Studies in Avian Biology No. 6:441–444, 1981.

DENSITY ESTIMATION USING LINE TRANSECT SAMPLING

P. V. Rao,[1] K. M. Portier,[1] and J. A. Ondrasik[2]

Abstract.—Line transect methods of estimating population density usually assume a fixed transect length. As a means of reducing the distance travelled by the observer, Rao and Ondrasik (1980) proposed a method based on a line transect of random length. In this method a length, L_0, and number, N_0, are fixed in advance and sampling is terminated as soon as either a distance L_0 is traversed or a number, N_0, of objects is sighted. A brief summary of the method including the density estimate and its variance is presented in the first part of the paper.

In the second part, a method of estimation of density for clustered populations is discussed. This method assumes (1) that the probability of sighting a cluster is a function of its size as well as its perpendicular distance from the transect line and (2) that not all objects in a cluster may be sighted by the observer. The estimate of the population density as well as estimates of other model parameters are obtained using maximum likelihood. The method is illustrated using artificially constructed data for a clustered population.

The use of line transect methods in estimating animal and plant population densities has received considerable attention in recent literature. Excellent reviews of the general subject area are found in Seber (1973), Eberhardt (1968) and Burnham et al. (1980).

With the few exceptions noted by Burnham et al. (1980 Appendix D), currently available density estimates from line transect methods use transects of fixed length and assume that sightings of objects are independent events. An obvious drawback of a sampling method based on a predetermined transect length is the possibility that it may be using an unnecessarily long transect to estimate density. Because cost of sampling is likely to increase with the length of the transect, it is desirable to consider estimation procedures based on random transect lengths, i.e., procedures which terminate as soon as a predetermined number of objects is sighted. Another drawback of most of the available line transect methods is due to the fact that many biological populations (e.g., coveys of quail, schools of porpoise and so on) aggregate into tight clusters. The assumption of independence is not reasonable for such populations, so density estimation procedures must account for the facts (1) that objects are sighted in groups and (2) that all objects in a group may not be seen by the observer.

The purpose of this paper is (1) to describe a recently developed method (Rao and Ondrasik 1980) that allows for the termination of sampling after a prespecified number of observations is made, and (2) to propose a model suitable for line transect sampling of clustered populations.

[1] Department of Statistics, University of Florida, Gainesville, Florida 32611.

[2] Boehringer Ingelheim Ltd., P. O. Box 368, Ridgefield, Connecticut 06877.

SAMPLING WITH RANDOM TRANSECT LENGTHS

A sampling plan which utilizes a predetermined length for the line transect will be referred to as a direct sampling plan. Many direct sampling density estimates use only the right angle distances of the objects and are based on the set of assumptions listed below (Seber 1973).

- A1. Objects are randomly and independently distributed over the area of interest at a rate (density) D objects per unit area.
- A2. Sightings of objects are independent events.
- A3. Objects are fixed, i.e., objects are immotile and no object is counted twice.
- A4. There exists a function $g(y)$ which is the probability of observing an object conditional on the existence of an object at right angle distance y from the transect.
- A5. $g(0) = 1$; i.e., objects on the transect line are observed with probability one.

Inverse Sampling

In contrast to the direct sampling plan, one can consider an inverse sampling plan. In an inverse sampling plan, observation is terminated as soon as a prespecified number, N_0, of objects is sighted by the observer.

Rao and Ondrasik (1980) developed an estimation procedure suitable for the inverse sampling plan. Following Burnham and Anderson (1976), they assume that the conditional probability density, $f(y)$, of the perpendicular distance y is unknown. Utilizing assumptions A1 to A5 they estimate the population density to be

$$\hat{D}_1 = \frac{(N_0 - 1)\hat{f}(0)}{2l} \qquad (1)$$

where l is the actual distance traversed and $\hat{f}(0)$ is an estimate of $f(0)$. Assuming the bias in $\hat{f}(0)$

to be relatively small, an approximation to the variance of $\hat{D}_1$ is given as

$$V(\hat{D}_1) = \frac{D^2}{N_0 - 1}[1 + \mathrm{C}\cdot\mathrm{V}(\hat{f}(0))], \qquad (2)$$

where $\mathrm{C}\cdot\mathrm{V}(\hat{f}(0))$ is the coefficient of variation of $\hat{f}(0)$. While any reasonable estimate of $f(0)$ can be used in (1), the Fourier series estimator suggested by Crain et al. (1978) appears to be the most desirable. The monograph by Burnham et al. (1980) contains many examples of calculation of the Fourier series estimate for $f(0)$.

