BIOMETRICS 52, 1223–1234
December 1996

Mantel–Haenszel-Type Inference for Cumulative Odds Ratios with a Stratified Ordinal Response

I-Ming Liu

Department of Statistics, National Chung Hsing University,
Taipei, Taiwan, R.O.C.

and

Alan Agresti

Department of Statistics, University of Florida,
Gainesville, Florida 32611, U.S.A.

SUMMARY

This article proposes a Mantel–Haenszel-type estimator of an assumed common cumulative odds ratio in a proportional odds model for an ordinal response with several $2\times c$ contingency tables. It is useful, for instance, for comparing two treatments on an ordinal response for data from several centers when the data are highly sparse. The estimator has behavior similar to the Mantel–Haenszel estimator of a common odds ratio for several 2×2 tables. It is consistent under the ordinary asymptotic framework in which the number of tables is fixed and, unlike the maximum likelihood (ML) estimator, also under sparse asymptotics in which the number of tables grows with the sample size. Simulations reveal a considerable difference between it and the ML estimator when each table has few observations. Efficiency comparisons suggest that little efficiency loss occurs compared to the ML estimator when the data are not sparse. Tests and estimators are presented for detecting and handling heterogeneity in the odds ratios, and generalizations are available for stratified $r\times c$ contingency tables.

1. Introduction

This article considers estimation of odds ratios for $2\times c\times K$ contingency tables in which the c columns are an ordinal response, the two rows are levels of an explanatory factor, and the K strata are levels of a control variable. Tables of this type often refer to a comparison of two treatments when data are collected from K centers of some type, such as medical clinics. Such data are often sparse. For instance, a study might use many clinics because of the time it takes each clinic to recruit many patients; the three-way table might then have many strata but few observations per stratum.

Table 1, analyzed by the first author during a summer internship at Merck Research Laboratories, is an example of this type. This table shows preliminary results from a double-blind, parallel-group clinical study conducted at a large number of centers. The purpose of the study was to compare an active drug with placebo in the treatment of patients suffering from asthma. Patients were randomly assigned to the treatments. At the end of the study, investigators described their perception of the patients' change in condition using the ordinal scale (better, unchanged, worse).

Let Y denote the response variable, with c ordered categories. Let X denote a binary explanatory variable, and let Z denote a control variable with K categories. Let π_{ijk} denote the probability that Y falls at level j, when X is at level i and Z is at level k. The jth cumulative probability is

Key words: Consistent estimators; Cumulative logit model; Maximum likelihood; Proportional odds; Sparse data.

$\pi^*_{ijk} = \pi_{i1k} + \ldots + \pi_{ijk}$, for $j = 1, \ldots, c$. We consider the cumulative logit model

$$\log\left(\frac{\pi^*_{ijk}}{1-\pi^*_{ijk}}\right) = \alpha_j + \gamma_k + \beta I_{\{i=1\}} \qquad i = 1, 2,\ j = 1, \ldots, c-1,\ k = 1, \ldots, K, \tag{1}$$

where $I_{\{i=1\}} = 1$ when $i = 1$ and $I_{\{i=1\}} = 0$ when $i = 2$. At each level of Z, the odds that Y falls below level j at level 1 of X are $\exp(\beta)$ times the odds at level 2 of X. This simple model having a common effect of X for all j has a *proportional odds* structure (McCullagh, 1980). We refer to $\theta = \exp(\beta)$ as the *cumulative odds ratio* for the X-Y conditional association.

When $c = 2$, the Mantel–Haenszel (MH) (Mantel and Haenszel, 1959) estimator of a common odds ratio is a popular way of summarizing association. It is used not only when a common odds ratio assumption seems plausible, but also as a summary measure when the association varies only mildly among the tables. An alternative estimator results from fitting a logit model with no treatment-by-center interaction. When that model holds but K is large and the data are sparse, the model-based maximum likelihood (ML) estimator tends to overestimate the true log odds ratio. In practice, this happens whenever the number of tables grows at the same rate as the sample size, so the number of parameters also grows (Neyman and Scott, 1948). When each stratum consists of a single matched pair, such as in case-control studies, the unconditional ML estimator of a common log odds ratio converges to double the true value (Andersen, 1980, p. 244).

We propose a simple extension of the Mantel–Haenszel estimator for the X-Y cumulative odds ratio in model (1), the goal being to improve on the ML estimator when the data are sparse. The proposed estimator is consistent for the two standard types of asymptotics—(a) when the sample size within each stratum increases and the number of strata is fixed, and (b) when the number of strata increases proportional to the overall sample size. We call limiting situation (a) the large-stra-

Table 1
Evaluations of patients suffering from asthma

		Response					Response		
Center	Drug	Better	Unchanged	Worse	Center	Drug	Better	Unchanged	Worse
1	Placebo	0	2	1	2	Placebo	0	1	0
	Active	1	1	0		Active	1	1	0
3	Placebo	1	1	0	4	Placebo	1	0	0
	Active	0	1	0		Active	1	1	0
5	Placebo	1	0	0	6	Placebo	1	0	0
	Active	1	0	0		Active	2	1	0
7	Placebo	0	1	0	8	Placebo	0	0	1
	Active	2	1	0		Active	0	1	0
9	Placebo	1	1	0	10	Placebo	0	2	0
	Active	1	1	0		Active	1	0	0
11	Placebo	2	0	0	12	Placebo	0	1	0
	Active	1	0	1		Active	1	0	0
13	Placebo	1	0	0	14	Placebo	0	1	0
	Active	1	0	0		Active	2	0	0
15	Placebo	1	0	0	16	Placebo	0	1	0
	Active	1	0	0		Active	1	0	0
17	Placebo	0	2	0	18	Placebo	0	1	0
	Active	1	1	0		Active	1	0	0
19	Placebo	1	0	0	20	Placebo	1	0	0
	Active	1	0	0		Active	1	0	0
21	Placebo	0	3	0	22	Placebo	0	2	0
	Active	0	1	0		Active	1	0	0
23	Placebo	1	0	0	24	Placebo	1	1	0
	Active	1	0	0		Active	1	0	0
25	Placebo	1	0	0	26	Placebo	0	1	1
	Active	1	0	0		Active	1	0	0
27	Placebo	0	1	0	28	Placebo	1	0	0
	Active	0	2	0		Active	1	1	0

ta case, and limiting situation (b) the sparse-strata case. We also provide an estimated standard error that is dually consistent. These results extend ones for the binary case by Gart (1962), Breslow (1981), and Robins, Breslow, and Greenland (1986).

