International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
www.ijmsi.org Volume 2 Issue 5 || May. 2014 || PP-09-13

Data Abnormality and Summated Ratings Scaling Method on Power of Test of Generalized Structured Component Analysis (GSCA) On Likert Scale Data, And Model Comparison between Regression Analysis and GSCA

Andi Julizar[1], WaegoHadi Nugroho[1], Solimun[1]
[1](*Department of Statistics, Brawijaya University, Indonesia)*

ABSTRACT : *GSCA is one of the preferred methods to analyze relationship between latent variables and is claimed that it doesn't need the data normality assumption. A latent variable is a variable whose value can't be measured or observed directly; instead it has to use some indicators as the accurate representations of the variable. Likert scale can be used to obtain such variable's values. Sometimes the discrete nature of Likert scale make the data didn't follow a normal distribution, which can affect the Power of test. On the other hand, based on psychometric theory, scaling method of summated ratings for transforming scores to scales is recommended before conducting a statistical analysis to make the measuring result values less biased. Many researches with Likert scale data was analyzed using regression analysis, instead of GSCA or other preferred methods for latent variables. This paper aims at confirming the Power of test of GSCA on Likert scale data with certain degree of data abnormality and between data without scaling and after scaling using previous researches which have Likert scale as their observational data. It turns out that even with the abnormality of data, GSCA still provide large Power of test. And there is no significant difference of Power between data without scaling and after scaling. Therefore, we don't need to normalize the data distribution and conducting scaling method since GSCA still gives a large Power of test. For the more accurate model, GSCA is more preferred since it has greater Power and smaller model's coefficient of variation than regression analysis.*

KEYWORDS: *GSCA, Likert scale, Normality, Power, Regression analysis, Scaling*

I. INTRODUCTION

In a research of the psychometrics, one often used variables such as attitudes or behaviors. Such variables cannot be observed or measured directly or are referred to as latent variables. Thus, the indicators are used so that the observed values of latent variables can be obtained, like using Likert scale which is the sum (or average) of responses from Likert items or scores [1]. The problem that may occur is the discrete nature of the scores which rarely makes the pattern of the observed data values follows a normal distribution. GSCA, as one of the preferred methods to analyze the relationship between latent variables beside Structural Equation Models (SEM) and Partial Least Squares (PLS) is claimed that it doesn't need the data normality assumption to be met [2], although the data abnormalities may affect the analysis' Power of Test [3]. Often we found that although the pattern of responses is different, but their sums of scores are the same with each other, which lead us to say that they have the same behavior or attitude. But the difference in the response pattern may indicate that the respondents have different behavior or attitude [4].

Therefore, based on the psychometric theory, the Summated Ratings (SMR) scaling needs to be done to the responses (scores) in order to obtain new responses (scales) so that the distance between the scales can be seen more clearly. And based on the sum of scales, the respondents can be placed on a continuum so that their behavior or attitude can be measured more objectively [5]. But, in fact, many studies that use Likert scales data did not perform the scaling first, instead the data, which still in the form of scores, are directly used in the analysis. And most of them were using Multiple Regression Analysis instead of GSCA or other preferred analyzes for latent variables. The Power of test is a probability to reject the null hypothesis (H_0) when H_0 is false. The value depends on the significance level (α), sample size (n), and effect size (ES). The importance of Power Analysis comes from the fact that most of empirical researches on the social and behavioral science begin with formulating and testing with hope of rejecting H_0 as a confirmation to the fact of the phenomenon under study [6]. The Power Analysis can be prospective (a priori) or retrospective (post hoc) one. Prospective analysis is used to determine the sample size in order to achieve the target Power, while retrospective analysis calculates Power by sample size (n) and effect size (ES). In this paper, we do retrospective analysis to find out the analysis' Power.

The relationship between α, n and ES toward Power are positive (Figure 1) where the increase of ES, n and α make the value of Power increase too [7]. The ES characterizes the model's goodness of fit, so the model fit index can be considered as ES [8].

