Randomization tests and computational software on statistic significance of community biodiversity and evenness^{*}

ZHANG Wen-Jun¹, QI Yan-Hong¹, Schoenly Kenneth George²

¹ Research Institute of Entomology and State Key Laboratory for Biological Control, School of Life Sciences, Department of Computer Science, Zhongshan University, Guangzhou 510275, P. R. China

² Department of Biological Sciences, California State University, Stanislaus, CA 95382, USA

Abstract : Diversity and evenness indices were widely used in community ecology and biodiversity researches. However, shortage of statistic tests on these indices restricted their reliability. To develop statistic test methods on diversity is one of the focuses in biodiversity researches. In present study, some randomization tests on statistic significance of diversity and evenness indices, confidence interval of diversity and evenness, and randomization test on statistic significance of between-community differences were presented. Shannon-Wiener diversity index, Simpson diversity index, McIntosh diversity index, Berger-Parker diversity index, Hurlbert diversity index, Brillouin diversity index, and corresponding evenness indices are included in the randomization test procedure. The web-based computational software for the statistic tests, BiodiversityTest[®], which is comprised of seven Java classes and an HTML file, is developed. It can be run on various operational systems and java-enabled web browsers and, may read ODBC linked databases such as MS Access, Excel, FoxPro, dBASE, etc. Rice arthropod diversity (15 sampling sites, 125 arthropod species, 17 functional groups) was recorded on September, 1996 in IRRI rice farm using RiceVac apparatus and bucket enclosure. The data were analysed using BiodiversityTest[®] with Shannon-Wiener index and Berger-Parker index respectively, and the results showed that the changes of diversity and evenness can be effectively detected by these tests. The randomization tests will correct the possible wrong conclusions aroused in direct comparison of arthropod diversity which was used in most of the researches up to now. The development of randomization tests on biodiversity will provide a quantitative tool for stricter statistic comparison of biodiversity between communities and present an absolute criterion for diversity measuring. BiodiversityTest[®] will make the computation realistic and accessible on Internet.

Key words : biodiversity indices, randomization test, confidence interval, statistic significance, internet software

Species diversity is a dual concept that includes the species richness in a community and evenness with which the individuals are divided among the species (Krebs, 1989). There are so many ways to index species diversity but there is no universal approach that can be recommended for all communities (Washington, 1984). Heterogeneity indices, usually refer to diversity indices, combine species richness and evenness in a single index. Nonparametric indices of diversity are commonly used, because no statistical distribution is assumed. Some indices place most weight on the rare species in the sample, such as Shannon-Wiener index, and some place most weight on the common

species such as Simpson index (Krebs, 1989; Hurlbert, 1971; McIntosh, 1967). Up to now, the biodiversity indices, owing to their simplicity, are widely used in community ecology but with carefulness when related conclusions are drawn, because there are no reliable statistic tests with these indices.

Some typical assumptions that have been made in many statistic tests are usually: (1) random sampling of individuals from the population of interest; (2) equal population standard deviations or means for the two groups tested, and (3) normal distributions for index values within groups (Manly, 1997). In community analysis, however, the data usually treat taxa as in-

dices and samples or communities as groups to be tested or compared , and complex relationships exist in different taxa. These assumptions are generally not met or only partially met. Randomization is an effective method for tests of hypothesis with community data and has been used in community studies (Manly , 1997 ; Sollow , 1993 ; Zhang & Schoenly , 1999a , 1999b , 1999c , 2001) , which has proved to be a manageable nonparametric test on statistic significance. In this study , some randomization test on significance of diversity and evenness indices , confidence interval of diversity and evenness , randomization test on significance of between-community differences are presented. The web-based computational software for the tests is developed to provide solutions. Rice arthropod diversity , which was recorded in IRRI farm , is analyzed using the tests and software developed.

