4

Estimating population vital rates

> *If I were an animal, I'd choose to be a skunk: live fearlessly, eat anything, gestate my young in just two months, and fall into a state of dreaming torpor when the cold bit hard. Wherever I went, I'd leave my sloppy tracks. I wouldn't walk so much as putter, destinationless, in a serene belligerence – past hunters, past death overhead, past death all around.*
>
> **Louise Erdrich (1993), in *The Georgia Review***

Introduction

How many leopards are in a National Park, and how fast are they dying and reproducing? Are deer in Colorado few and declining, or many and increasing? How is the proportion of male and female turtles changing due to global warming?

Population characteristics estimated from field data form the skeleton on which the body of applied population biology rests. These **vital rates**, within and among populations, are united by the famous BIDE equation:

$$N_{t+1} = N_t + B + I - D - E \tag{4.1}$$

Abundance (N) at time $t + 1$ equals the abundance the previous time step, t, plus the number of animals that arrive due to birth (B) or immigration (I), and minus those dying (D) or emigrating (E). Because males and females often have different dispersal and mortality patterns, sex ratio also affects and is affected by these vital rates.

In this chapter I will discuss within-population vital rates, including abundance and density, survival, reproduction, and sex ratio. Later in the book (Chapter 10) I will describe how to estimate immigration and emigration, key vital rates for the dynamics of multiple populations.

Estimating abundance and density

Abundance is probably the piece of information most sought after in wildlife population biology: it is key for determining harvest regulations, for deciding on protection

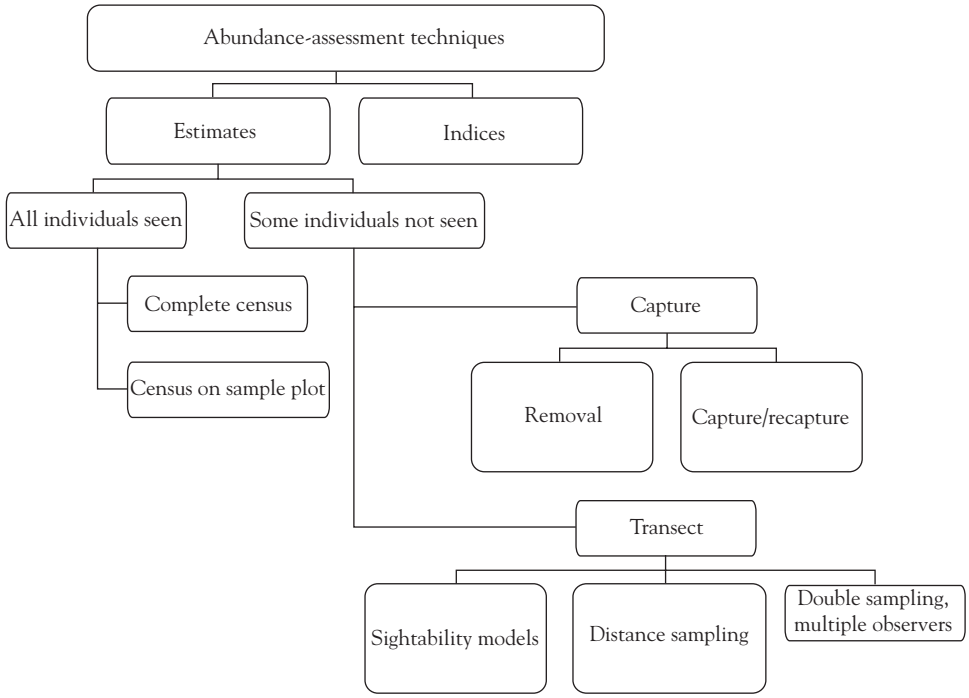


Fig. 4.1 Schematic of various measures mentioned in text to assess abundance of wildlife populations. The major groupings are based on indices or estimators, and whether all animals can be seen ($\hat{p} = 1.0$) or not ($\hat{p} < 1.0$). Modified from Lancia et al. (2005). Reproduced by permission of The Wildlife Society.

for species, and for evaluating the effects of predation or human actions, to name just a few. On the face of it, animal abundance might seem to be trivial: why not just count them? And yet, abundance estimation is one of the most mathematically sophisticated and conceptually challenging components of wildlife population biology. I will cover the basics of some of the most widely used approaches (Fig. 4.1). Throughout, I will use **abundance** and **population size** interchangeably to refer to the number of animals, and **density** to refer to abundance per unit area.

Abundance estimate versus census versus index

Here is one of the most important take-home messages of the chapter: all estimates of animal abundance, no matter how wild the math or intimidating the equations, can be reduced to a simple equation with profound implications:

$$\text{Estimate of abundance} = \hat{N} = \frac{\text{Count of animals}}{\text{Estimated probability of detection}} = \frac{\text{Count}}{\hat{p}} \qquad (4.2)$$

When all animals are detected ($\hat{p} = 1$), the count equals the estimate. But as the probability of detection declines, our estimate will increase.

This equation for abundance estimation has been called canonical (Williams et al. 2002), in the sense that it is a simple yet axiomatic and universally binding principle. The estimated probability of detection ($\hat{p}$) is a function of both detectability (e.g. capture, sighting) within the sampled area and the proportion of the area sampled relative to the total population. Setting aside for the moment the second of these components – the fact that time and money often limit us to sampling only subsections of a population's range (Chapter 2) – we will focus on the implications of detecting fewer than 100% of the animals within a sampled area.

