

Pollock's Closed Robust Design

Dr RW Rankin and KE Nicholson

Georgetown University, Murdoch University

October 29, 2017

Capture-Mark-Recapture

- ▶ secondary periods
- ▶ effective capture probabilities
- ▶ temporary emigration

Bayesian Modelling

- ▶ hierarchical Bayes
- ▶ hyper-priors
- ▶ shrinkage

Pollock

- ▶ do repeated *instantaneous* sampling -> secondary periods
- ▶ S_t secondary capture periods per primary period
- ▶ if we assume closure this increases the effective capture probability

$$p_t^* = 1 - \prod_{s=1}^{S_t} (1 - p_{s,t})$$

- ▶ solves some of parameter confounding

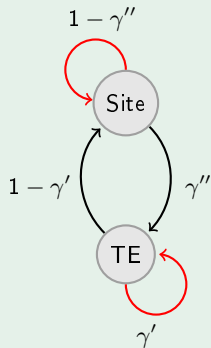
[Pollock, 1982]

0001001: are the zeros due to animal not being there? or missed-capture?

Kendall Model

repeated sampling also allows us to separate **temporary emigration** from missed-captures. [Kendall et al., 1995, Kendall and Nichols, 1995, Kendall et al., 1997]

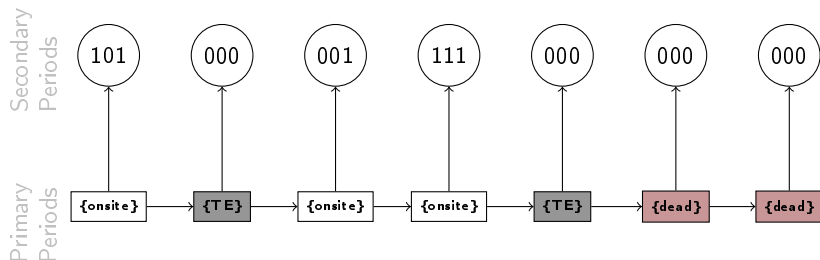
- ▶ γ'' probability of leaving the study area and becoming a temp. emigrant
- ▶ $1 - \gamma''$ probability of staying on-site
- ▶ γ' probability of staying as a temp. emigrant, conditional on already being a temp. emigrant
- ▶ $1 - \gamma'$ probability of return to the study-site



unobservable states

- ▶ TE is an **unobservable state**
- ▶ probability of capture is $0 | z = \{\text{TE}\}$
- ▶ (we've already encountered other unobservable states, like *unrecruited* and *dead*)

- ▶ example capture history ($y = \{10100000111100000000\}$) and plausible latent state sequence



HMM Matrices: Connection with CMR (PCRD)

HMM matrices for **Pollock's Closed Robust Design** (*full capture* model):

Transition Matrix

from $t - 1$ to t

$$\Phi_t = \begin{array}{c} \text{Unborn} \\ \text{Onsite} \\ \text{TE} \\ \text{Dead} \end{array} \begin{array}{c} \text{Unborn} \\ \text{Onsite} \\ \text{TE} \\ \text{Dead} \end{array} \begin{pmatrix} 1 - \psi_t & 0 & 0 & 0 \\ \psi_t & \phi_{t-1}(1 - \gamma_t'') & \phi_{t-1}(1 - \gamma_t') & 0 \\ 0 & \phi_{t-1}\gamma_t'' & \phi_{t-1}\gamma_t' & 0 \\ 0 & 1 - \phi_{t-1} & 1 - \phi_{t-1} & 1 \end{pmatrix}$$

- ▶ ϕ_{t-1} is the apparent survival between primary periods $t - 1$ and t
- ▶ γ_t'' is the probability of becoming a temporary emigrant between $t - 1$ and t
- ▶ γ_t' is the probability of staying as a temporary emigrant between $t - 1$ and t
- ▶ ψ_t are "recruitment" parameters between $t - 1$ and t

Emission Matrix

$$\Psi_{s,t} = \begin{array}{c} \text{Capture} \\ \text{No Capture} \end{array} \begin{array}{c} \text{Unborn} \\ \text{Onsite} \\ \text{TE} \\ \text{Dead} \end{array} \begin{pmatrix} 0 & p_{s,t} & 0 & 0 \\ 1 & 1 - p_{s,t} & 1 & 1 \end{pmatrix}$$

The capture histories are now 3D: [individual, secondary, primary]
(order is arbitrary)

```
Y.tt[2,,]  
  [,1] [,2] [,3] [,4] [,5]  
[1,]    0    0    0    0    0  
[2,]    0    0    0    0    0  
[3,]    0    0    0    0    0  
[4,]    0    0    0    1    NA  
[5,]    0    0    0    0    NA  
[6,]   NA   NA    0   NA   NA  
[7,]   NA   NA    0   NA   NA  
[8,]   NA   NA    0   NA   NA  
[9,]   NA   NA    0   NA   NA  
[10,]  NA   NA    0   NA   NA  
Y.tt[3,,]  
  [,1] [,2] [,3] [,4] [,5]  
[1,]    0    0    0    0    0  
[2,]    0    0    0    0    0  
[3,]    0    0    0    0    0  
[4,]    0    0    0    1    NA  
[5,]    0    0    0    0    NA  
[6,]   NA   NA    0   NA   NA  
[7,]   NA   NA    0   NA   NA  
[8,]   NA   NA    0   NA   NA  
[9,]   NA   NA    0   NA   NA  
[10,]  NA   NA    0   NA   NA
```


