

# Extended Diagonal Exponent Symmetry Model and Its Orthogonal Decomposition in Square Contingency Tables with Ordered Categories

Kiyotaka Iki, Akira Shibuya, Sadao Tomizawa

Department of Information Sciences, Tokyo University of Science, Noda, Japan  
Email: [iki@is.noda.tus.ac.jp](mailto:iki@is.noda.tus.ac.jp)

Received 31 March 2015; accepted 31 May 2015; published 3 June 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

For square contingency tables with ordered categories, this article proposes new models, which are the extension of Tomizawa's [1] diagonal exponent symmetry model. Also it gives the decomposition of proposed model, and shows the orthogonality of the test statistics for decomposed models. Examples are given and the simulation studies based on the bivariate normal distribution are also given.

## Keywords

Diagonal Exponent Symmetry, Ordinal Category, Orthogonal Decomposition, Quasi-Symmetry, Square Contingency Table

---

## 1. Introduction

Consider an  $R \times R$  square contingency table with the same row and column classifications. Let  $p_{ij}$  denote the probability that an observation will fall in the  $i$ th row and  $j$ th column of the table ( $i = 1, \dots, R; j = 1, \dots, R$ ). The symmetry (S) model is defined by

$$p_{ij} = \psi_{ij} \quad (i = 1, \dots, R; j = 1, \dots, R),$$

where  $\psi_{ij} = \psi_{ji}$ ; see Bowker [2]. Caussinus [3] considered the quasi-symmetry (QS) model defined by

$$p_{ij} = \alpha_i \beta_j \psi_{ij} \quad (i = 1, \dots, R; j = 1, \dots, R),$$

where  $\psi_{ij} = \psi_{ji}$ . The marginal homogeneity (MH) model is defined by

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

where  $p_{i\cdot} = \sum_{t=1}^R p_{it}$  and  $p_{\cdot i} = \sum_{s=1}^R p_{si}$ ; see Stuart [4]. Caussinus [3] gave the theorem that the S model holds if and only if both the QS and MH models hold.

Tomizawa [1] considered the diagonal exponent symmetry (DES) model defined by

$$p_{ij} = \begin{cases} \delta^{i+j} d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

By putting  $\delta = \gamma^{\frac{1}{2}}$  and  $d_{|j-i|} = \mu_{|j-i|} \gamma^{-1-\frac{1}{2}|j-i|}$ , this model is also expressed as

$$p_{ij} = p_{ji} = \mu_{|j-i|} \gamma^{i-1} \quad (i < j).$$

Note that the DES model implies the S model; thus the DES model implies the QS (MH) model. The DES model states that  $p_{i+1, j+1}$  ( $i \neq j$ ) is  $\delta^2$  times higher than  $p_{ij}$ ; in other words, for fixed distance  $k$  ( $k = 1, \dots, R-2$ ) from the main diagonal of the table,  $p_{i, i+k}$  increase (decrease) exponentially along every subdiagonal of the table as the value  $i$  increase ( $i = 1, \dots, R-k$ ).

Iki, Yamamoto and Tomizawa [5] considered the quasi-diagonal exponent symmetry (QDES) model defined by

$$p_{ij} = \begin{cases} \alpha^i \beta^j d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

A special case of the QDES model obtained by putting  $\alpha = \beta$  is the DES model. Note that the QDES model implies the QS model. Let  $X$  and  $Y$  denote the row and column variables, respectively. We define the mean equality (ME) model as  $E(X) = E(Y)$ . Iki et al. [5] gave the theorem that the DES model holds if and only if both the QDES and ME models hold.

Iki et al. [5] described the relationship between the QDES model and a joint bivariate normal distribution, and showed that the QDES model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution with equal marginal variances. We are interested in considering the new model which is appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution without equal marginal variances, and a decomposition using the proposed models.

The present paper proposes two models, and gives the decomposition using the proposed models. Also it shows the orthogonality of the test statistics for decomposed model.

## 2. New Models

Consider a model defined by

$$p_{ij} = \begin{cases} \delta_1^{i+j} \delta_2^{i^2+j^2} d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

A special case of this model obtained by putting  $\delta_2 = 1$  is the DES model. Thus we shall refer to this model as the extended diagonal exponent symmetry (EDES) model. The EDES model states that  $p_{i+1, j+1}$  ( $i \neq j$ ) is  $\delta_1^2 \delta_2^{2(i+j+1)}$  times higher than  $p_{ij}$ ; in other words, for fixed distance from the main diagonal of the table, the ratio of  $p_{i+1, j+1}$  to  $p_{ij}$  increases (decreases) exponentially along every subdiagonal of the table. Note that the EDES model implies the S model.

