WILD 502 - Estimating abundance & recruitment with mark-recapture methods
Ch. 18 of WNC & Ch. 12 of C&W

History:

- original Jolly-Seber model focused on $\hat{N}$
- CJS focused on $\hat{S}$ - heavily used in last 15-20 years
- Recently, more development and use of models that also focus on
  - $\hat{N}_i$
  - $\hat{f}_i$ **, recruitment rate**
  - $\hat{\lambda}_i$ **, population growth rate**

Utility:

Studies of population dynamics seek broader information than just survival rates. Consider a population with survival rate *S* and recruitment rate *f*. I think you can see how it would be useful &/or desirable to estimate those quantities for a population of interest, especially if we have accompanying measures of time-specific covariates, e.g., environmental conditions that might relate to those quantities.

$$N_{t+1} = N_t \cdot S_t + N_t \cdot f_t$$

$$\lambda_t = N_{t+1} / N_t$$

Data:

The data are like those seen for CJS models. Animals are captured, marked, released, and recaptured (or resighted) at discrete sampling occasions.

JS models use LLLL… capture history formats just like those used in CJS models.

Can also work with summary statistics, but this has fallen out of favor because:

- It's difficult to compute GOF statistics without individual histories
- Use of individual covariates requires individual histories

Modeling framework & assumptions:

In CJS studies, we did not concern ourselves with capture probabilities for unmarked animals: all probability statements for encounter histories were conditioned on the animal being captured and released, i.e., were for marked animals only.

In JS studies, we must model the capture process for unmarked animals as well. This is what allows us to estimate N and f. We assume that p_i is the same for marked and unmarked animals.

If you think about this, it makes good intuitive sense. If you want to estimate survival rate, you can just work with marked animals. But, if you want to know about population size &/or recruitment, you have to learn about the unmarked animals.

The implications are not trivial. You need to record the numbers of unmarked animals in surveys. If you are doing recaptures, this is probably done as a matter of course. In fact, it's typical in most studies, to mark and release all unmarked animals so this is a non-issue. But, if some unmarked animals are released unmarked, they need to be recorded. In studies relying on resightings, you also need to record the numbers of unmarked animals that are observed during surveys. This is *not* necessary in CJS studies.

The JS models focus on a single age class, typically adults. We cannot estimate abundance of young animals because we cannot estimate a capture probability for them. Remember back to multi-age CJS models: we estimate $p_i^{age>0}$ by knowing how many animals in age class v were alive at time i but were not detected. To do that, we worked with animals that were caught before and after occasion i. It's not possible to do such a thing for animals in the youngest age class: they didn't exist prior to the occasion in question (see page 443 of *WNC*, left-hand column). So, unless we're using closed models with repeated mark-recapture surveys within an occasion (which will be covered under Robust Design analyses next), we can't estimate abundance of the youngest age class.

The study area must remain a consistent size. If you don't then the concepts of population size and recruitment become nonsensical.

Assumptions:

- p_i is the same for marked and unmarked animals
- all those used in CJS models

Jolly-Seber Approach

Model Structure for situation with losses on capture:

- p_i – capture probability
- $q_i = 1 - p_i$ – probability of not being captured
- φ_i – survival probability
- χ_i – probability of not seeing an animal again after occasion i
- η_i – probability of release for marked animals (m_i) caught on occasion i
- η'_i – probability of release for unmarked animals (u_i) caught on occasion i

Unknown random variables

- N_i – total number of animals in the population exposed to sampling on occasion i
- M_i – total number of marked animals in the population just before sampling occasion i
- $U_i = N_i - M_i$ – total number of unmarked animals in the population just before occasion i
- n_i - total number of animals caught on occasion i
- m_i - total number of marked animals caught on occasion i
- u_i - total number of unmarked animals caught on occasion i
- B_i – the number of new animals joining the population between samples i & $i+1$; animals that enter the population between occasions i & $i+1$ but die before $i+1$ are excluded.

