Presented at the DLSU Research Congress 2019
De La Salle University, Manila, Philippines
June 19 to 21, 2019

Using Latent Class Analysis in Determining Risk among Filipino Young Adults and Adolescents

Coleen Kate S. Sison[1], Katrine Joyce A. Soriano [1*] and Rechel G. Arcilla[1,]
[1] *Mathematics and Statistics Department, De La Salle University*
**Corresponding Author: katrine_soriano@dlsu.edu.ph*

Abstract: Latent Class Analysis (LCA) is a person-centered approach in grouping respondents into classes of latent variables using observed categorical data. The analysis was used on six risk indicator variables: alcohol consumption in a day, cigarette smoking in a day, marijuana use in the past 30 days, number of sexual partners for the past 3 months, depression, and suicidal ideation, obtained from the database of a larger study by Sta. Maria, Arcinas and Arcilla in 2014 involving a random sample of 1757 students from state university and colleges of two provinces in Region IV-A of the Philippines. The main objective that the study intends to accomplish in preventing the continuous widespread of health compromising activities and in reducing the consequences of dysfunctional actions, is to determine which group of people, given certain situations, are more likely to experience or witness the severity of said costs. By doing so, the study can better aid the concerned officials in creating preventive policies directed at the group of young adults and adolescents that are more exposed to the risk behaviors. Through the comparison of the different class models generated, LCA has resulted to an optimum of four subgroups: High-Risk, Low-Risk, Poor Mental Health, and Legal Substance User. The probability that respondents in a given class have engaged in a dysfunctional action, and the probability of membership to each labeled class were established.

Key Words: Risk Behaviors; Latent Class Analysis; Young Adults and Adolescents; Philippines

1. INTRODUCTION

In the Philippines, 25% of the leading causes of deaths among young people (i.e., people aged 10 - 24 years) are results of their risky behaviors (Peltzer & Pengpid, 2016). Behaviors that led to their demise include substance use and poor mental well-being. For substance use, it was discovered that 23.5% of males and 8.2% of females smoke cigarettes. Suicidal ideation decreased among men from 2003 to 2011; however, it increased for women from 16.2% in 2003 to 21.8% in 2011. To stop the situation from becoming worse, studies revealed that health risk

behaviors must be detected and monitored (Centers for Disease Control and Prevention, 2012).

LCA has plenty of applications in different fields because it can cluster people based on their response patterns to categorical items (Dey, Chakraborty, Majumdar, & Mandal, 2016). Although there are existing researches concerning risk behaviors and latent class analysis, there is yet to be an investigation in the Philippines whereas the goal is to classify the respondents based on their patterns of dysfunctional activities.

To aid those concerned in reducing the mortality of young adults and adolescents near or inside the area from where the data was gathered, the study has determined the risk classes based on variables such as: alcohol use, cigarette use, marijuana use, number of sexual partners, depression, and suicidal ideation, using LCA of the SAS software. Additionally, the probability that respondents in a given class have engaged in a dysfunctional action, and the probability of membership to each labeled class were established. A better perspective on risk patterns can assist officials in creating policies directed at those who are potentially at-risk.

2. METHODOLOGY

The data used in this study was from a database of a larger study by Sta. Maria, Arcinas and Arcilla in 2014 regarding the risk and protective factors linked to the physical and emotional health of young Filipino adults and adolescents. A random sample of 1757 students in Region IV-A of the Philippines as approved by the De La Salle University Research Ethics Committee were chosen to participate in the survey.

Furthermore, the data consists of risk behaviors of Filipino young adults and adolescents among others collected using a standardized questionnaire. The risk behaviors used in the study that were previously mentioned are said to result in the greatest amount of morbidity, mortality and social problems among youth. The psychometric properties of scales used in the questionnaire were found to be reliable since resulting Cronbach's alpha were highly acceptable (P. Arcega, personal communication, December 10, 2018).

The chosen indicator variables were each transformed into binary variables as either 1 or 2, the former being risky and the latter being non-risky. For the *Alcohol* variable, individuals who had 5 or more alcoholic drinks in a day are considered to be at risk of alcohol-related problems (National Institute of Alcohol Abuse and Alcoholism, n.d.). For the *Cigarette* indicator, smoking less than 1 cigarette per day (CPD) is risky because evidence showed that compared to non-smokers, those who smoke even less than 1 CPD or 1 to 10 CPD have higher mortality risk (Inoue-Choi, Liao, Reyes-Guzman, Hartge, Caporaso, & Freedman, 2016). Similarly, for the *Marijuana* factor, a single consumption of marijuana is dangerous to the person since it is prohibited by the Republic Act No. 9165 (Comprehensive Dangerous Drug Act of 2002, 2002). For the *SexualPartners* variable, having 2 or more sexual partners for the past year puts a person at direct risk of having a sexually transmitted disease (Finer, Darroch & Singh, 1999). For the *Depression* indicator, the scale that was used for determining whether or not a person is depressed is the Zung Self-Rating Depression Scale (Zung, 1965). Participants were asked to answer 10 questions wherein they would have to rate how often they can relate to every situation, 1 being "a little of the time", 2 being "some of the time", 3 being "a good part of the time", and lastly 4 being "most of the time". The sum of the rating per item would indicate the severity of depression for each individual and a total number of points greater than or equal to 25 indicates depression (Zung, 1965). Lastly, for the *SuicidalIdeation* variable, a respondent is considered to be at risk when he/she has thought of committing suicide (American Foundation for Suicide Prevention, 2018).