Section 2 introduces the estimator. Section 3 summarizes a simulation study showing that the ML estimator tends to overestimate the odds ratio when the data are sparse. Section 4 discusses large-sample efficiency relative to the ML estimator for the large-strata case. Section 5 presents tests of the common odds ratio assumption, and Section 6 extends the odds ratio estimator to $r\times c\times K$ tables. The article's main emphasis is on $r = 2$ because the sparse-data bias of ML estimators diminishes as r increases. The consistency arguments are technical, and an appendix outlines the proofs.

2. An Ordinal Odds Ratio Estimator

We assume that each $2\times c$ table is formed by two independent multinomial samples $(X_{1jk}, X_{2jk}, j = 1,\dots,c)$, with sample sizes denoted by n_{1k} and n_{2k}. Let $N_k = n_{1k} + n_{2k}$, $k = 1,\dots,K$, and denote cumulative counts by $X^*_{ijk} = X_{i1k} + \cdots + X_{ijk}$. Let $R_{jk} = X^*_{1jk}(n_{2k} - X^*_{2jk})/N_k$, $S_{jk} = (n_{1k} - X^*_{1jk})X^*_{2jk}/N_k$, and $N = \Sigma_k N_k$. The large-strata case refers to all $N_k \to \infty$ with K fixed; the sparse-strata case refers to $K \to \infty$ as $N \to \infty$ with $\{N_k, k = 1,\dots,K\}$ bounded.

For model (1), the same odds ratio occurs for all collapsings of the response into the binary outcome $(\le j, > j)$, $j = 1,\dots,c-1$. Suppose we naively treat the 2×2 tables for the $c-1$ collapsings of each stratum as independent. Then the Mantel–Haenszel estimator of the common odds ratio in the resulting $(c-1)K$ separate 2×2 tables equals

$$\hat\theta = \frac{\displaystyle\sum_{k=1}^{K}\sum_{j=1}^{c-1} X^*_{1jk}(n_{2k} - X^*_{2jk})/N_k}{\displaystyle\sum_{k=1}^{K}\sum_{j=1}^{c-1}(n_{1k} - X^*_{1jk})X^*_{2jk}/N_k} = \frac{\displaystyle\sum_{k=1}^{K}\sum_{j=1}^{c-1} R_{jk}}{\displaystyle\sum_{k=1}^{K}\sum_{j=1}^{c-1} S_{jk}}. \tag{2}$$

The appendix shows that $\hat\theta$ is dually consistent for θ. This estimator also applies to a cumulative odds ratio for a more general model that replaces $\alpha_j + \gamma_k$ by α_{jk}.

The estimator $\hat\theta$ has the form $\Sigma_k \Sigma_j S_{jk}\hat\theta^*_{jk}/\Sigma_k \Sigma_j S_{jk}$, where $\hat\theta^*_{jk} = [X^*_{1jk}(n_{2k} - X^*_{2jk})]/[X^*_{2jk}(n_{1k} - X^*_{1jk})]$. Thus, when the common cumulative odds ratio assumption does not hold, $\hat\theta$ estimates a weighted average of stratum-specific cumulative odds ratios. When the heterogeneity is not severe and the directions of the odds ratio are the same, this is still a useful summary of the conditional association.

Though we motivated $\hat\theta$ by treating $(c-1)K$ separate 2×2 tables as independent, a standard error based on this would be inappropriate. Generalizing the method given by Robins et al. (1986) provides a relevant variance estimator,

$$\widehat{\mathrm{var}}[\log(\hat\theta)] = \frac{\displaystyle\sum_{k=1}^{K}\hat\xi_k(\hat\theta)}{\hat\theta^2(\displaystyle\sum_{k=1}^{K}\sum_{j=1}^{c-1} S_{jk})^2}, \tag{3}$$

where $\hat\xi_k(\theta) = \Sigma_{j=1}^{c-1}\hat\phi_{jjk}(\theta) + 2\Sigma_{j<j'}^{c-1}\hat\phi_{jj'k}(\theta)$, with

$$\hat\phi_{jsk}(\theta) = \frac{n_{1k}n_{2k}}{N_k^2}\left\{\frac{\theta(n_{1k} - X^*_{1sk})X^*_{2jk}}{n_{1k}}\left[1 + (\theta-1)\frac{X^*_{2sk}}{n_{2k}}\right] + \frac{X^*_{1jk}(n_{2k} - X^*_{2sk})}{n_{2k}}\left[\theta - (\theta-1)\,\frac{X^*_{1sk}}{n_{1k}}\right]\right\}, \qquad j \le s = 1,\dots,c-1.$$

The appendix provides consistency arguments for this estimator for the two limiting cases. Liu (1995) showed that the estimator takes the same value if one reverses, for each of the K tables, the order of the rows. Unfortunately, it is not invariant to reversing the order of the response categories. Following Robins et al. (1986), one can use the average of the estimators for the two orderings. In all cases we have studied, only trivial differences exist between the two estimates.

For Table 1, $\log(\hat{\theta}) = -1.153$ has a standard error estimate of .571. For each center, the estimated odds that the evaluation for the active drug falls below any fixed level are $\exp(1.153) = 3.17$ times the estimated odds for placebo. The ordinary unconditional ML estimator for model (1) equals $\hat{\beta} = -1.703$, with estimated standard error equal to .598.

When $c = 2$, $\hat{\theta}$ is the ordinary Mantel–Haenszel estimator. We refer to $\hat{\theta}$ and $\hat{\beta}$ with $c > 2$ as MH-type estimators. Mantel and Haenszel (1959) showed that a value of 0 for their statistic for testing conditional independence is equivalent to $\hat{\theta} = 1.0$. Mantel (1963) generalized the test statistic to a single-degree-of-freedom statistic for $r \times c \times K$ tables with ordered rows and column scores. When $r = 2$ and $c > 2$, it is straightforward to show that $\hat{\theta} = 1$ is equivalent to a value of 0 for Mantel's generalized ordinal statistic when one uses equally-spaced scores for the columns. Mantel's statistic is a natural one for testing the treatment effect. For Table 1, Mantel's statistic equals 4.84, with a P-value of .028. The value $\hat{\theta} = 3.17$ is a supplementary measure that describes departure from the null hypothesis.

