Biometrika (1974), **61**, 2, *p.* 215
Printed in Great Britain

Exploratory latent structure analysis using both identifiable and unidentifiable models

By LEO A. GOODMAN
Department of Statistics, University of Chicago

SUMMARY

This paper considers a wide class of latent structure models. These models can serve as possible explanations of the observed relationships among a set of m manifest polytomous variables. The class of models considered here includes both models in which the parameters are identifiable and also models in which the parameters are not. For each of the models considered here, a relatively simple method is presented for calculating the maximum likelihood estimate of the frequencies in the m-way contingency table expected under the model, and for determining whether the parameters in the estimated model are identifiable. In addition, methods are presented for testing whether the model fits the observed data, and for replacing unidentifiable models that fit by identifiable models that fit. Some illustrative applications to data are also included.

Some key words: Contingency tables; Latent structure; Log linear models; Maximum likelihood estimation; Tests of fit.

1. INTRODUCTION

This paper deals with the relationships among m polytomous variables, i.e. with the analysis of an m-way contingency table. These m variables are manifest variables in that, for each observed individual in a sample, his class with respect to each of the m variables is observed. We also consider here polytomous variables that are latent in that an individual's class with respect to these variables is not observed. The classes of a latent variable will be called latent classes.

Consider first a 4-way contingency table which cross-classifies a sample of n individuals with respect to four manifest polytomous variables A, B, C and D. If there is, say, some latent dichotomous variable X, so that each of the n individuals is in one of the two latent classes with respect to this variable, and within the tth latent class the manifest variables (A, B, C, D) are mutually independent, then this two-class latent structure would serve as a simple explanation of the observed relationships among the variables in the 4-way contingency table for the n individuals. There is a direct generalization when the latent variable has T classes. We shall present some relatively simple methods for determining whether the observed relationships among the variables in the m-way contingency table can be explained by a T-class structure, or by various modifications and extensions of this latent structure.

To illustrate the methods we analyze Table 1, a 2^4 contingency table presented earlier by Stouffer & Toby (1951, 1962, 1963), which cross-classifies 216 respondents with respect to whether they tend towards universalistic values (+) or particularistic values (−) when confronted by each of four different situations of role conflict. The letters A, B, C and D in

Table 1. *Observed cross-classification of* 216 *respondents with respect to whether they tend toward universalistic* (+) *or particularistic* (−) *values in four situations of role conflict* (*A*, *B*, *C*, *D*)

A	*B*	*C*	*D*	Observed frequency	*A*	*B*	*C*	*D*	Observed frequency
+	+	+	+	42	−	+	+	+	1
+	+	+	−	23	−	+	+	−	4
+	+	−	+	6	−	+	−	+	1
+	+	−	−	25	−	+	−	−	6
+	−	+	+	6	−	−	+	+	2
+	−	+	−	24	−	−	+	−	9
+	−	−	+	7	−	−	−	+	2
+	−	−	−	38	−	−	−	−	20

Table 1 denote the dichotomous responses when confronted by the four different situations. In addition, a second illustrative example in Table 4 below will also be discussed briefly. Our analysis of these data leads to conclusions that are different from those presented earlier.

2. The latent class model unrestricted

Suppose that the manifest polytomous variables A, B, C and D consist of I, J, K and L classes, respectively. Let π_{ijkl} denote the probability that an individual will be at level (i,j,k,l) with respect to the joint variable (A,B,C,D) $(i = 1,\ldots,I; j = 1,\ldots,J; k = 1,\ldots,K; l = 1,\ldots,L)$. Suppose that there is a latent polytomous variable X, consisting of T classes, that can explain the relationships among the manifest variables (A, B, C, D). This means that π_{ijkl} can be expressed as follows:

$$\pi_{ijkl} = \sum_{t=1}^{T} \pi_{ijklt}^{ABCDX}, \tag{1}$$

where

$$\pi_{ijklt}^{ABCDX} = \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X} \tag{2}$$

denotes the probability that an individual will be at level (i,j,k,l,t) with respect to the joint variable (A,B,C,D,X). Here π_t^X denotes the probability that an individual will be at level t with respect to variable X; also $\pi_{it}^{\bar{A}X}$ denotes the conditional probability that an individual will be at level i with respect to variable A, given that he is at level t with respect to variable X, and finally $\pi_{jt}^{\bar{B}X}$, $\pi_{kt}^{\bar{C}X}$ and $\pi_{lt}^{\bar{D}X}$ denote similar conditional probabilities. Formula (1) states that the individuals can be classified into T mutually exclusive and exhaustive latent classes, and formula (2) states that within the tth latent class the manifest variables (A, B, C, D) are mutually independent $(t = 1,\ldots,T)$. The meaning of the latent polytomous variable X will be clarified further when particular examples are considered later.

The following elementary formulae, (3)–(8), are required to obtain the subsequent results:

$$\sum_{t=1}^{T} \pi_t^X = 1, \quad \sum_{i=1}^{I} \pi_{it}^{\bar{A}X} = 1, \quad \sum_{j=1}^{J} \pi_{jt}^{\bar{B}X} = 1, \quad \sum_{k=1}^{K} \pi_{kt}^{\bar{C}X} = 1, \sum_{l=1}^{L} \pi_{lt}^{\bar{D}X} = 1, \tag{3}$$

$$\pi_t^X = \sum_{i,j,k,l} \pi_{ijklt}^{ABCDX}, \tag{4}$$

$$\pi_t^X \pi_{it}^{\bar{A}X} = \sum_{j,k,l} \pi_{ijklt}^{ABCDX}. \tag{5}$$

Formulae similar to (5) can be obtained for π_t^X multiplied by $\pi_{jt}^{\bar{B}X}$, $\pi_{kt}^{\bar{C}X}$ and $\pi_{lt}^{\bar{D}X}$. In addition, we obtain

$$\pi_{ijklt}^{ABCD\bar{X}} = \pi_{ijklt}^{ABCDX}/\pi_{ijkl}, \tag{6}$$

where $\pi_{ijklt}^{ABCD\bar{X}}$ denotes the conditional probability that an individual is in latent class t, given that he was at level (i, j, k, l) with respect to the joint variable (A, B, C, D). Using (6), we can rewrite (4) and (5) as

$$\pi_t^X = \sum_{i,j,k,l} \pi_{ijkl}\pi_{ijklt}^{ABCD\bar{X}}, \tag{7}$$

$$\pi_{it}^{\bar{A}X} = (\sum_{j,k,l} \pi_{ijkl}\pi_{ijklt}^{ABCD\bar{X}})/\pi_t^X. \tag{8}$$

Formulae similar to (8) can be obtained for $\pi_{jt}^{\bar{B}X}$, $\pi_{kt}^{\bar{C}X}$ and $\pi_{lt}^{\bar{D}X}$. Without loss of generality, we can assume that $\pi_t^X > 0$; we also assume that $\pi_{ijkl} > 0$.

