Journal of Statistics: Advances in Theory and Applications
Volume 21, Number 2, 2019, Pages 139-155
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jsata_7100122063

GENERALIZED DIAGONAL EXPONENT CONDITIONAL SYMMETRY MODELS FOR SQUARE CONTINGENCY TABLES WITH ORDERED CATEGORIES

KIYOTAKA IKI[1], DAISUKE NEMOTO[2] and SADAO TOMIZAWA[3]

[1]Department of Economics
Nihon University
Chiyoda-ku, Tokyo
Japan
e-mail: iki.kiyotaka@nihon-u.ac.jp

[2,3]Department of Information Sciences
Faculty of Science and Technology
Tokyo University of Science
Noda City, Chiba, 278-8510
Japan
e-mail: tomizawa@is.noda.tus.ac.jp

Abstract

For square contingency tables with ordered categories, this paper proposes a generalized diagonal exponent conditional symmetry model which indicates that in addition to the structure of conditional symmetry of the probabilities with respect to the main diagonal of the table, the log-odds of adjacent two

2010 Mathematics Subject Classification: 62H17.
Keywords and phrases: diagonal exponent symmetry, ordinal category, orthogonal decomposition, quasi-symmetry, square contingency table.
Received May 16, 2019

probabilities along subdiagonal of the table is the sum of polynomial of row value and polynomial of column value with same coefficients. This paper also gives the decomposition using the proposed model.

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

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

where $\psi_{ij} = \psi_{ji}$; see Bowker [3]. This describes a structure of symmetry of the probabilities $\{p_{ij}\}$ with respect to the main diagonal of the table. McCullagh [11] considered the conditional symmetry (CS) model, defined by

$$p_{ij} = \begin{cases} \gamma\psi_{ij} & (i < j), \\ \psi_{ij} & (i \geq j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. The CS model states that $p_{ij} (i < j)$ is γ times higher than p_{ji}. A special case of the CS model obtained by putting $\gamma = 1$ is the S model. The global symmetry (GS) model is defined by

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

see Read [12]. The GS model states that the probability that an observation will fall in one of the upper-right triangle cells above the main diagonal of the table is equal to the probability that it falls in one of the lower-left triangle cells below the main diagonal. Read [12] gave the theorem that the S model holds if and only if both the CS and GS models hold.

Caussinus [4] considered the quasi-symmetry (QS) model, defined by

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

where $\psi_{ij} = \psi_{ji}$. A special case of the QS model with $\{\alpha_i = \beta_i\}$ is the S model. Iki et al. [8] considered the k-th diagonal exponent symmetry $(\text{DES}(k))$ model, for a fixed k $(k = 1, \ldots, R-1)$, defined by

$$p_{ij} = \begin{cases} \prod_{t=1}^{k} \delta_t^{i^t + j^t} d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{1}$$

Note that (1) with $k = 1$ is the diagonal exponent symmetry (DES) model in Tomizawa [13]. The $\text{DES}(k)$ model states that, in addition to the structure of the S model, for fixed distance from the main diagonal of the table, the log-odds of $p_{i+1, j+1}$ to p_{ij} is the sum of polynomial of row value i and polynomial of column value j with same coefficients along every subdiagonal of the table (especially, when $k = 2$, the log-odds of them is a linear function of $i + j$). Note that the $\text{DES}(k)$ model implies the S model. Iki et al. [8] also considered the k-th quasi-diagonal exponent symmetry $(\text{QDES}(k))$ model, for a fixed k $(k = 1, \ldots, R-1)$, defined by

$$p_{ij} = \begin{cases} \prod_{t=1}^{k} \alpha_t^{i^t} \beta_t^{j^t} d_{|j-i|} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{2}$$

Note that (2) with $\alpha_t = \beta_t$ $(t = 1, \ldots, k)$ is the $\text{DES}(k)$ model, and (2) with $k = 1$ is the quasi-diagonal exponent symmetry (QDES) model in Iki et al. [10]. Note that the $\text{QDES}(k)$ model implies the QS model. Let X and Y denote the row and column variables, respectively. For a fixed k $(k = 1, \ldots, R-1)$, consider a model defined by $E(X^t) = E(Y^t)$ $(t = 1,$

$\dots, k$). We shall refer to this model as the k-th marginal moment equality (MME(k)) model. For a fixed k ($k = 1, \dots, R-1$), Iki et al. [8] gave the theorem that the DES(k) model holds if and only if both QDES(k) and MME(k) models hold.

