CHAPTER 18

Mark-resight models

Brett McClintock, *NOAA National Marine Mammal Laboratory, Seattle, Washington, USA*

Mark-resight methods constitute a slightly different type of data than found in traditional mark-recapture, but they are based on the same fundamental principle of explicitly accounting for imperfect detection towards reliably estimating demographic parameters (see White & Shenk 2001 for a thorough explanation of how these data are collected, and McClintock *et al.* 2009b, McClintock & White 2009, and McClintock & White 2012 for full details of the models). Like the other mark-recapture models in **MARK**, this approach models encounters (resightings) of marked individuals, but they also incorporate additional data via sightings of unmarked individuals into the estimation framework. Mark-resight data may be used to estimate abundance (N) in a fashion analogous to the closed capture models of Otis *et al.* (1978) (see also Chapter 14). When sampling is under the robust design, mark-resight data may be used to estimate abundance, apparent survival, and transition probabilities between observable and unobservable states in a fashion analogous to the closed capture robust design models of Kendall, Pollock & Brownie (1995) and Kendall, Nichols & Hines (1997) – see also Chapter 15.

These models assume some individuals have been marked prior to sampling, and sampling occasions consist of sighting surveys (instead of capture periods). The main advantage of this approach is that because costs associated with marking and recapturing can be minimized, it can in many circumstances be a less invasive and less expensive alternative to traditional mark-recapture as a means for monitoring. With limited funds and resources, mark-resight can be appealing to researchers because costs associated with capture are generally the most expensive aspects of mark-recapture studies. Not only can the financial burden of mark-recapture be discouraging for long-term population monitoring, but capture is also the most hazardous aspect for the animals and may unduly influence the attributes of scientific interest. If field-readable marks are feasible, mark-resight can substantially reduce stress to species because they can be observed at a distance with minimal disturbance after the initial marking period. This can be of particular concern when working with threatened, endangered, or sensitive species.

The methods require that the number of marked individuals in the population during sampling be known exactly or can at least be reliably estimated. If sampling during sighting occasions is without replacement (i.e., any single individual may only be sighted once per distinct occasion) and the number of marked individuals in the population available for resighting is known exactly, then the mixed logit-normal mark-resight model (McClintock *et al.* 2009b) may be employed to estimate N. If the mixed logit-normal model is appropriate but the population of interest within the study area is known to lack geographic closure (e.g., from telemetry data for the marked population), the immigration-emigration logit-normal model may be used to estimate N (or density). Alternatively, if sampling within sighting occasions is with replacement or the exact number of marked individuals in the population is unknown, the Poisson-log normal mark resight model (McClintock *et al.* 2009a) may be used to estimate N. If permanent field-readable marks are used but the number of marks is not known, then mark-resight data collected under the closed robust design may be analyzed with the Poisson-log normal model in a fashion analogous to the regular mark-recapture robust design (Chapter 15) for estimating apparent survival (φ), transition probabilities between observable and unobservable states (γ'' and γ'), and N (McClintock & White 2009).

These models were developed as reliable and more efficient alternatives to the mark-resight models previously available in Program **NOREMARK** (White 1996). Similar to the mark-recapture models in **MARK**, they provide a framework for information-theoretic model selection and multimodel inference based on AIC (Burnham & Anderson 2002) and the utilization of individual or environmental covariates on parameters. However, because the nature of mark-resight data is somewhat different than that of mark-recapture, a different format for the encounter history files has been developed to address this. Explanations of the various models and their **MARK** encounter history file formats are detailed below. The encounter history and results files referenced here accompany **MARK**. Following the explanations of the models and their **MARK** encounter history files, some general suggestions are provided for performing an analysis with these models in **MARK**. But first, a little more background on mark-resight.

18.1. What is mark-resight?

The basic premise behind mark-resight is fairly simple. First, some field-readable marks are introduced into the population. Then encounter data are collected (via non-invasive sighting surveys) on both the marked ***and*** unmarked individuals in the population. Lastly, the data are analyzed to estimate abundance (N) and/or related demographic parameters (φ, γ', γ''). Pretty simple, right? As usual, the complications lie in the particulars.

Initially, the focus of mark-resight was on utilizing radio-marked individuals to estimate closed population abundance. This dependency on radio-collars arose because of a need to know the exact number of marked individuals in the population. One of the simplest mark-resight models of abundance is the classic Lincoln-Petersen estimator:

$$\hat{N} = \frac{m_1 n_2}{m_2},$$

where m_1 is the number of marked animals in the population, n_2 is the total number of marked and unmarked animals seen, and m_2 is the number of marked animals seen (see Chapter 14 for more details on the Lincoln-Petersen and other closed population abundance estimators). Users of Program **NOREMARK** are probably familiar with other mark-resight models of abundance, such as the joint hypergeometric estimator (Bartmann *et al.* 1987), the Minta-Mangel estimator (Minta & Mangel 1989), the immigration-emigration joint hypergeometric estimator (Neal *et al*. 1993), and Bowden's estimator (Bowden & Kufeld 1995). Arnason, Schwarz & Gerrard (1991) developed a mark-resight model of abundance when the number of marked individuals in the population is unknown. These contributions were the motivation for developing a more general suite of mark-resight estimators that would fit into the flexible modeling framework that **MARK** provides.

There are several things to consider when deciding to use the mark-resight models in **MARK**. As with all mark-recapture studies, a population of interest must first be defined (both in space and time). For starters, we will assume this population is geographically and demographically closed, and abundance for a single period of time is the only item of interest. The simplest issue relevant to mark-resight is whether or not individuals in the population can possess field-readable marks. You're unlikely to use mark-resight on *Peromyscus*, but it has been applied to many different species including ursids, canids, badgers, ungulates, prairie dogs, snail kites, owls, robins, and grouse. Field-readable marks may come in many forms, including collars, bands, paint, dye, or natural patterns. The marks may be temporary (e.g., paint or dye on fur) or permanent, but no (unknown) marks may be lost during the sampling period of interest. An important distinction for the mark-resight models in **MARK** is whether the marks are individually identifiable or not. Much more information (and flexibility) can be attained through the use of individually identifiable marks, particularly if individual sighting probability heterogeneity is of concern. However, this methodology may still be employed if individually identifiable marks are not feasible (e.g., due to species or monetary constraints).

If field-readable marks are possible, then marked individuals must be introduced into the population before any sighting data can be collected. This is typically done via a capture event (but not necessarily). Whatever the marks and however they are introduced, **the most fundamental assumption of mark-resight is that the subset of the population selected for marking is representative of the entire population in terms of sighting probabilities**. A strategy typically employed to satisfy this condition is the use of a different method to randomly select marked individuals than is used for the sighting surveys. This may seem obvious, but mark-resight has often been applied (inappropriately) when the marked population was selected based on sightability.

One must also make sure that the marks themselves do not affect sightability. If the tags on the marked individuals are so eye-catching that they make the marked population much more sightable than the unmarked individuals, then the sighting probabilities will be overestimated (and population size therefore underestimated). For example, suppose fish were marked with bright fluorescent tags on the dorsal side (Fig. 18.1) and sighted from a bridge above. If the bright tags made the marked individuals sightable from depths where unmarked individuals were not, then there would be a problem.

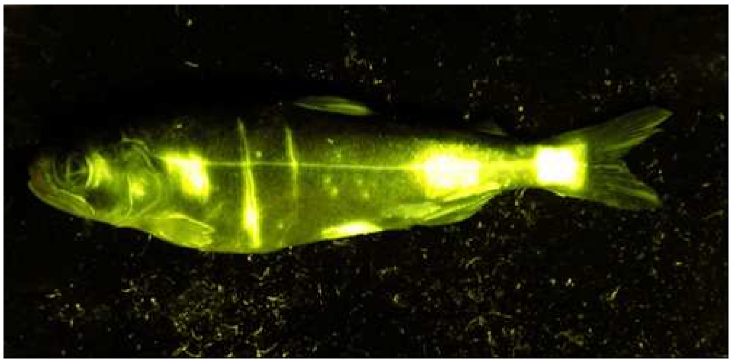

Figure 18.1: *Example of a 'tagged' organism which may be significantly more 'visible' than untagged individuals, which would potentially bias estimates of population size.*

Once marks have been introduced into the population, an important piece of information becomes 'how many marked individuals are alive and in the study area?'. If the number of marked individuals available for resighting is known exactly, this can be very useful information for estimation (particularly when individual sighting heterogeneity is a serious issue). The number of marks in the population is commonly determined via radio or GPS collars that emit a mortality signal. Another way this is accomplished is by conducting the marking period immediately prior to the collection of sighting data, such that it can be reasonably assumed that no marked individuals died or emigrated between the capture event and the sighting surveys. When marked individual mortality or movement cannot be monitored, and sufficient time has passed since the original introduction of marks, then the exact number of marks will usually be unknown.

The actual sighting data are collected during visual surveys within the study area. All sightings of marked and unmarked individuals in the population are recorded. If individually identifiable marks are used, then the individual identities of marked individuals are also recorded. **The sighting surveys themselves come in two basic flavors: sampling with or without replacement**. If sampling is without replacement, then each individual in the population can be seen at most once within each of the distinct sampling occasions (as in mark-recapture). However, in many circumstances sampling must be with replacement. This arises when sampling cannot be divided into distinct occasions where individuals can only be sighted once, such as when studying a highly mobile species or using camera traps. Sampling with replacement differs from other mark-recapture sampling because here sighting occasions need not be distinct, and consideration is given only to some closed period of sampling. When sampling with replacement, particular care must be taken to ensure that all sightings of a marked or unmarked individual are recorded, irrespective of marked status or any previous sightings. For example, suppose one were to stop counting resightings of a marked individual after it had already been resighted during the sampling period, but continued to record all unmarked sightings (because one cannot know whether or not an unmarked individual had been previously sighted during the sampling period). This would violate the must fundamental assumption of mark-resight because the marked individual sighting probabilities would not be representative of the unmarked individual sighting probabilities.

Sighting probabilities are modeled with mark-resight estimators just as capture probabilities are modeled with mark-recapture estimators. This means group, temporal, or individual covariates may be utilized to describe the detection process. Individual sighting heterogeneity is also an important issue because failure to account for it may result in underestimates of abundance (if the number of marks is unknown) and overestimates of precision. Individual heterogeneity may only be accounted for if marks are individually identifiable.

As is the case in most monitoring programs, let's now consider more than a single closed period of interest. We will adopt the terminology of the 'robust design' (Kendall, Pollock & Brownie 1995; Kendall, Nichols & Hines 1997; Chapter 15), where data are collected across both closed and open sampling periods (Fig. 18.2):

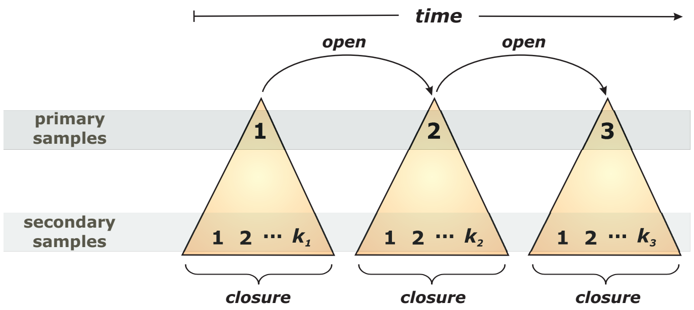

Figure 18.2: *Sampling structure of 'classical' robust design.*

The open periods refer to the encounter process between 'primary' sampling intervals, where each primary interval consists of 'secondary' sighting occasions. The time periods between the secondary sighting occasions within a primary interval must be of short enough duration for the assumption of closure to be satisfied (although this may in some circumstances be relaxed – see the next paragraph). As noted before, if sampling is with replacement, then we are not concerned with distinct secondary sighting occasions, but rather some closed period of secondary sampling during each of the primary intervals. New marks may be introduced to the population at any time during the open periods, but no marks may be added during the closed periods (except when using the immigration-emigration logit-normal model).

The issue of closure deserves a bit of attention before getting into the specifics on implementing the logit-normal, immigration-emigration logit-normal, and Poisson-log normal mark-resight models in **MARK**. When the population of interest is both geographically and demographically closed, then the estimates of abundance produced by all of the mark-resight models are exactly what we think they are: the population size residing within the study area during the period of interest. If the population is not geographically closed (i.e., individuals move in and out of the study area), then there are two notions of 'population' for the study area (Fig. 18.3). There is the population that actually resides within the study area during the period of interest (N), but there is also a 'super population' of individuals associated with the study area during the period of interest (N^*).

