Horvitz-Thompson Survey Sample Methods for Estimating Large-scale Animal Abundance

Michael D. Samuel
National Biological Survey
National Wildlife Health Research Center
Madison, Wisconsin

Edward O. Garton
Department of Fish and Wildlife
University of Idaho
Moscow

Introduction

Wildlife management depends on a fundamental knowledge of species population dynamics and on the ability to monitor population changes or responses to management. Animal abundance and the rate of change are two of the principal parameters for assessing the status of wildlife populations and determining the need for management. Survey sampling (or descriptive sampling) methods typically are used to obtain large-scale abundance estimates because the financial and logistic constraints allow sampling on only a fraction of the area occupied by the population (Cochran 1977, Eberhardt and Thomas 1991). However, most survey sampling methods fail to account for the fact that many animals on the sampled plots are not detected. In contrast, numerous methods have been developed to estimate the probability of detecting animals in small areas (Seber 1982, Lancia et al. 1994), but most have not been extended to large-scale surveys.

In addition to providing essential management information, considerable progress in large-scale ecological and environmental research can be made by combining field observations with controlled experiments (Eberhardt and Thomas 1991). These large-scale experimental field studies (Sinclair 1991) and evaluations of wildlife management (Macnab 1983) also should be conducted with a sampling framework that encompasses the population of concern (Eberhardt and Thomas 1991). To realistically evaluate landscape- and ecosystem-scale research, the spatial design of the sampling effort must be at least as meticulous as the design of experimental manipulations.

The basic components that must be included in large-scale animal abundance surveys are the delineation of the species (or population) range, the characterization of the geographic distribution of abundance, the selection of a spatial sampling design and the application of a survey technique to determine the abundance on each surveyed plot. These factors must be integrated to produce a statistically based estimate of the population size that meets the required level of precision. In this paper, we describe these basic survey components and present a Horvitz-Thompson survey sample method for wildlife populations. This method requires that the survey design consider the spatial distribution of animals and the effort needed to estimate animal abundance on sample plots.
Elements of Large-scale Surveys

Survey Objectives and Overview

The first step in planning a survey is to establish a clear set of survey objectives, including the goals of the survey, anticipated uses and level of desired precision. The precision of the abundance estimate is a critical component of planning a survey and evaluating its success. Robson and Regier (1964) provided some general guidelines for setting precision and accuracy objectives, but the circumstances and goals of each survey may be unique. When the objectives are established, a preliminary survey can be designed to meet the precision, accuracy, cost criteria and other critical features of the survey. Calculations can be made to determine the number of sample plots and the effort required to detect animals on sample plots. Because survey accuracy is affected by errors in spatial sampling (plot to plot variation) and the detection of animals on a sample plot, trade-offs between the number of sample plots and the effort spent detecting animals on those plots is necessary (Alho 1992). Finally, survey planning and optimization require informed guesses about the survey design characteristics, probability of detecting animals and survey costs. Because additional knowledge and experience are acquired each time the survey is conducted, a survey may require a recursive approach using previous results to further improve and optimize the design.

Population Range and Survey Boundary

After the survey objectives are established, the population boundary or species range must be determined. Because most survey methods require a geographic framework for estimating population abundance, a limit on the geographic area for the survey must be established. Precise delineations of the geographic boundaries for a population may be difficult for large-scale surveys when exact boundaries are unknown. Species range maps, broad ecosystem boundaries, habitat characteristics and other factors may provide useful guidelines for establishing the geographic scope of the survey. In the absence of such guidelines, a pre-survey may be necessary to determine the population boundary. The geographic boundary defines the extent of the population that will be surveyed. Therefore, changes in the geographic boundary have a direct effect on estimates of abundance.

Two practical problems in conducting large-scale surveys within fixed geographic boundaries are inadvertently excluding animals outside the boundary and wasting survey resources by sampling in areas with no animals. Animals that occur outside the survey boundary will not be included in the population estimate. Difficulties in identifying adequate geographic boundaries may be compounded for highly mobile species that respond primarily to suitable environmental conditions. For example, pintail ducks (Anas acuta) may drastically change annual breeding distributions in response to environmental conditions (Bellrose 1976: 267). Less obvious problems can arise when specific portions of the population have different use patterns by age or sex. If these problems exist, survey boundaries may have to be adjusted annually to account for changes in distribution.

