

Bayesian Test of Different Association Structures in Two-Way Contingency Tables

Zahra Saberi

Isfahan University of Technology

Received: 4/6/2016 Approved: 7/1/2017

Abstract. Bayesian methods for exact small-sample analysis with categorical data in $I \times J$ contingency tables are considered. Different structures of association are defined and tested concerning log odds ratios in these tables with fixed row margins. The conditional distribution of sufficient statistics for interesting parameters conditional on the sufficient statistics of other nuisance parameters in the model is obtained and used to eliminate the effect of nuisance parameters. The resulting distribution for the table is Fisher's multivariate noncentral hypergeometric distribution. For Bayesian approach, although computation under this distribution is complicated, a common Bayesian model is considered. Bayes factor is used as a measure of evidence for Bayesian testing of different association structures. The performance of our testing Bayesian approach is compared with that of the classical corrected likelihood ratio test by some simulation studies. Also the Bayesian test of "homogenous association" is applied on a real data set.

Keywords. Bayes factor; homogenous association; small-sample.

MSC 2010: 62F05; 62G07; 62J12.

1 Introduction

To test an association structure in two-way contingency tables by use of classical methods, there exist several test statistics such as Pearson's Chi-square test, the likelihood ratio test or test based on divergence measure (see for example Cressie and Read, 1984; Pardo, 2006). But when the sample sizes are small or cell expected values are smaller than 5 or when data are sparse, these approximation methods are not valid and other methods such as exact inference should be used (Cochran, 1954; Agresti, 2002). There are two classical approaches for exact inference of $I \times J$ contingency tables. These are unconditional and conditional approaches. Unconditional method is not what we concern in this paper.

Exact tests for $I \times J$ tables utilize the Fisher's multivariate noncentral hypergeometric distribution. Freeman and Halton (1951) defined the P-value as the probability of the set of tables with the given margins that are no more likely to occur than the table observed. Yates (1934) used X^2 , the corrected chi-squared statistic. The P-value is then the null value of $P(X^2 > x_{obs}^2)$ for observed value of x_{obs}^2 . Also Sharp (2008) presented an approximation for the G^2 statistic in order to improve the performance of it for testing independence in contingency tables with small samples.

Special algorithms and software for computing classical exact tests for $I \times J$ tables are widely available (e.g., Mehta and Patel, 1983). We recommend these tests when asymptotic approximations may be invalid. However, computing time increases exponentially as n , I , or J increase. However, one can use Monte Carlo to sample randomly from the set of tables with the given margins. The estimated P-value is then the sample proportion of tables having test statistic value at least as large as the value observed. As I and/or J increase, the number of possible values for any test statistic tends to increase. Thus, the conservativeness issue for conditional tests becomes problematic (Agresti, 2002). Problems arising for calculating P-value and also its conservative nature of rejecting H_0 lead us to think about Bayesian approaches which may solve these problems. We like to implement the Bayesian approaches which utilize Bayesian criteria such as Bayes factor or Bayesian P-value for testing hypothesis. These approaches should not be related to sample size and dimension of table. Such Bayesian approaches, considered in this paper, will be also compared with classical approaches, to see if they have better performance.

In contingency tables, the sample proportions are ordinary Maximum

Likelihood (ML) estimators of multinomial cell probabilities. When data are sparse, these can have undesirable features. For instance, for a cell with a sampling zero, 0.0 is usually an unappealing estimate. Early applications of Bayesian methods to contingency tables involved smoothing cell counts to improve estimation of cell probabilities with small samples. Previously Altham (1969, 1971) presented Bayesian analogs of small-sample frequentist test. An alternative approach using normal priors for logit received considerable attention in the 1970s by Leonard (1972). For a review of Bayesian inference for categorical data see Agresti and Hitchcock (2005) and Agresti (2014).

Here, we shall use a conditional Bayesian approach. The approach will be conditioning on all row totals of $I \times J$ tables, i.e. the experimental design will be the same as that of Fisher in his famous example for 2×2 tables. Under different structure of association, we condition as well as on column totals to eliminate the nuisance parameters. So, this paper may be regarded as an extension of Bayesian approach of test of different structure of association for $I \times J$ contingency tables (Saberi and Ganjali, 2013, 2015). Although Bayesian computation are very complex under the sample distribution which is multivariate noncentral hypergeometric distribution, we consider Bayesian approach about the parameters of interest [log odds ratios, a vector of dimension $(I - 1) \times (J - 1)$], then we present a Bayesian test of the desired hypothesis. For testing the hypothesis the Bayes factor and Bayesian P-value will be used as Bayesian evidence.

Section 2 gives exact conditional approaches in $I \times J$ tables. In Section 3 different models for association structure in two way contingency tables are presented, the details of our conditional Bayesian approach for testing different structures of association in $I \times J$ contingency tables are described in Section 4 where we use the Bayes factor as a summary of evidence. In Section 5, Bayesian test for different association structures are presented. Some simulation studies are given in Section 6. In section 7, some real applications are analyzed. In the end, some conclusions are given.

