Logit Models and Related Quasi-Symmetric Log-Linear Models for Comparing Responses to Similar Items in a Survey

ALAN AGRESTI

Sociological Methods & Research 1995 24: 68
DOI: 10.1177/0049124195024001004

Suppose that subjects respond to a battery of questions (items) of a similar nature in a survey, with each item having the same categorical scale. This article discusses models that express logits for the response distributions in terms of subject and item effects. The models, which generalize the Rasch model, have interpretations referring to subject-specific comparisons of the items. Recent literature shows that one can estimate item parameters using estimates of main effect parameters in corresponding quasi-symmetric log-linear models. We discuss this connection, giving primary attention to ordinal-response models using adjacent-category logits and cumulative logits. For the case of two items, we give expressions for models and corresponding parameter estimates that are the basis of simple tests of marginal homogeneity for square ordinal contingency tables.

Logit Models and Related Quasi-Symmetric Log-Linear Models for Comparing Responses to Similar Items in a Survey

ALAN AGRESTI
University of Florida

This article deals with modeling responses of subjects in a survey to a set of similar items that use the same categorical scale. Table 1, taken from the 1989 General Social Survey (Davis and Smith 1991) conducted by the National Opinion Research Center, illustrates the type of data. Subjects gave their opinions regarding government spending on the environment, health, assistance to big cities, and law enforcement, using a 3-point response scale (*too little*, *about right*, *too much*). Table 2 is similar, taken from the same survey. Subjects gave their opinions on early teens (aged 14-16) having sex relations before marriage and a man and a woman having sex relations

AUTHOR'S NOTE: This research was partially supported by a grant from the National Institutes of Health. Computations for the constrained maximum likelihood fit of the model (12) relating to the cumulative logit model were conducted using a FORTRAN program for generalized log-linear models written by Dr. Joseph Lang, Statistics Department, University of Iowa. The author thanks Atalanta Ghosh for computer aid in obtaining Table 3, and Kazuo Yamaguchi, Michael Sobel, and a referee for helpful comments.

SOCIOLOGICAL METHODS & RESEARCH, Vol. 24 No. 1, August 1995 68-95

TABLE 1: Opinions About Government Spending

		Cities								
		1 *Law Enforcement*			*2* *Law Enforcement*			*3* *Law Enforcement*		
Environment	*Health*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
1	1	62	17	5	90	42	3	74	31	11
		(62.0)	(17.9)	(3.4)	(85.8)	(34.0)	(5.7)	(77.3)	(27.6)	(11.1)
	2	11	7	0	22	18	1	19	14	3
		(13.1)	(5.2)	(0.9)	(24.9)	(14.8)	(2.1)	(20.1)	(10.0)	(4.4)
	3	2	3	1	2	0	1	1	3	1
		(1.8)	(0.6)	(0.3)	(3.1)	(1.5)	(0.7)	(6.0)	(3.3)	(1.3)
2	1	11	3	0	21	13	2	20	8	3
		(12.3)	(4.9)	(0.8)	(23.5)	(13.9)	(2.0)	(19.0)	(9.4)	(4.2)
	2	1	4	0	6	9	0	6	5	2
		(3.6)	(2.1)	(0.3)	(10.2)	(9.0)	(1.0)	(6.9)	(4.7)	(2.8)
	3	1	0	1	2	1	1	4	3	1
		(0.4)	(0.2)	(0.1)	(1.0)	(0.7)	(0.4)	(2.2)	(2.1)	(1.1)
3	1	3	0	0	2	1	0	9	2	1
		(1.6)	(0.6)	(0.2)	(2.7)	(1.4)	(0.6)	(5.3)	(2.9)	(1.1)
	2	1	0	0	2	1	0	4	2	0
		(0.4)	(0.2)	(0.1)	(1.0)	(0.7)	(0.4)	(2.1)	(2.0)	(1.0)
	3	1	0	0	0	0	0	1	2	3
		(0.1)	(0.1)	(0.0)	(0.3)	(0.3)	(0.2)	(0.6)	(0.7)	(3.0)

NOTE: Fitted values for ordinal quasi-symmetry model in parentheses. Categories are 1 = *too little*, 2 = *about right*, 3 = *too much*.
SOURCE: 1989 General Social Survey (Davis and Smith 1991).

TABLE 2: Opinions About Teenage Sex and Premarital Sex

	Premarital Sex			
Teen Sex	*1*	*2*	*3*	*4*
1	141 (141)	34 (34.5)	72 (72.4)	109 (109.0)
2	4 (1.8)	5 (4.9)	23 (22.8)	38 (37.5)
3	1 (0.6)	0 (1.8)	9 (8.9)	23 (22.9)
4	0 (0.1)	0 (0.3)	1 (1.5)	15 (15)

NOTE: Fitted values for cumulative logit model in parentheses. Categories are 1 = *always wrong*, 2 = *almost always wrong*, 3 = *wrong only sometimes*, 4 = *not wrong*.
SOURCE: 1989 General Social Survey (Davis and Smith 1991).

before marriage, using a 4-point response scale (*always wrong*, *almost always wrong*, *wrong only sometimes*, *not wrong at all*).

Two types of analyses for such data lead to distinct types of models. One type of analysis pertains to the dependence structure of the responses. For instance, one might study the strength of association between responses on various pairs of items, analyzing whether some pairs are more strongly associated than others or whether the pairwise association varies according to responses on other items. These matters relate to aspects of the joint distribution. Standard log-linear models are well designed for such investigations.

A second type of analysis relates to how the item response distributions differ. In Table 1, for instance, one might study whether subjects regarded spending as relatively higher on one item than the others. The analysis then involves comparing one-way marginal distributions. This article focuses on the second type of analysis. Our main emphasis concerns logit models that permit subject heterogeneity in the response distributions. For such models, we show how to estimate parameters that describe marginal effects by fitting certain log-linear models for the joint distribution.

We begin by reviewing the Rasch model, a popular item-response logit model that permits subject heterogeneity. We then review an interesting connection whereby one can estimate item parameters in the Rasch model using ordinary main effect estimates in a quasi-symmetric form of log-linear model. We discuss a generalized Rasch model for nominal responses, and show how to use a related quasi-symmetry model to make comparisons of marginal distributions. In particular, we see that the main effect parameters in quasi-symmetry models relate to simple odds-ratio comparisons of marginal distributions.

The main focus of the article concerns generalizations of the Rasch model that apply to ordinal response scales. For a model using adjacent-category logits, a related quasi-symmetry model is a special case of the ordinary model in which the main effect terms reflect the ordinality. An alternative generalization for ordinal responses uses cumulative logits.

For most cases, one can use standard software to fit the quasi-symmetric log-linear models. One can check model adequacy using goodness-of-fit tests and residuals for those log-linear models. We also discuss simpler representations of the ordinal models that occur when there are only two items. This leads to simple ways of testing marginal homogeneity for ordinal matched-pairs data in square contingency tables.

THE BINARY RASCH MODEL

We first introduce some notation. Suppose n subjects respond to T items that use the same categorical scale, having r categories. For subject i and item j, let ϕ_{hij} denote the probability of response in category h, $h = 1, \ldots, r$, so $\Sigma_h \phi_{hij} = 1$. The notation reflects potential within-subject and between-subject heterogeneity. The use of subject-specific notation allows us to consider models that differ from ordinary models whereby responses for different subjects are identically distributed. For instance, in the binary response case ($r = 2$), the usual *iid* assumption leads to the binomial sampling model; permitting different response probabilities for each subject is a way of allowing for additional variability (i.e., what is called *overdispersion*).