Combined Sampling

A disadvantage of the inverse sampling plan is the possibility that sampling may not terminate in a reasonable period of time. To overcome this drawback, Rao and Ondrasik (1980) developed the combined sampling plan in which sampling stops as soon as either a prescribed number, N_0, of objects is sighted or a prespecified length, L_0, of the transects is traversed. If n and l denote, respectively, the actual number of objects sighted and the actual distance traversed, then the combined sampling estimate of D has the form

$$\hat{D}_2 = \begin{cases} 0 & 0 \leq n \leq 1 \\ \dfrac{n\hat{f}(0)}{2L_0} & 1 < n < N_0 \\ \dfrac{(N_0 - 1)\hat{f}(0)}{2l} & n = N_0 \end{cases} \qquad (3)$$

An approximation to the variance of $\hat{D}_2$ is

$$V(\hat{D}_2) = \frac{D^2}{N_0 - 2}\Bigg\{1 + \mathrm{C}\cdot\mathrm{V}(\hat{f}(0)) - e^{-n}\mathrm{C}\cdot\mathrm{V}(\hat{f}(0))\left[\sum_{j=0}^{N_0-3}\left(\frac{j(N_0 - 2) - (j + 1)^2}{(j + 1)^2 j!}\right)n^j - (N_0 - 2)(2 - e^{-n})\right]\Bigg\} \qquad (4)$$

SAMPLING CLUSTERED POPULATIONS

Anderson et al. (1976) note that density estimates for clustered populations can easily be obtained if assumptions A1 to A5 hold for clusters of objects rather than for individuals. It is clear in this case that existing methods of density estimation are directly applicable to the estimation of cluster density. If the number of objects in every sighted cluster can be determined without error, then an estimate for the density of objects is

$$\hat{\Delta} = \hat{D}\bar{s} \qquad (5)$$

where $\hat{D}$ is an estimate of the cluster density and $\bar{s}$ is the average size of the observed clusters.

There are two reasons why the assumptions implied by the procedure suggested in the preceeding paragraph may not be reasonable when developing sampling models for clustered populations. First, the simple modification obtained by replacing the word "object" by the word "cluster" in A1 to A5 would imply that the probability of sighting a cluster depends only on its right angle distance. This may not be reasonable because the probability of sighting a larger cluster is likely to be greater than the probability of sighting a smaller cluster located at the same distance. Second, the sighting of a cluster may not necessarily mean that all of the objects comprising the cluster are seen and counted by the observer. A more reasonable assumption would be to let the probability of sighting an object belonging to a cluster depend on the distance to the cluster as well as the true cluster size.

Burnham et al. (1980, Appendix D) suggest a set of assumptions which imply that the *probability of sighting* a cluster depends on its size and distance. The set of assumptions listed below implicitly contains the assumption of Burnham et al. (1980) but also implies that the *number of objects seen* in a cluster depends on its (cluster) size and distance.

Assumptions

B1. The clusters are randomly and independently distributed over the area of interest at a rate (density) of D clusters per unit area.

B2. Sightings of clusters are independent events.

B3. Clusters are fixed, *i.e.*, clusters are immotile and no cluster is counted twice.

B4. The probability that a randomly chosen cluster is of size r is $p(r)$, $r = 1, 2, \ldots$

B5. There exists a non increasing function $h(y)$, with $h(0) = 1$ and $0 \leq h(y) \leq 1$, such that the probability of sighting s objects in a cluster, conditional on a cluster of size $r \geq s$ being located at right angle distance y is

$$p(s\,|\,r, y) = \binom{r}{s}[h(v)]^s[1 - h(y)]^{r-s} \qquad s = 0, 1, \ldots, r \qquad (6)$$

