Ridit Score Type Quasi-Symmetry and Decomposition of Symmetry for Square Contingency Tables with Ordered Categories

Kiyotaka Iki, Kouji Tahata and Sadao Tomizawa
Department of Information Sciences, Tokyo University of Science, Japan

Abstract: For square contingency tables with the same row and column ordinal classifications, this paper proposes the quasi-symmetry model based on the marginal ridits. The model indicates that the log-odds that an observation will fall in the \((i, j)\) cell instead of in the \((j, i)\) cell, \(i < j\), is proportional to the difference between the average ridit score of row and column marginal distributions for category \(j\) and that for category \(i\). This paper also gives a theorem such that the symmetry model holds if and only if both the proposed model and the marginal mean equality model hold. Examples are given.

Keywords: Quasi-Symmetry, Rank Score, Ridit, Square Contingency Table, Symmetry.

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 \((i, j = 1, \ldots, R)\). Caussinus (1965) considered the quasi-symmetry (QS) model, defined by

\[
p_{ij} = \mu \alpha_i \beta_j \psi_{ij}, \quad i, j = 1, \ldots, R,
\]

where \(\psi_{ij} = \psi_{ji}\). A special case of this model with \(\{\alpha_i = \beta_i\}\) is the symmetry (S) model (see, Bowker, 1948; Bishop, Fienberg, and Holland, 1975, p. 282; Tomizawa and Tahata, 2007). The marginal homogeneity (MH) model is defined by

\[
p_i = p_{i*}, \quad i = 1, \ldots, R,
\]

where \(p_i = \sum_{j=1}^{R} p_{ij}\) and \(p_{i*} = \sum_{s=1}^{R} p_{si}\) (Stuart, 1955; Bishop et al., 1975, p. 293). Also, Caussinus (1965) gave the theorem that the S model holds if and only if both the QS and MH models hold.
For square tables with ordered categories, Agresti (1983) proposed the linear diagonals-parameter symmetry (LDPS) model, defined by

\[ p_{ij} = \begin{cases} 
\delta^j - i \psi_{ij}, & \text{if } i < j, \\
\psi_{ij}, & \text{if } i \geq j,
\end{cases} \]

where \( \psi_{ij} = \psi_{ji} \). A special case of this model obtained by putting \( \delta = 1 \) is the S model. Note that the LDPS model implies the QS model.

Let \( u_1 < \cdots < u_R \) denote the ordered known scores assigned for both the rows and columns of same classifications. The generalized LDPS model with \( u_i \) instead of \( i \) is the ordinal quasi-symmetry (OQS) model (Agresti, 2002, p. 429).

Let \( X \) and \( Y \) denote the row and column variables, respectively. Refer to model of equality of marginal means, i.e., \( E(X) = E(Y) \), as the ME model. Yamamoto, Iwashita, and Tomizawa (2007) gave the theorem that the S model holds if and only if both the LDPS and ME models hold (see also Tahata, Yamamoto, and Tomizawa, 2008).

Let

\[ r_i^X = \sum_{k=1}^{i-1} p_k + \frac{p_i}{2}, \quad \text{and} \quad r_i^Y = \sum_{l=1}^{i-1} p_i + \frac{p_i}{2}, \]

for \( i = 1, \ldots, R \). The \( \{r_i^X\} \) and \( \{r_i^Y\} \) are the marginal ridits; see Bross (1958), Fleiss, Levin, and Paik (2003, pp. 198-205), and Tahata, Miyamoto, and Tomizawa (2008). The \( i \)th ridit for row (column) variable is the probability that an observation falls in the row (column) category \( i - 1 \) or below plus half the probability that it falls in the row (column) category \( i \). The ridits also may be expressed as

\[ r_i^X = \frac{F_{i-1}^X + F_i^X}{2}, \quad \text{and} \quad r_i^Y = \frac{F_{i-1}^Y + F_i^Y}{2}, \]

where \( F_i^X = \sum_{k=1}^{i} p_k \) and \( F_i^Y = \sum_{l=1}^{i} p_l \), \( i = 1, \ldots, R \), are the distribution functions of \( X \) and \( Y \), respectively. Suppose that the categories of the ordinal row and column variables represent intervals of an underlying continuous distribution. If the underlying distribution is uniform over each interval, then \( r_i^X \) (\( r_i^Y \)) would equal the probability that the row (column) value for a randomly selected individual falls below the midpoint of row (column) category \( i \) (Agresti, 1984, p. 168). Therefore, for the analysis of square contingency tables with the same row and column ordinal classifications, we are now interested in proposing a quasi-symmetry model with the ridit scores \( r_i^X \) and \( r_i^Y \), instead of scores \( u_i \). We are also interested in considering a decomposition of the S model using the ridit score type quasi-symmetry model.

