Intro
○

Modelling with CJS
○○○○

CJS Extensions
○○○○○○

Multiple random effects
○○○○○○

Conclusions
○○

Random Effects Modelling with Capture-Recapture Data

J. Marriner[1]   M. Howell[2]

[1]University of Glasgow

[2]STOR-i Centre for Doctoral Training

30/08/2024

Motivation

Why use capture-recapture?

Figure: Ringed Blackbird [7]

The Cormack-Jolly-Seber Model

We will be using a standard model known as the Cormack-Jolly-Seber (CJS) model. For this we assume;

1. The population is only open to animals leaving,

2. All emigration during the study is permanent,

3. Sampling is instantaneous,

4. No marks are lost during the study.

Intro
O

Modelling with CJS
○●○○

CJS Extensions
○○○○○○

Multiple random effects
○○○○○○

Conclusions
○○

Cormack-Jolly-Seber Structure

The CJS model conditions on the first capture of an animal and includes parameters;

1. p_i, the probability of **recapture** in period i, $i = 2, \ldots, K$
2. ϕ_i, the probability of **apparent/local** survival between occasions i and $i + 1$, $i = 1, \ldots, K - 1$

Occasion	1	2	3	4	5
Capture History	1	0	1	1	0

Capture Histories to Probabilities

What is the probability of capture history $\omega^* = \{1, 0, 1, 1, 0\}$?

$$1 \xrightarrow{\phi_1} 0 \xrightarrow{\phi_2} 1 \xrightarrow{\phi_3} 1 \longrightarrow 0$$

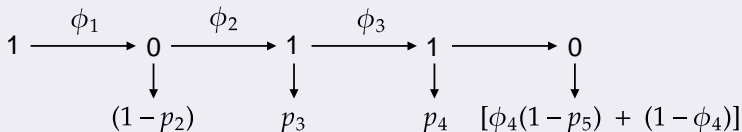$$(1 - p_2) \qquad p_3 \qquad p_4 \quad [\phi_4(1 - p_5) + (1 - \phi_4)]$$

Figure: Order of Parameter Occurence

Hence,

$$\mathbb{P}(\omega^*) = \phi_1(1 - p_2) \cdot \phi_2 p_3 \cdot \phi_3 p_4 \cdot (1 - \phi_4 p_5)$$

The final survival and capture probabilities are confounded.

Likelihood Formation in MARK/RMark

The programme MARK formulates the likelihood as;

$$\mathcal{L}(\phi, \mathbf{p} \mid \text{Capture Histories}) = \prod_{\omega \in \Omega} [\mathbb{P}(\omega)]^{n_\omega}$$

where

- Ω is the set of possible capture histories,
- n_ω is the number of animals with capture history ω,
- $\mathbb{P}(\omega)$ is the probability of capture history ω.

This is solved numerically by Newton-Rhapsom.

Adding Covariates

We can model parameters as a function of environmental or individual covariates;

$$\log\left(\frac{\phi_i}{1-\phi_i}\right) = \mathbf{x_i^T}\boldsymbol{\beta}$$

This can be substituted into the likelihood for numerical optimisation.

Random Effects

Suppose variations over time are due only to natural variance. i.e.

$$\log\left(\frac{\phi_i}{1 - \phi_i}\right) = \beta_0 + \epsilon_i$$

where:

- β_0 is an unknown mean,
- $\epsilon_i \sim N(0, \sigma^2)$ with σ^2 known as the process or environmental variance.

