12 Counting animals

12.1 Introduction

Knowledge of the size or density of a population is often a vital prerequisite to managing it effectively. Is the population too small? Is it too large? Is the size changing, and if so in what direction? To answer these questions, we may have to count the individual animals, or we may be able to obtain adequate information by way of an indirect indication of abundance. In any event, we need to know when a census is necessary and how it might be done.

Although *census* is strictly the total enumeration of the animals in an area, we use the word here in its less restrictive sense of an *estimate of population size or density*. Such an estimate may come from a total count, from a sampled count, or by way of an indirect method such as mark–recapture.

Closely related to the census is the *index*, a number that is not itself an estimate of population size or density but which has a proportional relationship to it. The number of whales seen per cruising hour is an index of whale density. It does not tell us the true density but it does allow comparison of densities between areas and between years. *Indices provide measures of relative density* and are used only in comparisons. They are particularly useful in tracking changes in rates of increase and decrease.

Almost all decisions on how a population might best be managed require information on density, on trends in density, or on both. There are many methods to choose from and these differ by orders of magnitude in their accuracy and expense. Hence, before any censusing is attempted, the wildlife manager should ask a number of questions:

- Do I need any indication of density and what question will that information answer?
- Is absolute density required or will an index of density suffice?
- Will a rough estimate answer the question or is an accurate estimate required?
- What is the most appropriate method biologically and statistically?
- How much will it cost?
- Do we have that kind of money?
- Would that money be better spent on answering another question?

The trick in obtaining a usable estimate of abundance is to choose the right method. What works in some circumstances is useless in others. Here we present a wide range of options and indicate the conditions under which each is most effective.

12.2 Total counts

The idea of counting every animal in a population, or in a given area, has an attractive simplicity to it. It is the method used by farmers to keep track of the size of their flocks. No arithmetic beyond adding is called for and the results are easily interpreted.

Wildlife Ecology, Conservation, and Management, Third Edition.
John M. Fryxell, Anthony R. E. Sinclair and Graeme Caughley.

Companion Website: www.wiley.com/go/Fryxell/Wildlife

This is why total counting was once very popular in wildlife management and why it is still the most popular method for censusing people.

Total counts have two serious drawbacks: they tend to be inaccurate and expensive. Nonetheless, they have a place. The hippopotami (*Hippopotamus amphibius*) in a clearwater stretch of river can be counted with reasonable facility from a low-flying aircraft. The number of large mammals in a 1 km^2 fenced reserve can be determined to a reasonable level of accuracy by a drive count. It takes much organization and many volunteers but it can be done. Every nesting bird can be counted in an adélie penguin (*Pygoscelis adeliae*) rookery, either from the ground or from an aerial photograph. That is an example of a "total count" providing an index of population size, because more than half the birds will be at sea on any given occasion.

Total counting of large mammals over extended areas was common in North America up to 1950. Gill *et al*. (1983) describe the system in Colorado:

> Biologists attempted to count total numbers of deer comprising the most important "herds" in the state. Crews of observers walked each drainage within winter range complexes and counted every deer they encountered. The sum of all counts over every drainage of a winter range was taken as the minimum population size of that herd (McCutchen 1938; Rasmussen and Doman 1943).

Total counting of large mammals from the air was a standard technique in Africa in the 1950s and early 1960s. Witness the total counts of large mammals on the 25 000 km^2 Serengeti–Mara plains (Talbot and Stewart 1964) and 20 000 km^2 Kruger National Park, South Africa:

> trends in population totals, spatial distribution, and social organization are obtained by means of surveys by fixed-wing aircraft. Due to the size of the Kruger National Park these (total count) surveys require three months to complete and are consequently undertaken only once annually (i.e. during the dry season from May to August (Joubert 1983)).

These massive exercises continued in Kruger until 1996, when they were abandoned due to cost. Similar methods are used to count pronghorn antelope in the United States (Gill *et al*. 1983). Total counts continue to be used on species that are highly clumped with wide spacing between clumps. For example, both African buffalo (*Syncerus caffer*) and African elephant (*Loxodonta africana*) live in widely dispersed large herds of several hundred animals in both Serengeti and Kruger National Parks, and total counting is still the best method of obtaining their population size. This is because the dispersion pattern of these species means that sample counts produce very high variances and hence wide confidence limits. A simulated transect sampling strategy for a known dispersion of buffalo showed that over 90% of the area had to be sampled before confidence limits were reduced to acceptable values (< 15% of the estimated total). Thus, total counting was more efficient because it was logistically easier than rigidly flown transects (Sinclair 1973). Similarly, the clumped distribution of pronghorn antelope (*Antilocapra americana*) in North Dakota produced such high variances from a variety of sampling strategies that Kraft *et al*. (1995) advised against using samples to estimate numbers.

12.3 **Sampled counts: the logic**

There are two important areas in which scientific thinking differs from everyday thinking: the selection of a random or unbiased sample and the choosing of an appropriate experimental control. Knowing how to sample and knowing how to design an experiment that gives an unambiguous answer are the two attributes distinguishing science from ideology. Sampling is the technique of drawing a subset of sampling units from the complete set and then making deductions about the whole from the part. It is used all the time in wildlife research and management, but often incorrectly.

The next section takes us through some of the mystery of sampling. It explores what actually happens when we sample a population in several different ways, making the point that the true estimate is independent of whatever mathematical calculations are applied to the data.

12.3.1 *Precision and accuracy*

If a large number of repeated estimates of density have a mean that does not differ significantly from the true density then each estimate is said to be *accurate* or *unbiased*. Accuracy is a measure of *bias error*. If that set of estimates has little scatter then the estimates are described as *precise* or *repeatable*. Precision is a measure of *sampling error*. A system of estimation may provide very precise estimates that are not accurate, just as a system may provide accurate but imprecise estimates. Ideally, both should be maximized, but often we must choose between one and the other according to what question is being asked. For example, is density below a critical threshold of one animal per square kilometer? Here we need an accurate measure of density and may be willing to trade off some precision to get it. But if we had asked whether present density was lower than that of last year then we would need two estimates, each of high precision. Their accuracy would be irrelevant so long as their bias was constant. Most questions require precision more than accuracy. Precision is obtained by rigid standardization of survey methods, by sampling in the most efficient manner, and by taking a large sample.

12.3.2 *Bias errors*

Bias errors derive from some systematic distortion in the counting technique, the observer's ability to detect animals, or the behavior of the animals. Often, but not always, the bias produces an undercount. Thus biases can accrue from sampling schemes that do not properly sample all habitats (e.g. using roads that avoid hills or riverine areas); from the observer missing animals on transects (because there are too many animals, because in counting one group the observer overlooks another, or simply because of observer fatigue); or from animals being hidden in thickets, under trees, or underwater.

The best way to measure bias error is to compare the census estimate with that from a known population. Pollock and Kendall (1987) review this method, along with the use of a subpopulation of marked animals, mapping with multiple observers, line transect sampling, and multiple counts on the same area. Visibility corrections have been calculated by comparing fixed-wing aerial surveys of waterfowl with ground counts (the known or unbiased population) (Arnold 1994; Bromley *et al*. 1995; Prenzlow and Lovvorn 1996). A similar approach was used to estimate bias in counts of wood stork (*Mycteria americana*) nests in Florida (Rodgers *et al*. 1995) and great blue heron (*Ardea herodias*) nests in south Carolina (Dodd and Murphy 1995). Moose usually live in dense habitats where they are difficult to see. Rivest *et al*. (1990) compared fixed-wing

surveys of moose with the more accurate subsample surveys by helicopter to correct for visibility bias, an approach also used for counts of chicks in osprey (*Pandion haliaetes*) nests (Ewins and Miller 1995).

12.3.3 *Sampling frames*

Before an area is surveyed to estimate the number of animals on it, it must be divided into *sampling units*, which cover the whole area and are non-overlapping. The sampling units might comprise areas of land if we count deer, or trees if we count nests, or stretches of river if we count beavers or crocodiles. To allow us to sample from this *frame list* of sampling units, the list must be complete for the whole area. Hence, the frame of units contains all the animals whose numbers we wish to estimate.

For purposes of explanation, we use the first example: sampling units of land. The survey area may be divided up into units in any way the surveyor desires: into quadrats, transects, or irregular sections of land, perhaps delimited by fences. The choice is a compromise between what is most efficient statistically and what is most efficient operationally.

12.3.4 *Sampling strategies*

Suppose that we wished to estimate the number of kangaroos or antelopes in a large area by counting animals on a sample of that area. Several strategies are open to us. We could sample quadrats or transects, we could select these sampling units systematically or randomly, and, if the latter, we could ensure that each sampled unit occurred only once in the sample (sampling without replacement) or that the luck of the draw allowed units to be selected more than once (sampling with replacement). The efficiency of these systems will be demonstrated with the hypothetical data of Table 12.1, which may be thought of as the number of kangaroos standing on each square kilometer of an area totaling 144 km^2. In all cases 1/3 of the area will be surveyed. We can test the accuracy of the method by determining whether the mean of a set of repeated estimates is significantly different from the true total of 1737 kangaroos. The precision of a sampling system is indicated by the spread of those repeated and independent estimates, and that spread will be measured by the standard deviation of those estimates:

$$s = \sqrt{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{N}\right)/(N-1)}$$

where x is an independent estimate of total numbers and N is the number of such repeated estimates.

