Ecology, 66(4), 1985, pp. 1204–1210
© 1985 by the Ecological Society of America

STATISTICAL INFERENCE ON MEASURES OF NICHE OVERLAP[1]

LAURENCE D. MUELLER[2] AND LEE ALTENBERG
Department of Biological Sciences, Stanford University, Stanford, California 94305 USA

Abstract. Estimates of measures of niche overlap are often reported without any indication of sampling variance or an accompanying confidence interval. We have investigated the delta, jackknife, and bootstrap methods for making statistical inferences on four measures of niche overlap: the coefficient of community, Morisita's index, Horn's index, and the Euclidian distance. Our qualitative conclusions are: (1) The bias of these estimators was usually $<10\%$ of the mean unless the sample size was small and the number of resource categories large. The jackknife and bootstrap can significantly reduce this bias. (2) The variance of the bootstrap and jackknife estimators was usually greater than that of the "standard" estimator. (3) Under a variety of circumstances, the population sampled may actually represent several unrecognized subpopulations. In such cases confidence intervals generated by the jackknife and delta techniques can be quite inaccurate, while the nonparametric confidence intervals derived from the bootstrap are highly accurate.

Key words: bootstrap; coefficient of community; computer simulation; delta method; Euclidian distance; Horn's index; jackknife; Morisita's index.

INTRODUCTION

Until recently little attention has been paid to the statistical properties of niche overlap measures (Petraitis 1979, Ricklefs and Lau 1980, Maurer 1982, Smith and Zaret 1982). Many of these overlap measures are complicated functions of the underlying observations. It is therefore not surprising that most ecologists have not attempted to estimate the variance or construct confidence intervals for these measures of overlap (Orians and Horn 1969, Pulliam and Enders 1971, Jaksic et al. 1981, Paine et al. 1981). Ricklefs and Lau (1980) have examined the bias and variance of four measures of niche overlap by sampling theoretical populations on a computer. They suggest that empirical workers conduct computer simulations using their own data to gain insights concerning the sampling variance and distribution of these overlap measures. In fact, in recent years the use of computer simulations has allowed the development of a growing body of "computer intensive" statistics (Efron 1979*b*), which sidestep the mathematical difficulties encountered in analyzing many statistics having non-normal distributions. We shall be examining one example of these techniques, the "bootstrap" estimator (Efron 1979*a*).

In addition we will derive analytic approximations for the variance of the overlap indices studied by Ricklefs and Lau. Using this variance estimate we will also describe a method for constructing confidence intervals. The methods involved in these analytic approximations come under the general heading of the "delta method." This method is based on Taylor series approximations. Smith and Zaret (1982) have used these same techniques to derive expressions for the bias of a variety of niche overlap measures. They show that, in general, bias will be greater when many resource categories are used rather than a few. Smith (1982) has also used the delta method to derive the approximate sampling variance of several measures of niche breadth and has shown these approximations to be quite accurate.

We will describe a third method, the jackknife, for both point and interval estimation. We examine statistical properties of all three methods by conducting extensive computer simulations. These computer simulations allow us to evaluate the accuracy of these methods, since each involves approximations. We also use the results from the computer simulations as a means of discerning if one method is superior to the other.

THE DELTA METHOD

We are interested in calculating the niche overlap of two species (populations) which utilize up to n different resources. Let the probabilities of resource utilization for the first species be given by $p_1, p_2, \ldots, p_n$ and for the second species $q_1, q_2, \ldots, q_n$. With these population quantities we can calculate measures of niche overlap in a variety of ways (Schoener 1970, Pianka 1974, Hurlbert 1978, Petraitis 1979, Ricklefs and Lau 1980, Maurer 1982). We will restrict our attention here to the four measures considered by Ricklefs and Lau (1980): coefficient of community (S_1), Morisita's index (S_2), Horn's index (S_3) and the Euclidian distance (S_4). These indices are calculated as follows,

$$S_1 = \sum_{i=1}^{n} \min\{p_i, q_i\}, \tag{1a}$$

$$S_2 = 2 \sum_i p_i q_i / \left(\sum p_i^2 + \sum q_i^2\right), \tag{1b}$$

[1] Manuscript received 26 August 1983; revised 10 May 1984; accepted 3 October 1984.

[2] Present address: Department of Zoology, Washington State University, Pullman, Washington 99164 USA.

$$S_3 = \left[\sum_i (p_i + q_i)\log(p_i + q_i) - \sum_i p_i \log p_i - \sum_i q_i \log q_i\right] \Big/ [2\log(2)], \quad (1c)$$

$$S_4 = 1 - \left\{\sum_i (p_i - q_i)^2/2\right\}^{1/2}. \quad (1d)$$

In practice we do not know the p_i's and q_i's, but we have sampled the population and obtained estimates $\hat{p}_i$ and $\hat{q}_i$ ($i = 1, \ldots, n$). One way of estimating S_j ($j = 1, \ldots, 4$) is to replace every p_i and q_i in Eq. 1 with $\hat{p}_i$ and $\hat{q}_i$. We will denote the estimators obtained in this fashion as $\hat{S}_j$.

