

Decomposition of Marginal Homogeneity into Logit and Mean Ridits Equality for Square Contingency Tables

Kiyotaka Iki, Kouji Tahata and Sadao Tomizawa
 Department of Information Sciences, Faculty of Science and Technology,
 Tokyo University of Science, Noda City, Chiba, 278-8510, Japan

Abstract: Problem statement: For square contingency tables with ordered categories, if the Marginal Homogeneity (MH) model holds, then the probability that X selected at random from the row marginal distribution is less than Y selected independently at random from the column marginal distribution equals the probability that X is greater than Y. However, the converse does not hold. We are interested in the condition in order that the converse holds. **Approach:** This study gave a decomposition theorem that the MH model holds if and only if both the marginal cumulative logistic model and the model of equality of mean ridits for the row and column marginal distributions hold. Examples are given. **Results:** For the data of cross-classification of ewes according to number of lambs born in consecutive years and cross-classification of unaided distance vision of women in British, the decomposition of the MH model is applied and the detailed analysis is given. **Conclusion:** When the MH model fits the data poorly, this decomposition is useful for seeing which of decomposed two models influences stronger.

Key words: Decomposition, marginal cumulative logistic model, marginal homogeneity, mean ridit, square contingency table

INTRODUCTION

Consider an $R \times R$ square contingency table with the same row and column ordinal classifications. Let X and Y denote the row and column variables, respectively and let $\Pr(X = i, Y = j) = p_{ij}$ ($i = 1, \dots, R; j = 1, \dots, R$). The Marginal Homogeneity (MH) model is defined by (Stuart, 1955; Bhapkar, 1966; Bishop *et al.*, 1975; Caussinus, 1965):

$$p_{i.} = p_{.i} \quad (i = 1, \dots, R)$$

Where:

$$p_{i.} = \sum_{t=1}^R p_{it}$$

$$p_{.i} = \sum_{s=1}^R p_{si}$$

This indicates that the row marginal distribution is identical to the column marginal distribution. The MH model also may be expressed as:

$$F_i^X = F_i^Y \quad (i = 1, \dots, R - 1)$$

Where:

$$F_i^X = \sum_{k=1}^i p_{k.}$$

$$F_i^Y = \sum_{k=1}^i p_{.k}$$

Let L_i^X and L_i^Y denote the marginal cumulative logit of X and Y, respectively. These are given as:

$$L_i^X = \log\left(\frac{F_i^X}{1 - F_i^X}\right)$$

and

$$L_i^Y = \log\left(\frac{F_i^Y}{1 - F_i^Y}\right)$$

for $i = 1, \dots, R - 1$.

The MH model may be further expressed as:

$$L_i^X = L_i^Y \quad (i = 1, \dots, R - 1)$$

As an extension of the MH model, the Marginal cumulative Logistic (ML) model is defined by McCullagh (1977) and Agresti (1984):

Corresponding Author: Kouji Tahata, Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan

$$L_i^X = L_i^Y + \Delta \quad (i=1, \dots, R-1)$$

This model states that the odds that X is i or below instead of i+1 or above, is $\exp(\Delta)$ times higher than the odds that Y is i or below instead of i+1 or above, for $i=1, \dots, R-1$. If $\Delta > 0$, X rather than Y tends to be i or below instead of i+1 or above, for $i=1, \dots, R-1$. A special case of this model obtained by putting $\Delta = 0$ is the MH model.

Miyamoto *et al.* (2005) and Tahata *et al.* (2007) gave the theorem that the MH model holds if and only if both the ML model and the marginal mean equality model (i.e., $E(X) = E(Y)$) hold (Tahata *et al.*, 2008).

On the other hand, Agresti (1983; 1984) considered the comparison between the marginal distributions using the measure defined by:

$$\tau = \sum_{i < j} \sum p_i \cdot p_j - \sum_{i > j} \sum p_i \cdot p_j = \Pr(X < Y) - \Pr(X > Y)$$

Where:

X = Selected at random from the row marginal distribution

Y = Selected independently at random from the column marginal distribution

This is a population value of the difference between discrete analogs of the Mann-Whitney statistics. This measure is positive when X is stochastically less than Y and negative when X is stochastically greater than Y. Note that τ can be expressed using the mean ridits; as described later.

We note that the MH model implies the structure of $\tau = 0$, thus, the MH model is not equivalent to $\tau = 0$. We are now interested in what structure is necessary to obtain the MH model in addition to the structure of $\tau = 0$.

The purpose of this study is to give a theorem that the MH model holds if and only if both the ML model and the structure of $\tau = 0$ hold.