2 Exact Conditional Approach in $I \times J$ Tables

This section reviews the conditional approach to exact inference for categorical data in $I \times J$ contingency tables with fixed row margins. This utilizes the distribution of the vector of sufficient statistics for the parameters of interest, conditional on sufficient statistics for the other parameters in the

model, (nuisance parameters).

First, consider a two-way contingency table having I rows and J columns. Suppose N_{ij} , $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$, is the variable which shows the number of events in the i th row and j th column, and $\mathbf{N}_i = (N_{i1}, N_{i2}, \dots, N_{iJ})'$, for $i = 1, 2, \dots, I$, denote independent multinomial random variables with respective indexes $N_{i+} = \sum_j N_{ij} = n_{i+}$ and probability vector parameters $\pi_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iJ})'$ ($\sum_{j=1}^J \pi_{ij} = 1$ for all i). The conditional distribution of $\mathbf{N} = \{N_{ij}\}$ given $N_{+j} = n_{+j}$ (where $N_{+j} = \sum_{i=1}^I N_{ij}$) for $j = 1, 2, \dots, J$ is given by:

$$f(N_{11}, \dots, N_{1J}, \dots, N_{I1}, \dots, N_{IJ} | n_{1+}, \dots, n_{I+}, n_{+1}, n_{+2}, \dots, n_{+J}; \Theta) = \frac{\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \prod_{i=1}^{I-1} \prod_{j=1}^{J-1} \theta_{ij}^{n_{ij}}}{\sum_{\{n_{ij}; \sum_i n_{ij} = n_{+j}; j=1, 2, \dots, J\}} [\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \prod_{i=1}^{I-1} \prod_{j=1}^{J-1} \theta_{ij}^{n_{ij}}]}; \quad (1)$$

where for $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$; $\theta_{ij} \geq 0$ and nonnegative integers n_{ij} are the observed values of random variables N_{ij} 's consistent with the marginal totals. Also in equation (1) n_{i+} 's, $i = 1, 2, \dots, I$, are row margins, n_{+j} 's, $j = 1, 2, \dots, J$, are the column margins, and θ_{ij} 's, $i = 1, 2, \dots, I - 1$; $j = 1, 2, \dots, J - 1$, are odds ratios of the i th row and j th column with respect to the I th row and J th column, which are defined as $\theta_{ij} = \frac{\pi_{ij}\pi_{IJ}}{\pi_{iJ}\pi_{Ij}}$ and take values 1 under the null hypothesis of independence. This is the multivariate Fisher's noncentral hypergeometric distribution (McCullagh and Nelder, 1989).

Now, in the next section different association structures in an $I \times J$ contingency table with fixed row margins will be discussed.

3 Different Models for Association in Two Way Contingency Tables

Structure I: Independence

First, consider a two-way contingency table having I rows and J columns. Suppose N_{ij} , $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$, is the variable which shows the number of events in the i th row and j th column, and $\mathbf{N}_i = (N_{i1}, N_{i2}, \dots, N_{iJ})'$, for $i = 1, 2, \dots, I$, denote independent multinomial random variables with respective indexes $N_{i+} = \sum_j N_{ij} = n_{i+}$ and probability vector parameters

$\pi_{\mathbf{i}} = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iJ})'$ ($\sum_{j=1}^J \pi_{ij} = 1$ for all i).

Two categorical response variables are defined to be independent if all joint probabilities be equal to the product of their marginal probabilities, i.e.,

$$\pi_{ij} = \pi_{i+}\pi_{+j}, \quad i = 1, \dots, I; \quad j = 1, \dots, J,$$

or, each conditional distribution of Y is identical to the marginal distribution of Y . Thus, two variables are independent when $\pi_{j|1} = \dots = \pi_{j|I}$ for $j = 1, \dots, J$, that is, the probability of any given column response is the same in each row. The conditional distribution of $\mathbf{N} = \{N_{ij}\}$ given $N_{+j} = n_{+j}$ (where $N_{+j} = \sum_{i=1}^I N_{ij}$) for $j = 1, 2, \dots, J$ under independence assumption is multivariate hypergeometric distribution with probability mass function,

$$Pr(N = n|n_{1+}, n_{2+}, \dots, n_{I+}, n_{+1}, n_{+2}, \dots, n_{+J}) = \frac{\prod_{j=1}^J n_{+j}! \prod_{i=1}^I n_{i+}!}{M! \prod_i \prod_j n_{ij}!},$$

where $M = \sum_i n_{i+} = \sum_j n_{+j}$ and n_{ij} 's are the observed counts of the table (see Agresti, 2002, page 97).

