Ch. 19

Distance sampling[1]

"There is no object so large but that at a great distance from the eye it does not appear smaller than a smaller object near."
-- Leonardo da Vinci

Questions to ponder:

- *Why is it necessary to record the distance to the animal during a survey?*
- *How can I estimate the effective strip width for a distance-based survey?*
- *What is a detection function?*
- *When should I use a point-count design, and when should I use line transects?*

Disappearing act

Ecologists often use visual surveys to count animals. And, as we explored in Chapter 14, there are many reasons to suspect that a count (c) does not equal the number of animals in a given area (N). Thus, we are usually very certain that $c \neq N$.

Distance sampling is a form of data collection intended for a specific type of analysis that uses detection functions to model the probability of detecting an animal, given a distance away from the transect or point-count location. Distance sampling is based on a basic concept—the probability of detecting an animal decreases as its distance from the observer increases. This concept could almost be referred to as a natural law! That is, it is difficult to imagine a situation in which our ability to detect an animal does not decrease at some point away from a transect. Even an elephant will eventually be too small to be seen!

When ecologists record survey data using distance sampling techniques, we see—over and over—that this "law" holds. Bigger animals may be seen for longer distances, but eventually even the number of large animals declines with distance away from the surveyor (Figure 19.1).

The approach for data collection when using distance sampling along a transect is to (1) establish the transect, (2) walk along the transect looking or listening for animals (or animal sign, e.g., dung or nests), and (3) recording the perpendicular distance from the animal to the transect line (x in Figure 19.2).

---

[1] *With thanks for content to Gary White and Michael Conroy*

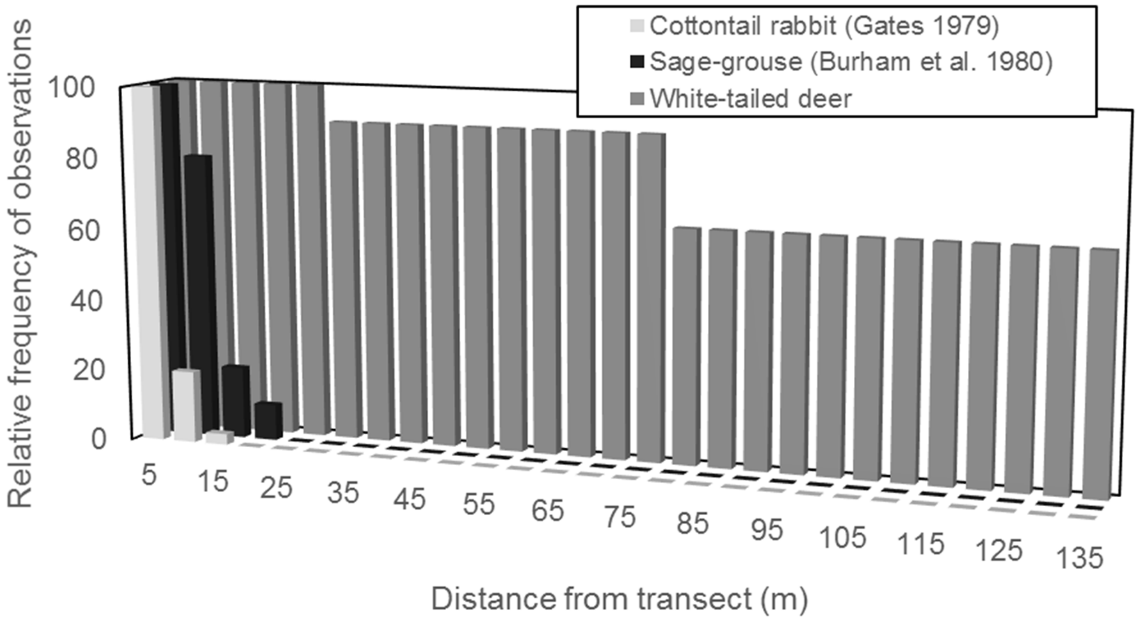

Figure 19.1: *Relative frequency of observations at given distances from a transect line for three species of wildlife during distance-based sampling: eastern cottontail (Sylvilagus floridanus) rabbits, greater sage-grouse (Centrocercus urophasianus), and white-tailed deer (Odocoileus virginianus). Figure after White 2007.*

A person may record the actual perpendicular distance (*x*) if that is possible, or the observer may record the radial distance to the animal (*r*) and the angle between the compass direction of the animal and the direction of the transect line (here, θ). Trigonometry can then be used to determine the perpendicular distance (*x*). Here, $\sin(\theta)=x/r$ so $x = \sin(\theta)/r$.

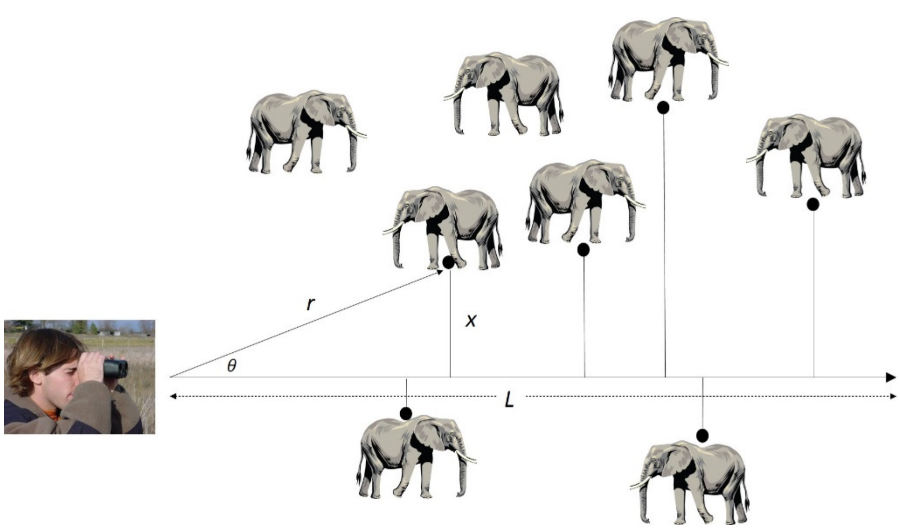

Figure 19.2: *Schematic showing perpendicular distance (x) recorded during a distance-based survey of detected (with distance) and undetected animals along a transect of length L.*

Building blocks

Distance sampling theory uses the set of perpendicular distances to estimate the density (*D*), or the number (*N*) of animals in a given area (*a*). We know that

$$D = \frac{N}{a}$$

But, we also know that **our count of animals is a sample**. Thus, we use *n* to represent the number of animals in our counted sample. And, if we look at a summary of our data, with observations binned into distance categories, we might see a pattern that looks something like the pattern in Figure 19.3b) at right. Simply put, it appears as if we have missed some animals at the farthest distances from the transect line. And, we can model that decline as a function—imagine placing a line through the tops of the bars in the histogram in Figure 19.3b. You would get the function shown in Figure 19.3c.

