Studies in Avian Biology No. 13:126–133, 1990.

Overview

ANALYTICAL CONSIDERATIONS FOR STUDY DESIGN

BARRY R. NOON AND WILLIAM M. BLOCK

Studies of the foraging behaviors of birds have been largely descriptive and comparative. One might then expect studies with similar objectives to have similar study designs but that is not the case. Papers in this symposium that focused specifically on study design contain a diversity of biological perspectives. Similarly, there is no accord among statisticians on experimental design and data analysis of multivariable systems. Further, biological and statistical considerations in study design are not always in agreement.

In this paper, we attempt to define the nature of foraging data and to discuss the arbitrary structure of much of the data that are collected. We then touch on the diversity of approaches to study design that appear in this symposium. Finally, we attempt to identify areas of contrasting opinion, offer our own perspectives on controversial issues, and suggest areas in need of further research.

THE NATURE OF FORAGING DATA

Most data on avian foraging are derived from field observations of foraging events that can be classified by one or more nominal attributes. If two or more attributes are recorded for each event, then the data are referred to as cross-classified. Events are now redefined according to each unique combination of attributes assigned to an observation. These classes of events have the property of being mutually exclusive and exhaustive. Given a sample of observations, the final data have the form of counts or frequencies with which certain events were observed. Data with this structure can be portrayed as cross-classified tables with each cell of a table representing the frequency with which a particular event was observed.

Occasionally, event frequencies are estimated across known time intervals, which makes it possible to estimate foraging rates as well as frequencies. If behavioral events are persistent and of sufficient duration, one can construct time budgets. Event-based and time-based approaches are combined when data are collected sequentially and represent a sequence of events. Time intervals can be of fixed or variable length; in the latter case they are dependent upon the cessation of an event. One can estimate event durations and rates from these data but, in addition, one can look specifically at the arrangement of events in the time series and estimate a number of conditional probabilities; for example, given that event A has occurred, what is the probability that it will be followed by event B? The conditional, or transition, probabilities can be arranged in a transition matrix. The event observed at time t is the row variable and the event observed at time $t + 1$, given the event at t, is the column variable. The probability of going from one event to another in a single time step is referred to as a Markov chain.

Regardless of the design of data collection, most foraging studies are event based and the data end up being represented by frequencies. As such, the data are counts of discrete random variables, and relationships among the event categories should be analyzed by discrete multivariate models (cf. Bishop et al. 1975).

The nominal attributes (such as tree species or substrate type) or factors involved in foraging can have many levels. If each event is classified according to bird species, sex, tree species, and foraging substrate, the potential number of mutually exclusive and exhaustive categories is large. A comparative study, for example, of the use of bark versus foliage of four tree species by both sexes of five bird species would result in 80 distinct event categories. Each observed foraging event is classified into the appropriate class for each of the four factors. As such, we can view each observed foraging event as a multinomial trial with a probability of falling in event category i given by p_i, where $i = 1, 2, \ldots, 80$. These probabilities can be estimated from the original frequency data by dividing the frequency of event i by the sum of the frequencies of all events. The data expressed in this form are still discrete, though no longer represented in integer form. If these probabilities are viewed as unbiased estimates of the true multinomial probabilities, assumed constant over the period of study, then the frequencies of each event category can be estimated by multiplying the total number of events (a constant) by the appropriate probability. This exercise will simply reproduce the original data indicating that its basic discrete nature has not been changed.

WAYS OF LOOKING AT THE SAME DATA: CONTINUOUS OR DISCRETE VARIABLES

Viewing the data as continuous random variables

Many authors have analyzed multinomial probabilities rather than event frequencies. That is, they have changed the representation of the data to appear as continuous rather than discrete random variables. Presumably the data have been standardized in this way, because some types of statistical models assume that the input data are continuous. Even so, the data are still discrete.

