

pointed out, however, biological populations and communities change in space and time in such a complex manner that the normal application of statistics to wildlife studies is very difficult. Nevertheless, the following discussion includes sampling schemes that can reasonably be expected to meet these assumptions, or that allow for uncertain population boundaries and selection of units of study with unequal probabilities of inclusion.

3.2 Simple Sampling Designs

Wildlife studies are limited by fundamental principles of inferential statistics when using sample survey data to make predictions about the population of interest. Within the population or study area boundaries, statistical inference is limited by the protocol by which study

sites and/or study specimens are selected. This limitation also extends to the protocol used for subsampling of units from sites and to the standard operating procedures (SOPs) by which observations are made or measurements are taken. Sampling plans can be arranged in four basic categories (Gilbert 1987): (1) haphazard sampling, (2) judgment sampling, (3) search sampling, and (4) probability sampling.

Sampling plans that are most likely to be used during observational studies are discussed below. Several of the more common designs are illustrated in Figure 3.1. Other common variations of probability sampling are discussed by Gilbert (1987, pp. 19–23), Johnson et al. (1989), and Gilbert and Simpson (1992).

3.2.1 Haphazard Sampling

Haphazard sampling is a common approach in wildlife studies. The use of the approach is

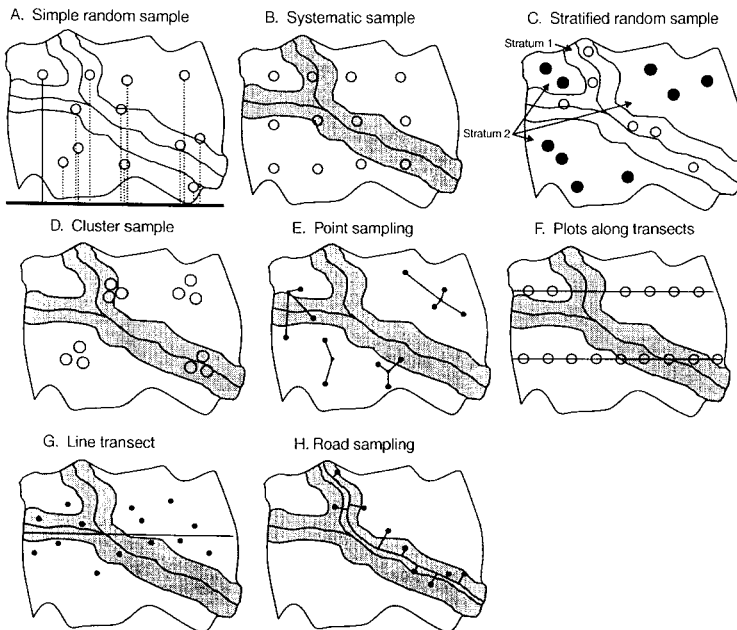


FIGURE 3.1. Typical sampling plans that are most likely to be used during observational studies, including judgment sampling (E) and haphazard sampling (H). (From Ratti and Garton 1994.)

often justified because of time, budget, and logistics. However, Gilbert (1987, p. 19) noted that:

Haphazard sampling embodies the philosophy of "any sampling location will do." This attitude encourages taking samples at convenient locations (say near the road) or times, which can lead to biased estimates of means and other population characteristics. Haphazard sampling is appropriate if the target population is completely homogeneous. . . . This assumption is highly suspect in most wildlife studies.

Certain kinds of surveys, such as report card harvest surveys, may have an element of haphazard sampling in them if the sample is self-selected by individuals volunteering to complete the survey. Haphazard sampling has little role to play in providing data for statistical inferences, because results are not repeatable. Information from haphazard sampling may be appropriate for preliminary reconnaissance of an area, but the information can only be used in making deductive arguments based on professional judgment. Examples of some common studies using haphazard sampling include roadside surveys for big game sex and age ratios, studies of stream flow using gauges permanently stationed at public access points along rivers, studies of contaminants at public access points, study areas adjacent to permanent field stations, and so on.