For binary responses, there is an enormous literature on models for subject-specific probabilities. Most of this literature refers to item-response models in which the two possible outcomes refer to correct and incorrect responses to a question. The best known item-response model is the Rasch model (Rasch 1961). It states that the probability of correct response by subject i on item j depends on subject abilities $\{\rho_i\}$ and item parameters $\{\tau_j\}$ through the equation

$$\phi_{1ij} = \rho_i\tau_j/(1 + \rho_i\tau_j). \tag{1}$$

Then $\phi_{2ij} = 1/(1 + \rho_i\tau_j)$, and we can express the two probabilities succinctly in a single formula as

$$\phi_{hij} = \frac{(\rho_i\tau_j)^{I(h)}}{1 + \rho_i\tau_j} \tag{2}$$

where $I(h)$ denotes an indicator of whether the response occurs in the first category (i.e., $I(h) = 1$ if $h = 1$, and $I(h) = 0$ if $h = 2$). The odds of a correct response has multiplicative form $\phi_{1ij}/\phi_{2ij} = \rho_i\tau_j$, corresponding to additive effects in a logit model,

$$\text{logit}(\phi_{1ij}) = \alpha_i + \beta_j. \tag{3}$$

The probability of a correct response increases as the subject's ability α_i or the item easiness β_j increases.

The term α_i common to all responses by subject i represents the effects of characteristics of that subject that could affect the response. This term is generally regarded as a latent variable. For a subject with given α_i, the Rasch model's local independence assumption states that responses on separate items are independent. Because of the variation in $\{\alpha_i\}$, responses on separate items by the same subject tend to be more similar than responses on separate items by different subjects. Thus, averaged over subjects, distributions of pairs of responses have positive associations. Besides allowing departures from the assumption that each subject has the same probability distribution, models with subject-specific terms can represent effects of omitted explanatory variables or measurement error in explanatory variables. Andersen (1980, chap. 6), Collett (1991, chap. 6), Morgan (1992, chap. 6), Hambleton, Swaminathan, and Rogers (1991), Goodman (1990), and Lindsay, Clogg, and Grego (1991) provide introductions to this and related models.

In fitting item response models, one's interest may focus on estimating the subject abilities, the item parameters, or both. Our discussion refers to estimation of the item parameters. A complication is that the total number of parameters increases as the sample size increases. As a consequence, the maximum likelihood (ML) estimators do not converge to the true parameter values. For $T = 2$ items, for instance, the ordinary ML estimator of $\beta_2 - \beta_1$ converges to double the true value (Andersen 1980, p. 244). On the other hand, replacing α_i by a common parameter α for all subjects usually results in a poor fit. It implies the simplistic log-linear model for the joint distribution whereby responses on the T items are mutually independent.

To estimate item parameters, there are two common ways of reducing the dimensionality of the parameter space. The first, a fixed-effects solution, treats the subject effects as nuisance parameters and eliminates them by a conditional argument. One conditions on the sufficient statistics for the subject effects and maximizes the resulting log likelihood, which depends only on the parameters of interest (e.g., $\{\beta_j\}$ in model [3]). Andersen's (1980) text is a good reference for the conditional ML approach. The second approach treats the subject terms as random effects, assuming some distribution for them, such as normal with unknown standard deviation. One integrates the likelihood with respect to that distribution to obtain a marginal likelihood. One then maximizes the marginal likelihood, which depends also on the parameters of the random effects distribution. For discussion of this marginal ML approach, see Bock and Aitkin (1981) and Hambleton et al. (1991).

A disadvantage often quoted for both the conditional ML and marginal ML approaches is their computational complexity. When the number of items is small, however, it is actually easy to obtain conditional ML estimates in Rasch-type logit models. In the next four sections, we review recent literature showing that conditional ML estimates are also ordinary ML estimates for main effect parameters in quasi-symmetric types of log-linear models.

THE RASCH MODEL AND QUASI-SYMMETRY

We continue with the case of T binary-response items. When we cross-classify the responses of subjects on T binary items, we obtain a 2^T contingency table. The jth dimension represents the two possible response outcomes for the jth item. We now discuss the connection between the Rasch model and a log-linear model for this contingency table. For ease of notation, we use $T = 3$, which occurs for data analyzed later in the section. Let (a, b, c) denote a potential response pattern for the three items, where each possible outcome is 1 or 2. Let π_{abc} denote the probability of this sequence for a randomly selected subject, with $\Sigma_a\Sigma_b\Sigma_c\pi_{abc} = 1$. This distribution refers to an averaging, for some population, of subject-specific probabilities. Let n_{abc} denote the number of subjects in the sample having response pattern (a, b, c), and let $m_{abc} = n\pi_{abc}$ denote its expected frequency. Log-linear models treat the cell counts $\{n_{abc}\}$ in the $2 \times 2 \times 2$ table as a multinomial sample of size n, with cell probabilities proportional to $\{m_{abc}\}$.

To provide intuition for the connection between the Rasch model and log-linear models, we show how the probabilities $\{\pi_{abc}\}$ for the $2 \times 2 \times 2$ cross-classification of the three items (averaged over subjects) relate to the subject-specific probabilities $\{\phi_{hij}\}$. For a particular subject i, the probability of the response sequence (a, b, c) equals $\phi_{ai1}\phi_{bi2}\phi_{ci3}$. For a randomly selected subject from a population of N subjects, the joint probability of these responses equals

$$\pi_{abc} = (\sum_i \phi_{ai1}\phi_{bi2}\phi_{ci3})/N.$$

Substituting the Rasch expression (2), we see that

$$\pi_{abc} = \tau_1^{I(a)}\tau_2^{I(b)}\tau_3^{I(c)} \sum_i \frac{\rho_i^{I(a)+I(b)+I(c)}}{N(1+\rho_i\tau_1)(1+\rho_i\tau_2)(1+\rho_i\tau_3)}. \tag{4}$$

The summation term in this expression takes the same value if we permute the indexes (a, b, c) in any way. Thus the expected frequencies satisfy the log-linear model

$$\log m_{abc} = \beta_1 I(a) + \beta_2 I(b) + \beta_3 I(c) + \lambda_{abc}, \tag{5}$$

where the interaction term λ_{abc} is identical for all permutations of its argument; that is, $\lambda_{112} = \lambda_{121} = \lambda_{211}$ and $\lambda_{122} = \lambda_{212} = \lambda_{221}$.

The parameters $\{\beta_1, \beta_2, \beta_3\}$ in (5) are identical to $\{\beta_1, \beta_2, \beta_3\}$ in the Rasch model (3). Model form (5) is the special case of the ordinary log-linear model

$$\log m_{abc} = \mu + \lambda_{1(a)} + \lambda_{2(b)} + \lambda_{3(c)} + \lambda_{12(ab)} + \lambda_{13(ac)} + \lambda_{23(bc)} + \lambda_{123(abc)}$$

for a $2 \times 2 \times 2$ table in which the two-factor terms are identical and symmetric in their indexes (e.g., $\lambda_{12(ab)} = \lambda_{13(ab)} = \lambda_{23(ab)} = \lambda_{12(ba)} = \lambda_{13(ba)} =$

$\lambda_{23(ba)}$), and the three-factor term is symmetric in its indexes, but the main effect terms are distinct. Because of the symmetry and identical values of the two-factor terms, without loss of generality we can set them equal to zero by absorbing them into the three-factor term, as in (5). The main effect terms in that model, which are the Rasch item effects, relate to the log-linear parameters by $\{\beta_j = \lambda_{j(1)} - \lambda_{j(2)}\}$. The indicator notation simply indicates coding that makes the parameters identifiable by setting the second parameter equal to zero for each main effect. Because model (5) exhibits symmetry in its interaction term but not in the main effects, it is called a quasi-symmetry model. Tjur (1982) showed the interesting result that the ordinary ML estimators of $\{\beta_j\}$ in the quasi-symmetry model (5) are identical to the conditional ML estimators of $\{\beta_j\}$ in the Rasch model (3).