To analyze data with this structure, most researchers have employed an ordination algorithm such as principal components analysis (PCA) or, less commonly, correspondence analysis (see Miles, this volume). Prior to analysis, the data are arranged in a matrix with each row representing a species and each column a probability associated with a distinct foraging variable. Assuming random sampling, entries in this matrix represent the probability of observing species i engaged in foraging behavior j. To visualize similarities and differences among species, it is useful to think of plotting the rows of this matrix in a j-dimensional space.

A frequent goal of principal components analysis is to plot the rows of the matrix in terms of linear combinations of the column variables. The coefficients defining the linear combination are functions of the eigenvectors estimated from an association matrix of the column variables (usually a correlation or covariance matrix). The scalar product of the jth eigenvector times the ith row of the probability matrix produces the score for the ith individual on the jth principal component. The weights assigned to the foraging variables are estimated so as to maximize the variance of the principal component scores. After the new scores are computed they are plotted according to bird species. The arrangement of species (= points) in this space, viewed in terms of their point-to-point distances, is used to infer similarities and differences among the species. The principal component axes are given biological interpretations in terms of the correlations among the scores and the original columns of the probability matrix.

Correspondence analysis, or reciprocal averaging (RA), is similar to PCA in that it is also based on an eigenanalysis of a two-way matrix (species by probabilities). However, in RA both the rows (species) and columns (foraging behaviors) are analyzed and ordinated simultaneously. The algorithm is referred to as reciprocal because the species ordination scores are averages of the column (foraging variables) ordination scores, and reciprocally, the variable ordination scores are averages of the species ordination scores (Gauch 1982:144). A further difference is that PCA is based on Euclidean distances, provides equal weight to all points, and the ordination is centered at the origin (for mean-corrected data). In contrast, RA is based on chi-square distances, weights are proportional to row and column sums, and the origin is at the center of gravity of the data (Gauch 1982a:147–148). However, the techniques are very similar in their goal of reducing the dimensionality of the original space, and providing some logical ordering of the species that can be given a biological interpretation. One of the most useful aspects of RA is the biplot. In a biplot, both row and column variables of the two-way table are simultaneously plotted with respect to the principal axes (Moser et al., this volume). The biological interpretation of the ordination is based on the relative positions of row and column variables (points) in the plot.

Treating the same data as discrete random variables

It seems somewhat arbitrary to take data that are originally portrayed as a multidimensional, cross-classified matrix and collapse them into two-way matrix of species by foraging variables for analysis by PCA or RA. In doing so we artificially create a series of quasi-independent variables and ignore relationships among the original factors. In light of this concern, RA is to be preferred to either PCA or its variants (e.g., factor analysis). It is possible to use RA complementary to traditional discrete multivariate analyses (van der Heijden and de Leeuw 1985) and to explore both two-way and multidimensional tables based on the original event frequencies (Greenacre 1984; Moser et al., this volume). RA can be used to explore multidimensional contingency tables by the use of dummy variables (Greenacre and Hastie 1987) or by structuring the event frequencies into Burt tables (Greenacre 1984:140–143). A Burt table contains each factor in both rows and columns of the table, thus containing all possible two-way tables (see Moser et al., this volume, for an example).

Since the original data can be arranged as a multiway contingency table, it seems logical to retain this structure for analysis. This is accomplished through the use of log-linear models which explicitly estimate the interdependencies among the factors. For illustration, we return to our previous example of a comparative foraging study of both sexes (S) of five species of birds (B) and their use of bark versus foliage substrates (I) on four species of tree (T). Each of the observed foraging events can be classified by bird species, sex, tree species, and substrate: these are the four factors. The model, presented below, of complete

TABLE 1. HYPOTHETICAL EXAMPLE OF AVIAN FORAGING DATA ILLUSTRATING VARIOUS LOGLINEAR MODELS AND THE INTERPRETATION OF MODEL PARAMETERS

Full model

$$\ln f_{ijkl} = u + B_i + S_j + T_k + I_l + BS_{ij} + BT_{ik} + BI_{il} + ST_{jk} + SI_{jl} + TI_{lk} + BST_{ijk} + BSI_{ijl} + BTI_{ikl} + STI_{jkl} + BSTI_{ijkl}$$

Parameters:

B_i = bird species	$i = 1, 2, \ldots, 5$
S_j = sex (male or female)	$j = 1, 2$
T_k = tree species	$k = 1, 2, \ldots, 4$
I_l = substrate	$l = 1, 2$
f_{ijkl} = cell frequency in the *ijkl* cell	

Model of complete independence

$$\ln f_{ijkl} = u + B_i + S_j + T_k + I_l$$

Parameter	Interpretation
u	Mean of the logarithms of the expected frequencies
B	One-way term for bird species
S	One-way term for sex
T	One-way term for tree species
I	One-way term for substrate
BS	Sample size effects: the same proportion of males and females were not sampled for all sexes
BT	Not all bird species are utilizing tree species in the same proportions
BI	Not all bird species are utilizing substrates in the same proportions
ST	The two sexes are not using tree species in the same proportions
SI	The two sexes are not using substrates in the same proportions
TI	The proportion of utilized substrates is not the same for all tree species (implicit bird species effect)
BST	The association between sex and tree species depends upon the level of bird species (i.e., males and females differ in the use of tree species according to which species they belong to)
BSI	The association between sex and substrate depends upon the level of bird species
BTI	The association between tree species and utilized substrates is dependent on the level of bird species
STI	The association between tree species and utilized substrates is dependent upon whether the bird is a male or a female
BSTI	The association between tree species and utilized substrates is dependent upon whether the bird is a male or a female and this three-way association is in turn dependent upon the level of bird species

association among the factors, would involve all interaction terms of order four or lower plus all individual factors (Table 1):

$$\ln f_{ijkl} = u + B_i + S_j + T_k + I_l + BS_{ij} + BT_{ik} + BI_{il} + ST_{jk} + SI_{jl} + TI_{lk} + BST_{ijk} + BSI_{ijl} + BTI_{ikl} + STI_{jkl} + BSTI_{ijkl}.$$

In contrast, the model of complete independence of the four factors would contain only the terms for the individual factors (Table 1):

$$\ln f_{ijkl} = u + B_i + S_j + T_k + I_l.$$

The full model contains 15 classes of parameters: four main effects terms, six two-way interaction terms, four three-way interaction terms, and one four-way term. In all, 80 parameters need to be estimated ($5 \times 2 \times 4 \times 2 = 80$). However, what we seek is the model with the fewest number of terms that adequately fits the data. By fit we mean that the chi-square statistic, based on the difference between observed and predicted frequencies, is not significant (e.g., $P > 0.05$). This model will lie somewhere between the model of complete independence and complete dependence. Inclusion of any interaction terms indicates some degree of dependence among the factors. In addition, to make interpretation easier, only hierarchial log-linear models are usually considered. For example, if any three-way interaction term is included in the model, then all two-way interaction terms involving those factors, and the individual factors, are also included in the model.

Model interpretation. Similar to linear models in the analysis of variance, there are alternative ways to block the factors to aid in interpretation. An example would be to define bird species (*B*) and sex (*S*) as explanatory or treatment variables and tree species (*T*) and substrate (*I*) as response variables. The parameter estimates by factor and interaction, and an interpretation of each parameter, are given in Table 1.

The interaction terms of primary importance are those involving some combination of explanatory and response variables. To illustrate the hierarchical nature of the models, if the highest order term required in the model was *BTI*, then the terms *BT*, *BI*, and *TI*, and *B*, *T*, and *I* would also be required for an adequate fit of observed and expected frequencies under the hierarchical principle.

Mixtures of continuous and categorical random variables

Foraging studies often involve a mixture of categorical and continuous random variables. For example, Sakai and Noon (this volume) recorded tree species and substrate types (categorical variables) as well as the height and distance from the trunk (continuous variables) of foraging flycatchers. They employed separate analyses, using different statistical models, of the two data types. However, one can use mixtures of variables in some analyses. For example, a PCA of mixed variable data sets is possible because the estimation of eigenvalues and eigenvectors is not dependent upon normality assumptions. Discriminant function analyses (DFA) can also be done with continuous and categorical variables, although logistic regression may be preferred in the two-group case because of its robustness to violations of the normality assumption (Press and Wilson 1978; for a contrasting opinion see Haggstrom 1983).