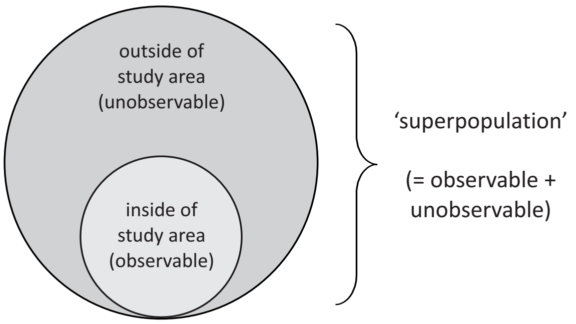

Figure 18.3: *Observable sampled population, as part of larger superpopulation.*

This distinction is important, because the latter is unsuitable for addressing questions related to population density. When geographic closure is violated, then the mixed logit-normal and Poisson-log normal mark-resight models produce estimates of N^*. For this reason, the immigration-emigration logit-normal model was developed as a means for estimating both N and N^* when geographic closure is violated. When demographic closure is violated (i.e., individuals may die or permanently emigrate independent of mark status), all of the models will produce estimates of the population size at the beginning of the sampling period of interest. Because the lack of geographic or demographic closure may induce non-negligible levels of individual sighting heterogeneity, we suggest that heterogeneity models be explored when these violations are suspected (this requires individually identifiable marks).

18.2. The mixed logit-normal mark-resight model

To be used when sampling is (i) without replacement within secondary sampling occasions, and (ii) the number of marked individuals in the population available for resighting is known exactly. Marks may or may not be individually identifiable. See McClintock *et al.* (2009b) for full details.

Data:

t = the number of primary sampling intervals

k_j = the number of secondary sampling occasions (without replacement) during the primary interval j

n_j = the exact number of marked individuals in the population during primary interval j

m_{ij} = $\sum_{s=1}^{n_j} \delta_{sij}$ = total number of marked individual sightings during secondary occasion i of primary interval j

T_{uj} = total number of unmarked individual sightings during primary interval j

δ_{sij} = Bernoulli random variable indicating sighting ($\delta_{sij} = 1$) or no sighting ($\delta_{sij} = 0$) of marked individual s on secondary occasion i of primary interval j (this only applies when individually identifiable marks are used)

ϵ_{ij} = total number of marks seen that were not identified to individual during secondary occasion i of primary interval j (this only applies when individually identifiable marks are used)

Parameters:

N_j = population size or abundance during primary interval j

p_{ij} = intercept (on logit scale) for mean resighting probability of secondary occasion i during primary interval j. If there is no individual heterogeneity ($\sigma_j = 0$), once back-transformed from the logit scale the real parameter estimate can be interpreted as the mean resighting probability

σ_j^2 = individual heterogeneity level (on the logit scale) during primary interval j (i.e., the variance of a random individual heterogeneity effect with mean zero)

Derived Parameter:

μ_{ij} = overall mean resighting probability for secondary occasion i of primary occasion j. This parameter is derived as a function of p_{ij}, σ_j^2, and ϵ_{ij}. Note that when $\sigma_j = 0$ and $\epsilon_{ij} = 0$, then the real parameter estimate for p_{ij} is identical to the derived parameter estimate for μ_{ij}.

18.2.1. No individually identifiable marks

If a known number of marks are in the population, but the marks are not individually identifiable, then the data for the mixed logit-normal model are t, k_j, n_j, m_{ij}, and T_{uj}. These are the same data as for the joint hypergeometric estimator (JHE) previously available in Program **NOREMARK** (White 1996), but the mixed logit-normal model can be a more efficient alternative because it can borrow information about resighting probabilities across primary intervals or groups (McClintock *et al.* 2009b). Note that because no information is known about individual identities, individual heterogeneity models cannot be evaluated with these data (i.e., $\sigma_j = 0$) and the probability of any individual being resighted on secondary occasion i of primary interval j is p_{ij}.

Suppose there is only one group and $t = 3$, $k_j = 4$, $n_1 = 30$, $n_2 = 33$, $n_3 = 32$, $m_{11} = 8$, $m_{21} = 9$, $m_{31} = 10$, $m_{41} = 5$, $m_{12} = 11$, $m_{22} = 10$, $m_{32} = 18$, $m_{42} = 9$, $m_{13} = 5$, $m_{23} = 10$, $m_{33} = 13$, $m_{43} = 8$, $T_{u1} = 96$, $T_{u2} = 68$, and $T_{u3} = 59$.

Although no individual identities are known, these data may be summarized into artificial individual encounter histories similar to those of the mark-recapture robust design. The total number of unmarked individuals seen $\left(T_{uj}\right)$ must be entered after the encounter histories under the heading '`Unmarked Seen Group=1`' such that the resulting encounter history file would be:

```
/* No Individual Marks 1 group */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

111111111111 5;
111011110111 3;
```

```
011011110110 1;
001011100110 1;
000010100010 1;
000000100010 2;
000000100000 5;
000000000000 12;
....00000000 2;
....0000.... 1;

Unmarked Seen Group=1;
96 68 59;

/* End Input File */
```

Let's look a bit more closely at the input file. First, remember that there are 3 primary sampling periods, each consisting of 4 secondary samples. To make this sampling scheme somewhat more 'visually obvious', we'll modify the encounter histories by putting a space between each primary sampling session:

```
/* No Individual Marks 1 group */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

1111 1111 1111 5;
1110 1111 0111 3;
0110 1111 0110 1;
0010 1110 0110 1;
0000 1010 0010 1;
0000 0010 0010 2;
0000 0010 0000 5;
0000 0000 0000 12;
.... 0000 0000 2;
.... 0000 .... 1;

Unmarked Seen Group=1;
96 68 59;

/* End Input File */
```

How do we generate the 'artificial encounter histories'? The process is actually straightforward. First, recall that $n_1 = 30$, $n_2 = 33$, $n_3 = 32$, where n_j is the exact number of marked individuals in the population during primary interval j. Now, consider the first primary sample. Recall that for the first primary sample, $m_{11} = 8$, $m_{21} = 9$, $m_{31} = 10$, $m_{41} = 5$, where m_{ij} is the total number of marked individual sightings during secondary occasion i of primary interval j. A little bit of mental math will demonstrate that $(8+9+10+5) \neq 30$. This is because individuals can be sighted on more than one secondary occasion (sampling with replacement among secondary samples, but not within an individual secondary sample). The same logic applies to each of the other primary periods.

Now, if you look at the encounter histories closely, you'll notice that the sums of the encounter history columns (when multiplied by the corresponding history frequency) equal m_{ij}, and the sums of the frequencies with non-missing entries (i.e., not '....') for each primary interval equals n_j.

In fact, this might give you a clue as to how you might generate the artificial encounter histories. The easiest, albeit somewhat laborious way to do it is to create a file (a matrix, actually) consisting of 12 columns, one for each of the sighting samples (3 primary × 4 secondary = 12), and 33 rows. Why 33 rows? Because 33 is the largest number of marked individuals observed during any one primary session (for this example, $n_2 = 33$). Now, simply fill in the matrix, entering a '1' for each marked individual sighted, and a '0' for each individual 'missed' (not sighted). How many '1' and '0' values need to be entered? Simple – for each column ij, you enter m_{ij} '1's.

So, given $m_{11} = 8$, $m_{21} = 9$, $m_{31} = 10$, $m_{41} = 5$, $m_{12} = 11$, $m_{22} = 10$, $m_{32} = 18$, $m_{42} = 9$, $m_{13} = 5$, $m_{23} = 10$, $m_{33} = 13$, $m_{43} = 8$, then simply enter m_{ij} '1' values for the corresponding column, starting from the top in each case:

```
1111 1111 1111
1111 1111 1111
1111 1111 1111
1111 1111 1111
1111 1111 1111
1110 1111 0111
1110 1111 0111
1110 1111 0111
0110 1111 0110
0010 1110 0110
0000 1010 0010
0000 0010 0010
0000 0010 0010
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
```

Next, we have to remember that for some primary sampling occasions, the number of marked individuals is < 33. Since each row corresponds to a potential marked individual, then for those primary periods where the number of marked individuals is < 33, some rows correspond to individuals which don't actually exist. To account for that, we'll use a '...' notation. Meaning, we'll enter a dot ('.') for unobserved or missing information. Recall that $n_1 = 30$ (so, $(33 - 30) = 3$ rows corresponding to 'missing individuals' for primary period 1), and $n_3 = 32$ (so, $(33 - 32) = 1$ row corresponding to 'missing individuals' for primary period 3).

Thus, our matrix would be modified as follows (look closely at the last 3 rows):

```
1111 1111 1111
1111 1111 1111
1111 1111 1111
1111 1111 1111
1111 1111 1111
1110 1111 0111
1110 1111 0111
1110 1111 0111
0110 1111 0110
0010 1110 0110
0000 1010 0010
0000 0010 0010
0000 0010 0010
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0010 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
0000 0000 0000
.... 0000 0000
.... 0000 0000
.... 0000 ....
```

Next, simply calculate the frequency (number) of the individual histories. If you look closely, you'll see there are 5 of the '`1111 1111 1111`' histories, 3 of the '`1110 1111 0111`' histories, and so on.

We take these calculated frequencies, and re-write the histories in summary form, removing the blank spaces/columns we inserted to make the sampling intervals more obvious, and put the encounter frequencies in a column to the right of the history (as is the standard format for **MARK** input files):

```
111111111111 5;
111011110111 3;
011011110110 1;
001011100110 1;
000010100010 1;
000000100010 2;
000000100000 5;
000000000000 12;
....00000000 2;
....0000.... 1;
```

This should look familiar – it is simply a piece of the larger input file (below), corresponding to the artificial histories for the marked individuals:

```
/* No Individual Marks 1 group */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

111111111111 5;
111011110111 3;
011011110110 1;
001011100110 1;
000010100010 1;
000000100010 2;
000000100000 5;
000000000000 12;
....00000000 2;
....0000.... 1;

Unmarked Seen Group=1;
96 68 59;

/* End Input File */
```

The final step simply involves entering the number of unmarked individuals, as shown above.

Constructing the `.INP` file based on 'artificial sighting histories' takes a bit of practice, but if you work through the preceding example, it shouldn't be too tough to grasp the basic idea.*

As a test of your understanding, consider the case if these single group data (above) were split into two groups, such that $n_1 = 17$, $n_2 = 19$, $n_3 = 18$, $m_{11} = 6$, $m_{21} = 6$, $m_{31} = 7$, $m_{41} = 4$, $m_{12} = 5$, $m_{22} = 5$, $m_{32} = 11$, $m_{42} = 5$, $m_{13} = 3$, $m_{23} = 7$, $m_{33} = 7$, $m_{43} = 7$, $T_{u1} = 48$, $T_{u2} = 40$, and $T_{u3} = 20$ for the first group, and $n_1 = 13$, $n_2 = 14$, $n_3 = 14$, $m_{11} = 2$, $m_{21} = 3$, $m_{31} = 3$, $m_{41} = 1$, $m_{12} = 6$, $m_{22} = 5$, $m_{32} = 7$, $m_{42} = 4$, $m_{13} = 2$, $m_{23} = 3$, $m_{33} = 6$, $m_{43} = 1$, $T_{u1} = 48$, $T_{u2} = 28$, and $T_{u3} = 39$ for the second group, a possible encounter history file would be:

```
/* No Individual Marks 2 groups */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

111111111111 3 0;
111111110111 1 0;
111011110111 1 0;
111000100111 1 0;
001000100111 1 0;
000000100000 4 0;
000000000000 6 0;
....00000000 1 0;
....0000.... 1 0;
111111111111 0 1;
111011111110 0 1;
```

* Since creating `.INP` files by hand is not a good use of your time, an example **R** script for creating the `.INP` files based on 'artificial encounter histories' is provided in the addendum. Tweak/adjust/improve as needed...

```
011011110110 0 1;
000011110010 0 1;
000011100010 0 1;
000010100010 0 1;
000000100000 0 1;
000000000000 0 6;
....00000000 0 1;

Unmarked Seen Group=1;
48 40 20;

Unmarked Seen Group=2;
48 28 39;
/* End Input File */
```

Notice here that the single group data has simply been split up into two group data. The encounter histories are followed by group frequencies just as in other **MARK** encounter history files for mark-recapture data. The twist is that the unmarked data must be entered separately for each group. Again, the sums of the encounter history columns (when multiplied by the corresponding group frequencies) equals m_{ij} for each group, and the sums of the frequencies with non-missing entries (i.e., not '....') for each primary interval equals n_j for each group.

The analysis using these encounter history data (`Logit_NoIndividualMarks_OneGroup.inp`) yielded the following results for the time-constant ($p_{ij} = p, \sigma_j = 0$) model in **MARK**:

```
                   Real Function Parameters of {p(.) sigma(.)=0 N(t)}

                                                            95% Confidence Interval
   Parameter                 Estimate       Standard Error      Lower           Upper
 -------------------------  --------------  --------------  --------------  --------------
   1:p Session 1             0.3064700       0.0236970       0.2620778       0.3547665
   2:sigma Session 1         0.0000000       0.0000000       0.0000000       0.0000000      Fixed
   3:N Session 1             108.02874       8.9593461       92.350040       127.65004
   4:N Session 2             88.188211       7.0136070       76.062792       103.72785
   5:N Session 3             79.846656       6.3659903       68.905724       94.031094
```

Note that σ_j must be fixed to zero for these data because heterogeneity models do not apply when marks are not individually identifiable. This is because no information is known about individual resighting probabilities, and the above encounter histories are artificial in that they don't actually refer to a real individual's encounter history (these artificial encounter histories are just a convenient and consistent way to enter the data into **MARK**). Because there is no individual heterogeneity in the model, the real parameter estimate of p may be interpreted as the overall mean resighting probability (0.31 in this case).

