Bayesian Ranking Responses in Multiple Response Questions

Hsiuying Wang
Institute of Statistics
National Chiao Tung University
Hsinchu, Taiwan
wang@stat.nctu.edu.tw

Abstract

Questionnaires are important tools for surveying in many studies. In previous studies, analyses of multiple response questions are not as established in depth as those for single response questions. Wang (2008a) proposed several methods for ranking responses in multiple response questions under the frequentist setup. However in many situations, prior information may exist for the ranks of the responses. Therefore, establishing a methodology combining updated survey data and past information for ranking responses is an essential issue in questionnaire data analysis. Based on several Bayesian multiple testing procedures, this study develops the Bayesian ranking methods by controlling the posterior expected false discovery rate. In addition, a simulation study is conducted to compare these approaches and to derive the appropriate rejection region for testing.

Key words: Dirichlet Prior, Bayes estimator, single response question, multiple response question, survey.

1 Introduction

Questionnaires are a widely-used tool for researchers in many fields to collect information, and they are especially important in marketing or management studies. There are two kinds of questions: single response questions and multiple response questions. Response models related to this problem are proposed in Thissen and Steinberg (1986,1984).

The analyses of multiple response questions are not as deeply established as those for single response questions, and approaches for analyzing multiple response questions have been lacking until recently. Umesh (1995) first discussed the problem of analyzing multiple response questions. Subsequently, Loughin and Scherer (1998), Decady and Thomas (1999) and Bilder, Loughin and Nettleton (2000) proposed several methods for testing marginal independence between a single response question and a multiple response question. Agresti and Liu (1999,2001) discussed the modeling of multiple response questions. These studies mainly focused on the analysis of the dependence between a single response question and a multiple response question. However, in practice, most researchers are also interested in ranking responses to a question according to the probabilities of responses being chosen. In fact, the ranking responses problem may be the primary issue in the analysis of a survey.

Here, we illustrate the problem by the example described in Wang (2008a), in which a company is designing a marketing survey to help develop an insect killer. The researchers list several factors, including high quality, price, packaging and smell which could affect the sales market and want to know the significance rank for these factors so they can design a product with lower cost and higher profit. To obtain the data, a group of individuals are surveyed about purchasing an insect killer by filling out questionnaires that include all the questions addressed to each respondent. The following is a multiple response question from the questionnaire:

Question 1: Which factors are important to you when considering the purchase of an indoor insect killer ? (1) price (2) high quality (3) packaging (4) smell (5) other.

Wang (2008a) proposed several approaches to solve this problem. However, these methodologies are derived under the frequestist setup, which cannot be adopted in a Bayesian framework. In real applications, empirical information may exist for the probabilities of responses being chosen. For related applications, readers can refer to Pammer, Fong and Arnold (2000), etc. When empirical information exists, an appropriate methodology that combines current data with prior information can provide a more objective ranking strategy than an approach based only on current data. Thus, this study proposes several methods for ranking the responses in a multiple response question under the Bayesian framework. The methodologies are an extension of the methods proposed in Muller, Parmigiani, Robert and Rousseau (2004). More details about Bayesian multiple testing and applications are discussed by Gopalan and Berry (1998), Do et al. (2005), Gonen, Westfall, Johnson (2003), Scott and Berger (2006), Muller, Parmigiani and Rice (2007) and Scott (2009). A related study about Bayesian ranking is Berger and Deely (2008), who rank items based on the posterior probability of the null hypothesis or the Bayes factor. Although this methodology provides a rule for ranking, it does not set up an error tolerance, and it cannot be directly applied to analyzing multiple response questions.

The conventional Bayesian multiple testing approach is to calculate the posterior probability, or Bayes factor, of the null hypothesis and then to reject or accept the null hypothesis based on the posterior probability or Bayes factor. The criterion for rejecting the null hypothesis is if the posterior probability or Bayes factor is greater than a critical value. The critical value selection in the conventional method is usually independent of the observations and sample size. When the sample size is large, the posterior probability can be adopted with more confidence to make the decision. When the sample size is small, a stricter threshold may need to be set to avoid large false discovery rate. Since conventional methods do not guide to select a critical value based on the observations, they cannot guarantee reaching an appropriate decision. In this study, we adopt the FDR procedure proposed in the literature for Bayesian multiple testing. The merits of the FDR associated with the loss function can provide a suitable criterion for the critical value selection (see Muller et al 2004). In addition, with the loss function, this procedure also can guarantee that the false discovery rate of the testing is less than a specified tolerance error that cannot be made by the conventional method.

In this study, we illustrate the use of the Bayesian model for analyzing multiple response questions and derive the exact and approximate Bayes estimator forms. The proposed method can provide a convenient means for researchers to directly adopt the formulas for ranking responses for multiple response questions.