3.2.2 Judgment Sampling

Judgment sampling is also very common in wildlife studies (see Fig. 3.1H). This form of sampling is based on the presumption that the wildlife scientist can select studies representative of the study area or population. According to (Gilbert 1987) judgment sampling means subjective selection of population units by an individual (the researcher) resulting in the following outcome:

If the [researcher] is sufficiently knowledgeable, judgment can result in accurate estimates of population parameters such as means and totals even if all population units cannot be visually assessed. But it is difficult to measure the accuracy of the estimated parameters. Thus, subjective sampling

can be accurate, but the degree of accuracy is difficult to quantify.

As in haphazard sampling, judgment sampling may be appropriate for preliminary reconnaissance of an area, but is of little use in providing data for statistical inferences, because results are not repeatable. Judgment sampling may have a role to play in understanding the mechanisms in force in a biological system. For example, several study areas may be selected to investigate the magnitude and duration of an environmental impact or the effect of some management action under a specific set of conditions. Judgment sampling can also be used to develop data for models of natural systems (see capture-recapture model discussion later in this chapter). However, statistical inferences from sites selected for study are strictly to the study sites selected and any inference beyond those sites is deductive, depending on the professional judgment of the individual making the selection and the rules by which the sites are selected. For example, judgment sampling is used in surveys for waterfowl abundance by the states and the U.S. Fish and Wildlife Service where surveys are restricted to known nesting and loafing areas.

3.2.3 Search Sampling

Search sampling requires historical knowledge or data indicating where the resources of interest exist. For example, a study of the effect of oil on nesting colonies of sea birds might be limited to known colonies; the search for eagle nests might be limited to areas with suitable nesting habitat; or, factors causing bird fatalities in a wind plant might be limited to the portion of the wind plant where bird use is common. Searching for *hot spots* when investigating environmental pollution, discussed more fully under the cost-cutting procedures section below, is a form of search sampling. The value of data resulting from this procedure depends on the accuracy of the information guiding where and when to search. The procedure is most effective when the search is guided by data generated by long-term inventories of the resource of interest.

3.2.4 Probability Sampling

Probability sampling is the selection of samples from a group of sample units, each with a known likelihood of being chosen (probability of selection). If a sample is selected with a known probability then the appropriate sampling theory can be applied to the data (Krebs 1989). *Randomization* is necessary to make probability or confidence statements concerning the magnitude and/or duration of an effect of a treatment or impact (Johnson et al. 1989). Examples of random sampling plans include *simple random sampling* (random sampling, see Fig. 3.1A), *stratified random sampling* (stratified sampling, see Fig. 3.1C), *random start systematic sampling* (systematic sampling, see Fig. 3.1B), and *sequential random sampling* (sequential sampling). These sampling plans (and others, especially for mobile animal populations) can be combined or extended to give a large number of possibilities. As these combinations become more complicated, additional guidance from a statistician may be necessary to select the best design and the appropriate analysis method (Johnson et al. 1989).

3.2.4.1 Simple Random Sampling

Simple random sampling requires that the location of each sample site (unit) be selected independently of all other sites (units). Typically in ecology studies, a given unit appears at most once in the sample when sampling *without replacement* (Krebs 1989). Samples can be replaced after measurements are taken so that sampling is *with replacement* but sampling without replacement results in a more precise estimate (Caughley 1977).

A simple random sample may be obtained by following the basic steps in the following list:

1. The population of sampling units is assumed to be finite.
2. Units selected in the sample can be located and the measurement of the attribute of interest is possible. Also, the error in measuring the attribute should be small compared with the differences in the attribute from unit to unit.

3. The study region or the study period, also known as the sampling frame, must be completely covered by distinct and nonoverlapping sampling units.

4. Sampling units need not be of equal size nor selected with equal probability, but differences in size and selection probability make parameter estimation formulas much more complicated.

5. Sample units are normally sampled without replacement.

Random sampling plans have straightforward mathematical properties, but random locations are often more clumped and patchy than expected. In studies with small sample sizes, which are common in wildlife studies, entire regions of interest may be underrepresented or overrepresented. Thus, random sampling is not always the best procedure. Random sampling should be used only if the area of interest is very homogeneous with respect to the variables and covariates of interest. Because this is seldom the case, researchers should try to avoid relying solely on simple random sampling.

