Lecture 11 – Multi-state models

Resources


Williams et al. Chapter 17.3-17.5


Overview

1. Further generalization of the original CJS model.
2. Similar in many respects to the age-structured open model – survival was transition to the next time period or age
3. Also similar in many respects to time variable covariates – survival in condition or state at time t equals transition to the next condition
4. Multi-state or multi-strata models
   a. Individuals classified into discrete states that could represent locations, ages, size classes, etc.
   b. Models permit stochastic transitions or movements among states
   c. Contrast with deterministic transitions among ages
   d. Description of the population requires sampling all states (e.g., locations)
   e. Estimate capture and survival probabilities of individuals in different states
   f. Additionally, estimate a suite of transition probabilities
   g. First developed by Arnason (1972, 1973) using moment estimators
   h. MLEs first developed by Hestbeck et al. (1991) for multiple locations
   i. Nichols et al. (1992) developed for a multiple phenotypic states (sizes)
   j. Brownie et al. (1993) and Schwartz et al. (1993) developed more generalized approaches
   k. Recently, Lebreton et al. suggested that multi-state models formed a unifying approach for all mark-recapture models
   l. Biological questions:
      1) Estimates of stage- (or age-) specific survival rates for use in population models
      2) Survival rates of breeding and nonbreeding states for examining costs of reproduction and fitness tradeoffs.
      3) Ecologists and managers have become more aware that animal distributions do not necessarily reflect habitat quality without measures of fitness-related
parameters such as survival and reproduction. Multi-strata models can be used to examine habitat quality as well as movement rates among habitats.

5. Markovian models
   a. State is stochastically determined as a function of state at some previous time – Markov process.
      a. Typically consider only the first order Markov process – state at t+1 is a function of state at t.
      b. Memory model – 2nd order Markov process – state at t+1 is a function of state at t and t-1. Theoretically, could be generalized to consider movement as a function of state at any prior time – not considered here.

6. Data structure
   a. Capture histories that indicate state at each occasion.
      a. Use letters or numbers instead of 1s (MARK uses either). 0s indicate failure to capture or observe on that occasion.
      b. Example:
         Capture history: N00BNB indicates individual observed in state N at occasion 1, not seen on occasions 2 and 3, observed in state 4 on occasion 4, observed in N on occasion 5, and finally in B on occasion 6.
   c. Summary statistics:
      Releases on occasion i
      \[
      \begin{bmatrix}
      R_i^1 \\
      R_i^2 \\
      R_i^3
      \end{bmatrix}
      \text{for } i = 1:K-1, \text{ and}
      \]
      Recaptures of animals \( m_{ij}^{rs} \) released in state r at time i and recaptured in state s at time j.
      \[
      m_{ij} = \begin{bmatrix}
      m_{ij}^{11} & m_{ij}^{12} & m_{ij}^{13} \\
      m_{ij}^{21} & m_{ij}^{22} & m_{ij}^{23} \\
      m_{ij}^{31} & m_{ij}^{32} & m_{ij}^{33}
      \end{bmatrix}, \text{ for } i = 1:K-1, \text{ and } j = i + 1:K,
      \]
      d. Can be reduced to a modified m-array.
      Ms-array example: Williams et al. (2002 p456)
7. Model structure

Parameters similar to those from the CJS model only now state specific:

\( \varphi^r_s \) - probability of survival in \( s \) at \( i + 1 \), given alive in \( r \) at \( i \). Note that this incorporates probability of survival in \( s \) from \( i \) to \( i+1 \) and transition from state \( r \) to \( s \).

\( p^r_r \) - probability of recapture (encounter) in \( r \) at \( i \).

