

SENSITIVITY ANALYSIS OF BAYES FACTOR FOR
CATEGORICAL DATA WITH EMPHASIS ON SPARSE
MULTINOMIAL DATA

U. Sangeetha¹, M. Subbiah², M.R. Srinivasan³, B. Nandram⁴,

¹*Department of Management Studies, SSN College of Engineering, Chennai.*

²*Department of Mathematics, L. N. Government College, Ponneri.*

³*Department of Statistics, University of Madras, Chennai.*

⁴*Department of Mathematical Sciences, Worcester Polytechnic Institute, 100
Institute Road, Worcester, MA 01609, United States.*

Abstract: This article considers hypothesis testing using Bayes factor in the context of categorical data models represented in two-dimensional contingency tables. The study includes multinomial model for a general $I \times J$ table data. Other data characteristics such as low as well as polarized cell counts and size of the tables are also considered. The objective is to investigate the sensitivity of Bayes factor taking these features into account so as to understand the performance of non-informative priors itself. Consistency has been studied based on different types of data and using Dirichlet prior with eight different choices for multinomial model followed by a bootstrap simulation. Study has emphasized the reasonable choice of values for the parameters that normally represents the underlying physical phenomena, though partially vague in nature.

Key words: Bayesian methods, Categorical data, Sparseness, Statistical computing.

1. Introduction

Categorical data are generally presented in a contingency table of size $I \times J$ with I representing the number of rows and J as number of columns. In most of the survey sampling, cell counts in $I \times J$ table follow a

multinomial distribution and Chi square test of association is widely applied to understand the association between the variables. However, in practice, it is impossible to have an agreement about the appropriateness of the model to be used unless there is a well-established theoretical framework or a mechanism underlying the problem. Therefore, it is imperative to take into account uncertainties in the model-building process, and so to start with, a set of competing models has to be considered with each model viewed as a different state of a random variable.

Marden (2000) has emphasized the need to expand the scope of hypotheses testing beyond p values and noted that Bayesian inference with automatic computing methods will be a promising approach. Statistical inference on parameter estimation based on Bayesian approach is conceptually straightforward in most of the problems. Once a prior distribution is defined and with a reasonable likelihood function, inference about parameters of interest is obtained from marginal posterior distributions and nuisance parameters are integrated. The role of probability in measuring the uncertainty at each stage of Bayesian estimation process is well established. In order to arrive at the posterior distribution of possible states of the model, prior distribution of each model is updated using the information contained in the data. Now, the inference is drawn from the entire posterior distribution of the most plausible model.

Johnson (2005) proposed methods based on Bayes factors for modeling the sampling distributions of standard test statistics. The study indicated the possible extensions to test statistics associated with categorical data. However, influence of priors and nature of subjectivity affects the sensitivity of results in a significant way. Vanpaemel (2010) has provided exhaustive list of studies that address the sensitivity of priors and has emphasized the importance of prior in sensitivity analysis. Particularly Hashemi (1997) and Nandram and Choi (2007) separated by a decade but similar in approach, have provided motivation to investigate deeply the Bayes factor for categorical data under multinomial design, which is ubiquitous in many social science survey sampling designs and problems.

The above studies provided varying recommendations in literature on the use of Bayes factors for contingency tables based on the study design and size of the table (Upton, 1982; 1992). Recently, Hitchcock (2009) emphasized that hypotheses testing problems related to $I \times J$ tables require elaborate studies in understanding the elegance of Yates correction.

Mirkin (2001) listed different ways to look at Chi square tests and observed that it is an appealing measure for studying the association. The relevance of study design and analyses based on 2×2 tables too need more careful attention (Upton, 1982; Campbell, 2007).

In this paper, algebraic form of Bayes factor, in the context of categorical data represented in two-dimensional contingency tables has been presented based on the underlying sampling design. The multinomial model for a general $I \times J$ table with two competing models of ‘no association’ and ‘association’ between categorical variables has been considered. The important aspects of this study relate to multinomial models on categorical data (i) to understand the sensitivity of priors and (ii) to incorporate the sparseness of data in the case of higher order $I \times J$ tables with zero and / or positive low counts, polarized cell counts and size of the tables. The sensitivity of priors has witnessed an active discussion in Bayes factor applications in both designs and sparseness of data has been well incorporated for 2×2 tables (Subbiah and Srinivasan, 2008).

In the following Section, a quick overview of Bayes factor in general and statistical details for categorical data is presented. In Section 3, comparative analysis based on data collected from literature that exemplify the above listed features are presented followed by a study using bootstrap simulation based on the data extracted and Section 4 has concluding remarks.