18.2.2. Individually identifiable marks

If marks are individually identifiable, encounter histories are constructed just as for robust design mark-recapture data with the tk_j possible encounters representing δ_{sij} for individual s during secondary occasion i of primary interval j. In other words, the encounter history is identical to the standard robust design mark-recapture encounter history introduced earlier (Chapter 15).

However, in the mark-resight context, it is possible to have an individual identified as marked, but not to individual identity. A marked individual may be encountered but not be identified to individual when the mark was seen but the unique pattern or characters that identify the individual were obscured or too far away to read.

These are the same data as could be used for Bowden's estimator (Bowden & Kufeld 1995) in Program **NOREMARK** (White 1996), but the logit-normal model can be more efficient because information about resighting probabilities may be borrowed across primary intervals, and it does not require investment in individual heterogeneity parameters unless deemed necessary by the data (McClintock *et al.* 2009b). If an individual was not known to be in the population during any primary interval j, then missing values (.) are included for all k_j secondary occasions of that interval in the encounter history. The total number of marks seen but not identified to individual during secondary occasion i of primary interval j (ϵ_{ij}) are entered sequentially $\left(\epsilon_{11}, \epsilon_{21}, \ldots, \epsilon_{k_1 1}, \ldots,\ \epsilon_{1t}, \epsilon_{2t}, \ldots, \epsilon_{k_t t}\right)$ with each entry separated by a space.

Using the data from the previous single group example but with ϵ = (0,0,0,0,1,1,1,0,0,3,0,1) entered after the unmarked data under the heading '**Marked Unidentified Group=1;**', one possible encounter history file would be:

```
/* Individual Marks 1 Group */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

001001000011 1;
000000100110 1;
010000000110 1;
0000........ 1;
....01101101 1;
000010000000 1;
001100100000 1;
001011100011 1;
000010000010 1;
010001100000 1;
000000000010 1;
001010010110 1;
101000100000 1;
....01001110 1;
010000100000 1;
11001000.... 1;
000100000000 1;
100000101011 1;
000011010000 1;
000100000000 1;
111000100001 1;
010000111001 1;
101000110000 1;
100001100010 1;
....00010000 1;
101000010010 1;
0000........ 1;
010000101000 1;
000110100000 1;
```

```
011000000000 1;
010011110010 1;
000010110000 1;
101100000001 1;
....00010110 1;
....11100100 1;

Unmarked Seen Group=1;
96 68 59;

Marked Unidentified Group=1;
0 0 0 0 1 1 1 0 0 3 0 1;

/* End Input File */
```

Note that the sums of each column $\sum_{s=1}^{n_j} \delta_{sij} = (m_{ij} - \epsilon_{ij})$. The last two encounter histories are for individuals that were not marked and known to be in the population until immediately prior to the second primary interval. The fourth encounter history from the top represents an individual who was marked and known to be in the population during the first primary interval (when it was resighted 0 times), but known to have not been marked and in the population during the second or third primary intervals. This could be because the individual was known to have died, emigrated, or lost its mark. Similar to other **MARK** encounter history files, the histories may pertain to multiple groups and include individual covariates.

Splitting the above data into two groups, the above encounter history file could look like:

```
/* Individual Marks 2 Groups */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

001001000011 0 1;
000000100110 1 0;
010000000110 1 0;
0000........ 1 0;
....01101101 1 0;
000010000000 0 1;
001100100000 1 0;
001011100011 0 1;
000010000010 0 1;
010001100000 0 1;
000000000010 0 1;
001010010110 1 0;
101000100000 1 0;
....01001110 1 0;
010000100000 1 0;
11001000.... 1 0;
000100000000 1 0;
100000101011 1 0;
000011010000 1 0;
000100000000 0 1;
111000100001 1 0;
010000111001 0 1;
```

```
101000110000 1 0;
100001100010 0 1;
....00010000 0 1;
101000010010 0 1;
0000........ 0 1;
010000101000 0 1;
000110100000 1 0;
011000000000 1 0;
010011110010 1 0;
000010110000 0 1;
101100000001 1 0;
....00010110 1 0;
....11100100 0 1;

Unmarked Seen Group=1;
48 40 20;

Unmarked Seen Group=2;
48 28 39;

Marked Unidentified Group=1;
0 0 0 0 0 1 1 0 0 1 0 1;

Marked Unidentified Group=2;
0 0 0 0 1 0 0 0 0 2 0 0;

/* End Input File */
```

Notice the encounter histories are followed by group frequencies the same way as they are in all other **MARK** encounter history files.

Because marks are individually identifiable, individual heterogeneity models may be explored with these data. Here, individual heterogeneity is modeled as a random effect with mean zero and unknown variance σ_j^2. These encounter history data (`Logit_IndividualMarks_OneGroup.inp`) yielded the following results for the time-constant individual heterogeneity $(p_{ij} = p, \sigma_j = \sigma)$ model in **MARK**:

```
                    Real Function Parameters of {p(.) sigma(.) N(t)}

                                                                    95% Confidence Interval
Parameter                  Estimate       Standard Error      Lower           Upper
-------------------------  --------------  --------------  --------------  --------------
   1:p Session 1           0.2786641       0.0273014       0.2284108       0.3351710
   2:sigma Session 1       0.4766088       0.2707817       0.1690244       1.3439241
   3:N Session 1           112.97626       10.415916       94.940988       136.02025
   4:N Session 2           87.429921       6.9734104       75.386318       102.89558
   5:N Session 3           77.935945       6.0515938       67.521842       91.403200
```

If one wanted to report an overall mean resighting probability for this model, then the derived parameter μ_{ij} ('`Mu-hat`') may be obtained (top of the next page).

```
                     Estimates of Derived Parameters
              Mean Resighting Rate Estimates of {p(.) sigma(.) N(t)}
```

```
                                                 95% Confidence Interval
Grp. Occ.      Mu-hat       Standard Error      Lower           Upper
---- ----   --------------  --------------  --------------  --------------
  1    1    0.2880297       0.0247720       0.2420014       0.3388985
  1    2    0.2880297       0.0247720       0.2420014       0.3388985
  1    3    0.2880297       0.0247720       0.2420014       0.3388985
  1    4    0.2880297       0.0247720       0.2420014       0.3388985
  1    5    0.3183328       0.0247720       0.2718623       0.3687242
  1    6    0.3183328       0.0247720       0.2718623       0.3687242
  1    7    0.3183328       0.0247720       0.2718623       0.3687242
  1    8    0.2880297       0.0247720       0.2420014       0.3388985
  1    9    0.2880297       0.0247720       0.2420014       0.3388985
  1   10    0.3817797       0.0247720       0.3345418       0.4313640
  1   11    0.2880297       0.0247720       0.2420014       0.3388985
  1   12    0.3192797       0.0247720       0.2727964       0.3696574
```

Even though the model included a constant p and σ for all occasions, there is some slight variation in μ_{ij} due to marked individuals not being identified to individual identity (ϵ_{ij}) on several occasions.

The time-constant model with no heterogeneity ($p_{ij} = p, \sigma_j = 0$) yields:

```
                    Real Function Parameters of {p(.) sigma(.)=0 N(t)}

                                                              95% Confidence Interval
Parameter                  Estimate       Standard Error      Lower           Upper
-------------------------  --------------  --------------  --------------  --------------
   1:p Session 1           0.2881305       0.0232879       0.2447124       0.3358270
   2:sigma Session 1       0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   3:N Session 1           112.98732       9.7939840       95.902227       134.50170
   4:N Session 2           87.446686       6.5355052       76.068609       101.83068
   5:N Session 3           77.954031       5.6720754       68.112315       90.477916

                     Estimates of Derived Parameters
             Mean Resighting Rate Estimates of {p(.) sigma(.)=0 N(t)}

                                                 95% Confidence Interval
Grp. Occ.      Mu-hat       Standard Error      Lower           Upper
---- ----   --------------  --------------  --------------  --------------
  1    1    0.2881305       0.0232879       0.2447124       0.3358270
  1    2    0.2881305       0.0232879       0.2447124       0.3358270
  1    3    0.2881305       0.0232879       0.2447124       0.3358270
  1    4    0.2881305       0.0232879       0.2447124       0.3358270
  1    5    0.3184336       0.0232879       0.2746235       0.3657090
  1    6    0.3184336       0.0232879       0.2746235       0.3657090
  1    7    0.3184336       0.0232879       0.2746235       0.3657090
  1    8    0.2881305       0.0232879       0.2447124       0.3358270
  1    9    0.2881305       0.0232879       0.2447124       0.3358270
  1   10    0.3818805       0.0232879       0.3373910       0.4284434
  1   11    0.2881305       0.0232879       0.2447124       0.3358270
  1   12    0.3193805       0.0232879       0.2755591       0.3666438
```

As before, when $\sigma_j = 0$, the real parameter estimate of p may be interpreted as the overall mean resighting probability ignoring unidentified marks (0.29 in this case), but μ_{ij} is an overall mean resighting probability that takes unidentified marks into account.

Notice that these results are different than the results from the same model when there were no individually identifiable marks. This is because the two versions (individually identifiable marks or not) of the mixed-logit normal model are only comparable when all marks are correctly identified to individual and σ_j is fixed to zero. Further, if one finds very little support for individual heterogeneity models (based on AIC$_c$) and has relatively many unidentified marks, it may be better to analyze the data as if there were no individually identifiable marks to begin with.

18.3. The immigration-emigration mixed logit-normal model

For use when the population of interest may not be geographically closed (i.e., individuals move in and out of the study area between secondary occasions of the primary sampling intervals). Because the study area is not closed, there is a 'super population' of individuals that use the area, but the population of interest may be that which actually resides within the study area at any given time (see Fig. 18.3). This distinction is important when density estimation is of concern. This model requires additional information on whether or not each marked individual was available for resighting within the study area for each secondary sampling occasion (e.g., from radio or GPS collars). One way this is commonly determined using radio-collars is by conducting an aerial survey locating all marked individuals immediately prior to each secondary sampling occasion, although the use of GPS collars may alleviate the need for such surveys.

Once the presence or absence of each marked individuals within the study area is determined, secondary resighting occasions are conducted only within the boundaries of the study area. As with the regular mixed logit-normal model, sampling must be without replacement within secondary sampling occasions, and the number of marked individuals in the population available for resighting must be known exactly for every secondary sampling occasion. Marks may or may not be individually identifiable (but individually identifiable marks are needed to investigate individual heterogeneity).

Unlike the regular mixed logit-normal or the Poisson-log normal models (where new marks may be introduced only during the open periods), new marks may be introduced at any time (other than *during* a secondary sampling occasion) when using the immigration-emigration mixed logit-normal model. See McClintock & White (2012) for full details.

Data:

t = the number of primary sampling intervals

k_j = the number of secondary sampling occasions (without replacement) during primary interval j

n_j = the exact number of marked individuals in the population during primary interval j

$m_{ij} = \sum_{s=1}^{n_j} \delta_{sij}$ = total number of marked individual sightings during secondary occasion i of primary interval j

T_{uij} = total number of unmarked individual sightings during secondary occasion i of primary interval j

δ_{sij} = Bernoulli random variable indicating sighting ($\delta_{sij} = 1$) or no sighting ($\delta_{sij} = 0$) of marked individual s on secondary occasion i of primary interval j (this only applies when individually identifiable marks are used)

T_{ij} = number of marked animals in the 'super population' during secondary occasion i of primary interval j. A marked individual is considered to be in the super population if it were located within the study area at least once during primary interval j

M_{ij} = number of marked animals that are actually in the study area during secondary occasion i of primary interval j

ϵ_{ij} = total number of marks seen that were not identified to individual during secondary occasion i of primary interval j (this only applies when individually identifiable marks are used)

Parameters:

N_j^* = 'super-population' size utilizing the study area at any time during primary interval j

$\bar{N}_j$ = mean population size within the study area during primary interval j. Because this quantity is generally of more interest (e.g., for density estimation) than the population size within the study area during secondary occasion i of primary interval j (N_{ij}), **MARK** uses the reparameterization $N_{ij} = \bar{N}_j + \alpha_{ij}$ where $\sum_{i=1}^{k_j} \alpha_{ij} = 0$

α_{ij} = the difference (relative to $\bar{N}_j$) in population size within the study area during secondary occasion i of primary interval j. Because of the imposed constraint $\sum_{i=1}^{k_j} \alpha_{ij} = 0$, only $(k_j - 1)$ of the α_{ij} must actually be estimated for primary interval j.

p_{ij} = intercept (on logit scale) for mean resighting probability of secondary occasion i during primary interval j. If there is no individual heterogeneity ($\sigma_j = 0$), once back-transformed from the logit scale the real parameter estimate can be interpreted as the mean resighting probability

σ_j^2 = individual heterogeneity level (on the logit scale) during primary interval j (i.e., the variance of a random individual heterogeneity effect with mean zero)

Derived Parameter:

μ_{ij} = overall mean resighting probability for secondary occasion i of primary interval j. This parameter is derived as a function of $p_{ij}, \sigma_j^2, M_{ij}$, and ϵ_{ij}. Note that when $\sigma_j = 0$ and $\epsilon_{ij} = 0$, then the real parameter estimate for p_{ij} is identical to the derived parameter estimate for μ_{ij}.