The purpose of this paper is (1) to propose the ridit score type quasi-symmetry model, and (2) to give a decomposition of the S model using the proposed model.

### 2 Ridit Score Type Quasi-Symmetry Model

Consider a square contingency table with the same row and column ordinal classifications. Let

\[ u_i = \frac{r_i^X + r_i^Y}{2}, \quad i = 1, \ldots, R. \]
Thus, \( v_i \) is the average of row and column ridits for category \( i \). Note that \( r^X_i < \cdots < r^X_R \) and \( r^Y_j < \cdots < r^Y_R \), thus, \( v_1 < \cdots < v_R \), and these are unknown.

Consider a model defined by

\[
p_{ij} = \begin{cases} 
\theta^{\psi_{ij} - \psi_{ij}}, & \text{if } i < j, \\
\psi_{ij}, & \text{if } i \geq j,
\end{cases}
\]

where \( \psi_{ij} = \psi_{ji} \). We shall refer to this model as the ridit score type quasi-symmetry (RQS) model. The RQS model may be expressed as

\[
\frac{p_{ij}}{p_{ji}} = \theta^{\psi_{ij} - \psi_{ij}}, \quad i < j.
\]

This indicates that the log-odds that an observation will fall in the \((i, j)\) cell instead of in the \((j, i)\) cell, \( i < j \), is proportional to the difference between \( v_j \) and \( v_i \).

A special case of the RQS model obtained by putting \( \theta = 1 \) is the S model. Also, the RQS model implies the QS model. Under the RQS model, \( \theta > 1 \) is equivalent to \( p_{ij} > p_{ji} \) for all \( i < j \), thus, \( \theta > 1 \) is equivalent to \( F^X_i > F^Y_i \) for all \( i = 1, \ldots, R - 1 \). Therefore, the parameter \( \theta \) in the RQS model would be useful for making inferences such as that \( X \) is stochastically less than \( Y \) or vice versa.

Let \( n_{ij} \) denote the observed frequency in the \( i \)th row and \( j \)th column of the table \((i, j = 1, \ldots, R)\), with \( n = \sum \sum n_{ij} \). Assume that a multinomial distribution applies to the \( R \times R \) table. Denote the row and column marginal counts by \( n_i = \sum_{j=1}^{R} n_{ij} \) and \( n_j = \sum_{i=1}^{R} n_{ij} \), \( i = 1, \ldots, R \), respectively.

The average ranks in category \( i \) are

\[
a^X_i = \frac{1}{2} \left( \sum_{k=1}^{i-1} n_k + 1 \right) + \frac{1}{2} \sum_{k=1}^{i} n_k \quad \text{and} \quad a^Y_i = \frac{1}{2} \left( \sum_{l=1}^{i-1} n_l + 1 \right) + \frac{1}{2} \sum_{l=1}^{i} n_l
\]

for \( i = 1, \ldots, R \) (Stuart, 1963; Agresti, 1984, p. 178). These are referred to the \( i \)th midrank. The \( i \)th midranks are related to the \( i \)th empirical ridits \( \tilde{r}^X_i \) and \( \tilde{r}^Y_i \) by

\[
a^X_i = n \tilde{r}^X_i + 1/2 \quad \text{and} \quad a^Y_i = n \tilde{r}^Y_i + 1/2,
\]

where \( \tilde{r}^X_i \) and \( \tilde{r}^Y_i \) denote \( r^X_i \) and \( r^Y_i \), respectively, with \( p_{st} \) replaced by \( \tilde{p}_{st} = n_{st}/n \).