Iki et al. [9] considered the diagonal exponent conditional symmetry (DECS) model, defined by

$$p_{ij} = \begin{cases} \delta^{i+j} d_{j-i} & (i \neq j), \\ \psi_{ii} & (i = j), \end{cases} \quad (3)$$

where $d_{j-i} = \gamma d_{i-j}$ $(i < j)$. Note that (3) with $\gamma = 1$ is the DES model. This model states that in addition to the structure of the CS model, $p_{i+1, j+1}$ $(i \neq j)$ is δ^2 times higher than p_{ij}. Under the DECS model, we see the structure of $p_{ij} / p_{ji} = \gamma$ $(i < j)$. Iki et al. [9] also considered the quasi-diagonal exponent conditional symmetry (QDECS) model, defined by

$$p_{ij} = \begin{cases} \alpha^{i} \beta^{j} d_{j-i} & (i \neq j), \\ \psi_{ii} & (i = j), \end{cases} \quad (4)$$

where $d_{j-i} = \gamma d_{i-j}$ $(i < j)$. Note that (4) with $\alpha = \beta$ is the DECS model, and (4) with $\gamma = 1$ is the QDES model. Under the QDECS model, we see the structure of $p_{ij} / p_{ji} = \gamma(\beta / \alpha)^{j-i}$ $(i < j)$. Iki et al. [9] gave the theorems as follows:

Theorem 1. *The DES model holds if and only if both the DECS and GS models hold.*

Theorem 2. *The DES model holds if and only if all the QDECS, GS, and MME*(1) *models hold.*

Theorem 3. *The DES model holds if and only if both the DECS and MME*(1) *models hold.*

We are now interested in considering the generalization of the DECS and QDECS models and Theorems 1, 2, and 3. The present paper proposes a generalized DECS and QDECS models, and gives the decomposition of the DES(k) model. It also shows the orthogonality of the test statistics for decomposed models.

2. New Models

We consider a generalized DECS model, for a fixed k $(k = 1, \dots, R-1)$, defined by

$$p_{ij} = \begin{cases} \prod_{t=1}^{k} \delta_t^{i^t + j^t} d_{j-i} & (i \neq j), \\ \psi_{ii} & (i = j), \end{cases}$$

where $d_{j-i} = \gamma d_{i-j}$ $(i < j)$. We shall refer to this model as the k-th diagonal exponent conditional symmetry (DECS(k)) model. Under the DECS(k) model, we see the structure of $p_{ij} / p_{ji} = \gamma$ $(i < j)$. Note that the DECS(k) model implies the CS model. The DECS(1) model is equivalent to the DECS model. A special case of the DECS(k) model with $\gamma = 1$ is the DES(k) model.

Moreover, consider a generalized QDECS model, for a fixed k $(k = 1, \dots, R-1)$, defined by

$$p_{ij} = \begin{cases} \prod_{t=1}^{k} \alpha_t^{i^t} \beta_t^{j^t} d_{j-i} & (i \neq j), \\ \psi_{ii} & (i = j), \end{cases}$$

where $d_{j-i} = \gamma d_{i-j}$ $(i < j)$. We shall refer to this model as the k-th quasi-diagonal exponent conditional symmetry (QDECS(k)) model. Under the

QDECS(k) model, we see the structure of $p_{ij} / p_{ji} = \gamma \prod_{t=1}^{k} (\beta_t / \alpha_t)^{j^t - i^t}$ $(i < j)$. The QDECS(1) model is equivalent to the QDECS model. A special case of the QDECS(k) model with $\alpha_t = \beta_t$ $(t = 1, \ldots, k)$ is the DECS(k) model. A special case with $\gamma = 1$ is the QDES(k) model. Also, a special case with $\alpha_t = \beta_t$ $(t = 1, \ldots, k)$ and $\gamma = 1$ is the DES(k) model.