Let circumflexes denote the maximum likelihood estimates of the corresponding parameters in the latent-class model. From (1) and (2), we obtain

$$\hat{\pi}_{ijkl} = \sum_{t=1}^{T} \hat{\pi}_{ijklt}^{ABCDX}, \tag{9}$$

where

$$\hat{\pi}_{ijklt}^{ABCDX} = \hat{\pi}_t^X \hat{\pi}_{it}^{\bar{A}X} \hat{\pi}_{jt}^{\bar{B}X} \hat{\pi}_{kt}^{\bar{C}X} \hat{\pi}_{lt}^{\bar{D}X}; \tag{10}$$

and from (6) we obtain

$$\hat{\pi}_{ijklt}^{ABCD\bar{X}} = \hat{\pi}_{ijklt}^{ABCDX}/\hat{\pi}_{ijkl}. \tag{11}$$

If p_{ijkl} denotes the observed proportion of individuals at level (i, j, k, l) with respect to the joint variable (A, B, C, D), standard methods prove that the maximum likelihood estimates satisfy the following system of equations:

$$\hat{\pi}_t^X = \sum_{i,j,k,l} p_{ijkl}\hat{\pi}_{ijklt}^{ABCD\bar{X}}, \tag{12}$$

$$\hat{\pi}_{it}^{\bar{A}X} = (\sum_{j,k,l} p_{ijkl}\hat{\pi}_{ijklt}^{ABCD\bar{X}})/\hat{\pi}_t^X, \tag{13a}$$

$$\hat{\pi}_{jt}^{\bar{B}X} = (\sum_{i,k,l} p_{ijkl}\hat{\pi}_{ijklt}^{ABCD\bar{X}})/\hat{\pi}_t^X, \tag{13b}$$

$$\hat{\pi}_{kt}^{\bar{C}X} = (\sum_{i,j,l} p_{ijkl}\hat{\pi}_{ijklt}^{ABCD\bar{X}})/\hat{\pi}_t^X, \tag{13c}$$

$$\hat{\pi}_{lt}^{\bar{D}X} = (\sum_{i,j,k} p_{ijkl}\hat{\pi}_{ijklt}^{ABCD\bar{X}})/\hat{\pi}_t^X. \tag{13d}$$

Compare (12) and (13a) with (7) and (8); recall that $\hat{\pi}_{ijklt}^{ABCD\bar{X}}$ in (12) and (13a)–(13d) was defined by (11), (10) and (9).

Let π denote the vector of parameters $(\pi_t^X, \pi_{it}^{\bar{A}X}, \pi_{jt}^{\bar{B}X}, \pi_{kt}^{\bar{C}X}, \pi_{lt}^{\bar{D}X})$ in the latent-class model, and let $\hat{\pi}$ denote the corresponding maximum likelihood estimate of the vector. To calculate $\hat{\pi}$, we apply the following iterative procedure. Start with an initial trial value for $\hat{\pi}$, say $\pi(0) = \{\pi_t^X(0), \pi_{it}^{\bar{A}X}(0), \pi_{jt}^{\bar{B}X}(0), \pi_{kt}^{\bar{C}X}(0), \pi_{lt}^{\bar{D}X}(0)\}$, which we discuss later. Then use formula (10) to obtain a trial value for $\hat{\pi}_{ijklt}^{ABCDX}$, replacing the terms on the right-hand side of (10) by the corresponding components in $\pi(0)$. We then use (9) to obtain a trial value for $\hat{\pi}_{ijkl}$, replacing the term on the right-hand side of (9) by the corresponding trial value; and we use (11) to obtain a trial value for $\hat{\pi}_{ijklt}^{ABCD\bar{X}}$, replacing the terms on the right-hand side of (11) by the corresponding trial values. Similarly, we use formula (12) to obtain a new trial value for $\hat{\pi}_t^X$, and (13a) to (13d) to obtain new trial values for $\hat{\pi}_{it}^{\bar{A}X}$, $\hat{\pi}_{jt}^{\bar{B}X}$, $\hat{\pi}_{kt}^{\bar{C}X}$ and $\hat{\pi}_{lt}^{\bar{D}X}$. Having thus obtained a new trial value for the vector $\hat{\pi}$, we repeat the procedure starting with the new trial value using in turn formulae (10), (9), (11), (12) and (13a)–(13d) to obtain the next trial value for $\hat{\pi}$. In this iterative procedure a latent class is deleted if the corresponding estimate tends to zero. The procedure will converge to a solution to the system of equations and to a corresponding likelihood. For some related but different results, see Haberman (1974).

From (12) and (13a)–(13d), we find that the components of $\hat{\pi}$ are such that

$$\sum_{t=1}^{T} \hat{\pi}_t^X = 1; \quad \sum_{i=1}^{I} \hat{\pi}_{it}^{\bar{A}X} = 1, \quad \sum_{j=1}^{J} \hat{\pi}_{jt}^{\bar{B}X} = 1, \quad \sum_{k=1}^{K} \hat{\pi}_{kt}^{\bar{C}X} = 1, \quad \sum_{l=1}^{L} \hat{\pi}_{lt}^{\bar{D}X} = 1. \tag{14}$$

Compare (14) with (3). Because of this, we can simplify the iterative procedure described above by estimating only $T-1$ of the π_t^X, say π_t^X for $t = 1, \ldots, T-1$, using (12) and by estimating only $I-1$ of the $\pi_{it}^{\bar{A}X}$, say, $\pi_{it}^{\bar{A}X}$ for $i = 1, \ldots, I-1$, using (13a), etc. The estimate of π_T^X can be obtained using (14); similarly for the estimate of $\pi_{It}^{\bar{A}X}$, etc.

The iterative procedure described above can be used when the manifest variables (A, B, C, D) are polytomous as well as in the special case when they are dichotomous. The earlier literature on maximum likelihood estimation in the latent class model (McHugh, 1956, 1958) dealt only with the special dichotomous case. In this case our procedure is easier to apply than are the formulae of McHugh.

We have concentrated on the case of four manifest variables (A, B, C, D). All the methods and results can be extended when there are m manifest variables ($m = 3, 4, \ldots$).

For the m-way contingency table, when the m manifest variables are dichotomous and $T \leqslant \frac{1}{2}(m+1)$, a determinantal method is available for calculating consistent estimates of the vector π of parameters in the latent class model under certain conditions; see Anderson (1954) and Lazarsfeld & Henry (1968, chapter 4).

The estimates so obtained are not asymptotically efficient (Anderson, 1959), except in the special case where $m = 3$ and $T = 2$. Even when the conditions specified in the earlier literature are satisfied, the determinantal method can yield estimates of π that are not permissible, e.g. where one or more of the components in the estimate of π lie outside the interval [0, 1]. When the components in this estimate all lie within the interval (0, 1), this estimate can be used as the initial trial value for the maximum likelihood estimate $\hat{\pi}$ in the iterative procedure described earlier in this section.

When the components in the initial trial value for $\hat{\pi}$ all lie within the interval (0, 1), the above iterative procedure will converge and yield a solution to (12) and (13a)–(13d). This solution will be either the maximum likelihood estimate $\hat{\pi}$, or some other solution to this system of equations, e.g. a terminal maximum, in which one or more of the components are 0 or 1. By trying various initial trial values for $\hat{\pi}$, we can compare the solutions obtained using the iterative procedure to see which solution minimizes the chi-squared statistic based upon the likelihood ratio

$$X^2 = 2 \sum_{i,j,k,l} f_{ijkl} \log (f_{ijkl}/\hat{F}_{ijkl}), \tag{15}$$

where

$$f_{ijkl} = np_{ijkl}, \quad \hat{F}_{ijkl} = n\hat{\pi}_{ijkl}, \tag{16}$$

with $\hat{\pi}_{ijkl}$ obtained from (9). The solution obtained which minimizes (15) yields the maximum likelihood estimate $\hat{\pi}$. For related comments, see Goodman (1974).

