Population Heterogeneity Representations of Log-linear, Latent Class, Rasch and Grade of Membership Models[1]

Elena A. Erosheva

October 30, 2000

Abstract

This paper employs a geometric approach for describing population heterogeneity representations of different discrete data models. The case of a 2×2 contingency table is considered in details. The paper provides geometric representations for log-linear, latent class, Grade of Membership and Rasch models in the marginal space (based on marginal probabilities) and in the full parameter space (based on cell probabilities). The approach generalizes 2×2 case to higher dimensions, i.e., to 2^J tables with $J \geq 3$. A discussion of similarities and differences among the models is included.

KEY WORDS: Geometric approach, 2×2 contingency table, Marginal probabilities, Cell probabilities, Local independence.

[1]This work was supported by Award #1R03 AG18986-01 from the National Institute on Aging. The presentation of the ideas in this paper owns much to discussions with Stephen Fienberg and Brian Junker.

1 Introduction

For many discrete data problems it is sufficient to assume that individuals are homogeneous in their responses and this naturally leads to log-linear type of models which allow the study of inter-dependences among observed variables. In the middle of the 20th century, however, researchers began to develop models that could be used to capture how individuals differ in their expected responses. In some cases, for example, in response to the question "have you ever considered buying a pick-up truck?", it may be something as simple as gender differences that matters. In other more complicated situations, the probability of an individual having a particular response pattern may depend on a latent quantity or a trait. Examples of such latent quantities are: an indicator of being a member of one of the latent classes (latent class model), an ability parameter (item response theory models), or a grade of membership score (Grade of Membership model). Most of the widely used latent class and item response theory models arose primarily in the social sciences and in education, and they are now relatively well understood statistically. The Grade of Membership (GoM) model arose largely in the context of health-related and demography applications. Although there are a substantial number of articles and book chapters describing these applications, only the limited number of statisticians have worked with the GoM model since 1970s, and the existing literature on the GoM approach has not been fully integrated with other parts of the statistical literature, e.g., with the literature on latent class models ([3]). Even though there exist theoretic formulae for the probability of response patterns as functions of the latent parameters for most latent variable statistical models, a geometric approach may bring sharper understanding of model capabilities with respect to describing population heterogeneity in responses. Recent publications that have considered geometric approaches are: Ramsey, [10], for the item response theory, and Woodbury, Manton, and Tolley, [12], for the Grade of Membership model.

This paper provides geometric representations for population heterogeneity of log-linear, latent class, Grade of Membership and Rasch models in the marginal space (based on marginal probabilities) and in the full parameter space (based on cell probabilities). The case of a 2×2 table is considered in details in Sections 3-6. Then, by using these geometric representations, in Section 7 we highlight similarities and differences between the Grade of Membership model, the latent class models, and the Rasch model (a representative of the item response theory models). Finally, we discuss some aspects of increased dimensionality in Section 8.

2 Preliminaries

Consider a two by two table, normalized in such a way that all cell entries $p_{lm} = P(Y_1 = l, Y_2 = m)$, $l, m = 1, 2$, add up to one. Here $\lambda_{lm} = P(Y_l = m)$, $l, m = 1, 2$, are the marginal probabilities. The marginal notation is not parsimonious but it does allow for easy generalization to the latent class and the GoM models.

$$
\begin{array}{cc|c}
\multicolumn{2}{c}{Y_2} & \\
p_{11} & p_{12} & \lambda_{11} \\
p_{21} & p_{22} & \lambda_{12} \\
\hline
\lambda_{21} & \lambda_{22} & 1
\end{array}
$$

with the two rows labelled Y_1.

There are different geometric representations based on the two sets of parameters: the cell probabilities, $\{p_{lm}\}$, and the marginal probabilities, $\{\lambda_{lm}\}$. Note, without any model restrictions, the marginal probabilities can be identified from the cell entries but not vice versa. Therefore, the geometric representation for the set of marginal probabilities can be derived from the geometric representation for the corresponding set of cell entries, but the converse in general is not true.

We will refer to the $(0,1) \times (0,1)$ square in the Cartesian plane with the basis vectors $(\lambda_{11}, \lambda_{21})$ as the *marginal space*. Since marginal probabilities for the same item necessarily add up to one, every possible set of margins for a two by two table corresponds to a point in the marginal space and vice versa. Ramsey [10] denotes the same construct as the *response probability space*. The marginal space is a simplification for the *space of all possible multinomial probabilities* used by Woodbury, Manton and Tolley in [12] to examine the Grade of Membership (GoM) model. The latter is build using as many independent basis vectors as there are marginal probabilities (e.g., four in the two by two case).

