Bayesian Score Merging for the Order Restricted RC Association Model

George Iliopoulos, Maria Kateri & Ioannis Ntzoufras
Department of Statistics and Insurance Science
University of Piraeus
Piraeus, Greece;
e-mails: {geh; mkateri}@unipi.gr
Department of Statistics
Athens University of Economics and Business
Athens, Greece;
e-mail: ntzoufras@aueb.gr

2007, 26th June, Lancaster University

Synopsis
1. Introduction.
2. Modeling Details.
3. RJMCMC Algorithm.
4. Illustrative example and results.
5. Discussion and further work.

1 Introduction

• Let \( y = (y_{ij}) \) be the frequencies and
• \( \Pi = (\pi_{ij}) \) be the probabilities
of an \( I \times J \) contingency table of two ordinal variables \( X \) and \( Y \) with \( I \) and \( J \) levels respectively.

Saturated log-linear model:

\[
\log \pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{XY} \quad i = 1, \ldots, I, \ j = 1, \ldots, J.
\]

\[
\log \pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \phi \mu_i \nu_j \quad \text{(Goodman, 1979, 1985)}
\]

where \( \mu = (\mu_1, \mu_2, \ldots, \mu_I) \) and \( \nu = (\nu_1, \nu_2, \ldots, \nu_J) \) be the scores assigned to the levels of \( X \) (rows) and \( Y \) (columns) respectively.

Interpretation of \( \phi \):

• \( \phi \) is an intrinsic association parameter.
• The above formulation reveals the analogies to the classical correspondence analysis (CA) or canonical correlation model.

Interpretation of \( \phi \): Odds ratio of successive categories if the score distances are equal to one since

\[
\log \left( \frac{\pi_{ij+1}}{\pi_{ij+1} \pi_{i+1,j}} \right) = \phi(\mu_{i+1} - \mu_i)(\nu_{j+1} - \nu_j).
\]
USUAL CONSTRAINTS

- Sum-to-zero constraints on row and column main effects ($\lambda_X^Y$ and $\lambda_Y^X$).
- Sum-to-zero constraints on row and column scores ($\mu_i$ and $\nu_j$).
- Two additional constraints on the row and column scores are needed in order to achieve the identifiability of the model (this due to the fact that (1) is multiplicative and not linear to its parameters).

\[ \sum_{i=1}^{I} \mu_i = \sum_{j=1}^{J} \nu_j = 0 \quad \text{and} \quad \sum_{i=1}^{I} \mu_i^2 = \sum_{j=1}^{J} \nu_j^2 = 1. \] (2)

Aim of this work

- Work with the order restricted RC model.
- Use the Bayesian approach to identify which scores $\mu_i, \mu_{i+1}$ and $\nu_j, \nu_{j+1}$ can be merged.
- Use Reversible jump MCMC to estimate posterior model probabilities (and odds) of each model
- Implement Bayesian model averaging

Why Use the Bayesian Approach in this Problem?

- They are not approximate and can be implemented even for samples with small size or with sparse contingency tables.
- Score merging in classical methods can be done using stepwise like methods and sequential implementation of significance tests (significance level is higher than the specified one, different model may selected if different starting points are selected).
- Using RJMCMC (or other varying dimension MCMC method) we automatically search the model space and estimate posterior model probabilities.
- Bayesian model averaging can be used in straightforward manner.

2 Modeling Details

- We focus on the order restricted version of the RC association model.
- $X$ and $Y$ ordinal $\Rightarrow$ natural to assume that the ordinal structure for scores

\[ \mu_1 \leq \mu_2 \leq \cdots \leq \mu_I \quad \text{and} \quad \nu_1 \leq \nu_2 \leq \cdots \leq \nu_J \]

- Which successive scores ($\mu_i, \mu_{i+1}$) and ($\nu_j, \nu_{j+1}$) are equal?
- In all models we assume that at least two row and two column scores are different.
We propose to use an alternative set of constraints:

\[ \mu_1 = \mu_{\min} < \mu_I = \mu_{\max} \text{ and } \nu_1 = \nu_{\min} < \nu_J = \nu_{\max} \]

- Row and column scores take values in the intervals \([\mu_{\min}, \mu_{\max}]\) and \([\nu_{\min}, \nu_{\max}]\) respectively.

