Log-linear models are useful for analyzing cross-classifications of counts arising in sociology, but it has been argued that in some cases, an alternative approach for formulating models—one based on simultaneously modeling univariate marginal logits and marginal associations—can lead to models that are more directly relevant for addressing the kinds of questions arising in those cases. In this article, the authors explore some of the similarities and differences between the log-linear models approach to modeling categorical data and a marginal modeling approach. It has been noted in past literature that the model of statistical independence is conveniently represented within both approaches to specifying models for cross-classifications of counts. The authors examine further the extent to which the two families of models overlap, as well as some important differences. The authors do not present a complete characterization of the conditions describing the intersection of the two families of models but cover many of the models for bivariate contingency tables and for three-way contingency tables that are routinely used in sociological research.

Specifications of Models for Cross-Classified Counts

Comparisons of the Log-Linear Models and Marginal Models Perspectives

MARK P. BECKER
University of Michigan
SUSAN MINICK
Indiana University
ILSOON YANG
Harvard School of Public Health

1. *INTRODUCTION*

Log-linear models provide a flexible and powerful set of tools for the analysis of cross-classifications of counts. They are routinely used

AUTHORS' NOTE: *We gratefully acknowledge support from the National Institutes of Health (CA53787) and Michael Sobel and Yu Xie for their critical reviews of earlier drafts of this work.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 26 No. 4, May 1998 511-529

to study a wide variety of social phenomena, and in many sociology graduate programs, they are part of the core body of material covered in research methods courses. An appealing feature of these models is that many of the fundamental concepts from linear regression analysis carry over with relatively straightforward modifications. Excellent coverage of the foundations and utility of these models is given in Agresti (1990). It has, however, been argued (e.g., McCullagh and Nelder 1989; Lang and Agresti 1994; Becker 1994; Sobel, Becker, and Minick 1997) that log-linear models are not ideally suited for addressing certain important substantive questions that arise in analyses of repeated measurements, and a better approach to modeling in these situations is based on the simultaneous modeling of associations and univariate marginal distributions (see also Koch et al. 1977; Plewis 1985).

The purpose of this article is to explore some of the similarities and differences between log-linear models and marginal models. It has been noted (see, e.g., Becker 1994) that the model of statistical independence is conveniently represented within both approaches to specifying models for cross-classifications of counts. In this article, we further examine the extent to which the two families of models overlap, as well as some important differences. We do not present a complete characterization of the intersection of the two families of models, but we do cover most of the fundamental models for two-way and three-way contingency tables. In addition, we consider log-multiplicative association (i.e., odds ratio) structures (e.g., Goodman 1979), which takes us outside of the log-linear family of models.

A subtle but critical distinction for comparing log-linear and marginal models is between models and parameterizations. A model is a space of probability distributions, and a parameterization of a model is an equation or system of equations for describing that space. Thus, the log-linear models framework is one approach for parameterizing models, and the marginal models framework is another. However, it is important to note that in the marginal parameterization of a model, there is an implied structure on the cell probabilities, a structure not made explicit in general. This is in contrast to the usual log-linear modeling approach. The two approaches to modeling are distinct in the sense that what is modeled explicitly is different in the two approaches, although the models themselves are not necessarily distinct. There are (1) marginal models that do not have a log-linear model parameterization, (2) log-linear models that do not have a marginal model parameterization, and (3) models that have both a marginal model parameterization and a log-linear model parameterization. In Section 2, we show, using the two-way cross-classifications, that the set of models (3) is nontrivial in the sense that many log-linear models used in statistical practice have marginal model parameterizations. Discussion of an important class of marginal models, those specifying a marginal shift, concludes the section. Three-way cross-classifications are considered in Section 3. The family of hierarchical log-linear models is compared to the family of hierarchical marginal models. It is shown that the intersection of these two families is large, but there are also important areas of nonintersection. Discussion of marginal models for sets of tables concludes the section. We conclude this article with a brief summary.

2. *TWO-WAY CROSS-CLASSIFICATIONS*

Let π_{ij} denote the probability associated with the cell in row i and column j of the $I \times J$ table ($i = 1, \ldots, I$; $j = 1, \ldots, J$), and let $\pi_{\cdot j} = \sum_{i=1}^{I} \pi_{ij}$ and $\pi_{i\cdot} = \sum_{j=1}^{I} \pi_{ij}$ denote the column and row marginal probabilities, respectively. The multinomial sampling model is assumed here. The characters R and C are used to denote the row variable and the column variable, respectively. Note that $I = J$ when there is a one-to-one correspondence between the categories of the row variable and the categories of the column variable.