In nature animals are elusive and sometimes downright uncooperative when we try to detect them. They avoid traps, turn their reflective eyes from spotlights, hide under trees when biologists fly overhead, and pass rub pads without leaving hairs behind. In short, detection probabilities will nearly always be less than 1.0 in wildlife studies, so eqn. 4.2 shows that if we use the count alone without accounting for detection probability the resulting measure of abundance will be too small. For example, if we count 40 rabbits on a spotlight transect and ignore the fact that detection probability is, say, 0.5, then we would report 40 animals when really $\hat{N}$ should be (40/0.5) = 80. Detection probability can change over time, space, and even among individuals, and these complexities are why so many articles and books describe mathematical ways to estimate this probability. As we will see, accounting for detectability is also relevant for estimating other vital rates such as survival and movement.

In wildlife applications, the word **census** is reserved for the special and unusual case where detection probability equals 1.0; that is, all animals are counted. This might occur when studying a visible species on a small island, or with surveys in a narrow open transect where all individuals are seen. Botanists quite appropriately census numbers by counting all the plants in a plot, but a true census is rare for wildlife (and sometimes even for plants, where seeds or young plants can be missed). In fact, even the census of humans conducted by many governments is actually not a census at all but rather a count index with unknown detection probability (Box 4.1).

An **index** is a field count of animals or their sign that (hopefully) contains information about the relative number or density, but is not in itself an abundance or density estimate. Examples include mammal captures uncorrected for detection probability, as well as pellet counts, bird-call counts, track counts, numbers of burrow entrances or lodges, harvest numbers at check stations, questionnaires of wildlife sightings, and many others. Because they are typically cheaper to implement than formal estimators of abundance or density, indices are usually favored when money is tight, the species is difficult to observe directly, and/or when the questions are of such broad scale that more intensive estimators are impractical.

As an indirect assay of abundance, the utility of indices must be judged against how well they track changes in absolute or relative abundance across time, space, habitat types, or management treatments. An index can reliably indicate trends over time or relative difference across space only if its relationship to true abundance remains linear and constant, or at least does not change systematically (Bart et al. 2004).

If the relationship between the index and abundance does vary, you will not know whether you are seeing real changes in abundance or changes in the index/abundance relationship (Nichols & Pollock 1983, Tallmon & Mills 2004). For example, fur-

Box 4.1 Example of a non-census: the US "census"

The US Constitution mandates a count of the population in each state every 10 years to apportion the 435 seats in the US House of Representatives and to distribute federal funds to the states. Although it is called a census, the logistics are just too daunting for it to have a detection probability of 1.0; even Thomas Jefferson, who initiated the first census, noted that some persons had been missed. Ignoring that detection probability is less than 1.0 and relying on the count alone means that it is really a US index, with no known relationship to true census population size.

Statisticians proposed for the year 2000 census a transition from index to population estimate. Prior to the traditional census (where an intensive count is followed by random sampling of non-responding housing units) a totally independent set of 750,000 housing units would be picked. After determining the number of housing units present in both counts, the abundance estimate would be:

$$(\text{Count 1} * \text{Count 2})/\text{number of matching housing units}$$

The census bureau calls this the one number census or dual system estimation, but we will recognize it as the Lincoln–Petersen estimator described later in this chapter.

The fascinating part of this story is that this straightforward proposal to move beyond an uncorrected count index has stirred enormous political debate, because of the possibility that different groups of citizens may have differing detection probability, which would adjust the estimates of numbers and thereby shift political power. Even the US Supreme Court ruled that a 1976 federal census law "directly prohibits the use of sampling in the determination of population for the purposes of apportionment." Apparently we have a way to go to educate some that a complete count – a census – is impossible with hundreds of millions of people, and that a solid sampling strategy is the best way to move from an index to a reliable population estimate.

Source: Wright (1998).

trapping data are probably a poor index of abundance because the number of trapped animals will in large part reflect trapper effort, which in turn will be driven by economics and social norms. Obviously, some control on the constancy of the relationship between the index and abundance can be exerted by the researcher; in a bird-call index survey, for example, sampling could be restricted to the same time of day or year, and steps taken to minimize observer bias. Also, some indices better lend themselves to testing for constancy of the relationship between index and abundance; for example, bird calls or counts of captured animals can be tested statistically for constant detection probability (MacKenzie & Kendall 2002).

Clearly, then, some indices will perform better than others at portraying changes over time or relative differences across habitats or treatments. However, a growing movement argues that instead of hoping that an abundance index will reliably indicate relative differences in abundance over time and space, it is better to either directly estimate abundance (see below) or to switch to a different state variable such as pro-

portion of area occupied (MacKenzie et al. 2005). Certainly, indices will almost always fail as descriptors of absolute abundance, because the relationship between index and abundance will rarely be both constant and known. In short, if the goal is to estimate how many individuals are actually in a population, an index will not do it; you need to use a statistically based estimator (Fig. 4.1), perhaps based on transect sampling or capture–mark–recapture (CMR).

Transect methods for estimating abundance

It is easy to envision counting animals on either side of a line that you sample by walking, driving, riding a horse, or flying. These transects have some known width. If all animals in a series of transects were detected, the abundance in the study area would simply be the number counted divided by the proportion of area sampled in all the transects (Thompson 2002). But if some animals are likely to be missed in a transect, the probability of detection must be estimated (eqn. 4.2). One way to do this is **double sampling** (or ratio estimates): incomplete counts are made over an extensive area (e.g. counts on transects in a helicopter or airplane) while a simultaneous complete census on the ground at a subset of the transects provides an estimate of detection probability (which equals mean aerial count/mean ground count) for the aerial count. Similarly, **multiple observers** can count animals, with the estimated detection probability based on overlap in observations (Lancia et al. 2005). Two of the most widely used approaches to estimate abundance on transects include distance sampling and sightability models.