The procedure described above calculated $\hat{\pi}_{ijkl}$ from the vector $\hat{\pi}$, using (10) and (9). Having thus obtained $\hat{\pi}_{ijkl}$, we consider next whether the vector $\hat{\pi}$ is uniquely determined by the $\hat{\pi}_{ijkl}$. If $\hat{\pi}$ is uniquely determined, we say it is identifiable. If $\hat{\pi}$ is uniquely determined by the $\hat{\pi}_{ijkl}$ within some neighbourhood of π, we say it is locally identifiable. We now give a useful sufficient condition for local identifiability.

In the earlier literature, identifiability, or local identifiability, was defined with respect to the vector π of parameters in the model for the π_{ijkl} rather than with respect to the corresponding maximum likelihood estimate $\hat{\pi}$ determined by the $\hat{\pi}_{ijkl}$. The method which we shall now present can be used to study whether π is locally identifiable and/or whether $\hat{\pi}$ is locally identifiable. The former problem will be discussed first.

On combining (1) and (2), our model is

$$\pi_{ijkl} = \sum_{t=1}^{T} \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X}, \tag{17}$$

Formula (17) describes a set of $IJKL$ functions that transform the parameters

$$(\pi_t^X, \pi_{it}^{\bar{A}X}, \pi_{jt}^{\bar{B}X}, \pi_{kt}^{\bar{C}X}, \pi_{lt}^{\bar{D}X})$$

into the π_{ijkl}. Because of (3), we need consider only $T-1$ of the π_t^X, say,

$$\pi_t^X \quad (t = 1, \ldots, T-1),$$

only $I-1$ of the $\pi_{it}^{\bar{A}X}$, say, $\pi_{it}^{\bar{A}X}$ for $i = 1, \ldots, I-1$, etc. The value of π_T^X can be obtained using (3); similarly for the value of π_{It}^{AX}, etc. Thus we need consider only

$$T-1+(I+J+K+L-4)\,T = (I+J+K+L-3)\,T-1$$

parameters. We shall call this set of parameters a 'basic set'. Similarly, since

$$\sum_{i,j,k,l} \pi_{ijkl} = 1, \tag{18}$$

we need consider only $IJKL-1$ of the π_{ijkl}, say, π_{ijkl} for $(i,j,k,l) \neq (I,J,K,L)$. We shall also call this set of π_{ijkl} a basic set. When

$$IJKL < (I+J+K+L-3)\,T, \tag{19}$$

the number of parameters in the basic set exceeds the corresponding number of π_{ijkl}, and so the parameters will not be identifiable in this case.

Next suppose that (19) is not satisfied, i.e. that the number of parameters in the basic set does not exceed the corresponding number of π_{ijkl}. In this case, for each π_{ijkl} in the basic set, we calculate the derivative of the function π_{ijkl} described by (17) with respect to the parameters in the basic set, thus obtaining a matrix consisting of $IJKL-1$ rows and $(I+J+K+L-3)\,T-1$ columns. For example, in the column pertaining to the derivative with respect to π_t^X,

$$\frac{\partial \pi_{ijkl}}{\partial \pi_t^X} = \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X} - \pi_{iT}^{\bar{A}X} \pi_{jT}^{\bar{B}X} \pi_{kT}^{\bar{C}X} \pi_{lT}^{\bar{D}X}, \tag{20}$$

for $t = 1, \ldots, T-1$; in the column pertaining to the derivative with respect to $\pi_{st}^{\bar{A}X}$,

$$\frac{\partial \pi_{ijkl}}{\partial \pi_{st}^{\bar{A}X}} = \begin{cases} \pi_t^X \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X} & (i = s), \\ -\pi_t^X \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X} & (i = I), \\ 0 & \text{otherwise}, \end{cases} \tag{21}$$

for $s = 1, \ldots, I-1$, etc. The second term on the right-hand side of (20), and the term on the second line on the right-hand side of (21), arise since

$$\pi_T^X = 1 - \sum_{t=1}^{T-1} \pi_t^X, \quad \pi_{It}^{\bar{A}X} = 1 - \sum_{i=1}^{I-1} \pi_{it}^{\bar{A}X};$$

see (3). By direct extension of a standard result about Jacobians, the parameters in the model will be locally identifiable if the rank of the matrix described above is equal to the number of columns, i.e. the number of parameters in the basic set. For a corresponding result in the special case where the variables are dichotomous, see McHugh (1956).

By replacing the π's by the corresponding $\hat{\pi}$'s in the preceding two paragraphs, the results can be applied to determine whether the maximum likelihood estimates of the parameters in the model are locally identifiable.

To test the null hypothesis that the T-class latent structure (17) is true, we can use the chi-squared statistic (15). Under this null hypothesis, the asymptotic distribution of the statistic (15) will be chi-squared with degrees of freedom

$$IJKL - 1 - \{(I+J+K+L-3)T - 1\} = IJKL - (I+J+K+L-3)T, \tag{22}$$

when the parameters in the latent structure are locally identifiable. When the parameters in the structure are not locally identifiable, various kinds of restrictions can be imposed upon the parameters in order to make them so. This would lead us to the analysis of restricted latent structures, rather than the unrestricted latent class model considered in the present section. We shall discuss various kinds of restricted structures later (§§4–6), but first we illustrate the application of the above techniques.

3. An example

Table 1 is concerned with whether respondents tend toward universalistic or particularistic values. Stouffer & Toby (1951, 1962, 1963) analyzed these data using a particular 5-class restricted latent structure, and they concluded that the underlying latent variable pertaining to universalistic versus particularistic values could be described by the 5 latent classes. In contrast to this conclusion, we shall show here that a much simpler model is congruent with these data.

Let H_1 denote the 2-class latent structure described in the preceding section, i.e. $T = 2$. To test the hypothesis that H_1 is true, we use the methods in that section to calculate the chi-squared statistic (15). It is 2·720 with 6 degrees of freedom. Thus this simple model fits the data very well indeed.

Table 2 gives both the likelihood ratio chi-squared (15) and the corresponding goodness-of-fit chi-squared for H_1 and for some other latent structure applied to Table 1; Table 3 gives the corresponding estimates obtained with the procedure introduced in §2. Having

Table 2. *Chi-squared values for some latent structures applied to Table* 1

Models	Number of latent classes	Degrees of freedom	Likelihood ratio chi-squared	Goodness-of-fit chi-squared
H_1	2	6	2·720	2·720
H_2, H_2', H_2''	3	2	0·387	0·423
H_3	4	0	0·000	0·000
H_4, H_4', H_4'', H_4''', H_4^{iv}	3	5	0·921	0·895
H_5	3	9	2·281	2·282
H_6	4	4	0·870	0·852
H_7	2	8	2·886	2·838
H_8	2	10	4·390	4·339
H_9	3	10	2·391	2·421

obtained such a good fit with H_1, we could stop with this model; but for purposes of comparison and illustration, we shall later consider the other models in Tables 2 and 3.

Model H_1 for Table 1 states that (*a*) there is a single latent dichotomous variable X pertaining to 'universalistic versus particularistic latent values', and (*b*) this latent variable alone can explain the observed relationships among all four manifest variables (A, B, C, D) in the table. From the estimated parameters for H_1 in Table 3, we see that, with respect to the joint manifest variable (A, B, C, D), the 'modal levels' are (1, 1, 1, 1) and (1, 2, 2, 2), for latent classes 1 and 2, respectively; and the latter latent class is modal, since $\hat{\pi}_2^X = 0{\cdot}721$. Thus, individuals in the modal latent class tend to be at manifest level 2, i.e. 'intrinsically' particularistic, except for the level on variable A, and individuals in the nonmodal latent class tend to be at manifest level 1, i.e. 'intrinsically' universalistic. The $\hat{\pi}_t^X$ for H_1 in Table 3 estimate the distribution of the latent variable X; and the other estimated parameters for this model can be used to estimate the effect of variable X upon each manifest variable (A, B, C, D).

