

Bayesian Test of Quasi-Independence in a Sparse Two-Way Contingency Table

Sang Gyu Kwak^a, Dal Ho Kim^{1,a}

^aDepartment of Statistics, Kyungpook National University

Abstract

We consider a Bayesian test of independence in a two-way contingency table that has some zero cells. To do this, we take a three-stage hierarchical Bayesian model under each hypothesis. For prior, we use Dirichlet density to model the marginal cell and each cell probabilities. Our method does not require complicated computation such as a Metropolis-Hastings algorithm to draw samples from each posterior density of parameters. We draw samples using a Gibbs sampler with a grid method. For complicated posterior formulas, we apply the Monte-Carlo integration and the sampling important resampling algorithm. We compare the values of the Bayes factor with the results of a chi-square test and the likelihood ratio test.

Keywords: Bayes factor, chi-squared test, likelihood ratio test, quasi-independent, zero cells.

1. Introduction

Sometimes, a few cells have zero counts in the contingency tables. When the categorical variables have many levels, the contingency tables tend to have zero cells. If someone wants test the independence of these tables, then they could use the chi-square test and likelihood ratio test. However, these tests become problematic for these tables because zero cells have expected counts of less than 5. When zero cells exist in a contingency table, the test of independence becomes a test of quasi-independence of the two categorical variables. In this article, we show how to use the Bayes factor (Kass and Raftery, 1995) for a test of quasi-independence when there are zero cells in a contingency table.

There is significant literature on contingency tables that have zero cells. Bishop *et al.* (1975) considered incomplete tables with zero cell counts known as structural zeros. These cells are called sampling zeros and random. It is easy to handle structural zeros by dropping them from the statistics. Brown and Fuchs (1983) discussed log linear models for sparse contingency tables with zero counts. In addition, Baker *et al.* (1985) also discussed structural and sampling zeros.

We can avoid the difficulty about sampling zeros by implementing a test of quasi-independence. This test reflects only the positive cells in the contingency tables. Bishop *et al.* (1975) discussed the notation of quasi-independence for two-way contingency tables that have structural zeros; however, we do not ignore the sampling zeros and use a modified notation with Bishop's one.

This paper derives the Bayes factor for the test of quasi-independence in a contingency table that has zero cells. The rest of the paper is organized as follows. In Section 2, we describe the Bayes factor for the test of quasi-independence in contingency tables with two categorical variables. In Section 3, we show how to compute the Bayes factor introduced in Section 2. In Section 4, we describe a numerical study with an example table. We also present the P -value of chi-square test and the result of a likelihood ratio test. In Section 5, we present some discussions.

¹ Corresponding author: Professor, Department of Statistics, Kyungpook National University, 80 Daehak-Ro, Buk-Gu, Deagu 702-701, Korea. E-mail: dalkim@knu.ac.kr

2. Bayes Factor

We consider two models, one is under the null hypothesis and the other one is under alternative hypothesis. Under the null hypothesis, the two categorical variables are independent. Let n_{ij} , $i = 1, \dots, I$, $j = 1, \dots, J$, denote the cell counts in the two-way contingency tables. Let π_{ij} denote cell probabilities, where $\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} = 1$, $p_i = \sum_{j=1}^J \pi_{ij}$ and $q_j = \sum_{i=1}^I \pi_{ij}$. The null hypothesis of independence is given by $\pi_{ij} = p_i q_j$, where $\sum_{i=1}^I p_i = 1$ and $\sum_{j=1}^J q_j = 1$. And the alternative hypothesis is given by $\pi_{ij} \neq p_i q_j$. We assume that each cell is independent in the IJ cells of the two-way contingency table. In addition, we assume that the total number for the cells consists of cells with sampling zeros and positive counts.