Distance sampling

Distance-sampling techniques are based on the idea that detection probability decreases with distance from the observer, so detections at various distances can be used to estimate detection probability as well as abundance and density (Buckland et al. 2001). The most common applications of distance sampling include sighting animals at various distances from a line during a transect count, and sighting (or listening to calls) from the center point of a circle. The essential data needed to conduct distance sampling are the measured perpendicular distances from the center line of the transect to each animal seen. To minimize flushing animals as the observer draws close, you can estimate the straight-line distance from where you first see the animal, then use trigonometry to calculate the perpendicular distance (Fig. 4.2).

The distance data are used to estimate the probability of detection ($\hat{p}$). It is easiest to first see how this works graphically (Fig. 4.3) and then summarize the calculus. If all animals could be seen equally at all distances, we could draw a **perfect sightability rectangle** with a $\hat{p}$ value of 1.0 across all distances. Because the sightability is assumed to decline with distance, the curve showing animals sighted at different distances would only occupy a portion of the perfect sightability rectangle. Therefore, the overall $\hat{p}$ in Fig. 4.3 is the proportion of the perfect sightability rectangle under the curve.

Mathematically, we first estimate the probability of observing an animal, given that it is found at distance x from the line $[g(x)]$. Then the computer software (such as the

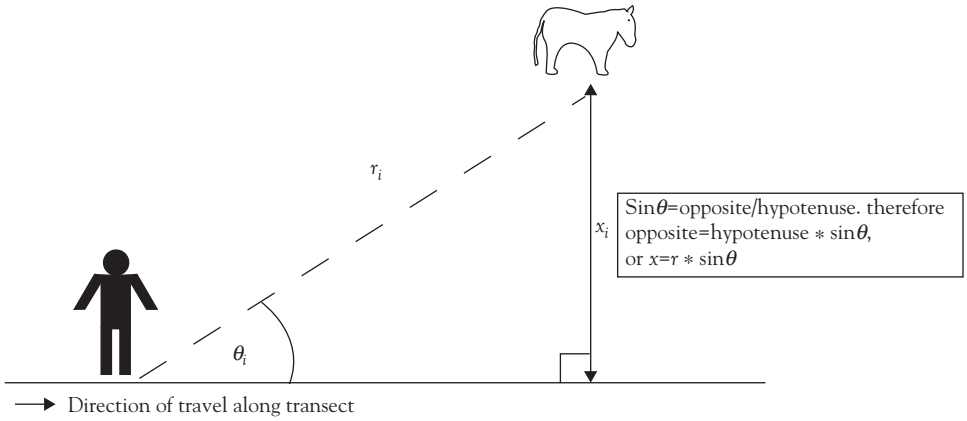


Fig. 4.2 How perpendicular distance for line-transect sampling can be calculated if the observer is not able to directly measure it. As the observer moves along the transect, they record the distance r_i from the line to where the animal is in the field. They also record the angle θ_i formed by the line of sight and the transect line. From these measures, the perpendicular distance (x_i) is $x_i = r_i(\sin \theta_i)$. Trigonometry really does have some interesting applications in applied biology!

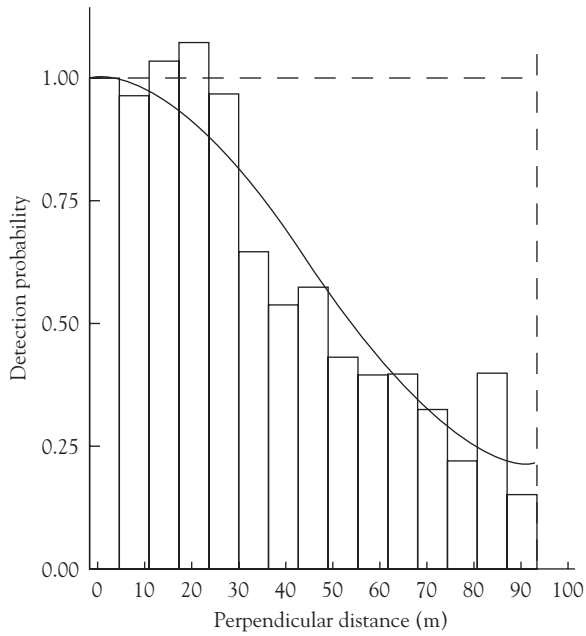


Fig. 4.3 Detection probabilities calculated from perpendicular distance-sighting data using line-transect sampling of wood ducks in forested habitat. The curve shows the fitted Fourier series estimator from the program DISTANCE. The dashed line shows what the overall detection probability would be if it were perfect at all distances (the perfect sightability rectangle); an intuitive estimator of overall detection probability for these data is the proportion of the perfect sightability rectangle under the curve. The increased detections at 15 and 20 m are assumed to be sampling anomalies. Modified from Kelley (1996). Reproduced by permission of The Wildlife Society.

program DISTANCE) integrates the detection function across all distances. In symbols, the probability of detection $\hat{p}_w$ is roughly the average of the detection probabilities across all distance categories from 0 to the maximum distance w:

$$\hat{p}_w = \frac{\left[\int_0^w g(x)\mathrm{d}x\right]}{w} \tag{4.3}$$

With an estimate of detection probability, abundance ($\hat{N}$) may be estimated using the canonical formula (eqn. 4.2), and with the known transect length (L) and width ($2w$; because you are sampling both sides of the transect line) density, D, can be estimated[1]:

$$D = \hat{N}/(2 * L * w) \tag{4.4}$$

Here are some key assumptions for distance sampling.