Randomization tests

Suppose that there are m taxa in a community , the number of individuals for each taxon is x_i , $i = 1 \ 2 \ \dots \ , m$. Let

$$s = \sum_{i=1}^m x_i$$

The Shannon-Wiener diversity index is expressed as

$$S_w = \sum_{i=1}^m [x_i/s \ln (x_i/s)]$$

Similarly , the Simpson diversity index , McIntosh diversity index , Berger-Parker diversity index , Hurlbert diversity index , and Brillouin diversity index are computed as

$$S_p = 1 - \sum_{i=1}^m [x_i(x_i - 1) / (s (s - 1))]$$

$$M_c = [s - (\sum_{i=1}^m x_i^2)^{1/2}] / (s - s^{1/2})$$

$$B_p = 1 - \max_i x_i / s$$

$$H_u = \sum_{i=1}^m [x_i (s - x_i) / (s (s - 1))]$$

$$B_r = \log [s ! / (x_1 ! x_2 ! \dots x_m !)] / s$$

The maximal values for S_w , S_p , M_c , B_p , H_u , and B_r are $\ln (m)$, $1 - (s - m) / (m (s - 1))$, $(s - s^{1/2}) / (s - s^{1/2})$, $1 - s (1 - 1/m) / (s - 1)$, and $\log [s ! / ((u !)^{m-v} ((u + 1) !)^v)] / s$, respectively , where $s/m = u + v/m$, and $u \ p$ are integers , $v < m$.

The evenness index based on Shannon-Wiener diversity index (S_p) , Simpson diversity index (S_p) , McIntosh diversity index (M_c) , Berger -Parker diversity index (B_p) , Hurlbert diversity index (H_u) , and Brillouin diversity index (B_r) can thus be calculated

as the ratio of indexd diversity to maximal diversity , and represented as E_w , E_p , E_m , E_b , E_h , E_r , where $0 \leq E_w \leq 1$, $0 \leq E_p \leq 1$, $0 \leq E_m \leq 1$, $0 \leq E_b \leq 1$, $0 \leq E_h \leq 1$, $0 \leq E_r \leq 1$.

Test on significance of diversity and evenness

For the test of statistic significance on diversity and evenness , the s individuals are randomly partitioned into m taxa , i. e. , make a randomization. Compute the diversity or evenness for this partition and compare whether the expected diversity or evenness is lager than the observed. For all randomizations , count the number of the expected > the observed , and obtain the percentage of the number *vs* randomization times as statistic p value. The threshold p value for the test may be defined as 0.1 , 0.05 , 0.01 , etc. If the computed p value is greater than the p threshold , the diversity or evenness may be considered to be statistically significant at this confidence level.

Confidence interval of diversity and evenness

In randomization procedure , the s individuals are randomly partitioned into m taxa each with the same proportion of individuals as the true proportion of individuals for this taxa. For all times of randomizations , compute the expected diversity or evenness , and their standard deviation. Through this procedure the confidence interval for expected diversity or evenness can be obtained for the community.

Test on statistic significance of between-community differences

For community i and j , $i = 1 \ 2 \ \dots \ n - 1 ; j \geq i$; where n is the number of communities , compute the observed diversity difference dw_{ij} , dp_{ij} , dm_{ij} , db_{ij} , dh_{ij} , dr_{ij} , and observed absolute diversity difference dw'_{ij} , dp'_{ij} , dm'_{ij} , db'_{ij} , dh'_{ij} , dr'_{ij} , $i = 1 \ 2 \ \dots \ , n - 1 ; j > i$. In the same way , observed evenness difference ew_{ij} , ep_{ij} , em_{ij} , eb_{ij} , eh_{ij} , er_{ij} , and observed absolute evenness difference ew'_{ij} , ep'_{ij} , em'_{ij} , eb'_{ij} , eh'_{ij} , er'_{ij} , $i = 1 \ 2 \ \dots \ , n - 1 ; j > i$, should be computed.

Use randomization technique , i. e. , for community i and j , $i = 1 \ 2 \ \dots \ , n - 1 ; j > i$; the two data sets are combined into a single set consisting of $s_1 + s_2$ individuals each labelled by its taxa , where s_1 , s_2 are individual number of community i and j respectively. The combined data set is randomly partitioned into two subsets

of $s_1 + s_2$ labelled individuals. For each subset , compute the S_{wi} , S_{wj} , S_{pi} , S_{pj} , etc. , and the expected dw_{ij} , dw'_{ij} , dp_{ij} , dp'_{ij} , etc. , and compare whether the expected dw_{ij} , dw'_{ij} , dp_{ij} , dp'_{ij} , etc. , are greater than the observed ones. After all times of randomizations finished , count the number of the expected > the observed , compute the percentage of the number *vs* randomization times as statistic *p* value. For the situation of the absolute difference , the *p* value is used to make two-sided test and , for the situation of the difference , the *p* value is used to make one-sided test. If the computed *p* value is greater than the *p* threshold , the community *i* and *j* should be considered as having significant statistic difference in diversity index , where $i = 1, 2, \dots, n-1 ; j > i$.

