Ch. 10

Cormack-Jolly-Seber: estimating apparent survival

"...Nature, red in tooth and claw"
-- Alfred, Lord Tennyson

"I've always believed in survival."
-- Hugh Leonard

Questions to ponder:

- *What is the difference between survival and apparent survival?*
- *How can survival be estimated from the information found in a set of capture histories?*
- *Is it possible to use a 2-occasion capture exercise to estimate survival?*

A basic framework for survival of marked animals

In the previous chapter, we covered the use of mark-recapture in a closed population form to estimate N, or population size. Now, we'll extend the general ideas of sampling without certainty of catching the animal again ($p<1.0$) to **open populations**—populations for which the time period of consideration allows births and deaths to occur. The type of data structure on which we focus here is the basic type of survival analyses that biologists use to assess survival of marked individuals which are <u>not</u> radio-marked—that is, individuals with leg bands, tags, or other marks. In contrast to Chapter 9, animals with these types of marks have **unknown fate**, as we cannot follow them and know their fate unless we capture them again.

We call this method "CJS" after Richard Cormack, George Jolly, and George Seber—three people who developed the method separately in 1964 (Cormack) and 1965 (Jolly and Seber).

The typical mark-recapture study for CJS-type analyses consists of a population identified for study and a method employed to capture animals. Almost any method can be used: nets, traps, pit falls, dart guns, or whatever can be used to capture a representative sample of the animal species of interest.

Those animals then receive some kind of mark that allows individual identification. Although population size can be estimated using marks that are not individually unique—such as paint splatters on iguanas or a fin clip of a fish—**CJS apparent survival estimation requires unique marks on individuals**. Visual sightings can be used for re-captures of animals if the proper kind of visual markers is used.

The mark-recapture project must be extended for at least 3 time periods. Later, we will discuss why you can't stop with two sampling occasions if you have interest in survival. During each time period, animals are captured, identified (if already marked) or newly marked, and then released back into the population, alive.

Time periods need to be long enough to make sense for the species of interest---some mortality needs the opportunity to occur during that time period. So, for most larger fish and wildlife, biologists might use month or three-month or six-month or annual intervals, as examples.

Once an animal is marked and released, three potential things can happen to it if we look forward to the next time period (Figure 10.1):

- First, an animal can survive through time period 1 and be captured at the start of time period two. That animal, has the capture history of "11".
- Second, an animal that survives through time period 1 may be alive during time period 2, but not captured in that time period. So, it would have a capture history of "10".
- And last, an animal may be released at the beginning of time period one and it may die of various causes during the first time period. It is not available for capture. So, it would also end up with a "10" capture history.

A key point is that, with CJS-type data, we cannot distinguish between the last two possibilities. That is, we have no way of knowing if the third animal died or not. It is not wearing a radio-tag. And, we did not find it dead. So, animals not captured again simply disappear. And, this is one reason that we need at least three time periods to estimate survival. If we stop at two time periods, we are left with an unsolvable dilemma—and no way to separate survival and capture rate for these animals with "10" histories (Figure 10.1).
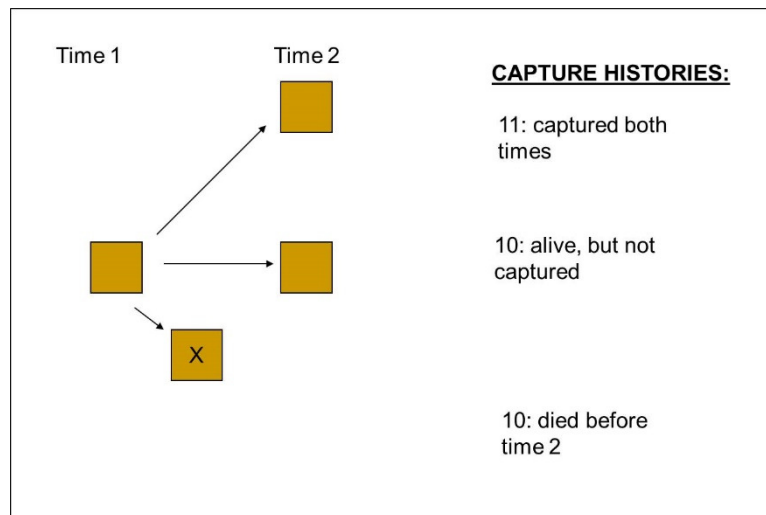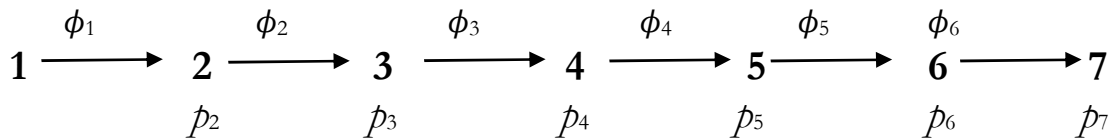

