MEASURES OF DEPARTURE FROM ORDINAL QUASI-SYMMETRY MODELS FOR SQUARE CONTINGENCY TABLES

Kouji Yamamoto*, Shuji Ando† and Sadao Tomizawa‡

ABSTRACT

For square contingency tables with ordered categories, Agresti (2002) considered the ordinal quasi-symmetry (OQS) model and Iki, Tahata and Tomizawa (2009) considered the ridit score type quasi-symmetry (RQS) model. The present paper proposes measures which represent the degree of departure from each of the OQS and RQS models. The proposed measures are expressed by using the Cressie-Read power-divergence or Patil-Taillie diversity index. These measures would be useful for comparing the degrees of departure from OQS and RQS in several tables. The measures are applied to the data of individual’s education and father’s or mother’s education in Japan.

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 \( (i = 1, \ldots, R; j = 1, \ldots, R) \). Bowker (1948) considered the symmetry (S) model, defined by

\[
p_{ij} = p_{ji} \quad (i \neq j).
\]

This indicates that the probability that an observation will fall in the \((i, j)\)th cell, \(i \neq j\), is equal to the probability that it falls in the \((j, i)\)th cell.

Caussinus (1965) considered the quasi-symmetry (QS) model, defined by

\[
p_{ij} = \mu_{i} \beta_{j} \psi_{ij} \quad (i = 1, \ldots, R; j = 1, \ldots, R),
\]

where \( \psi_{ij} = \psi_{ji} \). This model with \( \{\alpha_{i} = \beta_{i}\} \) is identical to the S model (Bishop, Fienberg and Holland, 1975, p. 282).

When we can assign ordered known scores \( u_{1} < \cdots < u_{R} \), Agresti (2002, p. 429) proposed the ordinal quasi-symmetry (OQS) model, defined by

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

*Center for Clinical Investigation and Research, Osaka University Hospital, 2-15, Yamadaoka, Suita, Osaka, 565-0871, Japan E-mail: yamamoto-k@hp-crc.med.osaka-u.ac.jp
†Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba 278-8510, Japan
‡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

Key words: Kullback-Leibler information; Measure; Ordered category; Power-divergence; Quasi-symmetry; Ridit; Shannon entropy
where $\psi_{ij} = \psi_{ji}$. The OQS model with $\{u_i = i\}$ is identical to the linear diagonals-parameter symmetry (LDPS) model, proposed by Agresti (1983). Also the OQS model with $\delta = 1$ is identical to the S model. The OQS model is a simple asymmetry model and the parameter $\delta$ indicates the degree of asymmetry. Denote the row and column variables by $X$ and $Y$, respectively. Under the LDPS (OQS) model, $\delta \geq 1$ is equivalent to $P(X \leq i) \geq P(Y \leq i)$ for every $i = 1, \ldots, R - 1$. Therefore the parameter $\delta$ in the LDPS (OQS) model would be useful for making inferences on whether $X$ is stochastically less than $Y$ or vice versa.

Let

$$r_i^X = \sum_{k=1}^{i-1} p_{ik} + \frac{p_{ii}}{2}, \quad r_i^Y = \sum_{l=1}^{i-1} p_{li} + \frac{p_{ii}}{2} \quad (i = 1, \ldots, R),$$

where $p_{ii} = \sum_{r=1}^{R} p_{ir}$ and $p_{ii} = \sum_{s=1}^{R} p_{si}$. The $\{r_i^X\}$ and $\{r_i^Y\}$ are the marginal ridits; see Bross (1958). When it may be difficult to assign known scores $\{u_i\}$ for the given data, Iki, Tahata and Tomizawa (2009) considered the ridit score type quasi-symmetry (RQS) model, defined by

$$p_{ij} = \begin{cases} 
\theta^{v_j - v_i} \psi_{ij} & (i < j), \\
\psi_{ij} & (i \geq j), 
\end{cases}$$

where $\psi_{ij} = \psi_{ji}$ and $v_i = (r_i^X + r_i^Y)/2$, for $i = 1, \ldots, R$. Note that $\{v_i\}$ are unknown scores. The RQS model with $\theta = 1$ is identical to the S model. Also the RQS model is identical to the LDPS model with $\{i\}$ replaced by the ridit scores $\{v_i\}$.

The RQS model may be expressed as

$$\log \left( \frac{p_{ij}}{p_{ji}} \right) = (v_j - v_i) \log \theta \quad (i < j).$$

Thus when the RQS model is applied to square table data with the same row and column classifications, this model indicates that the log odds ratio of symmetric cell probabilities is proportional to the difference between the (average) rank scores for categories.

