INFERENCES PROCEDURES FOR THE SYMMETRY IN A CONTINGENCY TABLE

Geng, Zhi
Graduate School of Engineering Sciences, Kyushu University

http://hdl.handle.net/2324/13394
INFERENCE PROCEDURES FOR THE SYMMETRY IN A CONTINGENCY TABLE

By

Zhi Geng*

Abstract

Several inference procedures are discussed for the partial symmetry and the complete symmetry in a two-dimensional contingency table. The inference procedures proposed in the present paper is firstly to select a model fitted to observed frequencies, and secondly to estimate the parameters of the selected model. Such procedures are also called those of the preliminary test estimation in the situation that an estimation follows a testing. We consider here the optimality of significance levels for the inference procedures based on the theory of minimax regret, and compare the optimal significance level with that based on AIC. Finally we propose a weighted estimate method for a contingency table, as a modification of the preliminary test estimation.

1. Introduction

The inference procedures proposed here carry out the selection of a suitable model and then estimate optimally parameters in the model, see Asano [2] and Kitagawa [5]. That is, a parsimonious model is at first selected in view of fitting to observed frequencies, then the parameters are estimated, depending on the selected model. The illustration for simplicity in model-building has been given in Bishop, Fienberg and Holland [3]. The overall variability of the estimates in the simpler model is smaller than that in the model with more parameters, see Altham [1]. We discuss here inference procedures for the symmetry in contingency tables. In a situation that a contingency table may be symmetric, we investigate it firstly, then may pool the frequencies if the symmetry is accepted. In this manner, we may avoid in some extent the risk of pooling asymmetrical frequencies in the table and do not lose so much profit by such a simplicity in model-building.

In section 2, two inference procedures are proposed for the partial symmetry and the complete symmetry in a contingency table. Section 3 gives the expected values and the mean square errors of the estimates based on the inference procedures. In section 4, the optimality of significance levels is discussed in detail, basing on the theory of the minimax regret. Finally, a weighted estimate procedure is proposed as a modified preliminary test estimation.

* Graduate school of Engineering Sciences, Kyushu University 39, Kasuga 816, Japan.
2. Inference Procedures for Symmetry in Contingency Tables

In an $I \times J$ two-dimensional contingency table, let $p_{ij}$ and $n_{ij}$ be the probability and the observed frequency in the cell $(i, j)$, respectively. Assume that the data follow the multinomial distribution with the sample size $N=\sum n_{ij}$. We consider a partial symmetry and a complete symmetry in the contingency table as below.

2.1 Partial symmetry in the contingency table

The partial symmetry is defined as $p_{ij}=p_{ji}$ for two cells $(i, j)$ and $(j, i)$, $i \neq j$. The null hypothesis $H_0$ for the partial symmetry is that $p_{ij}=p_{ji}$, and the alternative $H_1$ is that $p_{ij} \neq p_{ji}$.

Because a cell $(i, j)$ and the corresponding cell $(j, i)$ are only considered, we can test this hypothesis $H_0$ by the Fisher exact test, where $H_0$ is equivalent to a hypothesis $p_{ij}/(p_{ij}+p_{ji})=1/2$.

The maximum likelihood estimates are $\hat{p}_{ij}=n_{ij}/(2N)$ and $\hat{p}_{ji}=n_{ji}/N$ under $H_0$ and $H_1$, respectively. The estimate $\tilde{p}_{ij}$ based on the inference procedure is equal to $\hat{p}_{ij}^{(0)}$, if $\min(n_{ij}, n_{ji})>L$, otherwise $\tilde{p}_{ij}=\hat{p}_{ij}^{(0)}$, where $L=\max\{k: \sum_{i=j}^{m}(\binom{m}{i})\leq 2^{m-1}\alpha\}$, $m=n_{ij}+n_{ji}$ and $\alpha$ is a significance level.