- All animals directly on the line are seen (this assumption can be relaxed).
- Animals are counted only once, and do not move before being sighted.
- Perpendicular distances are measured exactly. It is permissible to lump estimates into categories (for example, 10-m intervals) to deal with uncertainty in distance estimation[2].
- Sightings are independent, such that one animal does not cause others to be more or less likely to be sighted (more complicated models can account for this, as in herding animals).

As a rough guideline for required sample sizes, Buckland et al. (2001) recommend measuring distance to at least 40, and preferably 60–80 animals.

A common variant of distance sampling uses a single point instead of a line. For example, in point counts of birds the observer records over a specified time period (e.g. 5 minutes) all birds that are seen or heard, along with their estimated distance. Rather than perpendicular distances, radial distances are recorded, but the approach again assumes a monotonic decline in detection with distance. The detection curve is then exactly like line-transect sampling, with similar assumptions and analysis. Considerable error can arise in estimating distance when you can hear but not see the birds (Nichols et al. 2000a).

Sightability or observation probability models

Given that accounting for detectability is critically important yet hard to do for every time and place, sightability models can be developed and the resulting detection probabilities used to estimate abundance in future surveys. Often coupled with aerial

[1]Note: I have derived the density estimate in a way that connects distance sampling to the canonical approach to estimating abundance. Buckland et al. (2001) provide an excellent treatment of variance estimators, and show how density can be directly estimated without going through these steps.

[2]In practice, the most important part of the sightability curve is near the center line of the transect. As such, Buckland et al. (2001) recommend having smaller or more numerous distance categories nearer the line than away from the line.

transect surveys, a known population (usually radio-marked) is observed while additional variables likely to influence sighting probability (for example, snow conditions, group size, animal activity, vegetation type, etc.) are recorded simultaneously. The sightability model is developed from the relevant variables coupled with the proportion of known animals detected. In subsequent surveys, observers record data on the sightability variables and use the formula to convert the raw counts into an abundance estimate with variance (Unsworth et al. 1994).

As an example, Samuel et al. (1987) used radio-marked elk to develop a sightability model for aerial surveys in Idaho. Observers flew over the sampling area and documented whether they actually observed radio-marked individuals. Sightability for a group of a given size was modeled with logistic regression, where the radio-marked elk were either observed or not, with covariates that affect the sightability. They found that sightability increased with group size and decreased with vegetation cover. Specifically:

$$\text{Sighting probability} = \hat{p} = \frac{1}{1+e^{-[1.22+1.55\ln(\text{group size})-0.05(\%\,\text{vegetation cover})]}} \tag{4.5}$$

Thus for a future survey following similar protocols and in similar conditions to the Idaho study, a survey detecting a group of five elk in 70% vegetation cover would have a sighting probability of 0.55 (arrived at by putting 5 and 70 into the equation above). Each group count divided by its sighting probability (eqn. 4.2) gives an estimate of abundance for that group, and the sum of all abundances gives the overall abundance estimate for the sampled area.

The appeal of this approach is that after the sightability model has been developed and tested, future efforts require only counts and data on the model variables, without the need to directly estimate detection probability. Although one must be aware that the sightability model may only work well under the particular conditions for which it was developed, a sightability model thoughtfully applied from one place to another is better than a guess at numbers not framed in statistical sampling (Box 4.2).

CMR methods for estimating abundance

A different class of abundance estimators relies on capturing and marking, and recapturing again (with **capturing** and **marking** defined broadly and not necessarily literally; see Box 4.3)[3]. A **capture history** for each animal is generated, with a 1 denoting capture on a sample occasion and a 0 denoting no capture.

Two broad classes of CMR model may be distinguished: **closed-population models**, where the population is assumed to experience neither losses (by death or emigration) nor additions (by birth or immigration) during the period sampled, and **open-population models**, where losses and additions occur and can, in fact, be measured. The more complicated open models build on concepts and definitions from closed models.

[3]I will not cover **removal** abundance estimators whereby captured animals are permanently removed from the population (these are often used for fisheries or game animals where the harvest can be used to help estimate abundance).

Box 4.2 Application of a sample-based population survey to resolve a debate over numbers of mule deer

The Colorado Division of Wildlife (CDOW) used harvest, sex, and age data to estimate a relatively stable population of approximately 7000–9000 mule deer through the 1990s for a population of deer residing in northwest Colorado. Some sportsmen in Colorado believed that mule deer in the state were in serious peril, and used casual methodology (e.g. personal observations, outfitter guesses) to estimate the number of deer in this particular population at closer to 1750. The discrepancy led some sportsmen to accuse the CDOW of misleading the public by inflating estimates of deer population size. In a mediation process, it was agreed that an intensive aerial survey system developed in Colorado would be used to estimate numbers of deer in this population and, additionally, a sightability model developed for mule deer in Idaho would be applied to counts of deer using the CDOW survey system (Freddy et al. 2004). The CDOW system used intensive censuses of deer on randomly selected sample units or quadrats and assumed 100% detectability of deer on each sampled quadrat. In the area having the contested numbers of deer, the CDOW method led to an estimated population size of 6782 ± 2497 (90% confidence interval), whereas adjusting the counts using the Idaho sightability model led to an estimate of $11{,}052 \pm 3503$. Thus the original CDOW estimate of approximately 7000 deer was supported by two approaches to estimating deer population size, but no statistical support could be found for the sportsmen's estimate of 1750. This case study demonstrates that intuitive guesses at wildlife abundance without a formal sampling framework can be very wrong. It also shows the potential of under-estimating numbers of deer even when intense counts are conducted on relatively small parcels of land having complex cover and terrain features, underscoring the need to estimate detectability through sightability models or other approaches.

