1

Introductory concepts

1.1 Introduction

Ecology is the study of the distribution and abundance of plants and animals and their interactions with their environment. Many studies of biological populations require estimates of population density (D) or size (N), or rate of population change $\lambda_t = D_{t+1}/D_t = N_{t+1}/N_t$. These parameters vary in time and over space as well as by species, sex and age. Further, population dynamics and hence these parameters often depend on environmental factors.

This book is a synthesis of the state-of-the-art theory and application of distance sampling and analysis. The fundamental parameter of interest is density (D = number per unit area). Density and population size are related as $N = D \cdot A$ where A is area. Thus, attention can be focused on D.

Consider a population of N objects distributed according to some spatial stochastic process, not necessarily Poisson, in a field of size A. A traditional approach has been to establish a number of plots or quadrats at random (e.g. circular, square or long rectangular) and **census** the population within these plots. Conceptually, if n objects are counted within plots of total area a, then an estimator of density, termed $\hat{D}$, is

$$\hat{D} = n/a$$

Under certain reasonable assumptions, $\hat{D}$ is an estimator of the parameter $D = N/A$. This is the finite population sampling approach (Cochran 1977) and was fully developed for most situations many years ago. This approach asks the following question:

> Given a fixed area (i.e. the total area of the sample plots), how many objects are in it (Fig. 1.1)?

Distance sampling theory extends the finite population sampling approach. Again, consider a population of N objects distributed according to some stochastic process in a field of size A. In distance sampling theory, a set of randomly placed lines (Fig. 1.2) or points (Fig. 1.3) is established and distances are measured to those objects detected by travelling the line or surveying the points. The theory allows for the fact that some, perhaps many, of the objects will go undetected. In addition, there is a marked tendency for detectability to decrease with increasing distance from the transect line or point. The distance sampling approach asks the following question:

> Given the detection of n objects, how many objects are estimated to be within the sampled area?

Two differences can be noted in comparing distance sampling theory with classical finite population sampling theory: (1) the size of the sample area is sometimes unknown, and (2) many objects may not be detected for whatever reason. One of the major advantages of distance sampling is that objects can remain undetected (i.e. it can be used when a census is not possible). As a particular object is detected, its distance to the randomly chosen line or point is measured. Thus, distances are sampled. Upon completion of a simple survey, n objects have been

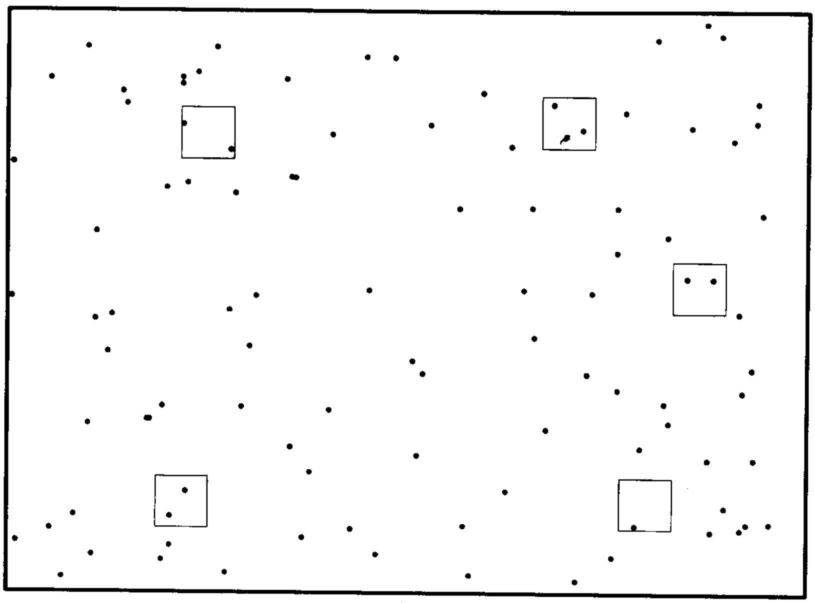

Fig. 1.1. Finite population sampling approach with five 1 m square quadrats placed at random in a population containing 100 objects of interest. $\sum a_i = 5$, $\sum n_i = 10$, and $\hat{D} = 2$ objects/m^2. In this illustration, the population is confined within a well-defined area.

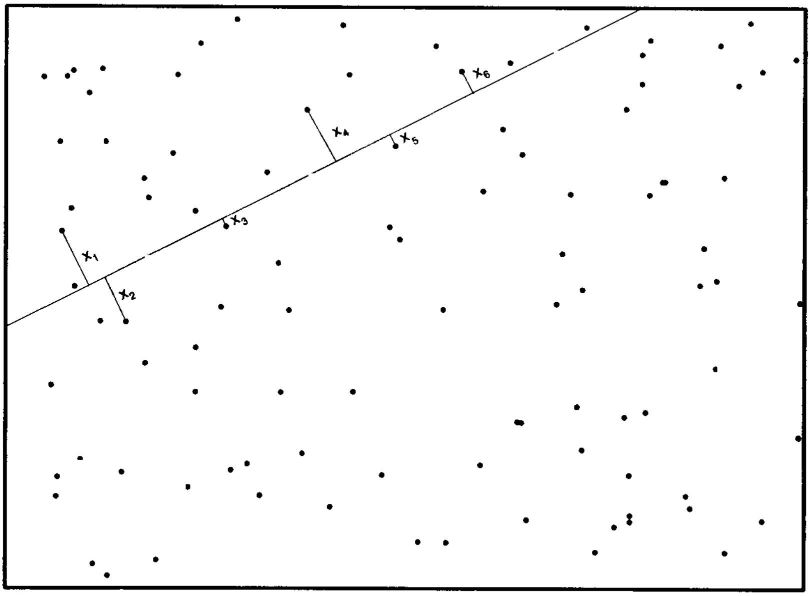

Fig. 1.2. Line transect sampling approach with a single, randomly placed, line of length L. Six objects ($n = 6$) were detected at distances $x_1, x_2, \ldots, x_6$. Those objects detected are denoted by a line showing the perpendicular distance measured. In practical applications, several lines would be used to sample the population.

detected and their associated distances $y_1, y_2, \ldots, y_n$ recorded. The variable y will be used as a general symbol for a distance measurement, while x will denote a perpendicular distance and r will denote a radial distance. Unbiased estimates of density can be made from these distance data **if certain assumptions are met**.

Distance sampling theory includes two main approaches to the estimation of density: line transects and point transects. Traditional sampling theory may be considered a special case of distance sampling theory. An application of point transect theory is the sampling method called a trapping web, which is potentially useful in animal trapping studies. Cue counting is another application of point transect theory and was developed for marine mammal surveys. Nearest neighbour and point-to-object methods are similar in character to point transects, but are generally less useful for estimating object density.