- ▶ the emission matrix is now **per secondary period**

i.e., an extra for loop in jags

```
for(t in 1:T){ # loop through primary periods
  for(s in 1:T2[t]){ # loop secondary periods
    # unrecruited: state 1
    em[1,1,t,s] <- 1 # no capture
    em[2,1,t,s] <- 0 # capture illegal
    # onsite: state 2
    em[2,2,t,s] <- 1/(1+exp(-1*(lp.mu+lp.t[t]+lp.tt[s,t]))) # capture
    em[1,2,t,s] <- 1-em[2,2,t,s] # no capture
    # TE: state 3
    em[1,3,t,s] <- 1 # 100% no capture
    em[2,3,t,s] <- 0 #
    # dead: state 4
    em[1,4,t,s] <- 1 # no capture
    em[2,4,t,s] <- 0 #
  } # s
} # t
```

- ▶ the latent state process is still per primary period
- ▶ the conditional likelihood is now **per secondary period**

i.e., an extra for loop in jags

```
# HMM PROCESS for t>1
for(t in 2:T){
  # LATENT STATE
  z[i,t] ~ dcat(tr[1:4, z[i,t-1], t-1])
  # EMISSION
  for(s in 1:T2[t]){ # loop through secondary periods
    y[i,s,t] ~ dcat(em[,z[i,t],t,s])
  }
} #t
} # i
```

All the PCRD Parameters are probabilities: can use Beta or logit-Normal or probit-Normal

- ▶ γ' : difficult to separate from ϕ (especially at low T)

avoid References priors or Jeffrey's Priors

The complexity of PCRD and MSCRD mean that it starts to make sense to use Hierarchical Bayesian models

Why HB?

- ▶ conceptual: build dependencies and **share information** among similar parameters [Cressie et al., 2009, Halstead et al., 2012]
- ▶ estimation: **shrinkage** $\phi(t) \rightarrow \phi(\cdot)$ [Royle and Link, 2002, Burnham and White, 2002, Rankin, 2016]
- ▶ type of **multi-model inference**: smooth over several “fixed-effects” specifications [Rankin et al., 2016]

Why not?

- ▶ shrinkage sometimes arbitrary, difficult to place prior information about “hyper-priors”

3 stages of model

$$\sigma_p \sim \mathcal{T}(0, s_0^2, \nu) \mathbb{I}[\sigma > 0]$$

$$\delta_{s,t} \sim \mathcal{N}(0, \sigma_p^2) \quad \forall s \in S_t, t \in T$$

$$y_{s,t} \sim \text{Bern} \left(\frac{1}{1 + \exp(-\delta_\mu - \delta_{s,t})} \right)$$

$\delta_{1,1}, \delta_{2,1}, \delta_{3,1}, \dots, \delta_{S_T, T}$ are random effects that come from a **distribution** the spread and dispersion among the δ effects: controlled by **hyperprior** σ_p .

- ▶ if $\sigma_p \rightarrow 0$, then $\delta_{s,t} \rightarrow 0$

$$\text{logit}^{-1}(-\delta_\mu - \delta_{s,t}) \rightarrow \text{logit}^{-1}(-\delta_\mu)$$

... i.e., the time-constant capture model $p(\cdot)$

- ▶ if $\sigma_p \gg 0$, then

$$\text{logit}^{-1}(-\delta_\mu - \delta_{s,t}) \rightarrow \hat{p}_{s,t}$$

... i.e., the fully-time-varying capture model $p(s, t)$

Intermediate values of $\sigma_p \gg 0$ results in a model something between time-constant p and time-varying $p_{s,t}$

- ▶ i.e., $p(s, t)$ is **shrunk** towards $p(\cdot)$

Bias-Variance Trade-off

- ▶ at low amount of data, the Hyperprior is relatively stronger and the shrinkage $p(s, t) \rightarrow p(\cdot)$ is more pronounced.
- ▶ at high amount of data, the data can drive the values of $p(s, t)$, and shrinkage is less pronounced.

less data = simpler models
more data = more complex models

How to control the value of the hyperprior σ_p ?

