2. In order to obtain the desired sample sizes, a sampling of data that was collected from the actual source was made on 10 sample groups. The original data has skewness = 0.258 and a kurtosis = - 0.526.
3. 50 sets of data were transformed through three data transformation methods and raw data were used for the research.
4. Test whether the data set has a normal distribution. If the curve is asymmetrical, the researcher would conduct data sampling again from the original data set.
5. Test for homogeneity of variance. After the data has been transformed by 4 transformation methods, the researcher will test for homogeneity of variance. If the data lacks homogeneity of variance, the researcher will conduct data sampling again from the original data set.
6. After testing for homogeneity of variance, the researcher will conduct an F-test by simultaneously conducting one-way analysis of variance and an analysis of power of the test. Then the researcher will change the level of statistical significance from .05 to .10 and the size of the sample group. This process is repeated for all 50 sets of data.

6. Research Result

The result of investigated power of one-way analysis of variance tests when sample sizes were different but the transformation methods and the level of statistically significant were constant.

Table 1 Power of the test of one-way analysis of variance with real data.

sample sizes	alpha .05				alpha .10			
	RAW	SQRT	LOGT	RECT	RAW	SQRT	LOGT	RECT
n = 30	0.449	0.450	0.449	0.437	0.545	0.544	0.545	0.537
n = 60	0.555	0.589	0.553	0.540	0.644	0.646	0.646	0.637
n = 90	0.721	0.732	0.743	0.736	0.812	0.821	0.832	0.824
n = 120	0.837	0.847	0.856	0.861	0.893	0.902	0.908	0.913
n = 150	0.915	0.915	0.917	0.914	0.951	0.949	0.951	0.949

From table 1, it can be seen that at the .05 level of statistical significance and when the size of the sample group is 30, power of the one-way analysis of variance test that is obtained from reciprocal transformation has the lowest value of 0.437. It is also found that at the .10 level of statistical significance and when the size of the sample group is 150, power of the analysis of variance tests obtained from raw data and logarithmic transformation has the highest value of 0.951.

The research results show the following:

1. At the .05 level of statistical significance and when the sizes of the sample groups are 30, 90, 120 and 150, power of the test obtained from the three data transformation methods and raw data increase in rather close amounts. However, when the size of the sample group is 60, power of the test that is obtained from square-root transformation has the highest value. Powers of the test that are obtained from logarithmic and reciprocal transformations have slightly lower values than the other methods. At the .10 level of statistical significance and when the sizes of the sample group are 30, 90, 120 and 150, power of the test obtained from the three data transformation methods and raw data increase in rather close amounts. In this regard, when the sample sizes are 60 and 90, the powers of tests increase noticeably.
2. At the .05 level of statistical significance and when the sizes of the sample group are 30, 90, 120 and 150, power of the test obtained from the three data transformation methods and raw data increase in rather close amounts. However, when the size of the sample group is 60, the power of the test that is obtained from square-root

transformation has the highest power. Meanwhile, powers of the test obtained from logarithmic and reciprocal transformations and raw data have close values. At the .10 level of statistical significance and when the sample sizes are 30, 60, 90, 120 and 150, power of the test obtained from the three data transformation methods and raw data increase in rather close amounts. When the size of the sample group is 120, power of the test obtained from raw data has the lowest value. Nonetheless, the figure is close to the value obtained from the other three data transformation methods.

3. When the level of statistical significance increases and the sizes of the sample group are 30, 60, 90, 120 and 150, power of the test obtained from raw data increase slightly. When the level of statistical significance increases and the sample sizes are 60, 120 and 150, power of the test obtained from square-root transformation increases slightly. Meanwhile, when the sizes of the sample group are 30 and 90, powers of the test increase noticeably.

4. When the level of statistical significance increases and the sizes of the sample groups are 120 and 150, power of the test obtained from logarithmic transformation increases slightly. However, when the sample sizes are 30, 60 and 90, power of the test increases noticeably. When the level of statistical significance increases and the sample sizes are 120 and 150, power of the test obtained from reciprocal transformation increases slightly. However, when the sizes of the sample group are 30, 60 and 90, power of the test increases noticeably.

Acknowledgement

This study has been fund by CHE and TRF as TRF grant no. MRG5180351 (Fiscal year 2009 - 2010).

References

Cohen, J. (1988). *Statistical Power Analysis of the Behavioral Science*. 2 nd ed.New Jersey : Larence Erlbaum.

Freeman, J.; & Modarres, R. (2004, September). Efficiency of Test for Independence After Box-Cox Transformation. *Journal of Multivariate Analysis*. Retrieved November 23, 2004, from http://www.sciencedirect.com/.

Hay, W. L. (1994). *Statistic*. 5th ed. Holt, Rinehart and Winston.

Howell, D. C. (1992). *Statistic Method for Psychology*. 3rd ed. Belmont, CA : Duxbury Pr.

Ho Yu, C. (2005). *An Overview of Remedial Tools for the Violation of Parametric Test Assumtion in the SAS System*. Retrieved September 23, 2005, from http://www.creative-wisdom.com/pub/parametric_WUSS2002.pdf

Kirk, R. E.(1995). *Experimental Design: Procedures for the Behavioral Science*. 3rd ed. California: Brooks/Cole.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Murphy, K. R.; & Myors, B. (2004). *Statistic Power Analysis*. 2nd ed. New Jersey : Lawrence Erlbaum.

Myer, J. L.; & Well, A. D. (2003). *Research Design and Statistic Analysis*. 2nd ed. New Jersey : Lawrence Erlbaum.

Mahapoonyanont, N. (2010). .T*he Causal Model of some Factors Affecting Critical Thinking Abilities.* Faculty of Education, Thaksin University.

Krataythong, N. (1999). *Data Transformation to Normal Distribution* .Thesis of Master Degree in Science in Statistics, Department of Statistics. Faculty of Commerce and Accoutancy, Chulalongkorn University.

Naiyapat, O. (2001). The power of statistical tests: the consideration for Determining the sample size for research. *Journal of Behavioral Sciences*. 7 (1, May): 1-19.

Runyon, R. P.; et al. (1996). *Fundamentals of Behavioral Statistic*. 8th ed. McGraw-Hill.

Rasmussen J. L.; & Dunlap W. P. (1991 winter). Dealing with Non-normal Data: Parametric Analysis of Transformed Data VS Nonparametric Analysis. *Education and Psychological Measurement*. 51(4): 809 – 820.

Wongmark, S. (2006). *A Study Of Power Of The Test Of One Way Anova After Transforming With Small Sample Size Data By Using The Monte–Carlo Simulation Technique And Real Dat*a. Thesis of Master Degree In Educational Research And Statistics, Srinakharinwirot University.

Stevens, J. P. (2002). *Applied Multivariate Statistic for The Social Science*. 4th ed. New Jersey : Lawrence Erlbaum.

Pagano, R. R. (2004). Understanding Statistics in The Behavioral Science. 7 th ed. Ontario : Thomson Learnin