For the well-known Danish occupational mobility data describing, say, the cross-classification of father's and son's occupational status categories, the quasi-uniform association (QU) model considered by Goodman (1979, 1981b) fits well. This paper introduces some parsimonious QU models in which the odds of row and/or column effect parameters pertaining to adjacent categories are given by an exponential function. For the Danish data, the parsimonious QU models introduced in this paper are preferable to the original QU model, and new interpretations are made.

Key words and phrases: British data; exponential; father-son pairs; local odds-ratio; odds; row (column) effect parameter; square contingency table.

1. Introduction

For the analysis of two-way contingency tables with ordered categories, Goodman (1979) considered various kinds of association models, e.g., the uniform association model (the U model), in terms of the local odds-ratios, which are generalizations of the independence model (called the null association model). Agresti (1983, 1984) considered the association models in their more general forms. Goodman (1979) also observed that regular multiplicative models for ordinal variables sometimes fit square contingency tables well when the cells on the main diagonal are ignored. He proposed a model having uniform local association for cells off that diagonal. This model was called the quasi-uniform association model (the QU model). Similarly, a model having null association for cells off the main diagonal of the square table is called the quasi-independence model (the QO model); see Goodman (1972), Bishop et al. (1975, p. 178), and Agresti (1984, p. 204). Note that for the square contingency table, the QO model implies the QU model.

The data in Table 1, taken directly from Goodman (1981b), describe the cross-classification of father's and son's occupational status categories in Denmark. These data have been analyzed by several statisticians, including Mosteller (1968), Kullback (1971), Haberman (1974, p. 222), Bishop et al. (1975, p. 100), and Goodman (1972, 1981b, 1984). Goodman (1981b) showed that for the data in Table 1, the QU model fits well, providing a dramatic improvement, though the QO model fits very poorly.

The purpose of this paper is to introduce some parsimonious QU models and to show that the parsimonious QU models are preferable to the original QU model for the data in Table 1. Before giving them, we shall simply describe the U and QU models in the next section.

* Department of Information Sciences, Faculty of Science and Technology, Science University of Tokyo, Noda City, Chiba 278 Japan
Table 1. Occupational status for Danish father-son pairs; from Goodman (1981b).
(The first, second, third, fourth parenthesized values are the MLEs of expected
frequencies under the QU model, PQU models I and II, and model (7.1), respectively).

<table>
<thead>
<tr>
<th>Father's status</th>
<th>Son's status</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>(1)</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>(15.5)</td>
<td>(15.7)</td>
</tr>
<tr>
<td></td>
<td>(17.5)</td>
<td>(18.2)</td>
</tr>
<tr>
<td></td>
<td>(17.0)</td>
<td>(14.8)</td>
</tr>
<tr>
<td></td>
<td>(19.8)</td>
<td>(13.6)</td>
</tr>
<tr>
<td>(2)</td>
<td>24</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>(27.8)</td>
<td>(105)</td>
</tr>
<tr>
<td></td>
<td>(26.5)</td>
<td>(105)</td>
</tr>
<tr>
<td></td>
<td>(25.2)</td>
<td>(105)</td>
</tr>
<tr>
<td></td>
<td>(23.3)</td>
<td>(105)</td>
</tr>
<tr>
<td>(3)</td>
<td>23</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>(23.6)</td>
<td>(84.2)</td>
</tr>
<tr>
<td></td>
<td>(23.9)</td>
<td>(80.5)</td>
</tr>
<tr>
<td></td>
<td>(20.2)</td>
<td>(94.1)</td>
</tr>
<tr>
<td></td>
<td>(22.4)</td>
<td>(84.5)</td>
</tr>
<tr>
<td>(4)</td>
<td>8</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>(8.2)</td>
<td>(45.6)</td>
</tr>
<tr>
<td></td>
<td>(9.2)</td>
<td>(47.9)</td>
</tr>
<tr>
<td></td>
<td>(6.8)</td>
<td>(49.7)</td>
</tr>
<tr>
<td></td>
<td>(10.2)</td>
<td>(50.1)</td>
</tr>
<tr>
<td>(5)</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>(1.4)</td>
<td>(12.6)</td>
</tr>
<tr>
<td></td>
<td>(1.5)</td>
<td>(12.1)</td>
</tr>
<tr>
<td></td>
<td>(1.2)</td>
<td>(14.1)</td>
</tr>
<tr>
<td></td>
<td>(2.2)</td>
<td>(14.0)</td>
</tr>
<tr>
<td>Total</td>
<td>79</td>
<td>263</td>
</tr>
</tbody>
</table>