Example for 2 states \( N \) and \( B \)

\[
Pr(\text{NN0B0} \mid \text{release in N at 1}) = \varphi^N_N p^N_2 N \left[ \varphi^N_N \left(1 - p^N_3 N\right) \varphi^B_B \left(1 - p^B_4^B\right) \varphi^B_B p^B_2^B \right] \\
\times \left(1 - \varphi^B_B p^B_4^B - \varphi^B_B \varphi^N_N \right)
\]

\[
Pr(\text{NN0B0} \mid \text{release in N at 1}) = 
\]

surviving and remaining in \( N \) during 1-2 and seen at 2 in \( N \),

surviving and remaining in \( N \) from 2-3 and not seen in \( N \) at 3 and surviving in \( N \) and moving to \( B \) in 3-4 and seen in \( B \) at 4, or

surviving in \( N \) and moving to \( B \) during 2-3 and not seen in \( B \) at 3 and surviving and remaining in \( B \) in 3-4 and being seen in \( B \) at 4.
Note that $\varphi_{i}^{rs}$ includes remaining in the study population, and survival in $r$ from $i-1$ to $i$, and movement between $r$ and $s$. For many questions of interest it is necessary to decompose these into separate rates of survival ($S_{i}^{r}$) and movement ($\psi_{i}^{rs}$):

$$\varphi_{i}^{rs} = S_{i}^{r} \psi_{i}^{rs}.$$  

It's also important to note that movement is conditional upon survival and the movement probabilities (rates) must sum to one.

$$\sum \psi_{i}^{rs} = 1$$

8. Model assumptions

The first two assumptions of the CJS model are modified as follows:

1) Every marked animal present in state $r$ immediately following sampling period $i$ has the same probability of recapture, and

2) Every marked animal present in state $r$ immediately following the sampling period $i$ has the same probability of surviving until $i + 1$ and moving to state $s$ by period $i + 1$.

Thus, homogeneity assumptions apply only within each state. In addition to the remaining CJS assumptions:

7) The state transition probabilities reflect a first-order Markov process, i.e., state at time $i + 1$ is dependent only on the state at time $i$.

9. Estimation
Estimation is by MLE with one of the transition rates estimated by one minus the sum of the remaining transition probabilities. The default in MARK is for the probability of remaining in r to be estimated as one minus the sum of the other transitions. However, the user can choose which rate is estimated by this difference and it’s useful to pick rates that are small for estimation and large rates to be determined by the difference method.

10. Alternative modeling
   a. Alternative models can be cast similar to the other models we’ve discussed by imposing constraints on parameters.
   b. Survival and transition can be modeled using time specific or individual covariates
   c. Multiple groups are possible
   d. Age models – two ways
      1) As with the CJS open models by constraining parameters
      2) Define each age as a separate state and only estimate the parameters that correspond to the age transitions by fixing all other transitions = 0.

Lebreton et al. (1999) have illustrated that many of the specialized models for age-specific parameters or multiple data types (e.g., live and dead, resighting and recovery) can be developed in the framework of multi-state models. Recently, innovative models have been developed for the situation when initial state is unknown and for the misclassification of state.

11. Model selection, estimator robustness, and model assumptions
    Model selection – as with other MLE approaches use AIC and LRTs.
    GOF – program U-care provides a generalized approach to estimating GOF and c-hat for multistate and robust design.

12. Memory models
13. Reverse time
14. Mark-recapture with auxiliary data
   a. Barker’s model
   b. Burnham’s model
15. Robust design

**Stochastic variation and simulation**

Resources:


Variance components

It is important to understand the sources of variation when developing models for simulations from parameter estimates. Generally there are two sources that are considered—process (environmental) variation and sampling variation.

1. Sampling variance (\(\text{var}(\hat{\theta}|\theta)\)) - a measure of precision and repeatability. Sampling variation occurs because only partial information is available about the population. If the sampling scheme is unbiased, the composition of the sample is a result of a random selection. The
observer has some control over this parameter estimate through the study design (e.g., sample size, stratification, selection of covariates).

2. Process variance ($\sigma^2$) - spatial or temporal deviance in population parameters due to environmental factors. These types of variation are not directly controlled by the observer, and are **characteristics of the population**.

For example, if the annual survival rate ($\theta$) of beach mice at Bon Secour NWR was measured for 10 ($n$) years. The survival rate would most certainly vary among years due to differences in environmental conditions and this would be considered temporal (process) variation.

Another example, the fecundity ($\theta$) of red-breasted sunfish is measured in a dozen ($n$) streams across south Alabama and Georgia to estimate spatial (process) variation in this population parameter.

