DYNAMICS OF THE RODENT COMMUNITY IN THE CHIHUAHUAN DESERT OF NORTH AMERICA IV. SIMULATION OF COMMUNITY VARIABLES *

ZENG Zongyong
(Department of biology, Sichuan University, Chengdu, 610064)

Abstract By Monte Carlo method computer simulation is conducted to delineate the overall, annual and monthly fluctuations, and the statistical characteristics of 4 variables of the rodent community in the Chihuahuan Desert of North America, including the number of species, biomass, species diversity and evenness. The results show that the means of 4 simulated community variables are 8.91, 1.544 kg/hm^2, 1.58 and 0.73. All of the values of the real community variables are between the 95% confidence intervals of 4 simulated community variables. The number of species and biomass of the simulated community show the similar seasonal fluctuation patterns. The number of species, biomass and diversity of the simulated community show the similar interannual fluctuation patterns. When the null hypothesis is tenable: a variable has a normal distribution, X^2 values show that the number of species in 962 samples, biomass in 109 samples, species diversity in 5 529 samples, species evenness in 6 654 samples are following the normal distribution, i.e. no one of 4 variables is normally distributed while the confidence level is 95%. The results suggest that when sample size is small, we cannot estimate the confidence intervals of these variables by mean ±1.96 s (s is the standard deviation of a variable).

Key words: Biomass; Community dynamics; Computer simulation; Monte Carlo method; Number of species; Probability distribution
CLC number: Q 958.1.155 Document code: A Artical ID: 1000-1050 (1999) 02-0107-12

Mark-recapture is a major way for studying ecology of small mammals. Mark-recapture in the wild is actually a random process. Captures or recaptures of individuals of a certain species population in a community are random events, while the number of captured or recaptured individuals of each species population and their body mass are random variables. On the other hand, we characterize a community often by its community parameters, for example, the number of species, biomass, species diversity and evenness. But we can only obtain the estimates of these community parameters by sampling a community, rather than their real values. The com-

* Received date: 1997-12-15, accepted date: 1998-12-10

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

munity variables are functions of the number of captured or recaptured individuals and their body mass. So these community variables are also random variables. We estimate the community parameters often by sampling the community. Even though the values of community variables by one sampling are the estimates of the corresponding community parameters, the resulting error is significant because of the serious changes of the physical environment and population densities of component species along the time dimension, and frequent movement of animals as well. So the estimates are biased[1]. The typical way should be that we first identify the theoretical distributions of each community variables. If a variable is normally distributed, and the mean and standard deviation are known, we can estimate theoretically the confidence interval of this variable by percentile (95 % or 99 % is often used) values for t-distribution when we have got at least two units of a sample. If the sample size is quite larger, for example we have got 50 sample units or more, the percentile values for normal distribution can be used to estimate the intervals of the variable with the mean and standard deviation and we do not need to know the frequency distribution of the variable in advance[2].

By a long-term monitoring of the Chihuahuan Desert rodent community we identified that 1) Biomass, and evenness were normally distributed variables; 2) frequency distribution of the number of species and diversity was skewed to the right and no common theoretical probability distribution could fit their fluctuations[1]. We got these results just based on a 92-month mark-recapture data base, we cannot make sure what are the confidence intervals of the community parameters and what types the real frequency distribution of the community parameters are. For this purpose we can simulate mark-recapture process of the rodent community in the Chihuahuan Desert. If only we make simulation of mark-recapture process to generate a large number of samples, can we draw conclusions on the statistical characteristics of the community variables.

This paper is to simulate by Monte Carlo method the mark-recapture process of the rodent community in the Chihuahuan Desert and characterize the overall, annual and monthly fluctuation patterns of the number of species, biomass, species diversity, and species evenness, and their probability distribution of the 4 community variables.

1 METHOD

1.1 The simulation of mark-recapture

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

The mark-recapture study of the rodent community in the Chihuahuan Desert of North America from November 1977 through June 1985 provides a 92-month data base for 17 species of rodents. Rodents were censused by mark-recapture with 980 live traps each month and the total study period covers 92 months. For more details of the mark-recapture study see Zeng et al [3,4].