2.2 Complete symmetry in the contingency table

The complete symmetry in an $I \times I$ contingency table is defined as $p_{ij}=p_{ji}$ for $1 \leq i < j \leq I$. We test the hypothesis $H_0: p_{ij}=p_{ji}$ for all $i < j$ against the alternative $H_1: \text{not } H_0$ by using McNemar's test statistic

$$X^2=\sum_{i<j}(n_{ij}-n_{ji})^2/(n_{ij}+n_{ji}),$$

which follows a chi-square distribution with $(I-1)/2$ degrees of freedom under $H_0$. The inference procedure for the complete symmetry is similar to that in section 2.1. The estimate $\tilde{p}_{ij}$ based on the procedure is equal to $\hat{p}_{ij}^{(0)}$ if $X^2 \leq X^2_{(I-1)/2}(\alpha)$, otherwise $\tilde{p}_{ij}=\hat{p}_{ij}^{(0)}$ for all $i \neq j$, where $X^2_{(I/2)}(\alpha)$ is the upper 100$\alpha$% point of the chi-square distribution with $f$ degrees of freedom.

3. Expected Values and Mean Square Errors of Estimates

The estimates of parameters in the contingency table depend on the methods of testing hypothesis and the significance level in the inference procedures. Therefore the expected values and the mean square errors (MSEs) of these estimates are given as below, for both procedures using the exact test and using McNemar's test, respectively.
3.1 For the partial symmetry

The estimate \( \hat{p}_{ij} \) proposed in section 2.1 can be expressed for the partial symmetry as follows,

\[
\hat{p}_{ij} = \begin{cases} 
\frac{(n_{ij} + n_{ji})}{(2N)}, & \text{if } \min(n_{ij}, n_{ji}) > L, \\
\frac{n_{ij}}{N}, & \text{otherwise}.
\end{cases}
\]  

(3.1)

According to the procedure, the expected value of the estimate of \( p_{ij} \) is

\[
E(\hat{p}_{ij}) = \sum_{\mathbf{n}} p(\mathbf{n}) \frac{(n_{ij} + n_{ji})}{(2N)} + \sum_{\mathbf{n}} p(\mathbf{n}) \frac{n_{ij}}{N},
\]

(3.2)

where \( Q_0 = \{ \mathbf{n} : \text{where } \mathbf{n} = [n_{11}, n_{12}, \ldots, n_{1t}, n_{21}, \ldots, n_{tt}] \text{ and } \min(n_{ij}, n_{ji}) > L \}, \) \( Q_1 = \{ \mathbf{n} : \text{where } \min(n_{ij}, n_{ji}) \leq L \}, \) and \( p(\mathbf{n}) \) is the probability of \( n_{ij} \) and \( n_{ji} \) occurring,

\[
p(\mathbf{n}) = C p_{ij}^k p_{ji}^{N-k} (1 - p_{ij} - p_{ji})^{N-n_{ij}-n_{ji}},
\]

where \( C = N! / [(N-n_{ij}-n_{ji})! n_{ij}! n_{ji}!] \).

From (3.2), we have

\[
E(p_{ij}) = \frac{1}{2} \left\{ p_{ij} + p_{ji} + N \sum_{k=0}^{\infty} \frac{(1-p_{ij}-p_{ji})^{N-k} \cdot \sum_{x=0}^{k} \binom{k}{x} (p_{ij} p_{ji})^x (p_{ij}^{1+x} - p_{ji}^{1+x})}{x! (k-x)!} \right\},
\]

(3.3)

where \( g(k) = \max \{ x : \sum_{i=0}^{x} \binom{k}{i} \leq 2^{k-1} \} \).

Similarly, the MSE of the estimate \( \hat{p}_{ij} \) in (3.1) can be shown as

\[
\text{MSE}(\hat{p}_{ij}) = \text{MSE}\left( \frac{n_{ij} + n_{ji}}{2N} \right) + N \sum_{k=0}^{\infty} \frac{(1-p_{ij}-p_{ji})^{N-k} \cdot \sum_{x=0}^{k} \binom{k}{x} (p_{ij} p_{ji})^x (2p_{ij}^{2+x} - p_{ji}^{2+x})}{x! (k-x)!} \cdot \frac{(k-2x) p_{ij}^{k+x} / N - k^2 (p_{ij}^{k+x} + p_{ji}^{k+x}) / (4N^2)}{(N-k)!} + x^2 p_{ji}^{k+x} / N^2 + (k-x)^2 p_{ij}^{k+x} / N^2 \right\},
\]

(3.4)

where \( \text{MSE}\left( \frac{n_{ij} + n_{ji}}{2N} \right) = (p_{ij} + p_{ji}) (1 - p_{ij} - p_{ji}) / (4N) + (p_{ij} - p_{ji})^2 / 4. \)

If \( p_{ij} = p_{ji} \), we have \( E(\hat{p}_{ij}) = p_{ij} \), i.e. \( \hat{p}_{ij} \) is an unbiased estimate of \( p_{ij} \).

