

Diagonal Exponent Conditional Symmetry Model for Square Contingency Tables with Ordered Categories

Kiyotaka Iki¹, Akira Shibuya¹ & Sadao Tomizawa¹

¹ Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Chiba, Japan

Correspondence: Kiyotaka Iki, Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan. Tel: 81-4-7124-1501. E-mail: iki@is.noda.tus.ac.jp

Received: February 24, 2016 Accepted: April 1, 2016 Online Published: June 11, 2016

doi:10.5539/ijsp.v5n4p38 URL: <http://dx.doi.org/10.5539/ijsp.v5n4p38>

Abstract

For square contingency tables with ordered categories, this article proposes new models which indicate that in addition to the structure of asymmetry of the probabilities with respect to the main diagonal of the table, the expected frequency has an exponential form along every subdiagonal of the table. Also it gives the new three kinds of decompositions using the proposed model and proves the orthogonality of the test statistics.

Keywords: conditional symmetry, diagonal exponent symmetry, ordered category, orthogonal decomposition, square contingency table

1. Introduction

Consider an $R \times R$ square contingency table with the same row and column classifications. We may be interested in symmetry or asymmetry about the main diagonal of the table rather than independence. Let p_{ij} denote the probability that an observation will fall in the i th row and j th column of the table ($i = 1, \dots, R; j = 1, \dots, R$). Caussinus (1965) 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 QS model with $\{\alpha_i = \beta_i\}$ is the symmetry (S) model (see, Bowker, 1948; Martin & Pardo, 2010; Kolassa & Bhagavatula, 2012). The conditional symmetry (CS) model is defined by

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

where $\psi_{ij} = \psi_{ji}$; see McCullagh (1978). The CS model states that p_{ij} ($i < j$) is γ times higher than p_{ji} . The CS model with $\gamma = 1$ is the S model.

The global symmetry (GS) model is defined by

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

see Read (1977). 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 (1977) gave the theorem that the S model holds if and only if both the CS and GS models hold.

Tomizawa (1992) considered the diagonal exponent symmetry (DES) model defined by

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

The DES model states that in addition to the structure of the S model, $p_{i+1, j+1}$ ($i \neq j$) is δ^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 & Tomizawa (2014) considered the quasi-diagonal exponent symmetry (QDES) model defined by

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

The QDES model with $\alpha = \beta$ is the DES model. The QDES model states that in addition to the structure of the QS model (instead of the S model), the expected frequency has an exponential form along every subdiagonal of the table. Under the QDES model, we see the structure of $p_{ij}/p_{ji} = (\beta/\alpha)^{j-i}$ ($i < j$).

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. (2014) gave the theorem that the DES model holds if and only if both the QDES and ME models hold. Other symmetry and asymmetry models have been described in Tahata & Tomizawa (2014).

We are interested in considering new models which indicate that in addition to the structure of the CS model (instead of the S model), the expected frequency has an exponential form along every subdiagonal of the table. The present paper proposes two new models and gives the new three kinds of decompositions of the DES model.

2. New Models

Consider a model defined by

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

where $d_{j-i} = \gamma d_{i-j}$ ($i < j$). 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} . Thus we shall refer to this model as the diagonal exponent conditional symmetry (DECS) model. Under the DECS model, we see the structure of $p_{ij}/p_{ji} = \gamma$ ($i < j$). The DECS model with $\gamma = 1$ is the DES model.

Next, consider a model defined by

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

where $d_{j-i} = \gamma d_{i-j}$ ($i < j$). We shall refer to this model as the quasi-diagonal exponent conditional symmetry (QDECS) model. Under the QDECS model, we see the structure of $p_{ij}/p_{ji} = \gamma(\beta/\alpha)^{j-i}$ ($i < j$). The QDECS model with $\gamma = 1$ is the QDES model. Also, QDECS model with $\alpha = \beta$ is the DECS model.

