

Testing Procedure for Item Response Probabilities of 2-Class Latent Model

Bushra Shamshad^{1a}, Junaid Sagheer Siddiqui^{1b}

RECEIVED ON 19.03.2019, ACCEPTED ON 03.05.2019

ABSTRACT

This paper presents Hotelling T^2 as a procedure for the testing of significance difference between the item response probabilities (ω_{ij} 's) of classes in a Latent Class Model (LCM). Parametric bootstrap technique is used in order to generate samples for ω_{ij} 's. These samples are based on the estimated parameters of 2-class latent model. The estimation of parameters in either situation is done using the Expectation Maximization (EM) algorithm through Maximum likelihood method. The hypothesis under consideration is whether the response probabilities (ω_{ij} 's) are equal against each item in both the classes. $\{H_0: \omega_{i1} = \omega_{i2} \text{ against } H_1: \omega_{i1} \neq \omega_{i2}\}$. If the test exhibits significant difference between response probabilities in both classes, it will be a clear indication of a presence of latent variable. We consider both training and testing data sets to develop the test. In order to apply Hotelling T^2 test the basic assumptions of normality and homogeneity of variance are also checked. Chi-square goodness of fit test is used for assessing normal distribution to be good fitted on the hypothesized (bootstrap samples) based on 2-class latent model parameters for each data and Bartlett test to check heterogeneity of variances in ω_{ij} 's. Moreover, our procedure produces a minimum standard error of estimates as compared to those obtained through the package in R.Gui environment.

Keywords: Latent Class Model; Hotelling T^2 ; Maximum Likelihood; Bootstrapping; EM Algorithm; goodness of fit test; Bartlett Test.

1. INTRODUCTION

Latent variables are the hidden factors or some underlying concept which cannot be observed directly therefore other factors called manifest variables are needed. These manifest (observable) variables are then served as an indicator for latent variable. A well-known method for predicting latent variable is LCM. It identifies subgroups with the help of information provided by the manifest variables. These subgroups are formed over the assumption of local independence often called conditional independence. Conditional probabilities corresponding to each manifest variable are used to unfold which manifest variable plays important role in

the given class of latent variable. The application of LCM includes stratification research in social sciences [1], epidemiology [2] eating disorder phenotypes [3], Gaming disorder [4, 5] and education data [6-8].

More formally, consider a p-dimensional vector \mathbf{x} of manifest variables $(x_1, x_2, \dots, x_p)'$ and a single latent variable y having j-levels (say, $j = 1, 2, \dots, k$), in practice j-levels will be much smaller than p. As only manifest variables can be observed, the main aim of LCM is to get information about y after \mathbf{x} has been observed. Equation (1) represents the model

$$f(\mathbf{x}) = \int_{R_y} h(y) \prod_{i=1}^p g_i(x_i|y) dy \quad (1)$$

¹ Department of Statistics, University of Karachi, Karachi, Pakistan

Email: ^abshamshad@uok.edu.pk (Corresponding Author), ^bjssdr123@yahoo.com

where $g(x|y) = \prod_{i=1}^p g_i(x_i|y)$ represent the assumption of local independence, in which y -levels must be chosen so that y is sufficient to explain the dependency among the x 's. For the K classes, equation (1) can be written as follows:

$$\phi(x_{1h}, x_{2h}, \dots, x_{ph}|\theta) = \sum_{j=1}^k \theta_j \prod_{i=1}^p \omega_{ij}^{x_{ih}} (1 - \omega_{ij})^{1-x_{ih}} \quad (2)$$

The probability density function $\phi(\mathbf{x}|\theta)$ depends on $\theta' = (\theta_1, \theta_2, \dots, \theta_k, \omega_1, \omega_2, \dots, \omega_k)$. The within class parameters for the manifest variables $\omega'_j = (\omega_{1j}, \omega_{2j}, \dots, \omega_{pj})$ for $j = 1, 2, \dots, k$, where ω_{ij} is the probability of a positive response on variable i for a person in category j ($i = 1, 2, \dots, p; j = 1, 2, \dots, k$) and let θ_j be the prior probability that a randomly chosen individual belongs to class j (i.e. $\sum_{j=1}^k \theta_j = 1$). The parameters involved in equation (2) are all probabilities and thus range from zero to one ($0 < \theta_j < 1$ and $0 < \omega_{ij} < 1$).

Estimation of model parameters involve iterative algorithm (i.e. EM) Algorithm). We have used Maximum Likelihood Estimation (MLE) method. Usually model summary statistics in LCM are Log-likelihood ratio test statistic $(-2\log\lambda)$, χ^2 , G^2 and information criteria(s) (AIC, BIC etc.). The use of these statistics in making a decision, is subjective, as to how many number of classes will be in the model. Nylund *et al.* [9] showed in their paper that in comparison with other information criteria's the accuracy of AIC decreases with an increase of the sample size and in many cases overestimate the number of latent classes.

The Log-likelihood ratio test statistic $(-2\log\lambda)$ is also inappropriate in making such decision as regularity condition violates. χ^2 and G^2 statistics are simply a check of how close the expected frequencies are with the observed frequencies, whereas, the expected frequencies are obtained by amalgamating the relative expected frequencies of all the classes present in the model. In addition, the basic assumption of having expected frequency ≥ 5 may also be violated if manifest variables are greater in number and the number of observations is small. Thus, motivates to develop an alternative test through which we can

check the adequacy of the model. Anderlucci and Hennig [10] compared an alternative "distance-based clustering" method for finding the true ("correct") number of classes suggested by Hennig and Liao [11] to latent class analysis.

