Editorial

Mixture Modeling: A Useful Analytical Approach for Drug Use Studies

Wang J*

Children's National Medical Center, The George Washington University School of Medicine, USA

EDITORIAL

The analytic methods often used in drug use studies, such as ANOVA, multiple regression, logistic regression, multilevel models, and structural equation modeling (SEM) including path analysis, factor analysis, and latent growth curve model, are variable-centered approaches. Those approaches assume that the study sample arises from a homogeneous population; and focus on relations among variables, such as effects of independent variables on dependent variables. When typology of drug use or pattern of growth trajectories of drug use practices is of interest, person-centered analytical approach should be applied. In recent years, mixture modeling - a person-centered analytic approach - has increasingly gained its popularity in many research fields, including drug use studies. Assuming a heterogeneous population, mixture modeling aims to identify a finite number of subpopulations, called latent classes, that are unknown a priori within the population under study [1,2-4]. Individuals within a latent class share characteristics thus are more similar than those between latent classes. Mixture modeling can be readily applied to both cross-sectional and longitudinal studies.

Variable-centered and person-centered analytical approaches are often integrated in a more generalized analytical framework. As such, one is able to 1) identify unobserved homogeneous classes of individuals based on response patterns; 2) examine the features of heterogeneity across the latent classes; 3) evaluate the effects of covariates on the latent class membership; 4) assess the relationship between the latent class membership and distal outcomes; and 5) study transitions of the latent class memberships over time and determine factors that influence such transitions. Such an analytical framework enables researchers to better understand the properties of the target population [1,4].

A variety of mixture models have been developed [1,4], including, but not limited to, latent class analysis (LCA), growth mixture model (GMM), latent transition analysis (LTA), factor mixture model (FMM), and multilevel mixture model. In this editorial article, I give a brief introduction to LCA and GMM that are most often used mixture models in cross-sectional and longitudinal data analyses, respectively.

The objective of LCA is to identify unobserved classes/groups in a target population using cross-sectional data. Individuals are more homogeneous within class, but heterogeneous between classes. This is similar to the traditional cluster analysis. However, LCA is a model-based approach to clustering. It identifies latent classes based on posterior membership probabilities rather than somewhat ad hoc dissimilarity measures, such as Euclidean distance. In addition, LCA determines the optimal number of classes based on formal statistical procedures, and it provides more interpretable results stated in terms of probabilities. Parallels can also be drawn between LCA and factor analysis. They both are latent variable models where observed indicators/items are used to measure the unobserved latent constructs or factors. The primary difference is that the latent variables in factor analysis are continuous and individuals lie along a spectrum on the underlying factors; in contrast, the latent variable in LCA is categorical, and individuals have estimated probabilities of membership in each latent class. While factor analysis groups observed indicators/items, LCA groups individuals or cases based on their responses to the items.

GMM is an extension of the latent growth curve model (LGCM) [1] that is widely used in longitudinal studies. Assuming a homogeneous population, outcome growth trajectories estimated from LGCM vary randomly around the overall mean growth trajectory. In contrast, GMM classifies individuals into groups with distinct outcome growth trajectories. From an intuitive perspective, we may consider that GMM is implemented in two steps: first, individual growth trajectories are estimated from LGCM, and then individuals are clustered, based on their estimated growth trajectories, into a finite number of classes captured by a categorical latent variable. Growth trajectories are similar within class but different across classes. Covariates can be readily included into the model to predict the membership of latent trajectory class, and distal outcome can be specified as a function of both the covariates and the trajectory class membership. A special case of GMM is the group-based trajectory mode [2], also known as latent class growth analysis (LCGA) [1]. Like GMM, LCGA identifies distinct classes/groups of growth trajectories and classifies individuals into different classes/groups, but it assumes no trajectory variation within class. Despite this limitation, the group-based trajectory model is less

◎SciMed Central

complicated and can be readily implemented in the well-known statistics package SAS [2].