In addition to obtaining an estimate of S_j we would like to be able to estimate the sampling variance of each S_j. It is possible to derive the approximate sampling variance of the $\hat{S}_j$ using the delta method (see Bishop et al. 1975:486–488). This method yields the following formulae,

$$\text{var}(\hat{S}_j) \cong \boldsymbol{h}_j^T \boldsymbol{D}_p \boldsymbol{h}_j + \boldsymbol{v}_j^T \boldsymbol{D}_q \boldsymbol{v}_j \quad (2)$$

where

$$\boldsymbol{h}_j^T = (\partial\hat{S}_j/\partial p_1, \ldots, \partial\hat{S}_j/\partial p_{n-1}),$$

$$\boldsymbol{v}_j^T = (\partial\hat{S}_j/\partial q_1, \ldots, \partial\hat{S}_j/\partial q_{n-1}),$$

and $\boldsymbol{D}_p$ and $\boldsymbol{D}_q$ are the covariance matrices of $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively. For the four indices we have,

1) Coefficient of community:

$$\partial\hat{S}_1/\partial p_i = A + B$$

where

$$A = \begin{cases} 1 \text{ if } \hat{p}_i < \hat{q}_i \\ \tfrac{1}{2} \text{ if } \hat{p}_i = \hat{q}_i \\ 0 \text{ if } \hat{p}_i > \hat{q}_i \end{cases}$$

$$B = \begin{cases} -1 \text{ if } 1 - \sum_{i=1}^{n-1} \hat{p}_i = \hat{p}_n > \hat{q}_n \\ -\tfrac{1}{2} \text{ if } \hat{p}_n = \hat{q}_n \\ 0 \text{ if } \hat{p}_n < \hat{q}_n \end{cases}$$

2) Morisita's index:

$$\partial\hat{S}_2/\partial p_i = 2[(\hat{q}_i - \hat{q}_n) - \hat{S}_2(\hat{p}_i - \hat{p}_n)] / \left(\sum \hat{p}_i^2 + \sum q_i^2\right),$$

3) Horn's index:

$$\partial\hat{S}_3/\partial p_i = \left\{\log\left[\frac{\hat{p}_i + \hat{q}_i}{\hat{p}_n + \hat{q}_n}\right] - \log[\hat{p}_i/\hat{p}_n]\right\} \Big/ [2\log(2)],$$

4) Euclidian distance:

$$\partial\hat{S}_4/\partial p_i = [(\hat{p}_i - \hat{q}_i) + (\hat{q}_n - \hat{p}_n)] / 2(1 - \hat{S}_4),$$

and all summations are for $i = 1, \ldots, n-1$. $\partial\hat{S}_j/\partial q_i$ are the same as $\partial\hat{S}_j/\partial p_i$ except $\hat{p}_i$, $\hat{p}_n$, $\hat{q}_i$, and $\hat{q}_n$, are replaced with $\hat{q}_i$, $\hat{q}_n$, $\hat{p}_i$, and $\hat{p}_n$, respectively.

If we assume that the count vector (the number of items recorded for resource category $1, 2, \ldots, n$, for each species) has a multinomial distribution, then the covariance matrix for $\hat{\boldsymbol{q}}^T = (\hat{q}_1, \ldots, \hat{q}_{n-1})$ is $\boldsymbol{D}_q$ and for $\hat{\boldsymbol{p}}^T = (\hat{p}_1, \ldots, \hat{p}_{n-1})$ is $\boldsymbol{D}_p$, where

$$D_{p,ii} = \hat{p}_i(1 - \hat{p}_i)/N_1,$$
$$D_{q,ii} = \hat{q}_i(1 - \hat{q}_i)/N_2,$$
$$D_{p,ij} = -\hat{p}_i\hat{p}_j/N_1,$$
$$D_{q,ij} = -\hat{q}_i\hat{q}_j/N_2,$$

and N_1 is the total number of items sampled for species 1, and N_2 is the same for species 2.

