

Biometrical Letters Vol. 58 (2021), No. 1, 59-68

# An anti-sum-symmetry model and its orthogonal decomposition for ordinal square contingency tables with an application to grip strength test data

# Shuji Ando

Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Katsushika-ku, Tokyo, 125-8585, Japan, e-mail: shuji.ando@rs.tus.ac.jp

## SUMMARY

For the analysis of  $R \times R$  square contingency tables, we need to estimate an unknown probability distribution with high confidence from obtained observations. For that purpose, we need to perform the analysis using a statistical model that fits the data well and has a simple interpretation. This study proposes two original models that have symmetric and asymmetric structures between the probability with which the sum of row and column variables is t, for t = 2, ..., R, and the probability with which the sum of row and column variables is 2(R + 1) - t. The study also reveals that it is necessary to satisfy the anti-global symmetry model, in addition to the proposed asymmetry model, in order to satisfy the proposed symmetry model. This decomposition theorem is useful to explain why the proposed symmetry model does not hold. Moreover, we show that the value of the likelihood ratio chi-squared statistic of the proposed symmetry model is equal to the sum of those of the decomposed models. We evaluate the utility of the proposed models by applying them to real-world grip strength data.

**Key words:** anti-diagonal cell, asymmetry, necessary and sufficient condition, symmetry, test statistic

## 1. Introduction

This study deals with  $R \times R$  square contingency tables with the same row and column ordinal classifications. Square contingency tables are usually obtained by cross-classifying for matched-pairs data of the ordinal categorical variable. For such data, we are generally interested in whether the probability that observations will fall in the (i, j)th cell, for i < j, of the table is equal to the probability that observations will fall in the (j, i)th cell of the table. In other words, we are interested in whether or not there exists symmetry of cell probability with respect to the main-diagonal cells of the table. The symmetry (S) model (Bowker, 1948) is useful for analyzing whether or not the above structure holds. Since the S model has severe restrictions, various models with weaker restrictions have been proposed, for example, the conditional symmetry (CS) model (McCullagh, 1978), the global symmetry (GS) model (Read, 1977), and so on. Read (1977) showed that (i) it is necessary to satisfy the GS model, in addition to the CS model, in order to satisfy the S model, and (ii) the value of the likelihood ratio chi-squared statistic of the S model is equal to the sum of those of the CS and GS models.

On the other hand, we are occasionally interested in whether (i) the probability that observations will fall in the (i, j)th cell, for i + j < R + 1, of the table is equal to the probability that observations will fall in the (R+1-i, R+1-j)th cell of the table, and (ii) the probability that observations will fall in the (i, j)th cell, for i+j < R+1, of the table is equal to the probability that observations will fall in the (R+1-j, R+1-i)th cell of the table. In other words, we are interested in whether or not there exists point-symmetry (i.e., the structure (i)) and symmetry (i.e., the structure (ii)) of cell probability with respect to the anti-diagonal cells of the table. The anti-point-symmetry (AP) (Tomizawa, 1985) and anti-symmetry (AS) (Iki, 2016) models are useful for analyzing whether the above structures (i) and (ii), respectively, hold or not. Since the AP model (or AS model) has severe restrictions, various models with weaker restrictions have been proposed, for example, the anti-conditional point-symmetry (ACP) model (or the anti- conditional symmetry (ACS) model) (Tomizawa, 1986), and the anti-global symmetry (AGS) model (Iki, 2016; Kurakami et al., 2017). Kurakami et al. (2017) showed that (i) it is necessary to satisfy the AGS model, in addition to the ACP model, in order to satisfy the AP model, and (ii) the value of the likelihood ratio chi-squared statistic of the AP model is equal to the sum of those of the ACP and AGS models.

Consider the data in Table 1. Table 1 is the data from a grip strength test on women aged 15–69; source: National Health and Nutrition Examination Survey 2011–2012 (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/MGX\_G.htm). The row variable X is the right hand grip strength level and the column variable Y is the left hand grip strength level, with the categories ordered from the highest level (1) to the lowest level (5). These

levels are categorized based on the Muscle Strength Procedure Manual of the National Health and Nutrition Examination Survey.