Constraints for the cell probabilities $p_{lm} \geq 0,\ l,m = 1,2$ and $\sum_{l,m} p_{lm} = 1$ define an object in the four-dimensional Cartesian space. Consider a projection from the four-dimensional space with orthogonal basis vectors corresponding to $(p_{11}, p_{12}, p_{21}, p_{22})$ onto a subspace with the basis corresponding to (p_{12}, p_{21}, p_{22}). Resulting object is a tetrahedron with the origin representing $p_{11} = 1$, and three other vertices representing $p_{12} = 1$, $p_{21} = 1$, $p_{22} = 1$ respectively. Fienberg and Gilbert [2] used an isomorphic tetrahedron to describe the geometry of the cell entries of a two by two table. It was also pointed out that there exists a one-to-one correspondence between points in the tetrahedron and all possible two by two tables. We will refer to the tetrahedron as the *full parameter space*. Under the independence assumption, the product of the pair of corresponding margins defines the cell entry. Using the independence assumption and the constraint that the cell entries add up to one, the equation for the *surface of independence* is obtained:

$$(p_{22})^2 - p_{22} \cdot (1 - p_{21} - p_{12}) + p_{21} \cdot p_{12} = 0.$$

The surface of independence is a part of the hyperboloid in the full parameter space (Figures 1, 2). It has one distinctive property that the rulings of the hyperboloid (two families of straight lines that define the surface [5]) are the lines of constant row and column margins.

Building a model for discrete data is equivalent to describing relationship among the cell probabilities in the full parameter space. Marginal representation is sufficient for a model if there exists a one-to-one correspondence between representation of the model in the full parameter space and corresponding representation in the marginal space. This

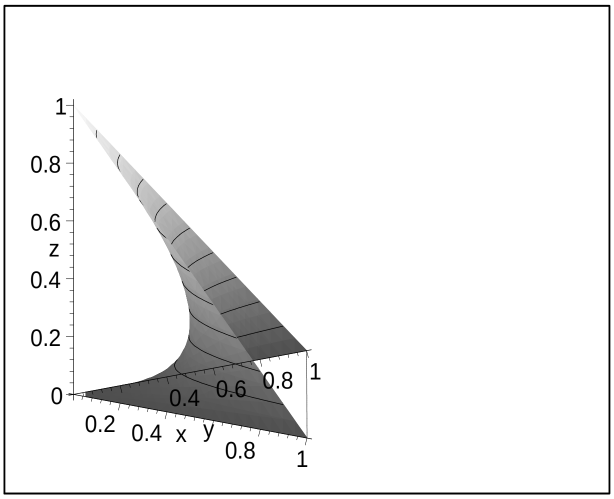

Figure 1: Surface of independence in the full parameter space $x = p_{12}, y = p_{21}, z = p_{22}$. Contour lines are given for a better perception of the 3dimensional-surface.

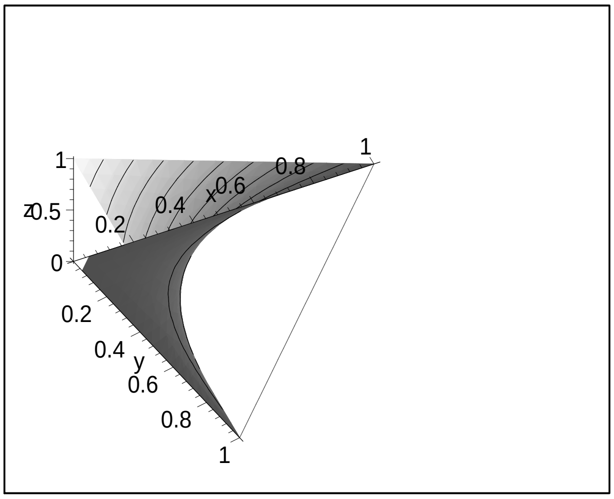

Figure 2: Surface of independence in the full parameter space $x = p_{12}, y = p_{21}, z = p_{22}$. Contour lines are given for a better perception of the 3dimensional-surface.

is the case for all models considered here. As we will show, however, understanding the geometry in the full parameter space may be required to compare different model structures. We will describe the Grade of Membership model by means of similarities and differences with the latent class model and a representative of the item response theory models, the Rasch model. We will begin with the simplest case, log-linear model of independence.