Figure 10.1: *Capture histories for marked animals with three possible transitions to capture and survival status in Time 2 after release in Time 1.*

We should also note at this point that the survival we estimate from CJS-type data analyses is labeled **'apparent survival.'**

A fourth alternative is actually possible here…that is, the animal could have remained alive, but left the study area. That would also result in a "10" capture history. There is definitely not a way to separate out the movement part of this from the mortality part when using the CJS framework. So, we assume that emigrated animals are dead, and we use the label 'apparent' survival in our publications to indicate that it is actually a conservative estimate of survival. That is, the true survival rate may actually be higher than the apparent survival estimate we obtain from CJS if movement rates out of our study site are high. To our CJS model, a situation in which animals are moving away would cause the capture histories to appear as if many animals are dying, and survival will be estimated to be low. The best we can do is to call it 'apparent' survival. *Note: Many editors will insist on this label in journals to distinguish CJS-estimated survival from known-fate survival or other estimates of survival.*

$$1 \xrightarrow{\phi_1} 2 \xrightarrow{\phi_2} 3 \xrightarrow{\phi_3} 4 \xrightarrow{\phi_4} 5 \xrightarrow{\phi_5} 6 \xrightarrow{\phi_6} 7$$

$$\quad p_2 \qquad p_3 \qquad p_4 \qquad p_5 \qquad p_6 \qquad p_7$$

Event timing in CJS models

It is worth a reminder of the timing of events that are assumed in the structure of a survival analysis. Animals are released IMMEDIATELY at the beginning of a time period. In fact, their capture and release are assumed to happen instantaneously and together the capture and release define the beginning of a time period.

So, the survival estimate, ϕ_1 ("phi 1"), here corresponds to the probability that an animal released at the beginning of time period 1 survives to the second time period. The capture probabilities are defined as the probability of being captured in a given time period, given that the animal is alive and available for capture. At the moment, this looks very sensible. It will become important to remember as our model structures grow more complex—for example as we head toward robust-design or multi-state models in future chapters.

NOTE: we will use the convention of other authors to distinguish between "true" survival, S, and apparent survival, ϕ. This distinction will become especially important in Chapter 13 for the robust design models.

Another thing to note, as we look at this CJS example, is that if you have 7 time periods, you will only obtain 6 survival rate estimates---and, as we will discuss momentarily---the last one is actually not estimable as a lone parameter. So, in effect, you will only obtain 5 survival rate estimates. This is important in study design and long-term planning. A student who wants to do a 3-year field project will actually only be able to estimate 2 annual survival rates, and the second is not completely estimable as a time-specific rate. We will come back to this concept later.
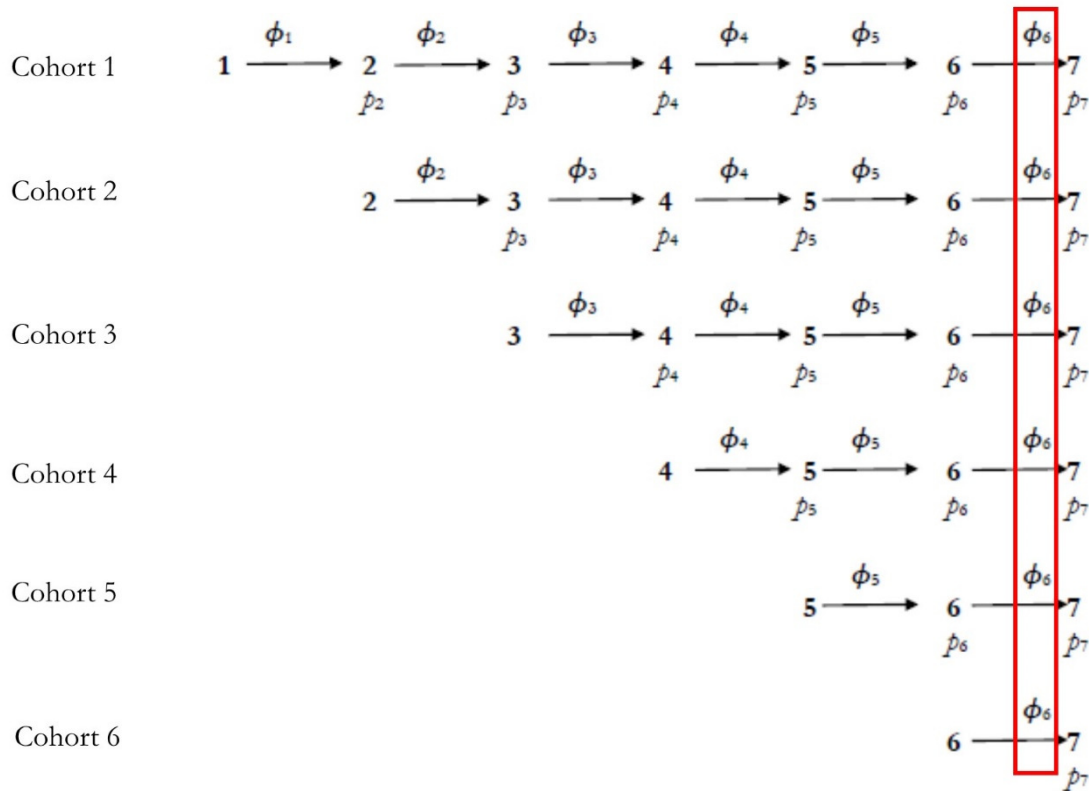
Cohorts working together

