USING SIGN COUNTS TO QUANTIFY ANIMAL ABUNDANCE

JOHN R. SKALSKI, Center for Quantitative Science, University of Washington, Seattle, WA 98195

Abstract: I present a statistical framework for sign counts that permits abundance and density estimation for closed populations (e.g., in respect to ingress and egress). The estimation technique requires the capture, marking, and release of animals so the subsequent sign they produce is distinguishable from unmarked individuals. I develop variance estimators using finite sampling theory for the cases of simple random sampling and stratified random sampling of field plots for animal sign in the landscape. The variance formulas are used to determine the effects of sampling effort on the subsequent precision of abundance estimates. I present sample size formulas to determine the joint levels of marking effort and areal sampling required for a prespecified level of sampling precision in sign-marking studies.

Animal sign (e.g., tracks, excavations, and scats) provide an indirect method for inferring animal density. For some elusive species, an index based on sign counts may be all that is available. When the rate of sign production by a species can be estimated, these indices can be calibrated to provide estimates of animal abundance or density. Pellet counts from ungulates are the single best example of such calibrations where the mean number of pellet groups per animal per day have been extensively reported (Eberhardt and Van Etten 1956, Neff 1968, Freddy and Bowden 1983).

In certain instances, sign counts can be calibrated by using sign-marking techniques. An increasingly popular technique is marking scat by implants or injections for use with pellet count surveys (Pelton and Marcum 1977, Kinningham et al. 1980, Davison 1981). Track counts can be amended to identify marked animals from the tracks left by toe-clipped individuals (Dell 1957, Justice 1961). Similarly, active scent posts or dwellings may be marked for identification with internal dyes that pass through the digestive system of animals (New 1958, 1959; Brown and Conaway 1961) or by surveillance with radio telemetry. In all cases, an estimate of sign production by the marked animals can be used to calibrate index counts. Unfortunately, many sign-marking studies have been mistakenly analyzed as mark-recapture investigations, or important error sources have been omitted in confidence interval estimation (Conner and Labisky 1985).

I present a statistical model for sign-marking studies that permits abundance and density estimation in closed and open populations. I also present variance estimators to determine the levels of animal trapping and areal sampling

necessary to achieve a predetermined level of precision. This statistical framework can be used to interpret index counts over the broader class of indirect methods where sign marking is not possible.

METHODS

General Survey Design

My presentation of the analysis of sign counts begins with a general description of survey design using sign-marking techniques along with a hypothetical example of radioisotope tagging in a black tailed jack rabbit (*Lepus californicus*) population. I use data from the jack rabbit example to illustrate abundance estimation and sample size calculations. Following the description of the general survey approach, I present formal development of abundance estimators. The form of the estimator and its associated variance are shown to be a function of the areal survey design. Two special cases are presented: first, simple random sampling and then, stratified random sampling of the study area. Variance formulas are presented which can be used to determine levels of marking and spatial sampling necessary to achieve a prescribed level of precision. In the special case where marked and unmarked sign are uniformly mixed, simplified variance formulas are presented to determine

				0/31								51/51			
					16/35										42/53
												8/113			
	0/81				0/34										
							0/52						0/26		
				0/68			19/52			23/30					
		0/74						98/108							
												32/63			
							9/39								

Fig. 1. A population survey where 10 animals from a population have been marked and subsequently released for abundance estimation using sign-marking techniques. A simple random sample of 16 of 160 subplots are surveyed for marked and unmarked sign. Values above the slash are the numbers of marked sign counted (x_{ni}, $i = 1, \ldots, k$) and below the slash are the numbers of total sign counted (x_i, $i = 1, \ldots, k$).

minimum sample sizes when preliminary survey data are not available. The statistical methods conclude with a discussion of density estimation from sign counts.

I consider a population survey consisting of a 2-phase sampling program. In the first phase, trapping is conducted to obtain a random sample of n of the N animals in a population. Before release, the n animals in the first sample are marked in a manner that subsequently transmits a recognizable mark to the animal sign (e.g., scat marking [Crabtree 1989]). For purposes of point estimation, differentiating marked and unmarked sign is sufficient. However, to estimate sampling error for confidence interval construction, sign marking identifiable to individual animals is necessary.

In the second phase of the population survey, an areal sample of k of K quadrats comprising the study plot is performed (Fig. 1). The purpose of this areal survey is to estimate the number of sign left in the study area by the marked and unmarked animals in the population. At the beginning of the sampling period, k of K quadrats are selected for canvassing. These areas are initially searched to remove the old sign left before the survey and are later revisited at the end of the study to inventory the sign accumulated during the survey.

The areal survey should be a probability sample that permits an unbiased estimate of the total sign production on the study area. Traditional survey sampling techniques can be used in designing the areal survey (Hansen et al. 1953, Cochran 1977, Jessen 1978). A survey design should be selected that accounts for prior knowledge about animal movements and areal use. A simple random sample of the study area may be suitable if animal movements are relatively homogeneous across the study site (Fig. 1). For species with a distinct preference for particular habitats or terrain, stratified random samples may be a more efficient sampling scheme. Areas such as ravines, roadways, and banks of waterways may be selected to reflect this prior knowledge. A systematic sample is a third possibility, but is not recommended because a consistent variance estimate is not always feasible. Wolter (1984) discusses the use of systematic sampling and notes that variances, computed as if the sampling were conducted as simple random sampling, are usually conservative. In the following, I consider only simple random sampling and stratified random sampling, with complete canvassing as a special case.