In this paper we have used Hotelling T^2 test for testing the differences between the item response probabilities (ω_{ij}) in a 2-class latent model, where ω'_{ij} s are the subsets of parameters within each group/class obtained by dividing a single population (mixture of two classes) on the basis of the basic assumption of local independence. There is a basic difference of purpose between the "usual Hotelling T^2 " statistics and "special approach we adopted to Hotelling T^2 ". In "usual Hotelling T^2 " the purpose is to check whether the means of one multivariate normal population is equal to the means of the other multivariate normal population. Whereas, "our special approach to Hotelling T^2 " involves bootstrapping the estimated parameters of a 2-class latent model to create a large number of surrogate samples. These bootstrap samples will serve the purpose of testing the item response conditional probabilities of the two classes of a latent variable to be significantly different from each other or not. In other words, our aim is to test whether the Item Response Probabilities (IRP) in class 1 are significantly different from those against the same item in class 2 of a single latent variable (i.e. $\omega_{i1} \neq \omega_{i2}$). Significantly different IRP's indicate the existence of latent variable while setting up an assumption of local independence. Thus, the IRP's can be "recognizable" via this testing procedure. That is, the sets of parameters can be clearly identified for the respective conditional posterior distribution in the log-likelihood function.

When testing two mean vectors (multivariate), Hotelling T^2 statistic can be used. The pooled variance covariance matrix is calculated similar to that obtained in univariate two sample mean testing. The Hotelling's T^2 statistic [12] for 2-sample is

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{y}^{(1)} - \bar{y}^{(2)}) \sim T^2(q, n_1 + n_2 - 2) \quad (3)$$

where $\bar{y}^{(1)}$ and $\bar{y}^{(2)}$ are averages of samples 1 and 2 of sizes n_1 and n_2 respectively in equation (3). S is the pool variance and T^2 statistic follows Hotelling T^2 distribution.

It is related to the central F distribution as follows:

$$\frac{n_1+n_2-q-1}{(n_1+n_2-2)q} T^2 \sim F(q, n_1 + n_2 - q - 1, \alpha) \quad (4)$$

As T^2 assumes that $(\bar{y}^{(1)} - \bar{y}^{(2)})$ follow multivariate normal. The data we used were investigated for normality and homogeneity of variance. We restrict our attention to 2-class model, further investigation for more than 2-class model is beyond the scope of this paper and can be considered as future work. We expect that the distribution of T^2 might follow Non central F distribution and Multivariate analysis of variance MANOVA may be used for a test of homogeneity of variance, in case of more than 2-class model.

1.1 Hotelling T^2 Test Approach for $(\omega_{ij}'s)$ IRP's (HAIRP):

We need repeated (estimated) values of the parameters $(\omega_{ij}'s)$ for which we want to carry the test described in this section. For this purpose, we initialized the parameters (we call them “Initial Estimated Parameters (IEP)”) obtained through applying 2-class latent model on given data. These IEP's are then used to simulate a hypothesized population and fitting 2-class latent model on it. The process is repeated several times (say “B” times) and the estimated model parameter of each hypothetical population is stored separately.

In case of 2-class model with 4 manifest items, there will be a total of 9 parameters

$$\theta' = (\theta_1, \omega_{11}, \omega_{21}, \omega_{31}, \omega_{41}, \omega_{12}, \omega_{22}, \omega_{32}, \omega_{42}).$$

During the computational process we will have to retain $\omega_1 = (\omega_{11}, \omega_{21}, \omega_{31}, \omega_{41})$ and $\omega_2 = (\omega_{12}, \omega_{22}, \omega_{32}, \omega_{42})$ in separate matrices each of order $B \times 4$.

The null hypothesis stated as “No class model” against the alternative hypothesis that “There exist at least two latent classes” or in other words;

H_0 Response probabilities in items in class1 are equal to those against the same item in the other class.

H_1 The two sets of item response probabilities

vectors in the respective classes are significantly different from each other.

1.1.1 Test Procedure (2-Class Latent Model; $j = 1, 2$)

Step1: $\omega_1 = \omega_2$ against $\omega_1 \neq \omega_2$; for four variables (items)

OR

$$H_0: \begin{pmatrix} \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} = \begin{pmatrix} \omega_{12} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix} \quad \text{vz.} \quad H_1: \begin{pmatrix} \omega_{11} \\ \omega_{21} \\ \omega_{31} \\ \omega_{41} \end{pmatrix} \neq \begin{pmatrix} \omega_{12} \\ \omega_{22} \\ \omega_{32} \\ \omega_{42} \end{pmatrix}$$

Step2: Level of significance = $\alpha = 0.05$

Step3: Test Statistics:

Since in our case sample 1 and 2 are both of size B

$$T^2 = \frac{B^2}{2B} (\bar{\omega}_1 - \bar{\omega}_2) S^{-1} (\bar{\omega}_1 - \bar{\omega}_2)$$

Step4: Critical Region:

$$\text{Reject } H_0 \text{ if } T^2 > \frac{(2B-2)}{2B-p-1} F_{(p, 2B-p-1, \alpha)}$$

We have created our own program codes in R.Gui for bootstrapping and estimating LCM (the codes are available from the authors upon request). Although we have written codes for the calculation of Hotelling T^2 test statistic too, but we used R package named “Hotelling” version 1.0 developed by Janes Curran available on the CRAN repository [13]. It includes functions to calculate a Hotelling T^2 test statistic and its p-value for the difference between two sets of (multivariate) normally distributed means. There is an option (command) in “hotelling.test()” function of whether to use a shrinkage method by “Shaefer and Strimmer’s James-Stein shrinkage estimator” to estimate the “large scale sample covariance matrices” or to use a simple pooled sample covariance matrix. By default the function uses simple pooled covariance matrix unless shrinkage is marked as TRUE in the function for the shrinkage estimates of sample covariance matrix. This package also includes “Aitchison’s additive log ratio” and “centered log ratio transformations” for compositional data [13]. We limit

use of this package only to “hotelling.test()” function with either shrinkage TRUE/FALSE.