1.1.1 Strip transects

Strip transects are long, narrow plots or quadrats and are typically used in conjunction with finite population sampling theory. Viewed differently,

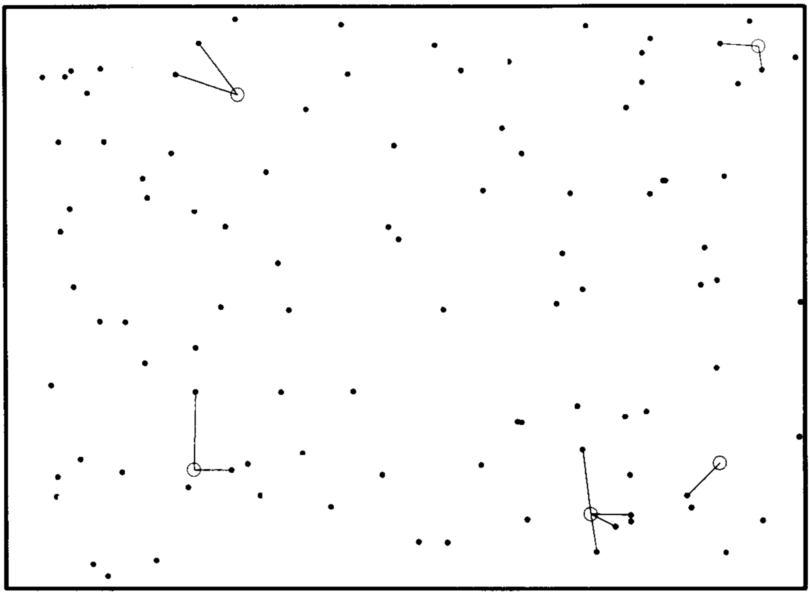

Fig. 1.3. Point transect sampling approach with five randomly placed points ($k = 5$), denoted by the open circles. Eleven objects were detected and the 11 sighting distances $r_1, r_2, \ldots, r_{11}$ are shown.

they represent a very special case of distance sampling theory. Consider a strip of length L and of width $2w$ (the width of the area censused). Then, it is **assumed** that all objects are detected out to distance w either side of the centreline, a complete census of the strip. No distances are measured; instead, the strong assumption is made that all objects in the strip are detected. Detections beyond w are ignored. Line and point transect surveys allow a relaxation of the strong assumptions required for strip (i.e. plot or quadrat) sampling (Burnham and Anderson 1984). Note the distinction here between a **census**, in which all objects in an area are counted, and a **survey**, where only some proportion of the objects in the sampled area is detected and recorded.

1.1.2 Line transects

Line transects are a generalization of strip transects. In strip transect sampling one assumes that the entire strip is censused, whereas in line transect sampling, one must only assume a narrow strip around the centreline is censused; that is, except near the centreline, there is no assumption that all objects are detected. A straight line transect is

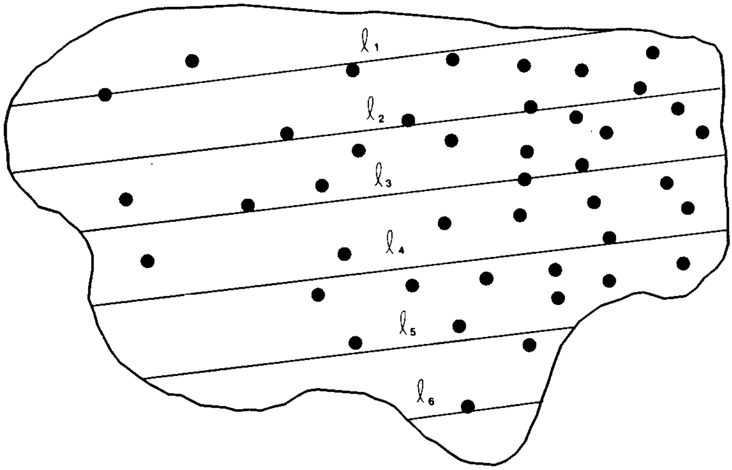

Fig. 1.4. A population of objects with a gradient in density is sampled with lines parallel to the direction of the gradient. In this case, there are $k = 6$ lines of length $l_1, l_2, \ldots, l_6$, and $\sum l_i = L$.

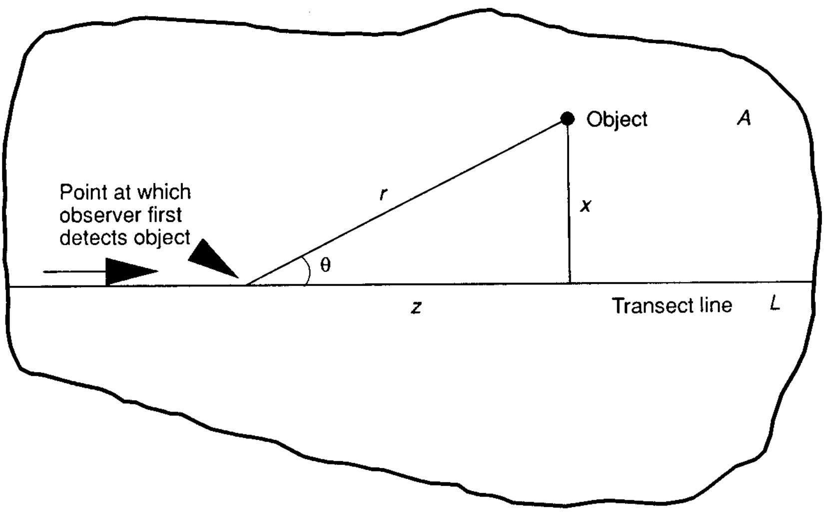

Fig. 1.5. Basic measurements that can be taken in line transect surveys. Here an area of size A is sampled by a single line of length L. If sighting distances r are to be taken in the field, one should also measure the sighting angles θ, to allow analysis of perpendicular distances x, calculated as $x = r \cdot \sin(\theta)$. The distance of the object from the observer parallel to the transect at the moment of detection is $z = r \cdot \cos(\theta)$.

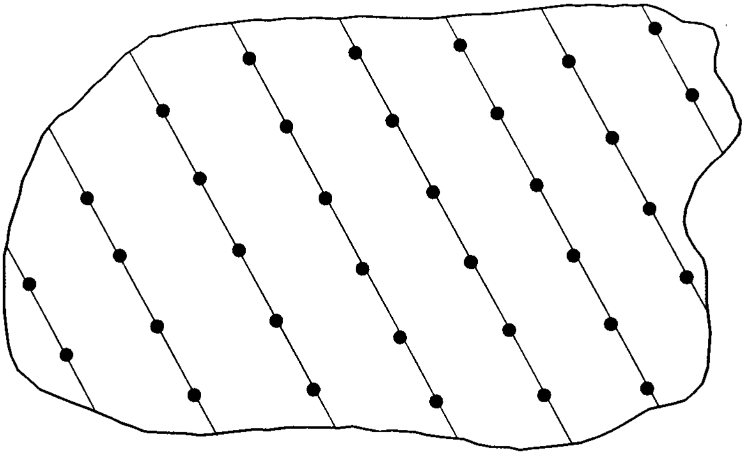

Fig. 1.6. Point transect surveys are often based on points laid out systematically along parallel lines. Alternatively, the points could be placed completely at random or in a stratified design. Variance estimation is dependent upon the point placement design.

traversed by an observer and perpendicular distances are measured from the line to each detected object (Fig. 1.2). The line is to be placed at random and is of known length, L. In practice, a number of lines of lengths $l_1, l_2, \ldots, l_k$ are used and their total length is denoted as L (Fig. 1.4). Objects away from the line may go undetected and, if distances are recorded accurately, reliable estimates of density can be computed.

It is often convenient to measure the sighting distance r_i and sighting angle θ_i, rather than the perpendicular distance x_i, for each of the n objects detected (Fig. 1.5). The x_i are then found by simple trigonometry: $x_i = r_i \cdot \sin(\theta_i)$. Methods exist to allow estimation of density based directly on r_i and θ_i. They are reviewed by Hayes and Buckland (1983), who show that they perform poorly relative to methods based on perpendicular distances, because they require more restrictive, and generally implausible, assumptions. In addition, observations made behind the observer (i.e. $\theta_i > 90°$) are problematic for models based on sighting distances and angles.

1.1.3 Point transects

The term point transect was coined because it may be considered as a line transect of zero length (i.e. a point). This analogy is only of limited conceptual use because there are several differences between line and point transect theory. Point transects are often termed variable circular plots in the ornithological literature, where the points are often placed at intervals along straight line transects (Fig. 1.6). We consider a series of k points positioned randomly. An observer measures the sighting (radial) distance r_i from the random point to each of the objects detected. Upon completion of the survey of the k points, one has distance measurements to the detected objects. Point transects are a generalization of traditional circular plot surveys. In circular plot sampling, an area of πw^2 is censused, whereas in point transect sampling, only the area close to the random point must be fully censused; a proportion of objects away from the random point but within the survey area remains undetected.

