Capture Recapture estimation using finite mixtures of arbitrary dimension

Richard Arnold
Yu Hayakawa
Paul Yip

International Biometrics Society
Taupo, 30 Nov-3 Dec 2009
Outline

- Capture-Recapture experiments and models
- Reversible Jump MCMC
- Application: Software reliability
Capture-Recapture

- $k$ repeated samples taken from a population
- Population size $N$ is unknown
- $D$ distinct individuals are seen
- Some seen on multiple occasions, some only once
- **Goal:** estimate population size $N$
Size of animal populations:
Samples are occasions on which animals are trapped, marked (for re-identification) and released.
$N$ is the number of animals in the population.
Capture-Recapture: Applications

- **Size of animal populations:**
  Samples are occasions on which animals are trapped, marked (for re-identification) and released. 
  $N$ is the number of animals in the population.

- **Software testing:**
  Samples are independent software testers detecting errors. 
  $N$ is the number of errors in the piece of software.
An unknown number \((N - D)\) of zero rows: to be estimated.
Capture-Recapture

Observations form the $D \times k$ capture matrix

$$X_{ij} = \begin{cases} 
1 & \text{if individual } i \text{ appears in sample } j \\
0 & \text{otherwise}
\end{cases}$$

Probability of capture

$$p_{ij} = \text{Probability individual } i \text{ is captured in sample } j$$

$$= p(X_{ij} = 1)$$

$$X_{ij} \sim \text{Bernoulli}(p_{ij})$$

$$p(X|P,N) = \frac{N!}{\prod x N_x!} \prod_{i=1}^{N} \prod_{j=1}^{k} p_{ij}^{x_{ij}} (1 - p_{ij})^{1 - x_{ij}}$$
1. All individuals are catchable

2. **Closed population**: no births/deaths or migration

3. No loss of marks (identifiable individuals)

4. No impact of sampling on capture probabilities (i.e. no behavioural responses)

   *(Models do exist for violations of 2-4)*
What makes this a difficult problem?
What makes this a difficult problem?

1. Heterogeneity
What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model $M_h$
What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model $M_h$
- **Samples differ** – some samples are more successful: model $M_t$
What makes this a difficult problem?

1. Heterogeneity

- **Individuals differ** – some are more easily caught: model $M_h$
- **Samples differ** – some samples are more successful: model $M_t$
- **Both sources** – model $M_{th}$
What makes this a difficult problem?

1. Heterogeneity

- Individuals differ – some are more easily caught: model $M_h$
- Samples differ – some samples are more successful: model $M_t$
- Both sources – model $M_{th}$

2. Identifiability

- $N$ is identifiable within but not between model classes. e.g. finite vs. beta (infinite) mixture models for $P$
1. **Individuals** – animals/errors differ in their catchability

Individuals belong to $A$ latent classes:

- Membership probabilities $\{\pi_a\}$
- Capture probabilities $\{\phi_a\}$

\[
f(p_i) = \sum_{a=1}^{A} \pi_a \delta(p_i - \phi_a)
\]

\[
\sum_{a=1}^{A} \pi_a = 1
\]
2. **Samples** – probability of capture varies between samples

- **Model** $M_{t[f]h}$: $k$ fixed effects

  $$p_{ij} = \theta_{aj}$$
  $$\logit(\theta_{aj}) = \logit(\phi_a) + \beta_j + \gamma_{aj}$$

  Individual $i$ is in latent class $a$
2. **Samples** – probability of capture varies between samples

- **Model** $M_{t[f]h}$: $k$ fixed effects

  \[
  p_{ij} = \theta_{aj} \\
  \logit(\theta_{aj}) = \logit(\phi_a) + \beta_j + \gamma_{aj}
  \]

  Individual $i$ is in latent class $a$

OR
2. **Samples** – probability of capture varies between samples

- **Model** $M_{t[f]h}$: $k$ fixed effects

  \[
  p_{ij} = \theta_{aj} \\
  \logit(\theta_{aj}) = \logit(\phi_a) + \beta_j + \gamma_{aj}
  \]

  Individual $i$ is in latent class $a$

  **OR**

- **Model** $M_{th}$ $B$ latent classes: membership probabilities $\{\lambda_b\}$ capture probabilities $\{\psi_b\}$

  \[
  p_{ij} = \theta_{ab} \\
  \logit(\theta_{ab}) = \logit(\phi_a) + \logit(\psi_b) - \logit(\phi_1) + \gamma_{ab}
  \]

  Individual $i$ is in latent class $a$ and
  Sample $j$ is in latent class $b
Identifiability

- Ordering constraints on mixture component probabilities
  
  \[ 0 < \phi_1 < \ldots < \phi_A < 1 \]  
  and  
  \[ \phi_1 = \psi_1 < \psi_2 < \ldots < \psi_B < 1 \]

  Implemented by ‘repulsive’ priors (following Green 1995)

  \[ \phi | A \sim \text{EOS}(\text{Uniform}(0, 1)^{2A+1}) \]