2. Bayes Factor

If there are several competing hypotheses or models about a system, then the set of models can be considered as mutually exclusive and exhaustive. A prior probability $p(H_i)$ ($i=1,2,\dots,N$) can be assigned to each hypothesis such that $\sum p(H_i) = 1$, with N denoting the number of hypotheses. After observing data y , the posterior probability of hypothesis H_i is

$$p(H_i | y) = \frac{p(H_i)p(y | H_i)}{\sum_{i=1}^N p(H_i)p(y | H_i)} .$$

where $p(y|H_i)$ is the marginal density which is the expected value of all possible likelihoods.

Then hypothesis i relative to j is of the form

$$\frac{p(H_i | y)}{p(H_j | y)} = \frac{\frac{p(H_i)p(y | H_i)}{\sum_{i=1}^N p(H_i)p(y | H_i)}}{\frac{p(H_j)p(y | H_j)}{\sum_{i=1}^N p(H_i)p(y | H_i)}} = \frac{p(H_i) p(y | H_i)}{p(H_j) p(y | H_j)}.$$

It could be observed that the posterior odds ratio is the product of the prior odds ratio and ratio of the marginal probabilities under each of the hypotheses. Then Bayes factor B_{ij} is defined as

$$B_{ij} = \frac{p(H_i | y)}{p(H_j | y)} = \frac{\frac{p(H_i)p(y | H_i)}{\sum_{i=1}^N p(H_i)p(y | H_i)}}{\frac{p(H_j)p(y | H_j)}{\sum_{i=1}^N p(H_i)p(y | H_i)}} = \frac{\text{posterior odds}}{\text{prior odds}}.$$

B_{ij} is not affected by prior specifications and $B_{ij} > 1$ can be interpreted as the hypothesis H_i to be more plausible than H_j in the light of y . However, the above interpretation holds only when both H_i and H_j are simple hypotheses. Berger and Delampady (1987), Kass (1993), Bernardo and Smith (1994), Kass and Raftery (1995), Goodman (1999), Delampady and Berger (1990), Lavine and Schervish (1999), Ghosh et al (2006) provide a better insight into the concept of Bayes factor. The explicit forms of Bayes factor for multinomial model is derived in the following sub section.

2.1 Bayes Factors for $I \times J$ multinomial model

In the case of general $I \times J$ tables, if x_{ij} ($i=1,2,\dots,I$; $j=1,2,\dots,J$) denotes the observed cell counts, with $r_i = \sum_j x_{ij}$ is the row total $c_j = \sum_i x_{ij}$ is the column total and $n = \sum \sum x_{ij}$ is the grand total then the Multinomial likelihood is

$$f(\mathbf{X} | \theta) = \frac{n!}{\prod_i \prod_j x_{ij}} \prod_i \prod_j \theta_{ij}^{x_{ij}} \text{ and } \sum_i \sum_j \theta_{ij} = 1.$$

Also, the conjugate prior (Gelman et al, 2002) for the proportion parameter vector $\theta = (\theta_{ij})$ could be a multivariate generalization of Beta distribution known as Dirichlet (α_{ij}) with $\alpha_{ij} > 0$ and density function is

$$\pi(\theta) = \frac{\Gamma \alpha}{\prod_i \prod_j \Gamma(\alpha_{ij})} \prod_i \prod_j \theta_{ij}^{\alpha_{ij}-1} \text{ where } \alpha = \sum \alpha_{ij}.$$

The pervasive inferential problem related to a categorical data summarized in contingency tables is testing the statistical independence of two categories of the categorical data. Model H_0 corresponds to the null hypothesis that there is no association between the two categories whereas Model H_1 takes that there is an association between the categories constituting $I \times J$ contingency table.

Then under H_0 , the prior distribution $\pi_0(\theta)$ for the parameter $\theta = (\theta_{ij})$ is based on the law of independence $\theta_{ij} = \Pi_i \Psi_j$ where

$$\Pi_i = \text{Dirichlet}(\gamma_i) \text{ and } \Psi_j = \text{Dirichlet}(\delta_j)$$

Also for the prior $\pi_1(\theta)$ for model H_1 is $\theta = (\pi_{ij}) \sim \text{Dirichlet}(\alpha_{ij})$. Hence the marginal likelihood under the model M_t ($t = 0, 1$) is $p(\mathbf{X} | H_t) = \int f(\mathbf{X} | \theta) \pi_t(\theta) d\theta$.