The area searched in strip and line transect sampling is $2wL$, whereas the area searched in circular plot and point transect sampling is $k\pi w^2$ (assuming, for the moment, that w is finite). In strip and traditional circular plot sampling, it is assumed that these areas are censused, i.e. all objects of interest are detected. In line and point transect sampling, only a relatively small percentage of the objects might be detected within the searched area (of width $2w$ for line transects or radius w for point transects), possibly as few as 10–30%. Because objects can remain undetected, distance sampling methods provide biologists with a powerful yet practical methodology for estimating density of populations.

1.1.4 Special applications

Distance sampling theory has been extended in two ways that deserve mention here: trapping webs and cue counts. These important applications are useful in restricted contexts and are direct applications of existing distance sampling theory. Two spatial modelling methods sometimes termed 'distance sampling' are more familiar to many botanists, but have limited use for estimating object density. These methods are point-to-object and nearest neighbour methods; they have some similarities to distance sampling as defined in this book, but differ in that there is no analogy to the detection function $g(y)$.

(a) *Trapping webs* Trapping webs (Anderson *et al.* 1983; Wilson and Anderson 1985b) represent a particular application of distance sampling theory and provide a new approach to density estimation for animal trapping studies. Traps are placed along lines radiating from randomly chosen points (Fig. 1.7); the traditionally used rectangular trapping grid cannot be used as a trapping web. Here 'detection' by an observer is replaced by animals being caught in traps at a known distance from the centre of a trapping web. Trapping continues for t occasions and only the data from the initial capture of each animal are analysed. Trapping webs provide an alternative to traditional capture–recapture sampling where density is of primary interest.

(b) *Cue counting* Cue counting (Hiby 1985) was developed as an alternative to line transect sampling for estimating whale abundance from sighting surveys. Observers on a ship or aircraft record all sighting cues within a sector ahead of the platform and their distance from the platform. The cue used depends on species, but might be the blow of a

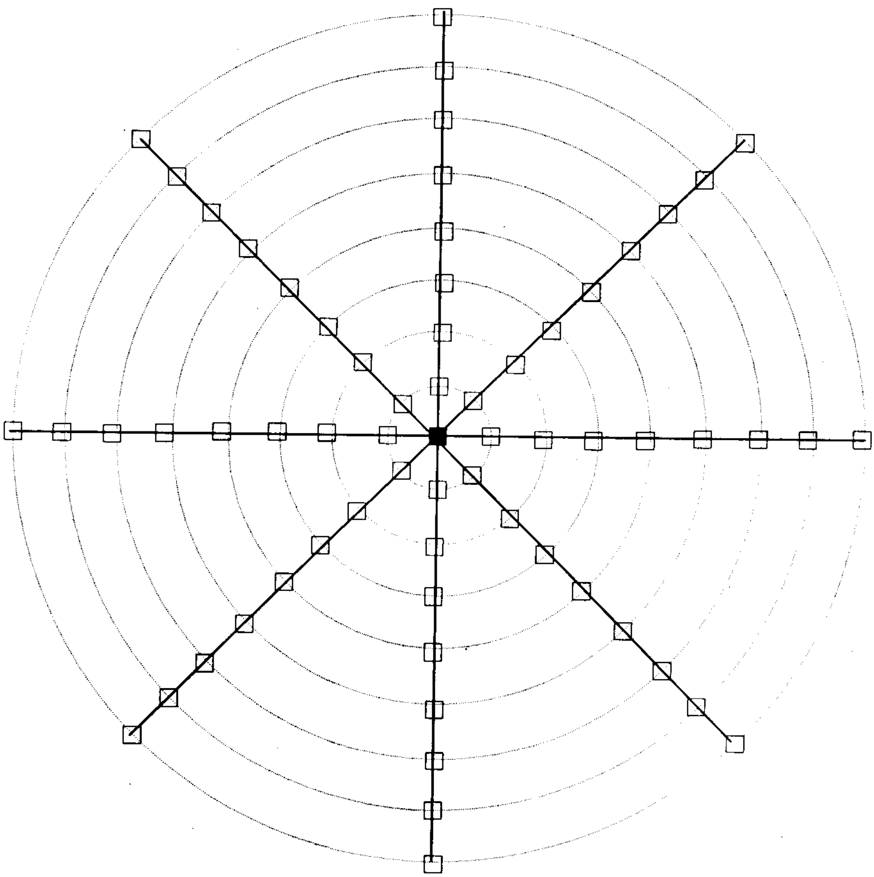

Fig. 1.7. Use of a trapping web to sample small mammal populations is an extension of point transect theory. Traps (e.g. live traps, snap traps or pitfall traps), represented as □, are placed at the intersections of the radial lines with the concentric circles.

whale at the surface. The sighting distances are converted into the estimated number of cues per unit time per unit area using point transect models. The cue rate (usually corresponding to blow rate) is estimated from separate experiments, in which individual animals or pods are monitored over a period of time.

(c) *Point-to-object methods* In point transect sampling, the distance of each detected object from the point is recorded. In point-to-object methods, the distance of the **nearest** object from the point is recorded (Clark and Evans 1954; Eberhardt 1967). The method may be extended, so that the distances of the n nearest objects to the point are recorded (Holgate 1964; Diggle 1983). Thus the number of detected objects from a point is predetermined, and the area around the point must be searched exhaustively to ensure that no objects are missed closer to the point than the farthest of the n identified objects. Generally the method is inefficient for estimating density, and estimators are prone to bias.

(d) *Nearest neighbour methods* Nearest neighbour methods are closely similar to point-to-object methods, but distances are measured from a random object, not a random point (Diggle 1983). If objects are randomly distributed, the methods are equivalent, whereas if objects are aggregated, distances under this method will be smaller on average. Diggle (1983) summarizes *ad hoc* estimators that improve robustness by combining data from both methods; if the assumption that objects are randomly distributed is violated, biases in the point-to-object and nearest neighbour density estimates tend to be in opposite directions.

1.1.5 The detection function

Central to the concept of distance sampling is the detection function $g(y)$:

$$\begin{aligned} g(y) &= \text{the probability of detecting an object, given that it is at distance} \\ &\quad\ y \text{ from the random line or point} \\ &= \text{prob } \{\text{detection} \mid \text{distance } y\}. \end{aligned}$$

The distance y refers to either the perpendicular distance x for line transects or the sighting (radial) distance r for point transects. Generally, the detection function decreases with increasing distance, but $0 \leq g(y) \leq 1$ always. In the development to follow we usually assume that $g(0) = 1$, i.e. objects on the line or point are seen with certainty (i.e. probability 1). Typical graphs of $g(y)$ are shown in Fig. 1.8. Often, only a small