Although the estimator I developed is analogous to a Lincoln Index (Seber 1982:59), the assumptions are quite different. These assumptions are (1) the population is closed to ingress and egress though mortality can occur if it is independent of marking, (2) all animals have equal probability of being caught in the first sample, (3) marking does not affect the rate of sign production of the animals, (4) the areal sample of the study site provides unbiased estimates of total sign production and marked sign production, (5) the mark is retained and transmitted to the sign throughout the course of the survey, and (6) all animal sign is correctly identified as being marked or unmarked and as sign

generated during the survey. In open populations subject to ingress and egress, a violation of assumption 2 is likely because the probability of capture is proportional to residence time.

RESULTS

Example of a Sign-Marking Study

For purposes of illustration, I present a sign-marking study with simple random sampling (Fig. 1). During an initial trapping period, 10 black-tailed jack rabbits were captured ($n = 10$), injected with unique radioisotope tracers, and released. A random sample of 16 of 160 subplots was then selected for initial canvassing to remove old pellets and for subsequent recanvassing 1 week after release. Results of the mock survey are reported in matrix format (Table 1), where the numbers of marked pellets found are cross-tabulated by subplot and by the marked animal contributing the sign (y_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n$). The matrix is augmented with additional information on the total numbers of sign (x_i, $i = 1, \ldots, k$) and marked sign (x_{ni}, $i = 1, \ldots, k$) found per subplot.

Abundance Estimation from Index Counts

The development of an abundance estimator from sign-marking data can be conceptualized using finite sampling theory. Consider the case where the study area is completely canvassed (i.e., all K quadrants sampled). Between the initial and final canvassings, the N animals of a population will produce X sign on the study plot. The mean production per animal is the $\bar{X} = X/N$. By taking a random sample of n animals that have a mean sign production denoted by $\bar{X}_n$, animal abundance can be estimated as:

$$\hat{N} = \frac{X}{\bar{X}_n} = \frac{nX}{X_n}, \qquad (1)$$

where X_n is the total sign production among the n marked animals. The abundance estimator equation (1) is a consistent estimator of N with an approximate expected value of

$$E_n(\hat{N}) \doteq N\left[1 + \left(\frac{1}{n} - \frac{1}{N}\right)\frac{S_x^2}{\bar{X}^2}\right], \qquad (2)$$

where

$$S_x^2 = \frac{\sum_{j=1}^{N}(X_j - \bar{X})^2}{N - 1}$$

is the variance in sign production on the study area between the N individuals of the population, and X_j is the sign produced on the study area by the jth individual in the population ($j = 1, \ldots, N$), under the assumption of n being a random sample of the N animals in the population. The positive bias (i.e., overestimation) of equation (2) becomes negligible as $n \rightarrow N$ and when the coefficient of variation in sign production among individuals (X_j, $j = 1, \ldots, N$) is small.

In practice, a complete canvassing and enumeration of X and X_n is unlikely. Instead, a fraction of the study area (i.e., k of K quadrats) will be sampled to obtain estimates of the parameters, i.e., $\hat{X}$ and $\hat{X}_n$. The more general form of the abundance estimator then becomes

$$\hat{N} = \frac{\hat{X}}{\hat{X}_n} = \frac{n\hat{X}}{\hat{X}_n}, \qquad (3)$$

where $\hat{X}$ and $\hat{X}_n$ are unbiased estimates of the total production of sign by the N and n animals on the study plot, respectively. The particular form for the estimators of total sign production (i.e., $\hat{X}$, $\hat{X}_n$) will depend on the sampling scheme used to survey the study area. Specific forms for equation (3) in the cases of simple random and stratified random sampling are presented below.

Simple Random Sample of Study Area.—In random sampling, the study area is divided into K equal size sampling units. A random sample of k of these K sampling units (or subplots) are then selected. Let x_i ($i = 1, \ldots, k$) be the number of sign found on the ith sampling unit, and let x_{ni} be the number of marked sign on the ith unit. The estimator of animal abundance based on simple random sampling and corresponding to equation (3) is

$$\hat{N} = \frac{\dfrac{K}{k}\sum_{i=1}^{k} x_i}{\dfrac{K}{nk}\sum_{i=1}^{k} x_{ni}} = \frac{n\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} x_{ni}} = \frac{nx.}{x_{n.}}. \qquad (4)$$

The variance estimate for equation (4) is found in stages (Appendix A) corresponding to the simple random sample of the study plot and the random sample of n of N animals for marking in the population. Variance formulas used in Conner and Labisky (1985) and Seber (1982:55, 376) ignore this second source of variation and consequently underestimate the true variance associated with sign-marking studies. To see this, note that equations (1) and (3) are consistent estimators only with regard to all possible selections of n of N animals in the population. The variance of $\hat{N}$ must therefore include the

variation that would be observed with alternative choices of n animals. The variance (Var) estimate of $\hat{N}$ with simple random sampling can be calculated as follows:

$$\widehat{\mathrm{Var}}(\hat{N}) = \frac{K^2\left(1 - \dfrac{k}{K}\right)}{k\hat{\bar{X}}^2}$$

$$\times \left\{ \left[1 + \frac{3\left(1 - \dfrac{n}{\bar{N}}\right)S_{\hat{X}_j}^2}{n\hat{\bar{X}}^2} \right] s_x^2 \right.$$

$$+ \frac{\hat{N}^2}{n^2}\left[1 + \frac{10\left(1 - \dfrac{n}{\bar{N}}\right)S_{\hat{X}_j}^2}{n\hat{\bar{X}}^2} \right] s_{x_n}^2$$

$$- \frac{2\hat{N}}{n}\left[1 + \frac{6\left(1 - \dfrac{n}{\bar{N}}\right)S_{\hat{X}_j}^2}{n\hat{\bar{X}}^2} \right]$$