2.1. *RECTANGULAR TABLES*

For a complete specification of a marginal model for the $I \times J$ table, $(I - 1)$ uniquely identified marginal row logits, $(J - 1)$ uniquely identified marginal column logits, and $(I - 1)(J - 1)$ uniquely identified that log-odds ratios must be specified. There are $(I - 1)(J - 1)$ nonredundant odds ratios for the association between R and C; for example, the local log-odds ratios are

$$\log[(\pi_{ij}\pi_{i+1,j+1})/(\pi_{i,j+1}\pi_{i+1,j})] = \alpha_{ij}^{RC}, i = 1, \cdots, I-1, j = 1, \cdots, J-1.$$

There are multiple definitions for logits when there are more than two categories for a variable. Baseline category logits generally are used for nominal variables, and cumulative logits, adjacent category logits, continuation-ratio logits, or baseline category logits are among the more frequently used logits for ordinal variables. See Agresti (1990) for definitions and discussion of various logits used in modeling categorical responses. For example, if adjacent category logits are used to characterize the marginal distributions, the $(I-1)$ adjacent category logits for the row variable are

$$\log(\pi_{i+1,\cdot}/\pi_{i\cdot}) = \alpha_i^R, i = 1, \cdots, I-1,$$

and the $(J-1)$ adjacent category logits for the column variable are

$$\log(\pi_{\cdot,j+1}/\pi_{\cdot j}) = \alpha_j^C, j = 1, \cdots, J-1.$$

The various formulations for marginal logits lead to equivalent models (i.e., the expected cell counts will be the same across models) if no restrictions are placed on the marginal distributions, but they lead to different models when the marginal distributions are restricted in some way (e.g., see discussion of marginal shift models below). Selecting between the various formulations of logits for ordinal responses is primarily a matter of identifying the parameterization(s) that best facilitates analysis of the substantive questions that the data are being used to address. See Sobel et al. (1997) for some illustrative comparisons.

Together the logits $\alpha_1^R, \cdots, \alpha_{I-1}^R, \alpha_1^C, \cdots, \alpha_{J-1}^C$ and log-odds ratios $\alpha_{11}^{RC}, \cdots, \alpha_{(I-1)(J-1)}^{RC}$ determine the probabilities of the $I \times J$ contingency table (Sinkhorn 1967). A marginal model without any restrictions on the marginal logits or odds ratios is a saturated model and hence equivalent to the log-linear model

$$\log(\pi_{ij}) = \lambda + \lambda_i^R + \lambda_j^C + \lambda_{ij}^{RC},$$

where the λ^R are row effects parameters, the λ^C are column effects parameters, the λ^{RC} characterize association in terms of local log-odds ratios, and suitable identifying restrictions, such as $\sum_i \lambda_i^R = \sum_j \lambda_j^C = \sum_i \lambda_{ij}^{RC} = \sum_j \lambda_{ij}^{RC} = 0$, are assumed. Note that we have specified the marginal model so that no identifying restrictions are necessary. The marginal association parameters α^{RC} are related to the log-linear model association parameters λ^{RC} through the identity $\alpha_{ij}^{RC} = \lambda_{ij}^{RC} + \lambda_{i+1,j+1}^{RC} - \lambda_{i,j+1}^{RC} - \lambda_{i+1,j}^{RC}$. In general, there is not a linear relationship between the marginal model logit parameters, α_i^R and α_j^C, and the log-linear model row and column effect parameters, λ_i^R and λ_j^C (see, e.g., Becker 1990, 1994; Sobel et al. 1997).

The saturated model is but one of many models routinely expressed as a linear model for the logarithm of the cell probabilities π_{ij} that have equivalent parameterizations as marginal models. Other important examples include the models of statistical independence, uniform association, row (*R*) association, column (*C*) association, $R + C$ association, homogeneous $R + C$ association, and quasi-symmetry. In addition, two useful log-nonlinear models are the $R \times C$ association and homogeneous $R \times C$ association models. Finally, models that fit diagonal cells exactly, such as quasi-independence, quasi-uniform association, and the model of agreement plus linear-by-linear interaction (Agresti 1988), can be expressed both as log-linear and as marginal models. With the exception of symmetry models, none of these models places restrictions on the univariate marginal distributions. Models that do not constrain the marginal distributions have marginal model equivalents because the same patterns for associations can be accommodated in either framework for specifying models. Symmetric probability models are also easily accommodated within both approaches to parameterizing models, but in the log-linear case, marginal homogeneity is conveniently handled only under symmetric probabilities. As demonstrated below, the marginal model approach to specifying a model is more flexible in this regard.