After suitable integration,

$$p(\mathbf{X} | H_1) = \frac{n!}{\prod_i \prod_j x_{ij}!} \frac{\prod_i \prod_j \Gamma(n_{ij} + \alpha_{ij})}{\Gamma(n + \alpha)} \frac{\Gamma(\alpha)}{\prod_i \prod_j \Gamma(\alpha_{ij})}$$

$$p(\mathbf{X} | H_0) = \frac{n!}{\prod_i \prod_j x_{ij}!} \frac{\Gamma(\gamma) \Gamma(\delta)}{\prod_i \Gamma(\gamma_i) \prod_j \Gamma(\delta_j)} \frac{\prod_i \Gamma(r_i + \gamma_i) \prod_j \Gamma(c_j + \delta_j)}{\Gamma(n + \gamma) \Gamma(n + \delta)}$$

where $\gamma = \sum \gamma_i$; $\delta = \sum \delta_j$.

Hence, the Bayes factor for comparing these two models is

$$\begin{aligned} B_{01} &= \frac{p(\mathbf{X} | H_0)}{p(\mathbf{X} | H_1)} \\ &= \frac{\prod \prod \Gamma(\alpha_{ij}) \Gamma \gamma \Gamma \delta \Gamma(n + \alpha) \prod \Gamma(r_i + \gamma_i) \prod \Gamma(c_j + \delta_j)}{\prod \prod \Gamma(n_{ij} + \alpha_{ij}) \Gamma \alpha \prod \Gamma \gamma_i \prod \Gamma \delta_j \Gamma(n + \gamma) \Gamma(n + \delta)} \end{aligned}$$

However computing B_{01} on log scale will alleviate the problem of overflow that may occur if it is computed directly. Kass and Raftery (1995) have provided appropriate guidelines for interpreting B_{01} and $\log(B_{01})$ as the degree of evidence for H_0 and is as follows;

- $1 < B_{01} < 3$ indicates ‘ H_0 is not worth more than a bare mention’
- $3 < B_{01} < 20$ indicates ‘ H_0 is positive’
- $20 < B_{01} < 150$ indicates ‘strong evidence for H_0 ’
- $150 < B_{01}$ indicates ‘very strong evidence for H_0 ’

3. Comparative Data Analysis

Statistical literature on theory and applications of inferential procedures associated with contingency tables of two-dimensional categorical data provide a few essential characteristics necessary for comparative studies. These include order of tables ($k=IJ$), sample size (n), zero counts, notable polarized cell counts and positive low counts (cell counts not more than 6).

All $I \times J$ data sets are extracted from Agresti (2002), a classical book for social sciences that illustrate most of the issues in categorical data. In line with the above features, tables vary in sizes from 3×3 to 6×4 rectangular tables; 5 % to 43 % of the cells have low counts; grand total spans over a range of 96 to 3600; cell counts vary from 14 to 711. Also, to quantify the polarized cell counts a metric $v = \text{range}/k$ has been used to indicate the nature of distribution of values. The summary of the selected characteristics for the data sets are presented in Table 1.

Table 1: Details of the eight data sets considered for the comparative study of Bayes factor associated with two dimensional contingency tables

Data	Size	Study Variables	Page No	n	v	% of zero counts	% of low counts
I	3 x 3	Education and religious belief	80	2726	60.00	0.00	0.00
II	4 x 4	Income and job Satisfaction	57	96	0.86	6.25	43.75
III	4 x 4	Opinions about premarital sex and teenage birth control	368	926	9.19	0.00	0.00
IV	5 x 5	First and second Purchase	446	521	6.200	4.00	20.00
V	3 x 3	Psychiatric thought and ascribed origin of Schizophrenia	83	282	9.89	0.00	11.11
VI	6 x 4	Mental health status and socio economic status	381	1660	5.00	0.00	0.00
VII	5 x 5	Father's and son's occupational status	447	3600	28.40	0.00	4.00
VIII	6 x 2	Daily average number of cigarettes and disease group	64	2714	46.92	0.00	0.00

v = range/k; n = sum of cell counts

It is necessary to consider three cases of Dirichlet parameters as prior distributions for multinomial model; α_{ij} for model H_1 and γ_i 's and δ_j 's for model H_0 . In all the cases, an equal value for these parameters has been considered with eight choices (0.5 to 2.5) that include either side of Uniform distribution ($\alpha_{ij} = 1$ in $\pi(\theta)$) and Table 2 presents the log Bayes factor in favor of H_0 .

It can be noted that Bayes factors have shown a consistent pattern except in one case (VI data set) within these two groups of choice of Dirichlet parameters. These data sets are of reasonable size with no zeros and the total is of moderately high value (1660) with reasonably non-polarized counts. This may be a significant observation that choice of Dirichlet parameter (the only prior parameters) for a well-behaved data set is more critical. However, sparseness or polarized count compel to select values either less than 1 or greater than 1 for all the Dirichlet distribution parameters so that Bayes factors are not much sensitive for making conclusions based on the usual recommendations.

