

Strategies for Modeling Two Categorical Variables with Multiple Category Choices

Christopher R. Bilder

Department of Statistics,
University of Nebraska-Lincoln, Lincoln, NE 68583
chris@chrisbilder.com, <http://www.chrisbilder.com>

Thomas M. Loughin

Department of Statistics, Kansas State University,
Manhattan, KS 66506

Key words: Alternating logistic regression, Correlated binary data, Generalized loglinear, Marginal independence, Multiple-response categorical variable, Pick any/c

1 Introduction

Surveys frequently have questions that ask subjects to "pick any" or "choose all that apply" from a set of items. Variables that summarize this type of data are called multiple-response (pick any/c) categorical variables. An example where multiple-response categorical variables (MRCVs) arise comes from the Kansas farmer data discussed in Bilder and Loughin (2004). Farmers were instructed to pick all sources of veterinary information that they use from among professional consultant, veterinarian, state or local extension service, magazines, and feed companies and representatives. Farmers were also instructed to pick all swine waste disposal methods they use from among lagoon, pit, natural drainage, and holding tank. Table 1 of Bilder and Loughin (2004) summarizes the responses. In this setting, it may be of interest to test for independence between the two MRCVs. Agresti and Liu (1999) first called this a test for simultaneous pairwise marginal independence (SPMI). The usual Pearson chi-square test for independence should not be used because of farmers' contributing multiple responses. The purpose of this research is to investigate and develop model-based procedures to test for marginal independence between two multiple-response categorical variables.

Little research has been done on testing for independence between two MRCVs. Non-model-based approaches using extensions of Pearson chi-square tests have been proposed by Thomas and Decady (2000) and Bilder and Loughin (2004). Agresti and Liu (1999, 2001) briefly suggest using generalized loglinear models fit by maximum likelihood estimation or the multivariate binomial logit-normal models of Coull and Agresti (2000). These model-based approaches are fully investigated in this paper. Approaches proposed here include using the alternating logistic regression model of Carey, Zeger, and Diggle (1993) and a new way to marginally fit the generalized loglinear model.

2 Notation

Consider the case of two MRCVs generically denoted as W , with r items, and Y , with c items. For example, W corresponds to waste storage (4 items) and Y corresponds to sources of veterinary information (5 items). Survey respondents contribute a vector of binary responses for both MRCVs indicating the items that are selected. For a randomly selected subject, s , let $\mathbf{W}_s = (W_{s1}, \dots, W_{sr})'$ denote the responses for W and $\mathbf{Y}_s = (Y_{s1}, \dots, Y_{sc})'$ denote the responses for Y where the individual item response random variables are $W_{si} = 0$ or 1 and $Y_{sj} = 0$ or 1 according to whether that item is selected by the subject (1 denotes the item is selected). Let $m_{ab(ij)}$ denote the number of ($W_i=a, Y_j=b$) responses where $a = 0$ or 1 and $b = 0$ or 1. The $E(m_{ab(ij)})$ is denoted by $\mu_{ab(ij)}$. The corresponding marginal probability of an a or b response is $P(W_i=a, Y_j=b) = \pi_{ab(ij)}$, and this is equal to $\mu_{ab(ij)}/n$ where n is the number of subjects.

Table 1 shows the Kansas farmer data organized corresponding to this notational format. For example, there are $m_{11(11)} = 34$ farmers who use professional consultant as their source of veterinary information and use lagoon as their waste storage method. Note that each (W_i, Y_j) combination forms a 2×2 marginal sub-table of responses to these items.

Bilder and Loughin (2004) examine the hypothesis of SPMI. This specifies the simultaneous pairwise independence of two groups of binary random variables

and is the extension of the concept of independence from single-response categorical variables to the MRCV case. Specifically, the hypotheses for a test of SPMI are

$$H_0: \pi_{ab(ij)} = \pi_{a\bullet(i\bullet)}\pi_{\bullet b(\bullet j)} \text{ for } i=1,\dots,r, j=1,\dots,c, a=b=0,1$$

$$H_a: \text{At least one equality does not hold,}$$

where $P(W_i=a) = \pi_{a\bullet(i\bullet)}$ and $P(Y_j=b) = \pi_{\bullet b(\bullet j)}$. The hypotheses can also be written in terms of odds ratios. SPMI exists if $OR_{WY,ij} = (\pi_{11(ij)}\pi_{00(ij)})/(\pi_{10(ij)}\pi_{01(ij)}) = 1$ for $i=1,\dots,r$ and $j=1,\dots,c$. In the context of Table 1, SPMI represents simultaneous independence in each of the rc 2×2 sub-tables.