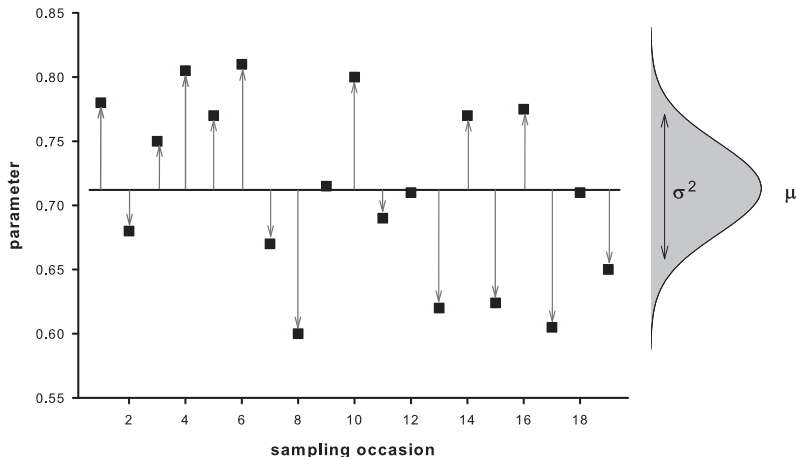This random effect motivates a switch to Bayesian inference.

Intro
○

Modelling with CJS
○○○○

CJS Extensions
○○●○○○

Multiple random effects
○○○○○○

Conclusions
○○

Random Effects Illustration

Figure: Process Variance of a Parameter from [1][p893]

Intro
o

Modelling with CJS
oooo

CJS Extensions
ooo●oo

Multiple random effects
oooooo

Conclusions
oo

Random Effects Simulation

We simulated capture-recapture data with a population size $N = 300$, capture occasions $K = 10$, a constant recapture rate $p = 0.5$ and survival rates such that;

$$\log\left(\frac{\phi_i}{1 - \phi_i}\right) = 1.4 + \epsilon_i$$

where;

▶ $\epsilon_i \sim N(0, 0.5)$

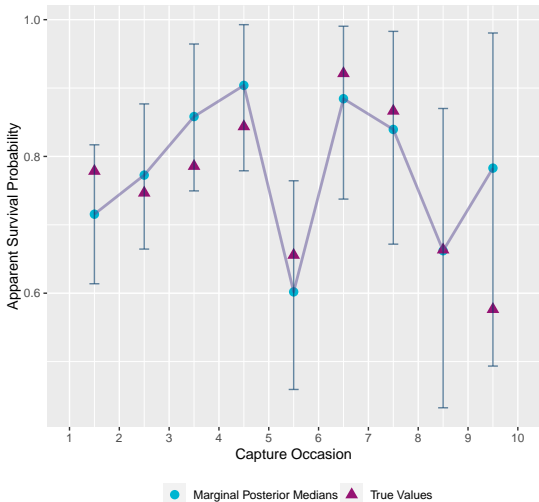We used MCMC to make posterior draws having used the uninformative priors from [5];

1. $\beta_0 \sim \text{Un}(-5, 5)$
2. $\sigma \sim \text{Un}(0, 3)$
3. $\text{logit}(p) \sim N(0, 10^3)$

MCMC Results

After 5000 burn in, a thinning rate of 5 applied to 100,000 posterior draws using WinBUGS, we have posterior estimates;

▶ $\hat{\beta}_0 = 1.402$ $(0.68, 2.6)$,

▶ $\hat{\sigma}^2 = 0.867$ $(0.068, 5.94)$,

▶ $\hat{p} = 0.467$ $(0.42, 0.51)$

Intro
○

Modelling with CJS
○○○○

CJS Extensions
○○○○○○●

Multiple random effects
○○○○○○

Conclusions
○○

Apparent Survival Results

Intro
○

Modelling with CJS
○○○○

CJS Extensions
○○○○○○

Multiple random effects
●○○○○○

Conclusions
○○

Multiple Random Effects

Multiple random effects used by [3] for multiple colonies of Puffins in the North Atlantic. Extended by [5] to describe **multiple species** of birds on the Isle of May.