Next, consider a model defined by

$$p_{ij} = \begin{cases} \alpha_1^i \alpha_2^{i^2} \beta_1^j \beta_2^{j^2} d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

A special case of this model obtained by putting  $\alpha_2 = \beta_2 = 1$  is the QDES model. Thus we shall refer to this model as the extended quasi-diagonal exponent symmetry (EQDES) model. A special case of the EQDES model obtained by putting  $\alpha_1 = \beta_1$  and  $\alpha_2 = \beta_2$  is the EDES model. The EQDES model states that  $p_{i+1,j+1}$  ( $i \neq j$ ) is  $\alpha_1 \alpha_2^{2i+1} \beta_1 \beta_2^{2j+1}$  times higher than  $p_{ij}$ ; in other words, for fixed distance from the main diagonal of the table, the ratio of  $p_{i+1,j+1}$  to  $p_{ij}$  increases (decreases) exponentially along every subdiagonal of the table. Note that the EQDES model implies the QS model.

Under the EQDES model, we can see

$$\frac{p_{i+1,j+1} / p_{ij}}{p_{j+1,i+1} / p_{ji}} = \left( \frac{\beta_2}{\alpha_2} \right)^{2d} \quad (i < j),$$

where  $d = j - i$ . This indicates that the odds that an observation will fall in the  $(i + 1, j + 1)$ th cell, instead of the  $(j + 1, i + 1)$ th cell is  $(\beta_2 / \alpha_2)^{2d}$  times higher than the odds that the observation will fall in the  $(i, j)$ th cell, instead of the  $(j, i)$ th cell. Also we can see

$$\frac{p_{ij}}{p_{ji}} = \xi_k^d \quad (i < j),$$

where  $k = i + j$  and  $\xi_k = (\beta_1 / \alpha_1)(\beta_2 / \alpha_2)^k$ . If  $\xi_k > 1$ , for corresponding  $i$  and  $j$ , the structure of  $p_{ij} > p_{ji}$  holds. Also if  $\xi_k < 1$ , the structure of  $p_{ij} < p_{ji}$  holds.

In Figure 1, we show the relationships among models. In figure,  $A \rightarrow B$  indicates that model A implies model B.

### 3. Decomposition

Refer to model of equality of marginal means and variances, i.e.,  $E(X) = E(Y)$  and  $\text{Var}(X) = \text{Var}(Y)$ , as the MVE model. This model is also expressed as  $E(X) = E(Y)$  and  $E(X^2) = E(Y^2)$ . We obtain the decomposition of the EDES model as follows:

Theorem 1. *The EDES model holds if and only if the EQDES and MVE models hold.*

Proof. If the EDES model holds, then the EQDES and MVE models hold. Assuming that both the EQDES and MVE models hold, then we shall show that the EDES model holds. Let  $\{p_{ij}^*\}$  denote the cell probabilities which satisfy both the EQDES and MVE models. Since the EQDES model holds, we see

$$\log p_{ij}^* = i \log \alpha_1 + i^2 \log \alpha_2 + j \log \beta_1 + j^2 \log \beta_2 + \log d_{|j-i|} \quad (i = 1, \dots, R; j = 1, \dots, R). \tag{1}$$

Let  $\pi_{ij} = d_{|j-i|} / c$  with  $c = \sum_{i=1}^R \sum_{j=1}^R d_{|j-i|}$ . We denote that  $\sum_{i=1}^R \sum_{j=1}^R \pi_{ij} = 1$  with  $0 < \pi_{ij} < 1$ . Then, since  $\{p_{ij}^*\}$  satisfy the EQDES and MVE models, we see

$$\log \left( \frac{p_{ij}^*}{\pi_{ij}} \right) = \log c + i \log \alpha_1 + i^2 \log \alpha_2 + j \log \beta_1 + j^2 \log \beta_2, \tag{2}$$

and

$$\mu_1^* = \mu_2^*, \nu_1^* = \nu_2^*, \tag{3}$$

where  $\mu_1^* = \sum_{i=1}^R i p_{i,i}^*$ ,  $\mu_2^* = \sum_{i=1}^R i^2 p_{i,i}^*$ ,  $\nu_1^* = \sum_{i=1}^R i p_{i,i}^*$ ,  $\nu_2^* = \sum_{i=1}^R i^2 p_{i,i}^*$  with  $p_{i,i}^* = \sum_{t=1}^R p_{it}^*$  and  $p_{\cdot,i}^* = \sum_{s=1}^R p_{si}^*$ .

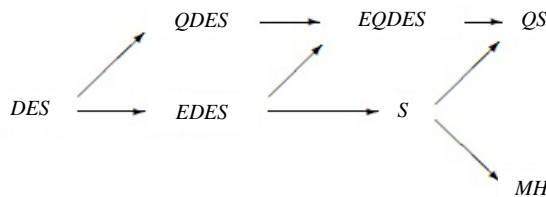


Figure 1. Relationships among models.

Then, we denote  $\mu_1^* (= \mu_2^*)$  by  $\mu_0$  and  $\nu_1^* (= \nu_2^*)$  by  $\nu_0$ .

Consider the arbitrary cell probabilities  $\{p_{ij}\}$  satisfying

$$\tilde{\mu}_1 = \tilde{\mu}_2 = \mu_0, \quad \tilde{\nu}_1 = \tilde{\nu}_2 = \nu_0, \tag{4}$$

where  $\tilde{\mu}_1 = \sum_{i=1}^R i p_i$ ,  $\tilde{\mu}_2 = \sum_{i=1}^R i p_{\cdot i}$ ,  $\tilde{\nu}_1 = \sum_{i=1}^R i^2 p_i$  and  $\tilde{\nu}_2 = \sum_{i=1}^R i^2 p_{\cdot i}$ .

