MEASUREMENT INVARIANCE TESTING OF LATENT CLASS MODELS USING RESIDUAL STATISTICS AND LIKELIHOOD RATIO TEST

Zsuzsa Bakk [1]

ABSTRACT: In latent class (LC) analysis a standard assumption is conditional independence, that is the indicators of the LC are independent of the covariates given the LCs. We compare the likelihood ratio based MIMIC test to residual statistics (BVR and $EPC_{interest}$) for identifying nonuniform direct effects (DEs) of covariates on the indicators of the LC model. The simulation study results show that the LR test and $EPC_{interest}$ correctly identifies direct effects more often than the BVR.

KEYWORDS: latent class analysis, measurement invariance, bivariate residuals, EPC, likelihood ratio test

1 Introduction

An often violated basic assumption of latent class modeling is the conditional independence assumption, also known as measurement equivalence. That is the association between the indicators of the LC model and the covariates are conditionally independent given the latent classes. Measurement equivalence can be tested by likelihood ratio based tests that compare the measurement equivalent model to models where direct effects (uniform or nonuniform) of covariates are allowed on the indicators of the LC model. An alternative approach for detecting missfit of the conditional independence model is to use residual statistics that can show violations of the conditional independence assumption. In this presentation we compare the power of the likelihood ratio based MIMIC model (Masyn, 2017) and that of two residual statistics ($\text{EPC}_{interest}$ and BVR) to detect the most complex type of measurement invariance, nonuniform direct effect. We first introduce the simple LC model, followed by a short presentation of the 3 approaches to detect missfit, compare them via a simulation experiment and conclude.

2 Latent class model

Consider the vector of responses $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iK})$, where Y_{ik} denotes the response of individual i on one of the K categorical indicator variables, with $1 \leq k \leq K$ and $1 \leq i \leq N$. Latent class (LC) analysis assumes that respondents belong to one of the T categories ("latent classes") of an underlying categorical latent variable X which affects the responses (Goodman, 1974). The measurement model for $\mathbf{Y}_i$ can then be written as:

$$p(\mathbf{Y}_i) = \sum_{t=1}^{T} p(X = t) \prod_{k=1}^{K} p(Y_{ik}|X = t). \quad (1)$$

The number of classes T is selected by comparing the goodness of fit of models with different values of T using model selection tools such as the AIC and BIC. Extending the model by a set of covariates $\mathbf{Z}$ affecting class membership leads to a model of a form:

[1] Methodology and Statistics, Leiden University, (e-mail: `z.bakk@fsw.leidenuniv.nl`)

$$p(\mathbf{Y}_i|Z) = \sum_{t=1}^{T} p(X=t|Z) \prod_{k=1}^{K} p(Y_{ik}|X=t).(2)$$

Usually the conditional class membership probabilities $P(X|Z)$ are parameterized using a multinomial logistic regression parametrization:

$$P(X=t|Z=z_i) = \frac{\exp(\alpha_t + \beta Z_i)}{1+\sum_{t=2}^{T}\exp(\alpha_t + \beta Z_i)}. \quad (3)$$

The model defined in Equation 2 assumes that given the LC variable X there is no direct relationship between Z and $\mathbf{Y}$ - a fairly common assumption in LV modeling known as measurement invariance. This assumption can be relaxed:

$$p(\mathbf{Y}_i|Z_i) = \sum_{t=1}^{T} p(X=t|Z_i) \prod_{k=1}^{K} p(Y_{ik}|X=t, Z_i).(4)$$

The most complex nonuniform DE can be parameterized as:

$$P(Y=y|Z=z_i) = \frac{\exp(\alpha_t + \beta \mathbf{X_t} + \beta Z_i|X)}{1+\sum_{t=2}^{T}\exp(\alpha_t + \beta \mathbf{X_t} + \beta Z_i|X)}. \quad (5)$$

The simpler uniform DE would mean dropping the class specific formulation of the effect of Z.

3 Identifying direct effects in LC models

3.1 Residual statistics

The BVR evaluates the residual association between each possible pair of observed variables (j, j') using a χ^2 test with 1 degree of freedom. The statistics can be formally defined as:

$$BVR_{jj'} = 1/P \sum_{j} \sum_{j'} \frac{(n_{jj'} - En_{jj'})^2}{n_{jj'}} \quad (6)$$

where the expected association $En_{jj'}$ for the covariate- indicator association is defined based on equation 2 in such a way that given the LC variable X there is no association between Z and Y. A downside of BVR is that the assumption of χ^2 distribution with 1 df does not hold (Oberski *et al.*, 2017).