Table 12.1 Simulated dispersion of kangaroos on a 1 × 1 km grid of 144 cells. Marginal totals give numbers on 1 × 2 km transects oriented both across and down the region.

1	2	7	4	7	14	9	18	24	22	19	15	142
0	1	5	6	12	11	9	15	20	21	27	28	147
2	3	5	6	10	13	16	20	160	14	19	21	147
1	4	4	6	9	13	14	17	20	16	25	20	149
2	2	5	7	10	12	16	19	20	16	18	22	149
2	4	5	6	9	12	16	22	18	18	21	23	156
0	2	5	8	4	7	11	13	17	16	21	30	134
1	0	4	9	8	10	11	16	14	20	17	17	127
0	4	2	7	8	11	11	11	12	19	22	21	128
0	2	5	8	8	12	16	20	24	25	23	25	168
1	0	4	9	8	8	8	17	17	14	18	22	126
2	5	7	6	12	12	13	15	20	21	20	23	156
12	29	58	82	105	135	150	203	222	222	250	269	1737

We will first sample 1 km^2 quadrats randomly with replacement: *sampling with replacement* (*SWR*). The quadrats are numbered from 1 to 144 and a sample of 48 of these is drawn randomly. Quadrats 27, 31, 50, and 53 are drawn twice and quadrat 7 three times, but since these are independent draws they are included in the sample as many times as they are randomly chosen. The quadrat is *replaced* in the frame list after each draw, allowing it the chance of being drawn again. The number of kangaroos in this sample of quadrats totals 523, and since we have sampled only a third of the quadrats we multiply the total by 3 to give an estimate of animals in the study area: 1569.

Note that this answer is wrong in the sense that it differs from the true total, known to be 1737 (i.e. it is not accurate). This disparity is called *sampling error*, which is quite distinct from *errors of measurement* resulting from failure to count all the animals on each sampled quadrat.

We now repeat the exercise by drawing a fresh sample of 48 units and get a sampled count of 493 kangaroos, which multiplies up to an estimate of 1479. The third and fourth surveys give estimates of 1836 and 1752. This exercise is repeated a thousand times with the help of a computer. The thousand independent estimates have a mean of $\overline{x} = 1741$, very close to the true total of 1737. We can be confident, therefore, that this sampling system produces accurate (i.e. unbiased) estimates. The thousand independent estimates have a standard deviation of $s = 153$, which tells us that there is a 95% chance that any one estimate will fall in the range $\overline{x} \pm 1.96s$ or 1741 ± 300, between 1441 and 2041. The standard deviation of a set of independent estimates is the measure of the efficacy of the sampling system and hence of the precision of any one of the independent estimates. It can be estimated from the quadrat counts of a single survey (see Section 12.5.1), and when estimated in this way it is called the *standard error of the estimate*. Hence the *standard error of an estimate is a calculation of what the standard deviation of a set of independent estimates is likely to be*.

With this background, we can now compare the efficiencies of several sampling systems.

12.3.5 *Sampling with or without replacement?*

When we use *sampling without replacement* (*SWOR*), a quadrat may be drawn no more than once. This is in contrast to the previous system, which allowed a quadrat to be selected by the luck of the draw any number of times. We draw a unit, check whether it has been selected previously, and if so reject it and try again. Having drawn 48 distinct units, we calculate density. The sampling is again repeated a thousand times, yielding 1000 independent estimates – each based on a draw of 48 units – of the total number of animals. We know the true total to be 1737. These 1000 estimates have a mean of 1743 and a standard deviation of 131, which is appreciably lower than the $s = 153$ accruing from SWR.

The gain in precision with SWOR reflects the slightly greater information on density carried by the 48 distinct quadrats of each survey. SWOR is always more precise than SWR for the same sampling fraction, the relationship being:

$$s(\text{SWOR}) = s(\text{SWR}) \times \sqrt{1 - f}$$

where f is the sampling fraction, in this case 0.333. The s(SWR) from the 1000 repeated surveys is 153, and from this we can estimate, without needing to run the simulation,

that the precision of the analogous SWOR system would be about:

$$s = 153 \times \sqrt{0.666} = 125$$

Our empirical *s*(SWOR) is 131, which is much the same as the $s = 125$ predicted theoretically.

However, it is not as simple as that. The quadrats chosen more than once in an SWR sample are not surveyed more than once, so the time taken for the survey is shorter. In the example, only about 41 of the 48 units drawn in an SWR sample would be distinct units, the other seven being repeats. To compare the precision of an SWOR sample with that of an SWR sample entailing the same groundwork, we would have to draw by SWR about 58 units. Ten are repeats, "free" units that do not need to be surveyed a second time. Intuitively, we would assume that the SWR sample of 48 distinct units and 10 repeats must give a more precise estimate than the SWOR sample with its 48 distinct units, none repeated. Not so. The smaller SWOR samples provide estimates that are more precise by a factor of $\sqrt{(1 - \frac{1}{2}f)}$. In all circumstances, SWOR is more precise than SWR (Raj and Khamis 1958). Precision is increased by rejecting the repeats and cutting the sample size back to that of the analogous SWOR sample.

Why then, if SWOR is always better, is SWR often used? First, when the sampling fraction is low (< 15%), the precision of the two systems of sampling is similar. At $f = 0.1$ there is only a 5% difference in precision, reflecting the low likelihood of repeats at low sampling intensity. Most sampling intensity in wildlife management is of this order. Second, it is often convenient to sample with replacement when an area is traversed repeatedly by aerial-survey transects. There is not the same necessity to ensure that no transect crosses another or overlaps it. This is a useful flexibility for an aerial survey in a strong crosswind or for a ground survey in thick forest.

12.3.6 *Transects or quadrats?*

A frame of transects is a good or a bad sampling system according to how it is oriented with respect to trends in density. The dispersion of Table 12.1 has a marked increase in density from left to right. The precision of the estimate of total numbers will be relatively high if the transects are oriented along this cline but low if oriented at right angles to it. This can be demonstrated empirically by sampling the column totals at one-third sampling intensity. Each column represents a transect and each survey comprises four transects, randomly chosen. A thousand independent surveys produce a standard deviation of estimates of 512 for SWR and 427 for SWOR. If these transects were oriented at right angles, so that the rows rather than the columns formed the transects, the standard deviation of estimates of a thousand independent surveys would be approximately 80 for SWR and 69 for SWOR. In this case, precision is increased enormously by swinging the orientation of the transects through 90°.

Transects should go across the grain of the country rather than along it, should cross a river rather than parallel it, and should go up a slope rather than hug the contour. They should be oriented such that each samples as much as possible of the total variability of an area. In essence, we must *ensure that the variation between transects is minimized and therefore that the precision of the estimate is maximized.*

Much the same principle adjudicates between the use of quadrats as against transects. So long as the frame of transects is oriented appropriately, the resultant estimate will be more precise than that from a set of quadrats whose area sums to that of

the transects. *The more clumped ithe distribution of the animals, the greater the gain in precision of transects over quadrats*. A quadrat is likely to land in a patch of either high or low density, whereas a transect is more likely to cut through areas of both. Table 12.2 shows that transects oriented along the cline in density of Table 12.1 provide estimates six times more precise than do quadrats of the same size and number.

12.3.7 *Random or nonrandom sampling?*

Sampling strategies grade from *strictly random* to *strictly systematic*. The region in between is described as *restricted random sampling*. One might decide, for example, to sample randomly but to reject a unit that abuts one previously drawn. Or one might break the area into zones (strata) and draw the same number of samples randomly from each zone. These two strategies depart from the requirement of strict random sampling, whereby each sampling unit has the same probability of selection. The extreme is systematic sampling, in which the choice of units is determined by the position of the first unit selected.

Systematic or restricted random sampling has several practical advantages over strict random sampling. First, it encourages or enforces SWOR, which, as we have seen, leads to a more precise estimate. Second, it reduces the disturbance of animals on a sampling unit caused by the surveying of an adjoining unit. This is particularly important in aerial survey, where the noise of the aircraft can move animals off one transect on to another. Third, any deviation from strictly random sampling tends to increase the precision of the estimate, because the sampled units together provide a more comprehensive coverage of total variability. Table 12.2 demonstrates this for our example. The standard deviation of a thousand independent surveys is lower for restricted random sampling than for random SWOR, and lower still for systematic sampling.