2. BOOTSTRAPPING RESPONSE PROBABILITIES OF MANIFEST ITEM IN A SPECIFIC CLASS

We used two kinds of data sets. “Classical/training data” used several times for applying LCM by different authors and scientists. Since, the model had been fitted on classical data, IEP are thus available. It includes Coleman’s panel data [14] and Mastery data [15]. The other is “real/test” data, which has been collected or obtained, for which IEP’s are to be obtained through applying LCM. These IEP’s are then used as parameters to simulate the hypothesized population repeated number of times. On each hypothesized population, LCM is again applied to obtain the estimates of ω_{ij} ’s. The values of these estimates are then stored in a matrix for testing differences among ω_{i1} and ω_{i2} ; $i = 1, 2, 3, 4$.

Efron and Tibshirani [16] suggested that 50 to 100 replications may be sufficient for standard error and bias estimation, whereas, for highly accurate assessments or calculation of p-value, 350 replications may be considered a higher size [17, 18]. We simulate the bootstrap samples of sizes 30, 200, 500, 1500 and 2000. In the program (codes) initial values are set to zero. Therefore, the first storage space does not fill during the 1st iteration. The program command starts storing the estimate from 2nd iteration onward. The resultant matrix with one lesser value is therefore of size 29, 199, 499, 999, 1499 and 1999.

The estimates of the likelihood function should be based on the global maximum. Bartholomew suggested using different starting points to search global maxima or maxima close to global one [19]. Linzer and Lewis in their paper on “poLCA” (a package in “R” software) recommended using model fitting command several time to find the global maxima [20]. We took care of using the estimates based on global maxima. The reason is that, if the solution is from any one of the local maxima that may create certain problems in achieving the true “situation”. Such as, one, in booting samples the algorithm might produce “NaNs” or in other words,

some of the booting sample may become redundant and could not be included for further analysis. “NaN’s” were due to the boundary problem where the solution exaggerates when item probabilities are very much closer to 0 and 1. The problem can be solved by adding a constraint that if IRP are very close to either 0 or 1 to avoid this problem we add or subtract 0.00001 to it respectively. Since the value added is so small, it does not affect the result rather one can have control on the boundary issue. Secondly, there is also a possibility (we have experienced) of occurrence of outliers in the bootstrap samples which could be identified in a chi-square quantile plot and a separate normal probability plot of each ω_{ij} ’s (not shown here). Presence of outliers may also create a problem in goodness of fit test. Although, it is observed that for HAIRP (Special approach to Hotelling’s T^2 statistic) outliers have no influence on the results.

2.1 Problem in Handling Item Response Probabilities

Initially LCA starts with an unconditional contingency table for assigning response profile to each class/category without any guidance and thus completes the table. As a result, class prevalence (θ_j) and item response probabilities are obtained for each of the latent class.

In the method of LCA during estimation recognition of “Class Identification Number (CIN)” is flexible. That is, due to no clear guidance to identify CIN, the algorithm assigned one set of estimates (class proportion and item response probabilities) to class 1, in the next iteration, are then assigned to class 2. If there are B iterations, (B/2) half of the CIN are assign to one specific class (say class 1) and the other half to class 2 in a random fashion.

Due to no proper recognition of CIN, the sampling distribution of IRP’s and θ appeared bimodal. Thompson used the term “fractured” for such bimodal sampling distribution. He obtained bootstrap samples for class response probabilities and class proportions through EM algorithm using log-linear model [21]. Linzer and Lewis also discussed the issue of CIN and suggest solution while using poLCA package [22].

2.2 Solution to the Problem of Handling Item Response Probabilities

In order to resolve such bimodal sampling distribution, that is, the inability of EM algorithm to consistently recognizing the latent classes, we propose the following method.

After, generating bootstrap samples “B” times through EM algorithm using Latent Class Model in case of 2-class model ($j = 1, 2$) with 4 items (manifest variables; $i = 1, 2, 3, 4$), we used the following steps for recognition of proper CIN (“fracturing” problem).

Step1: For any one item, calculate the difference in response probabilities for class 1 and 2. For example, we consider the difference in between the response probability of 2-classes for item 1. {i.e.; $D = (\omega_{11} - \omega_{12})$ }.

Step2: Create a coding variable **C** on the basis of “**D**”. Where for every positive difference C_j takes on a value “1” and “2” for the negative differences {if $D > 0, C = 1$ else $C = 2$ }.

Step3: Create another vector of the absolute difference of $3 - C$. {let $C2 = |3 - C|$ }

Step4: Stack ω_{i1} and ω_{i2} {i.e. create a new matrix **T** of ω_{ij} 's of dimension $B \times 4$ }.

Step5: Stack **C** and **C2** and name it “**G**”.

Step6: Unstack **T** on the basis of grouping variable **G**. The strategy is applied to each vector of the estimate. After such adjustments, bootstrap solutions are consistent with the initial estimates. That is, solution

will then consistently classify item response probabilities to the group they belong and the bimodal sampling distribution resolve to unimodal with means and variances accordingly. Thompson gave the solution of fracturing of estimates by identifying the most informative indicator for which the conditional probabilities given latent class membership differ most [21]. The command lines for such direction of indicators are also provided in his paper.

3. DATASETS

The first data used is “Mastery data” [15]. Bartholomew in his book discussed it and suggested a 2-class latent model. The two classes were named as “Master” and “Non-master” [19]. Originally 142 individuals were asked to take a test based on four problems (items) randomly selected from the area of problems on the multiplication of two-digit number by a 3 or 4 digit number, which involve use of carry operations. The response was marked as ‘1’ and ‘0’ for correct and incorrect solutions, respectively. The estimated 2-class latent model parameters (considered as IEP) are presented in Table 1. The classes are marked as “Master” and “Non-Master” based on item response probabilities in each class.