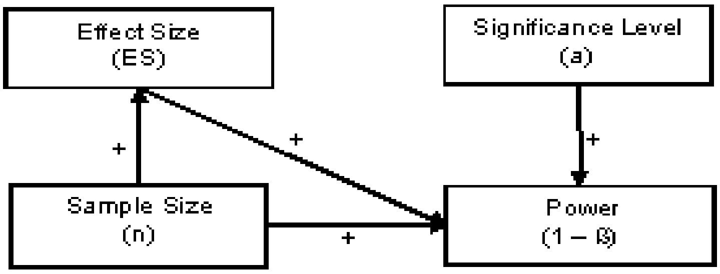

Figure 1: The Relationship of α, n and ES toward Power [7]

Data with normal distribution will produce greater Power than the data that did not follow a normal distribution. Given the normality of the data is also influenced by n, so the formula to obtain the value of Power is:

$$Power = 1 - \beta \propto \frac{ES\ \alpha \sqrt{n}}{s} \tag{1}$$

where s is square root of model's variance.

II. RESEARCH METHOD

Four data sets from theses and dissertations that used Likert scale data as their observations are used to compare their analysis' Power of Test with GSCA. The four data sets are grouped into 2 categories based on p-value of the Kolmogorov-Smirnov for testing the degree of data abnormality. Category 1 and 2 are data sets with response variable and one or more predictors which have the p-value of less than 1% ($p < 1\%$) and from 1% to 5% ($1\% < p \leq 5\%$), respectively.

Table 1. Data set sources

Data set	Research title	Writer	Remark
1.	Effect of Situational Leadership Styles of Work Productivity (Case Study in PT. Leces Paper Probolinggo)	Dona Era Septyana	Thesis BrawijayaUniversity Malang 2009
2.	Effect of Training and Development, Compensation, and Commitment to Employee's Performance (in Muhammadiyah University Malang)	Imam Muhtar	Thesis BrawijayaUniversity Malang 2005
3.	Effect of Cultural, Social, Price, Product, Promotion, and Service Location, Service of the Decision to Eat Fresh Fish (Case Study in Samarinda)	Helminuddin	Dissertation BrawijayaUniversity Malang 2005
4.	Analysis of Effect and Perception of Quality and Brand Image of Consumer's Buying Decision to Customer on Sanken Service Centre Pluit Branch	DewiFajar Indah	Thesis Bina Nusantara University Jakarta 2008

After that, each data set will be scaled with summated ratings method and be observed the increasing on p-value of normality test then it will be conducted using GSCA in order to obtain the fit indices and analysis' Power for each category and between data without scaling and after scaling. And the four data sets that are used to answer this paper's purposes are presented in Table 1.

III. ANALYSIS OF DATA

3.1. The data abnormality, scaling and Power

By using the Kolmogorov-Smirnov method for testing the normality of data, the obtained p-values from all variables in each data set are presented in Table 2. As seen in the table, some p-values increase after scaling is conducted, but the very small p-values are not significantly increased. The Increasing of the p-values are due to the decrease of Z which indicates the difference between the distribution of observational data and theoretical normal distribution is getting smaller. This explains why the data become closer to a normal distribution after summated ratings scaling is conducted.

Table 2. Data sets source

p-value category	Data set	Variables	Without scaling		After scaling	
			Z	p	Z	p
1 (p< 1%)	2	X_1	2.178	0,000	1.374	0,046
		X_2	1.894	0,002	1.479	0,025
		X_3	2.173	0,000	2.109	0,000
		Y	2.929	0,000	2.379	0,000
	3	X_1	3.602	0,000	3.813	0,000
		X_2	3.360	0,000	3.714	0,000
		X_7	3.656	0,000	3.316	0,000
		Y	3.458	0,000	3.634	0,000
2 ($1\% \leq p < 5\%$)	4	X_1	1.484	0,024	1.360	0,049
		Y	1.441	0,032	1.417	0,036
	1	X_3	1.506	0,021	1.223	0,100
		X_4	1.494	0,023	1.287	0,073
		Y	1.590	0,013	1.458	0,028

From each data set, the summary of GSCA results is presented in Table 3 where the sample size for each data set is 100 observations and $\alpha = 5\%$.