As an example, consider a multi-species study whose primary data have been arranged in a matrix with the rows partitioned by bird species and the columns representing foraging variables. Each row of this matrix is assumed to represent an independent foraging observation of an individual bird of a particular species. For each observation, bird species, tree species, behavior, substrate, bird height, and distance from the center of the plant are recorded. All but the last two variables are categorical. In general, any factor with k levels can be represented by $k - 1$ dummy (0/1) variables. If there are five possible tree species, then this variable is coded by four dummy, binary variables; four behaviors would be coded by three variables, and so on. (The sum of a set of 0/1 variables has approximately a normal distribution.) The species' groups are to be contrasted on the basis of the foraging variables by DFA.

A problem in discriminant analyses with both continuous and categorical variables is the procedure of selecting variables and thus the biological interpretation of the canonical variates. For example, some continuous variables may supply discrimination only if a particular discrete variable is already in the model (Daudin 1986). Several recent papers discuss the analysis of mixed variable data sets when group discrimination is the goal (Krzanowski 1980, Knoke 1982, Vlachonikolis and Marriott 1982, Daudin 1986) but reach no general consensus. Several authors have argued in favor of the location model approach to DFA, which involves aspects of log-linear analyses and parametric analysis of variance. This requires estimation of a large number of parameters and has not been implemented on any major statistical software package. Analyses of mixed variable data sets with standard statistical packages should be interpreted cautiously.

How are cross-classified categorical data best analyzed

It is possible to take cross-classified data and analyze them as discrete frequencies with log-linear models or to express the data as proportions for analysis by various ordination algorithms (e.g., PCA or factor analysis). But which method provides the clearest insights into the relationships among factors; and do different methods provide complimentary insights?

In the example discussed above of both sexes of five species of birds, a PCA ordination would be based on a matrix whose rows represent bird species-sex combinations (10 distinct categories) and whose columns represent all possible tree species by substrate combinations (8 distinct categories). Entries in this 10 × 8 matrix would represent the proportion of observations for species-sex combination i observed on tree species-substrate combination j. These entries can also be considered as conditional or multinomial probabilities. For example, entry ij would be interpreted as: given a random observation of species-sex combination i, what is the probability that it is foraging on tree species-substrate combination j. Biological inferences from the ordination of the rows of the matrix are based on distances among the rows plotted as points in the synthetic PC space and from the biological interpretations given to the PC axes. The statistical significance of interactions among the factors (bird species, sex, tree species, and substrate) is not explicitly examined. Rather, these methods of analysis lead to inferences about the similarities or differences among various species-sex combinations in terms of the measured tree species-substrate variables.

In contrast, log-linear analyses explicitly investigate the significance of interactions among the nominal factors and seek the simplest representation of the tabulated frequencies. The factors in these models can be viewed as possessing a treatment-response structure and the significance of any association between factors can be explicitly tested. Relationships among species-sex combinations would be inferred from a comparison of their parameter estimates (the BS_{ij} terms) or by a series of pairwise comparisons of species-sex by tree species and substrate contingency tables (see Raphael, this volume).

Ordination techniques, such as PCA or RA, are not primarily hypothesis testing procedures. Instead, they are most useful for exploring interrelationships among species or foraging variables. In contrast, log-linear models are often explicitly cast in an hypothesis testing context. This suggests that ordination analyses may be more valuable in the initial research into a species' or community's foraging patterns. Log-linear analyses may be used in a subsequent study to explicitly test for significant relationships among some subset of factors implicated by the initial, exploratory analyses.