2. The quasi-uniform association model

For the \( R \times R \) square contingency table with ordered categories, let \( f_{ij} \) denote the observed frequency in the \( i \)th row and \( j \)th column of the table \((i=1, 2, \ldots, R; j=1, 2, \ldots, R)\), and let \( F_{ij} \) denote the corresponding expected frequency under some model. As is usual in the statistical analysis of such tables, we assume here that a multinomial distribution applies to the \( R \times R \) table.

The U model for the expected frequencies \( F_{ij} \) is defined as

\[
F_{ij} = \alpha_i \beta_j \theta^{ij} \quad \text{for} \quad i=1, 2, \ldots, R; j=1, 2, \ldots, R,
\]

where \( \alpha_i, \beta_j, \) and \( \theta \) are parameters in the model (see Goodman 1979). Note that the U model was originally defined for the general \( R \times C \) contingency tables (see Goodman 1979). A special case of the U model obtained by setting \( \theta=1 \) is the usual independence model (called the null association model). For the \( 2 \times 2 \) subtables formed from adjacent rows (i.e., rows \( i \) and \( i+1 \)) and adjacent columns (i.e., columns \( j \) and \( j+1 \)) in the \( R \times R \) contingency table, let \( \Theta_{ij} \) denote the corresponding local odds-ratio (for \( i=1, 2, \ldots, R-1; j=1, 2, \ldots, R-1 \)) based on the expected frequencies. Thus,
\[ \Theta_{ij} = \frac{(F_{ij}F_{i+1,j+1})}{(F_{i,j+1}F_{i+1,j})} \].

This is the ratio of the odds \( F_{ij}/F_{i+1,j} \) that an observation will fall in row \( i \) rather than in row \( i+1 \), given that it is in column \( j \), and the odds \( F_{i,j+1}/F_{i+1,j+1} \) that an observation will fall in row \( i \) rather than in row \( i+1 \), given that it is in column \( j+1 \). The U model can then be expressed as

\[ \Theta_{ij} = \theta \quad \text{for} \quad i=1, 2, \ldots, R-1; \quad j=1, 2, \ldots, R-1. \]

Therefore, the U model indicates that all local odds-ratios are equal, that is, there is a uniform local association for all cells in the contingency table. Also, for the \( 2 \times 2 \) sub-tables formed from rows \( i \) and \( j (>i) \) and columns \( s \) and \( t (>s) \) in the \( R \times R \) table, let \( \Theta_{(i<s);(j<t)} \) denote the corresponding odds-ratio based on the expected frequencies. Thus,

\[ \Theta_{(i<s);(j<t)} = \frac{(F_{is}F_{jt})}{(F_{js}F_{it})}. \]

The U model can then also be expressed as

\[ \log \Theta_{(i<s);(j<t)} = (j-i)(t-s)\phi \quad \text{for} \quad 1 \leq i < j \leq R; \quad 1 \leq s < t \leq R. \]

Notice that \( \theta = \exp \{\phi\} \). For details of the U model, see Goodman (1979, 1981a, 1981b, 1984, 1985).

Next, the QU model for the expected frequencies \( F_{ij} \) is defined as

\[ F_{ij} = a_i b_j \Theta^{ij} \psi_{ij} \quad \text{for} \quad i=1, 2, \ldots, R; \quad j=1, 2, \ldots, R, \]

where \( \psi_{ij} = 1 \) for \( i \neq j \) (see Goodman 1979 and Agresti 1984, p. 204). A special case of the QU model obtained by setting \( \theta=1 \) is the QO model. Note that for the square contingency table, the U model implies the QU model. The QU model can be also expressed as

\[ (2.1) \quad \log \Theta_{(i<s);(j<t)} = (j-i)(t-s)\phi \quad \text{for} \quad 1 \leq i < j \leq R; \quad 1 \leq s < t \leq R; \quad i \neq s, i \neq t, j \neq s, j \neq t. \]

In other words, the log odds-ratio formed from a rectangular pattern of cells off the main diagonal of the square table is directly proportional to the product of the distance between the rows and the distance between the columns. Also the QU model implies that

\[ (2.2) \quad \Theta_{ij} = \theta \quad \text{for} \quad i=1, 2, \ldots, R-1; \quad j=1, 2, \ldots, R-1; \quad i \neq j-1, i \neq j, i \neq j+1. \]

In other words, the QU model implies that the local odds-ratios not containing cells on the main diagonal of the square contingency table are equal. However, condition (2.2) does not imply the QU model. Note that formula (2.1) implies (2.2). In passing, we note that \( a_i = 1 \), for example, can be inserted in the U and QU models without loss of generality.

