Advances and Applications in Statistics
Volume 17, Number 1, 2010, Pages 29-40
Published Online: September 28, 2010
This paper is available online at http://pphmj.com/journals/adas.htm

LIKELIHOOD INFERENCES FOR CORRELATED NOMINAL DATA WITHOUT KNOWING THE INTRA-CLUSTER CORRELATIONS

CHI-CHUAN YANG[1] and TSUNG-SHAN TSOU[1,2,3,*]

[1]Institute of Statistics
National Central University
Jhongli 320, Taiwan

[2]Institute of Systems Biology and Bioinformatics
National Central University
Jhongli 320, Taiwan

[3]Cathay General Hospital
Taipei 106, Taiwan
e-mail: tsou@mx.stat.ncu.edu.tw

Abstract

The multinomial distribution is the main statistical model for categorical data, including ordinal and nominal responses. A prerequisite assumption underlying this model is the independence between the individuals constituting the total number of the multinomial. This article shows that the multinomial model could be made asymptotically valid no matter how individuals are correlated. We focus on likelihood inference for

2010 Mathematics Subject Classification: 62F35, 62J12, 62P10.

Keywords and phrases: multinomial, robust likelihood, polytomous logistic regression.

This work is supported by grant NSC 98-2118-M-008-002-MY2 of the National Science Council, and National Central University-Cathay General Hospital Joint Research grant 98CGH-NCU-A1, Taiwan, R.O.C.

[*]Corresponding author

Received May 11, 2010

the regression parameter of interest associated with the polytomous logistic regression model. Simulations and several examples are used to demonstrate the efficacy of proposed parametric robust method.

1. Introduction

Correlated nominal data are commonly encountered in many research areas including biomedical or toxicology studies. The multinomial distribution is the conventional statistical model for nominal responses and polytomous logistic regression is the typical regression model for regression analysis uncovering the relationship between nominal cell probabilities and covariates (Agresti [1]).

The analysis of clustered correlated categorical data is generally strenuous because of lacking suitable joint distributions for the correlated responses. A popular model for correlated nominal data that one frequently resorts to is the Dirichlet-multinomial distribution. For instance, Paul et al. [9] used it to compare cell probabilities and derived explicit expressions for the exact Fisher information matrix for the parameters. Wilson and Chen [15] further considered generalized Dirichlet-multinomial distribution to cope with varying response rates over time.

Hedeker [4] suggested a mixed-effects multinomial logistic regression model for analyzing cluster nominal data. His method allows flexible choice of contrasts used to represent comparisons across the response categories. On the other hand, Zhang et al. [16] employed the Bayesian method to deal with multivariate nominal measures through multivariate multinomial probit models. Recently, Lee and Mercante [5] modeled longitudinal nominal data using a Markovian dependence structure.

The abovementioned parametric approaches, despite their usefulness, unavoidably suffer from the drawback of being sensitive to model misspecifications. Should any of the model assumptions involved fail, the model-based inference is generally fallacious.

The generalized estimating equations (GEE, Liang and Zeger [6]) methodology has become one of the most popular methods for analyzing correlated data, since its introduction in 1986. We can conveniently obtain consistent regression parameter estimates and their standard errors simply having the first moments correctly specified. In spite of its overwhelming popularity, the GEE, being a semi-parametric approach, fails to provide a likelihood function. Full likelihood inference tools, such as the likelihood ratio (LR) test, the score test and the goodness of fit test are usually unobtainable from GEE.

Tsou and Shen [14] proposed a robust likelihood approach for making inference about regression parameters for correlated ordinal data modeled by proportional odds model. They showed how we could fix the multinomial working model to encompass intra-cluster correlations. The naïve likelihood based on the working model, when properly adjusted, could be converted into an asymptotically valid likelihood for the regression parameter of interest. The validity of the resultant robust likelihood requires no knowledge of the true underlying correlation structures nor the joint distributions.

In this paper, our focus is on correlated nominal data. We demonstrate how to correct the multinomial model in order to develop a robust likelihood for the regression parameters. The rest of the paper is organized as follows: We give a brief introduction to the robust likelihood method by Royall and Tsou [11] in Section 2. Section 3 contains details, we need to correct the multinomial likelihood to become robust. Sections 4 and 5 contain simulations and real data analysis demonstrating the efficacy of the proposed method. Some concluding remarks are made in Section 6.