3.2.4.2 Stratified Sampling

Stratified sampling may be used to increase the likelihood that the sampling effort will be spread out over important subdivisions or strata of the study area, population, or study period (Fig. 3.2). Following stratification, units within strata are selected for study, usually by a random or systematic process. Similarly, units might also be stratified for subsampling. Ideally, strata should be homogeneous with respect to the variable of interest itself (e.g., animal density), but in practice, stratification is usually on variables that hopefully are highly correlated with the variable. For example, when estimating the density of deer, the wildlife biologist might stratify the study area into regions of high, medium, and low forest cover and sample each independently.

Strata must not overlap, all areas of interest must be included, and study sites (sampling units) must not belong to more than one stratum. Statistical inferences cannot be

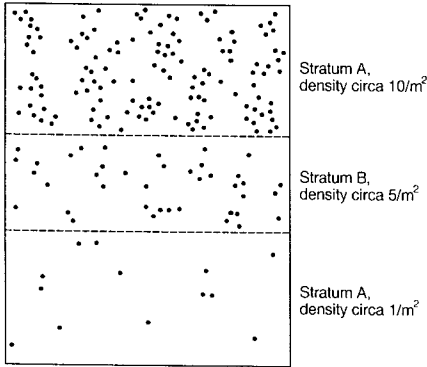


FIGURE 3.2. Illustration of three sampling strata with sampling effort spread out over important subdivisions or strata of the study area based on density of organisms. (From Krebs 1989.)

drawn toward differences in variables for any portion of strata unavailable for sampling. It may be possible to make professional judgments toward those areas, but conclusions will be made without the aid of statistical inference.

Stratification often will be used to estimate parameters within strata and for contrasting parameters among strata. For example, it may be of interest to investigate the impacts of a wind plant in different vegetation types (a potential stratification) where the objective is to make statistical inference to each vegetation type within the wind plant. This type of analysis is referred to as using "strata as domains of study... in which the primary purpose is to make comparisons between different strata" (Cochran 1977, p. 140). In this situation, the formulas for analysis and for allocation of sampling effort (Cochran 1977, pp. 140–141) are quite different from formulas appearing in introductory texts such as Scheaffer et al. (1990), where the standard objective is to minimize the variance of summary statistics for all strata combined (e.g., the entire wind plant).

Stratification offers several benefits to the researcher. If units are truly more homogeneous within strata, then the estimate of the

overall population mean will have a smaller standard error than can be obtained from a simple random sample of the same size. Having independent population estimates for each stratum may also be advantageous. Finally, stratification allows sampling of different parts of a population in different ways, making some cost savings possible.

In summary, a primary objective of stratification is improved precision based on optimal allocation of sampling effort into more homogeneous strata. In practice, it may be possible to create homogeneous strata with respect to one or a few primary indicators, but there are often many indicators measured, and it is not likely that the units within strata will be homogeneous for all of them. For example, one could stratify a study area based on vegetation and find that the stratification works well for indicators of effect associated with trees. But because of management (e.g., grazing), the grass understory might be completely different and make the stratification unsatisfactory for indicators of effect measured in the understory. Differences in variance among strata for the primary indicators may not occur or may not be substantially better than random sampling.

Factors used to stratify an area should be based on the spatial location of regions where the population is expected to be relatively homogeneous, the size of sampling units, and the ease of identifying strata boundaries. Strata should be relatively easy to identify by the methods that will be used to select strata and study sites within strata, and of obvious biological significance for the variables of interest. Typical strata for wildlife studies include physiography/topography, vegetation, land use, season of use, management action of interest, and so on. Spatial stratification can be valuable when few sites (units) are misclassified and the study is of relatively short duration. The length of study is of particular significance when stratifying on a variable that will likely change with time (e.g., land use). Stratified sampling works best when applied to short-term studies, reducing the likelihood that strata boundaries will change.

Samples can be allocated to strata in proportion to strata size or through some optimal allocation process. When using the stratification with proportional allocation, the samples are self-weighting in that estimates of the overall mean and proportion are the same as for estimates of these parameters from simple random sample. The stratified sampling variance estimate, on the other hand, is different from the simple random sample. Although proportional allocation is straightforward, it may not make the most efficient use of time and budget. If it is known that the variance within individual strata differ, then samples can be allocated to optimize sample size. Optimizing sample distribution can be used to either achieve a given level of precision for overall population estimates, or to gain the maximum precision at least cost. This approach and the method for estimating sample size are described in Cochran (1977). If variances and costs are about the same in all strata, then proportional allocation is the preferred approach to sample allocation. Krebs (1989) offered the following rules of thumb in stratified sampling: In a given stratum, take a large sample if