Structure II: Homogenous Association Model

Consider a two-way contingency table having I rows (for random variable X) and J columns (for random variable Y). An $I \times J$ contingency table has homogeneous association over the whole table when,

$$\theta_{11} = \theta_{12} = \dots = \theta_{(I-1)(J-1)} = \theta,$$

where $\theta > 0$ is an unknown parameter. Then the effect of X is the same at each category of Y . In this model, $\theta = 1$ is equivalent to independence model. Having this assertion, distribution of $\mathbf{N} = (N_{11}, \dots, N_{1J}, \dots, N_{I1}, \dots, N_{IJ})$ (under homogenous association model) would be changed to a distribution with one parameter by the following form,

$$f(N_{11}, \dots, N_{1J}, \dots, N_{I1}, \dots, N_{IJ} | n_{1+}, \dots, n_{I+}, n_{+1}, \dots, n_{+J}, N; \theta) = \frac{\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \theta^{\sum_{i=1}^{I-1} \sum_{j=1}^{J-1} n_{ij}}}{\sum_{\{n_{ij}; \sum_i n_{ij} = n_{+j}; j=1, 2, \dots, J\}} \left[\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \theta^{\sum_{i=1}^{I-1} \sum_{j=1}^{J-1} n_{ij}} \right]}$$

where for $\theta \geq 0$ and nonnegative integers n_{ij} , $\sum_i \sum_j n_{ij} = M$ (sample size). In order to interpret the homogenous association model in two way contingency tables, consider a 2×3 contingency table. If we consider X as a response variable and Y as a 3-level factor, homogenous association corresponds with,

$$\frac{\pi_{1|1}\pi_{2|3}}{\pi_{1|3}\pi_{2|1}} = \frac{\pi_{1|2}\pi_{2|3}}{\pi_{1|3}\pi_{2|2}},$$

where $\pi_{i|j} = p(X = i|Y = j)$. This means that,

$$\frac{\pi_{11}\pi_{23}}{\pi_{13}\pi_{21}} = \frac{\pi_{12}\pi_{23}}{\pi_{13}\pi_{22}},$$

or equivalently,

$$\frac{\frac{\pi_{11}}{\pi_{21}}}{\frac{\pi_{13}}{\pi_{23}}} = \frac{\frac{\pi_{12}}{\pi_{22}}}{\frac{\pi_{13}}{\pi_{23}}}.$$

This expression means that odds of the first and second categories of variable Y in respect to the third category of this variable are the same. Hence, homogenous association may be interpreted as: X has the same effect on the first and second category of Y in respect to the third category of Y .

Structure III: Extended Homogenous Association

For the third model of association, we consider the following structure for a given k ,

$$\theta_{ij} = \theta, \quad \text{for } i = 1, \dots, i^*; j = 1, \dots, J - 1,$$

$$\theta_{ij} = k\theta, \quad \text{for } i = i^* + 1, \dots, I - 1; j = 1, \dots, J - 1.$$

This assertion may be interpreted as saying that the first i^* levels of X (as row variable) has the same association (θ) on the variable (Y) with respect to the last category of Y whereas the association of the last $I - i^* - 1$ levels of X is θ multiplied by k . The conditional distribution of $\mathbf{N} = (N_{11}, \dots, N_{1J}, \dots, N_{I1}, \dots, N_{IJ})$ under this structure is given by,

$$f(N_{11}, \dots, N_{1J}, \dots, N_{I1}, \dots, N_{IJ} | n_{1+}, \dots, n_{I+}, n_{+1}, \dots, n_{+J}, N; k, \theta) =$$

$$\frac{\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \theta^{\sum_{i=1}^{I-1} \sum_{j=1}^{J-1} n_{ij}} k^{\sum_{i=i^*+1}^{I-1} \sum_{j=1}^{J-1} n_{ij}}}{\sum_{\{n_{ij}; \sum_i n_{ij} = n_{+j}; j=1, 2, \dots, J\}} \left[\prod_{i=1}^I \binom{n_{i+}}{n_{i1}, n_{i2}, \dots, n_{iJ}} \theta^{\sum_{i=1}^{I-1} \sum_{j=1}^{J-1} n_{ij}} k^{\sum_{i=i^*+1}^{I-1} \sum_{j=1}^{J-1} n_{ij}} \right]}$$

where $k > 0$, $\theta \geq 0$ and nonnegative integers n_{ij} are consistent with the marginal totals. Note that k is a known constant. In this model, $\theta = 1$ implied that the first i^* levels of row variable are independent from Y , but the last $I - i^* - 1$ levels of X and Y are correlated. This model is more appropriate to be used for analyzing contingency tables with ordinal variables.

4 The Bayesian Approach for Testing Different Structures of Association in $I \times J$ Contingency Tables Using Bayes Factor

In this section the Bayesian approaches for testing different structures of association in an $I \times J$ contingency table are presented and Bayes factor as a summary of evidence in $I \times J$ tables is described. Hypothesis testing, using classical methods, have high power when the sample size is large. However, in some applications, where the sample size is small, classical methods are so conservative. There are examples in which researchers intuitively see some evidence against the null hypothesis, but for any possible value of the test statistic that a classical method suggests, they have to say “there is no or weak evidence to reject H_0 ” (An example, in which this may happen is, Fisher’s exact test, Fisher, 1935, pp. 11-25).

Ganjali and Berridge (2008) present a Bayesian method for Fisher’s tea taster example. Saberi et al. (2010) describe a comparison between Bayesian and classical testing procedures for the association parameter in a 2×2 contingency table. Also Saberi and Ganjali (2013) present a conditional Bayesian approach for testing independence in two-way contingency tables. In this section we extend these works and present a Bayesian approach for testing different models of association in an $I \times J$ contingency table, where there are $(I - 1)(J - 1)$ association parameters.