- Sensible choices:
  - \(\mu_{\min} = \nu_{\min} = -1\) and \(\mu_{\max} = \nu_{\max} = 1\) [range similar to the parameters under constraints (2)]
  - We use: \(\mu_{\min} = \nu_{\min} = 0\) and \(\mu_{\max} = \nu_{\max} = 1\)

  * simplifies computations
  * \(\phi = \log \left( \frac{\mu_I}{\mu_{I+J}} \right)\)

- Posterior distributions of scores under (2) can be obtained by transforming MCMC output of the proposed parametrization.

Moreover the actual distinct unequal row and column scores will be denoted by the vectors \(\mu_\gamma\) and \(\nu_\delta\) of dimension \(I_I\) and \(\Delta_J\), respectively given by

\[
\mu_\gamma = \left( \{ \mu_i : \gamma_i = 1; i = 1, 2, \ldots, I \} \right) = \left( \mu_{\gamma(1)}, \mu_{\gamma(2)}, \ldots, \mu_{\gamma(I_1)} \right)^T
\]

\[
\nu_\delta = \left( \{ \nu_j : \delta_j = 1; j = 1, 2, \ldots, J \} \right) = \left( \nu_{\delta(1)}, \nu_{\delta(2)}, \ldots, \nu_{\delta(\Delta_J)} \right)^T.
\]

Then the original scores are given by

\[ \mu_i = \mu_\gamma(I_i) \text{ and } \nu_j = \nu_\delta(\Delta_j) \]

Let

\[ I_I = \sum_{k=1}^{i} \gamma_k \text{ and } \Delta_J = \sum_{k=1}^{j} \delta_k \]

be the distinct distinct scores under estimation until row \(i\) or column \(j\) respectively.

Moreover the actual distinct unequal row and column scores will be denoted by

\[
\begin{array}{cccc}
  i & 1, 2, & 3, 4, & 5 \\
  \mu_i & \mu_1 = \mu_2 = 0 & \mu_3 = \mu_4 = 0.6 & \mu_5 = 1 \\
  \gamma_i & 1, 0, & 1, 0, & 1 \\
  I_i & 1, 1, & 2, 2, & 3 \\
  \mu_\gamma(\ell) & 0 & 0.6 & 1 \\
  \nu_\delta(\ell) & \mu_\gamma(I_1) = \mu_\gamma(1) = 0, & \mu_\gamma(I_3) = \mu_\gamma(2) = 0.6, & \mu_\gamma(I_5) = \mu_\gamma(3) = 1 \\
  \mu_\gamma(I_2) = \mu_\gamma(1) = 0, & \mu_\gamma(I_4) = \mu_\gamma(2) = 0.6, \\
  \end{array}
\]
Consider the row and column score differences
\[ D_{\mu_i} = \mu_i - \mu_{i-1} \quad \text{and} \quad D_{\nu_j} = \nu_j - \nu_{j-1} \]
instead of the original parameters. Then
\[ \mu_i = \sum_{k=1}^{i} \gamma_k \mu_k \quad \text{and} \quad \nu_j = \sum_{k=1}^{j} \delta_k \nu_k; \quad i = 1, \ldots, I, \quad j = 1, \ldots, J. \]
For scores of range one \((R_{\mu} = \mu_{\text{max}} - \mu_{\text{min}} = 1) \Rightarrow \sum_{i=1}^{I} \gamma_i D_{\mu_i} = 1 \Rightarrow \text{we may use} \]
\[ D_{\gamma} = \{D_{\mu_i} : \gamma_i = 1\} \sim D(1, \Gamma_{\gamma}) \]
(Dirichlet prior of dimension \(I\) with all parameters equal to one) as non-informative prior for row score differences. Similarly, for column scores \( D_{\delta} = \{D_{\nu_j} : \delta_j = 0\} \sim D(1, \Delta_{\delta}). \)

