Some Initial Survival estimates for Steller sea lions in the Russian Far East using RMark

Tommy Garrison

University of Washington, QERM

March 30, 2008

Outline

- Data
- Cormack-Jolly-Seber Models
- Accounting for Age and other factors
- Software
- Some results

Assumptions

- Captured animals are representative of target population
- Number of releases known
- Fate of one individual independent of fate on individuals in other cohorts
  - Possible Violation, mother and pup tag
- The model is correct!

Framework

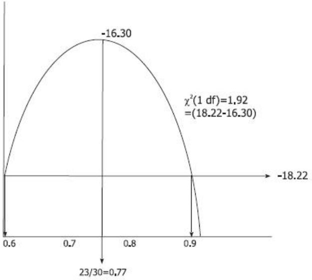
- ϕ_i = probability of survival from time i to time $i+1$
- p_i = probability of encounter at time i, conditional on being alive
- Encounter history: 1 = Encounter with marked individual, 0 = Not seen or dead
- For 3 occasions, the 4 possible outcomes
  - $111 = \phi_1 p_2 \phi_2 p_3$
  - $110 = \phi_1 p_2[\phi_2(1-p_3) + (1-\phi_2)] = \phi_1 p_2(1-\phi_2 p_3)$
  - $101 = \phi_1(1-p_2)\phi_2 p_3$
  - $100 = (1-\phi_1) + \phi_1(1-p_2)(1-\phi_2) + \phi_1(1-p_2)\phi_2(1-p_3)$
    $= 1 - \phi_1 p_2 - \phi_1(1-p_2)\phi_2 p_3$

Procedures

- All the possible outcomes (categories) are observed a certain amount of time and are associated with a specified unknown probability to be estimated, i.e. Multinomial Data
- Maximize $L(\phi_i, p_i|data)$, the multinomial likelihood function
- Use $-I^{-1}$, $I =$ Information Matrix (second partial derivatives of $L(\phi_i, p_i)$ evaluated at the MLE's), to obtain estimates of the variance and covariances

Profile likelihood

- Possible to get unreasonable 95% CI
- Adjust, by essentially determining where the $-2logL(\phi_i, p_i|data)$ intersects with the line χ^2_{df}
- *Automagically* get reasonable interval,
- Simulated Example



Basics

- CJS somewhat limiting
  - Limited to models where ϕ and p are time dependent or independent
  - More realistic to account for Age, Cohort (Categorical), Weight (Continuous)
- This can be achieved if model is stated as GLM
  - ex: $Y_{ij} = \beta_0 + \beta_i(\text{Sex}) + \beta_j(\text{Age}) + \epsilon_{ij}$, where $Y_{ij} \in \{0, 1\}$
  - $\equiv \phi_i = x_i^{'} \beta$
- For some $\pi_i \in (0, 1)$, but we have that $x_i^{'} \beta \in (-\infty, \infty)$
- Hope if data arises from an Exponential family distribution

Logit Example

- Link linear predictors to mean response
- $\pi_i \in (0, 1)$
- $\frac{\pi_i}{1-\pi_i} \in (0, \infty)$
- $log(\frac{\pi_i}{1-\pi_i}) \in (-\infty, \infty)$ [This is known as the Logit Link]
- So model is $log(\frac{\pi_i}{1-\pi_i}) = x_i^{'}\beta$
- Backtransform to get practical estimates of π_i

Multinomial Logit Model

- Default link for RMark (not Mark)
- Log-odds follow linear model
- $\nu_{ij} = log(\frac{\pi_{ij}}{\pi_{iJ}}) = \alpha_j + x_i^{'}\beta_j$ where α_j is a constant,β_j is a vector of regression coefficients, $j = 1, 2, ..., J - 1$
- Backtransform to get $\pi_{ij} = \frac{\exp(\nu_{ij})}{\sum_{k=1}^{J} \exp(\nu_{ik})}$
- By modeling survival and recapture probability, dependent on specified covariates with a GLM (using the Multinomial Link) we can get age, sex, etc. specific estimates!

Candidate Models

- All possible combinations of. . .
- Recapture probability
  - Constant, $p(\sim 1)$
  - Depends on Sex, $p(\sim \text{Sex})$
- Age-specific survival probability
  - $\phi(\sim \text{Age})$
  - $\phi(\sim \text{Age} + \text{Sex})$
  - $\phi(\sim \text{Age} + \text{Cohort})$
  - $\phi(\sim \text{Age} + \text{Year})$
  - $\phi(\sim \text{Age} + \text{Sex} + \text{Cohort})$
- Models with Interactions

Choosing Models

- Biological sense
- Another Information Criteria (AIC)
  - $\text{AIC} = -2 \log L(\phi_i, p_i) + K$
- Weighted Model Averaging
  - $w_i = \frac{\exp(\triangle \text{AIC}/2)}{\sum \exp(\triangle \text{AIC}/2)}$
- Goal prediction or description ?