  \{\phi_a\} are the even order statistics of \((2A + 1)\) draws from \(U(0, 1)\)

- Sum to zero constraints on fixed effects
  
  \[ \sum_j \beta_j = \sum_a \gamma_{aj} = \sum_j \gamma_{aj} = 0 \]  
  or

  \[ \sum_a \gamma_{ab} = \sum_b \gamma_{ab} = 0 \]

  Ensured by degenerate Normal priors, e.g.

  \[ \beta_j \sim \text{Normal}(0, \sigma^{'2}_{\beta}) \]  
  with  
  \[ \sum j \beta_j = 0 \]
## Priors

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
<th>Constants</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Geometric($1 - \eta$)</td>
<td>$\eta = 0.999$</td>
<td>★ ★</td>
</tr>
<tr>
<td>$A$</td>
<td>Geometric($1 - \rho_h$)</td>
<td>$\rho_h = 0.8$</td>
<td>★ ★</td>
</tr>
<tr>
<td>$B$</td>
<td>Geometric($1 - \rho_h$)</td>
<td>$\rho_h = 0.8$</td>
<td>★ ★</td>
</tr>
<tr>
<td>$\pi</td>
<td>A$</td>
<td>Dirichlet($\alpha_h$)</td>
<td>$\alpha_h = \frac{3}{2}$</td>
</tr>
<tr>
<td>$\phi</td>
<td>A$</td>
<td>EOS(Uniform($0,1)^{2A+1}$)</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_\beta$</td>
<td>InverseGamma($v_\beta, \kappa_\beta$)</td>
<td>$v_\beta = 3, \kappa_\beta = 40$</td>
<td>★</td>
</tr>
<tr>
<td>$\beta$</td>
<td>DegenNormal($k;0,\sigma^2_\beta$)</td>
<td></td>
<td>★</td>
</tr>
<tr>
<td>$\lambda</td>
<td>B$</td>
<td>Dirichlet($\alpha_t$)</td>
<td>$\alpha_t = \frac{3}{2}$</td>
</tr>
<tr>
<td>$\psi</td>
<td>B, \phi_1$</td>
<td>EOS(Uniform($\phi_1,1)^{2B-1}$)</td>
<td></td>
</tr>
<tr>
<td>$\zeta</td>
<td>A, (B)$</td>
<td>Bernoulli($p_\gamma$)</td>
<td>$p_\gamma = \frac{1}{2}$</td>
</tr>
<tr>
<td>$\gamma</td>
<td>\zeta, A, (B)$</td>
<td>DegenNormal($A,B;0,\sigma^2_\gamma$)</td>
<td>$\sigma_\gamma = 5$</td>
</tr>
<tr>
<td>$b</td>
<td>B, \lambda$</td>
<td>Discrete(${1, \ldots, B}; \lambda$)</td>
<td></td>
</tr>
</tbody>
</table>
Bayesian Estimation – RJMCMC

Estimation by Reversible Jump MCMC

- Draw from posterior

\[ p(\{m, \Theta_m\} | X) \]

- Reversible Jump MCMC allows dimension switching moves:
  Addition/deletion of mixture components.

- 4 types of MCMC move in model \( M_h \).
### RJMCMC Updates

#### MCMC Step

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_1^{(1)}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_2^{(1)}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_3^{(1)}$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### RJMCMC Updates

<table>
<thead>
<tr>
<th>MCMC Step</th>
<th>1</th>
<th>2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1^{(1)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(2)}$</td>
</tr>
<tr>
<td>$\theta_2^{(1)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(2)}$</td>
</tr>
<tr>
<td>$\theta_3^{(1)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(2)}$</td>
</tr>
</tbody>
</table>
RJMCMC Updates

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1^{(1)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(3)}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_2^{(1)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(3)}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_3^{(1)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(3)}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## RJMCMC Updates

### MCMC Step

<table>
<thead>
<tr>
<th>MCMC Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1^{(1)} )</td>
<td>( \theta_1^{(2)} )</td>
<td>( \theta_1^{(3)} )</td>
<td>( \theta_1^{(4)} )</td>
<td></td>
</tr>
<tr>
<td>( \theta_2^{(1)} )</td>
<td>( \theta_2^{(2)} )</td>
<td>( \theta_2^{(3)} )</td>
<td>( \theta_2^{(4)} )</td>
<td></td>
</tr>
<tr>
<td>( \theta_3^{(1)} )</td>
<td>( \theta_3^{(2)} )</td>
<td>( \theta_3^{(3)} )</td>
<td>( \theta_3^{(4)} )</td>
<td></td>
</tr>
</tbody>
</table>
### RJMCMC Updates