The key to understanding distance sampling is to understand that the area under the curve represents the size of the sample of animals that you saw (or heard). The area above the curve (Figure 19.3c, shaded) represents the number of animals that you missed—the animals that were present, but not detected. Obviously, if we do not correct for the missed animals, our density estimate is going to be biased (underestimated).

The correction provided by distance sampling is the probability of detection (here, $\hat{P}$):

$$\hat{D} = \frac{n}{a\hat{P}}$$

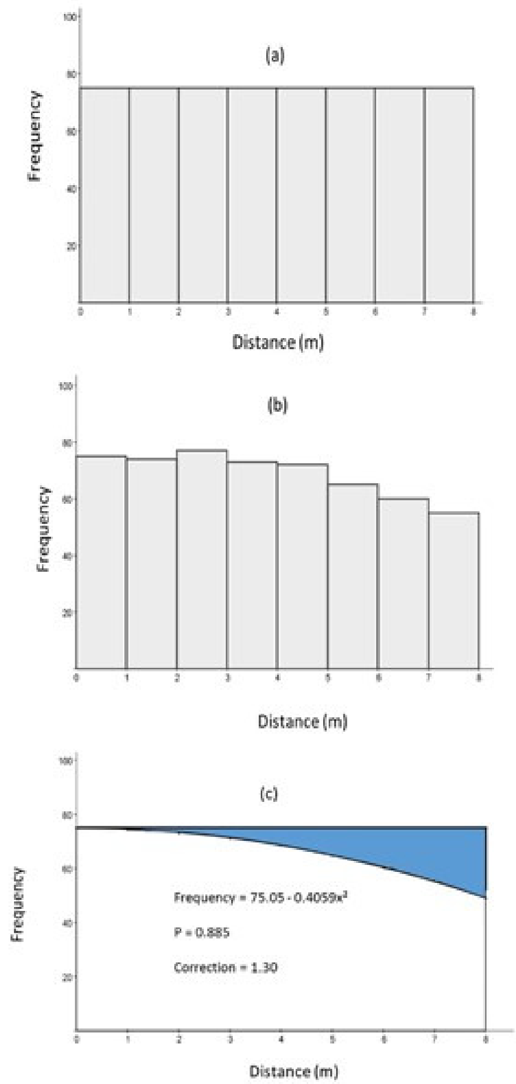

Figure 19.3: *survey results, in binned distance-from-line categories, expected with complete detectability (a) and incomplete detectability (b), as well as a detection function designed to describe the decline in detection as distance from line increases (c). Figure is after Buckland et al. 2001.*

Assumptions

The important **assumptions** inherent in analyses of distance-based data are as follows:

- Transect lines (or points) are randomly placed
- Objects directly on/at the line/point are detected with certainty (P = 1.0)
- Objects are detected at their initial location
- Measurements or groupings of data are correct
- Sightings of individuals are independent events

We assume random distribution of animals in space, relative to the placement of the transect line. Road surveys, for example, become problematic for distance surveys if animals are attracted to roads because of road edge habitats or provisioning from passing motorists (alternatively animals may avoid roads). Typically, it is better if transects are placed perpendicular (cutting across) possible gradients of habitat.

We assume that animals on the line or at the point are always detected, as we need a reference point for our detection curve. The curve for the probability of detection must start at 1.0 and decline in some manner. It is important for observers to collect exact measures, without error, because the estimation of the detection function is based upon this set of observations. As we will see, distance-sampling methods typically 'bin' or categorize distance data into groups, so the key is that animal observations are eventually placed into the correct distance category. Additionally, if surveys result in groups of animals (e.g., a covey of quail or a herd of deer) being detected at the same time, we will consider each group detection to be a single event. The user may indicate a single observation while also recording the size of the **cluster** of animals that was detected.

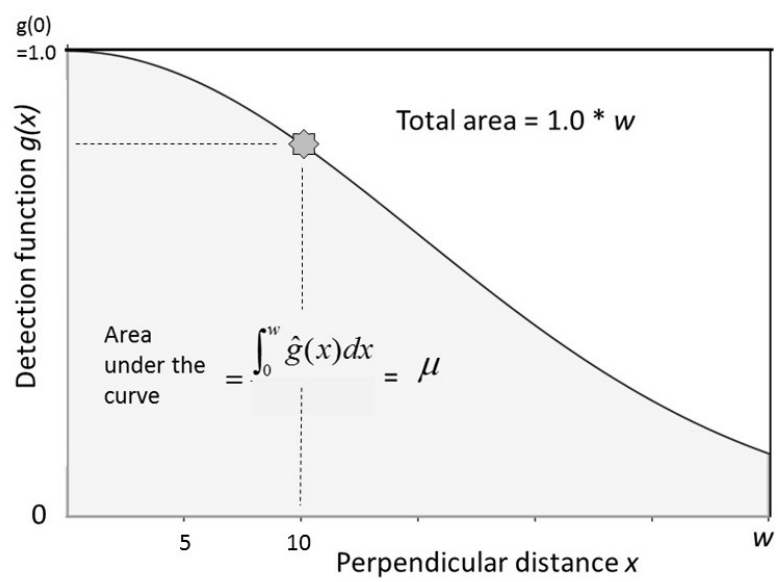

Figure 19.4: *Depiction of the detection function, g(x), declining from 1.0 to a minimum at, w, a distance from the transect. The value of g(10), or the detection probability at a distance of 10 units, is shown by dotted lines. Figure modified from Buckland et al. (2001).*

We can get a bit more specific in our analysis of the theoretical diagram (Figure 19.3c) for our detection function. First, we can name the function (the line) that describes the decline in probability of detection, g(*x*). Using this notation, we can refer to the probability at a certain distance (e.g., 10 meters) from the transect line as "g(10)", and we can see that the function would have a value (somewhere between 0 and 1.0) at the exact spot where x=10 in Figure 19.4. Second, if we imagine that our distances away from the line increase to some maximum distance, *w*, then we can also notice that the area of the entire rectangle (above and below the g(x) function) is 1.0*w—the rectangle has a height of 1.0 and a length of *w*. So, area=1.0*w = w.