Figure 10.2: *Example of 6 cohorts of released animals, captured and recaptured during a 7-period study to estimate apparent survival (ϕ). Apparent survival between periods and capture during periods (p) are shown. Vertical box during period 6 indicates the logic that all 6 cohorts provide information for an estimate of apparent survival for the population during that time period, but survival is not estimable as a time-specific (e.g., ϕ_6) parameter for the last interval.*

We've previously discussed the notion of a cohort. In a multi-year study, cohorts will eventually exist in the presence of other cohorts that are released in subsequent time periods, and in Figure 10.2 we see how a series of cohorts work together in a study that continues to capture new animals through time. In time period one, the first cohort is released. In time period two, a second cohort of animals is released. The same thing happens during all time periods.

Notice that all animals—regardless of which cohort they are in—are used to estimate time-specific parameters. For example ϕ_6 (or survival through the sixth time period) can be gotten by using data from all six cohorts.

Probability statements

One of our goals is to become more comfortable thinking about events in mark-recapture data as probability statements. Here, we have a table of cohorts that are released and animals captured again, for the first time, in subsequent time periods.

	Probability of recapture, first, in time period x		
Cohort released	**2**	**3**	**4**
R_1	$\phi_1 p_2$	$\phi_1(1\text{-}p_2) * \phi_2 p_3$	$\phi_1(1\text{-}p_2) * \phi_2(1\text{-}p_3) * \phi_3 p_4$
R_2		$\phi_2 p_3$	$\phi_2(1\text{-}p_3) * \phi_3 p_4$
R_3			$\phi_3 p_4$

Let us start with a look at a released cohort of animals at time one: R_1. It is possible that these animals will be next captured in time 2 (a capture history of "11" when they are recaptured for the first time). Or, they may skip capture in time 2 and be next captured in time 3 (capture history of "101" at first recapture). Or, an animal may go uncaptured through time 2 and 3 and be next captured in time 4 (capture history of "1001"). So, for R_1, we can write out several probability statements that go with these three possibilities.

It may help to walk through the logic behind the probabilities. First, what **binomial events** must happen to an animal released in time 1 and next captured in time 2 (Figure 10.1)? We know that it is has to remain alive to be captured again. So, it has to survive through time period 1 to be captured in time 2. And, it has to be captured in time two. Those two events (survival and capture) BOTH have to happen, so we multiply the probabilities: $\phi_1 * p_2$.

Similarly, if an animal is released in time 1 and captured next in time period 3, we can write the probability of that event. The animal has to survive through time 1, then it must NOT be caught in time period 2, and then it must survive through time period 2, and then (finally) be captured in time period 3. So, you should now envision a string of 4 probabilities for that 'event.' We used *1-p_2* to indicate that the animal is not captured in time 2. You can see a similar probability statement for time period 4. Work through it in your mind—it should make sense in terms of the order of binomial events that have to happen to be released in 1 and only seen again in time period 4: survive, not captured, survive, not captured, survive, and captured. We write that formally as: $\phi_1(1\text{-}p_2) * \phi_2(1\text{-}p_3) * \phi_3 p_4$

These probability statements are at the core of what the CJS models 'know' or assume about our system as we attempt to estimate these parameters from data that we collect.

CJS model assumptions

As with all data structures and models, CJS-type survival analyses come with some **assumptions**.

- Every animal has the same chance of capture, p
- Every animal has same probability of surviving (ϕ: phi) to the next sampling period
- Marks are not lost/overlooked
- Samples are instantaneous (short periods) and animals released immediately
- All emigration from the sample area is permanent (101 must indicate "1-p")
- Fates of animals (with regard to capture and survival) are independent of other animals

The first assumption simply means that we assume that our capture probability estimate applies to all individuals in the population—or at least within a group that we designate in our modeling with a specific estimate for p. For example, we could break the population into males and females, which could be allowed to have different estimates of p—but we assume then that all males have the same chance of capture.

