

HUTCHINSONIAN RATIOS AND STATISTICAL POWER

JONATHAN B. LOSOS^{1,2}, SHAHID NAEEM^{2,3} AND ROBERT K. COLWELL^{2,4}¹*Museum of Vertebrate Zoology and* ²*Department of Zoology, University of California, Berkeley, CA 94720*

Received January 26, 1988. Accepted May 22, 1989

Hutchinson's (1959) "Homage to Santa Rosalia or Why Are There so Many Kinds of Animals?", coupled with MacArthur and Levins' (1967) theory of limiting similarity, established the idea that competition prevents coexistence of species that are morphologically too similar (see also Carothers [1986]). Subsequently, numerous authors have reported either that the ratio of sizes of adjacent species in a size-ordered assemblage tends toward constancy or that all ratios exceed some minimum "limiting similarity."

Simberloff and Boecklen (1981) constructed null models to test 28 claims in the literature that species assemblages exhibited either unusual constancy or unusually large minimum ratios. In most cases, they found that observed size-ratio distributions were indistinguishable from those generated by the null model. Their paper has been widely cited in support of the position that little evidence exists for such "Hutchinsonian ratios."

Like many tests of null models proposed in the competition debate (see Grant and Abbott, 1980; Diamond and Gilpin, 1982; Case, 1983; Case and Sidell, 1983; Colwell and Winkler, 1984; Schoener, 1984), the methods used by Simberloff and Boecklen (1981) have been criticized for deficiency of statistical power (see Toft and Shea, 1983). In other words, the conjecture is that they may err too strongly on the side of accepting the null hypothesis when it is false (i.e., they commit frequent type-II errors). In the case of size ratios, sets of ratios that are more constant than expected by chance and minimum ratios that are larger than expected by chance would wrongly be declared merely random at an unacceptably high rate.

In this note, we confirm this conjecture and reanalyze Simberloff and Boecklen's (1981) findings by combining results from samples within studies, thereby increasing statistical power. We will show that evidence for Hutchinsonian ratios is, in many cases, quite strong, even on the basis of Simberloff and Boecklen's own tests.

Combining Results from Samples Within Studies

Under appropriate circumstances, results from independent tests of the same hypothesis may be combined (Sokal and Rohlf, 1981 pp. 779–782), which effectively increases sample size while avoiding the doubtful tactic of directly pooling data of diverse origins. Here, we apply this procedure to the results reported

by Simberloff and Boecklen (1981) from their tests for unusual constancy of morphological size ratios and unusually large minimum size ratios in assemblages of co-occurring species.

Constancy of Size Ratios.—When co-occurring species are ordered by body size, are the size ratios between adjacent species unexpectedly constant? Simberloff and Boecklen (1981) analyzed data from the literature to answer this question. To determine whether the data support a claim of ratio constancy, they employed the Barton-David statistic (Barton and David, 1956) to compare three sets of size ratios (smallest vs. largest, smallest vs. second-largest, and second-smallest vs. largest) for each species assemblage in a study and then tabulated the proportion of the assemblages in which a majority (two or all three) of these test statistics were significant. Simberloff and Boecklen (1981) considered that a claim of ratio constancy was generally supported in a study if, in a majority of the assemblages, a majority of the test statistics was significant.

Of the 21 studies analyzed by Simberloff and Boecklen (1981) to assess ratio constancy, only four support rejection of the null hypothesis by the above criteria (i.e., the hypothesis that size ratios are no more constant than expected at random) at the $P = 0.05$ level for type-I error. When they allow a more generous significance level ($P = 0.30$), the null hypothesis is rejected in an additional nine studies (due to an error in enumeration, Simberloff and Boecklen [1981] reported only seven; see their table 2 and our Table 1).

In 12 of the 21 studies, the original authors presented data for more than one assemblage of co-occurring species. The null hypothesis was rejected by Simberloff and Boecklen's criteria in only 1 of these 12 studies at the $P = 0.05$ level, and in another six studies at the $P = 0.30$ level.

With few exceptions, each assemblage within a study represents an independent test of the hypothesis that ratios are unusually constant. Consequently, for each study, it is appropriate to assess the overall probability that the assemblages taken together reveal significant patterns by combining probabilities of independent tests of the same hypothesis. By not doing so, Simberloff and Boecklen (1981) neglected a means of considerably lowering type-II error, while in the same spirit substituting the much less precise "majority rule" approach outlined above to "combine" results within studies.