Closed CMR models of abundance

The simplest and most well-known closed-population model is the Lincoln–Petersen (LP) estimator of abundance based on two sampling periods. The LP method has a deep history, with applications stretching back to estimates of the human population of France in 1786 and of waterfowl of North America in 1930 (Williams et al. 2002:290). I will explain the LP in depth, both because it is commonly used to estimate abundance and because it forms the basis for understanding most other CMR estimators.

LP estimator with two samples

Suppose you want to estimate the abundance (N) of mice on a grid. You open traps in the afternoon, and the next morning you check traps and mark some mice uniquely (n_1 marked mice). You release them, and then a short time later (say, the next day) repeat the process. In the second sample you capture a total of n_2 mice, of which m_2 of these are marked. The best way to understand the LP abundance estimator without having to memorize it is to remember eqn. 4.2 and think of the first capture and marking session as the count and the second as the means for estimating the probability of detection:

Box 4.3 Methods of marking animals

Individual marks for wildlife studies usually conjures images of bird leg bands, turtle shell notches, and mammal ear tags. Although these traditional methods continue to be useful and widely applied, a number of other approaches are now available to mark animals (Silvy et al. 2005). Radio transmitters can be implanted in animals as small as shrews, worn as backpacks in birds, and equipped with global positioning satellite location monitors for larger animals. Telemetry has the benefit of being able to help a researcher differentiate movement from mortalities. Passive integrated transponder (PIT) tags are rice-grain-sized glass and metal cylinders that are injected under the skin; they do not transmit a signal but individual tags are recorded by an electronic reader passed over the animal. For amphibians, elastomer (rubbery paint) of different colors can be injected under the skin. If the mark does need not be permanent or individuals do not need to be distinguished (e.g. in short studies using two-sample Lincoln–Petersen methods), options could include paint balls, dyes, or hair clipping.

In some cases animals can be distinguished with noninvasive methods whereby the animal does not have to be captured. Animals with stripes, spots, or other patterns can be individually identified (e.g. body patterns on species ranging from salamanders to tigers), as can animals with distinctive scars from wounds (e.g. manatees scarred by boat propellers; Langtimm et al. 1998). As discussed in Chapter 3, noninvasive individual identification can also be obtained via genotype marks from bits of hair, feather, feces, or other material.

In all cases the choice of tag should result from careful deliberation, weighing possible negative effects on the animal against the efficiency of the method and the information to be gained from the particular study design. The ideal mark is humane, does not affect the response being studied (e.g. abundance, survival, reproduction, or behavior), is not prone to misidentification, and lasts reliably for the length of the study. Some can be used for multiple purposes; for example, toe clips provide an instant DNA sample as well as a mark that is permanent for small mammals and for some (but not all) amphibians.

$$\hat{N} = \frac{n_1}{\hat{p}} = \frac{n_1}{\left(\frac{m_2}{n_2}\right)} \tag{4.6}$$

This rearranges to give the intuitive abundance estimate:

$$\hat{N} = \frac{n_1 n_2}{m_2} \tag{4.7}$$

Although eqn. 4.7 shows the intuitive form of the LP estimator, it turns out that it is negatively biased, so the operational forms of the LP formula for abundance (the one to use in actual application) and its variance are:

$$\hat{N} = \left[\frac{(n_1+1)(n_2+1)}{(m_2+1)}\right] - 1. \tag{4.8}$$

$$\mathrm{var}\left(\hat{N}\right) = \frac{(n_1+1)(n_2+1)(n_1-m_2)(n_2-m_2)}{(m_2+1)^2(m_2+2)} \tag{4.9}$$

The square root of the variance gives the SE of the estimate, and the 95% confidence interval of the abundance estimate is:

$$\hat{N} \pm 1.96(\text{SE}) \tag{4.10}$$

An example of abundance estimation using the LP method is shown in Box 4.4.

Three key assumptions underlie the LP and other closed-population estimators of abundance.

1. The population is closed, so that $\hat{N}$ applies to both capture occasions. If deaths or emigration occur between the two periods, the LP estimate refers only to the population at the time of the first sampling period[4]. Immigration or births can also violate closure, in which case the LP estimate refers to population size at the time of the second sample (see Kendall 1999). Closure in CMR studies is usually biologically reasonable if the trapping sessions are close together in time (e.g. consecutive nights of trapping).
2. Marks are not lost or overlooked by observers. If tags are lost between the two samples, the LP estimate will be positively biased (because fewer of the n_1 animals will be available to become m_2 animals, so $\hat{p}$ will be biased low and therefore $\hat{N}$ will be biased high).
3. All animals are equally likely to be captured in each sample. For any CMR study, capture probability might vary in three primary ways: over time, among individuals, or in response to the animal having been trapped (Box 4.5). For LP, a change in capture probability over time is not a problem (the first sample merely introduces marked animals to the population to facilitate an estimate of $\hat{p}$ at the next session), but individual heterogeneity or trap response can bias the LP estimator. As an example of the effects of individual heterogeneity, remember from Chapter 3 that noninvasive genotyping of hair or scat to mark individuals can in some cases lead to a shadow effect whereby certain individuals share the same genotype (the same mark). These shadows are essentially genotypes more likely to be detected, biasing $\hat{p}$ high which causes a negative bias in $\hat{N}$ (Mills et al. 2000b). Finally, a trap-shy response leads to a positive bias in $\hat{N}$ while a trap-happy response leads to a negative bias.