Statistical independence. The familiar model of statistical independence is $\pi_{ij} = \pi_{i.}\pi_{.j}$, or $\log(\pi_{ij}) = \lambda + \lambda_i^R + \lambda_j^C$. An equivalent marginal model arises from setting $\alpha_{ij}^{RC} = 0$ for all $i = 1, \ldots, I - 1; j = 1, \ldots, J - 1$. Note that in this case, there is a direct relationship between the univariate marginal logits and the log-linear model parameters that does not hold more generally; for example,

$$\alpha_i^R = \lambda_{i+1}^R - \lambda_i^R.$$

Uniform R, C, R + C, and R × C association. Goodman (1979) considered a family of simple models for the association in cross-classifications having ordered categories. He specified models for the basic set of local odds ratios $\alpha_{ij}^{RC} = (\pi_{ij}\pi_{i+1,j+1})/(\pi_{i,j+1}\pi_{i+1,j})$ and, in particular, the models of uniform association (e.g., $\alpha_{ij}^{RC} = \alpha^{RC}$ for all pairs i, j), row effects association, column effects association, and row + column effects association—which are all log-linear models for the odds ratios—and the row × column effects association. Because these models do not restrict the row and column marginal distributions, each has a marginal model representation arrived at by placing suitable restrictions on the α_{ij}^{RC}:

$$\text{Uniform: } \alpha_{ij}^{RC} = \alpha^{RC} \; i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

$$\text{Row effects: } \alpha_{ij}^{RC} = \alpha_i^{RC} \; i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

$$\text{Column effects: } \alpha_{ij}^{RC} = \alpha_j^{RC} \; i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

$$\text{Row + column effects: } \alpha_{ij}^{RC} = \alpha_{1i}^{RC} + \alpha_{2j}^{RC} \; i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

$$\text{Row} \times \text{column effects: } \alpha_{ij}^{RC} = \phi_{1i}^{RC}\phi_{2j}^{RC} \; i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

2.2. *SQUARE TABLES*

We now consider models for square tables, a setting where explicit modeling of the marginal distributions can be extremely profitable. The following models are restricted forms of models considered above, parameterized to exploit consideration of issues arising when there is a one-to-one correspondence between the row categories and the column categories.

Homogeneous row-column effects models. The first two models we consider are restricted forms of association models with row effects and column effects.

Homogeneous row + column effects: $\alpha_{ij}^{RC} = \alpha_i^{RC} + \alpha_j^{RC} i, j = 1, \cdots, I-1$.

Homogeneous row × column effects: $\alpha_{ij}^{RC} = \phi_i^{RC}\, \phi_j^{RC}\, i, j = 1, \cdots, I-1$.

Quasi-independence. This model is also a restricted form of a model considered above, and in fact it may be formulated for rectangular tables. A frequently employed form of the quasi-independence model is

$$\pi_{ij} = \begin{cases} \gamma_i \nu_j, & i \neq j \\ \gamma_i \nu_j \delta_i, & i = j \end{cases}.$$

Note that the γ_i and ν_j in the quasi-independence model are not the marginal probabilities of the independence model. A log-linear representation of the model is

$$\log(\pi_{ij}) = \lambda + \lambda_i^R + \lambda_j^C + \delta_i I(i = j),$$

where $I(\cdot)$ is the indicator function taking on the value 1 if the condition $i = j$ is satisfied, 0 otherwise. An equivalent marginal model is obtained by imposing the restrictions

$$\alpha_{ij}^{RC} = 0;\ |i - j| > 1$$

$$\alpha_{i,i+1}^{RC} = -\delta_{i+1},\ i = 1, \cdots, I-1$$

$$\alpha_{i+1,i}^{RC} = -\delta_{i+1},\ i = 1, \cdots, I-1$$

$$\alpha_{ii}^{RC} = \delta_i + \delta_{i+1},\ i = 1, \cdots, I-1$$

Symmetry and quasi-symmetry. The model of symmetry is $\pi_{ij} = \pi_{ji}$, with log-linear representation

$$\log(\pi_{ij}) = \lambda + \lambda_i + \lambda_j + \lambda_{ij},$$

where there is a single set of main effect parameters λ_i, as opposed to the two sets λ_i^R and λ_j^C in the preceding models, and $\lambda_{ij} = \lambda_{ji}$ for all i, j. The corresponding marginal model follows by setting $\alpha_i^R = \alpha_i^C$ for all i and $\alpha_{ij}^{RC} = \alpha_{ji}^{RC}$ for all i, j. The constraints on the univariate marginal logits are equivalent to marginal homogeneity (i.e., $\pi_{i.} = \pi_{.i}$ for all i), and the constraints on the local log-odds ratios specify symmetric association (i.e., $\alpha_{ij}^{RC} = \alpha_{ji}^{RC}$ for all i, j). The model of quasi-symmetry (Caussinus 1965) follows by relaxing the restriction of marginal homogeneity. The model of quasi-symmetry has figured prominently in the mobility literature (Sobel, Hout, and Duncan 1985), and more recently, Lang and Eliason (1997) and Sobel et al. (1997) have advocated the use of marginal models for modeling mobility.

