EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 4, Article Number 6439
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Anti-Linear-Diagonals-Parameter Symmetry Model and Orthogonal Decomposition of Anti-Symmetry for Square Contingency Tables with Ordinal Classifications

Shuji Ando

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, Japan

Abstract. For $R \times R$ square contingency tables with the same ordinal classifications for rows and columns, this study investigates models in which the relationship between the row and column variables is symmetric or asymmetric with respect to the anti-diagonal, rather than the main diagonal. The recently proposed anti-diagonals-parameter symmetry model includes $R-1$ asymmetric parameters and is capable of representing complex asymmetric structures. However, this model is saturated in the following cells: the cell in the first row and first column, the cell in the Rth row and Rth column, and all anti-diagonal cells. Since observed frequencies in square contingency tables tend to concentrate along the main diagonal, a model saturated only in the anti-diagonal cells may be more appropriate. We propose the anti-linear diagonals-parameter symmetry model, which captures asymmetry with respect to the anti-diagonal. The proposed model is saturated solely in the anti-diagonal cells and, like the anti-diagonals-parameter symmetry model, expresses how the degree of asymmetry varies according to the distance from the anti-diagonal. Furthermore, we demonstrate a decomposition of the anti-symmetry model using the proposed model and derive an orthogonal decomposition of the test statistic for the anti-symmetry model.

2020 Mathematics Subject Classifications: 62H17

Key Words and Phrases: Anti-diagonal, Asymmetry, Decomposition, Symmetry, Test statistic.

1. Introduction

For $R \times R$ square contingency tables with the same ordinal classifications for rows and columns, we denote by π_{ij}, for $(i,j) \in A$, the probability that an observed frequency falls in the (i,j)th cell of the table, where $A = \{(i,j) \mid i,j = 1,2,\ldots,R\}$. Assume that π_{ij}, for all $(i,j) \in A$, are positive.

For the analysis of square contingency tables, we often use a class of models defined by the following equation:

$$\pi_{ij} = \delta_{ij}\pi_{ji} \quad \text{for } (i,j) \in D,$$

DOI: https://doi.org/10.29020/nybg.ejpam.v18i4.6439

Email address: shuji.ando@rs.tus.ac.jp (S. Ando)

where δ_{ij}, for all $(i,j) \in D$, are unknown parameters, and $D = \{(i,j) \mid i,j = 1,2,\ldots,R;\ i < j\}$. When no restrictions are imposed on δ_{ij} for $(i,j) \in D$, the model above corresponds to the saturated model. Conversely:

(i) When the restrictions $\delta_{ij} = 1$ for $(i,j) \in D$ are imposed, the model corresponds to the symmetry (S) model [1].

(ii) When the restrictions $\delta_{ij} = \delta$ for $(i,j) \in D$ are imposed, the model corresponds to the conditional symmetry (CS) model [2].

(iii) When the restrictions $\delta_{ij} = \delta_{j-i}$ for $(i,j) \in D$ are imposed, the model corresponds to the diagonals-parameter symmetry (DPS) model [3].

(iv) When the restrictions $\delta_{ij} = \delta^{j-i}$ for $(i,j) \in D$ are imposed, the model corresponds to the linear DPS (LDPS) model [4].

The cells on the main diagonal of the table represent cases where the difference between the row variable X and the column variable Y is zero, which corresponds to the mean of $X - Y$. The S model exhibits a symmetric structure in the cell probabilities with respect to the main diagonal of the table. Conversely, the CS, DPS, and LDPS models exhibit asymmetric structures in the cell probabilities with respect to the main diagonal.

Next, we consider a class of models defined by the following equation:

$$\pi_{ij} = \delta_{ij}\pi_{j^*i^*} \quad \text{for } (i,j) \in E,$$

where δ_{ij}, for all $(i,j) \in E$, are unknown parameters, $i^* = R + 1 - i$, $j^* = R + 1 - j$, and $E = \{(i,j) \mid i,j = 1,2,\ldots,R;\ i+j < R+1\}$. The model above exhibits either a symmetric or asymmetric structure in the cell probabilities with respect to the anti-diagonal of the table. The cells on the anti-diagonal correspond to cases where the sum of X and Y equals $R + 1$, which represents the mean of $X + Y$.

When no restrictions are imposed on δ_{ij} for $(i,j) \in E$, the model above corresponds to the saturated model. Conversely:

(i) When the restrictions $\delta_{ij} = 1$ for $(i,j) \in E$ are imposed, the model corresponds to the anti-symmetry (AS) model [5].

(ii) When the restrictions $\delta_{ij} = \delta$ for $(i,j) \in E$ are imposed, the model corresponds to the anti-conditional symmetry (ACS) model [6].