3 Log-linear model of independence

Suppose the population responses come from a distribution with fixed marginal probabilities for two dichotomous variables Y_1, Y_2. The independence condition can be expressed as

$$p_{11} \cdot p_{22} = p_{12} \cdot p_{21}$$

This case is well understood as a log-linear model of independence. Given the set of margins, individual responses are assumed to be homogeneous. That is, every individual is assumed to have the same cell probabilities $p_{lm} = \lambda_{1l}\lambda_{2m}$, $l, m = 1, 2$, for the four response patterns. Geometrically, this situation can be represented by a single point with coordinates $(\lambda_{11}, \lambda_{21})$ in the marginal space. Similarly, in the full parameter space, the log-linear model of independence corresponds to a point on the surface of independence with coordinates (p_{12}, p_{21}, p_{22}).

4 Latent class model

Now assume that there are two distinct classes in the population. As an hypothetical example, consider a test with Y_1 being a math question and Y_2 being a history question. If the classes are students with math and history majors which rate differently difficulty of the questions, marginal probabilities are not homogeneous any more. Let $p^k_{lm} = P(Y_1 = l, Y_2 = m|Z = k)$, $l, m, k = 1, 2$, and $\lambda^k_{lm} = P(Y_l = m|Z = k)$, $l, m, k = 1, 2$, where Z is the latent class indicator. Then population responses for the two classes can be represented as

$$\begin{array}{cc|c} \multicolumn{3}{c}{Y_2} \\ p^1_{11} & p^1_{12} & \lambda^1_{11} \\ p^1_{21} & p^1_{22} & \lambda^1_{12} \\ \hline \lambda^1_{21} & \lambda^1_{22} & 1 \end{array} \quad Y_1 \qquad \begin{array}{cc|c} \multicolumn{3}{c}{Y_2} \\ p^2_{11} & p^2_{12} & \lambda^2_{11} \\ p^2_{21} & p^2_{22} & \lambda^2_{12} \\ \hline \lambda^2_{21} & \lambda^2_{22} & 1 \end{array}$$

The assumption that responses Y_1 and Y_2 are independent, conditionally on the latent class (e.g., student's major), is known as the *local independence assumption*:

$$p^k_{11} \cdot p^k_{22} = p^k_{12} \cdot p^k_{21}, \; k = 1, 2.$$

Lindsay, Clogg and Grego [7] call the general case of this model with J binomial responses and K latent classes the *conventional latent class model.* Notice, that a latent class model with two dichotomous responses and two latent classes is not identifiable. In this case, there are more parameters in the model than quantities to be fit.

Table 1: Example

	Y_2		
Y_1	0.03	0.07	0.1
	0.27	0.63	0.9
	0.3	0.7	1

	Y_2		
Y_1	0.48	0.32	0.8
	0.12	0.08	0.2
	0.6	0.4	1

Geometrically, marginal probabilities for the two classes correspond to two points M_1, M_2 in the marginal space, which, in turn, correspond to the points P_1, P_2 on the surface of independence in the full parameter space. To give an illustration, throughout this paper we will use the values for the cell and marginal probabilities (Table 1) from the numerical example given by Manton et al. in [8], page 334. The loci M_1, M_2 and P_1, P_2 for this example are shown on Figures 3 and 4 respectively.

The usual treatment of the latent class model also involves the assumption that each individual is a member of one and only one latent class. In the face of uncertainty, however, an individual can be assigned *probability* q of being classified in one of the latent classes. For example, in the context of latent class measurement models [4], a special interest is in estimation of a student-level parameter, in this case, the probability q that a student belongs to one of the two classes. If this can be achieved, heterogeneity of individual responses can now be described by the set of marginal class probabilities and the set of q, probabilities of belonging to one of the classes. Note that the latent class measurement model does not include *explicitly* individual-level parameters in the model specification. Probability q can be estimated *a posteriori* using empirical Bayes method. Therefore, population heterogeneity can be implicitly described by a latent class model.