Dividing these midranks by the sample size \( n \), yields the rank scores:

\[
s^X_i = \frac{1}{n} \left( \sum_{k=1}^{i-1} n_k + \frac{n_i + 1}{2} \right) = \frac{a^X_i}{n} \quad \text{and} \quad s^Y_i = \frac{1}{n} \left( \sum_{l=1}^{i-1} n_l + \frac{n_i + 1}{2} \right) = \frac{a^Y_i}{n}
\]

for \( i = 1, \ldots, R \). Thus the empirical ridits \( \tilde{r}^X_i \) (\( \tilde{r}^Y_i \)) and the rank scores \( s^X_i \) (\( s^Y_i \)) are essentially the same for large \( n \) (Agresti, 1984, p. 178; Freeman, 1987, p. 120).

Therefore, the average of row and column empirical ridits for category \( i \),

\[
\bar{u}_i = \frac{\tilde{r}^X_i + \tilde{r}^Y_i}{2} , \quad i = 1, \ldots, R ,
\]
is essentially the average of rank scores $s_i^X$ and $s_i^Y$ for category $i$ for large $n$. Also, we see that

$$\tilde{v}_j - \tilde{v}_i = \frac{\hat{r}_j^X + \hat{r}_j^Y}{2} - \frac{\hat{r}_i^X + \hat{r}_i^Y}{2} = \frac{s_j^X + s_j^Y}{2} - \frac{s_i^X + s_i^Y}{2}, \quad i < j.$$  

Thus, under the RQS model with $\{p_{st}\}$ replaced by $\{\hat{p}_{st}\}$, we see that the estimated log-odds that an observation will fall in the $(i, j)$ cell instead of in the $(j, i)$ cell, $i < j$, would be proportional to the difference between the average rank score for category $j$ and that for category $i$.

Further, the RQS model may be expressed as

$$\frac{p_{ij}}{p_{ji}} = \theta \frac{1}{2} \left[ r_j^X - r_i^X + r_j^Y - r_i^Y \right], \quad i < j.$$  

Therefore, this model applied to the data indicates that the log-odds that an observation will fall in the $(i, j)$ cell instead of in the $(j, i)$ cell, $i < j$, is proportional to the sum of difference between the ridit scores for categories $j$ and $i$ for row variable $X$ and the difference between the ridit scores for categories $j$ and $i$ for column variable $Y$.

The maximum likelihood estimates of expected frequencies under the RQS model could be obtained using the Newton-Raphson method in the log-likelihood equation (see Appendix). For the RQS model, $\{p_{ij}\}$ are determined by $R(R - 1)/2$ of $\{\psi_{ij}, i < j\}$, $(R - 1)$ of $\{\psi_{ii}\}$ (since $\sum \sum p_{ij} = 1$), and $1$ of $\theta$, thus a total of $R(R + 1)/2$. Therefore, the number of degrees of freedom for the RQS model is $(R^2 - 1) - R(R + 1)/2 = (R + 1)(R - 2)/2$, which is one less than that for the S model, and equal to that for the LDPS model.

### 3 Decomposition of the Symmetry Model

We obtain the decomposition of the S model as follows:

**Theorem 1:** The S model holds if and only if both the RQS and ME models hold.

**Proof:** If the S model holds, then the RQS and ME models hold. Assuming that both the RQS and ME models hold, then we shall show that the S model holds.

For $i = 1, \ldots, R - 1$, let

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

Then we have

$$F_i^X - F_i^Y = G_{1(i)} - G_{2(i)}, \quad i = 1, \ldots, R - 1,$$  

and

$$E(X) = R - \sum_{i=1}^{R-1} F_i^X, \quad E(Y) = R - \sum_{i=1}^{R-1} F_i^Y.$$  

Thus we see

\[ E(Y) - E(X) = \sum_{i=1}^{R-1} G_1(i) - \sum_{i=1}^{R-1} G_2(i). \]

If \( \theta = 1 \) in the RQS model, we see that the S model holds. If \( \theta > 1 \), we see

\[ \sum_{i=1}^{R-1} G_1(i) = \sum_{i=1}^{R-1} \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{st} > \sum_{i=1}^{R-1} \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{ts} = \sum_{i=1}^{R-1} G_2(i). \]

If \( \theta < 1 \), we see

\[ \sum_{i=1}^{R-1} G_1(i) = \sum_{i=1}^{R-1} \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{st} < \sum_{i=1}^{R-1} \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{ts} = \sum_{i=1}^{R-1} G_2(i). \]

Since the ME model holds, i.e., \( \sum_{i=1}^{R-1} G_1(i) = \sum_{i=1}^{R-1} G_2(i) \), we obtain \( \theta = 1 \). Namely, the S model holds. The proof is completed.

