<div align="right">

CHAPTER 9

</div>

Joint live encounter & dead recovery data

The first chapters in this book focussed on 'typical' open population mark-recapture models, where the probability of an individual being encountered (dead or alive) was defined by 2 parameters: the probability the animal survived and the probability that an animal alive in state (r) at time (i) is alive and encountered in state s at time ($i + 1$). In Chapter 8, we considered the situation where individuals are 'found dead' (or, recovered), as opposed to 'encountered alive'. In all cases, we model the probability of being encountered (either alive or dead) as a function of underlying parameters. However, up until now, we've only considered what we might call 'either or' models – the marked individuals is either resighted alive, or it isn't. The marked individual is found dead and reported, or it is not. And so forth.

However, clearly there can arise situations where one sort of encounter precludes (or determines or otherwise affects) the probability of an encounter of another kind. The simplest example of this (and the one we will focus on to begin this chapter) is the situation where an individual is found dead. Clearly, if the individual is dead, then it cannot be subsequently resighted as a living individual – the fact that it is dead precludes any other encounter process.

The technical issue (and the major theme of this chapter) is – 'how do we use this extra information in our analysis?'. In fact, this 'theme' of 'using extra information' is currently an area of very active research. As we will see, there are many situations in which data of various types (e.g., recoveries, recaptures, telemetry) can be used simultaneously, to provide estimates of parameters that are not estimable using only one source of data, to improve precision of parameter estimates beyond what can be achieved using data from one source only, and to provide some 'flexibility' in accommodating encounter data which might have been collected 'opportunistically' throughout the year. As we will see, the basic ideas for using data from various sources are merely extensions to what we've already discussed. Of course, the 'challenge' is (as always) in the details.

9.1. Combining live encounters + dead recoveries – first steps. . .

In 1993, Ken Burnham of Colorado State University published a seminal paper outlining an approach for combining dead recovery and live encounter data into a single analysis – we encourage you to read the original text. Here, we summarize some of the basic results presented in Burnham (1993), and discuss how the approach is implemented in program **MARK**.

First, we need to identify the basic elements of where the two types of data differ, and where they are the same. For both dead recoveries, and live encounters, there is some underlying probability that the marked individual will survive over a specified time interval, and that there is some probability that the marked individual will be encountered, either conditional on being alive and in the sample at the

end of the interval (in the case of a typical recapture analysis), or dead, recovered and reported during the interval (in the case of a typical dead recovery analysis). Clearly, an individual cannot be recaptured (or otherwise encountered) alive subsequent to being found and reported dead.

However, it is important to remember that the probability of recapture is conditional on (i) being alive, and (ii) remaining in the sampling region (or, as you'll recall from earlier chapters – permanent emigration and mortality are inexorably confounded in a standard recapture study). But what about recoveries? Clearly, it might be reasonable to assume that an individual is equally like to be recovered dead and reported regardless of whether it is in the sample or not. Of course, it is possible under some circumstances that the probability of mortality (recovery) and being reported is dependent on whether or not the marked individual is in or out of the sample area, but for now, we'll assume that the probability of recovery and reporting is independent of sampling location. In other words, we assume that dead recoveries can occur – and be reported – from 'anywhere'.

Given this assumption, then you might have already noted that the addition of recovery data to our analysis gives us something we didn't have before – the ability to separate 'true' mortality from 'apparent' mortality. In simplest terms, in a 'live-encounter' recapture study, the estimate of φ is an estimate of a product of two different events: survival, and fidelity. More formally, we can write

$$\varphi_i = S_i F_i,$$

where S_i is the probability of survival from (i) to ($i + 1$), and F_i is the probability of remaining in the sample area (F for 'fidelity') between (i) and ($i + 1$).

As we discussed in Chapter 8, S_i can be estimated directly using analysis of data from dead recoveries. As such, we might derive an estimate for F_i given estimates for S_i and φ_i (i.e., $\hat{F}_i = \hat{\varphi}_i/\hat{S}_i$). Of course, an *ad hoc* way to accomplish this would be to run separate recovery and recapture analysis, and take the estimates from each and 'do some algebra'. However, such an *ad hoc* approach provides no means for estimating the precision of the estimated fidelity parameter, nor the covariance of F_i with the other parameters.

Further, the assumption of *permanent* (as opposed to *temporary* or *transient*) emigration is not a prerequisite. The 'location' of a live individual during a capture occasion could be random with respect to whether or not it is in the sampling region (and thus at risk of being captured). In this case, the parameter F_i is the probability at time (i) that the individual is at risk of capture given that it is alive, p_i is the probability of capture given that the individual is alive and in the sample (i.e., at risk of capture).

What does this mean? Basically, it means that under a 'random' emigration model, φ_i is the true survival probability (S_i), and p_i is the product of F_i and the traditional conditional capture probability. Questions concerning permanent versus temporary emigration, availability for encounter, and so on, are treated more fully in Chapter 16 (which addresses the 'robust design'). Here, we consider one particular approach to separately estimating survival and fidelity.

9.1.1. Estimating fidelity rate: some key assumptions. . .

The combined 'live encounter-dead recovery' approach originated with considerations of sampling from harvested species. In such cases, the dead recoveries are assumed to occur over a spatial scale that is larger than the scale over which live encounters occur. This is shown in the diagram at the top of the next page. The dark, smaller area is the area where the marking and subsequent live encounters occur. Dead recoveries occur anywhere outside the darker area, but within the area bounded by the dotted line.
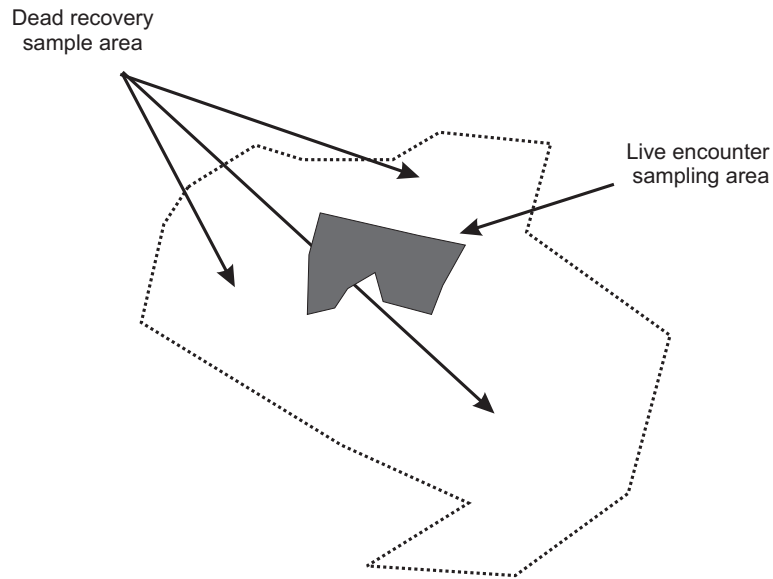
The key, though, is that the original assumption in Burnham (1993) was that all the dead recoveries occur outside the area where live encounters occur (i.e., outside the dark area, but within the dotted line). If this assumption is met, and if emigration from the sampling area is permanent, then you can partition φ as the product of true survival and fidelity, since in this case, all of the recoveries occur outside of the live encounter sampling area.

