A Parametric Approach Using Z-Test for Comparing 2 Means to Multi-Group Analysis in Partial Least Square Structural Equation Modeling (PLS-SEM)

Asyraf Afthanorhan[1*], Ahmad Nazim[2] and Sabri Ahmad[3]

[1]*Department of Mathematics, Faculty of Science and Technology, University Malaysia Terengganu, 21030 Kuala Terengganu, Malaysia.*
[2]*National Population and Family Development Board Malaysia, Ministry of Women, Family and Community Development, Building 12B Jalan Raja Laut, 50350 Kuala Lumpur, Malaysia.*
[3]*School of Informatics and Applied Mathematics, University Malaysia Terengganu, 21030 Kuala Terengganu, Malaysia.*

Authors' contributions

This work was carried out in collaboration between all authors. All authors read and approved the final manuscript.

Method Article

ABSTRACT

Aims: Multi-group analysis can be known as modeling the moderator variable since the strength of this method to moderates the influences of exogenous on endogenous variable. Basically, researchers interest to employ such method to extend their study to be more intense and practical.
Sampling: This paper work used stratified sampling which is one of the probability sampling. Thus, the parametric method can be applied as set of statistical assumption.
Methodology: Thus, multi-group analysis has become convenient to be practice in second generation modeling namely Structural Equation Modeling (SEM). Nevertheless, the limitation of Covariance Based Structural Equation Modeling (CB-SEM) causes the difficulties of researchers to further their studies. Hence, Partial Least Square Structural Equation Modeling (PLS-SEM) introduced to satisfy the necessity of researchers. Truthfully, multi-group analysis is not implement

_____

Corresponding author: E-mail: ash_18raft@yahoo.com;

in most of PLS-SEM software.

Results: Therefore, this paper intend to demonstrate a parametric approach using z-test to attain the probability level with the help of SmartPls 2.0. Generally, z-test comprised of two types of comparing for each groups namely means and population proportion. However, this paper work apply comparing of means to attain the value of z-scores. Consequently, the aimed of this paper work is success since the implementation of z-test approach to multi-group analysis.

Conclusion: In this case, categorical variable which is constituted for 2 groups namely gender group was the one to be moderator variable. The findings reveal that all the exogenous construct that link on endogenous construct fail to meet the required level of family wise error rate.

1. INTRODUCTION

Nowadays, Partial Least Square Structural Equation Modeling (PLS-SEM) often to be an attractive method due to the capability and benefit of methodology established. Previously, Covariance Based Structural Equation Modeling (CB-SEM) will be a main preferences for some researchers to carry on their research more deeply and comprehensive and of course to substitute the first generation of modeling. However, the statistical assumption is required to satisfy first before subsequent the further analysis [1,2]. Thus, PLS-SEM is an alternative method for any researchers that has faced difficulties to carry on their research. In particular, their ideology to justify analogous concept is unlimited since they do not have to rely on statistical conditions.

In seek of obsessing to create and proved their concept on the basis of PLS-SEM, they start to undertake research in many issues related on their expertise. One of the interested method is multi-group analysis. Indeed, some of the articles regarding on this approach has become an intense topics to employ such as from [3,4,5,6,7] and other prominent academicians.

Multi-group analysis can be known as moderating effect in which to examine whether the moderator variable has a potential to increase or decrease the influence of dependent variable [1,8]. Previously, multi-group analysis can be tested by using covariance based structural equation modeling but since this particular method has meet a limitation, thus, PLS-SEM is the only one method can be carry out. [3,4,6,7,8,9,10,11,] had explain the behavior of PLS-SEM and indeed has a strength to conduct the research work.

Normally, moderator variable would be a nominal or ordinal variable such as gender, level of education, status, citizenship, religion or others depend on the scope of the research. In this case, this paper work demonstrates the capability of gender as moderator variable in PLS-SEM. The aimed of this paper is to determine which group whether male or female is more important to contrive the side effect towards the form of dependent variable (Motivation).If the current result shows that the estimation of dependent variable is significant, means that, the moderator variable is exists.