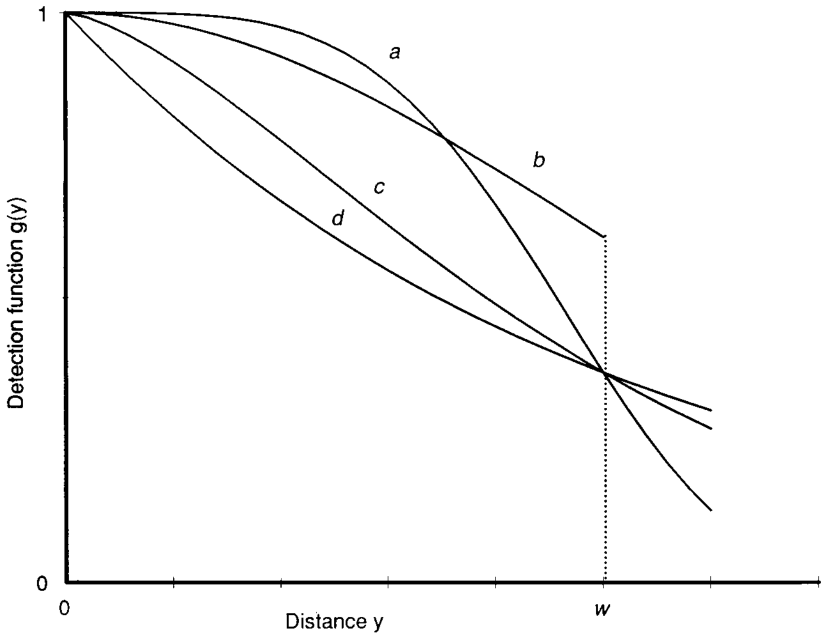

Fig. 1.8. Some examples of the detection function $g(y)$. Function b is truncated at w and thus takes the value zero for all $y > w$. Functions with shapes similar to a, b and c are common in distance sampling. Function d usually results from poor survey design and conduct, and is problematic.

percentage of the objects of interest are detected in field surveys. However, a proper analysis of the associated distances allows reliable estimates of true density to be made. The detection function $g(y)$ could be written as $g(y \mid v)$, where v is the collection of variables other than distance affecting detection, such as object size. We will not use this explicit notation, but it is understood.

1.1.6 Summary

Distance sampling is a class of methods that allow the estimation of density (D = number per unit area) of biological populations. The critical data collected are distances y_i from a randomly placed line or point to objects of interest. A large proportion of the objects may go undetected, but the theory allows accurate estimates of density to be made under mild assumptions. Underlying the theory is the concept of a detection function $g(y) = \text{prob}\ \{\text{detection} \mid \text{distance } y\}$. Detectability usually decreases with increasing distance from the random line or point.

1.2 Range of applications

1.2.1 Objects of interest

Studies of birds represent a major use of both point and line transect studies. Birds are often conspicuous by their bright coloration or distinctive song or call, thus making detection possible even in dense habitats. Surveys in open habitats often use line transects, whereas surveys in more closed habitats with high canopies often use point transects. Distance sampling methods have seen use in studying populations of many species of gamebirds, raptors, passerines and shorebirds.

Many terrestrial mammals have been successfully surveyed using distance sampling methods (e.g. pronghorn, feral pigs, fruit bats, mice, and several species of deer, rabbits, hares, primates and African ungulates). Marine mammals (several species of dolphin, porpoise, seal and whale) have been the subject of many surveys reported in the literature. Reptiles, amphibians, beetles and wolf spiders have all been the subject of distance sampling surveys, and fish (in coral reefs) and red crab densities have been estimated from underwater survey data.

Many inanimate objects have been surveyed using distance sampling, including birds' nests, mammal burrows, and dead deer and pigs. Plant populations and even plant diseases are candidates for density estimation using distance sampling theory. One military application is estimation of the number of mines anchored to the seabed in mine fields.

1.2.2 Method of transect coverage

Distance sampling methods have found use in many situations. Specific applications are still being developed from the existing theory. The versatility of the method is partially due to the variety of ways in which the transect line can be traversed. Historically, line transects were traversed on foot by a trained observer. In recent years, terrestrial studies have used trail bikes, all terrain vehicles, or horses. Transect surveys have been conducted using fixed wing aircraft and helicopters; 'ultralight' aircraft are also appropriate in some instances.

Transect surveys in aquatic environments can be conducted by divers with snorkels or scuba gear, or from surface vessels ranging in size from small boats to large ships, or various aircraft, or by sleds with mounted video units pulled underwater by vessels on the surface. Small submarines may have utility in line or point transect surveys if proper visibility can be achieved. Remote sensing may find extensive use as the technology develops (e.g. acoustic instruments, radar, remotely controlled cameras, multispectral scanners).

In general, the observer can traverse a line transect at a variable speed, travelling more slowly to search heavy cover. The observer may leave the line and walk an irregular path, keeping within w on each side of the line. However, the investigator must ensure that all objects on the line are detected, and that the recorded effort L is the length of the line, not the total distance the observer travels. Point transects are usually surveyed for a fixed time (e.g. 10 minutes per sample point).

1.2.3 Clustered populations

Distance sampling is best explained in terms of 'objects of interest', rather than a particular species of bird or mammal. Objects of interest might be dead deer, birds' nests, jackrabbits, etc. Often, however, interest lies in populations whose members are naturally aggregated into clusters. Here we will take clusters as a generic term to indicate herds of mammals, flocks of birds, coveys of quail, pods of whales, prides of lions, schools of fish, etc. A cluster is a relatively tight aggregation of objects of interest, as opposed to a loosely clumped spatial distribution of objects. More commonly, 'group' is used, but we prefer 'cluster' to avoid confusion with the term 'grouped data', defined below.

Surveying clustered populations differs in a subtle but important way between strip transect sampling and line or point transect sampling. In strip transect sampling, all individuals inside the strip are censused; essentially one ignores the fact that the objects occur in clusters. In contrast, in distance sampling with a fixed w, one records all clusters detected if the centre of the cluster is inside the strip (i.e. 0 to w). If the centre of the cluster is inside the strip, then the count of the size of the cluster must include all individuals in the cluster, even if some individuals are beyond w. On the other hand, if the centre of the cluster is outside the strip, then no observation is recorded, even though some individuals in the cluster are inside the strip.

In distance sampling theory, the clusters must be considered to be the object of interest and distances should be measured from the line or point to the geometric centre of the cluster. Then, estimation of the density of clusters is straightforward. The sample size n is the number of clusters detected during the survey. If a count is also made of the number of individuals (s) in each observed cluster, one can estimate the average cluster size, $E(s)$. The density of individuals D can be computed as a product of the density of clusters D_s times the average cluster size:

$$D = D_s \cdot E(s)$$

A complication arises if detection is a function of cluster size. This relationship is evident if most of the clusters detected at a substantial distance from the line or point are relatively large in size. Typically, the estimator of D_s is still unbiased, but using the mean cluster size $\bar{s}$ to estimate $E(s)$ results in a positive bias in the estimator (because the smaller clusters tend to go undetected toward w).

A well-developed general theory exists for the analysis of distance data from clustered populations. Here the detection probability is dependent on both distance from the line or point and cluster size (this phenomenon is called size-biased sampling). Several approaches are possible: (1) stratify by cluster size and apply the usual methods within each stratum, then sum the estimated densities of individuals; (2) treat cluster size as a covariate and use parametric models for the bivariate distance–cluster size data (Drummer and McDonald 1987); (3) truncate the distance data to reduce the correlation between detection distance and cluster size and then apply robust semiparametric line transect analysis methods; (4) first estimate cluster density, then regress cluster size on $\hat{g}(y)$ to estimate mean cluster size where detection is certain ($\hat{g}(y) = 1$); (5) attempt an analysis by individual object rather than cluster, and use robust inference methods to allow for failure of the assumption of independent detections. Strategy (3) is straightforward and generally quite robust; appropriate data truncation after data collection can greatly reduce the dependence of detection probability on cluster size, and more severe truncation can be used for mean cluster size estimation than for fitting the line transect model, thus reducing the bias in $\bar{s}$ further. We have also found strategy (4) to be effective.

1.3 Types of data

Distance data can be recorded accurately or grouped. Rounding errors in measurements often cause the data to be grouped to some degree, but they must then be analysed as if they had been recorded accurately, or grouped further, in an attempt to reduce the effects of rounding on bias. Distances are often assigned to predetermined distance intervals, and must then be analysed using methods developed for the analysis of frequency data.