We use the simulation method for mark-recapture[5] to simulate the dynamics of the rodent community in the Chihuahuan Desert. Suppose that the number of captured individuals of the jth species per hm^2 in the ith month is n_{ij} and 980 is the total number of live traps set each month. Then the capturing proportion of the individuals of the jth species is $p_{ij} = n_{ij}/980$, in which $i = 1, 2, ..., 92$; $j = 1, 2, ..., 17$. Here p_{ij} is the estimate of the probability that an individual of the jth species is captured in the ith month. We cumulate the capturing or recapturing probability of 17 species in the ith month to give a series of subintervals: $[0, p_{i1})$, $[p_{i1}, \sum_{j=1}^{2} p_{ij})$, $[\sum_{j=1}^{2} p_{ij}, \sum_{j=1}^{3} p_{ij})$, ……, $[\sum_{j=1}^{16} p_{ij}, \sum_{j=1}^{17} p_{ij})$. $\sum_{j=1}^{17} p_{ij}$ is generally less than or equal to 1. Then the computer generated pseudo random numbers.

Suppose $p_{i0} = 0$, if a random number R is in the interval $[\sum_{j=1}^{m} p_{ij}, \sum_{j=1}^{m+1} p_{ij}$, a captured individual of the jth species in the mth month and its mean body mass is, respectively, added to the record of the individual numbers and biomass of its population. In this way 980 random numbers represent 980 live traps set per month. The simulation results in 92-month consist a sample. By this way 10 000 samples are generated.

1.2　The overall, annual and monthly mean time series of 4 simulated community variables

Dividing the number of captured individuals and total biomass of each species by study plot area for each species produces the population density and biomass (unit: kg/hm^2). From the population density and biomass of each species in each month generated by simulation we attaine mean time series of the number of species, biomass, species and evenness (For more details of the definitions of the community variables, see Zeng's report[4]). An overall mean time series of each variable is consisted of 92 values for one sample, which describes the overall fluctuation patterns of the community variables during the 92-month trapping period. From the overall mean time series we compute year-to-year and month-to-month series to describe the annual and monthly fluctuation patterns of community variables. We also compute the 95% confidence intervals of 4 simulated community variables by mean ± 1.96 s, here s is the standard deviation of

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

variables.

1.3 Statistical hypothesis testing of the frequency distribution of the community variables

We calculate 7 statistics: mean, standard deviation, coefficient of variation, skewness, kurtosis, maximum, minimum that characterize the frequency distribution of each variable. From the means and standard deviation we also estimate the 95% confidence intervals for each variables. After that we teste the null hypothesis that a variable had a normal distribution by X^2 test.

2 RESULTS

2.1 The overall, annual and monthly mean time series of the simulated community variables

The overall mean time series of the simulated community variables, in which every value is the mean of 10 000 values in the same month of a same year, reveals the most possible fluctuation patterns of the community variables. We compute means and standard deviations of 4 community variables of 10 000 samples to obtain their overall mean and 95% confidence-interval time series (Fig. 1).

The 95% confidence intervals of the overall mean time series of the 4 simulated community variables are too narrow to be distinguishable from each other. Maybe it is because the sample size is large ($n = 10\ 000$). Both the number of species and species diversity show conspicuous increasing trends. The maximum of the mean number of species occurs in April and May of 1982 and the minimum in November of 1977, August and September of 1984. Before October of 1982, the number of species in more than half of the months is less than 9. After that time the number of species is almost always greater than 9. The mean biomass and evenness series of the simulated community do not show any trend. Biomass is the maximum in April of 1981, while minimum in September of 1981. Species diversity increases from January of 1979 through July of 1984. After July of 1979 it is always greater than 1.2 and gets its maximum in April of 1982 and the minimum in January of 1978. Species evenness gets its maximum in April of 1984 and the minimum April of 1981.