Statisticians do not like nonrandom sampling because the precision of the estimate cannot be calculated from a single survey. The formulae given in Section 12.5.1 for calculating the standard error of an estimate are correct only when sampling units are drawn at random, and they will tend to overestimate the true standard error when restricted random or systematic sampling is used. But not always. If a systematically drawn set of sampling units tends to align with systematically spaced highs and lows of density, the standard error calculated on the assumption of random sampling will be too low and the estimate of density will be biased.

Table 12.2 The effect of the sampling system on the precision of an estimate. All systems sample one-third of an area of 144 km^2 containing the dispersion of kangaroos simulated in Table 12.1. Each sampling system is run 1000 times to provide 1000 independent estimates of the true total of 1737.

Sampling system	Mean estimate	Standard deviation of 1000 estimates
Large quadrants, n = 4		
Random with replacement	1746	487
Random without replacement	1738	414
Small quadrants, n = 48		
Random with replacement	1741	153
Random without replacement	1743	131
Transects parallel to the density cline, n = 4		
Random with replacement	1732	80
Random without replacement	1734	69
Restricted random	1730	57
Systematic	1736	48

In practice, this tends not to happen. It is entirely appropriate to sample systematically or by some variant of restricted random sampling and to approximate the standard error of the estimate with the equation for random sampling. One can be confident that the estimate is unlikely to be biased and that the true standard error is unlikely to exceed that calculated.

12.3.8 *How not to sample*

There are a number of traps that sampling can lure one into, which can result in a biased estimate or an erroneous standard error. Suppose one decided to sample quadrats but, for logistical reasons, laid them out in lines, with the distance between the lines considerably greater than that between neighboring quadrats within the lines. The standard error of the estimate of density could not then be calculated by the usual formulae because the counts on those quadrats would not be independent. Density is correlated between neighboring quadrats, and this throws out the simple estimate of the standard error, which returns an erroneously low value. There are ways of dealing with the data from this design in order to yield an appropriate standard error (see Cochran 1977 for treatment of two-stage sampling and Norton-Griffiths 1973 for an example using the Serengeti wildebeest), but they are beyond the scope of this book. The simple remedy is to pool the data from all quadrats on each line, the line rather than the quadrat becoming the sampling unit. This procedure may appear to sacrifice information, but it does not (Caughley 1977a).

Another common mistake is to throw random points on to a map and to declare them centers of the units to be sampled, the boundary of each being defined by the position of the point. In this case, the requirement that sampling units cover the whole area and be non-overlapping is violated and the sampling design becomes a hybrid between SWR and SWOR, leading to difficulties in calculating a standard error. There is nothing wrong with choosing units to be sampled by throwing random points on a map so long as the frame of units is marked on the map first. The random points define the units to be selected. They do not determine where the boundaries of those units lie.

A third trap to watch for is a biased selection of units to be sampled. The most common source of this bias in wildlife management is the so-called "road count," in which animals are counted from a vehicle on either side of a road or track. Roads are not random samples of topography. They tend to run along the grain of the country rather than across it, they go around swamps rather than through them, they tend to run along vegetation ecotones, and they create their own environmental conditions, some of which attract animals and others of which repel them.

12.4 **Sampled counts: methods and arithmetic**

Sampled counts of animals fall easily into two categories. The first is the method of counting in sampling units with fixed boundaries. We might for example walk lines and count deer in the area within 100 m either side of the line of march, or count all the ducks in a sample of ponds, the shoreline of the pond providing a strict boundary to the sampling unit.

The alternative is unbounded sampling units (Buckland *et al*. 1993, 2001). Instead of restricting the counting to those animals within 100 m of a line of march, those outside the transect being ignored, we might count all the animals that we see. Since the observed density will fall away with distance from the observer, the raw counts are no longer an estimate of true density. They must therefore be corrected.

Of these two options (sampling units with boundaries and sampling units without boundaries), the first has immense advantages of simplicity and realism. If the transect

width is appropriately chosen, what the observer sees is what the observer gets. The mathematics of such sampling are simple, elegant, and absolutely solid. In contrast, the accuracy of a corrected density estimated from unbounded transects depends heavily upon which model is chosen for the analysis. There are many to choose from and they give markedly different answers for the same data. The advantage of unbounded transects lies in all the sightings being used, with none being discarded. Since the precision of an estimate is related tightly to the number of animals actually counted, any sampling scheme that increases the number of sightings also tends to increase the precision of the estimate. That is an advantage if the increased precision is obtained without the sacrifice of too much accuracy.

The choice of one system or another is often determined by density. If the species is rare then we might be tempted to use all the data we can get. If it is common, we might be content to use the more dependable sampling units with fixed boundaries, knowing that fewer things can go wrong.

12.4.1 *Fixed boundaries to sampling units*

The appropriate analysis depends on whether the sampling units are of equal or unequal size, and how they are selected. Formulae were originally developed by Jolly (1969), based on Cochran (1977) (see also Norton-Griffiths 1978).

Notation

y = the number of animals on a given sampled unit;
a = the area of a given sampled unit;
A = the total area of the region being surveyed;
n = the number of units sampled;
D (or d) = the estimate of mean density;
$SE(D)$ = the standard error of estimated mean density;
Y = the estimate of total numbers in the region of size A;
$SE(Y)$ = the standard error of the estimate of total numbers.

The simple estimate (for equal-sized sampling units)

The simple estimate is used when sampling units are of constant size, as when the region being surveyed is a rectangle, which can be subdivided into quadrats or transects. It will provide an unbiased, although imprecise, estimate, even when sampling units differ in size – but more appropriate designs are available for that case. We will explore this design at some length because most of the principles are shared with the others.

The region to be surveyed, of area A, is divided on a map or in one's head into an exhaustive set of non-overlapping sampling units, each of constant area a. Let us assume, for illustration, that the region is as given in Table 12.1, and that this region of $A = 144$ km^2 is to be sampled by $n = 4$ transects each of area $a = 12$ km^2. Sampling intensity is hence $na/A = 4 \times 12/144 = 0.333$.

In Table 12.1 the rows represent transects and the marginal totals the number of animals on each. Numbering the transects from 1 to 12 and selecting at random with replacement from this set, we draw transects 4, 8, 1, and 4. On surveying these transects, we obtain counts of:

Transect:	1	4	4	8
Count:	142	149	149	127

Note that transect 4 has been drawn twice, so in practice we survey only three transects, although the count from transect 4 enters the calculation two times.

Density is estimated as the sum of the transect counts (142 + 149 + 149 + 127) divided by the sum of the transect areas (12 + 12 + 12 + 12). Thus:

$$D = \Sigma y / \Sigma a = 567/48 = 11.81/\text{km}^2.$$

The precision of this estimate is indexed by its standard error SE(D), which is itself an estimate of what the standard deviation of many independent estimates of density would be, each derived from four transects drawn at random with replacement:

$$SE(D) = 1/a \times \sqrt{(\Sigma y^2 - (\Sigma y)^2/n)/(n(n-1))}$$

This is a slight approximation. To be exactly unbiased, it should be multiplied by a further term $\sqrt{1 - (\Sigma a)/A}$, but that usually makes so little difference that it tends to be ignored.

The calculation tells us that this hypothetical distribution of estimates, each of them made in the same way as we made ours, with the same sampling frame and the same sampling intensity, only the draw of sampling units being different, will have a standard deviation in the vicinity of ±0.43. In fact, this is likely to be an underestimate because it is based on only four sampling units, three degrees of freedom. With samples above 30 sampling units we can form 95% confidence limits of the estimate by multiplying by 1.96, but for smaller samples we must choose a multiplier from a Student's t -table corresponding to a two-tailed probability of 0.05 and the degrees of freedom (d.f.) of our sample. In the case of d.f. = 3, the multiplier is 3.182 and so the 95% confidence limits of our estimate of density are ±3.18 × 0.43 = ±1.37.

The number of animals Y in the surveyed region can now be calculated as the number of square kilometers in that region A multiplied by the estimated mean number per square kilometer D:

$$Y = AD = 144 \times 11.81 = 1701$$

This has a standard error of:

$$SE(Y) = \pm A \times SE(D) = \pm 144 \times 0.43 = \pm 62$$

Its 95% confidence limits are calculated as A multiplied by the 95% confidence limits of D:

$$\pm 144 \times 1.37 = \pm 197$$

We can check this against Table 12.2, which shows that the true total number Y is 1737, so the estimate with 95% confidence of $Y = 1701 \pm 197$ is entirely acceptable.

If the sampling is without replacement, the formula for SE(D) yields an overestimate. The standard error for SWOR is estimated by the formulation for the standard error with replacement multiplied by the square root of the proportion of the area not surveyed. This *finite population correction* (*FPC*) is:

$$FPC = \sqrt{1 - (\Sigma a)/A}$$

The simple estimate may validly be used even when sampling units are of unequal size. The constant a is then replaced by the mean area of sampling units. The precision of the estimate will be lower (i.e. the standard error will be higher) than that produced by the ratio method (see next subsection), but the estimate is unbiased and may be precise enough for many purposes.