### RJMCMC algorithm

1. Update model structure: Sample \((\gamma, \delta)\) using successive RJMCMC moves:
   - For \(i = 2, \ldots, I\), propose \((\gamma', \delta')\):
     - Split: if \((\gamma_i = 0) \rightarrow (\gamma_i = 1)\) then propose \((\mu_{i-1} = \mu_i) \rightarrow (\mu_{i-1} < \mu_i')\).
       - Generate \(u\) from \(q(u|\mu, \gamma, \gamma')\).
       - Set \(\mu_{y'}, u = g(\mu, u)\).
     - Merge: if \((\gamma_i = 1) \rightarrow (\gamma_i' = 0)\) then propose \((\mu_{i-1} < \mu_i) \rightarrow (\mu_{i-1}' = \mu_i')\).
       - Set \(\mu_{y'}, u = g^{-1}(\mu, u)\).
       - Obtain \(\mu'\) from \(\mu_{y'}\) via \(\mu_i = \mu_i(\Gamma_i)\) & accept/reject the proposed move.
   - The updating scheme for the components of \(\delta\) is similar.

2. Generate model parameters \((\lambda^X, \lambda^Y, \phi, \mu, \nu)\), given the model structure \((\gamma, \delta)\):
   - Sample row and column effects.
   - Sample \(\phi\) using a simple random walk Metropolis.
   - Use random walk on logits of column and row scores' differences.

### Prior Distributions on Scores

Equivalently, the scores are a priori distributed as ordered iid uniform random variables
\[ f(\mu_i) = \frac{(I_i - 2)!}{(\mu_{\text{max}} - \mu_{\text{min}}) I_i - 2} I(\mu_{\text{min}} < \text{ordered different} \mu'\text{'s} < \mu_{\text{max}}) \]
Similarly, for the column scores
\[ f(\nu_j) = \frac{(J_j - 2)!}{(\nu_{\text{max}} - \nu_{\text{min}}) J_j - 2} I(\nu_{\text{min}} < \text{ordered different} \nu'\text{'s} < \nu_{\text{max}}) \]

Normal with large variances for the rest of the parameters.
Bernoulli for \(\gamma_i\) and \(\delta_j\) with prior probabilities equal to 1/2.
In Split Move:

\[
\gamma_i = 1 \rightarrow \gamma_i' = 0, \quad i : 2 < \Gamma_i = \ell < \Gamma_1
\]

\[
\begin{align*}
&\ldots \leq \mu_\gamma(\ell - 2) < \mu_\gamma(\ell - 1) < \mu_\gamma(\ell) < \mu_\gamma(\ell + 1) \leq \ldots \\
\downarrow & \\
&\ldots \leq \mu_\gamma'(\ell - 2) < \mu_\gamma'(\ell - 1) < \mu_\gamma'(\ell) \leq \ldots \\
\downarrow & \\
\end{align*}
\]

Usual transformation: \(\mu_\gamma'(\ell - 1) = \frac{\mu_\gamma(\ell - 1) + \mu_\gamma(\ell)}{2}\)

and leave the rest of the scores unchanged

\[
\mu_\gamma'(k) = \begin{cases} 
\mu_\gamma(k) & \text{for } k < \ell - 1 \\
\mu_\gamma(k + 1) & \text{for } k > \ell - 1
\end{cases}
\]

Hence in Merge Move

\[
\gamma_i = 0 \rightarrow \gamma_i' = 1, \quad i : 2 \leq \Gamma_i = \ell < \Gamma_1
\]

\[
\begin{align*}
&\ldots \leq \mu_\gamma(\ell - 1) < \mu_\gamma(\ell) < \mu_\gamma(\ell + 1) \leq \ldots \\
\downarrow & \\
&\ldots \leq \mu_\gamma'(\ell - 1) < \mu_\gamma'(\ell) < \mu_\gamma'(\ell + 1) < \mu_\gamma'(\ell + 2) \leq \ldots \\
\downarrow & \\
&\mu_\gamma(\ell - 1) < \mu_\gamma(\ell) < \mu_\gamma(\ell + 1)
\end{align*}
\]

\[
\begin{align*}
&\mu_{\min} = \mu_\gamma'(1) < \mu_\gamma'(2) < \mu_\gamma(3) < \ldots \\
\downarrow & \\
&\mu_{\min} = \mu_\gamma'(1) < \mu_\gamma'(2) < \ldots \\
\downarrow & \\
&\mu_{\min} = \mu_\gamma')(1) < \mu_\gamma'(2) < \mu_\gamma(3) < \ldots
\end{align*}
\]

\[\text{Usual Transformation } \frac{\mu_{\min} + \mu_\gamma(2)}{2} < \mu_\gamma(3) < \ldots \]

\[\neq \mu_{\min}\]

(VIOLATES THE CONSTRAINT \(\mu_\gamma'(1) = \mu_{\min}\))

**PROBLEM**

The above transformation cannot be applied for merging/splitting the lowest or the highest scores.

