SUT Journal of Mathematics
Vol. 49, No. 2 (2013), 121–128

Sum-symmetry model and its orthogonal decomposition for square contingency tables with ordered categories

Kouji Yamamoto, Yayoi Tanaka and Sadao Tomizawa

(Received December 10, 2013)

Abstract. For the analysis of square contingency tables, many kinds of symmetry models have proposed. The present paper proposes a new kind of symmetry model, and gives a decomposition of the new model by introducing an extended model of it. Moreover, it shows the orthogonality of statistic for testing goodness-of-fit of the new model. Two unaided vision data analyses are also shown.

AMS 2010 *Mathematics Subject Classification.* 62H17.

Key words and phrases. Decomposition, orthogonality, square contingency table, unaided vision data.

§1. Introduction

For analyzing the data of a square contingency table such as Tables 1 and 2, one usually uses the well-known symmetry (S) model (Bowker [1]) and its many extended models, for example, the conditional symmetry (CS) model (McCullagh [2]) and so on. Table 1, taken directly from Tomizawa [5], is the data of the unaided distance vision of 3168 pupils aged 6-12 including about half girls at elementary schools in Tokyo, Japan, examined in June 1984. Table 2, taken directly from Stuart [4], is the data of the unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946. In Tables 1 and 2 the row variable is the right eye grade and the column variable is the left eye grade with the categories ordered from the lowest grade (1) to the highest grade (4).

Table 1: Unaided distance vision of 3168 pupils aged 6-12, including about half the girls at elementary schools in Tokyo, Japan examined in June 1984; from Tomizawa [5]. (The parenthesized values are the maximum likelihood estimates of expected frequencies under the SS model.)

	Left eye grade				
	Lowest	Second	Third	Highest	
Right eye grade	(1)	(2)	(3)	(4)	Total
Lowest (1)	92	16	7	12	127
	(92.00)	(15.50)	(6.00)	(10.78)	
Second (2)	15	75	42	10	142
	(15.50)	(75.00)	(37.72)	(15.50)	
Third (3)	5	33	138	96	272
	(6.00)	(37.22)	(138.00)	(111.00)	
Highest (4)	10	21	126	2470	2627
	(11.28)	(15.50)	(111.00)	(2470.00)	
Total	122	145	313	2588	3168

Table 2: Unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories from 1943 to 1946; from Stuart [4]. (The parenthesized values are the maximum likelihood estimates of expected frequencies under the CSS model.)

	Left eye grade				
	Lowest	Second	Third	Highest	
Right eye grade	(1)	(2)	(3)	(4)	Total
Lowest (1)	492	179	82	36	789
	(492.00)	(177.83)	(74.09)	(37.53)	
Second (2)	205	1772	362	117	2456
	(206.17)	(1772.00)	(377.40)	(111.60)	
Third (3)	78	432	1512	234	2256
	(85.91)	(417.31)	(1512.00)	(231.55)	
Highest (4)	66	124	266	1520	1976
	(63.76)	(129.40)	(268.45)	(1520.00)	
Total	841	2507	2222	1907	7477

It seems natural to see the degree of an individual's eye grade as the sum of the grades of both of right and left eyes. Thus, we define the degree of an individual's eye grade as the sum of his/her right and left eye grades throughout this paper.

In this paper, thus, we propose a new kind of symmetry model to represent such a structure, and give a decomposition of the new model by introducing an extended model of it. In addition, we show the orthogonality of the test statistics for decomposed models.

§2. New models

Consider a square $r \times r$ contingency table with row variable X, and column variable Y. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, \dots, r; j = 1, \dots, r$).

Consider a model defined by

$$\Pr(X + Y = t, X < Y) = \Pr(X + Y = t, X > Y) \quad (3 \le t \le 2r - 1).$$

We shall refer to this model as the sum-symmetry (SS) model. The SS model is also expressed as

$$\sum\sum_{(i,j)\in R(t)} p_{ij} = \sum\sum_{(i,j)\in R(t)} p_{ji} \quad (3 \le t \le 2r - 1),$$

where

$$R(t) = \{(i, j) \mid i + j = t, i < j\}.$$

We note that when $r = 3$, the SS model is equivalent to the S model.

If the SS model holds for the vision data like Tables 1 and 2, the probability that the degree of the eye grade for an individual whose left eye grade is greater than his/her right eye grade, is t ($3 \le t \le 2r - 1$), is equal to the probability that the degree of the eye grade for the individual whose right eye grade is greater than his/her left eye grade, is t.