For square contingency tables with nominal categories, Tahata, Miyamoto and Tomizawa (2004) considered a measure which represents the degree of departure from the QS model. For square tables with ordered categories, Yamamoto and Tomizawa (2008) proposed a measure which represents the degree of departure from the LDPS model. Each of the measures in Tahata et al. (2004) and in Yamamoto and Tomizawa (2008) is useful to represent what degree the departure from the model is toward the maximum departure from it. We point out that the test statistic (e.g., Pearson’s chi-squared statistic or likelihood ratio statistic) is used for testing the goodness-of-fit of the model and the measure (e.g., correlation coefficient) is used for measuring the degree of departure from the model toward the maximum departure from it.

For the analysis of square contingency tables with the same row and column classifications, if it is possible to assign integer scores (or equal-interval scores), we can apply the LDPS model and the measure in Yamamoto and Tomizawa (2008) for representing the degree of departure from the LDPS model. If it is possible to assign known scores $\{u_i\}$, we can apply the OQS model. We are then interested in proposing a measure to represent the degree of departure from the OQS model toward the maximum departure from it.
If it is difficult to assign integer scores, or known scores for the categories, we are interested in applying the model with the ridits scores being unknown, instead of the LDPS and OQS models, and also interested in proposing a measure to represent the degree of departure from the RQS model toward the maximum departure from it.

The purpose of this paper is to propose two measures which represent the degree of departure from each of the RQS and OQS models. These measures would be useful for comparing the degrees of departure from the corresponding model in several tables.

2. Measures of departure from the RQS and OQS models

2.1. Measure for the RQS model

For a fixed $d$ ($d = 1, \ldots, R - 1$), the RQS model may be expressed as

$$Q_{ij}^{(d)} = \frac{Q_{ij}^{(d)} + Q_{ji}^{(d)}}{2}$$

where

$$Q_{ij}^{(d)} = (p_{ij})^{\frac{d}{d-1}}, \quad Q_{ji}^{(d)} = (p_{ji})^{\frac{d}{d-1}}.$$  

Namely, for a fixed $d$, the ratios $\{Q_{ij}^{(d)} / Q_{ji}^{(d)}, i < j\}$ are constant. Note that $v_1 < \cdots < v_R$ are unknown scores defined in Section 1. When the RQS model does not hold, we shall consider the measure to represent the degree of departure from the RQS model. By the way, Tomizawa and Saitoh (1999) considered the measure to represent the degree of departure from McCullagh’s (1978) conditional symmetry model which indicates the constant of symmetric odds ratios, i.e., $\{p_{ij} / p_{ji} = \theta, i < j\}$. Note that the measure in Tomizawa and Saitoh (1999) is based on the power-divergence (including the well-known Kullback-Leibler information as a special case) and Patil and Taillie’s (1982) diversity index (including the well-known Shannon entropy as a special case). Since the RQS model indicates the constant of symmetric odds ratios $\{Q_{ij}^{(d)} / Q_{ji}^{(d)} = \theta, i < j\}$ for each $d$ ($d = 1, \ldots, R - 1$), as the measure to represent the degree of departure from the RQS model, we shall first consider the submeasure ($\psi^{(d)}_{ij}$ denoted below) representing the degree of departure from the constant of symmetric odds ratios and next consider the measure for the RQS model by the sum of submeasure for $d = 1, \ldots, R - 1$.

For a fixed $d$ ($d = 1, \ldots, R - 1$), let

$$\Delta_{ij}^{(d)} = \sum_{i<j} Q_{ij}^{(d)}, \quad \Delta_{ji}^{(d)} = \sum_{i<j} Q_{ji}^{(d)}.$$

Then for a fixed $d$ ($d = 1, \ldots, R - 1$), the RQS model is expressed as

$$Q_{ij}^{U(d)} = Q_{ji}^{L(d)} \quad (i < j).$$
or as

\[ Q_{ij}^{c(d)} = Q_{ji}^{c(d)} \left( \frac{1}{2} \right) \quad (i < j). \]

Assuming that \{p_{ij} + p_{ji} > 0\}, \{\Delta_{ij}^{(d)} > 0\} and \{\Delta_{ii}^{(d)} > 0\}, consider a measure to represent the degree of departure from the RQS model, defined by

\[ \psi^{(\lambda)} = \frac{1}{R - 1} \sum_{d=1}^{R-1} \psi_{d}^{(\lambda)} \quad (\lambda > -1), \]

where

\[ \psi_{d}^{(\lambda)} = \frac{\lambda(\lambda + 1)}{2^\lambda - 1} \sum_{i<j} \left( Q_{ij}^{U(d)} + Q_{ji}^{L(d)} \right) \frac{1}{2} I_{ij}^{(\lambda)} \left( \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\}; \{1/2, 1/2\} \right), \]