18.3.1. No individually identifiable marks

If a known number of marks are in the population, but the marks are not individually identifiable, then the data for the immigration-emigration mixed logit-normal model are $t, k_j, T_{ij}, M_{ij}, m_{ij}$, and T_{uij}. These are the same data as for the immigration-emigration joint hypergeometric estimator (IEJHE) previously available in Program **NOREMARK** (White 1996), but the immigration-emigration mixed logit-normal model can be a more efficient alternative because it can borrow information about resighting probabilities across primary intervals. Note that because no information is known about individual identities, individual heterogeneity models cannot be evaluated with these data (i.e., $\sigma_j = 0$) and the probability of any individual being resighted on secondary occasion i of primary interval j is p_{ij}.

Here we'll use vector notation because we must keep track of data for each secondary occasion of each primary interval, where any $\boldsymbol{x} = \{x_{11}, x_{21}, \ldots, x_{k_1 1}, x_{12}, x_{22}, \ldots, x_{k_2 2}, \ldots, x_{1t}, x_{2t}, \ldots, x_{k_t t}\}$.

Now, suppose there is only one group and

$$
\begin{aligned}
t &= 3, k_j = 4 \\
\boldsymbol{n} &= \{27, 22, 18, 29, 28, 23, 20, 32, 31, 19, 21, 33\} \\
\boldsymbol{T} &= \{28, 29, 30, 30, 30, 33, 33, 33, 33, 34, 34, 34\} \\
\boldsymbol{m} &= \{17, 15, 9, 8, 16, 14, 9, 13, 11, 14, 13, 16\} \\
\boldsymbol{T_u} &= \{264, 161, 152, 217, 217, 160, 195, 159, 166, 152, 175, 190\}
\end{aligned}
$$

These data show that marks were introduced into the population between secondary sampling occasions at some point for all three primary intervals. For example, one mark was introduced between the first ($T_{11} = 28$) and second ($T_{21} = 29$) secondary occasions of the first primary interval. Of these marked individuals in the super population using the study area, $n_{11} = 27$ and $n_{21} = 22$ marked individuals, respectively, were actually in the study area during these secondary sighting occasions of the first primary interval.

As before, these data may be summarized into artificial individual encounter histories similar to those of the mark-recapture robust design. Now, both the number of marked animals in the super population ($\boldsymbol{T}$) and the total number of unmarked individuals seen ($\boldsymbol{T_u}$) during each secondary occasion must be entered after the encounter histories under the headings '**`Marked Superpopulation Group=1`**' and '**`Unmarked Seen Group=1`**' such that the resulting encounter history file would be:

```
/* No Individual Marks 1 Group */
/* 12 occasions, 3 primary, 4 secondary each */

/* Begin Input File */

111111111111        8;
111011111111        1;
110011011111        2;
110011010111        2;
110011000101        1;
110010000001        1;
100010000001        1;
100000000000        1;
000000000000        1;
00.000000000        1;
00.000000.00        1;
00.000.00.00        1;
00.000.00..0        1;
0..000.00..0        1;
0..00..00..0        4;
...00..00..0        1;
...0...00..0        1;
.......00..0        2;
.......0...0        1;
...........0        1;

Marked Superpopulation Group=1;
28 29 30 30 30 33 33 33 33 34 34 34;

Unmarked Seen Group=1;
```

```
264 161 152 217 217 160 195 159 166 152 175 190;

/* End Input File */
```

Notice the sums of the encounter history columns (when multiplied by the corresponding frequency) equal m_{ij}, and the sums of the non-missing entries (i.e., not '.') for each column equal n_{ij}. If these two conditions are satisfied, then the data have been correctly manipulated into artificial encounter histories.

With no individually identifiable marks, only the parameters p_{ij}, $\bar{N}_j$, α_{ij}, and N^*_j should be estimated, and σ_j needs to be fixed to zero. The analysis using the encounter history data contained in (`IELNE_NoIndividualMarks.inp`) yielded the the following results for the fully time- and session-dependent model in **MARK**:

```
Real Function Parameters of {p(t*session) Nbar(session) alpha(t*session) Nstar(session)}
                                                               95% Confidence Interval
    Parameter                Estimate       Standard Error      Lower           Upper
 -------------------------  --------------  --------------  --------------  --------------
    1:p Session 1           0.5920048       0.0656170       0.4600336       0.7119197
    2:p Session 1           0.5336830       0.0849114       0.3696301       0.6907604
    3:p Session 1           0.5334205       0.0791715       0.3799858       0.6807828
    4:p Session 1           0.4660226       0.0528644       0.3652849       0.5696100
    5:sigma Session 1       0.0000000       0.0000000       0.0000000       0.0000000    Fixed
    6:Nbar Session 1        397.32353       37.779741       345.08949       499.99984
    7:alpha Session 1       77.488465       23.018541       32.372124       122.60481
    8:alpha Session 1       -67.485508      35.691739       -137.44132      2.4703031
    9:alpha Session 1       -95.435499      32.604481       -159.34028      -31.530716
   10:Nstar Session 1       494.93170       21.786836       460.75810       547.81022
   11:p Session 2           0.5299575       0.0619746       0.4091072       0.6473942
   12:p Session 2           0.5553905       0.0782237       0.4016923       0.6991743
   13:p Session 2           0.6222315       0.0715188       0.4756901       0.7493929
   14:p Session 2           0.3737708       0.0454746       0.2896954       0.4662306
   15:sigma Session 2       0.0000000       0.0000000       0.0000000       0.0000000    Fixed
   16:Nbar Session 2        385.24069       35.290100       331.28861       472.92637
   17:alpha Session 2       54.472940       22.564238       10.247032       98.698847
   18:alpha Session 2       -71.850164      29.657658       -129.97918      -13.721152
   19:alpha Session 2       -57.222270      25.778240       -107.74762      -6.6969177
   20:Nstar Session 2       475.22043       20.428033       443.38777       525.10066
   21:p Session 3           0.4143881       0.0495080       0.3217547       0.5134995
   22:p Session 3           0.7132910       0.0934712       0.5038977       0.8590295
   23:p Session 3           0.6569720       0.0793330       0.4899648       0.7924578
   24:p Session 3           0.4701110       0.0539388       0.3672353       0.5755906
   25:sigma Session 3       0.0000000       0.0000000       0.0000000       0.0000000    Fixed
   26:Nbar Session 3        346.12174       29.879666       300.03823       419.77196
   27:alpha Session 3       80.815452       21.159041       39.343731       122.28717
   28:alpha Session 3       -113.10159      28.082685       -168.14366      -58.059529
   29:alpha Session 3       -59.723073      25.125391       -108.96884      -10.477305
   30:Nstar Session 3       452.02721       20.950900       418.31832       501.69599
```

Here the mean population size using the study area during the first primary interval was $\hat{\bar{N}}_1 = 397.3$. The total 'super population' size associated with the study area during the first primary interval was $\hat{N}^*_1 = 494.9$. The estimates for α suggest the population within the study area fluctuated, with $\hat{N}_{11} = \hat{\bar{N}}_1 + \hat{\alpha}_{11} = 474.8$, $\hat{N}_{21} = \hat{\bar{N}}_1 + \hat{\alpha}_{21} = 329.8$, $\hat{N}_{31} = \hat{\bar{N}}_1 + \hat{\alpha}_{31} = 301.9$, and $\hat{N}_{31} = \hat{\bar{N}}_1 - \sum_{i=1}^{k_1-1} \hat{\alpha}_{i1} = 482.8$.

Suppose temporary emigration from the study area during primary interval j is constant and completely random. In this case, the expected population size within the study area doesn't change despite the fact that individuals freely move in and out.

Using these same data, this hypothesis may be explored by fixing $\alpha_{ij} = 0$ ($i = 1, \ldots, k_j - 1$) in **MARK**:

```
Real Function Parameters of {p(t*session) N(session) Nstar(session)}
                                                          95% Confidence Interval
    Parameter                Estimate       Standard Error      Lower           Upper
-------------------------  --------------  --------------  --------------  --------------
    1:p Session 1            0.6439324       0.0636854       0.5120143       0.7571063
    2:p Session 1            0.4029707       0.0440082       0.3204665       0.4913568
    3:p Session 1            0.3684690       0.0411356       0.2920867       0.4520701
    4:p Session 1            0.5153590       0.0531874       0.4119457       0.6174735
    5:sigma Session 1        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
    6:Nbar Session 1         436.60462       40.188783       376.61052       538.64137
    7:alpha Session 1        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
    8:alpha Session 1        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
    9:alpha Session 1        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   10:Nstar Session 1        532.11187       24.630622       494.78594       593.80284
   11:p Session 2            0.6078401       0.0608881       0.4844027       0.7188762
   12:p Session 2            0.4536115       0.0486381       0.3610701       0.5494736
   13:p Session 2            0.5320235       0.0548963       0.4246068       0.6365519
   14:p Session 2            0.4483689       0.0482075       0.3567993       0.5435784
   15:sigma Session 2        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   16:Nbar Session 2         383.50015       34.946464       330.10908       470.38476
   17:alpha Session 2        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   18:alpha Session 2        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   19:alpha Session 2        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   20:Nstar Session 2        480.30601       23.005205       444.82403       537.00999
   21:p Session 3            0.4922281       0.0511459       0.3936081       0.5914567
   22:p Session 3            0.4615808       0.0488057       0.3684460       0.5574769
   23:p Session 3            0.5228970       0.0535107       0.4185439       0.6252887
   24:p Session 3            0.5730778       0.0573259       0.4588866       0.6799768
   25:sigma Session 3        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   26:Nbar Session 3         359.57710       31.935979       309.89200       437.67227
   27:alpha Session 3        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   28:alpha Session 3        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   29:alpha Session 3        0.0000000       0.0000000       0.0000000       0.0000000     Fixed
   30:Nstar Session 3        466.75871       23.909342       429.24320       524.77071
```

When fixing $\alpha_{ij} = 0$, the model may still be used to estimate both the super population size (N_j^*) and the population size within the study area $\bar{N}_j = N_{ij}$ ($i = 1, \ldots, k_j$). For these data, however, the AIC$_c$ evidence strongly favors the previous model (ΔAIC$_c$= 67.4!).

18.3.2. Individually identifiable marks

As with the regular mixed logit-normal model with individually identifiable marks, the encounter histories are constructed with tk_j possible encounters representing δ_{sij} for individual s during secondary occasion i of primary interval j. If an individual is not yet marked or a marked individual is outside of the study area during secondary occasion i of primary interval j, then missing values (.) are included for that occasion in the encounter history. As before, the total number of marks seen but not identified to individual during secondary occasion i of primary interval j (ϵ_{ij}) are also entered into the encounter history file.

Using the same data from the previous example with one group and

$$
\begin{aligned}
t &= 3, k_j = 4 \\
\mathbf{n} &= \{27, 22, 18, 29, 28, 23, 20, 32, 31, 19, 21, 33\} \\
\mathbf{T} &= \{28, 29, 30, 30, 30, 33, 33, 33, 33, 34, 34, 34\} \\
\mathbf{m} &= \{17, 15, 9, 8, 16, 14, 9, 13, 11, 14, 13, 16\} \\
T_u &= \{264, 161, 152, 217, 217, 160, 195, 159, 166, 152, 175, 190\} \\
\epsilon &= \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}
\end{aligned}
$$

one possible encounter history file incorporating individually identifiable marks would be:

```
/* Individual Marks 1 Group */
/* 12 occasions, 3 primary, 4 secondary each */
/* Marked individuals off the study area or not yet marked indicated by "." */
/* in encounter history */

/* Begin Input File */

11010..00100        1;
.........1.0        1;
11.110110110        1;
0011.1001..1        1;
...00.1001.1        1;
0.0000.0000.        1;
1.11111001.0        1;
0..111.11..1        1;
0110000011.0        1;
111011111.11        1;
1111110101.1        1;
1110..100.11        1;
00.000000000        1;
1.001..10.00        1;
111011.0..11        1;
11.01..01..1        1;
.00010011110        1;
.....11011.1        1;
01.01.000.01        1;
000011101010        1;
110.10.11111        1;
1..00.0011.0        1;
11.010010..0        1;
......000111        1;
.01001010.10        1;
11.011.10100        1;
11.101110.10        1;
011000.10.00        1;
00.001.00001        1;
100000.000.0        1;
```

```
0..00.10...1        1;
1.011..11.11        1;
110011.00.10        1;
.....10.0111        1;

Marked Superpopulation Group=1;
28 29 30 30 30 33 33 33 33 34 34 34;

Unmarked Seen Group=1;
264 161 152 217 217 160 195 159 166 152 175 190;

Marked Unidentified Group=1;
0 0 0 0 0 0 0 0 0 0 0 0;

/* End Input File */
```

Note that the sums of each column $\sum_{s=1}^{n_{ij}} \delta_{sij} = (m_{ij} - \epsilon_{ij})$. The first encounter history describes a marked individual that was in the super population of marked individuals ($\boldsymbol{T}$) during all three primary intervals. This individual was outside the study area on the second and third secondary occasions of the second primary interval. The second encounter history from the top describes an individual that was not in the marked super population during the first and second primary intervals. This individual may not have been marked until sometime during the third primary interval or it may have already been marked but didn't use the study area during the first or second primary intervals. Either way, it's not included in T_{i1} or T_{i2}. We avoid needing to distinguish between these two possibilities in the encounter history by providing **MARK** with the known values for all T_{ij} under '**`Marked Superpopulation Group=1`**.'