$$\left. \times \, \mathrm{cov}(x_i, x_{ni}) \right\}$$

$$+ \hat{N}^2 \frac{\left(1 - \dfrac{n}{\bar{N}}\right)}{n\hat{\bar{X}}^2} s_{\hat{X}_j}^2 \qquad (5)$$

where

$$s_x^2 = \frac{\displaystyle\sum_{i=1}^{k} (x_i - \bar{x})}{k - 1} \qquad (6)$$

is variance in sign counts among sampling units,

$$s_{x_n}^2 = \frac{\displaystyle\sum_{i=1}^{k} (x_{ni} - \bar{x}_n)^2}{k - 1} \qquad (7)$$

is variance in marked sign counts amoung sampling units,

$$\mathrm{cov}(x_i, x_{ni}) = \frac{\displaystyle\sum_{i=1}^{k} (x_i - \bar{x})(x_{ni} - \bar{x}_n)}{k - 1} \qquad (8)$$

is covariance between total sign counts and marked sign counts on sampling units,

$$\bar{x} = \sum_{i=1}^{k} \frac{x_i}{k} \qquad (9)$$

is mean number of sign counts per sampling unit, and

$$\bar{x}_n = \sum_{i=1}^{k} \frac{x_{ni}}{k} \qquad (10)$$

is mean number of marked sign per sampling unit. To complete the evaluation of equation (5), the mean and variance in sign production between individuals on the study plot must be estimated ($\bar{X}$ and $S_{X_j}^2$. For convenience, let y_{ij} denote the number of sign found on the ith sampling unit from the jth marked animal ($j = 1, \ldots, n$). Then let

$$y_{.j} = \sum_{i=1}^{k} y_{ij}$$

denote the total number of marked sign found from the jth ($j = 1, \ldots, n$) animal. Similarly,

$$\sum_{j=1}^{n} y_{ij} = x_{ni}$$

illustrates the relationship between subplot counts and sign counts associated with individual animals. Then, an estimate of the mean sign production for the n marked animals is

$$\hat{\bar{X}} = \frac{1}{n}\sum_{j=1}^{n} \hat{X}_j = \frac{K}{nk}\sum_{j=1}^{n} y_{.j}. \qquad (11)$$

The variance in sign production among the n marked animals can then be estimated as

$$S_{\hat{X}_j}^2 = \frac{\displaystyle\sum_{j=1}^{n} (\hat{X}_j - \hat{\bar{X}})^2}{(n - 1)}$$

$$= \left(\frac{K}{k}\right)^2 \sum_{j=1}^{n} \frac{(y_{.j} - \bar{y})^2}{(n - 1)}, \qquad (12)$$

where

$$\bar{y} = \sum_{j=1}^{n} \frac{y_{.j}}{n}. \qquad (13)$$

Typically, equation (12) can be expected to overestimate the true variance in sign production between animals, for it includes the sampling error associated with the estimation of $\hat{X}_j$. As an alternative to the use of equation (12), the expression:

Table 1. Matrix of survey results indicating numbers of marked sign (y_{ij}) found per marked individual ($j = 1, \ldots, n$) on each subplot canvassed ($i = 1, \ldots, k$) along with total number of marked (x_{ni}, $i = 1, \ldots, k$) and unmarked (x_i, $i = 1, \ldots, k$) sign for each subplot canvassed.

Subplot	Individual										x_{ni}[a]	x_i[b]
	1	2	3	4	5	6	7	8	9	10		
1	0	0	0	0	0	0	0	0	0	0	0	31
2	0	0	0	0	51	0	0	0	0	0	51	51
3	0	16	0	0	0	0	0	0	0	0	16	35
4	0	0	0	0	0	0	0	0	0	42	42	53
5	0	0	0	0	8	0	0	0	0	0	8	113
6	0	0	0	0	0	0	0	0	0	0	0	81
7	0	0	0	0	0	0	0	0	0	0	0	34
8	0	0	0	0	0	0	0	0	0	0	0	52
9	0	0	0	0	0	0	0	0	0	0	0	26
10	0	0	0	0	0	0	0	0	0	0	0	68
11	0	0	0	19	0	0	0	0	0	0	19	52
12	0	0	0	0	0	0	0	0	23	0	23	30
13	0	0	0	0	0	0	0	0	0	0	0	74
14	61	0	0	37	0	0	0	0	0	0	98	108
15	0	0	0	0	0	0	0	32	0	0	32	63
16	9	0	0	0	0	0	0	0	0	0	9	39
$y_{.j}$[c]	70	16	0	56	59	0	0	32	23	42		
s_{yj}^2[d]	233.05	16	0	102.27	163.16	0	0	64	33.06	110.25		
$x_{n.}$[e]											298	
$x_{.}$[f]												910

[a] Marked sign found on the ith subplot surveyed.
[b] Total sign found on the ith subplot surveyed.
[c] Total marked sign detected from the jth marked animal.
[d] Sample variance in marked sign from the jth marked animal among subplots surveyed.
[e] Total marked sign found on all subplots surveyed.
[f] Total sign found on all subplots surveyed.

$$\hat{S}_{x_j}^2 = s_{x_j}^2 - \frac{K^2\left(1 - \frac{k}{K}\right)}{nk} \sum_{j=1}^{n} s_{yj}^2, \quad (14)$$

where

$$S_{yj}^2 = \sum_{i=1}^{k} \frac{(y_{ij} - \bar{y}_j)^2}{(k - 1)},$$

and

$$\bar{y}_j = \sum_{i=1}^{k} \frac{y_{ij}}{k},$$

should give an unbiased estimate of $S_{x_j}^2$. A negative value of equation (14) may arise, in which case equation (12) should be used.