Although the method of estimating the S_j just described is straightforward it may not be the best available. Before considering alternative methods of estimation we will make some analytic statements about the $\hat{S}_j$ estimators. In particular we would like to know if $E(\hat{S}_j) = S_j$, i.e., is the estimator unbiased? Horn's index is a concave function that has $\partial^2 S_j/\partial p_i^2 + \partial^2 S_j/\partial q_i^2 + 2\partial^2 S_j/\partial p_i \partial q_i \leq 0$ for all i. Consequently we can invoke Jensen's inequality to prove that $E(\hat{S}_3) \leq S_3$ (see Karlin and Taylor 1975:249). This result does not tell us the magnitude of the bias but it does say that it will always be negative when it is not 0. This fact explains the persistent negative bias in the simulations of Ricklefs and Lau (1980:Table 1). It is possible to estimate the magnitude of the bias for $\hat{S}_2$ and $\hat{S}_3$ by expanding these functions in a Taylor series about the points $(p_1, \ldots, p_n)$ and $(q_1, \ldots, q_n)$. This technique has been used by Smith and Zaret (1982) who give approximations for the bias of $\hat{S}_2$ and $\hat{S}_3$. An exact expression for the bias of $\hat{S}_1$ is also given in Smith and Zaret (1982). van Belle and Ahmad (1974) have previously derived the delta estimate of variance for Morisita's index.

The Jackknife

A second method of estimating the indices (Eq. 1) is the jackknife statistic (see Miller 1974 for a review). The jackknife has been applied to an ecological diversity index (Zahl 1977), population size estimation (Burnham and Overton 1979), and genetic distance estimates (Mueller 1979). The jackknife provides an easy method for estimating the overlap indices and their variance. The jackknife can also be useful in reducing bias. When the bias takes the form $a/N + O(N^2)$, then the jackknife will remove the first term of the bias.

Let $\hat{S}_{j,-i}$ be the same as before, except that the i^{th} observation has been deleted before the calculation. Since we assume we have made N_1 observations for species 1 and N_2 observations for species 2 there are a total of $N_1 + N_2$ observations to delete. $N_1 + N_2$ pseudovalues may then be defined as

$$s_i = (N_1 + N_2)\hat{S}_j - (N_1 + N_2 - 1)\hat{S}_{j,-i}.$$

The jackknife estimator and its variance are then simply

$$\tilde{S}_j = \sum_i s_i/(N_1 + N_2)$$

$$\text{var}(\tilde{S}_j) = \sum (s_i - \tilde{S}_j)^2/(N_1 + N_2)(N_1 + N_2 - 1).$$

The number of calculations may be reduced by making the following observation. Suppose there are only two categories and the number of observations for species k in category l is $n_{kl}(k, l = 1, 2)$. Obviously $n_{11} + n_{12} + n_{21} + n_{22} = N_1 + N_2$. It is apparent from Eq. 4 that although there are $N_1 + N_2$ pseudovalues, there will only be four different values of s_i, s_{11}, s_{12}, s_{21}, and s_{22} corresponding to deleting an observation in category 1 for species 1, category 2 for species 1, etc. Thus simplified expressions for $\tilde{S}_j$ and $\widehat{\text{var}}(\tilde{S}_j)$ are

$$\tilde{S}_j = \sum_k \sum_l s_{kl}n_{kl}/(N_1 + N_2), \quad (3)$$

$$\text{var}(\tilde{S}_j) = \sum_k \sum_l (s_{kl} - \tilde{S}_j)^2 n_{kl}/(N_1 + N_2)(N_1 + N_2 - 1). \quad (4)$$

These equations will apply to any number of categories.

The Bootstrap

The proliferation of high-speed computers has opened the way for new techniques in statistics that would have been unthinkable 30 yr ago (see Efron 1979*b* for a review). One of these techniques is called the bootstrap (Efron 1979*a*). The bootstrap requires, perhaps, several million more arithmetic operations than either the delta or jackknife methods. The reward for this increased computational effort is twofold: (1) no assumptions about the distribution of the observations are required, (2) statistical measures that may have untractable mathematical features may be examined. The bootstrap uses the sampled observations as an empirical estimate of the distribution function. This probability distribution simply puts mass $1/N$ at each observed point, if N is the total sample size. Repeated samples of size N are then chosen from this empirical distribution. This results in estimates of such things as the bias and variance of particular statistics and confidence intervals. If the distributions of the observed random variable are known exactly, there is less to be gained from the bootstrap. However, we expect that the benefits of the bootstrap will increase as the differences between the assumed distribution and the real one increase. We will illustrate this technique with the overlap measures already discussed.