Consider a mixture model with two latent classes and a person-specific probability q of belonging to the first class. If q could vary continuously from 0 to 1, then the corresponding cell probabilities for individual responses would vary according to

$$\begin{aligned} p_{lm} &= P(Y_1 = l, Y_2 = m) \\ &= q \cdot P(Y_1 = l, Y_2 = m | Z = 1) + (1 - q) \cdot P(Y_1 = l, Y_2 = m | Z = 2), \end{aligned}$$

where $l, m = 1, 2$. The values of p_{lm} that satisfy the above equation form a straight line segment P_1P_2 in the full parameter space (Figure 4) and a line segment M_1M_2 (Figure 3) in the marginal space. Unless the marginal probabilities for either the response Y_1 or the response Y_2 are the same for the two classes, the line segment P_1P_2 does not lie on the surface of independence. We refer to this line segment as the population heterogeneity manifold for the latent class model. We should note, however, that there is only a finite

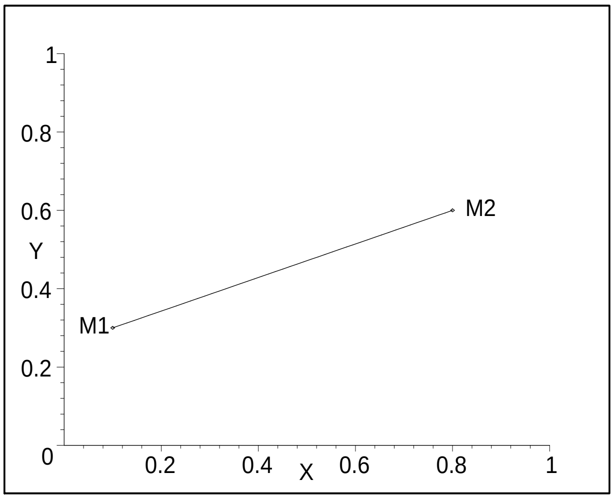

Figure 3: Latent class model heterogeneity manifold in the marginal space $x = \lambda_{11}, y = \lambda_{21}$.

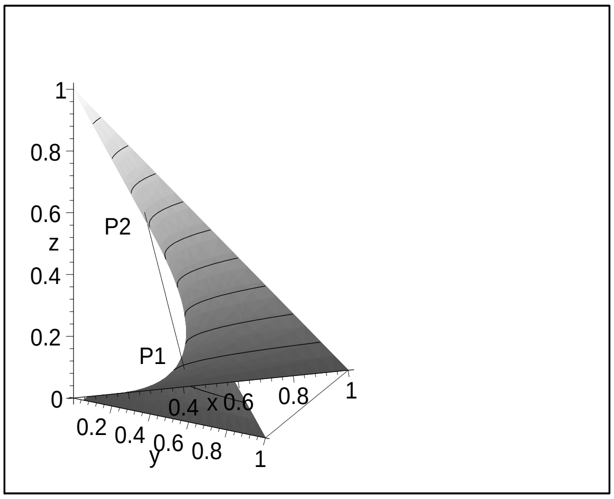

Figure 4: Latent class model heterogeneity manifold in the full parameter space $x = p_{12}, y = p_{21}, z = p_{22}$.

number of possible values of q that can be estimated from discrete data. The number of distinct values of q depends on a specific model used, but it can never be greater than the number of different patterns of the data records, 2^J in dichotomous case.

5 Rasch model

In the context of the item response theory (IRT), a number of models explicitly involve subject-specific parameters, which are usually called *ability parameters*. One of the most common examples is the Rasch model. The assumption here is that the marginal probability of response depends on the subject and item parameters through the logistic ogive function. For the Rasch model, no interaction between subject and item parameters is allowed. We will consider two cases, *latent class Rasch model* and *latent trait Rasch model*.

Suppose Y_1 and Y_2 are two dichotomous items with $Y_1 = 1$ and $Y_2 = 1$ being the correct responses. Denote the subject *ability* parameter by θ and the item *difficulty* parameters by b_1 and b_2 respectively. Then, given the item difficulty parameters, the two by two table for the Rasch model is

$$
\begin{array}{cccc|c}
 & & Y_2 & & \\
Y_1 & p_{11}(\theta) & & p_{12}(\theta) & \frac{\exp(\theta-b_1)}{1+\exp(\theta-b_1)} \\
 & p_{21}(\theta) & & p_{22}(\theta) & \frac{1}{1+\exp(\theta-b_1)} \\
\hline
 & \frac{\exp(\theta-b_2)}{1+\exp(\theta-b_2)} & & \frac{1}{1+\exp(\theta-b_2)} & 1
\end{array}
$$

And the independence assumption becomes

$$p_{11}(\theta) \cdot p_{22}(\theta) = p_{12}(\theta) \cdot p_{21}(\theta),\ k = 1, 2.$$

The assumption of two distinct classes in the population corresponds to two distinct values of $\theta = \theta_1, \theta_2$. This results in the latent class Rasch model. Because no interaction between the item and the subject parameters is allowed, it is technically a special case of the conventional latent class model described earlier.