In the design stage of an animal survey, an expression for $\text{Var}(\hat{N})$ is more useful for planning purposes than the $\hat{\text{Var}}(\hat{N})$ used in data analysis. In sign-marking studies, the form of $\text{Var}(\hat{N})$ will depend on the spatial distribution of the n marked animals and the animals' movement on the study plot. In the absence of specific information, a uniform mixing of marked and un-

marked animal sign may be assumed to calculate sample size (Appendices B and C). Under such an assumption, the variance of the abundance estimator becomes

$$\text{Var}(\hat{N}) \doteq \frac{N^2\left(1 - \frac{n}{N}\right)K\left(1 - \frac{k}{K}\right)}{n\bar{X}k}$$

$$+ \frac{N^2\left(1 - \frac{n}{N}\right)}{n\bar{X}^2} S_{x_j}^2. \quad (15)$$

Consequently, to predict the precision of the population survey, one must know prior estimates of the mean and variance in sign production among animals in addition to population size. The precision of the study is then calculated as a function of the fraction of animals marked in the population and the proportion of the study area canvassed for sign.

Analysis of Example.—The results of the survey of plots ($k/K = 16/160$) yielded $x_{.} = 910$ pellets, of which $x_{n.} = 298$ were marked from

the $n = 10$ animals released on the site (Table 1). The estimate of animal abundance from equation (1) is therefore

$$\hat{N} = \frac{10(910)}{298} = 30.54. \qquad (16)$$

The next step in the analysis is computing an estimate of the variance (i.e., $\widehat{Var}(\hat{N})$) of the abundance estimate. The variance in pellet counts among the subplots in the survey is

$$s_x^2 = \sum_{i=1}^{16} \frac{(x_i - \bar{x})^2}{15} = 704.25. \qquad (17)$$

Similarly, the variance in the numbers of marked pellets among subplots is

$$s_{x_n}^2 = \sum_{i=1}^{16} \frac{(x_{ni} - \bar{x}_n)^2}{15} = 715.58. \qquad (18)$$

The covariance between the total and number of marked pellets per plot is estimated to be

$$\mathrm{cov}(x_i, x_{ni}) = \frac{\left(\sum_{i=1}^{16} x_i x_{ni} - \dfrac{\sum_{i=1}^{16} x_i \cdot \sum_{i=1}^{16} x_{ni}}{16} \right)}{15}$$

$$= 264.75. \qquad (19)$$

The final intermediate statistics needed for variance estimation are estimates of mean pellet production per individual

$$\hat{\bar{X}} = \frac{160}{10(16)} \sum_{j=1}^{10} \sum_{i=1}^{16} y_{ij} = 298 \qquad (20)$$

and the variance in pellet production between marked individuals where

$$s_{\bar{x}_j}^2 = \frac{\left(\dfrac{160}{16}\right)^2 \sum_{j=1}^{10} (y_{\cdot j} - \bar{y})^2}{9}$$

$$= \frac{\left(\dfrac{160}{16}\right)^2 [(70 - 29.8)^2 + \ldots + (42 - 29.8)^2]}{9}$$

$$= 68{,}995.56. \qquad (21)$$

In this case of the jack rabbit data, $\hat{S}_{x_j}^2$ is negative, so equation (21) is the best estimate of the variance in pellet production. Substituting the intermediate calculations (eq 16–21) into equation (5), the variance estimate of $\widehat{Var}(\hat{N}) = 192.28$ results. The standard error of the estimate,

$\widehat{SE}(\hat{N}) = 13.87$, can be used to construct a confidence interval: $CI[\hat{N} \pm Z_{\alpha/2} \cdot SE(\hat{N})] = 1 - \alpha$, $CI[30.54 \pm 1.645(13.87)] = 1 - 0.10$, $CI(7.72 \le N \le 53.36) = 0.90$, or more correctly, $CI(10.0 \le N \le 53.36) = 0.90$, taking into account the known number of marked animals ($n = 10$) on the site. The relatively large coefficient of variation $[CV = 100\% \cdot \widehat{SE}(\hat{N})/\hat{N}]$ of 45.4% suggest greater trapping or areal sampling would be needed for reliable abundance estimation.

Stratifed Random Sample of Study Area.—Animal preferences for specific terrain or habitat suggest that a stratified random sample (Fig. 2) of a study site may often be appropriate. In this situation, the area is divided into L strata. These strata will often be associated with habitat or terrain suspected of having different levels of animal activity and consequently different densities of animal sign. Within each of these strata, the area is subdivided into K_h ($h = 1, \ldots, L$) sampling units of equal size. Between strata, the size and configuration of these sampling units can differ to take advantage of the most efficient method of canvassing each type of environment. Then, taking a random sample of k_h of K_h sampling units in each strata, animal abundance is estimated by

$$\hat{N} = \frac{n \sum_{h=1}^{L} K_h \bar{x}_h}{\sum_{h=1}^{L} K_h \bar{x}_{nh}}, \qquad (22)$$

where

$$\bar{x}_h = \sum_{i=1}^{k_h} \frac{x_{hi}}{k_h}$$

and

$$\bar{x}_{nh} = \sum_{i=1}^{k_h} \frac{x_{nhi}}{k_h}.$$

The estimated variance of $\hat{N}$ for a stratified random sample of the study site can be expressed as:

$$\widehat{Var}(\hat{N})$$

$$= \left[\frac{1}{\hat{\bar{X}}^2} + \frac{3\left(1 - \dfrac{n}{N}\right)s_{\bar{x}_j}^2}{n\hat{\bar{X}}^4} \right]$$

$$\times \sum_{h=1}^{L} K_h^2 \frac{\left(1 - \dfrac{k_h}{K_h}\right)}{k_h} s_{x_h}^2$$