The second classical data set is Coleman’s panel data [14]. This data is also used by Goodman [23, 24] and Bartholomew [19]. The data was collected from 3398 boys at two different points in time. Each time the

TABLE 1. INITIALLY ESTIMATED PARAMETERS (IEP) FOR A 2-CLASS LATENT MODEL OF MASTERY, COLEMAN’S PANEL AND KUTS DATA.					
Mastery Data: 2-class latent model parameters					
	Class prevalence	Item response probabilities			
		Item 1	Item 2	Item 3	Item 4
Class 1 (Master)	0.59	0.75	0.78	0.43	0.71
Class 2 (Non-master)	0.41	0.21	0.07	0.02	0.05
Coleman’s panel data: 2-class latent model parameters					
	Class prevalence	Item response probabilities			
Class 1 (Optimistic attitude)	0.4	0.77	0.64	0.89	0.67
Class 2 (Pessimistic attitude)	0.6	0.1	0.47	0.09	0.5
KUTS Panel Data: 2-class model parameters					
	Class prevalence	Item response probabilities			
		President	Vice President	Secretary	Treasurer
Class 1 (Rightist)	0.532	0.917	0.864	0.962	0.901
Class 2 (Mix)	0.468	0.037	0.043	0.112	0.132

individuals were asked two questions that are (1) “Whether or not they considered themselves to be in the leading crowd” (1 for a positive response and 0 for a negative response). (2) “Whether they thought that such membership involved can sometimes go against their principles” (1 for a negative response and 0 for a positive response). A restricted model with an assumption that there exist two latent variables, and that they altogether form four latent classes is the solution for the data [19]. On the basis of their response probabilities obtained through estimating 2-class model, we named the classes as being “optimistic” and “pessimistic” attitude groups. The LCM parameters (considered as IEP) are given in Table 1.

Real data which we have considered here is of election of Karachi University Teacher Society (KUTS) in year 1993-1994. A total of 434 teachers participated in 1993-94 KUTS election. We considered only the top four positions (President, Vice president, Secretary and Treasurer) in the analysis. The responses were coded as 1 = “Rightist” panel and 2 = “Mix” panel. A complete discussion regarding LCM on KUTS data is available in [25]. Table 1 shows 2-class latent model parameters for the data.

4. DISCUSSION AND RESULTS

In Mastery data (shown in Table 1), it should be noted that the two class sizes (estimated through LCM) are a ratio of 60:40 and the response probabilities of correctly answering an item for class 2 (non-master class) are too close to the lower boundary of the parameter space (i.e.; 0). Similar scenario is in the case of KUTS panel data that the estimated LC model parameters are too close to the boundaries (as high with maximum probability 0.96 and low with 0.03).

In case of Coleman’s Panel data, the estimated model parameters are not too close to the boundary of the parameter space. Although, ω_{12} and ω_{32} seem close to zero but they are not as low as less than 0.09 (shown in Table 1).

For each data considered and for each size, Hotelling T^2 reveals that there exist significant differences between the parameters of the two classes in a 2-class model, as the Hotelling T^2 test rejected the null hypothesis (i.e. $H_0: \omega_{i1} = \omega_{i2}$; for $i = 1, 2, \dots, 4$) with “0.000” p-value (Since the test statistics value(s)

are very high and results in a p-value too close to 0). Therefore, it can be concluded on the basis of Hotelling T^2 test, that latent variable exist in Mastery and Coleman’s panel data (shown in Table 2).

Bartlett test (Table 2, 3) shows that variances are heterogeneous in all bootstrap samples of different sizes for all the data set considered. It is observed that, where ω_{ij} ’s are not too close to the boundary of the parametric space, the chance of normal distribution to be well fitted is pretty high (shown in Table 4). It is also observed while bootstrapping ω_{ij} ’s that the normal distribution is good fitted in cases where the size of the bootstrap sample is relatively lower than 1000 (for $0 < \omega_{ij}$ ’s < 1 and particularly when ω_{ij} ’s are close to the boundary).

For each data considered and for each size, Hotelling T^2 reveals that there exist significant differences between the parameters of the two classes in a 2-class model, as the Hotelling T^2 test rejected the null hypothesis (i.e. $H_0: \omega_{i1} = \omega_{i2}$; for $i = 1, 2, \dots, 4$) with “0.000” p-value (Since the test statistics value(s) are very high and results in a p-value too close to 0). Therefore, it can be concluded on the basis of Hotelling T^2 test, that latent variable exist in Mastery and Coleman’s panel data (shown in Table 2).

In case of KUTS panel data we consider positive response when voters are in favor of the Rightist group. The result shows that there is a significant difference between ω_{ij} ’s of “Rightist” and “Mix” group. HAIRP reveals that there exist differences between the parameters of the two classes, since the test rejected the null hypothesis (i.e. $H_0: \omega_{i1} = \omega_{i2}$) with “0.000” p-value, no matter whether the covariance matrices are homogenous or heterogeneous (shown in Table 3). Therefore, it can be concluded on the basis of HAIRP that there may exist a latent variable in KUTS panel data.

4.1 Comparison of Standard Error of Estimates through Bayes LCA and HAIRP:

“BayesLCA” is a package in R.Gui environment, available at CRAN repository since 2012. The program was developed by Arthur White and Thomas Brendan Murphy. Along with model fitting it also provides with the bootstrap sampling for IRP’s and calculate standard errors. We have considered “BayesLCA” to compare the shape of the densities of

ω_{ij} 's with those obtained through HAIRP (a total of 6 (sizes (B)) \times 3 (data sets) = 18 cases). Both approaches approximately showed same distributional behavior for each ω_{ij} {i = 1, 2, 3, 4 and j = 1, 2}. However, estimates obtained through HAIRP give more accurate results than "BayesLCA". Comparison between HAIRP and "BayesLCA" is presented in Table 5 for all three data sets. For each size (B) the estimates of ω_{ij} 's are presented. The average estimates (mean) for each ω_{ij} obtained through HAIRP are much closer to the population parameters (IEP's) as compared to MAP obtained through "BayesLCA" (for each data). In most cases estimated model parameters obtained through "BayesLCA" are under or over estimated. Moreover, standard errors for ω_{ij} 's for all sizes (B) are

also very small than those obtained through package "BayesLCA". "BayesLCA" provides comparatively smaller standard errors of estimates for large samples ($B \geq 500$) as compared to smaller ones ($B < 500$).