Suppose we have the estimates $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$ from independent samples of size N_i $(i = 1, 2)$. We produce one bootstrap replicate as follows. Using a random number generator we take two independent multinomial samples of size N_i where the underlying probabilities for each sample are $\hat{\boldsymbol{p}}$ and $\hat{\boldsymbol{q}}$, respectively. We designate the sample so obtained $\hat{\boldsymbol{p}}_1^*$, $\hat{\boldsymbol{q}}_1^*$. We then substitute these values in Eq. 1 and obtain a bootstrap estimate $\hat{S}^*_{i,1}$. In the present study this process was repeated 1000 times. We obtain bootstrap estimates of the variance and bias of $\hat{S}_i$ as,

$$\widehat{\text{Bias}}\,(\hat{S}_i) = \hat{S}_i - \sum_{j=1}^{1000} \hat{S}_{i,j}^*/1000,$$

$$\widehat{\text{var}}\,(\hat{S}_i) = \sum \left(\hat{S}_{i,j}^* - \frac{\sum \hat{S}_{i,j}^*}{1000}\right)^2/999. \quad (5)$$

We use as our bootstrap estimators of the S_i,

$$\hat{S}_i^* = \hat{S}_i - \widehat{\text{Bias}}\,(\hat{S}_i). \quad (6)$$

In addition to estimating the bias of various overlap measures, we will use the 1000 replicates of each S_i to approximate the cumulative distribution function. A Bayesian justification for this approach is given in Efron (1981). Once the cumulative distribution function has been estimated, confidence intervals can be derived. To describe the procedure for estimating confidence intervals, we first introduce some notation which closely follows Efron (1981). Let our estimate of the cumulative distribution function for S_i, $\widehat{\text{CDF}}_i(t)$, be defined as,

$$\widehat{\text{CDF}}_i(t) = \text{prob}\,\{\hat{S}_i^* < t\} = \#\,\{\hat{S}_{i,j}^* < t\}/1000,$$

where "#" can be read "the number of times." The values of S_i, which is greater than or equal to, say, 2.5% of all observations, would be given by $\text{CDF}_i^{-1}(0.025)$. Thus our $1-2\alpha$ central confidence interval is

$$[\widehat{\text{CDF}}^{-1}(\alpha),\ \widehat{\text{CDF}}^{-1}(1-\alpha)]. \quad (7)$$

If there is substantial bias in $\hat{S}_i$, then the confidence interval generated by Eq. 7 will be biased in a similar fashion. Efron (1981) has described a method for correcting this bias which we now describe. Let $\Phi(\cdot)$ be the cumulative distribution function of a standard normal random variable. We can now define $z_o = \Phi^{-1}\,\widehat{\text{CDF}}\,(\hat{S}_i)$ and $\Phi(z_\alpha) = \alpha$. A bias-corrected $1-2\alpha$ confidence interval is then given by

$$[\widehat{\text{CDF}}^{-1}(\Phi(2z_0 + z_\alpha)),\ \widehat{\text{CDF}}^{-1}(\Phi(2z_0 + z_{1-\alpha}))]. \quad (8)$$

Obviously if the estimated bias is 0, then from Eq. 6, $\hat{S}_i^* = \hat{S}_i$ and the confidence intervals produced from Eqs. 7 and 8 will be the same.

Computer Simulations

It is highly unlikely that the estimators, $\hat{S}_j$ and $\tilde{S}_j$ and $\hat{S}_j^*$, are equivalent. A preference for one of these estimators over the other should be based on statistical criteria. The criteria we have chosen to examine are the bias, variance, and mean-squared error (MSE), which equals the variance plus the squared bias, of each estimator in addition to the associated confidence intervals. We would prefer the estimator with the smallest value of all these quantities. In practice the variance of $\tilde{S}_j$ and $\hat{S}_j$ will be used to construct a confidence interval. The confidence interval on $\hat{S}_j^*$ is determined nonparametrically. The accuracy of the confidence intervals on $\hat{S}_j$ and $\tilde{S}_j$ will depend on the accuracy of the variance estimate and on how closely the estimator is distributed as $t_{[N_1 + N_2 - 2(n-1)]}$.

To get some insight into the statistical properties of our three estimators we have carried out computer simulations which we now describe. Our first series of simulations considers just the delta and jackknife estimators. The bootstrap was not evaluated for all these examples simply because of time considerations. We considered two categories and p_1 took on values of 0.1, 0.2, 0.3, 0.4, and 0.5, while q_1 was set to 0.15, 0.35, 0.55, 0.75, and 0.95. All possible combinations of these p and q values were examined, yielding 25 different populations. The values of p_1 and q_1 for each of these 25 populations defined the true value of S_j. For each population the following process was repeated 2000 times.

1) A sample of N observations ($N = 20$, 60, or 200) was taken for each species to estimate p_1 and q_1.
2) $\hat{S}_j$, $\tilde{S}_j$, $\text{var}(\hat{S}_j)$, and $\text{var}(\tilde{S}_j)$ were calculated and a 95% confidence interval was constructed for each estimator, i.e., $\hat{S}_j \pm t_{2N-2,0.025}\,(\text{var}(\hat{S}_j))^{1/2}$.