The method used in the preceding paragraph to describe H_1 and its latent variable X can be applied in a similar way to the other models in Table 3; for some of these latent structures further insight into their meaning will be obtained by other means as well. We shall consider these latent structures here, in part in order to illustrate various problems that arise when determining whether the parameters in the model are identifiable, when moving from the simple 2-class model H_1 to the T-class models ($T = 3, 4, \ldots$), when moving from unrestricted to restricted latent structures, etc.

With respect to the estimated parameters in the 2-class model H_1, which are maximum likelihood estimates, the fact that they are locally identifiable can be established using either the method presented in the preceding section or a different determinantal method presented in the earlier literature. The latter method provides sufficient conditions for identifiability in the special case where the m manifest variables are dichotomous and $T \leqslant \frac{1}{2}(m+1)$ (see, e.g. Anderson, 1954), and these conditions were extended by Madansky (1960) to the dichotomous case where $T \leqslant 2^{\frac{1}{2}(m-1)}$. However, when considering the 3-class model H_2 in Table 3, the determinantal method in the earlier literature is not applicable, since in this case $T = 3 > 2^{\frac{3}{2}}$. Fortunately, we can apply to H_2 the method presented in the preceding section for studying local identifiability, and we find that H_2 has 14 parameters in its basic set, but the 15×14 matrix described there has rank 13. This model is not identifiable; models H_2' and H_2'' in Table 3 will produce the same estimated expected values $\hat{\pi}_{ijkl}$ as does H_2. Models H_2, H_2' and H_2'' produce the same $\hat{\pi}_{ijkl}$, but they provide different solutions

Table 3. *Estimated parameters in some latent structures applied to Table* 1

Model	Latent class t	$\hat{\pi}_t^X$	$\hat{\pi}_{1t}^{\bar{A}X}$	$\hat{\pi}_{1t}^{\bar{B}X}$	$\hat{\pi}_{1t}^{\bar{C}X}$	$\hat{\pi}_{1t}^{\bar{D}X}$
H_1	1	0·279	0·993	0·940	0·927	0·769
	2	0·721	0.714	0·330	0·354	0·132
H_2	1	0·208	0·997	0·973	0·986	0·884
	2	0·630	0·824	0·459	0·425	0·183
	3	0·162	0·404	0·052	0·256	0·067
H_2'	1	0·220	0·995	0·968	0·976	0·863
	2	0·672	0·806	0·428	0·407	0·170
	3	0·108	0·288	0·000	0·241	0·057
H_2''	1	0·193	0·998	0·980	1·000	0·913
	2	0·581	0·844	0·499	0·448	0·202
	3	0·226	0·481	0·095	0·269	0·075
H_3	1	0·226	0·999	0·965	0·970	0·852
	2	0·638	0·855	0·412	0·398	0·164
	3	0·102	0·189	0·104	0·148	0·034
	4	0·035	0·049	0·260	0·751	0·291
H_4	1	0·257	0·988	0·940	0·948	0·814
	2	0·627	0·812	0·401	0·364	0·136
	3	0·117	0·252	0·062	0·364	0·136
H_4'	1	0·257	0·988	0·940	0·948	0·814
	2	0·137	0·988	0·769	0·364	0·136
	3	0·607	0·664	0·253	0·364	0·136
H_4''	1	0·257	0·988	0·940	0·948	0·814
	2	0·103	0·988	0·940	0·364	0·136
	3	0·641	0·681	0·253	0·364	0·136
H_4'''	1	0·257	0·988	0·940	0·948	0·814
	2	0·091	1·000	1·000	0·364	0·136
	3	0·652	0·685	0·257	0·364	0·136
H_4^{iv}	1	0·257	0·988	0·940	0·948	0·814
	2	0·067	0·000	0·000	0·364	0·136
	3	0·676	0·796	0·383	0·364	0·136
H_5	1	0·175	1·000	1·000	1·000	1·000
	2	0·050	0·000	0·000	0·000	0·000
	3	0·775	0·796	0·420	0·437	0·175
H_6	1	0·253	0·991	0·953	0·945	0·811
	2	0·096	0·991	0·953	0·361	0·132
	3	0·009	0·684	0·256	0·945	0·811
	4	0·641	0·684	0·256	0·361	0·132
H_7	1	0·279	0·993	0·933	0·933	0·771
	2	0·721	0·714	0·342	0·342	0·132
H_8	1	0·231	0·986	0·986	0·986	0·841
	2	0·769	0·732	0·364	0·364	0·159
H_9	1	0·175	1·000	1·000	1·000	1·000
	2	0·050	0·000	0·000	0·000	0·000
	3	0·775	0·796	0·429	0·429	0·175

to the maximum likelihood equations; H_2' and H_2'' provide extreme solutions, whereas H_2 does not. Indeed, there is a one-dimensional continuum of models ranging between H_2' and H_2'' that will yield the same $\hat{\pi}_{ijkl}$. As $\hat{\pi}_1^X$ decreases from its value in H_2' to its value in H_2'', $\hat{\pi}_2^X$ decreases accordingly and all other estimated parameters increase. Note that $\hat{\pi}_{13}^{\bar{B}X} = 0$ in H_2', and $\hat{\pi}_{11}^{\bar{C}X} = 1$ in H_2''. If there is some need to select from among the various models that yield the same $\hat{\pi}_{ijkl}$, this can be done by the introduction of an *a priori* assumption of the kind considered in the following three paragraphs.

If we had assumed *a priori* that $\pi_{13}^{\bar{B}X} = 0$, and had estimated the other parameters in the model from the data, then there would have been 13, rather than 14, parameters in the basic set of estimated parameters, and they would have been identifiable. In the iterative procedure introduced in § 2, if the initial trial value for $\hat{\pi}_{13}^{\bar{B}X}$ had been taken as zero, all subsequent values of $\hat{\pi}_{13}^{\bar{B}X}$ obtained by (13*b*) would also be zero. The iterative procedure can be applied directly in this case.

The remarks in the preceding paragraph can be applied also in the case where we assume *a priori* that $\pi_{11}^{\bar{C}X} = 1$, and then estimate the other parameters in the model from the data. More generally, if we assume *a priori* that a given set of conditional probabilities, e.g. $\pi_{j3}^{\bar{B}X}$ for $j = 1, \ldots, J$, or $\pi_{k1}^{\bar{C}X}$ for $k = 1, \ldots, K$, is equal to a specified set of zeros and ones which satisfy the corresponding condition (3), then the iterative procedure described herein can be applied directly in this case, simply by taking the assumed values as the initial trial values for the corresponding estimated parameters in the iterative procedure. Still more generally, if a given set of conditional probabilities, e.g. $\pi_{i1}^{\bar{A}X}$ for $i = 1, \ldots, I$, is assumed known, where the assumed values are not necessarily zeros and ones but they must satisfy the corresponding condition (3), then the iterative procedure would be modified by replacing the corresponding iterative calculation, e.g. (13*a*) for $\hat{\pi}_{i1}^{\bar{A}X}$, by the assumed values of the corresponding parameters at each iteration.