3. Parsimonious quasi-uniform association models

For the \( R \times R \) table, we now introduce the first parsimonious QU model (denoted by PQU model I) for the expected frequencies \( F_{ij} \) defined as follows:

\[ F_{ij} = a_i b_j \Theta^{ij} \psi_{ij} \quad \text{for} \quad i=1, 2, \ldots, R; \quad j=1, 2, \ldots, R, \]

where \( \psi_{ij} = 1 \) for \( i \neq j \), and
\( \alpha_i = \mu^{i-1} \prod_{k=1}^{i-2} \gamma^k = \mu^{i-1} \gamma^{(i-2)(i-1)/2} \quad \text{for} \quad i = 1, 2, \ldots, R. \)

Note that \( \alpha_1 = 1 \) and \( \alpha_2 = \mu \), and also that the odds of row effect parameters pertaining to adjacent row categories (say, \( i \) and \( i+1 \)), i.e., \( \alpha_{i+1}/\alpha_i \), are equal to \( \mu \gamma^{i-1} \); that is to say, \( \{\alpha_{i+1}/\alpha_i\} \) is an exponential function for \( i = 1, 2, \ldots, R-1 \). Obviously, PQU model I implies the QU model. PQU model I can also be expressed in terms of both the pattern of odds-ratios and the pattern of odds based on the expected frequencies. In order to demonstrate this, consider the quantity \( \Omega_{ji}^{h,w} \) defined as follows:

\( \Omega_{ji}^{h,w} = F_{ij}/F_{i+1,j} \).

\( \Omega_{ji}^{h,w} \) expresses the odds that an observation will fall in row \( i \) rather than in row \( i+1 \), given that it was in column \( j \). Notice that the local odds-ratio \( \Theta_{ij} \) can now be expressed as follows:

\( \Theta_{ij} = \Omega_{ji}^{h,w} \Omega_{i+1,j}^{h,w} \).

PQU model I introduced here is equivalent to the condition that formula (2.1) holds, and in addition

\[
(3.1) \quad 1/\Omega_{ji}^{h,w} = \mu \gamma^{i-1} \theta^j \quad \text{for} \quad j \neq i, j \neq i+1;
\]

it is also equivalent to the condition that formula (2.1) holds, and in addition

\[
(3.2) \quad \Omega_{ji}^{h,w} \Omega_{j+1,i}^{h,w} = \gamma \quad \text{for} \quad j \neq i, j \neq i+1, j \neq i+2.
\]

Therefore, PQU model I states that in addition to the interpretation based on the pattern of (log) odds-ratios for the QU model described in Section 2, this model has an interpretation based on the pattern of odds described as follows. From (3.1), the odds \( \Omega_{ji}^{h,w} \) increase (or decrease) exponentially as row \( i \) (for \( i \neq j-1, i \neq j \)) increases, and this rate of increase (or decrease) is unaffected by the value of \( j \), but the intercept of the corresponding exponential curve may be affected by the value of \( j \); or from (3.2), the odds \( \Omega_{ji}^{h,w} \) that an observation will fall in row \( i \) rather than in row \( i+1 \), given that it was in column \( j \) (for \( j \neq i, j \neq i+1, j \neq i+2 \)), are always \( \gamma \) times higher than the odds \( \Omega_{j+1,i}^{h,w} \), that the observation will fall in row \( i+1 \) rather than in row \( i+2 \), given that it was in the same column \( j \).