5. CONCLUSION

In this paper we attempt to exhibit a test which can identify the difference between ω_{ij} 's of both classes of a single latent variable. We have demonstrated the problems faced, by selecting those data sets, in which the model parameters are on the boundary of the parameter space. Also, grouping formed in both classical data sets are almost of 60:40 ratio and 53:47 in KUTS panel data. Regardless of a little difference

TABLE 3: KUTS PANEL DATA: A SUMMARY TABLE OF BOOTSTRAP SAMPLES (OF SIZE "B").

				Hotelling's T ² test					
Bartlett Test				(Shrinkage =FALSE)		degrees of freedom		(Shrinkage =TRUE)	
B	K ² statistics	df	p-value	T ² statistics	P-value	v ₁	v ₂	T ² statistics	P-value
29	21.3651	7	0.0032	18633	0.000	4	53	21587	0.000
199	92.2596	7	2.20E-16	172920	0.000	4	393	178490	0.000
499	352.8211 *	7	2.20E-16	384530	0.000	4	993	397510	0.000
999	740.7012	7	2.20E-16	799770	0.000	4	1993	835420	0.000
1499	1026.375	7	2.20E-16	1186300	0.000	4	2993	1223000	0.000
1999	1343.526	7	2.20E-16	1562300	0.000	4	3993	1593200	0.000

NOTE: T^{2*} = Hotelling's T² (Shrinkage = FALSE) with (v₁, v₂) degrees of freedom; p-value is same for both T² and Wald statistics.

TABLE 4. A SUMMARY TABLE OF CHI-SQUARE TEST FOR ASSESSING NORMAL DISTRIBUTION TO BE GOOD FITTED ON THE HYPOTHESIZED (BOOTSTRAP SAMPLES) BASED ON IRP'S OF A 2-CLASS LM FOR COLEMAN'S PANEL, MASTERY AND KUTS PANEL DATA.

Goodness of fit (Normal Distribution)													
B		Coleman's panel data				Mastery data				KUTS panel data			
		IEP	χ^2 statistics	df	p-value	IEP	χ^2 statistic s	df	p-value	IEP	χ^2 statistic s	df	p-value
29	ω_{11}	0.77	2.633	8	0.955	0.75	11.91	7	0.10347	0.917	10.01	9	0.3493
	ω_{21}	0.64	2.339	9	0.985	0.78	7.456	7	0.38295	0.864	4.213	9	0.8968
	ω_{31}	0.89	2.359	8	0.968	0.43	6.736	6	0.34599	0.962	6.325	10	0.7872
	ω_{41}	0.67	3.63	10	0.963	0.71	8.689	7	0.27578	0.901	5.395	9	0.7986
	ω_{12}	0.1	6.226	10	0.796	0.21	3.522	7	0.83292	0.037	2.626	10	0.9889
	ω_{22}	0.47	8.541	12	0.742	0.07	8.053	7	0.3279	0.043	5.943	10	0.8201
199	ω_{32}	0.09	1.426	9	0.998	0.02	21.94	9	0.00906	0.112	6.541	8	0.5869
	ω_{42}	0.5	1.853	10	0.997	0.05	10.48	8	0.23266	0.132	4.986	8	0.7591
	ω_{11}	0.77	11.77	10	0.301	0.75	2.186	8	0.97478	0.917	7.731	13	0.8607
	ω_{21}	0.64	7.83	14	0.898	0.78	4.504	8	0.80901	0.864	17.19	11	0.1024
	ω_{31}	0.89	16.77	10	0.08	0.43	9.908	8	0.27154	0.962	9.451	13	0.7381

