The Condor 90:275–278

MINIREVIEW

ON SAMPLE SIZES AND RELIABLE INFORMATION[1]

Michael L. Morrison
Department of Forestry and Resource Management, University of California, Berkeley, CA 94720

During the past decade of reading journals such as *The Condor, Auk, Wilson Bulletin, Journal of Wildlife Management,* and *Ecology,* I can recall few articles that provided statistical evaluation or even qualitative assertion to justify the number of samples used in the analyses presented. I suspect review of other journals would reveal similar results. What conclusions can be drawn regarding the validity of the results and discussions from such studies? Given that inadequate sample sizes will bias the results of even the most carefully designed and conducted study, I am afraid that the answer to this question will be a sad commentary on the state of most of our investigations.

Data, of course, provide the foundation for all ornithological knowledge. Data can range from the handwritten field notes of a competent observer to values obtained through the use of complex apparatus. Whether data are in a qualitative form such as "I saw species *A* foraging consistently higher than species *B*," or a quantitative form such as "We found a significant ($P <$ 0.05, *t*-test) difference between the foraging heights of species *A* ($\bar{x} = 6.3$, SD = 2.1 m) and species *B* ($\bar{x} = 20.2$, SD = 1.4 m)," we have been provided data on a biological phenomenon. But upon what are we to base our trust in the biological reality of these results? Using my simple two-species example, which result would you tend to trust? I suspect most people would initially rely on the quantitative data analyzed by the *t*-test (or other appropriate analysis). Unfortunately, there is absolutely no reason to place any more trust in the quantitative assessment over the more qualitative one. If I had added, "On 132 occasions I saw . . ." to the qualitative description, and $n = 10$ to the quantitative analysis, I would certainly put my trust in the former data set. This example may seem somewhat trivial, but it is meant to make a simple point: without both the presentation *and* justification of one's sample size, no data set has much, if any, validity. Of course, my comments here assume a proper sampling design. The sampling design is an early, critical step in any study and, along with the proper statistical analyses, must be of a proper form to answer the question at hand. Regardless of the size of the sample, improperly collected data may be of little use. The question of study design, is, of course, a complex topic that will not be developed in this short review. Many texts on the subject are available (e.g., Cochran 1977, Williams 1978, Green 1979, Kerlinger 1986, Kish 1987).

I have often heard the comment: "but my sample size must have been adequate. . . . I found a significant difference . . ." to justify the associated sample size in a study. Erroneous and often contradictory conclusions may be reached with variations in sample size, which may result in fluctuating alpha levels. The probability of committing a Type I error (α; rejection of a null hypothesis when it is actually true) is inversely related to the probability of committing a Type II error (β; failing to reject the null hypothesis when it is in fact false), for a given n. Lower probabilities of committing a Type I error are associated with higher probabilities of committing a Type II error, and the only way to minimize both types of error is to increase n. Thus, for a given α, larger sample sizes will result in statistical tests with greater power ($1 - \beta$) (see Zar 1984:43–45). In theory, then, we have a strong basis for increasing our sample sizes.

Further, the biological interpretation of the results can vary as sample sizes increase and the resulting responses of an animal change. For example, take the use of terminal buds by Chestnut-backed (*Parus rufescens*) and Mountain (*P. gambeli*) chickadees shown in Figure 1: different conclusions could be reached regarding the percent use of terminal buds by each species, and the overlap in use between the species, based on the sample size used. Results can differ dramatically with even small changes in sample sizes (compare results for the Mountain Chickadee for samples of 30, 40, and 50 individuals).

Many formulas are available in basic statistics texts for determining proper sample sizes (e.g., Cochran 1977, Williams 1978, Green 1979, Sokal and Rohlf 1981). Many of these formulas require, however, an estimate of the population variance. Of course, the population variance is seldom known. Alternatively, one can collect data sequentially, with data being evaluated by these formulas at each step, and a decision being made on the adequacy of the samples and possible need for more data collection. Such sequential sampling procedures have been discussed by several workers (e.g., Kuno 1969, 1972; see also Green 1979:126–136). Sequential sampling provides a valuable, although seldom used, method of determining proper sample sizes without gross over-sampling. (Sequential techniques for estimating necessary sample sizes, as used in this paper, should not be confused with sequential sampling for classifying populations into categories; see Waters 1955.)

---

[1] Received 2 December 1987. Final acceptance 16 December 1987.