We would call this a 'definitional' assumption, then. The assumption is because of the manner in which we estimate the p's. We acknowledge that we have no way to account for heterogeneity in capture probabilities within a population in the basic models. So, if we THINK there are differences, we need to sample within those groups with an appropriate sample size for each group (such as the males and females) to account for differences that may exist.

If we ignored this assumption and had disproportionate sampling (e.g., ¾ of sample were males and ¼ females)…then our estimate of 'p' for the entire population would be biased by the males and might affect the estimates of survival of the animals.

The assumption about instantaneous sampling and release bears some comment. We assume the sampling and release is instantaneous, and this seems obvious. But, what if you were doing annual survival of fish in the Missouri River and your sampling season was three months long (June, July, and August)? Some fish would be released at the end of those summer months and only have to survive 9 months until the next sampling season while some fish would be released at the beginning of the summer and have to survive 12 months until the next sampling season. In this Missouri River fisheries example, we are probably not going to have many problems by simply acknowledging the assumption in our analyses. But, what if your sampling period each year was 6 months long? Now, fish sampled in the first month have to survive 5 months longer to be available for sampling during the next year than fish sampled in the last month—and all are represented in the same manner in your capture history with a "1" for that year. You can see the assumption of instantaneous sampling is being stretched, in this example. Or, to push this point even further, what if sampling occurred during the entire year!? Collapsing all of your data into one, annual interval is certainly going to be problematic because of the instantaneous

sampling assumption. One approach to avoid this problem would be to sample during two or more, 3-month time periods instead of 6-month or 12-month sampling periods. Generally, the sampling periods need to be much shorter than the time periods used in your survival model.

A note about the emigration assumption: we have already described that the CJS structure has no way to distinguish an animal that leaves the study area with one that dies. And, that is because of the design of our study—CJS-type data comes from a single study area. We aren't capturing animals in other places. So, all animals are 'dead' to the model when they emigrate. Because of that, animals that emigrate and then immigrate back to the area in a subsequent time period become problematic. A "101" capture history is assumed to mean that the animal was *on site* during time 2, but not captured. A lot of movement in and out starts to affect capture probabilities, and may mess up the estimates for this model structure. If you have a lot of movement, you might like to explore the robust-design or multi-state models in future chapters—to document and take advantage of movement information (which also means revising your study design to have multiple sites for multi-state data).

CJS models are relatively simple—they assume movement is not occurring. Remember that p is defined as the probability of capture, given that an animal is on-site and available for capture. You are breaking this assumption if your study animal has moved to Canada and your study site is in Mexico.

Last, the model structure treats all individuals as separate, independent creatures. We know that nature is messy and animals clump together and experience the same mortality risks as others, from time to time. For example, breeding pairs of songbirds or male deer in large groups during the non-breeding season, or fish in a school. Again, this assumption becomes a problem if *too much* pooling occurs. So, we need to be careful how we sample. Perhaps we mark only one individual from a given group if we are worried about the effects of the group on the individual.

These are all things that are important to consider as you design a study.

Logic behind the CJS model structure

How does the CJS model estimate survival from your data?

We can start with the closed-population logic seen in the Lincoln-Petersen estimator for closed populations (Chapter 8). In that 2-sample occasion, Lincoln and Petersen noted that there are two ratios that *should* be the same---the portion of the population that you capture during the first time period, n_1/N, and the portion of the second sample that is marked, m_2/n_2.

$$n_1/N = m_2/n_2$$

That is, all animals in the population have the same probability of capture, so the marked portion of the second sample should be equivalent to the portion that was sampled in the first sample. CJS simply extends that to an open situation, with the same logic: all individuals (whether previously marked or not) have the same probability of capture.

$$M_i/N_i = m_i/n_i$$

We introduce M_i to indicate (as with M_t in the Schnable estimator in Chapter 8) the number of animals that have been marked *previously* to the start of any capture period. And, m_i is defined as the number of marked animals captured in sample i, and n_i is the number of animals (marked + unmarked) in sample i.

So, Cormack, Jolly, and Seber postulated that the portion of the population that has been marked previously (M_i/N_i) should be equal to the portion of marked animals in the current sample (m_i/n_i).

The next step in this, and a fairly simple, straightforward one….is to note that survival, ϕ, is simply the portion of marked animals at the current time that are alive in the next time period.

$$\phi = M_{i+1} \ / \ M_i$$

So, we've marked 50 animals by time period 6 for example. By time period 7, 40 of them are alive. So, survival is simply 40/50 or 80%.

Ignoring <u>new</u> marks, the only way that M_i can change is to have animals die or leave the study area. Furthermore, we have no means of distinguishing emigrants from deaths, so they both are considered deaths.

This seems pretty logical. But, of course, the problem is that these animals are not radio-marked or tethered to our research station with a cable. We normally *do not* know when animals leave or die. We just know when we mark them and when we catch them. So, the simple logic gets a bit more complex, because **we have to use things we do know to estimate what we don't know**.