(https://wwwn.cdc.gov/Nchs/Nhanes/2011-									
$2012/\mathrm{MGX}\mathrm{_G.htm}).$									
	Left hand								
Right hand	(1)	(2)	(3)	(4)	(5)	Total			
Excellent $(1)$	380	229	49	17	2	677			
Very good $(2)$	34	204	198	80	23	539			
Good $(3)$	5	55	159	161	74	454			
Fair $(4)$	1	11	37	106	168	323			
Poor $(5)$	0	2	12	45	256	315			
Total	420	501	455	409	523	2308			

 Table 1. The table below is the data from a grip strength test on women aged 15–69; source: National Health and

Nutrition Examination Survey 2011–2012

Yamamoto et al. (2016) pointed out that it is natural to evaluate an individual's grip strength level as the sum of the levels of both right and left hands from these data. The sum of the levels of both right and left hands (i.e., X + Y) is categorized ordered from the highest level (2) to the lowest level (10).

Yamamoto et al. (2013) proposed the sum-symmetry (SS) and conditional sum-symmetry (CSS) models, which have symmetric and asymmetric structures between the probability with which X + Y is t under X < Y, for t = 2, ..., 2R - 1, and the probability with which X + Y is t under X > Y. The SS model represents a symmetric structure for the main diagonal cells of table. Moreover, Yamamoto et al. (2013) showed that (i) it is necessary to satisfy the GS model, in addition to the CSS model, in order to satisfy the SS model, and (ii) the value of the likelihood ratio chi-squared statistic of the SS model is equal to the sum of those of the CSS and GS models.

We point out that it may be more natural to evaluate whether or not symmetric and asymmetric structures of X + Y exist for the anti-diagonal cells of the table, rather than for the main diagonal cells. This is because, for grip strength data such as Table 1, more people are right-handed, and the grip strength of the dominant hand is usually higher than that of the non-dominant hand. Therefore, we are interested in whether the individual's grip strength level (i.e., X + Y) is symmetric for the midpoint R + 1 in the range [2, 2R] of X + Y, rather than whether X + Y is symmetric between X < Y and X > Y. If there is no clear reason to set a particular value as the reference value of X + Y, it is natural to set the midpoint as the reference value. In fact, the grip strength data in Table 1 are categorized such that the midpoint R + 1 in the range [2, 2R] of X + Y is the reference value.

This study proposes two original models that have symmetric and asymmetric structures between the probability with which X + Y is t, for  $t = 2, \ldots, R$ , and the probability with which X + Y is 2(R+1) - t. This study also reveals that it is necessary to satisfy the AGS model, in addition to the proposed asymmetry model, in order to satisfy the proposed symmetry model. Moreover, we show that the value of the likelihood ratio chi-squared statistic of the proposed symmetry model is equal to the sum of those of the proposed asymmetry and AGS models.

The remainder of this paper is organized as follows. Section 2 introduces the two original models in square contingency tables. Section 3 gives the orthogonal decomposition of the proposed symmetry model using the proposed asymmetry model. Section 4 evaluates the utility of the proposed models by applying them to real-world grip strength data. Section 5 closes with concluding remarks.

#### 2. Anti-sum-symmetry model

In this section, we propose two original models that have symmetric and asymmetric structures between the probability with which X + Y is t, for t = 2, ..., R, and the probability with which X + Y is 2(R + 1) - t. First, we propose the anti-sum-symmetry (ASS) model. The ASS model is defined by

$$\Pr(X+Y = t, X+Y < R+1) = \Pr(X+Y = 2(R+1) - t, X+Y > R+1),$$

for t = 2, ..., R. For grip strength data such as Table 1, when the ASS model holds, the degree of an individual's grip strength is symmetric with respect to the midpoint R + 1 in the range [2, 2R] of X + Y. Thus, we can conclude that the median of the degree of an individual's grip strength is equivalent to the midpoint R + 1. The number of degrees of freedom (df) for testing the goodness-of-fit of the ASS model is R - 1.