A few other practical points about the LP estimator are worth mentioning. First, because it is a two-sample estimator, marks are only applied once and checked once which means that animals do not have to be individually identified. This makes LP unusual among CMR estimators in that simple batch marks – perhaps a dab of paint on the back – are sufficient to identify the marked animals. The one-time mark also means that the second capture session can be based on any method that obtains an estimate of (m_2/n_2), including the use of hunters or anglers to obtain the sample.

A related feature arising from the two-sample construction of the LP estimator is that the mark–recapture can often be improved by using different methods for the two

[4]Deaths during the second trapping event will not affect LP. If a death occurs during the first session (due to trapping or handling), the dead animal(s) should not be included with the n_1 animals but rather added to the abundance after $\hat{N}$ is estimated; the variance estimate is unaffected (Williams et al. 2002:293).

Box 4.4 An example of abundance estimation using the LP method and noninvasive sampling

Eastern North Pacific humpback whales can be uniquely (and noninvasively) identified from natural markings including pigmentation, scars, and ridging of the flukes. Population closure over 1–3 years can be assumed because the whales have high site fidelity to distinct feeding aggregations off the coast of California, Oregon, and Washington before migrating to wintering grounds off Baja California, mainland Mexico, and Central America. Some humpback whale data and abundance estimates are shown below. Many animals were captured (photographed) several times in a year, so the number of photographs is much greater than the number of uniquely identified whales (n_1 and n_2; Calambokidis & Barlow 2004).

Years for n_1 and n_2	Number of identification photographs in first year	n_1	Number of identification photographs in 2nd year	n_2	m_2	$\hat{N}$	95% Confidence interval around $\hat{N}$
1991 and 1992	668	269	1023	398	188	569	537–601
1992 and 1993	1023	398	512	254	173	584	547–620
1993and 1994	512	254	402	244	108	572	512–633
1994 and 1995	402	244	661	331	100	804	704–904
1995 and 1996	661	331	564	331	144	759	690–829

n_1, The number of individuals identified in photographs in the first year; n_2, the number of individuals identified in photographs in the second year; m_2, the number of individuals in the second year that had been identified in the first year.

capture sessions. With **mark–resight** methods, potential logistical advantages accrue with using a different method (sighting) for the recapture, and trap response and individual heterogeneity should be reduced if capture probabilities of the two samples are independent. Finally, two samples does not necessarily mean only 2 days of trapping. Multiple days can be collapsed into two samples of unequal length, so for example the first three nights of trapping could be sample one and the second two nights sample two. Collapsing multiple days into two samples can help deal with population closure (Kendall 1999), and has the benefit of increasing sample size per trapping event. However, if you can conduct sampling over more than 2 days, other CMR models such as those discussed next may be more appropriate.

Box 4.5 Three ways that CMR studies can violate the assumption of equal catchability (and how field researchers try to minimize the violations)

1. **Time** may change capture probabilities for all animals in different capture sessions. Weather, moon phase, or time of year could cause capture probability to change. Researchers attempt to minimize this violation of equal catchability by closing down traps when weather patterns are likely to affect capture probability.
2. **Heterogeneity among individuals** indicates that different animals have different capture probabilities. It can be caused by an animal's age, sex, dominance status, home-range location relative to trap locations, and so on. A possible solution to heterogeneity is to estimate abundance separately for classes of animals thought to have different capture probabilities (e.g. males and females). Individual heterogeneity can also be minimized by making sure that each animal is likely to encounter more than one trap during the course of daily movements; although the traps do not need to be in a uniform grid pattern (e.g. Karanth & Nichols 1998), regular trap spacing is usually used with at least four per home range. It can also be minimized by using different techniques – such as marking, telemetry, sighting, and so on – on different occasions (for example, mark the first session, resight the second).
3. **Behavioral response** arises when an animal becomes more or less likely to be captured after the first capture. Trap-happy animals are more likely to be captured again (perhaps due to the novelty or security of the trap, or the allure of free food). Trap-shy animals are less likely to be trapped after first exposure. Trap happiness can be decreased by pre-baiting traps (which is also a good idea because it tends to increase capture probability in general), trap shyness by minimizing the time and severity of handling.

Closed-population estimates requiring three or more samples

Although the LP method is robust to some forms of unequal trappability it is not robust to all. If you are able to employ more than two capture sessions, then you will have a sufficient data stream to first test which forms (if any) of unequal trappability are apparent in the data, and then be able to employ an estimator robust to the identified deviations from equal catchability. You will not be doing this by hand; the approaches were originally codified in the computer program CAPTURE in the late 1970s (Otis et al. 1978, White et al. 1982), and more recently in the program MARK (Cooch 2001). For a particular data-set, the model whose assumptions of capture-probability structure most closely approximate the trapping data is chosen to provide the most accurate (least biased, most precise) estimate of abundance. Briefly, here are the models:

- Model M_O: equal catchability. Every animal has the same probability of capture for each sampling period in the study (no behavioral response, individual heterogeneity, or temporal variation).

- Model M_h: individual heterogeneity. Individuals have different capture probabilities.
- Model M_b: behavioral response. All animals initially have the same capture probability but after first capture may become trap happy or trap shy.
- Model M_t: time-variation model. Probabilities of capture change from trap period to trap period but within a period all animals have an equal chance of capture. Essentially an extension of the LP model.
- Models M_{bh}, M_{th}, M_{tb}, and M_{tbh}: various forms incorporating multiple deviations from equal trappability.