Table 2: Bayes factors (in natural log scale) for the evidence of null hypothesis (no association) in $I \times J$ data sets under multinomial sampling model.

Data No	Choices for Dirichlet Parameters							
	0.5	0.75	0.9	1	1.5	1.75	2	2.5
I	-23.734	-24.412	-24.654	-24.774	-25.074	-25.109	-25.096	-24.972
II	5.738	5.159	5.045	5.021	5.186	5.346	5.519	5.865
III	-45.260	-46.703	-47.199	-47.437	-47.963	-47.974	-47.885	-47.506
IV	-	-	-	-	-	-	-	-
	120.03	119.88	119.26	118.72	115.26	113.29	111.262	107.13
	7	5	5	9	8	7		2
V	-4.808	-4.761	-4.596	-4.455	-3.558	-3.053	-2.535	-1.492
VI	9.373	6.050	4.680	3.925	1.287	0.458	0.297	1.335
VII	-	-	-	-	-	-	-	-
	306.44	306.50	305.90	305.34	301.42	299.05	296.529	291.20
	1	9	6	4	0	1		2
VIII	-59.768	-59.739	-59.521	-59.326	-58.003	-57.215	-56.378	-54.616

Further, a simulation study has been carried out to supplement the findings. Since the study is focused to consider identified characteristics of $I \times J$ tables, bootstrap simulation are used for the computation of Bayes factor and testing the sensitivity of prior parameters. Based on each of the data sets, 1000 bootstrap samples are generated so that the noted features are expected to be consistently present in the samples. Estimates comprising mean together with its standard error (SE) and 95% limits for confidence interval as 2.5 and 97.5 percentiles are presented in Table 3. Also, Figures 1 and 2 depict the box plots to show the distribution of estimated Bayes factor from bootstrap samples based on data sets I–IV and V–VIII, respectively.

Table 3: Bootstrap estimates for the log Bayes factor in favor of null hypothesis of independence using eight distinct prior choices for Dirichlet parameters.

Data	Bootstrap Estimates	Different choices for Dirichlet prior used in Bootstrap estimation							
		0.5	0.75	0.9	1	1.5	1.75	2	2.5
I	Mean	194.04 1	203.7 76	190.426	194.009	188.858	188.901	-217.623	209.207
	SE(Mean)	9.125	8.906	9.645	9.437	9.441	9.672	9.892	9.427
	2.5 percentile	499.27 5	472.0 66	570.677	507.803	492.420	490.135	-544.224	522.162
	97.5 percentile	-10.706	25.45 4	-9.209	-2.497	-9.606	-9.631	-28.158	-9.658
II	Mean	-8.308	-9.226	-8.976	-9.154	-9.002	-8.458	-8.085	-7.542
	SE(Mean)	0.425	0.407	0.382	0.419	0.380	0.385	0.351	0.340
	2.5 percentile	-20.891	20.12 4	-20.297	-23.819	-19.703	-18.573	-17.715	-17.399
	97.5 percentile	3.778	1.315	1.399	1.640	2.818	2.838	0.527	0.594
III	Mean	-96.277	88.92 7	-91.993	-95.107	-86.625	-93.235	-98.776	-89.338
	SE(Mean)	3.476	3.693	3.976	4.149	3.942	3.867	3.792	3.514
	2.5 percentile	200.41 8	219.1 08	223.156	216.315	232.821	218.204	-228.757	199.759
	97.5 percentile	-17.078	16.26 2	-10.716	-14.755	-2.873	-11.668	-20.318	-12.512
IV	Mean	-72.783	71.17 9	-79.538	-78.764	-76.203	-75.599	-70.285	-79.941
	SE(Mean)	4.365	5.035	5.235	5.042	4.824	5.262	4.303	4.998
	2.5 percentile	243.83 0	279.5 48	293.427	260.655	267.460	277.431	-230.561	258.191
	97.5 percentile	0.246	0.097	-0.663	-1.030	-4.325	-1.107	-1.193	-1.536
V	Mean	-45.527	40.44 5	-47.852	-43.044	-47.209	-43.046	-40.528	-40.252
	SE(Mean)	2.472	2.296	2.237	2.489	2.631	2.482	2.377	2.426
	2.5 percentile	124.08 8	117.0 47	114.346	126.839	129.514	123.095	-128.226	115.307
	97.5 percentile	0.425	0.798	-0.721	2.359	-1.508	6.101	0.452	3.384