4.1 Bayes Factor as a Measure of Evidence

For testing $H_0 : \Theta \in \Theta_0$ versus $H_1 : \Theta \in \Theta_1$ using Bayesian approaches, Bayes factor is used to compare the two hypotheses H_0 and H_1 . The Bayes factor is a summary of the evidence provided by the data in favor of one

scientific theory as opposed to another as pointed out by Kass and Raftery (1995). The Bayes factor is the ratio of the posterior probabilities of the null and the alternative hypotheses over the ratio of the prior probabilities of the null and the alternative hypotheses, i.e.,

$$B_{01}^{\pi} = \frac{\frac{P(\theta \in \Theta_0 | \mathbf{x})}{P(\theta \in \Theta_1 | \mathbf{x})}}{\frac{P(\theta \in \Theta_0)}{P(\theta \in \Theta_1)}}.$$

This ratio evaluates the modification of the odds of H_0 against H_1 due to the observations and can naturally be compared to 1, although an exact comparison scale can only be based upon a loss function. In general, the Bayes factor depends on prior information, but is still proposed as an objective Bayesian answer, since it partly eliminates the influence of the prior modelling and emphasizes the role of the observations. Actually, it can be perceived as a Bayesian likelihood ratio since, if π_0 is the prior distribution under H_0 and π_1 the prior distribution under H_1 , B_{01} can be written as,

$$B_{01}^{\pi} = \frac{\int_{\Theta_0} f(\mathbf{x}|\theta)\pi_0(\theta)d\theta}{\int_{\Theta_1} f(\mathbf{x}|\theta)\pi_1(\theta)d\theta} = \frac{m_0(\mathbf{x})}{m_1(\mathbf{x})}, \quad (2)$$

where $m_0(\mathbf{x})$ and $m_1(\mathbf{x})$ are marginal distribution of vector of observations, \mathbf{x} , under H_0 and H_1 , respectively.

When $\Theta_0 = \{\Theta_0\}$, the point null case, the value of π_0 would be 0 with the above definition. To overcome this difficulty, π_0 and π_1 are chosen such that $\pi_0 + \pi_1 = 1$ and a prior, denoted by $g(\Theta)$, is defined on $\Theta \in \Theta_1$ which is proper. It follows that the Bayes factor reduces to

$$B_{01} = \frac{f(\mathbf{x}|\Theta_0)}{m(\mathbf{x})}, \quad (3)$$

where $m(\mathbf{x}) = \int_{\Theta \in \Theta_1} f(\mathbf{x}|\Theta)g(\Theta)d\Theta$.

Since the Bayes factor as defined in (2) or (3) can be interpreted as the comparative support of the data for H_0 versus H_1 , small values of B_{01} indicate evidence against H_0 . Guidelines for the use of Bayes factors as a measure of evidence in testing hypotheses are given in Jeffreys (1961) and in Kass and Raftery (1995). The former states that B_{01} between 0.1 and 0.3162 indicates moderate or substantial evidence against H_0 , B_{01} between 0.01 and 0.1 indicates strong evidence against H_0 , and B_{01} less than 0.01 indicates decisive evidence against H_0 .

5 Bayesian Test for Different Models of Association

In this section, the Bayesian approaches for testing of different association structures are discussed.

Let reparametrize the conditional distribution of vector \mathbf{N} using $\delta_{ij} = \ln(\theta_{ij})$, $i = 1, 2, \dots, I - 1$; $j = 1, 2, \dots, J - 1$, in order to provide a natural prior distribution for $\mathbf{\Delta} = (\delta_{11}, \dots, \delta_{1(J-1)}, \dots, \delta_{(I-1)(J-1)})$. When two variables are independent the vector of log odds ratios $\mathbf{\Delta}$ is $\mathbf{0}$. In general $\mathbf{\Delta}$ is symmetric about $\mathbf{0}$.

It is also known that the empirical estimates of the log odds ratios based on the observed data is approximately normally distributed in studies of even moderate sample sizes. This reparametrization helps us because we can use the multivariate normal distribution as a natural prior distribution for the vector of log odds ratios (McCullagh and Nelder, 1989).

In this paper, we consider the test of different association structures. In the following the Bayes approaches for doing these tests in these tables are presented.

5.1 Bayesian Approaches for Test of Independence

In this subsection, we consider the test of independence against any kind of association which is the test of $H_0 : \mathbf{\Delta} = \mathbf{0}$ against $H_1 : \delta_{ij} \neq 0$; at least for one (i, j) , $i = 1, 2, \dots, I - 1$; $j = 1, 2, \dots, J - 1$.

For test of independence in $I \times J$ contingency tables let us assume a multivariate normal distribution with mean $\underline{\mu}_0$ and covariance matrix Σ_0 , i.e.,

$$\pi(\mathbf{\Delta}) \sim N_{(I-1)(J-1)}(\underline{\mu}_0, \Sigma_0),$$

as our prior distribution, where $\underline{\mu}_0$ and Σ_0 are known.