From (2), (3) and (4), we see

$$\sum_{i=1}^R \sum_{j=1}^R (p_{ij} - p_{ij}^*) \log \left( \frac{p_{ij}}{\pi_{ij}} \right) = 0. \tag{5}$$

Using the Equation (5), we obtain

$$K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{p_{ij}^*\}, \{\pi_{ij}\}) + K(\{p_{ij}\}, \{p_{ij}^*\}),$$

where

$$K(\{a_{ij}\}, \{b_{ij}\}) = \sum_{i=1}^R \sum_{j=1}^R a_{ij} \log \left( \frac{a_{ij}}{b_{ij}} \right),$$

and  $K(\{a_{ij}\}, \{b_{ij}\})$  is the Kullback-Leibler information between  $\{a_{ij}\}$  and  $\{b_{ij}\}$ . Since  $\{\pi_{ij}\}$  being a function of  $\{p_{ij}^*\}$  is fixed, we see

$$\min_{\{p_{ij}\}} K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{p_{ij}^*\}, \{\pi_{ij}\}),$$

and then  $\{p_{ij}^*\}$  uniquely minimizes  $K(\{p_{ij}\}, \{\pi_{ij}\})$  (see Bhapkar and Darroch [6]).

Let  $p_{ij}^{**} = p_{ji}^*$  for  $1 \leq i, j \leq R$ . Then

$$\log p_{ij}^{**} = \log p_{ji}^* = j \log \alpha_1 + j^2 \log \alpha_2 + i \log \beta_1 + i^2 \log \beta_2 + \log d_{|j-i|}. \tag{6}$$

Noting that  $\{\pi_{ij} = \pi_{ji}\}$ , the Equation (6) is also expressed as

$$\log \left( \frac{p_{ij}^{**}}{\pi_{ij}} \right) = \log c + j \log \alpha_1 + j^2 \log \alpha_2 + i \log \beta_1 + i^2 \log \beta_2. \tag{7}$$

From (3), (4) and (7), we see

$$\sum_{i=1}^R \sum_{j=1}^R (p_{ij} - p_{ij}^{**}) \log \left( \frac{p_{ij}}{\pi_{ij}} \right) = 0. \tag{8}$$

Using the Equation (8), we obtain

$$K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{p_{ij}^{**}\}, \{\pi_{ij}\}) + K(\{p_{ij}\}, \{p_{ij}^{**}\}).$$

Since  $\{\pi_{ij}\}$  being a function of  $\{p_{ij}^{**}\}$  is fixed, we see

$$\min_{\{p_{ij}\}} K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{p_{ij}^{**}\}, \{\pi_{ij}\}),$$

and then  $\{p_{ij}^{**}\}$  uniquely minimizes  $K(\{p_{ij}\}, \{\pi_{ij}\})$ . Therefore, we see  $\{p_{ij}^* = p_{ij}^{**}\}$ . Thus,  $\{p_{ij}^* = p_{ji}^*\}$ .

From (1) and (6), for  $i < j$ , we see

$$\begin{aligned} \log \left( \frac{p_{ij}^*}{p_{ji}^*} \right) &= (i - j) \log \alpha_1 + (i^2 - j^2) \log \alpha_2 + (j - i) \log \beta_1 + (j^2 - i^2) \log \beta_2 \\ &= (j - i) \left( \log \frac{\beta_1}{\alpha_1} + (j + i) \log \frac{\beta_2}{\alpha_2} \right) = 0. \end{aligned}$$

Thus, we obtain  $\alpha_1 = \beta_1$  and  $\alpha_2 = \beta_2$ . Namely, the EDES model holds. The proof is completed.

#### 4. Orthogonality of Test Statistics

Let  $n_{ij}$  denote the observed frequency in the  $(i, j)$ th cell of the table  $(i = 1, \dots, R; j = 1, \dots, R)$  with  $n = \sum \sum n_{ij}$ , and let  $m_{ij}$  denote the corresponding expected frequency. Assume that  $\{n_{ij}\}$  have a multinomial distribution. The maximum likelihood estimates (MLEs) of  $\{m_{ij}\}$  under the EDES and EQDES models could be obtained using iterative procedures; for example, see Darroch and Ratcliff [7]. The MLEs of  $\{m_{ij}\}$  under the MVE model could be obtained using Newton-Raphson method to the log-likelihood equations.

Let  $G^2(M)$  denote the likelihood ratio chi-squared statistic for testing goodness-of-fit model  $M$ . The numbers of degrees of freedom (df) for the EDES and EQDES models are  $R^2 - 2R - 1$  and  $R^2 - 2R - 3$ , respectively.