In fact, as Cormack, Jolly, and Seber postulated, we actually do know quite a bit about a sample of marked animals…if we just look for evidence 'hidden' in their capture histories.

So, we have a group of 'sufficient statistics' as they are termed…essentially, some parameters that we need to estimate or calculate to estimate the parameter in which we have interest (i.e., survival).

Here is the list of these **sufficient statistics**. We can think about a set of captured animals with their capture histories as we review these:

- R_i : the number of animals released at time i. This is our **cohort**, and it becomes a critical group of animals.
- r_i : the number of animals in the cohort from time i that are captured again…at any point in the future. So, if we release animals in time period 4, r_4 would be animals released in time 4 that are captured again in time 5, 6, 7, 8, 9, 10, etc. until the end of the study.
- z_i : animals captured before i, not captured in i, but captured later after i. Admittedly, this is the oddest sufficient statistic, and one that may take some time to grasp. It is the

number of animals captured before the current time period which were *not* captured in the current time period *but were captured* in a later time period. This may be easier to visualize when we look at some capture histories below.

- M_i-m_i : the number of marked animals (M_i) that *are not* sampled in the current time period (remember, m_i is the number of marked animals that *are* captured in the current time period).

We can relate these sufficient statistics in an equation:

$$r_i/R_i = z_i/ (M_i-m_i)$$

This equation is an equality of two ratios, and we will explore this momentarily. But, for now, just recognize that we have an equation that has M_i in it (which we need to estimate survival across periods). And, all other parameters in that equation are *things that we can find in our data—* by looking at our capture histories. So, we are now able to estimate M_i for the current time period! This is the brilliance of Cormack, Jolly, and Seber. We can estimate M_i; and if we can do that for two time periods in a row, we can estimate survival!
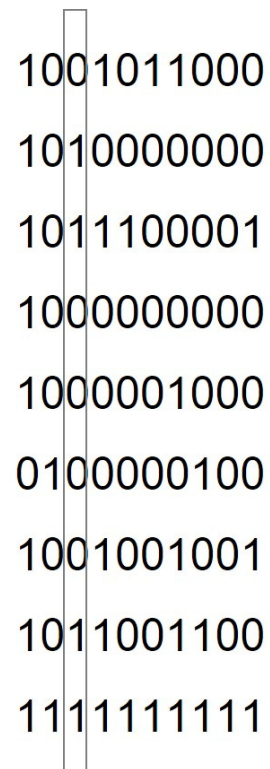
Sufficient statistics: an example

To help us look at the basic components of an apparent survival analysis-- R, *r* and *z*—let us examine a set of 9 capture histories. As a reminder, each row is a capture history for one individual over 10 capture occasions. A "1" indicates the individual was caught in time *i*, and a "0" indicates the individual was not captured in time *i*.

Let's look especially at time period 3, highlighted here by the red vertical box. We can see which of the 9 individuals were captured in that time period---the 2nd and 3rd individual, as well as the last two individuals in the set. No other individuals were captured during this period.

So, we can calculate $R_3 = 4$. We captured and released (alive) 4 individuals. That's the easiest statistic to calculate.

The value r_3 is equal to the number of R_3 that are captured again later. Of these 4 'critters' released in time 3, three are seen again—we can easily look to see if there is another "1" in the capture history after time 3. And, we can also see that the second individual in this set released in time 3 was never observed again. Therefore, $r_3 = 3$.

But, z_i is our 'trouble-maker'. This is the statistic that makes our head 'swim'. To review, z_3 is the number of animals *not caught in time 3* but *captured before and after time 3*. We start by looking for 0's in the capture histories during time 3—to indicate animals not caught in time 3.

1001011000

1010000000

1011100001

1000000000

1000001000

0100000100

1001001001

1011001100

1111111111

Figure 10.3: *A set of 9 capture histories used to estimate apparent survival. The capture status (0 or 1, not captured or captured) is emphasized for time period 3.*

There are 5 of them: the 1st, 4th, 5th, 6th, and 7th individuals in the set of 9. Now, we look at each one of those. Were they captured before time 3? Well, it turns out that…yes…they were all captured before (we can see a "1" in their history previously). And, finally, were they captured after time 3? Only individual 4 was not captured later. But, all the others were captured. So, z_3 = 4.

We note that z_i is a complicated 'little' statistic, but it turns out to be important for the logic of the CJS estimator. Make sure you understand z_i!