.

Based on equations 5 we can see that the test of measurement invariance often takes the form of restricting a set of parameters to 0. In our case this refers to $\beta Z|X$. Let us consider a restriction on a vector of such logit coefficients as $\psi = \mathbf{0}$. In a general form the $EPC_{interest}$ can be formulated as:

$$EPC_{interest} = \mathbf{P}(\frac{\partial\theta}{\partial\psi'})(\psi - \psi') \quad (7)$$

where $\mathbf{P}$ is a matrix selecting the parameters of interest and θ is the vector of free model parameters. $EPC_{interest}$ can be seen as a linear approximation of the relationship between the free and fixed parameters of interest (Oberski *et al.*, 2017).

3.2 Likelihood ratio based stepwise multiple indicator multiple cause (MIMIC) modeling

The likelihood ratio based MIMIC approach (Masyn, 2017) is a multistage approach where nested models are compared with the goal to find the least restrictive well fitting model. The approach starts by comparing the latent class model with covariate (see Eq 2) to the model including all possible nonuniform DEs (see Eq 5). In case the LR test of the 2 nested models shows better fit of the all-DE model, the assumption of no DE is rejected, and a stepwise approach follows to identify the source of misfit. In the 2nd step an item by item testing of non uniform DE is performed, followed by an item by item testing of uniform DE for items for which a non uniform DE was confirmed in step 2. The approach has in total 7 possible steps, but we focus only on first 2 steps that focus on identifying nonuniform DE.

4 Simulation study

Table 1. *Percentage of correctly(T) and wrongly (F) identified DE with BVR, EPC and LR test separately for the low and high DE condition per latent class separation condition averaged over all sample sizes*

Class	High DE						Low DE					
sep	BVR_T	BVR_F	EPC_T	EPC_F	LR_T	LR_F	BVR_T	BVR_F	EPC_T	EPC_F	LR_T	LR_F
high	,18	,00	,98	,16	,97	,18	,00	,00	,41	,10	,63	,16
med	,22	,00	,86	,20	,83	,29	,00	,00	,44	,12	,60	,29
low	,03	,00	,41	,15	,52	,34	,00	,00	,37	,15	,42	,27

To test the ability of the 3 approaches to identify the presence of non uniform DE we run a simulation study with a LC model with 3 equal sized classes (class 1 low on all indicators, class3 high on all, class 2 low on first 3, high on last 3 indicators) measured by 6 indicators and regressed on a covariate. A full factorial design crossing sample size (250,500,1000,200), class separation ($Y|X$: .70,.80,.90), strength of DE (low, $\beta Z|X = .25$; high, .75) was used. DE was allowed on items 1 and 6.

When comparing the LR test for all nonuniform DE vs no DE model in all simulated conditions the more complex model was chosen, as such results are not detailed. The results in Table 1 show that the BVR is not a good statistic to identify a nonuniform DE, while the performance of $EPC_{interest}$ and LR test is better, their ability to identify a DE strongly depends not only on the strength of the DE, but also on the quality of the measurement model. With weaker measurement models all statistics fail to have a nominal rate close to the 95%.

5 Discussion

In a simulation experiment we compared $EPC_{interest}$, BVR and LR tests to identify a nonuniform DE. The results show that the $EPC_{interest}$ and LR test are more reliable, yet only in a few conditions meat the nominal 95% true-positive rate while maintaining a high false positive rate (between .16% to .29%). The BVR test was the most unreliable. We can conclude that nonuniform DE in most conditions is under identified by all estimators.

References

GOODMAN, LEO A. 1974. The Analysis of Systems of Qualitative Variables When Some of the Variables Are Unobservable. Part I: A Modified Latent Structure Approach. *American Journal of Sociology*, 79–259.

MASYN, K. E. 2017. Measurement Invariance and Differential Item Functioning in Latent Class Analysis With Stepwise Multiple Indicator Multiple Cause Modeling. *Structural Equation Modeling*, **24**(2), 180–197.

OBERSKI, DANIEL L., VERMUNT, JEROEN K., & MOORS, GUY B. D. 2017. Evaluating Measurement Invariance in Categorical Data Latent Variable Models with the EPC-Interest. *Political Analysis*, **23**(4), 550–563.