Next, we introduce PQU model II for expected frequencies \( F_{ij} \) defined as follows:

\( F_{ij} = \alpha_i \beta_j \theta^{i,j} \psi_{ij} \quad \text{for} \quad i = 1, 2, \ldots, R; j = 1, 2, \ldots, R, \)

where \( \psi_{ij} = 1 \) for \( i \neq j \), and

\( \beta_j = \mu^{j-1} \prod_{k=1}^{j-1} \delta^k = \mu^{j-1} \delta^{(j+1)(j-1)/2} \quad \text{for} \quad j = 1, 2, \ldots, R. \)

Note that restrictions imposed instead on the row effect parameters \( \{\alpha_i\} \) in PQU model I are imposed on the column effect parameters \( \{\beta_j\} \) in PQU model II. Also note that PQU model II implies the QU model. In order to describe the pattern of odds based on the expected frequencies for PQU model II, consider the quantity \( \Omega_{ij}^{h,w} \) defined as follows:

\( \Omega_{ij}^{h,w} = F_{ij}/F_{i+1,j+1}. \)
Compare $\Omega_{ij}^{\theta}$ with $\Omega_{ji}^{\theta}$. Now notice that the local odds-ratio $\Theta_{ij}$ can be expressed as follows:

$$\Theta_{ij} = \Omega_{ij}^{\theta} / \Omega_{j,i+1}^{\theta}.$$  

PQU model II is equivalent to the condition that formula (2.1) holds, and in addition

$$(3.3) \quad 1/\Omega_{ij}^{\theta} = v^{j-i} \theta^i \quad \text{for} \quad i \neq j, i \neq j+1;$$

it is also equivalent to the condition that formula (2.1) holds, and in addition

$$(3.4) \quad \Omega_{ij}^{\theta} / \Omega_{j,i+1}^{\theta} = \delta \quad \text{for} \quad i \neq j, i \neq j+1, i \neq j+2.$$ 

Thus, we can give an interpretation of PQU model II similar to that of PQU model I described above, and we shall omit it here.

Consider now PQU model III for the expected frequencies $F_{ij}$ defined as follows:

$$F_{ij} = \alpha_i \beta_j \chi_{ij} \psi_{ij} \quad \text{for} \quad i = 1, 2, \ldots, R; j = 1, 2, \ldots, R,$$

where $\psi_{ij} = 1$ for $i \neq j$, and

$$\alpha_i = \mu^{i-1} \prod_{k=1}^{i-1} \gamma^k = \mu^{i-1} \gamma^{(i-2)(i-1)/2},$$

and

$$\beta_j = \nu^{j-1} \prod_{k=1}^{j-1} \delta^k = \nu^{j-1} \delta^{(j-2)(j-1)/2}.$$ 

Obviously, PQU model III implies both PQU models I and II.

Since all the various association models considered here have multiplicative forms, the maximum likelihood estimates (MLEs) of the expected frequencies under each model can be easily obtained using an iterative procedure, for example, the general iterative procedure for log-linear models of Darroch and Ratcliff (1972). We shall not go into these details here. Also, each model can be tested for goodness of fit by the likelihood ratio chi-squared statistic $G^2$ with the corresponding degrees of freedom (df). The numbers of df for the U and QU models are $R(R-2)$ and $R(R-3)$, respectively, and those for PQU models I and III are $(R+1)(R-3)$ and $(R+2)(R-3)$, respectively. Also, the number of df for PQU model II is equal to that for PQU model I.

4. Analysis of Danish occupational mobility data

We shall apply the various QU models considered in this paper to the data in Table 1. The QU model fits the data in Table 1 well, since the value of $G^2$ is 15.40 ($P > 0.1$) with 10 df. PQU models I and II also fit the data in Table 1 well, since the values of $G^2$ are 18.59 ($P \approx 0.1$) and 19.00 ($P > 0.05$), respectively, both with 12 df. However, PQU model III, unfortunately, does not fit the data in Table 1 so well, since the value of $G^2$ is 26.17 ($P \approx 0.025$) with 14 df. For testing the hypothesis that PQU model I holds under the assumption that the QU model holds true, the difference between the $G^2$ values for PQU model I and QU model is 3.19 ($P > 0.2$) with $12 - 10 = 2$ df. Therefore, this hypothesis is accepted. Hence we find that the odds of row effect parameters pertaining to adjacent categories, i.e., $\{\alpha_{i+1}/\alpha_i\}$, is an exponential function for $i = 1, 2, 3, 4$; thus, we prefer PQU
model I to the QU model for the data in Table 1. Comparing PQU model II to the QU model in the same way, we find that the odds of column effect parameters pertaining to adjacent categories, \( \{ \beta_{j+1}/\beta_j \} \), is an exponential function for \( i=1, 2, 3, 4 \); we thus prefer PQU model II to the QU model for the data in Table 1.