For this prior, the Bayes factor for testing independence is given by

$$\begin{aligned} B_{01} &= \frac{f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta} = \mathbf{0})}{\int_{\mathbf{\Delta}^*} f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta}) \pi(\mathbf{\Delta}) d\mathbf{\Delta}} \\ &= \frac{f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta} = \mathbf{0})}{\int_{\mathbf{\Delta}^*} |\Sigma_0^{-\frac{1}{2}}| f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta}) \varphi(\Sigma_0^{-\frac{1}{2}} (\mathbf{\Delta} - \underline{\mu}_0)) d\mathbf{\Delta}}, \end{aligned}$$

where $\varphi(\cdot)$ is the density function of the standard multivariate normal distribution. This form helps us to approximate the mean of $f(\mathbf{N} = \mathbf{n}|N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta})$ by Mont Carlo method, using sample means of the simulated values of the multivariate normal distribution with mean $\underline{\mu}_0$ and covariance matrix Σ_0 . So the denominator of B_{01} can be calculated and as the numerator is given by (1), the value of B_{01} can be obtained.

5.2 Bayesian Approaches for Test of Homogenous Association Model in Two Way Contingency Tables

In this subsection, we consider the test of homogenous association, i.e., ($H_0 : \delta_{11} = \delta_{12} = \dots = \delta_{(I-1)(J-1)} = \delta$) against $H_1 : \delta_{ij} \neq \delta$; at least for one (i, j) , $i = 1, 2, \dots, I - 1$; $j = 1, 2, \dots, J - 1$.

For Bayesian test of homogenous association model in $I \times J$ contingency tables, we would determine two priors distribution, one for δ under null hypothesis and the other for vector of log odds ratios, $\mathbf{\Delta}$. Let us assume a normal distribution with mean μ_0 and variance σ_0^2 as a prior distribution of δ under the null hypothesis, $\pi_0(\cdot)$, i.e.,

$$\pi_0(\delta|\mu_0, \sigma_0^2) \sim N(\mu_0, \sigma_0^2),$$

where μ_0 and σ_0^2 are considered as known values. Also we consider a multivariate normal distribution with mean $\underline{\mu}'_0$ and covariance matrix Σ'_0 as a prior distribution of $\mathbf{\Delta}$ in alternative hypothesis, $\pi_1(\cdot)$, i.e.,

$$\pi_1(\mathbf{\Delta}) \sim N_{(I-1)(J-1)}(\underline{\mu}_0, \Sigma_0),$$

where $\underline{\mu}'_0$ and Σ'_0 are known mean vector and covariance matrix, respectively.

For this prior, the Bayes factor (5) for testing homogenous association model is given by,

$$\begin{aligned} B_{01} &= \frac{\int_{-\infty}^{\infty} f(\mathbf{N} = \mathbf{n}|N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \delta) \pi_0(\delta) d\delta}{\int_{\Theta_1} f(\mathbf{N} = \mathbf{n}|N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta}) \pi_1(\mathbf{\Delta}) d\mathbf{\Delta}}, \\ &= \frac{\int_{\mathbb{R}} \sigma_0^{-1} f(\mathbf{N} = \mathbf{n}|N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \delta) \phi\left(\frac{1}{\sigma_0}(\delta - \mu_0)\right) d\delta}{\int_{\Theta_1} |\Sigma_0|^{-\frac{1}{2}} f(\mathbf{N} = \mathbf{n}|N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \mathbf{\Delta}) \varphi\left(\Sigma_0^{-\frac{1}{2}}(\mathbf{\Delta} - \underline{\mu}_0)'\right) d\mathbf{\Delta}}, \end{aligned}$$

where $\Theta_1 = \mathbb{R}^{(I-1)(J-1)}$. Also $\phi(\cdot)$ and $\varphi(\cdot)$ are the density function of

the standard univariate and multivariate normal distribution, respectively. These forms help us to approximate the means of $f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \delta)$ and $f(\mathbf{N} = \mathbf{n} | N_{1+}, \dots, N_{I+}, N_{+1}, \dots, N_{+J}, M, \Delta)$ by Mont Carlo method, using sample means of the simulated values of the univariate and multivariate normal distribution with (μ_0, σ_0^2) and $(\underline{\mu}_0, \Sigma_0)$ sets of parameters. So the numerator and denominator of B_{01} can be calculated and the value of B_{01} can be obtained.

Since the Bayesian test for the extended homogenous association model can be presented in the same manner as the homogenous model, we do not bring it here.

6 Simulation Studies

In this section, we present some simulation studies to consider the performance of the Bayesian approaches for testing of different association structures in a 2×3 contingency table with given margins. Four different values of $\Theta = (\theta_{11}, \theta_{12})$ and different values of m with assumption $n_{1+} = 2m$ and $n_{2+} = n_{+2} = n_{+1} = m$, are chosen. The results of these simulation studies are given in Tables 1 and 2. Also, in this section in order to compare the performance of Bayesian and classical approaches, we have done a simulation study that uses the classical testing methods of corrected Pearson chi-squared, corrected likelihood ratio chi-squared and Fisher's exact test for doing test of independence in two-way contingency tables. In order to correct Pearson chi-squared and likelihood ratio chi-squared test for testing independence in two-way contingency tables with small samples, we used Williams' correction (Williams, 1976) (we have used R software for simulation studies. Also data from different association structures are simulated by use of BiasedUrn package in R).

