Latent class representation of the Grade of Membership model

By ELENA A. EROSHEVA
Technical Report No. 492
Department of Statistics, University of Washington, Box 354322
Seattle, WA 98195-4322, U. S. A.

SUMMARY

Latent class and the Grade of Membership (GoM) models are two examples of latent structure models. Latent class models are discrete mixture models. The GoM model has been originally developed as an extension of latent class models to a continuous mixture. This note describes a constrained latent class model which is equivalent to the GoM model, and provides a detailed proof of this equivalence. Implications for model fitting and interpretation are discussed.

Some key words: Class membership, Contingency tables, Latent structure, Mixture models, Stochastic subject.

1 Introduction

Let $x = (x_1, x_2, \ldots, x_J)$ be a vector of polytomous manifest (observable) variables, where x_j takes on values $l_j \in \mathcal{L}_j = \{1, 2, \ldots, L_j\}$, $j = 1, 2, \ldots, J$, and L_j denotes the number of possible outcomes. Data of this structure recorded on a number of individuals are common in the social, behavioral, and health sciences. They are often analyzed via latent variable models which assume that individual responses differ according to some unobservable value. Formulating a latent variable model involves specifying a distribution of latent variables and a conditional distribution of manifest variables given latent variables; the former is usually chosen according to some plausible assumptions.

Latent class and the GoM models are two examples of latent structure models. Latent class models rely on a discrete distribution of latent variables over a number of categories, reflecting the assumption that individuals are full members of one of the latent classes (Lazarsfeld & Henry, 1968; Goodman, 1974). The GoM model assumes a continuous distribution of latent variables over a number of categories which reflects the original idea that individuals can be partial members in more than one class (Woodbury et al., 1978). Other examples of continuous mixture models are latent trait models, with the Rasch model being one of the most widely used (Rasch, 1960; Fischer & Molenaar, 1995).

Most intuitively, the difference between latent class and the GoM models should be similar to that between latent class and latent trait models in general. As described by

Bartholomew & Knott (1999), latent class models can be thought of as special cases of latent trait models in which latent variable distribution is constrained to discrete probability masses. Thus, comparing the GoM and latent class models, Manton et al. (1994b) concluded: "latent class model is nested in the GoM model structure...", but "...if we allow latent class model to have more classes, then it is potentially possible to "fit" the realized data set as well as with GoM" (p. 45). On the other hand, in his review of Manton et al. (1994b), Haberman (1995) suggested that the GoM model is a special case of latent class models since a set of constraints imposed upon a latent class model can specify a distribution of manifest variables which is identical to that provided by the GoM model.

This work explains in detail how the GoM model could be thought of as a generalization and as a special case of latent class models at the same time. Section 2 develops a common notation which illuminates the GoM model as a generalization of latent class models. Following Haberman (1995), section 3 derives a special case of a latent class model with constraints which is shown to be equivalent to the GoM model, therefore providing a latent class representation. Section 4 discusses implications for the GoM model interpretation. Finally, the discussion in Section 5 situates this work in the context of available literature.

2 The GoM Model as a Generalization of Latent Class Models

This section formally introduces latent class and the GoM models using a common notation. We first define distributions of latent variables and conditional distributions of manifest variables, and then derive marginal distributions of manifest variables.

Let $x = (x_1, x_2, \dots, x_J)$ be a vector of polytomous manifest variables taking on values $l_j \in \mathcal{L}_j = \{1, 2, \dots, L_j\}$, $j = 1, 2, \dots, J$. Then $\mathcal{X} = \prod_{j=1}^{J} \mathcal{L}_j$ is the set of all possible outcomes for vector x.

2·1 Latent Class Models

Let $y = (y_1, y_2, \dots, y_K)$ be a latent membership vector defined by

$$y_k = \begin{cases} 1, & \text{if member of class } k,\ k = 1, 2, \dots, K, \\ 0, & \text{otherwise,} \end{cases}$$

with probability density function

$$f(y) = \begin{cases} \pi_k, & \text{if } y_k = 1 \text{ and } y_l = 0,\ \ l \neq k,\ k = 1, 2, \dots, K \\ 0, & \text{otherwise.} \end{cases}$$