The orthogonality (asymptotic separability or independence) of the test statistics for goodness-of-fit of two models is discussed by, e.g., Darroch and Silvey [8] and Read [9]. We obtain as follow:

Theorem 2. *The test statistic  $G^2(\text{EDES})$  is asymptotically equivalent to the sum of  $G^2(\text{EQDES})$  and  $G^2(\text{MVE})$ .*

Proof. The EQDES model is expressed as

$$\log p_{ij} = i\beta_1^* + j\beta_2^* + i^2\beta_3^* + j^2\beta_4^* + \phi_{ij} \quad (i = 1, \dots, R; j = 1, \dots, R), \tag{9}$$

where  $\phi_{ij} = \gamma_{|j-i|}$  ( $i \neq j$ ). Let

$$p = (p_{11}, \dots, p_{1R}, p_{21}, \dots, p_{2R}, \dots, p_{R1}, \dots, p_{RR})^t,$$

$$\beta^* = (\beta_1^*, \beta_2^*, \beta_3^*, \beta_4^*, \beta_5^*)^t,$$

where “ $t$ ” denotes the transpose, and

$$\beta_5^* = (\gamma_1, \gamma_2, \dots, \gamma_{R-1}, \phi_{11}, \phi_{22}, \dots, \phi_{RR}),$$

is the  $1 \times (2R - 1)$  vector. The EQDES model is expressed as

$$\log p = X\beta^* = (X_1, X_2, X_3, X_4, X_5)\beta^*,$$

where  $X$  is the  $R^2 \times K$  matrix with  $K = 2R + 3$ ,  $X_1 = J_R \otimes 1_R$  (the  $R^2 \times 1$  vector),  $X_2 = 1_R \otimes J_R$  (the  $R^2 \times 1$  vector),  $X_3 = J_R^2 \otimes 1_R$  (the  $R^2 \times 1$  vector),  $X_4 = 1_R \otimes J_R^2$  (the  $R^2 \times 1$  vector), and  $X_5$  is the  $R^2 \times (2R - 1)$  matrix of 1 or 0 elements determined from (9),  $1_s$  is the  $s \times 1$  vector of 1 elements,  $J_R = (1, \dots, R)^t$ ,  $J_R^2 = (1^2, \dots, R^2)^t$ , and  $\otimes$  denotes the Kronecker product. The matrix  $X$  is full column rank which is  $K$ . In a similar manner to Haber [10], we denote the linear space spanned by the columns of the matrix  $X$  by  $S(X)$  with the dimension  $K$ .

Let  $U$  be an  $R^2 \times l_1$ , where  $l_1 = R^2 - K = R^2 - 2R - 3$ , full column rank matrix such that  $S(U)$  is the orthogonal complement of  $S(X)$ . Thus,  $U^t X = O_{l_1, K}$ , where  $O_{s,t}$  is the  $s \times t$  zero matrix. Therefore the EQDES model is expressed as

$$h_1(p) = 0_{l_1},$$

where  $0_s$  is the  $s \times 1$  zero vector, and  $h_1(p) = U^t \log p$ . The MVE model is expressed as

$$h_2(p) = 0_{l_2},$$

where  $l_2 = 2$ , and  $h_2(p) = Wp$ , with  $W = (\{1_R \otimes J_R - J_R \otimes 1_R\}, \{1_R \otimes J_R^2 - J_R^2 \otimes 1_R\})^t$  being the  $2 \times R^2$  matrix. Namely,  $W^t = (X_2 - X_1, X_4 - X_3)$ . Thus  $W^t$  belongs to  $S(X)$ . Hence  $WU = O_{l_2, l_1}$ . From Theorem 1, the EDES model is expressed as

$$h_3(p) = 0_{l_3},$$

where  $l_3 = l_1 + l_2 = R^2 - 2R - 1$ , and  $h_3 = (h_1^t, h_2^t)^t$ .

Let  $H_s(p) (s=1,2,3)$  denote the  $l_s \times R^2$  matrix of partial derivative of  $h_s(p)$  with respect to  $p$ , i.e.,  $H_s(p) = \partial h_s(p) / \partial p^t$ . Let  $\Sigma(p) = \text{diag}(p) - pp^t$ , where  $\text{diag}(p)$  denotes a diagonal matrix with  $i$ th component of  $p$  as  $i$ th diagonal component. Let  $\hat{p}$  denote  $p$  with  $\{p_{ij}\}$  replaced by  $\{\hat{p}_{ij} = n_{ij}/n\}$ . Then  $\sqrt{n}(\hat{p} - p)$  has asymptotically a normal distribution with mean  $0_{R^2}$  and covariance matrix  $\Sigma(p)$ . Using the delta method,  $\sqrt{n}(h_s(\hat{p}) - h_s(p))$  has asymptotically a normal distribution with mean  $0_{l_s}$  and covariance matrix

$$H_3(p)\Sigma(p)H_3(p)^t = \begin{bmatrix} H_1(p)\Sigma(p)H_1(p)^t & H_1(p)\Sigma(p)H_2(p)^t \\ H_2(p)\Sigma(p)H_1(p)^t & H_2(p)\Sigma(p)H_2(p)^t \end{bmatrix}.$$