3 Modeling MRCVs and Testing for SPMI

3.1 Generalized Loglinear Models

To model association between W_i and Y_j , separate loglinear models could be fit to each sub-table. However, it is more appropriate to fit one model across all sub-tables since simultaneous inference about the rc sub-tables is of the main interest. The loglinear model under SPMI is

$$\log(\mu_{ab(ij)}) = \lambda_{ij}^{WY} + \lambda_{a(ij)}^{W_i(WY)} + \lambda_{b(ij)}^{Y_j(WY)}$$

where the λ_{ij}^{WY} terms control the sample size to be n in each sub-table, the $\lambda_{a(ij)}^{W_i(WY)}$ terms control the row marginal counts in each sub-table, and the $\lambda_{b(ij)}^{Y_j(WY)}$ controls the column marginal counts in each sub-table. Appropriate restrictions are made on the model parameters to ensure identifiability. Fitting this model to a data set represented like the one in Table 1 creates predicted sub-table counts whose margins match those for the observed sub-tables. Furthermore, the predicted sub-table counts all have odds ratios of 1.

3.1.1 Maximum Likelihood Estimation

There are a few different choices for fitting the model. First, the model can be fit using maximum likelihood estimation with the generalized loglinear model methodology presented in Lang and Agresti (1994). These models are fit to a cross-classification of the multinomial counts for all possible W_s and Y_s . Bilder and Loughin (2004) refer to this cross-classification as a joint table. Pearson and likelihood

ratio goodness-of-fit statistics can then be used to test for SPMI.

With two MRCVs, the number of multinomial counts is 2^{r+c} and each of the corresponding multinomial probabilities need to be estimated under a set of model constraints. When r and/or c are not small, this can result in a large number of parameters that needs to be estimated. For example, there are $2^{4+5} = 512$ parameters that need to be estimated for the Kansas farmer data example. Furthermore, joint tables can be very sparse. For example, the Kansas farmer data example has 434 of its 512 observed multinomial counts equal to 0. As suggested in Agresti and Liu (2001) for a model under SPMI, the large dimension of the multinomial distribution with observed count sparseness can cause model convergence problems. These convergence problems are demonstrated in Section 5.

3.1.2 Marginal Estimation

To avoid fitting a model to the multinomial counts, the generalized loglinear model can be fit using a marginal modeling approach. In this case, the model is fit directly to the data as displayed in Table 1 without regard to the correlated counts among the sub-tables. This is similar to what Haber (1985, p. 2852-3) does for a different problem. Table II in Haber (1985) can be thought of as containing two of the four possible sub-tables. The estimated expected frequencies, $\hat{\mu}_{ab(ij)}$, from the model here are found through solving the usual likelihood equations of $\mathbf{X}'\hat{\boldsymbol{\mu}} = \mathbf{X}'\mathbf{m}$, where $\hat{\boldsymbol{\mu}}$ and \mathbf{m} are $4rc \times 1$ vectors of the corresponding $\hat{\mu}_{ab(ij)}$ and $m_{ab(ij)}$ quantities and \mathbf{X} is a matrix of 0's and 1's relating the model specified expected counts to the observed counts. Since the usual likelihood equations are being used, the model fitting can be performed using software such as PROC GENMOD in SAS or the glm function in R. Parameter estimates are called "pseudo" maximum likelihood estimates by Rao and Scott (1984) in a similar problem because the true likelihood equations are not used. The parameter estimates are consistent since they are functions of $\hat{\pi}_{ab(ij)} = m_{ab(ij)}/n$ which is a consistent estimator for $\pi_{ab(ij)}$.

Since the individual counts from Table 1 are not multinomial counts, the usual loglinear model goodness-of-fit statistics (Pearson and likelihood ratio) do not have asymptotic χ^2 distributions. Instead they are distributed as linear combinations of independent χ^2 random variables. Rao and Scott (1984) specifically discuss the asymptotic distributions for these types of goodness-of-fit statistics. The Pearson statistic is $\sum_{a,b,i,j} (m_{ab(ij)} - \hat{\mu}_{ab(ij)})^2 / \hat{\mu}_{ab(ij)}$ where $\hat{\mu}_{ab(ij)}$ results from fitting the model under SPMI. This statistic is the same as the modified Pearson statistic derived in Bilder and Loughin (2004) since $\hat{\mu}_{ab(ij)} = m_{a\bullet(i\bullet)} m_{\bullet b(\bullet j)} / n$ and the statistic has the same asymptotic distribution.