Likelihood – example for animals caught on occasion 1 in 3-occasion study = (u_1)

$$P[\chi_\omega] = \left[\frac{U_1!}{u_1!(U_1-u_1)!} p_1^{u_1}(1-p_1)^{U_1-u_1} \right]$$

$$\times \left\{ \frac{u_1!}{\prod_\omega (\chi_\omega)!(\chi_\omega^-)!} (\eta_1'\chi_1)^{\chi_{100}}(1-\eta_1')^{\chi_{\overline{100}}} (\eta_1'\phi_1 p_2\eta_2\chi_2)^{\chi_{110}}[\eta_1'\phi_1 p_2(1-\eta_2)^{\chi_{\overline{110}}}] \right.$$

$$\times [\eta_1'\phi_1(1-p_2)\phi_2 p_3\eta_3]^{\chi_{101}}[\eta_1'\phi_1(1-p_2)\phi_2 p_3(1-\eta_3)]^{\chi_{\overline{101}}}$$

$$\left. \times (\eta_1'\phi_1 p_2\eta_2\phi_2 p_3\eta_3)^{\chi_{111}}[\eta_1'\phi_1 p_2\eta_2\phi_2 p_3(1-\eta_3)]^{\chi_{\overline{111}}} \right\}$$

Three aspects of the process are dealt with in the likelihood (see equation 18.2 on page 497 of *WNC* for a version of the likelihood that uses the m_{ij}–array summary statistics and that has the 3 components separated).

1. The initial binomial term involves the capture of unmarked animals from the total number of unmarked animals on each occasion. This can be seen easily in the likelihood above. This component can be written as follows (here, we're not just working with u_1).

$$P_i(\{u_i\} | \{U_i\},\{p_i\}) = \prod_{i=1}^{K} \left[\frac{U_i}{u_i!(U_i-u_i)!} p_i^{u_i}(1-p_i)^{U_i-u_i} \right]$$

2. Losses on capture are handled for marked and unmarked animals. This can be seen by examining the equation and considering the meaning of η_i and η'_i in the equation. Remember that, as explained by *WNC* on page 497, χ_ω^- denotes the number of animals with capture history

ω that were <u>not</u> released following the final capture, e.g., 12_{110}^{-} would denote 12 that animals that were caught on occasions 1 and 2 were not released on occasion 2.

3. The recapture of marked animals that were released is modeled exactly the same way as was done in the CJS modeling.

Trap deaths or research removals can be accommodated throughout. For the example of 3 occasions & animals caught at time 1 (u_1), if all animals caught are released, the likelihood simplifies to:

$$P[\chi_\omega] = \left[\frac{U_1!}{u_1!(U_1 - u_1)!} p_1^{u_1} (1-p_1)^{U_1 - u_1} \right]$$

$$\times \left\{ \frac{u_1!}{\prod_\omega (\chi_\omega)!} (\chi_1)^{\chi_{100}} (\phi_1 p_2 \chi_2)^{\chi_{110}} [\phi_1 (1-p_2)\phi_2 p_3]^{\chi_{101}} (\phi_1 p_2 \phi_2 p_3)^{\chi_{111}} \right\}$$

$$\hat{N}_i = n_i / \hat{p}_i$$

- Need $\hat{p}_i$, so can estimate N_i for $i=2,\ldots$, K-1

$$\hat{B}_i = \hat{N}_{i+1} - \hat{\phi}_i (\hat{N}_i - n_i + R_i), \quad \text{where } R_i \text{ is the number released on occasion } i$$

- $-n_i + R_i$ simply accounts for the number of animals caught but not released back into the population.
- Need $\hat{N}_i$ and $\hat{N}_{i+1}$ so can estimate B_i for $i=2, \ldots$, K-2

Modifications:

- Partially open
  - Deaths-only = no immigrants, no births (or they can be identified and excluded)
  - Births-only = not likely but available if of interest
- Reduced parameters
  - Constancy in some or all parameters over time
  - Parameters vary as functions of time-specific covariates
  - Parameters vary as functions of group-specific covariates
  - Trap response models
  - Multiple-age models
  - Multi-state models
  - Survival rate depends on individual covariates
  - Capture probability as a function of individual covariates is harder but some work is possible