Note that  $1_{R^2}$  belongs to  $S(X)$  because  $X_5 1_{2R-1} = 1_{R^2}$ . Thus  $H_1(p)p = U^t 1_{R^2} = 0_{l_1}$ . Since  $H_1(p)\text{diag}(p) = U^t$  and  $H_2(p) = W$ , we see

$$H_1(p)\Sigma(p)H_2(p)^t = U^t W^t = O_{l_1, l_2}.$$

Thus, we obtain  $\Delta_3(p) = \Delta_1(p) + \Delta_2(p)$ , where

$$\Delta_s(p) = h_s(p)^t \left[ H_s(p)\Sigma(p)H_s(p)^t \right]^{-1} h_s(p). \tag{10}$$

Under each  $h_s(p) = 0_{l_s} (s=1,2,3)$ , the Wald statistic  $W_s = n\Delta_s(\hat{p})$  has asymptotically a chi-squared distribution with  $l_s$  degrees of freedom. From (10), we see that  $W_3 = W_1 + W_2$ . From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic, we obtain Theorem 2.

### 5. Examples

Example 1. Consider the data in **Table 1**, taken from Bishop, Fienberg and Holland [11], which describe the cross-classification of father's and son's occupational status categories in Denmark. The row is the father's status category and column is the son's status category. The categories are ordered from (1) to (5) (high to low). These data have also been analyzed by some statisticians; see for example, Kullback [12], Haberman [13], Goodman [14], and Yamamoto, Tahata and Tomizawa [15].

**Table 1.** Occupational status for Danish father-son pairs; from Bishop *et al.* [11]. (The parenthesized values are MLEs of expected frequencies under the EQDES model.)

Father's status	Son's status					Total
	(1)	(2)	(3)	(4)	(5)	
(1)	18 (18.00)	17 (21.13)	16 (18.26)	4 (6.85)	2 (4.49)	57
(2)	24 (22.01)	105 (105.00)	109 (95.43)	59 (59.12)	21 (15.91)	318
(3)	23 (18.50)	84 (92.83)	289 (289.00)	217 (206.96)	95 (91.94)	708
(4)	8 (6.31)	49 (52.26)	175 (188.06)	348 (348.00)	198 (215.58)	778
(5)	6 (3.51)	8 (11.93)	69 (70.92)	201 (183.00)	246 (246.00)	530
Total	79	263	658	829	562	2391

Note: Status (1) is high professionals, (2) white-collar employees of higher education, (3) white-collar employees of less high education, (4) upper working class, and (5) unskilled workers.

We see from **Table 3** that the EQDES and QS models fit these data well, although the other models fit poorly. The EQDES model is a special case of the QS model. We shall test the hypothesis that the EQDES model holds assuming that the QS model holds for these data. Since  $G^2(EQDES | QS) = G^2(EQDES) - G^2(QS) = 11.52$  with 6 df being the difference between the numbers of df for the EQDES and the QS models, this hypothesis is accepted at the 0.05 significance level. Therefore, the EQDES model would be preferable to the QS model.

Under the EQDES model, the MLEs of  $\alpha_1, \alpha_2, \beta_1$  and  $\beta_2$  are  $\hat{\alpha}_1 = 4.67, \hat{\alpha}_2 = 0.82, \hat{\beta}_1 = 4.05$  and  $\hat{\beta}_2 = 0.85$ , respectively. Therefore the probability that a father's and his son's status categories are  $i+1$  and  $j+1$ , respectively, is estimated to be  $\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1} = 13.11 \times (0.67)^i \times (0.72)^j$  times higher than the probability that those are  $i$  and  $j$ , respectively ( $i = 1, 2, 3, 4; j = 1, 2, 3, 4; i \neq j$ ). Since the values of  $\{\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1}\}$  for  $i \neq j$  and  $(i, j) \neq (4, 3)$  are greater than 1 and it for  $(i, j) = (4, 3)$  is less than 1 (see **Table 2**), the probability that a father's and his son's status categories are  $i+1$  and  $j+1$ , respectively, is estimated to be greater than the probability that those are  $i$  and  $j$ , respectively ( $i = 1, 2, 3, 4; j = 1, 2, 3, 4; i \neq j, (i, j) \neq (4, 3)$ ).

Also the MLEs of  $\xi_k$  ( $k = 3, \dots, 9$ ) are  $\hat{\xi}_3 = 0.96, \hat{\xi}_4 = 0.99, \hat{\xi}_5 = 1.03, \hat{\xi}_6 = 1.06, \hat{\xi}_7 = 1.10, \hat{\xi}_8 = 1.14$  and  $\hat{\xi}_9 = 1.18$ , respectively. Therefore, it is estimated that there is the structure of  $\{p_{ij} > p_{ji}\}$  for  $i < j$  with  $i + j = 5, \dots, 9$  and  $\{p_{ij} < p_{ji}\}$  for  $i < j$  with  $i + j = 3$  and 4.

We see from **Table 3** that the poor fit of the EDES model is caused by the influence of the lack of structure of the MVE model rather than the EQDES model.

Example 2. Consider the data in **Table 4** taken from Tomizawa [16]. These data are an unaided distance vision of 3168 pupils comprising nearly equal number of boys and girls aged 6 - 12 at elementary schools in Tokyo, Japan, examined in June 1984. These data have also been analyzed by Tomizawa [1], Tahata and Tomizawa [17], and Iki *et al.* [5]. The row is the right eye grade and column is the left eye grade.