with

\[ I_{ij}^{(\lambda)} (: : ) = \frac{1}{\lambda(\lambda + 1)} \left[ Q_{ij}^{c(d)} \left( \left( \frac{Q_{ij}^{c(d)}}{1/2} \right)^{\lambda} - 1 \right) + Q_{ji}^{c(d)} \left( \left( \frac{Q_{ji}^{c(d)}}{1/2} \right)^{\lambda} - 1 \right) \right]. \]

and the value at \( \lambda = 0 \) is taken to be the continuous limit as \( \lambda \to 0 \). Thus,

\[ \psi_{d}^{(0)} = \frac{1}{\log 2} \sum_{i<j} \left( Q_{ij}^{U(d)} + Q_{ji}^{L(d)} \right) I_{ij}^{(0)} \left( \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\}; \{1/2, 1/2\} \right), \]

where

\[ I_{ij}^{(0)} (: : ) = Q_{ij}^{c(d)} \log \left( \frac{Q_{ij}^{c(d)}}{1/2} \right) + Q_{ji}^{c(d)} \log \left( \frac{Q_{ji}^{c(d)}}{1/2} \right). \]

Note that \( I_{ij}^{(\lambda)} \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\}; \{1/2, 1/2\} \) is the power-divergence between \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\} and \{1/2, 1/2\}, and especially \( I_{ij}^{(0)} \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\}; \{1/2, 1/2\} \) is the Kullback-Leibler information between them. For more details of the power-divergence \( I_{ij}^{(\lambda)} (: : ) \), see Cressie and Read (1984), and Read and Cressie (1988, p. 15).

The \( \psi_{d}^{(\lambda)} \) may also be expressed as

\[ \psi_{d}^{(\lambda)} = 1 - \frac{\lambda 2^\lambda}{2^\lambda - 1} \sum_{i<j} \left( Q_{ij}^{U(d)} + Q_{ji}^{L(d)} \right) H_{ij}^{(\lambda)} \left( \{Q_{ij}^{c(d)}, Q_{ji}^{c(d)}\} \right) \quad (\lambda > -1), \]

where

\[ H_{ij}^{(\lambda)} (: ) = \frac{1}{\lambda} \left[ 1 - (Q_{ij}^{c(d)})^{\lambda+1} - (Q_{ji}^{c(d)})^{\lambda+1} \right]. \]
and the value at \( \lambda = 0 \) is taken to be the continuous limit as \( \lambda \to 0 \). Thus,

\[
\psi^{(0)}_d = 1 - \frac{1}{\log 2} \sum_{i<j} Q^{(d)}_{ij} \frac{Q^{(d)}_{ij} + Q^{(d)}_{ji}}{2} H^{(0)}_{ij} \left( \{ Q^{(d)}_{ij}, Q^{(d)}_{ji} \} \right),
\]

where

\[
H^{(0)}_{ij} (\cdot ; \cdot) = -Q^{(d)}_{ij} \log Q^{(d)}_{ij} - Q^{(d)}_{ji} \log Q^{(d)}_{ji}.
\]

Note that \( H^{(0)}_{ij} (\{ Q^{(d)}_{ij}, Q^{(d)}_{ji} \}) \) is Patil and Taillie’s (1982) diversity index for \( \{ Q^{(d)}_{ij}, Q^{(d)}_{ji} \} \), which includes the Shannon entropy when \( \lambda = 0 \).

Note that \( \psi^{(0)}_d \) lies between 0 and 1, namely, the measure \( \Psi^{(0)} \) lies between 0 and 1. Also for each \( \lambda \), (i) \( \psi_d^{(0)} = 0 \) for every \( d = 1, \ldots, R - 1 \), i.e., \( \Psi^{(0)} \) = 0 if and only if the RQS model holds, and (ii) \( \psi_d^{(0)} = 1 \) for every \( d = 1, \ldots, R - 1 \), i.e., \( \Psi^{(0)} \) = 1 if and only if the degree of departure from RQS is the largest in the sense that \( Q^{(0)}_{ij} = 0 \) or \( Q^{(0)}_{ji} = 0 \) (namely, \( p_{ij} = 0 \) (then \( p_{ji} > 0 \)) or \( p_{ji} = 0 \) (then \( p_{ij} > 0 \)) for \( i < j \). According to the weighted sum of power-divergence or the Patil and Taillie’s diversity index, the measure \( \Psi^{(0)} \) represents the degree of departure from RQS, and the degree increases as the value of \( \Psi^{(0)} \) increases.