In Table 1, for our simulation study of testing independence, the percentage of times that the various two-sided tests reject H_0 (or power) was recorded for all observed simulations from the noncentral ($\Theta \neq \mathbf{1}$) or central [$\Theta = (\theta_{11}, \theta_{12}) = (1, 1)$] multivariate Fisher's hypergeometric distribution, assuming various values of the true value Θ .

Recalling the guidelines discussed in Section 4.1 concerning the use of Bayes factors in testing hypotheses, it was desired to use an upper bound cutoff for the Bayes factor of at most 0.1 for rejecting H_0 .

The results of Table 1 show that, for small samples of $m = 2$ ($M = 6$, M is sample size), the Bayesian approach, Fisher's exact test and corrected

Pearson's chi-squared test are very conservative. When the sample sizes are increased, $m = 4, 8, 16$ and 32 , the performance of likelihood ratio test is better than the other classical approaches. But in comparison with results of Bayesian approach, the Bayesian approach has the highest power in rejecting H_0 in favor of H_1 . Also Bayesian approach is less conservative than the other approaches. So comparison between the results of all simulated tables implies that Bayesian approach has better performance of doing test of independence in $I \times J$ contingency tables. The more value of θ the more is the probability of rejecting H_0 and this is the highest when using the Bayesian approach (except when $m = 2$).

For test of homogenous association, classical test has not yet been introduced (since, there is not an estimate for θ under the null hypothesis). So, a simulation study have been done in order to study the performance of the Bayesian approach presented in Subsection 5.2. The results of Table 2 show that, for small samples $m = 2$ ($M = 6$, M is sample size), Bayesian approach is conservative. When the sample sizes are increased, for $m = 4, 8, 16$ and 32 and k is far from 1, the powers of rejecting H_0 in favor of H_1 are increased.

7 Real Applications

Application 1. Table 3 is a 2×3 contingency table which is extracted from a report on the relationship between aspirin use and heart attacks by the Physicians Health Study Research Group at Harvard Medical School. The Physicians Health Study was a 5-year randomized study of whether regular aspirin intake reduces mortality from cardiovascular disease. Every other day, physicians participating in the study took either one aspirin tablet or a placebo. The study was blind, i.e. those in the study did not know whether they were taking aspirin or a placebo. Of the 11,034 physicians taking a placebo, 18 suffered fatal heart attacks over the course of the study, whereas of the 11,037 taking aspirin, 5 had fatal heart attacks.

For this table corrected chi-squared test gives $\chi^2 = 26.9$ with P – value = 0.00000144. Also likelihood ratio test and corrected likelihood ratio test give respectively, $G^2 = 27.583$ (P – value = 1.01×10^{-6}) and $G_c^2 = 27.267$ (P – value = 1.19×10^{-6}). So classical approaches, show that there is strong evidence against the null hypothesis. Using Bayesian method testing independence, $H_0 : \delta_{11} = \delta_{12} = 0$ (where $\delta_{11} = \ln(\theta_{11})$, $\delta_{12} = \ln(\theta_{12})$), we found $B_{01} = 0.0036$. All tests show that there is strong evidence against the null hypothesis. So, treatment with aspirin has a strong effect on myocardial

infraction rate.

Application 2. Table 4, taken from Helmes and Fekken (1986) that classifies a sample of psychiatric patients by their diagnosis and by whether their treatment prescribed drugs.

For doing Bayesian test of homogenous association through the table ($H_0 : \delta_{11} = \delta_{12} = \delta_{13} = \delta_{14} = \delta$), we consider normal distribution with $\mu = 0$ and $\sigma^2 = 10^2$ as prior distribution of δ under H_0 and multivariate normal distribution with $\underline{\mu} = (\hat{\delta}_{11}, \dots, \hat{\delta}_{14})$ (where $\hat{\delta}_i = \frac{n_{1i}n_{52}}{n_{2i}n_{51}}$) and covariance matrix $\Sigma = 10^2I$ as prior distribution of vector $\Delta = (\delta_{11}, \delta_{12}, \delta_{13}, \delta_{14})$ under the alternative hypothesis. Using Bayesian approach for testing homogenous association over the whole table, we found $B_{01} = 1.95 \times 10^{-9}$, which indicates strong evidence against the null hypothesis.