When each row of the $2 \times c \times K$ table has only one observation, such as with matched pairs, the MH-type estimator (2) and the standard error estimator simplify dramatically. One can express the estimators directly in terms of counts in a $c \times c$ table that expresses the joint ratings of the two observations from each partial table. Let x_{ij} be the sample count for the cell in the ith row and the jth column of such a table; this is the number of matched pairs for which the first subject in the pair made response i and the second made response j. The estimate of $\log(\theta)$ and its variance estimate simplify to

$$\log(\hat{\theta}) = \log\left\{\left[\sum_{i<j}(j-i)x_{ij}\right] \Big/ \left[\sum_{i>j}(i-j)x_{ij}\right]\right\},$$

$$\widehat{\mathrm{var}}[\log(\hat{\theta})] = \sum_{i<j}(j-i)^2 x_{ij} \Big/ \left[\sum_{i<j}(j-i)x_{ij}\right]^2 + \sum_{i>j}(i-j)^2 x_{ij} \Big/ \left[\sum_{i>j}(i-j)x_{ij}\right]^2.$$

Agresti and Lang (1993) proposed this estimator for a proportional odds model for matched pairs.

When the true odds ratios are heterogeneous within or between strata, $\hat{\theta}$ often still provides a useful descriptive summary, but the variance estimator (3) may no longer be valid. Under the large-strata situation, the asymptotic variance of $\log(\hat{\theta})$ follows from results in Lemma 2 in Appendix A.2 and is given there.

Clayton (1974) provided a more complex estimator of $\log(\theta)$ that is a weighted average of estimators based on the separate collapsings for each partial table. The weights were chosen to minimize the variance of the estimator when $\theta = 1$ under the large-strata case. Clayton also proposed related estimators of Mantel–Haenszel form. It is unclear how to derive sparse-data standard errors for his estimators, as the estimated weights themselves are highly unstable in that case. In addition, the choice of weights is normally not critically important to the efficiency for these types of estimators (McCullagh and Nelder, 1989, p. 274).

3. Simulation Study

To investigate whether the ML estimator deteriorates relative to the MH-type estimator as sparseness increases, we conducted a simulation study with scenarios ranging from ones for which large-strata asymptotics should work well to ones for which sparse-strata asymptotics seem more appropriate. The results reported in Tables 2–4, for $K = 10$, are typical of those in Liu (1995) for a variety of cases that there is not space to report. These tables assume that model (1) holds with equal probabilities for the c-category response in the second row for each $2 \times c$ table, and with $\gamma_k = 0$ for all k. For each case we generated 10,000 $2 \times c \times K$ sample tables based on independent multinomial distributions.

For the simulated tables for each case, Table 2 summarizes the sample mean and the mean squared error for MH and ML estimators of β in model (1). Most of the standard errors for the sample means are less than .01. In comparing the estimators, a * symbol in this table indicates that that estimator is significantly better at the .05 level. The MH estimator tends to behave considerably better than the ML estimator for the sparse-strata cases ($n_{ik} \leq 3$). The ML estimate averages nearly double the true parameter when $n_{1k} = n_{2k} = 1$. The two estimators behave similarly for large-strata cases.

Table 3 shows the proportion of times that the ML estimator was closer than the MH estimator to β. The largest possible standard error of these estimates is $(.5 \times .5/10{,}000)^{1/2} = .005$. The MH

Table 2
The sample mean and the mean squared error (in parentheses) of ML and Mantel–Haenszel-(MH) type estimators of β, based on 10,000 *simulations with* 10 *strata*

			$c = 3$		$c = 5$		$c = 7$	
			$\theta = 2$	$\theta = 4$	$\theta = 2$	$\theta = 4$	$\theta = 2$	$\theta = 4$
Estimator	(n_{1k}, n_{2k})	β:	0.693	1.386	0.693	1.386	0.693	1.386
ML	20, 20		0.712	1.424	0.710	1.419	0.710	1.418
			(0.037)	(0.044)	(0.034)	(0.039)	(0.034)	(0.038)
	2, 3		0.867	1.740	0.837	1.681	0.828	1.667
			(0.512)	(0.724)	(0.434)	(0.583)	(0.418)	(0.547)
	1, 1		1.241	2.310	1.254	2.428	1.234	2.400
			(2.965)	(3.295)	(2.969)	(4.041)	(2.769)	(3.672)
MH	20, 20		0.699*	1.398*	0.699*	1.396*	0.699*	1.387*
			(0.036*)	(0.042*)	(0.033*)	(0.038*)	(0.032*)	(0.037*)
	2, 3		0.740*	1.482*	0.731*	1.470*	0.724*	1.470*
			(0.358*)	(0.486*)	(0.318*)	(0.432*)	(0.310*)	(0.409*)
	1, 1		0.666*	1.200*	0.742*	1.372*	0.759*	1.432*
			(0.741*)	(0.632*)	(0.838*)	(0.744*)	(0.870*)	(0.824*)

* The estimate is significantly better at the .05 level, based on the precision of simulation.

Table 3
Proportion of times that the ML estimate is closer than the Mantel–Haenszel-type estimate to β, based on 10,000 *simulations with* 10 *strata*

		$c = 3$			$c = 5$			$c = 7$		
(n_{1k}, n_{2k})	θ:	1	2	4	1	2	4	1	2	4
20, 20		.400	.462	.493	.417	.465	.499	.422	.472	.505
2, 3		.223	.340	.414	.287	.358	.434	.298	.367	.424
1, 1		.019	.199	.309	.113	.203	.290	.141	.220	.285

Table 4
Sample proportion estimates of the probability that the Wald statistic exceeds the $100(1 - \alpha)$ *percentage point of the chi-squared distribution, based on* 10,000 *simulations with* 10 *strata*

			$c = 3$		$c = 5$		$c = 7$	
Estimator	(n_{1k}, n_{2k})	α:	0.05	0.10	0.05	0.10	0.05	0.10
ML	20, 20		0.055	0.104	0.056	0.109	0.056	0.107
	2, 3		0.080	0.146	0.087	0.155	0.090	0.154
	1, 1		0.142	0.235	0.187	0.274	0.201	0.290
MH	20, 20		0.052	0.099	0.052	0.104	0.052	0.102
	2, 3		0.048	0.099	0.052	0.103	0.053	0.101
	1, 1		0.017	0.073	0.060	0.115	0.068	0.122

estimator becomes more preferable as the data become sparser and, to a lesser extent, as the association weakens.