First and second-order Rao-Scott (1984) adjustments to the Pearson and likelihood ratio test statistics create new statistics whose first and/or second moments are the same as a χ^2 random variable. Bilder and Loughin (2004) find that the first-order adjusted Pearson statistic does not hold the correct size in the SPMI test when there is strong pairwise association among the items of the same MRCV. They also find that the second-order adjusted Pearson statistic performs satisfactorily most of the time, but not quite as well as bootstrap procedures. Because of these findings, we propose a new bootstrap procedure to estimate the sampling distribution of the goodness-of-fit statistics. The resampling involves generating new vectors of binary responses using the algorithm of Gange (1995). Let X^2 be the Pearson or likelihood ratio goodness-of-fit statistic. A bootstrap estimate of the statistic's sampling distribution under SPMI requires generating new vectors of data, say, \mathbf{W}_s^* and \mathbf{Y}_s^* , satisfying SPMI. Resampling the observed $(\mathbf{W}_s, \mathbf{Y}_s)$ vector pairs together would not guarantee the resampling is being performed under SPMI. In the Gange algorithm, only marginal table representations of multinomial counts are needed in order to generate data from the full multinomial distribution. Thus, the following bootstrap procedure can be used to estimate the distribution of X^2 under SPMI:

1. Find $\hat{\mu}$ through using the usual likelihood equations and calculate X^2 .
2. Find the observed 2×2 tables for each $(W_i, W_{i'})$ ($i < i'$) and $(Y_j, Y_{j'})$ ($j < j'$) response pair.

3. With the predicted and observed counts from 1. and 2., use the algorithm of Gange (1995) to simulate B resamples of \mathbf{W}_s^* and \mathbf{Y}_s^* for $s=1, \dots, B$.
4. Fit the model again to each resample and calculate the goodness-of-fit statistic, say X_b^2 , for $b=1, \dots, B$.
5. Calculate the p-value as $B^{-1} \sum_{b=1}^B I(X_b^2 \geq X^2)$ where $I(\cdot)$ is the indicator function.

Notice that the proposed procedure does not rely on the multinomial counts being available, but rather all pairwise marginal count tables between items, both within and across W and Y . This is helpful in situations where only the pairwise marginal tables are available for analysis. Furthermore, the proposed resampling procedure is more general than the resampling procedure discussed in Bilder and Loughin (2004) because it can be used with models without the SPMI restriction. These types of models are to be discussed in a future manuscript.

The Gange algorithm performs the resampling under a special case of SPMI called joint independence (see Bilder and Loughin, 2004). Through simulations described in Section 5 and from past research where resampling has been done under joint independence (see Bilder, Loughin, and Nettleton, 2000; Bilder and Loughin, 2002; and Bilder and Loughin, 2004), there does not appear to be any detrimental effect due to resampling under a special case of the null hypothesis.

When there is a deviation from the model, the examination of standardized Pearson residuals can provide insight to where the deviation occurs. The standardized Pearson residuals can be derived in a similar manner as to what is outlined in Rao and Scott (1984, p. 49) with minor modifications similar to those given in Haber (1985, p. 2852). The details are excluded from here due to space constraints.

3.2 Marginal Logit Models Fit by GEE

Methodology

Another way to avoid having to estimate the entire multinomial distribution is to fit marginal logit models using generalized estimating equation (GEE)

methodology. This model is somewhat similar to the one proposed by Agresti and Liu (1999, p. 937) for the multiple marginal independence problem. The specification of SPMI here involves pairwise probabilities for W_i and Y_j . Therefore, instead of modeling only the marginal expectations of W_i and Y_j , the associations for W_i and Y_j pairs must also be modeled.

There have been many proposals for how to model the marginal expectations and pairwise associations of random variables using GEE methodology. See Pendergast et al. (1996, Section 5) for a review. We choose the alternating logistic regression model of Carey, Zeger, and Diggle (1993) for the SPMI problem. This procedure is easier to implement when there are a large number of binary responses than other procedures such as the one proposed by Prentice (1988).