After the 2000 trials we estimated the expected value for each estimator as the mean of the 2000 values of $\hat{S}_j$ $\tilde{S}_j$. With the expected values we can estimate the bias of each estimator as $E(\hat{S}_j) - S_j$ and $E(\tilde{S}_j) - S_j$. It should be noted that since we have only an estimate of $E(\hat{S}_j)$ and $E(\tilde{S}_j)$, some estimates of the bias may not be significantly different from 0. We tested this by calculating the variance of each estimator from the expected values and the 2000 observations of $\hat{S}_j$ and $\tilde{S}_j$. These variances we denoted as $\text{var}(\hat{S}_j)$ and $\text{var}(\tilde{S}_j)$. We used these to put confidence intervals on our estimates of bias.

To examine how close the estimates of variance Eq. 2 and Eq. 4 were to the values $\text{var}(\tilde{S}_j)$ and $\text{var}(\hat{S}_j)$, we calculated the mean of the 2000 values of $\text{var}(\tilde{S}_j)$ and $\text{var}(\hat{S}_j)$, which we denoted as $\overline{\text{var}}(\tilde{S}_j)$ and $\overline{\text{var}}(\hat{S}_j)$. Lastly, we kept track of the number of times our calculated confidence intervals included S_j. With this information we calculated the empirical confidence level of these intervals and compared them to the expected level of 95%.

It is clear from the analytic and numerical results of Smith and Zaret (1982) that the bias of these estimators is liable to be worse when the number of resource categories is large. We have carried out simulations similar to the ones described previously on two different simulated populations with 10 resource categories. These examples are identical to ones used by Smith and Zaret (1982) and the underlying p_i's and q_i's are given in Table 2. For each example we let $N = 20$, 60, and 200. Only the results for $N = 20$ are shown in Table 2 because the results for $N = 60$ and 200 are qualitatively similar.

Perhaps the most restrictive assumption used in both the jackknife and delta methods is that the count vector has a multinomial distribution. To examine how our estimators behave when the underlying distribution is not multinomial we conducted several additional simulations. In particular, we considered a population that was "contaminated" by a second subpopulation which samples the environment differently. For instance, suppose we were determining the overlap in diets of two species of lizards. Individuals were sampled at random and their gut contents examined. We might then classify the insects into various genera. Usually the results from each lizard would be pooled to get our final estimate of $\boldsymbol{p}$ and $\boldsymbol{q}$. It is possible, however, that some fraction of the individuals we sampled were actually choosing items with different probabilities from the rest of the population (see Mosteller and Tukey 1977: 17, for a brief discussion of contaminated distributions). We have studied this problem with the following sort of numerical simulations. We restricted ourselves to two categories. We assumed a base population for each species composed of individuals which sample the environment with probabilities $\boldsymbol{p}_{\text{I}}$ and $\boldsymbol{q}_{\text{I}}$. We called these individuals type I. We also assumed a fraction C of all individuals sampled the environment with probabilities $\boldsymbol{p}_{\text{II}}$ and $\boldsymbol{q}_{\text{II}}$. We called these individuals type II. Random samples from this population were generated as follows. Ten individuals of each species were sampled and the probability that type II individuals were included in the sample was C. Each individual was then given its own sample of 20 resource items; the probabilities used to generate these samples depended on whether a given individual was type I or type II. Since there were only two resource categories, each individual in the sample could be characterized by the frequency of category one items it contained in its sample of 20 resource items. For species one, let the frequencies of resource item one for each individual be $x_1, x_2, \ldots, x_{10}$ and for species two, $y_1, y_2, \ldots, y_{10}$. We conducted the bootstrap sampling on this set of data while keeping in mind that we can no longer distinguish type I from type II individuals. Each bootstrap replicate sampled 10 individuals, with replacement, for each species. From each individual a sample of 20 resource items was chosen: the probability of the item being from category one was determined by the par-

TABLE 1. Summary of 25 different simulations comparing the jackknife and delta estimators of four overlap measures: the coefficient of community (S_1), Morisita's index (S_2), Horn's index (S_3), and the Euclidian distance (S_4). For S_1 and S_4, the delta and jackknife estimators are equal for two categories, hence the entries of 50 for bias, variance, and MSE.

	Sample size											
	200				60				20			
	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4
Statistic*	Percent of all cases											
Bias	50	88	96	50	50	96	100	50	50	96	100	50
Variance	50	24	100	50	50	32	100	50	50	32	80	50
MSE	50	32	100	50	50	40	100	50	50	44	84	50
CI	48	64	100	48	66	48	100	66	86	12	100	86
Variance estimate†	32	100	88	32	32	100	80	48	32	100	68	48
Bias significantly >0‡												
Jackknife	24	20	28	24	28	28	28	28	36	40	44	36
Delta	24	60	96	24	28	68	100	28	36	84	100	36

* Percent of all cases in which the jackknife had a smaller bias, variance, and mean squared error than the delta estimator, and percent of all cases in which the confidence level of the jackknife measure was closer than the delta confidence level to 95%.

† Percent of all cases in which the jackknife estimate of variance was closer than the delta estimate to the observed value.