When PQU model I is applied to the data in Table 1, we obtain the following results. The MLE of \( \theta \) is 1.551. (Note that the MLE of \( \theta \) under the QU model is 1.565.) Therefore under the PQU model I applied to these data, for the local 2\times2 tables that do not contain cells on the main diagonal, the odds that a son’s status is \( j+1 \) instead of \( j \) is estimated to be 1.551 times higher when the father’s status is \( i+1 \) than when it is \( i \). In addition, for \( i<j \) and \( s<t \) with \( i\neq s \), \( i\neq t \), \( j\neq s \), and \( j\neq t \), the odds that a son’s status is \( t \) instead of \( s \) is estimated to be \((1.551)^{(j-i)(t-s)} \) times higher when the father’s status is \( j \) than when it is \( i \). Moreover, under PQU model I applied to the data in Table 1, the MLE of \( \gamma \) is \( \hat{\gamma} = 0.426 \), i.e., \( 1/\hat{\gamma} = 2.349 \). Therefore, under PQU model I, for \( j \neq i \), \( j \neq i+1 \), and \( j \neq i+2 \), when a son’s status is \( j \), the odds that the father’s status is \( i+1 \) instead of \( i+2 \) is estimated to be 2.349 times higher than the odds that the father’s status is \( i \) instead of \( i+1 \) (see formula (3.2)). In other words, under PQU model I, for \( j \neq i \) and \( j \neq i+1 \), when a son’s status is \( j \), the odds that the father’s status is \( i \) instead of \( i+1 \) increase exponentially as the value of \( i \) increases (i.e., as the father’s status category becomes higher), and the rate of increase is unaffected by son’s status \( j \) (see formula (3.1)). In addition, under PQU model I applied to the data in Table 1, the MLE of \( \Omega_{ij}^{ab} \) (\( =F_{ij}/F_{i+1,j} \)) is greater than 1 for \( i>j \) and less than 1 for \( i+1<j \) (see Table 1). Therefore, under PQU model I: (i) when a son’s status is \( j \) and the father has the higher status category, the father’s status tends to be \( i \) rather than \( i+1 \); and (ii) when a son’s status is \( j \) and the father has the lower status category, the father’s status tends to be \( i+1 \) rather than \( i \).

Next, we consider the results when PQU model II is applied to the data in Table 1. The MLE of \( \theta \) is then 1.579. This is very close to the MLE of \( \theta \) under each of PQU model I and the QU model applied to these data. Therefore, PQU model II applied to these data provides an interpretation based on the odds-ratios similar to that under PQU model I as described above, and we shall omit it. The MLE of \( \delta \) is \( \hat{\delta} = 0.467 \), i.e., \( 1/\hat{\delta} = 2.143 \). Therefore, for \( i\neq j \), \( i\neq j+1 \), and \( i\neq j+2 \), when a father’s status is \( i \), the odds that the son’s status is \( j+1 \) instead of \( j+2 \) is estimated to be 2.143 times higher than the odds that the son’s status is \( j \) instead of \( j+1 \) (see formula (3.4)). In other words, under PQU model II, for \( i\neq j \) and \( i\neq j+1 \), when a father’s status is \( i \), the odds that the son’s status is \( j \) instead of \( j+1 \) increase exponentially as the value of \( j \) increases (i.e., as the son’s status category becomes higher), and the rate of increase is unaffected by the father’s status \( i \) (see formula (3.3)). In addition, the MLE of \( \Omega_{ij}^{ab} \) (\( =F_{ij}/F_{i+1,j+1} \)) is greater than 1 for \( i<j \) and less than 1 for \( i+1<j+1 \) (see Table 1). Therefore, under PQU model II: (i) when a father’s status is \( i \) and the son has the higher status category, the son’s status tends to be \( j \) rather than \( j+1 \); and (ii) when a father’s status is \( i \) and the son has the lower status category, the son’s status tends to be \( j+1 \) rather than \( j \).

5. Note
Table 2, taken directly from Agresti (1984, p. 206), gives the fathers’ and sons’ occu-
Table 2. Occupational status for British father-son pairs; from Agresti (1984, p. 206).