8 Conclusion

In this paper, Bayesian methods for exact small-sample analysis with categorical data in $I \times J$ contingency tables considered. Different structures of association defined and tested concerning log odds ratios in these tables with fixed row margins. The performance of our testing Bayesian approach for independent and homogenous structures were compared with that of classical corrected likelihood ratio test by some simulation studies. The results of simulations for independent and homogenous structures showed that, for small samples, the Bayesian approach, Fisher's exact test and corrected Pearson's chi-squared test were very conservative. When the sample sizes were increased, the performance of likelihood ratio test was better than the other classical approaches. But in comparison with results of Bayesian approach, the Bayesian approach had the highest power in rejecting H_0 in favor of H_1 . Also Bayesian approaches was less conservative than the other approaches. For test of homogenous association, classical test has not yet been introduced. So, a simulation study was done in order to study the performance of the Bayesian approach for this test. The results showed that, for small samples, Bayesian approaches were conservative. When the sample sizes were increased, the powers of rejecting H_0 in favor of H_1 were increased. So, comparison between the results of all simulated tables implied that Bayesian approach had good performance of doing test of independence and homogenous association, in $I \times J$ contingency tables.

Table 1. Percentage of simulations out of 3000 in which $H_0 : \Theta = (1, 1)$ is rejected in favor of $H_1 : \theta_{ij} \neq 1$, at least for one (i, j) , for different values of m and $\Theta = (\theta_{11}, \theta_{12}) = (\theta, \theta)$ by use of classical approaches [*: Value of Θ for which H_0 is true].

		$\theta_{11} = \theta_{12} = \theta$				
		(1*)	(5)	(10)	(20)	(50)
		$\delta = \ln(\theta)$				
		(0*)	(1.6)	(2.3)	(2.99)	(3.91)
$m = 2$	Common Bayes	0	0	0	0	0
	Fisher's exact test (p – value < 0.1)	0	0	0	0	0
	Corrected Pearson's chi-squared	0	0	0	0	0
	Corrected likelihood ratio test	20.16	38.30	55.08	70.12	85.01
$m = 4$	Common Bayes	0.81	5.12	17.00	33.62	61.75
	Fisher's exact test (p – value < 0.008)	0.44	4.94	13.30	30.63	48.48
	Corrected Pearson's chi-squared	0.36	5.12	13.4	30.72	48.48
	Corrected likelihood ratio	0.50	6.00	15.76	32.22	57.96
$m = 8$	Common Bayes	0.70	27.12	57.20	77.62	95.00
	Fisher's exact test (p – value < 0.007)	0.28	8.06	22.14	44.36	73.58
	Corrected Pearson's chi-squared	0.34	9.64	22.14	45.26	73.78
	Corrected likelihood ratio	0.64	12.88	30.60	54.70	80.10
$m = 16$	Common Bayes	0.40	46.9	82.87	98.12	99.70
	Fisher's exact test (p – value < 0.004)	0.34	23.62	60.38	90.10	98.32
	Corrected Pearson's chi-squared	0.30	24.02	60.94	90.14	98.32
	Corrected likelihood ratio	0.32	25.80	62.66	90.54	98.40
$m = 32$	Common Bayes	0.20	85.12	99.01	99.99	99.99
	Fisher's exact test (p – value < 0.002)	0.18	52.76	92.78	99.86	99.99
	Corrected Pearson's chi-squared	0.18	54.14	93.48	99.84	99.99
	Corrected likelihood ratio	0.18	57.14	94.26	99.84	99.99

Table 2. Percentage of simulations out of 3000 in which $H_0 : \theta_1 = \theta_2 = \theta$ is rejected in favor of $H_1 : \theta_2 \neq \theta_1 (\theta_2 = k \times \theta_1)$ for different values of m and k and $\theta = 0.5$ by use of Bayesian methods [*: Value of k for which H_0 is true].

		k					
		(1*)	(2)	(4)	(6)	(8)	(10)
$m = 2$	Common Bayes	14.67	19.76	23.32	29.48	35.18	40.16
	Fisher's exact test (p – value < 0.007)	0	0	0	0	0	0
	Corrected Pearson's chi-squared	3.18	24.32	29.18	34.25	39.12	47.56
	Corrected likelihood ratio	8.20	35.08	49.58	58.64	69.21	81.05
$m = 4$	Common Bayes	10.66	26.41	30.83	37.46	43.16	51.73
	Fisher's exact test (p – value < 0.007)	1.05	24.35	28.93	34.18	39.15	47.41
	Corrected Pearson's chi-squared	3.18	24.32	29.18	34.25	39.12	47.56
	Corrected likelihood ratio	4.21	26.18	30.41	36.81	40.65	49.08
$m = 8$	Common Bayes	8.33	30.31	38.04	42.67	49.69	57.34
	Fisher's exact test (p – value < 0.007)	3.28	9.62	11.01	25.84	34.56	52.01
	Corrected Pearson's chi-squared	4.30	11.49	14.32	26.75	35.61	52.64
	Corrected likelihood ratio	6.18	14.28	21.10	34.27	40.94	55.74
$m = 16$	Common Bayes	3.10	37.28	47.69	53.86	60.99	68.32
	Fisher's exact test (p – value < 0.1)	1.58	27.56	46.64	52.34	59.83	66.31
	Corrected Pearson's chi-squared	2.14	28.91	46.81	52.85	59.74	66.52
	Corrected likelihood ratio test	2.51	30.42	47.01	53.74	60.25	67.94
$m = 32$	Common Bayes	1.20	43.36	52.64	63.29	86.04	91.66
	Fisher's exact test (p – value < 0.002)	0.82	26.42	44.23	55.78	85.66	91.53
	Corrected Pearson's chi-squared	0.98	26.68	44.98	55.89	85.48	91.56
	Corrected likelihood ratio	1.02	27.45	45.68	56.65	85.98	91.61

Table 3. Cross-Classification of Aspirin Use and Myocardial Infarction [Source: The Steering Committee of the Physicians' Health Study Research Group. Findings from the Aspirin Component of the Ongoing Physicians Health Study. *New Engl. J. Med.* **318**, 262-264, 1988].