The latent trait Rasch model has the same set of parameters as the latent class Rasch model but the ability parameter θ is treated as being continuous. If we let θ vary from $-\infty$ to ∞, the marginal probabilities $\lambda_{11} = P(Y_1 = 1)$ and $\lambda_{21} = P(Y_2 = 0)$ will vary from 0 to 1. Expressing one marginal probability as a function of another, heterogeneity in individual responses can be represented by a curvilinear segment in the marginal space:

$$\lambda_{21}(\lambda_{11}) \quad = \quad \frac{\lambda_{11}\exp(b_1 - b_2)}{1 - \lambda_{11} + \lambda_{11}\exp(b_1 - b_2)},$$

where λ_{11} and λ_{21} are the probabilities of correct responses, $\lambda_{11} \in (0, 1)$. This curve connects the point of complete ignorance, $(0, 0)$, with the point of full knowledge, $(1, 1)$. The curvature of the segment depends on the difference between the difficulty parameters

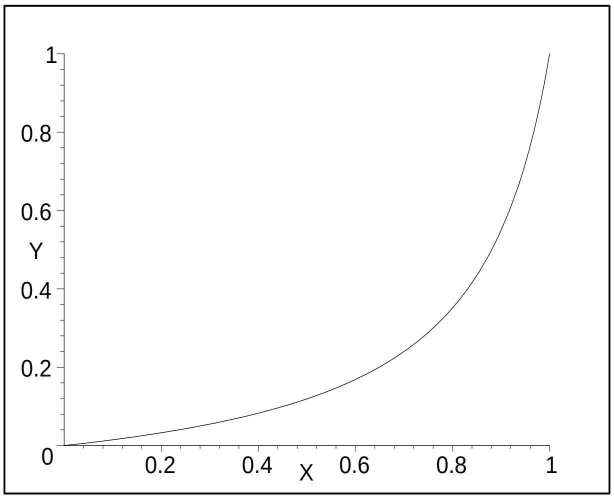

Figure 5: Latent trait Rasch model heterogeneity manifold in the marginal space $x = \lambda_{11}, y = \lambda_{21}$.

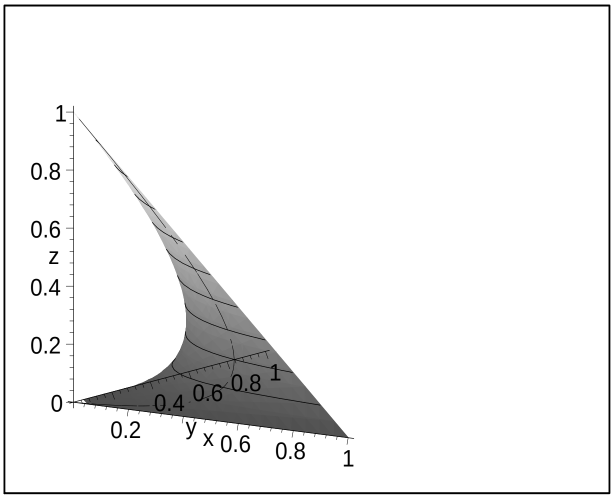

Figure 6: Latent trait Rasch model heterogeneity manifold in the full parameter space $x = p_{12}, y = p_{21}, z = p_{22}$.

of the items. When ability increases until a certain point, the marginal probability of a correct response to an easier item increases faster than the marginal probability of a correct response to a more difficult item. After a certain ability level is reached, the probability of a correct response for a harder item quickly catches up. Figure 5 shows an example of the Rasch model heterogeneity manifold in the marginal space for the two items such that $b_2 - b_1 = 2$. The curvilinear segment becomes a straight line in the marginal space when the items are of the same difficulty, $b_2 - b_1 = 0$.

Since the independence assumption is satisfied for every $\theta \in (-\infty, \infty)$, the corresponding heterogeneity manifold for the latent trait Rasch model in the full parameter space is a curve on the surface of independence that connects the origin $p_{11} = 1$ and the vertex $p_{22} = 1$. Changing θ from $-\infty$ to ∞ is equivalent to changing $\lambda_{11} = P(Y_1 = 1)$ from 0 to 1. Expressing cell probabilities as functions of λ_{11} for our convenience, the Rasch model heterogeneity manifold (Figure 6) is:

$$
\begin{aligned}
p_{12}(\lambda_{11}) &= \lambda_{11} \cdot \left(1 - \frac{\lambda_{11}\exp(b_1 - b_2)}{1 - \lambda_{11} + \lambda_{11}\exp(b_1 - b_2)}\right) \\
p_{21}(\lambda_{11}) &= (1 - \lambda_{11}) \cdot \left(\frac{\lambda_{11}\exp(b_1 - b_2)}{1 - \lambda_{11} + \lambda_{11}\exp(b_1 - b_2)}\right) \\
p_{12}(\lambda_{11}) &= (1 - \lambda_{11}) \cdot \left(1 - \frac{\lambda_{11}\exp(b_1 - b_2)}{1 - \lambda_{11} + \lambda_{11}\exp(b_1 - b_2)}\right),
\end{aligned}
$$

where $\lambda_{11} \in (0, 1)$. Since θ is involved in the model specification, the population heterogeneity manifold is explicitly described by the Rasch model.

6 Grade of Membership model

Another model which explicitly involves subject-level parameters is the *Grade of Membership* (GoM) model [9] that was developed for the analysis of categorical data on disease [11, 1]. GoM model postulates that individual marginal probabilities can be expressed as linear combinations of *extreme* marginal probabilities and a vector of grades of membership in corresponding *extreme profiles.* This model has also been called a fuzzy set model because an individual is assumed to be a partial member of several extreme classes simultaneously [9].

Suppose there are two extreme profiles. Using similar notation as for the latent class model, the two by two table for the GoM model becomes

$$
\begin{array}{cccc|c}
 & & Y_2 & & \\
Y_1 & p_{11}(\mathbf{g}) & & p_{12}(\mathbf{g}) & g_1\lambda_{11}^1 + g_2\lambda_{11}^2 \\
 & p_{21}(\mathbf{g}) & & p_{22}(\mathbf{g}) & g_1\lambda_{12}^1 + g_1\lambda_{12}^2 \\
\hline
 & g_1\lambda_{21}^1 + g_2\lambda_{21}^2 & & g_1\lambda_{22}^1 + g_2\lambda_{22}^2 & 1
\end{array}
$$

where $\mathbf{g} = (g_1, g_2)$ is the grade of membership vector, such that $g_1 + g_2 = 1,\ g_1 \geq 0,\ g_2 \geq 0$.

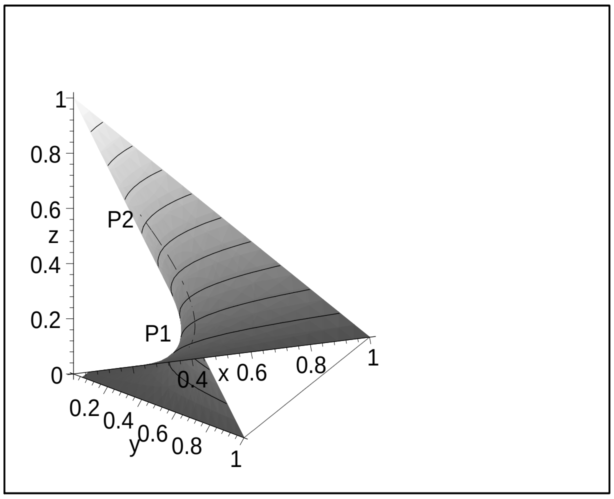

Figure 7: GoM model heterogeneity manifold in the full parameter space $x = p_{12}, y = p_{21}, z = p_{22}$.

The local independence assumption states that variables Y_1 and Y_2 are independent, given the grade of membership score:

$$p_{11}(\mathbf{g}) \cdot p_{22}(\mathbf{g}) = p_{12}(\mathbf{g}) \cdot p_{21}(\mathbf{g}), \; k = 1, 2. \tag{1}$$

It should be pointed out that, just like the latent class model considered above, the grade of membership model with two extreme profiles for two dichotomous responses is not identifiable.

Since the coordinates of the $\mathbf{g}$ vector add up to one, let $g_1 = g$ and $g_2 = 1 - g$. Then the population heterogeneity manifold for the GoM model can be obtained by varying g from 0 to 1. In the marginal space,

$$\begin{aligned} \lambda_{11}(g) &= g \cdot \lambda_{11}^1 + (1 - g) \cdot \lambda_{11}^2 \\ \lambda_{21}(g) &= g \cdot \lambda_{21}^1 + (1 - g) \cdot \lambda_{21}^2, \end{aligned}$$

where $\lambda_{11} = P(Y_1 = 1)$, $\lambda_{21} = P(Y_2 = 1)$, $g \in (0, 1)$, is a parametric representation of the line segment M_1M_2, which is identical to the population heterogeneity manifold for the latent class model (Figure 3). In the full parameter space, the heterogeneity manifold for the GoM model is a curve segment on the surface of independence connecting P_1 and P_2 (Figure 7):