We see from **Table 3** that the EDES and EQDES models fit these data well, although the MVE model fits poorly. The EDES model is a special case of the EQDES model. We shall test the hypothesis that the EDES

**Table 2.** Values of  $\{\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1}\}$ ,  $i \neq j$ , under the EQDES model applied to **Table 1**.

	$j=1$	2	3	4
$i=1$	-	4.52	3.24	2.32
2	4.22	-	2.17	1.56
3	2.83	2.03	-	1.04
4	1.89	1.36	0.97	-

**Table 3.** Likelihood ratio chi-squared values  $G^2$  for models applied to **Table 1** and **Table 4**.

Applied models	Table 1		Table 4	
	df	$G^2$	df	$G^2$
S	10	24.80*	6	9.69
QS	6	6.47	3	2.81
MH	4	18.26*	3	6.87
DES	15	131.45*	8	10.18
QDES	14	125.70*	7	8.71
ME	1	5.73*	1	1.46
EDES	14	25.48*	7	9.70
EQDES	12	17.99	5	3.12
MVE	2	7.47*	2	6.57*

\* means significant at the 0.05 level.

model holds assuming that the EQDES model holds for these data. Since

$$G^2(EDES | EQDES) = G^2(EDES) - G^2(EQDES) = 6.58$$

with 2 df being the difference between the numbers of df for the EDES and the EQDES models, this hypothesis is rejected at the 0.05 significance level. Therefore, the EQDES model would be preferable to the EDES model.

Under the EQDES model, the MLEs of  $\alpha_1, \alpha_2, \beta_1$  and  $\beta_2$  are  $\hat{\alpha}_1 = 0.34, \hat{\alpha}_2 = 1.13, \hat{\beta}_1 = 0.72$  and  $\hat{\beta}_2 = 0.97$ , respectively. Therefore the probability that a pupil's right eye grade and his or her left eye grade are  $i+1$  and  $j+1$ , respectively, is estimated to be  $\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1} = 0.26 \times (1.27)^i \times (0.94)^j$  times higher than the probability that those are  $i$  and  $j$ , respectively ( $i = 1, 2, 3; j = 1, 2, 3; i \neq j$ ). Since all values of  $\{\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1}\}$ ,  $i \neq j$ , are less than 1 (see **Table 5**), the probability that a pupil's right eye grade and his or her left eye grade are  $i+1$  and  $j+1$ , respectively, is estimated to be less than the probability that those are  $i$  and  $j$ , respectively ( $i = 1, 2, 3; j = 1, 2, 3; i \neq j$ ).

Also the MLEs of  $\xi_k$  ( $k = 3, \dots, 7$ ) are  $\hat{\xi}_3 = 1.35, \hat{\xi}_4 = 1.16, \hat{\xi}_5 = 0.99, \hat{\xi}_6 = 0.85$  and  $\hat{\xi}_7 = 0.73$ , respectively. Therefore, it is estimated that there is the structure of  $\{p_{ij} > p_{ji}\}$  for  $i < j$  with  $i + j = 3$  and 4 and  $\{p_{ij} < p_{ji}\}$  for  $i < j$  with  $i + j = 5, 6$  and 7.

### 6. Simulation Studies

Under the QDES model, we see the structure of  $p_{ij}/p_{ji} = (\beta/\alpha)^{j-i}$  which is the structure of Agresti's [18]

**Table 4.** Unaided distance vision of 3168 pupils comprising nearly equal number of boys and girls aged 6 - 12 at elementary schools in Tokyo, Japan, examined in June 1984; from Tomizawa [16]. (Upper and lower parenthesized values are MLEs of expected frequencies under the EDES and EQDES models, respectively.)

Right eye grade	Left eye grade				Total
	Best (1)	Second (2)	Third (3)	Worst (4)	
Best (1)	2470 (2470.00) (2470.00)	126 (110.80) (127.26)	21 (15.65) (17.88)	10 (11.00) (10.86)	2627
Second (2)	96 (110.80) (94.48)	138 (138.00) (138.00)	33 (37.72) (37.45)	5 (5.85) (4.92)	272
Third (3)	10 (15.65) (13.39)	42 (37.72) (37.77)	75 (75.00) (75.00)	15 (15.48) (13.10)	142
Worst (4)	12 (11.00) (11.14)	7 (5.85) (6.80)	16 (15.48) (17.95)	92 (92.00) (92.00)	127
Total	2588	313	145	122	3168

**Table 5.** Values of  $\{\hat{\alpha}_1 \hat{\alpha}_2^{2i+1} \hat{\beta}_1 \hat{\beta}_2^{2j+1}\}$ ,  $i \neq j$ , under the EQDES model applied to **Table 4**.

	$j=1$	2	3
$i=1$	-	0.29	0.28
2	0.40	-	0.35
3	0.51	0.48	-



linear diagonals-parameter symmetry model, and under the EQDES model, we see the structure of  $p_{ij}/p_{ji} = (\beta_1/\alpha_1)^{j-i} (\beta_2/\alpha_2)^{j-i^2}$  which is the structure of Tomizawa's [19] extended linear diagonals-parameter symmetry model. Also under the DES and EDES models, we see the structure of  $p_{ij}/p_{ji} = 1$  for  $i < j$ .