The uniform association model described above is actually a special case of the linear × linear interaction model, written in standard log-linear model form:

$$\log(\pi_{ij}) = \lambda + \lambda_i^R + \lambda_j^C + \beta \mu_i \nu_j,$$

where μ_i and ν_j are scores assigned to the row and column categories, respectively. The uniform association model arises when equal interval scores (e.g., $\mu_i = i$, $\nu_j = j$) are assigned. The linear × linear interaction model is useful for ordered row and column categories and is equivalent to the row × column interaction model when the scores are estimated from the data rather than assigned a priori.

We now turn to association models where extra parameters are specified for special features of diagonal cells. In some models, the diagonal cells are fitted exactly, in which case some authors refer to such models as blanking out the diagonal. We choose not to use such language because we think it is important to retain the diagonal cell counts for detailed examination of how they differ from the off-diagonal cells.

For square tables with ordered categories, it is often useful to combine a log-multiplicative association with extra parameters for the diagonal cells. One such model that has figured prominently in sociology is the quasi-uniform association model (Duncan 1979; Goodman 1979). This model combines uniform association with additional parameters for each of the main diagonal cells. The log-linear specification of this model is

$$\log(\pi_{ij}) = \lambda + \lambda_i^R + \lambda_j^C + \beta \mu_i \mu_j + \delta_i I(i = j).$$

Here, note that one set of scores is assigned to the rows and columns. The same principles may be applied to other association models, within either the log-linear or marginal frameworks. The marginal model representation of, for example, a homogeneous row × column association model with diagonal parameters is

$$\alpha_{ij}^{RC} = \phi_i^{RC}\phi_j^{RC};\ |i-j| > 1$$

$$\alpha_{i,i+1}^{RC} = \phi_i^{RC}\phi_j^{RC} - \delta_{i+1}$$

$$\alpha_{i+1,i}^{RC} = \phi_i^{RC}\phi_j^{RC} - \delta_{i+1}$$

$$\alpha_{ii}^{RC} = \phi_i^{RC}\phi_j^{RC} + \delta_i + \delta_{i+1}.$$

Becker (1990) shows that in some cases, a mover-stayer interpretation may by applied to the diagonal cell parameters δ_i. In some cases, it is not necessary that a separate parameter δ_i be specified for each diagonal cell. For example, Agresti (1988) considers the restricted case of quasi-uniform association where $\delta_i = \delta$ for all i in analyzing rater agreement data.

2.3. MARGINAL SHIFTS

All of the examples considered to this point illustrate the extent to which log-linear (and log-nonlinear) models and marginal models are not mutually exclusive of one another. In many cases, both the log-linear and marginal model representations are linear, even though the log-linear (marginal) model parameters are not necessarily linear combinations of the marginal (log-linear) model parameters. There are, in addition, models that have a linear representation in one of the approaches to modeling but not the other. We consider now certain models that are linear in terms of univariate logits and log-odds ratios but not in terms of logarithms of cell probabilities. The opposite situation is demonstrated below for the case of multiway tables.

Intermediate between marginal homogeneity and the general specification of two sets of univariate logits that are distinct from one another is the marginal shift model

$$\alpha_i^C = \alpha_i^R + \alpha.$$

The marginal shift model is appropriate only when the categories are ordered, but for ordered categorical data, it provides a simple yet powerful means for addressing important substantive questions regarding changes or differences in marginal distributions across items or across time. See Lang and Eliason (1997) and Sobel et al. (1997) for discussion of the application of such models to mobility tables. Marginal shift models are in fact models of uniform association but for the $2 \times I$ table of dependent responses formed by combining the row and column marginal distributions rather than for the full $I \times I$ table.

	Category		
Variable	*1*	. . .	*I*
Row	$\pi_{1\cdot}$	. . .	$\pi_{I\cdot}$
Column	$\pi_{\cdot 1}$	. . .	$\pi_{\cdot I}$

That is, for the table, the marginal shift model in terms of adjacent category logits corresponds to a uniform log-odds ratio model

$$\log\left(\frac{\pi_{i\cdot}\pi_{\cdot i+1}}{\pi_{i+1\cdot}\pi_{\cdot i}}\right) = \alpha.$$

Note that shift models may also be specified in terms of continuation-ratio logits or cumulative logits. In fact, if the rows in the $2 \times I$ table above were from independent samples or strata, then the marginal shift model combined with cumulative logits would be equivalent to the standard proportional odds model (see, e.g., Agresti 1990:322).

However, it is important to recognize that, given a model for association, a univariate shift model in terms of one type of logit is not equivalent to a univariate shift model that uses a distinctly different formulation for logits.

3. *THREE-WAY CROSS-CLASSIFICATIONS*

In this section, we do not require a one-to-one correspondence between the categories of the three variables. Relationships between log-linear models and marginal models are explored more generally, with the results then being applicable both to the case of perfectly square tables and to the case of partially square (i.e., a partial matching of the categories between the row and column variables) tables.