For analyzing the degree of departure from RQS, we first should check whether or not the RQS model holds by using a test statistic. Then, if it is judged that these is not a RQS model, we can consider a measure which represents the degree of departure from the OQS model. We shall consider the measure (denoted by \( \Phi^{(0)} \)) of \( \Psi^{(0)} \) (Section 3).

2.2. Measure for the OQS model

In a similar way to the measure \( \Psi^{(0)} \), we can consider a measure which represents the degree of departure from the OQS model. We shall consider the measure (denoted by \( \Phi^{(0)} \)) obtained by \( \Psi^{(0)} \) with unknown scores \( \{ v_i \} \) replaced by known scores \( \{ u_i \} \) (for \( \Phi^{(0)} \), see Appendix). We shall omit the details of properties for \( \Phi^{(0)} \) because the properties for \( \Phi^{(0)} \) are similar to those for \( \Psi^{(0)} \).

When known scores \( \{ u_i \} \) are integer scores (or equal-interval scores), the measure \( \Phi^{(0)} \) is identical to the measure for representing the degree of departure from the LDPS model, proposed by Yamamoto and Tomizawa (2008). We note that the LDPS model may be expressed as for a fixed \( d \) \( (d = 1, \ldots, R - 1) \),

\[
\left( \frac{p_{ij}}{p_{ji}} \right)^d = (\theta_d)^{j-i} \quad (i < j).
\]

When \( d = 1 \), the parameter \( \theta_d \) equals the odds \( p_{12}/p_{21} \) (or \( p_{i,i+1}/p_{i+1,i} \)), and when \( d = R - 1 \), the parameter \( \theta_{R-1} \) equals the odds \( p_{1,R}/p_{1,R-1} \). Yamamoto and Tomizawa (2008) considered the submeasure \( \phi^{(0)}_d \) for fixed \( d \) and considered the measure \( \Phi^{(0)} \) for representing the degree of departure from LDPS model by the sum of the submeasures \( \phi^{(0)}_d \) \( (d = 1, \ldots, R - 1) \). The measure \( \Psi^{(0)} \) representing the degree of departure from the RQS model is obtained by replacing integer scores \( \{ i \} \) by unknown ridit scores \( \{ v_i \} \).

3. Approximate confidence intervals for measures

Assuming that a multinomial distribution applies to the \( R \times R \) table, we shall consider an approximate standard error and large-sample confidence interval for the measure \( \Psi^{(0)} \), using
the delta method, of which descriptions are given by e.g., Bishop et al. (1975, Sec. 14.6). The sample version of $\Psi^{(\lambda)}$, i.e., $\hat{\Psi}^{(\lambda)}$, is given by $\Psi^{(\lambda)}$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, where $\hat{p}_{ij} = n_{ij}/n$ and $n = \sum n_{ij}$. Using the delta method, $\sqrt{n}(\hat{\Psi}^{(\lambda)} - \Psi^{(\lambda)})$ has asymptotically (as $n \to \infty$) a normal distribution with mean zero and variance $\sigma^2[\Psi^{(\lambda)}]$. The value of $\sigma^2[\Psi^{(\lambda)}]$ is

$$
\sigma^2[\Psi^{(\lambda)}] = \sum_{k=1}^{R} \sum_{l=1}^{R} (W^{(\lambda)\lambda}_{kl} p_{kl} - \left\{ \sum_{k=1}^{R} \sum_{l=1}^{R} W^{(\lambda)\lambda}_{kl} p_{kl} \right\}^2,
$$

where for $\lambda > -1; \lambda \neq 0$,

$$
W^{(\lambda)\lambda}_{kl} = \frac{\lambda^2}{2(R-1)(2\lambda-1)} \sum_{d=1}^{R-1} \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \left( A^{(d)}_{ij(kl)} + B^{(d)}_{ij(kl)} \right) H^{(\lambda)}_{ij}(\{\xi^{(d)}_{ij}, \xi^{(d)}_{ji}\}) - \frac{\lambda + 1}{\lambda} \left\{ A^{(d)}_{ij(kl)} \xi^{(d)}_{ij} \left( \xi^{(d)}_{ij} - \xi^{(d)}_{ji} \right) + B^{(d)}_{ij(kl)} \xi^{(d)}_{ij} \left( \xi^{(d)}_{ji} - \xi^{(d)}_{ij} \right) \right\},
$$