Consider now random variables  $U$  and  $V$  having a joint bivariate normal distribution with means  $E(U) = \mu_1$  and  $E(V) = \mu_2$ , variances  $\text{Var}(U) = \sigma_1^2$  and  $\text{Var}(V) = \sigma_2^2$ , and correlation  $\text{Corr}(U, V) = \rho$ . Then the joint bivariate normal density function  $f(u, v)$  satisfies

$$\frac{f(u, v)}{f(v, u)} = \exp \left[ \frac{1}{2(1-\rho^2)} \left\{ \frac{2\rho}{\sigma_1\sigma_2} (\mu_2 - \mu_1) + 2 \left( \frac{\mu_2}{\sigma_2^2} - \frac{\mu_1}{\sigma_1^2} \right) \right\} (v - u) \right] \cdot \exp \left[ \frac{1}{2(1-\rho^2)} \left( \frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2} \right) (v^2 - u^2) \right] \quad (u < v).$$

Namely,  $f(u, v)/f(v, u)$  has the form  $\zeta^{v-u} \eta^{v^2-u^2}$  for constant  $\zeta$  and  $\eta$ . Agresti [18] described relationship between the linear diagonals-parameter symmetry model and the joint bivariate normal distribution (see also Tomizawa [19]). We now consider the relationship between the QDES (DES) and EQDES (EDES) models and the joint bivariate normal distribution in terms of simulation studies.

**Table 6** gives the  $4 \times 4$  tables of sample size 5000 formed by using cut points for each variable at  $\mu_1, \mu_1 \pm 0.7\sigma_1$ , for underlying bivariate normal distribution with the conditions  $\rho = 0.3$ , and  $\mu_2 - \mu_1 = 0$  and  $\sigma_1^2 = \sigma_2^2$  (**Table 6(a)**),  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_1^2 = \sigma_2^2$  (**Table 6(b)**),  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_2^2 = 0.8\sigma_1^2$  (**Table 6(c)**) and  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_2^2 = 1.2\sigma_1^2$  (**Table 6(d)**).

**Table 6.** The  $4 \times 4$  tables of sample size 5000, formed by using cut points for each variable at  $\mu_1, \mu_1 \pm 0.7\sigma_1$ , from an underlying bivariate normal distribution with the conditions  $\rho = 0.3$  and (a)  $\mu_2 - \mu_1 = 0$  and  $\sigma_1^2 = \sigma_2^2$ , (b)  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_1^2 = \sigma_2^2$ , (c)  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_2^2 = 0.8\sigma_1^2$ , (d)  $\mu_2 - \mu_1 = 0.5$  and  $\sigma_2^2 = 1.2\sigma_1^2$ .

(a) $\mu_2 - \mu_1 = 0, \sigma_1^2 = \sigma_2^2$			
442	335	244	143
321	351	352	258
245	321	357	331
159	285	364	492
(b) $\mu_2 - \mu_1 = 0.5, \sigma_1^2 = \sigma_2^2$			
272	336	319	237
203	285	367	427
145	228	376	505
82	183	369	666
(c) $\mu_2 - \mu_1 = 0.5, \sigma_2^2 = 0.8\sigma_1^2$			
247	382	368	282
176	317	424	385
115	271	414	461
65	162	353	578
(d) $\mu_2 - \mu_1 = 0.5, \sigma_2^2 = 1.2\sigma_1^2$			
318	304	277	255
224	269	341	460
168	209	361	511
106	184	326	687

**Table 7.** Likelihood ratio chi-squared values  $G^2$  for models applied to **Tables 6(a)-6(d)**.

Applied		$G^2$			
models	df	Table 6(a)	Table 6(b)	Table 6(c)	Table 6(d)
DES	8	6.74	341.67*	506.78*	315.56*
EDES	7	6.73	333.26*	506.40*	302.20*
QDES	7	5.78	13.81	25.97*	34.92*
EQDES	5	3.60	4.59	1.98	9.74

\*means significant at the 0.05 level.

We see from **Table 7** that the EQDES model fits well for each of **Tables 6(a)-6(d)**, although the QDES model fits well for each of **Table 6(a)** and **Table 6(b)**, and fits poorly for each of **Table 6(c)** and **Table 6(d)**. The DES and EDES models fit well for **Table 6(a)** and fit poorly for each of **Tables 6(b)-6(d)**. Thus the EQDES model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution (without the equality of marginal variances), although the QDES model may be appropriate if it is reasonable to assume it with equal marginal variances, and the DES and EDES models may be appropriate if it is reasonable to assume it with both equal marginal means and equal marginal variances.

## 7. Concluding Remarks

Theorem 1 may be useful for seeing the reason for the poor fit when the EDES model fits the data poorly; in fact, see from Example 1, a poor fit of the EDES model would be caused by a poor fit of the MVE model rather than the EQDES model.