Because marks are individually identifiable, individual heterogeneity models may be explored with these data. The analysis using these encounter history data (`IELNE_NoIndividualMarks.inp`) yielded the following results for the fully time- and session-dependent model in **MARK**:

```
Real Function Parameters of
   {p(t*session) sigma(session) Nbar(session) alpha(t*session) Nstar(session)}

                                                           95% Confidence Interval
 Parameter                  Estimate       Standard Error     Lower           Upper
-------------------------- -------------- -------------- -------------- --------------
    1:p Session 1           0.6133918      0.0910340      0.4278053      0.7710052
    2:p Session 1           0.5616208      0.1194894      0.3310362      0.7683446
    3:p Session 1           0.5413694      0.1033785      0.3429405      0.7274912
    4:p Session 1           0.4574051      0.0730364      0.3213478      0.6001278
    5:sigma Session 1       1.0205809      0.4924485      0.4163080      2.5019584
    6:Nbar Session 1        394.61098      44.452039      337.29368      522.89418
    7:alpha Session 1       79.083762      23.173848      33.663020      124.50450
    8:alpha Session 1       -73.737498     35.569234      -143.45320     -4.0217987
    9:alpha Session 1       -93.058635     32.321863      -156.40949     -29.707783
   10:Nstar Session 1       494.08538      22.358088      459.06840      548.42536
   11:p Session 2           0.5324475      0.0802870      0.3770380      0.6818062
   12:p Session 2           0.5597357      0.0973820      0.3694818      0.7339233
   13:p Session 2           0.6357481      0.0878476      0.4534913      0.7859170
   14:p Session 2           0.3512142      0.0600856      0.2440611      0.4758014
   15:sigma Session 2       0.9087787      0.4205942      0.3832277      2.1550599
   16:Nbar Session 2        387.54071      40.782499      327.25659      492.00581
   17:alpha Session 2       54.176870      22.768204      9.5511885      98.802551
```

```
  18:alpha Session 2       -71.817931     29.415765      -129.47283     -14.163031
  19:alpha Session 2       -56.822380     26.126235      -108.02980     -5.6149591
  20:Nstar Session 2       477.60634      21.044542      445.03053      529.30608
  21:p Session 3           0.3978764      0.0894500      0.2411951      0.5787139
  22:p Session 3           0.8554374      0.1033849      0.5347726      0.9682158
  23:p Session 3           0.7842213      0.1143219      0.4915720      0.9317948
  24:p Session 3           0.4882646      0.0998619      0.3035819      0.6762078
  25:sigma Session 3       1.7968113      0.6252497      0.9262340      3.4856536
  26:Nbar Session 3        329.55557      33.183455      281.69496      416.44233
  27:alpha Session 3       79.218362      20.744845      38.558465      119.87826
  28:alpha Session 3       -110.37310     25.613884      -160.57631     -60.169884
  29:alpha Session 3       -58.532733     22.476479      -102.58663     -14.478834
  30:Nstar Session 3       432.33206      21.148660      398.57525      482.84169
```

For the model ignoring individual heterogeneity (i.e., where σ_j is fixed to 0):

```
Real Function Parameters of {p(t*session) Nbar(session) alpha(t*session) Nstar(session)}

                                                          95% Confidence Interval
 Parameter                  Estimate       Standard Error   Lower          Upper
--------------------------  --------------  --------------  --------------  --------------
    1:p Session 1          0.5920050      0.0656170      0.4600338      0.7119199
    2:p Session 1          0.5336829      0.0849114      0.3696300      0.6907603
    3:p Session 1          0.5334203      0.0791715      0.3799856      0.6807828
    4:p Session 1          0.4660228      0.0528644      0.3652850      0.5696102
    5:sigma Session 1      0.0000000      0.0000000      0.0000000      0.0000000     Fixed
    6:Nbar Session 1       397.32348      37.779732      345.08945      499.99977
    7:alpha Session 1      77.488323      23.018552      32.371960      122.60469
    8:alpha Session 1      -67.485442     35.691773      -137.44132     2.4704343
    9:alpha Session 1      -95.435342     32.604530      -159.34022     -31.530462
   10:Nstar Session 1      494.93154      21.786834      460.75796      547.81007
   11:p Session 2          0.5299575      0.0619746      0.4091071      0.6473944
   12:p Session 2          0.5553905      0.0782238      0.4016923      0.6991744
   13:p Session 2          0.6222314      0.0715188      0.4756901      0.7493929
   14:p Session 2          0.3737708      0.0454746      0.2896953      0.4662307
   15:sigma Session 2      0.0000000      0.0000000      0.0000000      0.0000000     Fixed
   16:Nbar Session 2       385.24068      35.290114      331.28859      472.92641
   17:alpha Session 2      54.472932      22.564274      10.246954      98.698911
   18:alpha Session 2      -71.850162     29.657677      -129.97921     -13.721115
   19:alpha Session 2      -57.222263     25.778259      -107.74765     -6.6968749
   20:Nstar Session 2      475.22042      20.428040      443.38775      525.10067
   21:p Session 3          0.4143880      0.0495079      0.3217546      0.5134994
   22:p Session 3          0.7132910      0.0934712      0.5038977      0.8590295
   23:p Session 3          0.6569718      0.0793330      0.4899648      0.7924576
   24:p Session 3          0.4701108      0.0539388      0.3672353      0.5755904
   25:sigma Session 3      0.0000000      0.0000000      0.0000000      0.0000000     Fixed
   26:Nbar Session 3       346.12180      29.879668      300.03828      419.77202
   27:alpha Session 3      80.815496      21.159034      39.343789      122.28720
   28:alpha Session 3      -113.10166     28.082677      -168.14371     -58.059612
   29:alpha Session 3      -59.723082     25.125383      -108.96883     -10.477330
   30:Nstar Session 3      452.02731      20.950896      418.31842      501.69607
```

The interpretation of the parameters remains the same as before. In this case, AIC$_c$ lends more support to the model including individual heterogeneity ($\Delta \text{AIC}_c = 7.0$). Notice that because all $\epsilon_{ij} = 0$ for these data, the estimates from the no-heterogeneity model with individually identifiable marks are the same as those for the same model when there were no individually identifiable marks.

18.4. The Poisson-log normal mark-resight model

For use when (i) the number of marked individuals in the population may be unknown, or (ii) sampling is with replacement within secondary sampling occasions (or there is no concept of a distinct secondary sampling occasion without replacement). Marks must be individually identifiable. See McClintock *et al.* 2009a and McClintock & White 2009 for full details.

Data:

t = the number of primary sampling intervals (through time, groups, or time and groups)

n_j = the exact number of marked individuals in the population during primary interval j

n_j^* = total number of marked individuals resighted at least once and known to be in the population

c_j = total number of individuals captured (e.g., for marking) immediately prior to primary interval j and therefore assumed to be present in the population during primary interval j, but not resighted during primary interval j

$c_j^* = n_j^* + c_j$ = total number of marked individuals captured immediately prior to primary interval j or resighted at least once during primary interval j. When the number of marks is known exactly, $c_j^* = n_j$. When the number of marks is unknown this is the minimum number of marked individuals known to be in the population

y_{sj} = Poisson random variable for the total number of times individual s was resighted during primary interval j

ϵ_j = total number of times an individual was sighted and identified as marked, but not identified to individual identity during primary interval j

T_{uj} = total unmarked individual sightings during primary interval j

Parameters:

U_j = number of unmarked individuals in the population during primary interval j

α_j = intercept (on log scale) for mean resighting rate during primary interval j. If there is no individual heterogeneity ($\sigma_j = 0$), once back-transformed from the log scale the real parameter estimate can be interpreted as the mean resighting rate for the entire population

σ_j^2 = individual heterogeneity level (on the log scale) during primary interval j, i.e., the additional variance due to a random individual heterogeneity effect with mean zero

φ_j = apparent survival between primary intervals j and $j+1$, $j = \{1, \ldots, t-1\}$

γ_j'' = probability of transitioning from an observable state at time j (e.g., on the study area) to an unobservable state at time $j+1$ (e.g., off the study area), $j = \{1, \ldots, t-1\}$. This is equivalent to transition probability ψ_j^{OU} of Kendall & Nichols (2002) – also, Chapter 15.

γ_j' = probability of remaining at an unobservable state at time $j+1$ (e.g., off the study area) when at an unobservable state at time j, $j = \{2, \ldots, t-1\}$. This is equivalent to $1 - \psi_j^{UO}$ of Kendall & Nichols (2002) – also, Chapter 15.

Derived Parameters:

λ_j = overall mean resighting rate for primary occasion j. This is a parameter derived as a function of α_j, σ_j^2, and ϵ_j. Note that when $\sigma_j = 0$ and $\epsilon_j = 0$, then the real parameter estimate for α_j is identical to the derived parameter estimate for λ_j.

p_j^* = overall probability of being sighted at least once during primary occasion j (see `-sidebar-`, below)

N_j = $U_j + n_j$ = total population size during primary occasion j. This is a *derived* parameter, because **MARK** actually estimates U_j in the model. If n_j is unknown, then N_j is derived as

$$N_j = U_j + \frac{n_j^*}{p_j^*}$$

where n_j^*/p_j^* is the number of marked individuals, and p_j^* is the probability of being sighted at least once within primary sampling occasion j (for details on p^*, and relation to λ, see the following `- sidebar -`).

`begin sidebar`

p^*, and interpreting λ

It is helpful for interpreting λ (a rate) to note that if there is no individual heterogeneity (i.e., $\sigma_j = 0$), then p^* (the probability of being sighted at least once within a primary sampling occasion) is related to λ as $p^* = 1 - e^{-\lambda}$.

If, however, $\sigma_j > 0$, then the approximation $p^* = 1 - e^{-\lambda}$ gets progressively worse as σ_j increases. Formally,

$$p_j^* = 1 - \int_{-\infty}^{\infty} \exp\Big(-\exp(\alpha_j + \exp(\sigma_j)z)\Big)\varphi(z)dz$$

where $\varphi(z)$ is the standard normal density and both σ_j and α_j are on the log scale (same scale as the beta parameters in **MARK**).

Fortunately, this 'nasty looking integral' can be accurately approximated using Gauss-Hermite quadrature:

$$p_j^* \approx 1 - \sum_{m=1}^{M} w_m \frac{\exp\Big(-\exp(\alpha_j + \sqrt{2}\exp(\sigma_j)v_m)\Big)}{\sqrt{\pi}}$$

where (w_m, v_m) are weights and nodes corresponding to M quadrature points.

The estimator for the number of marked individuals is then still $\hat{n} = n_j^*/p_j^*$, and abundance is still estimated as $\hat{N}_j = \exp(\hat{U}_j) + \hat{n}_j$, where $\hat{U}_j$ is on the log scale (which is the same scale as the β parameter in **MARK**). The variance of $\hat{N}$ can be estimated using the Delta method (Appendix B) as

$$\widehat{\text{var}}(\hat{N}_j) \approx \begin{bmatrix} \frac{\partial N_j}{\partial \alpha_j} & \frac{\partial N_j}{\partial \sigma_j} & \frac{\partial N_j}{\partial U_j} \end{bmatrix} \sum_j \begin{bmatrix} \frac{\partial N_j}{\partial \alpha_j} & \frac{\partial N_j}{\partial \sigma_j} & \frac{\partial N_j}{\partial U_j} \end{bmatrix}^{\mathsf{T}}$$

We'll leave it to you as an exercise to derive the partial derivatives in this approximation (actually, they're not too bad, and quite impressive looking when you're finished). Fortunately, **MARK** does all that 'heavy lifting' for you.

`end sidebar`

18.4.1. Closed resightings only

If interest is only in abundance estimates for different groups (or t primary intervals for group(s) with few or no marked individuals in common across the intervals), then the mark-resight Poisson-log normal model may be used in a fashion analogous to the closed mark-recapture models introduced in Chapter 14. Individual covariates may be used in modeling resighting probabilities. However, **because the data consist of the total number of times each marked individual was resighted, the encounter histories must be modified to reflect this different type of encounter data**. If the number of marks is known exactly, then n_j, y_{sj}, ϵ_j and T_{uj} are the same data used for Bowden's estimator (Bowden & Kufeld 1995) in **NOREMARK** (White 1996), but the Poisson-log normal model will often be more efficient because information about resighting probabilities may be borrowed across time or groups (McClintock *et al.* 2009a).