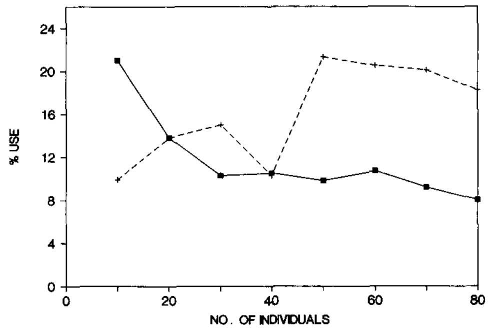

FIGURE 1. Influence of sample size on the observed use of terminal buds (coniferous and deciduous) by foraging Chestnut-backed (solid line) and Mountain (dashed line) chickadees in a mixed-conifer forest (Sierra Nevada, California) during September and October 1986; n = number of individuals observed (Brennan and Morrison, unpubl. data).

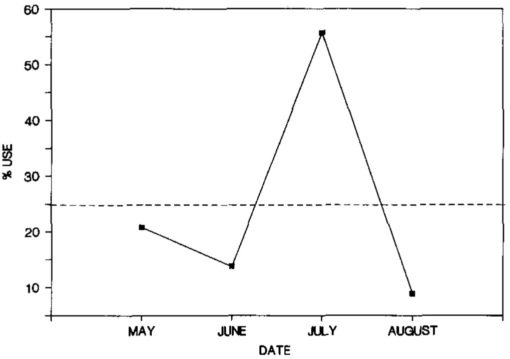

FIGURE 2. Use of white fir (*Abies concolor*) by foraging Mountain Chickadees in a mixed-conifer forest (Sierra Nevada, California) during 1986 (Brennan and Morrison, unpubl. data). Dashed line is mean of May–August.

Morrison (1984b) and Block et al. (1987) for habitat data and Morrison (1984a) for foraging data, used what we could call a posteriori (or "after the fact") sequential sampling to evaluate the adequacy of a large data set that was already collected. Here, an adequate sample was defined by the point at which the confidence interval and the point estimate showed little variation with increasing sample sizes. This examination could have been conducted throughout the collection process. These studies generally indicated that 40 or more sample plots or individuals—in the habitats studied—were required before reliable (i.e., stable) sample sizes resulted. For many variables, the n was much larger. Other workers (e.g., Johnson 1981, Carnes and Slade 1982) have reached similar conclusions using different methods. This does *not* mean that any study failing to meet these sample sizes is invalid; it simply indicates that many published studies should be suspect. This is especially true of multivariate analyses, in which violation of assumptions can markedly influence results (e.g., see Williams 1981, 1983). In discriminant function analysis, for example, the assumption of equality of variance-covariance matrices is usually violated in studies of avian habitat use. This does not mean the data are biologically meaningless, but without sample-size evaluations one simply does not know if the results are due to the biology of the animals or the inadequacy of the samples. As I found in an earlier study (Morrison 1984b), inadequate sample sizes may actually produce equal variance-covariance matrices.

Minimum sample sizes often vary among species being studied, and further, among variables and techniques used to study them. For example, Mosher et al. (1986) calculated minimum sample sizes separately for each species for each habitat variable in a study of raptor habitat models. Their criterion of acceptable sample size was that estimates were within 20% of the mean with 95% accuracy. They found wide variation among variables in the minimum sample sizes required, with the variance associated with some variables so high that adequate sample sizes were not realistic to achieve. For example, whereas only 28 plots were needed to determine precisely the number of trees/0.04 ha, 750 plots were needed to determine snag density.

In his text on sampling design, Green (1979:129) provides a brief overview of methods useful in determining the adequacy of sample sizes. He showed that ". . . for a wide range of field data the number of samples required to achieve a desired precision (D_o is independent of the . . . density, and is approximately equal to the inverse of the square of the desired precision." As an example, if the density must be estimated with a precision such that 0.95 confidence limits are ±20% of the mean, then $D_o \approx 0.10$ and $n \approx 100$. If ±40% is sufficient, then n drops to about 25. Green's example illustrates two key points. First, that the number of samples required is dependent upon the precision of the answer needed; a decision apparently seldom discussed in most papers. Second, that few studies reach an n of 100 (of anything), again severely questioning the validity of much published work.

Rare species or populations present special problems for study. Given our desire to obtain statistically independent samples, how does one collect adequate data on a population with few and possibly widely scattered individuals? In essence, one simply does not do so. In such cases, stratification of observations by time, and to the extent possible, space, combined with a thorough analysis of the influence of sample size on results, will at least allow the worker to place the correct amount of confidence in her/his results. Cochran (1977:76–77) provides further discussion on sampling rare items.