The simple estimate, with minor modification, can be used when the total area A is unknown. One of us was forced to this exigency while surveying from the air a population of rusa deer (*Cervus timorensis*) in Papua New Guinea. The deer lived on a grassed plain, the area of which could not be gauged with any accuracy from the available map. The remedy was to measure the length of the plain by timing the aircraft along it at constant speed, and then to run transects from one side of the plain to the other at right angles to that measured baseline. The area of a sampling unit is entered as $a = 1$, even though they are of different and unknown areas. D then comes out as average numbers per transect, rather than per unit area. Total numbers Y on the plain can be estimated by replacing A with N, where N is the total number of transects that could have been fitted into the area. This is simply the length of the baseline divided by the width of a single transect. A similar approach was used for censusing of wildebeest in the Serengeti (Norton-Griffiths 1973, 1978).

The ratio estimate (for unequal-sized sampling units)

This is the best method for a frame of sampling units of unequal size, as might be provided by a faunal reserve of irregular shape sampled by transects. Statistical texts warn that the estimate is biased when the number of units sampled is less than 30 or so, but the bias is usually so slight as to be of little practical importance. The number of units may be as low as two without generating a bias of more than a few per cent.

The appropriate formulae are given in Table 12.3 and the notation at the beginning of Section 12.5.1. That for the standard error looks quite different from that for the simple estimate but they are mathematical identities when the sampling units are of equal size. The ratio estimate is general, the simple estimate being a special case of it. Hence, if these analyses are to be programmed into a calculator or computer, the ratio method is the only one needed.

Table 12.3 Estimates and their standard errors for animals counted on transects, quadrants, or sections. The models are described in the text.

Model	Density	Numbers
Simple		
Estimate	$D = \Sigma y/\Sigma a$	$Y = A \times D$
Standard error of estimate (SWR)	$\mathrm{SE}(D)_1 = 1/a \times \sqrt{(\Sigma y^2 - (\Sigma y)2/n)/n(n-1)}$	$\mathrm{SE}(Y) = \mathrm{A} \times \mathrm{SE}(D)_1$
Standard error of estimate (SWOR)	$\mathrm{SE}(D)_2 = \mathrm{SE}(D)_1 \times \sqrt{1 - (\Sigma a)/A}$	$\mathrm{SE}(Y) = \mathrm{A} \times \mathrm{SE}(D)_2$
Ratio		
Estimate	$D = \Sigma y/\Sigma a$	$Y = \mathrm{A} \times D$
Standard error of estimate (SWR)	$\mathrm{SE}(D)_3 = n/\Sigma a \times \sqrt{(1/n(n-1))(\Sigma y^2 + D^2\Sigma a^2 - 2D\Sigma ay)}$	$\mathrm{SE}(Y) = \mathrm{A} \times \mathrm{SE}(D)_3$
Standard error of estimate (SWOR)	$\mathrm{SE}(D)_4 = \mathrm{SE}(D)_3 \times \sqrt{1 - (\Sigma a)/A}$	$\mathrm{SE}(Y) = \mathrm{A} \times \mathrm{SE}(D)_4$
PPS		
Estimate	$d = 1/n \times \Sigma(y/a)$	$Y = \mathrm{A} \times d$
Standard error of estimate (SWR)	$\mathrm{SE}(D) = \sqrt{(\Sigma y/a)^2 - (\Sigma(y/a))^2/n)/n(n-1)}$	$\mathrm{SE}(Y) = \mathrm{A} \times \mathrm{SE}(d)$

SWR, sampling with replacement; SWOR, sampling without replacement. Notation is given at the beginning of Section 12.5.1.

The PPS estimate

By the previous two methods, all sampling units in the frame have an equal chance of being selected. By the probability-proportional-to-size (PPS) method, the probability of selection is proportional to the size of the sampling unit. Suppose that the area to be surveyed is farmland. We might decide to declare the paddocks (or "pastures" or "fields," depending on which country you are in) as sampling units because the fences provide easily identified boundaries to those units.

If each sampling unit were assigned a number and the sample were chosen by lot, we would use the ratio method of analysis. However, we might decide instead to choose the sample by throwing random points on to a map. Each strike selects a unit to be sampled, the probability of selection increasing with the size of the unit.

The PPS estimate has the advantages that it is entirely unbiased and that the arithmetic (Table 12.3) is simple. Its disadvantage is that it can be used only when sampling with replacement and so it is not as precise as the ratio method used without replacement. Hence, this method should be restricted to surveys whose sampling intensity is below 15%. The PPS estimate is a mathematical identity of the simple estimate and the ratio estimate when units of equal size are sampled with replacement.

12.4.2 *Unbounded transects (line transects)*

The observer walks a line of specified length and counts all animals seen, measuring one or more subsidiary variables at each sighting (e.g. angle between the animal and the line of march; radial distance, the distance between the animal and the observer at the moment of sighting; the right-angle distance between the animal and the transect). If we know the shape of the sightability curve relating the probability of seeing an animal on the one hand to its right-angle distance from the line on the other, and if an animal standing on the line will be seen with certainty, it is fairly easy to derive an estimate of density from the number seen and their radial or right-angle distances. We seek a distance from the line at which the number of animals missed within that distance equals the number seen beyond it. True density is then the total seen divided by the product of twice that distance and the length of the line.

Therein lies the difficulty. That distance is determined by the shape of the sightability curve, which can seldom be judged from the data themselves. Consequently, the shape of the curve must be assumed to some extent, and the validity of the assumption determines the accuracy of the method.

We present here just two of the many models available, mainly to give some idea of their diversity. The first is the Hayne (1949) estimate, which is derived from the assumption that the surveyed animals have a fixed flushing distance and will be detected only when the observer crosses that threshold. If k is the number of animals detected and r the radial distance from a detected animal to the observer then:

$$D = (1/2L)\Sigma_k(1/r)$$

where L is the length of the line. Hence, density is the sum of the reciprocals of the radial sighting distances divided by twice the length of the line.

It is implicit in Hayne's model that $\sin\theta$, the sine of the sighting angle, is uniformly distributed between 0 and 1, and that the theoretically expected mean sighting angle is 32.7°. Hence, the reality of the model can be tested against the data. Eberhardt (1978) recommended tabulating the frequency of $\sin\ \theta$ in 10 intervals of 0.1 $(0.0 - 0.1, 0.1 - 0.2 \ldots 0.9 - 1.0)$ and testing the uniformity of the frequencies by chi-square. He gave a worked example for a survey of the side-blotched lizard

(*Uta stansburiana*). Robinette *et al*. (1974) and Burnham *et al*. (1980) suggested that most mean sighting distances tended to be around 40° or more, the latter authors being convinced that the Hayne estimate is used far too uncritically in wildlife management. Robinette *et al*. (1974) compared the accuracy of the Hayne estimate with that of eight other line transect models, showing that when applied to inanimate objects or to elephants it tended to overestimate considerably. However, Pelletier and Krebs (1997) found both the Hayne estimate and line transect estimates provided relatively unbiased results when compared with a known population of ptarmigan (*Lagopus* species) in the Yukon. Buckland *et al*. (1993) provides a starting point for reading further about line transect methods.

Our second example is a nonparametric method developed by Eberhardt (1978) from work by Cox (1969). First, we choose arbitrarily a distance, Δ, perpendicular from the line. Eberhardt's estimate of density is:

$$D = (3k_1 - k_2)/4L\Delta$$

where k_1 and k_2 are the number of animals seen on either side of the line transect at distances that fall within the intervals $0 - \Delta$ and $\Delta - 2\Delta$, respectively. Eberhardt (1978) considered that the method is most useful as a cross-check on the results of other methods, because its estimate is likely to be imprecise. Precision is enhanced by choosing a large value of Δ but accuracy is enhanced by choosing a small one (Seber 1982).

Much of the present use of line transects in wildlife management stems from the belief that they are somehow more "scientific" than strip transects, just as there was once a belief that quadrats were statistically superior to transects. There are rare situations in which transect sampling will not work and where line transect methodology might (e.g. in very thick cover). The unbounded line transect method has advanced considerably with the use of the computer software Distance (http://www.ruwpa.st-and.ac.uk/distance), developed by Buckland *et al*. (1993, 2001). Although the use of the software is not easy, it is currently the most powerful tool for line censuses. In particular, it is most useful for rare observations, although it does require at least 30 records in order to be reliable. In addition, time must be allowed to make the necessary estimates of perpendicular distance from the line to the animal (or groups of animals). If there are insufficient observations of a particular species (or other category) in a station or habitat then one can repeat the line survey until a sufficient number have been accumulated. The only proviso is that animals distribute themselves randomly with respect to the line and there is no spatial correlation between surveys. The method is particularly suitable for rare species such as carnivores and rare ungulates and birds. It is less suitable where there are large numbers of animals (e.g. ungulates on the Serengeti plains).