The variables cross-classified are denoted by R, C, and L, and they are indexed by i, j, and k, respectively. As a starting point, consider the saturated log-linear model for a three-way cross-classification:

$$\log(\pi_{ijk}) = \lambda + \lambda_i^R + \lambda_j^C + \lambda_k^L + \lambda_{ij}^{RC} + \lambda_{ik}^{RL} + \lambda_{jk}^{CL} + \lambda_{ijk}^{RCL}. \qquad (1)$$

There are $(I - 1)(J - 1)(K - 1)$ degrees of freedom for three-way association, $(I - 1)(J - 1)$ for the association between R and C, $(I - 1)(K - 1)$ for the association between R and L, $(J - 1)(K - 1)$ for the association between C and L, $I - 1$ degrees of freedom for summarizing the R "main effect," $J - 1$ for the C main effect, $K - 1$ for the L main effect, and a single parameter λ pertaining to the multinomial sampling constraint. Note that retention of certain terms is necessary in all reduced models arrived at by setting terms to zero when the three-way table represents a set of independent two-way multinomials. One feature of log-linear modeling is that the interpretation of the parameters retained in a model depends on the parameters that are omitted from the model. The fact that log-linear main effect parameters are linearly related to marginal logit parameters only under independence illustrated this point in the context of two-way tables. For an example in the context of multiway tables, consider that in the model of mutual independence (see, e.g., Agresti 1990:138), the main effect parameters can be interpreted in terms of univariate marginal probabilities or logits, but in the model of no-three-factor interaction (see, e.g., Agresti 1990:148) they cannot.

Becker (1994) presents a generalized log-linear model parameterization for multivariate categorical data where each set of model parameters is in terms of logits, log-odds ratios, or contrasts of log-odds ratios for the lowest order marginal distribution relevant for the particular parameter set. For the three-way table, there are $(I - 1)(J - 1)(K - 1)$ log-odds ratio contrasts for three-factor association, $(I - 1)(J - 1)$ log-odds ratios characterizing the bivariate marginal association between R and C, and so on. Note that the parameters in the marginal model parameterization, unlike those in a log-linear parameterization, have the same interpretation regardless of which terms are set to zero in the model. Just as pairwise association was parameterized the same way for both log-linear and marginal models for the two-way table, three-way association has the same parameterization under log-linear and marginal models for the three-way table. In general, for higher way tables, only the highest order association possible has the same parameterization under both approaches. For example, both models have the same parameterization of three-factor association in a three-way table, but the parameterizations of two-factor associations are different.

3.1. *HIERARCHICAL MODELS*

In the subsequent comparison of marginal models to log-linear models, we adopt a shorthand notation for expressing models (Agresti 1990); for example, the model of mutual independence is *(R, C, L)*, and the model of no-three-factor association is *(RC, RL, CL)*. In addition, we distinguish between log-linear and marginal specifications of models using subscripts to indicate how the so-called univariate effects and associations are parameterized. The subscript *LL* is used to denote log-linear parameterizations, and the subscript *M* is used to denote marginal model parameterizations. Given π_{ijk} as the probability in the *ijk*th cell of the three-way table, we use a subscript of "." to indicate summation over the replaced subscript; for example, $\pi_{ij\cdot} = \sum_{k=1}^{K} \pi_{ijk}$. Then $(RCL)_{LL}$ corresponds to (1), and $(RCL)_M$ corresponds to the following:

Univariate marginal logits:

$$\log(\pi_{i+1,\cdot\cdot}/\pi_{i\cdot\cdot}) = \alpha_i^R,\ i = 1, \cdots, I-1$$

$$\log(\pi_{\cdot,j+1,\cdot}/\pi_{\cdot j\cdot}) = \alpha_j^C,\ j = 1, \cdots, J-1$$

$$\log(\pi_{\cdot\cdot k+1}/\pi_{\cdot\cdot k}) = \alpha_k^L,\ k = 1, \cdots, K-1.$$

Bivariate log-odds ratios:

$$\log[(\pi_{ij\cdot}\pi_{i+1,j+1,\cdot})/(\pi_{i,j+1,\cdot}\pi_{i+1,j,\cdot})] = \alpha_{ij}^{RC},\ i = 1, \cdots, I-1, j = 1, \cdots, J-1$$

$$\log[(\pi_{i\cdot k}\pi_{i+1,\cdot,k+1})/(\pi_{i,\cdot,k+1}\pi_{i+1,\cdot,k})] = \alpha_{ik}^{RC}, i = 1, \cdots, I-1, k = 1, \cdots, K-1$$

$$\log[(\pi_{\cdot jk}\pi_{\cdot j+1,k+1})/(\pi_{\cdot j+1,k}\pi_{\cdot j,k+1})] = \alpha_{jk}^{CL}, j = 1, \cdots, J-1, k = 1, \cdots, K-1$$