Let \mathcal{N} denote the set of positive cells of the two-way table. Also let $R_i = \{(i, j) : n_{ij} > 0, j = 1, \dots, J\}$, $i = 1, \dots, I$ and $C_j = \{(i, j) : n_{ij} > 0, i = 1, \dots, I\}$, $j = 1, \dots, J$, where R_i is the set of positive counts in i^{th} row, and C_j is the set of positive counts in j^{th} column. Then, the null hypothesis of quasi-independence is given by the restriction,

$$\pi_{ij} = \begin{cases} p_i q_j, & (i, j) \in \mathcal{N}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

We apply the truncated multinomial distributions for the positive cells under the null hypothesis (*i.e.*, the restricted parameter space) and under the alternative hypothesis (*i.e.*, the whole parameter space), respectively. Under the restricted parameter space, we have

$$f(\mathbf{n} | \mathbf{p}, \mathbf{q}) = n! \prod_{(i,j) \in \mathcal{N}} \frac{1}{n_{ij}!} \left\{ \frac{p_i q_j}{\sum_{(i,j) \in \mathcal{N}} p_i q_j} \right\}^{n_{ij}}, \quad n_{ij} > 0, \quad \sum_{(i,j) \in \mathcal{N}} n_{ij} = n. \tag{2.2}$$

Under the whole parameter space, we have

$$f(\mathbf{n} | \boldsymbol{\pi}) = n! \prod_{(i,j) \in \mathcal{N}} \frac{1}{n_{ij}!} \left\{ \frac{\pi_{ij}}{\sum_{(i,j) \in \mathcal{N}} \pi_{ij}} \right\}^{n_{ij}}, \quad n_{ij} > 0, \quad \sum_{(i,j) \in \mathcal{N}} n_{ij} = n. \tag{2.3}$$

It is not easy to handle (2.2) and (2.3). Fortunately, (2.3) can be simplified using transformation. First, we string out the π_{ij} , and denote the new parameter $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_S)'$, where S is the positive cardinality of \mathcal{N} . Similarly string out n_{ij} , $(i, j) \in \mathcal{N}$, and denote the new vector by $\mathbf{n}^* = (n_1^*, \dots, n_S^*)$. Then, (2.3) is

$$p(\mathbf{n}^* | \boldsymbol{\gamma}) \propto \prod_{s=1}^S \left\{ \frac{\gamma_s}{\sum_{s=1}^S \gamma_s} \right\}^{n_s^*}, \quad \gamma_s > 0, \quad \sum_{s=1}^S \gamma_s < 1.$$

We transform $\delta_s = \gamma_s / \sum_{s=1}^S \gamma_s$, $s = 1, \dots, S - 1$ and $\Gamma = \sum_{s=1}^S \gamma_s$. So, $\gamma_s = \delta_s \Gamma$, $s = 1, \dots, S - 1$ and $\gamma_S = (1 - \sum_{s=1}^{S-1} \delta_s) \Gamma$. The Jacobian is Γ^{S-1} . Therefore, we can rewrite (2.3) by

$$p(\mathbf{n}^* | \boldsymbol{\delta}) \propto \left(\prod_{s=1}^{S-1} \delta_s^{n_s^*} \right) \left(1 - \sum_{s=1}^{S-1} \delta_s \right)^{n_S^*}.$$

Let $\mathbf{p} = (p_1, \dots, p_I)$, $\mathbf{q} = (q_1, \dots, q_J)$ and $\boldsymbol{\delta} = (\delta_1, \dots, \delta_S)$. A priori we assume that \mathbf{p} , \mathbf{q} and $\boldsymbol{\delta}$ are all independent, and we assume $\mathbf{p} | \boldsymbol{\mu}_1, \boldsymbol{\tau}_1 \stackrel{iid}{\sim} \text{Dirichlet}(\boldsymbol{\mu}_1, \boldsymbol{\tau}_1)$, $\mathbf{q} | \boldsymbol{\mu}_2, \boldsymbol{\tau}_2 \stackrel{iid}{\sim} \text{Dirichlet}(\boldsymbol{\mu}_2, \boldsymbol{\tau}_2)$ and

$\delta \mid \mu_3, \tau_3 \stackrel{iid}{\sim} \text{Dirichlet}(\mu_3 \tau_3)$ where $\mu_1 = (\mu_{11}, \dots, \mu_{1I})$, $\mu_2 = (\mu_{21}, \dots, \mu_{2J})$, and $\mu_3 = (\mu_{31}, \dots, \mu_{3S})$. For example, let us consider 3×3 table and n_{ij} , $i = 1, 2, 3$, $j = 1, 2, 3$ corresponding cell counts that a unit falls in the i^{th} row and j^{th} column. Let us suppose $n_{13} = n_{22} = n_{31} = 0$ then we have 3 zero cells and 6 positive cells in table. Thus, the positive cardinality (S) is six. So, $\mathbf{n}^* = (n_1^*, \dots, n_6^*)$, where $n_1^* = n_{11}$, $n_2^* = n_{12}$, $n_3^* = n_{21}$, $n_4^* = n_{23}$, $n_5^* = n_{32}$ and $n_6^* = n_{33}$, respectively. Therefore, the cell probability δ is $(\delta_1, \dots, \delta_6)$, where $\delta_6 = 1 - \sum_{s=1}^{6-1} \delta_s$. For hyper-parameters, we assume