Moreover, when \( p_{ij} = p_{ji} \), the MSE in (3.4) can be written as

\[
\text{MSE}(\hat{p}_{ij}) = \text{MSE}\left( \frac{n_{ij} + n_{ji}}{2N} \right) + N \sum_{k=0}^{\infty} \frac{(1-2p_{ij})^{N-k} \cdot \sum_{x=0}^{k} \binom{k}{x} p_{ij}^x (k-2x)^2}{x! (k-x)! (2N^2)} \]

\[
\geq \text{MSE}\left( \frac{n_{ij} + n_{ji}}{2N} \right).
\]

That is, the MSE of \( \hat{p}_{ij} \) is not less than the MSE of \( \hat{p}_{ij}^{(2)} \) when \( p_{ij} = p_{ji} \). The MSE of \( \hat{p}_{ij}^{(2)} \) is equal to \( p_{ij} (1 - p_{ij}) / N \). Also we can show in the similar way that the MSE of \( \hat{p}_{ij} \) is not larger than the MSE of \( \hat{p}_{ij}^{(2)} \) when \( p_{ij} = p_{ji} \).

As an illustration, the numerical evaluation is shown in Table 1, where the sample size is 20 and the significance level is 17% which is the optimal significance level.
given in section 4. In the table, MEAN, VAR and MSE show expected values, variances and mean square errors of estimate $\hat{\pi}_{ij}$, respectively. MEAN$_k$, VAR$_k$ and MSE$_k$ show those of $\hat{\pi}_{ij}^{(k)}$. The efficiency EFFIC of the estimate $\hat{\pi}_{ij}$ to the estimate $\hat{\pi}_{ij}^{(k)}$ is defined as $\text{MSE}_1/\text{MSE}$.

Table 1. TE procedure of a contingency table for the partial symmetry ($N=20$, $\alpha=0.17$)

<table>
<thead>
<tr>
<th>Pij</th>
<th>Pji</th>
<th>MEAN</th>
<th>VAR</th>
<th>MSE</th>
<th>MEAN0</th>
<th>VAR0</th>
<th>MSE0</th>
<th>EFFIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.1000</td>
<td>0.0027</td>
<td>0.0200</td>
<td>0.0000</td>
<td>0.0045</td>
<td>0.0045</td>
<td>1.6703</td>
</tr>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.1288</td>
<td>0.0043</td>
<td>0.0200</td>
<td>0.0051</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.8688</td>
</tr>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>0.1386</td>
<td>0.0073</td>
<td>0.0200</td>
<td>0.0030</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.5125</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.1324</td>
<td>0.0093</td>
<td>0.0200</td>
<td>0.0031</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.4346</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5</td>
<td>0.1204</td>
<td>0.0093</td>
<td>0.0200</td>
<td>0.0030</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.4623</td>
</tr>
<tr>
<td>0.1</td>
<td>0.6</td>
<td>0.1108</td>
<td>0.0080</td>
<td>0.0200</td>
<td>0.0026</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.5534</td>
</tr>
<tr>
<td>0.1</td>
<td>0.7</td>
<td>0.1045</td>
<td>0.0063</td>
<td>0.0200</td>
<td>0.0020</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.7067</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.1012</td>
<td>0.0051</td>
<td>0.0200</td>
<td>0.0011</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.7555</td>
</tr>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>0.1003</td>
<td>0.0047</td>
<td>0.0200</td>
<td>0.0016</td>
<td>0.0045</td>
<td>0.0045</td>
<td>0.9536</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>0.1712</td>
<td>0.0065</td>
<td>0.0200</td>
<td>0.0051</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.0930</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.2000</td>
<td>0.0048</td>
<td>0.0200</td>
<td>0.0030</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.6692</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.2278</td>
<td>0.0063</td>
<td>0.0200</td>
<td>0.0031</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.1260</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4</td>
<td>0.2722</td>
<td>0.0079</td>
<td>0.0200</td>
<td>0.0031</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.2046</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>0.3000</td>
<td>0.0061</td>
<td>0.0200</td>
<td>0.0030</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.7292</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>0.3269</td>
<td>0.0075</td>
<td>0.0200</td>
<td>0.0026</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.2837</td>
</tr>
<tr>
<td>0.3</td>
<td>0.5</td>
<td>0.3731</td>
<td>0.0084</td>
<td>0.0200</td>
<td>0.0026</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.3199</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>0.4000</td>
<td>0.0065</td>
<td>0.0200</td>
<td>0.0020</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.8459</td>
</tr>
<tr>
<td>0.4</td>
<td>0.5</td>
<td>0.4236</td>
<td>0.0081</td>
<td>0.0200</td>
<td>0.0011</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.3937</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.4764</td>
<td>0.0085</td>
<td>0.0200</td>
<td>0.0007</td>
<td>0.0088</td>
<td>0.0080</td>
<td>1.7351</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.5000</td>
<td>0.0060</td>
<td>0.0200</td>
<td>0.0000</td>
<td>0.0088</td>
<td>0.0080</td>
<td>2.0393</td>
</tr>
</tbody>
</table>