**Merge the Lowest Scores** \(\mu_\gamma(1)\) and \(\mu_\gamma(2)\)

\[(\gamma_i = 1 \rightarrow \gamma_i' = 0, \quad i : \Gamma_i = 2)\]

\[
\begin{align*}
&\mu_{\min} = \mu_\gamma(1) < \mu_\gamma(2) < \mu_\gamma(3) < \ldots \\
\downarrow & \\
&\mu_{\min} = \mu_\gamma'(1) < \mu_\gamma'(2) < \ldots \\
\downarrow & \\
&\mu_{\min} + \mu_\gamma(2) < \mu_\gamma(3) < \ldots \\
\end{align*}
\]

**Not Valid Since**

\[\mu_{\min} \neq \mu_{\min}\]
Using similar logic we apply the following transformations

\[
\begin{align*}
\mu_{\text{min}} &= \mu_\gamma(1) < \mu_\gamma(2) < \ldots < \mu_\gamma(\Gamma_1) = \mu_{\text{max}} \\
\mu_{\text{min}} + \frac{\mu_\gamma(2)}{2} &< \mu_\gamma(3) < \ldots < \mu_\gamma(\Gamma_1) = \mu_{\text{max}} \\
0 &< \mu_\gamma(3) - \frac{\mu_{\text{min}} + \mu_\gamma(2)}{2} < \ldots < \mu_{\text{max}} - \frac{\mu_{\text{min}} + \mu_\gamma(2)}{2} \\
0 &< \frac{\mu_\gamma(3) - \mu_{\text{min}}}{\mu_{\text{max}} - \mu_\gamma(2)} < \ldots < 1 \\
\mu_{\text{min}} &< \mu_{\text{min}} + \frac{2\mu_\gamma(3) - \mu_{\text{min}} - \mu_\gamma(2)}{2\mu_{\text{max}} - \mu_{\text{min}} - \mu_\gamma(2)} (\mu_{\text{max}} - \mu_{\text{min}}) < \ldots < \mu_{\text{max}} \\
\mu'_\gamma(1) &< \mu'_\gamma(2) < \ldots < \mu'_\gamma(\Gamma'_1)
\end{align*}
\]

**Final transformation**

\[
\mu'_\gamma(k) = \begin{cases} 
\mu_{\text{min}}, & k = 1, \\
\mu_{\text{min}} + (\mu_{\text{max}} - \mu_{\text{min}}) \frac{2\mu_\gamma(k + 1) - \mu_{\text{min}} - \mu_\gamma(2)}{2\mu_{\text{max}} - \mu_{\text{min}} - \mu_\gamma(2)}, & k > 1.
\end{cases}
\]

**Split the Lowest Score** \(\mu_\gamma(1)\) (reverse move)

\[
\begin{align*}
\left( \mu_{\text{min}} = \mu_\gamma(1) \quad < \quad \mu_\gamma(2) \quad < \ldots \right) \\
\downarrow \quad \downarrow \\
\left( \mu_{\text{min}} = \mu'_\gamma(1) \quad < \quad \mu'_\gamma(2) \quad < \quad \mu'_\gamma(3) \quad < \ldots \right)
\end{align*}
\]

- Set \(\mu'_\gamma(2) = u\).
- Generate \(u\) in the interval

\[
u \in \left( \mu_{\text{min}}, \mu_\gamma(2) + \frac{(\mu_\gamma(2) - \mu_{\text{min}})[\mu_{\text{max}} - \mu_\gamma(2)]}{\mu_\gamma(2) + \mu_{\text{max}} - 2\mu_{\text{min}}} \right).
\]
In Split Move \(\rightarrow |J| = \left(1 - \frac{1}{2} \frac{u - \mu_{\min}}{u - \mu_{\max}}\right)^{I'_I - 2}\)