Finally, we analyzed the asymptotic chi-squared approximations in the upper tail for Wald tests about β based on these estimators. Table 4 reports sample proportion estimates of the null probability that the Wald statistic exceeds the $100(1 - \alpha)$ percentage point of the chi-squared distribution with d.f. = 1. The largest possible standard error of these sample proportions is .005. For sparse-strata cases, the ML statistics are highly liberal. The MH statistics behave quite well.

4. Efficiency Comparisons

Though the MH-type estimator $\hat{\theta}$ behaves better than the ML estimator for sparse data, the ML estimator is asymptotically efficient for the large-strata case. The naive independence assumption used in motivating $\hat{\theta}$ is unrealistic, and the estimator could suffer some efficiency loss when the data are not sparse. We take some comfort from work by Liang and Zeger (1986) showing that such naive estimators for repeated measurement data can perform surprisingly well. For the large-strata case we studied the Pitman asymptotic relative efficiency (ARE) of the MH compared to the ML estimator, which is based on comparing their asymptotic variances. The ARE values are the same for the odds ratio scale and its log.

Table 5 illustrates Pitman ARE results for the null case ($\theta = 1$) when the response has three categories. The ARE does not vary much for different sample size structures. Though the ML estimator and its standard error do not have closed form for proportional odds models, an asymptotically equivalent test statistic for testing $\theta = 1$ is based on the efficient score test for model (1). When there is a single stratum, this test statistic is equivalent to the Wilcoxon test using midranks (McCullagh, 1980), and for K strata it corresponds to a stratified version of that test. Using this connection, one can obtain a closed form for the ARE in the null case. For the single-stratum case Liu (1995) used this to show that the proposed estimator has full efficiency when the probabilities alternate, such as (2/9, 5/9, 2/9). For the binary response case there is no efficiency loss (Breslow, 1981) under conditional independence.

Table 6 shows some ARE values for some cases with θ differing from 1 for a single stratum. Efficiencies for multi-strata cases were on the order of the average of the separate efficiencies for the single-stratum efficiencies of the component partial tables. Tables 5 and 6 and others in Liu (1995) for a variety of other cases suggest that the proposed estimator does not suffer a severe loss of efficiency, for large strata, compared to the ML estimator. The poorest ARE values that we observed, on the order of .90, tended to occur when the sample size is relatively small in a row for which the probabilities vary greatly, such as illustrated by the case having ARE = .909 in Table 6.

Of course, for large strata there is no need to use the MH-type estimator because the ML estimator behaves adequately. The main purpose of the MH-type estimator is to handle sparse-strata cases, for which ML may perform poorly.

5. Checks for Homogeneity of Cumulative Odds Ratios

The proposed estimator refers to a model having the proportional odds assumption that the odds ratios are constant across the strata and for different possible collapsings within each stratum. Ordinary Pearson or likelihood-ratio goodness-of-fit tests of model (1) provides large-strata tests of homogeneity of cumulative odds ratios within and across strata. For the sparse-strata case we now provide a test of homogeneity within strata, assuming a common odds ratio across the strata.

For $2 \times c \times K$ contingency tables, consider the collapsed response ($\leq j$, $> j$) for $j \in \{1, \ldots, c-1\}$. Suppose that this odds ratio is the same for each stratum, with common value denoted by θ_j. We test H$_0$: $\theta_1 = \theta_2 = \cdots = \theta_{c-1}$ against H$_a$: at least one of $\{\theta_j\}$ differs from the others. For the jth

Table 5

The Pitman ARE for the proposed estimator compared to the ML estimator, for K strata and three response categories when $\theta = 1$

K	n_{ik}	$\pi_{111} : \pi_{121} : \pi_{131}$	$\pi_{112} : \pi_{122} : \pi_{132}$	$\pi_{113} : \pi_{123} : \pi_{133}$	ARE(MH, ML)
1	I	1:1:1	N.A.	N.A.	1.000
	I	2:9:2	N.A.	N.A.	1.000
	I	1:1:2	N.A.	N.A.	0.990
	I	2:4:5	N.A.	N.A.	0.986
	I	1:1:10	N.A.	N.A.	0.934
	III	1:1:1	2:9:2	N.A.	0.986
2	II	1:1:1	2:8:2	N.A.	0.988
	III	1:1:1	2:8:2	N.A.	0.988
	II	1:1:1	1:1:10	N.A.	0.977
	III	1:1:1	1:1:10	N.A.	0.962
3	II	1:1:1	2:9:2	1:1:1	0.990
	IV	1:1:1	2:9:2	1:1:1	0.992
	II	1:1:1	2:9:2	2:9:2	0.986
	IV	1:1:1	2:9:2	2:9:2	0.986

Note: In the n_{ik} column we use I–IV to represent the different sample sizes structures, where I: n_{ik} is arbitrary; II: n_{ik} is equal for $i \in \{1,2\}$, $k \in \{1, \ldots, K\}$; III: $n_{11} : n_{21} : n_{12} : n_{22}$ = 1:2:3:4; and IV: $n_{11} : n_{21} : n_{12} : n_{22} : n_{13} : n_{23}$ = 1:2:1:1:1:4. The N.A. entry indicates 'not applicable.'