- 1 depends on the information in the data, and
- 2 the hyper-parameters of the half-student-t distribution s_0^2 and ν_0 .

small s_0 equals **more shrinkage**
big ν equals **more shrinkage** (long tails)

- 1 Shrinkage is the *bias* induced by random-effects: the distribution pulls the individual effects towards the distributions' shape and central tendency.
- 2 Shrinkage is desirable in low-information / small-sample size situations

Visualiation

- ▶ See a visualization of shrinkage here:

<http://colugos.blogspot.ca/2016/02/hierarchical-bayesian-automatic-occams.html>

- ▶ or navigate to

PART7_PCRD/HalfTdemo.gif
... and open with an internet-browser

PCRD Demo

We will build random effects for

- 1 mean (logit) capture probability **per primary period** $lp.t$ with hyper prior $\sigma.p.t$
- 2 mean (logit) capture probability **per secondary period** $lp.tt$ with hyper prior $\sigma.p.tt$

... using the half-student-t distribution (truncation at zero)

```
# JAGS hyperpriors
sigma.p.t ~ dt(0,pr.sigma.p.t[1],pr.sigma.p.t[2]) T(0,)
sigma.p.tt ~ dt(0,pr.sigma.p.tt[1],pr.sigma.p.tt[2]) T(0,)
```

Next, sample the $lp.t$ and $lp.tt$ and $lp.mu$

```
# hierarchical capture process
lp.mu ~ dnorm(pr.lp.mu[1],pr.lp.mu[2]) # mean capture (logit)
for(t in 1:T){ # loop through primary periods
  lp.t[t] ~ dnorm(0, pow(sigma.p.t,-2))
  for(s in 1:T2[t]){ # loop through secondary periods
    lp.tt[s,t] ~ dnorm(0, pow(sigma.p.tt,-2))
  }
}
```

Next, add them all together and back-transform to a probability

```
em[2,2,t,s] <- 1/(1+exp(-1*(lp.mu + lp.t[t] + lp.tt[s,t]))) # capture  
... or ...
```

```
em[2,2,t,s] <- ilogit(lp.mu+lp.t[t]+lp.tt[s,t]) # capture
```

Rather than each $p_{s,t}$ capture probability be *independent*, there is information sharing across primary and secondary periods

- ▶ outliers are pulled towards $\text{logit}^{-1}(\text{lp.mu})$

open the file `R_pcrd.R`

- ▶ **DEMONSTRATION** Individual-heterogeneity PCRD using random-effects
- ▶ **EXERCISE 1** play with hyperpriors-hyperparameters and learn their affect on inference for an HB-PCRD (hierarchical capture probability). Modify:

```
pr.sigma.p.t<-c(0.05^(-2), 3)  
pr.sigma.p.tt<-c(0.05^(-2), 3)
```

- 1 **EXERCISE 2** make the other parameters (γ'' , γ' , ϕ) into a HB model with random-effects for the temporal variation!

Make your own random effects

- ▶ make hyperpriors like `sigma.gamma1.t`
- ▶ make a global intercept: `lgamma1.mu`
- ▶ then make random effects `lgamma1[t] ~ dnorm(0,pow(sigma.gamma1.t,-2))`
- ▶ then back-transform to a probability:
`gamma1[t] <- ilogit(lgamma1.mu + lgamma1[t])`

- Kenneth P. Burnham and Gary C. White. Evaluation of some random effects methodology applicable to bird ringing data. *Journal of Applied Statistics*, 29(1-4):245–264, January 2002. ISSN 0266-4763, 1360-0532. doi: 10.1080/02664760120108755.
- Noel Cressie, Catherine A. Calder, James S. Clark, Jay M. Ver Hoef, and Christopher K. Wikle. Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3):553–570, 2009. ISSN 1051-0761. doi: 10.1890/07-0744.1.
- B. J. Halstead, G. D. Wylie, P. S. Coates, P. Valcarcel, and M. L. Casazza. 'Exciting statistics': The rapid development and promising future of hierarchical models for population ecology. *Animal Conservation*, 15(2): 133–135, March 2012. ISSN 1469-1795. doi: 10.1111/j.1469-1795.2012.00540.x.
- William L Kendall and James D Nichols. On the use of secondary capture-recapture samples to estimate temporary emigration and breeding proportions. *Journal of Applied Statistics*, 22(5-6):751 – 762, November 1995. ISSN 0266-4763. doi: 10.1080/02664769524595.
- William L Kendall, Kenneth H Pollock, and Cavell Brownie. A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics*, 51(1):293–308, 1995. doi: 10.2307/2533335.
- William L Kendall, James D Nichols, and James E Hines. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology*, 78(2):563 – 578, March 1997. ISSN 0012-9658. doi: 10.1890/0012-9658(1997)078[0563:ETEUCR]2.0.CO;2.
- Kenneth H Pollock. A capture-recapture design robust to unequal probability of capture. *The Journal of Wildlife Management*, 46:752–757, 1982. doi: 10.2307/3808568.
- Robert W. Rankin, Krista E. Nicholson, Simon J. Allen, Michael Krützen, Lars Bejder, and Kenneth H. Pollock. A full-capture Hierarchical Bayesian model of Pollock's Closed Robust Design and application to dolphins. *Frontiers in Marine Science*, 3(25), 2016. doi: 10.3389/fmars.2016.00025.
- Robert William Rankin. EM and component-wise boosting for Hidden Markov Models: A machine-learning approach to capture-recapture. *bioRxiv*, 2016. doi: 10.1101/052266.
- J. Andrew Royle and William A. Link. Random effects and shrinkage estimation in capture-recapture models. *Journal of Applied Statistics*, 29(1-4):329–351, January 2002. ISSN 0266-4763, 1360-0532. doi: 10.1080/02664760120108746.