1.3.1 Ungrouped data

Two types of ungrouped data can be taken in line transect surveys: perpendicular distances x_i or sighting distances r_i and angles θ_i. If sighting distances and angles are taken, they should be transformed to perpendicular distances for analysis. Only sighting distances r_i are used in the estimation of density in point transects. Trapping webs use the same type of measurement r_i, which is then the distance from the centre of the web to the trap containing animal i. The cue counting method also requires sighting distances r_i, although only those within a sector ahead of the observer are recorded. Angles (0 to 360° from some arbitrary baseline) are potentially useful in testing assumptions in point transects and trapping webs, but have not usually been taken. In cue counting also, angles (sighting angles θ_i) are not usually recorded, except to ensure that they fall between $\pm\,\phi$, where 2ϕ is the sector angle. In all cases we will assume that n distances $\{y_1, y_2, \ldots . y_n\}$ are measured corresponding to the n detected objects. Of course, n itself is usually a random variable, although one could design a survey in which searching continues until a pre-specified number of objects n is detected; L is then random and the theory is modified slightly (Rao 1984).

Sample size n should generally be at least 60–80, although for some purposes, as few as 40 might be adequate. Formulae are available to determine the sample size that one expects to achieve with a given level of precision (measured, for example, by the coefficient of variation). A pilot survey is valuable in predicting sample sizes required, and will usually show that a sample as small as 40 for an entire study is unlikely to achieve the desired precision.

1.3.2 Grouped data

Data grouping arises in distance sampling in two ways. First, ungrouped data y_i, $i = 1, \ldots, n$, may be taken in the field, but analysed after deliberate grouping into frequency counts n_i, $i = 1, \ldots, u$, where u is the number of groups. Such grouping into distance intervals is often done to achieve robustness in the analysis of data showing systematic errors such as heaping (i.e. rounding errors). Grouping the r_i and θ_i data by intervals in the field or for analysis in line transect surveys is not recommended because it complicates calculation of perpendicular distances, although techniques (e.g. 'smearing') exist to handle such grouped data.

Second, the data might be taken in the field only by distance categories or intervals. For example, in aerial surveys it may only be practical to count the number of objects detected in the following distance intervals: 0–20, 20–50, 50–100, 100–200, and 200–500 m. Thus, the exact distance of an object detected anywhere from 0 to 20 m from the line or point would not be recorded, but only that the object was in the first distance category. The resulting data are a set of frequency counts n_i by specified distance categories rather than the set of exact distances, and total sample size is equal to $n = \Sigma\, n_i$.

Distance categories are defined by boundaries called cutpoints c_i. For u such boundaries, one has cutpoints $0 < c_1 < c_2 < \ldots < c_u$. By convention, let $c_0 = 0$ and $c_u = w$, where w can be finite or infinite (i.e. unbounded). Typically in line transect sampling the intervals defined by the cutpoints will be wider near w and narrower near the centreline. However, in point transect sampling, the first interval may be quite wide because the area corresponding to it is relatively small. The sum of the counts in each distance category equals the total number of detections n, which is the sample size. In the example above, $u = 5$, and the cutpoints are $0 < c_1 = 20 < c_2 = 50 < c_3 = 100 < c_4 = 200 < c_5 = w = 500$. Suppose the frequency counts n_i are 80, 72, 60, 45 and 25, respectively. Then $n = \Sigma\, n_i = 282$ detections.

1.3.3 Data truncation

In designing a line transect survey, one can establish a distance $y = w$ whereby objects at distances greater than w are ignored. In this case, the width of the transect to be searched is $2w$, and the area searched is of size $2wL$. In point transects, a radius w can similarly be established, giving the total area searched as $k\pi w^2$. In the general theory, w may be assumed to be infinite so that objects may be detected at quite large distances. In such cases, the width of the transect or radius around the point is unbounded.

Distance data can be truncated (i.e. discarded) prior to analysis. Data can be truncated beyond some distance w to delete outliers that make modelling of the detection function $g(y)$ difficult (Fig. 1.9). For example, w might be chosen such that $\hat{g}(w) = 0.15$. Such a rule might eliminate many detections in some point transect surveys, but only relatively few detections in line transect surveys. A simpler rule might be to truncate 5–10% of the objects detected at the largest distances. If data are truncated in the field, further truncation may be carried out at the analysis stage if this seems useful.

General methodology is available for 'left-truncation' (Alldredge and Gates 1985). This theory is potentially useful in aerial surveys if visibility directly below the aircraft is limited and, thus, $g(0) < 1$. Quang and Lanctot (1991) provide an alternative solution to this problem. Selection of a model for the distance data is critical under left-truncation because estimation may be very model dependent. Other alternatives exist at the survey design stage and we hesitate to recommend left-truncation except in special circumstances, such as the case where there is evidence of a wide shoulder in the detection function.

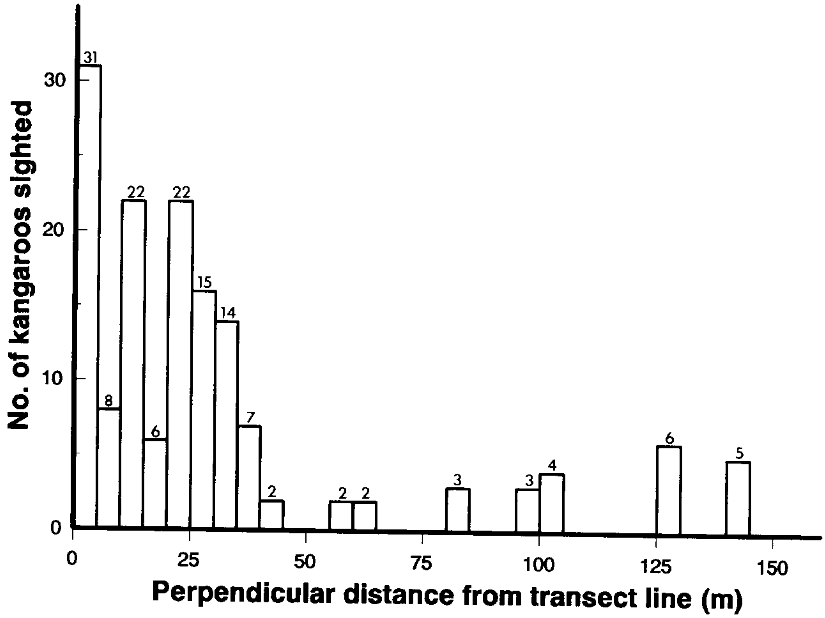

Fig. 1.9. Histogram of the number of eastern grey kangaroos detected as a function of distance from a line transect survey on Rotamah Island, Australia (redrawn from Coulson and Raines 1985). These data illustrate some heaping in the first, third and fifth distance classes, and the need to truncate observations beyond about 50 m.

1.3.4 Units of measurement

The derivation of the theory assumes that the units of y_i, L and D are all on the same measurement scale. Thus, if the distances y_i are measured in metres, then L should be in metres and density will be in numbers per square metre. In practice it is a simple but important matter to convert the y_i, l_i or D from any unit of measure into any other; in fact, computer software facilitates such conversions (e.g. feet to metres or acres to square kilometres or numbers/m^2 to numbers/km^2).

1.3.5 Ancillary data

In some cases, there is interest in age or sex ratios of animals detected, in which case these ancillary data must be recorded. Cluster size is a type of ancillary data. Size of the animal, its reproductive state (e.g.

kangaroos carrying young in the pouch), or presence of a marker or radio transmitter are other examples of ancillary data collected during a survey. Such ancillary information can be incorporated in a variety of ways. If probability of detection is a function of the ancillary variable, then it might be used to stratify the data, or it might enter the analysis as a covariate, to improve estimation.