<table>
<thead>
<tr>
<th>Father's status</th>
<th>Son's status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>(1)</td>
<td>50</td>
</tr>
<tr>
<td>(2)</td>
<td>28</td>
</tr>
<tr>
<td>(3)</td>
<td>11</td>
</tr>
<tr>
<td>(4)</td>
<td>14</td>
</tr>
<tr>
<td>(5)</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
</tr>
</tbody>
</table>

Table 2. Occupational status for British father-son pairs; from Agresti (1984, p. 206).

The data have been analyzed by several statisticians, including Mosteller (1968), Kullback (1971), Bishop et al. (1975, p. 100), Goodman (1981b, 1984), and Agresti (1984, pp. 205–207). For these data, the QU model fits well, since the value of $G^2$ is 14.08 ($P>0.1$) with 10 df. (Note that for these data, the value of $G^2$ for the QU model obtained by Goodman (1981b) is 14.08; however, that obtained by Agresti (1984, p. 207) is 13.9.) Also, for these data, PQU models I and II do not fit well, since the values of $G^2$ are 172.59 and 177.22, respectively, both with 12 df. Of course, PQU model III does not fit these data well, since the value of $G^2$ is 246.19 with 14 df. Therefore, for the British data, the QU model is preferable to all of the PQU models introduced in this paper.

6. Remark

In Section 4, for the data in Table 1 we saw the following at the 0.05 significance level:

1. The QU model was accepted.
2. PQU model I under the QU model was accepted.
3. PQU model II under the QU model was accepted.
4. PQU model III under the QU model was rejected.

Theoretically PQU model III is equivalent to the condition that both PQU models I and II hold. Unfortunately, however, the situation indicated by (2), (3), and (4) arises here. The referees point out that this situation may be caused by the following:

(i) For hypothesis (4), the type I error of the test may have arisen.
(ii) For hypotheses (2) and (3), PQU models I and/or II may not have held true.
(iii) The true model may differ from the QU model.

A similar situation often arises for the statistical test. For example, Table 7A in Goodman’s (1979) paper provides the following: for the data in (modified) Table 3 in Goodman (1979), the row-effect association (R) model and the column-effect association (C) model are accepted, however, the uniform association (U) model is rejected (at the 0.05 level), though the U model is theoretically equivalent to the condition that both the R and C models hold.

For the analysis of the Danish data, although various models would have to be applied as Goodman (1972) and Bishop et al. (1975, p. 323), the purpose of this paper is to show that the parsimonious QU models introduced here are preferable to the original
QU model, using the statistical test.

7. Other related models

We have seen that PQU model III is equivalent to the condition that formula (2.1) (i.e., the QU model) holds, and in addition formulae (3.2) and (3.4) hold, and that it did not fit the data in Table 1 very well. We are now interested in a model defined only by formulae (3.2) and (3.4). Although the proof is omitted, this model can be expressed in a multiplicative form as follows:

\[
F_{ij} = \begin{cases} 
\xi_{ij} \alpha_{ij} \beta_{ij} & \text{for } i < j, \\
\xi_{ij} \alpha_{ij} \beta_{ij} & \text{for } i > j, \\
\psi_{ii} & \text{for } i = j,
\end{cases}
\]

(7.1)

where

\[
\alpha_{ij} = \mu_{ij}^{-1} \prod_{k=1}^{i-2} \gamma_{ik} = \mu_{ij}^{-1} \gamma_{i-1}(i-1)/2,
\]

and

\[
\beta_{ij} = v_{ij}^{-1} \prod_{k=1}^{j-2} \gamma_{jk} = v_{ij}^{-1} \gamma_{j-1}(j-1)/2.
\]

A special case of model (7.1) obtained by setting \( \xi_i = \xi_s, \mu_i = \mu_s, \nu_i = \nu_s, \) and \( \theta_i = \theta_s \) is PQU model III. Note that model (7.1) is not a parsimonious QU model.