For $T = 3$, the complete symmetry model states that π_{abc} is identical for any permutation of (a, b, c). The quasi-symmetry model is a generalization of the complete symmetry model that permits different main effect parameters for each item, and hence marginal heterogeneity. For three items, the complete symmetry model is the special case of (5) in which $\beta_1 = \beta_2 = \beta_3$. The symmetric interaction term in (5) implies, for instance, that each pair of items has the same association, conditional on the response for the other item. There are other ways of defining quasi-symmetry that preserve distinct values for some of the higher-order terms (e.g., Bishop, Fienberg, and Holland 1975, p. 303), but the most common approach is this one (Darroch 1986).

In averaging over subjects in (4), we made no assumption about a parametric form for subject effects. This suggests that a nonparametric type of marginal ML estimation for the Rasch model should give item parameter estimates similar to those for main effect parameters in log-linear model (5). In fact, Tjur (1982) also proved that the ordinary ML estimators of $\{\beta_j\}$ in model (5) result from a slightly extended version of the likelihood obtained with nonparametric marginal ML. Thus one can use the quasi-symmetry model to obtain conditional ML and extended nonparametric marginal ML estimates. Under the assumption that the model holds, de Leeuw and Verhelst (1986) showed that a nonparametric marginal ML approach yields estimators that are identical to the conditional ML estimators with probability increasing to 1 as n increases for a fixed number of items. They assumed an unspecified continuous distribution for the subject parameter. Lindsay et al. (1991) strengthened this, showing the same result if the subject-effect distribution has at least $(T + 1)/2$ support points.

A NOMINAL RESPONSE MODEL AND QUASI-SYMMETRY

The Rasch model extends to a generalized subject-specific logit model for nominal responses (Andersen 1973). For r response categories, the model can be expressed using an set of $r - 1$ logits, such as the baseline category logits

$$\log(\phi_{hij}/\phi_{1ij}) = \alpha_{hi} + \beta_{hj},\ h = 2, \ldots, r. \tag{6}$$

For each of the $r - 1$ logits, there is additivity of subject and item effects. The related log-linear model for conditional ML estimation of the item effects is the general quasi-symmetry model for an r^T contingency table (Conaway 1989). That model has form

$$\log m_{ab\ldots t} = \lambda_{1(a)} + \lambda_{2(b)} + \ldots + \lambda_{T(t)} + \lambda_{ab\ldots t}, \tag{7}$$

where the interaction term is symmetric in its indexes.

The quasi-symmetry model most commonly occurs in the literature for square tables ($T = 2$). In that case, (7) simplifies to

$$\log m_{ab} = \lambda_{1(a)} + \lambda_{2(b)} + \lambda_{ab},$$

where $\lambda_{ab} = \lambda_{ba}$ for all a and b. For r^T tables, the complete symmetry model has log-linear form

$$\log m_{ab\ldots t} = \lambda_{ab\ldots t}$$

where the interaction parameter is symmetric in the indexes. It has a number of nonredundant interaction parameters equal to the number of combinations of T numbers chosen from $\{1, 2, \ldots, r\}$ with replacement, which is $\binom{r+T-1}{T} = (r + T - 1)!/T!(r - 1)!$. (Each such choice yields a separate interaction term, which is identical for all permutations of the indexes.) One obtains the quasi-symmetry model (7) by adding main effect terms to the symmetry model. Once we have $\binom{r+T-1}{T}$ interaction parameters, there are $(r - 1)(T - 1)$ nonredundant main effect parameters in the quasi-symmetry model. To make the parameters identifiable, one adds constraints such as

$$\lambda_{1(1)} = \lambda_{2(1)} = \ldots = \lambda_{T(1)} = 0$$

and

$$\lambda_{T(1)} = \lambda_{T(2)} = \ldots = \lambda_{T(r)} = 0.$$

The complete symmetry model is the special case in which the $(r - 1)(T - 1)$ nonredundant main effect parameters also equal zero.

Table 3 summarizes relationships among the quasi-symmetry, complete symmetry, and mutual independence models. The mutual independence model is the special case of (7) in which the interaction parameters are identical; that is, it is a log-linear model with a "grand mean" parameter and T sets of main effect parameters. When the quasi-symmetry model holds, complete symmetry is equivalent to marginal homogeneity (Caussinus 1966; Darroch 1981). The standard test of marginal homogeneity is based on comparing the fits of the quasi-symmetry and symmetry model, with $df = (r - 1)(T - 1)$.

The estimable parameters in the generalized Rasch model (6) are differences between parameters for separate items for each given response category. For instance, for two items j and k, relative sizes of estimated item effects in (6) are related to relative sizes of estimated main effect parameters in the quasi-symmetry model by

$$\hat{\beta}_{hk} - \hat{\beta}_{hj} = (\hat{\lambda}_{k(h)} - \hat{\lambda}_{k(1)}) - (\hat{\lambda}_{j(h)} - \hat{\lambda}_{j(1)}).$$

If one uses software (such as GLIM; Francis, Green, and Payne 1993) having parameter constraints whereby the log-linear main effect estimate for the baseline level is set to zero, then the ML estimate $\hat{\lambda}_{k(h)} - \hat{\lambda}_{j(h)}$ for the quasi-symmetry model equals the conditional ML estimate $\hat{\beta}_{hk} - \hat{\beta}_{hj}$ in the Rasch model.

To illustrate the use of quasi-symmetric models for estimating generalized Rasch item parameters, we use Table 4, taken from the 1991 General Social Survey (Davis and Smith 1991). White subjects in the sample were asked (a) Do you favor busing of (Negro/Black) and White school children from one school district to another? (b) If your party nominated a (Negro/Black) for president, would you vote for him if he were qualified for the job? (c) During the last few years, has anyone in your family brought a friend who was a (Negro/Black)

TABLE 3: Summary of Quasi-Symmetry, Complete Symmetry, Mutual Independence, and Ordinal Quasi-Symmetry Models

	Number of Nonredundant Parameters		
Model	*Main Effect*	*Interaction*	*Residual* df
Quasi-symmetry	$(r-1)(T-1)$	$\binom{r+T-1}{T}$	$r^T - \binom{r+T-1}{T} - (r-1)(T-1)$
Complete symmetry	0	$\binom{r+T-1}{T}$	$r^T - \binom{r+T-1}{T}$
Mutual independence	$(r-1)T$	1	$r^T - (r-1)T - 1$
Ordinal quasi-symmetry	$T-1$	$\binom{r+T-1}{T}$	$r^T - \binom{r+T-1}{T} - (T-1)$

TABLE 4: Opinions About Racial Items

		Home		
President	*Busing*	*1*	2	*3*
1	1	41 (41.0)	65 (68.0)	0 (0.1)
	2	71 (68.4)	157 (157.5)	1 (1.4)
	3	1 (0.8)	17 (15.5)	0 (0.3)
2	1	2 (1.6)	5 (3.7)	0 (0.0)
	2	3 (3.8)	44 (44.0)	0 (0.1)
	3	1 (0.4)	0 (1.3)	0 (0.0)
3	1	0 (0.1)	3 (2.3)	1 (0.1)
	2	0 (2.4)	10 (8.5)	0 (0.0)
	3	0 (0.6)	0 (0.0)	1 (1.0)

NOTE: Fitted values for quasi-symmetry model in parentheses. Categories are 1 = *yes*, 2 = *no*, 3 = *don't know*.
SOURCE: 1991 General Social Survey (Davis and Smith 1991).

home for dinner? Each item was answered on a 3-point response scale (*yes*, *no*, *don't know*). We denote the three items by *B* (busing), *P* (president), and *H* (home).