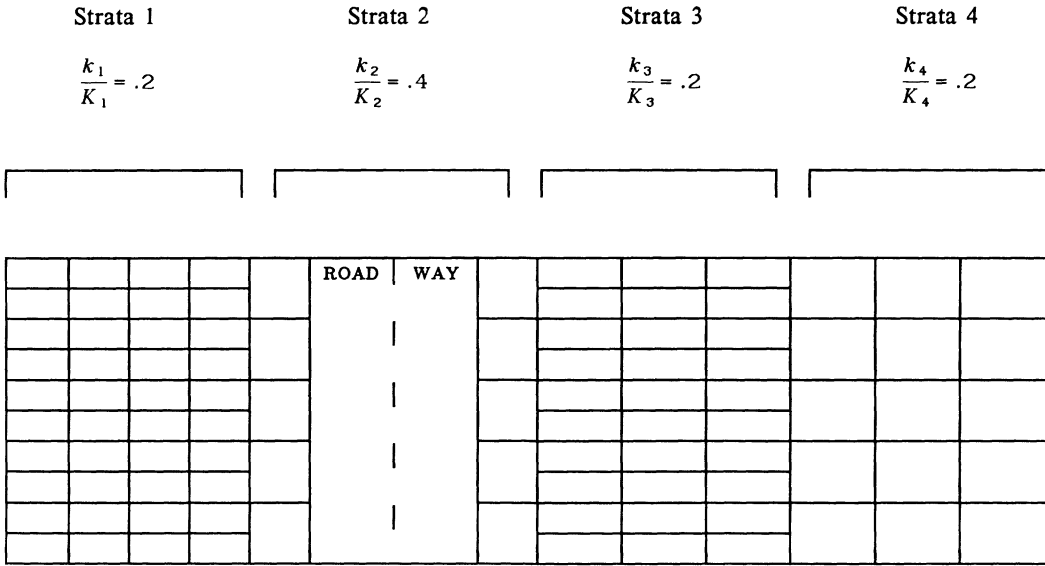| Strata 1	Strata 2	Strata 3	Strata 4
$\dfrac{k_1}{K_1} = .2$	$\dfrac{k_2}{K_2} = .4$	$\dfrac{k_3}{K_3} = .2$	$\dfrac{k_4}{K_4} = .2$

Fig. 2. A stratified random sample of a study area into strata (*L*) and the associated sampling fractions (k_h/K_h) for each stratum, where k_h is the number of subplots sampled out of K_h subplots per stratum.

$$+ \ \hat{N}^2 \left[\frac{1}{n^2 \hat{\bar{X}}^2} + \frac{10 \left(1 - \dfrac{n}{N}\right) s_{\bar{x}_j}^2}{n^3 \hat{\bar{X}}^4} \right]$$

$$\times \ \sum_{h=1}^{L} K_h^2 \frac{\left(1 - \dfrac{k_h}{K_h}\right) s_{x_{nh}}^2}{k_h}$$

$$- \ 2\hat{N} \left[\frac{1}{n \hat{\bar{X}}^2} + \frac{6 \left(1 - \dfrac{n}{N}\right) s_{\bar{x}_j}^2}{n^2 \hat{\bar{X}}^4} \right]$$

$$\times \ \sum_{h=1}^{L} K_h^2 \frac{\left(1 - \dfrac{k_h}{K_h}\right)}{k_h} \, \mathrm{cov}(x_{hi}, \, x_{nhi})$$

$$+ \ \frac{N^2 \left(1 - \dfrac{n}{N}\right)}{n \hat{\bar{X}}^2} s_{x_j}^2. \tag{23}$$

The quantities s_{xh}^2, s_{xnh}^2, and $\mathrm{cov}(x_{hi}, x_{nhi})$ in equation (23) are computed the same as equations (6–8), respectively, except now on a stratum-by-stratum basis. To obtain estimates of the number of sign produced by individual marked animals,

$$\hat{X}_j = \sum_{h=1}^{L} \frac{K_h}{k_h} \sum_{i=1}^{k_h} y_{hij}, \tag{24}$$

where y_{hij} is the number of sign belonging to

the *j*th marked animal ($j = 1, \ldots, n$) on the *i*th sampling unit ($i = 1, \ldots, k_h$) of the *h*th strata ($h = 1, \ldots, L$). The mean and variance of $\hat{X}_j$ for use in equation (23) are found using expressions (11) and (12), respectively. As an alternative to the use of equation (12) in stratified random sampling

$$\hat{S}_{x_j}^2 = s_{\bar{x}_j}^2 - \frac{1}{n} \sum_{j=1}^{n} \sum_{h=1}^{L} \frac{K_h^2 \left(1 - \dfrac{k_h}{K_h}\right)}{k_n} s_{y_{hj}}^2$$

can be used, provided the point estimate is positive.

To evaluate alternative allocations of effort, a variance expression analogous to equation (15) but for stratified random sampling, can be used where

$$\mathrm{Var}(\hat{N}) = \frac{N \left(1 - \dfrac{n}{N}\right)}{n \bar{X}^2} \sum_{h=1}^{L} K_h^2 \frac{\left(1 - \dfrac{k_h}{K_h}\right)}{k_h} \bar{X}_h$$

$$+ \ N^2 \left(1 - \frac{n}{N}\right) \frac{S_{x_j}^2}{n \bar{X}^2} \tag{25}$$

and where $\bar{X}_h$ is the mean sign production by the individuals in the *h*th stratum. An estimate of the mean number of sign in each strata must be known along with the mean and variance in sign production between animals to predict the sampling precision of a future survey.