‡ Percent of all cases in which the bias was significantly greater than 0 is listed for both the jackknife and delta method. Ties to within 10^{-9} are counted half.

ticular x_i or y_i. Thus, if individual j was chosen from species one, the probability of a resource item being in category one is x_j. The final samples have $N_1 = N_2 = 200$. We made 1000 replicates to get one bootstrap estimate of each S_j. When these estimates of S_j were made, a new cycle was begun by generating a new sample of 10 individuals for each species with the simultaneous classification of each individual as either type I or type II. A total of 1000 cycles was completed for each value of C in Table 3.

RESULTS

A summary of the results for the 25 two-category simulations which examined the jackknife and delta estimators only is given in Table 1. From examination of Table 1 it is clear that the jackknife is effective at reducing the bias for S_2 and S_3. This result is consistent with the form of the bias for S_2 and S_3 and the known properties of the jackknife. These same two overlap statistics also seem to have significant bias in a greater

TABLE 2. Niche overlap statistics with multiple categories. Given below are the percent bias, variance, mean squared error (MSE) and actual confidence level (C.L.) of supposed 95% confidence intervals for the delta ($\widehat{S}_i$) and jackknife ($\tilde{S}_i$) estimators.

Estimator	Example*	% bias	Var(S_i)	MSE	C.L.
$\widehat{S}_1$	E1	19.5 ± 1.0	0.0126	0.0221	85.5 ± 1.5
	E2	10.6 ± 1.1	0.0124	0.0147	92.9 ± 1.1
$\tilde{S}_1$	E1	6.5 ± 1.9	0.0450	0.0461	87.1 ± 1.8
	E2	15.1 ± 2.0	0.0428	0.0474	80.4 ± 1.7
$\widehat{S}_2$	E1	16.4 ± 1.3	0.0242	0.0320	91.4 ± 1.2
	E2	15.4 ± 1.2	0.0244	0.0321	92.9 ± 1.1
$\tilde{S}_2$	E1	2.6 ± 1.6	0.0371	0.0373	86.8 ± 1.5
	E2	1.6 ± 1.5	0.0357	0.0358	87.9 ± 1.4
$\widehat{S}_3$	E1	22.1 ± 0.8	0.0160	0.0423	61.4 ± 2.1
	E2	21.9 ± 0.8	0.0148	0.0387	58.9 ± 2.2
$\tilde{S}_3$	E1	3.5 ± 0.9	0.0227	0.0234	92.3 ± 1.2
	E2	4.5 ± 0.9	0.0225	0.0235	90.8 ± 1.3
$\widehat{S}_4$	E1	8.0 ± 0.4	0.00445	0.00767	86.4 ± 1.5
	E2	7.5 ± 0.4	0.00415	0.00715	89.6 ± 1.3
$\tilde{S}_4$	E1	0.4 ± 0.5	0.00627	0.00628	93.5 ± 1.1
	E2	0.04 ± 0.5	0.00581	0.00581	93.6 ± 1.1

* Two examples (E1 and E2) with 10 resource categories with $N = 20$ were tested with the following underlying probabilities:

Example 1. $\boldsymbol{p}^T = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)$
$\boldsymbol{q}^T = (0.4, 0.3, 0.1, 0.1, 0.05, 0.01, 0.01, 0.01, 0.01, 0.01)$;

Example 2. $\boldsymbol{p}^T = (0.2, 0.05, 0.05, 0.05, 0.1, 0.05, 0.2, 0.1, 0.1)$
$\boldsymbol{q}^T = (0.4, 0.3, 0.1, 0.1, 0.05, 0.01, 0.01, 0.01, 0.01)$.

portion of the populations and of a greater magnitude than S_1 or S_4.

The delta estimator had smaller variance than the jackknife for S_2, while the jackknife had the smaller variance for S_3, and for S_1 and S_4 they are equal. The jackknife always generated more accurate confidence intervals for S_3, and was generally more accurate for S_1 and S_4 than the delta estimator, but lost this lead with larger sample sizes. For S_2, the jackknife was better for large samples, but much worse for small samples. The jackknife appears to give more accurate variance estimates for S_2 and S_3, while the delta estimates seem superior for S_1 and S_4. In our simulations, the jackknife estimate of variance was usually biased upwards, which is consistent with the analytic predictions of Efron and Stein (1981). Although neither the jackknife nor the delta method is better for all measures of niche overlap, the jackknife estimator appears to be better for S_3, while the delta method may be slightly better for S_1 and S_4. The results for S_2 are equivocal.