Therefore, there are several approaches has been introduced to ascertain the scholar make their decision based on the reports. [9] had stress on this method using parametric and non-parametric test to demonstrate the readers of their strength, limitation, technique, and efficiency. Thus, this paper work decides to reveal other one approach of parametric test that involving on the large data which is z-test or z-score approach. Basically, z-test to be test once the raw data are considered as normal distribution that can be interpreted by using histogram, Kolmogorov-smirnov, Shapiro wilks and something else. In order to complement the potential of parametric approach on the large data, the authors decide to disclose the calculation of z-test when comparing of two means or groups.

The calculation of z-test would be illuminated in the subsequent subtopic with the formula provided. In this case, the scholars have to do a manual steps to generate the probability value.

2. MULTI-GROUP ANALYSIS

Multi-group analysis is usually to identify whether other factor can affect the influence of causal

effect between exogenous and endogenous construct [1,9]. Generally, the categorical variable such as gender, level education, status, citizenship, or other factor will be preferred to let these factors plays their role as moderator variable. In particular, moderator variable is certified to identify of whether the influence of exogenous and dependent variable can be affected by other factor in certain circumstances. However, the continuous variable also can be conducted if the researchers has a good reason to support their substantive concept.

In CB-SEM using Analysis Moment of Structural (AMOS) application, this particular method can be tested with the difference of chi-square test by comparing of structural model for two groups (example: Male versus Female). Chi-square test once consider as non-parametric test if conducted for one group solely but if the process is involving for more than one group, this approaches was automatically considered as parametric test. This is because the distribution of chi-square different test is the same thing as z-test. This approach has been justify by some of articles regarding on the chi-square difference test by [4,12,13].

Nevertheless, CB-SEM is quite cumbersome and complicated since this method relies on statistical assumption. In other words, if the requirement of CB-SEM is failed, the finding will be meaningless. In fact, CB-SEM have a lot of requirement that should be satisfied by all users if they intend to carry on their studies. Thus, PLS-SEM was established to solve the limitation of CB-SEM so as help the researchers to constantly continuous on their work.

In PLS-SEM, most of application currently are not employed of multi-group analysis unless SmartPls 3.0 [14]. However, this article is interested to implement of multi-group analysis using SmartPls 2.0 [5] with the introduced of z-test to determine the probability level. Previously, the authors demonstrate the parametric test using equal and unequal variance of t-test but since the limitation of this approaches has been discovered, thus, z-test was conducted for this research work. Hence, this aimed of the research work is to provide a parametric approach using z-test to multi-group analysis so as to extend the capability of aforementioned method. In statistics

discipline, z-test or z-score can be implement once the data attained is large sample size which is greater than 30.

Conversely, t-test can be implement if the data obtained was below than 30 [15,16,17]. Indeed, most of the articles stand their point to use of equal and unequal t-test approach for their research but the finding can be argue. Thus, the execution of z-test is conducted to multi-group analysis.

3. THEORETICAL FRAMEWORK

Theoretical framework is performed once the researchers identify which one of the factor that should be carry on for the multi-group analysis. Since the aimed of this research paper to guide the calculation of z-test approach, thus, the explanation regarding of each factors applied are neglected. In this case, the authors implement five constructs that has been approved to be a main problems. In particular, four out of five constructs will be an exogenous construct namely Benefit, Barrier, Challenge, and Motivation respectively as presented in Fig. 1.

The data were collected using stratified sampling technique because this method is more relevant and efficient rather than the simple random sampling [12]. The data collected including of four higher education institution chosen at area of Kuala Terengganu, Malaysia. Then, the authors collected absolute 453 respondents using these techniques. In this instances, the authors had addressed to use gender group to be implement in multi-group analysis, in particular, 159 and 193 respondents are represent for male and female group respectively.