An Example of Results – from page 507 of *WNC*

Sample period	Sampling dates	Abundance	Recruitment
1	6/27-7/1	--------	--------
2	8/1-8/5	74 (3.89)	17 (2.56)
3	8/29-9/2	59 (5.84)	21 (2.71)
4	10/3-10/7	62 (5.66)	19 (2.82)
5	10/31-11/4	55 (5.61)	--------
6	12/4-12/8	--------	--------

$$\hat{\lambda}_2 = 59/79 = 0.80$$

$$\hat{\lambda}_3 = 62/59 = 1.05$$

$$\hat{\lambda}_4 = 55/62 = 0.89$$

Multiple Formulations of the Same Process

Alternatives to the JS model above now exist. The primary difference among these is how they model the number of unmarked animals that are caught on each occasion.

In MARK, there are multiple parameterizations of the JS model:

- *POPAN* - a model developed by Schwarz and Arnason that parameterizes the Jolly-Seber model in terms of a super population (N), and the probability of entry into that superpopulation at each occasion
- Link-Barker – this is the Schwarz and Arnason model re-parameterized with the entry probabilities replaced by recruitment (*f*), where f_i is interpreted as a per capita recruitment rate, or the net new animals entering the population between occasion *i* and *i+1* per animal alive at occasion *i*.
- Burnham Jolly-Seber – Burnham developed a parameterization of the Jolly-Seber model that provides estimates of the rate of population change (lambda) and the population size on the first trapping occasion (N_i). Convergence is difficult for this likelihood and so it has not seen as much use as the others.
- Pradel models
  - Pradel-recruitment parameterization has similarities to the Link-Barker model in that it works with recruitment
  - Pradel-lambda has similarities to Burnham's JS model in that it works with lambda.

Choosing among the parameterizations – This topic is nicely covered in Ch. 12 of *C&W*, which was written by Carl Schwarz and Neil Arnason. Quoting from that chapter (page 12-11)

1. … Only certain of the formulations can be used in MARK if losses-on-capture occur in the experiment.
2. … more importantly, different formulations give you different types of information and can be used to test different hypotheses. All of the formulation should give the same estimates of survival and catchability as all formulations estimate these from recaptures of previously marked animals using a CJS likelihood component. Even though all the models give different types of estimates for growth or recruitment or births, it is always possible to transform the estimates from one type to another by simple transformation and the standard errors can be found using the delta method.
3. The major equivalents are between NET births, recruitment, and population growth parameters.

- $$f_i = B_i / N_i = Nb_i / N_i$$

- $$\lambda_i = \frac{N_{i+1}}{N_i} = \frac{N_i \phi_i + B_i}{N_i} = \phi_i + f_i$$

Table 12.5: *Summary of criteria to choose among the different JS formulations*

formulation	losses on capture	estimates available for			
		abundance	net births	recruitment	λ
POPAN	yes	yes	yes	no	no
Link-Barker	yes	no	no	yes	yes
Pradel-recruitment	no	no	no	yes	yes
Burnham JS	yes	yes	yes	no	yes
Pradel-λ	yes	no	no	no	yes

• The implementation of Burnham's JS model in **MARK** often does not converge, and is not recommended (although convergence problems may be minimized for some models if simulated annealing is used for the numerical optimization. - J. Laake & E.G. Cooch, *pers. obs.*)

• The standalone package of *POPAN* will estimate recruitment, and population growth as derived parameters

POPAN – a model developed by Schwarz and Arnason that parameterizes the Jolly-Seber model in terms of a super population (N), and the probability of entry into that superpopulation at each occasion. This parameterization will be our focus. For each group being considered in the analysis, we'll have 4 PIMs

- ϕ_i - apparent survival (t-1 estimates)
- p_i - capture probability given the animal is alive and on the study area, i.e., available for capture (*t* estimates)
- *pent_i* - probability of entry into the population for the occasion (*t-1* estimates for occasions 2, 3, …, *t*) [The probability of being in the population on the first occasion is *pent(0)* = 1 - sum(*pent*(i)). Convergence of this model is difficult to achieve unless you use the MLogit link for the *pent* parameters.]
- *N* - superpopulation size (1 estimate), where the superpopulation consists of all animals that were available for capture at any time during the study

The members of *N* are assumed to enter the sampled population at different times according to the entry probabilities. Recruitment (*B_i*) is distributed as a multinomial with parameters (*N, pent_0,…,pent_{K-1}*).