Last, if we go back to our calculus course (perhaps you had no idea how useful that course would be when you took it, right!?), we know that it is possible to use the method of **integration** to estimate the area under the curve. Integration is a way to sum up the area under the curve by imagining a bazillion (yes, a bazillion) very skinny (in fact, infinitely small) rectangles packed together under the curve. Each rectangle, at distance *x*, has a height that matches the value of the function, g(*x*). So, we integrate them together with this formula,

$$\int_0^w g(x)dx = \mu$$

And we will use μ to represent this integration. So, the area under the curve equals μ. *That will be important.*

Conceptually, if the area under the curve represents the animals you saw, and the area of the rectangle represents all of the animals, we can estimate the detection probability, *P*, as the proportion of the rectangle that is under the curve:

P = area under curve/area of rectangle.

There is a theoretical problem with our goal to use the detection function to correct for the animals that we missed during our survey—**you literally cannot observe g(*x*), the detection function**. That is, we cannot observe and record probabilities, directly. But, we can record individual observations, and we can accumulate the frequency of our sightings away from the transect line as a **probability density function, f(*x*),** as in Figure 19.5. And that distribution of observations, f(*x*), can be used to estimate the number of animals that we missed. The area under the function f(*x*) is always scaled to equal 1.0, which makes f(x) a probability density function—in contrast to the detection function g(*x*) for which the area under the curve is much greater than 1.0. An

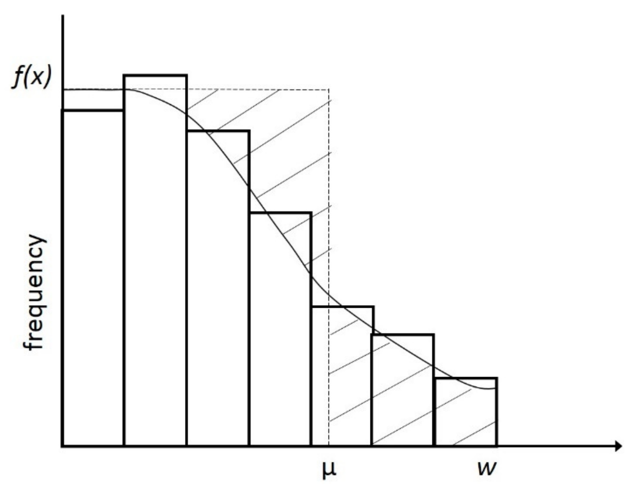

Figure 19.5: *The probability density function describing the distribution of perpendicular distance frequencies from a transect survey. The two shaded areas are equal in size so that the number of animals seen at distances >µ equals the number of animals not detected at distances <µ (figure after Buckland et al. 1998).*

example of a probability density function of which you may be familiar is the "bell curve", or normal distribution—100% of the students in a course fit under the function (the "curve").

As you can see in scenarios for line transects shown in Figures 19.4 and 19.5, line transect scenarios on previous pages, f(*x*), the distribution of observations, and g(*x*), the detection probability function, have the same shape. But, there are differences between f(*x*) and g(*x*). First, f(*x*) is built to describe how many animals are observed at distances away from the line—the function traces the histograms that summarize animal observations (Figure 19.6b), and the function is a **probability density function**.

In contrast, g(*x*) varies between 0 and 1.0 and describes the change in detection probability (Figures 19.4 and 19.6a). For point counts, the g(*x*) has the same general shape (or sometimes denoted g(*r*) for radial distances) as g(x) for transect counts (Figure 19.6a). But, f(*x*) (or f(*r*)) for point counts has a different shape (compare Figure 19.5 for transects and Figure 19.6b for point counts), because we expect fewer observations in the limited area near the point, because the center (e.g., 0-10 m from the point) of a 20-m radius circle has a smaller area than the area of the portion of the circle with radius 10-20 meters from the point. Detections increase with the area sampled away from the observer until some threshold distance whereupon detections begin to decline due to the effects of distance.

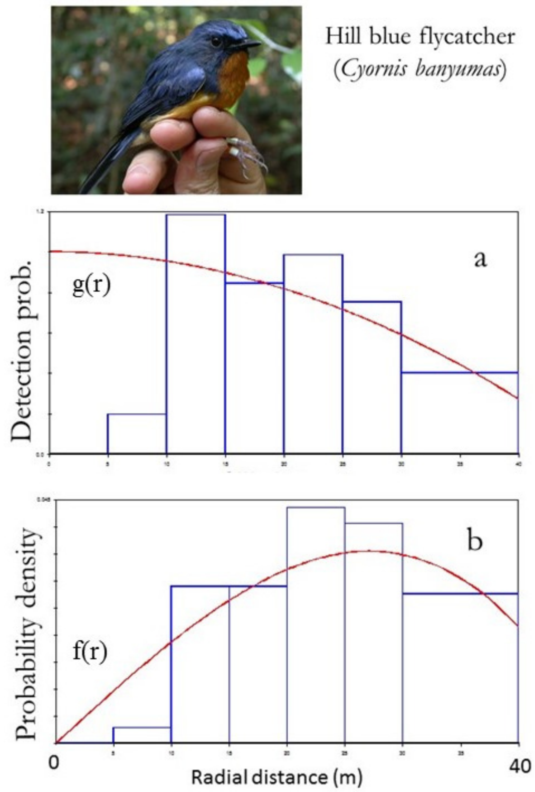

Figure 19.6: *Comparison of g(r), probability of detection at radial distance r (panel a), and f(r), the probability density function of observations at radial distances (panel b), for observations from point count surveys of hill blue flycatcher (Cyornis banyumas) in Thailand. Photo courtesy of G. Gale.*