Denote the conditional probability of manifest variable x_j taking on value l_j, given membership in kth class, by

$$\lambda_{kjl_j} = \mathrm{pr}(x_j = l_j | y_k = 1),\ k = 1, 2, \dots, K,\ j = 1, 2, \dots, J,\ l_j = 1, 2, \dots, L_j \quad (2{\cdot}1)$$

Since full membership vector y has exactly one nonzero component, the conditional probability for x_j, given the membership vector, can be rewritten as

$$\mathrm{pr}(x_j = l_j|y) \quad = \quad \lambda_{kjl_j} = \sum_{k=1}^{K} y_k \lambda_{kjl_j}.$$

The set of conditional probabilities must satisfy the following constraints:

$$\sum_{l_j \in \mathcal{L}_j} \lambda_{kjl_j} = 1, \quad k = 1, 2, \ldots, K;\ j = 1, 2, \ldots, J.$$

Employing the local independence assumption (Lazarsfeld & Henry, 1968), latent class models assume manifest variables are conditionally independent, given latent variables. Thus, conditional probability of observing response pattern l, given full membership vector y, is

$$f^{LCM}(l|y) = \mathrm{pr}(x = l|y) \quad = \quad \prod_{j=1}^{J} \left(\sum_{k=1}^{K} y_k \lambda_{kjl_j} \right),\ l \in \mathcal{X}.$$

Integrating out latent variable y, we obtain the marginal distribution of the manifest variables under the latent class model in the form of a discrete mixture:

$$f^{LCM}(l) = \mathrm{pr}(x = l) \quad = \quad \int f^{LCM}(l|y) f(y) dy = \sum_{k=1}^{K} \pi_k \prod_{j=1}^{J} \lambda_{kjl_j},\ l \in \mathcal{X}. \qquad (2{\cdot}2)$$

Equation (2·2) states that the probability of observing response pattern l is the sum of the probabilities of observing l from each of the latent classes, weighted by their relative sizes, π_k.

2·2 The GoM Model

Let $g = (g_1, g_2, \ldots, g_K)$ be a latent partial membership vector of K nonnegative random variables that sum to 1. Now K is the number of extreme profiles in the GoM model. Analogous to equation (2·1), denote extreme profile response probabilities by

$$\lambda_{kjl_j} \quad = \quad \mathrm{pr}(x_j = l_j|g_k = 1),\ k = 1, 2, \ldots, K;\ j = 1, 2, \ldots, J;\ l_j = 1, 2, \ldots, L_j \qquad (2{\cdot}3)$$

Parameters λ_{kjl_j} are conditional probabilities of responses in the extreme cases when one of the membership scores equals 1.

The main assumption of the GoM model is the convexity of conditional response probabilities. Given partial membership vector $g \in [0, 1]^K$, the conditional distribution of manifest variable x_j is given by a convex combination of the extreme profiles' response probabilities, i.e.,

$$\begin{aligned} \mathrm{pr}(x_j = l_j|g) \quad &= \quad \sum_{k=1}^{K} g_k \cdot \mathrm{pr}(x_j = l_j|g_k = 1) \\ &= \quad \sum_{k=1}^{K} g_k \lambda_{kjl_j}, \quad j = 1, 2, \ldots, J,\ l_j = 1, 2, \ldots, L_j. \qquad (2{\cdot}4) \end{aligned}$$

Applying the local independence assumption, we obtain the conditional probability of observing response pattern l as

$$f^{GoM}(l|g) = \text{pr}(x = l|g) = \prod_{j=1}^{J} \text{pr}(x_j = l_j|g) = \prod_{j=1}^{J} \left(\sum_{k=1}^{K} g_k \lambda_{kjl_j} \right), \ l \in \mathcal{X}.$$

Let us denote the distribution of g by $D(g)$. Integrating out latent variable g, we obtain the marginal distribution for response pattern l in the form of a continuous mixture

$$f^{GoM}(l) = \text{pr}(x = l) = \int f^{GoM}(l|g) dD(g) = \int \prod_{j=1}^{J} \left(\sum_{k=1}^{K} g_k \lambda_{kjl_j} \right) dD(g), \ l \in \mathcal{X}. \quad (2{\cdot}5)$$

Note that l_j appears as part of the index of the conditional probability in equations (2·1) and (2·3). Also note that since partial membership vector can have multiple non-zero entries, the integral in equation (2·5) does not simplify to a summation as does the integral in equation (2·2).