In Section 4, we also saw that the MLE of \( \gamma (=0.426) \) under PQU model I applied to the data in Table 1 was close to the MLE of \( \delta (=0.467) \) under PQU model II applied to the same data. Thus, we are also interested in a model defined as follows:

\[
F_{ij} = \begin{cases} 
\xi_{ij} \alpha_{ij} \beta_{ij} & \text{for } i < j, \\
\xi_{ij} \alpha_{ij} \beta_{ij} & \text{for } i > j, \\
\psi_{ii} & \text{for } i = j,
\end{cases}
\]

(7.2)

where

\[
\alpha_{ij} = \mu_{ij}^{-1} \prod_{k=1}^{i-2} \gamma_{ik} = \mu_{ij}^{-1} \gamma_{i-1}(i-1)/2,
\]

and

\[
\beta_{ij} = v_{ij}^{-1} \prod_{k=1}^{j-2} \gamma_{jk} = v_{ij}^{-1} \gamma_{j-1}(j-1)/2.
\]

Obviously model (7.2) is a special case of model (7.1) obtained by setting \( \gamma = \delta \). Model (7.2) is equivalent to the condition that formulae (3.2) and (3.4) with \( \gamma = \delta \) hold, though the proof is omitted.

When model (7.1) is applied to the data in Table 1, this model fits the data very well, yielding \( G^2 = 11.28 \) \( (P > 0.3) \) with \( R^2 - R - 10 = 10 \) df. In addition, model (7.2) fits the data well, yielding \( G^2 = 17.77 \) \( (P > 0.05) \) with \( R^2 - R - 9 = 11 \) df. For testing the hypothesis that model (7.2) holds under the assumption that model (7.1) holds true (namely the hypothesis that \( \gamma = \delta \) under the assumption), the difference between the \( G^2 \) values for models (7.2) and (7.1) is 6.49 \( (P = 0.01) \) with \( 11 - 10 = 1 \) df. Therefore, this hypothesis is rejected at the 0.05 significance level. Hence, for these data, there is strong evidence of
a difference between \( \gamma \) in formula (3.2) and \( \delta \) in formula (3.4), and model (7.1) is thus preferable to model (7.2). When model (7.1) is applied to the data in Table 1, the MLEs of \( \gamma \) and \( \delta \) are \( \hat{\gamma} = 0.473 \) and \( \hat{\delta} = 0.707 \) (i.e., \( 1/\hat{\gamma} = 2.114 \) and \( 1/\hat{\delta} = 1.414 \)). Therefore, model (7.1), when applied to these data, states that: (i) for \( j \neq i, j \neq i+1, \) and \( j \neq i+2 \), when a son's status is \( j \), the odds that the father's status is \( i+1 \) instead of \( i+2 \) is estimated to be 2.114 times higher than the odds that the father's status is \( i \) instead of \( i+1 \); and (ii) for \( i \neq j, i \neq j+1, \) and \( i \neq j+2 \), when a father's status is \( i \), the odds that the son's status is \( j+1 \) instead of \( j+2 \) is estimated to be 1.414 times higher than the odds that the son's status is \( j \) instead of \( j+1 \).

8. Model selection by Akaike's information criterion

As seen in Section 4, PQU models I and II were preferable to the original QU model for the data in Table 1. In addition, model (7.1) fitted the data very well (see Section 7). Therefore, we now wish to compare PQU models I and II with model (7.1). However, these models are not nested in each other, and hence a comparison based on the \( G^2 \) values cannot be done simply.

A quick method for choosing the best-fitting model among different models which include nonnested models is to use Akaike's (1974) information criterion (AIC), which is defined as

\[
AIC = -2 \text{(maximum log-likelihood)} + 2 \text{(number of free parameters)}
\]

for each model. For details see Sakamoto et al. (1986). This criterion gives the best-fitting model as the one with minimum AIC. Since we need only the difference between AIC's when two models are compared, we ignore a common constant of AIC and use a modified AIC defined as

\[
AIC^+ = G^2 - 2 \text{(number of degrees of freedom)}.
\]

Thus, for a certain body of data, the model with the minimum AIC+ (i.e., the minimum AIC) is the best-fitting model.

For the data in Table 1, the AIC+ values for the QU model, PQU models I and II, and model (7.1) are -4.60, -5.41, -5.00, and -8.72, respectively. Therefore, for these data, model (7.1) is preferable to PQU models I and II (and, of course, to the QU model) by the AIC criterion.

Acknowledgements

The author would like to express his sincere thanks to Professors K. Kunisawa, T. Makino and Dr. K. Shimizu, Science University of Tokyo, for their encouragement. He is also grateful to the referees for their valuable suggestions on improving this paper.

References

Contr. AC-19 716–723.