LCA has been widely used to study typology of drug use related phenomena in various drug using populations. For example, LCA was used to illustrate the typology of multi-drug use among MDMA users [5]; identify patterns of drug use practices among heroin and cocaine users [6]; classify youths/adolescents into homogenous groups based on their substance use [7]; identify distinct groups of cannabis users [8]; examine diagnostic classification for drug use disorders, such as drug abuse/dependence [9]. LCA has also been used to classify quality of life among opiate-dependent people [10], and to determine types of external barriers to substance abuse treatment [11].

GMM has been successfully applied to longitudinal studies on developmental trajectories of drug use practices. For example, GMM was used to identify distinct latent trajectory groups of cigarette smoking and alcohol use during emerging adulthood [12]; empirically demonstrate the chronic nature of heroin use level [13]. In addition, the special case of GMM - the group-based trajectory model or LCGA model or LCGA model - was applied to study growth trajectories of crack cocaine use [14,15].

In summary, the assumption of homogeneous population is often unrealistic. Ignoring potential population heterogeneity and focusing only on outcome overall mean could lead to misleading understanding and wrong conclusions. By modeling unobserved population heterogeneity, mixture modeling provides new insight in important areas of drug use studies, such as helping identify at-risk individuals and examining intervention impact on subgroups characterized by different drug use patterns or different types of growth trajectories of drug use practices. I hope that this editorial article will be helpful in inspiring further academic interest in applications of mixture models to drug use studies.

REFERENCES

1. Muthén B. Beyond SEM: General latent variable modeling. Behaviormetrika. 2002; 29: 81-117.

2. Nagin D S. Analyzing developmental trajectories: A semi-parametric, group-based approach, Psychological Methods. 1999; 4: 139-157.

3. Collins LM, Lanza ST. Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences. New York: Wiley; 2010.

4. Muthén L, Muthén B. Mplus User's Guide. LA, CA: Muthén L, Muthén; 1998-2012.

5. Carlson RG, Wang J, Falck RS, Siegal HA. Drug use practices among MDMA/ecstasy users in Ohio: a latent class analysis. Drug Alcohol Depend. 2005; 79: 167-179.

6. Harrell PT, Mancha BE, Petras H, Trenz RC, Latimer WW. Latent classes of heroin and cocaine users predict unique HIV/HCV risk factors. Drug Alcohol Depend. 2012; 122: 220-227.

7. Riehman KS, Stephens RL, Schurig ML. Substance use patterns and mental health diagnosis among youth in mental health treatment: a latent class analysis. J Psychoactive Drugs. 2009; 41: 363-368.

8. Fischer B, Rehm J, Irving H, Ialomiteanu A, Fallu JS, Patra J. Typologies of cannabis users and associated characteristics relevant for public health: a latent class analysis of data from a nationally representative Canadian adult survey. Int J Methods Psychiatr Res. 2010; 19: 110-124.

9. Wu LT, Woody GE, Yang C, Pan JJ, Blazer DG. Abuse and dependence on prescription opioids in adults: a mixture categorical and dimensional approach to diagnostic classification. Psychol Med. 2011; 41: 653-664.

10. De Maeyer J, van Nieuwenhuizen C, Bongers IL, Broekaert E, Vanderplasschen W. Profiles of quality of life in opiate-dependent individuals after starting methadone treatment: a latent class analysis. Int J Drug Policy. 2013; 24: 342-350.

11. Qi C, Kelly BC, Liu T, Liao Y, Hao W, Wang J. A latent class analysis of external barriers to drug treatment in China. J Subst Abuse Treat. 2013; 45: 350-355.

12. Huh J, Huang Z, Liao Y, Pentz M, Chou CP. Transitional life events and trajectories of cigarette and alcohol use during emerging adulthood: latent class analysis and growth mixture modeling. J Stud Alcohol Drugs. 2013; 74: 727-735.

13. Hser YI, Huang D, Chou CP, Anglin MD. Trajectories of heroin addiction: growth mixture modeling results based on a 33-year follow-up study. Eval Rev. 2007; 31: 548-563.

14. Falck RS, Wang J, Carlson RG. Crack cocaine trajectories among users in a midwestern American city. Addiction. 2007; 102: 1421-1431.

15. Wang J. Longitudinal study on growth trajectory of crack-cocaine use: Applications of group-based trajectory models. Advances and Applications in Statistical Sciences. 2010; 3: 285-310.