Consider an extension of the SS model as follows:

$$\Pr(X + Y = t, X < Y) = \Delta\Pr(X + Y = t, X > Y) \quad (3 \le t \le 2r - 1),$$

where the parameter Δ is unspecified. Then, this model may be expressed as

$$\Pr(X + Y = t | X < Y) = \Pr(X + Y = t | X > Y) \quad (3 \le t \le 2r - 1).$$

So we shall refer to this model as the conditional SS (CSS) model. The CSS model is also expressed as

$$\sum\sum_{(i,j)\in R(t)} p_{ij} = \Delta \sum\sum_{(i,j)\in R(t)} p_{ji} \quad (3 \le t \le 2r - 1).$$

A special case of the CSS model obtained by putting $\Delta = 1$ is the SS model. Also we note that when $r = 3$, the CSS model is equivalent to the CS model.

If the CSS model holds for the vision data, the probability that the degree of the eye grade for an individual whose left eye grade is greater than his/her right eye grade, is t ($3 \le t \le 2r-1$), is Δ times higher than the probability that the degree of the eye grade for the individual whose right eye grade is greater than his/her left eye grade, is t.

§3. Decomposition of the SS model

Read's [3] global symmetry (GS) model is defined by

$$\Pr(X < Y) = \Pr(X > Y).$$

This model is also expressed as

$$\sum\sum_{i<j} p_{ij} = \sum\sum_{i<j} p_{ji}.$$

Then we obtain the following theorem:

Theorem 1. *The SS model holds if and only if both the CSS and GS models hold.*

Proof. If the SS model holds, then the CSS and GS models hold. Assuming that both the CSS and GS models hold, then we shall show that the SS model holds. From the assumption that the CSS model holds, we have

$$\sum_{t=3}^{2r-1} \Pr(X+Y=t, X<Y) = \Delta \sum_{t=3}^{2r-1} \Pr(X+Y=t, X>Y),$$

thus, we see

$$\Pr(X < Y) = \Delta \Pr(X > Y).$$

Since the GS model holds, we obtain $\Delta = 1$. Namely the SS model holds, so the proof is completed. □

§4. Goodness-of-fit test

Let n_{ij} denote the observed frequency in the (i,j)th cell of the table ($i = 1,\dots,r; j = 1,\dots,r$). Assume that a multinomial distribution is applied to the $r \times r$ table. The maximum likelihood estimates (MLEs) of expected frequencies $\{m_{ij}\}$ under the SS, CSS and GS models, are expressed as the closed-forms as follows:

(a) The MLE of m_{ij} under the SS model is

$$\hat{m}_{ij} = \begin{cases} \dfrac{U(t)+L(t)}{2U(t)} n_{ij} & (i+j=t, i<j) \\ \dfrac{n_{ij}}{N} & (i=j) \\ \dfrac{U(t)+L(t)}{2L(t)} n_{ij} & (i+j=t, i>j), \end{cases}$$

where

$$U(t) = \sum\sum_{\substack{i+j=t \\ i<j}} n_{ij}, \quad L(t) = \sum\sum_{\substack{i+j=t \\ i<j}} n_{ji} \quad (3 \leq t \leq 2r-1).$$

(b) The MLE of m_{ij} under the CSS model is

$$\hat{m}_{ij} = \begin{cases} \dfrac{U}{U+L} \cdot \dfrac{U(t)+L(t)}{U(t)} n_{ij} & (i+j=t, i<j) \\ \dfrac{n_{ij}}{N} & (i=j) \\ \dfrac{L}{U+L} \cdot \dfrac{U(t)+L(t)}{L(t)} n_{ij} & (i+j=t, i>j), \end{cases}$$

where

$$U = \sum\sum_{i<j} n_{ij}, \quad L = \sum\sum_{i<j} n_{ji}.$$

(c) The MLE of m_{ij} under the GS model is

$$\hat{m}_{ij} = \begin{cases} \dfrac{U+L}{2U} n_{ij} & (i<j) \\ \dfrac{n_{ij}}{N} & (i=j) \\ \dfrac{U+L}{2L} n_{ij} & (i>j). \end{cases}$$

Each model can be tested for goodness-of-fit by, e.g., the likelihood ratio chi-squared statistic with the corresponding degrees of freedom (df). The likelihood ratio statistic for testing goodness-of-fit of model Ω is given by

$$G^2(\Omega) = 2\sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij} \log\left(\frac{n_{ij}}{\hat{m}_{ij}}\right),$$

where $\hat{m}_{ij}$ is the MLE of expected frequency m_{ij} under model Ω. The numbers of df for the SS, CSS and GS models are $2r-3$, $2(r-2)$, and 1, respectively.