Under the common notation developed in this section, latent membership vector y is a constrained version of latent membership vector g. Thus, the K-class latent class model is a special case of the K-profile GoM model with constraints placed on the distribution of the membership vector.

3 Latent class representation of the GoM model

Let us now relax the requirement of equality of the number of latent classes and extreme profiles. Following Haberman (1995), we construct a latent class model such that the marginal distribution of manifest variables is exactly the same as that under the GoM model.

Assume integer K denotes the number of extreme profiles and J denotes the number of manifest variables as before. Consider a vector of J polytomous latent variables $z = (z_1, z_2, \ldots, z_J)$, each taking on values from the set of integers $\{1, 2, \ldots, K\}$. Vector z can be thought of as defining latent classes. Denote by $\mathcal{Z} = \{1, 2, \ldots, K\}^J$ the set of all possible vectors z. Then $\mathcal{X} \times \mathcal{Z}$ is the index set for the cross-classification of the manifest variables x and latent classification variables z.

As apparent from equations (2·2) and (2·5), to obtain a latent class representation of the GoM model one must find a way to interchange the summation and the product operator in equation (2·5). The following lemma provides algebra which allows us to do so. Lemma 3.1 will also by instrumental in the development of the latent class model in this section.

Lemma 3.1 *For any two integers J and K, and for any two sets of real numbers $\{a_k, \ k = 1, 2, \ldots, K\}$ and $\{b_{kj}, \ k = 1, 2, \ldots, K, \ j = 1, 2, \ldots, J\}$,*

$$\prod_{j=1}^{J} \sum_{k=1}^{K} a_k b_{kj} \quad = \quad \sum_{z \in \mathcal{Z}} \prod_{j=1}^{J} a_{z_j} b_{z_j j}, \quad (3{\cdot}6)$$

where $z = (z_1, z_2, \ldots, z_J)$ is such that $z \in \mathcal{Z} = \prod_{j=1}^{J} \{1, 2, \ldots, K\}$.

Proof The left hand side of equation (3.1) is

$$(a_1b_{11} + a_2b_{21} + \ldots + a_Kb_{K1})(a_1b_{12} + a_2b_{22} + \ldots + a_Kb_{K2}) \ldots \\ \ldots (a_1b_{1J} + a_2b_{2J} + \ldots + a_Kb_{KJ}). \quad (3{\cdot}7)$$

Multiplying these J sums out, we get a summation in which each term has J multipliers of a's and corresponding, according to the k-index, J multipliers of b's. Each product of a's, as well as each product of b's, can therefore be indexed by a vector $z \in \mathcal{Z}$, where z_j would index the jth multiplier:

$$\prod_{j=1}^{J} \sum_{k=1}^{K} a_k b_{kj} \quad = \quad \sum_{z \in \mathcal{Z}} \prod_{j=1}^{J} a_{z_j} b_{z_j j}.$$

Thus, the order of the product and the summation can be interchanged by changing the space over which the summation is performed and by substituting z_j-indices instead of k-indices. ∎

To continue developing the latent class model, let us define a distribution over latent classes $z \in \mathcal{Z}$ conditional on the distribution of membership vector $g \in [0,1]^K$ as described in the following lemma.

Lemma 3.2 *If a K-dimensional vector of random variables $(g_1, g_2, \ldots, g_K)$ has a joint distribution $D(g)$ on $[0,1]^K$, such that $g_1 + g_2 + \ldots + g_K = 1$, then*

$$\pi_z \quad = \quad E_D \left(\prod_{j=1}^{J} g_{z_j} \right) \quad (3{\cdot}8)$$

is a probability measure on $\mathcal{Z}$.

Proof Values of π_z are nonnegative because they are defined by expected values of products of nonnegative random variables. By using properties of expectation and applying Lemma 3.1 with $a_k = g_k$ and $b_{kj} = 1$, for all k, for all j, one can show that the sum of π_z over $\mathcal{Z}$ equals 1. Since for all $z \in \mathcal{Z}$, $\pi_z \geq 0$ and $\sum_{z \in \mathcal{Z}} \pi_z = 1$, π_z is a probability measure on $\mathcal{Z}$. ∎

From the functional form of probabilities π_z, it follows that latent classification variables $z_1, z_2, \ldots, z_J$ are exchangeable.