In the analysis of Table 1 using a 3-class latent structure, if we introduce a single restriction of the kind described above, namely that a given conditional probability is equal to a specified number in the closed interval defined by the corresponding values in H_2' and H_2'', say that $\pi_{13}^{\bar{B}X} = 0$ or that $\pi_{11}^{\bar{C}X} = 1$, then the chi-squared statistic (15) yields a value of 0·387 with 2 degrees of freedom as noted in Table 2. The number of degrees of freedom is equal to the number of π_{ijkl} in the basic set minus the corresponding number of estimated parameters in this case.

We have now discussed the case where $T = 2$ and $T = 3$. With respect to the case where $T = 4$, condition (19) is satisfied, since $16 < 5 \times 4$, and so the parameters in the 4-class model will not be identifiable. When the m manifest variables are dichotomous, condition (19) can be replaced by the condition that $2^m < (m+1)T$. Model H_3 in Table 3 is an example of such a model. In the analysis of Table 1, model H_3 yields a chi-squared value of zero, as noted in Table 2. Since H_3 was a 'super-saturated' model, i.e. the number of parameters in the basic set for this model exceeds the corresponding number of π_{ijkl}, the chi-squared statistic will have zero degrees of freedom.

From Table 3 we see that $\hat{\pi}_4^X$ in H_3 is relatively small, as $\hat{\pi}_4^X = 0{\cdot}035$. Since H_3 fits the data perfectly, it would be desirable to consider also 3-class models obtained by the deletion of latent class 4 from H_3. In the estimation procedure of § 2 for 3-class models, as the components in the initial trial value for $\hat{\pi}$ we can use the estimated parameters in H_3 modified by the deletion of its fourth latent class, or by the absorption of this latent class in one or more of the other 3 classes. When this is done, we are led to 3-class models of the kind described earlier in this section, e.g. H_2, and also to 3-class models that provide terminal maxima rather than a global maximum.

Initial trial values in our estimation procedure for a 3-class model can be obtained either by (a) modifying the estimated parameters in 4-class models as indicated in the preceding paragraph, (b) modifying the estimated parameters in 2-class models, e.g. by inserting a third latent class in model H_1, (c) trial and error, or (d) using as initial values the estimated parameters in 3-class restricted latent structures of the kind discussed in §5, and removing the restrictions from the structures. The above remarks about initial values for 3-class models can be directly extended to T-class models for $T \geqslant 3$.

We have not yet discussed models H_4 to H_9 in Tables 2 and 3. These models will be discussed in §5 as particular examples of the kinds of restricted latent structures considered below.

4. Some restricted latent structures

As noted in §3, if a given set of conditional probabilities, e.g. $\pi_{i1}^{\bar{A}X}$ for $i = 1, \ldots, I$, is assumed known, then the estimation procedure introduced in §2 would be modified accordingly. Similarly, in the case where the set of probabilities π_t^X for $t = 1, \ldots, T$ is assumed known, this estimation procedure would be modified by using the known values of the π_t^X rather than $\hat{\pi}_t^X$ in the calculation of the $\hat{\pi}_{ijklt}^{ABCD\bar{X}}$ at each iteration, see (10), (9) and (11); but in order to ensure condition (14), the denominator on the right-hand side of (13a)–(13d) remains $\hat{\pi}_t^X$ defined by (12).

The estimation procedure introduced in §2 can also be modified in a straightforward way to accommodate the following kinds of T-class restricted latent structures.

(i) Models in which the following kind of condition is imposed upon the parameters:

$$\pi_{i1}^{\bar{A}X} = \pi_{i2}^{\bar{A}X} \quad (i = 1, \ldots, I). \tag{23}$$

(ii) More generally, models in which the T latent classes can be partitioned into α mutually exclusive and exhaustive subsets $\mathscr{T}_1^A, \ldots, \mathscr{T}_\alpha^A$, where $\alpha \leqslant T$, and/or into β mutually exclusive and exhaustive subsets $\mathscr{T}_1^B, \ldots, \mathscr{T}_\beta^B$, where $\beta \leqslant T$, such that

$$\pi_{it}^{\bar{A}X} = \pi_{it'}^{\bar{A}X} \quad (t, t' \in \mathscr{T}_a^A), \qquad \pi_{jt}^{\bar{B}X} = \pi_{jt'}^{\bar{B}X} \quad (t, t' \in \mathscr{T}_b^B), \tag{24}$$

where $a = 1, \ldots, \alpha$; $b = 1, \ldots, \beta$; $i = 1, \ldots, I$; $j = 1, \ldots, J$. Compare (24) with (23).

(iii) Models in which, in addition to condition (24), the following kind of condition is satisfied for certain specified pairs of subscripts, say (a, b), (a, a^*) and/or (b, b^*):

$$\pi_{it}^{\bar{A}X} = \pi_{jt'}^{\bar{B}X} \quad (t \in \mathscr{T}_a^B,\ t' \in \mathscr{T}_b^B);$$

$$\pi_{it}^{\bar{A}X} = \pi_{i^*t^*}^{\bar{A}X} \quad (t \in \mathscr{T}_a^A,\ t^* \in \mathscr{T}_{a^*}^A); \qquad \pi_{jt}^{\bar{B}X} = \pi_{j^*t^*}^{\bar{B}X} \quad (t \in \mathscr{T}_b^B,\ t^* \in \mathscr{T}_{b^*}^B), \tag{25}$$

where $i = 1, \ldots, I$, where there is a one-to-one correspondence between i and j, between i and i^*, and between j and j^*.

(iv) More generally, models in which the kinds of conditions described by (24) and (25), which were expressed in terms pertaining to variables A and B, are extended to other subsets of the m manifest variables in the m-way contingency table.

The above kinds of restricted latent structures are useful in the analysis of Table 1 and other contingency tables, e.g. Table 4 below, and they will be discussed further in §5 and examples given there.

To determine whether the estimated parameters in a restricted latent structure are locally identifiable, we can use a modified form of the method presented in §2; see (20) and (21). For example, if restriction (24) is imposed, then the T columns pertaining to the derivative with respect to $\pi_{st}^{\bar{A}X}$ $(t = 1, \ldots, T)$, which we described by (21), would be replaced by α columns, where the ath column $(a = 1, \ldots, \alpha)$ is the sum of the corresponding columns obtained from (21) pertaining to the derivative with respect to $\pi_{st}^{\bar{A}X}$ for $t \in T_a^A$; and a similar kind of replacement would be made pertaining to the derivative with respect to $\pi_{st}^{\bar{B}X}$. If the T classes are partitioned into α, β, γ and δ subsets with respect to variables A, B, C and D, such that conditions corresponding to (24) are satisfied for each of these variables, then the number of columns in the matrix described by (20) and (21) will be reduced from $T-1+(I+J+K+L-4)T$ to $T-1+(I-1)\alpha+(J-1)\beta+(K-1)\gamma+(L-1)\delta$. The parameters in the restricted latent structure will be locally identifiable if the rank of the modified matrix is equal to its number of columns.

Under the null hypothesis that the restricted latent structure is true, if the parameters in the latent structure are locally identifiable, then the asymptotic distribution of the statistic (15) will be chi-squared with degrees of freedom equal to one less than the difference between the number of cells in the 4-way table, or m-way table, and the number of columns in the modified matrix described above. Compare this with (22); the number of degrees of freedom is equal to the number of π_{ijkl} in the basic set minus the corresponding number of independent parameters estimated under the model, when the parameters are locally identifiable.