	Mean	-35.074	38.775	-38.567	-39.616	-42.416	-48.099	-45.291	-41.941
VI	SE(Mean)	1.884	1.957	2.064	1.789	1.756	1.950	1.678	1.732
	2.5 percentile	-97.896	99.581	104.364	-91.259	-98.098	122.688	-98.845	-93.290
	97.5 percentile	9.634	6.231	6.642	5.515	-1.254	-6.333	-5.021	-1.962
	Mean	1034.296	1049.616	1106.659	1085.007	1070.282	1103.873	1044.887	982.051
VII	SE(Mean)	27.509	32.974	32.583	27.412	30.265	30.007	30.327	30.290
	2.5 percentile	1911.889	2029.766	2249.898	1915.655	2048.573	2145.181	2099.729	1991.054
	97.5 percentile	430.798	343.333	410.322	396.549	363.783	375.844	-382.425	370.977
	Mean	470.470	475.680	493.897	493.070	457.516	452.399	-476.388	470.086
VIII	SE(Mean)	16.361	17.108	17.099	17.385	17.202	17.132	17.341	16.166
	2.5 percentile	977.974	1023.763	1028.482	1063.457	985.272	989.072	-919.026	937.623
	97.5 percentile	-74.202	60.188	109.046	-73.681	-51.111	-73.047	-57.593	-84.118

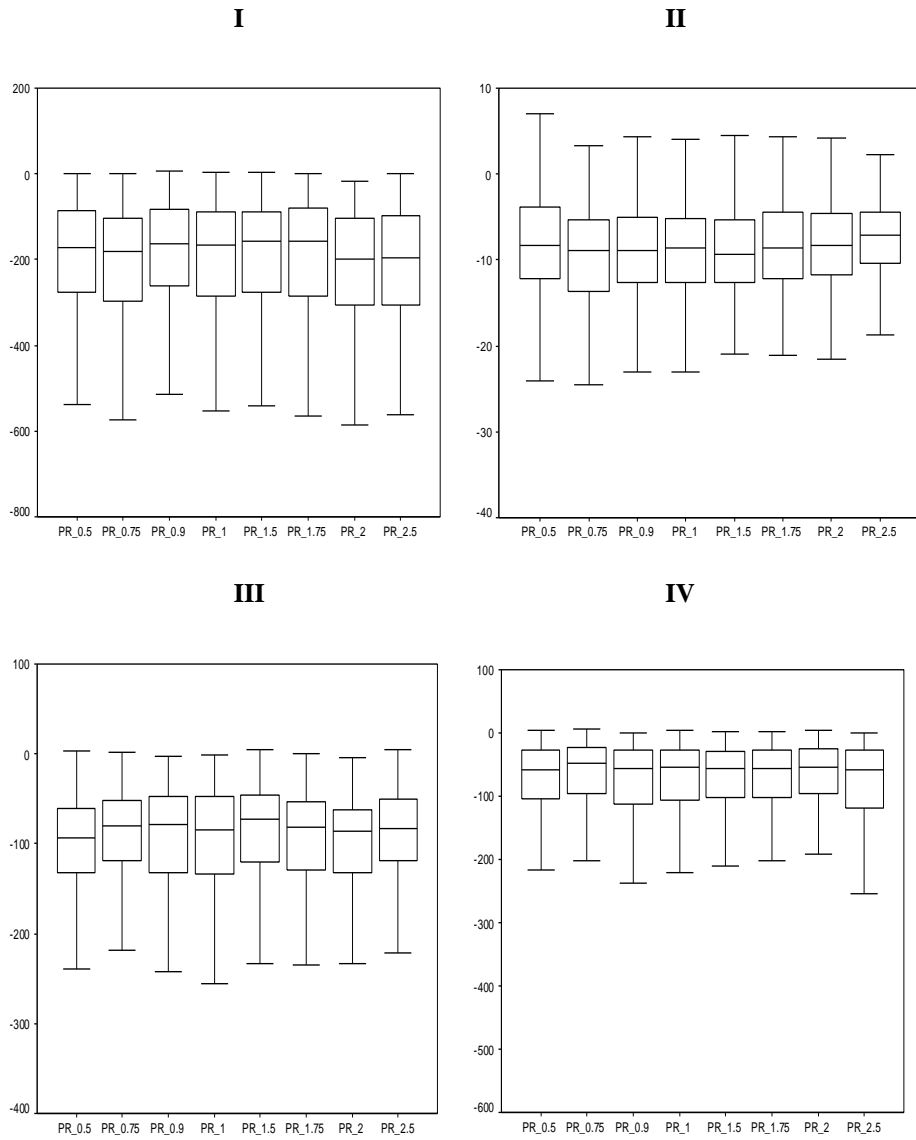


Figure 1: Box plots for estimated Bayes factors from the bootstrap samples based on data sets I – IV. Each plot corresponds to eight different choices of prior parameters.