Interestingly, most conclusions for the two-category comparisons hold for the multiple-category comparisons (Table 2). The jackknife outperforms the delta method for S_3 in all categories except variance. However, in spite of this, because of the substantial reduction in bias the jackknife has a smaller mean-squared error. The delta method does uniformly better than the jackknife for S_1 with neither estimator having an advantage in bias reduction. The one difference between the two- and multiple-category results is for S_4. With S_4 the jackknife performs better in all categories except the variance. The results for S_2 are again equivocal. Although the jackknife has a decided advantage in bias reduction, its variance and mean-squared error are larger than the delta method's. The data are not shown in Table 2, but results from additional computer simulation show that the delta estimator has more accurate confidence intervals at sample sizes of 20 and 60 but loses this advantage at the highest sample size (200).

In Table 3 we present the results of the simulations involving various degrees of contamination (0, 0.10, or 0.25). Since the results for the Euclidian distance were identical to the coefficient of community, only the latter results are presented. The most striking result in Table 3 is that even low levels of contamination can severely affect the accuracy of confidence intervals produced by either the jackknife or delta methods despite the large sample sizes. In marked contrast are the confidence intervals of the bootstrap, which are exceedingly accurate. There are also a few instances where the jackknife or delta estimators are biased while the bootstrap is unbiased (within the accuracy of our simulations). The bootstrap invariably has a higher variance and MSE: however, it is larger by only 3–11%.

TABLE 3. The percent bias, variance, mean squared error (MSE), and confidence level (C.L.) for the delta ($\widehat{S}_i$), jackknife ($\tilde{S}_i$) and bootstrap ($\widehat{S}_i^*$) estimators of the Euclidian distance. In all cases $N = 200$.†

Estimator	% bias	Var(S_i)	MSE	C.L.
A. Contamination = 0.				
$\widehat{S}_1$	0.06	0.00142	0.00142	95.5 ± 1.3
$\tilde{S}_1$	0.06	0.00142	0.00142	95.5 ± 1.3
$\widehat{S}_1^*$	0.4	0.00147	0.00147	98.9 ± 0.6
$\widehat{S}_2$	0.1	0.00235	0.00235	95.5 ± 1.3
$\tilde{S}_2$	0.1	0.00236	0.00236	95.6 ± 1.3
$\widehat{S}_2^*$	0.5	0.00244	0.00244	98.5 ± 0.8
$\widehat{S}_3$	0.4‡	0.00177	0.00177	95.1 ± 1.3
$\tilde{S}_3$	0.03	0.00176	0.00176	95.2 ± 1.3
$\widehat{S}_3^*$	0.1	0.00189	0.00189	98.5 ± 0.8
B. Contamination = 0.10.				
$\widehat{S}_1$	0.1	0.00535	0.00535	72.7 ± 2.8
$\tilde{S}_1$	0.1	0.00535	0.00535	72.7 ± 2.8
$\widehat{S}_1^*$	0.3	0.00534	0.00534	95.4 ± 1.3
$\widehat{S}_2$	0.008	0.00856	0.00856	72.3 ± 2.8
$\tilde{S}_2$	0.03	0.00866	0.00866	72.2 ± 2.8
$\widehat{S}_2^*$	0.2	0.00898	0.00898	94.9 ± 1.4
$\widehat{S}_3$	0.9‡	0.00363	0.00368	72.4 ± 2.8
$\tilde{S}_3$	0.6‡	0.00362	0.00364	72.8 ± 2.8
$\widehat{S}_3^*$	0.04	0.00380	0.00380	95.0 ± 1.4
C. Contamination = 0.25.				
$\widehat{S}_1$	0.005	0.00914	0.00914	65.5 ± 2.9
$\tilde{S}_1$	0.005	0.00914	0.00914	65.6 ± 2.9
$\widehat{S}_1^*$	0.4	0.00955	0.00956	95.1 ± 1.3
$\widehat{S}_2$	1.0‡	0.00939	0.00946	64.7 ± 3.0
$\tilde{S}_2$	0.7	0.00962	0.00965	64.4 ± 3.0
$\widehat{S}_2^*$	0.3	0.0106	0.0106	95.4 ± 1.3
$\widehat{S}_3$	0.9‡	0.00222	0.00229	66.6 ± 2.9
$\tilde{S}_3$	0.7‡	0.00221	0.00225	65.9 ± 2.9
$\widehat{S}_3^*$	0.1	0.00235	0.00235	95.5 ± 1.3

† We assume two categories of individuals that are sampled as follows:

Type I individuals: $p_1 = 0.80$, $q_1 = 0.15$
Type II individuals: $p_1' = 0.15$, $q_1' = 0.80$,

where p and q are the probabilities of two species utilizing resource 1.

‡ Significant bias.

DISCUSSION

Ricklefs and Lau (1980) suggested that computer simulations be used to estimate confidence intervals or conduct hypothesis tests on measures of niche overlap. This suggestion is in essence the bootstrap estimator we have examined. We have expanded the analysis of Ricklefs and Lau by examining the statistical properties of the bootstrap estimator in addition to jackknife and delta estimators.