Afterwards, LCA was the analysis used in the identification of distinct classes in accordance to the six indicator variables. The analysis can be generated through numerous software programs such as poLCA command using R Program (Linzer & Lewis, 2011), gsem command using STATA3 (Pitblado, 2017), Mplus (Li, 2017), Latent Gold (Vermunt & Magidson, 2013), PROC LCA command using SAS (Lanza et al., 2007), and many more. Both the poLCA and PROC LCA are easily accessible for the researchers, but they have chosen to proceed with the SAS command. This is because unlike the poLCA, PROC LCA has numerous papers/articles integrated from it, therefore such examples can be used as guides throughout the study.

3. RESULTS AND DISCUSSION

The six risk behavior variables mentioned were used in the estimation of latent class models. The model selection criteria are: G^2 statistic and its corresponding Degrees of Freedom (DF) and Likelihood Ratio Test (LRT) *p*-value, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Bootstrap Likelihood Ratio Test (BLRT) *p*-value. Through the analysis of these diagnostics and their results as shown in Table 1, the 4 Class Model is the chosen optimal model.

Table 1. Latent Class Model Diagnostics

Model	G^2 statistic	DF	LRT *p*-value	AIC	BIC	BLRT *p*-value
1 CLASS	621.98	57	0	633.98	666.81	.002
2 CLASS	157.73	50	<.001	183.73	254.85	.002
3 CLASS	126.21	43	<.001	166.21	275.62	.002
4 CLASS	43.93	36	.171	97.93	245.64	.344
5 CLASS	33.48	29	.259	101.48	287.48	.950
6 CLASS	30.55	22	.106	112.55	336.86	-

However, upon checking the Local Independence assumption for 4 Class Model through the Chi-Square Test for Independence and Fisher's Exact Test, it was deemed as violated. Due to this, models with less classes were considered but they are inadequate in taking into account all risk factors. Therefore, by aligning existing psychological theories about risk behaviors of the population, the 4 Class Model is still the chosen model, though prudence must be practiced in interpretations.

The item response conditional probability, or for the purpose of this study, the probability that a person from a specific latent class is considered to be at risk for a particular observed variable is illustrated in Table 2. One can say that members of Class 1 have high chances of participating in the risk behaviors indicated by variables: *Alcohol*, *Cigarette*, *Marijuana*, *SexualPartners*, and *Depression* with a low probability of considering suicide. Since almost all probabilities are greater than 50%, it can be considered as the High-Risk class.

Table 2. Item Response Conditional Probabilities for 4 Class Model

Indicator Variable	Class 1 (n=37)	Class 2 (n=1422)	Class 3 (n=120)	Class 4 (n=177)
Alcohol	**.6769***	.0810	.0489	**.7586**
Cigarette	**.9070**	.0158	.0517	.4855
Marijuana	**.8847**	.0007	.0001	.0001
Sexual Partners	**.5578**	.0374	.1139	.3579
Depression	**.5177**	.1601	**.6191**	.2550
Suicidal Ideation	.1301	.0016	**.6907**	.1708

*Item response >.50 in bold to facilitate interpretation

Class 2 has no item response conditional probability reaching 50% which implies that no risk behaviors showed high prevalence in this class (Table 2). The respondents classified into it are less likely to do activities that may endanger their health. Therefore, Class 2 is the Low-Risk class.

Class 3 showed high chances of respondent participation in *Depression* and *SuicidalIdeation* (Table 2). Both variables are considered as mental health issues that contribute to mortality (Pinto, Luna, Sivla Ade, Pinheiro, Braga, & Souza, 2014). Thus, Class 3 is the Poor Mental Health class.

Class 4 has a high probability of having risky alcohol consumption (Table 2). Since excessive alcohol drinking is a health risk behavior pertaining to substance use, this subgroup was labeled as the Legal Substance User.