and for $\lambda = 0$,

$$
W^{(0)}_{kl} = \frac{1}{2(R-1)\log 2} \sum_{d=1}^{R-1} \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \left( A^{(d)}_{ij(kl)} + B^{(d)}_{ij(kl)} \right) H^{(\lambda)}_{ij}(\{\xi^{(d)}_{ij}, \xi^{(d)}_{ji}\}) - A^{(d)}_{ij(kl)} \xi^{(d)}_{ij} \left( \log \xi^{(d)}_{ij} - \log \xi^{(d)}_{ji} \right) - B^{(d)}_{ij(kl)} \xi^{(d)}_{ij} \left( \log \xi^{(d)}_{ji} - \log \xi^{(d)}_{ij} \right),
$$

with

$$
A^{(d)}_{ij(kl)} = \frac{C^{(d)}_{ij(kl)} \Delta^{(d)}_{kl} - \xi^{(d)}_{ij} \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} C^{(d)}_{ij(kl)} \Delta^{(d)}_{kl}}{\sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \Delta^{(d)}_{kl}},
$$

$$
B^{(d)}_{ij(kl)} = \frac{D^{(d)}_{ij(kl)} \Delta^{(d)}_{kl} - \xi^{(d)}_{ij} \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} D^{(d)}_{ij(kl)} \Delta^{(d)}_{kl}}{\sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \Delta^{(d)}_{kl}},
$$

$$
C^{(d)}_{ij(kl)} = \begin{cases} Q^{(d)}_{ij} \left( -d \frac{E(i = k, j = l)}{(v_j - v_i)^2} \log p_{ij} + I(i = k, j = l) \frac{d}{(v_j - v_i)p_{ij}} \right) & (k < l), \\ -Q^{(d)}_{ij} \frac{d}{(v_j - v_i)^2} \log p_{ij} & \text{(otherwise)}, \end{cases}
$$

$$
D^{(d)}_{ij(kl)} = \begin{cases} Q^{(d)}_{ij} \left( -d \frac{E(i = k, j = l)}{(v_j - v_i)^2} \log p_{ij} + I(i = l, j = k) \frac{d}{(v_j - v_i)p_{ij}} \right) & (k > l), \\ -Q^{(d)}_{ij} \frac{d}{(v_j - v_i)^2} \log p_{ij} & \text{(otherwise)}, \end{cases}
$$

$$
E_{i(kl)} = \frac{1}{2} \left\{ I(k \leq i - 1) + \frac{1}{2} I(k = i) + I(l \leq i - 1) + \frac{1}{2} I(l = i) \right\},
$$

and the indicator function $I(\cdot)$. 

— 32 —
Let $\hat{\sigma}^2[\Psi^{(\lambda)}]$ denote $\sigma^2[\Psi^{(\lambda)}]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. Then $\hat{\sigma}[\Psi^{(\lambda)}]/\sqrt{n}$ is an estimated standard error for $\Psi^{(\lambda)}$, and $\Psi^{(\lambda)} \pm z_p \hat{\sigma}[\Psi^{(\lambda)}]/\sqrt{n}$ is an approximate 100$(1-p)$ percent confidence interval for $\Psi^{(\lambda)}$, where $z_{p/2}$ is the percentage point from the standard normal distribution corresponding to a two-tail probability equal to $p$. Although the details are omitted here, we can discuss the measure $\Phi^{(\lambda)}$ similarly.

4. An example

Consider the data in Table 1, taken from the database that is on the web site (http://srdq.hus.osaka-u.ac.jp) by Seiyama: Social Stratification and Social Mobility Research Group (1995). These data describe the cross-classification of individual’s education and father’s education (Table 1a) or mother’s education (Table 1b) in Japan, examined in 1995. For Table 1, it seems difficult to assign specific known scores for the levels of individual’s, father’s, or mother’s education.

Table 1: The individual’s education and father’s or mother’s education data in Japan from the database that is on the web site (http://srdq.hus.osaka-u.ac.jp).

<table>
<thead>
<tr>
<th>(a) Individual’s education and father’s education data</th>
<th>Father’s education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual’s education</td>
<td>First (1)</td>
</tr>
<tr>
<td>First (1)</td>
<td>380</td>
</tr>
<tr>
<td>Middle (2)</td>
<td>740</td>
</tr>
<tr>
<td>High (3)</td>
<td>170</td>
</tr>
<tr>
<td>Total</td>
<td>1290</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Individual’s education and mother’s education data</th>
<th>Mother’s education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual’s education</td>
<td>First (1)</td>
</tr>
<tr>
<td>First (1)</td>
<td>427</td>
</tr>
<tr>
<td>Middle (2)</td>
<td>782</td>
</tr>
<tr>
<td>High (3)</td>
<td>189</td>
</tr>
<tr>
<td>Total</td>
<td>1398</td>
</tr>
</tbody>
</table>