The grand equality

Let's return to our equation from earlier—the critical step in the CJS estimation process, as we try to estimate M_i:

$$r_i / R_i = z_i / (M_i - m_i)$$

If you are a visual learner, the flow chart in Figure 10.4 may help you appreciate the logic in this equation. Why are these two ratios set aside as equal and what is the logic?
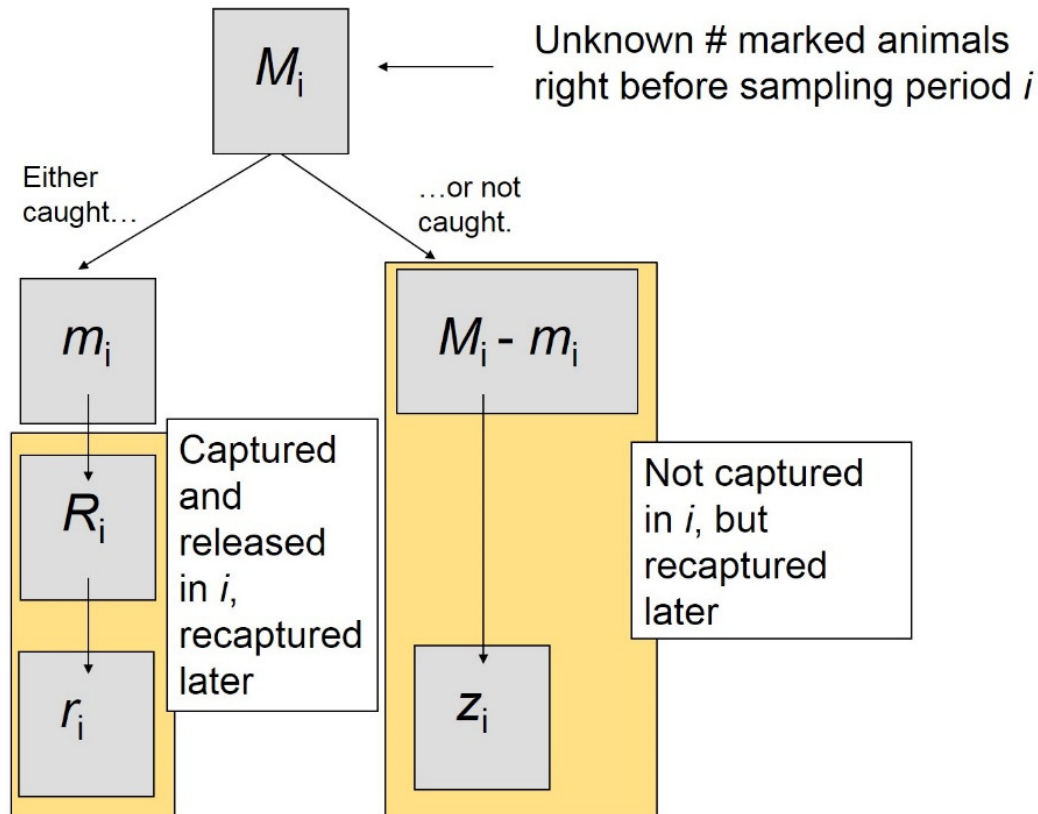

Figure 10.4: *Depiction of the 'grand equality' that allows estimation of apparent survival from CJS-type data. The unknown number of marked animals alive before a sampling period (M_i) will either be caught (m_i) or not caught ($M_i - m_i$) in the sampling period, i. The probability of capture in subsequent time periods is the same for both groups of marked animals, regardless of capture in time period i.*

As we evaluated Figure 10.4, consider M_i, the number of marked animals you've marked in the population to this point that are still alive. For our discussion, let's say we say we are in time period 6.

So, consider M_6. You know how many animals you have marked, but you don't know how many have died (and thus you don't know M_6, the number of marked animals alive before time period 6).

We only know how many we've marked.

So, of those unknown number of marked animals that are still alive out there, there are only two things (a binomial event) that can happen to them during each time period! They can either be caught or they can be 'not caught'.

For time period 6, then let's divide the 'pile' of unknown marked animals into two 'heaps of critters'—those caught (these will go on the left) and those not caught (these go on the right) during time period 6 (Figure 10.4). We can now start to travel down the two forks of the diagram, and if you are especially alert you will notice that the two sides of the diagram also contain statistics that match the two sides of our equation, above!

Let's follow the side that we know something about—because we caught them during time period 6. These are designated as m_6. We are going to release R_6 of those in time 6 and we are going to recapture a certain number of them (r_6) after time period 6. Thus, the left side of our equation is that ratio: r_6/R_6. Essentially, this is the long-term capture probability of the released animals from time period 6 until the end of the study.

Now, let's go to the right side of the diagram and the right side of our equation, above. These are the marked animals that we did not catch during time 6. Again, we don't know how many we didn't catch, because we assume some of our marked animals have died without evidence during the past few time periods. But, we can enumerate these dead animals (unknown in number) as $M_6 - m_6$. We don't know what that number is, but we can stipulate it, theoretically and (importantly), we do know m_6.