Note that none of these unbounded methods can be used in aerial survey. They are all anchored by the assumption that all animals on the line of march (equivalent to the inner strip marker of aerial survey) are tallied by the observer. This assumption does not hold for aerial survey because the ground under the inner strip marker is at a distance from the observer, because an animal under a tree on that line may be missed, and because an observer cannot watch all parts of the strip at once and may therefore miss animals in full view on the line. In addition, the speed of the aircraft makes the measurement of distances from the observer unfeasible.

The assumption that all animals on the line are counted can be relaxed if the probability of detecting animals on the line can be estimated. This is particularly important for marine mammals, where only a fraction of a group or pod is on the surface at any one time. The probability of detection on the line for harbor porpoises (*Phocoena phocoena*) was estimated to be only 0.292, which illustrates just how many remain unseen. Furthermore, this estimate was made by experienced observers; for inexperienced observers, the sighting probability was only 0.079; that is, some 90% were missed. This shows the importance of training and experience (Laake *et al*. 1997).

The biologist must decide whether the statistical power of line transects justifies their practical application. Can the difficulty of measuring sighting distances and the unreliability of the resultant estimates be justified when an alternative with fewer problems is available? The line transect was originally introduced to circumvent the difficulty of counting all animals on a transect or quadrat. It cured that problem by replacing it with several others. Perhaps we should give some thought to ways of treating the original problem without introducing new ones. If animals are difficult to see on a transect of fixed width, why not walk two people abreast down the boundaries? If that does not work, put a third person between them. And so on.

12.4.3 *Stratification*

The precision of an estimate is determined by sampling intensity and by the variability of density among sampling units. Suppose there were two distinct habitats in the survey area and that from our knowledge of the species we could be sure that it would occur commonly in one and rarely in the other. If we surveyed those two sub-areas separately and estimated a separate total of animals for each, the combined estimate for the whole area would be appreciably more precise than if the area had been treated as an undifferentiated whole.

The process is called stratification and the sub-areas strata. By this strategy, we divide an area of uneven density into two or more strata within which density is much more even. The strata are treated as if they were each a total area of survey and the results are later combined. The estimate from each stratum will be called Y_h, which has a standard error of $SE(Y_h)$. Total numbers Y are estimated by $Y = \Sigma Y_h$. The standard error is the square root of the sum of the variances of the contributing stratal estimates. The variance of an estimate is the square of its standard error. Here it is designated Var(est) to distinguish it from the variance of a sample designated s^2. Calculate:

$$\mathrm{Var}(Y_h) = (\mathrm{SE}(Y_h))^2$$

for each stratum and then:

$$\mathrm{SE}(Y) = \sqrt{\Sigma Var(Yh)}$$

to give the standard error of the combined estimate of total numbers.

Optimum allocation of sampling effort

If our aim is to get the most precise estimate of Y as opposed to a precise estimate of each Y_h, sampling intensity should be allocated between strata according to the expected standard deviation of sampled unit counts in each stratum. This requires a pilot survey or at least approximate knowledge of distribution and density gained on a previous survey. Often we have nothing more than aerial photographs or a vegetation

map to give us some idea of the distribution of habitat, and only knowledge of the animal's ecology to guide us in predicting which habitats will hold many animals and which will hold few. This scant information is in fact sufficient to allow an allocation of sampling effort between strata that will not be too far off the optimum. The important point to understand is that for almost all populations the standard deviation of counts on sampling units rises linearly with density. From this can be derived the rule of thumb that the *number of sampling units put into a stratum should be directly proportional to what* Y_h *is likely to be*.

At first thought, this is a daunting challenge – to guess each Y_h before we have estimated it – but it is easier if we break it down into components. First, guess the density in each stratum. It does not matter too much if this is wrong, even badly wrong, because all we need to get roughly right is the ratios of densities between strata. Second, multiply each guessed density by the mapped area of its stratum to give a guess at numbers in the stratum. Third, divide each by the total area to give the proportion of total sampling effort that should be allocated to each stratum. Table 12.4 shows the calculation for a degree block that can be divided into three strata from a vegetation map and to which a total of 10 hours of aerial survey have been allocated.

12.4.4 *Comparing estimates*

If the sampling units are drawn independently of each other, the estimates of density from two surveys may be compared. The surveys may be of two areas, or of the same area in two different years, or of the same area surveyed in the same year by two teams or by different methods. A quick and dirty comparison is provided by the normal approximation, which is adequate if each survey covered more than 30 sampling units. The two estimates are significantly different when:

$$(\text{est}_1 - \text{est}_2)\sqrt{Var(est_1) + Var(est_2)} \quad > \quad 1.96$$

If sample sizes are too low, or if more than two surveys are being compared, the determination of significance should be made by one-factor analysis of variance. If the surveys are not independent, as when the same transects are run each year, a comparison may still be made by analysis of variance but with TRANSECTS now declared a factor in a two-factor analysis. Chapter 14 goes further into this and other uses of analysis of variance.

12.4.5 *Merging estimates*

If a comparison shows that two or more independent estimates of the same population are not significantly different we may wish to merge them to provide an estimate that is more precise than the originals. This procedure is quite distinct from stratification, where estimates from different populations are combined to give an overall estimate. Merging is restricted to the same population estimated more than once. We must make

Table 12.4 Allocation of $E = 10$ hours of aerial survey among strata to maximize the precision of the estimate of animals in the total area.

Stratum (h)	Area (A_h), km^2	Guessed density (D_h)	Guessed numbers ($Y_h = A_h D_h$)	Proportion of total effort ($P_h = Y_h/\Sigma Y_h$)	Hours allocated ($E_h = P_h E$)
1	2000	1	2000	0.03	0.3
2	7000	5	35 000	0.52	5.2
3	3000	10	30 000	0.45	4.5
	12 000		67 000	1.00	10.0

sure that environmental (e.g. different seasons) and biological (e.g. significant mortality or emigration) conditions do not differ between censuses. Merging is particularly powerful in obtaining a reduced confidence interval from a series of individual censuses, each with very wide confidence intervals. If we obtain a single estimate with a wide confidence interval (say, because too few samples were counted) then it will often pay to repeat the census as soon as possible and merge the two results.

There are two methods. The first is quick and dirty, to be used only when the individual estimates were made with about the same sampling intensity. The merged estimate $\hat{Y}$ can then be calculated as:

$$\hat{Y} = (Y_1 + Y_2 + Y_3 + \ \ldots \ + Y_N)/N$$

where there are N surveys. It has a variance of:

$$\mathrm{Var}(\hat{Y}) = (\mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + \mathrm{Var}(Y_3) + \quad \ldots \ + \mathrm{Var}(Y_N))/N^2$$

Thus, the merged estimate is simply the mean of the individual estimates, and its variance is the mean of the individual estimate variances divided by their number.

$\mathrm{SE}(\hat{Y})$ is the square root of $\mathrm{Var}(\hat{Y})$. From these, the merged density estimate is $D = \hat{Y}/A$, which has a standard error of $\mathrm{SE}(D) = \mathrm{SE}(\hat{Y})/A$.

A second, more appropriate method, particularly for surveys utilizing markedly different intensities of sampling, is provided by Cochran (1954), who also considers more complex merging. Here the contribution of an individual estimate to the merged one is weighted according to its precision. Letting $w = 1/\mathrm{Var}(Y)$:

$$\hat{Y} = (w_1 Y_1 + w_2 Y_2 + w_3 Y_3 + \ \ldots \ + w_N Y_N)/(w_1 + w_2 + w_3 + \ \ldots \ + w_N)$$

with a variance of:

$$\mathrm{Var}(\hat{Y}) = 1/(w_1 + w_2 + w_3 + \ \ldots \ + w_N)$$

12.5 Indirect estimates of population size

This section outlines some of the methods available for calculating the size of a population using techniques that do not necessarily depend on accurate counts of animals. The line transect method could well come under this head but is placed in "Sampled Counts" because it requires accurate counting of animals on the line.

12.5.1 *Index–manipulation–index method*

If we obtain two indices of population size, I_1 and I_2, the former before and the latter after a known number of animals C was removed, the population's size can be estimated for the time of the first index by:

$$Y_1 = I_1 C/(I_1 - I_2)$$

The proportion removed is estimated as $p^* = (I_1 - I_2)/I_1$ and the proportion of those remaining as $q^* = 1 - p^*$. Following Eberhardt (1982), the variance of the estimate of Y can be approximated by:

$$\mathrm{Var}(Y_1) \ \approx \ {Y_1}^2 \, (q^*/p^*)^2 \, (1/I_1 + 1/I_2)$$

from which $SE(Y_1) = \sqrt{Var(Y1)}$. Eberhardt (1982) gives three examples from populations of feral horses. The data from his Cold Springs population were:

$I_1 = 301$;
$I_2 = 76$;
$C = 357$;
$p^* = 0.748$.