To specify the conditional distribution of the manifest variables given the latent variables, we need two additional assumptions. First, assume that x_j is independent of z_a for $a \neq j$, given z_j. That is,

$$\begin{aligned} \mathrm{pr}\,(x_j = l_j | z) \quad &= \quad \mathrm{pr}\,(x_j = l_j | z_1, z_2, \ldots, z_J) \\ &= \quad \mathrm{pr}\,(x_j = l_j | z_j)\,, \end{aligned} \quad (3{\cdot}9)$$

where $z_j \in \{1, 2, \ldots, K\}$ is the value of the latent classification variable, and $l_j \in \mathcal{L}_j$ is the observed value of manifest variable x_j. In essence, equation (3·9) postulates that manifest variable x_j is directly influenced only by the jth component of the latent classification vector z.

Second, assume that conditional probabilities in equation (3·9) are given by

$$\mathrm{pr}\,(x_j = l_j | z_j) \;=\; \lambda_{z_j j l_j},\; z_j \in \{1, \ldots, K\};\; j = 1, 2, \ldots, J;\; l_j = 1, 2, \ldots, L_j \quad (3\cdot10)$$

where the set of λs is the same as the set of the extreme profile probabilities for the GoM model. These structural parameters must also satisfy the constraints:

$$\sum_{l_j \in \mathcal{L}_j} \lambda_{z_j j l_j} = 1, \quad \text{for all } z \in \mathcal{Z}, j \in \{1, 2, \ldots, J\}.$$

Assuming further that manifest variables are conditionally independent given latent classification variables, we obtain the probability of observing response pattern l for the latent class model proposed by Haberman (HLCM) (1995) as

$$\begin{aligned} f^{HLCM}(l) &= \mathrm{pr}\,(x_1 = l_1, x_2 = l_2, \ldots, x_J = l_J) \\ &= \sum_{z \in \mathcal{Z}} \{\mathrm{pr}(Z = z)\mathrm{pr}\,(x_1 = l_1, x_2 = l_2, \ldots, x_J = l_J | z)\} \\ &= \sum_{z \in \mathcal{Z}} \left\{ \pi_z \left(\prod_{j=1}^{J} \mathrm{pr}(x_j = l_j | z_j) \right) \right\} \\ &= \sum_{z \in \mathcal{Z}} \left\{ E_D \left(\prod_{j=1}^{J} g_{z_j} \right) \left(\prod_{j=1}^{J} \lambda_{z_j j l_j} \right) \right\},\; l \in \mathcal{X}. \end{aligned} \quad (3\cdot11)$$

The probability of observing response pattern l in equation (3·11) is the sum of the conditional probabilities of observing l from each of the latent classes, weighted by the latent class probabilities. The probability of latent class z is the expected value of a J-fold product of the membership scores.

Next lemma proves the equivalence between the marginal probabilities of the observed response patterns for the GoM and HLCM models.

Lemma 3.3 $f^{GoM}(l) = f^{HLCM}(l) \;\; l \in \mathcal{X}$.

Proof Consider the marginal probability of an arbitrary response pattern $l \in \mathcal{X}$ for the GoM model provided by equation (2·5). Applying lemma 3.1 with $a_k = g_k,\; b_{kj} = \lambda_{kjl_j}$, and using properties of expectation, we obtain

$$f^{GoM}(l) \;=\; \sum_{z \in \mathcal{Z}} \left\{ E_D \left(\prod_{j=1}^{J} g_{z_j} \right) \left(\prod_{j=1}^{J} \lambda_{z_j j l_j} \right) \right\}.$$

Hence, we have

$$f^{GoM}(l) \;=\; f^{HLCM}(l) \;\text{ for all } l \in \mathcal{X}. \quad (3\cdot12)$$

The marginal probability distribution placed by the GoM model on manifest variables coincides with the marginal distribution placed by the latent class model with constraints. ∎

Table 1: *Low-dimensional example: Extreme profile probabilities for the GoM model.*

item j	λ_{1j}	λ_{2j}
item 1	0.08	0.77
item 2	0.14	0.96
item 3	0.03	0.90

It follows that the GoM model can be reformulated as a latent class model with a distribution on the latent classes given by a functional form of the distribution of membership scores.