Now, we see our 'friend', z_6. z_6 is a statistic that tells us something about the animals not captured in the current time period (time 6 in our example). We know that there are some animals out there not captured, because we can look at the capture histories and see that there were animals out there that were captured before time 6. And, for z of them we know that they were also alive after time period 6…because we can see they were captured later. By definition, if you were an animal captured before time 6 and after time 6, but not during time 6, I know you were alive in time 6…even though I did not record your capture. By logic and the assumptions of the CJS model structure, you had to be alive in time 6 if you were captured before and after time 6!

So, that number of 'known alive' critters not captured during time 6 can be thought of as a proportion of marked animals that are alive in time 6 but not caught.

> **The key:** *given the animal is alive, it is assumed to have the same probability of being captured in a future time period, regardless of whether it was captured in time i.*

Furthermore, that proportion should be equal to the portion of animals *known to be alive* in time 6 and recaptured later….because they have the same chance of being captured again. Is your head 'swimming' yet? It's relatively simple when you work through it, but it takes a bit of time to get your head around this…We warned you!

Now that we can manually find the statistics we need from our capture histories, we can be grateful that our software packages will do this for us in the future! And, we can rearrange the equality to estimate the parameter of interest, M_i:

$$M_i = m_i \quad + \quad (R_i * z_i)/r_i$$

For another reminder, we get survival estimates by using the estimates of M in subsequent time periods:

$$\phi = M_{i+1} \, / \, M_i$$

The logic of the formula shown here is that the number of alive, marked animals in the population (M_i) at a point in time is equal to the number of marked animals that we capture in the given time period (m_i), plus the number of marked animals that we don't catch, but who are still alive and still in our study area: $(R_i * z_i)/r_i$.

We really can get a lot of information from looking at capture histories, can't we? You didn't think it was magic, did you?!

Maximum "Eye-klihood" Estimation?

In a previous chapter, we discussed maximum likelihood estimation, or MLE. The general idea of MLE is that the process finds the most likely estimate for a parameter, given your data. Although at times this sounds like wizardry, our goal here is to cut through the haze and provide a glimpse of the logic behind the formulae. So, let's try a simple logic exercise—use your eyes to suggest some general answers.

Suppose a biologist conducts a seven-year, CJS-type capture-recapture study of three different populations of animals (they don't even have to be the same species). A portion of the capture histories (*we will assume the portion that is shown is representative of the entire captured sample*) are shown in Figure 10.5.

Study #1	Study #2	Study #3
10100000	11100000	10100101
10010000	11110000	10010101
01010001	01110101	01010101
01001000	01011000	01001001
01100000	01110000	01100010
01000000	01100000	01000001
10000000	11000000	10000010
00100100	00110100	00100101
00101100	00111100	00101101
00010100	00011100	00010101
00011000	00011100	00011010

Figure 10.5: *Three sets of capture histories from different hypothetical studies to provide visual appraisal of relative survival and capture probabilities between studies.*

Look at the capture histories closely. There are some differences that should start to appear to your eye. Look especially for evidence that your data provide for the length of time that animals remain alive, and also look for evidence of capture probability.

Start by comparing Study #1 with Study #2. What do you see? With regards to survival, do animals in one population appear to live longer than the other? No. Animals in study #1 are typically in the data set for about 3-4 years before dropping out (although the third individual in the set lived for at least 7 years, correct?). For animals in study #2 they appear to be roughly the same as #1. What is different between study #1 and study #2? You should see that there are many more instances of "11", "111", or "1111" in Study #2—that is, capture probability seems to be higher in Study #2, because animals are routinely captured many times in sequence after their initial capture. Very simply, there are fewer "0" entries in study #2, which means that the biologist had a relatively higher probability of capturing her study animals.

To our eye, we can see that, most likely (to use the MLE phrasing), capture probability is higher for Study #2. And, although our eye cannot tell us whether $p = 0.36$ or $p = 0.48$ or $p = 0.88$, we do know that $p_{study\#1} < p_{study\#2}$.

Now, let's compare Study #1 with Study #3. What do your eyes see here?
Is capture probability different? No, p doesn't appear to vary (at least noticeably) between Study #1 and Study #3. But, what about survival? Would you expect survival to be different if you conducted an analysis of the two data sets?

Yes, survival certainly seems to be much higher in Study #3. In Study #3, we see evidence (based on our capture histories) that animals typically survive 5, 6, or 7 years. And, as we noted before, the capture histories in Study #1 show evidence for animals surviving 3 or 4 years after initial capture (with that one exception of 7 years).

So, to our eye, we can postulate that survival is "most likely" higher for the species of animal that was captured in Study #3, relative to the species captured in Study #1.

We'll call this *Maximum Eye-klihood Estimation*—using our eye to look at our data and using our brain to make relative comparisons. "MEE" is a technique at which you should become adept. You should always *look* at your capture histories *before* submitting them to some package of software for analysis!