The number of marks available for each of the groups or t primary intervals may be known or unknown. The encounter history file contains individual encounter histories composed of the y_{sj} resightings, the frequencies and group(s) to which each encounter history pertains, the T_{uj} unmarked sightings and group(s) to which they pertain, the ϵ_j unidentified marks and the group(s) to which they pertain, and whether or not the number of marks is known exactly for each group. **Instead of the familiar 0's and 1's of other MARK encounter histories, these histories simply contain the y_{sj} for each marked individual** s. Two character spaces are allocated to allow $y_{sj} > 9$. Note that this coding does not allow $y_{sj} > 99$. For reasons that will become clear in the next section covering the robust design Poisson-log normal model, **entries for which $y_{sj} = 0$ are entered using '+0' instead of '00'**. Further, (unlike the logit-normal model and mark-recapture robust design), because the Poisson-log normal model does not condition on distinct secondary resighting occasions, **the number of encounter occasions entered into MARK when creating a new analysis is the number of primary occasions**.

For instance, suppose in a very simple example that there were two groups and $t = 1$ primary interval with known $n_1 = 3$, $y_{11} = 2$, $y_{21} = 3$, $y_{31} = 0$, $T_{u_1} = 11$, and $\epsilon_1 = 2$ for the first group, and $n_1 = 3$, $y_{11} = 0$, $y_{21} = 0$, $y_{31} = 12$, $T_{u_1} = 5$, and $\epsilon_1 = 3$ for the second group. The encounter history file would be:

```
/* Poisson log-normal mark-resight */
/* Occasions=1 groups=2 */

/* Begin Input File */

02 1 0;
03 1 0;
+0 1 0;
+0 0 1;
+0 0 1;
12 0 1;

Unmarked Seen Group=1;
11;

Unmarked Seen Group=2;
5;

Marked Unidentified Group=1;
2;

Marked Unidentified Group=2;
```

```
3;

Known Marks Group=1;
3;


Known Marks Group=2;
3;

/* End Input File */
```

The columns following the encounter histories are the frequencies for the two groups, just as would be done in other **MARK** encounter history files. Under '**`Unmarked Seen`**', the T_{uj} are entered separately for each group. The '**`Marked Unidentified`**' data (ϵ_j) are entered in the same fashion separately for each group. Similarly, the '**`Known Marks`**' headings contain the n_j for each group.

Using the same example, but now with the number of marks being unknown for the second group, the encounter history file must be modified to reflect that n_2 is unknown and $y_{s2} = 0$ is no longer observed:

```
/* Poisson log-normal mark-resight */
/* occasions=1 groups=2 */

/* Begin Input File */

02 1 0;
03 1 0;
+0 1 0;
12 0 1;

Unmarked Seen Group=1;
11;

Unmarked Seen Group=2;
5;

Marked Unidentified Group=1;
2;

Marked Unidentified Group=2;
3;

Known Marks Group=1;
3;

Known Marks Group=2;
0;

/* End Input File */
```

Here, the encounter histories for $y_{12} = 0$ and $y_{22} = 0$ have been removed because they cannot be observed if the number of marked individuals in the population (n_2) is unknown. Further, under '`Known Marks;`' there is now a '`0`' for the second group. **By including a '`0`' for the second group's '`Known Marks`',**

MARK knows the number of marks is unknown and will use the zero-truncated Poisson-log normal model.

It is possible that the number of marks may be unknown for a given group, but some marking was conducted immediately prior to the primary sampling interval of interest. Here, some additional information is known about the minimum number of marks in the population because those (previously marked or newly marked) individuals captured during the marking period are known to have been present and available for resighting (even if they were not resighted during the interval of interest).

For example, suppose it's 2015 and we're conducting a study of bighorn sheep. A student had previously conducted a telemetry study on the same population in 2011, so those radios are no longer transmitting, but some unknown number of those individuals are still alive and marked, and potentially sightable. At the onset of the 2015 study, there are n_{old} individuals that still have those marks from 2011. Immediately prior to the primary sampling interval in 2015, we decide to introduce some additional marks, say $n_{new} = 20$. The total number of marks (n) is unknown in 2015, because $n = n_{old} + n_{new}$, although there is clearly an upper bound on n, since n_{old} can't be greater than the number of marks in the 2011 study. When entering '**Known Marks**' in the `.INP` file, there are two options: (1) enter n if the exact number of marks is known; or (2) enter '`0`' to indicate that the exact number of marks is unknown. If '`Known Marks=0`', then the encounter histories are used to tally up the minimum number of marks known to be in the population ($c^* = n_{new}$+ any of the n_{old} individuals sighted at least once), and things proceed from there.

Suppose this sort of situation were the case in the above example, such that the second individual of the second group was captured and marked immediately prior to resighting surveys but never resighted. This information (although not used in the zero-truncated likelihood) may be included in the encounter history file to make the lower bound for $N_2 \geq c_2^*$:

```
/* Poisson log-normal mark-resight */
/* occasions=1 groups=2 */

/* Begin Input File */

02 1 0;
03 1 0;
+0 1 0;
+0 0 1;
12 0 1;

Unmarked Seen Group=1;
11;

Unmarked Seen Group=2;
5;


Marked Unidentified Group=1;
2;

Marked Unidentified Group=2;
3;

Known Marks Group=1;
3;
```

```
Known Marks Group=2;
0;

/* End Input File */
```

Because the '**Known Marks;**' is still '0' for the second group, **MARK** knows the actual number of marks is unknown and to use the zero-truncated model for the second group, but $c_2^* = 2$ (instead of $n_2^* = 1$) will be used in establishing the lower bound for N_2. When the number of marks is unknown, the information provided by such encounters via capture events will become more useful when considering the robust design Poisson-log normal model in the next section.

Now to analyze a more realistic data set where the number of marks was known for the first group but not for the second. No marking occurred immediately prior to resighting surveys for the second group, so $c_2^* = n_2^*$, and therefore no '+0' encounter histories are included for the second group. For the first group, $n_1 = 60$, $T_{u_1} = 1{,}237$, and $\epsilon_1 = 10$. For the second group, $n_1^* = 33$, $T_{u_1} = 588$, and $\epsilon_1 = 5$:

```
/* Poisson log-normal mark-resight */
/* Occasions=1 groups=2 */

/* Begin Input File */

02 1 0;
03 1 0;
03 1 0;
01 1 0;
01 1 0;
01 1 0;
02 1 0;
09 1 0;
05 1 0;
01 1 0;
01 1 0;
01 1 0;
03 1 0;
03 1 0;
02 1 0;
06 1 0;
04 1 0;
02 1 0;
03 1 0;
01 1 0;
02 1 0;
01 1 0;
03 1 0;
04 1 0;
03 1 0;
03 1 0;
05 1 0;
03 1 0;
04 1 0;
04 1 0;
+0 1 0;
04 1 0;
```

```
01 1 0;
03 1 0;
02 1 0;
01 1 0;
03 1 0;
02 1 0;
03 1 0;
05 1 0;
06 1 0;
03 1 0;
01 1 0;
04 1 0;
07 1 0;
03 1 0;
+0 1 0;
06 1 0;
+0 1 0;
04 1 0;
+0 1 0;
02 1 0;
02 1 0;
02 1 0;
02 1 0;
05 1 0;
02 1 0;
01 1 0;
04 1 0;
+0 1 0;
02 0 1;
02 0 1;
04 0 1;
01 0 1;
02 0 1;
01 0 1;
01 0 1;
01 0 1;
04 0 1;
03 0 1;
01 0 1;
05 0 1;
02 0 1;
02 0 1;
05 0 1;
02 0 1;
01 0 1;
05 0 1;
01 0 1;
02 0 1;
07 0 1;
01 0 1;
03 0 1;
05 0 1;
03 0 1;
```

```
03 0 1;
04 0 1;
02 0 1;
03 0 1;
05 0 1;
02 0 1;
02 0 1;
02 0 1;

Unmarked Seen Group=1;
1237;

Unmarked Seen Group=2;
588;

Marked Unidentified Group=1;
10;

Marked Unidentified Group=2;
5;

Known Marks Group=1;
60;

Known Marks Group=2;
0;

/* End Input File */
```

The analysis for these data (`Poisson_TwoGroups.inp`) yielded the following results for the most general model:

```
                   Real Function Parameters of {alpha(g)sigma(g)U(g)}

                                                                95% Confidence Interval
   Parameter                Estimate       Standard Error      Lower           Upper
 -------------------------  --------------  --------------  --------------  --------------
   1:alpha                  2.6274189       0.2483643       2.1839589       3.1609248
   2:alpha                  2.3834952       0.3632005       1.7711208       3.2076012
   3:sigma                  0.2782579       0.1405534       0.1093112       0.7083213
   4:sigma                  0.2316744       0.2787288       0.0362715       1.4797580
   5:U                      426.66770       37.155745       359.83441       505.91416
   6:U                      227.09486       29.801418       175.78405       293.38314
```

In most situations, these real parameter estimates may not be of interest. The derived parameters for abundance (N) and mean resighting rate (λ) are typically what we want:

```
                     Estimates of Derived Parameters
             Population Estimates of {alpha(g)sigma(g)U(g)}

                                                  95% Confidence Interval
```

```
Grp. Occ.   N-hat            Standard Error     Lower            Upper
---- ----   --------------   --------------   --------------   --------------
  1     1   486.66770        37.155822        419.12029        565.10136
  2     1   263.21721        30.821410        209.40169        330.86314

        Mean Resighting Rate Estimates of {alpha(g)sigma(g)U(g)}

                                                95% Confidence Interval
Grp. Occ.   Lambda-hat       Standard Error     Lower            Upper
---- ----   --------------   --------------   --------------   --------------
  1     1   2.8977973        0.2355306        2.4716992        3.3973507
  2     1   2.5867444        0.3200561        2.0315747        3.2936257
```

Here are the results for the model with no group effects on α_j or σ_j:

```
                   Real Function Parameters of {alpha(.)sigma(.)U(g)}

                                                              95% Confidence Interval
   Parameter                 Estimate       Standard Error      Lower            Upper
-------------------------  --------------  --------------  --------------  --------------
   1:alpha                 2.5449662       0.2037816       2.1758646       2.9766800
   2:sigma                 0.2670036       0.1248112       0.1117130       0.6381611
   3:U                     440.94680       32.590191       381.55642       509.58148
   4:U                     211.45044       17.316388       180.14242       248.19966



                     Estimates of Derived Parameters
             Population Estimates of {alpha(.)sigma(.)U(g)}

                                                95% Confidence Interval
Grp. Occ.   N-hat            Standard Error     Lower            Upper
---- ----   --------------   --------------   --------------   --------------
  1     1   500.94680        32.590259        441.03409        568.99842
  2     1   246.99366        17.749865        214.58185        284.30115

        Mean Resighting Rate Estimates of {alpha(.)sigma(.)U(g)}

                                                95% Confidence Interval
Grp. Occ.   Lambda-hat       Standard Error     Lower            Upper
---- ----   --------------   --------------   --------------   --------------
  1     1   2.8039855        0.1882162        2.4586823        3.1977840
  2     1   2.7779927        0.1902567        2.4294158        3.1765839
```

Here are the results for the model with no group effect on α_j and $\sigma_j = 0$:

```
                 Real Function Parameters of {alpha(.)sigma(.)=0 U(g)}

                                                              95% Confidence Interval
   Parameter                 Estimate       Standard Error      Lower            Upper
-------------------------  --------------  --------------  --------------  --------------
   1:alpha                 2.6488895       0.1731506       2.3306735       3.0105529
```

```
  2:sigma                 0.0000000       0.0000000       0.0000000       0.0000000      Fixed
  3:U                     439.16724       29.754643       384.61298       501.45959
  4:U                     210.59709       15.810414       181.81833       243.93104


                    Estimates of Derived Parameters
          Population Estimates of {alpha(.)sigma(.)=0 U(g)}

                                                   95% Confidence Interval
Grp. Occ.    N-hat           Standard Error      Lower           Upper
---- ----    --------------  --------------  --------------  --------------
  1     1    499.16724       29.754705       444.17194       560.97181
  2     1    246.10883       16.203557       216.34382       279.96896


     Mean Resighting Rate Estimates of {alpha(.)sigma(.)=0 U(g)}

                                                   95% Confidence Interval
Grp. Occ.    Lambda-hat      Standard Error      Lower           Upper
---- ----    --------------  --------------  --------------  --------------
  1     1    2.8155562       0.1731506       2.4961207       3.1758707
  2     1    2.7896881       0.1750062       2.4672233       3.1542988
```

Note that to run models without individual heterogeneity, σ_j must be fixed to zero. When $\sigma = 0$, the real parameter estimate of α may be interpreted as the overall mean resighting rate ignoring unidentified marks, but λ is an overall mean resighting rate that takes unidentified marks into account.