The effective strip width concept

So, how do we estimate density if we don't know the area that we surveyed? Distance-sampling allows us to estimate the **effective area** that we surveyed, starting with the basic equation for density as the maximum likelihood estimator:

$$\hat{D} = \frac{N}{a}$$

And, as we previously noted, we are not able to count all N, and we are not completely certain about the area that we have counted, either. But, for a distance sampling research design with a transect of length L, and a theoretical width on either side of the transect, w, and an estimated detection probability, we can estimate density as:

$$\hat{D} = \frac{n}{2Lw\hat{P}}$$

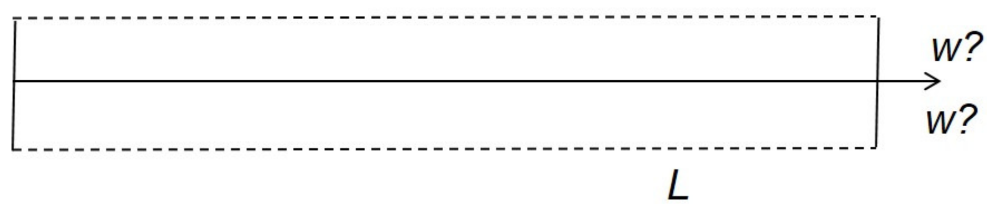

Figure 19.7: *Potential area surveyed along a line transect of length L and width, w. to each side of transect line.*

Distance sampling employs a concept described as the effective strip width, or more precisely effective "half-width." And, we should describe that to understand the logic of the method that distance sampling uses to estimate $\hat{D}$.

We start with the detection probability, P, which is the probability of detecting an animal in a strip surrounding a transect with a half-width (Figure 19.7) of w. Of course, w is unknown, as we do not need to enforce a fixed-width (or fixed-radius for point counts) rule when using distance sampling.

We have also already established that you can estimate P by dividing the area under the function, g(x) by the area of the 1*w rectangle (and remember that we've established that μ is going to represent the integration of the area under the curve). Thus

$$\hat{P} = \frac{\int_0^w g(x)dx}{w} = \frac{\mu}{w}$$

So, if we replace P in our formula for density, we find:

$$\hat{D} = \frac{n}{2Lw\frac{\hat{\mu}}{w}} = \frac{n}{2L\hat{\mu}}$$

Through simplification, we see that μ ends up standing in the place of the "strip width" of the rectangle around the transect line (a=2Lw, now a=2Lμ). So, we can refer to μ as the effective strip width—the realized width on either side of our transect line after we account for our detection probability (P).

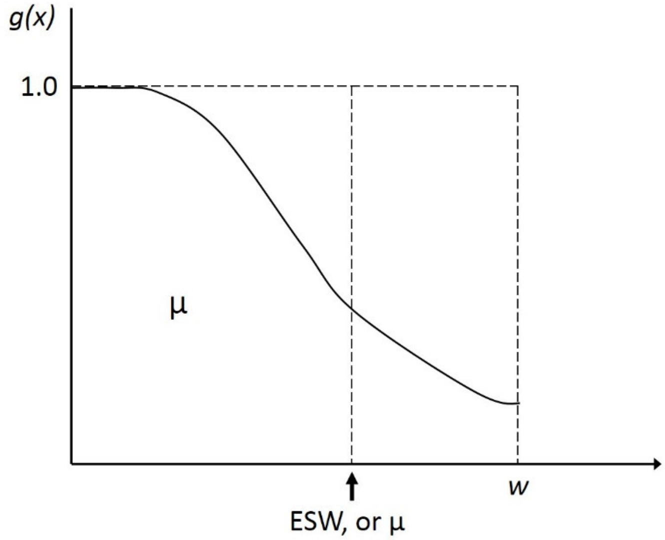

Figure 19.8: *The concept of μ: first, μ is the area under the detection function g(x) and takes a value between 0 and 1.0*w=w; second, μ is also the effective strip width, and takes a linear value between 0 and w (figure after Buckland et al. 1998).*

Brain warp

Notice that μ was first defined as the *area* under the curve.

Now, we also see that μ is a *width* (a point). Wrap your head around that! So, μ is one dimensional <u>and</u> two dimensional!

To visualize how μ, the effective strip width, can be seen as a distance <u>and</u> an area, consider Figure 19.8. We can define the effective strip width, μ, as the distance at which the number of animals missed within that distance is equal to the number of animals detected beyond that point. So, if we split our function, g(x) at the ESW, or μ, the area under the curve should equal the area above the curve for the other half of g(x).

Now, let's summarize the logic of distance sampling:

- We don't know the width of our transect (*w*) because of incomplete detection.
- We <u>do</u> know the distribution of our observations.
- We can fit a curve to observations that estimates probabilities of detecting at any distance, *x*. We can create a probability density function.
- That allows us to find ESW, or μ.
- We can use μ as our "w" to estimate D, because μ represents the width at which we have a count that is equivalent to of all animals in our sample (equal to our raw count).

The detection function

It should now be clear that the shape of the detection function is going to be important, as it will affect the effective half-width of the strip. And, although program Distance (the primary program used for distance sampling analyses) provides short-cuts to create the linear models used in the analysis, those equations are important to using and interpreting results from distance sampling.

Distance allows the fitting of complex detection functions of the form:

$$g(x) = key(x)[1 + series(x)]$$

Each function is based on a **key function** that determines the general shape of the function.

The key functions used by Distance are (Figure 19.9):

- **Half normal**—describes a gentle decline in probability of detection that follows the well-known 'bell curve' shape
- **Uniform**—describes a scenario in which probability of detection does not fall
- **Negative exponential**—describes an immediate drop in probability of detection at a very fast rate (an exponential rate!)
- **Hazard**—describes a scenario in which the probability of detection remains high until a threshold point, at which detection drops precipitously.