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

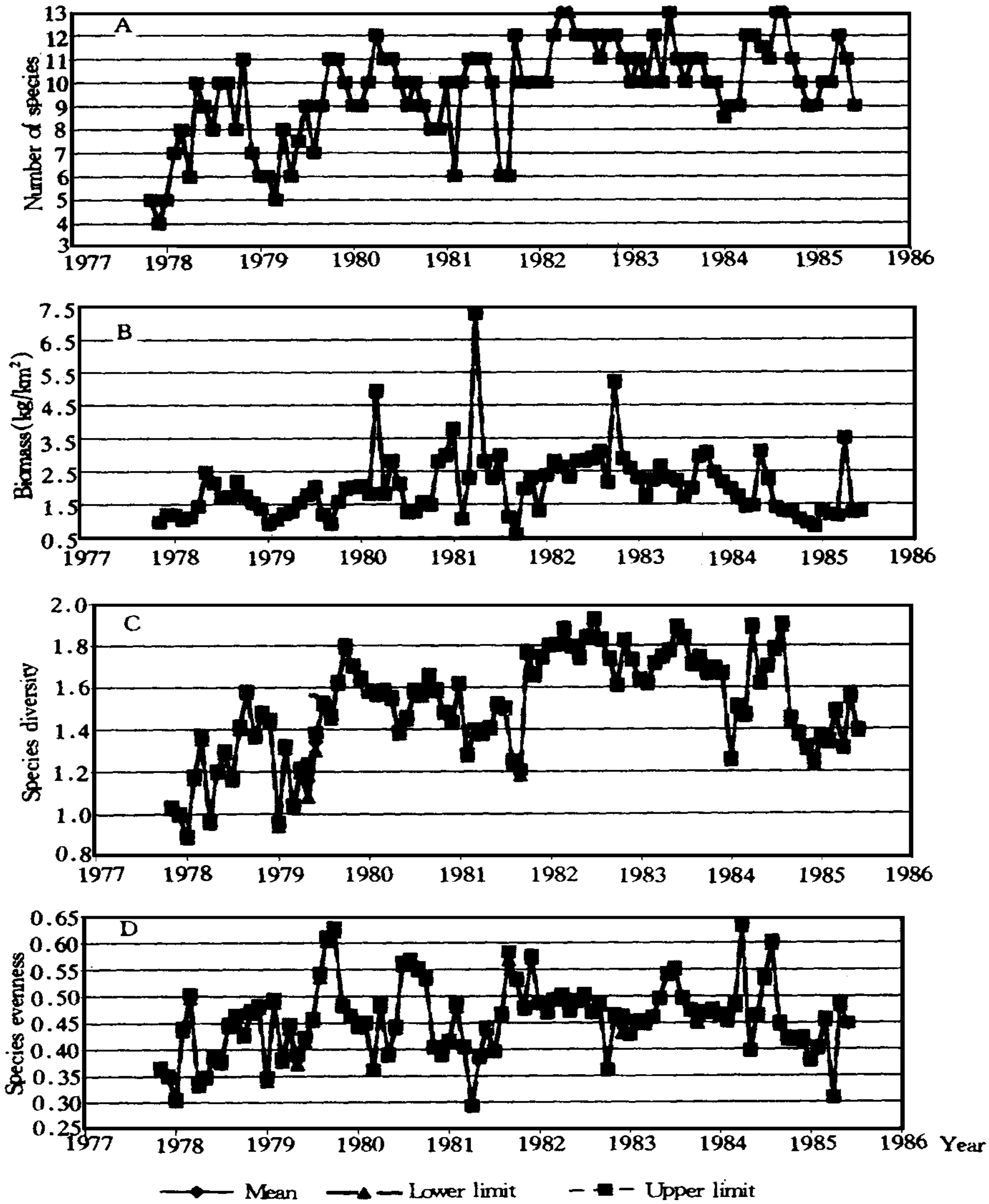

Fig. 1　The overall mean time series of means, upper and lower limits of 95 % confidence intervals of 4 simulated community variables

The upper and lower limits of the 95 % confidence intervals of the year-to-year mean time series of the 4 simulated community variables are distinct (Fig. 2). Both the number of species and biomass are the minimum in 1979, while the maximum in 1982. After 1982, they are decreasing. Both

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd.　All rights reserved.

