N-mixture models for estimating abundance from repeated counts

Nathan Rudd
Bureau of Reclamation, MSCP
Objectives

• Definition: N-mixture models
• Development & Assumptions
• Example: Marsh Birds
• Advantages & Caveats
• References & Resources
Context: Imperfect Detection, Hits = 1
50 groups, N ~ 80-95
Context: Imperfect Detection, Hits = 0, 50 groups, N ~ 80-95
Context: Imperfect Detection, Hits = 2, 50 groups, N ~ 80-95
Context: Imperfect Detection, Hits = 6, 200 groups, N ~ 365-380
Context: Imperfect Detection, Hits = 7, 200 groups, N ~ 365-380
Context: Imperfect Detection, Hits = 9, 200 groups, N ~ 365-380
N-mixture models

Combines probability models for:

- True *abundance* (unknown)
  e.g., Poisson, Negative Binomial
  \( N_i | \lambda_i \sim \text{Poi}(\lambda_i) \)

- Observation > Detection (\( p \))
  \( c_{ij} | N_i \sim \text{Bin}(N_i, p_{ij}) \)
  Binomial or Multinomial
Estimating Abundance, Unmarked Animals

Presence/absence data, closed population

Royle (2004)
Closed population

Chandler et al. (2011)
Closed, Temporary emigration

Dail & Madsen (2011)
Open population
Recruitment, Survivorship

Others ref’d in Dénes et al. 2015
Assumptions, Royle 2004

- Closed Population
- No double counts
- Independent detections
- Equal $p$ for all individuals within sample
- Parametric model (binomial, Poisson)
Performance? Simulation Studies

- Varying sample size ($R$), # of visits ($T$) and detection probability ($p$).
- Royle (2004) unbiased even at low $R$ (20) except for low $p$ (e.g., small + bias, $R(p) = 20$ (0.25))
- Larger $T$ and $p$ counter small $R$
- Most other models tested with $R \geq 100$, $T \geq 3$, $p \geq 0.25$
- Dail & Madsen Open model, $R = 20, 100$. Less biased than Royle binomial (closed) model at low $R$.
- General problem with Negative Binomial, esp. at low $p$. 
Performance? Field Data

- Higher abundance relative to territory maps of birds, distance sampling estimates of desert tortoises (Kéry et al. 2005, Zylstra et al. 2010)
- But not compared to direct observations of lizards (Doré et al. 2011).
- Temporary emigration less biased than Royle closed and multinomial (removal sampling) models when compared with spot mapping data on Chestnut sided Warbler.
- Capture-recapture estimates ~ 2x higher than distance and N-mixture models. In simulations, N-mixture biased high when p < 0.5 (Courturier et al. 2013).
Model Fitting

Maximum Likelihood

Presence (limited), R package unmarked
Requires $K$, upper integration boundary (> max count/point) – problem with Neg. Binomial

Bayesian

WinBugs, JAGS
Priors, Markov Chain Monte Carlo
Kéry et al. (2009) – extends basic Royle model for varying $N$ over years.
unmarked

- Range of model types
  Occupancy, N-mixture, Distance
- Utilities for correct file format
- Goodness of fit (bootstrap w/ parboot) but also Nmix.gof.test {AICcmodavg}
- Open models and GOF **VERY SLOW**
Example code
Royle (2004) model

library(unmarked)
library(AICcmodavg)
ccla2013 <- read.csv("clra2013_np.csv")
unmf13 <- unmarkedFramePCount(y=ccla2013)
(fm13 <- pcount(~1 ~1, umf13, mixture="P"))
(bts <- backTransform(fm13, type="state"))
confint(bts)
(btd <- backTransform(fm13, type="det"))
confint(btd)
Interpreting Results

\[ \lambda = \text{average abundance at site} \]

Area expansion to get population total or simple multiplication by # of sampling sites, \( R\lambda \)

Difficult to interpret without fixed radius plot.
## Clapper Rail Example

**N = 52, 3 surveys per year**

<table>
<thead>
<tr>
<th>Model &amp; Year</th>
<th>AIC</th>
<th>Est. N</th>
<th>Est. Total</th>
<th>95% LCL</th>
<th>95% UCL</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Royle &amp; Nichols, P/A, 2013</td>
<td>184.5</td>
<td>1.4</td>
<td>73.8</td>
<td>45.0</td>
<td>120.9</td>
<td>0.42</td>
</tr>
<tr>
<td>2014</td>
<td>187.8</td>
<td>1.3</td>
<td>66.4</td>
<td>39.2</td>
<td>112.6</td>
<td>0.39</td>
</tr>
<tr>
<td>Royle 2004, Poisson, 2013 data</td>
<td>434.9</td>
<td>3.0</td>
<td>155.2</td>
<td>110.8</td>
<td>217.3</td>
<td>0.35</td>
</tr>
<tr>
<td>Zero-inflated Poisson</td>
<td>409.6</td>
<td>6.1</td>
<td>316.2</td>
<td>181.8</td>
<td>549.5</td>
<td>0.25</td>
</tr>
<tr>
<td>Negative Binomial</td>
<td>380.8*</td>
<td>32.5*</td>
<td>1518.4</td>
<td>574.7</td>
<td>3296.1</td>
<td>0.04</td>
</tr>
<tr>
<td>2014, Poisson</td>
<td>530.0</td>
<td>4.3</td>
<td>223.9</td>
<td>153.7</td>
<td>326.2</td>
<td>0.27</td>
</tr>
<tr>
<td>Zero-inflated Poisson</td>
<td>465.4</td>
<td>14.4</td>
<td>751.0</td>
<td>310.5</td>
<td>1816.5</td>
<td>0.13</td>
</tr>
<tr>
<td>Chandler Poisson, 2013-14</td>
<td>866.7</td>
<td>3.5</td>
<td>181.5</td>
<td>146.0</td>
<td>225.7</td>
<td>0.31</td>
</tr>
<tr>
<td>Dail &amp; Madsen Open Pop, Poisson, 2013-14</td>
<td>868.5</td>
<td>3.4</td>
<td>(182.6)</td>
<td>136.4</td>
<td>227.7</td>
<td>0.31</td>
</tr>
<tr>
<td>Dail &amp; Madsen, ZIP</td>
<td>823.0</td>
<td>6.4</td>
<td>(337.5)</td>
<td>244.1</td>
<td>450.8</td>
<td>0.23</td>
</tr>
</tbody>
</table>
Clapper Rail Example
N = 52, 3 surveys per year, 2006-2014
Advantages

- Simple models
- Relatively inexpensive data
- Missing values OK
- Covariates for both $N$ and $\rho$
- For a few species, Open Population (Dail & Madsen) model fits BBS data OK, ‘reasonable’ results based on other information.
Caveats

• New models, simulations of limited scenarios
• Need working knowledge of R and unmarked for getting CI’s
• Limited field testing
• Problems with Negative Binomial (sparse data)
• Poor fit on BBS data for many species. Extend to use distance sampling data?
Resources


R package *AHMbook* – complements *unmarked*

Google Groups: “unmarked”

Workshops – Training postings at Google group or phidot.org
Questions, Comments?

nrudd@usbr.gov