1.4 Known constants and parameters

1.4.1 Known constants

Several known constants are used in this book and their notation is given below:

A = area occupied by the population of interest;

k = number of lines or points surveyed;

l_i = length of the ith transect line, $i = 1, \ldots, k$;

L = total line length $= \Sigma\, l_i$;

and w = the width of the area searched on each side of the line transect, or the radius searched around a point transect, or the truncation point beyond which data are not used in the analysis.

1.4.2 Parameters

In line and point transect surveys there are only a few unknown parameters of interest. These are defined below:

D = density (number per unit area);

N = population size in the study area;

$E(s)$ = mean cluster size in the population (not the same as, but often estimated by, the sample mean $\bar{s}$ of detected objects);

$f(0)$ = the probability density function of distances from the line, evaluated at zero distance;

$h(0)$ = the slope of the probability density function of distances from the point, evaluated at zero distance;

and $g(0)$ = probability of detection on the line or point, usually assumed to be 1. For some applications (e.g. species of whale which spend substantial periods underwater and thus avoid detection, even on the line or point), this parameter must be estimated from other types of information.

Density D may be used in preference to population size N in cases where the size of the area is not well defined. Often an encounter rate n/L is computed as an index for sample size considerations or even as a crude relative density index.

1.5 Assumptions

Statistical inference in distance sampling rests on the validity of several assumptions. First, the survey must be competently designed and conducted. No analysis or inference theory can make up for fundamental flaws in survey procedure. Second, the physical setting is idealized:

1. Objects are spatially distributed in the area to be sampled according to some stochastic process with rate parameter D (= number per unit area).
2. Randomly placed lines or points are surveyed and a sample of n objects is detected, measured and recorded.

It is not necessary that the objects be randomly (i.e. Poisson) distributed. Rather, it is critical that the line or point be placed randomly with respect to the distribution of objects. Random line or point placement ensures a representative sample of the relevant distances and hence a valid density estimate. The use of transects along trails or roads does not constitute a random sample and represents poor survey practice. In practice, a systematic grid of lines or points, randomly placed in the study area, suffices.

Three assumptions are essential for reliable estimation of density from line or point transect sampling. These assumptions are given in order from most to least critical:

1. Objects directly on the line or point are always detected (i.e. they are detected with probability 1, or $g(0) = 1$).
2. Objects are detected at their initial location, prior to any movement in response to the observer.
3. Distances (and angles where relevant) are measured accurately (ungrouped data) or objects are correctly counted in the proper distance category (grouped data).

Some investigators include the assumption that one must be able to identify the object of interest correctly. In rich communities of songbirds, this problem is often substantial. Marine mammals often occur in mixed schools, so it is necessary both to identify all species present and to count the number of each species separately. In rigorous theoretical developments, assumption (2) is taken to be that objects are immobile. However, slow movement relative to the speed of the observer causes few problems in line transects. In contrast, responsive movement of animals to the approaching observer can create serious problems. In point transects, undetected movement of animals is always problematic because the observer is stationary.

The effects of partial failure of these assumptions will be covered at length in later sections, including the condition $g(0) < 1$; estimation in this circumstance is one of the main areas of current methodological development. All of these assumptions can be relaxed under certain circumstances. These extensions are covered in the following chapters. We note that no assumption is made regarding symmetry of $g(y)$ on the two sides of the line or around the point, although extreme asymmetry would be problematic. Generally, we believe that asymmetry near the line or point will seldom be large, although topography may sometimes cause difficulty. If data are pooled to give a reasonable sample size, such problems can probably be ignored.

1.6 Fundamental concept

It may seem counterintuitive that a survey be conducted, fail to detect perhaps 60–90% of the objects of interest in the survey plots (strips of dimension L by $2w$ or circular plots of size πw^2), and still obtain accurate estimates of population density. The following two sections provide insights into how distances are the key to the estimation of density when some of the objects remain undetected. We will illustrate the intuitive ideas for the case of line transect sampling; those for point transects are similar.

Consider an arbitrary area of size A with objects of interest distributed according to some random process. Assume a randomly placed line and grouped data taken in each of eight 1-foot distance intervals from the line on either side, so that $w = 8$. If all objects were detected, we would expect, on average, a histogram of the observations to be uniform as in Fig. 1.10a. In other words, on average, one would not expect many more or fewer observations to fall, say, within the seventh interval than the first interval, or any other interval.

In contrast, distance data from a survey of duck (*Anas* and *Aythya* spp.) nests at the Monte Vista National Wildlife Refuge in Colorado,

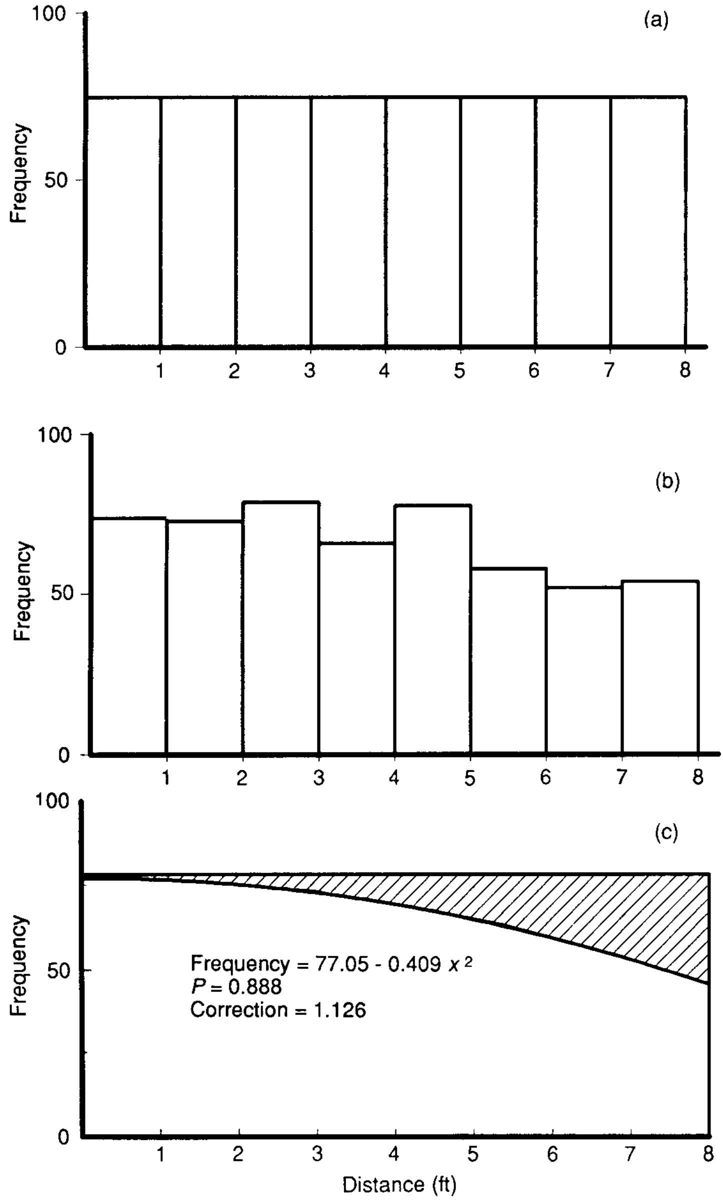

Fig. 1.10. Conceptual basis for line transect sampling: (a) the expected number of objects detected in eight distance classes if no objects were left undetected; (b) real data where a tendency to detect fewer objects at greater distances from the line can be noticed; (c) simple methods can be used to estimate the proportion of the objects left undetected (shaded area). The proportion detected, P, can be estimated from the distance data.