(iii) When the restrictions $\delta_{ij} = \delta_{R+1-(i+j)}$ for $(i,j) \in E$ are imposed, the model corresponds to the anti-diagonals-parameter symmetry (ADPS) model [7].

Given that the ADPS model includes $R-1$ asymmetric parameters (i.e., $\delta_1, \delta_2, \ldots, \delta_{R-1}$), while the ACS model includes only one (i.e., δ), the ADPS model offers greater flexibility than the ACS model in representing complex asymmetric structures. However, the ADPS model is saturated in the following cells: the cell in the first row and first column, the cell in the Rth row and Rth column, and all anti-diagonal cells. Observed frequencies in square contingency tables tend to concentrate along the main diagonal. In light of this tendency, it may be advantageous to consider a model that, like the ACS model, is saturated only on the anti-diagonal, but that offers greater flexibility in capturing asymmetric structures. To address this issue, the present paper proposes a new model that is saturated only in the anti-diagonal cells of the table. Like the ADPS model, the proposed model allows the degree of asymmetry to vary depending on the distance from the anti-diagonal.

The remainder of this paper is organized as follows. In Section 2, we propose a new model. Section 2 also presents the decomposition of the AS model using the proposed model, while Section 3 describes the decomposition of the test statistic for the AS model. Section 4 demonstrates the advantages of the proposed model through an application to a real dataset. Finally, Section 5 provides concluding remarks.

2. Proposed model

This study proposes the anti-linear diagonals-parameter symmetry (ALDPS) model, which is defined by the following equation:

$$\pi_{ij} = \delta^{R+1-(i+j)}\pi_{j^*i^*} \quad \text{for } (i,j) \in E,$$

where δ is unknown parameter. Note that the ALDPS model is saturated only in the anti-diagonal cells of the table. Similar to the ADPS model, the ALDPS model allows the degree of asymmetry to vary depending on the distance from the anti-diagonal (i.e., $R+1-(i+j)$). However, while the ADPS model includes $R-1$ asymmetric parameters, the ALDPS model includes only a single asymmetric parameter. As a special case of the ADPS model, the ADPS model with $\delta_{R+1-(i+j)} = \delta^{R+1-(i+j)}$ for $(i,j) \in E$ coincides with the ALDPS model. Therefore, the ADPS model offers greater flexibility than the ALDPS model in representing complex asymmetric structures. Nevertheless, the ADPS model is saturated in the following cells: the cell in the first row and first column, the cell in the Rth row and Rth column, and all anti-diagonal cells. Given that observed frequencies in square contingency tables tend to concentrate along the main diagonal, the ALDPS model may be considered more advantageous than the ADPS model in practice.

As a special case of the ALDPS model, the ALDPS model with $\delta^{R+1-(i+j)} = 1$ for $(i,j) \in E$ coincides with the AS model. When the AS model holds, the ALDPS model necessarily holds as well, although the converse is not always true. In Section 3, we identify a counterpart to the ALDPS model for the purpose of decomposing the AS model.

We discuss the relationship between the ALDPS model and the bivariate normal distribution. Let V and W be random variables that follow a joint bivariate normal distribution with means $E(V) = \mu_1$ and $E(W) = \mu_2$, variances $\mathrm{Var}(V) = \mathrm{Var}(W) = \sigma^2$, and correlation $\mathrm{Corr}(V,W) = \rho$. The joint bivariate normal density $f(v,w)$ is given by:

$$f(v,w) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{4\sigma^2(1-\rho)}\{(v-w)-(\mu_1-\mu_2)\}^2 - \frac{1}{4\sigma^2(1+\rho)}\{(v+w)-(\mu_1+\mu_2)\}^2 \right].$$

Therefore, the density $f(v,w)$ satisfies the following identity:

$$\frac{f(v,w)}{f(w^*,v^*)} = \exp\left[\frac{1}{\sigma^2(1+\rho)}\{\eta - (\mu_1+\mu_2)\}\{\eta - (v+w)\}\right] \quad \text{for } v < w,$$

where $v^* = \eta - v$ and $w^* = \eta - w$ for any η. This identity resembles the functional form of the ALDPS model. Therefore, when we aim to capture the linear structure in the log odds $\log\{f(v,w)/f(w^*,v^*)\}$, the ALDPS model provides an appropriate framework.

In terms of simulation studies, Tables 1a and b, taken from Tomizawa, Miyamoto and Ashihara [8], give the 4×4 tables of sample size 10,000 formed by using cut points for each variable at $\mu_1, \mu_1 \pm 0.6\sigma$, for underlying bivariate normal distribution with the conditions $\sigma_1^2 = \sigma_2^2 = \sigma^2, \mu_2 = \mu_1 + 0.4$, and the correlations $\rho = 0$ (Table 1a) and $\rho = 0.3$ (Table 1b).