Based on the Kolmogorov-smirnov result, the data collected are normally distributed and subsequent timely to employ the z-test approaches in this research. Indeed, there are several techniques can be applied for testing the normality of data but the authors decide to apply Shapiro-wilks since this method perceived is the most powerful to identify the normality of data [18]. The framework of this particular analysis is presented as follows:
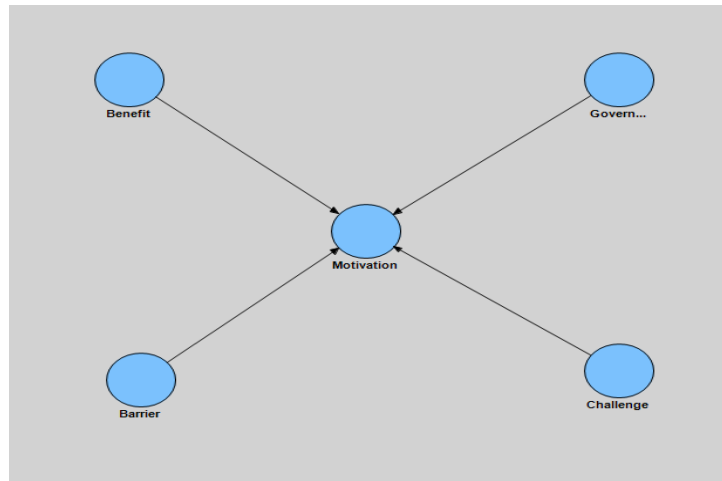
Fig. 1. Theoretical Framework

4. Z-test For Comparing Means to Multi-Group Analysis in PLS-SEM

In statistics discipline, z-test have two types that can be explained for one group and two groups. In the basic of multi-group analysis, two groups and above are considered to carry on this particular method. In the accordance of [19], the usage of z-test is adequate to put into practice of multi-group analysis for large sample size. In other words, the t-test for large sample size is inadequate to be practice since the elementary of t-test is limited for small sample size. Z-test for two population proportion is used when the researchers want to know whether two populations or groups (example: male and female; urban and rural; single and married, indigenous and non-indigenous) are significant difference on some influences of exogenous towards endogenous constructs. Thus, there are two requirement that should be stressed such as:

1. A random sample of each of the population groups to be compared
2. The sample size should be greater than 30
3. The data should be in categorical
4. The equation of z-test is present as below:

$$\text{Z-test} = \frac{(p1-p2)-0}{\sqrt{p(1-p)(\frac{1}{n1}+\frac{1}{n2})}}$$

Where:

a. P_1: Population for Male
b. P_2: Population for Female
c. p: Probability of gender group
d. n_1: Total population for Male
e. n_2: Total population for Female

Formula for p as follows:

$$p = \frac{x1+x2}{n1+n2}$$

Where:

a) X_1: The number of Male
b) X_2: The number of Female
c) N_1: Total population for Male
d) N_2: Total population for Female

However, this paper works is decided to use the methodology of comparing means including of two groups (Male. and Female). Basically, this approach illuminated once the studies had established the value of standard error or standard deviation. If the researchers fail to perform the standard error in their research, z-test to comparing between population proportions can be used as aforementioned above. Then, the formula of z-test of comparing means is presented as follows:

$$\text{Z-test} = \frac{(p1-p2)-0}{\sqrt{p(1-p)(\frac{1}{n1}+\frac{1}{n2})}}$$

Where:

a) P_1: Population for Male
b) P_2: Population for Female
c) p: Probability of gender group
d) n_1: Total population for Male
e) n_2: Total population for Female

Now, the authors will guide the readers on how to execute the multi-group analysis using Smart Pls 2.0 and the calculation to attain the

probability level (family wise error rate) as follows:

The first thing is the scholars should execute the full data by PLS algorithm and Bootstrapping approach. PLS algorithm is aimed to provide the outer loading and outer weight for researchers to identify the strength of each manifest variables (items). Bootstrap technique is used to generate the standard error for each constructs, thus, t-test will be performed.

Then, identify the moderator variable that will be implement. In this case, the authors use gender (male and female) to be a moderator variable.

Then, split the data into two model that represent structural model for Group 1 (male) and Group 2 (female)

Execute both structural model as the full model previously.

Copy the result of male and female respectively in other word document so that these result will be extended to carry on the subsequent steps.

The calculation of z-test will be present at the subtopic of findings.

