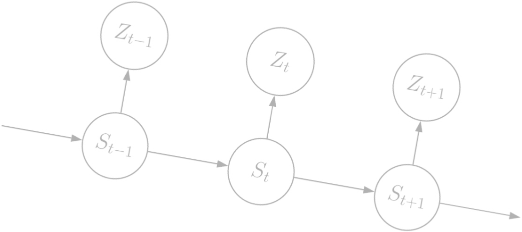
# Maximum likelihood estimation of mark-recapture-recovery models in the presence of continuous covariates

*Roland Langrock & Ruth King*

*University of St Andrews*




- Mark-recapture-recovery (MRR) data:
  - repeated surveyings of a population (e.g. yearly)
  - animals are tagged at initial encounter
  - observations: seen/not seen, survival status (alive, "recent death", "long dead"), potentially covariates
- Aim: gain understanding of the underlying system
  - → estimate survival rates
  - → detect population trends
  - → etc.
- here we deal with a statistical problem arising from **missing individual-specific time-varying covariate values** (e.g., body mass of an animal)

- Mark-recapture-recovery (MRR) data:
  - repeated surveyings of a population (e.g. yearly)
  - animals are tagged at initial encounter
  - observations: seen/not seen, survival status (alive, "recent death", "long dead"), potentially covariates
- Aim: gain understanding of the underlying system
  - → estimate survival rates
  - → detect population trends
  - → etc.
- here we deal with a statistical problem arising from

  **missing individual-specific time-varying covariate values**

  (e.g., body mass of an animal)

## *MRR data in the absence of covariates*

*Example MRR encounter history:*

$$1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 2$$

0: not seen
1: seen alive
2: recovered dead

*Associated likelihood:*

$$\mathcal{L} = \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 (1 - p_4) \phi_4 p_5 (1 - \phi_5) \lambda_6$$

$\phi_t$: prob. of surviving from $t$ to $t+1$
$p_t$: prob. of being seen when alive at time $t$
$\lambda_t$: prob. of recovery at time $t$ if death in $(t-1, t]$

## *MRR data in the absence of covariates*

*Example MRR encounter history:*

1 1 0 0 1 2

0: not seen
1: seen alive
2: recovered dead

*Associated likelihood:*

$$\mathcal{L} = \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 (1 - p_4) \phi_4 p_5 (1 - \phi_5) \lambda_6$$

$\phi_t$: prob. of surviving from $t$ to $t+1$
$p_t$: prob. of being seen when alive at time $t$
$\lambda_t$: prob. of recovery at time $t$ if death in $(t-1, t]$

## *MRR data in the absence of covariates*

Likelihood for encounter history $x_1, \ldots, x_T$ if survival process, $s_1, \ldots, s_T$, is *not fully known*:

$$\mathcal{L} = \sum_{\tau \in \mathcal{S}^c} \sum_{s_\tau \in \{1,2,3\}} \prod_{t=2}^{T} f(s_t | s_{t-1}) f(x_t | s_t) \, ,$$

where

$$s_t = \begin{cases} 1 & \text{if alive at time } t; \\ 2 & \text{if dead at time } t, \text{but was alive at time } t-1; \\ 3 & \text{if dead at time } t, \text{and was dead already at time } t-1, \end{cases}$$

and $\mathcal{S}^c = \{ t \,|\, s_t \text{ is unknown} \}$.

## *HMM formulation of MRR data in the absence of covariates*

In hidden Markov-type matrix product form:

$$\mathcal{L} = \delta \left( \prod_{t=2}^{T} \mathbf{\Gamma}_{t-1} \mathbf{Q}(x_t) \right) \mathbf{1}_3 \ ,$$

$$\text{where} \quad \mathbf{\Gamma}_t = \begin{pmatrix} \phi_t & 1-\phi_t & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{and} \quad \mathbf{Q}(x_t) = \begin{cases} \text{diag}(1-p_t, 1-\lambda_t, 0) & \text{if } x_t = 0; \\ \text{diag}(p_t, 0, 0) & \text{if } x_t = 1; \\ \text{diag}(0, \lambda_t, 0) & \text{if } x_t = 2. \end{cases}$$