Sample Size Calculations for Population Surveys

For an observational study with the objective of abundance estimation (or equivalently, density estimation), the precision of a survey estimate of $\hat{N}$ can be written as

$$P\left(\left|\frac{\hat{N} - N}{N}\right| < \epsilon\right) \geq 1 - \alpha \qquad (26)$$

where ϵ is the maximum relative error expected with a probability of $1 - \alpha$. Assuming approximate normality of the abundance estimate from equation (3), the anticipated precision of a population survey can be estimated from the cumulative normal distribution where

$$P\left(\left|\frac{\hat{N} - N}{N}\right| < \epsilon\right) \doteq 1 - 2\phi\left(\frac{-\epsilon}{CV(\hat{N})}\right) \qquad (27)$$

and where

$$CV(\hat{N}) = \frac{\sqrt{Var(\hat{N})}}{N}.$$

For simple random sampling inspection of equations (5) and (15) indicates that the $CV(\hat{N})$ is a function of 2 design parameters $\left(P_N\left[= \frac{n}{N} = 1 - Q_N\right], P_K\left[= \frac{k}{K} = 1 - Q_K\right]\right)$ and 3 population parameters $(N, \bar{X}, \text{ and } CV(X_j))$ that must be available prior to performing sample size calculations. With prior estimates of $\hat{\bar{X}}$ and $\widehat{CV}(X_j)$, equation (27) can be used to determine levels of P_N and P_K needed to be within $\epsilon 100\%$ of the true value of N, $(1 - \alpha)100\%$ of the time.

Example of Sample Size Calculations.—With the sample results from the jack rabbit example, alternative allocations of sampling effort can be determined such that future surveys of the jack rabbit population may have a prescribed level of precision. The variance of $\hat{N}$ is a function of the sampling fractions of the number of animals marked and the number of subplots canvassed. Substituting the numerical results (eqs 16–21) into equation (5) for $n = 10$ and k unspecified,

$$\widehat{Var}(\hat{N}) = 2551.89\left(\frac{1}{k} - \frac{1}{160}\right) + 48.74,$$

the level of spatial sampling can be determined for prespecified levels of $(1 - \alpha)$ and ϵ from equation (27).

To be within 50% of the true value of N 80% of the time, the number of subplots to randomly sample is calculated as

$$0.8 = 1 - 2$$
$$\times \phi\left(\frac{-30.54(0.50)}{\sqrt{2551.89\left(\frac{1}{k} - \frac{1}{160}\right) + 48.74}}\right)$$

$$\phi^{-1}(0.10) = \frac{-15.27}{\sqrt{\left(\frac{2551.89}{k} + 32.79\right)}}$$

$$-1.2817 = \frac{-15.27}{\sqrt{\left(\frac{2551.89}{k} + 32.79\right)}}$$

$$k = 23.4,$$

or, rounding up to the next largest integer, $k = 24$.

To be within ± 40% of the true value of N 80% of the time, $k = 44$ subplots must be sampled when $n = 10$ animals are marked. For $\epsilon = 0.35$ and $1 - \alpha = 0.80$, $k = 70$ subplots must be sampled. However, for $\epsilon = 0.25$ and $1 - \alpha = 0.80$, $k = 947$ subplots require sampling. In other words, canvassing the entire site ($k = 160$) will not yield the desired precision when only 10 animals are sign-marked. Both n and k must be adjusted to yield a precision of $\epsilon = 0.25$ and $1 - \alpha = 0.80$. With $n = 15$ sign-marked animals and $k = 57$, a precision of $\epsilon = 0.25$ and $1 - \alpha = 0.80$ is achievable. These sample size calculations illustrate the importance of both the intensity of marking and spatial subsampling on the overall performance of a sign-marking survey.

The marked sign in the jack rabbit example were not randomly mixed across the site, but appear in clusters (Fig. 1). Sample size calculations in this case will be underestimated using the simple variance formula (15) which assumes uniform mixing of marked and unmarked sign. In contrast to the above calculations, the variance of $\hat{N}$ assuming formula (15) is

$$Var(\hat{N}) = 33.68\left(\frac{1}{k} - \frac{1}{160}\right) + 48.74,$$

where the first term associated with the spatial variance in marked sign is underestimated. A Chi-square test of homogeneity can be used to test for uniform mixing of marked and unmarked sign across subplots. In the case of jack rabbit data, the assumption of homogeneity is rejected, $(\chi_{15}^2 > 573.98, P \approx 0)$, and variance formula (5) must be used in sample size calculations.

Relationship Between Abundance and Density Estimation

Different survey methods are typically required to estimate absolute abundance versus population density when using mark-recapture techniques (White et al. 1982). However, the distinction disappears in sign-marking studies. Define density (D) as

$$D = \frac{1}{A} \sum_{j=1}^{N} P_j, \qquad (28)$$

where A = areal size of the study site, N = total number of animals using the site for some fraction of the survey period (previously estimated by $\hat{N}$), and P_j = fraction of survey period the jth animal $(j = 1, \ldots, N)$ used the site. From equation (28), density is a function of the cumulative temporal use of a study plot of size A by the N individuals of the population. Assuming the number of sign left on the site by the jth animal may be roughly proportional to its residence time on the study plot, then

$$E(X_j) = R_j P_j, \qquad (29)$$

where R_j is the total sign production by the jth animal regardless of location. It follows from equation (29) that a natural estimate of P_j is then

$$\hat{P}_j = \frac{X_j}{R_j}$$

such that animal density can be written as

$$D = \frac{1}{A} \sum_{j=1}^{N} \frac{X_j}{R_j}.$$

Assuming all animals have the same expected deposition $E(R_j) = R$, then density can be written as

$$D = \frac{\sum_{j=1}^{N} X_j}{AR} = \frac{X}{AR} \qquad (30)$$

for subsequent estimation.