3.2 For the complete symmetry

The estimates $\hat{\pi}_{ij}$'s for the complete symmetry can be similarly expressed as in (3.1). But in this situation, we use McNemar's test rather than the exact test. Thus the regions $Q_0$ and $Q_1$ are changed into

$$Q_0 = \{ n : \sum_{i<j} (n_{ij} - n_{ji})^2/(n_{ij} + n_{ji}) < \chi^2_{(1)}(\alpha/2) \}$$

$$Q_1 = \{ n : \sum_{i<j} (n_{ij} - n_{ji})^2/(n_{ij} + n_{ji}) \geq \chi^2_{(1)}(\alpha/2) \}$$

To evaluate the expected values and MSEs, we first introduce two variables

$$X = \sum_{i<j} \left( n_{ij} - (n_{ij} + n_{ji})/2 \right)^2 / (n_{ij} + n_{ji})/2 $$

$$Y = \sum_{i<j} \left( n_{ij} - (n_{ij} + n_{ji})/2 \right)^2 / (n_{ij} + n_{ji})N/2 $$

**Theorem 1.** Two random variables $X$ and $Y$, defined by (3.5) and (3.6), have a same asymptotic distribution, when $E(X) < \infty$.

**Proof.** Let $\hat{\pi}_{ij}$ be the maximum likelihood estimate $n_{ij}/N$. Therefore, by the property of the MLEs, $\hat{\pi}_{ij} - \pi_{ij} = O(1/\sqrt{N})$. Using the Taylor's theorem, we get the expansion of $X$ at $\pi_{ij}$'s,

$$X = \sum_{i<j} \frac{N(\hat{\pi}_{ij} - \pi_{ij})^2}{(\pi_{ij} + \pi_{ji})^2} + \sum_{i<j} \frac{N(\hat{\pi}_{ij} - \pi_{ij})(\pi_{ij} + 3\pi_{ji})}{(\pi_{ij} + \pi_{ji})^2} (\hat{\pi}_{ij} - \pi_{ij})$$
Inference procedures for the symmetry in a contingency table

\[ N(p_{ij} - p_{ji})(p_{ij} + p_{ji})^4 \left( \frac{4N}{(p_{ij} + p_{ji})^3} \right) \]

\[ \cdot \left[ p_{ij}(p_{ij} - p_{ji}) - p_{ij}(p_{ji} - p_{ij}) \right]^2 + O(1/N^3) \]

As \( N \to \infty \), the expected value \( E(X) \) tends to infinite except that \( (p_{ij} - p_{ji})^2 < O(1/N) \).

Similarly, we expand \( Y \) and calculate the difference between \( X \) and \( Y \) under the condition \( (p_{ij} - p_{ji})^2 < O(1/N) \),

\[ X-Y = \sum_{i<j} N(p_{ij} - p_{ji})^2 \left( \frac{p_{ij} + p_{ji}}{N} \right) \left( p_{ij} - p_{ji} \right) + \sum_{i<j} N(p_{ij} - p_{ji})^2 \left( p_{ij} - p_{ji} \right) \]

\[ \cdot \left[ p_{ij} - p_{ji} + p_{ji} \left( 3p_{ij} - p_{ji} \right) - p_{ji} \left( 3p_{ji} - p_{ij} \right) \right] = O(1/N^3) \]

that is, the difference tends to zero as \( N \to \infty \). Therefore, the result follows directly from Rao [6].