In Merge Move \(\rightarrow u = \mu(y)\) and
\[|J| = \left[\left(1 - \frac{1}{2} \frac{u - \mu_{\min}}{u - \mu_{\max}}\right)^{I'_I - 2}\right]^{-1} = \left(1 - \frac{1}{2} \frac{\mu(y) - \mu_{\min}}{\mu(y) - \mu_{\max}}\right)^{3 - I_I} .
\]

Reminder:
- \(I_I\) is the number of scores of the current model (In split “smaller”, In merge: “larger” model)
- \(I'_I\) is the number of scores of the proposed model (In split “larger”, In merge: “smaller” model)

\[\mu_{\min} = \mu(y)_{\Gamma_I(1)} < \ldots < \mu(y)_{\Gamma_I(2)} < \mu(y)_{\Gamma_I(1)} < \mu(y)_{\Gamma_I} = \mu_{\max},\]
\[\Downarrow\]
\[\mu_{\min} = \mu(y)_{\Gamma_I(1)} < \ldots < \mu(y)_{\Gamma_I(2)} < \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(1)} - \mu_{\min}\right) / 2 \Downarrow\]
\[0 < \ldots < \mu(y)_{\Gamma_I(2)} - \mu_{\min} < \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(1)} - \mu_{\min}\right) / 2 \Downarrow\]
\[0 < \ldots < \left(\mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) \cdot \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) / 2 \Downarrow\]
\[\mu_{\min} < \ldots < \mu_{\min} + 2 \left(\mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) < \mu(y)_{\Gamma_I} .\]

Merge the Highest Scores \(\mu(y)(\Gamma_I - 1)\) and \(\mu(y)(\Gamma_I)\)

\[(\gamma_i = 1 \rightarrow \gamma'_i = 0, \ i : \Gamma_i = \Gamma_I)\]

\[\mu_{\min} = \mu(y)(1) < \ldots < \mu(y)(\Gamma_I - 2) < \mu(y)(\Gamma_I - 1) < \mu(y)_{\Gamma_I} = \mu_{\max}\]
\[\Downarrow\]
\[\mu_{\min} = \mu(y)(1) < \ldots < \mu(y)(\Gamma_I - 2) < \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(1)} - \mu_{\min}\right) / 2 \Downarrow\]
\[0 < \ldots < \mu(y)_{\Gamma_I(2)} - \mu_{\min} < \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(1)} - \mu_{\min}\right) / 2 \Downarrow\]
\[0 < \ldots < \left(\mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) \cdot \left(\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) / 2 \Downarrow\]
\[\mu_{\min} < \ldots < \mu_{\min} + 2 \left(\mu(y)_{\Gamma_I(2)} - \mu_{\min}\right) < \mu(y)_{\Gamma_I} .\]

Final transformation

\[\mu'(\gamma)(k) = \begin{cases} 
\mu_{\min} + 2 \left(\mu_{\max} - \mu_{\min}\right) \cdot \frac{\mu(y)(k) - \mu_{\min}}{\mu(y)_{\Gamma_I(1)} + \mu(y)_{\Gamma_I(1)} - 2 \mu_{\min}}, & k \leq \Gamma_I - 1 = \Gamma_I - 2 \\
\mu_{\max}, & k = \Gamma_I' = \Gamma_I - 1.
\end{cases}\]

\[k = \Gamma_I' = \Gamma_I - 1. \tag{5}\]
Split the Highest Score $\mu_\gamma(I_i)$ (reverse move)

$$\gamma_i = 0 \rightarrow \gamma'_i = 1, \ i : I_i = I_I$$

- $\mu_{\min} = \mu_\gamma(1) < \ldots < \mu_\gamma(I_I - 1) < \mu_\gamma(I_I) = \mu_{\max}$

- $\mu_{\min} = \mu'_{\gamma}(1) < \ldots < \mu'_{\gamma}(I_I - 1) < \mu'_{\gamma}(I_I) < \mu'_{\gamma}(I_I + 1) = \mu_{\max}$

- Generate $u$ in the interval

$$u \in \left(0, 2 \frac{(\mu_{\max} - \mu_{\min})(\mu_{\max} - \mu_\gamma(I_I - 1))}{\mu_{\max} - \mu_{\min}} + \frac{\mu_{\max} - \mu_\gamma(I_I - 1)}{\mu_{\max} - \mu_{\min}} \right)$$

- and set $\mu'_{\gamma}(I_I - 1) = \mu'_{\gamma}(I_I) = \mu_{\max} - u$.