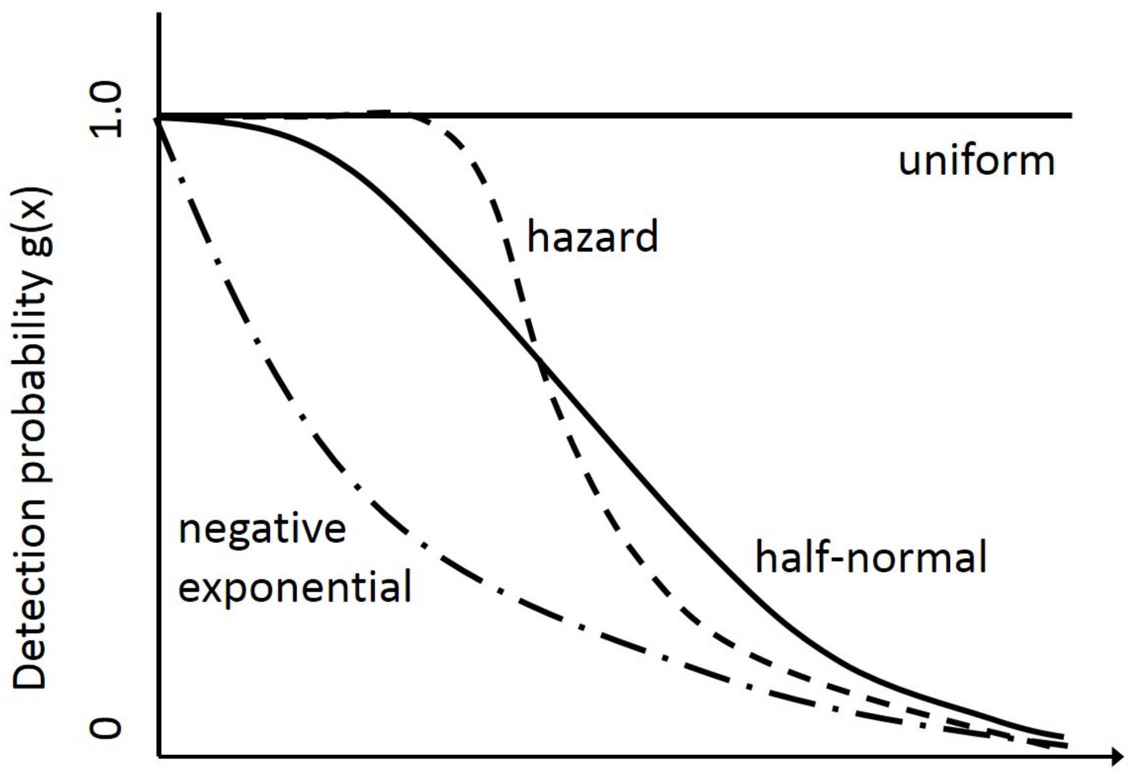

Figure 19.9: *General shapes of 4 key functions used by program Distance to describe the decline in detection as a function of distance from the transect line or point.*

Program Distance uses the four basic key functions, or shapes (Figure 19.9), because they are generally flexible shapes with fairly simple mathematical structure (and thus have low variance), and each function meets the criteria that we must have g(0) = 1. That is, the functions can all have some type of “shoulder” at $x = 0$ before declining as distance increases away from the transect line.

Then, Distance uses a **series adjustment** that provides a “tweak” to the basic shape of the function—especially at the far end. These adjustments (Figure 19.10) allow the g(x) function to better fit the data. However, the adjustments also add parameters to the function, and thus make it more complex—so, we use AIC to assess the models.

- Cosine
- Simple polynomial
- Hermite polynomial

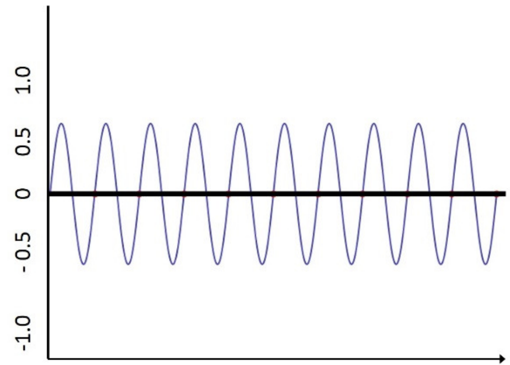

Figure 19.10: *An example of a cosine function used as a series adjustment to key functions by Distance.*

We can look at an example of how a key function can be modified to provide better fit to distance-based survey data. For example, if we were to try to use a uniform key function without any adjustments (Figure 19.11a) , our eyes tell us that it does not do a good job at describing the manner in which the probability of detection appears to decline with distance away from the transect line.

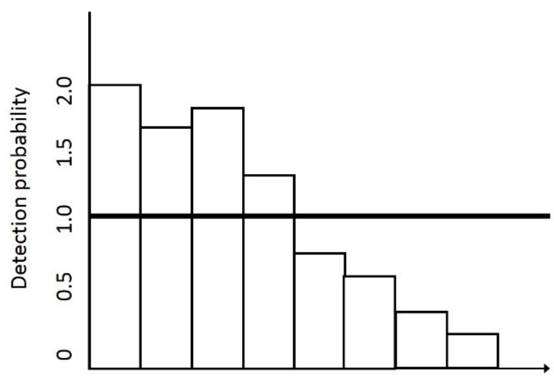

Figure 19.11a: *A uniform distribution fitted to survey observations (Figure after White 2007).*

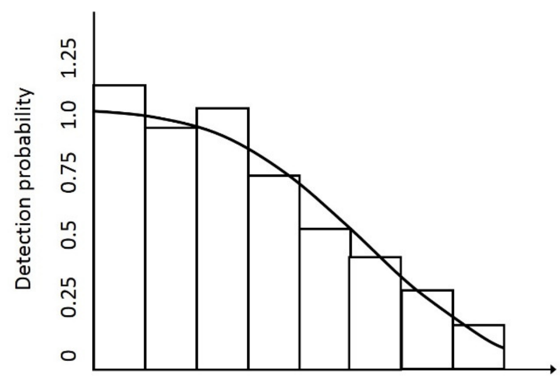

Figure 19.11b: *Observations from a survey fitted with a uniform key function with a cosine adjustment term (after White 2007).*

But, we can ask the software to add one or more cosine adjustment terms to the linear model describing the shape of the detection function. And, for example, when we add one cosine adjustment term (Figure 19.10) to the uniform key function, we get an adjusted uniform detection function (Figure 19.11b).