First we are interested in applying the S model, which indicates that, e.g., for the data in Table 1, the probability that the levels of individual’s and his/her father’s (or mother’s) educations are $i$ and $j$, respectively, $i < j$, is equal to the probability that those are $j$ and $i$, respectively. We see from Table 2 that the S model fits each of Tables 1a and 1b poorly. Hence, we are next interested in applying the models of asymmetry instead of symmetry. We are interested in applying the LDPS model if it is possible to assign integer scores (or equal-interval scores) or in applying the OQS model if it is possible to assign known scores. However, it would be difficult to assign integer scores, or known scores for the levels of education, ‘First’ (junior high school), ‘Middle’ (high school), and ‘High’ (university). For instance, it may not be appropriate to assume that the difference between the levels of ‘Middle’ and ‘First’ is equal to the difference between the levels of ‘High’ and ‘Middle’. Therefore for these education data, we are interested in applying the RQS model with
Table 2: Value of likelihood ratio statistic $G^2$ under the S and RQS models applied to Tables 1a and 1b.

<table>
<thead>
<tr>
<th>Applied models</th>
<th>Degrees of freedom</th>
<th>$G^2$</th>
<th>Table 1a</th>
<th>Table 1b</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>3</td>
<td>966.82*</td>
<td>1334.56*</td>
<td></td>
</tr>
<tr>
<td>RQS</td>
<td>2</td>
<td>85.46*</td>
<td>2.97</td>
<td></td>
</tr>
</tbody>
</table>

* means significant at the 0.05 level.

unknown scores $\{v_i\}$.

We see from Table 2 that the RQS model fits the data in Table 1a poorly, but it fits the data in Table 1b well. Under the RQS model applied to Table 1b, the maximum likelihood estimates of (average) rank scores for levels of education are estimated to be $\bar{v}_1 = 0.206$ (for 'First'), $\bar{v}_2 = 0.629$ (for 'Middle'), $\bar{v}_3 = 0.923$ (for 'High'). Thus we see that the probability that the levels of an individual's and his/her mother's educations are $i$ and $j$ ($i < j$), respectively, is estimated to be $\theta^{i-j}$ times higher than the probability that those are $j$ and $i$, respectively.

Next, we shall measure to what degree the departure from the RQS model is toward the maximum departure from the RQS, using measure $\Psi^{(\lambda)}$, because it is impossible to measure it by the test statistic. Also, we see from $\Psi^{(0)}$ that for Table 1a, the degree of departure from RQS is estimated to be 52.2 percent of the maximum degree of departure from RQS (see Table 3a).

On the other hand, the confidence intervals for $\Psi^{(\lambda)}$ applied to the data in Table 1b

Table 3: Estimate of $\Psi^{(\lambda)}$, estimated approximate standard errors for $\Psi^{(\lambda)}$, and approximate 95% confidence intervals for $\Psi^{(\lambda)}$, applied to Tables 1a and 1b.

(a) For Table 1a

<table>
<thead>
<tr>
<th>Values of $\lambda$</th>
<th>Estimated measure</th>
<th>Standard error</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.4</td>
<td>0.425</td>
<td>0.056</td>
<td>(0.316, 0.534)</td>
</tr>
<tr>
<td>0</td>
<td>0.522</td>
<td>0.056</td>
<td>(0.412, 0.632)</td>
</tr>
<tr>
<td>0.6</td>
<td>0.585</td>
<td>0.055</td>
<td>(0.477, 0.692)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.601</td>
<td>0.054</td>
<td>(0.495, 0.707)</td>
</tr>
<tr>
<td>1.6</td>
<td>0.606</td>
<td>0.054</td>
<td>(0.500, 0.712)</td>
</tr>
</tbody>
</table>

(b) For Table 1b

<table>
<thead>
<tr>
<th>Values of $\lambda$</th>
<th>Estimated measure</th>
<th>Standard error</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.4</td>
<td>0.010</td>
<td>0.039</td>
<td>(−0.067, 0.087)</td>
</tr>
<tr>
<td>0</td>
<td>0.014</td>
<td>0.055</td>
<td>(−0.093, 0.122)</td>
</tr>
<tr>
<td>0.6</td>
<td>0.018</td>
<td>0.068</td>
<td>(−0.115, 0.151)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.019</td>
<td>0.072</td>
<td>(−0.121, 0.160)</td>
</tr>
<tr>
<td>1.6</td>
<td>0.020</td>
<td>0.073</td>
<td>(−0.124, 0.163)</td>
</tr>
</tbody>
</table>
Measures of Departure from Ordinal Quasi-symmetry Models

include zero for all \( \lambda \) (see Table 3b), these would indicate that there is a structure of RQS in Table 1b; or, if this is not the case, then it indicates that the degree of departure from RQS is slight.