Here, we have used Fisher's method for combining probabilities (Sokal and Rohlf, 1981 pp. 779–782),

which takes advantage of the fact that $2 \sum_{i=1}^k \ln P_i$ exhibits a chi-square distribution ($d.f. = 2k$), where P_i is the probability of type-I error for the i th independent test.

³ Present address: Department of Biology, University of Michigan, Ann Arbor, MI 48109.

⁴ Present address: Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269-3042.

TABLE 1. Claims of ratio constancy based on studies of multiple assemblages. Unless otherwise noted, references are those making the claim of ratio constancy. If the claim was made for data other than their own, the source of the data is presented in parentheses. Results from Simberloff and Boecklen (1981) are presented as [number of assemblages in which a majority of test statistics were significant]/[number of assemblages]. Symbols: y = claim supported; n = claim not supported.

Reference	Claim supported			
	Simberloff and Boecklen (1981)		Present paper	
	$P < 0.05$	$P < 0.30$	$P < 0.05$	$P < 0.30$
McNab, 1971 (Phillips, 1968)	2/11	5/11	y	y
McNab, 1971 ^a (Goodwin and Greenhall, 1961)	0/2	1/2	n	n
Brown, 1973	2/14	12/14	y	y
Brown, 1975	1/2	2/2	y ^b	y
May, 1978 (Cody, 1974)	0/11	2/11	n	n
Terborgh et al., 1978	1/5	5/5	y	y
MacArthur, 1972; Brown, 1975 (Storer, 1966) ^c	1/1	1/1	y	y
Barbour, 1973 ^a	0/3	2/3 ^d	n	n
Brown, 1975 ^a (Roughgarden, 1974)	0/2	0/2	n	n
Smith, 1978	0/2	0/2	n	n
Brown, 1975 (Evans, 1970)	0/2	1/2	n	y
Inouye, 1977	1/3	1/3	y	y

^a Author did not actually make a claim of ratio constancy. McNab (1971) discussed ratio constancy but never claimed that these data displayed it. Barbour (1973) claimed that there were consistent differences in absolute lengths, not ratios, between species (though Brown [1975] did make a claim of ratio constancy for these data). Brown (1975) never mentions Roughgarden (1974).

^b Both assemblages have the same smallest and second-largest species. If the computation for $G_{1,n-1}$ uses this combination only once, the combined probability changes from $P < 0.05$ to $P < 0.10$.

^c Data for males and females are presented by Storer (1966) and are combined using Fisher's test.

^d Simberloff and Boecklen (1981) incorrectly state this as 1/3 (see their tables 1 and 2).

When the results for each of the 12 multiassemblage studies presented by Simberloff and Boecklen (1981) are analyzed by combining results within studies, six (instead of one) of the 12 claims are significant at the $P = 0.05$ level, and a seventh is significant at the $P = 0.30$ level (Table 1). We use the criterion of Simberloff and Boecklen that two or three of the three ratio comparisons (smallest vs. largest, smallest vs. second-largest, and second-smallest vs. largest) must be significant to consider the results of a given study significant at the $P = 0.05$ or $P = 0.30$ level. The three comparisons cannot be combined, because they are not independent.

Minimum Size Ratios.—Many studies have purported to demonstrate that an unusually large minimum size ratio exists among members of an assemblage. Simberloff and Boecklen (1981) used the Irwin

statistic (Irwin, 1955) to test 18 such claims against the null hypothesis that the minimum in each assemblage is no larger than expected at random. Only one of the 18 claims was generally sustained at the $P = 0.05$ level; another 12 were supported at the 0.30 level.

Again, Fisher's method for combining probabilities can be employed in those studies that involved multiple assemblages. Of ten such studies, Simberloff and Boecklen found one generally supported at the $P = 0.05$ level and five more supported at the $P = 0.30$ level (Table 2). In contrast, after combining results from different assemblages within studies, we found that six of the ten studies were significant at the $P = 0.05$ level and that another was significant at the $P = 0.30$ level. Hutchinson's (1959) data on finches from three of the Galápagos Islands were not combined,

TABLE 2. Claims of unusually high minimum size ratios based on studies of multiple assemblages (format as in Table 1).