Table 3. Summary of GSCA results

p-value category	Data set	Without scaling			After scaling		
		FIT	SRMR	Power	FIT	SRMR	Power
1 (p< 1%)	2	0.412	0.142	0.99990227	0.400	0.141	0.99986161
	3	0.502	0.178	0.99985806	0.485	0.174	0.99979126
2 ($1\% \leq p < 5\%$)	4	0.419	0.164	0.99997432	0.402	0.154	0.99841010
	1	0.444	0.233	0.99990333	0.439	0.226	0.99989141

Based on Table 3, paired sample t-test is conducted to conclude whether there is a significantly difference between FIT, SRMR and Power of data without scaling and data after scaling. The test is applied using bootstrap method since there are only 4 observations. The results are FIT from data without scaling is significantly different with FIT from data after scaling where scaling method makes the FIT smaller, which is not good, but SRMR and Power did not differ significantly between data without scaling and after scaling. And the Powers of p-value from category 1 and category 2 are also similar, which are quite large.

The large value of Power is due to the using of bootstrap method to estimate standard errors which makes the probability to obtain small standard errors and greater t-statistics becomes large, which likely to allow in rejecting the null hypothesis which states that the estimate coefficients are not different from zero. The using of bootstrap method for estimating standard errors makes GSCA does not require the assumption of data normality to be met. The idea of bootstrap is the sample data is a representation of the population in which the sample was drawn, makes the resampling data are what would be obtained if more unit samples is taken from the population. Thus the precision of the statistics still can be obtained and consequently, the estimation of the parameters has adequate Power to decide their significance.

3.2. Regression analysis and GSCA

Table 4 presents the results of estimation of parameters (path coefficients) from regression analysis and GSCA which the standard error for each path coefficient for each analysis is estimated using bootstrap with 100 replications.

From the estimation results, it is seen that the data set 1 and 4 give the same decisions related to the significance of the estimate path coefficient, while from the data set 2 and 3, GSCA provide more significant coefficients than regression analysis.

Table 4. Parameters estimation results between regression analysis and GSCA

Data set	Variable	Regression analysis			GSCA		
		Coefficient	SE	T	Coefficient	SE	T
1	X_1	-0.03	0.08	-0.38	-0.06	0.12	-0.47
	X_2	0.41	0.13	3.15*	0.54	0.17	3.15*
	X_3	-0.03	0.10	-0.30	-0.06	0.17	-0.36
	X_4	-0.29	0.10	-2.90*	-0.30	0.12	-2.58*
2	X_1	0.08	0.10	0.80	0.08	0.11	0.79
	X_2	0.07	0.04	1.75	0.11	0.04	2.58*
	X_3	0.25	0.06	4.17*	0.41	0.00	348.14*
3	X_1	0.08	0.05	1.60	0.12	0.02	6.85*
	X_2	0.02	0.07	0.29	0.10	0.05	2.13*
	X_3	0.01	0.03	0.33	0.00	0.01	0.30
	X_4	-0.02	0.05	-0.40	-0.17	0.03	-5.93*
	X_5	0.05	0.04	1.25	0.09	0.02	5.86*
	X_6	0.07	0.03	2.33*	0.38	0.01	29.80*
	X_7	0.23	0.07	3.29*	0.35	0.03	10.54*
4	X_1	-0.14	0.11	-1.27	-0.15	0.05	-2.87*
	X_2	0.39	0.12	3.25*	0.36	0.02	16.37*

*) significant at 5% level

To determine which model is more accurate, the Power and model's coefficient of variation can be used to describe the suitability of the model. Table 4 shows that the GSCA produce larger t-test statistic value than regression analysis. This indicates that the coefficients of the model can be predicted with GSCA more accurately and more have significant effect on the response variable.