In the same way , for each subset , compute the Ew_i , Ew_j , Ep_i , Ep_j , etc. , and the expected Ew_{ij} , Ew'_{ij} , Ep_{ij} , Ep'_{ij} , etc. Compare whether the expected Ew_{ij} , Ew'_{ij} , Ep_{ij} , Ep'_{ij} , etc. , are greater than the observed ones or not. Following the procedures above , the statistic significance for difference of evenness index between community *i* and *j* can be computed , $i = 1, 2, \dots, n-1 ; j > i$.

Software description

Details and functions

The software is developed using Java Development Kit 1.1.8. It is an Internet on-line computational tool and comprised of seven classes and an HTML file(http://www.browso.com/sites/reference/mathsoft/pattern_recognition/)

(1) BiodiversityTest[®] class. This class is a main computational class which calls other classes to complete different tasks. After it is loaded to web browser , an interface is displayed to accept parameter inputs and choices. The parameters and choices to be input include the computational types (i. e. , diversity or evenness) , the number of communities , the number of taxa , the number of randomizations , the type of biodiversity indices , the database name of community file , and the table name of community file.

(2) Randomize class. This class is designed to generate a randomized vector , based on a known vector with the same dimension.

(3) DiverEvenness class. This class is used to compute the diversity index and evenness index of a community. Six indices , as described above , can be

computed in the class.

(4) ReadDatabase class. The connection to ODBC database through JDBC-ODBC bridge is made in this class. The data are read in this class and provided to be used by BiodiversityTest[®] class.

(5) ResultShow class. This class is developed to display computational results and to save the results. When it is called by BiodiversityTest[®] class , a window will be launched on which the elapsed time and progress of computation and results are displayed. The results can be saved by clicking a button on this window , in which situation a window for saving results will be launched.

(6) Hint class. A window for this class will be launched when it is called by BiodiversityTest[®] class. The messages , such as community file arrangement and some references , are displayed in this window.

(7) WarningShow class. This class is called by BiodiversityTest[®] class and ResultShow class. A window will be launched and error messages occurred in program running and result saving will be displayed on it when it is called.

(8) BiodiversityTest[®].html file. This file is used to load the BiodiversityTest[®] class to web browser and in which the parameters such as the size of windows and components can be adjusted.

In community file , the number of rows is the number of communities and the number of columns is the number of taxa.

The results to be displayed include :(1) the actual number of taxa for each community ; (2) the number of individuals for each community ; (3) the observed diversity index for each community , the expected diversity index for each community , the standard deviation of diversity index for each community , upper limit and lower limit of expected diversity index for each community at 95% and 99% of confidence degree , the significance of diversity for each community at 95% and 90% confidence degree , respectively ; (4) the observed evenness index for each community , the expected evenness index for each community , the standard deviation of evenness index for each community , and upper limit and lower limit of expected evenness index for each community at 95% and 99% of confidence degree , the significance of evenness for each community at 95% and 90% confidence degree , respectively ; (5) the two-tailed results : the communi-

ty pairs with significant difference of diversity or evenness at 95% and 90% confidence degree ; (6) the one-tailed results : the community pairs with significant difference of diversity or evenness at 95% and 90% confidence degree.

Software features

The software was characterized by the following features (Zhang & Qi , 2002a , 2002b) :

(1) It is stored in web server and can be run on various java-enabled browsers such as HotJava and IE. The users do not need to install and store it on client computer. It will disappear after the user close the web browser ; (2) it is a platform-independent software which can be run on different operational systems , computers and workstations ; (3) it is easily updated and distributed , no version problem ; (4) higher safety and strong robustness of the software are major guarantees for safe running , and (5) it can read data from various databases such as MS Access , Excel , dBASE , FoxPro , Oracle , etc.