### 4 Examples

**Example 1:** The data in Table 1, taken from Stuart (1955), are constructed from unaided distance vision of 7477 women aged 30–39 employed in Royal Ordnance factories in Britain from 1943 to 1946. These data have been analyzed by many statisticians, e.g., including Stuart (1955), Caussinus (1965), Bishop et al. (1975, p. 284), McCullagh (1978), Goodman (1979), Tomizawa (1993), and Tomizawa and Tahata (2007), etc.

Table 1: Unaided distance vision of women from Stuart (1955). The upper and lower parenthesized values are the maximum likelihood estimates of expected frequencies under the LDPS and RQS models, respectively.

<table>
<thead>
<tr>
<th>Right eye grade</th>
<th>Best (1)</th>
<th>Second (2)</th>
<th>Third (3)</th>
<th>Worst (4)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best (1)</td>
<td>1520</td>
<td>266</td>
<td>124</td>
<td>66</td>
<td>1976</td>
</tr>
<tr>
<td></td>
<td>(1520.00)</td>
<td>(263.37)</td>
<td>(133.35)</td>
<td>(59.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1520.00)</td>
<td>(263.18)</td>
<td>(133.99)</td>
<td>(58.80)</td>
<td></td>
</tr>
<tr>
<td>Second (2)</td>
<td>234</td>
<td>1512</td>
<td>432</td>
<td>78</td>
<td>2256</td>
</tr>
<tr>
<td></td>
<td>(236.63)</td>
<td>(1512.00)</td>
<td>(418.23)</td>
<td>(88.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(236.81)</td>
<td>(1511.95)</td>
<td>(420.64)</td>
<td>(88.07)</td>
<td></td>
</tr>
<tr>
<td>Third (3)</td>
<td>117</td>
<td>362</td>
<td>1772</td>
<td>205</td>
<td>2456</td>
</tr>
<tr>
<td></td>
<td>(107.65)</td>
<td>(375.77)</td>
<td>(1772.00)</td>
<td>(202.27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(107.01)</td>
<td>(373.35)</td>
<td>(1772.03)</td>
<td>(200.00)</td>
<td></td>
</tr>
<tr>
<td>Worst (4)</td>
<td>36</td>
<td>82</td>
<td>179</td>
<td>492</td>
<td>789</td>
</tr>
<tr>
<td></td>
<td>(42.88)</td>
<td>(71.47)</td>
<td>(181.73)</td>
<td>(492.00)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(43.20)</td>
<td>(71.93)</td>
<td>(184.01)</td>
<td>(492.03)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1907</td>
<td>2222</td>
<td>2507</td>
<td>841</td>
<td>7477</td>
</tr>
</tbody>
</table>
Table 3 gives the values of likelihood ratio statistic $G^2$ for testing the goodness-of-fit of each model. The S, MH and ME models fit these data poorly, however, the QS, LDPS, and RQS models fit these data well. Namely, we see that the QS model with the ridit scores, i.e., the RQS model, and the QS model with the integer scores, i.e., the LDPS model, both fit these data well.

From Theorem 1, we see that the poor fit of the S model is caused by the influence of the lack of structure of the ME model rather than the RQS model.

Also, we see from Table 4a that under the RQS model, the differences between the maximum likelihood estimates of ridits are $\hat{r}_2^X - \hat{r}_1^X = 0.283$, $\hat{r}_3^X - \hat{r}_2^X = 0.315$, $\hat{r}_4^X - \hat{r}_3^X = 0.217$, and $\hat{r}_2^Y - \hat{r}_1^Y = 0.276$, $\hat{r}_3^Y - \hat{r}_2^Y = 0.316$, $\hat{r}_4^Y - \hat{r}_3^Y = 0.224$. These seems to be close to a constant. Namely, $\hat{v}_{i+1} - \hat{v}_i$ for $i = 1, 2, 3$, seems to be close to an equal-interval (Table 4). Thus, the RQS model applied to these unaided vision data would have a similar structure to the LDPS model. Indeed, for these data, the $G^2$ value for the RQS model is close to that for the LDPS model, and the maximum likelihood estimates of expected frequencies under the RQS model are close to the corresponding those under the LDPS model (see Table 1).