$$\pi(\mu_1, \tau_1) = \frac{(I-1)!}{(1+\tau_1)^2}, \quad \pi(\mu_2, \tau_2) = \frac{(J-1)!}{(1+\tau_2)^2} \quad \text{and} \quad \pi(\mu_3, \tau_3) = \frac{(S-1)!}{(1+\tau_3)^2}.$$

Note that this hyper-parameters are noninformative and proper prior. Recall that $\mathbf{p} \mid \mu, \tau \sim \text{Dirichlet}(\mu\tau)$ has the density $f(\mathbf{p} \mid \mu, \tau) = (\prod_{i=1}^k p_i^{\mu_i \tau - 1}) / D(\mu\tau)$, $0 < p_i < 1$, $\sum_{i=1}^k p_i = 1$ where $D(\mu\tau) = \prod_{i=1}^k \Gamma(\mu_i \tau) / \Gamma(\tau)$, $0 < \mu_i < 1$, $\tau > 0$.

Thus, the Bayes factor for test of quasi-independence is

$$\text{BF} = \frac{p(\mathbf{n} \mid H_0)}{p(\mathbf{n}^* \mid H_a)} = \frac{\iint p(\mathbf{n} \mid \mathbf{p}, \mathbf{q}) \pi(\mathbf{p}) \pi(\mathbf{q}) d\mathbf{p} d\mathbf{q}}{\int p(\mathbf{n}^* \mid \delta) \pi(\delta) d\delta}, \tag{2.4}$$

where

$$\pi(\mathbf{p}) = \iint \pi(\mathbf{p} \mid \mu_1, \tau_1) \pi(\mu_1, \tau_1) d\mu_1 d\tau_1, \tag{2.5}$$

$$\pi(\mathbf{q}) = \iint \pi(\mathbf{q} \mid \mu_2, \tau_2) \pi(\mu_2, \tau_2) d\mu_2 d\tau_2, \tag{2.6}$$

$$\pi(\delta) = \iint \pi(\delta \mid \mu_3, \tau_3) \pi(\mu_3, \tau_3) d\mu_3 d\tau_3, \tag{2.7}$$

H_0 is the null hypothesis and H_a is the alternative hypothesis.

The Bayes factor is used to quantify the difference between a model with association and one without association. It is a very useful tool to test for association between two categorical variables (Kass, 1993); however, the calculation is not easy in a complicated situation where there are many zero cells in a contingency table. Instead of using a Metropolis-Hastings algorithm, we will use Gibbs sampling with grid method. This is a useful way to draw parameters because we can avoid finding out some candidate function and calculating the accepting probability. In addition, it could be extended to more complex situation as well.

3. Computation

The problem that remains is how to calculate (2.4). First, the numerator of (2.4) is given by

$$p(\mathbf{n} \mid H_0) = \iiint A_1 \cdot A_2 \cdot \frac{1}{(1+\tau_1)^2} \cdot \frac{1}{(1+\tau_2)^2} d\mu_1 d\tau_1 d\mu_2 d\tau_2, \tag{3.1}$$

where

$$A_1 = \frac{D(\mathbf{n}^{(1)} + \mu_1 \tau_1)}{D(\mu_1 \tau_1)} \cdot \frac{D(\mathbf{n}^{(2)} + \mu_2 \tau_2)}{D(\mu_2 \tau_2)}, \tag{3.2}$$

$$A_2 = \iint \prod_{(i,j) \in \mathcal{N}} \left\{ \frac{1}{\sum_{(i,j) \in \mathcal{N}} p_i q_j} \right\}^{n_{ij}} \cdot \frac{\prod_{i=1}^I p_i^{n_i + \mu_i \tau_1 - 1}}{D(\mathbf{n}^{(1)} + \mu_1 \tau_1)} \cdot \frac{\prod_{j=1}^J q_j^{n_j + \mu_j \tau_2 - 1}}{D(\mathbf{n}^{(2)} + \mu_2 \tau_2)} d\mathbf{p} d\mathbf{q}, \tag{3.3}$$