All animals present on the 1st sampling period are "new" with respect to sampling and so

$$B_0 = N \qquad\qquad N = \sum_{i=0}^{K-1} B_i$$

An Example – taken directly from Ch. 12 of *C&W*

```
/* Estimating salmon numbers returning to spawn in Chase River 1989 */
/* These are the male salmon with two groups.  */
/* Group1 = adults . group2=jacks */
/* Survey conducted over 10 weeks. Weeks 1 & 2 pooled. weeks 9 & 10 pooled */
11000000 -1 0  ;  /* tagnum=1  */
10000000 0 1   ;  /* tagnum=3  */
10000000 0 1   ;  /* tagnum=4  */
10000000 0 1   ;  /* tagnum=6  */
11000000 0 1   ;  /* tagnum=8  */
10110000 0 1   ;  /* tagnum=9  */
10000000 0 1   ;  /* tagnum=10 */
10000000 1 0   ;  /* tagnum=11 */
10000000 0 1   ;  /* tagnum=12 */
   ... additional histories follow ....
```

Model = p., phi(t), pent(t)

```
                          Chase 1989 Adults Only

             Real Function Parameters of {p(*), phi(t

 Parameter                    Estimate          Standard Error
 --------------------         ------------      --------------
    1:Phi                     0.5843217         0.1007203
    2:Phi                     0.9264599         0.1759781
    3:Phi                     0.7973302         0.1328374
    4:Phi                     0.9216908         0.1312724
    5:Phi                     0.5792390         0.1372571
    6:Phi                     0.3306517         0.1258056
    7:Phi                     0.6180676         0.1405468
    8:p                       0.3153771         0.0422789
    9:pent                    0.0449635         0.0506509
   10:pent                    0.3325037         0.0718152
   11:pent                    0.1107655         0.0718261
   12:pent                    0.1742893E-06     0.1128910E-03
   13:pent                    0.0279008         0.0411325
   14:pent                    0.1313499         0.0355726
   15:pent                    0.2496515E-06     0.7215834E-04
   16:N                       331.99832         29.077189
```

pent(0) = 1-(.045+.333+.111+0+.028+.131+0) = 1-0.648 = 0.352

```
 Grp. Occ.    B-hat            Standar
 ---- ----    --------------   -------
   1    1     14.927808        16.7742
   1    2     110.39068        25.7066
   1    3     36.773956        23.4912
   1    4     0.5786375E-04    0.03748
   1    5     9.2630338        13.6037
   1    6     43.607953        12.9183
   1    7     0.8288389E-04    0.02395
              Population Estimates of

 Grp. Occ.    N-hat            Standar
 ---- ----    --------------   -------
   1    1     117.03475        23.9943
   1    2     67.202663        15.8427
   1    3     171.72480        28.7789
   1    4     164.92468        22.8121
   1    5     150.16623        25.9079
   1    6     91.611254        19.2550
   1    7     71.254152        16.0931
   1    8     31.707588        10.7750
```

$$\hat{B}_0 = \hat{N}_1 = 117.04 = 0.352 * 332$$

$$\hat{B}_1 = \hat{N}\hat{b}_1 = 332 \times .044 = 14.9$$

$$\hat{N} = \sum_{i=0}^{7} \hat{B}_i = 117.04 + 14.93 + ... + 43.61 + 0.00 = 332$$

$$\hat{N}_1 = \hat{B}_0$$

$$\hat{N}_2 = \hat{N}_1\hat{\phi}_1^t + \hat{B}_1$$

# of spawning males present on each sampling

$$\hat{N}_2 = (\hat{N}_1 - loss_1)\hat{\phi}_1^{1.5} + \hat{B}_1 = (117.03 - 0)0.584^{1.5} + 14.92 = 67.2 \text{[1]}$$

[1]. *The time intervals between occasions 1 & 2 and 6 & 7 were 1.5 months apart, i.e., t=1.5 vs. 1.0*