2. Robust Likelihood

The idea of the robust likelihood methodology proposed by Royall and Tsou [11] has been explored and extended to various inferential problems. We only sketch the theory and will not reiterate it in details. Interested readers can refer themselves to Tsou and Shen [14].

Consider observations $y_1, ..., y_n$ regarded as realizations of independent random variables $Y_1, ..., Y_n$. Let $l(\theta, \phi)$ denote the log likelihood function based on the working model f, where θ and ϕ are, respectively, the parameter of interest and the fixed-dimensional vector of nuisance parameters. Let θ_0 and ϕ_0 denote the limiting values of the working model-based (or simply model-based) maximum likelihood (ML) estimators, say, $\hat{\theta}$ and $\hat{\phi}$. The convergence referred to here and henceforth is of the weak mode.

Now suppose $y_1, ..., y_n$ are in reality generated from, say, $\{h_i = h(\cdot; \tau_i(\theta, \lambda)), i = 1, ..., n\}$, where λ represents the nuisance parameter (or fixed-dimensional parameters) for h. Here we assume regularity conditions on f and h that ensure the asymptotic normality of the ML estimators.

Assume that the true model is actually h while θ_0 remains the true value of the parameter of interest. Royall and Tsou [11] demonstrated that working likelihoods that fulfill this condition can be converted to be asymptotically legitimate under model misspecification.

Denote the elements of the model-based Fisher information matrix by $[(I_{\theta\theta}, I_{\theta\phi})^t, (I_{\phi\theta}, I_{\phi\phi})^t]^t$ calculated by the second derivatives of the log likelihood function. Its counterpart, calculated according to the products of the first derivatives of the log likelihood, is denoted by $[(V_{\theta\theta}, V_{\theta\phi})^t, (V_{\phi\theta}, V_{\phi\phi})^t]^t$. For instance, $I_{\theta\phi} = \lim_{n\to\infty} E[-\partial^2 l(\theta_0, \phi_0)/\partial\theta\partial\phi^t]/n$ and

$$V_{\theta\phi} = \lim_{n\to\infty} E[\partial l(\theta_0, \phi_0)/\partial\theta \times \partial l(\theta_0, \phi_0)/\partial\phi^t/n],$$

where E denotes expectation taken under h.

Now define $A = I_{\theta\theta} - I_{\theta\phi}I_{\phi\phi}^{-1}I_{\phi\theta}$ and $B = V_{\theta\theta} - 2I_{\theta\phi}I_{\phi\phi}^{-1}V_{\phi\theta} + I_{\theta\phi}I_{\phi\phi}^{-1}V_{\phi\phi}I_{\phi\phi}^{-1}I_{\phi\theta}$. These two quantities play the key roles for the modification of the working likelihood and can be estimated by their empirical versions, say, $\hat{A}$ and $\hat{B}$, by simply substituting $\hat{\theta}$ and $\hat{\phi}$ for their respective targets and plugging in second moment estimates where needed. In the meantime, let $\phi(\theta)$ denote the constrained ML estimate of ϕ given θ. Then the function $(\hat{A}/\hat{B})l(\theta, \phi(\theta))$ is the robust likelihood function for θ, see Royall and Tsou [11] and also Royall [10].

The adjusted robust LR test, the robust score test and other likelihood-based statistical tools could be obtained by operating on this robust likelihood function. For example, the adjusted robust LR test statistic $2(\hat{A}/\hat{B})\{l(\hat{\theta}, \phi(\hat{\theta})) - l(\theta_0, \phi(\theta_0))\}$ is asymptotically χ_1^2 distributed, whereas the naïve version of the statistic $2\{l(\hat{\theta}, \phi(\hat{\theta})) - l(\theta_0, \phi(\theta_0))\}$ is χ_1^2 only if the model assumption is correct.

It is emphasized that A^{-1} is the model-based variance of $\sqrt{n}\hat{\theta}$ in the presence of ϕ (Cox and Hinkley [3]). In contrast $A^{-1}BA^{-1}$ is the robust variance of $\sqrt{n}\hat{\theta}$, that is, asymptotically legitimate under model misspecifications (Stafford [12]).