Testing Procedure for Item Response Probabilities of 2-Class Latent Model

	ω_{41}	0.67	6.896	13	0.907	0.71	10.94	8	0.20523	0.901	11.98	12	0.4472
	ω_{12}	0.1	6.864	13	0.909	0.21	2.187	8	0.97472	0.037	3.597	14	0.9974
	ω_{22}	0.47	6.233	14	0.96	0.07	37.53	8	9.20E-06	0.043	155.2	13	2.00E-26
	ω_{32}	0.09	6.602	11	0.83	0.02	102.2	8	1.50E-18	0.112	7.314	12	0.8362
	ω_{42}	0.5	7.664	15	0.937	0.05	74.75	8	5.54E-13	0.132	12.08	11	0.3576
499	ω_{11}	0.77	7.012	12	0.857	0.75	6.505	11	0.8376	0.917	20.55	13	0.0823
	ω_{21}	0.64	19.28	15	0.201	0.78	13.45	11	0.26511	0.864	5.8	13	0.9531
	ω_{31}	0.89	10.61	12	0.562	0.43	56.26	11	4.60E-08	0.962	26.04	15	0.0376
	ω_{41}	0.67	10.64	15	0.777	0.71	38.96	10	2.60E-05	0.901	17.17	14	0.2471
	ω_{12}	0.1	6.47	13	0.927	0.21	8.415	11	0.67574	0.037	62.67	14	4.00E-08
	ω_{22}	0.47	44.11	15	1.00E-04	0.07	72.74	11	3.70E-11	0.043	26.66	15	0.0316
	ω_{32}	0.09	13.99	13	0.375	0.02	597.6	13	2.00E-119	0.112	7.896	13	0.8503
	ω_{42}	0.5	19.84	17	0.282	0.05	226.1	12	1.30E-41	0.132	1.473	13	0.9999
999	ω_{11}	0.77	126.1	13	1.00E-20	0.75	13.03	12	0.36695	0.917	88.13	14	9.00E-13
	ω_{21}	0.64	14.41	17	0.638	0.78	79.15	12	6.00E-12	0.864	17.1	14	0.2507
	ω_{31}	0.89	13.65	13	0.399	0.43	826.5	12	4.00E-169	0.962	19.34	17	0.3094
	ω_{41}	0.67	9.354	16	0.898	0.71	371.8	11	5.70E-73	0.901	24.72	15	0.0539
	ω_{12}	0.1	9.347	15	0.859	0.21	1125	11	3.00E-234	0.037	37.73	17	0.0027
	ω_{22}	0.47	21.28	17	0.214	0.07	3.00E+11	11	0.000	0.043	29.17	16	0.0228
	ω_{32}	0.09	34.85	14	0.002	0.02	504.7	6	8.00E-106	0.112	11.42	15	0.7222
	ω_{42}	0.5	13.42	17	0.708	0.05	266	6	5.60E-49	0.132	38.74	14	0.0004
1499	ω_{11}	0.77	4.514	14	0.991	0.75	31.63	12	0.00158	0.917	18.77	15	0.2243
	ω_{21}	0.64	10.46	17	0.883	0.78	174.6	12	5.30E-31	0.864	11.23	15	0.7363
	ω_{31}	0.89	17.91	14	0.211	0.43	436.6	12	6.50E-86	0.962	27.14	17	0.056
	ω_{41}	0.67	13.84	17	0.678	0.71	6.265	13	0.93596	0.901	20.47	15	0.1547
	ω_{12}	0.1	20.73	15	0.146	0.21	415.1	12	2.30E-81	0.037	38.9	17	0.0018
	ω_{22}	0.47	10.65	18	0.908	0.07	1.00E+08	12	0.000	0.043	67.72	17	5.00E-08
	ω_{32}	0.09	28.12	15	0.021	0.02	676.9	13	0.000	0.112	14.21	15	0.5097
	ω_{42}	0.5	11.79	19	0.894	0.05	1.00E+12	12	0.000	0.132	22.54	15	0.0944
1999	ω_{11}	0.77	12.13	14	0.596	0.75	153.2	13	4.70E-26	0.917	28.91	16	0.0245
	ω_{21}	0.64	11.26	17	0.843	0.78	9265	13	0.000	0.864	8.708	16	0.925
	ω_{31}	0.89	22.88	14	0.062	0.43	56.92	13	1.90E-07	0.962	30.19	18	0.0356
	ω_{41}	0.67	16.35	16	0.429	0.71	406.7	13	8.50E-79	0.901	25.99	16	0.0541
	ω_{12}	0.1	34.14	16	0.005	0.21	2278	13	0.000	0.037	80.05	18	8.00E-10
	ω_{22}	0.47	5.099	18	0.999	0.07	3.00E+10	13	0.000	0.043	66.52	17	8.00E-08
	ω_{32}	0.09	33.54	15	0.004	0.02	7.00E+09	14	0.000	0.112	8.547	15	0.9
	ω_{42}	0.5	4.386	19	0.999	0.05	1109	13	0.000	0.132	32.07	15	0.0063

Testing Procedure for Item Response Probabilities of 2-Class Latent Model

TABLE 5. A COMPARISON OF THE RESULTS OBTAINED THROUGH HAIRP WITH BAYESLCA FOR COLEMAN'S PANEL, MASTERY AND KUTS PANEL DATA (MEAN EST IMATES (MEAN), STANDARD ERRORS (S.E.) THROUGH HAIRP) AND (MAXIMUM A POSTERIORI (MAP) AND STANDARD ERROR (S.E.) THROUGH BAYESLCA)