The average sign production, R, in equation (30) can be estimated from the n marked animals by selecting a plot size, A, sufficiently large to encompass the home ranges of all n tagged animals. In which case, R is estimated by $\hat{\bar{X}}_n$, and density is estimated by

$$\hat{D} = \frac{\hat{X}}{A\hat{\bar{X}}_n} = \frac{\hat{N}}{A}, \qquad (31)$$

the naive density estimator (White et al. 1982:120). Hence, in situations where the n individ-

uals are closed to ingress and egress, density and abundance estimation are essentially equivalent when using sign-marking data. Variance expressions for density estimates can be calculated from equations (5) and (23) by noting the relationship

$$\widehat{\text{Var}}(\hat{D}) = \frac{1}{A^2} \widehat{\text{Var}}(\hat{N}). \qquad (32)$$

The assumptions for density estimation (eq 31) are essentially the same as listed for abundance estimation. However, the assumption of closure can now be relaxed to include just the n marked animals.

DISCUSSION

Guidance on the design and analysis of index studies is surprisingly sparse in wildlife and statistical literature (Eberhardt 1978, Giles 1969). Typically, regression or ANOVA procedures are employed under the assumptions of normality and linear response models. However, the actual distributional properties and sampling models may be quite different or are unknown. In the development of estimators for sign-marking techniques, finite sampling theory is used, thereby avoiding distributional or model assumptions. The subsequent techniques are therefore nonparametric and are robust to varying environmental or population conditions. The variance estimators can be used with knowledge about rates of sign production, areal density, and spatial variation in sign to determine the sampling effort required for prespecified levels of sampling precision.

Another application of sign-marking techniques is the comparison of animal density between spatially or temporally distinct sites, or the assessment of treatment effects on wild populations. In situations where an environmental effect may cause a demographic change, a behavioral change should also be suspected. For instance, manipulation of cover might cause a change in carrying capacity and/or the movement and diet of animals. As a result, the frequency of track counts may subsequently change, but pellet counts may or may not be affected. A prudent choice of which animal sign to count should reduce the likelihood of biased contrasts between treatments.

Alternatively, sign marking can be used to validate the assumption of homogeneous calibration or to convert index counts to absolute abundance. Without tests of assumptions and

valid variance estimates, management decisions based on sign counts could be wrong. By modifying index techniques, by improving the reliability of sign counts, and by including error variances in risk assessment, management decisions under uncertainty can be improved.

LITERATURE CITED

BROWN, L. N., AND C. H. CONAWAY. 1961. Dye excretion as a method for determination of small mammal home range. Am. Midl. Nat. 66:128–137.

COCHRAN, W. G. 1977. Sampling techniques. Third ed. John Wiley & Sons Ltd., New York, N.Y. 428pp.

CONNER, M. C., AND R. F. LABISKY. 1985. Evaluation of radioisotope tagging for estimating abundances of raccoon populations. J. Wildl. Manage. 49:326–332.

CRABTREE, R. L. 1989. Socio-demographics of an unexploited coyote population. Ph.D. Thesis, Univ. Idaho, Moscow. 70pp.

DAVISON, R. P. 1981. The effect of exploitation on some parameters of coyote populations. Ph.D. Thesis, Utah State Univ., Logan. 153pp.

DELL, J. 1957. Toe clipping varying hares for track identification. N.Y. Fish and Game J. 4:61–68.

DESAPIO, R. 1978. Calculus for the life sciences. Freeman, San Francisco, Calif. 740pp.

EBERHARDT, L. L. 1978. Appraising variability in population studies. J. Wildl. Manage. 42:207–238.

———, AND R. C. VAN ETTEN. 1956. Evaluation of the pellet group count as a deer census method. J. Wildl. Manage. 20:70–74.

FREDDY, D. J., AND D. C. BOWDEN. 1983. Sampling mule deer pellet-group densities in juniper-pinyon woodland. J. Wildl. Manage. 47:476–485.

GILES, R. H., JR., editor. 1969. Wildlife management techniques. Third ed. The Wildlife Society, Washington, D.C., 633pp.

HANSEN, M. H., W. N. HURWITZ, AND M. G. MADOW. 1953. Sample survey methods and theory. Vol. 1. John Wiley & Sons Ltd., New York, N.Y. 638pp.

JESSEN, R. J. 1978. Statistical survey techniques. John Wiley & Sons Ltd., New York, N.Y. 520pp.

JUSTICE, K. E. 1961. A new method for measuring home ranges of small mammals. J. Mammal. 42:462–470.

KINNINGHAM, M. J., M. R. PELTON, AND D. C. FLYNN. 1980. Use of the pellet count technique for determining densities of deer in the southern Appalachians. Proc. Southeast. Assoc. Fish Wildl. Agencies 34:508–514.

NEFF, D. J. 1968. The pellet-group count technique for big game trend, census and distribution: a review. J. Wildl. Manage. 32:597–614.

NEW, J. G. 1958. Dyes for studying the movements of small mammals. J. Mammal. 39:416–429.