- Parameter Index Matrix (PIM)

RMark

- Formula Based
- Process data with some basic commands
- Create a design matrix
- Example call to `mark`

```
> p.sex <- list(formula= ~sex)
> Phi.age.cohort <- list(formula= ~age + cohort)
> m1 <- mark(ssl.m.process,ssl.m.ddl,
+ model.parameters=list(Phi=Phi.age.cohort,p=p.sex.age)
```

Typical Output (Truncated)

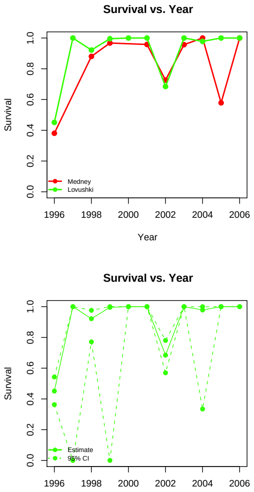
```
Output summary for CJS model
Name : Phi(~age + cohort)p(~sex)

Npar :  21  (unadjusted=18)
-2lnL:  2696.425
AICc :  2739.123  (unadjusted=2732.9402)

Beta
                  estimate           se           lcl           ucl
Phi:(Intercept) -0.4868053    0.2128575    -0.9040060    -0.0696046
Phi:age1         0.3994463    0.2623180    -0.1146970     0.9135895
...
Phi:age8        15.2147940  988.5269000 -1922.2980000 1952.7276000
Phi:age9         0.8039507    0.9148112    -0.9890793     2.5969806
Phi:age10       17.2084980 1834.3955000 -3578.2068000 3612.6238000
...
p:sexM          -0.2097346    0.1422148    -0.4884756     0.0690064

Real Parameter Phi
Group:sexF
          1996      1998      1999      2001      2002      2003      2004      2005      2006
1996 0.3806464 0.8804031 0.9674378 0.9581046 0.7264748 0.9568190 0.9999996 0.5786284 0.9999999
1998           0.5629123 0.6575593 0.9841919 0.7036992 0.9795593 0.8476911 0.9789176 0.9999998
1999                     0.5344496 0.9322057 0.9822998 0.6791793 0.9771257 0.8322464 0.9764094
2001                               0.6387592 0.7250056 0.9549136 0.9884367 0.7653015 0.9850292
2002                                         0.6965578 0.7738896 0.9649066 0.9910691 0.8089103
2003                                                   0.7449847 0.8132832 0.9722155 0.9929688
2004                                                             0.7479329 0.8156372 0.9726332
2005                                                                       0.3573682 0.4532968
2006                                                                                 0.8439105
```

Discussion

- Lots of Information that's hard to concisely summarize
- Different model might tell different story, what's the *correct* model
- Note the unreasonable parameter estimates
- This data is very sparse
  - Juvenile don't typically return to recapture site for several years until large enough to compete with older males, i.e. 10000000000010101
  - Of 100 sea lions in 1996, less than 10 were recaptured

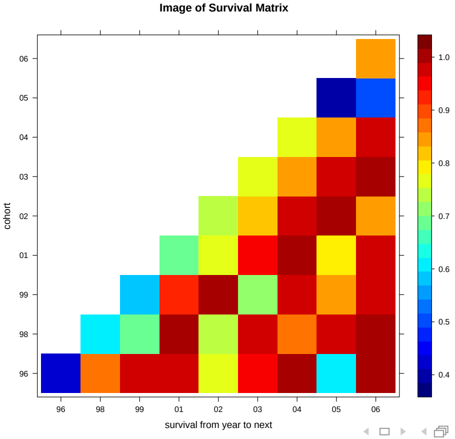
Sex, Age Cohort Model

- $\phi(\sim \text{Age} + \text{Cohort})p(\sim \text{Sex})$, 1996 Medney cohort survival



Visual Summary

Since this is a QERM presentation, I had to include some neat plot! This is a nice quick way to summarize/visualize that enormous survival matrix. $\phi(\sim \text{Age} + \text{Cohort})p(\sim \text{Sex})$ model, Medney Island.



More to do. . .

- Model Diagnostics. . . overdispersion, deviance, etc.
- Pick a final model!
- Bin Age classes, less parameters
- Fix recapture probabilities
- Time scales, analyze Island effect