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

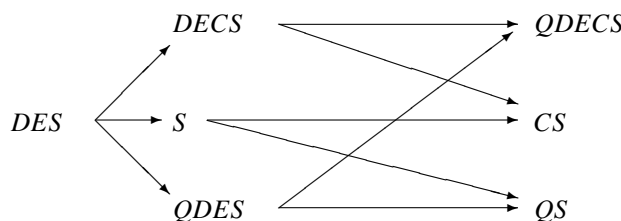


Figure 1. Relationships among models.

3. Decompositions and Orthogonality of Test Statistics

We obtain the new three kinds of decompositions of the DES model as follows:

Theorem 1 *The DES model holds if and only if all the QDECS, GS and ME models hold.*

Theorem 2 *The DES model holds if and only if both the DECS and GS models hold.*

Theorem 3 *The DES model holds if and only if both the DECS and ME models hold.*

The proofs of these theorems are given in Appendix 1.

Consider the model that has the structure of both the GS and ME models. We shall refer to this model as the GSME model. From Theorem 1, we can obtain the following the corollary:

Corollary 1 *The DES model holds if and only if the QDECS and GSME 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 and QDECS models could be obtained using iterative procedures; for example, see Darroch & Ratcliff (1972). The MLEs of $\{m_{ij}\}$ under the GSME 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 of model M . The numbers of degrees of freedom (df) for the DECS and QDECS models are $R^2 - 2R - 1$ and $R^2 - 2R - 2$, respectively.

The orthogonality (asymptotic separability or independence) of the test statistics for goodness-of-fit of two models is discussed by, e.g., Darroch & Silvey (1963) and Read (1977). We obtain the following theorems for the orthogonality.

Theorem 4 *The test statistic $G^2(DES)$ is asymptotically equivalent to the sum of $G^2(QDECS)$ and $G^2(GSME)$.*

Theorem 5 *The test statistic $G^2(DES)$ is asymptotically equivalent to the sum of $G^2(DECS)$ and $G^2(GS)$.*

The proof of Theorem 4 is given in Appendix 2. We shall omit the proof of Theorem 5 because it is obtained in a similar way to the proof of Theorem 4.

4. An Example

Consider the data in Table 1 taken from Agresti (2002, p. 462). These data are insomniac patient's reported time (in minutes) to fall asleep after going to bed. The response is the patient's reported time at baseline (before treatment) and following two weeks of treatment (hypnotic drug).

Table 1. Insomniac patient's reported time (in minutes) to fall asleep after going to bed; from Agresti (2002, p. 462). (The upper and lower parenthesized values are MLEs of expected frequencies under the DECS model and the special DECS model with $\delta = 1$, respectively.)

| Initial | Follow-up | | | | Total |
|---------|--------------------------|--------------------------|--------------------------|-----------------------|-------|
| | < 20 | 20-30 | 30-60 | > 60 | |
| < 20 | 7 (7.00) (7.00) | 4 (1.81) (1.88) | 1 (1.69) (1.72) | 0 (0.94) (0.94) | 12 |
| 20-30 | 11 (15.58) (16.12) | 5 (5.00) (5.00) | 2 (1.87) (1.88) | 2 (1.75) (1.72) | 20 |
| 30-60 | 13 (14.53) (14.78) | 23 (16.12) (16.12) | 3 (3.00) (3.00) | 1 (1.94) (1.88) | 40 |
| > 60 | 9 (8.06) (8.06) | 17 (15.03) (14.78) | 13 (16.67) (16.12) | 8 (8.00) (8.00) | 47 |
| Total | 40 | 49 | 19 | 11 | 119 |

We see from Table 2 that the CS, DECS and QDECS models fit these data well, although the other models fit poorly. Since the DECS model is a special case of the QDECS model, we shall test the hypothesis that the DECS model holds (i.e., the hypothesis of $\alpha = \beta$) assuming that the QDECS model holds. Since $G^2(DECS|QDECS) = G^2(DECS) - G^2(QDECS) = 1.54$ with 1 df being the difference between the numbers of df for the DECS and the QDECS models, this hypothesis is accepted at the 0.05 significance level. Similarly, the hypothesis that the DECS model holds assuming that the CS model holds is accepted for these data. Therefore, the DECS model would be preferable to the CS and QDECS models.