Open CMR models of abundance

In many cases, the length of the study makes it impossible to assume that the population is closed to additions and losses. In the mid 1960s Richard Cormack, followed closely and independently by George Jolly and George Seber, developed a modeling framework to provide estimates of vital rates in open populations. Abundance estimates in open populations are therefore based on **Jolly–Seber** (JS) models, while survival estimates are called Cormack–Jolly–Seber (CJS) models.

The Jolly–Seber model for estimating abundance is analogous to the LP estimate, except that the pieces are generalized to more than two sessions (with subscript *i* referring to the session) and capture probability focuses on marked animals only, to account for the open population (Pollock et al. 1990). As with LP, a total of n_i animals are caught at time *i*. Although m_i marked animals are captured at time *i*, in an open population with a probability of detection of less than 1.0 we know that some of the animals previously marked have died or emigrated. Therefore $\hat{M}_i$, the number of marked animals alive and in the population just before occasion *i*, is estimated using information on animals captured both before and after *i* (and therefore known to be alive at *i*)[5]. Thus, the Jolly–Seber abundance estimate for sample *i* is:

[5]While I am trying to avoid gory details I do not want to leave a big black box. $\hat{M}_i$, the number of marked animals in the population at time *i*, is estimated based on the following (Pollock et al. 1990).

- R_i, The number of the n_i animals caught at *i* that are successfully released with marks (this could be less than n_i if there were losses during capture).
- r_i, The number of the R_i subsequently recaptured after *i*.
- m_i, The number of marked animals caught in the *i*th sample.
- z_i, The number of marked animals not captured at *i* but recaptured after *i*.

If we assume the probability of ever seeing again a marked animal that has just been released is the same as the chance of seeing again a marked animal that was not captured at time *i* then $\frac{r_i}{R_i} = \frac{z_i}{(M_i - m_i)}$, which rearranges to: $\hat{M}_i = m_i + (R_i z_i / r_i)$. So in words, we estimate the number of marked animals alive in a sample based on both the marked animals captured (m_i) and the number of animals not captured but known to be alive because they were captured later ($R_i z_i / r_i$).

$$\hat{N}_i = \frac{n_i \hat{M}_i}{m_i} \tag{4.11}$$

Robust design

In 1982, wildlife biometrician Ken Pollock made the simple but profound observation that closed- and open-population models could be combined to take advantage of each of their strengths. Closed-population models can be robust to unequal capture probabilities among individuals or over time, but are only valid over relatively short periods of time during which no additions or losses to the population occur. Open-population models allow estimates of abundance in the face of gains and losses, but are not as precise, or as easily accommodating to unequal catchability in the form of individual heterogeneity or behavioral responses (Lebreton et al. 1992, Burnham & Anderson 2002). Pollock's suggestion was to use a **robust design** such that a long-term study of an open population is implemented as a sequence of short-term studies of closed populations (Pollock et al. 1990). The robust design is implemented with several primary sampling occasions, between which the population is likely to be open to gains and losses (Fig. 4.4). Each primary session includes several secondary sampling periods (preferably four or more), analyzed using closed models. For example, there may be four nights of mark–recapture trapping once a month for 5 months (Fig. 4.4). Abundance is estimated with closed-population models within each month. Cormack–Jolly–Seber survival estimates (described below) are based on pooled data across the secondary periods (e.g. each animal is recorded if it was captured at least once during a four-night set of secondary periods).

The robust design provides a powerful engine for estimating vital rates (Box 4.6). In addition to good estimates of survival and abundance, with two age classes you can estimate new individuals entering the population from both immigration and *in situ* reproduction (Nichols & Pollock 1990, Nichols & Coffman 1999), as well as temporary movement into and out of the study area (Kendall et al. 1997, Bailey et al. 2004a, 2004b; Chapter 10 will show the application of this method to estimating dispersal).

A note on density estimation in capture-mark-recapture studies

Often the density of animals per unit area is of more interest than abundance per se. Intuitively, density is simply the abundance estimate divided by the area of the trapping grid. However, animals whose home range barely overlaps the trapping grid, or animals that come onto the grid from outside its perimeter, make the effective size of the trapping grid larger than the actual grid. The size of the **effective trapping grid** depends on the animal, the grid shape, and the study design, so it should be estimated for each study.

A practical approach assumes that animals off the grid are just as likely to move toward the grid as away, and that the distance moved can be indexed by the maximum distance between captures for animals on the grid. A boundary strip equal to half

Box 4.6 Some insights using the robust design

The robust design provides a powerful framework for estimating abundance and survival within populations, and for connectivity (movement) among populations. I will save examples of quantifying movement for Chapter 10. Here are two case studies to show how detection and survival can be quantified with robust design.

Case study 1: how many salamanders are missed during monitoring counts? (Bailey et al. 2004a, 2004b)