Reference	Claim supported			
	Simberloff and Boecklen (1981)		Present paper	
	$P < 0.05$	$P < 0.30$	$P < 0.05$	$P < 0.30$
McNab, 1971 (Phillips, 1968)	3/11	5/11	y	y
McNab, 1971 ^a (Goodwin and Greenhall, 1961)	0/2	1/2	n	n
Brown, 1973	3/14	11/14	y	y
Brown, 1975	1/2	2/2	y	y
May, 1978 (Cody, 1974)	1/11	3/11	n	n
Terborgh et al., 1978	1/5	5/5	y	y
Brown, 1975 ^a (Barbour, 1973)	0/3	0/3	n	n
Inouye, 1977	1/3	3/3	y	y
Uetz, 1977	0/2	2/2	n	y
Robison, 1975	3/3	3/3	y	y

^a Author did not make a claim of unusually large minimum size ratio.

because Simberloff and Boecklen (1981) make the plausible claim that such data were "dredged" (sensu Selvin and Stuart, 1966) from a larger data set.

Discussion on Combining Results.—Two objections could be raised to using Fisher's method. Simberloff (1983), Simberloff and Boecklen (1981), Simberloff and Connor (1981), and Boecklen and NeSmith (1985) argue that only positive findings are published; consequently, combining probabilities across separate studies is inappropriate, because negative findings are less likely to be reported. This objection, however, is far less valid for combining probabilities computed for independent assemblages within studies. It is unreasonable to assume that authors omit data from assemblages within a study that do not support their claim. Examination of the papers reveals, furthermore, that "negative" data are indeed reported alongside those that support the authors' claims.

Using Fisher's method might also be inappropriate if the assemblages within studies were composed of the same species, because the samples then might not be independent (Brown and Bowers, 1985). We examined the original data and determined the number of cases in which comparisons between different assemblages included the same species (i.e., the same pair of species constituted the smallest, second smallest, second largest, or largest ratio in both assemblages). We found only a few such cases, and for each we recalculated the combined probabilities omitting redundancies. One study reversed from significance to nonsignificance at the $P = 0.05$ level in the ratio-constancy results (Table 1), but there were no other changes in those results, and there were no changes for the minimum-ratio analysis.

We then checked the five purported claims of ratio constancy that were not supported even at the $P = 0.30$ level. In three of these papers (McNab, 1971; Barbour, 1973; Roughgarden, 1974; see Table 1), no such claims could be discovered. Likewise, no claims of unusually large minimum size ratios were made for two of the studies analyzed by Simberloff and Boecklen (McNab, 1971; Brown, 1975); neither rejected the null hypothesis, even at $P = 0.30$.

In summary, statistical inference based on individual small assemblages tends strongly toward incorrect acceptance of the null hypothesis. Our reanalysis of the cases that Simberloff and Boecklen (1981) analyzed reveals that, although some claims (and nonclaims) of ratio constancy are unfounded, nearly half (47%) are supported at the $P = 0.05$ level, and a large majority (more than 70%) are supported at the $P = 0.30$ level that Simberloff and Boecklen (1981) considered.

Statistical Power of the Irwin Test for Minimum Size Ratios

The Irwin test detects an unexpectedly large minimum size ratio by comparing the minimum ratio for a real-world species assemblage with the corresponding minimum ratio expected for a sample drawn from a log-uniform distribution. Several authors (Case et al., 1983; Colwell and Winkler, 1984; Schoener, 1984; Tonkyn and Cole, 1986) suggested that, because the test is based on a log-uniform rather than a log-normal distribution of species' sizes, it overestimates the expected minimum size ratio. These critics point out that

body sizes (or other homologous sets of measurements) are almost invariably log-normally distributed within clades or assemblages in nature, as Eadie et al. (1987) have recently confirmed in a large survey. If the log-normal distribution is, indeed, the more biologically appropriate null model for size distributions, then points in the distribution will be clumped around the mode. Consequently, the expected minimum size ratio drawn from such a distribution would be smaller than one drawn from a log-uniform distribution. Thus, an empirical minimum size ratio larger than expected for a log-normal distribution might nonetheless test out *smaller* than expected for the log-uniform distribution. Further, the variance of the log-normal distribution should affect the outcome of the test.

Boecklen and NeSmith (1985) claimed to have refuted this conjecture, however, by demonstrating that the results from the Irwin test were nearly indistinguishable from those produced by a simulated draw from a log-normal distribution (the variance of which was not reported). Further, although they showed analytically that changing the variance of the log-normal distribution does not alter the expected distribution of the Barton-David statistic for ratio constancy, they merely asserted, without proof, that "the same result will obtain for the Irwin test statistic" (Boecklen and NeSmith, 1985 p. 697). Eadie et al. (1987), however, showed Boecklen and NeSmith to be mistaken, by confirming the earlier conjecture that, if the log-normal distribution is substituted for the log-uniform distribution, the variance of the log-normal distribution is positively correlated with the expected minimum size difference.