The theoretical (total) variance is:

$$\text{var}(\hat{\theta}) = \frac{\sigma^2 + \text{var}(\hat{\theta} | \theta)}{n}$$

and it's unbiased estimator is:

$$\hat{\text{var}}(\hat{\theta}) = \frac{\sum_{i=1}^{n} (\hat{\theta}_i - \hat{\theta})^2}{n(n-1)}.$$

In order to estimate the process variation we use a weighted procedure with weights equal to the reciprocal

$$w_i = \frac{1}{\sigma^2 + \text{var}(\hat{\theta}_i | \theta)},$$

$$\hat{\theta} = \frac{\sum_{i=1}^{n} w_i \hat{\theta}_i}{\sum_{i=1}^{n} w_i}$$

theoretical variance

$$\text{var}(\hat{\theta}) = \frac{1}{\sum_{i=1}^{n} w_i}$$

and the empirical estimate of the variance

$$\hat{\text{var}}(\hat{\theta}) = \frac{\sum_{i=1}^{n} w_i (\hat{\theta}_i - \hat{\theta})^2}{\left(\sum_{i=1}^{n} w_i\right) (n-1)}.$$

This ultimately leads to the following equation which can be solved for $\sigma^2$: 
3. Example:

Estimates of clutch size and standard error are obtained from a population of nesting birds over ten-year period as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean clutch size ($\hat{y}_i$)</th>
<th>$\hat{se}($ $\hat{y}_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.04</td>
<td>0.39</td>
</tr>
<tr>
<td>2</td>
<td>5.64</td>
<td>0.29</td>
</tr>
<tr>
<td>3</td>
<td>5.82</td>
<td>0.42</td>
</tr>
<tr>
<td>4</td>
<td>4.99</td>
<td>0.43</td>
</tr>
<tr>
<td>5</td>
<td>5.93</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>5.56</td>
<td>0.15</td>
</tr>
<tr>
<td>7</td>
<td>5.84</td>
<td>0.08</td>
</tr>
<tr>
<td>8</td>
<td>5.40</td>
<td>0.11</td>
</tr>
<tr>
<td>9</td>
<td>5.86</td>
<td>0.45</td>
</tr>
<tr>
<td>10</td>
<td>5.67</td>
<td>0.26</td>
</tr>
</tbody>
</table>

$\hat{y}$ = 5.68

The $\hat{se}($ $\hat{y}_i$) in this example include only sampling variation. Also, in this example $\hat{se}(\hat{y})$ is calculated using the empirical formula:

$$\hat{se}(\hat{y}) = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - \hat{y})^2}{n(n-1)}} = 0.096,$$

which includes both temporal variation and sampling variation. We can then iteratively solve for $\sigma^2$ to estimate the temporal variance in clutch size after controlling for sampling variance using the formulae:

$$w_i = \frac{1}{\sigma^2 + \text{var}(\hat{y}_i)}$$

and

$$\frac{1}{n-1} \sum_{i=1}^{n} w_i (\hat{y}_i - \hat{y})^2 = 1.$$
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.93</td>
<td>0.45</td>
<td>0.064</td>
<td>4.735</td>
<td>0.301</td>
</tr>
<tr>
<td>6</td>
<td>5.56</td>
<td>0.15</td>
<td>0.013</td>
<td>30.113</td>
<td>0.404</td>
</tr>
<tr>
<td>7</td>
<td>5.84</td>
<td>0.08</td>
<td>0.026</td>
<td>54.886</td>
<td>1.448</td>
</tr>
<tr>
<td>8</td>
<td>5.40</td>
<td>0.11</td>
<td>0.075</td>
<td>44.051</td>
<td>3.301</td>
</tr>
<tr>
<td>9</td>
<td>5.86</td>
<td>0.45</td>
<td>0.036</td>
<td>4.673</td>
<td>0.166</td>
</tr>
<tr>
<td>10</td>
<td>5.67</td>
<td>0.26</td>
<td>0.000</td>
<td>12.740</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\[ \hat{y} = 5.68 \quad \text{Sum 178.478} \quad 9.000 \quad 1007.598 \]