Rice arthropod diversity

A grid of 15 sampling sites was established in the plot at the IIRRI farm on September , 1996. Sampling sites were spaced 2 m apart. Each sampling site was suction-sampled for arthropods every 2 weeks in the growing season using a RiceVac apparatus (Domingo &

Schoenly , 1998) and a bucket enclosure (0. 16 m²). Laboratory sorting treated different life stages (immature , adults) separately. On the sampling date 125 arthropod species were recorded.

Result and analysis

According to the conclusion from Tilman *et al.* (1997) , the functionalities of forest ecosystem were determined not only by biological species but in a more realistic sense , by functional group or higher taxonomical group of species. Andow (1991) examined various arthropod communities in crop farms and found that the biodiversity linked to biological species or lower taxonomical units has only 60% to 70% of correlation with the yield increase of crops.

More conclusions indicated that the biodiversity linked to the functional groups or higher taxonomical groups could reflect the development trend and sustainable production of agri-ecosystem (Schoenly & Zhang , 1999a). In view of these facts , the sampling set for rice arthropod is lumped into functional groups or higher taxonomical groups along the direction of arthropod species or subspecies (hold the number of sampling sites 15) , using the biodiversity utility software LUMP (Schoenly & Zhang , 1999a ; <http://www.irri.org/BiodiversitySoftware/index.html>). The results show that 125 arthropod species are lumped into 17 functional groups for 15 sites (Table 1). The lumping

Table 1 Sampled number of individuals of rice arthropod functional groups (IIRRI 1996)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Site 1	63	1	2	9	3	2	2	2	0	1	0	39	9	0	1	0	0
Site 2	33	0	1	7	1	0	6	5	1	1	0	19	20	0	0	1	0
Site 3	24	0	1	6	2	1	3	1	0	0	0	0	11	0	0	0	0
Site 4	48	0	0	6	0	2	2	1	1	0	0	4	12	0	1	0	0
Site 5	18	0	0	10	1	1	3	5	0	0	1	5	2	0	1	0	0
Site 6	88	0	1	12	1	0	4	3	0	0	1	10	15	0	0	0	0
Site 7	40	0	1	6	1	1	1	0	0	2	0	2	12	0	1	0	0
Site 8	50	0	0	12	0	1	2	1	0	0	0	6	7	0	0	0	0
Site 9	60	0	0	11	0	0	1	2	0	0	0	8	6	0	0	0	0
Site 10	69	0	2	10	0	0	2	1	1	0	1	20	12	0	0	0	0
Site 11	51	1	3	19	0	1	4	0	0	1	1	8	12	1	0	0	0
Site 12	50	0	3	5	1	1	3	1	1	1	1	9	5	0	0	0	1
Site 13	63	1	1	7	0	1	0	1	0	1	0	20	11	0	0	0	0
Site 14	35	0	1	1	1	1	3	1	2	0	0	13	10	0	0	0	0
Site 15	39	0	1	4	0	0	0	1	0	0	0	6	0	0	0	0	0

* Columns 1 ~ 17 denote functional groups : 1 : external plant feeder ; 2 : terrestrial blood sucker ; 3 : terrestrial flyer ; 4 : terrestrial crawler , walker , jumper or hunter ; 5 : planktonic swimmer and diver ; 6 : terrestrial web-builder ; 7 : tourist ; 8 : collector(filterer , suspension feeder) 9 : collector(gather , deposit feeder) ; 10 : flying adult that is searching ,ovipositing ,or larvipositing ; 11 : predator and parasitoid ; 12 : shredder ,chewer of coarse particulate matter ; 13 : neustonic swimmer and diver ; 14 : pollen feeder ; 15 : mixed herbivores ; 16 : mixed predators ; and 17 : koinobiont

Table 2 Randomization tests of rice arthropod diversity with Shannon-Wiener index *

	Observed diversity	Expected diversity	Upper limit	Lower limit	Site pairs with significant difference at 95% (one-tailed test)														
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Site 1	1.523	1.474	1.483	1.467						✓		✓	✓						✓
Site 2	1.778	1.724	1.731	1.716	✓			✓		✓	✓	✓	✓	✓			✓		✓
Site 3	1.482	1.404	1.415	1.393								✓	✓						✓
Site 4	1.296	1.239	1.251	1.229															✓
Site 5	1.811	1.693	1.705	1.682	✓			✓		✓	✓	✓	✓	✓			✓	✓	✓
Site 6	1.229	1.196	1.205	1.189															✓
Site 7	1.355	1.276	1.288	1.264															✓
Site 8	1.189	1.154	1.164	1.145															
Site 9	1.059	1.031	1.040	1.022															
Site 10	1.316	1.277	1.286	1.269															✓
Site 11	1.568	1.510	1.519	1.501						✓		✓	✓				✓		✓
Site 12	1.504	1.424	1.436	1.412									✓						✓
Site 13	1.258	1.215	1.224	1.206															✓
Site 14	1.492	1.419	1.430	1.408						✓			✓						✓
Site 15	0.811	0.771	0.783	0.759															