For the SPMI testing problem, the alternating logistic regression model simultaneously estimates $\text{logit}(\pi_{1\bullet(i\bullet)}) = \beta_i$ and $\text{logit}(\pi_{\bullet 1(j\bullet)}) = \alpha_j$ where β_i denotes the effect of W_i and α_j denotes the effect of Y_j . The log odds ratios for each pairwise combination of items are also estimated. This involves odds ratios for each W_i and $W_{i'}$ pair ($i < i'$), each Y_j and $Y_{j'}$ pair ($j < j'$), and each W_i and Y_j pair. A Wald statistic is constructed to test whether $\log(\text{OR}_{WY,ij}) = 0$ for all W_i and Y_j pairs. If SPMI is rejected, the individual estimates for the $\log(\text{OR}_{WY,ij})$ can be examined to determine which item pairs deviate from SPMI.

As for the log odds ratios for the W_i and $W_{i'}$ pairs and Y_j and $Y_{j'}$ pairs, the most nonrestrictive assumption is to estimate each of them individually. This means a total of ${}_{r+c}C_2$ log odds ratio parameters are estimated. Because only rc odds ratios are needed for the W_i and Y_j pairs, various assumptions about the other odds ratios could be made. For example, the odds ratios for each W_i and $W_{i'}$ pair could be set to be equal and those for each Y_j and $Y_{j'}$ pair could also be set to be equal. This assumption reduces the number of parameters that are estimated and can help reduce the frequency of model convergence problems (to be discussed in Sections 4 and 5). In many realistic situations, however, these odds ratios actually are unequal. Applying false

assumptions may adversely affect the fit of the model and the quality of the resulting test.

The marginal logit model can be fit in SAS (version 7 or higher) using PROC GENMOD. The ALRCON8 SAS/IML module (available at <http://ftp.sas.com/techsup/download/stat/alrcon8.html>) can be used to perform the SPMI test.

3.3 Other Modeling Approaches

Agresti and Liu (2001) suggest that the multivariate binomial logit-normal model of Coull and Agresti (2000) could also be used to test for SPMI. A $(r+c) \times 1$ vector of random effects can be used for the r items of W and the c items of Y . The null hypothesis model under SPMI specifies the correlations between the W and Y random effects to be 0. The alternative hypothesis model does not specify the correlations. Using maximum likelihood estimation, the fit of these two models can be compared with a likelihood ratio test statistic to test for SPMI. The main advantage for using this type of random effects model over simpler ones is it allows for negative dependence to occur between item responses. However, this needed advantage makes the model more difficult to fit. Test data sets with $r=c=2$ were simulated and the multivariate binomial logit-normal model was fit using PROC NLMIXED in SAS. In every case, the Gauss-Hermite quadrature procedure had extreme difficulty performing the four-dimensional integration. Coull (personal communication, 2002) has had similar difficulties with fitting this type of model to MRCV data. Future research could look into other numerical integration techniques that would allow the model to be fit.

Another modeling approach involves fitting a loglinear model to a joint table where the rows are considered one single-response variable representing all possible combinations of W_i responses and the columns are considered one single-response variable representing all possible combinations of Y_j responses. In this setting, suppose the row variable is denoted by R with 2^r levels and the column variable is denoted by C with 2^c levels. Using the common loglinear model notation, the model (R, C) could be fit to the data in this form. A test

for independence between R and C can be constructed using a likelihood ratio test comparing the fit of this model to (RC). The resulting test is for joint independence which is discussed in Section 2 of Bilder and Loughin (2004). Joint independence is a special case of SPMI so it is possible to reject joint independence while satisfying SPMI. Thus, this procedure is not recommended for testing SPMI.

4 Application to the Kansas Farmer Data

The modeling procedures discussed in Sections 3.1 and 3.2 are applied to the Kansas farmer data set and tests for SPMI are performed. Convergence for the generalized loglinear model fit by maximum likelihood estimation is not obtained after 150 iterations. To help alleviate possible problems with the 0 multinomial counts, 0.5 was added to each cell of the joint table and the generalized loglinear model was refit to the data set; however, convergence was still not obtained. The convergence criterion used is the difference in likelihood ratio test statistics for successive iterations being less than 0.001 and the norm of the difference between estimated joint table cell counts for successive iterations being less than 0.001.