Next, we propose the anti-conditional sum-symmetry (ACSS) model.

The ACSS model is defined by

$$\Pr(X+Y = t, X+Y < R+1) = \Delta \Pr(X+Y = 2(R+1) - t, X+Y > R+1),$$

for t = 2, ..., R. The ACSS model with  $\Delta = 1$  is equivalent to the ASS model. For grip strength data such as Table 1, when the ACSS model holds, the degree of an individual's grip strength is asymmetric with respect to the midpoint R + 1. Thus, we can conclude that (i) when  $\Delta > 1$ , the median of the degree of an individual's grip strength is smaller than the midpoint R + 1, and (ii) when  $\Delta < 1$ , it is larger than the midpoint R + 1. The number of df for testing the goodness-of-fit of the ACSS model is R - 2.

Interpretations such as those of the ASS and ACSS models cannot be obtained from the SS model or the CSS model. Thus, the ASS and ACSS models are useful for application to real-world data such as grip strength data.

#### 3. Orthogonal decomposition of the anti-sum-symmetry model

In this section, we show that (i) it is necessary to satisfy the AGS model, in addition to the ACSS model, in order to satisfy the ASS model; (ii) the value of the likelihood ratio chi-squared statistic of the ASS model is equal to the sum of those of the ACSS and AGS models.

The AGS model (Iki, 2016; Kurakami et al., 2017) is defined by

$$\Pr(X + Y < R + 1) = \Pr(X + Y > R + 1).$$

We see that the AGS model has weaker restrictions than the ASS model from the following equalities:

$$Pr(X + Y < R + 1) = \sum_{t=2}^{R} Pr(X + Y = t, X + Y < R + 1) \text{ and}$$
$$Pr(X + Y > R + 1) = \sum_{t=2}^{R} Pr(X + Y = 2(R + 1) - t, X + Y > R + 1).$$

Let U(M) denote that the model M holds. We obtain the following decomposition theorem. Note that the number of df for the ASS model is equal to the sum of those for the ACSS and AGS models.

**Theorem 3.1.** The following necessary and sufficient condition holds:

 $U(ASS) \Leftrightarrow U(ACSS) \land U(AGS).$ 

*Proof.* It is clear that the necessary condition  $U(ASS) \Rightarrow U(ACSS) \land U(AGS)$  holds. We need to show that the sufficient condition  $U(ASS) \leftarrow U(ACSS) \land U(AGS)$  also holds. Since the ACSS model holds, the following equality holds:

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

Since the AGS model holds, we obtain  $\Delta = 1$ . The proof is complete.  $\Box$ 

Theorem 3.1 is useful for evaluating the cause of the fact that the ASS model does not hold for given data.

We denote by  $n_{ij}$  the observed frequency in the (i, j)th cell of the table, and by  $p_{ij}$  (=  $\Pr(X = i, Y = j)$ ) the probability that an observation will fall in the (i, j)th cell of the table (i = 1, ..., R; j = 1, ..., R). We assume multinomial sampling over the cells of the square contingency table; that is, the observed frequencies  $\{n_{ij}\}$  have a multinomial distribution with parameters that are the cell probabilities  $\{p_{ij}\}$ .

Each model can be tested for goodness-of-fit by means of, for example, the likelihood ratio chi-square statistic (denoted by  $G^2$ ) with the corresponding df. The test statistic  $G^2$  of the model M is given by

$$G^{2}(\mathbf{M}) = 2\sum_{i=1}^{R}\sum_{j=1}^{R}n_{ij}\log\left(\frac{n_{ij}}{\hat{e}_{ij}}\right),$$

where  $\hat{e}_{ij}$  is the maximum likelihood estimate (MLE) of the expected frequency  $e_{ij}$  under the model M.

We obtain the following orthogonality of the test statistic.

**Theorem 3.2.** The following equality holds:

$$G^{2}(ASS) = G^{2}(ACSS) + G^{2}(AGS).$$

*Proof.* Although details are omitted, the  $\hat{e}_{ij}$  under the ASS, ACSS and AGS models are provided as (1), (2) and (3), respectively.