* 1000 randomizations for diversity analysis and Shannon-Wiener diversity index is chosen. Upper and lower limits are both at 99% confidence degree

procedure will also be advantageous to avoid such problem as the true number of species is always impossible to be determined for taxa-rich communities. For example , counts of macroarthropods alone in Philippine and Indonesia rice fields exceeded 500 to 700 species (Schoenly & Zhang ,1999a , 1999b).

Take 1000 randomizations for arthropod diversity analysis and the Shannon-Wiener index is chosen as diversity index. The results computed with BiodiversityTest[®] indicate that there are seemingly little observed diversity (evenness) differences among 15 sampling sites. Expected diversity indices and confidence intervals overall reflect the trends of true situations (Table 2) , i. e. , suppose that the observed number of taxa is the true number of taxa in the community and , the observed distribution of individuals of taxa is the true distribution of individuals. The confidence interval produced by 1000 randomizations represents the interval that produced by 1000 samplings and each with the same total number of individuals as the observed total number of individuals. The computed confidence degrees of diversity (which are omitted in Table 2) , as we can also find in Table 1 , indicate that the individuals are not homogeneously distributed among functional groups for every sampling site , in other words , predominant functional groups , which determine the dynamic trend of rice arthropod community , exist in every site. For example , external plant feeders , as one of the predominant functional groups , are the major groups which damage rice crop and which are also the preys of many of other predatory arthropods.

Evenness indices indicate a similar trend as com-

pared with diversity indices (the results on evenness indices are omitted in Table 2).

For one-tailed test , the arthropod diversity of community site 15 is significantly lower than that of site 1 to site 7 , and site 10 to site 14 , at 95% confidence degree. The arthropod diversity of community site 9 is significantly lower than that of site 1 to site 3 , site 5 , site 11 and site 12 , and site 14 at 95% confidence degree.

If the direct comparisons is made for different communities , as frequently done by most of the studies , the unreliable conclusions are almost unavoidably drawn. For examples , it seems that the arthropod diversity for community site 2 (1.778) was significantly greater than community site 1 (1.523) , however , the randomization test shows a negative result , i. e. , there is no significant statistic difference between site 2 and site 1 based on Shannon-Wiener index. For the site pairs with obvious difference in arthropod diversity , such as site 1 *vs* site 4 , site 1 *vs* site 8 and site 9 , the tests draw the same conclusions as the direct comparisons.

The conclusions drawn from Table 3 are overallly in coincidence with that from Table 2. For example , the community site 15 holds the lowest arthropod diversity , site 2 and site 5 hold the second position. However , because of different diversity indices chosen , a little difference exists in the two result sets. This small discrepancy indicates that for special communities the corresponding biodiversity index should be used as described above.

Table 3 Randomization tests of rice arthropod diversity with Berger-Parker index *

	Observed diversity	Expected diversity	Upper limit	Lower limit	Site pairs with significant difference at 95% (one-tailed test)														
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Site 1	0.529	0.531	0.536	0.526				✓		✓	✓	✓	✓		✓	✓		✓	
Site 2	0.653	0.653	0.659	0.648	✓		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	
Site 3	0.510	0.509	0.517	0.501						✓		✓	✓					✓	
Site 4	0.377	0.379	0.385	0.372														✓	
Site 5	0.617	0.615	0.623	0.607				✓		✓	✓	✓	✓		✓	✓	✓	✓	
Site 6	0.348	0.349	0.353	0.344														✓	
Site 7	0.403	0.401	0.407	0.394														✓	
Site 8	0.367	0.369	0.376	0.363														✓	
Site 9	0.318	0.318	0.324	0.313															
Site 10	0.415	0.418	0.423	0.413														✓	
Site 11	0.500	0.498	0.504	0.493				✓		✓		✓				✓		✓	
Site 12	0.390	0.391	0.397	0.385														✓	
Site 13	0.406	0.407	0.413	0.402														✓	
Site 14	0.485	0.486	0.493	0.479						✓			✓					✓	
Site 15	0.235	0.231	0.238	0.224															