An inspection of the assumptions B1 to B5 shows that B1, B2, and B3 are directly obtained from A1, A2, and A3. From assumption B5, the probability of sighting a cluster conditional on a cluster of size r being located at distance y is seen to be

$$1 - p(0|r,y) = 1 - [1 - h(y)]^r, \qquad (7)$$

which clearly depends on y and r. In particular,

TABLE 1
CONSTRUCTED DATA FOR INVERSE SAMPLING OF CLUSTERED POPULATION ($N_0 = 25$, $l = 25$ KM)

Perpendicular distance (y_i) (meters)	Observed cluster size (s_i)
1	1
3	2
7	1
10	1
2	3
5	5
4	1
7	2
15	1
22	1
6	1
3	6
2	1
12	1
28	3
9	2
18	1
36	7
17	6
5	1
4	1
3	1
8	2
3	4
13	1

the probability of sighting a cluster of size r at $y = 0$ is $1 - [1 - h(0)]^r = 1$.

The form of the cluster detection function is easily derived. Let $g(y)$ denote the probability of sighting a cluster conditional on the right angle distance y. Then

$$g(y) = \sum_{r=1}^{\infty} (1 - [1 - h(y)]^r)p(r)$$
$$= 1 - \sum_{r=1}^{\infty} [1 - h(y)]^r p(r) \quad (8)$$

If every cluster in the population has size 1, then the probability distribution of cluster size satisfies

$$p(1) = 1,$$

and

$$g(y) = 1 - [1 - h(y)]\,p(1) = h(y)$$

Thus $h(y)$ may be regarded as the probability of detecting a single object at distance y.

Under assumption B1 (see Seber 1973), the expected number of clusters seen in a transect of length l is θl, where,

$$\theta = 2cD \quad (9)$$

TABLE 2
MAXIMUM LIKELIHOOD ESTIMATES OF MODEL PARAMETERS

Parameter	Estimate	Standard error
Δ[a]	58.6 per km^2	21.0 per km^2
θ	1.00 per km	.04 per km
α	.709	.071
λ	59.3 per km	11.5 per km
D[a]	17.04 per km^2	4.32 per km^2
c[b]	.029	.005

[a] Estimate calculated using $\hat{\alpha}$, $\hat{\gamma}$ and $\hat{\theta}$.
[b] Estimate calculated using $\hat{\alpha}$ and $\hat{\gamma}$.

is the expected number of cluster sightings per unit length of the transect and

$$c = \int_0^{\infty} g(y)\,dy. \quad (10)$$

ESTIMATION OF DENSITY

Maximum likelihood estimation of the density of objects, Δ, is possible when $p(r)$ and $h(y)$ are completely specified. Clearly, the appropriate form of the likelihood function will depend upon the sampling plan. For example, suppose the sampling plan calls for sampling until N_0 clusters are sighted. If $(s_1, y_1), (s_2, y_2), \ldots, (s_{N_0}, y_{N_0})$ denote the sizes and right angle distances and l denotes the actual length of the transect traversed, then the likelihood function of the sample can be shown to have the form

$$L = \left(\frac{\theta^{N_0} l^{N_0 - 1} \exp(-\theta l)}{c^{N_0}(N_0 - 1)!}\right) \prod_{i=1}^{N_0} p(s_i \mid y_i) g(y_i), \quad (11)$$

where $p(s \mid y)$ is the conditional probability of sighting s objects at distance y:

$$p(s \mid y) = \sum_{r=1}^{\infty} p(s \mid r, y)p(r)$$
$$= \sum_{r=s}^{\infty} \binom{r}{s} [h(y)]^s [1 - h(y)]^{r-s} p(r) \quad (12)$$

Note that, in addition to θ and c, the likelihood function will contain parameters appearing in the specification of $p(r)$ and $h(y)$. The joint likelihood will have to be maximized using an appropriate iterative procedure.

EXAMPLE

Since real data to which the likelihood given by Eq. (11) is appropriate are not readily available in the literature, an artificially constructed data set will be used to illustrate the maximum likelihood estimation procedure. Suppose that

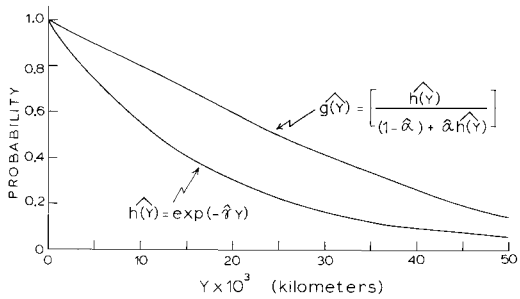

FIGURE 1. Estimated detection function $g(\hat{y})$ and function $h(\hat{y})$.