The adjusted function (Figure 19.11b) appears to fit the distribution of our observations much better. And, we note that it looks somewhat similar to the half-normal function. However, the detection probability does not fall towards 0 as quickly as the half normal would. For the other adjustment terms, the simple polynomial and hermite polynomial, the adjustments are applied similarly to cosine function except using a polynomial or a hermite polynomial series adjustment. The two polynomial adjustments often produce identical results, but the adjusted functions will differ in the amount and gradualness of the curvature relative to the cosine adjustment depending on the data being fitted.

Step-down model selection

Program Distance (Thomas et al. 2010) provides a robust example of the **step-down model selection** process that we discussed in Chapter 4. The purpose of this process is to discard models with uninformative parameters. Before exploring whether density of animals varies between geographic strata or other groupings, Distance compares the competing models (*key functions*) that describe the decline in detection probability. In fact, Distance starts the selection process with an examination of the best way to "tweak" the key function—as we saw above with the cosine adjustment of the uniform key function.

Thus, Distance first uses AIC (or AIC_c, see Chapter 4) to select the best adjustment term for one or more key functions. Then, the second step is to take the best form of each key function (now modified with an adjustment term) to compare with one or more of the other key functions. Now, Distance uses a second AIC assessment to compare the half-normal with, for example, the hazard and the uniform key functions. And, the last step is to compare the models that represent various hypotheses about density—using the best description of detection probability.

Distance does not evaluate the density hypotheses by comparing all key functions with each possible adjustment term in a simultaneous model comparison. *No, Distance discards uninformative models along the three-step process.* It is also important to remember that AIC cannot be used to compare between different datasets. For example, AIC cannot be used to compare between models if the data are grouped or truncated differently. In such cases the coefficient of variation is useful for deciding where to truncate or how to group the data before proceeding with model selection.

Point-counts

To this stage, we've mostly discussed the application of distance sampling to transect surveys. But, many people use point counts to survey for birds or butterflies or other species. The theory of distance sampling can be applied to point-counts as well. Some research biologists will refer to such efforts as **variable circular plots** (Figure 19.12), meaning that there is no fixed radius used for the survey. If an animal is seen or heard, the

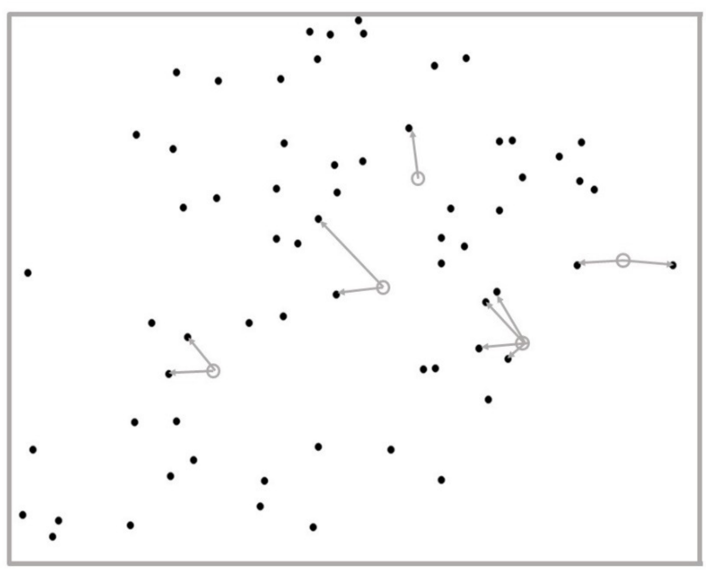

Figure 19.12: *Depiction of 5 point counts conducted on a study site with 11 observed animals; relative radial distances are shown (Figure modified from Buckland et al. 2001).*

distance to that animal is recorded. Thus, our data will consist of sets of radial distances from the point to the animal.

Transects vs. point-counts

When should you use line transects and when should you use point-counts? We can provide some general suggestions for your consideration as you design a study.

Line transects are generally best for:

- Sparsely distributed populations for which sampling needs to be efficient (e.g., whales, deer, grassland birds). You might consider using line transects if you think that you will often not count any individuals at a given point-count location.
- Populations that occur in well-defined clusters, and at low or medium density.
- Populations that are detected through a flushing response, where the act of walking the transect helps flush the animals.

Point-counts are generally best for:

- Patchily distributed populations, as you can easily stratify points to patches.
- Populations that occur in difficult terrain, or with problematic access—walking a straight transect as you survey for animals in these locations may be hazardous to your health if you cannot keep track of where you should be stepping!
- Associating survey data with associated habitat characteristics—a point is located at a discrete location at which you can measure vegetation. Associating vegetation with transects is more difficult, as transects are continuous.
- Populations that occur at medium or high densities—points are not as effective at low densities, because you obtain very few observations from each point.

Details to consider

Not everyone's data will follow a nice continuously declining function. In the songbird example depicted in Figure 19.13a, you can see that the surveyors in Nebraska detected western meadowlarks a considerable distance from the line. But, there were also gaps where animals were not seen at some distance categories. If you try to imagine fitting one of the basic key functions to this data, you realize it may not be easy—the fit may be poor.

More critically, these outlier observations can affect the shape of the detection functions fairly dramatically. **Right truncation** is recommended (Buckland et al. 2001), which means that the outliers should be removed before analysis begins (Figure 19.13). Although some ecologists use a rule, such as removing 5% of the observations to remove the outliers, there are good reasons to use a **visual assessment** to determine which outliers that should be removed. It is possible, for example, that a portion of the top 5% of the distances will not be in the 'outlier' group, and thus you will have removed some of your sample for no good reason. *We always recommend looking at summaries of your data before jumping into the analyses.*