B		Coleman's panel data					Mastery data					KUTS panel data				
		EP	HAIRP		BayesLCA		IEP	HAIRP		BayesLCA		IEP	HAIRP		BayesLCA	
			Mean	S.E.	AP	S.E.		mean	S.E.	MAP	S.E.		mean	S.E.	MAP	S.E.
29	ω_{11}	0.77	0.77	0.00	0.94	0.13	0.75	0.77	0.00	0.80	0.16	0.29	0.92	0.00	0.83	0.12
	ω_{21}	0.64	0.64	0.00	0.46	0.19	0.78	0.79	0.00	0.82	0.14	0.86	0.87	0.00	0.89	0.10
	ω_{31}	0.89	0.89	0.00	0.73	0.13	0.43	0.45	0.01	0.51	0.18	0.96	0.96	0.00	1.00	0.00
	ω_{41}	0.67	0.67	0.00	0.89	0.09	0.71	0.71	0.00	0.61	0.17	0.90	0.90	0.00	1.00	0.00
	ω_{21}	0.10	0.10	0.00	0.00	0.00	0.21	0.21	0.00	0.31	0.20	0.04	0.04	0.00	0.07	0.07
	ω_{22}	0.47	0.47	0.00	0.52	0.17	0.07	0.09	0.00	0.10	0.14	0.04	0.04	0.00	0.00	0.00
	ω_{23}	0.09	0.09	0.00	0.24	0.13	0.02	0.02	0.00	0.00	0.02	0.11	0.11	0.00	0.09	0.09
199	ω_{42}	0.50	0.50	0.00	0.54	0.05	0.05	0.05	0.00	0.05	0.05	0.13	0.13	0.00	0.13	0.03
	ω_{11}	0.77	0.77	0.00	0.99	0.02	0.75	0.75	0.00	0.74	0.05	0.92	0.92	0.00	0.95	0.03
	ω_{21}	0.64	0.64	0.00	0.62	0.06	0.78	0.79	0.00	0.67	0.05	0.86	0.86	0.00	0.87	0.04
	ω_{31}	0.89	0.89	0.00	0.93	0.07	0.43	0.44	0.00	0.34	0.05	0.96	0.96	0.00	0.96	0.03
	ω_{41}	0.67	0.67	0.00	0.64	0.06	0.71	0.71	0.00	0.61	0.05	0.90	0.90	0.00	0.89	0.04
	ω_{21}	0.10	0.10	0.00	0.09	0.05	0.21	0.20	0.00	0.00	0.00	0.04	0.04	0.00	0.03	0.02
	ω_{22}	0.47	0.47	0.00	0.48	0.05	0.07	0.07	0.00	0.07	0.05	0.04	0.04	0.00	0.07	0.03
499	ω_{32}	0.09	0.09	0.00	0.11	0.03	0.02	0.02	0.00	0.02	0.02	0.11	0.11	0.00	0.16	0.04
	ω_{42}	0.50	0.50	0.00	0.54	0.05	0.05	0.05	0.00	0.05	0.05	0.13	0.13	0.00	0.13	0.03
	ω_{11}	0.77	0.77	0.00	0.93	0.06	0.75	0.76	0.00	0.71	0.03	0.92	0.92	0.00	0.91	0.02
	ω_{21}	0.64	0.64	0.00	0.66	0.04	0.78	0.78	0.00	0.79	0.04	0.86	0.86	0.00	0.84	0.03
	ω_{31}	0.89	0.89	0.00	0.92	0.06	0.43	0.43	0.00	0.41	0.04	0.96	0.96	0.00	0.97	0.01
	ω_{41}	0.67	0.67	0.00	0.72	0.05	0.71	0.72	0.00	0.64	0.04	0.90	0.90	0.00	0.90	0.02
	ω_{12}	0.10	0.10	0.00	0.12	0.04	0.21	0.21	0.00	0.26	0.03	0.04	0.04	0.00	0.04	0.02
999	ω_{22}	0.47	0.47	0.00	0.45	0.03	0.07	0.07	0.00	0.04	0.03	0.04	0.04	0.00	0.04	0.01
	ω_{32}	0.09	0.09	0.00	0.17	0.03	0.02	0.02	0.00	0.01	0.01	0.11	0.11	0.00	0.12	0.03
	ω_{42}	0.50	0.50	0.00	0.50	0.03	0.05	0.05	0.00	0.08	0.03	0.13	0.13	0.00	0.08	0.02
	ω_{11}	0.77	0.77	0.00	0.85	0.05	0.75	0.76	0.00	0.75	0.02	0.92	0.92	0.00	0.94	0.01
	ω_{21}	0.64	0.64	0.00	0.62	0.03	0.78	0.79	0.00	0.78	0.02	0.86	0.86	0.00	0.87	0.02
	ω_{31}	0.89	0.89	0.00	0.91	0.05	0.43	0.43	0.00	0.42	0.02	0.96	0.96	0.00	0.96	0.01
	ω_{41}	0.67	0.67	0.00	0.66	0.03	0.71	0.72	0.00	0.70	0.03	0.90	0.90	0.00	0.90	0.02
1499	ω_{12}	0.10	0.10	0.00	0.09	0.04	0.21	0.21	0.00	0.18	0.02	0.04	0.04	0.00	0.03	0.01
	ω_{22}	0.47	0.47	0.00	0.47	0.02	0.07	0.08	0.00	0.04	0.02	0.02	0.04	0.00	0.06	0.01
	ω_{32}	0.09	0.09	0.00	0.12	0.03	0.02	0.03	0.00	0.02	0.01	0.11	0.11	0.00	0.12	0.02
	ω_{42}	0.50	0.50	0.00	0.51	0.02	0.05	0.05	0.00	0.05	0.02	0.13	0.13	0.00	0.12	0.02
	ω_{11}	0.77	0.77	0.00	0.85	0.05	0.75	0.76	0.00	0.78	0.02	0.92	0.92	0.00	0.92	0.01
	ω_{21}	0.64	0.64	0.00	0.68	0.02	0.78	0.79	0.00	0.77	0.02	0.86	0.86	0.00	0.87	0.01
	ω_{31}	0.89	0.89	0.00	0.90	0.04	0.43	0.44	0.00	0.46	0.02	0.96	0.96	0.00	0.96	0.01
1999	ω_{41}	0.67	0.67	0.00	0.70	0.02	0.71	0.72	0.00	0.70	0.02	0.90	0.90	0.00	0.90	0.01
	ω_{12}	0.10	0.10	0.00	0.11	0.02	0.21	0.21	0.00	0.19	0.02	0.04	0.04	0.00	0.02	0.01
	ω_{22}	0.47	0.47	0.00	0.49	0.02	0.07	0.07	0.00	0.08	0.02	0.04	0.04	0.00	0.05	0.01
	ω_{32}	0.09	0.09	0.00	0.15	0.03	0.02	0.02	0.00	0.04	0.01	0.11	0.11	0.00	0.12	0.01
	ω_{42}	0.50	0.50	0.00	0.49	0.02	0.05	0.06	0.00	0.08	0.02	0.13	0.13	0.00	0.10	0.01
	ω_{11}	0.77	0.77	0.00	0.79	0.03	0.75	0.76	0.00	0.74	0.02	0.92	0.92	0.00	0.90	0.01
	ω_{21}	0.64	0.64	0.00	0.63	0.02	0.78	0.79	0.00	0.77	0.02	0.86	0.86	0.00	0.85	0.01
1999	ω_{31}	0.89	0.89	0.00	0.83	0.03	0.43	0.44	0.00	0.43	0.02	0.96	0.96	0.00	0.96	0.01
	ω_{41}	0.67	0.67	0.00	0.68	0.02	0.71	0.72	0.00	0.70	0.02	0.90	0.90	0.00	0.90	0.01
	ω_{12}	0.10	0.10	0.00	0.05	0.02	0.21	0.21	0.00	0.23	0.02	0.04	0.04	0.00	0.03	0.01
	ω_{22}	0.47	0.47	0.00	0.46	0.02	0.07	0.08	0.00	0.07	0.02	0.04	0.04	0.00	0.04	0.01
	ω_{32}	0.09	0.09	0.00	0.09	0.03	0.02	0.02	0.00	0.03	0.01	0.11	0.11	0.00	0.11	0.01
ω_{42}	0.50	0.50	0.00	0.51	0.02	0.05	0.06	0.00	0.04	0.01	0.13	0.13	0.00	0.12	0.01	

between the class proportions of two classes, item response probabilities of class 1 are significantly different from those of the second class (i.e. $\omega_{i1} \neq \omega_{i2}$), in each case considered. HAIRP proposed, as

demonstrated earlier, leads us to accept the existence of latent variable forming groups or clusters and its existence indicates significant difference between the classes, while setting up local independence

assumption. Moreover, our adopted bootstrapping procedure provides better estimates along with their minimal standard errors as compared to the one obtained through package “BayesLCA”.