Clearly, this is not always going to be the case – especially for non-game species. In such cases, interpretation of the various parameters is potentially not so simple. If all dead recoveries and live encounters occur in the same sampling area, then clearly realized fidelity F for the marked sample is 1. So, you can use the same approach, except you fix $F = 1$. But, in general, you need to think hard about what the parameters mean, and how the sampling assumptions (and whether or not they are met) influence the interpretation of the parameter estimates.

9.2. Live + dead encounters: underlying probability structure

For now, we'll assume that if an individual emigrates from the sampling area, it is a 'permanent' emigration, such that if the analysis were based entirely on recapture data, an emigrated individual would appear 'dead'. (*Note*: the situation when emigration from the sampling is not permanent, but instead is *temporary*, is considered in Chapter 16, where we develop the 'robust design'). How would such a model be parameterized?

If you think about it for a moment, you should realize that the fate of this individual is governed by several probabilities: S (the probability of surviving the interval), r (the probability of being found dead and reported – the recovery probability using the Seber parametrization discussed in Chapter 8), F (the probability of fidelity to the sampling region – i.e., remaining in the sample. $(1 - F)$ is the probability of permanently emigrating), and p (the probability of live encounter, conditional on being alive and in the sampling region).

The 'fate diagram' for the 'permanent emigration' model is shown at the top of the next page. Note that the sequencing of survival before emigration is arbitrary – we could just as easily (and equivalently) place the fidelity 'decision' first. However, the ordering does affect how the probability expressions corresponding to a given encounter history are written.
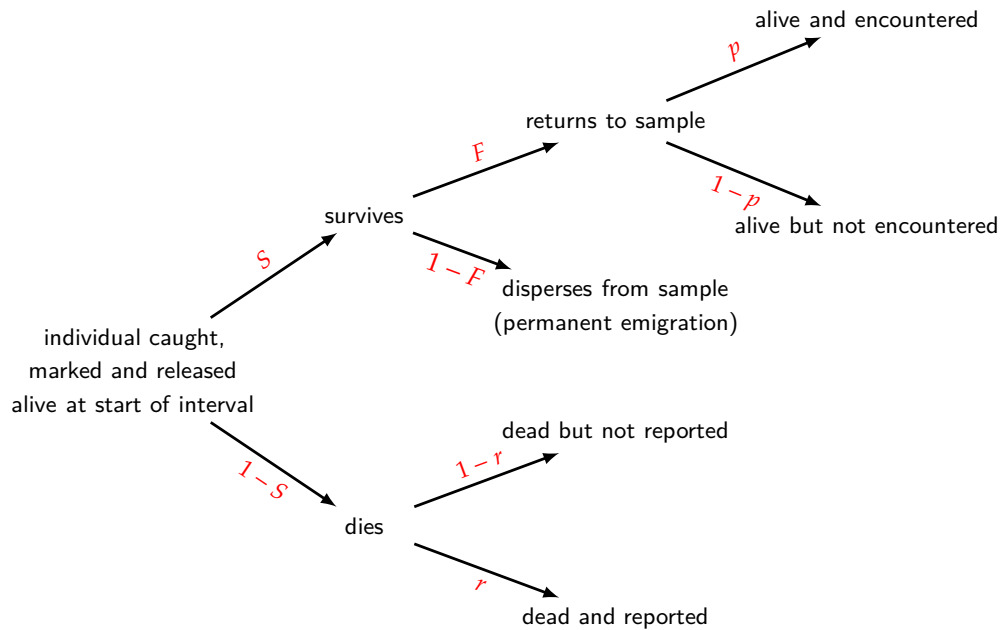
It is important to keep time-scale in mind when considering this diagram. First, live encounters occur *at* (i), $(i+1)$,...,$(i+n)$, while recoveries occur *between* (i), $(i+1)$,...,$(i+n)$. Second, the number of estimated live recapture-based or dead recovery-based parameters depends on when you end your study. You might recall from Chapter 2 ('data formatting') that joint live-dead analysis uses the '**LD**' encounter history format, where the **L** refers to 'live' and **D** to 'dead' encounters, respectively. Thus, a history of '1011' refers to an individual marked and released on the first occasion, that survives the interval between occasion 1 and occasion 2 (since it was subsequently encountered) and then recaptured at occasion 2 and recovered dead during the interval following occasion 2. The '**LD**' format assumes that each recapture occasion is followed by a recovery interval. If your study terminates with the final recapture occasion, and you do not collect recovery data during the following interval, you need to 'code the terminal recovery column (the terminal '**D**' column) as zero, and fix the recovery probability for this interval to 0.

There are several important points to make here. First, the assumption that the terminal live encounter (recapture) occasion is followed by a period (interval) when the individual may be recovered (i.e., encountered dead) means that for standard **LD** analysis where all parameters are fully time-dependent, the number of parameters for survival (S) and recovery (r) will be one greater than the number of parameters for fidelity (F) and recapture (p). So, the PIMs for survival and recovery parameters will have one more column than will the PIMs for fidelity and recapture. The second point concerns the structural equivalence of the Seber 'S and r' parametrization and the CJS parametrization for live encounter studies (see Chapter 8). As such, you'll need to remember that, as with a CJS analysis of live encounter data, although there are more S parameters than F parameters, the last S is not estimable in a fully time specific model. Its presence is an artifact of modeling recoveries as $(1 - S)r$ (using the Seber convention) instead of as f (using the Brownie convention).

Consideration of the probability statements corresponding to various encounter histories will give you a better idea of what is going on. In the table at the top of the next page, we indicate the history in both **LD** format (using **L** and **D** to indicate live encounter or dead recovery, respectively), and the actual binary (0 or 1) coding used in the input file. Each probability statement refers to a different path by which the encounter history could be realized, and assumes that survival occurs before the fidelity 'decision' (as shown in the fate diagram on the preceding page). The total probability of the particular

LD history	binary history	probability
LD00	1100	$(1 - S_1)r_1$
L0L0	1010	$S_1 F_1 p_2 S_2 + S_1 F_1 p_2 (1 - S_2)(1 - r_2)$
L00D	1001	$S_1 F_1 (1 - p_2)(1 - S_2) r_2 + S_1 (1 - F_1)(1 - S_2) r_2$
L0LD	1011	$S_1 F_1 p_2 (1 - S_2) r_2$
L000	1000	$(1 - S_1)(1 - r_1) + S_1 \left[(1 - F_1) + F_1 (1 - p_2) \right] \left[S_2 + (1 - S_2)(1 - r_2) \right]$

encounter history is the sum of the individual probabilities.