$\mathbf{n}^{(1)} = (n_{1.}, \dots, n_{I.})$ and $\mathbf{n}^{(2)} = (n_{.1}, \dots, n_{.J})$. Second, the denominator of (2.4) is given by

$$p(\mathbf{n}^* | H_a) = \iint \frac{D(\mathbf{n}^* + \boldsymbol{\mu}_3 \tau_3)}{D(\boldsymbol{\mu}_3 \tau_3)} \cdot \frac{1}{(1 + \tau_3)^2} d\boldsymbol{\mu}_3 d\tau_3, \tag{3.4}$$

where $\mathbf{n}^* = (n_1, \dots, n_S)$. Next, (3.1) and (3.4) are still difficult to perform integration mathematically, so we use the Monte-Carlo integration (Robert and Casella, 1999) for computation. In addition, we draw $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \tau_1, \tau_2, \tau_3)$ using grid method. Grid method provides a good approximation for the complicated conditional posterior density by a discrete probability mass function on some grid points. Let $\boldsymbol{\mu}_{(1i)}, \boldsymbol{\mu}_{(2j)}, \boldsymbol{\mu}_{(3s)}, i = 1, \dots, I, j = 1, \dots, J, s = 1, \dots, S$ denote the vector of all components of $\boldsymbol{\mu}$ excluding the $\mu_{1i}, \mu_{2j}, \mu_{3s}$, respectively. Then, we obtain the conditional posterior densities of $(\mu_{1i} | \boldsymbol{\mu}_{(1i)}, \tau_1, \mathbf{n}), (\mu_{2j} | \boldsymbol{\mu}_{(2j)}, \tau_2, \mathbf{n}), (\mu_{3s} | \boldsymbol{\mu}_{(3s)}, \tau_3, \mathbf{n})$. Note that the conditional posterior distribution of $\mu_{1I}, \mu_{2J}, \mu_{3S}$ is obtained using $\mu_{1I} = 1 - \sum_{i=1}^{I-1} \mu_{1i}, \mu_{2J} = 1 - \sum_{j=1}^{J-1} \mu_{2j}, \mu_{3S} = 1 - \sum_{s=1}^{S-1} \mu_{3s}$, respectively. Therefore, each conditional posterior densities of $\mu_{1i}, \mu_{2j}, \mu_{3s}$ to execute the iterate Gibbs sampler are as follows:

$$\pi(\mu_{1i} | \boldsymbol{\mu}_{(1i)}, \tau_1, \mathbf{n}) \propto \frac{D(\mathbf{n}^{(1)} + \boldsymbol{\mu}_1 \tau_1)}{D(\boldsymbol{\mu}_1 \tau_1)} \cdot A_2, \quad \pi(\mu_{2j} | \boldsymbol{\mu}_{(2j)}, \tau_2, \mathbf{n}) \propto \frac{D(\mathbf{n}^{(2)} + \boldsymbol{\mu}_2 \tau_2)}{D(\boldsymbol{\mu}_2 \tau_2)} \cdot A_2, \tag{3.5}$$

and

$$\pi(\mu_{3s} | \boldsymbol{\mu}_{(3s)}, \tau_3, \mathbf{n}) \propto \frac{D(\mathbf{n}^* + \boldsymbol{\mu}_3 \tau_3)}{D(\boldsymbol{\mu}_3 \tau_3)}. \tag{3.6}$$

Since we expect τ 's to be small, we respectively transform τ_1, τ_2, τ_3 to $\rho_1 = 1/(1 + \tau_1), \rho_2 = 1/(1 + \tau_2), \rho_3 = 1/(1 + \tau_3)$ to draw ρ_1, ρ_2 and ρ_3 efficiently using a grid method. And the Jacobian is $1/\rho_1^2, 1/\rho_2^2, 1/\rho_3^2$ respectively. The each conditional posterior densities of ρ_1, ρ_2, ρ_3 to execute the iterate Gibbs sampler are follows:

$$\pi(\rho_1 | \boldsymbol{\mu}_1, \mathbf{n}) \propto \frac{D(\mathbf{n}^{(1)} + \boldsymbol{\mu}_1 \frac{1-\rho_1}{\rho_1})}{D(\boldsymbol{\mu}_1 \frac{1-\rho_1}{\rho_1})} \cdot A_3, \quad \pi(\rho_2 | \boldsymbol{\mu}_2, \mathbf{n}) \propto \frac{D(\mathbf{n}^{(2)} + \boldsymbol{\mu}_2 \frac{1-\rho_2}{\rho_2})}{D(\boldsymbol{\mu}_2 \frac{1-\rho_2}{\rho_2})} \cdot A_3, \tag{3.7}$$

and

$$\pi(\rho_3 | \boldsymbol{\mu}_3, \mathbf{n}) \propto \frac{D(\mathbf{n}^* + \boldsymbol{\mu}_3 \frac{1-\rho_3}{\rho_3})}{D(\boldsymbol{\mu}_3 \frac{1-\rho_3}{\rho_3})}, \tag{3.8}$$

where

$$A_3 = \iint \prod_{(i,j) \in \mathcal{N}} \left\{ \frac{1}{\sum_{(i,j) \in \mathcal{N}} p_i q_j} \right\}^{n_{ij}} \cdot \frac{\prod_{i=1}^I p_i^{n_i + \mu_{1i} \frac{1-\rho_1}{\rho_1} - 1}}{D(\mathbf{n}^{(1)} + \boldsymbol{\mu}_1 \frac{1-\rho_1}{\rho_1})} \cdot \frac{\prod_{j=1}^J q_j^{n_j + \mu_{2j} \frac{1-\rho_2}{\rho_2} - 1}}{D(\mathbf{n}^{(2)} + \boldsymbol{\mu}_2 \frac{1-\rho_2}{\rho_2})} dpdq. \tag{3.9}$$

For the grid method, we introduce the specific cases for $\boldsymbol{\mu}_1 = (\mu_{11}, \dots, \mu_{1I}), \sum_{i=1}^I \mu_{1i} = 1$ and $\rho_1, 0 \leq \rho_1 \leq 1$. First, we consider μ_{11} where $0 \leq \mu_{11} \leq 1 - \sum_{i'=1, i' \neq 1}^{I-1} \mu_{1i'}$. We draw μ_{11} using the grid method by following these steps:

Table 1: The number of three types of fish caught in three different lakes

	SALMON	TROUT	BASS
KEUKA LAKE	0	2	0
SENECA LAKE	0	4	0
CAYUGA LAKE	13	5	1

- Step 1. Divide 100 intervals $(0, 1 - \sum_{i'=1, i' \neq 1}^{I-1} \mu_{i'})$ and take the point values $T_k, k = 1, \dots, 101$
- Step 2. Calculate 100 mid-points $(M_k, k = 1, \dots, 100)$ for each intervals
- Step 3. Input M_k to the conditional posterior density and calculate the values $(a_k, k = 1, \dots, 100)$
- Step 4. Calculate $b_k = a_k/A$, where $A = \sum_{k=1}^{100} a_k$
- Step 5. Generate $u_1 \sim \text{Uniform}(0, 1)$
- Step 6. Select k^{th} interval made by Step 1, which satisfies $b_k \leq u_1 < b_{k+1}$
- Step 7. Generate $u_2 \sim \text{Uniform}(T_k, T_{k+1})$ and $\mu_{11} = u_2$

We also draw μ_{12} with the above algorithm. But we should use the updated μ_{11} when we calculate $1 - \sum_{i'=1, i' \neq 2}^{I-1} \mu_{1i'}$. We obtain $\mu_{1l} = 1 - \sum_{i=1}^{I-1} \mu_{1i}$, where $\mu_{11}, \dots, \mu_{1l-1}$ are updated value. For ρ_1 , we divide 100 intervals between 0 and 1 in Step 1.

Unfortunately, we cannot integrate (3.3) and (3.9) formulas. To solve this, we apply the sampling importance resampling(SIR) algorithm. We take approximation conditional posterior densities that does not have A_2 and A_3 in (3.5) and (3.7), respectively. The ratio of SIR algorithm given by A_2, A_3 , respectively. Then, we calculate A_2 by follow manners: Draw $\mathbf{p}^{(h)}$ and $\mathbf{q}^{(h)}$ from Dirichlet($\mathbf{n}^{(1)} + \mu_1 \tau_1$) and Dirichlet($\mathbf{n}^{(2)} + \mu_2 \tau_2$), respectively. We calculate

$$\hat{A}_2 = H^{-1} \times \sum_{h=1}^H \left[\prod_{(i,j) \in \mathcal{N}} \left\{ \frac{1}{\sum_{(i,j) \in \mathcal{N}} p_i q_j} \right\}^{n_{ij}} \right]. \tag{3.10}$$

We also calculate A_3 using similar manners.