* 500 randomizations for diversity analysis and Berger-Parker diversity index is chosen. Upper and lower limits are both at 99% confidence degree

Discussion

There is a general problem with all indices of evenness. They all assume that the total number of taxa in the whole community is known. However , this number is always impossible to determine for taxa-rich communities. Since the observed taxa number are always less than true taxa numbers in the community , the evenness indices are often overestimated(Krebs ,1989). Routledge (1980) recommended that evenness indices should not be used in ecological work unless the number of taxa in the whole community is known. It is suggested that in some temperate zone communities and in well studied tropical communities , this may be possible (Krebs ,1989). However , lumping species into functional groups or higher taxonomical groups , as indicated in this paper , will be an alternative solution to avoid the problem. Species richness can be easily determined in simple communities with few species. In many cases , the species list will be larger with the larger sample size. The rarefaction technique allows us to adjust a series of samples to a common sample size (i. e. , the number of individuals) , so that species richness can be compared among samples (Krebs ,1989 ; Schoenly & Zhang , 1999a , 1999b , 1999c). For quadrat sampling , a jackknife estimate of species richness can be made based on the number of species that occur in only one quadrat(Heltshe & Forrester ,1985 ; Palmer ,1991 ;Zhang & Schoenly , 1999c).

It should be noted that the specified number of randomizations should depend on the number of individuals in community , and generally it may be taken as

1000 , or 500 times , etc.

References

Andow D A, 1991. Vegetational diversity and arthropod population response. *Annual Review of Entomology*, **36**: 561 ~ 586
Domingo I and K Schoenly, 1998. An improved suction apparatus for sampling invertebrate communities in flooded rice. *International Rice Research Notes*, **23**(2): 38 ~ 39
Heltshe J F and N E Forrester, 1985. Statistical evaluation of the jackknife estimate of diversity when using quadrat samples. *Ecology*, **66**: 107 ~ 111
Hurlbert S H, 1971. The concept of species diversity: a critique and alternative parameters. *Ecology*, **52**: 577 ~ 585
Krebs C J, 1989. *Ecological Methodology*. Happer Collins Publishers, Inc. New York
Manly B F J, 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Second edition). Chapman & Hall, London, UK
May R M, 1975. Patterns of species abundance and diversity. In: M L Cody and J M Diamond (eds.) , *Ecology and Evolution of Communities*, Belknap Press, Harvard University, Cambridge, Mass. 81 ~ 120
McIntosh R P, 1967. An index of diversity and the relation of certain concepts to diversity. *Ecology*, **48**: 392 ~ 404
Palmer M W, 1991. Estimating species richness: the second-order jackknife reconsidered. *Ecology*, **72**: 1512 ~ 1513
Peet R K, 1974. The measurement of species diversity. *Annual Review of Ecological System*, **5**: 285 ~ 307
Routledge R D, 1980. Bias in estimating the diversity of large, uncensused communities. *Ecology*, **61**: 276 ~ 181
Schoenly K G and W J Zhang, 1999a. IIRRI Biodiversity Software Series. I. LUMP, LINK, AND JOIN: utility programs for biodiversity research. *IIRRI Technical Bulletin* No. 1. International Rice Research Institute. Manila, Philippines, 23
Schoenly K G and W J Zhang, 1999b. IIRRI Biodiversity Software Series. V. RARE, SPPDISS, and SPPRANK: programs for detecting between-sample difference in community structure. *IIRRI Technical Bulletin* No. 5. International Rice Research Institute. Manila, Philippines, 17
Solow A R, 1993. A simple test for change in community struc-

ture. *Journal of Animal Ecology*, **62**: 191 ~ 193

Tilman D, J Knops, D Wedin, P Reich, M Ritchie and E Siemann, 1997. The influence of functional diversity and composition on ecosystem processes. *Science*, **277**: 1300 ~ 1302

Washington H G, 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. *Water Research*, **18**: 653 ~ 694