Take a moment to make sure you see where these probability statements come from – they're clearly more complicated than those for recapture or recovery analyses alone. For example, consider history '1001' (corresponding to 'L00D'). There are two ways this history could be achieved: the individual clearly survives the first interval (since it is recovered during the second interval, but it could either (1) remain in the sample, and not be seen at occasion 2, and then die and be recovered, or (2) leave the sample area (after which $p = 0$; it cannot be recaptured if it is outside the sample area), and then die and be recovered. The key to remember is that (in theory) a dead recovery can occur whether the individual is in the sample area or not – the same cannot be said for live encounters, which require the marked individual be in the sampling region.

9.3. Combined recapture/recovery analysis in MARK: marked as adult + young

We now address the more practical issue of using program **MARK** to simultaneously analyze dead recovery and live encounter data. To demonstrate the mechanics, we simulated a data set (**ld1.inp**) with the following structure. We assumed 2 age-classes (young, and adult), with the young age class spanning the first year of life. Within an age class, parameter values were held constant over time – any variation in the parameter values was among age classes. We assume that individuals are marked as both young and as adults at each occasions (recall from Chapter 8 that we suggested that for recovery analyses, we must at least mark adults as well as young. But what about for combined recovery-recapture analyses? More on that later.). The parameter values used in simulating the data set were:

age class	S	p	r	F
young	0.4	0.5	0.6	0.6
adult	0.8	0.5	0.6	0.9

Thus, in the simulated data set, younger individuals had lower survival ($S_y < S_a$), and were less likely to remain in the sampling region during the first year of life (conditional on remaining alive; $F_y < F_a$). Recapture and recovery probabilities were equal for both age classes. There were 8 'occasions' in the simulated data set (where each 'occasion' consists of the recapture event followed by the full year after the recapture event during which recoveries might occur), and 1,500 individuals newly marked and released in each age class on each occasion. The simulated data were constructed using the 'probability sequence' depicted in the figure shown on the previous page.

Start **MARK**, and select the data file **ld1.inp**. We noted already that there were 8 occasions in the data set. However, to remind ourselves about the somewhat different structure of an '**LD**' data set, let's

have a look at the `.INP` file shown at the top of the next page).

Again, given the **LD** data format, for 8 capture occasions, we have 8 corresponding intervals over which recoveries can occur – thus, 16 columns in total. Each consecutive pair of values denotes the encounter history for a given year ('11' – seen and recovered in the same year, '10' – seen but not recovered in a given year, and '01' – not seen but recovered in a given year). Note also that we have two 'groups' individuals marked as young (first frequency column), and marked as adults (second frequency column), respectively. For example, 5 individuals marked as young and 23 individuals marked as adult had encounter history '1010110000000000'.

Next, we need to pick the '**Data Type**'. The third item on the list in the model specification window is '**Joint Live and Dead Encounters**'. When you click the radio button, a window will pop up asking you to select from among the 4 different live encounter-dead recovery data types are currently implemented in **MARK**. For the present example, we want to select '**Joint Live and Dead Encounters (Burnham)**', as shown below:
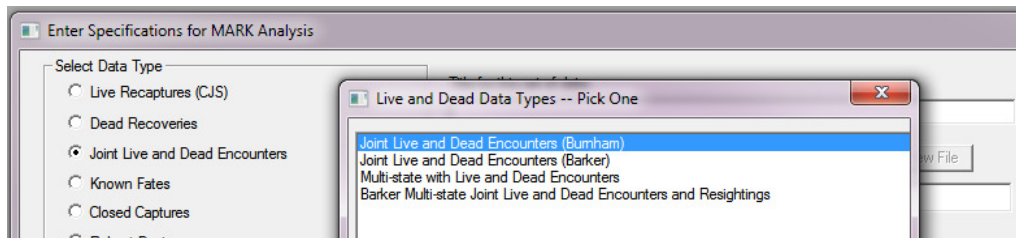
Before proceeding, one important note: Burnham originally used the 'classic' Brownie parametrization (Brownie *et al*. 1985) for the 'dead recovery' side of things – specifically, he used the Brownie parametrization of survival (S) and recovery probability (f). Recall from Chapter 8 that under this parametrization, recovery probability is defined as the product of the kill probability (K), the retrieval probability (c), and the reporting probability (λ) – in other words, f is the probability that the marked individual will be harvested (killed), the mark retrieved and then reported.

You may also recall from Chapter 8 that **MARK** also allows you to specify a different parametrization (the Seber parameterization) for recovery probability (r, rather than f), wherein

$$f_i = r_i(1 - S_i).$$

Chapter 9. Joint live encounter & dead recovery data

For the joint analysis of dead recovery and live encounter data, program **MARK** uses the Seber parameterization (based on parameters '*S* and *r*'), and <u>not</u> the Brownie parametrization originally used by Burnham, primarily to take advantage of increased modeling flexibility this parametrization provides. It is important to keep this in mind.