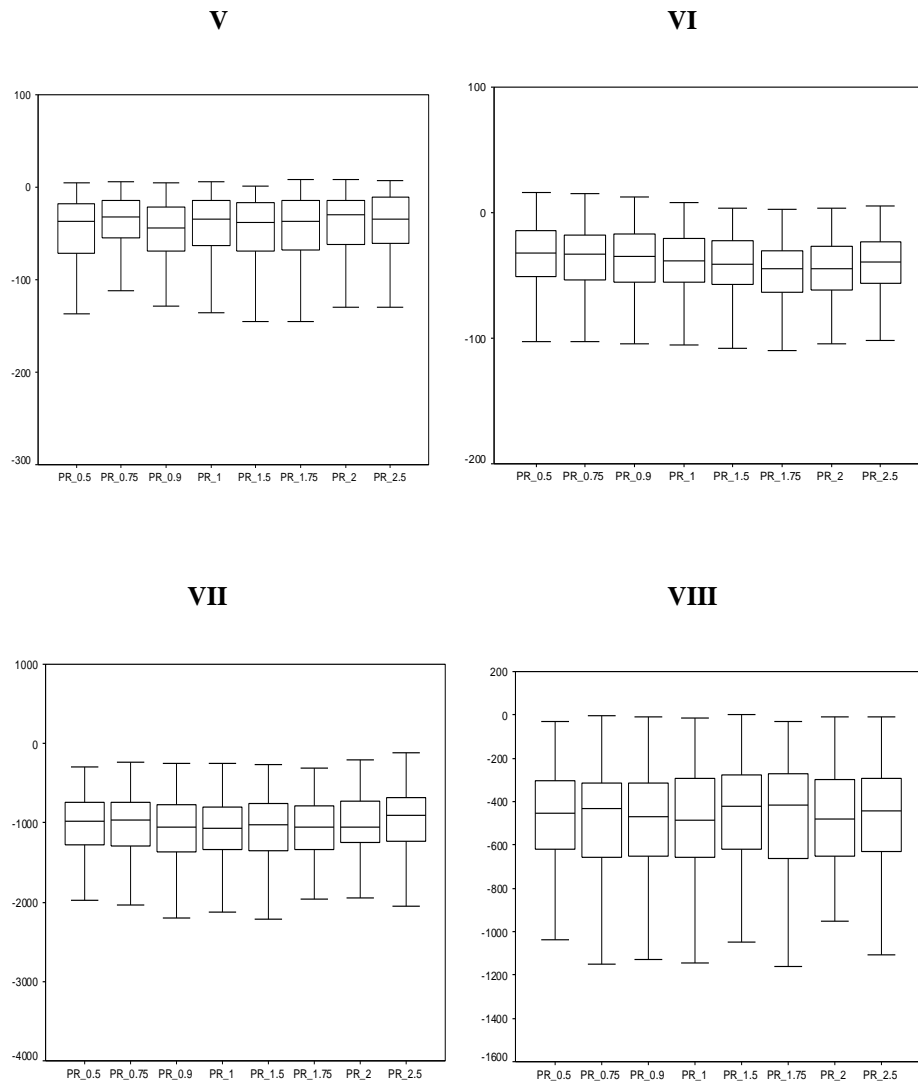


Figure 2: Box plots for estimated Bayes factors from the bootstrap samples based on data sets V – VIII. Each plot corresponds to eight different choices of prior parameters

From the box plots it could be observed that values log Bayes factor are exactly on one side of zero with an exemption in data sets II and VI. Though they could be considered as less-polarized data set, low and / or zero counts would cause the changes in the estimates. This is further

evident in the numerical summaries which indicate the changes in terms of positive and negative estimated values.

Interestingly, such changes are evident when prior values are near to 1 and this tends to provide a method on choice of prior parameters based on the study which deals with rare or non-rare phenomena. Further, among other data sets with moderately polarized and notable low counts such as data set VI, such directional changes in the estimated values are visible from Table 3 but may not be fully captured by box plots. All other tables with sufficiently large cell counts do not yield any such feature irrespective of polarized cell counts.