Thus, the population at the time of the first index is estimated as:

$$Y_1 = (301 \times 357)/(301 - 76) = 478$$

with the following estimated variance:

$$\text{Var}(Y_1) \approx 478^2\,(0.252/0.748)^2\,(1/301 + 1/76) = 428$$

from which $SE(Y_1) = \sqrt{428} = 21$.

The index–manipulation–index method assumes that the population is closed (no births, deaths, immigration, or emigration) between the estimation of the first and second indices. This assumption is approximated when the entire experiment is run over a short period.

12.5.2 *Change-of-ratio method*

If a population can be divided into two classes, say males and females or juveniles and adults, and one class is significantly reduced or increased by a known number of animals, the size of the population can be estimated from the change in ratio. Kelker (1940, 1944) introduced this method to estimate the size of deer populations manipulated by bucks-only hunting.

The two classes are designated x and y. Before the manipulation, there is a proportion p_1 of x individuals in the population; this becomes p_2 after the manipulation, which removes or adds C_x x individuals (additions are positive, removals negative) and C_y y individuals. $C = C_x + C_y$. The size of the population before the manipulation may be estimated as:

$$Y_1 = (C_x - p_2 C)/(p_2 - p_1)$$

As with the index–manipulation–index method, Kelker's method assumes that the population is closed. Hence, the two surveys to estimate the class proportions must be run close together. Additionally, all removals or additions must be recorded and the two classes must be equally amenable to survey.

Cooper *et al*. (2003) have extended this approach using likelihood estimates of the ratios. When harvesting is highly skewed towards a single sex or age class, the change in these ratios provides information about the exploitation rate, and when combined with absolute numbers removed also provides information on absolute abundance.

12.5.3 *Mark–recapture*

Mark–recapture is a special case of the change-of-ratio method. A sample of the population is marked and released and a subsequent sample is taken to estimate the ratio of marked to unmarked animals in the population. From data of this kind we can estimate the size of the population, and with further elaboration (individual markings, multiple recapturing occasions) the rate of gain and loss.

The huge number of mark–recapture models available has been reviewed adequately by Blower *et al.* (1981) and in detail by Seber (1982) and Krebs (1999). Bowden and Kufeld (1995) present some methods for estimating confidence limits for general mark–recapture calculations, using the example of Colorado moose (*A. alces*). Here we outline the range of methods, provide an introduction to the most simple cases, and emphasize their pitfalls, as well as mentioning some recent advances which might circumvent these pitfalls.

Petersen–Lincoln models

A sample of M animals is marked and released. A subsequent sample of n animals is captured, of which m are found to be marked. If Y is the unknown size of the population then clearly:

$$M/Y = m/n$$

within the limits of sampling variation. With rearranging, this allows an estimate of population size as:

$$Y = Mn/m$$

Intuitively obvious as this is, it is not quite right because of a statistical property of ratios that leads on average to a slight overestimation. This bias may be corrected by (Bailey 1951, 1952):

$$Y = (M(n + 1))/(m + 1)$$

which has a standard error of approximately:

$$\mathrm{SE}(Y) = \sqrt{[(M^2(n + 1)(n - m))/((m + 1)^2(m + 2))]}$$

These formulae are for "direct sampling," where the number of animals to be recaptured is not decided upon prior to recapturing. There are further variants for SWR and for inverse sampling (see Seber 1982).

Except in the unlikely case of half or more of the population being marked, the distribution of repeated independent estimates of population size is always strongly skewed to the right – a positive skew. (The direction of skew is the direction of the longest tail.) Fig. 12.1 shows this effect from a computer simulation of 1000 estimates of a population of 500 animals containing 100 marked individuals. Each estimate is derived from a capturing of 50 animals. Apart from demonstrating the skew of estimates, this figure makes the point that only a limited number of estimated values is possible. With $Y = 500$ and $M = 100$, the probability of a given animal being marked is 0.2, so the expected number of marked animals in a sample of 50 is 10. This would give a population estimate of $Y = 464$. If nine were recaptured, the estimate would be $Y = 510$. No estimate between 464 and 510 is possible.

Since the estimates are skewed, the confidence limits of an estimate are also skewed and cannot easily be calculated from the standard error. Blower *et al.* (1981) recommended an approximating procedure. Let $a = m/n$. In a large sample, the 95%

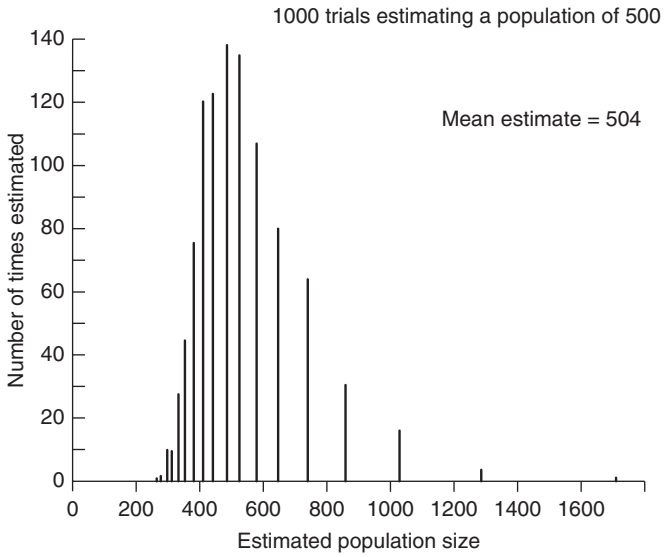

Fig. 12.1 Simulated replications of estimates of a population of 500 individuals by mark–recapture where 100 are marked and 50 captured. Note the positive skew of estimates and the fact that only a limited number of estimated values are possible.

confidence limits of a are approximately $\pm 1.96\sqrt{(a(1-a)/n)}$. Since $Y = M/a$, the upper and lower 95% confidence limit of a can each be divided into M to give upper and lower 95% confidence limits of Y.

The Petersen estimate is the most simple of a family of estimation procedures. If animals are marked on more than one occasion and recaptured on more than one occasion then it is possible to estimate gains and losses from the population, in addition to its size. Seber (1982) describes most of the options.

The Petersen estimate depends on these assumptions:

1 All animals are equally catchable.
2 No animal is born or migrates into the population between marking and recapturing.
3 Marked and unmarked animals die or leave the area at the same rate.
4 No marks are lost.

Assumption **2** is not needed when marked animals are recaptured on more than one occasion, but the others are common to all elaborations of the Petersen estimate. The least realistic is the assumption of equal catchability, which is routinely violated by almost any population the wildlife manager is called upon to estimate (Eberhardt 1969). For this reason, the Petersen estimate and its elaborations (Bailey's triple catch, Schnabel's estimate, the Jolly–Seber estimate, and many others) are of limited utility in wildlife management.

Frequency-of-capture models

Petersen models work only when all animals in the population are equally catchable. Frequency-of-capture models are not constrained in this way but will work only if the population is closed: if there are no losses from or gains to the population over the interval of the experiment. This is easy enough to approximate by running the exercise over a short period.

Animals are captured on a number of occasions, usually on successive nights, and marked individually at the first capture. At the end of the experiment, each individual

caught at least once can be scored according to the number of times it was captured. The data come in the form:

Number of times caught (i):	1	2	3	4	5	6	7	8	…	18
Number of animals (f_i):	43	16	8	6	0	2	1	0	…	0

which are from Edwards and Eberhardt (1967), who trapped a penned population of wild cottontail rabbits for 18 days. Of these, 43 were caught once only, 16 twice, 8 three times, and so on. $\Sigma f_i = 76$ gives the number of rabbits caught at least once and so the population must be at least that large. If we could estimate f_0, the number of rabbits never caught, we would have an estimate of population size:

$$Y = f_0 + 76$$

Traditionally, this has been attempted by fitting a zero-truncated statistical distribution (Poisson, geometric, negative binomial) to the data and thereby estimating the unknown zero frequency. Eberhardt (1969) exemplifies this approach. More complex mark–recapture models use sophisticated analytical techniques to cope with variation in the probability of capture due to time (seasonal trends, changes in weather), variation among individual animals (site fidelity, sex differences, dominance relationships), prior trapping history (capture-shyness and capture-proneness), and various combinations of these (Pollock 1974; Burnham and Overton 1978; Otis *et al*. 1978). The fit of each model can be tested against the data and an objective decision can be made as to which is the most appropriate, often using information theory (see Chapter 15). The computations are too lengthy to be attempted by hand, but several software programs are freely available on the Web: CAPTURE (White *et al*. 1982), SURGE (Lebreton *et al*. 1992), and MARK (White and Burnham 1999).

Estimation of density

All previously reviewed mark–recapture methods yield a population size Y that can be converted to a density D only when the area A relating to Y is known. In most studies, Y itself is meaningless because the "population" is not a population in the biological sense but the animals living on and drawn to a trap grid of arbitrary size.