Class membership probability, or the probability that a respondent is a member of a certain latent class was also identified for each class. Results showed that respondents are 72.84% likely to belong to the Low-Risk class. The Legal Substance User class has the second to the highest probability of 14.94% while High-Risk class has the least probability of 2.77%.

4. CONCLUSIONS

Based on the model selection criteria (i.e., BLRT *p-value,* G^2 statistic, AIC and BIC), the 4-class model was chosen because it adhered to the requirements of a good model fit and are aligned with existing theories about risk behaviors. Class 1 or the High-Risk class exemplified the respondents who were at-risk of all the factors except *SuicidalIdeation*. Class 2 comprised those who were most likely to be safe from the health-compromising effects of the risk behaviors hence, it was called Low-Risk subgroup. Class 3 or Poor Mental Health class, having high probabilities for *Depression* and *SuicidalIdeation* displayed conditions of individuals whose state of mind is unhealthy. Lastly, Class 4 showed a large percentage of riskiness with respect to alcohol thus, it was named Legal Substance User. Furthermore, a respondent has the highest probability of 72.84% of being categorized in the Low-Risk class, 14.94% chance of being included in the Legal Substance User class, 9.45% chance of belonging to the Poor Mental Health class and 2.77% probability of being a member of the High-Risk class.

5. ACKNOWLEDGMENTS

The researchers would like to express their immense gratitude to Dr. Madelene Sta. Maria and her team for providing the data for the analysis, and for providing theoretical assistance.

6. REFERENCES

American Foundation for Suicide Prevention. (2018). *Risk Factors and Warning Signs.* Retrieved from https://afsp.org/about-suicide/risk-factors-and-warning-signs/.

Centers for Disease Control and Prevention (2012). Youth Risk Behavior Surveillance - United States, 2011. *MMWR 61*(4), 1-162.

Comprehensive Dangerous Drug Act of 2002, R.A. No. 9165 (2002).

Dey, A., Chakraborty, A., Majumdar, K., & Mandal, A. (2016). Application of latent class analysis to estimate susceptibility to adverse health outcomes based on several risk factors. *International Journal of Community Medicine and Public Health 3*(12), 3423-3429.

Finer, L., Darroch, J. & Singh, S. (1999). Sexual Partnership Patterns as a Behavioral Risk Factor for Sexually Transmitted Diseases. *Family Planning Perspectives, 31*(5), 228. doi:10.2307/2991570.

Inoue-Choi, M., Liao, L., Reyes-Guzman, C., Hartge, P., Caporaso, N., & Freedman, N. (2016) Association of Long-term, Low-Intensity Smoking with All-Cause and Cause-Specific Mortality in the National Institutes of Health-AARP Diet and Health Study. *JAMA Internal Medicine.* *DOI:10.1001/jamainternmed.2016.7511.*

Lanza, S., Collins, L., Lemmon, D., & Schafer, J. (2007). PROC LCA: A SAS Procedure for Latent Class Analysis. *Structural Equation Modeling, 14*(4), 671–694.

Li, C. (2017). *Latent Class Analysis in Mplus.* Applied Psychometric Strategies Lab Applied Quantitative and Psychometric Series, University of Kentucky, Lexington, Kentucky.

Linzer, D. & Lewis, J. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software 42*(10), 1-29. URL http://www.jstatsoft.org/v42/i10/.

National Institute of Alcohol Abuse and Alcoholism (n.d.) *What's at-risk or heavy drinking? - Rethinking Drinking - NIAAA.* Retrieved from https://www.rethinkingdrinking.niaaa.nih.gov/How-much-is-too-much/Is-your-drinking-pattern-risky/Whats-At-Risk-Or-Heavy-Drinking.aspx.

Peltzer, K. & Pengpid, S. (2016). Health Risk Behavior among In-School Adolescents in the Philippines: Trends between 2003, 2007 and 2011, A Cross-Sectional Study. *International Journal of Environmental Research and Public Health 13*(2), 73. DOI:10.3390/ijerph13010073.

Pinto, A.C., Luna, I.T., Sivla Ade, A., Pinheiro, P.N., Braga, V.A., & Souza, A.M. (2014). Risk factors

associated with mental health issue in adolescents: an integrative review. *Rev Esc Enferm USP 48*(3), 552-561.

Pitblado, J. (2017). *A Journey to Latent Class Analysis (LCA)*. 2017 Nordic and Baltic Stata Users Group Meeting, Stockholm, Sweden.

Vermunt, J. K., & Magidson, J. (2013). *Latent class cluster analysis*. In: Hagenaars, J.A. and McCutcheon, A.L., Eds., Applied Latent Class Analysis, Cambridge University Press, Cambridge, pp.89-106.

Zung, W.W. (1965). A self-rating depression scale. *Archives of General Psychiatry 12*, 63-70.