5. FINDINGS

As aforementioned, the authors promised to demonstrate the calculation of z-test population proportion using gender group as categorical variable to be moderator factor. Since the procedure of multi-group analysis is given, thus,

the researchers should have the same matter to attain the probability level. The result of Table 1 and Table 2 can be obtained by Smart Pls 2.0 as presented below. All of these results can be obtained in default report in Smart Pls 2.0.

Table 1 is the result of male that comprised of original sample, sample mean, standard error and t-statistics. All of these result was appeared once the researchers execute the bootstrap technique. Based on the T-statistics given, there are two exogenous construct namely Benefit and Government have meet the required level of probability level which is greater than 1.96 (p-value ≤ 0.05). Conversely of other factors since they failed to meet the required level. By inspecting through for each column of original sample, Benefit factor is the most importance factor to influence on Motivation factor. Thus, it can be perceived that this group strongly agree that Benefit factor is the most significant that should be addressed.

Table 2 is a result for female group that comprised of original sample, sample mean standard error and t-statistics same as male group. Based on the t-statistics performed, there are one variable that present the different hypothesis which is Barrier factor. For male group, they disagree that Barrier factor is significant on Motivation factor differ for female group. This group seems belief this factor has a significant impact on Motivation. However, both group are agree to decide Benefit and Government factors are have significant impact on Motivation. Afterwards, both of these results will be computed in a new table as Table 3:

Table 1. Path coefficient for male

Male	Original sample (O)	Sample mean (M)	Standard error (STERR)	T statistics (|O/STERR|)
Barrier -> Motivation	0.078119	0.0813	0.052801	1.479489
Benefit -> Motivation	0.688298	0.6868	0.064688	10.640206***
Challenge -> Motivation	0.012209	0.0274	0.054873	0.222503
Government -> Motivation	0.124517	0.1250	0.061750	2.016464**

*: p-value ≤ 0.01; **: p-value ≤ 0.05; ***: p-value ≤ 0.01*

Table 2. Path coefficient for female

Female	Original sample (O)	Sample mean (M)	Standard error (STERR)	T statistics (|O/STERR|)
Barrier -> Motivation	0.0776	0.0785	0.0398	1.9520*
Benefit -> Motivation	0.6922	0.6890	0.0468	14.7998***
Challenge -> Motivation	0.0134	0.0214	0.0381	0.3518
Government -> Motivation	0.1197	0.1237	0.0447	2.6780***

*: p-value ≤ 0.01; **: p-value ≤ 0.05; ***: p-value ≤ 0.01*

Table 3. Multi-group analysis using Z-test

Z-Test Approach	Male	Female	Z-Score	P-value
Barrier -> Motivation	0.0813	0.0785	0.1051	0.9124
Benefit -> Motivation	0.6868	0.6890	0.0088	0.99202
Challenge -> Motivation	0.0274	0.0214	0.402	0.68916
Government -> Motivation	0.1250	0.1237	0.04	0.9681

Table 3 is a result of z-test approach that comprised of male, female, z-score and p-value for each construct provided. Based on the p-value result, all of these construct failed to meet the required level of 0.05. Thus, it can be indicates that male and female do not have a significant different between exogenous and endogenous construct. The calculation of z-test is performed as follows:

5.1 Hypothesis

H_0: Male do not has a significant different with female between Barrier and Motivation
H_1: Male has a significant different with female between Barrier and Motivation

5.2 Test Hypothesis

$$Z\text{-test} = \frac{(p1-p2)-0}{\sqrt{p(1-p)(\frac{1}{n1}+\frac{1}{n2})}}$$

$$Z\text{- test} = \frac{(0.0813-0.0785)-0}{\sqrt{(\frac{0.0398}{159}+\frac{0.05281}{193})}};$$

Z-test = 0.1051

5.3 Conclusion

1. Z-test will be considered significant different once the absolute value of z-test is greater than 1.96 (p-value < 0.05). In this case, the findings shows that the absolute z-test is below than 1.96. Thus, one can be conclude that the male do not has a significant different with female between Barrier and Motivation.
2. However, some of the researchers fail to contribute the p-value to provde a better understanding for the readers. Thus, this paper works shows the step by step to attain p-value using the calculator as follows:

   a) Press "Mode" two times
   b) Press "1" which is represent for SD

c) Then, press "Shift" and "3" that will present [P (Q(R(→ t]
d) And last press "1". In this stage, the readers can include the value of z-test. In this case, the authors type 0.1051 as presented above in calculator.