Under the DECS model, the MLEs of parameters of γ and δ are $\hat{\gamma} = 0.116$ and $\hat{\delta} = 1.017$. Since $\hat{\delta}$ is close to 1, we are now interested in a special DECS model obtained by putting $\delta = 1$. For this model we obtain the likelihood ratio statistic $G^2(DECS \text{ with } \delta = 1) = 10.30$ with 8 df. Thus the special DECS model with $\delta = 1$ also fits these data well. Moreover, we shall test the hypothesis that the special DECS model with $\delta = 1$ holds (i.e., the hypothesis of $\delta = 1$) assuming that the DECS model holds for these data. Since $G^2(DECS \text{ with } \delta = 1|DECS) = G^2(DECS \text{ with } \delta = 1) - G^2(DECS) = 0.05$ with 1 df, this hypothesis is accepted at the 0.05 significance level. Therefore the special DECS model with $\delta = 1$ may be preferable to the DECS model.

Under the special DECS model with $\delta = 1$, the MLE of parameter γ is $\hat{\gamma} = 0.116$. Thus, under the special DECS model with $\delta = 1$, the probability that a patient's reported time at baseline and his or her reported time at following two weeks of treatment are i and j ($i > j$), respectively, is estimated to be $\hat{\gamma}^{-1} = 8.621$ times higher than the probability that those are j and i , respectively. Thus, since $\hat{\gamma}^{-1} > 1$, the patient's reported time at following two weeks of treatment is faster than the patient's reported time at baseline. Also under this model, the probability that a patient's reported time at baseline and his or her reported time at following two weeks of treatment are $i + 1$ and $j + 1$, respectively, is estimated to be equal to the probability that those are i and j , respectively.

We see from Table 2 and Theorem 1 that the poor fit of the DES model is caused by the influence of the lack of structure of the GS and ME models rather than the QDECS model. Similarly, from Theorem 2, the poor fit of the DES model is caused by the influence of the lack of structure of the GS model rather than the DECS model. Also, from Theorem 3, the poor fit of the DES model is caused by the influence of the lack of structure of the ME model rather than the DECS model.

Table 2. Likelihood ratio chi-squared values G^2 for models applied to Table 1.

| Applied models | df | G^2 |
|----------------|----|--------|
| S | 6 | 74.55* |
| CS | 5 | 5.62 |
| DES | 8 | 79.18* |
| DECS | 7 | 10.25 |
| QDECS | 6 | 8.71 |
| GS | 1 | 68.93* |
| ME | 1 | 66.64* |
| GSME | 2 | 71.19* |

* means significant at the 0.05 level.

5. Concluding Remarks

We have proposed the DECS and QDECS models, and given the three kinds of decompositions of the DES model. These decompositions may be useful for seeing the reason for the poor fit of the DES model.

The $G^2(DES)$ is asymptotically equivalent to the sum of values $G^2(QDECS)$ and $G^2(GS ME)$ as described by Theorem 4. However, we point out that for the decomposition in Theorem 1, the $G^2(DES)$ is not asymptotically equivalent to the sum of values $G^2(QDECS)$, $G^2(GS)$ and $G^2(ME)$ because the sum of values $G^2(GS)$ and $G^2(ME)$ is not asymptotically equivalent to the $G^2(GS ME)$.

Acknowledgements

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

Appendix 1

Proof of Theorem 1. If the DES model holds, then the QDECS, GS and ME models hold. Assuming that all the QDECS, GS and ME models hold, then we shall show that the DES model holds. Since the QDECS model holds, we have

$$p_{st} - p_{ts} = (\gamma\theta^{t-s} - 1)p_{ts} \quad (s < t), \tag{A.1}$$

where $\theta = \beta/\alpha$. Since (A.1) and the GS model hold, we see

$$\sum_{s < t} (\gamma\theta^{t-s} - 1)p_{ts} = 0,$$

namely,

$$\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} (\gamma\theta^k - 1)p_{s+k,s} = 0. \tag{A.2}$$