We shall next describe some simple kinds of restrictions that would perforce make the parameters in the latent structure unidentifiable. Consider first the case where the T-class model is such that

$$\pi_{i1}^{\bar{A}X} = \pi_{i2}^{\bar{A}X}, \quad \pi_{j1}^{\bar{B}X} = \pi_{j2}^{\bar{B}X}, \quad \pi_{k1}^{\bar{C}X} = \pi_{k2}^{\bar{C}X}, \quad \pi_{l1}^{\bar{D}X} = \pi_{l2}^{\bar{D}X}. \tag{26}$$

In this case, formula (17) can be replaced by

$$\pi_{ijkl} = \sum_{t=2}^{T} \Theta_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X}, \tag{27}$$

where

$$\Theta_t^X = \begin{cases} \pi_1^X + \pi_2^X & (t = 2), \\ \pi_t^X & (t = 3, \ldots, T). \end{cases} \tag{28}$$

Thus, we can collapse latent classes (1) and (2) to obtain an equivalent latent structure having $T-1$ classes rather than T classes; and the parameters π_1^X and π_2^X will not be identifiable unless additional restrictions, other than condition (3), are imposed upon them.

Consider next the case where the T-class model is such that

$$\pi_{j1}^{\bar{B}X} = \pi_{j2}^{\bar{B}X}, \quad \pi_{k1}^{\bar{C}X} = \pi_{k2}^{\bar{C}X}, \quad \pi_{l1}^{\bar{D}X} = \pi_{l2}^{\bar{D}X}. \tag{29}$$

In this case, formula (17) can be replaced by

$$\pi_{ijkl} = \sum_{t=2}^{T} \Theta_t^X \Theta_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X}, \tag{30}$$

where Θ_t^X is defined by (28) and

$$\Theta_{it}^{\bar{A}X} = \begin{cases} (\pi_1^X \pi_{i1}^{\bar{A}X} + \pi_2^X \pi_{i2}^{\bar{A}X})/\Theta_2^X & (t = 2), \\ \pi_{it}^{\bar{A}X} & (t = 3, \ldots, T). \end{cases} \tag{31}$$

Thus, here too we can collapse classes (1) and (2); and the parameters $\pi_1^X, \pi_2^X, \pi_{i1}^{\bar{A}X}, \pi_{i2}^{\bar{A}X}$ will not be identifiable unless additional restrictions, other than condition (3), are imposed upon them

Consider next the case where the T-class model is such that

$$\pi_{k1}^{\bar{C}X} = \pi_{k2}^{\bar{C}X}, \quad \pi_{l1}^{\bar{D}X} = \pi_{l2}^{\bar{D}X}. \tag{32}$$

In this case, (17) can be replaced by

$$\pi_{ijkl} = \sum_{t=1}^{T} \Theta_t^X \Theta_{it}^{\bar{A}X} \Theta_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X} \pi_{lt}^{\bar{D}X}, \tag{33}$$

where the Θ's are equal to the corresponding π's for $t = 3, \ldots, T$; and where

$$\sum_{t=1}^{2} \Theta_t^X \Theta_{it}^{\bar{A}X} \Theta_{jt}^{\bar{B}X} = \sum_{t=1}^{2} \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X}, \tag{34}$$

for $i = 1, \ldots, I; j = 1, \ldots, J$. When $T > 2$, (34) imposes IJ restrictions on the Θ's; and because of (3) the number of Θ's that we need to consider is $2(I+J-2)+2 = 2(I+J-1)$. The $\Theta_{it}^{\bar{A}X}$ and $\Theta_{jt}^{\bar{B}X}$ are required to satisfy the same kind of condition as (3). Thus, the number of restrictions on the Θ's will be less than the number of Θ's that we need to consider whenever $IJ < 2(I+J-1)$, which will be the case whenever variable A or B is dichotomous or when both these variables are trichotomous. In these cases, the Θ's or the corresponding π's will not be identifiable unless additional conditions are imposed upon them.

The remarks in the preceding three paragraphs can be directly generalized to the m-way table. Conditions (26), (29) and (32) can be expressed more generally as follows: the corresponding conditional probabilities in latent classes 1 and 2 are equal for each of the m variables, see (26), or for $m-1$ of these variables, see (29), or for $m-2$ of these variables, see (32). Still more generally, we can consider the case where the corresponding conditional probabilities are equal in a given subset of the T latent classes, for the $m, m-1$, or $m-2$ variables.

5. Some applications of restricted models

We return now to the analysis of Table 1. Examination of the estimated parameters in Table 3 for model H_2, and/or H_2' and H_2'', would lead us to consider a number of restricted models of the kind presented in the preceding section. For example, if we impose the restriction that

$$\pi_{12}^{\bar{D}X} = \pi_{13}^{\bar{D}X}, \quad \pi_{12}^{\bar{C}X} = \pi_{13}^{\bar{C}X}, \tag{35}$$

then the chi-squared statistic (15) yields a value of 0·921. Model H_4 in Table 3 is an example of a restricted model that will yield this chi-squared value. Since (35) is the same kind of restriction as (32), model H_4 will not be identifiable. In this case, the number of Θ's that we need to consider in (34) exceeds by 2 the number of restrictions on the Θ's. If the matrix method presented in §4 for studying local identifiability is applied to H_4, we find that the 15×12 matrix described there has rank 10. Examination of the estimated parameters in H_4, and/or in some other models that yield the same $\hat{\pi}_{ijkl}$, would lead us to consider imposing the additional restriction that

$$\pi_{11}^{\bar{A}X} = \pi_{12}^{\bar{A}X}. \tag{36}$$

Model H_4' in Table 3 is an example of a model that satisfies (35) and (36). Examination of the estimated parameters in H_4' leads us to impose the additional restriction that

$$\pi_{11}^{\bar{B}X} = \pi_{12}^{\bar{B}X}, \tag{37}$$

which yields model H_4''. The estimated parameters in H_4'' are identifiable. For this model, the chi-squared (15) yields a value of 0·921 with 5 degrees of freedom. The $\hat{\pi}_{ijkl}$ for H_4, H_4' and H_4'' are the same.

Model H_4'' is a simple generalization of the 2-class model. In H_4'', latent class 2 is inserted between the other two classes. Latent class 2 is the same as latent class 1 with respect to two variables, A and B, and it is the same as latent class 3 with respect to the other two variables, C and D. Model H_4'' has one more parameter, and one less degree of freedom, than H_1.

If we had limited ourselves to models in which the parameters are identifiable, we could have begun our exploration with H_1, and then considered models of the H_4'' type, i.e. models in which latent class 2 is the same as latent class 1 with respect to two variables, e.g. A and B, or A and C, or A and D, and is the same as latent class 3 with respect to the other two variables.