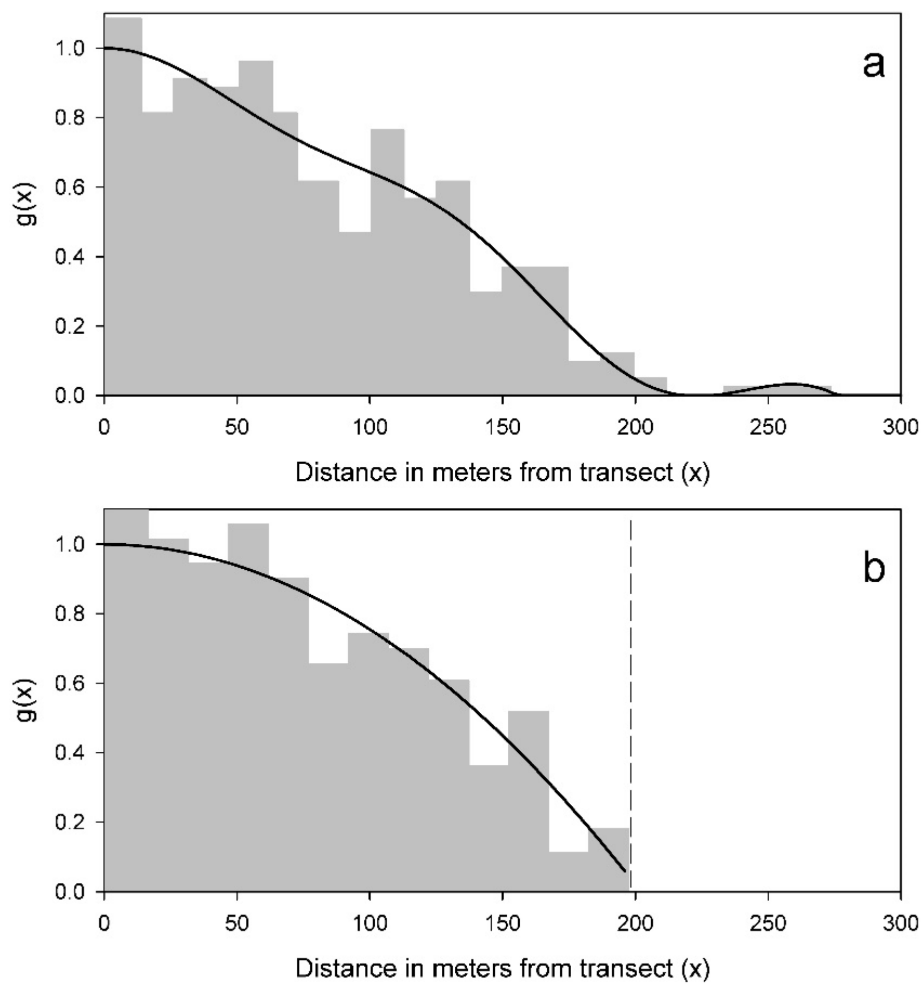

Figure 19.13: *Detection functions, g(x), for observations of western meadowlarks (Sturnella neglecta) from line transects in grasslands in Nebraska, USA (Kempema 2007). The full data set is shown at top (a), and the right-truncated data set is shown below (b, the data has also been grouped into larger bins).*

If right truncation is a good thing, perhaps **left truncation** is also a good thing? *We suggest not.*

Some ecologists might be tempted to use left truncation when working with point-counts. One potential problem that is encountered when using distance sampling and point-counts of grassland birds for example, is that the observer can create a "halo effect" (inset, Figure 19.14). As the observer walks to the point-count location, birds may move away from the point. Then, the distribution of the observations may not meet our expectations—we may have fewer observations that we expected near the 0-point (Figure 19.14).

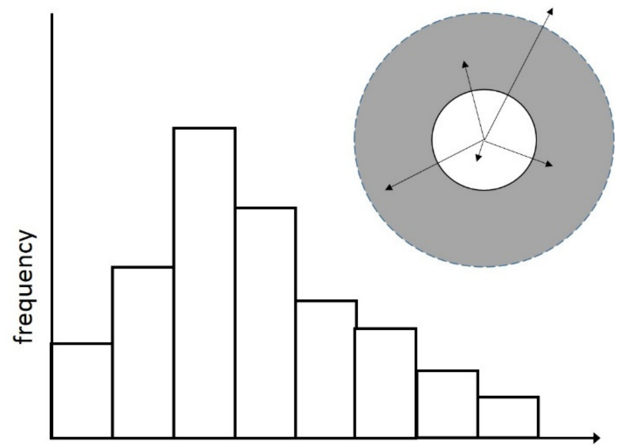

Figure 19.14: *Observations from a point-count survey, showing effects of the observer's presence at the point; inset, the effects of left truncation on the area included in the estimate of density.*

Left truncation removes the observations from the left side of the figure here. And it reduces the effective diameter of the circle by creating a 'doughnut' shape—a hole is created in the middle of the point count.

The problem with this approach is that you still have the observations immediately to the right of the left truncation zone. And, those observations consist of some animals that left the truncated portion of the point-count in response to the presence of the observer. As they moved outwards, they increased (biased) the density higher in the outer zones—there are now more animals in the outer zone than there should be. And, the density estimate will be biased high. For this reason, left truncation is not recommended. Left truncation is only advocated for a very limited set of circumstances where observations close to the line are obscured or problematic (see Buckland et al. 2001).

If you are faced with this problem, you may wish to consider using ***N*-mixture modeling** (Chapter 18) as an approach to compare relative abundance, rather than distance sampling. The *N*-mixture modeling design is just based on the count—the movement of the animals is not a problem in this type of study design.

Distance needs data: a lot of data

In our primer, we've tried to avoid explicitly making our discussions based on the software you might use to analyze your data. However, a great majority of people attempting a distance-based approach will use program Distance. When you receive your results from Distance, you will be given a table of parameter estimates. These parameters are mostly "nuisance" parameters—parameters that had to be estimated to allow estimation of density ($\hat{D}$).

Here are the parameter estimates from a fox squirrel (*Sciurus niger*) survey in Lincoln, Nebraska, USA by University of Nebraska-Lincoln students. Students conducted point-counts on campus.