р

$$\hat{e}_{ij} = \begin{cases} \frac{A_{i+j} + B_{i+j}}{2A_{i+j}} n_{ij} & (i+j < R+1), \\ n_{ij} & (i+j = R+1), \\ \frac{A_{i+j} + B_{i+j}}{2B_{i+j}} n_{ij} & (i+j > R+1), \end{cases}$$
(1)

where

$$A_{i+j} = \sum_{(k,l)\in C_{i+j}} n_{kl}, \quad B_{i+j} = \sum_{(k,l)\in D_{i+j}} n_{kl},$$

$$C_{i+j} = \begin{cases} \{(k,l)|k+l = i+j, k+l < R+1\} & (i+j < R+1), \\ \{(k,l)|k+l = 2(R+1) - (i+j), k+l < R+1\} & (i+j > R+1), \\ \end{cases}$$
 and

$$D_{i+j} = \begin{cases} \{(k,l)|k+l = 2(R+1) - (i+j), k+l > R+1\} & (i+j < R+1), \\ \{(k,l)|k+l = i+j, k+l > R+1\} & (i+j > R+1). \end{cases}$$

$$\hat{e}_{ij} = \begin{cases} \frac{(A_{i+j} + B_{i+j})E}{A_{i+j}(1+E)} n_{ij} & (i+j < R+1), \\ n_{ij} & (i+j = R+1), \\ \frac{A_{i+j} + B_{i+j}}{B_{i+j}(1+E)} n_{ij} & (i+j > R+1), \end{cases}$$
(2)

where

$$E = \frac{\sum_{t=2}^{R} \sum_{(k,l) \in C_{t}} n_{kl}}{\sum_{t=2}^{R} \sum_{(k,l) \in D_{t}} n_{kl}} = \frac{\sum_{k+l < R+1} n_{kl}}{\sum_{k+l > R+1} n_{kl}}.$$

$$\hat{e}_{ij} = \begin{cases} \frac{1+E}{2E} n_{ij} & (i+j < R+1), \\ n_{ij} & (i+j = R+1), \\ \frac{1+E}{2} n_{ij} & (i+j > R+1). \end{cases}$$
(3)

The  $n_{ij}/\hat{e}_{ij}$  in the ASS model is equal to the product of those in the ACSS and AGS models. Therefore, the value of the test statistic  $G^2$  for the ASS model is equal to the sum of those for the ACSS and AGS models. The proof is complete.  $\hfill \Box$ 

From Theorem 3.2, we see that the value of the likelihood ratio chisquared statistic for testing the goodness-of-fit of the ASS model assuming that the ACSS model holds (i.e.,  $G^2(ASS|ACSS)$ ) is equal to the value of  $G^2(AGS)$ , because  $G^2(ASS|ACSS) = G^2(ASS) - G^2(ACSS)$ .

# 4. Application to real-world grip strength data

Consider the data set in Table 1, which presents a cross-classification of grip strength levels for right and left hands. These data are from a grip strength test on women aged 15–69; source: National Health and Nutrition Examination Survey 2011–2012 (https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/MGX\_G.htm).

Table 2 gives the values of  $G^2$  for the ASS, ACSS and AGS models. This table shows that the ACSS model fits well, but the other models fit poorly. Table 3 shows the MLEs of the expected frequencies under the ACSS model.

**Table 2.** Values of the likelihood ratio chi-squarestatistic  $G^2$  for each model applied to the data of<br/>Table 1

Applied models	Degrees of freedom	$G^2$
ASS	4	$43.98^{*}$
ACSS	3	3.50
AGS	1	$40.48^{*}$

\* indicates significance at the 0.05 level.

Under the ACSS model, the MLE of  $\Delta$  is 1.33. In the ACSS model, the probability with which X + Y is t, for  $t = 2, \ldots, R$  is estimated to be 1.33 times higher than the probability with which X + Y is 2(R + 1) - t. Moreover, since  $\Delta > 1$ , the median of the degree of an individual's grip strength is estimated to be smaller than the midpoint R + 1. Therefore, it may be necessary to revise the criteria for the grip strength level, since we can infer that the women's grip strength tends to be higher than the criteria.