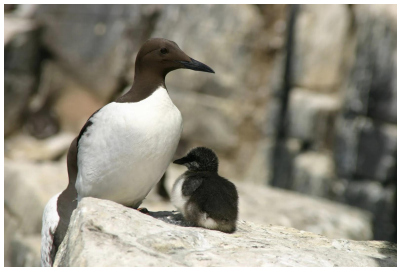Figure: Puffins on the Isle of May

Isle of May Birds

Figure: Two Guillemots [7]

Figure: Two Razorbills [7]

Multi-Species Model

We set up likelihoods for each species using the CJS model and proceed to fit the model;

$$\log\left(\frac{\phi_{i,s}}{1 - \phi_{i,s}}\right) = f_s(\mathbf{x_{i,s}^T}) + \delta_i + \epsilon_{i,s}$$

where

- $\delta_i \sim N(0, \sigma_\delta^2)$, constant across all species,
- $\epsilon_{i,s} \sim N(0, \sigma_s^2)$ exclusive to each species.
- $f_s(\mathbf{x_{i,s}^T})$ is a unique covariate link function for each species.

Measuring Synchrony

For each species, we define the **Intra-class correlation coefficient**;

$$ICC_s = \frac{\hat{\sigma}_\delta^2}{\hat{\sigma}_\delta^2 + \hat{\sigma}_s^2}$$

▶ Quantifies the environmental variance shared between species,
▶ High synchrony $\Rightarrow ICC_s \to 1$,
▶ Low synchrony $\Rightarrow ICC_s \to 0$

What generates Synchrony?

We can fit a model for each species with and without covariates. This gives;

- $\hat{\sigma}_s^2(\text{res})$ the residual species variance from the covariate model
- $\hat{\sigma}_s^2(\text{total})$ the species variance from the intercept only model

What generates Synchrony?

We can fit a model for each species with and without covariates. This
gives;

- $\hat{\sigma}_s^2(\text{res})$ the residual species variance from the covariate model
- $\hat{\sigma}_s^2(\text{total})$ the species variance from the intercept only model

For each species we define;

$$C_s = 1 - \frac{\hat{\sigma}_s^2(\text{res})}{\hat{\sigma}_s^2(\text{total})}$$

- Covariates cause synchrony $\Rightarrow C_s \to 1$
- Covariates do not synchrony $\Rightarrow C_s \to 0$

Isle of May results

The results of the study on Isle of May birds were;

- Puffins: $ICC_1 = 0.894$ $(0.304, 0.999)$, and $C_{\text{puffins}} = 0.81$
- Guillemots: $ICC_2 = 0.787$ $(0.350, 0.996)$, and $C_{\text{guillemots}} = 0.425$
- Razorbills: $ICC_3 = 0.785$ $(0.205, 0.998)$, and $C_{\text{razorbills}} = 0.595$

This indicates common random effects can describe a large amount of variation between species.

Benefits and Limitations

Benefits:

▶ Detects links between populations,

▶ Describes expected variation and environmental effects,

▶ Combats parameter redundancy,

▶ Reduces parameter number

Limitations:

▶ Requires data for multiple populations,

▶ Does not describe the synchrony source fully,

▶ Does not describe pairwise effects.

Further Extensions

▶ Up to 4 random effects included in [2],

▶ Used for birth rates in an Integrated Population Model by [4],

There are further opportunities for;

▶ Different parameters and model types,

▶ Inclusion of predators or prey,

▶ Assessing synchrony over time

Thank You!

References I

E. G. Cooch and G. C. White.
Program MARK, A Gentle Introduction.
13th edition, 2014.

M. Ghislain, T. Bonnet, U. Godeau, O. Dehorter, O. Gimenez, and P.-Y. Henry.
Synchrony in adult survival is remarkably strong among common temperate songbirds across france.
Ecology, 105(6):e4305, 2024.

V. Grosbois, M. P. Harris, T. Anker-Nilssen, R. H. McCleery, D. N. Shaw, B. J. T. Morgan, and O. Gimenez.
Modeling survival at multi-population scales using mark–recapture data.
Ecology, 90(10):2922–2932, 2009.

References II

J. Lahoz-Monfort, M. Harris, S. Wanless, S. Freeman, and B. Morgan.
Bringing it all together: Multi-species integrated population modelling of a breeding community.
Journal of Agricultural, Biological and Environmental Statistics, 22, 04 2017.

J. J. Lahoz-Monfort, B. J. T. Morgan, M. P. Harris, S. Wanless, and S. N. Freeman.
A capture–recapture model for exploring multi-species synchrony in survival.
Methods in Ecology and Evolution, 2(1):116–124, 2011.