USA (Anderson and Pospahala 1970) are shown in Fig. 1.10b as a histogram. Approximately 10 000 acres of the refuge were surveyed using $L = 1600$ miles of transect, and an area $w = 8$ feet on each side of the transect was searched. A total of 534 nests was found during 1967 and 1968 and the distance data were grouped for analysis into 1-foot intervals. Clearly there is evidence from this large survey that some nests went undetected in the outer three feet of the area searched. Visual inspection might suggest that about 10% of the nests were missed during the survey. Note that the intuitive evidence that nests were missed is contained in the distances, here plotted as a histogram.

Examination of such a histogram suggests that a 'correction factor', based on the distance data, is needed to correct for undetected objects. Note that such a correction factor would be impossible if the distances (or some other ancillary information) were not recorded. Anderson and Pospahala (1970) fitted a simple quadratic equation to the midpoints of each histogram class to obtain an objective estimate of the number of nests not detected (Fig. 1.10c). Their equation, fitted by least squares, was

$$\text{frequency} = 77.05 - 0.4039x^2$$

The proportion (P) of nests detected was computed as the unshaded area in Fig. 1.10c divided by the total area (shaded + unshaded). (The areas were computed using calculus, but several simpler approximations could be used.) The estimated proportion of nests detected from 0 to 8 feet can be computed to be 0.888, suggesting a correction factor of 1.126 (= 1/0.888) be applied to the total count of $n = 534$. Thus, the estimated number of nests within eight feet of the sample transects was $n/P = 601$, and because the transects sampled 5.5% of the refuge, the estimate of the total number of nests on the refuge during the 2-year period was $601/0.055 = 10\,927$. This procedure provides the intuition that distances are important in reliable density estimates even if most of the objects are not detected. The Anderson–Pospahala method is no longer recommended since superior analysis methods are now available, but it illustrates the principle underlying the theory. The next two chapters will put this intuitive argument on a more formal basis.

1.7 Detection

When a survey has been conducted, n objects will have been detected. Considerable confusion regarding the meaning of n exists in the literature. Here an attempt is made to factor n into its fundamental components.

Burnham *et al.* (1981), Dawson (1981) and Morgan (1986: 9–25) give a discussion of these issues. Ramsey (1981) provides an informative example.

The number detected, n, is a confounding of true density and probability of detection. The latter is a function of many factors, including cue production by, or characteristics of, the object of interest, observer effectiveness, and the environment. Of these factors, one could hope that only the first, density, influences the count. While this might be ideal, it is rarely true.

1.7.1 Cue production

The object of interest often provides cues that lead to its detection by the survey observer. Obvious cues may be a loud or distinctive song or call. A splash made by a marine mammal or flock of sea birds above a school of dolphins are other examples of cues. Large size, bright or contrasting colouring, movement or other behaviour may be causes for detection. These cues are frequently species-specific and may vary by age or sex of the animal, time of day, or season of the year. Thus, the total count n can vary for reasons unrelated to density (Mayfield 1981; Richards 1981; Bollinger *et al.* 1988). Most often, the probability of detection of objects based on some cue diminishes as distance from the observer increases.

1.7.2 Observer effectiveness

Observer variability is well known in the literature on biological surveys. Interest in the survey, training and experience are among the dominant reasons why observers vary widely in their ability to detect objects of interest. However, both vision and hearing acuity may be major variables which are often age-specific (Ramsey and Scott 1981a; Scott, *et al.* 1981). Fatigue is a factor on long or difficult surveys. Even differing heights of observers may be important for surveys carried out on foot, with tall observers detecting objects at a higher rate. Generally, the detection of objects decreases with increasing distance due to observer effectiveness.

1.7.3 Environment

Environmental variables often influence the number of objects detected (Best 1981; Ralph 1981; Verner 1985). The habitat type and its phenology are clearly important (Bibby and Buckland 1987). Physical conditions often inhibit detection: wind, precipitation, darkness, sun angle, etc. Cue production varies by time of day, which can have a tenfold effect in the detectability of some avian species (Robbins 1981; Skirvin 1981). Often, these variables interact to cause further variability in detection and the count n.

Distance sampling provides a general and comprehensive approach to the estimation of population density. The distances y_i allow reliable estimates of density in the face of variability in detection due to factors such as cue production, observer effectiveness and environmental differences. The specific reasons why an object was not detected are unimportant. Furthermore, it seems unnecessary to research the influence of these environmental variables or to standardize survey protocol for them, if distances are taken properly and appropriate analysis carried out. Distance sampling methods fully allow for the fact that many objects will remain undetected, as long as they are not on the line or point. For example, in Laake's stake surveys (Burnham *et al.* 1980) only 27–67% of the stakes present were detected and recorded by various surveyors traversing the line. Still, accurate estimates of stake density were made using distance sampling theory.

1.8 History of methods

1.8.1 Line transects

In the 1930s, R.T. King recognized that not all animals were seen on strip transect surveys and presumably tried to estimate an **effective** width of the transect. He recognized that distances were useful and used the average sighting distance $\bar{r}$ as the effective width surveyed (Leopold 1933; Gates 1979). The early literature tried to conceptualize the idea of effective area sampled. Finally, Gates (1979) provided a formal definition for the effective strip width (μ): the distance for which unseen animals located closer to the line than μ equals the number of animals seen at distances greater than μ. Then, $D = n/A'$, where $A' = 2\mu L$ and is the estimated area 'effectively' sampled. Note that μ is actually one-half the effective strip width, i.e. only one side of the line.

Kelker (1945) took an alternative approach that is still sometimes used. Instead of trying to retain the total sample of n distances and estimate the 'area' effectively sampled, Kelker determined a strip width Δ on each side of the transect centreline, within which all animals were probably seen. The value of Δ was judged subjectively from an inspection of the histogram of the perpendicular distance data. Once Δ was chosen, density was estimated as a strip transect with $W = \Delta$ and n the number of objects detected from 0 to Δ on each side of the line transect. Distance data exceeding Δ were not used further.

No attempts were made to formulate a firm conceptual and mathematical foundation for line transects until Hayne's paper in 1949. All estimators then in use were *ad hoc* and generally based on either the concept of the effective strip width or the related idea of determining a strip width narrow enough such that no animals were undetected in that strip. Variations of these approaches are still being used and sometimes 'rediscovered' today, even though better methods have existed for many years.

Hayne (1949) provided the first estimator that has a rigorous justification in statistical theory. While Hayne's method rests on only the use of sighting distances r_i, the critical assumption made can only be tested using the sighting angles θ_i. Hayne's (1949) method is poor if $\bar{\theta}$ is not approximately 32.7° and may not perform well even if $\bar{\theta}$ falls close to this value, i.e. not a robust method.

After Hayne's (1949) paper, almost no significant theoretical advances appeared until 1968. During that 20 year period, line transect sampling was used frequently, and on a variety of species. The assumptions behind the method were sharpened in the wildlife literature and some evaluations of the method were presented (e.g. Robinette *et al.* 1956).

In 1968, two important papers were published in which some of the basic ideas and conceptual approaches to line transect sampling finally appeared (Eberhardt 1968; Gates *et al.* 1968). Gates *et al.* (1968) published the first truly rigorous statistical development of a line transect estimator, applicable only to untruncated and ungrouped perpendicular distance data. They proposed that $f(x)$ be a negative exponential form, $f(x) = a \cdot \exp(-ax)$, where a is an unknown parameter to be estimated. Under that model, $f(0) = a$. Gates *et al.* (1968) developed the optimal estimator of a based on a sample of perpendicular distances and provided an estimator of the sampling variance. For the first time, rigorous consideration was given to questions such as optimal estimation under the model, construction of confidence intervals, and tests of assumptions. The one weakness was that because the assumed detection function was very restrictive and might easily be inappropriate, the resulting estimate of density could be severely biased.