Three-factor log-odds ratio contrasts:

$$\log\left(\frac{\pi_{ij,k+1}\pi_{i+1,j+1,k+1}}{\pi_{i,j+1,k+1}\pi_{i+1,j,k+1}}\right) - \log\left(\frac{\pi_{ijk}\pi_{i+1,j+1,k}}{\pi_{i,j+1,k}\pi_{i+1,jk}}\right) = \alpha_{ijk}^{RCL},$$

$$i = 1, \cdots, I-1, j = 1, \cdots, J-1, k = 1, \cdots, K-1.$$

Reduced models follow from setting parameters to 0; for example, $(RC, RL, CL)_{LL}$ follows from $\lambda_{ijk}^{RCL} = 0$ for all i, j, k, and $(RC, RL, CL)_M$ follows from $\alpha_{ijk}^{RCL} = 0$ for all i, j, k.

Since the definitions for pairwise association parameters are different in the log-linear and marginal model parameterizations, there is the question of when the model structures under the log-linear and marginal model parameterizations correspond to the same model and when they correspond to different models. That is, since a model is defined by the space of possible multinomial distributions, the parameterization of a model is not unique, and hence it is of interest to know when and under what conditions a model defined in terms of a log-linear parameterization has an equivalent parameterization in terms of a marginal model. Here we limit our attention to so-called hierarchical models, but the implications for reduced models that constrain the form of associations (e.g., uniform association) are straightforward.

There are nine models in families of hierarchical models for three-way cross-classifications, and they may be grouped according to the degree of association or dependence modeled. The nine models are mutual independence, (R, C, L); models where only a single pair of variables are dependent, (RC, L), (RL, C), (R, CL); models where only a single pair of variables are conditionally (LL) or marginally (M) independent, (RC, CL), (RL, CL), (RC, RL); the model of no-three-factor association, (RC, RL, CL); and the saturated model, (RCL).

An immediate consequence of the sufficient conditions for establishing when marginal and partial associations are identical (see, e.g., Agresti 1990:146) is that the following equivalences hold between log-linear models and marginal models:

$$(R, C, L)_{LL} \equiv (R, C, L)_M$$

$$(RC, L)_{LL} \equiv (RC, L)_M$$

$$(R, CL)_{LL} \equiv (R, CL)_M$$

$$(RL, C)_{LL} \equiv (RL, C)_M.$$

Models in which only a single pair of variables are conditionally independent (i.e., log-linear models) are not equivalent to models in which a single pair of variables are marginally independent; for example, $(RC, RL)_{LL} \neq (RC, RL)_M$. Under $(RC, RL)_{LL}$, the variables C and L are conditionally independent given R, whereas under $(RC, RL)_M$, as the term CL is absent, they are marginally independent. The lack of equivalence between models including exactly two bivariate associations is demonstrated by constructing a 2 × 2 × 2 cross-classification where two of the variables are conditionally independent but not marginally independent; see section 5.2.4 of Agresti (1990) for an example.

If a log-linear model does not satisfy the sufficient conditions for equal marginal and partial associations, does it follow that there is not a marginal model equivalent to that model? In other words, does failure to satisfy the sufficient conditions for collapsibility imply that a log-linear model does not have an equivalent marginal model parameterization? The answer is no. The trivial case is the saturated model, in which the equivalence $(RCL)_{LL} \equiv (RCL)_M$ holds since neither model is restricted in any way. A more interesting case is the model of no-three-factor interaction. The models $(RC, RL, CL)_{LL}$ and $(RC, RL, CL)_M$ are equivalent since in both cases the only constraint specified under the two parameterizations, no-three-factor interaction, is the same under both parameterizations. That is, $\lambda_{ijk}^{RCL} = 0$ for all i, j, k and $\alpha_{ijk}^{RCL} = 0$ for all i, j, k are equivalent constraints. It does not follow that the estimated marginal log-odds ratios are identical to the estimated conditional log-odds ratios in the model of no-three-factor interaction; that is, the αs are not equal to a linear contrast of the λs. Consider the following 2 × 2 × 2 cross-classification as an example.

TABLE 1: No Three-Factor Association Example

	L = 1		L = 2	
	C = 1	C = 2	C = 1	C = 2
$R = 1$	7	4	1	1
$R = 2$	2	8	1	7

The conditional odds ratios for the *RC* association are 7, implying $\lambda_{ijk} = \alpha_{ijk} = 0$, but the marginal odds ratio for the *RC* association is 8. The table must be a perfect table in the sense of Darroch (1962) for equality to hold (see also Agresti 1990).