Type-I Error.—Eadie et al. (1987) did not address the adequacy of the Irwin test as a method of detecting unexpectedly large minimum size ratios. We undertook this task by conducting a series of simulations, drawing three-, four-, five-, and six-species communities (a large number of the assemblages in the studies examined by Simberloff and Boecklen [1981] contained 3–6 species) at random from a log-normal distribution with variance (of the logs) increasing from 0.02 to 0.10 in 0.01 increments. We arbitrarily used a mean of 100; as Eadie et al. (1987) noted, changes in the mean do not affect minimum ratios.

For each sample drawn, we calculated the critical value that the minimum ratio would have to exceed in order to reject the null hypothesis at the $P = 0.05$ level by the Irwin test, and we determined whether the minimum ratio in the sample exceeded this critical level. The Irwin test statistic had to be recalculated for each draw, because the statistic depends not only on the number of "species" drawn but also on the smallest and largest "species sizes" in each sample.

For each combination of species number and variance of the sampling distribution, we conducted 10 runs of 100 draws each. Because each of these draws was made at random, any rejection of the null hypothesis is a type-I error. Given that the Irwin statistics we used were computed for a type-I error rate of $P = 0.05$, the proportion of samples in which the null hypothesis is rejected should also be 0.05 if the test is accurate for a log-normal distribution. We found that the mean proportion of samples for which the Irwin test rejected the (true) null hypothesis ranged from 0.043 to 0.069 (Table 3), over the span of variances and

TABLE 3. Probability of type-I error (\pm SD) of the Irwin test at the $P < 0.05$ level when applied to randomly drawn samples from a log-normal distribution. Values shown represent probability values $\times 100$.

Variance of distribution	Number of species drawn			
	3	4	5	6
0.02	5.8 \pm 0.54	4.8 \pm 1.03	6.0 \pm 0.75	5.4 \pm 0.80
0.03	4.3 \pm 0.45	6.4 \pm 0.82	4.8 \pm 0.68	5.5 \pm 0.79
0.04	5.8 \pm 0.72	6.5 \pm 0.42	6.9 \pm 0.81	5.8 \pm 0.86
0.05	5.9 \pm 0.51	5.3 \pm 0.89	6.2 \pm 0.75	5.5 \pm 0.36
0.06	4.8 \pm 0.72	4.7 \pm 0.50	6.5 \pm 1.01	4.7 \pm 0.63
0.07	5.6 \pm 0.76	5.4 \pm 0.55	5.5 \pm 0.79	5.8 \pm 0.68
0.08	6.0 \pm 0.80	5.5 \pm 0.36	5.5 \pm 0.45	4.5 \pm 0.65
0.09	5.8 \pm 0.44	6.0 \pm 1.12	6.6 \pm 0.91	4.8 \pm 0.70
0.10	5.1 \pm 0.95	6.5 \pm 0.61	5.8 \pm 0.99	5.7 \pm 0.80

species numbers tested. Therefore, for any practical purpose, the type-I error rate of the Irwin test is accurate enough when the test is applied to log-normal distributions, even though the test is based on the log-uniform distribution. Boecklen and NeSmith (1985) had already demonstrated this result for, apparently, a single log-normal distribution of unreported variance. Our results extend the analysis to a range of defined variances.

Type-II Error.—In fact, the more important part of the original conjecture concerned type-II error (i.e., accepting a false null hypothesis), rather than type-I error. Type-II errors underestimate the degree to which minimum ratios are, in fact, larger than expected by chance.

We undertook, therefore, a direct analysis of the statistical power of the Irwin test (i.e., its ability to detect nonrandomly large minimum ratios in samples from log-normal distributions). In other words, we asked whether the Irwin statistic could detect patterns produced by some deterministic process that limits the similarity of co-occurring species.

Simulations were conducted exactly as above, but any draw containing a minimum ratio smaller than a specified threshold value was discarded and replaced with a new draw. In other words, a limit to similarity was imposed. The threshold ratio was varied between 1.1 and 1.5, and the variance of the log-normal distribution from which samples were drawn was varied as before. Combinations of very small variances and very large thresholds, however, were not tested, because the rate of rejection of draws was so excessive that simulations took too long to be practical.

Because only those samples that pass the threshold criterion are retained, the null hypothesis that the minimum size ratio in each sample is no greater than expected for a random sample from a log-normal distribution is consistently false. Therefore, any sample for which the (false) null hypothesis is accepted by the Irwin test represents a type-II error.