Table 6

The Pitman ARE for the proposed estimator compared to the ML estimator, for one stratum and three response categories when θ differs from one

θ	n_{ik}	$\pi_{111} : \pi_{121} : \pi_{131}$	$\pi_{211} : \pi_{221} : \pi_{231}$	ARE(MH, ML)
2	I	1:1:1	2:3:5	0.995
	II	1:1:1	2:3:5	0.999
	III	1:1:1	2:3:5	0.972
	I	2:1:1	5:4:6	0.998
	II	2:1:1	5:4:6	0.983
	III	2:1:1	5:4:6	0.997
	I	10:1:1	65:12:14	0.944
	II	10:1:1	65:12:14	0.929
	III	10:1:1	65:12:14	0.963
4	I	1:1:1	1:2:6	0.976
	II	1:1:1	1:2:6	0.999
	III	1:1:1	1:2:6	0.909
	I	2:1:1	7:8:20	0.999
	II	2:1:1	7:8:20	0.978
	III	2:1:1	7:8:20	0.960
	I	10:1:1	75:24:36	0.955
	II	10:1:1	75:24:36	0.921
	III	10:1:1	75:24:36	0.986

Note: In the n_{ik} column we use I, II, and III to represent the different sample sizes structures, where I: $n_{11} : n_{21}$ = 1:1; II: $n_{11} : n_{21}$ = 1:10; and III: $n_{11} : n_{21}$ = 10:1.

collapsed table, denote the Mantel–Haenszel estimator by $\hat{\theta}_j$. Under the null hypothesis, $\{\log(\hat{\theta}_j), j = 1, \ldots, c-1\}$ have an asymptotic multivariate normal distribution with a common mean. Thus, one can construct a Wald test statistic,

$$\hat{\mathbf{\Delta}}'(\hat{\mathbf{\Sigma}})^{-1}\hat{\mathbf{\Delta}},$$

where $\hat{\mathbf{\Delta}}' = [\log(\hat{\theta}_2) - \log(\hat{\theta}_1), \ldots, \log(\hat{\theta}_{c-1}) - \log(\hat{\theta}_1)]$ and $\hat{\mathbf{\Sigma}}$ is the estimated covariance matrix for $\{\log(\hat{\theta}_j) - \log(\hat{\theta}_1), j = 2, \ldots, c-1\}$. This quadratic form has asymptotically a chi-squared null distribution with $c-2$ degrees of freedom. Its value is invariant to the choice of the baseline category for forming $\hat{\mathbf{\Delta}}$. Under the null hypothesis, dually consistent estimators for the variance and covariance of $\{\log(\hat{\theta}_j)\}$ are

$$\widehat{\text{cov}}[\log(\hat{\theta}_j), \log(\hat{\theta}_s)] = C_{js} = \frac{\sum_{k=1}^{K} \hat{\phi}_{jsk}(\hat{\theta})}{\hat{\theta}^2 \left(\sum_{k=1}^{K} S_{jk}\right)\left(\sum_{k=1}^{K} S_{sk}\right)}, \qquad j \leq s = 1, \ldots, c-1,$$

where $\hat{\phi}_{jsk}(\theta)$ is given in (3). Thus, $\widehat{\text{cov}}[\log(\hat{\theta}_j) - \log(\hat{\theta}_1), \log(\hat{\theta}_s) - \log(\hat{\theta}_1)] = C_{js} - C_{1j} - C_{1s} + C_{11}$ for all $j \leq s \in \{2, \ldots, c-1\}$. As mentioned before, if such a test provides evidence of heterogeneity but the degree of heterogeneity is mild, one might still use $\hat{\theta}$ to summarize the association, using the standard error mentioned in Section 2 that permits heterogeneity.

For Table 1, $\log(\hat{\theta}_1) = -1.206$ and $\log(\hat{\theta}_2) = -.903$. The test of homogeneous odds ratios for the two collapsings has test-statistic value .06, based on d.f. = 1. The proportional odds assumption seems plausible. The small test-statistic value results from a large estimated standard error for $\log(\hat{\theta}_2)$, since only four strata contribute to its calculation.

For sparse strata with the odds ratios treated as fixed effects, it does not seem realistic to expect to construct a powerful test of homogeneous odds ratios across the strata. However, a referee has noted that one could use a random effects approach by assuming that the log cumulative odds ratios follow some distribution; the test of homogeneity is then a test that the variance of that distribution equals zero. More generally, an alternative way to permit heterogeneity of odds ratios is based on a mixed model that has components allowing for heterogeneity within and across strata. This model would generalize a binary response model given by Liu and Pierce (1993) that permitted between-strata heterogeneity of odds ratios.

6. Extensions to Several Rows

Generalizations of the MH cumulative odds ratio estimator apply to tables having more than two rows. Suppose, for instance, that the explanatory variable has r nominal categories, and consider the model

$$\log\left(\frac{\pi^*_{ijk}}{1-\pi^*_{ijk}}\right) = \alpha_j + \gamma_k + \beta_i,$$

for $i \in \{1,\ldots,r\}$, $j \in \{1,\ldots,c-1\}$, and $k \in \{1,\ldots,K\}$. For this model the cumulative odds ratio for the ith and the hth rows equals $\theta_{ih} = \exp(\beta_i - \beta_h)$ in each stratum.

Denote the MH estimator for $\log(\theta_{ih})$ based on rows i and h alone by L_{ih}. An alternative estimator uses the full $r\times c\times K$ contingency table. It is

$$\bar{L}_{ih} = (L_{i+} - L_{h+})/r,$$

where $L_{i+} = \Sigma^r_{i'=1} L_{ii'}$. Mickey and Elashoff (1985) defined analogous estimators for $2\times c\times K$ tables with c nominal response categories. Since L_{ih} is dually consistent for $\log(\theta_{ih}) = \beta_i - \beta_h$, $\bar{L}_{ih} = \Sigma_{i'}(L_{ii'} - L_{hi'})/r$ is dually consistent as well for $\Sigma_{i'}[(\beta_i - \beta_{i'}) - (\beta_h - \beta_{i'})]/r = \beta_i - \beta_h$. For the large-strata case, Liu (1995) showed that the Pitman ARE for L_{ih} compared to $\bar{L}_{ih}$ is 1 when $K = 1$ and independence applies, but otherwise $\bar{L}_{ih}$ is more efficient than L_{ih}.

Liu (1995) also provided an estimate when the explanatory variable has r ordinal categories. The asymptotic variances for these estimators are complex. [See Liu (1995) or a technical report available from the authors for details on this, as well as consistency arguments.] For sparse-data examples with $r > 2$, the ML estimates do not tend to be as dramatically larger than the MH-type estimates as when $r = 2$. For instance, for $r \times 2 \times K$ tables with one observation per row in each table, the bias of ML estimators has order $r/(r-1)$ as K increases. This follows from standard results for the Rasch model, which model (1) simplifies to in this binary-response case (Anderson, 1980, p. 244).

7. Applications and Extensions

We have seen that the ML estimator of a cumulative odds ratio can be unreliable when the data are highly sparse. We recommend using the MH estimator instead of the ML estimator when the sample size for most strata is on the order of 5 or less.