SETS OF TABLES

An important special case of the three-way cross-classification is the case in which two of the three categorical variables are viewed as responses, and the third is a group or layer variable (see, e.g., Clogg 1982). The relevant sampling model in this case is the product-multinomial where *K* different *IJ*-dimensional multinomials are considered. Assume that the variable *L* is the group variable. In this case, the fully marginal parameterization is not particularly interesting, especially as compared to one in terms of logits for the conditional distribution of *R* given *L*, logits for the the conditional distribution of *C* given *L*, and the log-odds ratios for the association between *R* and *C* given *L*. Let $\pi_{ij|k}$ denote the conditional probability $Pr(R = i, C = j \mid L = k)$, $\pi_{i\cdot|k} = \sum_{j=1}^{J} \pi_{ij|k}$, and $\pi_{\cdot j|k} = \sum_{i=1}^{I} \pi_{ij|k}$. A full specification of the saturated model for a set of K $I \times J$ tables is the following:

Conditional univariate logits:

$$\log(\pi_{i+1,\cdot|k}/\pi_{i\cdot|k}) = \alpha_{i|k}^{R|L}, \; i = 1, \cdots, I-1, \; k = 1, \cdots, K$$

$$\log(\pi_{\cdot,j+1|k}/\pi_{\cdot j|k}) = \alpha_{j|k}^{C|L}, \; j = 1, \cdots, J-1, \; k = 1, \cdots, K.$$

Conditional log-odds ratios:

$$\log[(\pi_{ij|k}\pi_{i+1,j+1|k})/(\pi_{i,j+1|k}\pi_{i+1,j|k})] = \alpha_{ij|k}^{RC|L},$$

$$i = 1, \cdots, I - 1, j = 1, \cdots, J - 1, k = 1, \cdots, K.$$

This model is equivalent to the saturated log-linear model in the absence of any restrictions on the logits and log-odds ratios. All of the models considered above for square two-way tables may be applied across the levels of L when there is a one-to-one correspondence between the categories of R and C to arrive at reduced models of particular interest for square tables. In making comparisons across the levels of L, there are additional reduced models with homogeneity restrictions that can be used to focus attention on similarities and differences between the groups. The general model is one of heterogeneity since the pattern and strength of the associations between R and C are unrestricted across the levels of L, and the marginal distributions of R and C are unrestricted across the levels of L. The use of homogeneity restrictions in the modeling of association across sets of multiway cross-classifications is considered in Clogg (1982) and Becker and Clogg (1989). The general principles covered therein may be applied directly to the models considered here, both for modeling association and for modeling marginal distributions.

Sobel et al. (1997) use the above formulation of marginal models for sets of tables to make comparisons of structural mobility cross-nationally. Using these models, they are able to consider models considered previously by Yamaguchi (1987) and Xie (1992), as well as new models employing marginal logits for an explicit parameterization of structural mobility. As demonstrated above, most of the association structures employed in the usual log-linear and log-multiplicative frameworks are easily incorporated into the marginal modeling framework. Thus, Sobel et al. were able to move beyond the traditional frameworks for mobility, retaining all of the desirable features of those approaches. Marginal models facilitate explicit modeling of differences in occupational distributions, and hence modeling of structural mobility is more straightforward with this approach. As documented in Sobel et al., previous attempts to do this using log-linear and log-multiplicative models parameterizations are flawed. More generally, the key point of departure between marginal models and models for logarithms of cell probabilities in sets of tables is that the marginal model framework makes it easy to characterize substantively interesting marginal distributions in terms of layers (e.g., nations).

The above model formulation for sets of tables may also be used as an approach for specifying latent class models where the conditional independence assumption is relaxed. Yang and Becker (1998) develop such models, and they demonstrate how they may be applied to tables where it is plausible to view the manifest table as the result of mixing different population structures (e.g., independence and exchangeability). Not only does one attain more general latent class models than have been considered heretofore, but they also retain the ability to work with and model the latent marginal probabilities that are central to much of latent class modeling.

4. SUMMARY

In this article, we have discussed similarities and differences between standard log-linear models and marginal models. In the two-way classification, it was shown that many familiar log-linear models can also be written in the marginal models framework. In addition, marginal shift models for square tables were given as an important example where marginal models and standard log-linear models differ. The ready facility for constraining and comparing marginal distributions is what makes marginal models particularly well suited to modeling repeated measurements.

In comparing log-linear models and marginal models for three-way cross-classifications, we considered the degree to which the families of hierarchical models for the two parameterizations intersect. Standard collapsibility arguments were given as the reason that models where at least two pairs of variables are conditionally independent are equivalent to corresponding models of marginal independence. The equivalence of models for no three-factor interaction also was established. However, even when two models are equivalent, it does not necessarily follow that the marginal and conditional odds are equivalent. Equivalence follows only in the special case when the table is a perfect table. Models where only a single pair of variables are conditionally independent do not, in general, have marginal models equivalents.