**MCMC Step**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ₁⁽¹⁾</td>
<td>θ₁⁽²⁾</td>
<td>θ₁⁽³⁾</td>
<td>θ₁⁽⁴⁾</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>θ₂⁽¹⁾</td>
<td>θ₂⁽²⁾</td>
<td>θ₂⁽³⁾</td>
<td>θ₂⁽⁴⁾</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>θ₃⁽¹⁾</td>
<td>θ₃⁽²⁾</td>
<td>θ₃⁽³⁾</td>
<td>θ₃⁽⁴⁾</td>
<td>…</td>
<td></td>
</tr>
</tbody>
</table>
**RJMCMC Updates**

### MCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_1^{(1)})</td>
<td>(\theta_1^{(2)})</td>
<td>(\theta_1^{(3)})</td>
<td>(\theta_1^{(4)})</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>(\theta_2^{(1)})</td>
<td>(\theta_2^{(2)})</td>
<td>(\theta_2^{(3)})</td>
<td>(\theta_2^{(4)})</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>(\theta_3^{(1)})</td>
<td>(\theta_3^{(2)})</td>
<td>(\theta_3^{(3)})</td>
<td>(\theta_3^{(4)})</td>
<td>…</td>
<td></td>
</tr>
</tbody>
</table>

### RJMCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(m^{(1)})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:1}^{(1)})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:2}^{(1)})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### RJMCMC Updates

#### MCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_1^{(1)})</td>
<td>(\theta_1^{(2)})</td>
<td>(\theta_1^{(3)})</td>
<td>(\theta_1^{(4)})</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>(\theta_2^{(1)})</td>
<td>(\theta_2^{(2)})</td>
<td>(\theta_2^{(3)})</td>
<td>(\theta_2^{(4)})</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>(\theta_3^{(1)})</td>
<td>(\theta_3^{(2)})</td>
<td>(\theta_3^{(3)})</td>
<td>(\theta_3^{(4)})</td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

#### RJMCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(m^{(1)})</td>
<td>(m^{(2)})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:1}^{(1)})</td>
<td>(\theta_{m^{(2)}:1}^{(2)})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:2}^{(1)})</td>
<td>(\theta_{m^{(2)}:2}^{(2)})</td>
<td>(\theta_{m^{(2)}:3}^{(2)})</td>
<td>(\theta_{m^{(2)}:4}^{(2)})</td>
<td></td>
</tr>
</tbody>
</table>
### RJMCMC Updates

#### MCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>$\theta_1^{(1)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(3)}$</td>
<td>$\theta_1^{(4)}$</td>
<td>...</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>$\theta_2^{(1)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(3)}$</td>
<td>$\theta_2^{(4)}$</td>
<td>...</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>$\theta_3^{(1)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(3)}$</td>
<td>$\theta_3^{(4)}$</td>
<td>...</td>
</tr>
</tbody>
</table>

#### RJMCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>$m^{(1)}$</td>
<td>$m^{(2)}$</td>
<td>$m^{(3)}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\theta^{(1)}_{m^{(1)}:1}$</td>
<td>$\theta^{(2)}_{m^{(2)}:1}$</td>
<td>$\theta^{(3)}_{m^{(3)}:1}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\theta^{(1)}_{m^{(1)}:2}$</td>
<td>$\theta^{(2)}_{m^{(2)}:2}$</td>
<td>$\theta^{(3)}_{m^{(3)}:2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\theta^{(1)}_{m^{(1)}:3}$</td>
<td>$\theta^{(2)}_{m^{(2)}:3}$</td>
<td>$\theta^{(3)}_{m^{(3)}:3}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\theta^{(1)}_{m^{(1)}:4}$</td>
<td>$\theta^{(2)}_{m^{(2)}:4}$</td>
<td>$\theta^{(3)}_{m^{(3)}:4}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## RJMCMC Updates

### MCMC Step

<table>
<thead>
<tr>
<th>MCMC Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_1^{(1)})</td>
<td>(\theta_1^{(2)})</td>
<td>(\theta_1^{(3)})</td>
<td>(\theta_1^{(4)})</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>(\theta_2^{(1)})</td>
<td>(\theta_2^{(2)})</td>
<td>(\theta_2^{(3)})</td>
<td>(\theta_2^{(4)})</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>(\theta_3^{(1)})</td>
<td>(\theta_3^{(2)})</td>
<td>(\theta_3^{(3)})</td>
<td>(\theta_3^{(4)})</td>
<td>…</td>
<td></td>
</tr>
</tbody>
</table>

### RJMCMC Step

<table>
<thead>
<tr>
<th>RJMCMC Step</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m^{(1)})</td>
<td>(m^{(2)})</td>
<td>(m^{(3)})</td>
<td>(m^{(4)})</td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:1}^{(1)})</td>
<td>(\theta_{m^{(2)}:1}^{(2)})</td>
<td>(\theta_{m^{(3)}:1}^{(3)})</td>
<td>(\theta_{m^{(4)}:1}^{(4)})</td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(1)}:2}^{(1)})</td>
<td>(\theta_{m^{(2)}:2}^{(2)})</td>
<td>(\theta_{m^{(3)}:2}^{(3)})</td>
<td>(\theta_{m^{(4)}:2}^{(4)})</td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(2)}:3}^{(2)})</td>
<td>(\theta_{m^{(2)}:3}^{(3)})</td>
<td>(\theta_{m^{(2)}:3}^{(4)})</td>
<td>(\theta_{m^{(4)}:3}^{(4)})</td>
<td></td>
</tr>
<tr>
<td>(\theta_{m^{(2)}:4}^{(2)})</td>
<td>(\theta_{m^{(2)}:4}^{(3)})</td>
<td>(\theta_{m^{(2)}:4}^{(4)})</td>
<td>(\theta_{m^{(4)}:4}^{(4)})</td>
<td></td>
</tr>
</tbody>
</table>
### RJMCMC Updates

#### MCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1^{(1)}$</td>
<td>$\theta_1^{(2)}$</td>
<td>$\theta_1^{(3)}$</td>
<td>$\theta_1^{(4)}$</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>$\theta_2^{(1)}$</td>
<td>$\theta_2^{(2)}$</td>
<td>$\theta_2^{(3)}$</td>
<td>$\theta_2^{(4)}$</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>$\theta_3^{(1)}$</td>
<td>$\theta_3^{(2)}$</td>
<td>$\theta_3^{(3)}$</td>
<td>$\theta_3^{(4)}$</td>
<td>…</td>
<td></td>
</tr>
</tbody>
</table>

#### RJMCMC Step

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m^{(1)}$</td>
<td>$m^{(2)}$</td>
<td>$m^{(3)}$</td>
<td>$m^{(4)}$</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>$\theta_{m^{(1)}:1}$</td>
<td>$\theta_{m^{(2)}:1}$</td>
<td>$\theta_{m^{(3)}:1}$</td>
<td>$\theta_{m^{(4)}:1}$</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>$\theta_{m^{(1)}:2}$</td>
<td>$\theta_{m^{(2)}:2}$</td>
<td>$\theta_{m^{(3)}:2}$</td>
<td>$\theta_{m^{(4)}:2}$</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>$\theta_{m^{(1)}:3}$</td>
<td>$\theta_{m^{(2)}:3}$</td>
<td>$\theta_{m^{(3)}:3}$</td>
<td>…</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{m^{(2)}:4}$</td>
<td>$\theta_{m^{(3)}:4}$</td>
<td>…</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(1) Shift a support point
(2) Exchange probability between points
(3) Split a support point

\[ \pi_{a-1} \rightarrow \pi_{a} \rightarrow \pi_{a+1} \]

\[ \phi_{a-1} \rightarrow \phi_{a} \rightarrow \phi_{a+1} \]

\[ s, (1-u_1)s, u_2(1-s) \]
(4) Merge two support points
Estimation of finite mixture models affected if the MCMC sampler can’t mix because of

- artificial identifiability constraints,
  (e.g. so that individuals/samples become persistently misallocated) or
- updating protocols making one mode inaccessible from another.
Label switching and poor mixing

Estimation of finite mixture models affected if the MCMC sampler cannot mix because of

- artificial identifiability constraints, (e.g. so that individuals/samples become persistently misallocated) or
- updating protocols making one mode inaccessible from another.

In our case:

- Mixture components are labelled by a single parameter: have a natural ordering (e.g. preventing misallocations);
- Dimension-switching RJMCMC allows components to be deleted and reappear – the chain moves rapidly around model space.
$k = 6$ reviewers tested switches and found $D = 43$ errors (Basu 1998)

<table>
<thead>
<tr>
<th>Error</th>
<th>Reviewer $j$</th>
<th>$x_i$</th>
<th></th>
<th>Error</th>
<th>Reviewer $j$</th>
<th>$x_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 0 0 0 0 0 0</td>
<td>1</td>
<td>23</td>
<td>0 0 0 0 0 0 1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0 0 0 1 0 1</td>
<td>2</td>
<td>24</td>
<td>0 0 0 0 1 0 1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0 0 0 0 1 0</td>
<td>1</td>
<td>25</td>
<td>0 0 0 1 1 0 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1 0 0 0 0 0</td>
<td>1</td>
<td>26</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0 0 0 1 0 0</td>
<td>1</td>
<td>27</td>
<td>0 0 0 1 0 1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0 0 0 1 0 0</td>
<td>1</td>
<td>28</td>
<td>0 0 1 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0 0 0 1 0 0</td>
<td>1</td>
<td>29</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1 0 0 0 0 0</td>
<td>1</td>
<td>30</td>
<td>1 0 0 1 1 0 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1 0 0 0 0 0</td>
<td>1</td>
<td>31</td>
<td>1 0 0 1 0 1 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0 0 0 1 0 0</td>
<td>1</td>
<td>32</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>33</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1 0 0 0 0 0</td>
<td>1</td>
<td>34</td>
<td>0 0 1 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>35</td>
<td>0 0 1 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1 0 0 0 0 0</td>
<td>1</td>
<td>36</td>
<td>0 0 0 0 0 1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>37</td>
<td>0 0 1 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>38</td>
<td>1 0 0 0 0 1 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1 1 0 1 1 1</td>
<td>1</td>
<td>39</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1 0 0 0 0 1</td>
<td>2</td>
<td>40</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0 1 0 0 0 0</td>
<td>1</td>
<td>41</td>
<td>0 0 0 0 1 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0 1 0 0 0 0</td>
<td>1</td>
<td>42</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>43</td>
<td>1 0 0 0 0 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1 0 0 1 0 0</td>
<td>2</td>
<td>44</td>
<td>n_j = 60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 1

$$\log L = -91.9596431614535$$

$\pi$

$\theta$

Rh(2)