3. The Multinomial Model

Let $Z_{i,j}$, the *j*th subunit of the *i*th cluster, indicate the nominal outcome, taking on nominal values 1 to *H* for $j = 1, 2, ..., q_i$, $i = 1, 2, ..., n$. Let $\pi_{i,h}$ represent the probability $P(Z_{i,j} = h | \boldsymbol{x}_i)$, given *p*-dimensional covariate $\boldsymbol{x}_i$ in the *i*th cluster. Now consider the polytomous logistic regression model,

$$\log\left(\frac{\pi_{i,h}}{\pi_{i,1}}\right) = \tau_{i,h} = \boldsymbol{x}_i^t \boldsymbol{\beta}_h, \quad h = 2, 3, ..., H,$$

where $\boldsymbol{\beta}_h = (\beta_{h,0}, \beta_{h,1}, ..., \beta_{h,p-1})^t$, $\boldsymbol{x}_i = (x_{i,0}, x_{i,1}, ..., x_{i,p-1})^t$ and $x_{i,0} \equiv 1$. Here $\tau_{i,1}$ is zero so that

$$\pi_{i,h} = \tau_{i,h} \bigg/ \sum_{h=1}^{H} \tau_{i,h}, \quad h = 1, 2, ..., H.$$

Now let $Y_{i,j,h}$ be the binary indicator, whose value is 1 if $Z_{i,j} = h$ and 0 otherwise, for $h = 1, 2, ..., H$. Note that $\sum_{h=1}^{H} Y_{i,j,h} = 1$ and $(Y_{i,j,1}, Y_{i,j,2}, ..., Y_{i,j,H})^t$ follows a multivariate distribution with probabilities $(\pi_{i,1}, \pi_{i,2}, ..., \pi_{i,H})^t$.

Define the partial sum $Y_{i,+,h} = \sum_{j=1}^{q_i} Y_{i,j,h}$. Under the multinomial working model, the probability function for $(Y_{i,+,1}, Y_{i,+,2}, ..., Y_{i,+,H})^t$ is $q_i! \prod_{h=1}^{H} \pi_{i,h}^{y_{i,+,h}} / y_{i,+,h}!$ and the log likelihood function ℓ is proportional to

$$\sum_{i=1}^{n} \sum_{h=1}^{H} y_{i,+,h} \log(\pi_{i,h}). \tag{1}$$

Notice that when $Z_{i,j}$, $j = 1, ..., q_i$ are dependent, (1) is no longer a valid likelihood function.

Notice that the score functions for $\boldsymbol{\beta}_k$, $k = 2, 3, ..., H$ are

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}_k} = \sum_{i=1}^{n} \boldsymbol{x}_i \sum_{h=1}^{H} y_{i,+,h} \frac{1}{\pi_{i,h}} \frac{\partial \pi_{i,h}}{\partial \tau_{i,k}} = \sum_{i=1}^{n} \boldsymbol{x}_i (y_{i,+,k} - q_i \pi_{i,k}). \tag{2}$$

An immediate result is that the expected values of $\partial\ell/\partial\boldsymbol{\beta}_k$ equal zero so long as $q_i\pi_{i,k}$ is the correct mean of $Y_{i,+,k}$, regardless of whether or not $Z_{i,j}$, $j = 1, \ldots, q_i$ are independent. This guarantees that the solutions of $\partial\ell/\partial\boldsymbol{\beta}_k = 0$, say $\hat{\pi}_{i,k}$, remain consistent as long as $E(Y_{i,+,k}) = q_i\pi_{i,k}$ (McCullagh [7]). We can hence modify (1) to become legitimate for general correlated nominal responses.

We can also easily see that

$$\frac{1}{\pi_{i,h}}\frac{\partial\pi_{i,h}}{\partial\tau_{i,k}} = -\pi_{i,k}, \text{ for } h = 1, 2, \ldots, H, k = 2, 3, \ldots, H, k \neq h,$$

$$\frac{1}{\pi_{i,h}}\frac{\partial\pi_{i,h}}{\partial\tau_{i,h}} = 1 - \pi_{i,h}, \text{ for } h = 2, \ldots, H, \text{ and } \pi_{i,1} = 1 - \sum_{h=2}^{H}\pi_{i,h}.$$

Routine calculations lead to three $p \times p$ matrices,

$$I_{\beta_k\beta_s} = -\lim_{n\to\infty}\sum\nolimits_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^t q_i\pi_{i,k}\pi_{i,s}/n,$$

$$I_{\beta_k\beta_k} = \lim_{n\to\infty}\sum\nolimits_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^t q_i\pi_{i,k}(1 - \pi_{i,k})/n$$

and

$$V_{\beta_k\beta_u} = \lim_{n\to\infty}\sum\nolimits_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^t Cov(Y_{i,+,k}, Y_{i,+,u})/n, \text{ for } u, k \neq s \in \{2, \ldots, H\}.$$

These then constitute the elements of A and B for any chosen regression parameter of interest.