ACKNOWLEDGMENT:

Authors are indebted to two referees for valuable comments and suggestions.

REFERENCES:

- [1] Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., Le, R. B., Friedman, S., Miles, A. “A New Model of Social Class? Findings from the BBC’s Great British Class Survey Experiment”, *Sociology*, Vol. 47, No. 2, pp. 219–250.
- [2] Collins, L. M., and Lanza, S. T., *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*, John Wiley and Sons Inc, 2010.
- [3] Keel, P. K., Fichter, M., Quadflieg, N., Bulik, C. M., Baxter, M. G., Thornton, L., ... Kaye, W. H, Application of a Latent Class Analysis to Empirically Define Eating Disorder Phenotypes. *Archives of General Psychiatry*, Vol. 61, No. 2, pp. 192-200, 2004.
- [4] Myrseth, H., Notelaers, G., “A Latent Class Approach for Classifying the Problem and Disordered Gamers in a Group of Adolescence”, *Frontiers in Psychology*, Vol. 9, pp. 2273, 2018.
- [5] Carras, M. C., and Kardefelt-Winther, D, “When addiction symptoms and life problems diverge: a latent class analysis of problematic gaming in a representative multinational sample of European adolescents”, *Eur .Child Adoles. Psychiatry*, Vol. 27, pp. 513–525, 2018.
- [6] Aitkin. M. D, Anderson et al., "Statistical modelling of data on teaching styles." *Journal of the Royal Statistical Society, Series A (General)*, Vol. 144, No. 4, pp. 419-461, 1981.
- [7] J.R. Bergan, J. R., “Latent-Class models in educational research”. In E. W. Gordon (ed.), *Review of Research in Education* 10, American Educational Research Association, Washington, DC, 1983.
- [8] Dayton, C. M., “Educational applications of latent class analysis”. *Measurement and Evaluation in Counseling and Development*, Vol. 24, pp. 131-141, 1991.
- [9] Nylund, K. I., Asparouhov, T., and Muthén. B., “Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study”. *Structural Equation Modeling*, Vol. 14, No. 4, pp. 535–569, 2007.
- [10] Anderlucci, L., and Hennig, C., “The clustering of categorical data: a comparison of a model-based and a distance-based approach”, *Communication in Statistics—Theory and Methods*, volume. 43, pp. 704-721, 2014.
- [11] Hennig, C., and Liao, T. F., “Comparing latent class and dissimilarity based clustering for mixed type variables with application to social stratification (with discussion)”. *Journal of the Royal Statistical Society, Series C*, vol. 62, pp. 309-369, 2013.
- [12] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, New York: Wiley, 1984.
- [13] Curran, J., “Hotelling: Hotelling's T-squared Test and Variants”. R package version 1.0-0, 2011. Available: <http://CRAN.R-project.org/package=Hotelling>.
- [14] Coleman, J. S., *Introduction to Mathematical Society*. New York: Free Press, 1964.
- [15] Macready, G. B., and Dayton, C. M., “The use of probabilistic models in the assessment of mastery”. *Journal of Educational Statistics*, Vol. 2, pp. 99-120, 1977.
- [16] Efron, B., and Tibshirani, R., “The bootstrap method for standard errors, confidence intervals, and other measures of statistical accuracy”. *Statistical Science*, Vol. 1 (1), pp. 1-35, 1986.
- [17] McLachlan, G. J., “On bootstrapping the likelihood ratio test statistic for the number of

- components in a normal mixture”, Applied Statistics, vol. 36, pp. 318-324, 1987.
- [18] McLachlan, G. J., and Krishnan, T. The EM algorithm and Extensions. Hoboken, N.J: Wiley-Interscience, 2008.
- [19] Bartholomew, D. J., Latent Variable Models and Factor Analysis. Charles Griffin & Co. Ltd: London, 1987.
- [20] Linzer D.A and Lewis J. B. poLCA: Polytomous Variable Latent Class Analysis. R package version 1.3, 2007-2011. Available:<http://CRAN.R-project.org/package=poLCA>.
- [21] Thompson, D. M., “Latent class analysis in SAS: Promise, problems, and programming”. Invited paper at Statistics and Data Analysis Section, SAS Global Forum 2007. Retrieved February 10, 2009 from SAS, Inc. Web site: <http://www2.sas.com/proceedings/forum2007/192-2007.pdf>.
- [22] Linzer, D. A., and Lewis, J. B., “poLCA: An R Package for Polytomous Variable Latent Class Analysis”. Journal of Statistical Software, vol. 42, No. 10, pp. 1-29, 2011. Available: www.jstatsoft.org/v42/i10/.
- [23] Goodman, L. A., “Exploratory latent structure analysis using both identifiable and unidentifiable models”. Biometrika, Vol. 61(2), pp. 215-231, 1974a.
- [24] Goodman, L. A., “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, Vol. 79. No. 5, pp. 1179-1259, 1974b.
- [25] Shamshad, B., and Siddiqi, J. S., “Exploration of groups through latent structural model”. Journal of Basic and Applied Science, Vol. 8, No. 1, pp. 145-150, 2012.