$N=43$
Step 2

\[ \log L = -67.92296466689119 \]

\[
\begin{array}{c}
\pi \\
0.0 & 0.4 & 0.8 \\
\hline
0.0 & 0.2 & 0.4 & 0.6 & 0.8 & 1.0
\end{array}
\]

\( \theta \)

Rh(4)
N= 44
Step 3

$$\log L = -58.1418394830594$$

Rh(5)
N = 46
Step 4

\[
\log L = -60.4784899550032
\]

Rh(6)
N = 44
Step 5

log L = -58.7074675287983

Rh(7)
N = 45
Step 6

\[ \log L = -60.0611555129325 \]

Diagram showing points at various \( \theta \) values with corresponding \( \pi \) values for Rh(6), \( N = 46 \).
Step 7

\[ \log L = -58.0130422680327 \]

Rh(5)
N = 47
Step 8

\[ \log L = -55.1097509120916 \]

![Graph showing \( \theta \) and \( \pi \) with Rh(5) and N=44]
Step 9

$$\log L = -49.5727659877251$$

$\pi$ $0.0$ $0.4$ $0.8$

$\theta$

$0.0$ $0.2$ $0.4$ $0.6$ $0.8$ $1.0$

$Rh(6)$

$N= 46$
Step 10

\[
\log L = -46.5488986935539
\]

- $\theta$
- $\pi$
- $\log L = -46.5488986935539$
- $\text{Rh(6)}$
- $N = 45$
Step 11

\[ \log L = -45.664424798022 \]

Graph:
- \( \pi \) on the vertical axis
- \( \theta \) on the horizontal axis
- Points at \( \theta \) values 0.2, 0.4, 0.6, 0.8, 1.0 with corresponding \( \pi \) values
- \( Rh(6) \)
- \( N = 46 \)
Step 12

\[ \log L = -45.0363060198825 \]

\[ \pi \]

\[ N = 47 \]

\[ \text{Rh(6)} \]
Step 13

\[ \log L = -44.6636665499824 \]

\[ \pi \]

\[ N = 47 \]

\[ \text{Rh}(7) \]
$\log L = -47.0076330888468$
Step 15

$log L = -47.6636985662387$

Rh(7)
N = 48
Step 16

$log L = -45.737454851479$

![Graph showing data points with labels Rh(8) and N=44]
Step 17

\[ \log L = -42.4371645695473 \]

Rh(7)
N = 48
log L = −41.4345231830313
Step 19

\[ \log L = -39.3970709062431 \]

Diagram showing

- \( \pi \)
- \( 0.0 \), \( 0.4 \), \( 0.8 \)
- \( \theta \)
- \( 0.0 \), \( 0.2 \), \( 0.4 \), \( 0.6 \), \( 0.8 \), \( 1.0 \)
- Rh(5)
- \( N = 48 \)
Step 20

\[ \log L = -40.5227716139258 \]
Step 21

\[ \log L = -37.5693882100671 \]

\begin{center}
\begin{tikzpicture}
\draw[very thick, ->] (0,0) -- (1.5,0) node[below] {$\theta$};
\draw[very thick, ->] (0,0) -- (0,1.5) node[left] {$\pi$};
\filldraw[black] (0.2,0.7) circle (2.5pt);
\filldraw[black] (0.4,0.3) circle (2.5pt);
\filldraw[black] (0.6,0.1) circle (2.5pt);
\filldraw[black] (0.8,0.05) circle (2.5pt);
\filldraw[black] (1.0,0.2) circle (2.5pt);
\node at (1.5,1.0) {Rh(7)};
\node at (1.5,0.8) {N= 44};
\end{tikzpicture}
\end{center}
Step 22

\[ \log L = -36.4306162111048 \]

\[ \text{Rh(7)} \]
\[ N = 47 \]
Step 23

\[ \log L = -33.0389452678949 \]

\[ \theta \]

\[ \pi \]

\[ N = 46 \]

Rh(5)
Step 24

\[ \log L = -32.9077337310434 \]