3. This procedure is repeated for the subsequent path analysis.

6. CONCLUSION

In this study, the authors address z-test approach to carry on the multi-group analysis. As aforementioned, z-test is appropriate to be conducted since the data used are large sample size. Thus, the requirement of equal and unequal variance t-test that has been proposed by [20] are limited. This is because the t-test is normally be practice for small sample size, thereby, this method is inadequate for large sample size. Thus, this paper work decide to demonstrate the application of z-test with the help of Smart Pls 2.0 software.

By regarding on the findings, the z-test was success to be implemented for multi-group analysis since this test are achieve the requirement needs for z-test population proportion. Besides, the procedure to employ is also ease to be conducted along of the p-value. On the use of z-test approach towards four exogenous constructs, all of these variables failed to meet the requirement of family wise error rate (p-value ≤ 0.05). Hence, one can be conclude that there are no significant different between male and female on each of causal effect.

7. LIMITATION

Limitation is often happen when carry on the research paper and should be improved for the future research. In this case, the authors had claims the parametric approach using z-test is appropriate to be conducted for the large sample size. However, these particular methods become meaningless if the scholars intend to advance their research using the small sample size. This

is because the formula in z-test is modified to the large sample and if the scholars interested to apply the small sample size, the t-test is required.

In addition, this research is limited to four higher education institutions at Kuala Terengganu only. Thus, these findings provided are not holistic to represent for the whole countries since the sample size prevailed is represent at Kuala Terengganu. In other states, the findings would be totally different since the issue address is particular for the students in Kuala Terengganu.

Besides, the implementation of z-test in this paper is limited to compare between 2 groups that involve of two groups in particular male and females group so as address in objective research. But if the scholars interested to apply the moderator variable comprised of more than 2 groups, this method would be fail to meet. Thus, the scholars have to seek other approaches that manage to analyze the analysis more than 2 groups

8. CONTRIBUTION

Contribution is important to let the readers are more understanding to the benefit of the z-test approach. There are three genuine that can be a main factors on the research paper. The first thing is the scholar would know how to conduct the multi-group analysis using the z-test approach. Previously, we had been learn on how to calculate the z-test but we are still ambiguities of their purpose. In this paper, the result was serve the z-test to compare between 2 groups in PLS-SEM in order to attain the probability level.

The second things is the scholar can expand their research to identify whether of male or female are more benefited on dependent variable (Motivation). In particular, the researchers manage to identify which group has provide less contribution (Female) based on method applied. Thus, we can provide a thoroughly analysis to focus on female group in the future research.

In addition, the researchers manage to use this particular approach for other discipline such as marketing, social science, management, economic and something else since this method is relevant to be implemented. Besides, the findings obtained would be more understanding to help the managerial to make their decision to improve their flaws.

9. RECOMMENDATION

Once objective research are achieve, there might have suggestion or recommendation to improve the strength of research. In this case, the authors had address the issues regarding on the limitation of potential t-test approach to multi-group analysis. In fact, z-test also had significant to be implemented. The first thing is the requirement that should be achieve are the normality of data. In statistics discipline, z-test is considered as parametric test since the distribution is bell-shaped curved. Thereby, the normality using Shapiro-wilks, Kolmogorov-Smirnov, histogram and other related test should be conducted first. In accordance of [7], Shapiro-wilks are powerful once the data obtained are small sample size whereas Kolmogorov-Smirnov test are useful for large sample size.