Table 1: The 4×4 tables of sample size 10,000, formed by using cut points for each variable at $\mu_1, \mu_1 \pm 0.6\sigma$, from an underlying bivariate normal distribution with the conditions $\mu_2 = \mu_1 + 0.4, \sigma_1^2 = \sigma_2^2 = \sigma^2$ and $\rho = 0, 0.3$

(a) For $\rho = 0$				(b) For $\rho = 0.3$			
428	526	671	1174	696	666	678	785
358	416	561	951	384	436	587	836
374	405	544	875	269	388	554	1008
405	509	658	1145	216	366	615	1516

Source: Tomizawa, Miyamoto and Ashihara [8]

The ALDPS model fits these data well yielding the likelihood ratio chi-squared values $G^2 = 5.89$ (for Table 1a) and $G^2 = 4.35$ (for Table 1b) with both 5 degrees of freedom. Therefore, the ALDPS model may be appropriate for a square table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances.

3. Separation of the anti-symmetry model via the anti-linear diagonals-parameter symmetry model

In this section, we present a decomposition of the AS model using the ALDPS model. The objective is to identify a counterpart to the ALDPS model that enables the decomposition of the AS model. As such a counterpart, we introduce the anti-marginal equity (AME) model [9], which is defined by the following equation:

$$E(X) = E(R+1-Y),$$

where

$$E(X) = \sum_{(i,j)\in A} i\pi_{ij} \quad \text{and} \quad E(R+1-Y) = \sum_{(i,j)\in A} (R+1-j)\pi_{ij} = \sum_{(i,j)\in A} j^*\pi_{ij}.$$

The decomposition of the AS model is derived using the ALDPS and AME models, as stated in the following theorem.

Theorem 1. *The AS model holds if and only if both the ALDPS and AME models hold simultaneously.*

Proof. We first show the necessity. Assume that the AS model holds, i.e., $\pi_{ij} = \pi_{j^*i^*}$ for $(i,j) \in E$. Then, the ALDPS model clearly holds. Consider the following equation:

$$E(X-(R+1-Y)) = \sum_{i+j<R+1} (i-j^*)\pi_{ij} + \sum_{i+j=R+1} (i-j^*)\pi_{ij} + \sum_{i+j>R+1} (i-j^*)\pi_{ij}. \tag{1}$$

In equation (1), the second term on the right-hand side equals zero since $i - j^* = 0$ when $i + j = R + 1$. Therefore, under the AS model, equation (1) becomes:

$$\begin{aligned} E(X-(R+1-Y)) &= \sum_{i+j<R+1} (i-j^*)\pi_{ij} + \sum_{i+j>R+1} (i-j^*)\pi_{ij} \\ &= -\sum_{i+j<R+1} (j^*-i)\pi_{ij} + \sum_{i+j>R+1} (i-j^*)\pi_{ij} \\ &= -\sum_{i+j>R+1} (i-j^*)\pi_{j^*i^*} + \sum_{i+j>R+1} (i-j^*)\pi_{ij} \\ &= -\sum_{i+j>R+1} (i-j^*)\pi_{ij} + \sum_{i+j>R+1} (i-j^*)\pi_{ij} \\ &= 0. \end{aligned}$$

Hence, we obtain $E(X) = E(R + 1 - Y)$, which implies that the AME model holds. Therefore, the necessary condition is satisfied.

We now show the sufficiency. Assume that both the ALDPS and AME models hold. Then we have:

$$\begin{aligned} E(X-(R+1-Y)) &= -\sum_{i+j>R+1} (i-j^*)\pi_{ij}\delta^{R+1-(i+j)} + \sum_{i+j>R+1} (i-j^*)\pi_{ij} \\ &= \sum_{i+j>R+1} (i-j^*)\pi_{ij}(1-\delta^{R+1-(i+j)}) \\ &= 0. \end{aligned}$$

Since $\pi_{ij} > 0$ for $(i,j) \in A$, it follows that $\delta = 1$. Therefore, the ALDPS model reduces to the AS model, and the sufficiency is established.

When the AS model does not fit the observed data well, Theorem 1 provides insight into the possible reasons for the poor fit.

4. Separation of test statistic for the anti-symmetry model

Let N denote the sample size, i.e., $N = \sum\sum_{(i,j)\in A} n_{ij}$, where n_{ij} for $(i,j) \in A$ is the observed frequency in the (i,j)th cell of the table. We assume that the observed frequencies n_{ij} for $(i,j) \in A$ follow a multinomial distribution with probability vector $\boldsymbol{\pi}$, where

$$\boldsymbol{\pi} = (\pi_{11}, \pi_{12}, \ldots, \pi_{1R}, \pi_{21}, \pi_{22}, \ldots, \pi_{2R}, \ldots, \pi_{R1}, \pi_{R2}, \ldots, \pi_{RR})^{\top}.$$

The symbol "$\top$" denotes the transpose of a vector or matrix.