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

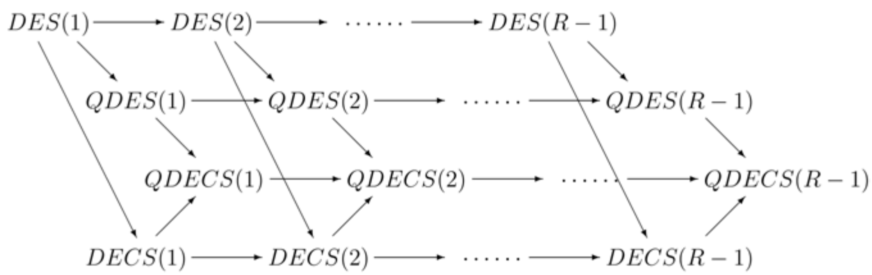

Figure 1. Relationships among models.

3. Decomposition and Orthogonality of Test Statistics

We obtain the decompositions of the DES(k) model as follows:

Theorem 4. *For a fixed* k $(k = 1, \ldots, R-1)$, *the* DES(k) *model holds if and only if both the* DECS(k) *and GS, models hold.*

Theorem 5. *For a fixed* k $(k = 1, \ldots, R-1)$, *the* DES(k) *model holds if and only if all the* QDECS(k), *GS, and* MME(k) *models hold.*

Theorem 6. *For a fixed* k $(k = 1, \ldots, R-1)$, *the* DES(k) *model holds if and only if both the* DECS(k) *and* MME(1) *models hold.*

Proof of Theorem 4. For a fixed k $(k = 1, \ldots, R-1)$, if the DES(k) model holds, then the DECS(k) and GS models hold. Assuming that both the DECS(k) and GS models hold, then we shall show that the DES(k) model holds. Since the DECS(k) and GS models hold, we see

$$\begin{aligned}\sum\sum_{s<t} p_{st} - \sum\sum_{s<t} p_{ts} &= \sum\sum_{s<t} \Big(\prod_{l=1}^{k} \delta_l^{s^l + t^l}\Big) d_{t-s} - \sum\sum_{s<t} \Big(\prod_{l=1}^{k} \delta_l^{s^l + t^l}\Big) d_{s-t} \\ &= \sum\sum_{s<t} \Big(\prod_{l=1}^{k} \delta_l^{s^l + t^l}\Big) \gamma d_{s-t} - \sum\sum_{s<t} \Big(\prod_{l=1}^{k} \delta_l^{s^l + t^l}\Big) d_{s-t} \\ &= (\gamma - 1) \sum\sum_{s<t} \Big(\prod_{l=1}^{k} \delta_l^{s^l + t^l}\Big) d_{s-t} \\ &= 0.\end{aligned}$$

Thus, we obtain $\gamma = 1$. Namely, the DES(k) model holds. The proof is complicated.

Proof of Theorem 5. For a fixed k $(k = 1, \ldots, R-1)$, if the DES(k) model holds, then the QDECS(k), GS, and MME(k) models hold. Assuming that all the QDECS(k), GS, and MME(k) models hold, then we shall show that the DES(k) model holds. Let $\{\tilde{p}_{ij}\}$ denote the cell probabilities which satisfy all the QDECS(k), GS, and MME(k) models. Since the QDECS(k) model holds, we see

$$\log \tilde{p}_{ij} = \begin{cases} \log \gamma + \sum_{l=1}^{k} (i^l \log \alpha_l + j^l \log \beta_l) + \log d_{i-j} & (i < j), \\ \sum_{l=1}^{k} (i^l \log \alpha_l + j^l \log \beta_l) + \log d_{j-i} & (i > j), \\ \log \psi_{ii} & (i = j). \end{cases} \quad (5)$$

Let

$$\pi_{ij} = \frac{d_{-|j-i|}}{c} \ (i \neq j) \quad \text{and} \quad \pi_{ii} = \frac{\psi_{ii}}{c} \ (i = j),$$

with

$$c = \sum_{\substack{i=1 \\ (i \neq j)}}^{R} \sum_{j=1}^{R} d_{-|j-i|} + \sum_{i=1}^{R} \psi_{ii}.$$