In addition, we shall compare the degrees of departure from RQS in Tables 1a and 1b using the confidence intervals for \( \Psi^{(\lambda)} \). For any given \( \lambda (> -1) \), the values in the confidence interval for \( \Psi^{(\lambda)} \) applied to the data in Table 1a are greater than the values in the corresponding confidence interval for \( \Psi^{(\lambda)} \) applied to the data in Table 1b. We point out that for any \( \lambda \), the values in the confidence interval do not overlap for Table 1a and for Table 1b. Thus the degree of departure from the RQS model in Table 1a is greater than that in Table 1b.

5. Discussions

We point out that the submeasure \( \psi_d^{(\lambda)} \) in \( \hat{\Psi}^{(\lambda)} \) can be expressed as

\[
\psi_d^{(\lambda)} = \frac{\lambda(\lambda + 1)}{2^{\lambda+1}} \cdot \frac{1}{2} \left[ f^{(\lambda)} \left( \{Q_{ij}^{U(d)}\}; \{E_{ij}^{(d)}\} \right) + f^{(\lambda)} \left( \{Q_{ji}^{L(d)}\}; \{E_{ji}^{(d)}\} \right) \right],
\]

where

\[
E_{ij}^{(d)} = \frac{Q_{ij}^{U(d)} + Q_{ji}^{L(d)}}{2},
\]

\[
I^{(\lambda)} \left( \{a_{ij}\}; \{b_{ij}\} \right) = \frac{1}{\lambda(\lambda + 1)} \sum_{i<j} a_{ij} \left[ \left( \frac{a_{ij}}{b_{ij}} \right)^\lambda - 1 \right].
\]

Then it is easily seen that the submeasure \( \psi_d^{(0)} \) can be expressed as

\[
\psi_d^{(0)} = \frac{1}{2} \log 2 \min_{\{D_{ij}^{(d)}\}} \left[ I^{(0)} \left( \{Q_{ij}^{U(d)}\}; \{D_{ij}^{(d)}\} \right) + I^{(0)} \left( \{Q_{ji}^{L(d)}\}; \{D_{ji}^{(d)}\} \right) \right],
\]

where

\[
\sum_{i<j} D_{ij}^{(d)} = 1 \quad \text{and} \quad D_{ij}^{(d)} > 0.
\]

We note that \( E_{ij}^{(d)} \) in \( \psi_d^{(\lambda)} \) is the value of \( D_{ij}^{(d)} \) such that the sum of Kullback-Leibler (KL) distance (i.e., the KL distance between \( \{Q_{ij}^{U(d)}\} \) and \( \{D_{ij}^{(d)}\} \) with a structure of RQS and the KL distance between \( \{Q_{ji}^{L(d)}\} \) and \( \{D_{ji}^{(d)}\} \)) is a minimum. We note that readers may be interested in (1) with \( I^{(0)} \) replaced by the power-divergence \( I^{(\lambda)} \); however, it is difficult to obtain the values of \( \{D_{ij}^{(d)}\} \) such that the sum of the corresponding two power-divergence is a minimum, and difficult to obtain the maximum value of such a measure.

Readers may also be interested in which value of \( \lambda \) is preferred for a given table. However, it would be difficult to discuss it. For the analysis of data, it seems to be important and safe that for comparing the degrees of departure from the RQS in several tables, the user calculates the values of \( \hat{\Psi}^{(\lambda)} \) for various values of \( \lambda \) and discusses the degree of departure from the RQS in terms of them, rather than calculating \( \hat{\Psi}^{(\lambda)} \) for only one specified value of \( \lambda \). However, if the analyst wants to choose one value of \( \lambda \), the case of \( \lambda = 0 \), i.e., \( \hat{\Psi}^{(0)} \) may be recommended in terms of expression (1).
Finally, we shall compare the measure $\hat{\Psi}(\lambda)$ with the test statistic (e.g., likelihood ratio statistic). Consider the artificial data in Table 4. When the RQS model is applied to the data in Tables 4a, 4b and 4c, the values of the likelihood ratio statistic are 8.73, 31.39 and 50.97, respectively, with 2 degrees of freedom. The estimated values of $\{v_i\}$, $i = 1, 2, 3$, obtained under the saturated model are 1/6, 3/6 and 5/6 for all Tables 4a, 4b and 4c. On the other hand, for any fixed $\lambda (> -1)$, the value of $\hat{\Psi}(\lambda)$ is greater for Table 4a than for Table 4b, and it is greater for Table 4b than for Table 4c (see Table 5). In terms of $(\hat{p}_{ij}/\hat{p}_{ji})$, $i < j$ (see Table 4), it seems natural to conclude that the degree of departure from RQS is greater for Table 4a than for Table 4b, and it is greater for Table 4b than for Table 4c. Therefore, $\hat{\Psi}(\lambda)$ may be preferable to the test statistic for comparing the degree of departure from RQS in several tables.