Table 5 shows likelihood-ratio (G^2) and Pearson (X^2) goodness-of-fit statistics for several log-linear models. We will not take the numerical values too literally because the sampling design for this survey is not simple random sampling and because the data are sparse. However, these statistics give us indexes for summarizing fit and comparing fits of different models. The quasi-symmetry model gives a much

TABLE 5: Goodness of Fit of Log-Linear Models for Table 4

Model	*Likelihood-Ratio Statistic*	*Pearson Statistic*	df
Mutual independence	66.9	266.5	20
Complete symmetry	454.1	430.9	17
Quasi-symmetry	16.3	24.7	13

better fit than the symmetry model (quasi-symmetry with marginal homogeneity) or the mutual independence model (quasi-symmetry with no interactions). Table 4 shows the fit for the quasi-symmetry model. The primary lack of fit results from a single cell in which the observed count of 1 and fitted value of 0.053 contributes 17.0 to the Pearson statistic.

Table 6 shows the main effect estimates for the quasi-symmetry model, using category 1 as the baseline response and *H* as the baseline item. There were relatively few responses in the third category (don't know), and the main information Table 6 presents is that the estimated odds of outcome "yes" instead of "no" are much higher for *P* than for *B* or *H*; that is, White subjects are much more likely to vote for a Black person for president than to favor school busing or to have brought a Black person home for dinner. We use these estimates from the quasi-symmetry model to estimate item effects for the generalized Rasch model that expresses the logit for a response category in terms of additive subject and item effects. These help us summarize differences among the marginal distributions of the items. For instance, for each subject, the estimated odds of a yes rather than no response for *P* are $\exp(3.734 + 0.005) = 42$ times those for *B*.

To help readers replicate these results, Table 7 expresses the quasi-symmetry model for this $3 \times 3 \times 3$ table in the log-linear representation $\log \mathbf{m} = \mathbf{X}\beta$ for a vector of expected frequencies $\mathbf{m}$, a design matrix $\mathbf{X}$, and a vector of parameters β. Here, m_{abc} denotes the expected frequency for *B* at level *a*, *P* at level *b*, and *H* at level *c*, and we use the constraints (as in GLIM) whereby the main effect parameters are zero for category 1 and for item *H*. Note that cells having indexes that are permutations of each other all have the same interaction term. For instance, m_{123}, m_{132}, m_{213}, m_{231}, m_{312}, m_{321} all use interaction term λ_{123}.

TABLE 6: Estimates of Item Parameters for Table 4

	President		*Busing*		*Home*	
Category	*Estimate*	SE	*Estimate*	SE	*Estimate*	SE
1	0.000	—	0.000	—	0.000	—
2	–3.734	0.325	0.005	0.164	0.000	—
3	0.537	0.816	2.429	0.787	0.000	—

TABLE 7: Log-Linear Representation of Quasi-Symmetry for Table 4

m_{111}	0	0	0	0	1	0	0	0	0	0	0	0	0	0	$\lambda_{B(2)}$
m_{112}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	$\lambda_{B(3)}$
m_{113}	0	0	0	0	0	0	1	0	0	0	0	0	0	0	$\lambda_{P(2)}$
m_{121}	0	0	1	0	0	1	0	0	0	0	0	0	0	0	$\lambda_{P(3)}$
m_{122}	0	0	1	0	0	0	0	1	0	0	0	0	0	0	λ_{111}
m_{123}	0	0	1	0	0	0	0	0	1	0	0	0	0	0	λ_{112}
m_{131}	0	0	0	1	0	0	1	0	0	0	0	0	0	0	λ_{113}
m_{132}	0	0	0	1	0	0	0	0	1	0	0	0	0	0	λ_{122}
m_{133}	0	0	0	1	0	0	0	0	0	1	0	0	0	0	λ_{123}
m_{211}	1	0	0	0	0	1	0	0	0	0	0	0	0	0	λ_{133}
m_{212}	1	0	0	0	0	0	0	1	0	0	0	0	0	0	λ_{222}
m_{213}	1	0	0	0	0	0	0	0	1	0	0	0	0	0	λ_{223}
m_{221}	1	0	1	0	0	0	0	1	0	0	0	0	0	0	λ_{233}
log m_{222} =	1	0	1	0	0	0	0	0	0	0	1	0	0	0	λ_{333}
m_{223}	1	0	1	0	0	0	0	0	0	0	0	1	0	0	
m_{231}	1	0	0	1	0	0	0	0	1	0	0	0	0	0	
m_{232}	1	0	0	1	0	0	0	0	0	0	0	1	0	0	
m_{233}	1	0	0	1	0	0	0	0	0	0	0	0	1	0	
m_{311}	0	1	0	0	0	0	1	0	0	0	0	0	0	0	
m_{312}	0	1	0	0	0	0	0	0	1	0	0	0	0	0	
m_{313}	0	1	0	0	0	0	0	0	0	1	0	0	0	0	
m_{321}	0	1	1	0	0	0	0	0	1	0	0	0	0	0	
m_{322}	0	1	1	0	0	0	0	0	0	0	0	1	0	0	
m_{323}	0	1	1	0	0	0	0	0	0	0	0	0	1	0	
m_{331}	0	1	0	1	0	0	0	0	0	1	0	0	0	0	
m_{332}	0	1	0	1	0	0	0	0	0	0	0	0	1	0	
m_{333}	0	1	0	1	0	0	0	0	0	0	0	0	0	1	

AN ORDINAL MODEL USING ADJACENT-CATEGORIES LOGITS

We next discuss two types of subject-specific logit models for ordered categorical responses. These use the two most popular types of logit transforms for ordinal data—adjacent-categories logits (Goodman 1983) and cumulative logits (McCullagh 1977). For both approaches, estimates again result from fitting a corresponding quasi-symmetry model. We use notation for $T = 4$ items, corresponding to Table 1.

The first ordinal model has the adjacent-categories logit representation

$$\log(\phi_{h+1,\, ij}/\phi_{hij}) = \alpha_{hi} + \beta_j. \tag{8}$$

This is a special case of model (6) in which the item effects have the ordinal structure $\beta_{h+1,\, j} - \beta_{hj} = \beta_j$ for all h; that is, $\{\beta_{hj}\}$ is linear in h. In model (8), for each subject, the odds of outcome $h + 1$ instead of outcome h for item j are $\exp(\beta_j - \beta_k)$ times the odds for item k. The item effects are assumed to be identical for each pair of adjacent categories. A somewhat simpler model decomposes α_{hi} in (8) into $\delta_h + \alpha_i$ (Andersen 1973; Andrich 1978; Duncan 1984; Hout, Duncan, and Sobel 1987; Agresti 1993a). We will not discuss that model because of our primary interest in estimating item effects rather than subject effects, and also because the more general structure in (8) relates more closely to symmetry and quasi-symmetry models. For instance, the condition of symmetry is the special case of (8) with equal item parameters.

Generalizing Tjur's (1982) argument, Agresti (1993a) noted that conditional ML estimates and extended nonparametric marginal ML estimates of the item effects in this model are identical to the ordinary ML estimates obtained in fitting the log-linear model

$$\log m_{abcd} = a\beta_1 + b\beta_2 + c\beta_3 + d\beta_4 + \lambda_{abcd}, \tag{9}$$

where λ is permutationally invariant. This is a special case of the quasi-symmetry model that has linear structure for the main effects. It treats the main effects as variates, with equally spaced scores, rather than qualitative factors. A generalization of (9) using arbitrary monotone scores $\{\upsilon_h\}$ in the linear structure relates to replacing β_j by $(\upsilon_{h+1} - \upsilon_h)\beta_j$ in logit model (8). We refer to model (9) as the ordinal quasi-symmetry model.