Under the RQS model, the maximum likelihood estimate of $\theta$ is $\hat{\theta} = 1.459$, which is greater than 1. Noting that $\hat{v}_j - \hat{v}_i$ are positive for all $i < j$, we see that $\hat{\theta}^{(j)-v}$ for $i < j$ are greater than 1. Therefore, under the RQS model, the probability that a woman’s right eye grade is $j$ and her left eye grade is $i$ ($j > i$) is estimated to be $\hat{\theta}^{(j)-v}$ ($> 1$) times higher than the probability that the woman’s right eye grade is $j$ and her left eye grade is $i$. In addition, since $\hat{\theta} > 1$, the marginal probability that a woman’s right eye is $i$ or below ($i = 1, 2, 3$) is estimated to be greater than the marginal probability that the woman’s left eye is $i$ or below. Thus, under the RQS model, a woman’s right eye is estimated to be better than her left eye.

**Example 2:** Table 2 taken directly from Agresti (1984, p. 206) is the father’s and son’s occupational mobility data in Britain. These data have been analyzed by some statisticians, e.g., including Bishop et al. (1975, p. 100), Agresti (1984, pp. 205-206) and Yamamoto et al. (2007).

We see from Table 3 that the S, LDPS, MH and ME models fit these data poorly, however, the QS and RQS models fit these data well. Thus, we see that the QS model with the ridit scores, i.e., the RQS model, fits these data well, although the QS model with the integer scores, i.e., the LDPS model, fits these data poorly.

We see from Theorem 1 that the poor fit of the S model is caused by the influence of the lack of structure of the ME model rather than the RQS model.

We see from Table 4b that under the RQS model, the differences between the maximum likelihood estimates of ridits are $\hat{r}_2^X - \hat{r}_1^X = 0.092$, $\hat{r}_3^X - \hat{r}_2^X = 0.148$, $\hat{r}_4^X - \hat{r}_3^X = 0.283$, $\hat{r}_3^X - \hat{r}_4^X = 0.335$, and $\hat{r}_2^Y - \hat{r}_1^Y = 0.082$, $\hat{r}_3^Y - \hat{r}_2^Y = 0.132$, $\hat{r}_4^Y - \hat{r}_3^Y = 0.275$, $\hat{r}_3^Y - \hat{r}_4^Y = 0.353$. Thus, the differences between the ridits and $\hat{v}_{i+1} - \hat{v}_i$ for $i = 1, \ldots, 4$ are unlikely to be constant (Table 4). Therefore, the RQS model applied to the father-son pair data in Table 2 would not have a similar structure to the LDPS model. Indeed, for these data, the $G^2$ value for the RQS model is not close to that for the LDPS model, and the RQS model fits these data well, however, the LDPS model fits these data poorly.

Under the RQS model, the maximum likelihood estimate of $\theta$ is $\hat{\theta} = 1.753$, which is
Table 2: Occupational status for father/son pairs; from Agresti (1984, p. 206). The parenthesized values are the maximum likelihood estimates of expected frequencies under the RQS model.

<table>
<thead>
<tr>
<th>Father’s status</th>
<th>Son’s status</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td>50</td>
<td>45</td>
<td>8</td>
<td>18</td>
<td>8</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(50.03)</td>
<td>(37.38)</td>
<td>(10.09)</td>
<td>(18.27)</td>
<td>(6.79)</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td>28</td>
<td>174</td>
<td>84</td>
<td>154</td>
<td>55</td>
<td>495</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(35.60)</td>
<td>(173.80)</td>
<td>(84.00)</td>
<td>(169.74)</td>
<td>(58.74)</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td>11</td>
<td>78</td>
<td>110</td>
<td>223</td>
<td>96</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(8.88)</td>
<td>(77.66)</td>
<td>(109.66)</td>
<td>(219.66)</td>
<td>(98.47)</td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td>14</td>
<td>150</td>
<td>185</td>
<td>714</td>
<td>447</td>
<td>1510</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(13.75)</td>
<td>(134.13)</td>
<td>(187.76)</td>
<td>(714.20)</td>
<td>(420.76)</td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td></td>
<td>3</td>
<td>42</td>
<td>72</td>
<td>320</td>
<td>411</td>
<td>848</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.22)</td>
<td>(38.28)</td>
<td>(69.41)</td>
<td>(347.01)</td>
<td>(411.71)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>106</td>
<td>489</td>
<td>459</td>
<td>1429</td>
<td>1017</td>
<td>3500</td>
</tr>
</tbody>
</table>

Table 3: Likelihood ratio values $G^2$ for models applied to the data in Tables 1 and 2.

