

Jackknifing and Bootstrapping: Important "New" Statistical Techniques for Ornithologists

SCOTT M. LANYON¹

The purpose of this commentary is to inform ornithologists about two statistical tools: jackknifing (Bradley et al. 1984, Gibson et al. 1984, Lanyon 1985a, Lanyon and Lanyon 1986) and the logically similar bootstrapping. The following brief overview of these procedures supplements the review of data analysis by James and McCulloch (1985), which, for lack of examples in ornithology, did not discuss either jackknifing or bootstrapping. More complete descriptions of these techniques may be found elsewhere (for general introductions see Sokal and Rohlf 1981, Diaconis and Efron 1983, Efron and Gong 1983; and for technical discussions see Quenouille 1956; Tukey 1958; Miller 1964, 1974; Bissell and Ferguson 1975; Clarkson 1979; Efron 1979, 1982).

Jackknifing is new to ornithological research but was first described 30 years ago (Quenouille 1956). The delayed use of this technique, as well as of bootstrapping (Efron 1979), in ornithology can be attributed directly to the cost of computer time. Both approaches are iterative, requiring massive numbers of calculations, and consequently have been prohibitively expensive. The attractiveness of jackknifing and bootstrapping is that they provide investigators with an important, and previously unattainable, type of information: estimates of dispersion for statistics of unknown or poorly known distribution.

Measures of dispersion (e.g. standard deviation, standard error, kurtosis, range, etc.) are critical to studies that employ statistical procedures because these measures help indicate the accuracy of locational statistics (e.g. mean, mode, median). Without an estimate of the accuracy of sample statistics, both the investigator and the reader are at a disadvantage when interpreting data. The importance of dispersion measures is reflected by the fact that in the ornithological literature sample means are rarely reported without accompanying measures of dispersion. In volume 102 of *The Auk*, more than 75% of papers reporting means included measures of dispersion. The many other types of location statistics with measurements of dispersion, however, are rarely reported.

Genetic distance measures, indices of diversity, factor loadings in principal components analysis, discriminant coefficients in discriminant function analysis, and correlation and regression coefficients are examples of statistics of unknown or poorly known distribution, and, with very few exceptions, all are reported without measures of dispersion. The func-

tion of jackknifing and bootstrapping is to generate a distribution for such statistics that permits the calculation of dispersion estimates. The distribution is generated by resampling the original data set. The way in which the data set is resampled differentiates the two approaches.

In bootstrapping the original data set is sampled randomly but with replacement to produce "pseudoreplicate" data sets. Each pseudoreplicate consists of the same number of elements as the original data set but may not include all the original elements (some elements may appear more than once, others not at all). This resampling may be repeated thousands, even millions, of times, and each iteration produces a new pseudoreplicate from which statistics may be calculated. In contrast, jackknifing produces a limited number of pseudoreplicate data sets, each of which contains all but one of the original data elements. For a data set with 20 elements, 20 pseudoreplicate data sets will be generated, each lacking a different data element. Because jackknifing requires far fewer iterations, it is thought of as a means of approximating bootstrapping (Efron 1979).

For both approaches, a new estimate of the statistic is calculated for each pseudoreplicate data set, resulting in a set of estimates of a particular statistic. These values can then be used to estimate a location and dispersion for that statistic. For example, in an analysis of allelic frequencies, we can calculate a matrix of genetic distances between taxa for each pseudoreplicate. When all pseudoreplicates have been analyzed, there will exist an array of estimates for each pairwise distance. These estimates can be used to estimate a mean and standard deviation for the genetic distance between each pair of taxa. The resultant distance matrix is far more informative than the original because it includes an estimate of the accuracy of each pairwise distance.

The nature of the observations to be jackknifed or bootstrapped is an important consideration in these techniques. For most applications the individual data points are the elements that are manipulated. For some applications, however, valuable insights may be gained by resampling other types of data elements such as variables or taxa (Felsenstein 1985, Lanyon 1985b).

The use of jackknifing and bootstrapping should enable investigators to learn more about their data than was previously possible because of the information on the dispersion of sample statistics. I hasten to add that this situation does not imply that investigators will be able to *conclude* more from their data. In fact, this is almost certainly not the case. Use of jackknifing and bootstrapping has already demon-

¹ Division of Birds, Field Museum of Natural History, Roosevelt Road at Lake Shore Drive, Chicago, Illinois 60605 USA.

strated that some investigators overanalyze their data. Rather, these techniques provide a means of determining when to *stop* analyzing data. That is, investigators can identify the point in a statistical analysis beyond which the results are more strongly influenced by sampling error than by the underlying biological phenomena.

Jackknifing and bootstrapping should be used in all cases where a statistic is generated and the distribution for that statistic is unknown or too complicated for the more conventional methods of dispersion estimation. Just as means reported without estimates of dispersion are unacceptable in scientific publications, so too should other statistics be unacceptable when presented alone. This is especially true in an era when original data are rarely published. In the absence of the original data, the reader has no way to estimate the accuracy of the reported statistics unless measures of dispersion are included. In the paragraphs to follow I present three examples of appropriate applications for these statistical techniques.