Zhang W J and K G Schoenly, 1999a. IRRI Biodiversity Software Series. II. COLLECT1 and COLLECT2: programs for calculating statistics of collectors' curves. *IRRI Technical Bulletin* No. 2. International Rice Research Institute. Manila, Philippines, 15

Zhang W J and K G Schoenly, 1999b. IRRI Biodiversity Software Series. III BOUNDARY: a program for detecting boundaries in ecological landscapes. *IRRI Technical Bulletin* No. 3. International Rice Research Institute. Manila, Philippines, 17

Zhang W J and K G Schoenly, 1999c. IRRI Biodiversity Software Series. IV. EXTSP1 and EXTSP2: programs for comparing and performance-testing eight extrapolation-based estimators of total taxonomic richness. *IRRI Technical Bulletin* No. 4. International Rice Research Institute. Manila, Philippines, 13

Zhang W J and K G Schoenly. 2001. A randomization test and software to compare ecological communities. *International Rice Research Notes*, **26**(2): 48 ~ 49

Zhang W J and Y H Qi, 2002a. Statistic test and network software on statistic significance of biological diversity and evenness. *Modern Computer*, **145**: 6 ~ 9

Zhang W J and Y H Qi, 2002b. Functional link artificial neural network and agri-biodiversity analysis. *Biodiversity Science*, **10**(3): 345 ~ 350

This project was supported by the National Natural Science Foundation of China, through 30170184, "Quantitative Algorithms and Softwares for Insect IPM-related Agrobiodiversity"; Foundation of Scientific Research for Personnel from Abroad, through Ministry of Education of China, "Development of Computational Softwares and Algorithms for IPM-related Agrobiodiversity"; and Asian Development Bank, through RETA 5711, "Exploiting Biodiversity for Sustainable Pest Management".

生物多样性和均匀度显著性的随机化检验及计算软件

张文军¹ 齐艳红^{2 3} Schoenly, Kenneth George⁴

1 (中山大学生物防治国家重点实验室, 广州 510275)
2 (中山大学生命科学学院, 广州 510275) 3 (中山大学计算机科学系, 广州 510275)
4 (Department of Biological Sciences, California State University, Stanislaus, CA 95382, USA)

摘要: 多样性指数和均匀度以其简单易用而被广泛应用于群落生态学和生物多样性等研究中,然而由于缺乏合适的统计检验方法等原因,其分析的可信性往往较低,因而限制了其应用。鉴于生物多样性研究中广泛应用主观和直接的比较法,有必要建立和使用较为严格的多样性统计检验。本研究建立和应用了如下随机化检验方法:单群落多样性指数和均匀度的显著性检验,单群落多样性指数和均匀度的置信区间,群落间多样性和均匀度的差异显著性检验。随机化方法已被成功地应用于群落生态学研究,其原理是:随机排序某一向量中的元素,或随机交换两向量中的对应元素。计算该随机化数据的多样性和均匀度,重复该过程多次,统计和计算显著性检验的 *p* 值,由此可确定多样性和差异的统计显著性。同时,研制了相应的 Internet 计算软件 BiodiversityTest[®]。该软件由 7 个 Java 类和 1 个 HTML 文件组成,可运行于多种操作系统和网络浏览器上,可读取多种类型的 ODBC 数据库文件如 MS Access, Excel, FoxPro, Dbase 等。该软件中包括 Shannon-Wiener 多样性指数, Simpson 多样性指数, McIntosh 多样性指数, Berger-Parker 多样性指数, Hurlbert 多样性指数以及 Brillouin 多样性指数。基于 Shannon-Wiener 多样性指数和 Berger-Parker 多样性指数,用 BiodiversityTest[®] 软件对水稻田节肢动物群落多样性(15 个地点, 17 个功能群, 125 个节肢动物种)进行了比较和分析。结果显示,两组结果可较好地反映水稻田节肢动物群落多样性的差异显著性,这些检验方法可有效地反映多样性指数和均匀度的变化。与水稻田节肢动物群落间多样性的直接比较法相比,该随机化检验方法能获得更客观的结果。本算法与软件有助于改进生物多样性研究中使用的某些不甚严格的分析方法,为随机化检验方法在生物多样性研究中的进一步应用提供了一种可用的工具。

关键词: 生物多样性, 统计显著性, 随机化检验, 置信区间, Internet 软件