![Graph showing \( \pi \) vs. \( \theta \) with \( N = 52 \) and \( \text{Rh}(6) \)]
Step 25

\[ \log L = -30.4732050323583 \]

\[ \theta \]

\[ \pi \]

\[ Rh(7) \]
\[ N= 44 \]
log L = $-28.4096812545929$

Rh(8)
N= 47
log L = −27.59453640333

Rh(7)
N= 46
Step 28

$$\log L = -27.1528912375878$$

![Graph](image-url)
Step 29

$$\log L = -26.2582030338909$$

![Graph showing data points for \( \pi \) and \( \theta \) with \( N = 52 \).]
Step 30

$log L = -25.7864676274672$

Rh(3)
N= 51
log L = −22.3220481167324

Rh(3)
N= 48
Step 32

\[ \log L = -19.8536269710666 \]

\[ \theta \]

\[ \pi \]

\[ Rh(4) \]

\[ N= 49 \]
Step 33

\[ \log L = -22.6319674062755 \]

\[ \text{Rh}(3) \]
\[ N = 51 \]
log L = \(-20.6214301093238\)
Step 35

\[
\log L = -10.2673835002677
\]

\[Rh(3)\]

\[N= 59\]
Step 36

$log L = -10.819486250473$

Rh(4)
N = 64
Step 37

\[ \log L = -11.5513213071656 \]
Step 38

\[ \log L = -9.16791808262374 \]

\[ \text{Rh}(3) \]
\[ N = 62 \]
Step 39

\[ \log L = -11.3187838098980 \]

\[ \theta \]

\[ \pi \]

\[ 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \]

\[ N = 55 \]

Rh(2)
Step 40

\[ \log L = -6.50771712565023 \]

![Graph showing \( \pi \) vs. \( \theta \) with a dot at \( \theta = 0.2 \) and \( N = 66 \).]
Step 41

\[ \log L = -7.77292212428978 \]

\[ \pi \]

\[ \theta \]

\[ \text{Rh(1)} \]

\[ N = 79 \]
log $L = -9.07442912991769$

Rh(2)

N = 58
log L = $-6.7347692978596$
log $L = -6.85775153090216$

![Graph showing $\pi$ and $\theta$ with Rh(3) and N=80 annotations.](image)
Step 45

\[ \log L = -6.44686591772984 \]

\[ \pi \]

Rh(1)

\[ N = 68 \]

\[ \theta \]
Step 46

\[ \log L = -8.34608316367388 \]

\[ \theta \]

\[ \pi \]

\[ \rho(1) \]

\[ N = 106 \]
Step 47

\[ \log L = -7.24947491612883 \]

\[ \pi \]

Rh(1)
N= 93

\[ 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \]

\[ 0.0 \quad 0.4 \quad 0.8 \]

\[ \theta \]

\[ \text{log L} = -7.24947491612883 \]
Step 48

\[ \log L = -8.68619170427243 \]

\[ \theta \]

\[ \pi \]

\[ N = 77 \]

\[ \text{Rh(1)} \]
Step 49

\[ \log L = -7.03769307699849 \]

\( \theta \)

\( \pi \)

\( \text{Rh}(1) \)

\( N = 80 \)
Step 50

\[ \log L = -6.54071271014857 \]

Rh(1)
N = 66
Step 51

\[ \log L = -6.99200544598219 \]

The graph shows a plot of \( \pi \) against \( \theta \) with a point at \( \pi = 0.2 \) and \( \theta = 0.2 \), indicating \( \text{Rh(1)} \) and \( N = 89 \).
Step 52

\[ \log L = -6.92828015307384 \]

\[ \pi \]

\[ \theta \]

Rh(1)  
N = 65
Step 53

\[
\log L = -7.0832078891851
\]

Rh(2)
N = 76
Step 54

\[ \log L = -6.47554471024264 \]

\[ Rh(1) \]
\[ N = 72 \]
log L = -6.7014630905897

Rh(1)
N = 65
Step 56

$$\log L = -5.64197053930326$$

Rh(2)
N= 81
Step 57

\[ \log L = -5.96373888768323 \]

\[ \pi \]

\[ N = 69 \]
log L = -8.84591665324345

\[ \pi \]

N = 67

Rh(1)
Step 59

\[ \log L = -7.64672016924314 \]

Rh(1)
N = 57
log L = -6.73174374308653

Rh(1)
N= 74
log $L = -6.55142574603127$

$\pi$

$\theta$

$N = 72$

Rh(1)
Step 62

\[
\log L = -6.69541020421036
\]

\[
\pi
\]

\[
\theta
\]

Rh(1)

N = 83
\[
\log L = -6.8475622252277
\]
Step 64

\[ \log L = -5.22857102139116 \]

\[ \pi \]

\[ \theta \]

Rh(2)

N= 83
Step 65

\[ \log L = -4.81063005314064 \]

\[ \theta \]

\[ \pi, \mu, \theta \]

\[ \text{Rh(3)} \]

\[ N = 77 \]
Step 66

\[ \log L = -5.46413707072475 \]

\[ \pi \\
0.0 \ 0.4 \ 0.8 \]

\[ \theta \\
0.0 \ 0.2 \ 0.4 \ 0.6 \ 0.8 \ 1.0 \]

\[ \text{Rh}(2) \]
\[ N=93 \]
log L = \(-7.88242747736024\)
Step 68

$$\log L = -6.48263347003243$$

Diagram showing a plot with markers labeled $\pi$ and $\theta$. The plot indicates data points with $N=91$.
Step 69

\[ \log L = -7.91704677251013 \]