Because large-scale surveys typically are expensive, excluding areas that have no animals within the geographic boundary of the survey may be cost effective. There are no conceptual reasons why the survey areas must be contiguous or why some...
areas within the geographic boundary cannot be excluded. Exclusions may be based on known geographic distributions, unsuitable habitats or other reliable indicators that no (or relatively few) animals will be present. Again, animals outside the survey boundaries will not be included in the population estimate.

**Spatial Sampling Patterns**

Few species of animals have a uniform distribution of abundance across the landscape. Most species respond to the favorable or unfavorable distribution of environmental characteristics, thereby creating patterns or gradients in species abundance. These patterns of abundance can be used to improve the reliability (precision) of surveys through stratification. The goal of stratification is to produce survey areas with similar levels of abundance so that variance among sample plots within each strata is minimized. However, it is not commonly recognized that the estimated abundance of animals after corrections for detection probability is used to calculate the variance within each strata. Assignments of plots to different strata must be based on how detection will affect the estimate of abundance on each plot and is especially important when habitat characteristics, which may affect detection probabilities, are used to define different strata. Usually at least three to six strata are defined (Eberhardt and Thomas 1991); a higher number of strata levels can facilitate reliable population estimates. For most wildlife surveys, the a priori knowledge of population distribution is too coarse to define more than three strata. On a landscape-scale, predictive habitat or environmental characteristics may be useful for identifying potential strata (Ratti and Garton 1994). The revision of strata boundaries or even levels as more knowledge is gained about the actual species distribution patterns is not unusual. Similar to the problem of population boundaries, highly mobile species may require a flexible annual adjustment of strata boundaries, depending on changing distribution patterns.

If populations are distributed randomly or information on population distribution is limited, a simple random sample may be the best choice. This approach requires that every sample plot in the population has an equal chance of being selected and that the procedure for selecting plots must truly be random (Ratti and Garton 1994). A more general form of random sampling arises when sample plots are selected with unequal probabilities. This usually occurs when plots have different sizes and random coordinates are used to select plots. In this case, each plot has a sampling probability proportional to the size of the plot (PPS sampling).

When population abundance follows well-defined gradients, other survey designs, including systematic sampling, may be more appropriate. Systematic sampling also may be a useful survey design when the objective is to determine the pattern of abundance. For randomly distributed populations, systematic sampling may provide estimates that are similar to a simple random sample (Ratti and Garton 1994). Other survey designs such as cluster sampling may be useful when the logistics of estimating abundance make travel between survey units expensive and when the population densities on adjacent sample plots are heterogeneous. If densities on adjacent plots are similar, cluster sampling will not increase the precision of the population estimate, compared to a simple random sample. Additional information on multi-stage sampling for wildlife studies can be found in Bart and Notz (1994) and more advanced statistical details are available in Cochran (1977) and Scheaffer et al. (1990).
Abundance on Sample Plots

The fundamental problem with determining animal abundance, even in relatively small areas, is that many animals will not be detected. In many circumstances, there is clear evidence that a large portion of the population will not be detected even with the most sophisticated methods (Caughley 1977: 35). Our inability to determine the actual number of animals in a particular area has given rise to numerous survey methods to estimate abundance when only a portion of the animals actually are observed. These methods include a plethora of popular survey techniques, such as capture-recapture, line transect, point counts, aerial surveys and catch-effort (Bibby et al. 1992, Lancia et al. 1994). This variety of survey techniques has developed to accommodate differences in species biology and behavior, habitats used, logistic considerations, seasons, sample plot size, and even researcher preferences. Further complications can arise when animals occur in groups because an assumption for many survey techniques is that each animal is observed independently. Group size also may be confounded with detection probabilities (Cook and Martin 1974, Samuel and Pollock 1981, Drummer and McDonald 1987, Samuel et al. 1987), producing biased population estimates when this effect is not considered. Few survey techniques currently permit the estimation of detection probabilities of groups of animals.

In general, most survey techniques attempt to determine the probability of detecting animals in an area and convert this probability and the number of observed animals into an estimate of actual abundance. Ideally, the survey design accommodates a variety of methods for determining animal detection probabilities on sample plots. Surveys where more than a small portion of the animals are undetected cannot provide estimates of abundance, unless detection probabilities are determined. Techniques that do not account for undetected animals should be considered only indices of abundance.