It is evident that both the jackknife and bootstrap are effective at reducing bias. It appears, however, that

there is a cost to reducing bias in the form of increased variance of the estimator.

A basic assumption in this and other work (Ricklefs and Lau 1980, Smith 1982, Smith and Zaret 1982) is that the count vector of niche categories has a multinomial distribution. We have considered deviations from this assumption by examining populations in which one multinomial distribution is contaminated with a second to various degrees. There are a variety of circumstances where such distributions are likely to be found. If there is genetically based variation for resource preference, then the resulting population might be a mixture of several multinomial populations. Contaminated distributions might also arise if individual resource use is dependent on some form of preconditioning or experience. In any case, the bootstrap has a tremendous advantage when there is a contaminated distribution. This advantage is due to the nonparametric confidence intervals that the bootstrap generates. Consequently, if it seems unwise to make the multinomial assumption for a given data set, the bootstrap would be the best method for estimating niche overlap. Implementation of the bootstrap requires access to computing facilities, while the delta and jackknife techniques can be carried out with calculators.

ACKNOWLEDGMENTS

We thank Bradley Efron, Greg Murphy, Jack Reeves, Frank Yellin, Brian Maurer, and two anonymous referees for helpful comments on this work. This research was supported by National Institutes of Health (NIH) grant GM 28016 to M. W. Feldman, NIH-National Service Award GM 07310 from the National Institute of General Medical Sciences and an NIH training grant in integrative biology.

LITERATURE CITED

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, Massachusetts, USA.

Burnham, K. P., and W. S. Overton. 1979. Robust estimation of population size when capture probabilities vary among animals. Ecology **60**:927–936.

Efron, B. 1979*a*. Bootstrap methods: another look at the jackknife. Annals of Statistics **6**:1–26.

———. 1979*b*. Computers and the theory of statistics: thinking the unthinkable. Society for Industrial and Applied Mathematics Review **21**:460–480.

———. 1981. Nonparametric standard errors and confidence intervals. Canadian Journal of Statistics **9**:139–172.

Efron, B., and C. Stein. 1981. The jackknife estimate of variance. Annals of Statistics **9**:586–596.

Hurlbert, S. L. 1978. The measurement of niche overlap and some relatives. Ecology **59**:67–77.

Jaksic, F. M., H. W. Greene, and J. L. Yanez. 1981. The guild structure of a community of predatory vertebrates in central Chile. Oecologia (Berlin) **49**:21–28.

Karlin, S., and H. M. Taylor. 1975. A first course in stochastic processes. Second edition. Academic Press, New York, New York, USA.

Maurer, B. A. 1982. Statistical inference for MacArthur-Levins niche overlap. Ecology **63**:1712–1719.

Miller, R. G. 1974. The jackknife—a review. Biometrika **61**:1–15.

Mosteller, F., and J. W. Tukey. 1977. Data analysis and regression. Addison-Wesley, Menlo Park, California, USA.

Mueller, L. D. 1979. A comparison of two methods for making statistical inferences on Nei's measure of genetic distance. Biometrics **35**:757–763.

Orians, G. H., and H. Horn. 1969. Overlap in foods and foraging of four species of blackbirds in the potholes of central Washington. Ecology **50**:930–938.

Paine, T. D., M. C. Birch, and P. Svihra. 1981. Niche breadth and resource partitioning by four sympatric species of bark beetles (Coleoptera: Scolytidae). Oecologia (Berlin) **48**: 1–6.

Petraitis, P. S. 1979. Likelihood measures of niche breadth and overlap. Ecology **60**:703–710.

Pianka, E. R. 1974. Niche overlap and diffuse competition. Proceedings of the National Academy of Sciences (USA) **71**:2141–2145.

Pulliam, H. R., and F. Enders. 1971. The feeding ecology of five sympatric finch species. Ecology **52**:557–566.

Ricklefs, R. E., and M. Lau. 1980. Bias and dispersion of overlap indices: results of some Monte Carlo simulations. Ecology **61**:1019–1024.

Schoener, T. W. 1970. Non-synchronous spatial overlap of lizards in patchy habiats. Ecology **51**:408–418.

Smith, E. P. 1982. Niche breadth, resource availability, and inference. Ecology **63**:1675–1681.

Smith, E. P., and T. M. Zaret. 1982. Bias in estimating niche overlap. Ecology **63**:1248–1253.

van Belle, G., and I. Ahmad. 1974. Measuring affinity of distributions. Pages 651–668 *in* F. Proschan and R. J. Serfling, editors. Reliability and biometry. Society for Industrial and Applied Mathematics (SIAM) Publications, Philadelphia, Pennsylvania, USA.

Zahl, S. 1977. Jackknifing an index of diversity. Ecology **58**: 907–913.