§5. Orthogonality of test statistics

Theorem 2. *The test statistic $G^2(SS)$ is equal to the sum of $G^2(CSS)$ and $G^2(GS)$.*

Proof. From Section 4, we see that the $(n_{ij}/\hat{m}_{ij})$ under the SS model is equal to the product of the $(n_{ij}/\hat{m}_{ij})$ under the CSS model and that under the GS model. Therefore, the proof is completed. □

§6. Examples

6.1. Example 1

Consider the data in Table 1. From Table 3 we see that the SS model fits these data well. Also the CSS and GS models fit these data well. According to the test (at the 0.05 level) based on the difference between the likelihood ratio chi-squared values of two nested models, the SS model may be preferable to the other models.

Under the SS model, the probability that the degree of the eye grade for a pupil whose left eye grade is greater than his/her right eye grade, is t ($3 \leq t \leq 2r-1$), is estimated to be equal to the probability that the degree of the eye grade for the pupil whose right eye grade is greater than his/her left eye grade, is t.

Table 3: Likelihood ratio statistic G^2 for models applied to the data in Tables 1 and 2.

Applied	For Table 1			For Table 2		
Model	df	G^2	p-value	df	G^2	p-value
SS	5	9.673	0.085	5	15.299*	0.009
CSS	4	7.817	0.099	4	3.403	0.493
GS	1	1.856	0.173	1	11.896*	< 0.001

*means significant at 5% level

6.2. Example 2

Consider the data in Table 2. The SS model fits the data in Table 2 poorly. Also, the GS model fits these data poorly, however, the CSS model fits these data well. Therefore, it is seen from Theorem 1 that the poor fit of the SS model is caused by the influence of the lack of structure of the GS model rather than the CSS model.

Under the CSS model, the MLE of Δ is 0.86. Thus, under the CSS model, the probability that the degree of the eye grade for a woman whose right eye grade is greater than her left eye grade, is t $(3 \leq t \leq 2r-1)$, is estimated to be 1.16 $(= 0.86^{-1})$ times higher than the probability that the degree of the eye grade for the woman whose left eye grade is greater than her right eye grade, is t. Namely, a woman's right eye is estimated to be better than her left eye.

§7. Concluding Remarks

In this paper we have proposed the decomposition for the SS model into the CSS and GS models. This decomposition (i.e., Theorem 1) may be useful for seeing a reason for the poor fit of the SS model.

In addition, we point out that the likelihood ratio chi-squared statistic for testing goodness-of-fit of the SS model assuming that the CSS model holds true, is $G^2(SS) - G^2(CSS)$ and this is equal to the likelihood ratio chi-squared statistic for testing goodness-of-fit of the GS model, i.e., $G^2(GS)$ (from Theorem 2). We observe that for the data in Tables 1 and 2 the value of $G^2(SS)$ is exactly equal to the sum of the values $G^2(CSS)$ and $G^2(GS)$.

References

[1] A. H. Bowker, *A test for symmetry in contingency tables*, Journal of the American Statistical Association **43** (1948), 572–574.

[2] P. McCullagh, *A class of parametric models for the analysis of square contingency tables with ordered categories*, Biometrika **65** (1978), 413–418.

[3] C. B. Read, *Partitioning chi-square in contingency tables: a teaching approach*, Communications in Statistics-Theory and Methods **6** (1977), 553–562.

[4] A. Stuart, *A test for homogeneity of the marginal distributions in a two-way classification*, Biometrika **42** (1955), 412–416.

[5] S. Tomizawa, *Analysis of data in square contingency tables with ordered categories using the conditional symmetry model and its decomposed models*, Environmental Health Perspectives **63** (1985), 235–239.

Kouji Yamamoto
Department of Medical Innovation, Osaka University Hospital
2-15, Yamadaoka, Suita, Osaka, 565-0871, Japan
E-mail: yamamoto-k@hp-crc.med.osaka-u.ac.jp

Yayoi Tanaka
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
2641, Yamazaki, Noda, Chiba, 278-8510, Japan

Sadao Tomizawa
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
2641, Yamazaki, Noda, Chiba, 278-8510, Japan
E-mail: tomizawa@is.noda.tus.ac.jp