We note that $\sum_{i=1}^{R} \sum_{j=1}^{R} \pi_{ij} = 1$ with $0 < \pi_{ij} < 1$. Then, since $\{\tilde{p}_{ij}\}$ satisfy the QDECS(k), GS, and MME(k) models, we see

$$\log\left(\frac{\tilde{p}_{ij}}{\pi_{ij}}\right) = \begin{cases} \log c + \log \gamma + \sum_{l=1}^{k} (i^l \log \alpha_l + j^l \log \beta_l) & (i < j), \\ \log c + \sum_{l=1}^{k} (i^l \log \alpha_l + j^l \log \beta_l) & (i > j), \\ \log c & (i = j), \end{cases} \tag{6}$$

$$\tilde{\delta}_U = \tilde{\delta}_L \quad \text{and} \quad \tilde{\mu}_X^l = \tilde{\mu}_Y^l \quad (l = 1, \ldots, k),$$

where

$$\tilde{\delta}_U = \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \tilde{p}_{ij}, \quad \tilde{\delta}_L = \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \tilde{p}_{ji},$$

$$\tilde{\mu}_X^l = \sum_{i=1}^{R} \sum_{j=1}^{R} i^l \tilde{p}_{ij}, \quad \tilde{\mu}_Y^l = \sum_{i=1}^{R} \sum_{j=1}^{R} j^l \tilde{p}_{ij}.$$

Then, we denote $\tilde{\delta}_U (= \tilde{\delta}_L)$ and $\tilde{\mu}_X^l (= \tilde{\mu}_Y^l)$ by δ_0 and μ_0^l, respectively. Namely,

$$\tilde{\delta}_U = \tilde{\delta}_L = \delta_0 \quad \text{and} \quad \tilde{\mu}_X^l = \tilde{\mu}_Y^l = \mu_0^l \ (l = 1, \ldots, k). \tag{7}$$

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

$$\delta_U = \delta_L (= \delta_0) \quad \text{and} \quad \mu_X^l = \mu_Y^l \ (= \mu_0^l)\ (l = 1, \ldots, k), \tag{8}$$

where $\delta_U, \delta_L, \{\mu_X^l\}$, and $\{\mu_Y^l\}$ denote $\tilde{\delta}_U, \tilde{\delta}_L, \{\tilde{\mu}_X^l\}$, and $\{\tilde{\mu}_Y^l\}$ with $\{\tilde{p}_{ij}\}$ replaced by $\{p_{ij}\}$, respectively. From the Equations (6), (7), and (8), we see

$$\sum_{i=1}^{R}\sum_{j=1}^{R}(p_{ij} - \tilde{p}_{ij})\log\left(\frac{\tilde{p}_{ij}}{\pi_{ij}}\right) = 0. \tag{9}$$

Using the Equation (9), we obtain

$$K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{\tilde{p}_{ij}\}, \{\pi_{ij}\}) + K(\{p_{ij}\}, \{\tilde{p}_{ij}\}),$$

where

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

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

$$\min_{\{p_{ij}\}} K(\{p_{ij}\}, \{\pi_{ij}\}) = K(\{\tilde{p}_{ij}\}, \{\pi_{ij}\}),$$

and then $\{\tilde{p}_{ij}\}$ uniquely minimizes $K(\{p_{ij}\}, \{\pi_{ij}\})$ (see Bhapkar and Darroch [2]).

Let $\tilde{p}_{ij}^* = \tilde{p}_{ji}$ $(i = 1, \ldots, R;\ j = 1, \ldots, R)$. Then

$$\log \tilde{p}_{ij}^* = \log \tilde{p}_{ji} = \begin{cases} \sum_{l=1}^{k}(j^l \log \alpha_l + i^l \log \beta_l) + \log d_{i-j} & (i < j), \\ \log \gamma + \sum_{l=1}^{k}(j^l \log \alpha_l + i^l \log \beta_l) + \log d_{j-i} & (i > j), \\ \log \psi_{ii} & (i = j). \end{cases} \tag{10}$$