Let $\hat{e}_{ij}$ for $(i,j) \in A$ denote the maximum likelihood estimator (MLE) of the expected frequencies e_{ij} under the model of interest. The MLE $\hat{e}_{ij}$ can be obtained by maximizing the log-likelihood function subject to the constraints imposed by the model. Specifically, under the ALDPS model, $\hat{e}_{ij}$ for $(i,j) \in A$ can be obtained by maximizing the following Lagrangian with respect to $\{\pi_{ij}\}$, ϕ, $\{\psi_{ij}\}$, and δ, using the Newton–Raphson method:

$$L = \sum_{(i,j)\in A} n_{ij} \log \pi_{ij} - \phi \left(\sum_{(i,j)\in A} \pi_{ij} - 1 \right) - \sum_{(i,j)\in E} \psi_{ij} \left(\pi_{ij} - \delta^{R+1-(i+j)} \pi_{j^* i^*} \right).$$

The likelihood ratio statistic for testing the goodness-of-fit of model M is given by

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

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

Table 2 shows the degrees of freedom for testing the goodness of fit for each model. The number of degrees of freedom for the ALDPS model is given by $(R+1)(R-2)/2$, which is $R-2$ fewer than that for the ADPS model. Note that the number of degrees of freedom for the AS model equals the sum of that for the ALDPS and AME models.

Table 2: The number of degrees of freedom for testing goodness-of-fit for each model

Models	Degrees of freedom
AS	$R(R-1)/2$
ACS	$(R+1)(R-2)/2$
ADPS	$(R-1)(R-2)/2$
ALDPS	$(R+1)(R-2)/2$
AME	1

Assume that model M_3 holds if and only if both models M_1 and M_2 hold, where the number of degrees of freedom for the model M_3 equals the sum of that for the models M_1 and M_2. Darroch and Silvey [10] described that (i) when the following asymptotic equivalence holds:

$$G^2(\mathrm{M}_3) \simeq G^2(\mathrm{M}_1) + G^2(\mathrm{M}_2), \tag{2}$$

if both models M_1 and M_2 are accepted (at the α significance level) with high probability, then the model M_3 would be accepted; however, (ii) when the equation (2) does not hold, such an incompatible situation that both models M_1 and M_2 are accepted with high probability but the model M_3 is rejected with high probability is quite possible. In fact, Darroch and Silvey [10], Tahata, Ando and Tomizawa [11] showed such interesting examples.

We now demonstrate that Theorem 1 satisfies the asymptotic equivalence given in equation (2). Accordingly, the following theorem is obtained.

Theorem 2. *For $R \times R$ square contingency tables, the following asymptotic equivalence holds:*

$$G^2(\mathrm{AS}) \simeq G^2(\mathrm{ALDPS}) + G^2(\mathrm{AME}).$$

Proof. The ALDPS model may be expressed as

$$\log \pi_{ij} = \{R + 1 - (i + j)\}\beta_1 + \phi_{ij} \qquad i = 1, \dots, r; j = 1, \dots, r, \tag{3}$$

where $\phi_{ij} = \phi_{j^*i^*}$. Let

$$\boldsymbol{\beta} = (\beta_1, \boldsymbol{\beta}_2)^\top,$$

where

$$\boldsymbol{\beta}_2 = (\phi_{11}, \phi_{12}, \dots, \phi_{1r}, \phi_{21}, \phi_{22}, \dots, \phi_{2,R-1}, \dots, \phi_{RR}),$$

is the $1 \times R(R+1)/2$ vector of ϕ_{ij} for $i+j \leq R+1$. Then, the ALDPS model is expressed as

$$\log \boldsymbol{\pi} = \boldsymbol{X}\boldsymbol{\beta} = (\boldsymbol{X}_1, \boldsymbol{X}_2)\boldsymbol{\beta},$$

where $\boldsymbol{X}$ is the $R^2 \times K$ matrix with $K = (R^2 + R + 2)/2$ and

$$\boldsymbol{X}_1 = (R+1)\mathbf{1}_{R^2} - (\mathbf{1}_R \otimes \boldsymbol{J}_R + \boldsymbol{J}_R \otimes \mathbf{1}_R); \quad \text{the } R^2 \times 1\text{vector},$$

and $\boldsymbol{X}_2$ is the $R^2 \times R(R+1)/2$ matrix of 1 or 0 elements, determined from the equation (3), $\mathbf{1}_s$ is the $s \times 1$ vector of 1 elements and $\boldsymbol{J}_R = (1, 2, \dots, R)^\top$, and the symbol "$\otimes$" represents the Kronecker product. Note that $\boldsymbol{X}_2\mathbf{1}_{R(R+1)/2} = \mathbf{1}_{R^2}$ holds. Note that the matrix $\boldsymbol{X}$ is full column rank which is K.