	Myocardial Infarction			
	Fatal Attack	Nonfatal Attack	No Attack	Total
Placebo	18	171	10845	11034
Aspirin	5	99	10933	11037
Total	23	270	21778	22071

Table 4. A sample of psychiatric patients by their diagnosis and by whether their treatment prescribed drugs [Source: Helmes, E. and Fekken, G.C. (1986). Reprinted with Permission, *J. Clin. Psychol.*, **42**, 569-576].

Diagnosis	Drug	NoDrug
Schizophrenia	105	8
Affective disorder	12	2
Neurosis	18	19
Personality disorder	47	52
Special symptoms	0	13

Acknowledgement

I would like to thank the Editor and referees for their helpful comments and considerations which improved the quality of the manuscript.

References

- Agresti, A. (2002). *Categorical Data Analysis*, Wiley, New York.
- Agresti, A. (2014). Two Bayesian/frequentist Challenges for Categorical Data Analyses. *Metron*, **72**, 125-132.
- Agresti, A. and Hitchcock, D.B. (2005). Bayesian Inference for Categorical Data Analysis. *Statistical Methods and Applications*, **14**, 297-330.
- Altham, P.M.E. (1969). Exact Bayesian Analysis of a 2×2 Contingency Table, and Fisher's Exact Significance Test. *Journal of the Royal Statistical Society, Series B., Methodological*, **31**, 261-269.
- Altham, P.M.E. (1971). The Analysis of Matched Proportions. *Biometrika*, **58**, 561-576.
- Cochran, W.G. (1954). Some Methods of Strengthening Common χ^2 Tests. *Biometrics*, **10**, 417-451.
- Cressie, N. and Read, T.R.C. (1984). Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society. Series B (Methodological)*, **46**, 440-464.
- Fisher, R.A. (1935). The Logic of Inductive Inference (with Discussion). *Journal of the Royal Statistical Society, Series A.*, **98**, 39-82.
- Freeman, G.H. and Halton, J.H. (1951). Note on an Exact Treatment of Contingency Goodness-of-fit and Other Problems of Significance. *Biometrika*, **38**, 141-149.

Ganjali, M. and Berridge, D. (2008). Conditional Bayesian Hypothesis Testing for 2×2 Tables. *Pakistan Journal of Statistics*, **24**, 37-44.

Helmes, E. and Fekken, G.C. (1986). Effects of Psychotropic Drugs and Psychiatric Illness on Vocational Aptitude and Interest Assessment. *Journal of Clinical Psychology*, **42**, 569-576.

Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press.

Kass, R. and Raftery, A. (1995). Bayes Factors. *Journal of the American Statistical Association*, **90**, 773-795.

Leonard, T. (1972). Bayesian Methods for Binomial Data. *Biometrika*, **59**, 581-589.

Mehta, C.R. and Patel, N.R. (1983). A Network Algorithm for Performing Fishers Exact Test in $r \times c$ Contingency Tables. *Journal of the American Statistical Association*, **78**, 427-434.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, **2nd edn**. London: Chapman and Hall.

Pardo, L. (2006). *Statistical Inference based on Divergence Measures*. London: Chapman and Hall.

Saberi, Z., Ganjali, M., Berridge, M. and Ofoghi, P. (2010). A Comparison of Bayesian and Classical Testing Procedures for the Association Parameter in a 2×2 Contingency Table. *Journal of Applied Probability and Statistics*, **5**, 53-67.

Saberi, Z. and Ganjali, M. (2013). A Conditional Bayesian Approach for Testing Independence in Two-way Contingency Tables. *Kuwait journal of science*, **40**, 103-113.

Saberi, Z. and Ganjali, M. (2015). Bayesian Test of Independence and Conditional Independence of Two Ordinal Variables. *Journal of Statistical Theory and Applications*, **14**, 156-168.

Sharp, S. (2008). Tests for Independence in Two-way Contingency Tables with Small Samples. *Journal of Modern Applied Statistical Methods*, **7**, 180-186.

Williams, D.A. (1976). Improved Likelihood Ratio Tests for Complete Contingency Tables. *Biometrika*, **63**, 33-37.

Yates, F. (1934). Contingency Tables Involving Small Numbers and the χ^2 Test. *Supplement to the Journal of the Royal Statistical Society*, **1**, 217-235.

Zahra Saberi

Department of Mathematical Sciences,

Isfahan University of Technology,

Isfahan 84156-83111, Iran.

email: *z-saberi@cc.iut.ac.ir*