18.4.2. Full-likelihood robust design

If interest is in apparent survival (φ), transition probabilities between observable and unobservable states (γ' and γ''), and abundance (N) for one or more groups through time, then a mark-resight robust design analogous to the mark-recapture robust design of Kendall, Pollock & Brownie (1995) and Kendall, Nichols & Hines (1997) may be employed (see Chapter 15). Full details on the mark-resight robust design model may be found in McClintock & White (2009). In contrast to the modeling of recapture probabilities in the mark-recapture robust design utilizing the full likelihood closed capture models of Otis *et al.* (1987), the mark-resight robust design may incorporate individual covariates in modeling resighting probabilities.

The encounter history files are similar to those from the previous '**`Closed Resightings`**' model (section 18.4.1), but now the open period encounter process for individuals with permanent field-readable marks may be modeled through time across t primary sampling intervals in a robust design. For instance, if an individual s was encountered $y_{s1} = 4$ times during the first primary interval and $y_{s2} = 2$ times during the second primary interval, then the encounter history would be '`0402`'. Each encounter history over t primary samples will contain $2t$ characters, again allowing two characters for each y_{sj}. Because the number of marks can be known or unknown for any given primary interval, the primary intervals must again be identified as such under the '**`Known Marks`**' heading in the encounter history file. In the individual encounter histories, **a '`+0`' indicates that the individual was known to be a marked individual available for resighting during primary interval j but never resighted**. Therefore, when the number of marks is unknown, the total number of '`+0`' entries during primary interval j is equal to c_j as defined above. **A '`-0`' indicates a previously encountered individual that was not encountered**

(via capture or resighting) during primary interval j, and only applies when the number of marks is unknown (i.e., when the number of marks is known a '`-0`' is impossible).

Lastly, a '..' indicates a marked individual who has not yet been encountered prior to and during primary interval j, or an individual that is known to no longer be in the marked population (due to removal, mortality, or permanent emigration) during and after primary interval j. As in the regular CJS model in **MARK**, any '..' contributes no information to the estimation of parameters. When n_j is known, '`+0`' contributes information towards estimation of survival, transition probabilities, resighting probabilities, and abundance. When n_j is unknown, '`+0`' contributes information towards estimating survival and transition probabilities, but makes no contribution to the estimation of resighting probabilities or abundance (but it does affect the minimum lower bound for N_j as described in the previous section). A '`-0`' contributes no information to the estimation of resighting probabilities or abundance (it is only a valid entry when the number of marks is unknown), and is equivalent to a '`0`' in the regular CJS encounter history for **MARK**. It therefore only contributes to the estimation of survival and transition probabilities. As before, the encounter histories are followed by group frequencies in the usual **MARK** encounter history file. The entries for '**`Unmarked Seen`**', '**`Marked Unidentified`**', and '**`Known Marks`**' are the same as described earlier and are entered separately for each group.

In the following example encounter history file with a single group and $t = 4$ primary intervals, the number of marks are known for the first and second primary intervals, but unknown for the third and fourth. **Because the model does not condition on distinct secondary resighting occasions, the number of encounters that are input into MARK is equal to the number of primary occasions** ($t = 4$ in this case). Capturing for marking occurred immediately prior to the first, second, and third occasion, but not the fourth occasion, so $n_4^* = c_4^*$.

Here, $n_1 = 45$, $T_{u_1} = 1{,}380$, $\epsilon_1 = 8$, $n_2 = 67$, $T_{u_2} = 1{,}120$, $\epsilon_2 = 10$, $n_3^* = 56$, $T_{u_3} = 1{,}041$, $\epsilon_3 = 9$, $n_4^* = 52$, $T_{u_4} = 948$, and $\epsilon_4 = 11$:

```
/* Poisson log-normal Mark-resight -- 4 occasions, 1 group */

/* Begin Input File */

....+002 1;
..06-0-0 1;
04060202 1;
+0010402 1;
070602-0 1;
04020606 1;
..020101 1;
060602-0 1;
..04-004 1;
040401-0 1;
03010103 1;
02030503 1;
..03+0-0 1;
070503-0 1;
04+00104 1;
01010401 1;
06060103 1;
02010602 1;
..0403-0 1;
..020306 1;
020202-0 1;
```

```
..050201 1;
02010103 1;
031002-0 1;
+0+00704 1;
01030102 1;
01010302 1;
..02-0-0 1;
..020210 1;
020301-0 1;
02+00503 1;
02+0+0-0 1;
02020302 1;
..080201 1;
..040603 1;
030304-0 1;
02020202 1;
..030107 1;
04050402 1;
+0050101 1;
..030605 1;
05+00101 1;
..04-003 1;
06020204 1;
..03-004 1;
..010201 1;
04+00303 1;
04040204 1;
01+00201 1;
0403-004 1;
01+00103 1;
..020307 1;
01060701 1;
..040101 1;
03040301 1;
..0404-0 1;
03050101 1;
05040202 1;
03010202 1;
05+00302 1;
01020202 1;
01+0+0-0 1;
01070202 1;
..050105 1;
02040205 1;
02010301 1;
..03-010 1;
..01+0-0 1;

Unmarked Seen Group=1;
1380 1120 1041 948;

Marked Unidentified Group=1;
8 10 9 11;
```

```
 Known Marks Group=1;
 45 67 0 0;

/* End Input File */
```

Let's look at the first 4 encounter histories a bit more closely. Here, we'll add some blank columns in the history to more clearly indicate the 4 different primary sampling periods.

```
.. .. +0 02 1;
.. 06 -0 -0 1;
04 06 02 02 1;
+0 01 04 02 1;
```

The first encounter history indicates this individual was not captured for marking until immediately prior to the third primary occasion, and the '`+0`' for the third sampling period indicates that it was not resighted (although known to be a marked individual available for resighting during this sampling occasion). This individual was then resighted twice during the fourth occasion.

The second encounter history from the top indicates that this individual was only known to be marked and in the population during the second primary occasion (when it was resighted 6 times). Recall that for this example, the number of marks are known for the first and second primary intervals, but unknown for the third and fourth (see below) . So, for this second history, because the number of marks is known for the first primary interval, this individual must have been marked between the first and second primary intervals. As indicated by '`-0`', this individual was never encountered again when the number of marks was unknown during the third and fourth primary intervals.

The third encounter history from the top indicates and individual known to be marked and available for resighting for all $t = 4$ occasions. This individual was resighted four, six, two, and two times during the first, second, third and fourth intervals, respectively.

The fourth encounter history from the top indicates an individual who was known to be marked and available for resighting for all $t = 4$ occasions. The '`+0`' entry for the first primary occasion indicates that it was known to be marked and available for resighting, but never resighted. This individual was then resighted one, four, and two times during the second, third, and fourth intervals, respectively.

Now, let's consider the final encounter history – again, we've added some blank columns in the history to more clearly indicate the 4 different primary sampling periods.

```
.. 01 +0 -0 1;
```

The final encounter history describes an individual that was not marked until immediately prior to the second primary occasion, and during the second occasion it was resighted one time. It was then captured immediately prior to (but never resighted during) the third occasion. Because the number of marks was unknown for the third occasion, this '`+0`' primarily contributes information to the estimation of survival and transition probabilities (as described in the previous section). As indicated by '`-0`' this individual was then never resighted during the fourth occasion (and could not have been captured immediately prior to the occasion because no capturing took place). Because no individuals were captured (e.g., for marking) immediately prior to the fourth occasion (and the number of marked individuals was unknown), no '`+0`' appears in the entries for this occasion. Because no marked individuals were known to have left the population (due to removal, mortality, or permanent emigration), no '`..`' entries appear after an individual's first encounter.

The '**`Unmarked Seen`**' entry tells **MARK** that 1,380 unmarked sightings occurred during the first primary interval, 1,120 during the second, 1,041 during the third, and 948 during the fourth. The '**`Marked`**

`Unidentified`' entry follows the same pattern. The '**`Known Marks`**' entry tells **MARK** that n_j is known for the first and second primary intervals ($n_1 = 46$, $n_2 = 60$), but unknown for the third and fourth (as indicated by '`0`' for these occasions).

As a simple two group example, suppose for the first group that $n_1 = 10$, $T_{u_1} = 800$, $\epsilon_1 = 4$, $n_2 = 14$, $T_{u_2} = 950$, $\epsilon_2 = 2$, $n_3^* = 11$, $T_{u_3} = 500$, $\epsilon_3 = 6$, $n_4^* = 8$, $T_{u_4} = 1201$, and $\epsilon_4 = 3$. For the second group, $n_1 = 11$, $T_{u_1} = 459$, $\epsilon_1 = 2$, $n_2^* = 14$, $T_{u_2} = 782$, $\epsilon_2 = 5$, $n_3^* = 15$, $T_{u_3} = 256$, $\epsilon_3 = 0$, $n_4^* = 11$, $T_{u_4} = 921$, and $\epsilon_4 = 1$. With capturing (e.g., for marking) occurring for both groups immediately prior to the first and second occasions, a possible encounter history file would be:

```
/* Poisson log-normal Mark-resight -- 4 occasions, 2 groups */

/* Begin Input File */
04060202 1 0;
..06-0-0 1 0;
+0010402 1 0;
070602-0 1 0;
04020606 1 0;
..020101 1 0;
060602-0 1 0;
..04-004 1 0;
040401-0 1 0;
03010103 1 0;
02030503 1 0;
..03-0-0 1 0;
070503-0 1 0;
04+00104 1 0;
01010401 0 1;
06060103 0 1;
02010602 0 1;
..0403-0 0 1;
..020306 0 1;
020202-0 0 1;
..050201 0 1;
02010103 0 1;
031002-0 0 1;
+0-00704 0 1;
01030102 0 1;
01010302 0 1;
..02-0-0 0 1;
..020210 0 1;
020301-0 0 1;
02+00503 0 1;

Unmarked Seen Group=1;
800 950 500 1201;

Unmarked Seen Group=2;
459 782 256 921;

Marked Unidentified Group=1;
4 2 6 3;
```

```
Marked Unidentified Group=2;
2 5 0 1;

Known Marks Group=1;
10 14 0 0;

Known Marks Group=2;
11 0 0 0;

/* End Input File */
```

Here, the encounter histories are followed by two columns for group frequencies in the usual **MARK** encounter history file manner. The entries for '**Unmarked Seen**', '**Marked Unidentified**', and '**Known Marks**' are entered separately for each group. The entries under '**Known Marks**' tell **MARK** that the number of marks was known for the first and second primary occasions of the first group ($n_1 = 10$, $n_2 = 14$) and for only the first primary occasion of the second group ($n_1 = 11$). Again, no '-0' can appear for a primary occasion where the number of marks is unknown.

Notice that a '+0' appears in the encounter history for the last individual of the second group, but that the number of marks for this primary occasion was unknown. This indicates that this individual happened to be caught (e.g., during marking) immediately prior to the second primary occasion, but was never resighted. Hence, for the second group during the second primary interval , $n_2^* = 14$ and $c_2^* = 15$.