In a similar manner to Haber [12], and Lang and Agresti [13], we denote the linear space spanned by the columns of the matrix $\boldsymbol{X}$ by $\mathcal{S}(\boldsymbol{X})$ with the dimension K. Let $\boldsymbol{U}$ be an $R^2 \times d_1$ matrix, where $d_1 = R^2 - K = (R+1)(R-2)/2$, full column rank matrix such that the linear space spanned by the columns of $\boldsymbol{U}$, i.e., $\mathcal{S}(\boldsymbol{U})$, is the orthogonal complement of the space $\mathcal{S}(\boldsymbol{X})$. Thus, $\boldsymbol{U}^\top\boldsymbol{X} = \boldsymbol{O}_{d_1,K}$ where $\boldsymbol{O}_{d_1,K}$ is the $d_1 \times K$ zero matrix. Therefore, the ALDPS model is expressed as

$$h_1(\boldsymbol{\pi}) = \mathbf{0}_{d_1},$$

where $\mathbf{0}_{d_1}$ is the $d_1 \times 1$ zero vector and

$$h_1(\boldsymbol{\pi}) = \boldsymbol{U}^\top \log \boldsymbol{\pi}.$$

The AME model may be expressed as

$$h_2(\boldsymbol{\pi}) = 0,$$

where

$$h_2(\boldsymbol{\pi}) = \boldsymbol{W}\boldsymbol{\pi},$$

with

$$\boldsymbol{W} = ((R+1)\mathbf{1}_{R^2} - (\mathbf{1}_R \otimes \boldsymbol{J}_R + \boldsymbol{J}_R \otimes \mathbf{1}_R))^\top; \quad \text{the } 1 \times R^2 \text{vector.}$$

Namely, $\boldsymbol{W}^\top = \boldsymbol{X}_1$. Thus $\boldsymbol{W}^\top$ belongs to the space $\mathcal{S}(\boldsymbol{X})$, i.e., $\mathcal{S}(\boldsymbol{W}^\top) \subset \mathcal{S}(\boldsymbol{X})$. Hence $\boldsymbol{W}\boldsymbol{U} = \mathbf{0}_{d_1}^\top$. From Theorem 1, the AS model may be expressed as

$$h_3(\boldsymbol{\pi}) = \mathbf{0}_{d_3},$$

where $d_3 = d_1 + d_2 = R(R-1)/2$ with $d_2 = 1$,

$$h_3 = (h_1, h_2)^\top.$$

Note that $h_s(\boldsymbol{\pi})$, $s = 1, 2, 3$, are the vectors of order $d_s \times 1$, and d_s, $s = 1, 2, 3$, are the numbers of degrees of freedom for testing goodness-of-fit of the ALDPS, AME and AS models, respectively.

Let $H_s(\boldsymbol{\pi})$, $s = 1, 2, 3$, denote the $d_s \times R^2$ matrix of partial derivatives of $h_s(\boldsymbol{\pi})$ with respect to $\boldsymbol{\pi}$, i.e., $H_s(\boldsymbol{\pi}) = \partial h_s(\boldsymbol{\pi})/\partial \boldsymbol{\pi}^\top$. Let $\boldsymbol{\Sigma}(\boldsymbol{\pi}) = \text{diag}(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}^\top$, where $\text{diag}(\boldsymbol{\pi})$ denotes a diagonal matrix with ith component of $\boldsymbol{\pi}$ as ith diagonal component. Let $\hat{\boldsymbol{\pi}}$ denote $\boldsymbol{\pi}$ with $\{\pi_{ij}\}$ replaced by $\{\hat{\pi}_{ij}\}$, where $\hat{\pi}_{ij} = n_{ij}/N$. Then $\sqrt{N}(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi})$ has asymptotically a normal distribution with mean $\mathbf{0}_{R^2}$ and covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\pi})$. Using the delta method, $\sqrt{N}(h_3(\hat{\boldsymbol{\pi}}) - h_3(\boldsymbol{\pi}))$ is asymptotically distributed to a normal distribution with mean $\mathbf{0}_{d_3}$ and covariance matrix

$$H_3(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_3(\boldsymbol{\pi})^\top = \begin{bmatrix} H_1(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_1(\boldsymbol{\pi})^\top & H_1(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_2(\boldsymbol{\pi})^\top \\ H_2(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_1(\boldsymbol{\pi})^\top & H_2(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_2(\boldsymbol{\pi})^\top \end{bmatrix}.$$