Model (9) tends to fit well when there are location shifts in the item distributions; that is, when observations for one item tend to be shifted downward or upward relative to responses for another item. It fits poorly if the items have quite different dispersion, such as when the responses fall mostly in one category for one item but are dispersed among all categories for another item. The estimates of $\{\beta_j\}$ have the same order as the sample mean responses (for equally spaced scores) in the T one-way margins of the r^T table. The complete symmetry model is the special case of the ordinal quasi-symmetry model in which $\beta_1 = \ldots = \beta_T$. When model (9) fits well, one can test marginal homogeneity using a likelihood-ratio test with $df = T - 1$, based on comparing its fit to that of the complete symmetry model. Table 3 compares its main effect and interaction structure to models discussed in the previous section.

To illustrate the ordinal subject-specific logit model (8), we analyze Table 1 on government spending items. Table 8 shows the goodness of fit of several log-linear models. The ordinal quasi-symmetry model (9) fits relatively well ($G^2 = 64.90$, $df = 63$). Table 1 also displays fitted values for this model. There is a dramatic improvement compared to the complete symmetry model (which has $G^2 = 638.24$, $df = 66$), at the expense of only adding three independent parameters.

Denote main effects in the ordinal quasi-symmetry model for the four items by $\beta_C, \beta_L, \beta_H, \beta_E$. The estimated main effects and asymptotic standard errors, using the constraint $\hat{\beta}_E = 0$, are $\hat{\beta}_C = 1.941$ (asymptotic $SE = 0.118$), $\hat{\beta}_L = 0.372$ (asymptotic $SE = 0.104$), $\hat{\beta}_H = 0.059$ (asymptotic $SE = 0.108$). One can use these main effect estimates to describe marginal heterogeneity by interpreting them as estimates of the corresponding item parameters in the generalization (8) of the Rasch model. Aid for cities received substantially less support than aid for the other items. For instance, for each subject, the estimated odds that the response is "too much" rather than "about right," or "about right" rather than "too little," are exp(1.941) = 7.0 times as high for cities as for the environment. All asymptotic standard errors of differences of estimates are about 0.11. To compare all 6 pairs of item parameters while maintaining a bound of .05 on the overall probability of Type I error, we used .05/6 = 0.0083 for the α level for each comparison. This analysis indicates significant differences between all pairs except β_H and β_E.

One can use software for log-linear models to fit quasi-symmetry models. To illustrate, Table 9 shows the use of GLIM for obtaining the

TABLE 8: Goodness of Fit of Log-Linear Models for Table 1

Model	*Likelihood-Ratio Statistic*	*Pearson Statistic*	df
Mutual independence	124.3	277.6	72
Complete symmetry	638.2	711.6	66
Ordinal quasi-symmetry	64.9	70.5	63
Quasi-symmetry	58.0	61.7	60

results just quoted for the ordinal quasi-symmetry model, and for fitting the more general quasi-symmetry model discussed in the previous section. The %GL terms generate levels; for instance, L has three levels in blocks of size 1, so L takes levels (1, 2, 3, 1, 2, 3, 1, . . .) for the cells in the order their counts are entered. The factor assigned the name "symm" generates the symmetric interaction terms, one for each possible subset of indexes $\{a, b, c, d\}$, because each interaction term having the same index is equal. For instance, the second entered cell count has indexes (1, 1, 2, 1) on (E, H, L, C), and the fourth entered count has indexes (1, 1, 1, 2), so they share the same symmetry parameter. In the second fit statement, the coefficients of the terms denoted by C, L, H, and E are the ordinal item parameters. After being declared as factors, they are treated as nominal factors rather than ordinal variates in subsequent model statements.

AN ORDINAL MODEL USING CUMULATIVE LOGITS

Perhaps the most popular model form for ordinal responses uses cumulative logits. For subject i and item j, denote the cumulative probability at category h by $\gamma_{hij} = \phi_{1ij} + \ldots + \phi_{hij}$, $h = 1, \ldots, r$. The cumulative logit alternative to model (8) has form

$$\log[\gamma_{hij}/(1 - \gamma_{hij})] = \alpha_{hi} - \beta_j, \tag{10}$$

$h = 1, \ldots, r - 1$, $i = 1, \ldots, n$, $j = 1, \ldots, T$. We attach a negative sign to the item parameter so that relatively larger values of β_j correspond to a tendency to make higher responses on the scale for that item. For each subject, the odds that the response for item a falls above any fixed level are $\exp(\beta_a - \beta_b)$ times the odds for item b. This model has the

TABLE 9: GLIM Code for Fitting Quasi-Symmetry Models to Table 1

```
$units 81
$data count $read
 62   17   5    90   42   3    74   31   11
 11    7   0    22   18   1    19   14    3
  2    3   1     2    0   1     1    3    1
 11    3   0    21   13   2    20    8    3
  1    4   0     6    9   0     6    5    2
  1    0   1     2    1   1     4    3    1
  3    0   0     2    1   0     9    2    1
  1    0   0     2    1   0     4    2    0
  1    0   0     0    0   0     1    2    3
$calc C = %GL(3,3): L = %GL(3,1): H = %GL(3,9): E = %GL(3,27) $
$ass symm =
1,2,3,2,4,5,3,5,6,2,4,5,4,7,8,5,8,9,3,5,6,5,8,9,6,9,10,2,4,5,4,7,8,5,8,9,4,7,8,7,11,12,8,12,13,5,8,
9,8,12,13,9,13,14,3,5,6,5,8,9,6,9,10,5,8,9,8,12,13,9,13,14,6,9,10,9,13,14,10,14,15 $
$fac symm 15
$yvar count $err pois
$fit symm ! Fits complete symmetry model
$fit symm + C + L + H + E $ ! Fits ordinal quasi-symmetry
$fac C 3 L 3 H 3 E 3 $ ! nominal main effects
$fit C + L + H + E $ ! Fits mutual independence
$fit + symm $ ! Fits quasi-symmetry
$dis e r ! Displays estimates, fitted values, and residuals
$dis s ! Displays standard errors of difference estimates
```

proportional odds property, for which the T-item effects $\{\beta_j\}$ are identical at each h. For $r = 2$, this model and the one in the previous section simplify to the Rasch model. McCullagh (1977) discussed a related model for $T = 2$.

One could also assume ordinal structure for subject effects, decomposing $\{\alpha_{hi}\}$ in (10) into

$$\alpha_{hi} = \delta_h - \alpha_i.$$

The resulting model holds if for each pairing of subject i and item j, there is an underlying continuous variable that has a logistic distribution with mean $\alpha_i + \beta_j$, and the observed outcome falls in category h when the underlying continuous variable falls between the "cutpoints" δ_{h-1} and δ_h. The more general model (10) permits different subjects to use different cutpoints for determining the response category, with the region between $\alpha_{h-1,\,i}$ and α_{hi} determining category h for subject

i. For instance, for a given value for an underlying continuous scale, one subject may regard it as part of category "good," while a second subject regards it as "very good." Allowing such generality for cutpoints implies that one cannot distinguish between location differences for different subjects. Hence one can scale the cutpoints so there is a common mean for the underlying distribution for each subject's responses on a given item. Because our focus is on estimating item effects rather than cutpoints or subject effects, we discuss the more general model (10). An advantage of that model is that it has symmetry embedded as the special case of equal item parameters.