From (A.2), we see

$$\gamma = \frac{\sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} p_{t+\ell,t}}{\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} p_{s+k,s} \theta^k}. \tag{A.3}$$

The ME model can be expressed as

$$\sum_{i=1}^{R-1} G_{1(i)} = \sum_{i=1}^{R-1} G_{2(i)}, \tag{A.4}$$

where

$$G_{1(i)} = \sum_{s=1}^i \sum_{t=i+1}^R p_{st}, \quad G_{2(i)} = \sum_{s=1}^i \sum_{t=i+1}^R p_{ts}.$$

From (A.1) and (A.4), we see

$$\sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R (\gamma\theta^{t-s} - 1)p_{ts} = 0,$$

namely,

$$\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} k(\gamma\theta^k - 1)p_{s+k,s} = 0. \tag{A.5}$$

From (A.5), we see

$$\gamma = \frac{\sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} \ell p_{t+\ell,t}}{\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} k p_{s+k,s} \theta^k}. \tag{A.6}$$

From (A.3) and (A.6), we obtain

$$\left(\sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} p_{t+\ell,t}\right) \left(\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} k p_{s+k,s} \theta^k\right) - \left(\sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} \ell p_{t+\ell,t}\right) \left(\sum_{k=1}^{R-1} \sum_{s=1}^{R-k} p_{s+k,s} \theta^k\right) = 0,$$

namely,

$$\sum_{k=1}^{R-1} \sum_{\ell=1}^{R-1} \sum_{s=1}^{R-k} \sum_{t=1}^{R-\ell} (k - \ell) p_{s+k,s} p_{t+\ell,t} \theta^k = 0. \tag{A.7}$$

The equation (A.7) is also expressed as

$$(\theta - 1) \sum_{m=2}^{R-1} \left(\theta^{m-1} \sum_{g=m}^{R-1} \sum_{s=1}^{R-g} \sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} (g - \ell) p_{s+g,s} p_{t+\ell,t} \right) = 0. \tag{A.8}$$

In addition,

$$\sum_{g=m}^{R-1} \sum_{s=1}^{R-g} \sum_{\ell=1}^{R-1} \sum_{t=1}^{R-\ell} (g - \ell) p_{s+g,s} p_{t+\ell,t} = \sum_{g=m}^{R-1} \sum_{\ell=1}^{m-1} \sum_{s=1}^{R-g} \sum_{t=1}^{R-\ell} (g - \ell) p_{s+g,s} p_{t+\ell,t} + \sum_{g=m}^{R-1} \sum_{\ell=m}^{R-1} \sum_{s=1}^{R-g} \sum_{t=1}^{R-\ell} (g - \ell) p_{s+g,s} p_{t+\ell,t}. \tag{A.9}$$

The first term on the right-hand side of (A.9) is positive and the second term equals zero. Thus (A.9) is positive. Therefore, noting that $\theta > 0$, from (A.8) we obtain $\theta = 1$, i.e., $\alpha = \beta$. Thus, from (A.2) we obtain $\gamma = 1$. Namely, the DES model holds. The proof is complicated.

Proof of Theorem 2. If the DES model holds, then the DECS and GS models hold. Assuming that both the DECS and GS models hold, then we shall show that the DES model holds. Since the DECS and GS models hold, we see

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

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

Proof of Theorem 3. If the DES model holds, then the DECS and ME models hold. Assuming that both the DECS and ME models hold, then we shall show that the DES model holds. Since the DECS and ME models hold, we see

$$\begin{aligned} \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R p_{st} - \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R p_{ts} &= \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R \delta^{s+t} d_{t-s} - \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R \delta^{s+t} d_{s-t} \\ &= \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R \delta^{s+t} \gamma d_{s-t} - \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R \delta^{s+t} d_{s-t} \\ &= (\gamma - 1) \sum_{i=1}^{R-1} \sum_{s=1}^i \sum_{t=i+1}^R \delta^{s+t} d_{s-t} \\ &= 0. \end{aligned}$$