Horvitz-Thompson Population Estimator

In the standard sample survey methods, the probabilities of selecting sample plots must be predetermined. This implies that an exhaustive sampling frame of non-overlapping units can be listed and randomly selected (see selection schemes above) with know probabilities for each plot. For most wildlife surveys, this requirement may be met for sample plots of land that can be completely and uniquely identified. In contrast, a sampling frame of individual animals cannot be developed without a priori knowledge of the number of animals on each sample plot. To overcome the requirement of a sampling frame for animals, Steinhorstad Samuel (1989) and Samuel et al. (1992) developed a modified Horvitz-Thompson sample survey estimator of animal abundance that incorporates the probability of detecting animals during aerial surveys. In the original development, the term “sighting probability” is used for aerial surveys; however, this approach applies to the general probability of detecting animals on a sampled plot. By using this general approach to animal detection, the modified Horvitz-Thompson method provides a comprehensive framework for the design of large-scale abundance surveys. The population estimator (Steinhorstad and Samuel 1989) uses the detection probabilities to provide an unbiased estimate of abundance. The general abundance estimator is:
where

\[ t = \sum_{k=1}^{t} \frac{1}{p_k} \sum_{i=1}^{n_k} \frac{m_{ik}}{\pi_{ik}}, \]

\( t \) = the estimated total population,
\( p_k \) = the probability of selecting the \( k \)th sample plot,
\( n_k \) = the number of groups (\( \geq 1 \) animal) detected on sample plot \( k \),
\( m_{ik} \) = the number of animals in the \( i \)th detected group on sample plot \( k \), and
\( \pi_{ik} \) = the probability of detecting group \( m_{ik} \) during the survey.

In this general form, the Horvitz-Thompson estimator allows for unique probabilities of sampling each plot and different probabilities of detecting each animal (or group) on a sample plot. The number of detected animals are adjusted by the detection probability and the probability of sampling a plot to produce an estimated total abundance. Lancia et al. (1994) presented a simplified version of Equation 1 and discussed adjustments for the detection probability and the sampling proportion.

The Horvitz-Thompson estimator incorporates three sources of survey error (Steinhorst and Samuel 1989, Samuel et al. 1922) from not surveying all the sample plots, not detecting all animals on a sample plot and from estimating the probability of detecting animals. The general equation for the variance is:

\[ S_t^2 = S_{D_t}^2 + S_{V_t}^2 + S_{\pi_t}^2 \]

where

\( S_t^2 \) = the variance of the estimated population,
\( S_{D_t}^2 \) = variance attributed to the sampling design,
\( S_{V_t}^2 \) = variance attributed to not detecting all animals (visibility), and
\( S_{\pi_t}^2 \) = variance attributed to estimating the probability of detecting animals.

The variation in spatial sampling (\( S_{D_t}^2 \)) often is the largest portion of the total variance (\( S_t^2 \)). In the Horvitz-Thompson approach, any sampling design for plots can be accommodated, but designs that reduce the variability in spatial sampling (e.g., stratified sampling) are more efficient because they provide more precise estimates of the population. Spatial variation also can be reduced by increasing the proportion of sample plots. In a similar manner, population variance (\( S_t^2 \)) can be reduced with survey techniques that maximize the probability of detecting animals (reducing \( S_{V_t}^2 \)) and minimize the variation from detection probabilities (\( S_{\pi_t}^2 \)).

In the Horvitz-Thompson method, variance for the probability of detecting animals (\( S_{\pi_t}^2 \)) must be based on the model for estimating detection probability. This method allows flexible estimators with separate detection probabilities for each observed animal; however, such heterogeneity may increase variance in the \( S_{\pi_t}^2 \) component and in the population estimate. A variety of survey techniques for estimating detection probabilities are available to biologists and include line transect (Burnham et al. 1980), capture-recapture (Otis et al. 1978), circular plots (Reynolds et al. 1980), visibility models (Samuel et al. 1987, Otten et al. 1993), catch-effort (Alho 1992) and other approaches (see Lancia et al. 1994). Care should be used in selecting the most efficient method(s) for detecting animals on the selected plots. Improved precision can be achieved with survey methods that can incorporate homogeneous detection probabilities of animals in a sample plot or, better yet, of animals in many sample plots. The need to improve efficiency in estimating detection probabilities
was a principal motivation for developing general visibility models for elk (*Cervus elaphus*) surveys (Samuel et al. 1987).