Figure 1 presents the proportion of samples for which the Irwin statistic incorrectly accepted the null hypothesis (i.e., the rate of type-II error (the complete matrix of simulation results is available from J. B. L. upon request). The results show that the rate of type-II error can be very high indeed but that it decreases with larger thresholds, larger numbers of species in the

sample, and lower variances of the log-normal distribution.

The qualitative effect of the threshold is expected; as with any statistic, as the size of the actual deviation from the value expected under the null hypothesis increases (the threshold, in this case), the power of the test to detect that deviation increases (type-II error decreases). Unfortunately, however, the minimum ratios in nature that are likely to be tested fall in a region of rather low statistical power.

Hutchinson (1959) proposed that sympatric species must differ in size dimensions by a factor of 1.28. Many studies have accepted this value, claimed to confirm it, or proposed different values of limiting similarity (reviewed in Simberloff [1983]). Most studies support a value of limiting similarity at or below 1.3 (but see Terborgh et al. [1978]), and at these values, the Irwin statistic is very weak. Even at higher thresholds, the Irwin statistic is reasonably powerful only under some circumstances.

The effects of species number and variance are equally troubling. Type-II error is fairly low for small variances, but the variance of real distributions is what counts. Eadie et al. (1987) collected data on the variance of the log-normal distribution from 33 studies encompassing 439 assemblages of species. From their data, we extracted the standard deviations calculated from single assemblages, discarding the values reported from data pooled from more than one assemblage. A summary appears in Table 4. Of the resulting 42 size distributions, only five have variances less than 0.03, while 26 have variances greater than 0.10. The expected type-II error for the latter studies is quite high. Consequently, little reliance can be placed on the Irwin test for many realistic cases.

Conclusions

Despite the inherently large probability of type-II error of the Irwin test for many realistic cases, we have shown that Simberloff and Boecklen's (1981) results nevertheless indicate the existence of unusually large minimum size ratios for many cases when results within studies are combined. Likewise, combining results within studies from the Barton-David test for nonrandom constancy of size ratios also produced a considerable increase in the number of cases judged signifi-

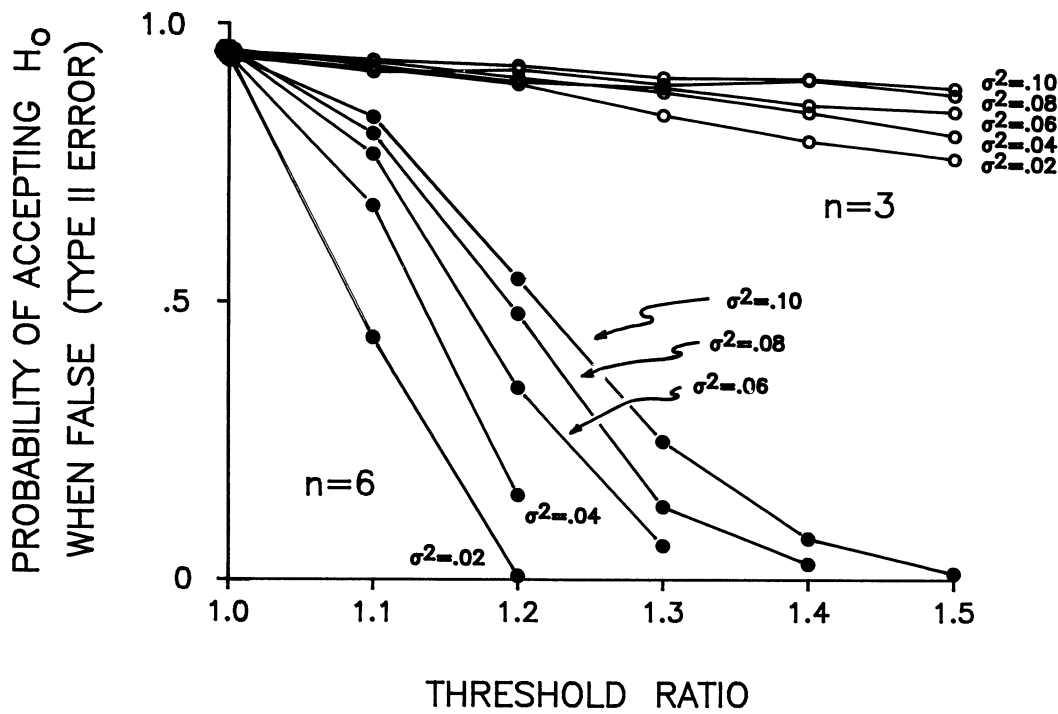


FIG. 1. Probability (P) of type-II error for the Irwin test when applied to samples drawn from log-normal distributions. If presented as $1 - P$, a standard statistical power curve would result. Draws containing ratios less than the threshold ratio were discarded. Results are presented for draws of three- and six-species sets from distributions with selected variances (σ^2). The complete matrix of simulation results is available from the authors upon request.

cant. More recent studies of size ratios (e.g., Gilpin and Diamond, 1982; Case, 1983; Case and Sidell, 1983; Case et al., 1983; Schoener, 1984; Brandl and Topp, 1985) have utilized appropriate null models and have found strong evidence for the existence of unusually large minimum size ratios.