The method presented in the earlier literature for obtaining consistent estimates of the parameters in H_1 can be extended as follows to obtain consistent estimates of the parameters in H_4'', which can be used as the components in the initial trial value for $\hat{\pi}$, under H_4'', in our maximum likelihood estimation procedure. Under H_4'', the latent structure for the 3-way marginal table $\{ABC\}$, obtained by ignoring variable D, can be expressed as

$$\pi_{ijk} = \sum_{t=1}^{3} \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X}, \tag{38}$$

where

$$\pi_{i1}^{\bar{A}X} = \pi_{i2}^{\bar{A}X}, \quad \pi_{j1}^{\bar{B}X} = \pi_{j2}^{\bar{B}X}. \tag{39}$$

Restriction (39) in the 3-way table is the same kind of restriction as (29) in the 4-way table; i.e. the corresponding conditional probabilities in latent classes 1 and 2 are equal for $m-1$ variables in the m-way table. Thus latent classes 1 and 2 can be collapsed to obtain a 2-class model for table $\{ABC\}$. By estimating the parameters in the 2-class model, we obtain consistent estimates of the following parameters in the 3-class model H_4'': π_3^X, $\pi_{13}^{\bar{C}X}$, and also $\pi_{1t}^{\bar{A}X}$ and $\pi_{1t}^{\bar{B}X}$ $(t = 1, 2, 3)$. A similar analysis of the 3-way marginal table $\{BCD\}$ yields consistent estimates of π_1^X, $\pi_{11}^{\bar{B}X}$, and also $\pi_{1t}^{\bar{C}X}$ and $\pi_{1t}^{\bar{D}X}$ $(t = 1, 2, 3)$. The consistent estimates thus obtained from $\{ABC\}$ and $\{BCD\}$ can be used to provide consistent estimates for all of the parameters in the 3-class model H_4'' for the 4-way table $\{ABCD\}$. Instead of using $\{ABC\}$ we could have used $\{ABD\}$; similarly, instead of $\{BCD\}$ we could have used $\{ACD\}$.

The method described in the preceding paragraph can be directly extended to more general restricted models for the m-way table (Goodman, 1974).

We arrived at the identifiable model H_4'' considered above by imposing restrictions (36) and (37) upon H_4. Other kinds of restrictions could have been imposed upon H_4 to obtain other identifiable models that would leave $\hat{\pi}_{ijkl}$ unchanged. See, for example, models H_4''' and H_4^{iv} in Table 3, which impose upon H_4 restrictions (40a) and (40b), respectively, for

$t = 2$. In addition to condition (35) imposed by H_4, the imposition of any one of the following restrictions would still leave $\hat{\pi}_{ijkl}$ unchanged

$$(\pi_{1t}^{\bar{A}X}, \pi_{1t}^{\bar{B}X}) = (1, 1) \quad (t = 2 \quad \text{or} \quad t = 3), \tag{40a}$$

$$(\pi_{1t}^{\bar{A}X}, \pi_{1t}^{\bar{B}X}) = (0, 0) \quad (t = 2 \quad \text{or} \quad t = 3), \tag{40b}$$

$$(\pi_{1t'}^{\bar{A}X}, \pi_{1t''}^{\bar{A}X}) = (1, 0) \quad ((t', t'') = (2, 3) \quad \text{or} \quad (3, 2)), \tag{40c}$$

$$(\pi_{1t'}^{\bar{B}X}, \pi_{1t''}^{\bar{B}X}) = (1, 0) \quad ((t', t'') = (2, 3) \quad \text{or} \quad (3, 2)). \tag{40d}$$

More generally, any set of Θ's that satisfy (34), and that are in the closed interval $[0, 1]$, would leave $\hat{\pi}_{ijkl}$ unchanged; and these Θ's can be used by assuming that $\pi_{12}^{\bar{A}X}$ and $\pi_{12}^{\bar{B}X}$, or any two conditional probabilities from among the $\pi_{1t}^{\bar{A}X}$ and $\pi_{1t}^{\bar{B}X}$, for $t = 2, 3$, would be equal to the corresponding two Θ's.

Model H_4'' had four restrictions imposed upon it, see (35)–(37), and the removal of any one of the four restrictions, e.g. (37) as in H_4', would leave $\hat{\pi}_{ijkl}$ unchanged. Furthermore, the removal of two particular restrictions, namely (36) and (37) as in H_4, or the two restrictions included in (35) would also leave $\hat{\pi}_{ijkl}$ unchanged. If restrictions (36) and (37) are imposed but the two restrictions in (35) are not, then the imposition of any one of the sets of restrictions corresponding to (40a)–(40d), with A and B replaced by C and D in these formulae, would also leave $\hat{\pi}_{ijkl}$ unchanged. A generalization corresponding to the one at the end of the preceding paragraph can be made here too.

Examination of the estimated parameters in H_4^{iv} leads us to consider next model H_5 in Table 3, in which the following conditions are imposed upon the 3-class latent-class model

$$\pi_{11}^{\bar{A}X} = \pi_{11}^{\bar{B}X} = \pi_{11}^{\bar{C}X} = \pi_{11}^{\bar{D}X} = 1, \quad \pi_{12}^{\bar{A}X} = \pi_{12}^{\bar{B}X} = \pi_{12}^{\bar{C}X} = \pi_{12}^{\bar{D}X} = 0. \tag{41}$$

Examination of the estimated parameters in either H_2, H_2', H_2'' or H_4, or a more direct examination of Table 1 could also lead to consideration of a model equivalent to H_5, with labelling of latent classes 2 and 3 in H_5 interchanged. Model H_5 states that, with respect to the joint variable (A, B, C, D), individuals in latent classes 1 and 2 will be at levels (1, 1, 1, 1) and (2, 2, 2, 2), respectively, with probability one; and as usual the manifest variables (A, B, C, D) are mutually independent for the individuals in latent class 3. Thus, when levels (1, 1, 1, 1) and (2, 2, 2, 2) are deleted from the 4-way table $\{ABCD\}$, the manifest variables (A, B, C, D) will be quasi-independent under H_5; see, e.g., Goodman (1968). We could have arrived at H_5 by carrying out an analysis of quasi-independence in Table 1 with levels (1, 1, 1, 1) and (2, 2, 2, 2) deleted, followed by the insertion of latent classes 1 and 2 to account for the observed frequencies at levels (1, 1, 1, 1) and (2, 2, 2, 2), respectively.

We consider next model H_6 in Table 3, in which the following conditions are imposed upon the 4-class latent class model

$$\begin{aligned} &\pi_{11}^{\bar{A}X} = \pi_{12}^{\bar{A}X}, \quad \pi_{13}^{\bar{A}X} = \pi_{14}^{\bar{A}X}, \quad \pi_{11}^{\bar{B}X} = \pi_{12}^{\bar{B}X}, \quad \pi_{13}^{\bar{B}X} = \pi_{14}^{\bar{B}X}, \\ &\pi_{11}^{\bar{C}X} = \pi_{13}^{\bar{C}X}, \quad \pi_{12}^{\bar{C}X} = \pi_{14}^{\bar{C}X}, \quad \pi_{11}^{\bar{D}X} = \pi_{13}^{\bar{D}X}, \quad \pi_{12}^{\bar{D}X} = \pi_{14}^{\bar{D}X}. \end{aligned} \tag{42}$$

This model is an extension of H_1 and H_4''. In H_6, latent classes 2 and 3 are inserted between the two latent classes of H_1 which we now call, for convenience, latent classes 1 and 4, with latent class 2 the same as latent class1 with respect to two variables, A and B, and the same as latent class 4 with respect to the other two variables, C and D, and with latent class 3 the same as latent class 4 with respect to the former two variables, A and B, and the same as latent class 1 with respect to the latter two variables, C and D. Model H_6 has two more parameters and two less degrees of freedom than H_1, and one more parameter and one less degree of freedom than H_4''.