Based on the conventions from Cohen (1992), the t-statistic can also be used to measure ES by calculating the variation percentage of variables that can be explained, by the formula:

$$r^2 = \frac{t^2}{t^2 + df} \tag{2}$$

where r^2 is the ES and df is the degree of freedom.

Equation (2) shows that the t- statistic is directly proportional to the ES, and by Equation (1), ES is proportional to the Power, so the greater the value of the t- statistic, the greater the Power of the analysis.

The model's coefficient of variation (CV) for each data set is shown as in Table 5 which is calculated by the formula:

$$CV = {}^{s}/_{\bar{y}} \tag{3}$$

where s and $\bar{y}$ is square root of model's variance and mean scores of response variable respectively.

The GSCA's coefficient of variation is smaller than regression analysis', where the lower the coefficient of variation, the smaller the error relative to the prediction values. Based on the Power and model's coefficient of variation, GSCA models provide more accurate predictions (closer to the true value) rather than regression analysis for latent variables using Likert scale data. In other words, GSCA has the greater probability to obtain the significant coefficients.

Table 5. Model's coefficient of variation of regression analysis and GSCA

Data set	Mean of Y	RMSE	SRMR	Coefficient of variation	
				Reg. analysis	GSCA
1	18	1.537	0.233	0.085	0.013
2	16,46	1.116	0.142	0.068	0.009
3	13,17	1.793	0.178	0.136	0.014
4	8	1.797	0.164	0.225	0.021

IV. CONCLUSION

The summated ratings scaling method gives the greater probability to make the observational data distribution closer to a normal distribution since the Z-statistics of Kolmogorov-Smirnov method is getting smaller, but makes the total variance of all variables explained the particular model specification (FIT) decrease. And the Likert scale data is robust to use with GSCA as Power of test between the results of the data analysis with the scores (before scaling) and the scales (after scaling) is relatively the same, which is quite large. So, the summated ratings scaling method is not recommended to be conducted on a Likert scale data if GSCA is used as an analytical tool. Then there is no need to normalize data since GSCA gives no significantly difference of Power of test with different degrees of data abnormality. If the data were highly skewed, then GSCA still can be used without having small Power of test.

This is due to the bootstrap method is used to estimate standard errors of the coefficients which resulting small sampling error, even though the data did not follow a normal distribution. This makes the parameter estimation becomes more precise, hence the estimation has adequate Power to decide whether an estimate parameter is not different from zero. GSCA is the more preferred method than regression analysis to give estimates of path coefficients (parameters) because it has the greater Power and the smaller model's coefficient of variation. Thus, GSCA gives parameter estimation more accurate than regression analysis and has greater probability to obtain significant coefficients.

REFERENCES

[1] J. S. Uebersax, Likert Scales: Dispelling the Confusion, [Online], *http://www.john-uebersax.com/stat/likert.htm*, 2006.
[2] H. Hwang and Y. Takane, Generalized Structured Component Analysis, *Psychometrika*, *69(1)*, 2004, 81-99.
[3] M. Mendes, The effects of Non-Normality on Type III Error for Comparing Independent Means, *Journal of Applied Quantitative Methods*, *2(4)*, 2007, 444-454.
[4] S. M. Smith and G. S. Albaum, *Fundamentals of Marketing Research*, 1st Ed. (Thousand Oaks: SAGE Publications, Inc., 2004).
[5] S. Azwar, *Dasar-DasarPsikometri*, 1st Ed. (Yogyakarta: PustakaPelajar Offset, 2013).
[6] J. Cohen, Statistical Power Analysis, *Current Directions in Psychological Science*, *1(3)*, 1992, 98-101, Sage Publication, Inc.
[7] H. M. Park, Hypothesis Testing and Statistical Power of a Test, *Working Paper*, University Information Technology Services (UITS) Center for Statistical Power and Mathematical Computing, Indiana University, 2010.
[8] B. Thompson, A Suggested Revision to the Forthcoming 5th Edition of the APA Publication Manual – Effect Size Section, [Online], *http://people.cehd.tamu.edu/~bthompson/apaeffec.htm*, 2000.