Of course, adequate sampling in time and space is necessary for even the most abundant species. We know that birds dramatically alter their use of resources between seasons (e.g., Conner 1981, Hutto 1981, Morrison et al. 1985, Morrison and With 1987). Even *within* a season, we know that birds will change resource use. For example, we (Brennan and Morrison, unpubl. data) have found that chickadees alter their use of tree species within different periods of a year. Lumping data across the breeding period, for example—a practice often justified to achieve "adequate sampling"—clouds such a relationship as the plot in Figure 2 indicates. Here, the study population never performed the "average" use. Thus, the sampling period, as well as the size of the sample itself, becomes critical. The period is often measured in weeks or a few months, not seasons. Necessary sample sizes may also vary between periods within a year. When and where data are collected are usually considered "design" problems. It is clear, however, that adequate sample sizes must accompany implementation of any design.

Many studies, whether or not they contain adequate samples, are properly designed and executed. The important step usually omitted, however, is the exploratory stage of data analysis. As discussed by James and McCulloch (1985), the primary idea of exploratory analysis is to avoid premature refinement of a problem. James and McCulloch, however, did not explicitly include sample-size evaluations in their discussion. They implied such evaluations in their statements on the importance of violations of assumptions and the need for replication. The previous citations on sequential sampling relate to the idea of exploratory data analysis. I am certainly guilty of designing a study, collecting what I think are adequate amounts of data, and then publishing the results. Only later, after being inspired by the discussion of adequate sample sizes given by Johnson (1981), did I *go back* to the data set to see if indeed enough data had been collected (e.g., Morrison 1984a, 1984b).

In this review I am not calling for more rigorous hypothesis testing, experimentation, or even more "eloquent" statistical designs. Nor am I advocating the accumulation of huge natural history data sets collected without regard to specific questions; over-kill is not the answer. Large sample sizes are not good in and of themselves. They are advocated to avoid the probability of selecting samples that are not representative of the population from which they are drawn, and to allow the principle of randomization an opportunity to function (see Kerlinger 1986:119). What I am calling for is an evaluation and justification of the size and nature (i.e., meeting of assumptions) of the data analyzed. Hopefully, such an evaluation will begin during the collection phase. The ready availability of statistical packages on even small (e.g., micro-) computers renders evaluation of the influence of sample size on results of even complex analyses an easy process. I see little if any reason for the publication of most, if not all, scientific work that does not include an evaluation-justification phase that describes the rationale for the number of samples used.

I thank W. M. Block, L. A. Brennan, and K. A. With for helpful comments and suggestions on an earlier version of this review, M. L. Morton for the opportunity to express my views on sample sizes, and L. Merkle for manuscript preparation.

LITERATURE CITED

Block, W. M., K. A. With, and M. L. Morrison. 1987. On measuring bird habitat: influence of observer variability and sample size. Condor 89: 241–251.

Carnes, B. A., and N. A. Slade. 1982. Some comments on niche analysis in canonical space. Ecology 63:888–893.

Cochran, W. G. 1977. Sampling techniques. 3rd ed. John Wiley and Sons, New York.

Conner, R. N. 1981. Seasonal changes in woodpecker foraging patterns. Auk 98:562–570.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York.

Hutto, R. L. 1981. Temporal patterns of foraging activity in some wood warblers in relation to the availability of insect prey. Behav. Ecol. Sociobiol. 9:195–198.

James, F. C., and C. E. McCulloch. 1985. Data analyses and the design of experiments in ornithology, p. 1–63. *In* R. F. Johnston [ed.], Current Ornithology. Vol. 2. Plenum Press, New York.

Johnson, D. H. 1981. How to measure habitat—a statistical perspective, p. 53–57. *In* D. E. Capen [ed.], The use of multivariate statistics in studies of wildlife habitat. U.S.D.A. For. Serv. Gen. Tech. Rep. RM-87.

Kerlinger, F. N. 1986. Foundations of behavioral research. 3rd ed. Holt, Rinehart, and Winston, New York.

Kish, L. 1987. Statistical design for research. John Wiley and Sons, New York.

Kuno, E. 1969. A new method of sequential sampling to obtain estimates with a fixed level of precision. Res. Popul. Ecol. 11:127–136.

Kuno, E. 1972. Some notes on population estimation by sequential sampling. Res. Popul. Ecol. 14:58–73.

Morrison, M. L. 1984a. Influence of sample size and sampling design on analysis of avian foraging behavior. Condor 86:146–150.