A common analytical approach in biology is principal components analysis (PCA). In PCA a set of multivariate axes (factors) is determined that minimizes the dispersion of samples around the axes. The statistics produced by this analysis, factor loadings, describe the correlation between the variables and these axes. Before the use of jackknifing and bootstrapping there was no reliable way to estimate the accuracy of these factor loadings. Yet many published studies have depended on the relative magnitude of these values. Now it is possible to obtain not only loadings for each variable, but an estimate of their dispersion as well. Through the application of these techniques it is possible to limit discussions to only those variables characterized by factor loadings that are significantly different from zero (Diaconis and Efron 1983, Gibson et al. 1984).

A second, broad category of statistics that would benefit from the determination of measures of dispersion is indices. Indices are extremely common in the biological literature: indices of genetic distance, diversity, niche breadth, etc. In all cases it has been impossible to determine how representative these values are of the population, species, or community under investigation. Once again these resampling techniques provide a means to answer this question. By estimating the dispersion for each index, it is possible to determine which indices are significantly different and which are not. Thus, as demonstrated by Zahl (1977), jackknifing provides the investigator with a better understanding of the biological processes under investigation than is possible with a single measure of an index of diversity.

Finally, these techniques have been applied to phylogeny reconstruction (Felsenstein 1985, Lanyon 1985b). This application differs from the foregoing in that it is not a statistic that is being estimated, but rather a pattern: a phylogenetic pattern. The gener-

ation of bootstrap or jackknife estimates of a phylogeny (in the form of pseudoreplicate trees or cladograms) permits investigators to determine how well each branching point is supported by the data. Furthermore, the generation of a suite of alternative trees allows the measurement of branch length dispersion. Once again, these resampling techniques permit investigators to determine what subsets of the results are worthy of further discussion.

Jackknifing and bootstrapping are not panaceas. They have limitations and inherent assumptions, as do all statistical procedures, and investigators would do well to be aware of these. However, these techniques represent an important step in refining the process of data analysis.

I thank James Koepfel and Douglas Stotz for their helpful suggestions on an earlier draft of this commentary.

LITERATURE CITED

- BISSELL, A. F., & R. A. FERGUSON. 1975. The jackknife—toy, tool, or two-edged weapon? *Statistician* 24: 79–100.
- BRADLEY, D. W., R. E. LANDRY, & C. T. COLLINS. 1984. The use of jackknife confidence intervals with the Richards curve for describing avian growth patterns. *Bull. Southern California Acad. Sci.* 83: 133–147.
- CLARKSON, D. B. 1979. Estimating the standard errors of rotated factor loadings by jackknifing. *Psychometrika* 44: 297–314.
- DIACONIS, P., & B. EFRON. 1983. Computer-intensive methods in statistics. *Sci. Amer.* 249: 116–130.
- EFRON, B. 1979. Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7: 1–26.
- . 1982. The jackknife, the bootstrap, and other resampling plans. CBMS-NSF regional conference series in applied mathematics No. 38. Philadelphia, Pennsylvania, Society for Industrial and Applied Mathematics.
- , & G. GONG. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. *Amer. Statist.* 37: 36–48.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- GIBSON, A. R., A. J. BAKER, & A. MOEED. 1984. Morphometric variation in introduced populations of the Common Myna (*Acridotheres tristis*): an application of the jackknife to principal components analysis. *Syst. Zool.* 33: 408–421.
- JAMES, F. C., & C. E. MCCULLOCH. 1985. Data analysis and the design of experiments in ornithology. Pp. 1–63 in *Current ornithology*, vol. 2 (R. F. Johnston, Ed.). New York, Plenum Press.
- LANYON, S. M. 1985a. Molecular perspective on higher-level relationships in the Tyrannoidea (Aves). *Syst. Zool.* 34: 404–418.

- . 1985b. A consensus technique for detecting internal inconsistencies in distance data. *Syst. Zool.* 34: 397-403.
- LANYON, W. E., & S. M. LANYON. 1986. Generic status of Euler's Flycatcher: a morphological and biochemical study. *Auk* 103: 341-350.
- MILLER, R. G. 1964. A trustworthy jackknife. *Ann. Math. Stat.* 35: 1594-1605.
- . 1974. The jackknife—a review. *Biometrika* 61: 1-15.
- QUENOUILLE, M. 1956. Notes on bias in estimation. *Biometrika* 43: 353-360.
- SOKAL, R. R., & F. J. ROHLF. 1981. *Biometry. The principles and practice of statistics in biological research*, 2nd ed. New York, W. H. Freeman Co.
- TUKEY, J. 1958. Bias and confidence in not quite large samples. (Abstr.) *Ann. Math. Stat.* 29: 614.
- ZAHL, S. 1977. Jackknifing an index of diversity. *Ecology* 58: 907-913.

Received 11 September 1986, accepted 12 September 1986.