What process or processes are responsible for these patterns is another question. Eadie et al. (1987) argue that Hutchinsonian ratios are an artifact of the underlying log-normal distribution of species' sizes. How-

ever, their results do not provide a method for testing whether ratios are either too constant or too large to be explained as the result of random processes. They demonstrate that the expected ratios between species ordered by size lie in a narrower range than one might naïvely assume and that the range includes the minimum size ratios often cited as evidence of community structure.

Boecklen and NeSmith (1985) demonstrated that, even for samples from a log-normal distribution, the Barton-David and Irwin tests maintain a correct level of type-I error. We have extended their result for the Irwin test by showing it to be robust, for practical purposes, to changes in variance of the log-normal distribution. Thus, although Eadie et al. (1987) are doubtless correct that many ratios in the range 1.2–2.0 are simply artifacts of sampling from a log-normal distribution with a low variance, artifactual ratios can nevertheless be separated from nonrandom size ratios in the same range by use of these statistical tests, with the proviso that results from related but distinct assemblages should be combined and with the realization that type-II error may nonetheless be high.

Several recent studies (e.g., Case, 1983; Case and Sidell, 1983; Case et al., 1983; Schoener, 1984; Grant, 1986) have provided strong corroborative biogeographical, ecological, and behavioral evidence to bolster the suggestion that competition is the process causing many of these nonrandom patterns. In most cases, however, detailed studies are still required to elucidate

TABLE 4. Variances of log-normal distributions in nature (calculated from Eadie et al. [1987]).

Number of species per assemblage	Number of assemblages	Variance		
		Mean	Minimum	Maximum
2	5	0.033	0.004	0.063
3	4	0.102	0.048	0.212
4	3	0.649	0.078	1.638
5	4	0.700	0.053	2.250
6	6	0.556	0.020	1.513
7	6	0.153	0.008	0.250
8	2	0.117	0.048	0.185
9	5	1.037	0.152	2.756
11	2	0.135	0.102	0.168
14	3	0.251	0.048	0.436
16	2	0.178	0.019	0.336

the processes responsible for creating unusually large minimum and constant size ratios.

Acknowledgments

We thank M. Kirkpatrick for helping with the randomization algorithm. Thanks also to B. Stein for programming assistance and to D. Ribble for computational assistance. The paper profited from discussion with E. Losos, R. Ostfeld, and T. Schoener. IBM kindly provided the computers used in this analysis. R.K.C. was supported by the U.S. National Science Foundation, grant number BSR-8604929.