**THEOREM 2.** When \( E(Y) < \infty \), the random variable \( Y \), defined by (3.6), follows asymptotically the noncentral chi-square distribution with \( I(I-1)/2 \) degrees of freedom and the noncentrality parameter

\[ \delta = \sum_{i<j} N(p_{ij} - p_{ji})^2 / (p_{ij} + p_{ji}). \]  

**Proof.** Let \( p_{ij} = p_{ji} + \mu_{ij} \) and \( y_{ij} = (n_{ij} - n_{ji}) / \left( (n_{ij} + n_{ji}) N \right)^{1/2} \). It is known that \( y_{ij} \) has the asymptotic normal distribution with the mean \( \mu_{ij} \) and the variance \( \sigma_{ij}^2 \), where

\[ \mu_{ij} = (p_{ij} - p_{ji}) \left( N / (p_{ij} + p_{ji}) \right)^{1/2}, \]

\[ \sigma_{ij}^2 = 1 - (p_{ij} - p_{ji})^2 / (p_{ij} + p_{ji}). \]

Since \( (p_{ij} - p_{ji})^2 = O(1/N) \) for \( E(Y) = \sum[N(p_{ij} - p_{ji})^2 / (p_{ij} + p_{ji}) + 1/2] < \infty \), we have \( \sigma_{ij}^2 \to 1 \). Thus \( Y \), the sum of \( y_{ij} \) over all \( i < j \), follows asymptotically the noncentral chi-square distribution with \( I(I-1)/2 \) degrees of freedom and the noncentrality parameter \( \delta \) given in (3.7). □

**THEOREM 3.** Let a \( J \)-dimensional vector \( W \) follow the multinormal distribution \( N(0, I) \), a \( J \times J \) matrix \( A \) be positive definite and \( 0(\theta) \) be an arbitrary real function. Then

\[ E[\Phi(W'W)W] = \theta E[\Phi(\xi_{J+2, \theta}^2)], \]

\[ E[\Phi(W'W)W'AW] = E[\Phi(\xi_{J+2, \theta}^2)] \text{tr} A + E[\Phi(\xi_{J+4, \theta}^2)] \theta' A \theta, \]

\[ E\left[ \Phi\left( \sum_{i=1}^{K} w_i^2 \right) w_i^2 \right] = E[\Phi(\xi_{J+2, \theta}^2)] + \theta^2 E[\Phi(\xi_{J+4, \theta}^2)], \]

where \( \xi_{K, \theta}^2 \) denotes the noncentral chi-square distribution with \( K \) degrees of freedom and the noncentrality parameter \( \theta \).

**Proof.** The theorem is based on Stein [9] and Sclove, et al. [8]. □

Suppose the following indicator function
Since $X$ follows $X_{j,(i-1)/2}$, by theorems 1 and 2 where $\delta$ is given in (3.7), the expected values and MSEs of the estimates based on McNemar’s test can be given respectively as

$$E(\hat{p}_{ij}) = E\left[ I\{x_2^a \leq X \leq x_1^b \} \right] = p_{ij},$$

$$\text{MSE}(\hat{p}_{ij}) + \text{MSE}(\hat{p}_{ji}) = \frac{p_{ij}(1-p_{ij})}{N}.$$
4. The Optimal Significance Level

Let the minimax regret be the criterion for the optimal significance level. The significance levels based on the criterion are compared with those based on AIC. The regret $R$ is now defined as

$$R = \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}) - \min\left\{ \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}), \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}) \right\}.$$  

Then the optimal significance level $\alpha_{\text{opt}}$ is determined such that the maximum regret $\max_p R$ is minimized.