Noting that $\{\pi_{ij} = \pi_{ji}\}$, we see

$$\log\left(\frac{\tilde{p}_{ij}^*}{\pi_{ij}}\right) = \begin{cases} \log c + \sum_{l=1}^{k} (j^l \log \alpha_l + i^l \log \beta_l) & (i < j), \\ \log c + \log \gamma + \sum_{l=1}^{k} (j^l \log \alpha_l + i^l \log \beta_l) & (i > j), \\ \log c & (i = j). \end{cases} \quad (11)$$

From Equations (7), (8), and (11), we see

$$\sum_{i=1}^{R}\sum_{j=1}^{R} (p_{ij} - \tilde{p}_{ij}^*) \log\left(\frac{\tilde{p}_{ij}^*}{\pi_{ij}}\right) = 0. \quad (12)$$

Using the Equation (12), we obtain

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

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

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

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

From Equations (5) and (10), for $i < j$, we see

$$\begin{aligned} \log\left(\frac{\tilde{p}_{ij}}{\tilde{p}_{ji}}\right) &= \log \gamma + \sum_{l=1}^{k} \left((i^l - j^l) \log \alpha_l + (j^l - i^l) \log \beta_l\right) \\ &= \log \gamma + \sum_{l=1}^{k} \left((j^l - i^l) \log \frac{\beta_l}{\alpha_l}\right) \\ &= 0. \end{aligned} \quad (13)$$

Thus

$$\log\left(\frac{\tilde{p}_{i,j+1}}{\tilde{p}_{j+1,i}}\right) - \log\left(\frac{\tilde{p}_{ij}}{\tilde{p}_{ji}}\right) = \sum_{l=1}^{k}\left(\{(j+1)^l - j^l\}\log\frac{\beta_l}{\alpha_l}\right)$$

$$= 0. \tag{14}$$

From Equation (14), we obtain $\alpha_l = \beta_l (l = 1, \ldots, k)$, and from Equation (13), we obtain $\gamma = 1$. Namely, the DES(k) model holds. The proof is completed.

Proof of Theorem 6. For a fixed k $(k = 1, \ldots, R-1)$, if the DES(k) model holds, then the DECS(k) and MME(1) models hold. Assuming that both the DECS(k) and MME(1) models hold, then we shall show that the DES(k) model holds. The MME(1) model is also expressed as

$$\sum_{u=1}^{R-1}\sum_{s=1}^{R-u} up_{s,s+u} = \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} up_{s+u,s}.$$

Since the DECS(k) and MME(1) models hold, we see

$$\sum_{u=1}^{R-1}\sum_{s=1}^{R-u} up_{s,s+u} - \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} up_{s+u,s}$$

$$= \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} u(\prod_{l=1}^{k}\delta_l^{s^l+(s+u)^l})d_{(s+u)-s} - \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} u(\prod_{l=1}^{k}\delta_l^{(s+u)^l+s^l})d_{s-(s+u)}$$

$$= \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} u(\prod_{l=1}^{k}\delta_l^{s^l+(s+u)^l})\gamma d_{-u} - \sum_{u=1}^{R-1}\sum_{s=1}^{R-u} u(\prod_{l=1}^{k}\delta_l^{(s+u)^l+s^l})d_{-u}$$

$$= (\gamma - 1)\sum_{u=1}^{R-1}\sum_{s=1}^{R-u} u(\prod_{l=1}^{k}\delta_l^{s^l+(s+u)^l})d_{-u}$$

$$= 0.$$

Thus, we obtain $\gamma = 1$. Namely, the DES(k) model holds. The proof is complicated.

Consider the model that has the structure of both the GS and MME(k) models. We shall refer to this model as the GSMME(k) model. From Theorem 5, we can obtain the following the corollary:

Corollary 1. *For a fixed* k $(k = 1, \dots, R-1)$, *the DES*(k) *model holds if and only if both the QDECS*(k) *and GSMME*(k) *models hold.*

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 DECS(k) and QDECS(k) models could be obtained by using iterative procedures; for example, see Darroch and Ratcliff [5] and Agresti ([1], p. 242). The MLEs of $\{m_{ij}\}$ under the MME(k) and GSMME(k) models could be obtained by 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 of model M. The numbers of degrees of freedom for the DECS(k), QDECS(k), MME(k), and GSMME(k) models are $R^2 - 2R - k$, $R^2 - 2R - 2k$, k, and $k + 1$, 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 [6] and Read [12]. We obtain the theorems as follows:

Theorem 7. *For a fixed* k $(k = 1, \dots, R-1)$, *the test statistic* $G^2(DES(k))$ *is asymptotically equivalent to the sum of* $G^2(DECS(k))$ *and* $G^2(GS)$.