In Split move

- Determinant of the Jacobian: $|J| = \left(1 - \frac{1}{2} \frac{\mu_{\max} - \mu_{\min}}{\mu_{\max} - \mu_{\min}} \right)^{I_I - 2}$

- $I_I$ is the number of scores in the smaller (current) model.

In the Merge move

- $u = \mu_{\max} - \mu_\gamma(I_I - 1)$ and

- Det. of Jacobian:

$$|J| = \left(1 - \frac{1}{2} \frac{\mu_{\max} - \mu_{\min}}{\mu_{\max} - \mu_{\min}} \right)^{2-I_I} = \left(1 - \frac{1}{2} \frac{\mu_{\max} - \mu_{\min}}{\mu_{\max} - \mu_{\min}} \right)^{I_I - 1}$$

- Here:

  * $I_I$ is the number of scores in the “bigger” (current) model.

  * $I'_I$ is the number of scores in the “smaller” (proposed) model.

Split the Highest Score $\mu_\gamma(I_I)$ (reverse move)

$$\gamma_i = 0 \rightarrow \gamma'_i = 1, \ i : I_i = I_I$$

Final transformation

$$\mu'_{\gamma}(k) = \begin{cases} 
\frac{\mu_\gamma(k) - u}{2} & \frac{\mu_{\gamma}(k) - \mu_{\min}}{\mu_{\max} - \mu_{\min}} = k - I_I - 2 = I_I - 1 \\
\mu_{\max} - u, & k = I_I' - 1 = I_I \\
\mu_{\max}, & k = I_I' = I_I + 1.
\end{cases} \quad (6)$$

Additional Details

- In practice we have used $\mu_{\min} = \nu_{\min} = 0$ and $\mu_{\max} = \nu_{\max} = 1$.

- When $I_I = 2$ then only two scores are different and set equal to $\mu_{\min}$ and $\mu_{\max}$. No further splitting is allowed. Similar is the case for column scores $\nu_j$.

- Rescaled Beta proposals can be used for proposing values for $u$.

- In practice we have used Uniform proposal which has been proved sufficient for datasets we have implemented the methodology.

- Further investigation is needed in order to construct proposals leading to more efficient RJMCMC schemes.
4 Illustrative Example.

Classical dataset of Maxwell (1961) concerning the severity of dreams’ disturbance of 223 boys aged from 5 to 15 years.

<table>
<thead>
<tr>
<th>Disturbance (from low to high)</th>
<th>Age Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>5–7</td>
<td></td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>8–9</td>
<td></td>
<td>10</td>
<td>15</td>
<td>11</td>
<td>13</td>
<td>49</td>
</tr>
<tr>
<td>10–11</td>
<td></td>
<td>23</td>
<td>9</td>
<td>11</td>
<td>7</td>
<td>50</td>
</tr>
<tr>
<td>12–13</td>
<td></td>
<td>28</td>
<td>9</td>
<td>12</td>
<td>10</td>
<td>59</td>
</tr>
<tr>
<td>14–15</td>
<td></td>
<td>32</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>44</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>100</td>
<td>42</td>
<td>41</td>
<td>40</td>
<td>223</td>
</tr>
</tbody>
</table>