Unfortunately, the cumulative logit model does not have reduced sufficient statistics, so the standard conditional ML approach for estimating item parameters is unavailable. For the case $T = 2$, we now present (from Agresti and Lang 1993) a way of eliminating the nuisance parameters by noting a corresponding model for the $r \times r$ table of observed counts. Let

$$L_{ab} = \log\{\frac{P(Y_{i1} > a, y_{i2} \leq b)}{P(Y_{i1} \leq a, y_{i2} > b)}\}.$$

By the assumed independence of (Y_{i1}, Y_{i2}) for a given subject, each joint probability in this expression factors as the product of marginal probabilities. Hence $L_{ab} = \text{logit}(\gamma_{bi2}) - \text{logit}(\gamma_{ai1})$, which equals $(\alpha_{bi} - \alpha_{ai}) - (\beta_2 - \beta_1)$ for model (10). Thus

$$L_{ab} + \mathrm{L}_{ba} = 2(\beta_1 - \beta_2), \tag{11}$$

for all $a \leq b$. This expression applies to the $r \times r$ table of probabilities for each subject, but it is straightforward to show that the same relationship holds for the $r \times r$ joint distribution averaged over subjects; that is,

$$\log\left(\frac{\Sigma_{a' > a}\Sigma_{b' \leq b}\pi_{a'b'}}{\Sigma_{a' \leq a}\Sigma_{b' > b}\pi_{a'b'}}\right) + \log\left(\frac{\Sigma_{a' > b}\Sigma_{b' \leq a}\pi_{a'b'}}{\Sigma_{a' \leq b}\Sigma_{b' > a}\pi_{a'b'}}\right) = 2(\beta_1 - \beta_2) \tag{12}$$

for all $a \leq b$. In fact, when $T = 2$, equation (12) characterizes the joint distribution corresponding to model (10).

Representation (12) suggests a way to estimate the difference in item parameters for the cumulative logit model (10) applied to two items. One can maximize the multinomial likelihood for the $r \times r$ observed table, subject to the constraint (12) holding for all $r(r-1)/2$ combinations of $a \leq b$. The special case with no item effect (i.e., constraining the sum of log-odds to equal 0 for all $a \leq b$) is an alternative characterization of the complete symmetry model. One obtains the estimated item effect using methods for maximizing a multinomial likelihood subject to constraints. Methodology for doing this has been available for some time (e.g., Aitchison and Silvey 1958), and several individually written programs for doing so exist, but the method is not available in ordinary log-linear model software. In SAS (PROC CATMOD), it is possible to obtain a weighted least squares (WLS) fit of model (12) applied to all $a \leq b$. In the next section, we show how to obtain a good estimate of the item effect with a simple closed-form expression that relates to fitting the quasi-symmetry model for a set of collapsed tables.

To illustrate the cumulative logit model, we analyze Table 2. Here the nature of the response categories (*always wrong*, *almost always wrong*, *wrong only sometimes*, *not wrong at all*) makes the use of equally spaced response scores questionable, and it is not obvious what scores are appropriate. The cumulative logit model does not require such a choice. For these data, the ML fit of the model (12) used to obtain the estimated item effect for the cumulative logit model has $G^2 = 6.86$ and $X^2 = 5.46$, based on $df = 5$. By contrast, the symmetry model has $G^2 = 378.4$ and $X^2 = 282.9$ ($df = 6$). The ML estimate of $\beta_2 - \beta_1$ is 4.46 (asymptotic $SE = 0.43$). Responses regarding teen sex tended to be much more conservative than those regarding premarital adult sex. For instance, for each subject, the estimated odds that response on teenage sex is $\leq h$ are $\exp(4.46) = 87$ times the corresponding estimated odds for premarital adult sex.

Table 10 shows the code for using SAS (CATMOD) to fit model (12), using WLS. In this procedure, one expresses the model in the form $\mathbf{C} \log \mathbf{A}\pi = \mathbf{X}\beta$, for a vector of cell probabilities π. The $\mathbf{A}$ matrix consists of 0 and 1 elements and forms the quadrants of terms used in (12). After taking the logarithms of these quadrant probabilities, one forms the appropriate contrasts of them with the $\mathbf{C}$ matrix. A disadvantage of WLS is that small constants must be added to zero cells when any of the quadrant totals of sample counts are zero. One should check the dependence of the WLS estimate on that constant using a

TABLE 10: SAS Code for WLS Fitting of Model (12) to Table 2

```
data clogit;
input cell $ count @@;
if count=0 then count=1e-2;
cards;
   1    141     2    34     3    72     4    109
   5      4     6     5     7    23     8     38
   9      1    10     0    11     9    12     23
  13      0    14     0    15     1    16     15
;
proc catmod order=data; weight count;
response
   2 -2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  2 -2  0  0  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  0  0  2 -2  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  0  0  0  0  1 -1  0  0  1 -1  0  0  0  0  0  0,
   0  0  0  0  0  0  0  0  1 -1  0  0  0  0  1 -1  0  0,
   0  0  0  0  0  0  0  0  0  0  0  0  1 -1  0  0  1 -1
  log
   0  1  1  1  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  0  0  1  0  0  0  1  0  0  0  1  0  0  0,
   0  0  1  1  0  0  1  1  0  0  0  0  0  0  0  0,
   0  0  0  0  0  0  0  0  1  1  0  0  1  1  0  0,
   0  0  0  1  0  0  0  1  0  0  0  1  0  0  0  0,
   0  0  0  0  0  0  0  0  0  0  0  0  1  1  1  0,
   0  0  1  1  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  0  0  1  1  0  0  1  1  0  0  1  1  0  0,
   0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0,
   0  0  0  0  1  1  1  0  1  1  1  0  1  1  1  0,
   0  1  1  1  0  1  1  1  0  0  0  0  0  0  0  0,
   0  0  0  0  0  0  0  0  1  0  0  0  1  0  0  0,
   0  0  0  1  0  0  0  1  0  0  0  0  0  0  0  0,
   0  0  0  0  0  0  0  0  1  1  1  0  1  1  1  0,
   0  1  1  1  0  1  1  1  0  1  1  1  0  0  0  0,
   0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0,
   0  0  1  1  0  0  1  1  0  0  1  1  0  0  0  0,
   0  0  0  0  0  0  0  0  0  0  0  0  1  1  0  0;
model cell=( 2, 2, 2, 2, 2, 2)/pred covb;
run;
```

sensitivity analysis by performing the analysis with different choices of the constant. For Table 2, adding a constant of 0.01 to the empty cells gave a WLS item effect estimate of 3.31 (asymptotic $SE = 0.32$), whereas adding 0.1 gave an estimate of 3.65 (asymptotic $SE = 0.35$). When the estimate is unstable, it often helps to fit the model only to the $r - 1$ constraints in (12) for which $a = b$ (McCullagh 1977), for which there is less of a tendency to have zero quadrant counts. (One would do this using the SAS code in Table 10 by deleting the last three rows and the last 12 columns of the **C** matrix, the last 12 rows of the **A** matrix, and the last three elements in the **X** matrix specified in the model statement; one then gets an estimate of 3.99, with asymptotic $SE = 0.41$.) When the cell counts are large, the WLS estimate and its standard error are nearly identical to those obtained using constrained ML.

Similar substantive results occur in using the adjacent-categories-logit item effects model (8) for these data. The conditional ML estimated effect is 2.628 (asymptotic $SE = 0.353$). The estimated effect is smaller than with the cumulative logit model because the log-odds ratio refers to adjacent response categories rather than the entire scale. The related ordinal quasi-symmetry model fits well and also gives very strong evidence that attitudes are much more conservative toward teen sex than adult sex. Table 11 describes the fit of this and other models. The ordinal models fit essentially as well as the general quasi-symmetry model, but much better than the mutual independence or complete symmetry models. Compared to general quasi-symmetry, they have the advantage of simpler interpretation.