———. 1959. Additional uses of dyes for studying the movements of small mammals. J. Mammal. 23:348–351.

PELTON, M. R., AND L. C. MARCUM. 1977. Potential use of radioisotopes for determining densities of black bears and other carnivores. Pages 221–236 *in* R. L. Phillips and C. Jonkel, eds. Proc. of the 1975 predator symposium. Mont. For. Conserv. Exp. Stn., Univ. Montana, Missoula.

SEBER, G. A. F. 1982. The estimation of animal abundance and related parameters. Second ed. Macmillan Publ. Co., New York, N.Y. 654pp.

WHITE, G. C., D. R. ANDERSON, K. P. BURNHAM, AND L. L. OTIS. 1982. Capture-recapture and removal methods for sampling closed populations. LA-8781-NERP. Los Alamos Natl. Lab., Los Alamos, N.M. 235pp.

WOLTER, K. M. 1984. An investigation of some estimators of variance for systematic sampling. J. Am. Stat. Assoc. 79:781–790.

APPENDIX A

Estimated Variance

For simple random sampling the variance of the abundance estimate is found by first conditioning on the n animals tagged and then taking expectation over all $\binom{N}{n}$ choices where

$$\text{Var}(\hat{N})$$

$$= E_n[\text{Var}(\hat{N}|n)] + \text{Var}_n[E(\hat{N}|n)]$$

$$= \frac{n^2 K^2 \left(1 - \dfrac{k}{K}\right)}{k}$$

$$\times E_n\left[S_x^2 \left(\frac{1}{X_n^2}\right) + S_{x_n}^2\left(\frac{X^2}{X_n^4}\right) \right.$$

$$\left. - 2\,\text{cov}(x_i, x_{ni})\left(\frac{X}{X_n^3}\right)\right]$$

$$+ \text{Var}_n\left(\frac{nX}{X_n}\right)$$

$$= n^2 K^2 \frac{\left(1 - \dfrac{k}{K}\right)}{k}$$

$$\times \left\{ S_x^2 E_n\left(\frac{1}{X_n^2}\right) \right.$$

$$+ X^2 E_n(S_{x_n}^2) E_n\left(\frac{1}{X_n^4}\right)$$

$$\left. - 2X E_n[\text{cov}(x_i, x_{ni})] E_n\left(\frac{1}{X_n^3}\right) \right\}$$

$$+ n^2 X^2 \text{Var}_n\!\left(\frac{1}{X_n}\right). \tag{33}$$

Using Taylor series approximations (DeSapio 1978:491–523) for terms of the order $(1/X_n)^a$, the variance can be expressed as

$\text{Var}(\hat{N})$

$$\doteq \frac{K^2\!\left(1 - \dfrac{k}{K}\right)}{k\bar{X}^2}$$

$$\times \left\{\left[1 + \frac{3\!\left(1 - \dfrac{n}{N}\right)S_{x_j}^2}{n\bar{X}^2}\right]S_x^2\right.$$

$$+ \frac{N^2}{n^2}\left[1 + \frac{10\!\left(1 - \dfrac{n}{N}\right)S_{x_j}}{n\bar{X}^2}\right]E_n(S_{x_n}^2)$$

$$- \frac{2N}{n}\left[1 + \frac{6\!\left(1 - \dfrac{n}{N}\right)S_{x_j}^2}{n\bar{X}^2}\right]$$

$$\left. \times E_n[\text{Cov}(x_i,\ x_{ni})]\right\}$$

$$+ \frac{N^2\!\left(1 - \dfrac{n}{N}\right)}{n\bar{X}^2}S_{x_j}^2. \tag{34}$$

The values $E_n(S_{x_n}^2)$ and $E_n[\text{Cov}(x_i,\ x_{ni})]$ will depend on the dispersion pattern of the population (Appendix C). For purposes of variance estimation, sample values may be substituted into equation (34) to obtain equation (5). In the case of stratified random sampling, the above approach is used within each stratum with the resulting variance (eq 23) being a weighted sum of the within-strata variances.

APPENDIX B
Estimated Variance under Random Mixing

For sample size calculations, the dispersion of animal sign may be approximated by a random mixing model. Under such an assumption, the probability a randomly selected sign is marked is $P = \dfrac{n}{N}$ for $P = 1 - Q$, so

$$E_n(S_{x_n}^2) = \frac{X}{K}PQ + P^2 S_x^2, \tag{35}$$

and

$$E_n[\text{Cov}(x_i,\ X_{ni})] = PS_x^2. \tag{36}$$

Using the further simplification of

$$E\!\left(\frac{1}{X^n}\right) \doteq \frac{1}{E(X)^n}$$

in the derivation of equations (33–34), the variance formula simplifies to equation (15). A similar argument can be used to derive equation (25).

APPENDIX C
Variance and Covariance

The general expressions for the expected values of the variance of x_{ni} and the covariance of x_i and x_{ni} as a function of the numbers of animals marked (e.g., eq 34) can be found using indicator variables. It can be shown that

$$E_n[S_{x_n}^2] = \frac{n(N - n)}{N(N - 1)}\sum_{j=1}^{N} S_{x_j}^2$$

$$+ \frac{n(n - 1)}{N(N - 1)}S_x^2 \tag{37}$$

where $S_{x_j}^2$ = variance in sign for the jth animal among subplots ($i = 1, \ldots, k$). Similarly,

$$E_n[\text{Cov}(x_i,\ x_{ni})] = \frac{n}{N}S_x^2. \tag{38}$$

Although equations (37) and (38) are free of assumptions concerning the distribution of animals on the study site, the prodigious parameterization of equation (37) limits its usefulness in sample size calculations.