Results: Most frequently visited models

<table>
<thead>
<tr>
<th>k</th>
<th>Model (scores)</th>
<th>Post. prob.</th>
<th>$P_{O_{k}}$</th>
<th>AIC</th>
<th>BIC</th>
<th>DIC</th>
<th>$p_{m}$</th>
<th>$d_{m}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.1620</td>
<td>1.00</td>
<td>1265.0</td>
<td>1295.7</td>
<td>1265.0</td>
<td>9.0</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.1540</td>
<td>1.05</td>
<td>1265.9</td>
<td>1300.0</td>
<td>1265.1</td>
<td>9.6</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0877</td>
<td>1.85</td>
<td>1267.6</td>
<td>1301.6</td>
<td>1266.3</td>
<td>9.4</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0725</td>
<td>2.23</td>
<td>1268.6</td>
<td>1306.1</td>
<td>1266.4</td>
<td>9.9</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0609</td>
<td>2.66</td>
<td>1269.0</td>
<td>1306.5</td>
<td>1266.4</td>
<td>9.7</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0579</td>
<td>2.80</td>
<td>1267.6</td>
<td>1301.7</td>
<td>1266.5</td>
<td>9.4</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0541</td>
<td>2.99</td>
<td>1269.0</td>
<td>1306.5</td>
<td>1266.7</td>
<td>9.9</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>$\mu_1 = \mu_2 &lt; \mu_3 = \mu_4 &lt; \mu_5$</td>
<td>0.0522</td>
<td>3.10</td>
<td>1268.3</td>
<td>1302.4</td>
<td>1266.8</td>
<td>9.2</td>
<td>10</td>
</tr>
</tbody>
</table>

Single RJMCMC (R RESULTS): 100,000 iterations + additional burn-in of 10,000 iterations.

Results: Marginal Probabilities $f(\gamma_i = 1|y)$ and $f(\delta_j = 1|y)$

<table>
<thead>
<tr>
<th>Row Scores Probability</th>
<th>Posterior</th>
<th>Column Scores Probability</th>
<th>Posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(\gamma_1 = 1</td>
<td>y)$ = 0.285</td>
<td>$f(\delta_2 = 1</td>
<td>y)$ = 0.996</td>
</tr>
<tr>
<td>$f(\gamma_4 = 1</td>
<td>y)$ = 0.391</td>
<td>$f(\delta_4 = 1</td>
<td>y)$ = 0.484</td>
</tr>
</tbody>
</table>

Single RJMCMC (R RESULTS): 100,000 iterations + additional burn-in of 10,000 iterations.

Posterior Distributions of $\phi$ over models with highest posterior probabilities.
Some Comments on the Results

- Negative association between age and severity of dreams’ disturbance ($\phi < 0$).
- Age:
  - Categories 2-3 (8–9, 10–11 years old) and 4-5 (12–13, 14–15 years old) ⇒ different in terms of the association (marginal post.prob. = 0.94 and 0.96 respectively).
  - Categories 1-2 (5–7, 8–9 years old) and 3-4 (10–11, 12–13 years old) ⇒ indistinguishable concerning the association (mild evidence with marginal post.prob. = 0.715 and 0.609 respectively).
- Severity of dreams’ disturbance: More uncertainty is involved:
  - Clear evidence that the first category differs than the rest $[f(\delta_2 = 1|y) = 0.996]$.
  - Model with the highest posterior probability ⇒ all the other three scores equal ($\nu_2 = \nu_3 = \nu_4$).
  - Model with the 2nd highest posterior probability ⇒ $\nu_2 < \nu_3 < \nu_4$.
- The algorithm was highly mobile visiting 69, 86 and all 105 models in 10, 100 iterations 400 thousand iterations respectively.

Comparison to Previous Results

- RJMCMC indicated a more parsimonious model (according to highest posterior probability) than the one (2nd in rank) indicated by our previous analysis (see Iliopoulos et al. 2007).
- Agresti et al. (1987) proposed an order restricted C model under which $\hat{\nu}_1 < \hat{\nu}_2 = \hat{\nu}_3 < \hat{\nu}_4$.
- Ritov and Gilula (1993) suggested an order restriction model with $\hat{\nu}_1 < \hat{\nu}_2 = \hat{\nu}_3 < \hat{\nu}_4$ and $\hat{\mu}_1 = \hat{\mu}_2 < \hat{\mu}_3 = \hat{\mu}_4$, which is the second highest probability according to our method.

5 Work in progress and future work

1. Comparison of the above models with the Uniform association, Independence and Saturated models [use different prior for $\phi$].
2. Incorporate selection between unrestricted RC, Row, Column association models (can we use similar parametrization?)
3. Use similar approach in unrestricted RC model for merging/grouping scores
4. Expand methodology to high dimensional tables
5. Use different priors for scores; for example power prior and imaginary data.
Publications by the same Group


Related Work


References