The generalized loglinear model fit using the marginal modeling approach did not have the same convergence problems. The parameter estimates converged and the resulting Pearson goodness-of-fit statistic was 62.31. The bootstrap p-value was 0.0002 using B=10,000 resamples and the second-order Rao-Scott adjusted p-value is 3.07×10^{-5} . Both procedures show strong evidence against SPMI. The likelihood ratio test statistic produced similar results. To investigate the cause of the deviation from SPMI, the standardized Pearson residuals are calculated and given in Table 2. The (W_i, Y_j) pairwise combinations that show deviations from SPMI are highlighted in the table. The largest deviation from SPMI occurs in the lagoon and professional consultant sub-table.

For the marginal logit model, there are ${}^9C_2 = 36$ different log odds ratios to estimate for the item pairs. Unfortunately, the parameter estimation procedure does not converge. To reduce the number of parameters to

be estimated, the odds ratios for each W_i and $W_{i'}$ pair are set to be equal and each Y_j and $Y_{j'}$ pair are set to be equal. This reduces the number of log odds ratios to ${}^9C_2 - {}^4C_2 - {}^5C_2 + 2 = 22$ where the 2 represents the one log odds ratio for all W_i and $W_{i'}$ pairs and one for all Y_j and $Y_{j'}$ pairs. The observed odds ratios for each W_i and $W_{i'}$ pair are between 0 and 0.34. The observed odds ratios for each Y_j and $Y_{j'}$ pair are between 0.71 and 2.73 with a few 95% confidence intervals for the true odds ratios not overlapping. This suggests that the assumption of equality of odds ratios may not be valid, but it is made here to illustrate the method. The Wald test p-value is approximately 3.3×10^{-7} indicating evidence against SPMI. To investigate the cause of the deviation from SPMI, individual Wald tests and the corresponding p-values using a χ_1^2 approximation are calculated. Using a 0.05 significance level, the significant pairwise combinations are the same as those found with the standardized Pearson residuals in Table 2. If a Bonferroni adjusted significance level of $0.05/20 = 0.0025$ is used instead, only the lagoon and professional consultant pair has a smaller p-value.

5 Simulation Study

To investigate how well the model-based approaches test for SPMI, a simulation study is performed. The data is simulated using the data generation algorithm of Gange (1995). Under SPMI, all OR_{W_i, Y_j} are set to 1. Odds ratios between each W_i and $W_{i'}$ pair ($i < i'$) denoted by $OR_{W_{ii'}}$ and each Y_j and $Y_{j'}$ pair ($j < j'$) denoted by $OR_{Y_{jj'}}$ are controlled as well, but not necessarily at a level of 1. For each set of simulation settings, 500 simulated data sets are generated and the testing procedures are applied to each data set using a 0.05 significance level. These simulated data sets are the same as those used for the simulation study in Bilder and Loughin (2004). The proportion of times a test rejects SPMI is recorded to estimate the type I error rate. The number of resamples used with the bootstrap is 1,000.

Table 3 gives the estimated type I error rates. Model convergence rates are also given for the models where convergence was not always obtained. The estimated

type I error rates are always calculated conditionally on the data sets where convergence is obtained. The approximate 95% expected range of estimated type I error rates for testing procedures holding the correct size is $0.05 \pm 1.96\sqrt{0.05(1-0.05)/C}$ where C is the number of simulated data sets for which a test is performed. For the testing procedures where convergence is always obtained, $C = 500$ provides a 95% expected range of (0.031, 0.069).

For the generalized loglinear model fit by maximum likelihood estimation, the same convergence criterion as described in Section 4 is used here. In the $r=c=2$ simulations, 0.5 is added to each cell of the joint table for the first 10 iterations to help with convergence. The convergence rates are between 87% and 99%. Adding 0.5 to each cell for all iterations improves the convergence considerably, but the estimated type I error rates are even more conservative than those given in Table 3 so they are not shown here. Only the likelihood ratio test statistic results are given in Table 3 since it had estimated type I error rates within the 95% expected range more often than the Pearson statistic; however, the estimated type I error rates are still lower than the 95% expected range for the smaller sample size simulations. For the $r=c=5$ simulations, the model is fit to a test set of twenty simulated data sets using $OR_{W_{i,r}} = OR_{Y_{j,r}} = 2$ and $n = 500$. For each simulated data set, convergence is not achieved. Due to these problems, the entire set of $r=c=5$ simulations are not performed and maximum likelihood estimation should be avoided in this setting.