Notice that the above I and V terms could be conveniently estimated by substituting ML estimates for the regression parameters and replacing second moments, such as $Cov(Y_{i,+,k}, Y_{i,+,u})$ with the empirical version $(Y_{i,+,k} - q_i\hat{\pi}_{i,k}) \cdot (Y_{i,+,u} - q_i\hat{\pi}_{i,u})$. Here $\hat{\pi}_{i,u}$ and $\hat{\pi}_{i,k}$ are the ML estimates of $\pi_{i,u}$ and $\pi_{i,k}$, respectively. We can hence derive $\hat{I}$ and $\hat{V}$, the empirical versions of above I and V, which then lead to consistent estimate $\hat{A}$ and $\hat{B}$ of A and B, respectively.

4. Simulation Studies

In this section, we use simulations to demonstrate the merit of robust likelihood method. The number of response levels H is set to be 3 and the number of clusters n employed includes 30, 50, 100 and 200. Polytomous logistic regression models, $\log(\pi_{i,h}/\pi_{i,1}) = \beta_{h,0} + x_i\beta_{h,1}$, $h = 2, 3$, $i = 1, 2, ..., n$ are ascertained, with x_i independently and uniformly sampled from the interval $[-1, 1]$. The cluster size q_i is generated uniformly from integers in [2, 5]. The coefficient $\beta_{3,1}$, designated as θ for convenience, is the parameter of interest and $(\beta_{2,0}, \beta_{2,1}, \beta_{3,0}, \beta_{3,1})^t$ = $(-0.5, 0.4, -0.5, 0.6)^t$ are the true values employed.

The correlated responses $(y_{i,+,1}, y_{i,+,2}, y_{i,+,3})^t$ are generated from the Dirichlet-multinomial distribution with probability density function

$$q_i!\frac{\Gamma(\kappa)}{\Gamma(\kappa + q_i)}\prod_{h=1}^{3}\frac{\Gamma(y_{i,+,h} + \kappa\pi_{i,h})}{y_{i,+,h}!\,\Gamma(\kappa\pi_{i,h})}.$$

The parameter $\kappa(> 0)$ is related to the implied intra-cluster correlation and the second moments by $Var(Y_{i,+,h}) = (q_i + \kappa)q_i\pi_{i,h}(1 - \pi_{i,h})/(1 + \kappa)$ and

$$Cov(Y_{i,+,h}, Y_{i,+,k}) = -(q_i + \kappa)q_i\pi_{i,h}\pi_{i,k}/(1 + \kappa), \text{ for } h \neq k \in \{1, 2, 3\},$$

respectively. The values 10, 5, 1 and 0.2 are selected for κ in the simulations. Note that the larger κ is, the less $Y_{i,+,h}$ and $Y_{i,+,k}$ are correlated.

Table 1 displays the empirical type I error probabilities committed by the naïve and the robust LR tests, denoted by α_{na} and α_{adj}, respectively, for testing the null hypothesis $H_0 : \theta = 0.6$ with a nominal level of 0.05. Also, included in the tabulation are the average of the 3,000 simulated $\hat{\theta}$ values and the corresponding sample variance, denoted by $m(\hat{\theta})$ and $s^2(\hat{\theta})$, respectively. We also exhibit the average of the robust and the naïve asymptotic variance estimates of $\hat{\theta}$, denoted by $\hat{V}_{adj}(\hat{\theta})$ and $\hat{V}_{na}(\hat{\theta})$, respectively. Recall that the former and the latter are calculated based on $\hat{B}/(n\hat{A}^2)$ and $1/(n\hat{A})$, respectively.