$$\begin{aligned} p_{12}(g) &= (g\lambda_{11}^1 + (1 - g)\lambda_{11}^2) \cdot (g(1 - \lambda_{21}^1) + (1 - g)(1 - \lambda_{21}^2)) \\ p_{21}(g) &= (g(1 - \lambda_{11}^1) + (1 - g)(1 - \lambda_{11}^2)) \cdot (g\lambda_{21}^1 + (1 - g)\lambda_{21}^2) \\ p_{12}(g) &= (g(1 - \lambda_{11}^1) + (1 - g)(1 - \lambda_{11}^2)) \cdot (g(1 - \lambda_{21}^1) + (1 - g)(1 - \lambda_{21}^2)), \end{aligned}$$

where $g \in (0, 1)$.

7 Similarities and Differences

From the geometric representations for the case of two dichotomous responses and two latent classes, some similarities and differences may be seen between the GoM model, the Rasch model, and the latent class model.

First, we note that the Rasch model assumes implicitly the existence of two extreme classes: $\theta = -\infty$ for complete ignorance and $\theta = \infty$ for full knowledge. In the simplest case described here, the loci of these two classes are fixed. Once we introduce other item parameters, the points for the complete ignorance and full knowledge may have other locations in the marginal space [10] or, equivalently, on the surface of independence. Because the IRT models assume positive association requirement, however, not all locations of extreme profiles are possible for the IRT-type models. For example, the complete ignorance and full knowledge points can never be located at the vertices $p_{12} = 1, p_{21} = 1$. This distinguishes the Rasch-type models from the conventional latent class models, where no restrictions are placed on the location of the extreme profiles.

The ability parameter θ in the Rasch model plays a role similar to the grade of membership parameter g in the GoM model. They both appear explicitly in formulation of the models and in the local independence assumption. Therefore, they both can be referred to as *state* parameters, the terminology used to describe the grade of membership scores [9, 12]. There are distinctions between their roles. For example, θ is *extrinsic* to the manifold in the sense that the relationship of position on the domain of θ, $(-\infty, \infty)$ in our case, is only indirectly related to position within the manifold itself [10], whereas g for the GoM model is *intrinsic* to the manifold, as it measures the distance along the manifold from its beginning, assuming the manifold is normalized to have a unit length. Further, θ for the Rasch model is a unidimensional parameter, whereas g for the GoM model in general is multidimensional.

Next we consider geometric representations for the GoM and the latent class models. The population heterogeneity manifold for the GoM model in the marginal space is identical to the one obtained above for the latent class model. Aside from the fact that the GoM model provides an explicit representation of population heterogeneity with respect to the pure types and the latent class model can do it only implicitly, these manifolds are clearly different in the full parameter space. To illustrate the nature of the difference, we use a numerical example with two dichotomous variables given by Manton et al. [8]. In this example, they compared the latent class and the GoM model with two latent classes by assuming that posterior probability of being in the first class for the latent class model is equal to the GoM score for the first class (and equal to 0.6). They showed that the latent class and the GoM models yield identical predictions of the *marginal* probabilities but different estimates of the *joint* probabilities. Figure 8 shows the latent class and the GoM model heterogeneity manifolds for this example. The points on the manifolds correspond to $g = q = 0.6$ respectively.

From the geometric representation of the GoM and the latent class models, it is clear that a mixture of latent classes can not be equivalent to the GoM model in general. That

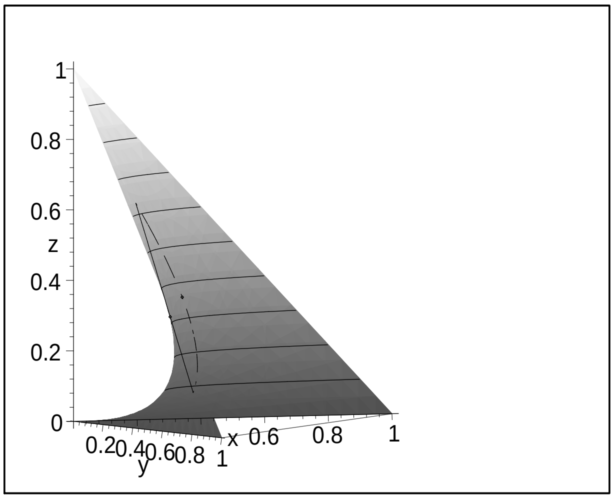

Figure 8: Illustration for the numerical example. Straight line is the heterogeneity manifold for the latent class model. Curve on the surface is the heterogeneity manifold to the GoM model. Points correspond to $g = q = 0.6$.