LITERATURE CITED

- BARBOUR, C. D. 1973. A biogeographical history of *Chirotoma* (Pisces: Atherinidae): A species flock from the Mexican plateau. *Copeia* 1973:533-556.
- BARTON, D. E., AND F. N. DAVID. 1956. Some notes on ordered random variables. *J. Roy. Stat. Soc. B* 18:79-94.
- BOECKLEN, W. J., AND C. NESMITH. 1985. Hutchinsonian ratios and log-normal distributions. *Evolution* 39:695-698.
- BRANDL, R., AND W. TOPP. 1985. Size structure of *Pterostichus* spp. (Carabidae): Aspects of competition. *Oikos* 44:234-238.
- BROWN, J. H. 1973. Species diversity of seed-eating desert rodents in sand dune habitats. *Ecology* 54:775-787.
- . 1975. Geographical ecology of desert rodents, pp. 315-341. In M. L. Cody and J. M. Diamond (eds.), *Ecology and Evolution of Communities*. Harvard Univ. Press, Cambridge, MA.
- BROWN, J. H., AND M. A. BOWERS. 1985. Community organization in hummingbirds: Relationships between morphology and ecology. *Auk* 102:251-269.
- CAROTHERS, J. H. 1986. Homage to Huxley: On the conceptual origin of minimum size ratios among competing species. *Amer. Natur.* 128:440-442.
- CASE, T. J. 1983. Sympatry and size similarity in *Cnemidophorus*, pp. 297-325. In R. B. Huey, E. R. Pianka, and T. W. Schoener (eds.), *Lizard Ecology: Studies of a Model Organism*. Harvard Univ. Press, Cambridge, MA.
- CASE, T. J., J. FAABORG, AND R. SIDELL. 1983. The role of body size in the assembly of West Indian bird communities. *Evolution* 37:1062-1074.
- CASE, T. J., AND R. SIDELL. 1983. Pattern and chance in the structure of model and natural communities. *Evolution* 37:832-849.
- CODY, M. L. 1974. *Competition and the Structure of Bird Communities*. Princeton Univ. Press, Princeton, NJ.
- COLWELL, R. K., AND D. W. WINKLER. 1984. A null model for null models in biogeography, pp. 344-359. In D. R. Strong, Jr., D. Simberloff, L. G. Abele, and A. B. Thistle (eds.), *Ecological Communities: Conceptual Issues and the Evidence*. Princeton Univ. Press, Princeton, NJ.
- DIAMOND, J. M., AND M. E. GILPIN. 1982. Examination of the "null" model of Connor and Simberloff for species co-occurrences on islands. *Oecologia* 52:64-74.
- EADIE, J. M., L. BROEKHOVEN, AND P. COLGAN. 1987. Size ratios and artifacts: Hutchinson's rule revisited. *Amer. Natur.* 129:1-17.
- EVANS, H. E. 1970. Ecological-behavioral studies of the wasps of Jackson's Hole, Wyoming. *Bull. Mus. Comp. Zool.* 140:451-511.
- GILPIN, M. E., AND J. M. DIAMOND. 1982. Factors contributing to non-randomness in species co-occurrences on islands. *Oecologia* 52:75-84.
- GOODWIN, G. G., AND A. M. GREENHALL. 1961. A review of the bats of Trinidad and Tobago. *Bull. Amer. Mus. Nat. Hist.* 122:187-302.
- GRANT, P. R. 1986. *Ecology and Evolution of Darwin's Finches*. Princeton Univ. Press, Princeton, NJ.
- GRANT, P. R., AND I. ABBOTT. 1980. Interspecific competition, island biogeography and null hypotheses. *Evolution* 34:332-341.
- HUTCHINSON, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? *Amer. Natur.* 93:145-159.
- INOUE, D. W. 1977. Species structure of bumblebee communities in North America and Europe, pp. 35-40. In W. J. Mattson (ed.), *The Role of Arthropods in Forest Ecosystems*. Springer-Verlag, N.Y.
- IRWIN, J. O. 1955. A unified derivation of some well-known frequency distributions of interest in biometry and statistics. *J. Roy. Stat. Soc. A* 118:389-404.
- MACARTHUR, R. H. 1972. *Geographical Ecology*. Harper and Row, N.Y.
- MACARTHUR, R. H., AND R. LEVINS. 1967. The limiting similarity, convergence, and divergence of coexisting species. *Amer. Natur.* 101:377-385.
- MAY, R. M. 1978. The dynamics and diversity of insect faunas. *Symp. Roy. Entomol. Soc. Lond.* 9:188-204.
- M McNAB, B. K. 1971. The structure of tropical bat faunas. *Ecology* 52:352-358.
- PHILLIPS, C. J. 1968. Systematics of the megachiropteran bats in the Solomon Islands. *Univ. Kans. Publ. Mus. Nat. Hist.* 16:777-837.
- ROBISON, R. A. 1975. Species diversity among agnostoid trilobites. *Fossils and Strata* 4:219-226.
- ROUGHGARDEN, J. 1974. Species packing and the competition function with illustrations from coral reef fish. *Theoret. Popul. Biol.* 5:163-186.
- SCHOENER, T. W. 1984. Size differences among sympatric, bird-eating hawks: A worldwide survey, pp. 254-281. In D. R. Strong, Jr., D. Simberloff, L. G. Abele, and A. B. Thistle (eds.), *Ecological Communities: Conceptual Issues and the Evidence*. Princeton Univ. Press, Princeton, NJ.
- SELVIN, H. C., AND A. STUART. 1966. Data-dredging procedures in survey analysis. *Amer. Statist.* 20:20-23.
- SIMBERLOFF, D. S. 1983. Sizes of coexisting species, pp. 404-430. In D. J. Futuyma and M. Slatkin (eds.), *Coevolution*. Sinauer, Sunderland, MA.
- SIMBERLOFF, D. S., AND W. BOECKLEN. 1981. Santa Rosalia reconsidered: Size ratios and competition. *Evolution* 35:1206-1228.
- SIMBERLOFF, D. S., AND E. F. CONNOR. 1981. Missing species combinations. *Amer. Natur.* 118:215-239.
- SMITH, C. L. 1978. Coral reef fish communities: A compromise view. *Env. Biol. Fish* 3:109-128.