The variance formulas for the estimator both in the homogeneous and heterogeneous strata cases are rather complex. An alternative standard error may be obtained using a parametric bootstrap. In the homogeneous case, for instance, one can base this estimate on the variability in the $\log(\hat{\theta})$ values obtained by generating independent multinomial samples of the given sizes from the rows of the partial tables with the structure of model (1), using the ML estimates of $\{\alpha_j\}$ and $\{\gamma_k\}$ and the MH-type estimate of β. In our experience, the ML, MH, and bootstrap standard errors are similar unless the data are sparse.

For very sparse data, the probability of obtaining an infinite MH or ML estimate of $\log(\theta)$ is nonnegligible. In such cases the parametric bootstrap estimate is not well defined, and in fact the standard error for a limiting distribution itself has limited meaning. We constructed a bootstrap standard error by conditioning on those cases in the bootstrap simulations that have a finite MH estimate for the log odds ratio, since the usual asymptotic standard error also ignores the fact that the exact distribution of the estimator has finite mass on an infinite value. For very sparse data, the bootstrap standard error is then usually anywhere from 5% to 30% larger than the MH estimate. This serves as a warning that the asymptotic variance estimate (3) for the MH estimate may be overly optimistic. In fact, some limited simulations suggest that in such cases the bootstrap tends to overestimate and the MH approach tends to underestimate the appropriate ASE value.

In practice it is important to investigate and describe the heterogeneity across strata of cumulative odds ratios. We are developing an analysis using a generalized linear mixed model, as Liu and Pierce (1993) did in the binary case. Another possible extension is to allow for additional risk factors in the model and for dependencies in the data, as discussed by Liang (1987) for the binary case with an estimating equations approach.

For binary responses ($c = 2$), an alternative dually consistent estimator results from fitting the logit model using conditional maximum likelihood. For $c > 2$, however, the cumulative logit model has no reduced sufficient statistics for the parameters, and conditional ML estimation does not apply. For $K = 1$, McCullagh and Nelder (1989, p. 276) proposed a 'pseudo conditional likelihood' estimator using the conditional noncentral hypergeometric distributions for $\{X^*_{1jk}\}$. Computations in their method involve inverses of conditional cell means. For many data sets, including Table 1 and matched-pairs data having one observation in each row, some conditional cell means equal zero. Their measure is most appropriate for nonsparse-data cases. Though one cannot use conditional ML to eliminate the strata parameters $\{\gamma_k\}$, one could do this by treating them as a random effect. One would use numerical integration to integrate out these terms with respect to their distribution, thus yielding a likelihood for a marginal model (see Hedeker and Gibbons, 1995).

One could also consider a different model structure for the ordinal response, such as an adjacent-categories logit model. Liu (1995) proposed Mantel–Haenszel-type estimators for the odds ratio for this case and developed their properties. Another possibility is continuation-ratio logits (Thompson, 1977). We have not discussed these two model types in this article partly because they have an optimal alternative—the conditional maximum likelihood estimator. See McCullagh (1980) and Greenland (1994) for discussions regarding issues in choosing among ordinal models.

A FORTRAN program is available for calculating the MH-type estimators and their standard errors. Readers can request it from the first author through e-mail (ILIU@STAT2.NCHULC.EDU.TW).

Acknowledgements

The authors thank two referees and an associate editor for helpful comments that improved this paper. The authors also thank Dr Sudeep Kundu for making the data available to us. Prof Agresti's work was partially supported by a grant from the National Institutes of Health.

Résumé

Cet article propose un estimateur de type Mantel-Haenszel d'un odds ratio cumulatif supposé commun dans un modèle à chances proportionnelles pour une réponse ordinale avec plusieurs $2 \times c$ tableaux de contingence. Il est utile, par exemple, pour comparer deux traitements sur une réponse ordinale pour des données provenant de plusieurs centres quand les données sont fortement disséminées. L'estimateur a un comportement similaire à celui de Mantel-Haenszel d'un odds ratio commun pour plusieurs tableaux 2×2. Il est consistant dans le cadre asymptotique ordinaire dans lequel le nombre de tableaux est fixé et également, à la différence de l'estimateur du maximum de vraisemblance (ML), sous des conditions asymptotiques "disséminées" pour lesquelles le nombre de tableaux augmente avec la taille de l'échantillon. Les simulations révèlent une différence considérable entre cet estimateur et l'estimateur ML quand chaque tableau a peu d'observations. Les comparaisons d'efficacité suggèrent une petite perte d'efficacité par rapport à l'estimateur ML quand les données ne sont pas disséminées. Des tests et des estimateurs sont présentés pour détecter et traiter l'hétérogénéïté des odds ratios, et des généralisations sont déduites pour des tableaux de contingence $r \times c$ stratifiés.

References

Agresti, A. and Lang, J. B. (1993). A proportional odds model with subject-specific effects for repeated ordered categorical responses. *Biometrika* **80,** 527–34.

Andersen, E. B. (1980). *Discrete Statistical Methods with Social Science Applications.* New York: North–Holland.

Breslow, N. (1981). Odds ratio estimators when the data are sparse. *Biometrika* **68,** 73–84.

Clayton, D. G. (1974). Some odds ratio statistics for the analysis of ordered categorical data. *Biometrika* **61,** 525–531.

Gart, J. J. (1962). On the combination of the relative risks. *Biometrics* **18,** 594–600.

Greenland, S. (1994). Alternative models for ordinal logistic regression. *Statistics in Medicine* **13,** 1665–1677.

Guilbaud, O. (1983). On the large-sample distribution of the Mantel–Haenszel odds-ratio estimator. *Biometrics* **39,** 523–525.

Hedeker, D. R. and Gibbons, R. D. (1995). A random-effects ordinal regression model for multilevel data. *Biometrics,* in press.

Liang, K. Y. (1987). Extended Mantel–Haenszel estimating procedure for multivariate logistic regression models. *Biometrics* **43,** 289–299.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73,** 13–22.

Liu, I. M. (1995). Mantel–Haenszel-type inference for odds ratios with ordinal responses. Unpublished Ph.D. dissertation, University of Florida, Gainesville.

Liu, Q. and Pierce, D. A. (1993). Heterogeneity in Mantel–Haenszel-type models. *Biometrika* **80,** 543–556.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel–Haenszel procedure. *Journal of the American Statistical Association* **58,** 690–700.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22,** 719–748.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society, Series B* **42,** 109–142.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models.* London: Chapman and Hall.