From Theorem 2, we point out that the  $G^2(MVE)$  can be easily calculated using the  $G^2(EDES)$  and  $G^2(EQDES)$ ; in fact, see from **Table 3**, the value of  $G^2(EDES)$  is very close to the value of the sum of  $G^2(EQDES)$  and  $G^2(MVE)$ .

From Simulation studies, the EQDES model may be appropriate for a square ordinal table if it is reasonable to assume an underlying bivariate normal distribution without equal marginal means and equal marginal variances; although the QDES model may be appropriate if it is reasonable to assume it with equal marginal variances.

## Acknowledgements

The authors would like to thank the editor and the referee for their helpful comments.

## References

- [1] Tomizawa, S. (1992) A Model of Symmetry with Exponents along Every Subdiagonal and Its Application to Data on Unaided Vision of Pupils at Japanese Elementary Schools. *Journal of Applied Statistics*, **19**, 509-512. <http://dx.doi.org/10.1080/02664769200000046>
- [2] Bowker, A.H. (1948) A Test for Symmetry in Contingency Tables. *Journal of the American Statistical Association*, **43**, 572-574. <http://dx.doi.org/10.1080/01621459.1948.10483284>
- [3] Caussinus, H. (1965) Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse*, **29**, 77-182.
- [4] Stuart, A. (1955) A Test for Homogeneity of the Marginal Distributions in a Two-Way Classification. *Biometrika*, **42**, 412-416. <http://dx.doi.org/10.1093/biomet/42.3-4.412>
- [5] Iki, K., Yamamoto, K. and Tomizawa, S. (2014) Quasi-Diagonal Exponent Symmetry Model for Square Contingency Tables with Ordered Categories. *Statistics and Probability Letters*, **92**, 33-38. <http://dx.doi.org/10.1016/j.spl.2014.04.029>
- [6] Bhapkar, V.P. and Darroch, J.N. (1990) Marginal Symmetry and Quasi Symmetry of General Order. *Journal of Multivariate Analysis*, **34**, 173-184. [http://dx.doi.org/10.1016/0047-259X\(90\)90034-F](http://dx.doi.org/10.1016/0047-259X(90)90034-F)
- [7] Darroch, J.N. and Ratcliff, D. (1972) Generalized Iterative Scaling for Log-Linear Models. *Annals of Mathematical Statistics*, **43**, 1470-1480. <http://dx.doi.org/10.1214/aoms/1177692379>

- [8] Darroch, J.N. and Silvey, S.D. (1963) On Testing More than One Hypothesis. *Annals of Mathematical Statistics*, **34**, 555-567. <http://dx.doi.org/10.1214/aoms/1177704168>
- [9] Read, C.B. (1977) Partitioning Chi-Square in Contingency Table: A Teaching Approach. *Communications in Statistics-Theory and Methods*, **6**, 553-562. <http://dx.doi.org/10.1080/03610927708827513>
- [10] Haber, M. (1985) Maximum Likelihood Methods for Linear and Log-Linear Models in Categorical Data. *Computational Statistics and Data Analysis*, **3**, 1-10. [http://dx.doi.org/10.1016/0167-9473\(85\)90053-2](http://dx.doi.org/10.1016/0167-9473(85)90053-2)
- [11] Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. The MIT Press, Cambridge.
- [12] Kullback, S. (1971) Marginal Homogeneity of Multidimensional Contingency Tables. *Annals of Mathematical Statistics*, **42**, 594-606. <http://dx.doi.org/10.1214/aoms/1177693409>
- [13] Haberman, S.J. (1974) *The Analysis of Frequency Data*. The University of Chicago Press, Chicago.
- [14] Goodman, L.A. (1981) Association Models and the Bivariate Normal for Contingency Tables with Ordered Categories. *Biometrika*, **68**, 347-355. <http://dx.doi.org/10.1093/biomet/68.2.347>
- [15] Yamamoto, K., Tahata, K. and Tomizawa, S. (2012) Some Symmetry Models for the Analysis of Collapsed Square Contingency Tables with Ordered Categories. *Calcutta Statistical Association Bulletin*, **64**, 21-36.
- [16] Tomizawa, S. (1985) Analysis of Data in Square Contingency Tables with Ordered Categories Using the Conditional Symmetry Model and Its Decomposed Models. *Environmental Health Perspectives*, **63**, 235-239. <http://dx.doi.org/10.1289/ehp.8563235>
- [17] Tahata, K. and Tomizawa, S. (2006) Decompositions for Extended Double Symmetry Models in Square Contingency Tables with Ordered Categories. *Journal of the Japan Statistical Society*, **36**, 91-106. <http://dx.doi.org/10.14490/jjss.36.91>
- [18] Agresti, A. (1983) A Simple Diagonals-Parameter Symmetry and Quasi-Symmetry Model. *Statistics and Probability Letters*, **1**, 313-316. [http://dx.doi.org/10.1016/0167-7152\(83\)90051-2](http://dx.doi.org/10.1016/0167-7152(83)90051-2)
- [19] Tomizawa, S. (1991) An Extended Linear Diagonals-Parameter Symmetry Model for Square Contingency Tables with Ordered Categories. *Metron*, **49**, 401-409.