Table 4: Artificial data ($n$ is sample size).

<table>
<thead>
<tr>
<th></th>
<th>(a) $n = 21$</th>
<th>(b) $n = 123$</th>
<th>(c) $n = 453$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 5 1</td>
<td>10 1 30</td>
<td>50 100 50</td>
</tr>
<tr>
<td></td>
<td>1 1 5</td>
<td>10 1 30</td>
<td>100 50 1</td>
</tr>
<tr>
<td></td>
<td>5 1 1</td>
<td>30 10 1</td>
<td></td>
</tr>
</tbody>
</table>

Note: $\hat{p}_{21}^{(a)} = 5$, $\hat{p}_{31}^{(a)} = 0.2$, $\hat{p}_{32}^{(a)} = 5$.

Table 5: The values of $\hat{\Psi}(\lambda)$ applied to Tables 4a, 4b and 4c.

<table>
<thead>
<tr>
<th>Values of $\lambda$</th>
<th>Table 4a</th>
<th>Table 4b</th>
<th>Table 4c</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4</td>
<td>0.483</td>
<td>0.260</td>
<td>0.096</td>
</tr>
<tr>
<td>0</td>
<td>0.541</td>
<td>0.296</td>
<td>0.116</td>
</tr>
<tr>
<td>0.6</td>
<td>0.583</td>
<td>0.319</td>
<td>0.130</td>
</tr>
<tr>
<td>1.0</td>
<td>0.595</td>
<td>0.325</td>
<td>0.133</td>
</tr>
<tr>
<td>1.6</td>
<td>0.598</td>
<td>0.327</td>
<td>0.134</td>
</tr>
</tbody>
</table>
Acknowledgements

The authors would like to express their thanks to anonymous referees for their helpful suggestions and comments.

Appendix

For a fixed $d$ ($d = 1, \ldots, R - 1$), let

$$q_{ij}^{(d)} = (p_{ij})^{rac{d}{u_{ij} - u_{ji}}}, \quad q_{ji}^{(d)} = (p_{ji})^{rac{d}{u_{ij} - u_{ji}}}, \quad (i < j),$$

and

$$\delta_U^{(d)} = \sum_{i<j} q_{ij}^{(d)}, \quad \delta_L^{(d)} = \sum_{i<j} q_{ji}^{(d)}.$$

Assuming that $\{p_{ij} + p_{ji} > 0\}$, $\{\delta_U^{(d)} > 0\}$, $\{\delta_L^{(d)} > 0\}$, consider a measure to represent the degree of departure from the OQS model, defined by

$$\Phi(\lambda) = \frac{1}{R - 1} \sum_{d=1}^{R-1} \phi_d^{(\lambda)} \quad (\lambda > -1),$$

where

$$\phi_d^{(\lambda)} = \frac{\lambda(\lambda + 1)}{2^\lambda - 1} \sum_{i<j} \left( \frac{q_{ij}^{(d)}}{q_{ji}^{(d)}} \right)^{\frac{\lambda}{2}} I_{ij}^{(\lambda)} \left( \left\{ q_{ij}^{c(d)}, q_{ji}^{c(d)} \right\}; \{1/2, 1/2\} \right),$$

with

$$I_{ij}^{(\lambda)}(\cdot, \cdot) = \frac{1}{\lambda(\lambda + 1)} \left[ \left( \frac{q_{ij}^{c(d)}}{q_{ji}^{c(d)}} \right)^{\frac{\lambda}{2}} - 1 \right] + \frac{\lambda}{2} \left( \frac{q_{ij}^{c(d)}}{q_{ji}^{c(d)}} \right)^{\frac{\lambda}{2}} - 1 \right],$$

and

$$q_{ij}^{U(d)} = \frac{q_{ij}^{(d)}}{\delta_U^{(d)}}, \quad q_{ij}^{L(d)} = \frac{q_{ij}^{(d)}}{\delta_L^{(d)}}, \quad q_{ij}^{c(d)} = \frac{q_{ij}^{U(d)}}{q_{ij}^{U(d)} + q_{ij}^{L(d)}}, \quad q_{ij}^{c(d)} = \frac{q_{ij}^{L(d)}}{q_{ij}^{U(d)} + q_{ij}^{L(d)}}, \quad (i < j),$$

and the value at $\lambda = 0$ is taken to be the continuous limit as $\lambda \to 0$.

REFERENCES


(Received: August 4, 2010, Accepted: February 10, 2011)