Once you've confirmed that you've set up the specifications for 8 occasions, two attribute groups (marked as adults and marked as young), and have correctly selected the joint data type (Burnham), click the '**OK**' button to continue. As usual, you're presented with the open PIM for the survival parameter (survival is always the first parameter **MARK** considers). To see the other parameters, open up the PIM chart (shown at the top of the next page). Again, note immediately that the model is specified by 4 parameters (8 'blue' boxes in total – 4 for marked as young, 4 for marked as adult). **MARK** indexes them starting with survival (*S*), then recapture probability (*p*), then recovery probability (*r*), and finally the fidelity parameter (*F*). As your reading of previous chapters has hopefully made clear, it is essential to know the sequence that **MARK** uses in 'treating' (indexing) the parameters in the model. Also, confirm that the number of columns in the PIMs for *F* and *p* is one less than the number of columns for *S* and *r* – make sure you understand why!
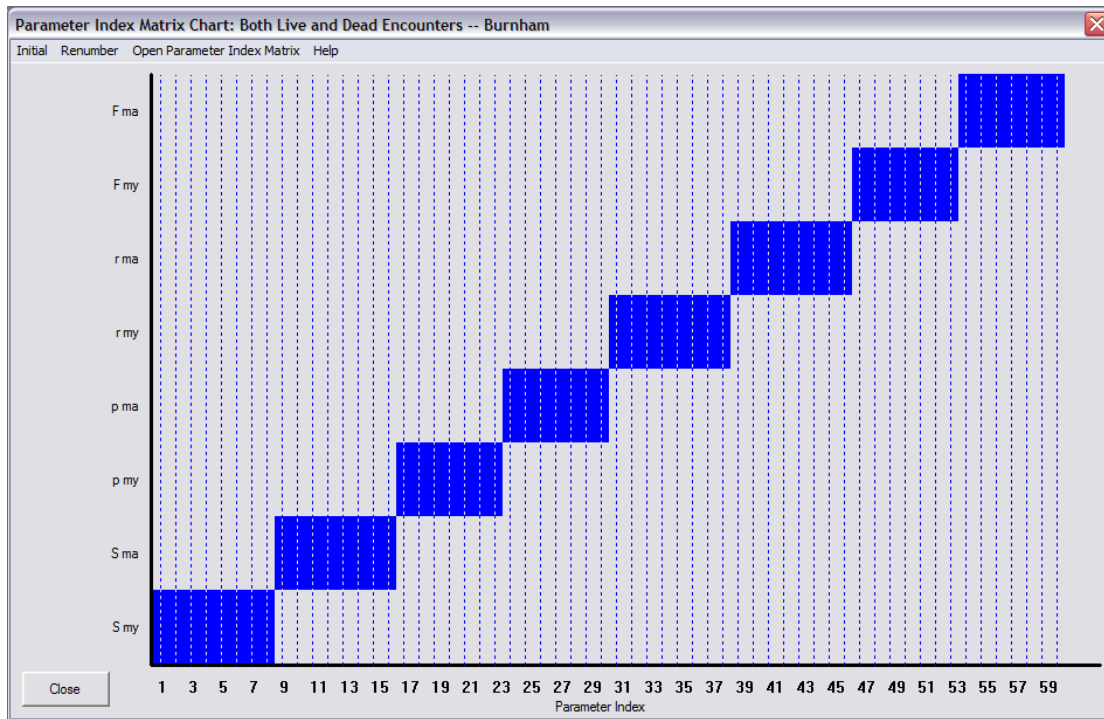
Since our purpose here is not to conduct a 'real analysis' (since the data are completely artificial), we'll only run the 'true' model (under which the data were simulated) – model $\{S_{a(\cdot),j(\cdot/\cdot)} \, p \cdot r \cdot F_{a(\cdot),j(\cdot/\cdot)}\}$. Recall that the quickest way to accomplish this is to modify the PIM chart directly. This is especially true in this case since the 'age structure' we want to impose on the survival, recovery and fidelity parameters is 'constant' (i.e., no temporal variation within age class). This is easily accomplished by right-clicking each of the blue-boxes in the PIM chart (note also that if there had been temporal variation within age class, we'd have to resort to modifying the PIM for each parameter independently, outside the PIM chart – the right-click menu accessible through the PIM chart only allows you to specify a 'constant' age model).

Right-click each of the parameters we want to add age-structure to (*S*, *r* and *F* respectively, for

individuals marked as young) , and select 2 as the maximum age (since there are only 2 age classes – young, spanning one year, and adult, spanning the rest of the years in the study). For the parameters corresponding to individuals marked as adults, and for recapture rate for both age classes, we want a constant value – right-click the 'blue-box' for these parameters and select '`constant`'.

Also, for recapture, make the recapture probability the same for both age classes by 'stacking' the boxes over each other. Finally, note that survival and fidelity probabilities for 'adults' is the same, regardless of whether the individual was marked as young or adult. Thus, the blue box for 'adult' survival and fidelity should overlap the corresponding parameter in the blue box for individuals marked as you. Then, somewhere in the PIM chart, but not over one of the blue-boxes, right-click and select '`renumber with overlap`' – remember that this eliminates 'gaps' in the parameter indexing, but allows for overlapping for some parameters.

Your PIM chart should look like the following – make sure you see the correspondence between this PIM chart and the true model:
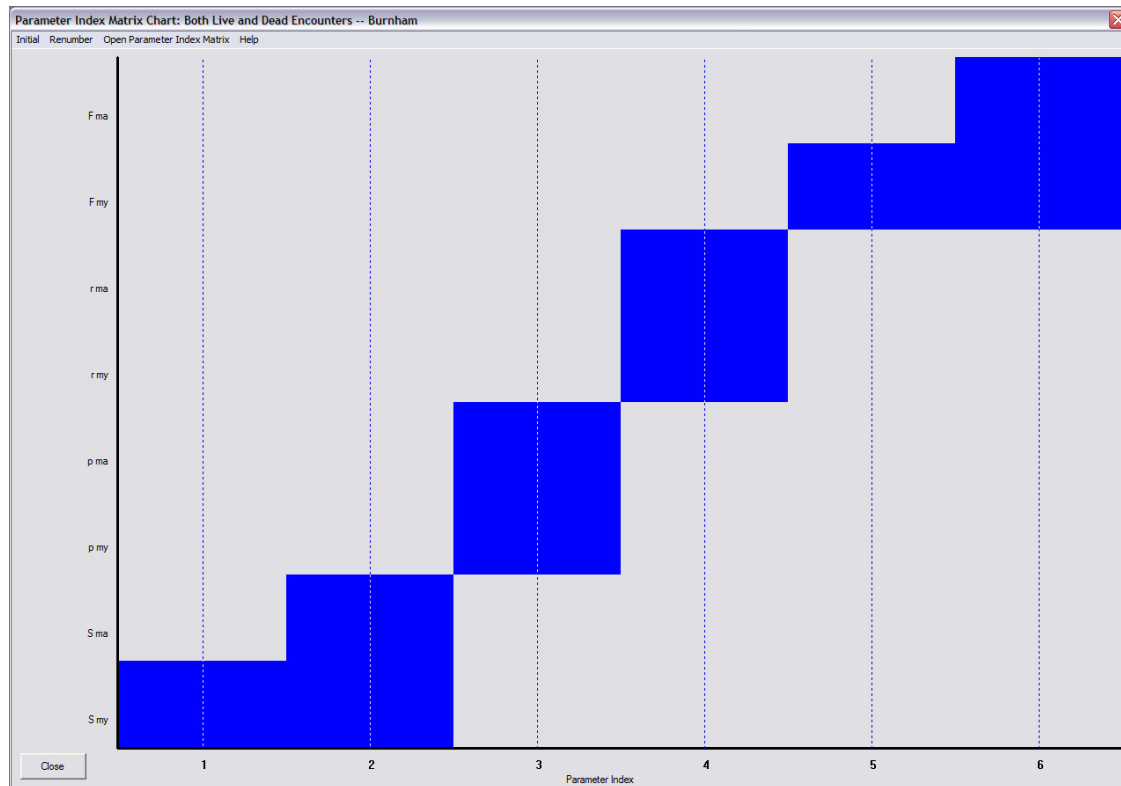
Go ahead and run this model, and add the results to the browser. The parameter estimates for this model (we've added a couple of blank lines to highlight the respective parameters) are shown at the top of the next page. Parameter 1 is the juvenile survival probability (i.e., age $0 \rightarrow 1$ year), for individuals marked as juveniles (obviously!). Parameter 2 is the survival probability for adults (including individuals marked as young, and as adults). Both estimates are quite close to the 'true' values of 0.4 and 0.8, respectively. Parameters 3 and 4 are the common recapture and recovery probabilities, which are also very close to the 'true' values of 0.5 and 0.6, respectively. Finally, parameters 5 and 6 are the fidelity probabilities for juveniles and adults, and again, both are close to the 'true' fidelity probabilities of 0.6 and 0.9.

```
                     Chapter 9 -  live-dead (LD1.inp example)

         Real Function Parameters of {S(a20./.)p(.)r(.)F(a20./.) - true model}
                                                  95% Confidence Interval
        Parameter              Estimate      Standard Error      Lower            Upper
        ---------------------  --------------  --------------  --------------   --------------
            1:S               0.3915571       0.0070583       0.3778132        0.4054750
            2:S               0.7949066       0.0040304       0.7868941        0.8026934

            3:p               0.5137565       0.0048175       0.5043105        0.5231928

            4:r               0.5977562       0.0060602       0.5858230        0.6095748

            5:F               0.5773108       0.0138017       0.5500537        0.6041051
            6:F               0.8811859       0.0052713       0.8704594        0.8911352
```