R. S. McCrea and B. J. T. Morgan.
Analysis of capture-recapture data.
Chapman & Hall/CRC Interdisciplinary Statistics Series. CRC Press, Boca Raton, FL, 2015.

References III

📄 I. of May Bird Observatory.
Isle of may bird observatory blog.
https://isleofmaybirdobs.org/.

M-arrays

We can describe our CR-data using a series of sufficient statistics within an m-array. For example, given 5 sampling occasions;

Number Released	Number Recaptured				Never Recaptured
R_1	$m_{1,2}$	$m_{1,3}$	$m_{1,4}$	$m_{1,5}$	$m_{1,0}$
R_2		$m_{2,3}$	$m_{2,4}$	$m_{2,5}$	$m_{2,0}$
R_3			$m_{3,4}$	$m_{3,5}$	$m_{3,0}$
R_4				$m_{4,5}$	$m_{4,0}$

where;

- R_i is the number captured in occasion i and subsequently released.
- $m_{i,j}$ is the number released from occasion i and first recaptured in period j

Here, $m_{i,0} = R_i - \sum_{j=i+1}^{K} m_{i,j}$

M-array probabilities

For notation we use;

- ϕ_i - The probability of survival from period i to $i+1$
- p_i - The probability of recapture in period i given the unit is alive

The probability of each $m_{i,j}$ cell in the m-array is then defined as;

$$\nu_{ij} = \left\{ \prod_{k=i}^{j-1} \phi_k \prod_{k=i+1}^{j-1} (1 - p_k) \right\} p_j \quad \text{for } i < j$$

Similarly the probability of never being recaptured after first capture in i is;

$$\chi_i = 1 - \sum_{j=i+1}^{T} \nu_{ij}$$

Multinomial distributions in M-arrays

To form a likelihood, we consider each row in the M-array as a multinomial distribution. For example;

$$\mathcal{L}_1(\phi, \mathbf{p} \mid R_1, \mathbf{m}_{1.}) = \binom{R_1}{m_{1,j}} \prod_{j=2}^{5} \nu_{1j}^{m_{1,j}} \cdot \chi_1^{R_1 - \sum_{k=2}^{5} m_{1,k}}$$

where we have the multinomial coefficient;

$$\binom{R_1}{m_{1,j}} = \frac{R_1!}{(m_{1,2})!(m_{1,3})!(m_{1,4})!(m_{1,5})!(R_1 - \sum_{j=2}^{5} m_{1,j})!}$$

Multinomial Likelihood

By combining the likelihoods of each row, we get the total likelihood;

$$\mathcal{L}(\phi, \mathbf{p} \mid \mathbf{R}_i, \mathbf{m_{ij}}) \propto \prod_{i=1}^{T-1} \prod_{j=i+1}^{T} \nu_{ij}^{m_{i,j}} \cdot \chi_i^{R_i - \sum_{j=i+1}^{T} m_{ij}}$$

This gives us the full log-likelihood from which explicit MLEs may be found;

$$\log \mathcal{L}(\phi, \mathbf{p} \mid \mathbf{R}_i, \mathbf{m_{ij}}) = \sum_{i=1}^{T} \log(\mathcal{L}_i)$$

Final Probability

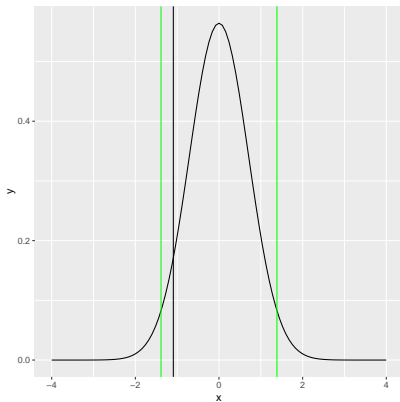▶ The least amount of data is in the final period,
▶ ϵ_9 was the 6th percentile.

Figure: Central 95% of N(0,0.5) in Green

Graph with MLEs