We now briefly discuss the cumulative logit model (10) for the case of arbitrary T. That model implies that for each fixed h, the ordinary Rasch model holds for a collapsing of the response into binary outcomes ($\leq h$, $> h$). Thus, when the model holds, one can estimate $\{\beta_j\}$ by obtaining item estimates for the Rasch model applied to the 2^T table corresponding to the collapsed binary response scale, for any h (e.g., one could do this by fitting the quasi-symmetry model to the 2^T table). To obtain more efficient estimates, Agresti and Lang (1993) suggested the following approach. Consider simultaneously the $(r - 1)$ separate 2^T contingency tables, in which the h^{th} table is the cross-classification of responses for the h^{th} binary collapsing of the ordinal response scale, $h = 1, \ldots, r - 1$. They fitted the quasi-symmetry model simultaneously to the $r - 1$ separate 2^T tables. The proportional odds assumption for model (10) implies that the same main effect parameters apply to each of the $(r - 1)$ quasi-symmetry submodels. Their fitting process takes into account the dependence due to each table classifying the same subjects. The process is based on maximizing a

TABLE 11: Goodness of Fit of Models for Table 2

Model	*Likelihood-Ratio Statistic*	*Pearson Statistic*	df
Mutual independence	94.9	78.5	9
Complete symmetry	378.4	282.9	6
Ordinal quasi-symmetry	5.4	4.0	5
Cumulative logit	6.9	5.5	5
Quasi-symmetry	2.6	2.5	3

multinomial likelihood for the original r^T table, subject to the constraint that this form of quasi-symmetry model holds simultaneously for each collapsed 2^T table.

ANALYSIS OF ORDINAL MATCHED PAIRS

This section considers separately the special case of a bivariate response ($T = 2$), which occurs for matched-pairs data. In this case, quasi-symmetry models used to obtain subject-specific item estimates have simple logit representations, and additional ways exist of getting the estimates.

The ordinary quasi-symmetry model for $T = 2$ is

$$\log m_{ab} = \lambda_{1(a)} + \lambda_{2(b)} + \lambda_{ab},$$

where $\lambda_{ab} = \lambda_{ba}$. This has logit form

$$\log\{m_{ab}/m_{ba}\} = \beta_b - \beta_a$$

for all a and b, where $\beta_h = \lambda_{2(h)} - \lambda_{1(h)}$. The subject-specific logit model (8) for adjacent categories relates to a special case (9) of the quasi-symmetry model, for which this logit representation is especially simple. Letting $\beta = \beta_2 - \beta_1$, we have

$$\log\{m_{ab}/m_{ba}\} = \beta(b - a). \qquad (13)$$

In fact, we can also obtain the estimate of this effect using software for logistic regression models. We then treat $\{n_{ab}, a < b\}$ as independent binomial variates with sample sizes $\{n_{ab} + n_{ba}\}$. For the binary case

($r = 2$), $\hat{\beta} = (n_{12}/n_{21})$ is the resulting conditional ML estimate (Cox 1958) for the Rasch model.

Model (13) is a special case of a family of models satisfying

$$\log\{m_{ab}/m_{ba}\} = \delta_{a-b}, \tag{14}$$

discussed by Goodman (1979). Model (14) is called a diagonals-parameter symmetry model, and model (13) is called a linear diagonals-parameter symmetry model (Agresti 1983). Given that model (13) holds, marginal homogeneity is equivalent to symmetry, which is the case $\beta = 0$. For ordinal responses, one can base simple tests of marginal homogeneity on model (13). A Wald test uses as a test statistic the ratio of $\hat{\beta}$ to its asymptotic standard error, which is a by-product of the Newton-Raphson algorithm for fitting log-linear models. The likelihood-ratio test uses the difference between the G^2 statistics for the symmetry model and model (13). Rao's efficient score test is based on the difference in sample means for the marginal distributions for equally spaced category scores. Specifically, let $\{p_{ij}\}$ denote the sample proportions in the observed $r \times r$ table. One can form a z test statistic by the ratio of $d = [\Sigma_i\, i(p_{i+} - p_{+i})]$ to its estimated standard error, which is the square root of $(1/n)[\Sigma_i\Sigma_j(i-j)^2 p_{ij} - d^2]$.

For the cumulative logit model (10) with $T = 2$, it is possible to construct a simple estimate of the item effect $\beta = \beta_2 - \beta_1$ using the fact that the model implies a Rasch model for each of the $r - 1$ collapsings of the response to a binary variable. For each collapsing, we can use the off-diagonal cells of the 2×2 table to get an estimate in the form of the binary conditional ML estimate, $\log(n_{12}/n_{21})$. We can obtain a nearly efficient estimator by combining these, adding the numerators and adding the denominators before taking their ratio and their logarithm (Agresti and Lang 1993). In terms of the cell counts in the full $r \times r$ table, the resulting estimate of the effect is

$$\tilde{\beta} = \log\{[\sum_{i<j}(j-i)n_{ij}]/[\sum_{i>j}(i-j)n_{ij}]\}. \tag{15}$$

It is usually very similar to the constrained ML estimator, and often more stable than the WLS estimator.

This estimator is simple to compute and intuitively appealing. The numerator of the estimate weights each cell count above the main diagonal by its distance from that diagonal, whereas the denominator weights each cell count below the main diagonal by its distance. Its estimated asymptotic variance equals

$$\hat{V}(\tilde{\beta}) = \frac{\Sigma_{i<j}(j-i)^2 n_{ij}}{[\Sigma_{i<j}(j-i)n_{ij}]^2} + \frac{\Sigma_{i>j}(i-j)^2 n_{ij}}{[\Sigma_{i>j}(i-j)n_{ij}]^2}. \quad (16)$$

Applying these formulas to Table 2, we get an effect estimate of

$$\tilde{\beta} = \log\left(\frac{(34+23+38)+2(72+38)+3(109)}{(4+0+1)+2(1+0)+3(0)}\right) = \log(642/7) = 4.519$$

and a variance estimate of

$$\hat{V}(\tilde{\beta}) = \frac{(34+23+38)+4(72+38)+9(109)}{(642)^2} + \frac{(4+0+1)+4(1+0)+9(0)}{(7)^2} = 0.1874,$$

for which the asymptotic $SE = 0.433$. By comparison, the constrained ML values are $\hat{\beta} = 4.465$ and the asymptotic $SE = 0.434$.

We can base another simple Wald test of marginal homogeneity for ordinal matched-pairs data on the ratio of $\hat{\beta}$ to its estimated standard error. Like the test based on the ordinal quasi-symmetry model, it is sensitive to location shifts in the marginal distributions. The likelihood-ratio test uses the increase in the G^2 values between the fits of the constrained model (12) corresponding to the cumulative logit model and the simpler model (complete symmetry) with $\beta_2 - \beta_1 = 0$.

REMARKS

We have introduced two types of a quasi-symmetric log-linear model for ordinal responses. One type implies that the ordinary quasi-symmetry model for a 2^T table holds for each of the $(r-1)$ pairs of adjacent response categories, with the same main effect parameters applying to each pair. The other type implies that the ordinary quasi-symmetry model for a 2^T table holds for each of the $(r-1)$ binary collapsings of the response, again with the same main effect parameters applying to each collapsing. By fitting the models, one can obtain item estimates for subject-specific ordinal logit models. Specifically, we have shown that main effect parameters for quasi-symmetric models relate to meaningful odds ratios for comparing marginal distributions in subject-specific models.

A factor that can influence the choice of ordinal model relates to whether one prefers odds-ratio interpretations referring to the entire response scale (as in the cumulative logit model) or to pairs of categories (as in the adjacent-categories logit). The cumulative logit model has the advantage of a natural connection to a model for an underlying continuous response, with the same item effects no matter how one collapses that response into discrete categories (McCullagh 1977). Samejima (1969), Andrich (1978), Masters (1982), Duncan (1984), and Tutz (1990) described closely related types of subject-specific models for ordinal responses.

One can extend the models to incorporate covariates. For instance, one might stratify a sample by gender and analyze whether the same item effects apply to men and women. One could do this by comparing the fits of two models, one assuming homogeneous item effects and the other permitting heterogeneous item effects. The related quasi-symmetry models also have homogeneous or heterogeneous main effects, with the symmetric interaction term having different parameters for each gender. Agresti (1993b) gave examples of this type and illustrated the use of GLIM and SAS for fitting the models. There is need for further work on marginal ML solutions for the ordinal models. It remains to be seen whether results of Lindsay et al. (1991) about connections between nonparametric marginal ML and conditional ML estimates for the binary case generalize to ordinal responses.