Wile this concordance is perhaps unremarkable given we fit the 'true' generating model to the simulated data, for comparison, let's re-run these same data through **MARK**, using '`Dead recoveries (Seber)`' and '`Live recaptures (CJS)`'. We can do this by first editing the `ld1.inp` file for each data type. Can you think of how we would do this? If you can, then you clearly have a fair understanding of the **LD** data format. For the live encounters only analysis, you could extract all the **L** columns, yielding an 8 occasion `.INP` file. If you have some facility with programming, this is in fact relatively straightforward.

For example if you had the following **LD** encounter history

```
101010000010
```

then you could use the basic ideas represented in the following **R** script to extract every odd element (since for an **LD** history, the odd elements represent of the live encounters):

```
LD <- "101010000010"
L_hist <- paste(unlist(strsplit(LD,""))[seq(1,nchar(LD),2)],collapse="")

print(L_hist)
 [1] "111001"
```

Alternatively, you could simply make all of the **D** columns have a 0 value (i.e., discarding all of the dead recovery data), and use a 'multi-state approach', setting the probability of encounter in the 'dead state' to 0. For this latter approach, we treat dead recoveries as an 'unobservable state' (we cover multi-state models in Chapter 10).

What about for the 'dead recoveries only' analysis? Recall from Chapter 2 and Chapter 8 that there are, in fact, two ways to code recovery data. The 'classic' approach is to use a recovery matrix. However, while traditional, the recovery matrix 'lumps' individuals, and inhibits the ability to constrain estimates as functions of individual covariates (covered in Chapter 11). Conveniently, the encounter history format for 'dead recoveries only' is also an **LD** format – except that all **L**'s after the initial marking event are coded as '0' (check Chapter 2 to make sure you understand this).

For example, given the following **LD** history, which has both live encounters and one dead recovery, which occurs after the live encounter during **LD** interval 5:

```
101010001100
```

then you could use the basic ideas represented in the following **R** script to change every odd element, after the initial marking event, to a '0' (since for an **LD** history, the odd elements represent the live encounters):
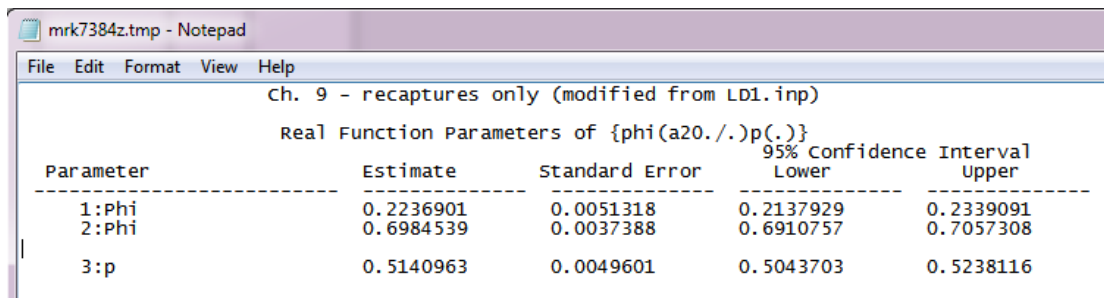
```
LD <- "101010001100"

LD_list <- as.list(substring(LD, seq(1,nchar(LD),1), seq(1,nchar(LD),1)) )
LD_list[seq(3,length(LD_list),2)] <- "0"

D_only <- paste(LD_list, collapse='')

print(D_only)
  [1] "100000000100"
```

For our analysis of **ld1.inp**, modified as described to include 'live encounters only', we fit model $\{\varphi_{./.}p_.\}$. For our analysis of **ld1.inp** modified to include 'dead recoveries only', we fit model $\{S_{./.}r_.\}$ – using the Seber parametrization.

Start with the recaptures only analysis. The results for model $\{\varphi_{./.}p_.\}$ are shown below:

```
mrk7384z.tmp - Notepad
File  Edit  Format  View  Help
              Ch. 9 - recaptures only (modified from LD1.inp)

           Real Function Parameters of {phi(a20./.)p(.)}
                                                    95% Confidence Interval
      Parameter                Estimate      Standard Error      Lower            Upper
  --------------------------  -------------  --------------  --------------  --------------
        1:Phi                  0.2236901      0.0051318        0.2137929       0.2339091
        2:Phi                  0.6984539      0.0037388        0.6910757       0.7057308

        3:p                    0.5140963      0.0049601        0.5043703       0.5238116
```

Hmmm...are these estimates right? The estimate of $\hat{\varphi}_Y = 0.2237$ doesn't seem particularly close to the value of $S_y = 0.4$ used in the simulation – is there a problem? No! The key is remembering that, under the assumption of permanent emigration, $\varphi_i = S_i F_i$. In the simulation, $F_y = 0.6$, and thus the expectation for $\varphi_y = S_y F_y = (0.4 \times 0.6) = 0.24$, which is quite close to the estimated value of 0.2237.