We see that $H_1(\boldsymbol{\pi}) = \boldsymbol{U}^\top \mathbf{1}_{R^2} = \mathbf{0}_{d_1}$ since $\mathbf{1}_{R^2} \subset \mathcal{S}(\boldsymbol{X})$, $H_1(\boldsymbol{\pi})\text{diag}(\boldsymbol{\pi}) = \boldsymbol{U}^\top$ and $H_2(\boldsymbol{\pi}) = \boldsymbol{W}$. Therefore we obtain

$$H_1(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})H_2(\boldsymbol{\pi})^\top = \boldsymbol{U}^\top \boldsymbol{W}^\top = \mathbf{0}_{d_1}.$$

Thus we obtain $\Delta_3(\hat{\boldsymbol{\pi}}) = \Delta_1(\hat{\boldsymbol{\pi}}) + \Delta_2(\hat{\boldsymbol{\pi}})$, where

$$\Delta_s(\hat{\boldsymbol{\pi}}) = h_s(\hat{\boldsymbol{\pi}})^\top \left[H_s(\hat{\boldsymbol{\pi}})\boldsymbol{\Sigma}(\hat{\boldsymbol{\pi}})H_s(\hat{\boldsymbol{\pi}})^\top \right]^{-1} h_s(\hat{\boldsymbol{\pi}}). \tag{4}$$

Under each $h_s(\boldsymbol{p}) = \mathbf{0}_{d_s}$ $(s = 1, 2, 3)$, the Wald statistic $N\Delta_s(\hat{\boldsymbol{\pi}})$ has asymptotically a chi-squared distribution with d_s degrees of freedom. From the equation (4), we see that $N\Delta_3(\hat{\boldsymbol{\pi}}) = N\Delta_1(\hat{\boldsymbol{\pi}}) + N\Delta_2(\hat{\boldsymbol{\pi}})$. From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic [10], we obtain Theorem 2. The proof is completed.

Table 3: The grip strength test dataset for men between the ages of 15 and 69

	Left hand					
Right hand	(1)	(2)	(3)	(4)	(5)	Total
Highest (1)	215	124	46	14	2	401
(2)	37	143	165	74	16	435
(3)	7	45	156	166	51	425
(4)	2	20	62	226	210	520
Lowest (5)	1	2	16	61	495	575
Total	262	334	445	541	774	2356

Source: Yamamoto, Aizawa and Tomizawa [14]

5. Real data analysis

We examine the dataset presented in Table 3, which is taken directly from Yamamoto, Aizawa, and Tomizawa [14]. This dataset summarizes the results of grip strength measurements conducted on a cohort of 2,356 men aged 15 to 69 years, as part of the National Health and Nutrition Examination Survey (NHANES) conducted in 2011–2012. Grip strength for the right and left hands is categorized according to the NHANES manual.

In the case of grip strength data, such as that presented in Table 3, symmetry or asymmetry between right and left hand grip strength may be of less concern. This is because the dominant hand typically exhibits greater grip strength than the non-dominant hand, with the majority of individuals being right-handed, as reported by an Intage Group self-administered survey (`https://gallery.intage.co.jp/smartphone-operation/`). Furthermore, Iki [5] and Ando [7, 15, 16] analyzed grip strength data using models in which the relationship between the row and column variables is symmetric or asymmetric with respect to the anti-diagonal of the table, rather than the main diagonal. Accordingly, we are interested in applying the AS, ACS, ADPS, and ALDPS models to the dataset in Table 3.

When models M_1 and M_2 are nested (i.e., model M_1 is more parsimonious than model M_2), the likelihood ratio statistic for testing whether model M_1 holds under the assumption that model M_2 is true is given by

$$G^2(\mathrm{M}_1 \mid \mathrm{M}_2) = G^2(\mathrm{M}_1) - G^2(\mathrm{M}_2).$$

Under the null hypothesis, the statistic $G^2(\mathrm{M}_1 \mid \mathrm{M}_2)$ asymptotically follows a chi-squared distribution with the number of degrees of freedom equal to the difference in the number of degrees of freedom between models M_1 and M_2; see, for example, Agresti [17, Sec. 3.4.4]. When models M_1 and M_2 are not nested, the statistic $G^2(\mathrm{M}_1 \mid \mathrm{M}_2)$ cannot be used for model comparison. In such cases, alternative criteria are required. Among the most widely used are the Akaike Information Criterion (AIC) [18] and the Bayesian Information Criterion (BIC) [19], both of which are used to identify the best-fitting model among competing candidates. The model with the minimum AIC or BIC is considered the best-fitting model. Since only the difference between AIC (or BIC) values is required for comparison, the constant part can be omitted. Therefore, we define the modified AIC and

BIC as follows:

$$\mathrm{AIC}^{+} = G^2 - 2 \times (\text{number of degrees of freedom}),$$
$$\mathrm{BIC}^{+} = G^2 - \log N \times (\text{number of degrees of freedom}),$$

where N denotes the sample size. The model with the smallest AIC^{+} or BIC^{+} is selected as the best-fitting model among the models under consideration.