Theorem 8. *For a fixed* k $(k = 1, \dots, R-1)$, *the test statistic* $G^2(DES(k))$ *is asymptotically equivalent to the sum of* $G^2(QDECS(k))$ *and* $G^2(GSMME(k))$.

Proof of Theorem 7. For a fixed k $(k = 1, \ldots, R-1)$, the DECS(k) model is expressed as

$$\log p_{ij} = \begin{cases} \gamma^* + (i+j)\beta_1^* + (i^2 + j^2)\beta_2^* + \cdots + (i^k + j^k)\beta_k^* + d_{i-j}^* & (i < j), \\ (i+j)\beta_1^* + (i^2 + j^2)\beta_2^* + \cdots + (i^k + j^k)\beta_k^* + d_{j-i}^* & (i > j), \\ \psi_{ii}^* & (i = j), \end{cases} \tag{15}$$

$$(i = 1, \ldots, R;\ j = 1, \ldots, R).$$

Let

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

$$\beta = (\gamma^*, \beta_1^*, \beta_2^*, \ldots, \beta_k^*, \phi)^t,$$

where "t" denotes the transpose, and

$$\phi = (d_{-1}^*, d_{-2}^*, \ldots, d_{-(R-1)}^*, \psi_{11}^*, \psi_{22}^*, \ldots, \psi_{RR}^*)$$

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

$$\log p = X\beta = (X_0, X_1, X_2, \ldots, X_k, X_{k+1})\beta,$$

where X is the $R^2 \times L$ matrix with $L = 2R + k$, $X_0 = (v_1, \ldots, v_R)^t$, $X_l = J_R^l \otimes 1_R + 1_R \otimes J_R^l$ $(l = 1, \ldots, k)$, and X_{k+1} is the $R^2 \times (2R-1)$ matrix of 1 or 0 elements determined from (15); and where v_p is the $1 \times R$ vector of 0 for the first p elements or 1 for the others, 1_s is the $s \times 1$ vector of 1 elements, $J_R^l = (1^l, \ldots, R^l)^t$ and $\otimes$ denotes the Kronecker product. The matrix X has full column rank. The rank of X is L. In a similar manner to Haber [7], we denote the linear space spanned by the columns of the matrix X by $S(X)$ with the dimension L.

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

$$H_2(p) = 0_{l_2},$$

where $l_2 = 1$, and $H_2(p) = Wp$, with

$$W = (X_0 - (1_{R^2} - X_0 - \sum_{k=1}^{R} w_k))^t,$$

being the $1 \times R^2$ matrix; and where $w_i (i = 1, \ldots, R)$ is the $R^2 \times 1$ vector, being the corresponding column vectors in X_{k+1} shouldering ψ_{ii}^*. Note that $X_{k+1} 1_{2R-1} = 1_{R^2}$. Thus W^t belongs to $S(X)$. Hence $WU = O_{l_2, l_1}$. From Theorem 1, the DES(k) model is expressed as

$$H_3(p) = 0_{l_3},$$

where $l_3 = l_1 + l_2 = R^2 - 2R - k + 1$, and $H_3 = (H_1^t, H_2)^t$.

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 $\Sigma(p) = \mathrm{diag}(p) - pp^t$, where $\mathrm{diag}(p)$ 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, $\sqrt{n}\,(H_3(\hat{p}) - H_3(p))$ has asymptotically a normal distribution with mean 0_{l_3} and covariance matrix

$$h_3(p)\Sigma(p)h_3(p)^t = \begin{bmatrix} h_1(p)\Sigma(p)h_1(p)^t & h_1(p)\Sigma(p)h_2(p)^t \\ h_2(p)\Sigma(p)h_1(p)^t & h_2(p)\Sigma(p)h_2(p)^t \end{bmatrix}.$$

Since $h_1(p)p = U^t 1_{R^2} = 0_{l_1}$, $h_1(p)\mathrm{diag}(p) = U^t$ 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

$$\Delta_s(p) = H_s(p)^t [h_s(p)\Sigma(p)h_s(p)^t]^{-1} H_s(p). \tag{16}$$

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 Equation (16), we see that $W_3 = W_1 + W_2$. From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic, we obtain Theorem 7.