In fact, if we take $\hat{S}_y = 0.3916$, and multiply it by $\hat{F}_y = 0.5773$, we get $(0.3916 \times 0.5773) = 0.2261$, which is very close to our estimate of $\hat{\varphi}_Y = 0.2237$. Similarly, the estimate of $\hat{\varphi}_a = 0.6985$, is quite close to the expectation of $\varphi_a = (0.8 \times 0.9) = 0.72$, and the product of $(\hat{S}_a = 0.7949) \times (\hat{F}_a = 0.8812) = 0.7005$. The estimate for recapture probability ($\hat{p} = 0.51$) is also close to the true parameter value ($p = 0.50$).

Thus, as expected, a recaptures-only analysis provides robust estimates for apparent (or local) survival (φ), and recapture rate. However, since permanent emigration and mortality are confounded in a recaptures-only analysis, the estimate of φ represents only a minimum estimate of true survival, and will generally be lower than the true survival probability (by a factor corresponding to the fidelity rate).

What about the recoveries-only analysis? Clearly, estimates of survival from a recoveries only analysis are anticipated to be more accurate estimates of 'true' survival, since they deal directly with dead individuals (i.e., are not confounded by emigration). The estimates from our analysis (shown at the top of the next page) are close to the 'true' values for both survival and recovery probabilities, as expected.

It is worth noting that these estimates are not identical to those from the joint live-dead analysis presented a few pages back. For example, our recoveries only estimate for $\hat{S}_y = 0.3684$ is somewhat lower than the estimate from the joint live-dead analysis, $\hat{S}_y = 0.3916$. This is to be expected, since restricting the analysis to 'dead recoveries only' ignores the additional information from the 'live encounter data' which is included in the joint analysis.

```
mrk8663z.tmp - Notepad

File   Edit   Format   View   Help
                 Ch. 9 - recoveries only (modified from LD1.inp)

             Real Function Parameters of {S(a20./.)r(.)}
                                                      95% Confidence Interval
     Parameter              Estimate      Standard Error      Lower           Upper
  --------------------------  --------------  --------------  --------------  --------------
       1:S                    0.3710960      0.0082901      0.3549984      0.3874850
       2:S                    0.7785659      0.0056335      0.7673275      0.7894104

       3:r                    0.5773389      0.0070297      0.5635039      0.5910537
```

9.4. Marked as young only: combining live encounters + dead recoveries

As we discussed at length in Chapter 8, there are significant difficulties in recovery analysis with data from individuals marked as young only (this is probably a good point to go back and review). Does the 'extra information' from live encounters help us at all?

We can explore this by simply deleting the frequency column for adults from the **ld1.inp** file, and running the analysis. For the recoveries only analysis, using just the data from individuals marked as young, our estimates of survival and recovery probability are:

```
mrk3136z.tmp - Notepad

File   Edit   Format   View   Help
              Chapter 9 - recoveries only - young only (modified from LD1.inp)

             Real Function Parameters of {S(a20./.)r(.)}
                                                      95% Confidence Interval
     Parameter              Estimate      Standard Error      Lower           Upper
  --------------------------  --------------  --------------  --------------  --------------
       1:S                    0.3485340      0.0095993      0.3299630      0.3675769
       2:S                    0.7309589      0.0143136      0.7019991      0.7580742

       3:r                    0.5611687      0.0079891      0.5454543      0.5767614
```

Superficially, these estimates for $\hat{S}_y$ and $\hat{S}_a$ don't seem too bad, but they are clearly different from the true values of $S_y = 0.4$ and $S_a = 0.8$.

However, our main objective here is to see how much improvement there is, if any, if you use the combined live encounter-dead recovery models, if the data are restricted to marked as young only? Here are the estimates from fitting the 'true' live encounter-dead recovery model to the 'marked as young only' joint live capture-dead recovery data:

```
mrk5453z.tmp - Notepad

File   Edit   Format   View   Help
             Ch. 9 - joint live-dead - LD1.inp restricted to young only.

             Real Function Parameters of {S(a20./.)p(.)r(.)F(.)}
                                                      95% Confidence Interval
     Parameter              Estimate      Standard Error      Lower           Upper
  --------------------------  --------------  --------------  --------------  --------------
       1:S                    0.3665133      0.0071175      0.3526782      0.3805721
       2:S                    0.7800113      0.0084236      0.7630573      0.7960763

       3:p                    0.4599018      0.0103172      0.4397571      0.4801785

       4:r                    0.5832488      0.0068552      0.5697539      0.5966200

       5:F                    0.7727018      0.0116419      0.7490800      0.7947093
```

Chapter 9. Joint live encounter & dead recovery data

In this case, the estimates are much closer to the true parameter estimates. This suggests that by combining data from dead recoveries and live recaptures, everything is estimable, with less bias and better precision (whereas with a recoveries only analysis from individuals marked as young, everything is not estimable, or if it is, with significant bias and loss of precision).

As noted by Brownie, depending on what assumptions are made concerning constancy of some parameters (as in our example), a few of the parameters in an analysis of recoveries from individuals marked as young are estimable (for example, if we assume constant adult recovery probabilities, adult survival probabilities are estimable, but juvenile survival and both juvenile and adult recovery probabilities are confounded). So, think carefully. But – as this simple example demonstrates, there is great utility in combining data from different sources.

9.5. Barker's joint live 'encounter' + dead recovery model

Consider a study where once a year, over (say) a few days each summer, you capture and recapture, mark, and release individuals in some population of interest. These data might form the basis of a typical mark-recapture analysis. But, suppose that during the interval between summer sampling occasions, you potentially 'encounter' marked individuals from your study population. For example, you might have reports from the public ('citizen science') that someone has 'encountered' one of your marked individuals. This encounter might be a 'dead recovery'. The preceding parts of this chapter addressed the joint use of live-encounter and dead recovery data (*sensu* Burnham 1993).

Richard Barker (Barker 1997, 1999) extended Burnham's (1993) live/dead model to the case where both dead encounters + live resightings are potentially reported during the open period between live recapture occasions (as shown in the following diagram). The population is not assumed to be closed while encounters are being obtained and recoveries of tags from dead animals (dead resightings) may be reported.
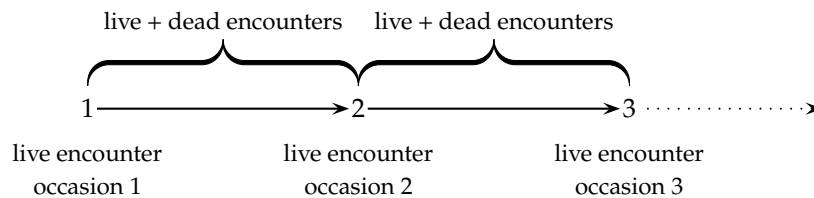
Because the encounter interval between live encounter sampling occasions is open, it is possible for an animal to be resighted alive several times within an interval then be reported dead. For these animals, only the last dead sighting is used in the model, the earlier live resightings in that period are ignored. Therefore the status of an animal on resighting (live or dead) is determined on the last occasion on which it was resighted in the open interval.