Mickey, R. M. and Elashoff, R. M. (1985). A generalization of the Mantel–Haenszel estimator of partial association for $2 \times J \times K$ tables. *Biometrics* **41,** 623–635.

Neyman, J. and Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrika* **16,** 1–22.

Robins, J., Breslow, N., and Greenland, S. (1986). Estimators of the Mantel–Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics* **42,** 311–323.

Thompson, W. A. (1977). On the treatment of grouped observations in life studies. *Biometrics* **33,** 463–470.

Received March 1995; *revised December* 1995; *accepted March* 1996.

Appendix

In certain places in this appendix, the calculations are long and tedious and we refer to the dissertation by Liu (1995) for the complete details. A technical report is also available that provides further details.

A.1 Proof of Sparse-Strata Consistency

We assume that θ is positive and finite, and express $\hat{\theta}-\theta = \Sigma_k \Sigma_j (R_{jk} - \theta S_{jk})/\Sigma_k \Sigma_j S_{jk}$. Because for the sparse-strata case $\{R_{jk}\}$ and $\{S_{jk}\}$ are bounded random variables and $\mathrm{E}(R_{jk} - \theta S_{jk}) = 0$, the Chebyshev weak law of large numbers implies that $\hat{\theta}$ is consistent for θ. Let var^a, E^a, cov^a represent asymptotic variances, asymptotic expectations, and asymptotic covariances. From the Central Limit Theorem and Slutsky's Theorem, $\hat{\theta} - \theta$ is asymptotically normally distributed with

$$\lim_{K\to\infty} K\,\mathrm{var}^a(\hat{\theta}) = \frac{\lim_{K\to\infty} \sum_{k=1}^{K} \mathrm{var}\left(\sum_{j=1}^{c-1} (R_{jk} - \theta S_{jk})\right) /K}{\left[\lim_{K\to\infty} \sum_{k=1}^{K} \mathrm{E}\left(\sum_{j=1}^{c-1} S_{jk}\right)/K\right]^2}$$

$$= \frac{\lim_{K\to\infty} \sum_{k=1}^{K} \left[\sum_{j} \mathrm{var}(R_{jk} - \theta S_{jk}) + \sum_{j\neq j'} \mathrm{cov}(R_{jk} - \theta S_{jk},\, R_{j'k} - \theta S_{j'k})\right] /K}{\left[\lim_{K\to\infty} \sum_{k=1}^{K} \sum_{j=1}^{c-1} \mathrm{E}(S_{jk})/K\right]^2}. \quad (4)$$

Consider $K\,\widehat{\mathrm{var}}(\hat{\theta}) = [\Sigma_{k=1}^{K} \hat{\xi}_k(\hat{\theta})/K]/[\Sigma_{k=1}^{K} \Sigma_{j=1}^{c-1} S_{jk}/K]^2$. Lemma 1 shows that $\hat{\xi}_k(\theta)$ $(= \Sigma_{j=1}^{c-1} \hat{\phi}_{jjk}(\theta) + 2\,\Sigma_{j<j'}^{c-1} \hat{\phi}_{jj'k}(\theta))$ is an unbiased estimator of $\Sigma_{j=1}^{c-1} \mathrm{var}(R_{jk} - \theta S_{jk}) + \Sigma_{j<j'=1}^{c-1} [2 \times \mathrm{cov}(R_{jk} - \theta S_{jk},\, R_{j'k} - \theta S_{j'k})]$. Thus, $\Sigma_k \hat{\xi}_k(\theta)/K$ is consistent for estimating the numerator of (4) by the Chebyshev weak law of large numbers for summing nonidentically distributed, bounded, and unbiased random variables. Also, since $\hat{\theta} \xrightarrow{p} \theta$, $\Sigma_k \hat{\xi}_k(\hat{\theta})/K$ is also consistent. Furthermore, since $\{S_{jk}\}$ are bounded random variables, we can use the same arguments to conclude that $\Sigma_k \Sigma_j S_{jk}/K \xrightarrow{p} \lim_{K\to\infty} \Sigma_k \Sigma_j \mathrm{E}(S_{jk})/K$. Thus, the denominator of $K\,\widehat{\mathrm{var}}(\hat{\theta})$ is a consistent estimator of the denominator of (4). Therefore, $K\,\widehat{\mathrm{var}}(\hat{\theta})$ is consistent for (4).

Lemma 1. $\mathrm{E}(\hat{\phi}_{jsk}(\theta)) = \mathrm{cov}(R_{jk} - \theta S_{jk},\, R_{sk} - \theta S_{sk})$ *for* $j \le s = 1, \ldots, c-1$.

Proof. After collapsing the response into the three outcomes ($\leq j$, $> j$ and $\leq s$, $> s$), the variance and covariance can be calculated from a $2\times3\times K$ table. Let $\underline{X}^*_{ijk} = X_{i(j+1)k} + \cdots + X_{ick}$ and $\underline{\pi}^*_{ijk} = \pi_{i(j+1)k} + \cdots + \pi_{ick}$. Define

$$Z_{jsk} = \left(\frac{1}{N_k^2}\right)(\theta-1)^2 n_{1k}n_{2k}\pi^*_{1jk}\underline{\pi}^*_{1sk}\pi^*_{2jk}\underline{\pi}^*_{2sk}$$

and

$$D_{jsk} = \frac{n_{1k}n_{2k}}{N_k^2}\{n_{2k}[1+(\theta-1)(1-\underline{\pi}^*_{2sk})]\theta\underline{\pi}^*_{1sk}\pi^*_{2jk} + n_{1k}[\theta-(\theta-1)(1-\underline{\pi}^*_{1sk})]\pi^*_{1jk}\underline{\pi}^*_{2sk}\}.$$

Liu (1995) showed that $\mathrm{cov}(R_{jk}-\theta S_{jk},\, R_{sk}-\theta S_{sk}) = Z_{jsk}+D_{jsk}$ and $\mathrm{E}(\hat{\phi}_{jsk}(\theta)) = Z_{jsk}+D_{jsk}$ for $j \leq s$, which generalize the arguments given by Robins et al. (1986) for binary responses.□

A.2 Proof of Large-Strata Consistency

The estimator $\hat{\theta}$ given in (2) has the form $\hat{\theta} = (\Sigma_k\, \hat{t}_k\hat{\theta}^*_k)/(\Sigma_k\, \hat{t}_k)$, where $\hat{\theta}^*_k = [\Sigma_{j=1}^{c-1} X^*_{1jk}(n_{2k} - X^*_{2jk})]/[\Sigma_{j=1}^{c-1} X^*_{2jk}(n_{1k} - X^*_{1jk})]$, and $\hat{t}_k = \Sigma_j\, S_{jk}$. Under the large-strata case, by the weak law of large numbers, $\hat{\theta}^*_k$ is consistent for θ under the common cumulative odds ratio assumption. Consequently, $\hat{\theta}$ is consistent for θ.

