

CHAPTER 20

Young survival from marked adults

Paul Lukacs & Victoria Dreitz, *University of Montana*

The survival probability of juvenile animals is often of interest, but juvenile animals may be difficult to mark or marking may impact survival. In such cases, a count of the young animals with an attending adult may be the only data that can be obtained. If the adult is marked and highly detectable (such as with a radio transmitter), then survival of individual young can be estimated from the counts of young in a family group. This chapter describes the **MARK** implementation of the ‘**Young Survival from Marked Adults**’ model developed in Lukacs *et al.* (2004).

The ‘young survival model’ is an extension of the Cormack-Jolly-Seber (CJS) model (which is discussed in detail in Chapters 1 → 7). The model contains two types of parameters, apparent survival (φ) and detection probability (p). Data for the CJS model consist of a string of ones and zeros indicating an individual animal is detected or not detected, whereas the data for the young survival model are a string of counts on young in a family group. Encounter histories are entered in the input file at the family group level rather than the individual young animal level. Despite this, it is important to note that the parameters refer to individual young animals not to the group in which they were detected. When there is only one young per family group, the young survival model reduces to the CJS model.

The young survival model is based on extending the CJS model to all possible combinations of events that could have occurred given the count of young observed. For example, in a clutch of three plover chicks that are known to be alive at hatching, consider a case where only two chicks are observed at the next sampling occasion. There are six events that could have occurred to produce that observation.

- First, two chicks survived and both were detected. This event occurs with probability $\varphi^2(1-\varphi)p^2$. The event can happen in $\binom{3}{2} = 3$ ways (because each of the 3 chicks starting the interval could have been the one that died), therefore the total probability of observing the event is $3\varphi^2(1-\varphi)p^2$.
- Second, all three chicks may have survived but only two were detected. This event occurs with probability $\varphi^3p^2(1-p)$. The event also can happen in $\binom{3}{2} = 3$ ways (again because each of the 3 surviving chicks may have been the one not detected), therefore the total probability of observing that event is $3\varphi^3p^2(1-p)$.

The probability of observing two chicks given three chicks were alive at the previous occasion is therefore $3\varphi^2(1-\varphi)p^2 + 3\varphi^3p^2(1-p)$. Based on the combinatorics, it is easy to see that the number of possible events grows rapidly as the family group size increases. Therefore, precision decreases rapidly as the number of events increases.

20.1. Assumptions

The young survival model assumes that the number of young at the start of observation is known (if, for example, the number of eggs hatched out of a nest is counted exactly). If the initial size of the family group is not known and the entire family group is not observed at least once for each group in the data set, parameter estimates will be biased. Family groups can vary in size, but the initial size of each group must be known, and is assumed in **MARK** to be the maximum count observed (even if not the first count). **MARK** will function properly if the size of the family group is observed at least once during the study, but to assure parameter estimates are unbiased studies should be designed to know the family group size at the beginning of observation.

The young survival model also assumes that there is no brood mixing. Therefore, the number of young can never exceed the number the brood began with and new individuals cannot join the family group. Within a family group, young are assumed to be exchangeable, meaning that each individual animal has the same survival and detection probability as its siblings on a given occasion.

The parameter ϕ in the young survival model is defined as an *apparent* survival probability (the probability of an individual being alive and available for detection at occasion i given the individual was alive at $i-1$) rather than true survival (the probability of being alive at occasion i given the individual was alive at $i-1$) just as it is in the CJS model. In cases where separation of young from the attending adult results in death of the young, ϕ will essentially equal true survival. Study design should be carefully considered to separate fledging from mortality. Design considerations for this aspect will be highly species-dependent.

20.2. Data

Data for the young survival model consist of counts of young detected per family group on each occasion. The **MARK** input file uses the two digit count format similar to that of the Poisson log-normal mark-resight models (Chapter 19). Given this format, 0 to 99 young could be counted per group (although precision decreases greatly with more than 5 or so young per family group). Consider a case where a nest has 3 eggs hatch and 3, 2, 3, 0, and 1 young are counted over 5 occasions. The encounter history for this brood would be

```
0302030001      1;
```

Note that all counts must be entered as two digits with a leading 0 for values from 0-9 (for example 3 would be coded as '03'). If there are occasions where the marked attending adult is not located, and therefore the young have no chance of being located, two dots ('..') can be used to represent that information. For example, if the adult is not detected (or not sampled) on occasion 3, then the encounter history would be

```
0302..0001      1;
```

In all cases, a frequency (count of number of family groups with this encounter history) per encounter history (1 in this example) and a semicolon are needed as in other **MARK** data types.