The model selected through step-down model selection was the hazard rate key function with no adjustment terms:

```
      Effort              :    58.00000
      # samples           :    58
      Width               :    169.0000
      # observations      :    67

 Model
    Hazard Rate key, k(y) = 1 - Exp(-(y/A(1))**-A(2))


              Point        Standard    Percent Coef.        95 Percent
  Parameter   Estimate      Error      of Variation     Confidence Interval
  ---------  -----------  -----------  --------------  ----------------------
    A( 1)    30.97         4.493
    A( 2)     3.350        0.4096
    h(0)      0.001016     0.000203      19.99           0.000684    0.001509
    p         0.068902     0.013776      19.99           0.046400    0.102320
    EDR      44.361        4.4346        10.00          36.351      54.137
```

We show you these results to impress upon you that distance-sampling requires the estimation of many parameters—not just density. Here, *A(1)* and *A(2)* are parameters used in the non-linear model (hazard function) for probability of detection (Figure 19.15). And, this key function has only two parameters—if adjustment terms had been selected, we would see parameter estimates for the parameters used to adjust the key function, which would make our model even more complex.

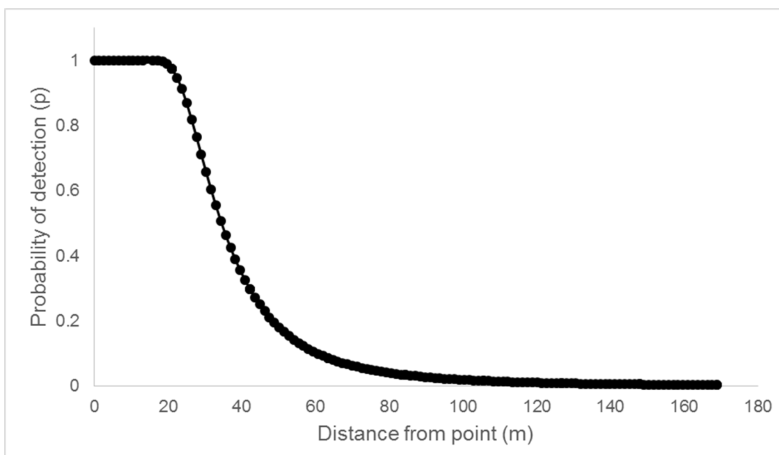

Figure 19.15: *Detection curve (probability of detection, g(x)) for a fox squirrel survey conducted in September 2014 by students at the University of Nebraska-Lincoln, USA. The curve is shaped as a hazard function, which was selected as the model to describe the decline in detection as distance increased from the point count location (goodness of fit:* X^2_2 *= 0.21, P = 0.90).*

Then, we see three parameters, *h(0)* (a measure of the effective detection area), *p* (probability of detection for the study), and EDR (the effective detection radius for the squirrel point counts) that can be estimated after f(x) and g(x) are established. Then, as a third step, density is estimated using EDR and *p*. The probability of detection for animals near points was low, and the detection curve (Figure 19.15) declined much more rapidly than students anticipated before conducting the survey—trees, shrubs, and buildings on campus most likely combined with the small size and camouflage of the squirrel such that students had less than a 20% chance of seeing a squirrel that was more than 40 meters from the survey point. In our example, the density for fox squirrels, per hectare, was estimated as:

```
                 Estimate      %CV     df     95% Confidence Interval
                  ------------------------------------------------------
           D      1.8685      20.68   73.70      1.2427        2.8094
```

It should now be clear that our estimate for density, and its variance, are both a function of several other parameters. If you do not have enough observations for a species, it will be difficult to find a well-fitting description of the decline in detection probability away from your lines or points. The uncertainty in the key function for detection will manifest itself as high variance associated with your detection probability, p. And, high variance in p will result in high variance for D.

Therefore, study design and forethought are critical for distance-sampling. The University of Nebraska squirrel survey worked well to provide a reasonably precise estimate of density. But, you must plan to conduct enough effort to collect adequate data (see Buckland et al. 2001 for advice on planning). In general "adequate" data entirely depends on the quality of the detection function- if the data are messy even a sample of 120 might not be large enough but on the other hand a well-formed detection function might give a reasonable estimate with as few as 30 detections.

Distance-sampling has one clear advantage over all other types of survey design and analysis—**it is the only method to provide *bona fide* (direct) density estimates**. All other methods estimate abundance (N), which may be transformed to density (D) estimates only after making some rather important, and often very ill-informed, assumptions about the area from which the estimate of N was made.

So, the trade-off for the use of distance sampling is between amount of data needed for the parameter estimates and the precision and bias of the density estimate provided. You must evaluate this tradeoff for your research questions.

Availability…?

Surveys used to estimate density should account for detectability (the probability that an animal is detected if it is present and available) and availability (the probability that the animal is available to be detected). Animals below ground (see Chapters 13 and 14) or quietly hidden in vegetation are unavailable for sampling, as it is impossible to see or hear them. Distance sampling, in its most basic form, only deals with detectability, and most users of distance sampling typically assume an availability of 1.0. This is often unrealistic (Diefenbach et al. 2007), but clearly depends on the species of interest. Availability for some species is often as low as 0.1, and unadjusted distance sampling may underestimate density by as much as a factor of 4. We refer the reader to discussion of a number of methods to account for availability of singing/calling birds by Diefenbach et al. (2007).

Conclusion

Distance sampling uses a maximum-likelihood estimation process to estimate density. The method addresses a basic "law" regarding observations during surveys—the probability of detection for an animal declines the farther the animal is from the observer. Distance sampling is rigorous and robust if assumptions are met. In its simplest form, distance sampling does not provide an estimate of the probability of availability, but it deals very effectively with detectability. Distance sampling can be applied to many species and types of surveys.

References

Buckland, S.T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1998. Distance sampling, in Encyclopedia of Biostatistics, P. Armitage and T. Colton, eds, John Wiley & Sons, Ltd, Chichester.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, New York.

Diefenbach, D. R., M. R. Marshall, J. A. Mattice, and D. W. Brauning. 2007. Incorporating availability for detection in estimates of bird abundance. Auk 124:96–106.

Kempema, S. L. F. 2007. The influence of grazing systems on grassland bird density, productivity, and species richness on private rangeland in the Nebraska Sandhills. MS Thesis, University of Nebraska-Lincoln, Lincoln, NE.

Thomas, L., S.T. Buckland, E.A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R.B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology 47: 5-14.

White, G. 2007. FW663 -- Sampling & Analysis of Vertebrate Populations. On-line: http://warnercnr.colostate.edu/~gwhite/fw663/ Accessed 15 February 2015.

For more information on topics in this chapter

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

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

Citing this primer

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

Caught Napping Publications
Lincoln, Nebraska USA