In the multiple response question Question 1, there are a total of $2^5 - 1 = 31$ possible answers because we exclude the case in which respondents do not select any response. The 31 random variables constitute a multinomial distribution with multinomial proportions $p \in P = \{p_{i_1 i_2 i_3 i_4 i_5}, i_j = 0 \text{ or } 1 \text{ and } 0 < \sum_{j=1}^{5} i_j \leq 5\ \}$, where i_j cannot be simultaneously equal to 0. Note that the requirement of a multiple response question is that at least one response is selected. This is not equivalent to a true-false question with five items. If we allow respondents to select no item or to select all items, it would be equivalent to the five true-false items question. The method developed in this study can be extended to this situation.

We can also consider the parameter space under the frequentist framework instead of the Bayesian framework. Wang (2008a) provides examples showing that the conventional testing approaches do not have the property of ranking consistency. This property is a reasonable criterion to reflect the validity of the testing approach. Under the frequentist framework, it is still unknown if a satisfactory approach exists to rank responses with the property of ranking consistency. In this study, in addition to proposing a ranking approach under the Bayesian framework, a Bayesian ranking consistency property is introduced and the proposed method is shown to be Bayesian ranking consistent.

In the Bayesian framework, we assume that we have prior information on the parameter space P and we rank the responses based on a survey study under this prior information. This problem is related to the usual Bayesian multiple testing problem if we consider a single response question. However, the application is more complicated when analyzing multiple response questions. Muller et al. (2004) proposed several criteria for the Bayesian multiple testing. Miranda-Moreno, Labbe and Fu (2007) applied Muller et al. method to hotspot identification in an engineering study. Wang (2008b) carried out a related study estimating the proportions in a multinomial distribution. In this paper, we investigate these Bayesian multiple testing procedures and extend the approaches to rank the responses for multiple response questions.

References

[1] Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rates: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B, 57, 289-300.

[2] Agresti, A. and Liu, I.M. (1999) Modeling a categorical variable allowing arbitrarily many category choices. Biometrics 55, 936-943.

[3] Agresti, A. Liu, I.M. (2001) Strategies for modeling a categorical variable allowing multiple category choices. Sociological Methods and Research 29, 403-434.

[4] Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis (2nd ed.), New York: Springer-Verlag.

[5] Bickel, P.J., and Doksum, K.A. (2007), Mathematical Statistics: Basic Ideas and Selected Topics, Vol. 1., 2nd ed., Pearson Prentice Hall, Upper Saddle River, NJ.

[6] Bilder, C. R., Loughin, T. M.and Nettleton, D. (2000) Multiple marginal independence testing for pick any/c variables. Communications in Statistics-Simulation and Computation , 29, 1285-1316.

[7] Decady, Y. J. and Thomas, D. H. (2000). A simple test of association for contingency tables with multiple column responses. Biometrics 56, 893-896.

[8] Gopalan, R. and Berry, D. A. (1998). Bayesian multiple comparisons using Dirichlet process priors. Journal of the American Statistical Association 93, 1130V1139.

[9] Do, K., Muller, P. and Tang, F. (2005). A Bayesian mixture model for differential gene expression. Journal of the Royal Statistical Society, Series C, 54, 627-644.

[10] Pammer, S., Fong, D. K. H. and Arnold, S. F. (2000). Forecasting the Penetration of a New Product: A Bayesian Approach. Journal of Business and Economic Statistics, 18, no. 4, 428-435.

[11] Gonen, M., Westfall, P. H. and Johnson, W. O. (2003). Bayesian Multiple Testing for Two-Sample Multivariate Endpoints. Biometrics, 59, 76-82.

[12] Loughin, T. M. and Scherer, P. N. (1998). Testing for association in contingency tables with multiple column responses. Biometrics 54, 630-637.

[13] Muller P, Parmigiani G, and Rice K. (2007). "FDR and Bayesian decision rules." In Bayesian Statistics 8. (Bernardo, J. et al. ed.) Oxford University Press.

[14] Miranda-Moreno, L. F., Labbe, A. and Fu, L. (2007). Bayesian multiple testing procedures for hotspot identification. Accident Analysis and Prevention, 39, 1192-1201.

[15] Muller, P., Parmigiani, G., Robert, C. and Rousseau, J. (2004). Optimal sample size for multiple testing: The case of gene expression microarrays. Journal of the American Statistical Association, 99, no.468, 990-1001.

[16] Scott, J. (2009). "Nonparametric Bayesian multiple testing for longitudinal performance stratification." Annals of Applied Statistics.

[17] Scott, J.G. and Berger, J.O. (2006). An exploration of aspects of Bayesian multiple testing. J. Stat. Plann. Inference 136, no. 7, 2144-2162.

[18] Thissen, D. and Steinberg L. (1984). A response model for multiple choice items. Psycometrika 49, , 501- 519.

[19] Thissen, D. and Steinberg L. (1986). A taxonomy of item response models. Psycometrika 51, 567–577.

[20] Umesh, U. N. (1995). Predicting nominal variable relationships with multiple responses. Journal of Forecasting 14, 585-596.

[21] Wang, H. (2008a). Ranking responses in multiple response questions. Journal of Applied Statistics, 35, 465-474.

[22] Wang, H. (2008b) Exact confidence coefficients of simultaneous confidence intervals for multinomial proportions. Journal of Multivariate Analysis, 99, 896-911.