From Theorem 3.1, we can infer that the cause of the fact that the ASS model does not hold for the data of Table 1 is the AGS model rather than the ACSS model.

	Left hand							
Right hand	(1)	(2)	(3)	(4)	(5)			
Excellent $(1)$	380	229	49	17	2			
	(362.54)	(236.26)	(48.72)	(17.66)	(2)			
Very good $(2)$	34	204	198	80	23			
	(35.08)	(202.83)	(205.74)	(80)	(21.91)			
Good $(3)$	5	55	159	161	74			
	(4.97)	(57.15)	(159)	(153.35)	(74.57)			
Fair $(4)$	1	11	37	106	168			
	(1.04)	(11)	(35.24)	(106.82)	(161.42)			
Poor $(5)$	0	2	12	45	256			
	(0)	(1.90)	(12.09)	(43.24)	(273.46)			

Table 3. Maximum likelihood estimates of expected frequencies under theanti-conditional sum-symmetry (ACSS) model applied to the data set in<br/>Table 1

Note: Estimates under the ACSS model are shown in parentheses in the second line.

#### 5. Concluding remarks

This study has proposed two original models (the ASS and ACSS models) that have symmetric and asymmetric structures between the probability with which X + Y is t, for t = 2, ..., R, and the probability with which X + Y is 2(R+1) - t. These models are useful for evaluating whether or not the level of X + Y is symmetric with respect to the midpoint R + 1 in the range [2, 2R] of the X + Y. The proposed models are useful for application to data where one is interested in the structure of X + Y, such as grip strength data, because interpretations such as those of the proposed models cannot be obtained from the existing models.

In addition, this study has revealed that (i) it is necessary to satisfy the AGS model, in addition to the ACSS model, in order to satisfy the ASS model (Theorem 3.1), and (ii) the value of the likelihood ratio chi-squared statistic of the ASS model is equal to the sum of those of the ACSS and AGS models (Theorem 3.2). Theorem 3.1 is useful for evaluating the cause of the fact that the ASS model does not hold for given data, as shown in Section 4. From Theorem 3.2, we note that the value of the likelihood ratio chi-squared statistic for testing the goodness-of-fit of the ASS model assum-

ing that the ACSS model holds (i.e.,  $G^2(ASS|ACSS)$ ) is equal to the value of  $G^2(AGS)$ , because  $G^2(ASS|ACSS) = G^2(ASS) - G^2(ACSS)$ .

# Acknowledgements

The author would like to thank the anonymous reviewers and the editors for their comments and suggestions to improve this paper.

# References

- Bowker, A. H. (1948): A test for symmetry in contingency tables. Journal of the American Statistical Association 43: 572–574.
- Iki, K. (2016): Analysis of the grip strength data using anti-diagonal symmetry models. Open Journal of Statistics 6: 590–593.
- Kurakami, H., Negishi, N., Tomizawa, S. (2017): On decomposition of pointsymmetry for square contingency tables with ordered categories. Journal of Statistics: Advances in Theory and Applications 17: 33–42.
- McCullagh, P. (1978): A class of parametric models for the analysis of square contingency tables with ordered categories. Biometrika 65: 413–418.
- Read, C. B. (1977): Partitioning chi-square in contingency tables: A teaching approach. Communications in Statistics-Theory and Methods 6: 553–562.
- Tomizawa, S. (1985): The decompositions for point-symmetry models in two-way contingency tables. Biometrical Journal 27: 895–905.
- Tomizawa, S. (1986): Four kinds of symmetry models and their decompositions in a square contingency table with ordered categories. Biometrical Journal 28: 387–393.
- Yamamoto, K., Aizawa, M., Tomizawa, S. (2016): Measure of departure from sumsymmetry model for square contingency tables with ordered categories. Journal of Statistics: Advances in Theory and Applications 16: 17–43.
- Yamamoto, K, Tanaka, Y., Tomizawa, S. (2013): Sum-symmetry model and its orthogonal decomposition for square contingency tables. SUT Journal of Mathematics 49: 121–128.