We can view latent variable X in H_6 as the joint latent variable (Y, Z), where latent variables Y and Z are dichotomous, and where latent level t in H_6 ($t = 1, 2, 3, 4$) describes the joint latent level (r, s) with respect to (Y, Z), with $(r, s) = (1, 1), (1, 2), (2, 1), (2, 2)$ corresponding to $t = 1, 2, 3, 4$, respectively. Under H_6 in Table 3, variables A and B are affected by the level of latent variable Y but not Z, and variables C and D are affected by the level of latent variable Z but not Y. For further insight into the meaning of H_6 and its latent variables (Y, Z), see the corresponding interpretation of model II, a model of the H_6 type, in §6 below.

Having indicated above with H_6 how to obtain a latent structure containing two latent variables, Y and Z, we can use similar methods to obtain structures containing more than two latent variables. Model H_6, and models similar to it, will be found useful in the analysis of many m-way tables; see e.g., §6 below.

When we introduced H_6 above, it was first described as an extension of H_1, and we introduced H_4'' still earlier in a similar way. The various remarks following the introduction of H_4'' can be directly extended to H_6. As an example of a different identifiable model that yields the same $\hat{\pi}_{ijkl}$ as H_6, we can consider the 2-class model applied to the 3-way table consisting of variables A, B, and the joint variable (C, D), or variables C, D, and the joint variable (A, B), under certain conditions which we omit here to save space (Goodman, 1974).

Model H_6, and some of the other models considered earlier in the section, provide examples of latent structures that impose the kinds of conditions described by (24) expressed in terms pertaining to various subsets of the manifest variables. As examples of models that impose the kinds of conditions described by (25), we consider next H_7, H_8 and H_9 in Table 3, which impose conditions (43), (44) and (45), respectively, upon models H_1, H_7 and H_5, respectively.

$$\pi_{11}^{\bar{B}X} = \pi_{11}^{\bar{C}X}, \quad \pi_{12}^{\bar{B}X} = \pi_{12}^{\bar{C}X}, \tag{43}$$

$$\pi_{11}^{\bar{A}X} = \pi_{11}^{\bar{B}X}, \quad \pi_{11}^{\bar{D}X} = \pi_{22}^{\bar{D}X}, \tag{44}$$

$$\pi_{13}^{\bar{B}X} = \pi_{13}^{\bar{C}X}. \tag{45}$$

Models H_7 and H_8 can be interpreted in the same way as H_1, but there are fewer independent parameters in the former models. Similarly, model H_9 can be interpreted in the same way as H_5, but there are fewer independent parameters in H_9. Note that H_7 and H_9 state that both variables B and C are affected in the same way by the latent variable X; and H_8 states other things as well; see (44). From Table 2 we see that these models fit the data very well indeed.

6. Another example

Table 4 (Coleman, 1964, p. 171) cross-classifies schoolboys interviewed at two successive points in time. Variable A denotes self-perceived membership in the 'leading crowd', in it or out of it, at the time of the first interview, variable B denotes attitude concerning the leading crowd, favourable or unfavourable, expressed at the first interview and variables C and D denote the corresponding membership and attitude at the second interview. For brevity, we consider here only two models: (I) model H_1 in which there is one latent dichotomous variable, and (II) a model of the H_6 type in which there are two latent dichotomous variables, say V and W, where latent variable V can affect manifest variables A and C, and latent variable W can affect manifest variables B and D. The chi-squared values for models I and II are given in Table 5, together with the corresponding estimated parameters. The improvement in fit obtained with model II is dramatic.

Table 4. *Observed cross-classification of* 3398 *schoolboys, in interviews at two successive points in time, with respect to two dichotomous variables:* (1) *self-perceived membership in 'leading crowd', and* (2) *favourableness of attitude concerning the 'leading crowd'*

		Second interview			
	Membership	+	+	−	−
	Attitude	+	−	+	−
First interview					
Membership	Attitude				
+	+	458	140	110	49
+	−	171	182	56	87
−	+	184	75	531	281
−	−	85	97	338	554

Table 5. *Chi-squared values for two latent structures applied to Table* 4 *and the corresponding estimated parameters**

Model	Number of latent variables	Degrees of freedom	Likelihood ratio chi-squared	Goodness-of-fit chi-squared
I	1	6	249·502	251·171
II	2	4	1·270	1·281

Model	Latent class t	$\hat{\pi}^{X}_{t}$	$\hat{\pi}^{\bar{A}X}_{1t}$	$\hat{\pi}^{\bar{B}X}_{1t}$	$\hat{\pi}^{\bar{C}X}_{1t}$	$\hat{\pi}^{DX}_{1t}$
I	1	0·401	0·769	0·645	0·889	0·674
	2	0·599	0·101	0·467	0·090	0·499
II	1	0·272	0·754	0·806	0·910	0·832
	2	0·128	0·754	0·267	0·910	0·302
	3	0·231	0·111	0·806	0·076	0·832
	4	0·368	0·111	0·267	0·076	0·302

* For model II, the symbol X denotes the joint latent variable (V, W), and the four latent classes correspond to the four levels with respect to this joint variable; i.e. (1, 1), (1, 2), (2, 1) and (2, 2), respectively.

Model II for Table 4 states that (*a*) there are two latent dichotomous variables, V and W, pertaining to latent self-perceived membership in the leading crowd and latent attitude concerning the leading crowd, respectively; and (*b*) these two latent variables alone can explain the observed relationships among the manifest variables (A, B, C, D). The relationship between the two latent variables can be estimated using the $\hat{\pi}^{X}_{t}$ for model II in Table 5; the other estimated parameters for this model can be used to estimate the effects of each latent variable upon the corresponding manifest variables. For a comparison of this model with the models presented by the present author in his earlier article (1973) analyzing Table 4 above, see Goodman (1974).

The research was partly supported by the Division of the Social Sciences, National Science Foundation. For helpful comments the author is indebted to D. F. Andrews, O. D. Duncan, S. Haberman and A. Madansky.

References

Anderson, T. W. (1954). On estimation of parameters in latent structure analysis. *Psychometrika* **19**, 1–10.

Anderson, T. W. (1959). Some scaling models and estimation procedures in the latent class model. In *Probability and Statistics*, Ed. U. Grenander, pp. 9–38. New York: Wiley.

Coleman, J. S. (1964). *Introduction to Mathematical Sociology*. Glencoe, Illinois: Free Press.

Goodman, L. A. (1968). The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries. *J. Am. Statist. Assoc.* **63**, 1091–131.

Goodman, L. A. (1973). The analysis of multidimensional contingency tables when some variables are posterior to others: A modified path analysis approach. *Biometrika* **60**, 179–92.

Goodman, L. A. (1974). The analysis of systems of qualitative variables when some of the variables are unobservable. I. A modified latent structure approach. *Am. J. Sociol.* **79**, 1179–259.

Haberman, S. H. (1974). Log-linear models for tables derived by indirect observation: Maximum likelihood equations. *Ann. Statist.* **2**. To appear.

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

Madansky, A. (1960). Determinantal methods in latent class analysis. *Psychometrika* **25**, 183–98.

McHugh, R. B. (1956). Efficient estimation and local identification in latent class analysis. *Psychometrika* **21**, 331–47.

McHugh, R. B. (1958). Note on 'Efficient estimation and local identification in latent class analysis'. *Psychometrika* **23**, 273–4.

Stouffer, S. A. & Toby, J. (1951). Role conflict and personality. *Am. J. Sociol.* **56**, 395–406. Reprinted in S. A. Stouffer (1962), *Social Research to Test Ideas* (New York: Free Press). Reprinted in part in R. M. Riley (1963), *Sociological Research: A Case Approach* (New York: Harcourt, Brace and World).

[*Received May* 1973. *Revised December* 1973]