The maximum difference in the sums of the MSEs between $\hat{p}_{ij}$'s and $\hat{p}_{ij}^{(0)}$'s is

$$R_0 = \max_p \left\{ \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}) - \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}^{(0)}) \right\},$$

and that between $\hat{p}_{ij}$'s and $\hat{p}_{ij}^{(0)}$'s is

$$R_1 = \max_p \left\{ \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}) - \sum_{i \neq j} \text{MSE}(\hat{p}_{ij}^{(0)}) \right\}.$$

For the complete symmetry of $I \times I$ contingency tables, $R_0$ and $R_1$ can be easily obtained as follows,

$$R_0 = \max_p \sum_{i \neq j} \left\{ \left[ \frac{p_{ij} + p_{ji}}{2N} - (p_{ij} - p_{ji})^2 \right] \Pr(\chi_f^{p} \leq \chi_a^{f+2}) \right\}$$

$$+ \left[ \frac{p_{ij} - p_{ji}}{2N} \right] \Pr(\chi_f^{p} \leq \chi_a^{f+2})/2 \right\},$$

$$R_1 = \max_p \sum_{i \neq j} \left\{ \left[ \frac{p_{ij} - p_{ji}}{2N} - \frac{p_{ij} + p_{ji}}{2N} \right] \Pr(\chi_f^{p} \leq \chi_a^{f+2}) \right\}$$

$$- \left[ \frac{p_{ij} - p_{ji}}{2N} \right] \Pr(\chi_f^{p} \leq \chi_a^{f+2})/2 \right\},$$

where $f = I(I-1)/2$. With the significance level $\alpha$ decreasing, $R_0$ decreases but $R_1$ increases. The optimal significance level $\alpha_{\text{opt}}$ can be determined so that $R_0$ is equated to $R_1$.

Sakamoto, Isikuro and Kitagawa [7] proposed an AIC method for model selection for contingency tables. The difference between AIC values under $H_1$ and $H_0$ is given as

$$\text{AIC}(H_1) - \text{AIC}(H_0) = -\chi^2 + 2[I(I-1)/2],$$

where $\chi^2 = 2 \sum_{i \neq j} n_{ij} \log[2n_{ij}/(n_{ij} + n_{ji})]$. Thus the significance level $\alpha_{\text{AIC}}$ based on AIC is determined so that $\chi^2_{11-1/(1-\alpha)} = I(I-1)$.

To compare $\alpha_{\text{opt}}$ with $\alpha_{\text{AIC}}$, we consider the significance levels for $2 \times 2$ contingency tables in Table 3. Table 4 gives the corresponding regrets, and it is shown that $\alpha_{\text{AIC}}$ shows the larger maximum regret than that based on $\alpha_{\text{opt}}$. 
5. Weighted Estimate for Symmetry in Contingency Tables

The inference procedures proposed above may be formulized in the following way. Let \( \hat{\theta}^{(0)} \) and \( \hat{\theta}^{(1)} \) be the maximum likelihood estimates under \( H_0 \) and \( H_1 \), respectively. Suppose \( T \) is a test statistic and \( \Phi(T) \) is a decision function defined as

\[
\Phi(T) = \begin{cases} 
0, & \text{if } H_0 \text{ is accepted}, \\
1, & \text{if } H_0 \text{ is rejected}.
\end{cases}
\]

Then the estimate \( \hat{\theta} \) based on the inference procedures can be written as

\[
\hat{\theta} = \Phi(T)\hat{\theta}^{(1)} + [1 - \Phi(T)]\hat{\theta}^{(0)}
\]

(5.1)

For the normal distribution, Huntsberger [4] tried to generalize \( \Phi(T) \) to a real function and gave a weighted estimate method. According to this weighted estimate method, let us investigate estimates of probabilities in contingency tables for the symmetry.

In an \( I \times I \) contingency table, the weighted estimate is

\[
\hat{\pi}_{ij} = \frac{\Phi n_{ij}}{N} + (1 - \Phi)(n_{ij} + n_{ji})/(2N), \quad \text{for all } i \neq j,
\]