Although the model is complicated and involves 4 sets of nuisance parameters for the recapture and resighting/recovery process, the additional data from resightings and tag recoveries can lead to substantial gains in precision on survival probability estimates. Currently, the model does not allow estimation of abundance or recruitment. For a detailed description of the models, see Barker (1997, 1999).

Parameters in the model are:

p_i = the probability an animal at risk of capture at (i) is captured at (i)

r_i = the probability an animal that dies over the interval from (i) to $(i+1)$ is found dead and the band reported

R_i = the probability an animal that survives from (i) to $(i+1)$ is resighted (alive) some time between (i) and $(i+1)$

R'_i = the probability an animal that dies over the interval from (i) to $(i+1)$, without being found dead, is resighted alive in the interval from (i) to $(i+1)$ before it died

F_i = the probability an animal at risk of capture at (i) is at risk of capture at $(i+1)$

F'_i = the probability an animal not at risk of capture at (i) is at risk of capture at $(i+1)$ (*NB*: this differs from the definition in Barker 1997)

The resighting parametrization used in **MARK** differs from that described by Barker (1997). An advantage of the parametrization used by **MARK** is that it enforces certain internal constraints that arise because the joint probability $\Pr(A \cap B)$ should always be less than or equal to the unconditional probabilities $\Pr(A)$ and $\Pr(B)$. For example, the **MARK** parametrization ensures that the probability an animal is resighted alive over the interval from (i) to $(i+1)$, and survives from (i) to $(i+1)$, is less than the probability it is resighted alive over the interval from (i) to $(i+1)$. It also ensures that $\Pr$(resighted alive and dies over the interval from (i) to $(i+1)$ without being reported), is $<$ the $\Pr$(dies over the interval from (i) to $(i+1)$ without being reported). These internal constraints are not enforced by the other parametrization.

9.6. Barker Model – 'movement'

Between trapping sessions, animals are permitted to leave the study area then return. If an animal is in the study area then it is considered 'at risk of capture'. If it leaves the study area it is considered 'not at risk of capture'. Animals that are at risk of capture at time (i), leave the study area with probability $(1 - F_i)$. Thus F_i has the same interpretation as in Burnham's (1993) live-dead model as the fidelity to the study area. Animals not at risk of capture are permitted to return to the study area with probability F'_i. In Barker (1997) F'_i was the probability that an animal out of the study area at (i) remained out of the study area at $(i+1)$, but the definition has been changed in the interest of having a parametrization in common with the robust design model (see Chapter 16).

constraints	model
$F'_i = 0$	permanent emigration
$F_i = 1, F' = 0^a$	random emigration
$R_i = R'_i = 0, F'_i = 0$	Burnham's (1993) model under permanent emigration
$R_i = R'_i = 0, F_i = 1, F'_i = 0$	Burnham's (1993) model under random emigration
$r_i = R_i - R'_i = 0, F_i = 1, F'_i = 0$	Cormack-Jolly-Seber model (CJS)
$p_i = 0, R_i = R'_i = 0, F_i = 1, F'_i = 0$	Model M_1 ($\{S_t r_t\}$) from Brownie *et al.* (1985)

Under this parametrization there are 3 types of emigration:

Random $(F'_i = F_i)$, **Permanent** $(F'_i = 0)$, & **Markov** (no constraint.)

A complication is that in the random emigration model the parameters $F_i = F_i'$ are confounded with the capture probability p_{i+1}. By making the constraint $F_i = F_i' = 1$ in **MARK** the random emigration model is fitted, but now the interpretation of p_i is the joint probability that an animal is at risk of capture and is caught, $F_{i-1}p_i$.

Under Markov emigration there tends to be serious confounding of the movement and capture probabilities. In a model with time-dependent capture probabilities, it is usually necessary to constrain $F_i = F$. and $F_i' = F'$. for all i. Even then, the Markov emigration model may perform poorly. In practice the parameters F and F' are usually estimable only if the movement model is markedly different from the random emigration model, that is, if there is a large difference between F_i and F_i'.

To illustrate the meaning of the emigration parameters, suppose the animal is captured during the first trapping session, not captured during the second trapping session, and then captured during the third trapping session. One of several encounter histories (again, using the **LD** history format – more on formatting Barker model histories later in this section) that would demonstrate this scenario would be: '100010'. The probability of observing this particular encounter history can be broken into 4 factors:

$P_1 = $ Pr(animal survives from time 1 to time 3 | released at 1)

$P_2 = $ Pr(animal is not resighted between 1 and 3 | released at 1 and survives to 3)

$P_3 = $ Pr(animal is not captured at 2 but is captured at 3 | released at 1 and survives from 1 to 3 without being resighted)

$P_4 = $ Pr(encounter history after trapping period 3 | events up to trapping period 3)

For describing movement, the relevant factor is P_3. An animal captured at time 1 is known to be at risk of capture at time 1. Because it was captured at time 3 we also know it was at risk of capture at time 3.

There are two possible histories that underlie this observed history:

1. The animal was at risk of capture at time 2 and was not captured, but was captured at time 3

2. The animal left the study area between time 1 and 2 but then returned and was captured.

Because we do not know which one actually occurred we instead find the probability that it was either of the two, which is:

$$P_3 = \left[(1 - F_1)F_2' + F_1(1 - p_2)F_2\right]p_3.$$

The complicated term in the square brackets represents the probability that the animal was not captured during the second trapping session but is at risk of capture at time 3. The first product within the brackets $(1 - F_1)F_2'$ is the joint probability that the animal emigrated between the first 2 trapping sessions (with probability $1 - F_1$) and then immigrated back onto the study area during the interval between the second and third trapping sessions (with probability F_2').

However, a second possibility exists for why the animal was not captured – it could have remained on the study area and not been captured. The term F_1 represents the probability that it remained on the study area between time 1 and 2 and the term $(1 - p_2)$ is the probability that it was not captured at time 2. The final term F_2 represents the probability that the animal remained on the study area so that it was available for capture during the third trapping session.

9.6.1. Formatting encounter histories for the Barker model

Encounter histories for the Barker model are coded as '`LDLD`'. Because animals can be encountered in this model as either alive or dead during the interval between capture occasions (see figure at the start of section 9.5), 2 different codes are required in the encounter histories to provide information:

- A '1' in the **D** portion of an encounter history means that the animal was reported dead during the interval. A '2' in the **D** portion of an encounter history means that the animal was reported alive during the interval.