MATERIALS AND METHODS

Let:

$$r_i^X = \frac{F_{i-1}^X + F_i^X}{2}$$

and

$$r_i^Y = \frac{F_{i-1}^Y + F_i^Y}{2}$$

for, $i=1, \dots, R$, where $F_0^X = F_0^Y = 0$, $F_R^X = F_R^Y = 1$. The $\{r_i^X\}$ and $\{r_i^Y\}$ are the marginal ridits; Bross (1958). The mean ridit for the distribution of Y when the distribution of X is the identified distribution for calculating the ridits is:

$$R_X(Y) = \sum_{j=1}^R r_j^X p_j$$

Similarly, we have:

$$R_Y(X) = \sum_{i=1}^R r_i^Y p_i$$

and also $R_X(X) = R_Y(Y) = 0.5$, where:

$$R_X(X) = \sum_{i=1}^R r_i^X p_i, \quad R_Y(Y) = \sum_{j=1}^R r_j^Y p_j.$$

Then, τ is expressed as Agresti (1984):

$$\tau = R_X(Y) - R_Y(X)$$

We shall refer to the structure of $\tau = 0$ as the marginal Mean Ridits equality (MR) model.

We see:

$$\begin{aligned} F_i^X - F_i^Y &= \sum_{s=1}^i \sum_{t=1}^R p_s \cdot p_t - \sum_{s=1}^i \sum_{t=1}^R p_s \cdot p_t \\ &= \sum_{s=1}^i \sum_{t=i+1}^R p_s \cdot p_t - \sum_{s=i+1}^R \sum_{t=1}^i p_s \cdot p_t \end{aligned}$$

for $i=1, \dots, R-1$. Therefore, the MH model may be expressed as:

$$H_{1(i)} = H_{2(i)} \quad (i=1, \dots, R-1)$$

Where:

$$H_{1(i)} = \sum_{s=1}^i \sum_{t=i+1}^R p_s \cdot p_t = F_i^X (1 - F_i^Y)$$

$$H_{2(i)} = \sum_{s=i+1}^R \sum_{t=1}^i p_s \cdot p_t = (1 - F_i^X) F_i^Y$$

This indicates that the probability that the row variable X selected at random from the row marginal distribution is in category i or below and the column variable Y selected independently at random from the column marginal distribution is in category i+1 or above is equal to the probability that such X is in

category $i+1$ or above and such Y is in category i or below. Using $\{H_{1(i)}\}$ and $\{H_{2(i)}\}$, the ML model may also be expressed as:

$$H_{1(i)} = \theta H_{2(i)} \quad (i = 1, \dots, R-1)$$

This model with $\theta = 1$ is the MH model. We obtain the following theorem.

Theorem 1: The MH model holds if and only if both the ML and MR models hold.

Proof: If the MH model holds, then the ML and MR models hold. Assuming that both the ML and MR models hold, then we shall show that the MH model holds.

Since the ML model holds, we have:

$$F_i^X - F_i^Y = H_{1(i)} - H_{2(i)} = (\theta - 1)H_{2(i)} \quad (i = 1, \dots, R-1)$$

If $\theta = 1$, we see that the MH model holds. If $\theta > 1$, we have:

$$F_i^X > F_i^Y \quad (i = 1, \dots, R-1)$$

Noting that $R_X(Y) + R_Y(X) = 1$, if $\theta > 1$, we have:

$$\begin{aligned} \tau &= 2R_X(Y) - 1 \\ &= 2 \sum_{j=1}^R \left(\frac{F_{j-1}^X + F_j^X}{2} \right) p_{.j} - 1 \\ &> 2 \sum_{j=1}^R \left(\frac{F_{j-1}^Y + F_j^Y}{2} \right) p_{.j} - 1 \\ &= 2R_Y(Y) - 1 \end{aligned}$$

Since $R_Y(Y) = 0.5$, we have $\tau > 0$. Similarly, if $\theta < 1$, we have:

$$F_i^X < F_i^Y \quad (i = 1, \dots, R-1)$$

Thus, if $\theta < 1$, we have $\tau < 0$. Since the MR model holds, i.e., $\tau = 0$, we obtain $\theta = 1$. Namely, the MH model holds. The proof is completed.

Assume that a multinomial distribution is applied to the $R \times R$ table. The maximum likelihood estimates of expected frequencies under the MH, ML and MR models could be obtained using the Newton-Raphson method in the log-likelihood equation. The numbers of degrees of freedom for testing the goodness-of-fit of the MH, ML and MR models are $R-1$, $R-2$ and 1 , respectively.

RESULTS

Example 1: The data in Table 1, are taken directly from Bishop *et al.* (1975). These data indicate that ewes are cross-classified by the number of lambs born to them in two consecutive years, 1952 and 1953.

The MH model fits these data poorly, yielding the likelihood ratio chi-squared statistic $G^2 = 18.65$ with 2 degrees of freedom. Also, the ML model fits these data poorly, yielding $G^2 = 18.55$ with 1 degree of freedom. However, the MR model fits these data well, yielding $G^2 = 0.78$ with 1 degree of freedom.

From Theorem 1, we see that the poor fit of the MH model is caused by the influence of the lack of structure of the ML model rather than the MR model.

Example 2: The data in Table 2, taken from Stuart (1955), are constructed from unaided distance vision of 7477 women aged 30-39 employed in royal ordnance factories in Britain from 1943-1946. These data have been analyzed by many statisticians, for example, including Stuart (1955), Caussin (1965), Bishop *et al.* (1975), Agresti (1984).