1. The stratum is larger.
2. The stratum is more variable internally.
3. Sampling is cheaper in the stratum.

Stratification has some inherent problems. In any stratification scheme, some potential study sites will be misclassified in the original classification (e.g., a dark area classified as a pond on the aerial photo was actually a parking lot). Stratification is often based on maps that are inaccurate. Map problems result in misclassified sites that have no chance of selection in the field SOPs used by investigations. Misclassified portions of the study area can be adjusted once the errors have been found but the analysis of data becomes much more complicated, primarily because of differences in the probability of selection of study units in the misclassified portions of the study area.

Short-term studies usually lead to additional research questions and longer-term study and more complicated analysis of sub-

populations (Cochran 1977, pp. 142–144) that cross strata boundaries and strata may change (e.g., the cornfield has become grassland). In long-term studies, investigators are likely to be happiest with the stratification procedure at the beginning of the study. Benefits of stratification on characteristics such as vegetative cover type, density of prey items, land use, etc. diminish quickly with time as these phenomena change.

A fundamental problem is that strata normally are of unequal sizes and, thus, units from different strata have different weights (importance values) in any overall analysis. The formulas for computing an overall mean and its standard error based on stratified sampling are relatively complex (Cochran 1977, pp. 87–95). Formulas for the analysis of subpopulations (subunits of a study area) that belong to more than one stratum (Cochran 1977, pp. 142–144) are even more complex for basic statistics such as means and totals. The influence of these unequal weights in subpopulations is often unknown for many analyses, such as ordination or multidimensional scaling, and is usually ignored and assumed to be equal.

It might be necessary to stratify with little prior knowledge of the study area; but if possible, stratification should be limited to geographic features based on excellent maps, and the minimum number of strata should be used (preferably no more than three or four). Covariates that are potentially correlated with the magnitude and duration of a potential treatment effect should be measured on the study sites (or on subsampling units within sites). Some analyses, such as ordination and multidimensional scaling, may require additional original mathematical research for justification of their use.

3.2.4.3 Systematic Sampling

Systematic sampling is possible when a population can be listed in ascending or descending order of some characteristic (e.g., group size) or occupies a well-defined spatial area. A systematic sample from an ordered list would consist of sampling every k th item in the list.

A spatial sample typically utilizes a systematic grid of points. Systematic sampling distributes the locations of samples (units) uniformly through the list or over the area (site). A relatively unbiased systematic sample can be achieved by adopting a random starting rule. That is, the first point in a systematic sample is placed at a truly random point with the remainder of the points located in reference to this *random* point.

Systematic sampling is attractive for two primary reasons:

1. It is easier than random sampling.
2. A systematic sample may appear more representative and thus more precise than a random sample, because it gives uniform coverage of the whole of the population of interest.

Mathematical properties of systematic samples are not as straightforward as for random sampling, but the statistical precision generally is better (Scheaffer et al. 1990).

Systematic sampling has been criticized for two basic reasons. First, the arrangement of points may follow some unknown cyclic pattern in the response variable. This problem is addressed a great deal in theory, but is seldom a problem in practice. If there are known cyclic patterns in the area of interest, the patterns should be used to advantage to design a better systematic sampling plan. For example, in a study of the cumulative effects of proposed wind energy development on passerines and shore birds in the Buffalo Ridge area of southwestern Minnesota, Strickland et al. (1996) implemented a grid of sampling points resulting in observations at varying distances from the intersection of roads laid out on section lines.

Second, in classical *finite sampling theory* (Cochran 1977), variation is assessed in terms of how much the result might change if a different random starting point could be selected for the uniform pattern. For a single uniform grid of sampling points (or a single set of parallel lines) this is impossible, and thus variation cannot be estimated in the classical sense. Various model-based approximations have

been proposed for the elusive measure of variation in systematic sampling (Wolter 1984). Sampling variance can be estimated by replicating the systematic sample. For example, in a study requiring a 10% sample it would be possible to take multiple smaller samples (say a 1% sample repeated 10 times), each with a random starting point. Inference to the population mean and total can be made in the usual manner for simple random sampling.