For a geometric interpretation of factor relationships, ordination analyses are preferred to log-linear analyses. However, if log-linear analyses are done along with RA analyses of combinations of factors, complementary inferences can arise. Van der Heijden and de Leeuw (1985) argue that log-linear analyses yield insights into factor relationships whereas RA analyses provide insights into associations among levels within factors. To illustrate, one could initially analyze the multiway foraging data by log-linear algorithms to estimate the simplest model that adequately fits the observed frequencies. If the model contained significant interaction terms, then these terms could be examined in combination with the treatment factor by correspondence analysis. That is, one or more two-way tables of frequencies, in which the columns of the table represent all possible combinations of levels of factors within a significant interaction term, would be examined for association with the treatment factor and interpreted geometrically. This approach is illustrated by Moser et al. (this volume) and van der Heijden and de Leeuw (1985). A lucid discussion of the geometry of correspondence analysis is presented by Greenacre and Hastie (1987).

We have not seen a comparison of ordination algorithms and log-linear models on the same data set, but suspect that similar inferences about the relationships among factors would be drawn. An explicit comparison of these contrasting methods of analysis is an important area for future investigation. At this time it is not clear if one method is to be preferred over the other and whether more information is extracted from the data by conducting both analyses. However, the complimentary relationship among log-linear and correspondence analyses in the exploration of categorical variables appears most promising at this time.

SEQUENTIAL OR POINT OBSERVATIONS OF FORAGING BEHAVIORS

Two methods of recording foraging events are commonly used. Point samples record the first event observed (or the first recorded after a fixed waiting period to avoid recording only conspicuous behaviors). Sequential samples consist of sequences of events recorded during a fixed or variable time interval. The debate over the use of sequential or point observations focuses, in part, on the issue of statistical independence. Independence of observations is critical for the valid use of most statistical distributions, and thus for tests of hypotheses. Let the events y_1, y_2, and y_3 be mutually exclusive and exhaustive. Define y_1 equal to the event that a bird forages on a leaf, y_2 that it forages on a twig, and y_3 that it forages on bark. Further, let events y_1, y_2, and y_3 occur with probabilities p_1, p_2, p_3, and with the sum $(p_i) = 1.0$. Assuming only first-order correlations, we say that events y_1 and y_3 are statistically independent if the probability of y_3 occurring at time $t + 1$, given that y_1 occurred at time t, is equal to p_3. That is, the conditional probability of an event is equal to its marginal probability. We infer events y_1 and y_3 to be statistically dependent if the probability of observing event y_3 at $t + 1$, given y_1 at t, is not equal to p_3. Tests to examine dependencies in categorical and continuous data are discussed in Hejl et al. (this volume).

When foraging events are recorded in sequence, there is often a tendency for observations close together in either time or space to be more similar than events separated by longer time intervals or distances. Several authors in this volume have addressed issues of temporal dependency, but there has been little discussion of spatial dependency. An exception is Block (this volume), who sampled so as to ensure spatial independence of foraging observations within the same season. Spatial associations may actually be more prevalent, because so many studies are conducted when birds are spatially restricted. For example, subsequent observations of territorial birds, even if separated by long time intervals, may be significantly dependent because territories are likely to encompass different ranges of foraging possibilities and in different proportions. This is an area in need of further research.

Because most statistical models require random and independent observations, many researchers have recorded point observations. Such a sampling design may fulfill the independence assumption, but random sampling is difficult to achieve because the probability of obtaining a foraging observation differs among and within species. An argument, however, in favor of recording sequential foraging acts can be made because most of our data sets are sparse. Maurer et al. (this volume) have estimated that most foraging studies record fewer than 1% of the behaviors occurring during the period of study. Given the size of our sample relative to the sampling frame, we should attempt to collect as much information as possible and to record sequential observations. Such an approach, however, will necessitate recording data so that the temporal sequence of behaviors is documented. This information is needed to estimate the conditional probabilities (given that species i is engaged in foraging act j at time t, what is the likelihood that it will be engaged in act k at time $t + 1$) that form the elements of the first-order transition matrices.

We propose that researchers start with the assumption that sequential behaviors of the same individual are usually dependent (see Hejl et al., this volume). Further, we believe that estimates of the magnitude and direction of these dependencies will yield important insights into a species' foraging ecology and lead to improved predictive models. We support the argument of Raphael (this volume) in favor of Markov analyses, which estimate both the stationary distribution vector of foraging acts (however defined) and model building via log-linear algorithms. The latter analyses allow explicit tests for symmetry (i.e., the likelihood of the transition from behavior j to k equal to that from k to j) as well as comparisons of the transition matrices of different bird species (see Raphael, this volume, for details).