The details of the machinery are easy to see from a simple example of the GoM model with two extreme profiles. Suppose that three dichotomous items have extreme profile probabilities as in Table 1. Given membership scores $g \sim D(g)$, the latent class representation of the GoM model has $2^3 = 8$ latent classes determined by latent classification vector z. Table 2 provides the latent class, as indicated by the values of z, and the corresponding conditional response probabilities. The first latent class has the conditional response probabilities from the first extreme profile for all items. The second latent class has the conditional response probabilities for items 1 and 2 from the first extreme profile, and from the second extreme profile for item 3. Going through all the permutations in this fashion, we obtain the eight latent classes, where response probabilities for the last class coincide with those of the second extreme profile.

Table 2: *Low-dimensional example: Latent class representation of the GoM model. Latent class and conditional response probabilities.*

latent class	z	item 1	item 2	item 3	π_z
1	(1,1,1)	0.08	0.14	0.03	$E_D(g_1 g_1 g_1)$
2	(1,1,2)	0.08	0.14	0.90	$E_D(g_1 g_1 g_2)$
3	(1,2,1)	0.08	0.96	0.03	$E_D(g_1 g_2 g_1)$
4	(1,2,2)	0.08	0.96	0.90	$E_D(g_1 g_2 g_2)$
5	(2,1,1)	0.77	0.14	0.03	$E_D(g_2 g_1 g_1)$
6	(2,1,2)	0.77	0.14	0.90	$E_D(g_2 g_1 g_2)$
7	(2,2,1)	0.77	0.96	0.03	$E_D(g_2 g_2 g_1)$
8	(2,2,2)	0.77	0.96	0.90	$E_D(g_2 g_2 g_2)$

4 Implications for interpretation

The standard GoM interpretation states that membership score g_k "represents the degree to which the element ... belongs to the kth set" (see Manton et al., 1994b, pg.

3). For example, when a two-profile GoM model is fitted to discrete responses on J questions from a disability survey, two estimated extreme profiles might be interpreted as 'healthy' and 'disabled'. Then, membership scores would show how healthy the subject is, relative to the 'disabled' and 'healthy' profiles. Although intuitively appealing to many medical and social science researchers, this interpretation might be confusing to statistical audiences. The latent class representation of the GoM model provides a way to operationalize this interpretation.

To proceed, let us first recall two general rationales for interpreting latent structure models described by Holland (1990), the random sampling and the stochastic subject. The former rationale assumes that latent quantities define a probability of a correct response among subjects with that value of the latent variable. Under this rationale, latent parameters are employed in order to obtain legitimate values for probabilities of observable response patterns in a population. Thus, following the random sampling rationale, interpretations of the GoM model and of the latent class representation are the same since they place identical probability structures on observed responses.

In contrast, the stochastic subject rationale assumes that human behavior is inherently random, and a latent quantity determines response probabilities of the subject with that value of the latent variable. Adopting this rationale, under the standard formulation of the GoM model, each of J marginal response probabilities for a subject is given by a linear combination of the extreme profiles' response probabilities weighted by the subject-specific GoM scores. The probability of observing a J-dimensional response pattern is a J-fold product of these linear combinations. Each individual's responses on the manifest variables are being generated by a multinomial process with fixed subject-specific probabilities. The fact that these probabilities are provided by convex combinations motivates the standard partial membership interpretation for the GoM membership scores.

Given J questions and K extreme profiles, the total number of latent classes in the latent class representation of the GoM model is K^J. Each subject is considered to be a complete member in one of these latent classes. To apply the stochastic subject rationale to the latent class representation of the GoM model, recall that a distribution on the latent classes is determined through membership scores. On an individual level, subject's membership score g_k determines the expected proportion of questions that the subject answers as if he was a full member of the kth extreme profile. Then, specific combinations of questions answered from each profile correspond to particular latent classes in the latent class representation of the GoM model. Notice the apparent conditional exchangeability of the manifest variables in this interpretation. Taking the health survey as an example, a subject with the membership score $g_2 = 1/3$ would answer on average a third of the survey questions as a person from the 'disabled' profile, and two thirds as a person from the 'healthy' profile. Specific combinations of survey questions that come from 'healthy' and 'disabled' profiles define the latent classes.