The MH model fits these data poorly, yielding the likelihood ratio chi-squared statistic $G^2 = 11.99$ with 3 degrees of freedom. Also, the MR model fits these data poorly, yielding $G^2 = 11.94$ with 1 degree of freedom. However, the ML model fits these data well, yielding $G^2 = 0.39$ with 2 degrees of freedom.

We see from Theorem 1 that the poor fit of the MH model is caused by the influence of the lack of structure of the MR model rather than the ML model.

Table 1: Cross-classification of ewes according to number of lambs born in consecutive years; from Bishop *et al.* (1975)

Number of lambs 1953	Number of lambs 1952			Total
	0	1	2	
0	58 (57.73)	52 (47.94)	1 (0.88)	111
1	26 (28.35)	58 (58.17)	3 (2.87)	87
2	8 (9.27)	12 (12.72)	9 (9.07)	29
Total	92	122	13	227

Note: The parenthesized values are the maximum likelihood estimates of expected frequencies under the MR model

Table 2: Unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943-1946; from Stuart (1955)

Right eye grade	Left eye grade				Total
	Best (1)	Second (2)	Third (3)	Worst (4)	
Best (1)	1520	266	124	66	1976
Second (2)	234	1512	432	78	2256
Third (3)	117	362	1772	205	2456
Worst (4)	36	82	179	492	789
Total	1907	2222	2507	841	7477

DISCUSSION

In Example 1, we have seen that the MR model fits well, however, the MH model fits poorly. The readers may be interested in why this situation arises. We point out that testing the MR model is not equivalent to testing the MH model. From Theorem 1, the structure of uniformity of ratio $\{H_{1(i)} / H_{2(i)}\}$ is necessary to obtain the MH model in addition to the MR model. Under the MR model, the maximum likelihood estimates of $\{H_{1(i)}\}$ and $\{H_{2(i)}\}$ are $\hat{H}_{1(1)} = 0.272$, $\hat{H}_{1(2)} = 0.049$, $\hat{H}_{2(1)} = 0.223$ and $\hat{H}_{2(2)} = 0.129$ and thus $\hat{H}_{1(1)} / \hat{H}_{2(1)} = 1.221$ and $\hat{H}_{1(2)} / \hat{H}_{2(2)} = 0.378$; so these indicate that there is not the structure of uniformity of ratio $\{H_{1(i)} / H_{2(i)}\}$ in these data. Therefore it would be natural that the MH model does not fit the data in Table 1 well.

CONCLUSION

This study gave the theorem that the MH model holds if and only if both the ML and MR models hold. When the MH model fits the data poorly, this theorem would be useful for seeing which of the lack of the MR model and the lack of the ML model influences stronger. Indeed, for the data in Table 1, the poor fit of the MH model is caused by the poor fit of the ML model rather than the MR model (Example 1). For the data in Table 2, the poor fit of the MH model is caused by the poor fit of the MR model rather than the ML model (Example 2).

REFERENCES

- Agresti, A., 1983. Testing marginal homogeneity for ordinal categorical variables. *Biometrics*, 39: 505-510. <http://www.jstor.org/stable/2531022>
- Agresti, A., 1984. *Analysis of Ordinal Categorical Data*. 1st Edn., Wiley, New York, ISBN: 978-0-471-89055-3, pp: 304.
- Bhapkar, V.P., 1966. A note on the equivalence of two test criteria for hypotheses in categorical data. *J. Am. Stat. Assoc.*, 61: 228-235. <http://www.jstor.org/stable/2283057>
- Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1975. *Discrete Multivariate Analysis: Theory and Practice*. 1st Edn., The MIT Press, Cambridge, ISBN: 13: 978-0-262-52040-9, pp: 568.
- Bross, I.D.J., 1958. How to use ridity analysis. *Biometrics*, 14: 18-38. <http://www.jstor.org/stable/2527727>
- Caussinus, H., 1965. Contribution to the statistical analysis of correlation tables. *Ann. Fac. Sci. Univ. Toulouse*, 29: 77-182. http://www.numdam.org/item?id=AFST_1965_4_29_77_0
- McCullagh, P., 1977. A logistic model for paired comparisons with ordered categorical data. *Biometrika*, 64: 449-453. <http://www.jstor.org/stable/2345320>
- Miyamoto, N., K. Niibe and S. Tomizawa, 2005. Decompositions of marginal homogeneity model using cumulative logistic models for square contingency tables with ordered categories. *Aus. J. Stat.*, 34: 361-373. <http://www.stat.tugraz.at/AJS/ausg054/054Tomizawa.pdf>
- Stuart, A., 1955. A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, 42: 412-416. <http://www.jstor.org/stable/2333387>
- Tahata, K., H. Kobayashi and S. Tomizawa, 2008. Conditional marginal cumulative logistic models and decomposition of marginal homogeneity model for multi-way tables. *J. Stat. Applied*, 3: 239-252. <http://www.mdpppl.com/journals/journal-of-statistics-and-applications/content.html?issue=108>
- Tahata, K., S. Katakura and S. Tomizawa, 2007. Decompositions of marginal homogeneity model using cumulative logistic models for multi-way contingency tables. *Revstat: Stat. J.*, 5: 163-176. <http://www.ine.pt/revstat/pdf/rs070202.pdf>