## *HMM formulation of MRR data in the absence of covariates*

In hidden Markov-type matrix product form:

$$\mathcal{L} = \boldsymbol{\delta} \left( \prod_{t=2}^{T} \boldsymbol{\Gamma}_{t-1} \mathbf{Q}(x_t) \right) \mathbf{1}_3 \ ,$$

$$\text{where} \quad \boldsymbol{\Gamma}_t = \begin{pmatrix} \phi_t & 1-\phi_t & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{and} \quad \mathbf{Q}(x_t) = \begin{cases} \text{diag}(1-p_t, 1-\lambda_t, 0) & \text{if } x_t = 0; \\ \text{diag}(p_t, 0, 0) & \text{if } x_t = 1; \\ \text{diag}(0, \lambda_t, 0) & \text{if } x_t = 2. \end{cases}$$

## *MRR data in the presence of covariates*

The "interesting" case: what if survival probabilities $\phi_i$ are affected by individual-specific continuous covariates (e.g., body mass)?

*Example MRR encounter history:*

$$1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 2$$

*Associated likelihood:*

$$\mathcal{L} = \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 (1 - p_4) \phi_4 p_5 (1 - \phi_5) \lambda_6$$

## *MRR data in the presence of covariates*

The "interesting" case: what if survival probabilities $\phi_i$ are affected by individual-specific continuous covariates (e.g., body mass)?

*Example MRR encounter history:*

$$1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 2$$

Animal not seen $\rightarrow$ covariate value unknown

*Associated likelihood:*

$$\mathcal{L} = \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 (1 - p_4) \phi_4 p_5 (1 - \phi_5) \lambda_6$$

## *MRR data in the presence of covariates*

The "interesting" case: what if survival probabilities $\phi_i$ are affected by individual-specific continuous covariates (e.g., body mass)?

*Example MRR encounter history:*

$$1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 2$$

Animal not seen → covariate value unknown

*Associated likelihood:*

$$\mathcal{L} = \phi_1 p_2 \phi_2 (1 - p_3) \phi_3 (1 - p_4) \phi_4 p_5 (1 - \phi_5) \lambda_6$$

covariate unknown → likelihood can't be computed

## *Maximum likelihood?*

Idea: model covariate process and integrate over possible values

Bonner, Morgan and King (2010, Biometrics):

*"except when few values are missing, the large number of integrals [...] will make it impossible to perform maximum likelihood estimation"*

## *Most popular existing estimation methods*

- Trinomial approach
  - → conditioning on only observed covariate values
  - → closed-form likelihood, but throwing away information
- Bayesian imputation approach
  - → assume model for covariate process, for example AR(1)
  - → impute missing covariate values within MCMC
  - → model selection difficult
  - → prior distributions need to be specified

## *ML approach*

Assuming some (Markovian) model for the covariate process, e.g.,

$$y_t = y_{t-1} + \phi(\mu - y_{t-1}) + \sigma\epsilon_t \ ,$$

the likelihood for an encounter history $x_1, \ldots, x_T$ is

$$\mathcal{L} = \int \ldots \int \sum_{\tau \in \mathcal{S}^c} \sum_{s_\tau \in \{1,2,3\}} f(y_1)^{I_{\{1 \in \mathcal{W}^c\}}} \times \prod_{t=2}^{T} f(s_t | s_{t-1}, y_{t-1}) f(x_t | s_t) f(y_t | y_{t-1}) d\mathbf{y}_{\mathcal{W}^c} \ ,$$

where $\mathcal{W}^c = \{t \,|\, y_t \text{ is unobserved}, t \in \mathcal{S}, s_t \neq 2, 3\}$.

$\Rightarrow$ multiple integral, cannot be evaluated....

## *ML approach*

Assuming some (Markovian) model for the covariate process, e.g.,

$$y_t = y_{t-1} + \phi(\mu - y_{t-1}) + \sigma\epsilon_t \ ,$$

the likelihood for an encounter history $x_1, \ldots, x_T$ is

$$\mathcal{L} = \int \ldots \int \sum_{\tau \in \mathcal{S}^c} \sum_{s_\tau \in \{1,2,3\}} f(y_1)^{I_{\{1 \in \mathcal{W}^c\}}} \times \prod_{t=2}^{T} f(s_t | s_{t-1}, y_{t-1}) f(x_t | s_t) f(y_t | y_{t-1}) d\mathbf{y}_{\mathcal{W}^c} \ ,$$

where $\mathcal{W}^c = \{t \,|\, y_t \text{ is unobserved}, t \in \mathcal{S}, s_t \neq 2, 3\}$.