5 Discussion

The GoM model was developed in the late 1970s. GoM applications now cover a wide spectrum of studies, ranging from studying depression (Davidson et al., 1989) and schizophrenia (Manton et al., 1994a) to analyzing complex genotype-phenotype relations (Manton et al., 2004). However, the model remains to be relatively unfamiliar to statistical audiences. Despite a multitude of published large-scale GoM applications, there is a lack of statistical publications that explore basic GoM properties and provide simple examples. The articles by Potthoff et al. (2000), Wachter (1999), Erosheva (2005) are the only exceptions, as far as I can ascertain. The goals of this work are twofold: first, to provide details on a close relationship between more familiar latent class models and the GoM model, and, second, to point out implications of this relationship for statistical audiences as well as for applied researchers in other fields.

This work explains how the GoM model could be thought of as a generalization and as a special case of latent class models at the same time. By formulating latent class and the GoM models within a common notational framework, we clearly illustrate the GoM model as a generalization of latent class models when the number of latent classes is restricted to be the same as the number of extreme profiles. On the other hand, when the equality requirement on the number of classes and profiles is relaxed, we show the GoM model as a special case of latent class models using the key algebraic equality which allows us to interchange a summation and a product operators.

Although the GoM model as a generalization of latent class models was mentioned frequently in various publications of Woodbury, Manton and colleagues, the details of that relationship have not been formalized. The relationship in the other direction was suggested by Haberman (1995) in a brief review but deserves a more in depth explanation.

Note that Manton and colleagues provided a different form of an algebraic equality that is similar to equation (3·12); see Tolley & Manton (1992, pg. 91) or Manton et al. (1994b, pg. 53), for example. By using this equality the authors correctly concluded that the marginal probability of observed responses under the GoM model depends on the order-J moments of the membership scores, however, they did not make a step further to consider the equivalence of the GoM and latent class models. More recent publications also suggest that equivalence between the GoM and latent class models is still a subject of a debate (e.g., see Manton et al., 2004, pg. 396).

Models that allow for specification of continuous latent constructs are increasingly popular among researchers in social, behavioral, and health sciences since many latent variables of interest can be thought of as having fine gradations. When substantive theory justifies distinct latent categories as well as continuous latent variables, approaches that describe heterogeneity of individuals with respect to those discrete categories often focus on class membership probabilities. To give a few examples, Foody et al. (1992) emphasize the utility of posterior probabilities of class membership in the area of remote sensing; Muthen & Shedden (1999) model the class membership probability as a function of covariates in a study of alcohol dependence; Roeder et al. (1999) address a similar issue by modelling uncertainty in latent class assignments in a criminology case study. The GoM model also addresses the issue of uncertainty in class membership but by a different approach, the one which incorporates degrees of membership as incidental parameters in the model. Understanding the partial membership structure of the GoM model is essential when comparing GoM analytic capabilities with those of alternative approaches.

Discrete approximations to continuous distributions can be useful for purely computational reasons. For example, Luceno (1999) approximates continuous univariate random variables by discrete random variables sharing several low order moments. The approximations help the author to avoid costly computation involved in using simulation studies. Standard methods of estimating the GoM model such as those in Manton et al. (1994b) do not rely on the GoM representation as a discrete mixture model and have questionable properties (Haberman, 1995). Under Bayesian approach, Erosheva (2003) developed a GoM estimation algorithm which is based on the structure provided by the latent class representation. The Bayesian estimation approach has several advantages over likelihood-based estimation procedures for the GoM model (Erosheva, 2002, 2003).

Understanding the duality of the GoM model makes it easier to establish direct connections with models from other areas. For example, although a clustering model with admixture developed for genetic data (Pritchard et al., 2000) and the standard GoM model seem quite different, deep similarities between these two independently developed models become obvious when one considers the latent class representation. Moreover, the latent class representation of the GoM model allows us to develop a general class of mixed membership models. The generalization is flexible enough to accommodate models for other data structures such as data that come from text documents (Erosheva et al., 2004). See Erosheva (2002) for a description of general class of mixed membership models and more on connections between the GoM and other closely related models.

The GoM model is not unique in its dual representation as a discrete and a continuous mixture model. Lindsay et al. (1991) provide special cases which exhibit the same property in the case of conditional estimation for the Rasch model. The key quality of the Rasch model which allows for such dual representation in some special cases is the form of sufficient statistics. For the discrete data GoM model, on the other hand, the functional form of the conditional response probabilities is solely responsible for the existence of the latent class representation.