The young survival model can incorporate covariates for more flexible modeling. Family group-specific covariates can be included just as individual covariates are included after the frequency column in the input file as in other **MARK** data types. Given individual young are not handled nor individually identifiable in this data collection scheme, only covariates measured on the whole family group will be available. Environmental covariates (covariates acting on all family groups) can be included in the design matrix (presented in detail in Chapter 6).

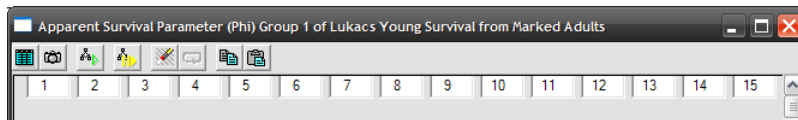
20.3. Model Implementation

To demonstrate the implementation of the young survival model in **MARK**, we use an example of mountain plover (*Charadrius montanus*) data simplified and modified from Dreitz (2009). The data consist of encounter histories of 31 occasions, one group and three covariates: year (three years coded in two covariates, with year 1 the '00' code)] and sex (0 = female, 1 = male) of the attending adult. The maximum number of chicks per family group is three in this case, because mountain plovers seldom or never have more than three eggs in a nest. Below are four encounter histories from the file `plover_mark.inp` showing the input file format.

```
0300..00.....00....02..00....01..00..00..01..01..00..01.. 1 0 1 0 ;
0301000203....03.....01..01..010001..00....01..01..00..01..03 1 0 1 0 ;
01..00..00..01..01..00..0100..00..00..00.....00..00..00..00.. 1 0 1 1 ;
02000000....00....00.....01....01..01..00..00..... -1 0 1 0 ;
```

The fourth record has a '-1' in the frequency column indicating a family group that was censored at the last '00' because of a radio failure on the attending adult. However, because all the succeeding potential encounters are coded as '.', the '-1' value is equivalent to the usual '1'. Losses on capture in this data type are coded as '.' for the encounters after the loss, rather than as negative values of the frequency.

Model implementation for the young survival model follows that of the CJS model directly. The primary difference between implementing the young survival model and the CJS model in **MARK** is in the PIM format. The CJS model has triangular PIMs with rows representing cohorts while the young survival model has a single row in each PIM.



The single row PIM forces age-specific models to be built in the design matrix rather than the PIM, but otherwise does not limit the capability of **MARK**. On the whole, the single row PIM makes more sense than a triangular PIM for the young survival model because data generally consist of multiple occasions within a single year. Multiple years of data are better input as multiple attribute groups or as individual covariates indicating the year.

A year-specific survival model can be built in the design matrix with covariates as we demonstrate here with the plover data.

Design Matrix Specification (B = Beta)				
B1 Phi Intercept 2004	B2 Phi 2005	Pam	B3 Phi 2006	B4 p
1	y04	1:Phi	y05	0
1	y04	2:Phi	y05	0
1	y04	3:Phi	y05	0
1	y04	4:Phi	y05	0
1	y04	5:Phi	y05	0
1	y04	6:Phi	y05	0
1	y04	7:Phi	y05	0
1	y04	8:Phi	y05	0

In this case y_{05} and y_{06} are binary covariates indicating data that occur in each year, when both covariates are 0 the data are from 2004.

Many young animals have a lower survival probability immediately after they are born, but survival increases quickly thereafter. One way to model this is to place a trend on the first few occasions after birth. In our plover example, all family groups are started on the first occasion, so the implementation of the trend is simple.

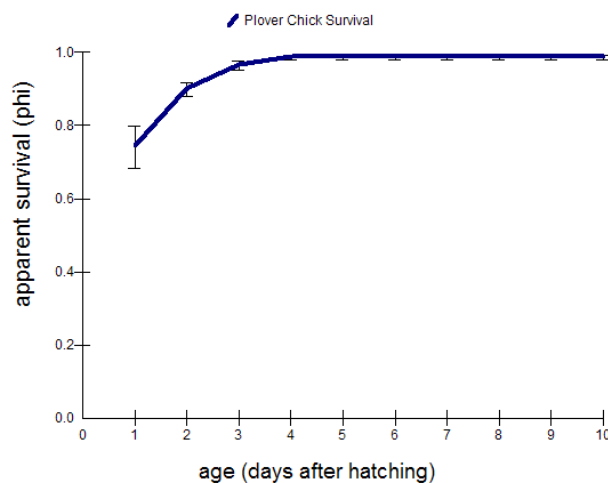
Design Matrix Specification (B = Beta)				
B1 Phi Intercept	Parm	B2 Phi age	B3 p	
1	1:Phi	3	0	
1	2:Phi	2	0	
1	3:Phi	1	0	
1	4:Phi	0	0	
1	5:Phi	0	0	
1	6:Phi	0	0	
1	7:Phi	0	0	
1	8:Phi	0	0	

If family groups enter the data set on different occasions then a covariate would be needed to indicate the occasion of first observation. Likewise, occasions before the family group is entered into the analysis have to be coded with the '.' notation, because the '00' notation would suggest that the family group was sampled, and no young counted.