<table>
<thead>
<tr>
<th>Applied models</th>
<th>Table 1</th>
<th>Degree of freedom</th>
<th>$G^2$</th>
<th>Table 2</th>
<th>Degree of freedom</th>
<th>$G^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>6</td>
<td>19.25*</td>
<td>10</td>
<td>37.46*</td>
<td>6</td>
<td>4.66</td>
</tr>
<tr>
<td>QS</td>
<td>3</td>
<td>7.27</td>
<td>10</td>
<td>17.13*</td>
<td>3</td>
<td>6.56</td>
</tr>
<tr>
<td>LDPS</td>
<td>5</td>
<td>7.28</td>
<td>9</td>
<td>12.67</td>
<td>5</td>
<td>3.65</td>
</tr>
<tr>
<td>RQS</td>
<td>5</td>
<td>7.32</td>
<td>9</td>
<td>12.67</td>
<td>5</td>
<td>3.65</td>
</tr>
<tr>
<td>MH</td>
<td>3</td>
<td>11.99*</td>
<td>4</td>
<td>32.80*</td>
<td>3</td>
<td>11.98</td>
</tr>
<tr>
<td>ME</td>
<td>1</td>
<td>11.98*</td>
<td>1</td>
<td>20.28*</td>
<td>1</td>
<td>11.98</td>
</tr>
</tbody>
</table>

* means significant at 5% level.

greater than 1. We see that $\hat{\theta}_{ij} - \hat{\theta}_i$ for $i < j$ are greater than 1. Therefore, under the RQS model, the probability that the status category for the father in a pair is $i$ and that for his son is $j$ ($> i$), is estimated to be $\hat{\theta}_{ij} - \hat{\theta}_i$ ($> 1$) times higher than the probability that the status category for the father is $j$ and that for his son is $i$. In addition, since $\hat{\theta} > 1$, the marginal probability that the status category for the father is $i$ or below ($i = 1, \ldots, 4$) is estimated to be greater than the marginal probability that the status category for his son is $i$ or below. Namely, the status category for the father rather than that for his son tends to be $i$ or below ($i = 1, \ldots, 4$).

5 Concluding Remarks

The RQS model applied to the data is based on the marginal ridits. The RQS model may be appropriate if it is supposed that the categories of the ordinal row and column variables represent intervals of an underlying continuous distribution.
Table 4: Maximum likelihood estimates of ridits \( \{ \hat{r}_i^X \} \), \( \{ \hat{r}_i^Y \} \) and \( \{ \hat{v}_i \} \) under the RQS model applied to the data in Tables 1 and 2.

(a) For Table 1

| \( i \) | \( \hat{r}_i^X \) | \( \hat{r}_i^Y \) | \( \hat{v}_i \) | \( \hat{v}_{i+1} - \hat{v}_i \) |
|-------|---------------|---------------|-------------|-----------------
| 1     | 0.132         | 0.128         | 0.130       | 0.279           |
| 2     | 0.415         | 0.404         | 0.409       | 0.316           |
| 3     | 0.730         | 0.720         | 0.725       | 0.221           |
| 4     | 0.947         | 0.944         | 0.946       | —               |

(b) For Table 2

| \( i \) | \( \hat{r}_i^X \) | \( \hat{r}_i^Y \) | \( \hat{v}_i \) | \( \hat{v}_{i+1} - \hat{v}_i \) |
|-------|---------------|---------------|-------------|-----------------
| 1     | 0.018         | 0.016         | 0.017       | 0.087           |
| 2     | 0.110         | 0.098         | 0.104       | 0.140           |
| 3     | 0.258         | 0.230         | 0.244       | 0.279           |
| 4     | 0.541         | 0.505         | 0.523       | 0.344           |
| 5     | 0.876         | 0.858         | 0.867       | —               |

The readers may be interested in why one should use the RQS model instead of the QS model. The QS model indicates the structure that the odds-ratios are symmetric with respect to the main diagonal of the table. However, under the QS model we cannot infer that \( X \) is stochastically less than \( Y \) or vice versa. On the other hand, the RQS model implies the QS model and the parameter \( \theta \) in the RQS model would be useful for making inferences such as that \( X \) is stochastically less than \( Y \) or vice versa (see Sections 2 and 4). Also, the QS model is considered for nominal categorical data, and the RQS model should be considered for ordinal categorical data, because the RQS model is not invariant under the same arbitrary permutations of the row and column categories.