![Graph showing \( \pi \) vs. \( \theta \) with points at \( \pi \approx 0.2 \) and \( \theta \approx 0.1 \), labeled as \( \text{Rh}(2) \) with \( N = 99 \).]
Step 70

\[ \log L = -7.41664781432527 \]

\[ \pi \]

\[ \theta \]
log L = −7.36978157074631
log L = \(-7.42029372486004\)
log L = -6.50030323887847

Rh(1)
N = 77
Step 74

\[ \log L = -7.08379337128636 \]

\[ \pi \]

0.0 0.4 0.8

\[ \theta \]

0.0 0.2 0.4 0.6 0.8 1.0

Rh(1)
N= 61
Step 75

\[ \log L = -6.98594579627081 \]

Diagram:
- \( \pi \) vs \( \theta \)
- Point labeled \( \text{Rh(1)} \)
- \( N = 63 \)
Step 76

$$\log L = -8.07461582671763$$

Rh(1)
N= 87
log L = -8.69088322869459

Rh(3)
N = 137
\[
\log L = -9.06991579820229
\]

Rh(3)

\[ N = 185 \]
log L = -6.59579706710454

Rh(2)
N= 141
Step 80

\[ \log L = -6.72727099921829 \]

-log L = -6.72727

\begin{align*}
\pi & \quad 0.0 \quad 0.4 \quad 0.8 \\
0.0 & \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0
\end{align*}

Rh(2)

N = 142
\[ \log L = -6.37796903283765 \]
log $L = -8.43700588880688$
log $L = -6.2297006532936$
\log L = -6.59598781624027

\pi

\theta

Rh(2)
N= 119
Step 85

\[ \log L = -6.53657057131011 \]

Graph with coordinates: 
- \( \pi \) vs \( \theta \)
- Points at \( \pi = 0, 0.4, 0.8 \)
- \( \theta \) ranging from 0.0 to 1.0
- \( \text{Rh(2)} \) with \( N = 106 \)
Step 86

$$\log L = -6.49950133785143$$

\[ \pi \quad 0.0 \quad 0.4 \quad 0.8 \]

\[ 0.0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1.0 \]

Rh(1)

N= 69
Step 87

$\log L = -5.86507082741511$

$\pi$

$\theta$

Rh(2)

$N = 68$
\[ \log L = -5.15535623326514 \]

- The plot shows a point at \( \theta \approx 0.2 \) with \( \pi \approx 0.8 \).
- The point is labeled with \( \text{Rh}(2) \) and \( N = 99 \).
Step 89

\[ \log L = -7.52915403726294 \]

**Diagram:**

\( \pi \)

\( \theta \)

\( \text{Rh}(1) \)

\( N = 97 \)
log L = −5.42621622335881

Rh(2)
N= 81
\[ \log L = -7.53051945806166 \]
Step 92

\[ \log L = -6.48125416952934 \]

\[ \theta \]

Rh(1)

N = 67
Step 93

$\log L = -5.56786282597483$

$\pi$

$0.0$ $0.2$ $0.4$ $0.6$ $0.8$ $1.0$

$\theta$

$N = 94$

$\text{Rh}(2)$
log L = -7.61592304565889

Rh(2)
N= 137
Step 95

$$\log L = -7.14847973079398$$

$\theta$

Rh(3)

$N = 148$
\[ \log L = -6.96778206285407 \]
Step 97

$log L = -7.71918326510837$

$\pi$

$0.0$ $0.2$ $0.4$ $0.6$ $0.8$ $1.0$

$\theta$

$0.0$ $0.4$ $0.8$

$Rh(3)$

$N = 137$
\[ \log L = -9.8181414985707 \]
log $L = -6.64868771633988$
Step 100

\[ \log L = -5.83175113202486 \]

\[ Rh(2) \]
\[ N = 111 \]
$M_{th}$: 2-way finite mixture
**Application: AT&T Switch Testing**

$M_{t[f]h}$: 1-way finite mixture with fixed sample effects

![Bar chart showing probability distribution for $M_{t[f]h}$ with different values of $A$.]
## Posterior Model Probabilities