- A '1' in the **L** portion of an encounter history means that the animal was alive on the study area during a live capture occasion.

The following are valid encounter histories for a 5-occasion example: '`1010101002`'. The animal was captured on the first occasion, and recaptured again on the 2nd, 3rd, and 4th occasions. It was not captured on the 5th occasion, but was seen alive during the last interval.

Consider the history '`0000120100`'. In this case, the individual was captured on the 3rd occasion, and seen alive during the 3rd interval. It was reported dead during the 4th interval. Note that there can be multiple occasions with a '1' in the **L** columns, and multiple occasions with a '2' in the **D** columns, but only one **D** column can have a '1'.

9.7. Live encounters, dead recoveries & multi-state models

The multi-state model with live and dead encounters is a generalization of the multi-state model that allows inclusion of recoveries of marks from dead animals. Multi-state models are covered in great detail in the next chapter (10), but we'll introduce some of the basic concepts here.

Suppose you are conducting a study of some harvested species that is either marked and/or recovered (dead) in any one of three discrete geographical locations. Let the locations (stratum) have the less-than-inspired names of 'region A', 'region B', and 'region C'. Each individual, given that it is alive, will do so on one of these three islands. You mount survey operations in all three locations simultaneously. On each occasion, you record encounters with marked individuals, the type of encounter (live, or dead), and in which region the individual was encountered. The information contained in these encounter data allow us to not only estimate survival, but movement/fidelity among regions.

9.7.1. Barker model: assumptions

In addition to the usual assumptions of the multi-state model, this model assumes that apart from group and time effects, the reporting rate of marks from dead animals depends only on the stratum (location) that the animal was in at the immediately preceding live-capture occasion. In some applications, it may be reasonable to also assume that the state of the animal at the time of the dead recovery can be used to determine the state of the animal at the previous live-recapture occasion. This assumption is not included in the model so any such information is ignored.

Model structure and likelihood

If there are S strata, define the following:

$\qquad \varphi_h$ is an $(S \times S)$ matrix with (s,t)'th element = Pr(animal alive at time (h) in stratum (s)
$\qquad\qquad$ is alive at time $(h + 1)$ in stratum (t)

ψ_h is an $(S \times S)$ matrix of transition probabilities with (s,t)'th element = Pr(animal moves from (s) to (t) | alive at (h) and $(h + 1)$

P_h is an $(S \times 1)$ matrix with s'th element = Pr(animal alive at time (h) in stratum (s) is captured)

S_j is an $(S \times 1)$ vector with s'th element = Pr(animal alive at time (j) in stratum (s) is alive at time $(j + 1)$)

r_j is an $(S \times 1)$ vector with (s)'th element = Pr(animal in stratum s that dies between (j) and $(j + 1)$ is found and reported)

$\mathbf{D}(x)$ is a diagonal matrix with vector $\mathbf{x}$ along the diagonal, 1 = a $(s \times 1)$ vector of ones

Y_h is an indicator variable that = 1 if the animal was caught at time h, and 0 otherwise.

Note that $\varphi_h = D(Sh)\psi_h$. The animals in the study can be categorized according to whether their last encounter was as a live recapture or as a dead recovery

Animals last encountered by dead recovery

For an animal first released in stratum (s) at time (i), that was found dead between samples (j) and $(j + 1)$, and was last captured alive at in stratum (t) at time (k) the likelihood, conditional on the first release, is factored into two parts:

1. Pr(encounter history between (i) and (including) (k) | first released at time (i) in stratum (s) is the s,t'th element of the matrix formed by taking the product from $h = i$ to $h = (k-1)$:

$$\prod \psi_h \varphi_h \mathbf{D}(P_{h+1}) + (1 - Y_h)\varphi_h D(1 - P_{h+1}).$$

We take the (s,t)'th element because we know that the animal was in stratum s at time (i) and in stratum (t) at time (k).

2. Pr(not caught between (k) and (including) (j) and found dead between (j) and $(j + 1)$ | released at time (k) in stratum (t) is the sum across the (t)'th row of the matrix formed by taking the product from $h = k$ to $h = (j - 1)$:

$$\prod \varphi_h \mathbf{D}(1 - P_{h+1})\mathbf{D}(1 - S_j)\mathbf{D}(r_j).$$

Although we know that the animal was in stratum (t) at time (k), we do not know which stratum the animal was in at time (j). However it must have been in one of the strata and therefore we can find the probability we require by taking the sum across the (t)'th row of this matrix.

Animals last encountered by live recapture

For an animal first released in stratum (s) and sample (i) and last encountered by live-recapture in stratum (t) and sample (j), the likelihood, conditional on the first release, is factored into the two parts:

1. Pr(encounter history between (i) and (including) (j) | first released at time (i) in stratum (s) is the (s,t)'th element of the matrix

$$\varphi_{j-1}D(P_j) \prod Y_h \varphi_h D(P_{h+1}) + (1 - Y_h)\varphi_h D(1 - P_{h+1}),$$

where the product is taken from $h = i$ to $h = (j - 2)$.

    2. Pr(not encountered again | released alive at (j) in stratum (t). This is found by finding the probability that the animal (i) encountered at least once after sample (j) using the above expressions, and then subtracting this probability from 1.

Parameter identifiability

If the capture occasions are indexed up to sample t and the dead recovery occasions up to sample l, then in addition to the parameters that can be estimated using the multi-strata model, we can also estimate ψ_{t-1}, P_t, S_{t-1} and r_j ($j = 1, ..., t$). If $l > t$ then (complicated) confounded products of stratum-specific survival and reporting probabilities can also be estimated.

9.8. Summary

One of the recent trends in analysis of data from marked individuals is the increasing focus on using data from a variety of sources. In this chapter, we've looked at simultaneously using data from live encounters and dead recoveries. However, the principles are general – we could also use live encounters combined with known-fate telemetry data, and so on. These developing approaches will increasingly allow us to address several interesting questions which previously were not possible when data from only a single source was used. Next up, multi-state models. . .

9.9. References

Barker, R. J. (1997) Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics*, **53**, 666-677.

Barker, R. J. (1999) Joint analysis of mark-recapture, resighting and ring-recovery data with age-dependence and marking-effect. *Bird Study*, **46**, 82-91.

Brownie, C., Anderson, D. R., Burnham, K. P., and Robson, D. S. (1985) *Statistical Inference from Band-Recovery Data – A Handbook*, 2nd Edition. Resource Publication No. 156. Washington, D.C. – Fish & Wildlife Service, U.S. Department of the Interior.

Burnham, K. P. (1993) A theory for combined analysis of ring recovery and recapture data. In '*Marked Individuals in the Study of Bird Population*' (J.-D. Lebreton & P. M North, Eds) – Birkhäuser-Verlag, Basel