$\Rightarrow$ multiple integral, cannot be evaluated....

## ML approach

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$$
\begin{aligned}
\mathcal{L} \approx \sum_{\kappa \in \mathcal{W}^c} \sum_{j_\kappa = 1}^{m} \sum_{\tau \in \mathcal{S}^c} \sum_{s_\tau \in \{1,2,3\}} & f(y_1 \in B_{j_1})^{I_{\{1 \in \mathcal{W}^c\}}} \\
& \times \prod_{t=2}^{T} \Big[ f(s_t | s_{t-1}, y_{t-1})^{I_{\{(t-1) \in \mathcal{W}\}}} f(s_t | s_{t-1}, b_{j_\kappa}^*)^{I_{\{(t-1) \in \mathcal{W}^c\}}} f(x_t | s_t) \\
& \times f(y_t | y_{t-1})^{I_{\{t \in \mathcal{W}, (t-1) \in \mathcal{W}\}}} f(y_t | b_{j_{t-1}}^*)^{I_{\{t \in \mathcal{W}, (t-1) \in \mathcal{W}^c\}}} \\
& \times f(y_t \in B_{j_t} | y_{t-1})^{I_{\{t \in \mathcal{W}^c, (t-1) \in \mathcal{W}\}}} f(y_t \in B_{j_t} | b_{j_{t-1}}^*)^{I_{\{t \in \mathcal{W}^c, (t-1) \in \mathcal{W}^c\}}} \Big]
\end{aligned}
$$

## *ML approach*

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$> m^{|\mathcal{W}^c|}$ summands!!!

## *ML approach*

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$> m^{|\mathcal{W}^c|}$ summands!!! ☹

## *ML approach*

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$$\mathcal{L} \approx \boldsymbol{\delta}^{(m)} \left( \prod_{t=g+1}^{T} \boldsymbol{\Gamma}_{t-1}^{(m)} \mathbf{Q}^{(m)}(x_t) \right) \mathbf{1}_{m+2}$$

## *ML approach*

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$$\mathcal{L} \approx \boldsymbol{\delta}^{(m)} \left( \prod_{t=g+1}^{T} \boldsymbol{\Gamma}_{t-1}^{(m)} \mathbf{Q}^{(m)}(x_t) \right) \mathbf{1}_{m+2}$$

- with $\boldsymbol{\Gamma}_t^{(m)}$ and $\mathbf{Q}^{(m)}(x_t)$ suitably defined
- corresponds to an efficient HMM-type recursive scheme

(idea: augment "alive" survival state by dividing it into $m$ distinct states, associated with the different intervals the covariate value may lie in)

## ML approach

- idea: discretize covariate space, thereby reducing $\int$ to $\sum$
- split "essential range" into $m$ intervals
- $j$th interval: $B_j = [b_{j-1}, b_j)$, with midpoint $b_j^*$

$$\mathcal{L} \approx \delta^{(m)} \left( \prod_{t=g+1}^{T} \mathbf{\Gamma}_{t-1}^{(m)} \mathbf{Q}^{(m)}(x_t) \right) \mathbf{1}_{m+2} \qquad ☺$$

- with $\mathbf{\Gamma}_t^{(m)}$ and $\mathbf{Q}^{(m)}(x_t)$ suitably defined
- corresponds to an efficient HMM-type recursive scheme

(idea: augment "alive" survival state by dividing it into $m$ distinct states, associated with the different intervals the covariate value may lie in)