Our suggestion in favor of collecting sequential data is in contrast to that of Hejl et al. (this volume), Bell et al. (this volume), and Recher and Gebski (this volume), who suggested that point observations generally yield more precise parameters for estimating the probabilities of events. If sequences are recorded, then Hejl et al. recommended bootstrap or jackknife methods, because they are less time-consuming than Markov analyses, do not require assumptions about the order of the transitions, and provide estimates with smaller standard errors. However, these studies focused on estimating the mean probabilities of foraging events. We argue, from biological and not statistical grounds, that the transition probabilities themselves are as important in gaining insights to the behavior of foraging birds as are the expected probabilities. We recommend methods that provide both types of estimates.

SAMPLE SIZE REQUIREMENTS

In this symposium approaches to estimate sample sizes range from qualitative interpretations of graphs (Brennan and Morrison) to quantitative calculations of sample sizes based on different target levels of absolute or relative precision (L. Petit et al.). Suggested minimum sample sizes range from 40 to 500 independent observations to an extreme figure of 20,000!

Despite a diversity of approaches, all foraging studies must state what behavioral parameters will be estimated and with what levels of precision. The latter will require at least preliminary knowledge of the species' foraging variability. If the study is comparative, then determining what precision levels can be obtained is essential to estimate the power of any between-species comparisons. For species with variable foraging repertoires, sample size requirements may be so large that the researcher will need to be satisfied with tests of lower power. In this case, only differences among the most disparate species may be detected.

Log-linear analyses

Many papers in this symposium used log-linear models in analyzing categorical foraging data. Recall that the test-statistics for fitting log-linear models are only asymptotically chi-square distributed, and that some minimal sample size is needed for valid statistical inference. For a fixed sample size, the more cell frequencies that are estimated, the more questionable are the probability levels associated with the computed chi-square values. An indication of an inadequate sample size is an excess of small expected cell frequencies. Cochran (1954) suggested that no expected cell frequencies should be <1, and $<20\%$ of the cells should have frequencies <5. A rough guideline is that one should collect about five times as many observations as there are cells in the table (Raphael, this volume). If the table contains one or more rows or columns of all zeroes, the degrees of freedom associated with the test-statistic must be adjusted (Bishop et al. 1975:116).

Surprisingly, an analysis can be affected by too many observations. The result is that most models will fail to fit the data. If too large a sample is taken, any possible model structure will provide a poor fit no matter how minor the discrepancies. This occurs because chi-squares are proportional to the total sample size. If too large a sample is a problem, then the appropriate model may be selected by a stepwise procedure. For example, the magnitude of reduction of the sum of squares of the differences between observed and expected proportions can be computed each time an additional term is added to the model. Terms producing a large decrease in the sum of squares should be considered for inclusion in the final model.

A need to limit the number of factors

A large number of observations is needed to analyze a cross-classified table of even moderate size, because of the number of parameters that need to be estimated. Three factors with four levels each would require the estimation of 64 parameters. In contrast, a multiple regression model with three independent variables and no interaction terms would require, at most, the estimation of seven parameters. Because the number of possible sources of variation in avian for-

TABLE 2. FACTORS AND NUMBERS OF LEVELS CONSIDERED IN A STUDY OF THE FORAGING BEHAVIOR OF THE WESTERN AND HAMMOND'S FLYCATCHERS (FROM SAKAI AND NOON, THIS VOLUME)

Factor	Number of levels
Observers	4
Years	2
Age of forest	3
Stage of breeding cycle	4
Behavior	3
Tree species	6
Substrates	4

Total number of cells = 4 × 2 × 3 × 4 × 3 × 6 × 4 = 6912.

aging behavior is staggering, one cannot estimate all sources of variation, all significant interactions among factors, or investigate all possible factor levels.