Table 4 provides the values of G^2 and AIC^{+} for the AS, ACS, ADPS, ALDPS, and AME models applied to the dataset in Table 3. From Table 4, we observe that the goodness-of-fit for both the ADPS and ALDPS models is well and shows a substantial improvement over the ACS model. The value of the test statistic $G^2(\mathrm{ALDPS} \mid \mathrm{ADPS})$ is 2.30. Therefore, the ALDPS model is preferable to the ADPS model. Furthermore, in terms of AIC^{+}, the ALDPS model is identified as the best-fitting model among those applied to the dataset in Table 3.

Table 4 provided the values of G^2 and AIC^{+} for each AS, ACS, ADPS, ALDPS, and AME model applied to the dataset in Table 3. From Table 4, we see that the goodness-of-fit of the ADPS model and ALDPS modell are well and dramatically improves compared to the ACS model. The value of test statistics $G^2(\mathrm{ALDPS} \mid \mathrm{ADPS})$ is 2.30. Therefore, the ALDPS model is preferable to the ADPS model. Moreover, in terms of AIC, the ALDPS model is the best-fitting model among the models applied to the dataset in Table 3.

Table 4: Results of goodness-of-fit test applied each model to the dataset of Table 3

Applied models	Degrees of freedom	G^2	p-value	AIC^{+}
AS	10	167.37*	< 0.01	147.37
ACS	9	43.92*	< 0.01	25.92
ADPS	6	5.35	0.50	-7.35
ALDPS	9	7.65	0.57	-10.35
AME	1	160.07*	< 0.01	158.07

The symbol $*$ represents significance at the 0.05 level.

Table 5 presents the estimated expected frequencies $\hat{e}_{ij}$ under the ALDPS and ADPS models. The maximum likelihood estimate of δ in the ALDPS model is 0.82. These results reveal a clear discrepancy in the ratio of men classified as having high versus low grip strength. It can therefore be concluded that the current criteria used to determine grip strength levels may be inadequate. In the future, it is anticipated that the classification criteria for grip strength will be revised.

Given that observations in square contingency tables tend to concentrate in the main diagonal cells, it may be desirable to consider a model that is saturated only on the anti-diagonal cells. In fact, the ratio of $(n_{11} + n_{55})$ to N in the dataset in Table 3 is approximately 30%. The ADPS model is saturated in the first row and first column cell, the Rth row and Rth column cell, as well as in the anti-diagonal cells of the table. For the dataset in Table 3, under the ADPS model, $\hat{e}_{11}$ and $\hat{e}_{55}$ are equal to n_{11} and n_{55}, respectively. In contrast, the ALDPS model is saturated only in the anti-diagonal cells.

Table 5: Maximum likelihood estimates of expected frequencies under the anti-diagonals-parameter symmetry (ADPS) model and anti-linear diagonals-parameter symmetry (ALDPS) model applied to the dataset in Table 3.The parenthesized values in lines 2 and 3 are the maximum likelihood estimates of expected frequencies under the ADPS and ALDPS models, respectively.

	Left hand					
Right hand	(1)	(2)	(3)	(4)	(5)	Total
Highest (1)	215	124	46	14	2	401
	(215)	(124.48)	(38.88)	(14.36)	(2)	
	(222.66)	(119.31)	(39.12)	(13.54)	(2)	
(2)	37	143	165	74	16	435
	(36.52)	(147.90)	(158.49)	(74)	(15.64)	
	(35.01)	(148.82)	(149.35)	(74)	(16.46)	
(3)	7	45	156	166	51	425
	(9.22)	(51.23)	(156)	(172.51)	(58.12)	
	(9.28)	(48.28)	(156)	(181.65)	(57.88)	
(4)	2	20	62	226	210	520
	(1.92)	(20)	(55.77)	(221.10)	(209.52)	
	(1.80)	(20)	(58.72)	(220.18)	(214.69)	
Lowest (5)	1	2	16	61	495	575
	(1)	(2.08)	(13.78)	(61.48)	(495)	
	(1)	(2.20)	(13.72)	(62.99)	(487.34)	
Total	262	334	445	541	774	2356

Accordingly, for the same dataset, under the ALDPS model, $\hat{e}_{11}$ and $\hat{e}_{55}$ are not equal to n_{11} and n_{55}, respectively.