When the subject parameter is assumed to be identical for all subjects (i.e., there is no subject heterogeneity), the logit models discussed in this article also describe population-averaged effects. Such effects relate to the response of a randomly selected subject on one item, relative to the response of another randomly selected subject on another item. In general, though, if a subject-specific model holds of logistic form, the implied population-averaged model for the marginal distributions is not necessarily of logistic form. When there is subject heterogeneity, estimated effects in population-averaged models are usually weaker than estimated effects in subject-specific models (see Neuhaus, Kalbfleisch, and Hauck 1991 for discussion of this issue).

Parameters in population-averaged models describe differences between marginal distributions purely in terms of the marginal probabilities. By contrast, parameters in subject-specific models describe differences in a way that directly incorporates the parametric form for the dependence of the responses in the joint distribution. In our models, for instance, the marginal structure is implied by the quasi-symmetric structure of the joint distribution. For an example of the population-averaged approach, see Becker (1994).

The subject-specific models we have discussed are rather simplistic, and the related quasi-symmetry models will fit well in a limited range of situations. Even when quasi-symmetric models show lack of fit, however, they usually fit much better than complete symmetry or mutual independence models. They may fit poorly if there are severe violations of the local independence assumption, or the assumption that each item effect is the same for every subject. The first of these violations may happen for repeated measurement of an item over time, whereby there is some residual dependence between observations nearby in time (Conaway 1989, 1992). Violations of the latter assumption sometimes occur when marginal distributions show differences in dispersion as well as location. Nevertheless, the models address components of relationships not analyzed by standard log-linear analyses of associations. In practice, they should often provide useful comparisons of response distributions for items of a similar nature.

REFERENCES

Agresti, A. 1983. "A Simple Diagonals-Parameter Symmetry and Quasi-Symmetry Model." *Statistics & Probability Letters* 1:313-6.

———. 1993a. "Computing Conditional Maximum Likelihood Estimates for Generalized Rasch Models Using Simple Log-Linear Models With Diagonals Parameters." *Scandinavian Journal of Statistics* 20:63-71.

———. 1993b. "Distribution-Free Fitting of Logit Models With Random Effects for Repeated Categorical Responses." *Statistics in Medicine* 12:1969-87.

Agresti, A. and J. B. Lang. 1993. "A Proportional Odds Model With Subject-Specific Effects for Repeated Ordered Categorical Responses." *Biometrika* 80:527-34.

Aitchison, J. and S. D. Silvey. 1958. "Maximum Likelihood Estimation of Parameters Subject to Restraints." *Annals of Mathematical Statistics* 29:813-28.

Andersen, E. B. 1973. "Conditional Inference for Multiple-Choice Questionnaires." *British Journal of Mathematical and Statistical Psychology* 26:31-44.

———. 1980. *Discrete Statistical Models With Social Science Application.* Amsterdam: North-Holland.

Andrich, D. 1978. "A Rating Formulation for Ordered Response Categories." *Psychometrika* 43:561-73.

Becker, M. 1994. "Analysis of Cross-Classifications of Counts Using Models for Marginal Distributions: An Application to Trends in Attitudes on Legalized Abortion." *Sociological Methodology* 24:229-65.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete Multivariate Analysis.* Cambridge, MA: MIT Press.

Bock, R. D. and M. Aitkin. 1981. "Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm." *Psychometrika* 46:443-59.

Caussinus, H. 1966. "Contribution à l'analyse statistique des tableaux de correlation." *Annales de la Faculte des sciences de l'Université de Toulouse* 29:77-182.

Collett, D. 1991. *Modelling Binary Data.* London: Chapman and Hall.

Conaway, M. 1989. "Analysis of Repeated Categorical Measurements With Conditional Likelihood Methods." *Journal of the American Statistical Association* 84:53-62.

———. 1992. "The Analysis of Repeated Categorical Measurements Subject to Nonignorable Non-Response." *Journal of the American Statistical Association* 87:817-24.

Cox, D. R. 1958. "Two Further Applications of a Model for Binary Regression." *Biometrika* 45:562-5.

Darroch, J. N. 1981. "The Mantel-Haenszel Test and Tests of Marginal Symmetry: Fixed Effects and Mixed Models for a Categorical Response." *International Statistical Review* 49:285-307.

———. 1986. "Quasi Symmetry. Pp. 469-73 in *Encyclopedia of Statistical Sciences,* Vol. 7, edited by S. Kotz and N. L. Johnson. New York: Wiley.

Davis, J. A. and T. W. Smith. 1991. *General Social Surveys, 1972-1991* [Machine-readable data file]. J. A. Davis, principal investigator; T. W. Smith, director and co-principal investigator. Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor].

de Leeuw, J. and N. Verhelst. 1986. "Maximum Likelihood Estimation in Generalized Rasch Models." *Journal of Educational Statistics* 11:183-96.

Duncan, O. D. 1984. "Rasch Measurement: Further Examples and Discussion." Pp. 367-403 in *Surveying Subjective Phenomena,* Vol. 1, edited by C. F. Turner and E. Martin. New York: Russell Sage.

Francis, B., M. Green, and C. Payne, eds. 1993. *The GLIM System Release 4 Manual.* Oxford: Clarendon.

Goodman, L. A. 1979. "Multiplicative Models for Square Contingency Tables With Ordered Categories." *Biometrika* 66:413-8.

———. 1983. "The Analysis of Dependence in Cross-Classifications Having Ordered Categories, Using Log-Linear Models for Frequencies and Log-Linear Models for Odds." *Biometrics* 39:149-60.

———. 1990. "Total-Score Models and Rasch-Type Models for the Analysis of a Multidimensional Contingency Table, or a Set of Multidimensional Contingency Tables, With Specified and/or Unspecified Order for Response Categories." *Sociological Methodology* 20:249-94.

Hambleton, R. K., H. Swaminathan, and H. J. Rogers. 1991. *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage.

Hout, M., O. D. Duncan, and M. E. Sobel. 1987. "Association and Heterogeneity: Structural Models of Similarities and Differences." *Sociological Methodology* 17:145-84.

Lindsay, B., C. C. Clogg, and J. Grego. 1991. "Semiparametric Estimation in the Rasch Model and Related Exponential Response Models, Including a Simple Latent Class Model for Item Analysis." *Journal of the American Statistical Association* 86:96-107.

Masters, G. N. 1982. "A Rasch Model for Partial Credit Scoring." *Psychometrika* 47:149-74.

McCullagh, P. 1977. "A Logistic Model for Paired Comparisons With Ordered Categorical Data." *Biometrika* 64:449-53.

Morgan, B. J. T. 1992. *Analysis of Quantal Response Data*. London: Chapman and Hall.

Neuhaus, J. M., J. D. Kalbfleisch, and W. W. Hauck. 1991. "A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data." *International Statistical Review* 59:25-35.

Rasch, G. 1961. "On General Laws and the Meaning of Measurement in Psychology." Pp. 321-33 in *Proceedings of the 4th Berkeley Symposium in Mathematical Statistics and Probability*, Vol. 4, edited by J. Neyman. Berkeley: University of California Press.

Samejima, F. 1969. "Estimation of Latent Ability Using a Response Pattern of Graded Scores." *Psychometrika* 17 (monograph supplement).

Tjur, T. 1982. "A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model." *Scandinavian Journal of Statistics* 9:23-30.

Tutz, G. 1990. "Sequential Item Response Models With an Ordered Response." *British Journal of Mathematical and Statistical Psychology* 43:39-55.

Alan Agresti is a professor in the Department of Statistics, University of Florida. His research interests are in categorical data analysis, particularly for ordered categories. His books include Categorical Data Analysis *(Wiley, 1990) and* Statistical Methods for the Social Sciences, 2nd Edition *(with B. Finlay; Macmillan, 1986).*