The second things is z-test is appropriate for large sample which is greater than 30. If the sample size less than 30, the result will be meaningless due to the deficiency of small sample size. The last things is about the moderator effect (gender) applied is might inefficient due to the failed of research hypothesis. Thus, the future research should be highlight the other factor that has a potential to discriminant of each categorical variable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Afthanorhan A, Ahmad S, Safee S. Moderated mediation using covariance based structural equation modeling. Journal of Advances in Natural and Applied Sciences. 2014;8(8):108-115.
2. Afthanorhan WM, ABW. Hierarchical component using reflective-formative measurement model in partial least square structural equation modeling (PLS-SEM). International Journal of Mathematics and Statistics Invention (IJMSI). 2014;2(2):55-71.
3. Hair JF, Ringle C. M, Sarstedt M. PLS-SEM: Indeed a silver bullet. The Journal of Marketing Theory and Practice. 2011; 19(2):139-152.
4. Hair JF, Sarstedt M, Ringle CM, Mena JA. An assessment of the use of partial least

squares structural equation modeling in marketing research. Journal of the Academy of Marketing Science. 2012; 40(3):414-433.

5. Ringle CM, Wende S, Will A. SmartPLS 2.0 (beta); 2005.

6. Sarstedt M, Wilczynski P. More for less? A comparison of single-item and multi-item measures. Die Betriebswirtschaft. 2009; 69(2):211.

7. Sarstedt M, Henseler J, Ringle CM. Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. Advances in International Marketing. 2011;22:195-218.

8. Afthanorhan WMABW, Ahmad S. Modelling The multigroup moderator-mediator on motivation among youth in higher education institution towards volunteerism program. International Journal of Scientific & Engineering Research (IJSER). 2013;4(7):5.

9. Afthanorhan A, Nazim A, Ahmad SA. Parametric approach to partial least square structural equation modeling of multigroup analysis (PLS-MGA). International Journal of Economic, Commerce, and Management. 2014;2(10): 15.

10. Henseler J. PLS-MGA: A non-parametric approach to partial least squares-based multi-group analysis. In Challenges at the interface of data analysis, computer science, and optimization Springer Berlin Heidelberg. 2012;495-501.

11. Henseler J, Ringle CM, Sinkovics RR. The use of partial least squares path modeling in international marketing. Advances in international marketing. 2009;20:277-319.

12. Awang Z, Ahmad JH, Zin NMModelling job satisfaction and work commitment among lecturers: A case of UiTM Kelantan. Journal of Statistical Modeling and Analytics. 2010;1(2):45-59.

13. Afthanorhan WMABW, Ahmad S. Path analysis in covariance-based structural equation modeling with amos 18.0.

European Journal of Business and Social Sciences (EJBSS). 2014;2(6):10.

14. Henseler J, Dijkstra TK, Sarstedt M, Ringle CM, Diamantopoulos A, Straub DW, Calantone RJ. Common beliefs and reality about pls comments on rönkkö and evermann (2013). Organizational Research Methods; 2014. 1094428114526928.

15. Afthanorhan WMABW, Ahmad S, Mamat I. Pooled Confirmatory Factor Analysis (PCFA) using structural equation modeling on volunteerism program: A step by step approach. International Journal of Asian Social Science. 2014;4(5):642-653.

16. Ahmad S, Afthanorhan WMABW. The importance-performance matrix analysis in partial least square structural equation modeling (PLS-SEM) with smartpls 2.0 M3. International Journal of Mathematics Research. 2014;3(1):1-14.

17. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proceedings of the National Academy of Sciences. 1984;81(11):3443-3446.

18. Afthanorhan WMAW, Razali, Hidayah. Improving energy conservation using six sigma methodology at Faculty of Computer and Mathematical Sciences (FSKM), Universiti Teknologi Mara (UITM), Shah Alam. Asian Journal of Economic Modeling. 2014;2(2):52-68.

19. Chin WW, Dibbern J. A permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross of information system services between Germany and the USA, in. Vinzi VE, Chin WW, Henseler J, Wang H (eds). Handbook of partial least squares: Concepts, Methods and Applications in Marketing and Related Fields, Berlin: Springer; 2009.

20. Keil M, Tan BC, Wei KK, Saarinen T, Tuunainen V, Wassenaar A. A cross-cultural study on escalation of commitment behavior in software projects. Mis Quarterly. 2000;299-325.