In contrast, Eberhardt (1968) conceptualized a fairly general model in which the probabilities of detection decreased with increasing perpendicular distance. He reflected on the shape of the detection function $g(x)$, and suggested both that there was a lack of information about the appropriate shape and that the shape might change from survey to survey. Consequently, he suggested that the appropriate approach would be to adopt a family of curves to model $g(x)$. He suggested two such families, a power series and a modified logistic, both of which are fairly flexible parametric functions. His statistical development of these models was limited, but important considerations had been advanced.

Since 1968, line transect sampling has been developed along rigorous statistical inference principles. Parametric approaches to modelling $g(x)$ were predominant, with the notable exception of Anderson and Pospahala (1970), who rather inadvertently introduced some of the basic ideas that underlie a non-parametric or semiparametric approach to the analysis of line transect data. Emlen (1971) proposed an *ad hoc* method that found use in avian studies.

A general model structure for line transect sampling based on perpendicular distances was presented by Seber (1973: 28–30). For an arbitrary detection function, Seber gave the probability distribution of the distances $x_1, \ldots, x_n$ and the general form of the estimator of animal density D. This development was left at the conceptual stage and not pursued to the final step of a workable general approach for deriving line transect estimators, and the approach was still based on the concept of an effective strip width.

More work on sighting distance estimators appeared (Gates 1969; Overton and Davis 1969). There was a tendency to think of approaches based on perpendicular distances as appropriate for inanimate or non-responsive objects, whereas methods for flushing animals were to be based on sighting distances and angles (Eberhardt 1968, 1978a). This artificial distinction tended to prevent the development of a unified theory for line transect sampling. By the mid-1970s, line transect sampling remained a relatively unexplored methodology for the estimation of animal density. Robinette *et al.* (1974) reported on a series of field evaluations of various line transect methods. Their field results were influential in the development of the general theory.

Burnham and Anderson (1976) pursued the general formulation of line transect sampling and gave a basis for the general construction of line transect estimators. They developed the general result $\hat{D} = n \cdot \hat{f}(0)/2L$, wherein the parameter $f(0)$ is a well-defined function of the distance data. The key problem of line transect data analysis was seen to be the modelling of $g(x)$ or $f(x)$ and the subsequent estimation of $f(0)$. The nature of the specific data (grouped or ungrouped, truncated or untruncated) is irrelevant to the basic estimation problem. Consequently, their formulation is applicable for the development of any parametric or semiparametric line transect estimator. Further, the general theory is applicable to point transect sampling with some modification (Buckland 1987a).

Burnham and Anderson's (1976) paper heralded a period of new statistical theory. Major contributions published during the 1976–80 period include Schweder (1977), Crain *et al.* (1978, 1979), Pollock (1978), Patil *et al.* (1979b), Quinn (1979), Ramsey (1979), Seber (1979) and Quinn and Gallucci (1980). Other papers developing methodology during this short period include Anderson *et al.* (1978, 1979, 1980), Eberhardt (1978a, b, 1979), Sen *et al.* (1978), Burnham *et al.* (1979), Patil *et al.* (1979a) and Smith (1979). Anderson *et al.* (1979) provided guidelines for field sampling, including practical considerations. Burdick (1979) produced an advanced method to estimate spatial patterns of abundance from line transect sampling where there are major gradients in population density. Laake *et al.* (1979) and Gates (1980) produced comprehensive computer software packages, TRANSECT and LINETRAN respectively, for the analysis of line transect data.

Gates (1979) provided a readable summary of line transect sampling theory and Ramsey's (1979) paper presents a more mathematical treatment of parametric approaches. Hayes (1977) gave an excellent summary of methodology and provided many useful insights at that time.

Burnham *et al.* (1980) published a major monograph on line transect sampling theory and application. This work provided a review of previous methods, gave guidelines for field use, and identified a small class of estimators that seemed generally useful. Usefulness was based on four criteria: **model robustness, pooling robustness**, a **shape criterion**, and **estimator efficiency**. Theoretical and Monte Carlo studies led them to suggest the use of estimators based on the Fourier series (Crain *et al.* 1978, 1979), the exponential power series (Pollock 1978), and the exponential quadratic model.

Since 1980, more theory has been developed on a wide variety of issues. Seber (1986) and Ramsey *et al.* (1988) give brief reviews. Major contributions during the 1980s include Butterworth (1982a, b), Patil *et al.* (1982), Hayes and Buckland (1983), Buckland (1985), Burnham *et al.* (1985), Johnson and Routledge (1985), Quinn (1985), Drummer and McDonald (1987), Ramsey *et al.* (1987), Thompson and Ramsey (1987) and Zahl (1989). Other papers during the decade include Buckland (1982), Stoyan (1982), Burnham and Anderson (1984), Anderson *et al.* (1985a, b) and Gates *et al.* (1985). Several interesting field evaluations where density was known have appeared since 1980, including Burnham *et al.* (1981), Hone (1986, 1988), White *et al.* (1989), Bergstedt and Anderson (1990) and Otto and Pollock (1990). In addition, other field evaluations where the true density was not known have been published, but these results are difficult to interpret.

A great deal of statistical theory has been developed since 1976, but new theory may have started to decrease by the late 1980s. Field studies using line transect sampling have increased and new applications have appeared in the literature. No attempt to discuss all of the recent developments will be given in this chapter. At the present time, there are several good models for fitting $g(x)$. There now exist sound approaches for analysing grouped or ungrouped data with truncated or untruncated transect widths, under various extensions (e.g. clustered populations). Estimation based on sighting distances and angles has been shown to be problematic and we recommend transforming such data to perpendicular distances prior to analysis. Current areas of development include estimation when $g(0) < 1$, when responsive movement to the observer occurs, when the objects occur in clusters, leading to size-biased sampling, and when there is covariate information on factors such as sighting conditions or habitat.

1.8.2 Point transects

Point transect sampling has had a much shorter history. The method can be traced to the paper by Wiens and Nussbaum (1975) and their application of what they called a variable circular plot census. They drew heavily on the paper on line transects by Emlen (1971). Ramsey and Scott (1979) provided a statistical formalism for the general method and noted several close relationships to line transect sampling. Following the 'effective area' thinking, they noted 'The methods are similar in spirit to line transect methods, in that the total number of detections divided by an estimate of the area surveyed is the estimate of the population density.' Ramsey and Scott (1979) provided a summary of the assumptions and derived a general theory for density estimation, including sampling variances. This represented a landmark paper at the time.

Reynolds *et al.* (1980) presented additional information on the variable circular plot method. Burnham *et al.* (1980) and Buckland (1987a) also noted the close links between line transects and point transects (i.e. variable circular plots). Buckland (1987a) developed other models, evaluated the Fourier series, Hermite polynomial and hazard-rate estimators, and provided an evaluation of the efficiency of binomial models (where objects of interest are grouped into two categories, within or beyond a specified distance c_1). The general theory for line and point transects is somewhat similar because they both involve sampling distances. Thus, the term point transect will be used rather than the variable circular plot 'census'.

1.9 Program DISTANCE

The computation for most estimators is arduous and prone to errors if done by hand. Estimators of sampling variances and covariances are similarly tedious. Data should be plotted and estimates of $f(y)$ should usually be graphed for visual comparison with the observed distance data.

Program DISTANCE (Laake *et al.* 1993) was developed to allow comprehensive analyses of the type of distance data we discuss here. The program is written in FORTRAN and runs on any IBM PC compatible microcomputer with 640 K of RAM. A math coprocessor is desirable, but not required. Program DISTANCE allows researchers to focus on the biology of the population, its habitat and the survey operation; one can concentrate on the results and interpretation, rather than on computational details. Almost all the examples presented in this book were analysed using program DISTANCE; the distance data and associated program commands for some of the examples are available as an aid to data analysts. The program is useful both for data analysis and as a research tool. Only occasional references to DISTANCE are made throughout this book because a comprehensive manual on the program is available (Laake *et al.* 1993).