is, the cell probabilities from the GoM model can not be matched by using a mixture of the latent classes with specified marginal probabilities. This contradicts the claim first made by Haberman in his review of *Statistical Applications Using Fuzzy Sets* [3] and then by Woodbury, Tolley and Manton in [13]. In principle, however, for a given data set the two models can be indistinguishable in the same way as the latent class Rasch model can be identical (produce identical ability estimates) to the Rasch model with a continuous latent trait [7]. From the geometric representation, a necessary condition for the GoM model to be indistinguishable from a latent class model is that the number of latent classes should be equal to the number of distinct estimated vectors of the grade of membership scores (the vectors are considered distinct if their loci in the full parameter space are distinct). A question requiring further investigation is: under which conditions the estimated cell probabilities would be identical and the GoM and the latent class model parameters would be simultaneously identifiable (not identifiable) from the same data?

8 Increasing Dimensionality

The geometric representation can be extended to 2^J table for $J \geq 3$. For example, for $J = 3$, the marginal space is a unit cube. Even though we can not picture the independence manifold in four-dimensional space, it can be shown that a line which connects two arbitrary points on the manifold does not belong to the manifold unless at least one pair of margins is the same for the two classes that are determined by the two points. In the marginal space, the GoM [latent class] model population heterogeneity manifold is a line segment in the cube connecting the two extreme profiles [latent classes]. In the full parameter space, it is a line and a curve on the independence manifold for the latent class and the GoM models respectively. When J increases, the marginal space becomes a J-dimensional cube and the full parameter space becomes an $(2^J - 1)$-dimensional polyhedron. For the case of two extreme profiles [latent classes] we still be looking on a curve and its position relative to the manifold of independence.

The kind of geometric representation described here can be used to gain some insights on the structure of other models that deal with population heterogeneity. For example, it can be extended to the class of constrained latent class models, and, specifically, a multidimensional item response model described by Hoijtink and Molenaar [6].

References

[1] Jonathan Clive, Max A. Woodbury, and Ilene C. Siegler. Fuzzy and Crisp Set-Theoretic-Based Classification of Health and Disease. A Qualitative and Quantitative Comparison. *Journal of Medical Systems*, 1983.

[2] Stephen E. Fienberg and John P. Gilbert. The Geometry of a Two by Two Contingency Table. *Journal of the American Statistical Association*, 1970.

[3] Shelby J. Haberman. Book review of 'Statistical Applications Using Fuzzy Sets', by Kenneth G. Manton, Max A. Woodbury, and Larry S. Corder. *Journal of the American Statistical Association*, 1995.

[4] Ton Heinen. *Latent Class and Discrete Latent Trait Models: Similarities and Differences*. Sage, Newbury Park, CA, 1996.

[5] D. Hilbert and S. Cohn-Vossen. *Geometry and the Imagination*. Chelsea Publishing Company. New York., 1952.

[6] Herbert Hoijtink and Ivo W. Molenaar. A Multidimensional Item Response Model: Constrained Latent Class Analysis Using Gibbs Sampler and Posterior Predictive Checks. *Psychometrika*, 1997.

[7] Bruce Lindsay, Clifford C. Clogg, and John Grego. 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*, 1991.

[8] Kenneth G. Manton, Max A. Woodbury, Eric Stallard, and Larry S. Corder. The Use of Grade-of-membership Techniques to Estimate Regression Relationships. *Sociological Methodology*, 22:321–381, 1992.

[9] Kenneth G. Manton, Max A. Woodbury, and H. Dennis Tolley. *Statistical Applications Using Fuzzy Sets*. Wiley, 1994.

[10] J. O. Ramsay. A Geometrical Approach to Item Response Theory. *Behaviormetrika*, 1996.

[11] Max A. Woodbury and Kenneth G. Manton. A New Procedure for Analysis of Medical Classification. *Methods of Information in Medicine*, 1982.

[12] Max A. Woodbury, Kenneth G. Manton, and H. Dennis Tolley. Convex Models of High Dimensional Discrete Data. *Annals of the Institute of Statistical Mathematics*, 49:371–393, 1997.

[13] Max A. Woodbury, Dennis H. Tolley, and Kenneth G. Manton. Grade of Membership Analysis: Models to Analyze Categorical Data Based on a Convex Geometry. Unpublished Manuscript. Undated.