(5.2)

where \( \Phi \) is a parameter to be determined so that MSE(\( \hat{\pi}_{ij} \)) is minimized. Since

\[
\text{MSE(\( \hat{\pi}_{ij} \)) = \frac{\left[ (N-1)\rho_{ij}^2 + \rho_{ji}^2 \right]}{(4N)} + \left[ (1 - \Phi^2)(N-1)/(2N) + \Phi - 1 \right] \rho_{ij} \rho_{ji},}
\]

taking the partial derivative of MSE(\( \hat{\pi}_{ij} \)) with respect to \( \Phi \), we have

\[
\frac{\delta \text{MSE(\( \hat{\pi}_{ij} \))}}{\delta \Phi} = \rho_{ij}(\Phi + 1)[(N-1)\rho_{ij} + 1]/(2N) - \rho_{ji}(1 - \Phi)
\]

\[
\cdot [(N-1)\rho_{ji} + 1]/(2N) - \rho_{ij}^2 + \rho_{ij}\rho_{ji}[1 - (N-1)\Phi/N].
\]

Finally setting it to zero, we obtain that

\[
\Phi = 1 - 2\rho_{ij}(1 + \rho_{ji} - \rho_{ij})/[(N-1)(\rho_{ij} - \rho_{ji})^2 + \rho_{ij} + \rho_{ji}].
\]

(5.3)
Inference procedures for the symmetry in a contingency table

Since $\Phi$ is a function of unknown parameters $p_{ij}$ and $p_{ji}$, substituting $n_{ij}/N$ for $p_{ij}$, we obtain

$$\hat{\Phi} = 1 - 2n_{ij}(N + n_{ji} - n_{ij})/[(N - 1)(n_{ij} - n_{ji})^2 + N(n_{ij} + n_{ji})].$$

Replacing $\Phi$ in (5.2) with the above estimate $\hat{\Phi}$, we obtain

$$\hat{p}_{ij} = \frac{n_{ij}}{N} + \frac{n_{ij}[1-(n_{ij}-n_{ji})/N](n_{ji}-n_{ij})}{(N-1)(n_{ij}-n_{ji})^2 + N(n_{ij}+n_{ji})}.$$

The expected value and MSE of the estimate $\hat{p}_{ij}$ are obtained in the exact formulas

$$E(\hat{p}_{ij}) = p_{ij} + \sum_{k=0}^{N-1} \frac{N!}{(N-k)!} (1-p_{ij}-p_{ji})^{N-k} \left( \sum_{x=0}^{k} \frac{\hat{p}_{ij} \hat{p}_{ji}^{k-x}}{x!(k-x)!} x[1-(2x-k)/N](k-2x)/(N-1)(2x-k)^2 + kN] \right),$$

$$\text{MSE}(\hat{p}_{ij}) = p_{ij}(1-p_{ij})/N + 2 \sum_{k=0}^{N-1} \frac{N!}{(N-k)!} (1-p_{ij}-p_{ji})^{N-k} \left( \sum_{x=0}^{k} \frac{\hat{p}_{ij} \hat{p}_{ji}^{k-x}}{x!(k-x)!} x[1-(2x-k)/N](k-2x)/2[(N-1)(2x-k)^2 + kN] \right).$$

In Table 5, we give the numerical evaluation for the weighted estimates with the sample size 20.