Some survey techniques evaluate heterogeneity in detection probabilities during surveys on single sample plots (e.g., heterogeneous capture-recapture models). However, methods for evaluating and combining detection probabilities across multiple plots have not received much attention. One exception is the development of statistical methods for testing capture-recapture models among different populations (Skalski and Robson 1992). These tests also may be applicable to testing capture-recapture model similarity among sample plots. When models among plots are similar, more precise detection probabilities can be estimated by pooling results across plots for more precise population estimates. Similar improvements may be achieved with general methods to model capture-recapture probabilities (Alho 1990), catch-effort models (Alho 1992) or line transect methods (Burnham et al. 1980). Whatever approach is used, alternative survey techniques and detection probability estimates must be considered thoroughly during planning and analyzing large-scale abundance surveys. Special care also must be given to ensure that the assumptions (e.g., closed population, homogeneous detection probabilities, tag loss, etc.) for the selected survey technique can be met (Seber 1982). If detection probabilities cannot be determined in a timely manner (violating the closure assumption), open-population models (Pollock et al. 1990, Lancia et al. 1994) may have to be used. In the latter case, models that incorporate movement between sample plots (Hestbeck et al. 1991) also should be considered.

**Survey Examples**

In this section, we provide brief examples of some of the problems that may be encountered in large-scale surveys. We use elk population estimates to illustrate the importance of spatial and temporal variation in detection probability. Preliminary results from Canada goose surveys are used to illustrate some of the recursive aspects of survey design. Experiments on detection probability for duck surveys are used to speculate about the effects of animal behavior on survey results.

Elk populations have been monitored in portions of northcentral Idaho by helicopter survey during the last 10 years. Detection probabilities have been estimated with a visibility model (Samuel et al. 1987), with additional refinements as further data were collected (E. O. Garton unpublished data). Average visibility rates of bull and cow elk have differed (Samuel et al. 1992), primarily because bulls occur in smaller groups and more dense cover that makes them less visible than cows during aerial surveys. Incorporating heterogeneous visibility based on group size and vegetation allowed us to more accurately assess the total bull population and bull:cow ratios for improved herd management. In addition, winter conditions have varied considerably during the decade of conducting surveys. In particular, annual changes in snow conditions influenced the spatial distribution, habitat use and grouping behavior of elk in the survey area. During mild winters, animals are more dispersed, in smaller groups and in denser vegetation. These annual changes influenced the average visibility of animals. The spatial and temporal changes in detection probabilities are beyond the control of wildlife biologists and emphasize the danger in assuming a constant rate of detection. Use of an average detection probability would have severely biased
population estimates of bulls and decreased the probabilities of detecting population changes from elk harvest and habitat management.

At one time, giant Canada geese (*Branta canadensis maxima*) were believed to be extinct (Bellrose 1976). However, the race was rediscovered and increased under protection, propagation and vigorous transplant programs. Recently, aerial surveys were initiated to assess the population of these birds in the Mississippi Flyway, where they have become a nuisance in some locations. Intensive helicopter surveys of sample plots were conducted during the nesting season to maximize detection of breeding pairs, nests and nonbreeding groups. Initially, 1-square mile (2.59 km²) sample plots were surveyed in a stratified random design and random plots that did not contain viable goose habitat (absence of water on aerial photos) were not sampled. Subsequent survey refinements were attempts to reduce population variation with 2.25-square mile (5.83 km²) sample plots and reduce helicopter transport costs by sampling additional plots from a surrounding cluster. Preliminary results indicate the number of geese are more consistent on larger plots than on smaller plots. However, cluster sampling has not proved effective because the densities of geese on sample plots in the surrounding area are similar. Thus, sampling nearby areas provides little new information about goose abundance. Results from the 1993 survey indicate that ≥ 800,000 giant Canada geese now are present in the Mississippi Flyway. Further survey refinements are needed to improve the precision of population estimates, determine detection probability, and improve the efficiency of the survey design and conduct.

Smith (1993) recently conducted experiments with decoys to estimate duck detection probabilities during helicopter and fixed-wing aerial surveys. They concluded that visibility varied by habitat characteristics, distance from the transect and group size. From simulations, they concluded that changes in habitat-use patterns could produce large changes in overall visibility and confound population monitoring.