For example, Sakai and Noon (this volume) used seven factors (Table 2) in their log-linear model. Considering the levels of all factors there were a total of 6912 cells for each bird species. This value greatly exceeded the total number of data points. The authors had decided *a priori* to pool across forest age because their objective was to estimate foraging patterns across the range of forest types occupied by the species. However, after recognizing the limitations imposed by the size of their data set, they chose to pool across observers and years as well. This probably masked statistically significant interactions and lost information on the joint distribution of some factors. Whether insights into significant biological interactions were lost is unclear.

Our point is that pooling is necessary and justifiable in almost all studies. When possible, interactions among factors that are of minimal biological interest should be controlled in the experimental design and data collection phases, and not in the analysis phase. Our zeal to partition sources of variation as finely as possible needs to be tempered with the recognition that one of our primary objectives is to understand a complex system in terms of a small set of key factors. We are interested in models that can describe and predict the average outcome of samples, not the outcome of individual observations.

MARKOV ANALYSES

We are aware of little published information on sample size requirements for Markov analyses. From unpublished simulation studies conducted by R. M. Fagen (Fagen in Colgan 1978: 107–108), some general guidelines have been proposed. If we let k equal, for example, the number of substrate categories considered, and assuming a first-order Markov model, then a sample of $2k^2$ foraging events is too few, $10k^2$ almost always adequate, and $5k^2$ a borderline value. Thus, if 10 substrate categories are considered, the minimum number of foraging events required is 500.

MULTIVARIATE ANALYSES

Estimates of sample size requirements for multivariate studies are considerably more complex than for univariate studies. We are still concerned with the precision of parameter estimates and the power to reject false null hypotheses, but in addition, one must consider the number of variables, the covariance structure of the data, the number of groups, and the sample size per group. There are "rules of thumb" but few are based on either analytical or simulation studies (e.g., Morrison 1984b). An example of a sample size effect, similar to univariate parameter estimates, is that the confidence interval around a principal component's variance (i.e., its eigenvalue) is a function of the reciprocal of the square root of its sample size (Neff and Marcus 1980: 37). Estimates of confidence intervals, as a function of different sample sizes, can be computed by resampling methods such as the jackknife or bootstrap (Efron 1982; Efron and Gong 1983; Miles, this volume). These computer-intensive methods to variance estimation have considerable application to foraging data.

A clear exception to the lack of information on sample size requirements is the recent study of Williams and Titus (1988). Based on a large scale simulation study, they have developed the following sampling rule: "For discriminant analysis of ecological systems with homogeneous dispersions, choose the total number of samples per group to be at least three times the number of variables to be measured." More guidelines such as these are needed. In their absence, researchers can empirically estimate the variance of many multivariate parameters (i.e., eigenvalues, factor loadings) by the use of jackknife and bootstrap methods. If the resulting confidence intervals on these parameters are too broad for study objectives, then larger sample sizes will be required.

CONCLUSIONS

We believe the papers presented in this symposium represent a significant advancement in the design and analysis of studies of avian foraging behavior. An explicit concern for precise and unbiased parameter estimates, and the necessary sampling design and sample sizes to achieve these goals, should become a regular part of all study designs. In addition, analytical techniques such as log-linear models, Markov processes, and correspondence analysis have become part of the repertoire for the analysis of foraging data. While most of these statistical techniques are not new to the ecological sciences, their application to studies of avian foraging behavior is novel. An additional advancement is the use of computer-intensive methods such as the jackknife and bootstrap. Diversity indices, factor loadings, eigenvalues, discriminant coefficients and other statistics that are regularly computed in foraging studies are usually done without estimates of their variances. Through intensive resampling of the original data, jackknife and bootstrap methods allow estimates of the standard errors of these statistics, yielding better or more appropriate insights into the variability of the systems under study.

Many issues require further work: the variable structure of foraging data and whether it is best analyzed by discrete or continuous multivariate models; the analysis of mixtures of continuous and categorical data; and whether we should sample so as to ensure independent observations or explicitly estimate the dependencies of foraging behaviors. We encourage investigators to address these and related issues in their future research efforts.