For the generalized loglinear model fit through the marginal modeling approach, the estimated type I error rates are all within the expected range. This suggests the procedure is holding the correct size. Additional simulations (not shown) where the joint independence hypothesis was not satisfied were also performed since the resampling is performed under joint independence (the special case of SPMI). Again, the procedure appears to hold the correct size for these simulations.

Convergence problems are also evident for the marginal logit models. Whenever a table summarizing the pairwise responses between two items has a 0 cell

count in it, the model which tries to estimate all ${}_{r+c}C_2$ log odds ratio parameters does not converge. This is true not only for W_i and Y_j response pairs, but also for W_i and $W_{i'}$ pairs and Y_j and $Y_{j'}$ pairs. Similar to what was done in Section 4 for these models, the odds ratios for each W_i and $W_{i'}$ pair are set to be equal and each Y_j and $Y_{j'}$ pair are set to be equal to reduce the number of parameters estimated. In the $r=c=2$ simulations, the number of log odds ratios estimated is unchanged by this strategy, but the number estimated for the $r=c=5$ simulations is reduced from ${}_{10}C_2 = 45$ to ${}_{10}C_2 - {}_5C_2 - {}_5C_2 + 2 = 27$. This helps to increase the convergence rate, but achieving convergence still is a problem for $n \leq 300$. For the models that did converge in the $r=c=2$ simulations, the estimated type I error rates are close to the nominal level except for the smaller sample sizes. For most of the $r=c=5$ simulations, the method does not hold the correct size. Since the model assumes equality of each association for the W_i and $W_{i'}$ pairs and Y_j and $Y_{j'}$ pairs, additional simulations (not shown) for $r=c=5$ marginal tables are also performed with $OR_{W_{i,r}} = OR_{Y_{j,r}} = 2$ for eight of the odds ratios and $OR_{W_{i,r}} = OR_{Y_{j,r}} = 25$ for twelve of the odds ratios to examine sensitivity to the false odds ratio assumption. Estimated type I error rates are similar to those where the odds ratios are the same for each W_i and $W_{i'}$ pair and each Y_j and $Y_{j'}$ pair.

6 Discussion

This paper shows some of the common approaches for modeling correlated binary data have limited use when one wants to test for SPMI. Of the models presented here, only the generalized loglinear model fit using a marginal modeling strategy appears to provide a general method for fitting the model and testing for SPMI. Both this modeling approach and the non-model-based approaches recommended in Bilder and Loughin (2004) consistently work and hold the correct size for the SPMI test.

ACKNOWLEDGEMENTS

This research was supported in part by NSF grants SES-0207212 and SES-0233321.

REFERENCES

- Agresti, A. and Liu, I.-M. (1999). Modeling a categorical variable allowing arbitrarily many category choices. *Biometrics* 55, 936-943.
- Agresti, A. and Liu, I.-M. (2001). Strategies for modeling a categorical variable allowing multiple category choices. *Sociological Methods & Research* 29, 403-434.
- Bilder, C. R. and Loughin, T. M. (2002). Testing for conditional multiple marginal independence. *Biometrics* 58, 200-208.
- Bilder, C. R. and Loughin, T. M. (2004). Testing for Marginal Independence Between Two Categorical Variables with Multiple Responses. To appear in *Biometrics*.
- Bilder, C. R., Loughin, T. M., and Nettleton, D. (2000). Multiple marginal independence testing for pick any/c variables. *Communications in Statistics: Simulation and Computation* 29(4), 1285-1316.
- Carey, V., Zeger, S. L., and Diggle, P. (1993). Modelling multivariate binary data with alternating logistic regressions. *Biometrika* 80, 517-526.
- Coull, B. A. and Agresti, A. (2000). Random effects modeling of multiple binomial responses using the multivariate binomial logit-normal distribution. *Biometrics* 56, 73-80.
- Gange, S. J. (1995). Generating multivariate categorical variates using the iterative proportional fitting algorithm. *The American Statistician* 49, 134-138.
- Haber, M. (1985). Log-linear models for correlated marginal totals of a contingency table. *Communications in Statistics: Theory and Methods* 14(12), 2845-2856.
- Lang, J. and Agresti, A. (1994). Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *Journal of the American Statistical Association* 89, 625-632.
- Pendergast, J. F., Gange, S. J., Newton, M. A., Lindstrom, M. J., Pakta, M., and Fisher, M. R. (1996). A survey of methods for analyzing clustered binary response data. *International Statistical Review* 64, 89-118.
- Prentice, R. L. (1988). Correlated binary regression with covariates specific to each binary observation. *Biometrics* 44, 1033-1048.
- Rao, J. N. K. and Scott, A. J. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *The Annals of Statistics* 12(1), 46-60.
- Thomas, D. R. and Decady, Y. J. (2000). Analyzing categorical data with multiple responses per subject. *Statistics Society of Canada Proceedings of the Survey Methods Section*, 121-130.