4. Numerical Study

We use a following example for the numerical study. One might be interested in the number of fish of different species caught in three different lakes. The data is given in Table 1 (Francoise, 1999). There are nine cells in the table and four cells have zero counts.

Let $a_i = \sum_{j=1}^J n_{ij}, i = 1, \dots, I, b_j = \sum_{i=1}^I n_{ij}, j = 1, \dots, J$ and $\hat{\lambda}_{ij} = a_i b_j / n$ denote the maximum likelihood estimates under the hypothesis of independence. Then, the Pearson chi-squared statistic is given by $X^2 = \sum_{i,j} (n_{ij} - \hat{\lambda}_{ij})^2 / \hat{\lambda}_{ij}$ and the likelihood ratio statistic is given by $G^2 = 2 \sum_{i,j} n_{ij} \log(n_{ij} / \hat{\lambda}_{ij})$. The value of X^2, G^2 is 10.048, 12.396, respectively. The asymptotic P -value of X^2, G^2 for the test of independence is 0.040, 0.015, respectively. Thus, we can conclude that this table is dependent. In the table, the number of cells with less than 5 expected counts is seven (77.8 %) among nine cells. Therefore, these results may not be a valid test because it uses an asymptotic method, which invokes the Central Limit Theorem, and is not reliable when the data are sparse, skewed or heavily tied (Agresti, 2002).

Next, we calculate Bayes factor (2.4) in following manner. We draw each element of μ 's and τ 's from the conditional posterior density with Gibbs sample using grid method and SIR algorithm. We

calculate the Bayes factor with (3.1) and (3.4) using Monte-Carlo integration. For our numerical study we use 2,000 iterates to “burn in” the Gibbs sampling with grid method for μ 's and τ 's, and take every fifth iterate thereafter to obtain 1,000 iterates. This is a very conservative procedure, but it virtually removes the autocorrelation for all elements of μ 's and τ 's. Further calculation with 10,000 iterates showed virtually no changes in the relevant estimates. The Bayes factor for this table equals 0.0001, which is a strong evidence for dependence. In addition, we checked Fisher's exact test with P -value 0.0093, which also shows a strong evidence for dependence.

5. Discussion

In this paper, we studied the Bayes factor for the test of quasi-independence in a contingency table that has zero cells; in addition, we implemented a numerical study with an example table. Finally, we compared the result of our algorithm with the P -value of chi-square test and the result of likelihood ratio test. It appears that the table that we used in the numerical study is dependent with three kinds of results.

Our computation has several valuable results. First, we compute the Bayes factor for the test of quasi-independence using a grid method. Our algorithm does not need to use a Metropolis-Hastings, although the density of parameter does not have any closed form. This is a quite useful way to draw parameters because we can avoid finding out some candidate function and we do not need to calculate the accepting probability. Second, our algorithm can process any number $I (\geq 1)$ of clusters and $J (\geq 2)$ of column. This method can be extended for an application of test of quasi-independence with a sparse contingency tables. Finally, when we test the quasi-independence with a sparse contingency table, we make a further comparison to the results of the chi-square test, the likelihood ratio test and the Bayes factor.

References

- Agresti, A. (2002). *Categorical Data Analysis*, 2nd Ed., Wiley, New York.
- Baker, R. J., Clarke, M. R. B. and Lane, P. W. (1985). Zero entries in contingency tables, *Computational Statistics and Data Analysis*, **3**, 33–45.
- Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis*, MIT Press, Cambridge, MA.
- Brown, M. B. and Fuchs, C. (1983). On maximum likelihood estimation in sparse contingency tables, *Computational Statistics and Data Analysis*, **1**, 3–15.
- Francoise, V. (1999). *StatNews*, **01**, Spare Contingency Tables, Cornell Statistical Consulting Unit.
- Kass, R. E. (1993). Bayes factors in practice, *Statistician*, **42**, 551–560.
- Kass, R. E. and Raftery, A. E. (1995). Bayes factor, *Journal of the American Statistical Association*, **90**, 773–795.
- Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods*, Springer, New York.

Received April 18, 2012; Revised April 24, 2012; Accepted April 30, 2012