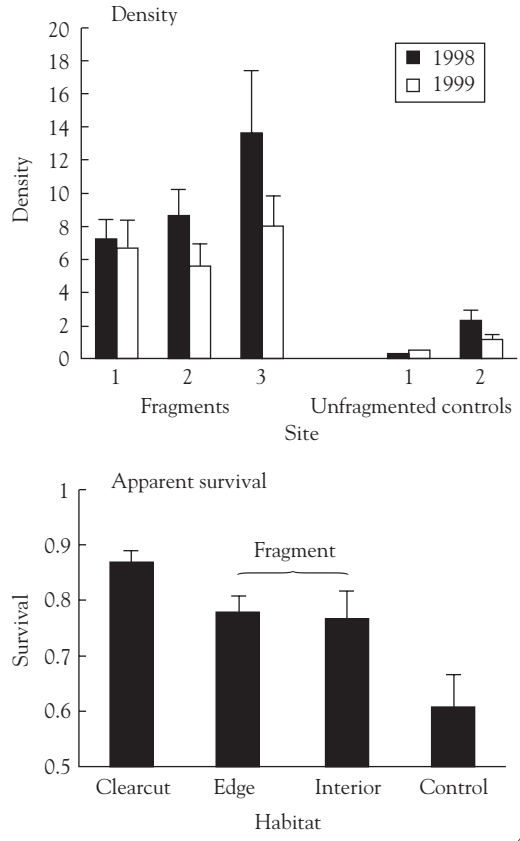
Concerns about global amphibian declines have led to widespread efforts to monitor amphibians over time and space. Plethodon salamanders have received special attention because they may be particularly susceptible to human-caused stressors, and therefore may be good indicator species (Chapter 13). But how good are raw counts of salamanders as a tool for population monitoring? Working in Great Smoky Mountains National Park, Larissa Bailey and colleagues were interested in the likelihood of being able to actually detect salamanders present in an area, and in the stability of detection probability over time and space. For salamanders that spend a lot of time underground, the probability of capture or detection is a product of the probability that the salamander is near the surface and thus available to be captured multiplied by the probability of catching a salamander given that it is near the surface during a set of secondary samples. The first bit is the probability that the salamander has not temporarily emigrated, or is otherwise temporarily unavailable for sampling; this is a problem for lots of species that pop up briefly but then seem to disappear for a while (e.g. marine mammals that only visible when they come to the surface or snow geese that are only detectable when they are breeding). The second bit is simply the capture probability.

The robust design of sampling involved four primary periods 6–10 days apart and lasting for 3–4 consecutive daily secondary samples, repeated for 3 years. What did they find? The average probability of a salamander being available near the surface (that is, not temporarily emigrated to the soil depths) was only 13%, and the average probability of catching a salamander near the surface was 30%. Together, that means that the probability of detecting a salamander that occurs in a particular plot is only 4%. Furthermore, the capture probability varied across years, across habitat type, and across species. Therefore, actual population size could decline quite a bit with relatively little change in a count-based index. Conversely, a count index could fluctuate a lot because of changes in detection probability, while actual abundance (what we care about) changed very little. Therefore, count indices are not reliable in this case and if abundance is of interest, formal CMR estimators should be used.

Box 4.6 Continued

Case study 2: survival of voracious deer mice on clearcuts and forest fragments (Tallmon et al. 2003)

Deer mice can be voracious seed predators, and are known to respond positively to many human perturbations, so it is of interest how forest fragmentation affects their density and survival. In a study in southwest Oregon, David Tallmon and colleagues used a robust design to trap four primary periods in each of two summers. The first primary session consisted of eight consecutive nights of secondary samples, whereas the other three primary periods were four consecutive nights each. A 16-day interval separated primary sessions (except for one year where 20 days separated two sessions). From this robust design coupled with 340 mouse captures it was possible to calculate density in forest fragments compared with unfragmented controls during the last primary session of each summer (closed models), as well as apparent survival in different fragmentation habitat types over 20-day intervals (see figure). So, deer mice love forest fragmentation!



Estimates of deer mouse population density and survival in a fragmented landscape. From Tallmon et al. (2003). Reproduced by permission of the ESA.

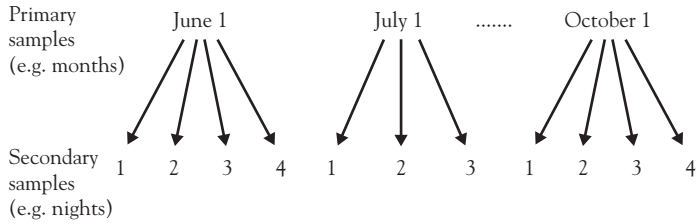


Fig. 4.4 Pollock's robust design. Here I use five monthly primary sampling periods (June–October, with the dots indicating August and September) and three- or four-night secondary periods. Notice that the number of secondary periods can differ among primary periods, which is nice because rain shuts down trapping efforts, trucks break down, field assistants get sick, and so on. Apparent survival (and recruitment and movement) can be estimated between primary periods (e.g. monthly survival using Cormack–Jolly–Seber methods), and abundance and capture probabilities can be estimated using closed models across secondary periods (e.g. abundance for each month).

the mean maximum distance moved is added all the way around the grid to provide an effective grid size (Wilson & Anderson 1985, Karanth & Nichols 1998)[6]. The density estimate is based on estimated abundance divided by estimated effective area, with its variance accounting for uncertainty in both abundance and effective grid size.

Survival estimation

Three main classes of survival estimator can be distinguished by whether or not all animals can be relocated (known fate) or whether survivors are recorded (CMR) or deaths are recorded (band recovery or return)[7]. The computer program MARK (Cooch 2001) is the workhorse for all of these analyses.

Known-fate models

In cases where a method such as radiotelemetry allows for certainty in relocating or detecting an animal that is alive and part of the study, **known-fate models** or **complete follow-up models** can be used. The simplest way to think about survival estimation using known-fate models is to start with the ideal case where the surviving animals (x) relative to the total number (n) are known unambiguously. The estimated

[6]Alternative approaches for estimating effective grid size include **nested grid** analysis and direct measurement using radio-collared animals (Lancia et al. 2005).

[7]Many other methods combine and extend the categories (see Further reading at the end of the chapter, and Lebreton et al. 1992, Winterstein et al. 2001).