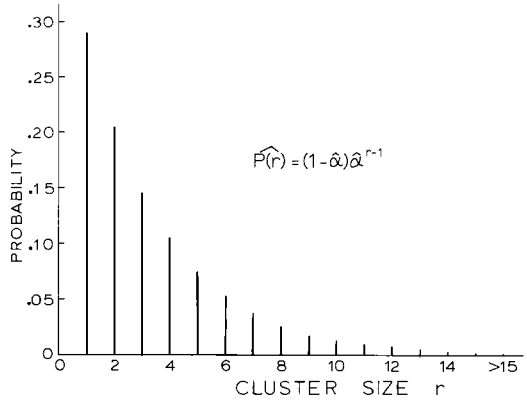

FIGURE 2. Estimated cluster size distribution.

a hypothetical inverse sampling plan designed to observe $N_0 = 25$ clusters resulted in the sighting distances y (in meters) and observed cluster sizes s presented in Table 1. Assume that a distance of $l = 25$ km was required to sight 25 clusters.

Assuming the geometric distribution

$$p(r) = (1 - \alpha)\alpha^{r-1} \quad r = 1, 2, \ldots \tag{13}$$

for cluster size and the exponential form

$$h(y) = \exp(-\gamma y) \quad y \geq 0, \gamma > 0 \tag{14}$$

for $h(y)$ in Eq. (8), it is easily seen that the cluster detection function has the form

$$g(y) = \exp(-\gamma y)/[(1 - \alpha) + (\alpha\exp(-\gamma y] \tag{15}$$

Similar calculations using (13) and (14) in (12) shows that

$$p(s|y) = (1 - \alpha)\alpha^{s-1}[g(y)]^s/ [(1 - \alpha) + \alpha\exp(-\gamma y)]. \tag{16}$$

Substituting Eq. (15) and Eq. (16) into Eq. (11) yields the likelihood function in terms of the parameters c, θ, α and γ. However, these parameters are not independent because substitution for $g(y)$ in Eq. (10) from Eq. (15) gives

$$c = -(\alpha\gamma)^{-1} \ln(1 - \alpha) \tag{17}$$

Therefore, the likelihood function, Eq. (8), must be maximized with respect to θ, α and γ. The estimate of the cluster density is (see Eq. (9)).

$$\hat{D} = \frac{\hat{\theta}}{2\hat{c}} \tag{18}$$

Finally, the estimate of Δ is obtained by noting the relationship

$$\Delta = DE(S),$$

where $E(S)$ is the expected cluster size. For the geometric distribution specified in Eq. (13) the expected cluster size is $(1 - \alpha)^{-1}$. Therefore the maximum likelihood estimate of Δ is

$$\hat{\Delta} = (1 - \hat{\alpha})^{-1}\hat{D}$$

where $\hat{D}$ is as in Eq. (18).

The maximization of the likelihood may be carried out using the FORTRAN based MAXLIK program (Kaplan and Elston 1978) designed to numerically find maximum likelihood estimates and their standard errors. Table 2 gives the estimates and their standard errors, based on data in Table 1.

The forms of $h(y)$ and the detection function $g(y)$, inserting the maximum likelihood estimates, $\hat{\gamma}$ and $\hat{\alpha}$, are given in Figure 1. As expected, the probability of sighting a cluster ($g(y)$) is greater than the probability of sighting an individual ($h(y)$) for all distances y.

Given, $\hat{\alpha}$, the estimated distribution of true cluster sizes is given in Figure 2. From this it is clearly seen that more than half of the clusters should have less than four individuals in them.

DISCUSSION

In conclusion it must be noted that the combined sampling method and the cluster sampling model presented in this paper are in a preliminary stage of their development. Many problems of practical importance have yet to be solved. For example, guidelines for the specification of L_0 and N_0 in a combined sampling plan need to be carefully formulated. Sensitivity of the cluster sampling model to errors in the specification of $p(r)$ and $h(y)$ must be investigated, and the possibility of developing a robust density estimator should be looked into. We are currently exploring solutions to some of these problems.