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

Appendix 2

Proof of Theorem 4. The QDECS model is expressed as

$$\log p_{ij} = \begin{cases} \gamma^* + i\alpha^* + j\beta^* + d_{i-j}^* & (i < j), \\ i\alpha^* + j\beta^* + d_{j-i}^* & (i > j), \\ \psi_{ii}^* & (i = j). \end{cases} \tag{A.10}$$

Let

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

where “ t ” denotes the transpose, and

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

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

$$\log p = X\beta = (X_0, X_1, X_2, X_3)\beta,$$

where X is the $R^2 \times L$ matrix with $L = 2R + 2$, $X_0 = (v_1, \dots, v_R)^t$ (the $R^2 \times 1$ vector), $X_1 = J_R \otimes 1_R$ (the $R^2 \times 1$ vector), $X_2 = 1_R \otimes J_R$ (the $R^2 \times 1$ vector), and X_3 is the $R^2 \times (2R - 1)$ matrix of 1 or 0 elements determined from (A.10); 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, $J_R = (1, \dots, R)^t$ and \otimes denotes the Kronecker product. The matrix X is full column rank which is L . In a similar manner to Haber (1985), 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 $l_1 = R^2 - L = R^2 - 2R - 2$, full column rank matrix such that the linear space spanned by the columns of U , i.e., $S(U)$, 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 QDECS 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 GSME model is expressed as

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

where $l_2 = 2$ and $h_2(p) = Wp$ with

$$W = \begin{pmatrix} (2X_0 - 1_{R^2} + w_1 + w_2 + \dots + w_R)^t \\ (X_2 - X_1)^t \end{pmatrix}; \text{ the } 2 \times R^2 \text{ matrix,}$$

where w_i ($i = 1, \dots, R$) is the $R^2 \times 1$ vector, being one of column vectors in X_3 shouldering ψ_{ii}^* . Note that $X_3 1_{2R-1} = 1_{R^2}$. Thus W^t belongs to $S(X)$, i.e., $S(W^t) \subset S(X)$. Hence $WU = O_{l_2, l_1}$. From Corollary 1, the DES model is expressed as

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

where $l_3 = l_1 + l_2 = R^2 - 2R$, 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 \widehat{p} denote p with $\{p_{ij}\}$ replaced by $\{\widehat{p}_{ij} = n_{ij}/n\}$. Then $\sqrt{n}(\widehat{p} - p)$ has asymptotically a normal distribution with mean 0_{R^2} and covariance matrix $\Sigma(p)$. Using the delta method, $\sqrt{n}(h_3(\widehat{p}) - h_3(p))$ 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 $H_1(p)p = U^t 1_{R^2} = 0_{l_1}$, $H_1(p)\text{diag}(p) = U^t$ 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

$$\Delta_s(p) = h_s(p)'[H_s(p)\Sigma(p)H_s(p)']^{-1}h_s(p). \quad (\text{A.11})$$

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 (A.11), we see that $W_3 = W_1 + W_2$. From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic, we obtain Theorem 4.

References

- Agresti, A. (2002). *Categorical Data Analysis, 2nd ed.* Wiley, New York.
- 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>
- 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.
- Darroch, J. N. & 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>
- Darroch, J. N. & 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>
- 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)
- Iki, K., Yamamoto, K. & 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>
- Kolassa, J. E. & Bhagavatula, H. G. (2012). Accurate approximations to the distribution of a statistic testing symmetry in contingency tables. *Institute of Mathematical Statistics*, *8*, 181-189. <http://dx.doi.org/10.1214/11-IMSCOLL812>
- Martin, N. & Pardo, L. (2010). A new measure of leverage cells in multinomial loglinear models. *Communications in Statistics-Theory and Methods*, *39*, 517-530. <http://dx.doi.org/10.1080/03610920903139991>
- McCullagh, P. (1978). A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika*, *65*, 413-418. <http://dx.doi.org/10.1093/biomet/65.2.413>
- 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>
- Tahata, K. & Tomizawa, S. (2014). Symmetry and asymmetry models and decompositions of models for contingency tables. *SUT journal of Mathematics*, *50*, 131-165.
- 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>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).