**Future Needs**

Little attention has been given to either the practical problems of developing large-scale surveys of wildlife species or the unique statistical problems associated with wildlife population estimation. In general, practical and theoretical work are needed in at least three areas. First, new survey techniques or modifications are needed to assess spatial heterogeneity in detection probabilities. These methods should incorporate testing for heterogeneous detection and the means to efficiently combine detection probabilities across sample plots. Comprehensive development of statistical methods may be difficult because many different survey techniques currently are in use; however, for the capture-recapture techniques, model tests among populations (Skalski and Robson 1992) may be adaptable to testing among sample plots.

Statistical procedures for selecting the most appropriate model from the detection data have received considerable attention (Burnham and Anderson 1992). However, the practical effects of different detection models need further consideration in the context of population estimation. In general, detection models with more heterogeneity produce less biased but more variable population estimates. Although completely unbiased estimates of wildlife populations may be impractical, decision rules are needed to evaluate the relative merit of biased, more precise estimates compared with less biased, less precise estimates. One possible approach is to compare the
population mean square error (MSE = bias² + variance) of different detection models. Because MSE consists of variance and bias², comparisons can be made among uncorrected population estimates and estimates corrected for different amounts of detection heterogeneity (Figure 1). These MSEs may be scaled by estimated population size (e.g., CVs) to standardize the comparisons. In addition, biologists should consider the importance of improving accuracy and precision by devoting more resources to increasing detection and estimating detection probabilities.

A third potential area for improvement in wildlife surveys is the development of predictive associations between animal abundance and environmental characteristics. On a landscape-scale, species abundance consistently may be related to particular habitat characteristics that are favorable to the species. Potential relationships between landscape patterns and population abundance could be evaluated on a portion of the sample plots. Consistent relationships could be used to develop regression methods to predict abundance on sample plots and to incorporate predictions into an overall population estimate. Data from large-scale geographic information systems should be useful for investigating landscape patterns (Turner 1990) and evaluating species relationships (Palmeirim 1988). Large-scale approaches based on techniques such as cokriging (Stein and Corsten 1991) also deserve investigation.

Summary

Large-scale surveys to estimate animal abundance can be useful for monitoring population status and trends, for measuring responses to management or environmental alterations, and for testing ecological hypotheses about abundance. However, large-scale surveys may be expensive and logistically complex. To ensure resources are not wasted on unattainable targets, the goals and uses of each survey should be specified carefully and alternative methods for addressing these objectives always should be considered. During survey design, the importance of each survey error component (spatial design, proportion of detected animals, precision in detection) should be considered carefully to produce a complete statistically based survey. Failure to address these three survey components may produce population estimates that are inaccurate (biased low), have unrealistic precision (too precise) and do not satisfactorily meet the survey objectives. Optimum survey design requires trade-offs in these sources of error relative to the costs of sampling plots and detecting animals on plots, considerations that are specific to the spatial logistics and survey methods. The Horvitz-Thompson estimators provide a comprehensive framework for considering all three survey components during the design and analysis of large-scale wildlife surveys.

Problems of spatial and temporal (especially survey to survey) heterogeneity in detection probabilities have received little consideration, but failure to account for heterogeneity produces biased population estimates. The goal of producing unbiased population estimates is in conflict with the increased variation from heterogeneous detection in the population estimate. One solution to this conflict is to use an MSE-based approach to achieve a balance between bias reduction and increased variation. Further research is needed to develop methods that address spatial heterogeneity in detection, evaluate the effects of temporal heterogeneity on survey objectives and optimize decisions related to survey bias and variance.
Figure 1. Comparison of hypothetical population size estimates of a true population of 100 animals. Estimate $P_1 (t = 70, S^2_t = 196)$, which is significantly lower than the true population, includes sampling variance ($S^2_{b_1}$) but ignores variation related to detecting animals. Estimate $P_2 (t = 90, S^2_t = 400)$ includes sampling variance and variance components related to animal detection ($S^2_{v_1}$ and $S^2_{m_1}$). Although the two estimates are not significantly different, mean square error is greater for $P_1$ (bias$^2 + S^2_t = 1,096$) than for $P_2$ (500). With the mean square error approach, correction for animal detection provides improved accuracy with undue sacrifice of precision.

Finally, managers and researchers involved in the survey design process must realize that obtaining the best survey results requires an interactive and recursive process of survey design, execution, analysis and redesign. Survey refinements will be possible as further knowledge is gained on the actual abundance and distribution of the population and on the most efficient techniques for detection animals.

References