However, if a point estimate like mean value is compared for a data set over different choices of priors then no directional changes can be observed. Such a pattern prevails among all the data sets of distinct characteristics. This is consistent with the values of corresponding standard errors too and hence reporting point estimate may not be fully sufficient to study the sensitivity of estimates over various choices of parametric values. The entire comparison is to demonstrate the sensitivity of priors in the values of Bayes factors attributed to data characteristics, especially when they are sparse in nature. Hence, sparse data occurrence could be reckoned by the researcher as a priori based on the problem under consideration and the way respondents might behave to the choices of variables. This partially elicited information will help to set the range for the prior parameters to estimate as well as consider the sensitivity analysis.

4. Conclusion

Two contrasting recommendations regarding Bayes factors for $I \times J$ tables and a 2×2 table provide ample scope to investigate the data model, sparseness and pattern of cell counts as it could affect the estimates and thereby the conclusions. Sensitivity of priors on Bayes factor has been accepted in principle; yet understanding the nature of data representing the physical phenomena is important. In such cases, even a controlled vague prior would reflect partial information which tends to reduce the extreme sensitivity as exhibited in Nandram and Choi (2007). Three examples considered in Nandram and Choi (2007) have ignored the distinct features of $I \times J$ tables as listed in the present study and could be well influenced by the specific choice of Dirichlet parameters.

Subsequently, Nandram et al (2013) have provided a test of independence related to data from a two-stage cluster sampling design; simulation study has revealed the fact that Bayes factor will not be sensitive to small changes in the uniform prior. This finding has been attributed to the nature of cell counts that are expected to be larger than zero but one or two cells can have zero counts.

The main objective of the paper is to carry out the inevitable sensitivity analysis for Bayes factor but with reasonable parameter choice related to the problem of interest as observed in Vanpaemel (2010). The present work considered the problem on sparse multinomial tables not only limited to zero counts but also low and / or polarized counts as polarized counts tend to affect estimation of multinomial probabilities in the sense of aberrations (May and Johnson, 2000).

Also, to substantiate the findings, Bootstrap samples have been used to study the sensitivity mainly to preserve the expected features of data / underlying phenomena. Computational tool has been provided for a variety of Dirichlet parameters beyond the usual uniform distribution. The procedure envisages a more flexible approach to handle priors for null hypothesis and replication of these procedures can be done through R codes presented in the Appendix.

Kruschke (2010) has pointed out the need to attempt mildly informed or consensually informed prior distributions rather than objective priors. The present study has made an attempt to divide the parameter space appropriately for choosing prior distributions in obtaining Bayes factor to test the independence related to data from a $I \times J$ contingency table. Such recommendations are suitable for a more realistic sensitivity analysis for Bayes factor computed for various choices of prior values. Hence the researchers are encouraged to adopt Bayes Factor with plausible priors that may be partially informative as a result of theoretical background of the problem. However, a more concrete way to define the distance between the cell counts could be attempted to study the effect on Bayes factors and redefine the recommendations in the analysis of contingency tables.

Acknowledgment

The authors would like to thank the referees and editor for the valuable comments and suggestions which helped to improve the article.

References

- Agresti, A. (2002). *Categorical Data Analysis 2/e*. New Jersey: John Wiley & Sons, Inc. Berger, J.O. and Delampady, M. (1987). Bayesian testing of precise hypotheses (with discussion). *Statistical science*, **2**, 317-348.
- Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian Theory*. New York: John Wiley & Sons.
- Campbell, I. (2007). Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. *Statistics in Medicine*, **26**, 3661-3675.
- Delampady, M. and Berger, J.O. (1990). Lower bounds on Bayes factors for multinomial distributions with application to chi-squared tests of fit. *The Annals of Statistics*, **18**, 1295-1316.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2002). *Bayesian Data Analysis*. London, UK: Chapman & Hall.
- Ghosh, J.K., Delampady, M. and Samanta, T. (2006). *An Introduction to Bayesian Analysis*, New York: Springer-Verlag.
- Goodman, S.N. (1999). Toward Evidence – Based Medical Statistics. 2: The Bayes Factor. *Annals of Internal Medicine*, **130**, 1005-1013.
- Hashemi, L., Nandram, B. and Goldberg, R. (1997). Bayesian analysis for a single 2×2 table. *Statistics in Medicine*, **16**, 1311 – 1328.
- Hitchcock, D.B. (2009). Yates and Contingency Tables: 75 Years Later. *Statistics Electronic Journal for History of Probability and Statistics*, **5**, 1-14.
- Johnson, V.E. (2005). Bayes factors based on test statistics. *Journal of the Royal Statistical Society. B*, **67**, 687-701.
- Kass, R.E. (1993). Bayes Factors in Practice. *The Statistician*, **42**, 551-560.
- Kass, R.E. and Raftery, A.E. (1995). Bayes Factors. *Journal of the American Statistical Association*, **90**, 773-795.
- Kruschke, J.K. (2010). What to believe: Bayesian methods for data analysis, *Trends in Cognitive Sciences* **14**, 293–300.
- Lavine, M. and Schervish, M.J. (1999). Bayes factor: What they are and what they are not. *The American Statistician*, **53**, 119-122.