Estimable parameters

It is necessary to pause and discuss situations for which you cannot estimate a parameter (or may have trouble doing so).

First, if your data are sparse (few captures) and if you are using a time-specific model (asking for time specific estimates of survival), you may see bogus estimates of parameters or estimates for their variance in your analysis' output. That is, if you have 0's for many of the sufficient statistics for a given time period (r_i, z_i, or m_i), that will cause problems in the time-specific estimates. What can you do? If you believe a time-specific model is biologically or otherwise important, then try to estimate the time-specific parameters. But, be ready to collapse survival to a 'constant' parameter (that is, not time-specific) if you see problems with the estimates or variance for estimates in the time-specific models.

Note that this problem has implications for sampling as well. We need to have sufficient effort to get enough captures and recaptures of marked individuals during each time period.

Secondly, because of the structure of the models, it is *never* possible to estimate the last time period's survival probability in a CJS-type model structure. It turns out that you can only estimate the product of the capture rate and survival, which we loosely label as a 'return rate'— or the probability of seeing an animal again (surviving and being captured).

This may or may not be a problem to you. But, this problem does affect short-term, time-specific analyses that are often used by MS or PhD students who really want time-specific survival estimates.

To showcase why it's not possible to estimate the last period's survival rate, let's take the example of a two-time-period study. Now, we already said earlier that we need three time periods for a CJS study…to be able to get a "z" statistic. So this example will 'poke a nail in the coffin', so to speak.

> *Let's assume you are a MS student, and you have two field seasons to 'make your mark' on the world by estimating annual survival for a rare species. You work very hard to release 100 animals in time period one and you work equally hard to capture 60 of those marked animals during the second time period.*

Here's the problem encountered by any software package that employs a maximum likelihood estimator to get your survival estimates. There are literally an infinite number of combinations of values for ϕ and p that will result in 60 animals being sampled:

- You could have 100% survival and 60% capture probability.
- Or, you could have 80% survival and 75% capture probability.
- Or, you could have 60% survival and 100% capture probability.
- Etc.

The MLE is going to literally blow up. Well, perhaps not "literally", but MLE cannot work in this situation. There is no 'most likely value given the data'.

The solution: you need more data from additional time periods to estimate ϕ and p separately. Or, you can change to a model with constant recapture rate to obtain the last survival rate (now, p is the same across all time periods, and the other time periods are providing the information needed to estimate p—leaving the model with enough information to estimate ϕ for the last period).

However, in a model with time-specific survival and recapture probabilities, we get an estimate that the product of the two (ϕ and p) is 60%. That might provide useful information, to some extent, but it's not the apparent survival estimate that we wanted. This is important to consider when you design a study! Essentially, you always need one additional year of capture data to estimate annual survival and recapture rate(s).

Conclusion

Open populations are often more biologically relevant than closed populations because we are often interested in the probability of survival over time. The time periods used in CJS-type models however, need to fit the life history of the species of interest—longer periods for species with longer life spans and shorter periods for species with shorter life spans. CJS-models assume that capture probability is not 100%, so the models include parameters for capture probability and survival probability during each time period. A set of capture histories is used as the input data for CJS-type analyses, and we can calculate several 'sufficient statistics' from summaries of the capture histories that allow eventual estimation of survival.

CJS-type models estimate 'apparent survival' because the models do not include any movement parameters; thus, an animal that emigrates is considered to have died. Apparent survival estimates are considered conservative estimates of survival, as some animals (assumed to have died) are probably still alive but outside of the study area. If there appears to be a lot of movement, you may wish to consider the use of either the robust-design or multi-state models (Chapters 12 and 13) to document and take advantage of movement information.

For more information on topics in this chapter

Amstrup, S. C., T. L. McDonald, and B. F. J. Manly. 2005. Handbook of capture-recapture analysis. Princeton Univ. Press: Princeton, NJ.

Conroy, M. J., and J. P. Carroll. 2009. Quantitative Conservation of Vertebrates. Wiley-Blackwell: Sussex, UK.

Schwarz, C. J., and A. N. Arnason. 2014. Chapter 12: Jolly-Seber models in MARK. *In* Program MARK: a gentle introduction, 12th edition, Cooch, E. and G. White, eds. Online: *http://www.phidot.org/software/mark/docs/book/pdf/chap12.pdf*

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press: San Diego.

Citing this primer

Powell, L. A., and G. A. Gale. 2015. Estimation of Parameters for Animal Populations: a primer for the rest of us. Caught Napping Publications: Lincoln, NE.

An American alligator (Alligator mississippiensis) is tagged with an individual ID tag on a raised tail scute by biologist Russ Walsh (Mississippi Wildlife, Fisheries and Parks) during a night capture exercise. Photo by Tim Hiller (used with permission).