Table 1. The efficacy of robust likelihood

N	κ	$m(\hat{\theta})$	$s^2(\hat{\theta})$	$\hat{V}_{na}(\hat{\theta})$	$\hat{V}_{adj}(\hat{\theta})$	α_{na}	α_{adj}
30	10	0.6338	0.4203	0.3258	0.3685	0.0830	0.0813
	5	0.6285	0.4901	0.3288	0.4262	0.1050	0.0807
	1	0.6461	0.8680	0.3489	0.7190	0.2053	0.0913
	0.2	0.7138	1.3570	0.4030	1.0970	0.2697	0.0767
50	10	0.6136	0.1829	0.1447	0.1697	0.0770	0.0597
	5	0.6025	0.2121	0.1456	0.1951	0.1033	0.0660
	1	0.6440	0.3681	0.1510	0.3189	0.2080	0.0767
	0.2	0.6511	0.5646	0.1578	0.4623	0.2957	0.0797
100	10	0.6021	0.0851	0.0674	0.0825	0.0877	0.0627
	5	0.6207	0.0992	0.0678	0.0964	0.1010	0.0507
	1	0.6173	0.1727	0.0686	0.1618	0.2137	0.0637
	0.2	0.6228	0.2292	0.0699	0.2310	0.2717	0.0507
200	10	0.6127	0.0370	0.0299	0.0371	0.0777	0.0520
	5	0.6034	0.0455	0.0300	0.0436	0.1187	0.0573
	1	0.6027	0.0710	0.0302	0.0718	0.2013	0.0553
	0.2	0.6056	0.1011	0.0304	0.1012	0.2857	0.0520

Obviously, the ML estimate under model misspecifications is close to the true value, 0.6, owing to the property of consistency. Moreover, the adjustment $\hat{A}/\hat{B}$ does correct the naïve likelihood so as to providing legitimate statistics. This is evident by seeing that the robust variance estimate, $V_{adj}(\hat{\theta})$ is close to the sample variance of $\hat{\theta},$ as n increases. Similarly, the empirical type I error probability of the robust LR test is close to the nominal level 0.05. The accuracy and approximation improves as the sample size increases as well. On the contrarily, the corresponding statistics derived from the naïve likelihood are way off their respective nominal levels.

5. Real Examples

In this section, we analyze real data sets to show the performance of our parametric robust approach. We use LR_{na} and LR_{adj} to, respectively, stand for the naïve and the robust LR test statistics for $H_0 : \theta = 0$.

Example 1. Moore and Tsiatis [8] analyzed a low-iron rat teratology data set. The purpose of study was to assess the effects of chemical agents or dietary regimens on fetal development in laboratory rats. The 48 female rats were dieted by iron-deficient food and classified into four groups according to iron supplement (GRP=1: untreated group, GRP=2: rat received injections on day 7 or day 10 only, GRP=3: rat received injections on days 0 and 7, GRP=4: rat injected with iron supplement weekly). Each litter of female rat was counted the total number of fetuses and number of dead fetuses. The hemoglobin levels (HB) of the mothers were also recoded.

Moore and Tsiatis [8] fitted several logistic regression models to assess the effects of HB and GRP separately. We, nonetheless, use our robust approach to study the joint effects of the two covariates. First, we code GRP simply to differentiate rats receiving injections $(1 : \text{GRP} = 2, 3, 4)$ or not $(0 : \text{GRP} = 1)$.

The probability of dead fetuses was estimated by the following logistic regression model $\text{logit}(\hat{\pi}) = 2.03 - 0.19\text{HB} - 2.65\text{GRP}.$ More statistics are reported in Table 2.

Example 2. This example concerned with the developmental effects resulting from exposure to hydroxyurea. The data set contained litters (clusters) of mice with varying sizes. Each female rat was randomly assigned to one of the treatment groups, including control, low dose (L), medium dose (M) and high dose (H). Three outcomes were recorded (the numbers of implantation sites (I), malformations (II) and dead/resorbed fetuses (III)). We only consider three dose groups (L, M and H) due to zero response in malformations under control group, like the way Chen et al. [2] did.

Chen et al. [2] proposed using the Dirichlet-trinomial distribution to incorporate the within-cluster association. Before proceeding, we first define two dummy indicators x_1 and x_2 for the treatment groups: $L(x_1 = 0, x_2 = 0)$, $M(x_1 = 1, x_2 = 0)$ and $H(x_1 = 0, x_2 = 1)$. Our method suggested that $\log(\hat{\pi}_{\text{II}}/\hat{\pi}_{\text{I}}) = -2.37 + 0.24x_1$

$+0.17x_2$ and $\log(\hat{\pi}_{III}/\hat{\pi}_I) = -2.25 + 1.49x_1 + 1.80x_2$. Naïve and robust statistics are also included in Table 2.