Seber (1982) and Anderson *et al*. (1983) reviewed the methods currently used to estimate A as a prelude to determining density. Most rely on Dice's (1938) notion of a boundary strip around the trapping grid such that the *effective trapping area A* is the grid area plus the area of the boundary strip. Most of these methods are *ad hoc* and subject to numerous problems, or require large quantities of data to produce satisfactory estimates, or require supplementary trapping beyond the trapping grid.

Anderson *et al*. (1983) circumvented this problem with a method of mark–recapture that provides a direct estimate of density. The traps are laid out not in a grid but at equal intervals along the spokes of a wheel. Trap density therefore falls away progressively from the center of the web. The method pivots upon the assumption that the high density of traps at the center guarantees that all animals at the center will be captured. This is analogous to the assumption of line transect methodology that all animals are tallied on the line itself. The data collected as "distance of first capture from the center of the web" are analyzed almost exactly as if they were from a line transect (Buckland *et al*. 1993, 2001). This analysis can be run on the computer program Distance (Laake *et al*. 1993).

12.5.4 *Incomplete counts*

The problem of estimating the size of a population from "total counts" known to be inaccurate has been approached from three directions. One family of methods requires a set of replicate estimates, the second requires two estimates, and the third provides an estimate known with confidence to be below true population size.

Many counts

Hanson's (1967) method assumes that all animals have the same probability of being seen but that this probability is less than one. Hence, whether a given animal is seen or not on a given survey is a draw from a binomial distribution. It follows from the mathematics of the binomial distribution that $Y = \overline{x}\, x/(\overline{x} - s^2)$, where Y is the population size, $\overline{x}$ the mean of a set of (incomplete) counts, and s^2 the variance of those counts.

This method is not recommended, because of the restriction that all animals have the same sightability. In practice, sightability varies by individuals and between surveys. The variance of a set of replicate counts tends to be greater than their mean (a binomial variance is always less than the mean), indicating that the method is unworkable.

A modification of the method to circumvent this restriction was suggested by Caughley and Goddard (1972). It requires repeated counts made at two levels of survey efficiency (e.g. two sets of aerial surveys, one flown at 50 m and the other at 100 m altitude). However, Routledge (1981) showed by simulation that this method yields a very imprecise estimate unless the number of surveys is prohibitively large, and hence we do not recommend it.

The nonparametric *method of bounded counts* (Robson and Whitlock 1964) provides a population estimate from a set of replicate counts as twice the largest minus the second largest. Routledge (1982) dismissed this method also (as do we), because in most circumstances it greatly underestimates the true number.

Two counts

Caughley (1974) showed that if the counts of two observers of equivalent efficiency are divided into those animals (or groups of animals) seen by only one observer and those seen by both, the size of the population can be estimated. Henny *et al*. (1977) and Magnusson *et al*. (1978) extended the method to allow for the two observers being of disparate efficiency.

Essentially the method is a Petersen estimate, although animals are neither marked nor captured. Suppose that the entities being surveyed are stationary and that their individual positions can be mapped. Magnusson *et al*. (1978) surveyed crocodile nests and Henny *et al*. (1977) the nests of ospreys. If the area is surveyed independently twice, perhaps once from the ground and once from the air, the entities can be divided into four categories:

1 S_1 = the number seen on the first survey but missed on the second.
2 S_2 = the number seen on the second survey but missed on the first.
3 B = the number tallied by both surveys.
4 M = the number missed on both surveys.

This is equivalent to a mark–recapture exercise. The first survey maps (marks) a set of entities, each of which may or may not be seen (recaptured) on the second. But unlike a true mark–recapture exercise, the model is symmetrical and the first and second surveys are interchangeable.

If P_1 is the probability of an entity being seen on the first survey and P_2 the probability of its being seen on the second then:

$$P_1 = B/(B + S_2);$$
$$P_2 = B/(B + S_1);$$
$$M = S_1 S_2/B;$$
$$Y = ((B + S_1)\,(B + S_2))/B.$$

where Y is an estimate of the size of the population.

The last equation may be corrected for statistical bias (Chapman 1951) to:

$$Y = [((B + S_1 + 1)\,(B + S_2 + 1))/(B + 1)] - 1$$

which has a variance given by Seber (1982) of:

$$\mathrm{Var}(Y) = (S_1\,S_2\,(B + S_1 + 1)(B + S_2 + 1))/((B + 1)^2\,(B + 2))$$

Magnusson *et al*. (1978) reported that although the method is based on the assumptions that the two surveys are independent and that there is a constant probability of seeing an entity on a given survey (equal catchability), the second is not critical. The population estimate is close enough even when the probability of being seen varies greatly between individuals.

Caughley and Grice (1982) extended the method to moving targets, dropping the requirement that the position of stationary entities must be mapped so that they could be identified as seen or not seen at the two surveys. Groups of emus (*Dromaius novaehollandiae*) were tallied simultaneously but independently by two observers seated in tandem on one side of an aircraft. Their counts of $S_1 = 7$, $S_2 = 3$, and $B = 10$ yielded $P_1 = 0.77$ and $P_2 = 0.59$, the population estimate being $Y = 22$ emu groups on the 843 km^2 of transects that they surveyed together, a density of 0.03 groups/km^2.

This method of simultaneous and fully independent tallying carries two dangers, one technical, the other statistical. First, the two observers must not unconsciously cue each other to the presence of animals in their field of view and ideally should be screened from each other. Second, the chances of "marking" and "recapturing" an entity should be uncorrelated, but they are not, because marking and recapturing occur at the same instant, the search images transmitted to each observer being nearly identical. Caughley and Grice (1982) showed by simulation that the effect of the close correlation was to underestimate density but that the underestimation became serious only when the mean of P_1 and P_2 was less than 0.5.

Known-to-be-alive

Most estimates of population size require that the manager makes a leap of faith. There is seldom any certainty whether the population fits the assumptions of the model, whether the estimate is wildly inaccurate, or whether the confidence limits have much to do with reality. The more complex the model, the greater the uncertainty. Many ecologists, particularly those working on small mammals, have decided that the work needed to achieve an unbiased estimate is not worth the effort. They would prefer an estimate that, although perhaps inaccurate, is inaccurate in a predictable direction and does not depend on a set of assumptions of dubious reliability. Hence the *known-to-be-alive* estimate, the number of animals that the researcher knows with certainty to be in the study area. These estimates for small-mammal populations are

usually made by trapping an area at high intensity over a short period. Each animal is marked at first capture, the estimated population size being simply the number of first captures. Such estimates are acknowledged as underestimates but they have the advantage of yielding a real number, not an abstract concept, to work with.

Known-to-be-alive estimates are often the most appropriate in wildlife management. There are several problems of conservation and of harvesting for which an overestimate of density might lead to inappropriate management action. An underestimate, on the other hand, should simply produce inefficient but entirely safe management. The penalty for a poor estimate is often distributed asymmetrically around the true population size. It is not good to overestimate the number of individuals of an endangered species. It is not safe to apply a harvesting quota known to be safe for a given population size to one that is much smaller than you thought. Where the undesirable consequences of an overestimate are considerably greater than those accruing from an underestimate, the known-to-be-alive number is often the most appropriate estimate to work with.

12.6 Indices

An index of density is some attribute that changes in a predictable manner with changes in density. It may be the density of bird nests, the density of tracks of brown bears, or the number of minke whales (*Balaenoptera acutorostrata*) seen per cruising hour. A common index is the pellet or fecal dropping count (often used in studies of deer). This was used for endangered marsh rabbits (*Sylvilagus palustris*) in Florida, where pellet counts were closely correlated with radiotelemetry estimates (Forys and Humphrey 1997). Active burrow entrances were used for ground squirrel populations (van Horne *et al*. 1997) and call counts for mourning dove (*Zenaida macroura*) densities (Sauer *et al*. 1994). The North American Breeding Bird Survey is a standardized method in which some 2000 routes are sampled in June each year and the number of singing birds of each species is scored (Droege and Sauer 1989).

These indices reveal something about the density of birds, mammals, or whales. Without knowing anything about the proportional relationship between the index and the abundance of the animal, we could be confident that if the index halved or doubled it would reflect roughly a halving or doubling of animal density. Formally, this holds only when the relationship between index and density is a straight line that passes through the point of zero index and zero density.

Indices of density, if comparable, are useful for comparing the density of two populations or for tracking changes in the density of one population from year to year. Often a comparison is all we need. The relevant question may be not how large is the population but has it declined or increased under a particular regime of management. In such circumstances, the accuracy of an index is irrelevant; precision is paramount.