<table>
<thead>
<tr>
<th>A</th>
<th>( p(A, \zeta \mid X, M_t \mid h) )</th>
<th>( p(A, B, \zeta \mid X, M_{th}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B = )</td>
<td>( \zeta = 0 )</td>
<td>( \zeta = 1 )</td>
</tr>
<tr>
<td>1</td>
<td>58.3</td>
<td>27.5</td>
</tr>
<tr>
<td>2</td>
<td>29.0</td>
<td>8.9</td>
</tr>
<tr>
<td>3</td>
<td>9.2</td>
<td>1.3</td>
</tr>
<tr>
<td>4</td>
<td>2.4</td>
<td>0.2</td>
</tr>
<tr>
<td>5</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>
### Posterior Means

| A | $E[N|X, M_{t[f|n]}]$ | $E[N|X, M_{th}]$ |
|---|-----------------|-----------------|
|   | $\zeta = 0$   | $\zeta = 1$   | $\zeta = 0$ | $\zeta = 1$ |
| 1 | 67.8  | 78.5 | 65.9 | 64.7 | 64.3 | 64.2 | 64.2 | 79.4 | 79.2 | 76.8 | 76.9 | 75.9 |
| 2 | 78.6  | 103.6| 72.1 | 70.5 | 70.1 | 70.0 | 70.0 | 76.1 | 81.4 | 82.9 | 83.4 | 80.0 |
| 3 | 79.6  | 100.0| 71.2 | 69.8 | 69.2 | 68.3 | 70.4 | 75.6 | 85.8 | 77.8 | 73.6 | 72.8 |
| 4 | 78.1  | 71.0 | 70.8 | 67.7 | 67.5 | 68.9 | 75.6 | 78.8 | 75.6 | 67.0 | 69.7 |
| 5 | 75.5  | 68.4 | 67.1 | 69.5 | 64.7 | 64.0 | 67.5 | 78.8 | 75.6 | 67.0 | 69.7 |
| 6 | 75.0  | 70.2 | 64.9 | 62.0 | 67.8 | 58.0 | 65.0 | 91.3 | 78.2 | 64.5 | 66.1 |
| 7 | 76.4  | 62.0 |       |       |       |       |       |       |       |       |       |
| 8 | 84.0  |       |       |       |       |       |       |       |       |       |       |
| 9 | 83.0  |       |       |       |       |       |       |       |       |       |       |
| 10| 64.0  |       |       |       |       |       |       |       |       |       |       |
Population size estimates $\hat{N}$ in fixed and variable dimension models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Med.</th>
<th>95% CI</th>
<th>Npar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_0$</td>
<td>78.5</td>
<td>76</td>
<td>(63,101)</td>
<td>2</td>
</tr>
<tr>
<td>$M_{t_2}$</td>
<td>65.9</td>
<td>66</td>
<td>(54,97)</td>
<td>10</td>
</tr>
<tr>
<td>$M_{t_3}$</td>
<td>64.7</td>
<td>64</td>
<td>(55,77)</td>
<td>12</td>
</tr>
<tr>
<td>$M_{t_4}$</td>
<td>64.3</td>
<td>62</td>
<td>(60,64)</td>
<td>14</td>
</tr>
<tr>
<td>$M_{t_2}+h_2$</td>
<td>72.1</td>
<td>70</td>
<td>(63,77)</td>
<td>12</td>
</tr>
<tr>
<td>$M_{t_2}\times h_2$</td>
<td>79.4</td>
<td>73</td>
<td>(55,122)</td>
<td>13</td>
</tr>
<tr>
<td>$M_{th}$</td>
<td>68.7</td>
<td>66</td>
<td>(51,103)</td>
<td></td>
</tr>
<tr>
<td>$M_{t[f]}$</td>
<td>67.8</td>
<td>66</td>
<td>(52,94)</td>
<td>8</td>
</tr>
<tr>
<td>$M_{t[f]}+h_2$</td>
<td>78.6</td>
<td>74</td>
<td>(54,129)</td>
<td>10</td>
</tr>
<tr>
<td>$M_{t[f]}\times h_2$</td>
<td>72.4</td>
<td>69</td>
<td>(52,114)</td>
<td>15</td>
</tr>
</tbody>
</table>
Posterior distribution of $N$
Application: AT&T Switch Testing

- Models $M_{th}$ and $M_{tfh}$ both predict of the order 70 errors in total (95% credible intervals [51,103] and [52,114])

- $\sim$ 30 errors not so far detected

- Evidence for heterogeneity among reviewers, but not amongst faults

- Bayes Factor $BF(M_{th} : M_{tfh}) = 0.42$: neither model class convincingly favoured.
Skills of the 6 reviewers

Capture Probability

![Graph showing the skills distribution of 6 reviewers.](image)
Application: Snowshoe Hares

![Graph showing distribution of N given X for different values of A (1, 2, 3). The x-axis represents N, and the y-axis represents p(N|X). There are two lines representing M_{th} and M_{th} with different styles.](image-url)
Interval estimates – 3 examples

![Graph showing interval estimates for Hares, AT&T, and Rabbits]
Interval estimates – 3 examples

N

0 50 100 150 200

Hares

AT&T

Rabbits

N

1 2 3 4 5

1 2 3 4 5

1 2 3 4 5
RJMCMC is a practicable means for model selection/averaging in Capture-Recapture with finite mixtures.

Doesn’t solve non-identifiability between model classes

Priors regularise the likelihood – adding extra components does not affect estimates much.