Example 3. This data set contains records of eye-testing results of 3,242 men and 7,477 women employees in Royal Ordnance factories in 1943-46. Both eyes of a person (cluster) were recorded according to four grades, 1~4. We used "grade 1" as the baseline category and found $\log(\hat{\pi}_{i,2}/\hat{\pi}_{i,1}) = 1.11 - 0.51x_i$, $\log(\hat{\pi}_{i,3}/\hat{\pi}_{i,1}) = 1.01 - 0.55x_i$ and $\log(\hat{\pi}_{i,4}/\hat{\pi}_{i,1}) = 0.87 - 0.12x_i$, where $x_i = 1$ for man and $x_i = 0$ for woman.

Obviously, results tabulated in Table 2 indicate that, contrary to the conclusion from the naïve LR test statistic, our adjusted LR test statistic found no significant difference between sex for eye grade 4 and grade 1. This is identical to the conclusion by Stuart [13] who used some rank statistic to compare gender difference.

Despite Stuart [13] and our adjusted LR test statistic lead to the same conclusion, our robust approach provides not only valid likelihood ratio test results, but also consistent point estimates and valid standard errors. Most importantly, our method supplies a legitimate likelihood function which makes available full likelihood inferences that are unobtainable from any nonparametric or semi-parametric means.

Table 2. Naïve and robust statistics for real data analysis

	$\hat{\theta}$	$\hat{V}_{na}(\hat{\theta})$	$\hat{V}_{adj}(\hat{\theta})$	LR_{na}	LR_{adj}
Teratology data	–2.65	0.2327	0.6086	35.15	13.44
Hydroxyurea data	1.80	0.0475	0.0615	96.31	74.35
Eyes data	–0.12	0.0024	0.0039	5.87	3.58

6. Conclusion

We have proposed a robust likelihood means for regression analysis of correlated nominal data. The validity of our robust likelihood requires no knowledge of the true underlying joint distributions of the cluster data. This accomplishment is not accessible by nonparametric or semi-parametric methods.

References

[1] A. Agresti, An Introduction to Categorical Data Analysis, 2nd ed., John Wiley, New York, 2010.

[2] J. J. Chen, R. L. Kodell, R. B. Howe and D. W. Gaylor, Analysis of trinomial responses from reproductive and developmental toxicity experiments, Biometrics 47 (1991), 1049-1058.

[3] D. R. Cox and D. V. Hinkley, Theoretical Statistics, Chapman & Hall, New York, 1974.

[4] D. Hedeker, A mixed-effects multinomial logistic regression model, Statistics in Medicine 22 (2003), 1433-1446.

[5] K. Lee and D. Mercante, Longitudinal nominal data analysis using marginalized models, Comput. Statist. Data Anal. 54 (2010), 208-218.

[6] K. Y. Liang and S. L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986), 13-22.

[7] P. McCullagh, Quasi-likelihood functions, Ann. Statist. 11 (1983), 59-67.

[8] D. F. Moore and A. Tsiatis, Robust estimation of the variance in moment methods for extra-binomial and extra-Poisson variation, Biometrics 47 (1991), 383-401.

[9] S. R. Paul, U. Balasooriya and T. Banerjee, Fisher information matrix of the Dirichlet-multinomial distribution, Biom. J. 47 (2005), 230-236.

[10] R. M. Royall, On the probability of observing misleading statistical evidence, J. Amer. Statist. Assoc. 95 (2000), 760-768.

[11] R. M. Royall and T. S. Tsou, Interpreting statistical evidence by using imperfect models: robust adjusted likelihood functions, J. Roy. Statist. Soc. Ser. B 65 (2003), 391-404.

[12] J. E. Stafford, A robust adjustment of the profile likelihood, Ann. Statist. 24 (1996), 336-352.

[13] A. Stuart, The estimation and comparison of strengths of association in contingency tables, Biometrika 40 (1953), 105-110.

[14] T. S. Tsou and C. W. Shen, Parametric robust inferences for correlated ordinal data, Statistics in Medicine 27 (2008), 3550-3562.

[15] J. R. Wilson and S. C. Chen, Dirichlet-multinomial model with varying response rates over time, J. Data Science 5 (2007), 413-423.

[16] X. Zhang, W. J. Boscardin and T. R. Belin, Bayesian analysis of multivariate nominal measures using multivariate multinomial probit models, Comput. Statist. Data Anal. 52 (2008), 3697-3708.