<table>
<thead>
<tr>
<th>pij</th>
<th>pji</th>
<th>mean</th>
<th>var</th>
<th>mse</th>
<th>mean1</th>
<th>var1</th>
<th>mse1</th>
<th>effic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.0926</td>
<td>0.033</td>
<td>0.034</td>
<td>0.1000</td>
<td>0.0020</td>
<td>0.1000</td>
<td>0.0045</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.1775</td>
<td>0.061</td>
<td>0.066</td>
<td>0.1500</td>
<td>0.0026</td>
<td>0.1500</td>
<td>0.0051</td>
</tr>
<tr>
<td>0.1</td>
<td>0.3</td>
<td>0.1134</td>
<td>0.050</td>
<td>0.052</td>
<td>0.2000</td>
<td>0.0030</td>
<td>0.2000</td>
<td>0.0013</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
<td>0.1155</td>
<td>0.056</td>
<td>0.058</td>
<td>0.2500</td>
<td>0.0031</td>
<td>0.2500</td>
<td>0.0026</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5</td>
<td>0.1153</td>
<td>0.059</td>
<td>0.061</td>
<td>0.3000</td>
<td>0.0030</td>
<td>0.3000</td>
<td>0.0043</td>
</tr>
<tr>
<td>0.1</td>
<td>0.6</td>
<td>0.1144</td>
<td>0.060</td>
<td>0.062</td>
<td>0.3500</td>
<td>0.0026</td>
<td>0.3500</td>
<td>0.0065</td>
</tr>
<tr>
<td>0.1</td>
<td>0.7</td>
<td>0.1132</td>
<td>0.060</td>
<td>0.061</td>
<td>0.4000</td>
<td>0.0020</td>
<td>0.4000</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.1122</td>
<td>0.059</td>
<td>0.059</td>
<td>0.4500</td>
<td>0.0020</td>
<td>0.4500</td>
<td>0.0020</td>
</tr>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>0.1116</td>
<td>0.057</td>
<td>0.059</td>
<td>0.5000</td>
<td>0.0000</td>
<td>0.5000</td>
<td>0.0016</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>0.1775</td>
<td>0.061</td>
<td>0.066</td>
<td>0.1500</td>
<td>0.0026</td>
<td>0.1500</td>
<td>0.0051</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.1946</td>
<td>0.056</td>
<td>0.056</td>
<td>0.2000</td>
<td>0.0050</td>
<td>0.2000</td>
<td>0.0030</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4</td>
<td>0.2109</td>
<td>0.063</td>
<td>0.064</td>
<td>0.2500</td>
<td>0.0031</td>
<td>0.2500</td>
<td>0.0056</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2</td>
<td>0.2797</td>
<td>0.076</td>
<td>0.080</td>
<td>0.2500</td>
<td>0.0031</td>
<td>0.2500</td>
<td>0.0056</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>0.2964</td>
<td>0.068</td>
<td>0.069</td>
<td>0.3000</td>
<td>0.0030</td>
<td>0.3000</td>
<td>0.0030</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>0.3123</td>
<td>0.073</td>
<td>0.075</td>
<td>0.3500</td>
<td>0.0026</td>
<td>0.3500</td>
<td>0.0051</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3</td>
<td>0.3817</td>
<td>0.081</td>
<td>0.084</td>
<td>0.3500</td>
<td>0.0026</td>
<td>0.3500</td>
<td>0.0051</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>0.4145</td>
<td>0.073</td>
<td>0.075</td>
<td>0.4500</td>
<td>0.0011</td>
<td>0.4500</td>
<td>0.0036</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4</td>
<td>0.4386</td>
<td>0.075</td>
<td>0.078</td>
<td>0.4500</td>
<td>0.0011</td>
<td>0.4500</td>
<td>0.0036</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.5000</td>
<td>0.062</td>
<td>0.062</td>
<td>0.5000</td>
<td>0.0000</td>
<td>0.5000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

To compare these methods each other, the maximum regrets are given for 2×2 contingency tables with several sample sizes in Table 6.
Table 6. The maximum regrets

<table>
<thead>
<tr>
<th>N</th>
<th>$\alpha_{\text{opt}}$</th>
<th>$\alpha_{\text{sic}}$</th>
<th>Weighted est.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R_0=R_1$</td>
<td>$R_0$</td>
<td>$R_1$</td>
</tr>
<tr>
<td>10</td>
<td>0.0299</td>
<td>0.0286</td>
<td>0.0326</td>
</tr>
<tr>
<td>20</td>
<td>0.0150</td>
<td>0.0143</td>
<td>0.0163</td>
</tr>
<tr>
<td>50</td>
<td>0.0060</td>
<td>0.0057</td>
<td>0.0065</td>
</tr>
</tbody>
</table>

The weighted estimate method brings about the least maximum regret, although it gives no information on the model.

6. Conclusion

We have proposed several inference procedures for the symmetry in contingency tables, and have discussed the optimal significance level. Moreover, a weighted estimate method has been suggested which generalizes the decision function to a real function. In this paper, the maximum likelihood estimates are only considered as the basic estimates, i.e., $\hat{p}_{ij}$'s and $\hat{p}_{ij}$'s are used as the basic estimates under $H_0$ and $H_1$, respectively. Also, we may use other approaches with some advantages to find the basic estimates, e.g., the kernel smoothing technique and Bayesian methods may be preferable in sparse tables. Moreover, it may be possible to apply these procedures to the analyses of independence and loglinear models for contingency tables.

Acknowledgements

The author would like to express his thanks to Professor Ch. Asano for his valuable comments and suggestions.

References


Received October 2, 1987
Communicated by Ch. Asano