- Marden, J.I. (2000). Hypothesis Testing: from p values to Bayes factors. *Journal of the American Statistical Association*, **95**, 1316-1320.
- May, W.L. and Johnson W.D. (2000). Constructing two-sided simultaneous confidence intervals for multinomial proportions for small counts in a large number of cells. *Journal of statistical software*, **5**(6).
- Mirkin, B. (2001). Eleven Ways to Look at the Chi-Squared Coefficient for Contingency Tables. *The American Statistician*, **55**, 111-120.
- Nandram, B. and Choi, J.W. (2007). Alternative Tests of Independence in Two-Way Categorical Tables. *Journal of Data Science*, **5**, 217-237.
- Nandram, B., Bhatta, D., Sedransk, J. and Bhadra, D. (2013). A Bayesian test of independence in a two-way contingency table using surrogate sampling *Journal of Statistical Planning and Inference*, **143**, 1392–1408.
- Subbiah, M. and Srinivasan, M.R. (2008). Classification of 2×2 sparse data with zero cells. *Statistics & Probability Letters*, **78**, 3212 – 3215.
- Upton, G.J.G. (1982). A comparison of alternative tests for the 2×2 comparative trial. *Journal of the Royal Statistical Society. A*, **145**, 86 – 105.
- Upton, G.J.G. (1992). Fisher's Exact Test. *Journal of the Royal Statistical Society. A*, **155**, 395 – 402.
- Vanpaeml, W. (2010). Prior sensitivity in theory testing: An apologia for the Bayes factor. *Journal of Mathematical Psychology* **54**, 491–498.

Received June 25, 2013; accepted September 28, 2013.

U. Sangeetha
Department of Management Studies,
SSN College of Engineering, Chennai.

M. Subbiah
Department of Mathematics,
L. N. Government College, Ponneri.

M.R. Srinivasan
Department of Statistics,
University of Madras, Chennai.

B. Nandram
Department of Mathematical Sciences,
Worcester Polytechnic Institute, 100 Institute Road,
Worcester, MA 01609, United States.

Appendix

R code for obtaining Bayes factor for the evidence of null or alternative hypothesis related to test of independence for contingency tables is presented in this Appendix. Input together with the respective symbols used in the code include number of rows(r) and columns (c) of the given contingency table, cell counts as a row vector (ns), and the prior parameters for no association (alp) and association (gam and del) models. Bayes factors in log scale favoring both statements together with the interpretation criterion will be provided; exponentiated version is also given as output.

```
#####Input the size of the contingency table

r=4                                #no of rows
c=4                                #no of coulms
#####Input the original cell counts as a row vector
ns=c(1,3,10,6,2,3,10,7,1,6,14,12,0,1,9,11) #Cell counts row wise
data=matrix(ns,r,c,byrow=TRUE)
rt=0                                #Calculating row totals
for (i in 1:r)
{
rt[i]=sum(data[i,])
}
ct=0                                #Calculating column
totals
for (j in 1:c)
{
ct[j]=sum(data[,j])
}
n=sum(ns)                            #Grand Total
alp = 0
k=r*c
for (k1 in 1:k)                      #Prior for Dirichlet
parameters-Independent Model
{
alp[k1]=0.5
}
a=sum(alp)
```

```

#####Prior for Dirichlet parameters-Association Model-Row
gam = 0
for (k2 in 1:r)
{
gam[k2]=0.5
}
g=sum(gam)
#####Prior for Dirichlet parameters-Association Model-Col
del = 0
for (k3 in 1:c)
{
del[k3]=0.5
}
d=sum(del)
#####Terms for Log Bayes Factor Formula
t1=ns+alp
t2=n+g
t3=n+d
t4=n+a
t5=rt+gam
t6=ct+del
#####Numerator and Denominator of Log Bayes Factor Formula
nr=sum(lgamma(t1))+lgamma(a)+sum(lgamma(gam))+sum(lgamma(
del))+lgamma(t2)+lgamma(t3)
dr=sum(lgamma(alp))+lgamma(g)+lgamma(d)+sum(lgamma(t4))+su
m(lgamma(t5))+sum(lgamma(t6))
lbf=nr-dr                                #Log Bayes Factor B10
lbf
lbf1=-lbf                                #Log Bayes Factor B01
lbf1
exp(lbf)                                  #Bayes Factor B10
exp(lbf1)                                  #Bayes Factor B01

```