- SOKAL, R. R., AND F. J. ROHLF. 1981. *Biometry: The Principles and Practice of Statistics in Biological Research*, 2nd Ed. Freeman, N.Y.
- STORER, R. W. 1966. Sexual dimorphism and food habits in three North American accipiters. *Auk* 83: 423-436.
- TERBORGH, J., J. FAABORG, AND H. J. BROCKMANN. 1978. Island colonization by Lesser Antillean birds. *Auk* 95:59-72.
- TOFT, C. A., AND P. J. SHEA. 1983. Detecting community-wide patterns: Estimating power strengthens statistical inference. *Amer. Natur.* 122:618-625.
- TONKYN, D. W., AND B. J. COLE. 1986. The statistical analysis of size ratios. *Amer. Natur.* 128:66-81.
- UETZ, G. W. 1977. Coexistence in a guild of wandering spiders. *J. Anim. Ecol.* 46:531-542.

Corresponding Editor: N. Chr. Stenseth

Evolution, 43(8), 1989, pp. 1826-1827

IS LIFETIME DATA ALWAYS NECESSARY FOR EVALUATING THE "INTENSITY" OF SELECTION?

TAKAYOSHI NISHIDA

Entomological Laboratory, College of Agriculture, Kyoto University, Kyoto 606, JAPAN

Received April 13, 1989. Accepted May 25, 1989

Lifetime reproductive success has been regarded as the most accurate measure of fitness, and it has been argued that measurement of selection parameters (such as selection gradient, selection differential, and opportunity for selection) should be made based on lifetime data (Clutton-Brock, 1983, 1988; Endler, 1986). This argument is indeed strictly true; however, it is often logistically or practically impossible to collect lifetime data in the field, especially for highly mobile animals. In this paper, I show that lifetime data are not always required to obtain a precise estimate of selection parameters.

I measured lifetime reproductive success of individ-

uals of both sexes in the coreid bug, *Colpula lativentris* (Nishida, 1987). This bug forms a compact aggregation composed of copulating pairs, on stems of its host plant, *Polygonum cuspidatum*. Males search for mates, and females oviposit on the ground near the host plants. Because a single copulation lasts an average of three days and because the bugs are rather sedentary, lifetime mating success could be easily measured. A detailed description of the mating system and behaviors is given in Nishida (1988, 1989).

Figure 1 shows the ratio of the opportunity for selection based on a 10-day cumulative observation period to the opportunity for selection based on lifetime reproductive success. As the observation period increased, the value of this ratio decreased rapidly, almost reaching one after 40 days of observation and remaining constant after that. Thus, a 40-day observation period was sufficient to evaluate the opportunity for selection on lifetime reproductive success in both males and females.

These results suggest that lifetime data are not always required for measuring selection, especially for multiply mating insects. However, if reproduction at an early age results in a reduction of future reproduction or if there is a strong negative trade-off between fitness components, lifetime data will be required. Nonetheless, I believe that, for most insects that mate multiply in a single mating season, the precise measurement of selection parameters does not require lifetime data. In the future, by utilizing a large body of data concerning relations between selection parameters based on both lifetime data and short-term data, it should prove possible to determine empirically what period is necessary for collection of data in order to derive a valid estimation of selection parameters.

ACKNOWLEDGMENTS

I thank E. Kuno, T. Inoue, and E. Kasuya for helpful discussion.

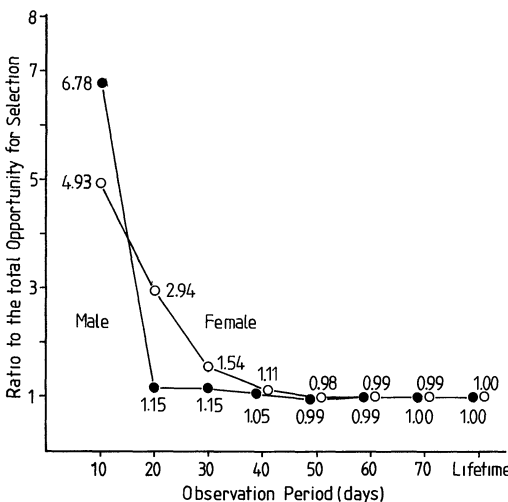


FIG. 1. The ratio of the opportunity for selection based on a 10-day cumulative observation period to the opportunity for selection based on lifetime reproductive success.