## *Trinomial vs. full ML approach – a simulation experiment*

Survival probability: $\text{logit}(\phi_t) = \beta_0 + \beta_1 y_t$

Table: *Sample mean estimates (ME), sample mean widths (CW) of the estimated 95% confidence intervals and coverage probabilities (CC) of the confidence intervals, in four simulation scenarios.*

| | | inte. ($\beta_0 = -3$) | | | slope ($\beta_1 = 0.2$) | | |
|---|---|---|---|---|---|---|---|
| Sc | Meth. | ME | CW | CC | ME | CW | CC |
| $p = 0.95$ | Tri | -3.00 | 1.39 | 0.96 | 0.20 | 0.08 | 0.94 |
| $\lambda = 0.95$ | full ML | -3.00 | 1.33 | 0.94 | 0.20 | 0.07 | 0.93 |
| $p = 0.9$ | Tri | -3.01 | 1.69 | 0.95 | 0.20 | 0.12 | 0.95 |
| $\lambda = 0.3$ | full ML | -3.01 | 1.37 | 0.95 | 0.20 | 0.08 | 0.95 |
| $p = 0.3$ | Tri | -3.09 | 3.08 | 0.97 | 0.20 | 0.14 | 0.97 |
| $\lambda = 0.9$ | full ML | -3.00 | 1.46 | 0.94 | 0.20 | 0.08 | 0.95 |
| $p = 0.3$ | Tri | -2.98 | 3.73 | 0.98 | 0.20 | 0.20 | 0.95 |
| $\lambda = 0.3$ | full ML | -2.99 | 1.92 | 0.95 | 0.20 | 0.11 | 0.95 |

## *Application to Soay sheep MRR data*


- capture histories for 1344 female Soay sheep, 1985–2009
- four different age groups: lambs ($< 1$), yearlings (1–2), adults (2–7) and seniors ($> 7$)
- Survival probability:

$$\text{logit}(\phi_t) = \beta_{a_t,0} + \beta_{a_t,1}\text{weight}_t$$

- covariate process model:

$$\text{weight}_t = \text{weight}_{t-1} + \phi_{a_t}\left(\mu_{a_t} - \text{weight}_{t-1}\right) + \sigma_{a_t}\epsilon_t$$

- year-dependent recapture and recovery probabilities

# Application to Soay sheep MRR data – estimated survival probability

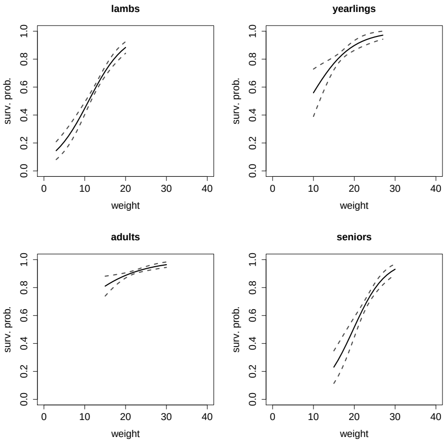



Figure: *Solid lines: ML estimates. Dashed lines: 95% pointwise confidence intervals, obtained using the delta method.*

# Application to Soay sheep MRR data – fitted covariate process

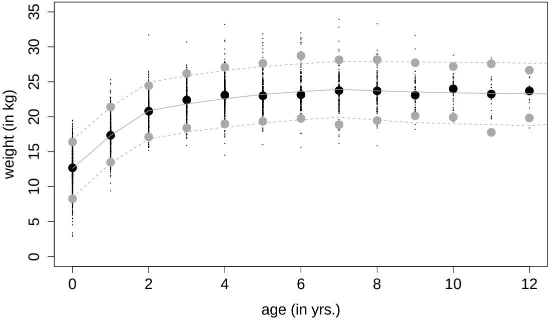



Figure: *Empirical 5% and 95% quantiles (big grey dots) and empirical medians (big black dots) of body masses, and model-derived 5% and 95% quantiles (dashed grey lines) and medians (solid black lines) of body mass distributions at those ages.*

## *Remarks, future work & references*

- HMM-based discretization strategy applicable to essentially any state-space model
- multiple covariates computationally challenging → apply more sophisticated numerical integration strategies
- other state processes can be considered: e.g., additional states (Arnason-Schwarz model), semi-Markov components


Bonner, R., Morgan, B.J.T., King, R., 2010. Continuous covariates in mark-recapture-recovery analysis: a comparison of methods. *Biometrics*, 66, 1256–1265.

Langrock, R., King, R., 2013. Maximum likelihood estimation of mark-recapture-recovery models in the presence of continuous covariates. *Annals of Applied Statistics*, in press.

Langrock, R., 2011. Some applications of nonlinear and non-Gaussian state-space modelling by means of hidden Markov models. *Journal of Applied Statistics*, 38, 2955–2970.