The age-specific model estimates suggest that survival is substantially lower immediately after hatching, but increases with age. The β parameter is strongly negative for the slope on the age parameter.

Parameter	LOGIT Link Function Beta	Standard Error	Parameters of $\{\phi(\text{age}) p(\cdot)\}$	
			95% Confidence Interval	
			Lower	Upper
1:Phi Intercept	4.4528516	0.2809240	3.9022406	5.0034626
2:Phi age	-1.1261975	0.1218076	-1.3649404	-0.8874546
3:p	-0.3606089	0.0631196	-0.4843233	-0.2368945

The real parameters for the $\{\phi_{age} p.\}$ model show the same effect on the probability scale.



20.4. Parameters and Sample Size

Parameter identifiability in the young survival model is the same as that in the CJS model. The last p and φ in a time-specific model are confounded. The use of fully time-specific models will be rare with this data type because occasions will often refer to days or other short time intervals.

Effective sample size for the young survival model is based on the number of “family group releases” (counts >0 and not ‘.’) in the encounter histories. This matches the implementation of the CJS model in **MARK**, but differs from the way that effective sample size was computed in Lukacs *et al.* (2004). Lukacs *et al.* (2004) used the number of encounter histories. The appropriate value for effective sample size in mark-recapture models remains an unresolved issue and more research on the topic is warranted. **MARK** allows the user to specify the value of effective sample size if desired by selecting the option under the ‘Adjustments | Effective Sample Size’ menu.

20.5. Relationship with CJS and Multi-state Models

The young survival model has its roots in the multi-state and CJS models. When all family groups have only a single young, the ‘CJS’ and ‘young survival’ models are equivalent. This can be demonstrated with the cell probability for a ‘0101’ encounter in the young survival model assuming a group size of 1 and the cell probability of a ‘11’ encounter history for a CJS model:

$$\begin{aligned}\Pr[0101_{YS}] &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \varphi & (1-\varphi) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \varphi p \\ &= \Pr[11_{CJS}]\end{aligned}$$

The relationship between the multi-state model and the young survival model is a bit less obvious. To demonstrate it, let us consider a case of two young per family group. First, we must define the states in terms of the number of young alive at any given occasion. Our three states are: *A*: 2 young live, *B*: 1 young live, *C*: 0 young live. Note that the states are defined by the true number of young alive not the observed number of young. When state definitions are based on whether or not an animal is alive, survival is modeled with the transition matrix and the multistate model survival probability is set to 1. Survival transitions can only proceed in one direction; therefore the lower left side of the matrix has all zero elements. The transition matrix takes the following form:

$$\Psi = \begin{bmatrix} \psi^{AA} & \psi^{AB} & \psi^{AC} \\ 0 & \psi^{BB} & \psi^{BC} \\ 0 & 0 & \psi^{CC} \end{bmatrix} = \begin{bmatrix} \varphi^2 & 2\varphi(1-\varphi) & (1-\varphi)^2 \\ 0 & \varphi & (1-\varphi) \\ 0 & 0 & 1 \end{bmatrix}$$

The capture probability matrix is more complicated. The states are defined by the true number of live young, but the observations are of young counted (some may be missed). Therefore, the only observation with a known state is that of a 2. There is state uncertainty in observations of 1 and 0. The detection probability matrix must account for the state uncertainty. For an observation of 1 young, the matrix is:

$$D(p) = \begin{bmatrix} 2p(1-p) & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix handles the two possible states that an observation of 1 young could imply: either one young alive and detected, or two young alive with one detected and one not detected. A standard multi-state model would simply have a p for the state observed or $(1-p)$ for all states if the animal was missed.

20.6. Summary

The young survival from marked adults data type in **MARK** allows apparent survival of young animals in a family group with a marked attending adult to be estimated. The model only requires counts of young in the family group rather than individually marked young. When designing a study, it is important to consider trade-offs of counting young versus marking young. As long as the marking process does not affect the survival of the young, estimates of survival from marked young will be more precise than that from an equal number of unmarked young. It may be possible to count more family groups of unmarked young than can be marked; in that case weighing the tradeoff requires a more careful examination of the benefits of marks versus a larger sample size.

20.7. References

- Dreitz, V. J. (2009) Parental behaviour of a precocial species: implications for juvenile survival. *Journal of Applied Ecology*, **46**, 870-878.
- Lukacs, P. M., Dreitz, V. J., Knopf, F. L., and Burnham, K. P. (2004) Estimating survival probabilities of unmarked dependent young when detection is imperfect. *Condor*, **106**, 927-932.