Finally, it is worth emphasizing one more time that the developed latent class representation of the GoM model places identical probability structure on observable variables and hence can not possibly be distinguished from the continuous mixture GoM model on the basis of data.

Acknowledgement

The author is grateful to Professors Stephen E. Fienberg and Brian Junker, Carnegie Mellon, for inspiring and helpful discussions on this topic. This work was partly supported by a grant from National Institutes of Health.

References

Bartholomew, D. J. & Knott, M. (1999). *Latent Variable Models and Factor Analysis.* Arnold.

Davidson, J., Woodbury, M. A., Zisook, S. & Giller, E. (1989). Classification of depression by grade of membership: A confirmation study. *Psychological Medicine* **19**, 987–998.

Erosheva, E. A. (2002). *Grade of Membership and Latent Structure Models with Application to Disability Survey Data*. Ph.D. thesis, Carnegie Mellon University.

Erosheva, E. A. (2003). *Bayesian estimation of the Grade of Membership model*, chap. In *Bayesian Statistics 7*, Ed. J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, pp. 501-510. Oxford University Press.

Erosheva, E. A. (2005). Comparing latent structures of the Grade of Membership, Rasch and latent class models. *To appear in Psychometrika* .

Erosheva, E. A., Fienberg, S. E. & Lafferty, J. (2004). Mixed-membership models of scientific publications. *Proceedings of the National Academy of Sciences* **101** (Suppl.1), 5220–5227.

Fischer, G. H. & Molenaar, I. W. (1995). *Rasch Models: Foundations, Recent Developments, and Applications*. New York: Springer-Verlag.

Foody, G. M., Campbell, N. A., Trodd, N. M. & Wood, T. F. (1992). Derivation and applications of probabilistic measures of class membership from the maximum-likelihood classification. *Photogrammetric Engineering and Remote Sensing* **58**, 1335–1341.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61**, 215–231.

Haberman, S. J. (1995). Book review of 'Statistical Applications Using Fuzzy Sets', by Manton, K. G., Woodbury, M. A., and Corder, L. S. *Journal of the American Statistical Association* **90**, 1131–1133.

Holland, P. W. (1990). On the sampling theory foundations of the item response theory models. *Psychometrika* **55**, 557–601.

Lazarsfeld, P. F. & Henry, N. W. (1968). *Latent Structure Analysis*. Boston, MA: Houghton Mifflin.

Lindsay, B., Clogg, C. C. & Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association* **86**, 96–107.

Luceno, A. (1999). Discrete approximations to continuous univariate distributions – an alternative to simulation. *Journal of the Royal Statistical Society* **61(2)**, 345–352.

Manton, K. G., Gu, X., Huang, H. & Kovtun, M. (2004). Fuzzy set analyses of genetic determinants of health and disability status. *Statistical Methods in Medical Research* **13**, 395–408.

Manton, K. G., Woodbury, M. A., Anker, M. & Jablensky, A. (1994a). Symptom profiles of psychiatric disorders based on graded disease classes: an illustration using data from the WHO International Pilot Study of Schizophrenia. *Psychological Medicine* **24**, 133–144.

Manton, K. G., Woodbury, M. A. & Tolley, H. D. (1994b). *Statistical Applications Using Fuzzy Sets.* Wiley-Interscience.

Muthen, B. & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* **55**, 463–469.

Potthoff, R. G., Manton, K. G., Woodbury, M. A. & Tolley, H. D. (2000). Dirichlet generalizations of latent-class models. *Journal of Classification* **17**, 315–353.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests.* Neisen and Lydiche, Copenhagen; expanded English edition (1980), University of Chicago Press.

Roeder, K., Lynch, K. G. & Nagin, D. S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association* **94**, 766–776.

Tolley, H. D. & Manton, K. G. (1992). Large sample properties of estimates of a discrete Grade of Membership model. *Annals of the Institute of Statistical Mathematics* **44**, 85–95.

Wachter, K. W. (1999). Grade of membership models in low dimensions. *Statistical Papers* **40**, 439–457.

Woodbury, M. A., Clive, J. & Garson, A. (1978). Mathematical typology: A Grade of Membership technique for obtaining disease definition. *Computers and Biomedical Research* **11**, 277–298.