Guilbaud (1983) proposed a large sample variance of the Mantel–Haenszel estimator for $2\times2\times K$ tables in which the individual odds ratios need not be equal. For the ordinal response case, Lemma 2, proved by Liu (1995) using the delta method, generalizes it and presents the asymptotic distribution when the odds ratios for all strata and collapsings need not be equal.

LEMMA 2. *As $N \to \infty$ with $Na_{ik} = n_{ik}$, where $a_{ik} \neq 0$ for $i \in \{1,2\}$ and $k \in \{1,\ldots,K\}$, $\sqrt{N}(\hat{\theta}-\theta)$ is asymptotically normal with mean 0 and variance*

$$V_{\hat{\theta}} = \frac{1}{\left(\sum_{l=1}^{K} t_l\right)^2}\left\{\sum_{k=1}^{K}\frac{t_k^2}{a_{1k}}\left[\sum_{a=1}^{c}\pi_{1ak}\left(\frac{\theta^*_k}{A_k}\sum_{b=a+1}^{c}(b-a)\pi_{2bk} - \frac{\theta}{B_k}\sum_{b=1}^{a-1}(a-b)\pi_{2bk}\right)^2\right]\right.$$
$$\left. + \sum_{k=1}^{K}\frac{t_k^2}{a_{2k}}\left[\sum_{b=1}^{c}\pi_{2bk}\left(\frac{\theta^*_k}{A_k}\sum_{a=1}^{b-1}(b-a)\pi_{1ak} - \frac{\theta}{B_k}\sum_{a=b+1}^{c}(a-b)\pi_{1ak}\right)^2\right]\right\}$$
$$- \sum_k (a_{1k}^{-1} + a_{2k}^{-1})[w_k(\theta^*_k-\theta)]^2,$$

where $A_k = \Sigma^c_{a<b}\,(b-a)\pi_{1ak}\pi_{2bk} = \Sigma_j\, \pi^*_{1jk}(1-\pi^*_{2jk})$, $B_k = \Sigma^c_{a>b}\,(a-b)\pi_{1ak}\pi_{2bk} = \Sigma_j\, \pi^*_{2jk}(1-\pi^*_{1jk})$, $t_k = (a_{1k}^{-1}+a_{2k}^{-1})^{-1}[\Sigma_{j=1}^{c-1}\pi^*_{2jk}(1-\pi^*_{1jk})]$, $w_k = t_k/\Sigma_k\, t_k$, $\theta^*_k = \Sigma_{j=1}^{c-1}\pi^*_{1jk}(1-\pi^*_{2jk})/\Sigma_{j=1}^{c-1}\pi^*_{2jk}(1-\pi^*_{1jk})$, and $\theta = \Sigma_k\, w_k\theta^*_k$, for $k \in \{1,\ldots,K\}$.

Under the common cumulative odds ratio assumption, the asymptotic variance simplifies to

$$N\,\mathrm{var}^a(\hat{\theta}) = \frac{\theta^2\sum_k t_k^2/\omega_k}{\left(\sum_k t_k\right)^2} \tag{5}$$
$$= \frac{\sum_k\left\{\sum_j N\,\mathrm{var}^a\left[(R_{jk}-\theta S_{jk})/N\right] + \sum_{j\neq j'} N\,\mathrm{cov}^a[(R_{jk}-\theta S_{jk})/N, (R_{j'k}-\theta S_{j'k})/N]\right\}}{\left[\sum_{k=1}^{K}\sum_{j=1}^{c-1}\mathrm{E}^a(S_{jk}/N)\right]^2},$$

where $\omega_k^{-1} = (1/a_{1k})[\Sigma_{a=1}^{c}\, \pi_{1ak}\eta_{1ak}^2] + (1/a_{2k})[\Sigma_{b=1}^{c}\, \pi_{2bk}\eta_{2bk}^2]$, $\eta_{1ak} = (1/A_k)[\Sigma_{b=a+1}^{c}\,(b-a) \times \pi_{2bk}] - (1/B_k)[\Sigma_{b=1}^{a-1}\,(a-b)\pi_{2bk}]$, and $\eta_{2bk} = (1/A_k)[\Sigma_{a=1}^{b-1}\,(b-a)\pi_{1ak}] - (1/B_k)[\sum_{a=b+1}^{c}(a-b)\pi_{1ak}]$. Following the argument in Robins et al. (1986), one can prove that

$$\frac{\theta^2 t_k^2}{\omega_k} = \lim_{N\to\infty} \frac{\displaystyle\sum_{j=1}^{c-1} D_{jjk} + \sum_{j<j'} 2D_{jj'k}}{N}.$$

Because Z_{jsk}/N is $O(1/N)$ and D_{jsk}/N is $O(1)$, it follows that D_{jsk}/N is the asymptotically nonnegligible portion of $\text{cov}(R_{jk} - \theta S_{jk},\, R_{sk} - \theta S_{sk})/N$ for $j \le s$. Also, since X_{1jk}/n_{1k} and X_{2jk}/n_{2k} are consistent for π_{1jk} and π_{2jk}, respectively, it follows by the construction of $\hat{\phi}_{jsk}(\theta)$ that $\hat{\phi}_{jsk}(\theta)/N$ converges to $\lim_{N\to\infty} D_{jsk}/N$. Furthermore, since $\Sigma_{j=1}^{c-1}\, S_{jk}/N$ converges to t_k, $N\,\widehat{\text{var}}(\hat{\theta}) = [\Sigma_{k=1}^{K}\, \hat{\xi}_k(\hat{\theta})/N]/[\Sigma_{k=1}^{K}\, \Sigma_{j=1}^{c-1}\, S_{jk}/N]^2$ is consistent for the asymptotic variance given in (5).