We shall omit the proof of Theorem 8 because it is obtained in a similar way to the proof of Theorem 7.

4. Concluding Remarks

We have proposed the DECS(k) and QDECS(k) models, and given the three kinds of decompositions of the DES(k) model. These decompositions may be useful for seeing the reason for the poor fit of the DES(k) model when the DES(k) model fits the data poorly.

We point out that $G^2(DES(k))$ is not asymptotically equivalent to the sum of $G^2(QDECS(k))$, $G^2(GS)$ and $G^2(MME(k))$ because the sum of $G^2(GS)$ and $G^2(MME(k))$ is not asymptotically equivalent to $G^2(GSMME(k))$, however, the $G^2(DES(k))$ is asymptotically equivalent to the sum of $G^2(QDECS(k))$ and $G^2(GSMME(k))$ (see Theorem 8).

We note that the DECS$(R-1)$ model implies the CS model, and the difference between the numbers of degrees of freedom for the DECS$(R-1)$ and the CS model is $(R-1)(R-4)/2$. The DECS$(R-1)$ model is equivalent to the CS model when $R = 4$.

References

[1] A. Agresti, Analysis of Ordinal Categorical Data, New York: Wiley, 1984.

[2] V. P. Bhapkar and J. N. Darroch, Marginal symmetry and quasi symmetry of general order, Journal of Multivariate Analysis 34(2) (1990), 173-184.

DOI: https://doi.org/10.1016/0047-259X(90)90034-F

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

[4] H. Caussinus, Contribution à l'analyse statistique des tableaux de corrélation, Annales de la Faculté des Sciences de l'Université de Toulouse 29(4) (1965), 77-182.

[5] J. N. Darroch and D. Ratcliff, Generalized iterative scaling for log-linear models, Annals of Mathematical Statistics 43(5) (1972), 1470-1480.

DOI: https://doi.org/10.1214/aoms/1177692379

[6] J. N. Darroch and S. D Silvey, On testing more than one hypothesis, Annals of Mathematical Statistics 34(2) (1963), 555-567.

DOI: https://doi.org/10.1214/aoms/1177704168

[7] M. Haber, Maximum likelihood methods for linear and log-linear models in categorical data, Computational Statistics and Data Analysis 3 (1985), 1-10.

DOI: https://doi.org/10.1016/0167-9473(85)90053-2

[8] K. Iki, A. Shibuya and S. Tomizawa, Generalized diagonal exponent symmetry model and its orthogonal decomposition for square contingency tables with ordered categories, Journal of Statistics: Advances in Theory and Applications 14(2) (2015), 207-220.

DOI: http://dx.doi.org/10.18642/jsata_7100121584

[9] K. Iki, A. Shibuya and S. Tomizawa, Diagonal exponent conditional symmetry model for square contingency tables with ordered categories, International Journal of Statistics and Probability 5 (2016), 38-44.

DOI: https://doi.org/10.5539/ijsp.v5n4p38

[10] K. Iki, K. Yamamoto and S. Tomizawa, Quasi-diagonal exponent symmetry model for square contingency tables with ordered categories, Statistics and Probability Letters 92 (2014), 33-38.

DOI: https://doi.org/10.1016/j.spl.2014.04.029

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

DOI: https://doi.org/10.1093/biomet/65.2.413

[12] C. B. Read, Partitioning chi-square in contingency table: A teaching approach, Communications in Statistics-Theory and Methods 6(6) (1977), 553-562.

DOI: https://doi.org/10.1080/03610927708827513

[13] S. Tomizawa, 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(4) (1992), 509-512.

DOI: https://doi.org/10.1080/02664769200000046

■