Morrison, M. L. 1984b. Influence of sample size on discriminant function analyses of habitat use by birds. J. Field Ornithol. 55:330–335.

Morrison, M. L., and K. A. With. 1987. Interseasonal and intersexual resource partitioning in Hairy and White-headed woodpeckers. Auk 104:225–233.

Morrison, M. L., I. C. Timossi, K. A. With, and P. N. Manley. 1985. Use of tree species by forest birds during winter and summer. J. Wildl. Manage. 49:1098–1102.

Mosher, J. A., K. Titus, and M. R. Fuller. 1986. Developing a practical model to predict nesting habitat of woodland hawks, p. 31–35. *In* J. Verner, M. L. Morrison, and C. J. Ralph [eds.], Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Univ. Wisconsin Press, Madison.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry. 2nd ed. W. H. Freeman and Co., San Francisco.

Waters, W. E. 1955. Sequential sampling in forest insect surveys. For. Sci. 1:68–79.

Williams, B. 1978. A sampler on sampling. John Wiley and Sons, New York.

Williams, B. K. 1981. Discriminant analysis in wildlife research: theory and applications, p. 59–71. *In*

D. E. Capen [ed.], The use of multivariate statistics in studies of wildlife habitat. U.S.D.A. For. Serv. Gen. Tech. Rep. RM-87.

Williams, B. K. 1983. Some observations on the use of discriminant analysis in ecology. Ecology 64: 1283–1291.

Zar, J. H. 1984. Biostatistical analysis. 2nd ed. Prentice-Hall, Englewood Cliffs, NJ.

The Condor 90:278

NEWS AND NOTES

58TH ANNUAL MEETING OF THE COOPER ORNITHOLOGICAL SOCIETY

The 58th Annual Meeting will be held at the Asilomar Conference Center, Pacific Grove (Monterey), California, on 17–22 March 1988. Walter D. Koenig is chair of the Local Committee. A symposium on "Food Exploitation by Terrestrial Birds," sponsored by the C.O.S., Pacific Southwest Forest and Range Experiment Station (U.S.D.A. Forest Service), Western Foundation of Vertebrate Zoology, and the University of California–Berkeley (Dept. Forestry and Resource Management) has been organized. Approximately 65 abstracts have been accepted for presentation at the Symposium; contributors are from throughout North America and several foreign countries. For information contact Koenig (Hastings Reservation, Star Rt. Box 80, Carmel Valley, CA 93924; 408-659-2664) or Morrison (Dept. Forestry, University of California, Berkeley, CA 94720; 415-642-5344).

ANNUAL MEETING OF ASSOCIATION OF FIELD ORNITHOLOGISTS

The Association of Field Ornithologists (formerly NEBBA) will hold its annual meeting 13–15 May 1988 at the Vermont Institute of Natural Sciences, Woodstock, Vermont. Housing and the Saturday evening banquet will be at the nearby Kedron Valley Inn. The meeting will include invited and contributed papers, workshops and field trips. For information about the meeting, contact: Sarah B. Loughlin, AFO Local Committee Chair, Vermont Institute of Natural Science, Woodstock, VT 05091 (802/457-2779). For information about the scientific program, contact: Peter F. Cannell, Program Committee Chair, Division of Birds, NHB 116, Smithsonian Institution, Washington, DC 20560 (202/357-2334).

NORTHEAST RAPTOR MANAGEMENT SYMPOSIUM AND WORKSHOP

The Northeast Raptor Management Symposium and Workshop, hosted by the National Wildlife Federation's Institute for Wildlife Research, will be held 16–18 May 1988 at the Hotels at Syracuse Square, Syracuse, New York. The Symposium will feature technical papers on the status and management of northeastern raptors and land use issues which impact raptor populations. Interactive workshops will encourage participants to discuss raptor management issues in the region and develop management recommendations. For more information, contact the National Wildlife Federation, 1412 Sixteenth St., N.W., Washington, DC 20036-2266 or call (703) 790-4264.

TECHNICAL PAPERS SOUGHT BY THE NORTH AMERICAN BLUEBIRD SOCIETY

The North American Bluebird Society (NABS) is soliciting manuscripts for their Research Series. The NABS Research Series is intended to serve as an outlet for technical papers on any aspect of the biology and conservation of North American cavity nesting birds. Manuscripts will be peer reviewed and are not constrained by length. Accepted manuscripts will be published at no cost to the authors. Inquiries and manuscripts should be sent to : Jeffrey D. Brawn, Smithsonian Tropical Research Institute, APO, Miami, FL 34002-0011, USA.