Table 1. Pairwise responses for the Kansas farmer data.

	Sources of veterinary information										
	Professional consultant		Veterinarian		State or local ext. service		Magazine		Feed company & rep.		
	1	0	1	0	1	0	1	0	1	0	
Lagoon	1	34	109	54	89	50	93	63	80	41	102
Pit	0	10	126	36	100	45	91	68	68	52	84
Natural drainage	1	17	63	33	47	34	46	43	37	37	43
Holding tank	0	27	172	57	142	61	138	88	111	56	143
Waste Storage	1	6	79	23	62	30	55	49	36	34	51
Waste Storage	0	38	156	67	127	65	129	82	112	59	135
Waste Storage	1	1	12	4	9	4	9	6	7	2	11
Waste Storage	0	43	223	86	180	91	175	125	141	91	175

Table 2. Standardized Pearson residuals for the loglinear model fit through marginal modeling.

	Sources of veterinary information											
	Professional consultant		Veterinarian		State or local ext. service		Magazine		Feed company & rep.			
	1	0	1	0	1	0	1	0	1	0		
Lagoon	1	3.91	-3.91	2.04	-2.04	0.33	-0.33	-1.00	1.00	-1.70	1.70	
	0	-3.91	3.91	-2.04	2.04	-0.33	0.33	1.00	-1.00	1.70	-1.70	
Pit	1	1.48	-1.48	1.97	-1.97	1.84	-1.84	1.44	-1.44	2.79	-2.79	
	0	-1.48	1.48	-1.97	1.97	-1.84	1.84	-1.44	1.44	-2.79	2.79	
Natural drainage	1	-3.11	3.11	-1.26	1.26	0.29	-0.29	2.38	-2.38	1.53	-1.53	
	0	3.11	-3.11	1.26	-1.26	-0.29	0.29	-2.38	2.38	-1.53	1.53	
Holding tank	1	-1.06	1.06	-0.12	0.12	-0.26	0.26	-0.06	0.06	-1.64	1.64	
	0	1.06	-1.06	0.12	-0.12	0.26	-0.26	0.06	-0.06	1.64	-1.64	

<-1.96, >1.96
<-2.576, >2.576

Table 3. Estimated type I error and model convergence rates using the same simulation settings as in Bilder and Loughin (2004); the marginal probabilities are $(\pi_{1\bullet(1\bullet)}, \pi_{1\bullet(2\bullet)})' = (0.4, 0.5)'$ and $(\pi_{\bullet 1(1\bullet)}, \pi_{\bullet 1(2\bullet)})' = (0.2, 0.3)'$ for $r=c=2$ and $(\pi_{1\bullet(1\bullet)}, \dots, \pi_{1\bullet(5\bullet)})' = (\pi_{\bullet 1(1\bullet)}, \dots, \pi_{\bullet 1(5\bullet)})' = (0.1, 0.2, 0.3, 0.4, 0.5)'$ for $r=c=5$. Shaded cells correspond to estimated type I error rates outside of the 95% expected range.

r = c	OR _{w,ir} = OR _{v,ij'}	n	Gen. loglinear model MLE	Gen. loglinear model marg.	Marginal logit model GEE	Gen. loglinear model MLE convergence	Marginal logit model GEE convergence
			2	12	0.015	0.046	0.085
2	2	25	0.033	0.054	0.074	0.904	0.648
		50	0.062	0.064	0.044	0.966	0.964
		100	0.065	0.058	0.052	0.992	0.998
		12	0.005	0.056	0.097	0.872	0.062
25	25	25	0.034	0.062	0.025	0.954	0.318
		50	0.056	0.064	0.060	0.962	0.704
		100	0.049	0.054	0.056	0.970	0.958
		100		0.040	0.556		0.468
2	2	300		0.040	0.137		0.946
		500		0.062	0.060		0.990
		100		0.050	0.735		0.544
25	25	300		0.042	0.172		0.940
		500		0.048	0.109		0.990