We concluded our discussion of three-way tables by presenting an alternative parameterization of a marginal model when the table is viewed as a collection of two-way tables. This case is an important starting point for exploiting the power of marginal models in comparing marginal distributions across populations or groups. Occupational mobility and latent class analysis were given as areas where this has already been done successfully.

REFERENCES

Agresti, Alan. 1988. "A Model for Agreement Between Ratings on an Ordinal Scale." *Biometrics* 44:539-48.

———. 1990. *Categorical Data Analysis*. New York: John Wiley.

Becker, Mark P. 1990. "Quasi-Symmetric Models for the Analysis of Square Contingency Tables." *Journal of the Royal Statistical Society, Series B* 52:369-78.

———. 1994. "Analysis of Cross-Classifications of Counts Using Models for Marginal Distributions: An Application to Trends in Attitudes on Legalized Abortion." Pp. 229-65 in *Sociological Methodology 1994*, edited by P. V. Marsden. Oxford, UK: Blackwell.

Becker, Mark P. and Clifford C. Clogg. 1989. "Analysis of Sets of Two-Way Contingency Tables Using Association Models." *Journal of the American Statistical Association* 84:142-51.

Caussinus, Henri. 1965. "Contribution à L'analyse Statistique des Tableaux de Correlation." *Annals of the Faculty of Science, University of Toulouse* 29:77-182.

Clogg, Clifford C. 1982. "Some Models for the Analysis of Association in Multi-Way Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 77:803-15.

Darroch, John N. 1962. "Interactions in Multi-Factor Contingency Tables." *Journal of the Royal Statistical Society, Series B* 24: 251-63.

Duncan, Otis Dudley. 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 84:793-803.

Goodman, Leo A. 1979. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 74:537-52.

Koch, Gary G., J. Richard Landis, Jean L. Freeman, Daniel H. Freeman Jr., and Robert G. Lehnen. 1977. "A General Methodology for the Analysis of Experiments With Repeated Measurement of Categorical Data." *Biometrics* 33:133-58.

Lang, Joseph B. and Alan Agresti. 1994. "Simultaneously Modeling Joint and Marginal Distributions of Multivariate Categorical Responses." *Journal of the American Statistical Association* 89:625-32.

Lang, Joseph B. and Scott R. Eliason. 1997. "The Application of Association-Marginal Models to the Study of Social Mobility." *Sociological Methods and Research* 26:183-213.

McCullagh, Peter and John R. Nelder. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.

Plewis, Ian. 1985. *Analysing Change: Measurement and Explanation using Longitudinal Data*. New York: John Wiley.

Sinkhorn, R. 1967. "Diagonal Equivalence to Matrices With Prescribed Row and Column Sums. *American Mathematical Monthly* 74:402-5.

Sobel, Michael E., Mark P. Becker, and Susan M. Minick. 1997. "Origins, Destinations, and Association in Occupational Mobility." Submitted for publication.

Sobel, Michael E., Michael Hout, and Otis Dudley Duncan. 1985. "Exchange, Structure, and Symmetry in Occupational Mobility." *American Journal of Sociology* 91:359-72.

Xie, Yu. 1992. "The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables." *American Sociological Review* 57:380-95.

Yamaguchi, Kazuo. 1987. "Models for Comparing Mobility Tables: Toward Parsimony and Substance." *American Sociological Review* 52:482-94.

Yang, I. and Mark P. Becker. 1998. "Latent Class Marginal Models for Cross-Classifications of Counts." Forthcoming in *Sociological Methodology 1998*, edited by A. Raftery. Oxford, UK: Blackwell.

*Mark P. Becker is an associate professor in the Department of Biostatistics and associate dean for Academic Affairs, School of Public Health, and a faculty associate in the Institute for Social Research at the University of Michigan. His general research interests include nonlinear statistical models, numerical optimization, social stratification and mobility, leadership, and decision making. Recent publications include "Latent Class Marginal Models for Cross-Classifications of Counts," with Ilsoon Yang (*Sociological Methodology, *forthcoming), and "Latent Variable Modeling of Diagnostic Accuracy," with Ilsoon Yang (*Biometrics, *forthcoming).*

Susan Minick is an assistant professor in the Department of Medicine, Division of Biostatistics, at Indiana University, Indianapolis. Her research interests include the development of methods for the analysis of correlated categorical outcomes in areas such as rater agreement and social mobility. She has recently finished her dissertation on the simultaneous modeling of marginal distributions and log-multiplicative associations for ordered multivariate categorical data.

Ilsoon Yang is a postdoctoral research fellow in the Department of Biostatistics, Harvard School of Public Health. Her main research interests are in methods for analyzing survival data, categorical data, and longitudinal data. In particular, recently she has developed a class of semiparametric regression models for the analysis of doubly censored data and for the analysis of multivariate recurrent event data. She is also involved in collaborative research with the AIDS Clinical Trials Group.