An analysis using the single group data (Poisson_RobustDesign_OneGroup.inp) yielded the following results for the random emigration model $\{\varphi(.)\gamma''(.) = \gamma'(.)\alpha(t)\ \sigma(t)U(t)\}$:

```
      Real Function Parameters of {phi(.) gamma'(.)=gamma'(.) alpha(t) sigma(t) U(t)}

                                                                95% Confidence Interval
   Parameter                Estimate       Standard Error      Lower           Upper
 -------------------------  --------------  --------------  --------------  --------------
   1:alpha                  2.7638408       0.2886637       2.2534628       3.3898122
   2:alpha                  2.6470841       0.2695821       2.1692136       3.2302279
   3:alpha                  2.1173163       0.2745082       1.6439392       2.7270036
   4:alpha                  2.1254054       0.3281373       1.5732477       2.8713520
   5:sigma                  0.2368147       0.1786795       0.0635331       0.8827081
   6:sigma                  0.4564778       0.1114859       0.2847935       0.7316598
   7:sigma                  0.3925358       0.1552277       0.1859589       0.8285937
   8:sigma                  0.5348317       0.1257812       0.3394039       0.8427864
   9:U                      456.73003       43.067154       379.81489       549.22102
  10:U                      362.54432       34.740271       300.59433       437.26168
  11:U                      427.89101       45.664583       347.33475       527.13045
  12:U                      358.01017       44.974968       280.14293       457.52102
  13:Phi                    0.9857548       0.0182401       0.8443633       0.9988683
  14:Gamma'                 0.0552683       0.0363436       0.0147309       0.1862693


                        Estimates of Derived Parameters
 Population Estimates of {phi(.) gamma'(.)=gamma'(.) alpha(.) sigma(.) U(t)}

                                                   95% Confidence Interval
 Grp. Occc.   N-hat            Standard Error      Lower           Upper
 ---- ----   --------------   --------------  --------------  --------------
```

```
  1     1   524.94217      28.178946      472.55342      583.13891
  1     2   460.37288      24.092419      415.52193      510.06500
  1     3   425.58023      23.324431      382.26492      473.80369
  1     4   383.16101      21.077938      344.02552      426.74845


Mean Resighting Rate Estimates of {phi(.) gamma'(.)=gamma'(.) alpha(.) sigma(.) U(t)}
                                              95% Confidence Interval
Grp. Occ.   Lambda-hat      Standard Error      Lower          Upper
---- ----   --------------  --------------  --------------  --------------
  1     1   2.8737886      0.1396905      2.6127801      3.1608711
  1     2   2.8452646      0.1396905      2.5843816      3.1324827
  1     3   2.8458811      0.1412053      2.5823048      3.1363607
  1     4   2.8932760      0.1416843      2.6286365      3.1845582
```

For model $\{\varphi(.)\gamma''(.) = \gamma'(.)\alpha(.)\sigma(.)U(t)\}$:

```
      Real Function Parameters of {Phi(.) gamma'(.)=gamma'(.) alpha(.) sigma(.) U(t)}

                                                                95% Confidence Interval
Parameter                  Estimate       Standard Error      Lower           Upper
-------------------------  --------------  --------------  --------------  --------------
   1:alpha                 2.4536985      0.1478956      2.1805245      2.7610956
   2:sigma                 0.4376083      0.0655452      0.3268107      0.5859693
   3:N                     524.49384      28.499239      471.81002      583.68075
   4:N                     460.04989      24.370049      415.11342      510.78703
   5:N                     426.24093      23.102678      383.69402      474.39761
   6:N                     379.16926      20.875980      340.74421      422.70778
   7:Phi                   0.9858690      0.0178497      0.8499082      0.9988380
   8:Gamma'                0.0751540      0.0287552      0.0348592      0.1545672

                    Estimates of Derived Parameters
Population Estimates of {phi(.) gamma'(.)=gamma'(.) alpha(.) sigma(.) U(t)}

                                              95% Confidence Interval
Grp. Occ.   N-hat           Standard Error      Lower          Upper
---- ----   --------------  --------------  --------------  --------------
  1     1   524.94217      28.178946      472.55342      583.13891
  1     2   460.37288      24.092419      415.52193      510.06500
  1     3   425.58023      23.324431      382.26492      473.80369
  1     4   383.16101      21.077938      344.02552      426.74845

Mean Resighting Rate Estimates of {phi(.) gamma'(.)=gamma'(.) alpha(.) sigma(.) U(t)}

                                              95% Confidence Interval
Grp. Occ.   Lambda-hat      Standard Error      Lower          Upper
---- ----   --------------  --------------  --------------  --------------
  1     1   2.8737886      0.1396905      2.6127801      3.1608711
  1     2   2.8452646      0.1396905      2.5843816      3.1324827
  1     3   2.8458811      0.1412053      2.5823048      3.1363607
  1     4   2.8932760      0.1416843      2.6286365      3.1845582
```

Here, AIC$_c$ indicates much more support for the simpler model (1,042.3 versus 1,050.0). Notice that a significant population decline would be inferred from the latter model (but not the former), one of the advantages of borrowing information across primary intervals that the Poisson-log normal model provides over other previously available mark-resight estimators.

18.5. Which mark-resight model? Decision table

As described in this chapter, there a variety of mark-resight models available to you in **MARK**: (1) the logit-normal estimator (*LNE*); (2) the immigration-emigration log-normal estimator (*IELNE*); and (3) the (zero-truncated) Poisson log-normal estimator [(*Z*)*PNE*].

The following summary table provides some guidance by comparing several of the important differences in the underlying assumptions:

estimator	*geographic closure*	*sampling with replacement*	*number of known marks*	*individually identifiable marks*
LNE	required	not allowed	required	not required
IELNE	not required	not allowed	required	not required
(*Z*)*PNE*	required	allowed	not required	required

Geographic closure is only required for *LNE* and (*Z*)*PNE* within primary sampling intervals. As described at the end of Section 18.1, closure assumptions may often be relaxed, but abundance estimates should be carefully interpreted under these circumstances.

18.6. Suggestions for mark-resight analyses in MARK

1. To start an analysis from scratch (after an encounter history file has been created), select the '**`Mark-Resight`**' data type. You will then be asked to select from several different models: '**`Logit-Normal`**, '**`Immigration-Emigration Logit-Normal`**,' or '**`Poisson-log normal`**'.
   - For the '**`Logit-Normal`**' and '**`Immigration-Emigration Logit-Normal`**' models one doesn't specify whether or not individual marks were used. This is left to the user to keep track of (by not running any individual heterogeneity models if marks were not individually identifiable).
   - For '**`Poisson-log normal`**' one doesn't need to specify robust design or not. If there are multiple primary occasions for the group(s), then **MARK** will automatically set up an analysis that includes the open period parameters (φ, γ'', and γ').
2. Because convergence with these models is sensitive to the starting values (particularly for N and σ), initial values (on the log scale) should always be manually provided in the '**`Run`**' window when using the design matrix. This means that if $N = 100$ and $\sigma = 0.5$, then $\log(N) = 4.6$ and $\log(\sigma) = -0.69$ should be provided as initial values. **MARK** provides its own initial values that usually work when running a model from the PIMs, so we suggest that an analysis begin with simple PIM models from which the initial values may then be provided for running more complex models and for when utilizing the design matrix.

   If convergence issues remain after following this strategy, we suggest trying a series of initial values covering the suspected range of the parameter(s) and possibly other '**`Run`** `window`' options such as '`Use Alt. Opt. Method`' or '`Do not standardize design matrix`'. It is typically fairly obvious when N does not converge correctly ('garbage' estimates, SE, or AIC$_c$), but it can be more tricky with σ. Sometimes the regular **MARK** optimization method can converge to a local maximum where $\hat{\sigma}$ is almost zero. Caution should be taken before concluding that such an estimate is reliable.

3. Even when using the sin link from the PIMs, **MARK** will sometimes get the parameter count wrong for the α parameters in the immigration-emigration logit-normal model. Extra care should be taken when using the model to verify the number of estimable parameters (e.g., for AIC$_c$ calculation) is correct. We hope to have this issue resolved in the future.
4. The σ parameter must be fixed to zero in the '`Run`' window to examine a model that ignores individual heterogeneity in resighting probabilities.
5. When using the (immigration-emigration) logit-normal model, **MARK** by default assigns the log link to σ and N, and applies whatever link is specified in the '`Run`' window to p.
6. When using the Poisson model, **MARK** by default assigns the log link to α, σ, and N, and applies whatever link is specified in the '`Run`' window to φ, γ'', and γ' (if using the robust design).

18.7. References

Arnason, A. N., Schwarz, C. J., and Gerrard, J. M. 1991. Estimating closed population size and number of marked animals from sighting data. *Journal of Wildlife Management*, **55**, 716-730.

Bartmann, R. M., White, G. C., Carpenter, L. H., and Garrott, R. A. 1987. Aerial mark-recapture estimates of confined mule deer in pinyon-juniper woodland. *Journal of Wildlife Management*, **51**, 41-46.

Bowden, D. C., and Kufeld, R. C. 1995. Generalized mark-resight population size estimation applied to Colorado moose. *Journal of Wildlife Management*, **59**, 840-851.

Burnham, K. P., and Anderson, D. R. 2002. *Model Selection and Multi- model Inference: A Practical Information-Theoretic Approach*. 2nd edition. Springer-Verlag, New York, USA.

Kendall, W. L., and Nichols, J. D. 2002. Estimating state-transition probabilities for unobservable states using capture-recapture/resighting data. *Ecology*, **83**, 3276-3284.

Kendall, W. L., Nichols, J. D., and Hines, J. E. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology*, **78**, 563-578.

Kendall, W. L., Pollock, K. H., and Brownie, C. 1995. A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics*, **51**, 293-308.

McClintock, B. T., and White, G. C. 2009. A less field-intensive robust design for estimating demographic parameters with mark-resight data. *Ecology*, **90**, 313-320.

McClintock, B. T., White, G. C., Antolin, M. F., and Tripp, D. W. 2009a. Estimating abundance using mark-resight when sampling is with replacement or the number of marked individuals is unknown. *Biometrics*, **65**, 237-246.

McClintock, B. T., White, G. C., Burnham, K. P., and Pryde, M. A. 2009b. A generalized mixed effects model of abundance for mark-resight data when sampling is without replacement. In *Modeling Demographic Processes in Marked Populations*, D. L. Thomson, E. G. Cooch and M. J. Conroy, eds. Springer, New York, New York, USA, pp. 271-289.

McClintock, B. T., and White, G. C. 2012. From NOREMARK to MARK: software for estimating demographic parameters using mark-resight methodology. *Journal of Ornithology*, **152** (Suppl 2), S641-S650.

Minta, S., and Mangel, M. 1989. A simple population estimate based on simulation for capture-recapture and capture-resight data. *Ecology*, **70**, 1738-1751.

Neal, A. K., White, G. C., Gill, R. B., Reed, D. F., and Olterman, J. H. 1993. Evaluation of mark-resight model assumptions for estimating mountain sheep numbers. *Journal of Wildlife Management*, **57**, 436-450.

Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, **62**.

White, G. C. 1996. NOREMARK: population estimation from mark-resighting surveys. *Wildlife Society Bulletin*, **24**, 50-52.

White, G. C., and Shenk, T. M. 2001. Population estimation with radio-marked animals. In *Radio Tracking and Animal Populations*, J. Millspaugh and J. M. Marzluff, eds. Academic Press, pp. 329- 350.

Addendum – formatting mark-resight input files

As noted at various points in this chapter, generating the input (.INP) file for mark-resight analysis is somewhat more complex than for general 'mark-recapture' analysis (as discussed in Chapter 2) – so much so, that frequently the biggest 'rate-limiting step' forpeople with mark-resight data is formatting the .INP file. **MARK** has no capability of generating .INP files, for mark-resight, or any other form of analysis. This is something you will need to do for yourself. In this short addendum, we provide several **R** scripts which can be used to generate the .INP files used in several of the examples presented in this chapter.

Note: Since there are any number of software environments you could use to accomplish the task of generating .INP files, we state for the record that we are going to demonstrate creating .INP files using **R**, not as a point of advocacy for using **R** (in fact, there are other software environments that can accomplish generating .INP files using fewer lines of code), but owing more to its increasing ubiquity. We also acknowledge from the outset that while there may be (and undoubtedly are) more elegant ways to accomplish some of the steps in the process using **R**, our goal is to present scripts that are relatively transparent in terms of 'what they're doing' (at least to **R** users), and thus, relatively easy to customize to specific purposes.

LNE & IELNE – no individually identifiable marks

The following script will generate 'artificial encounter histories' for the single-group .INP files presented for the LNE (section 18.2.1) and the IELNE (section 18.3.1) – both situations involves studies where there are no individually identifiable marks. The script is more or less identical in either case – the only change you need to make are to the **n** and **m** vectors.

```
#
# short script to generate artificial encounter histories for LNE and IELNE
# one groups examples - no individually identifiable marks...
#

# enter n vector (n - exact number of marked individuals in pop during primary interval)
# Here, we enter n for each secondary within each primary...
n <- c(30,30,30,30,33,33,33,33,32,32,32,32)  # LNE - section 18.2.1
# n <- c(27,22,18,29,28,23,20,32,31,19,21,33) # IELNE - section 18.3.1

# calc total number of sample periods (primary x secondary)
tot_per=length(n);

# enter m vector (m = total number of marked individual sightings during
# secondary occasion within primary interval)
m <- c(8,9,10,5,11,10,18,9,5,10,13,8); # LNE - section 18.2.1
# m <- c(17,15,9,8,16,14,9,13,11,14,13,16); # IELNE - section 18.3.1

# find largest element of n vector - need this later...
n_max=max(n);

# initialize encounter history matrix
eh <- matrix(0,n_max,tot_per);

# fill in 1's for each secondary sample...
```

```
for (i in 1:tot_per)
 {  eh[1:m[i],i]=1; }

# fill in dots as needed for each secondary sample
eh <- data.frame(eh);
for (i in 1:tot_per)
  { if (n[i]<n_max) eh[(n[i]+1):n_max,i]="."; }

# write out EH matrix...this will need to be edited to include unmarked seen individuals for
# each primary period...

eh <- as.matrix(eh);
summ_eh <- apply(eh,1,paste,collapse="")
summ_eh <- as.data.frame(table(summ_eh));
summ_eh <- summ_eh[rev(rownames(summ_eh)),];
summ_eh$end <- ";";

write.table(summ_eh,file="LNE_EH.inp",sep=" ",quote=F,col.names=F,row.names=F);
```

The script generates the 'artificial encounter histories'. You will still need to manually edit the file output by the script, to enter additional information needed for the particular estimator (e.g., for the LNE, you would need to add information about `Unmarked Seen Group=1;` and `Marked Unidentified Group=1;`). We'll leave it to you as an exercise to figure out how to modify the script to accommodate > 1 group.