In accordance with Theorem 1, for the dataset in Table 3, the poor goodness-of-fit of the AS model can be attributed to the AME model rather than to the ALDPS model. From Table 4, we observe that $G^2(\mathrm{AS}) = 167.37$, and the sum of $G^2(\mathrm{ALDPS})$ and $G^2(\mathrm{AME})$ is 167.72. According to Theorem 2, since the following asymptotic equivalence holds:

$$G^2(\mathrm{AS}) \simeq G^2(\mathrm{ALDPS}) + G^2(\mathrm{AME}),$$

we conclude that $G^2(\mathrm{AS})$ is nearly equal to the sum of $G^2(\mathrm{ALDPS})$ and $G^2(\mathrm{AME})$ in this dataset.

6. Conclusion

This study proposed the ALDPS model, which captures the asymmetric structure of cell probabilities with respect to the anti-diagonal of the table. We also examined the relationship between the ALDPS model and the bivariate normal distribution. Furthermore, we provided a decomposition of the AS model using the ALDPS model (Theorem 1), as well as a decomposition of the test statistic for the AS model (Theorem 2).

The ALDPS model was shown to offer substantial advantages through its application to the real dataset in Table 3. The results supported the hypothesis that males with high grip strength levels are less prevalent than those with low grip strength levels. Moreover, the findings suggested that the current criteria used to classify grip strength levels may be inadequate.

Readers may be interested in an extended ALDPS model, which would be suitable for a square contingency table under the assumption of an underlying bivariate normal distribution without requiring equality of the marginal variances. Following the approach of Tomizawa [20], extensions of the ALDPS model can also be considered. A detailed discussion of such extensions is left for future research.

Acknowledgements

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

References

[1] H Bowker. A test for symmetry in contingency tables. *Journal of the American Statistical Association*, 43(244):572–574, 1948.

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

[3] A Goodman. Multiplicative models for square contingency tables with ordered categories. *Biometrika*, 66(3):413–418, 1979.

[4] A Agresti. A simple diagonals-parameter symmetry and quasi-symmetry model. *Statistics & Probability Letters*, 1(6):313–316, 1983.

[5] K Iki. Analysis of the grip strength data using anti-diagonal symmetry models. *Open Journal of Statistics*, 6(04):590, 2016.

[6] S Tomizawa. Four kinds of symmetry models and their decompositions in a square contingency table with ordered categories. *Biometrical Journal*, 28(4):387–393, 1986.

[7] S Ando. Anti-diagonals-parameter symmetry model for square contingency tables with ordinal classifications. *European Journal of Pure and Applied Mathematics*, 19(2):6079, 1–14, 2025.

[8] S Tomizawa, N Miyamoto, and N Ashihara. Measure of departure from marginal homogeneity for square contingency tables having ordered categories. *Behaviormetrika*, 30(2):173–193, 2003.

[9] H Kurakami, N Negishi, and S Tomizawa. On decomposition of point-symmetry for square contingency tables with ordered categories. *Journal of Statistics: Advances in Theory and Applications*, 17(1):33–42, 2017.

[10] N Darroch and D Silvey. On testing more than one hypothesis. *The Annals of Mathematical Statistics*, 34:555–567, 1963.

[11] K Tahata, S Ando, and S Tomizawa. Ridit score type asymmetry model and decomposition of symmetry for square contingency tables. *Model Assisted Statistics and Applications*, 6:279–286, 2011.

[12] M Haber. Maximum likelihood methods for linear and log-linear models in categorical data. *Computational Statistics and Data Analysis*, 3:1–10, 1985.

[13] J Lang and A Agresti. Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *Journal of the American Statistical Association*, 89(426):625–632, 1994.

[14] K Yamamoto, M Aizawa, and S Tomizawa. Measure of departure from sum-symmetry model for square contingency tables with ordered categories. *Journal of Statistics: Advances in Theory and Applications*, 16:17–43, 2016.

[15] S Ando. An anti-sum-symmetry model and its orthogonal decomposition for ordinal square contingency tables with an application to grip strength test data. *Biometrical Letters*, 58(1):59–68, 2021.

[16] S Ando. Anti-sum-asymmetry models and orthogonal decomposition of anti-sum-symmetry model for ordinal square contingency tables. *Austrian Journal of Statistics*, 52(1):72–86, 2023.

[17] A Agresti. *An Introduction to Categorical Data Analysis.* Wiley, Hoboken, New Jersey, 3 edition, 2018.

[18] H Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[19] G Schwarz. Estimating the dimension of a model. *Annals of statistics*, 6(2):461–464, 1978.

[20] S Tomizawa. An extended linear diagonals-parameter symmetry model for square contingency tables with ordered categories. *Metron*, 49:401–409, 1991.