Systematic sampling works very well in the following situations:

1. Analyses of *observational* data conducted as if random sampling had been conducted (effectively ignoring the potential correlation between neighboring locations in the uniform pattern of a systematic sample)
2. Encounter sampling with unequal probability (Overton et al. 1991, Otis et al. 1993)
3. The model-based analysis commonly known as *spatial statistics*, wherein models are proposed to estimate treatment effects using the correlation between neighboring units in the systematic grid (see, for example, kriging [Johnson et al. 1989, ch. 10])

The design and analysis in case 1, above, is often used in evaluation of indicators of a treatment response (e.g., change in density) in relatively small, homogeneous study areas or small study areas where a gradient is expected in measured values of the indicator across the area. Ignoring the potential correlation and continuing the analysis as if it is justified by random sampling can be defended (Gilbert and Simpson 1992), especially in situations where a conservative statistical analysis is desired (e.g., impact assessment). Estimates of variance treating the systematic sample as a random sample will tend to overestimate the true variance of the sample (Hurlbert 1984, Scheaffer et al. 1990, Thompson 1992). Thus, systematic sampling in relatively small impact assessment study areas following Gilbert and Simpson's (1992) formulas for analysis makes a great deal of sense. This applies whether

systematic sampling is applied to compare two areas (assessment and reference), the same area before and following the incident, or between strata of a stratified sample.

A uniform grid of points or parallel lines may not encounter rare units. To increase the likelihood of capturing some of these rare units, scientists may stratify the sample such that all units of each distinct type are joined together into strata and simple random samples are drawn from each stratum. Nevertheless, stratification works best if the study is short term, no units are misclassified, and no units change strata during the study. In longer-term studies, such as the U.S. Environmental Protection Agency's (EPA's) long-term Environmental Monitoring and Assessment Program (EMAP), as described by Overton et al. (1991), systematic sampling has been proposed to counter these problems. Even in longer studies, unequal probability sampling is almost inescapable, but the problems associated with misclassified units and units that change strata over time can largely be avoided.

For long-term monitoring studies, or when problems with misclassification and changes in land use are anticipated, one should consider systematic sampling strategies. As an example, in the case of the Buffalo Ridge site in Minnesota (Strickland et al. 1996), a stratification of the study and reference areas by vegetation type was considered. However, the two major vegetation types present on the study area were fallow lands, reserved from tillage under the federal Conservation Reserve Program (CRP), and lands being actively farmed. These vegetation types are both very influential on bird use and also likely to change within a year or so. Thus, stratification on vegetation might be desirable in year 1 of the study, only to be completely inappropriate in year 2 or 3. In anticipation of this problem, a systematic grid of points with a random starting point was established covering each study area. For a given year, bird use within vegetation types on the assessment and reference areas may be compared statistically as if the points were randomly located,

even if vegetation types change from year to year.

3.3 Single-Level Sampling

The simplest form of *probability sampling* is sampling at a single level or scale. That is, the study area is divided into a set of potential units from which a sample is taken. For example, a study area could be divided into a grid of sample plots all of the same size from which a simple random sample is drawn. The organisms of interest in each cell in the selected sample are then counted.

The above discussion deals primarily with the spatial and/or temporal allocation of study effort, including the location of study plots, the timing of surveys, and in plotless surveys the allocation human resources. In studies involving counts of animals and plants on study plots, the size and shape of plots are also important components of the study method. Field designs that promote similar detection probabilities of measured items result in comparisons with smaller sampling error Skalski et al. (1984). Plot size makes little difference if organisms are randomly or uniformly distributed throughout the study area, but the use of a larger number of smaller plots increases precision with aggregated organisms (Green 1979). Since aggregations of organisms are the norm in nature, it generally is better to use a larger number of smaller plots located throughout the study area or stratum.

Cost, logistics, the behavior of the organism being studied, and the patterning (Krebs 1989) of the organism will all influence plot size. Use of larger plots usually allows the researcher to cover more area at a lower unit cost (e.g., cost per hectare sampled). Plots can be so small that measurement error increases dramatically (e.g., too many decisions regarding whether the study subject is in or out of the plot) or the variance of the sample increases because the detection of the organism is rare, resulting in a data set with a lot of zeros. As a rule, the smallest plot size practical should be selected.