Moreover, the RQS model rather than the LDPS (OQS) model may be appropriate when we cannot assign the integer scores (or known scores \( u_1 < \cdots < u_R \)) to the categories for both row and column of same classifications.

Each of S, QS, LDPS (OQS), MH and ME models is saturated on the main diagonal cells of the table, but the RQS model is unsaturated on them. Thus, under the RQS model, the estimated expected frequencies on the main diagonal are always not equal to the observed frequencies on the main diagonal (see Tables 1 and 2). The RQS model may be useful when we want to utilize the information on the main diagonal.

The decomposition of the S model into the RQS and ME models, given by Theorem 1, would be useful for seeing the reason for its poor fit when the S model fits the data poorly. Indeed, for the data in Table 1, the poor fit of the S model is caused by the poor fit of the ME model rather than the RQS model, i.e., by the reason that the mean of grade of the right eye is different from the mean of grade of the left eye (see Example 1).

**Acknowledgements**

The authors would like to thank a referee for the helpful comments.
Appendix

We consider the maximum likelihood estimates of expected frequencies \( \{m_{ij}\} \) under the RQS model. We must maximize the Lagrangian

\[
L = \sum_{i=1}^{R} \sum_{j=1}^{R} n_{ij} \log p_{ij} - \lambda \left( \sum_{i=1}^{R} \sum_{j=1}^{R} p_{ij} - 1 \right) - \sum_{i<j} \phi_{ij} (p_{ij} - \theta^{v_j - v_i} p_{ji})
\]

with respect to \( \{p_{ij}\}, \lambda, \{\phi_{ij}\}, \) and \( \theta \). Setting the partial derivations of \( L \) equal to zero, we obtain the equations: for \( s < t \),

\[
m_{st} = n_{st} - \phi_{st} \frac{m_{st}}{n} + \sum_{i<j} \phi_{ij} \frac{m_{ji}}{n} (\log \theta) \left[ \frac{\partial(v_j - v_i)}{\partial p_{st}} \right] \theta^{v_j - v_i} \frac{m_{st}}{n},
\]

\[
m_{ts} = n_{ts} + \phi_{st} \theta^{v_i - v_j} \frac{m_{ts}}{n} + \sum_{i<j} \phi_{ij} \frac{m_{ji}}{n} (\log \theta) \left[ \frac{\partial(v_j - v_i)}{\partial p_{ts}} \right] \theta^{v_j - v_i} \frac{m_{ts}}{n},
\]

\[
m_{ss} = n_{ss} + \sum_{i<j} \phi_{ij} \frac{m_{ji}}{n} (\log \theta) \left[ \frac{\partial(v_j - v_i)}{\partial p_{ss}} \right] \theta^{v_j - v_i} \frac{m_{ss}}{n},
\]

\[
m_{st} = \theta^{v_i - v_j} m_{ts},
\]

\[
\sum_{i<j} \phi_{ij} m_{ji} (v_j - v_i) \theta^{v_j - v_i - 1} = 0,
\]

where

\[
\left[ \frac{\partial(v_j - v_i)}{\partial p_{kl}} \right] = \frac{1}{2} \left\{ I(j, k) + I(j, l) - I(i, k) - I(i, l) \right\},
\]

\[
I(a, b) = \begin{cases} 
1 & \text{if } a > b, \\
1/2 & \text{if } a = b, \\
0 & \text{if } a < b,
\end{cases}
\]

\[
m_{kl} = np_{kl}.
\]

Using the Newton-Raphson method, we can solve the above equations with respect to \( \{m_{st}\}, \{\phi_{ij}\} \) and \( \theta \). Thus we obtain the maximum likelihood estimates of \( \{m_{ij}\} \) and the parameter \( \theta \) under the RQS model.

References


Authors’ Address

Kiyotaka Iki, Kouji Tahata and Sadao Tomizawa
Department of Information Sciences
Faculty of Science and Technology
Tokyo University of Science
Noda City
Chiba 278-8510
Japan

E-Mails: iki.kiyotaka@gmail.com, kouji_tahata@is.noda.tus.ac.jp and tomizawa@is.noda.tus.ac.jp