Let us compare an aerial survey designed to yield an estimate of absolute density with one designed to yield an index of density, as was conducted for pronghorn antelope in Colorado (Pojar *et al*. 1995). The first maximizes accuracy, the second precision. The "accurate" survey would probably inspect small quadrats by circling at a low but varying height above the ground. This is a good way to see animals but it is difficult to standardize between pilots. The "precise" survey would sample transects from a fixed height above ground at a constant speed. Since there is no requirement that all animals be counted on the sampled units, only a fixed proportion being sought, the survey variables are set according to how easily they may be standardized. Groundspeed is higher than for an "accurate" survey, to allow the pilot to maintain constant groundspeed safely even with a strong tailwind. Height above ground is set higher so that

the inevitable variations in height will be proportionally less than at low level; ±10 m around a height of 30 m results in large variations in search image. The same variation around 90 m has little effect. We might choose a transect width of 50 m per observer for an accurate survey but 200 m for a precise survey. The precision of the estimate is approximately proportional to the square root of the number of animals actually tallied (Eberhardt 1978) and so, although proportionally fewer will be seen on a 200 m strip, we choose the wider one in order to increase the absolute number that we see.

Consistency and rigid standardization of techniques are crucial when estimating an index. A good observer is not one who gets a high tally but one who has a consistent level of concentration and who produces results of high repeatability.

All the rules of sampling and of analysis hold as well for indices as for absolute counts of animals. Remember however that indices are useful only in comparisons and, therefore, the quantity to be estimated is the difference between two indices. The variance of an estimate of difference is the sum of the variances of the two estimated indices. As a rule of thumb, we should measure the two indices with a precision such that each standard error is less than a third of the difference we anticipate. Hence, an index must often be estimated much more precisely than is a one-off estimate of population size or density.

Errors in indices can be estimated by comparing results with a known population, similar to the way we estimate bias errors in counts (see Section 12.4.2; Eberhardt and Simmons 1987). For example, the number of sightings of fallow deer (*Dama dama*) in France along a transect (the index) was calibrated against a known population. The sighting index was found to be an effective standardized method for detecting trends in the population (Vincent *et al*. 1996).

12.6.1 *Known-to-be-alive used as an index*

Although known-to-be-alive is sometimes used as a one-off estimate of population size, it is more often used to track trends in population size. The operating rules governing these two uses are quite different. In the first exercise we seek the most accurate estimate we can get. In the second we seek consistency of method among several estimates, such that their bias is held constant. In the first case we put in as much work as possible. In the second we put in precisely the same amount of sampling on each surveying or capturing occasion. Otherwise, the trend in the estimates may reflect no more than variation in capturing effort.

A variant of this aberration of effort, very common in ecological research, is to boost the number known to be alive (because they were caught) on a given occasion by the number of individuals not caught on that occasion but which must have been there because they were caught on both previous and subsequent occasions. Although the accuracy of the estimate of absolute numbers is thereby enhanced, the consistency of the string of estimates is lowered. Estimates for the earlier occasions are inflated relative to those of later occasions, the rate of increase being underestimated if density is rising and the rate of decrease being overestimated if density is falling.

12.7 **Harvest-based population estimates**

In many cases, the only information that is available on wildlife or fishery abundance derives from harvest statistics. Indeed, many fisheries are managed almost exclusively on the basis of catch statistics, so the subject has attracted a great deal of attention from fisheries scientists. Harvest data vary widely in detail and in quality, depending largely on the time and resources that are used in collecting them, but they can be enormously useful if applied properly.

The simplest, but least reliable, index of abundance derived from harvests is a simple assumption that population size is proportionate to total harvest. This logic has been used to interpret long-term fur harvest records from the Hudson's Bay Company in order to study the well-known 10-year cycle of snowshoe hares and lynx, for example (Elton 1924; MacLulich 1937). There are several drawbacks to such an approach. Most importantly, in our example it assumes that the number of fur trappers and the efficiency with which they trap remains constant over time. This is highly unlikely given that economic climate varies over time, trapping techniques often evolve, and trapping is only possible in areas that haven't been turned over to other land use activities, such as agriculture. Nonetheless, such data represent our longest-standing records of crude abundance in the ecological record and have led to important scientific insights.

A much more robust method is to combine harvest totals with the effort expended to calculate *catch per unit effort* (*CPUE*). The logic for this approach is quite straightforward, being based on the same sort of behavioral processes that influence predation (see Chapter 10). Let us imagine that over the course of each full day of hunting a grouse hunter can effectively search 3 km^2. If we further assume that hunters are 100% successful in killing any animals that are encountered and that there is no limit on the number of animals hunted, then hunts would be expected to increase proportionately with grouse density. Following this logic, if the population density of grouse is 1 per km^2 then over the course of 5 days a single hunter would be expected to harvest 15 grouse. If the population density of grouse were to double, that number would be expected to jump to 30 grouse. In practice, harvesters cannot kill all animals encountered, and we don't know how efficiently they search the environment. Nonetheless, we would still expect harvest to scale proportionately with effort and abundance in the following manner:

$$H = qEN$$

where H is the harvest, E is effort, and q is the so-called *catchability* coefficient, which relates effort to the efficiency of search by hunters or fisherman. By rearranging this equation, we can show that:

$$N = \frac{H}{qE}$$

A minor modification (called the *random search equation*) is sometimes applied to accommodate the fact that a large number of hunters hunting over a large number of days will often tend to revisit sites that have already been well travelled. Under these circumstances, the relationship between harvest, abundance, and effort is better represented by a gentle curve than a straight line:

$$H = N(1 - e^{-qE})$$

or equivalently:

$$\frac{H}{N} = 1 - e^{-qE}$$

As an example, Fig. 12.2 shows that the proportion of white-tailed deer harvested from a game management area in Ontario varies in curvilinear fashion with the total

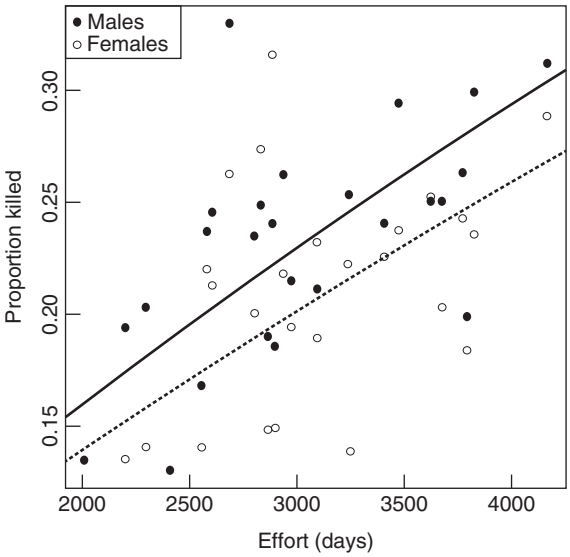

Fig. 12.2 Proportion of white-tailed deer harvested by hunters in Canada in relation to the number of days of hunting effort (males shown with filled symbols, females by open symbols). Lines indicate the best-fit models based on the random search formula (males solid line, females dotted line).

number of days of hunting (Fryxell *et al*. 1991), well approximated by the random search equation.

A more explicit calculation of abundance is possible when age-specific data are also available with respect to harvest. This methodology, known as *virtual population analysis* or *catch-at-age analysis*, is based on keeping track of the fates of specific cohorts (groups of individuals born in the same year) as they work their way through the population. For example, let $H_{i,t}$ represent the catch of age i individuals in year t. Let us further assume that all individuals have a constant probability of survival p and that no individuals live beyond $\tau = 5$ years of age. We start by recognizing that at minimum we know that $H_{i,t}$ individuals were alive before the harvest occurred at time t, hence age-specific abundance $N_{i,t} = H_{i,t}$. We similarly estimate the number of individuals that were alive the year prior to the harvest:

$$N_{i-1,t-1} = H_{i-1,t-1} + H_{i,t}/p$$

In other words, the number of individuals in the same cohort that were a year younger the previous year is determined by the number that show up in the harvests in both years, scaled by the proportion $(1/p)$ that survived long enough to be captured. Similar logic can be used to estimate abundance from age-specific harvest data using the generic formula:

$$N_{i,t} = H_{0,t} + \Sigma_i^{\tau} H_i/p^i$$

This procedure is sufficient to estimate abundance in completed cohorts (i.e. those that have worked their way to the maximum age τ). For incomplete cohorts, we need to estimate abundance in each age class for the final year, usually from CPUE:

$$N_{i,t} = \frac{H_{i,t}}{qE_t}$$

and then use these values in the last year of the dataset to calculate abundance in previous years:

$$N_{i-k,t-k} = N_{i,t} + \Sigma_{j=1}^{k} H_{i-j,t-j}/p^{j}$$

While the formulae look daunting because of all the subscripts, in practice this repetitive calculation is straightforward to apply within a spreadsheet.

12.8 Summary

Animal numbers can be estimated by total counts, sampled counts, mark–recapture, or various indirect methods. In each case, the usefulness of the method is determined by how closely its underlying assumptions are matched by the realities of what the animals do and how difficult they are to see, trap, or detect. The range of methods provided should allow wildlife managers to choose one that will be adequate in any given circumstance.