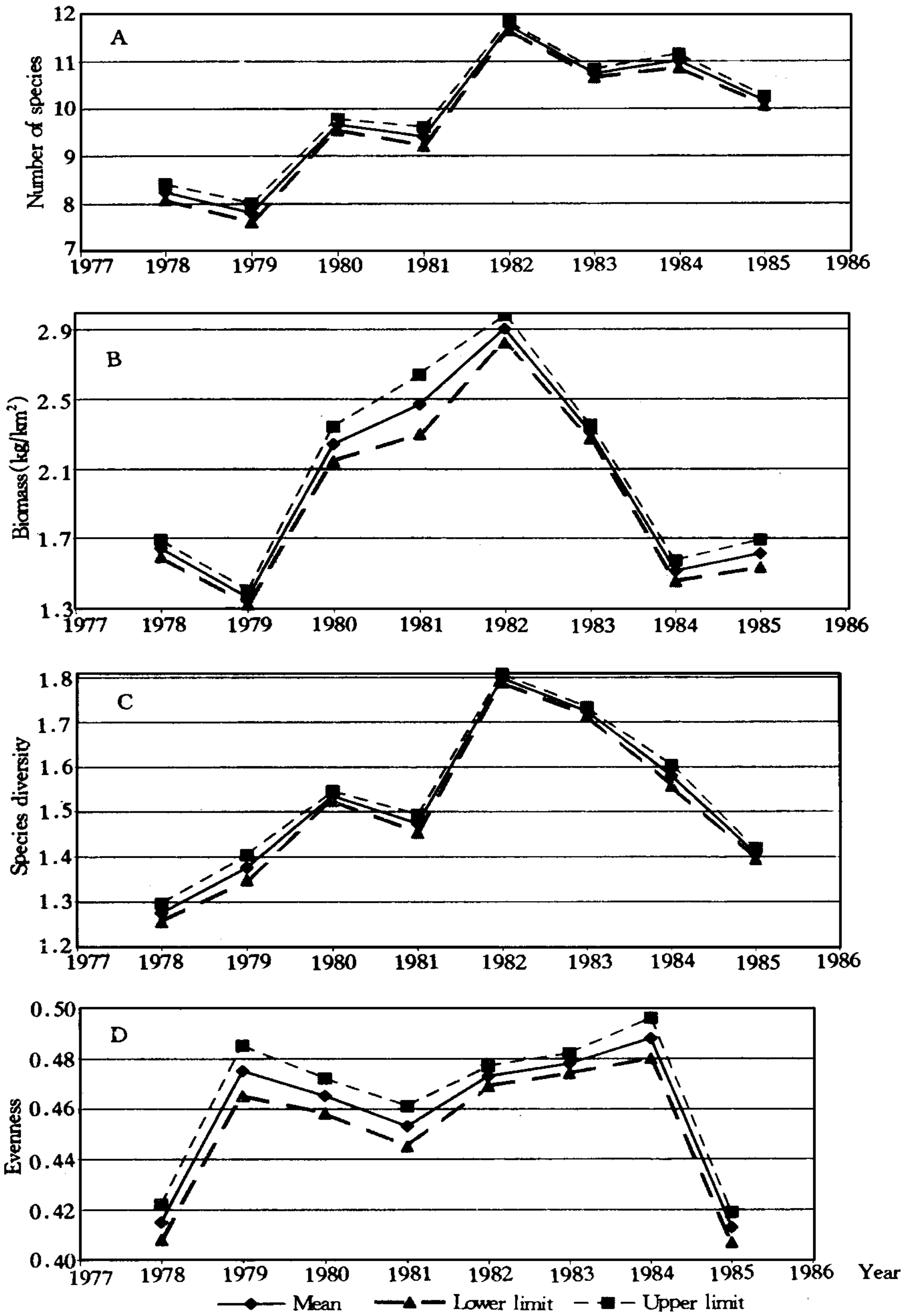

Fig. 2 The year-to-year mean time series of means, upper and lower limits of 95 % confidence intervals of 4 simulated community variables

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

species and evenness get their minimum in 1978, but species diversity gets its maximum in 1982, while evenness in 1984.

The 95% confidence intervals of the month-to-month mean time series of 4 simulated community variables are also distinct (Fig. 3). From April through June, the number of species is always greater than 10 and gets its

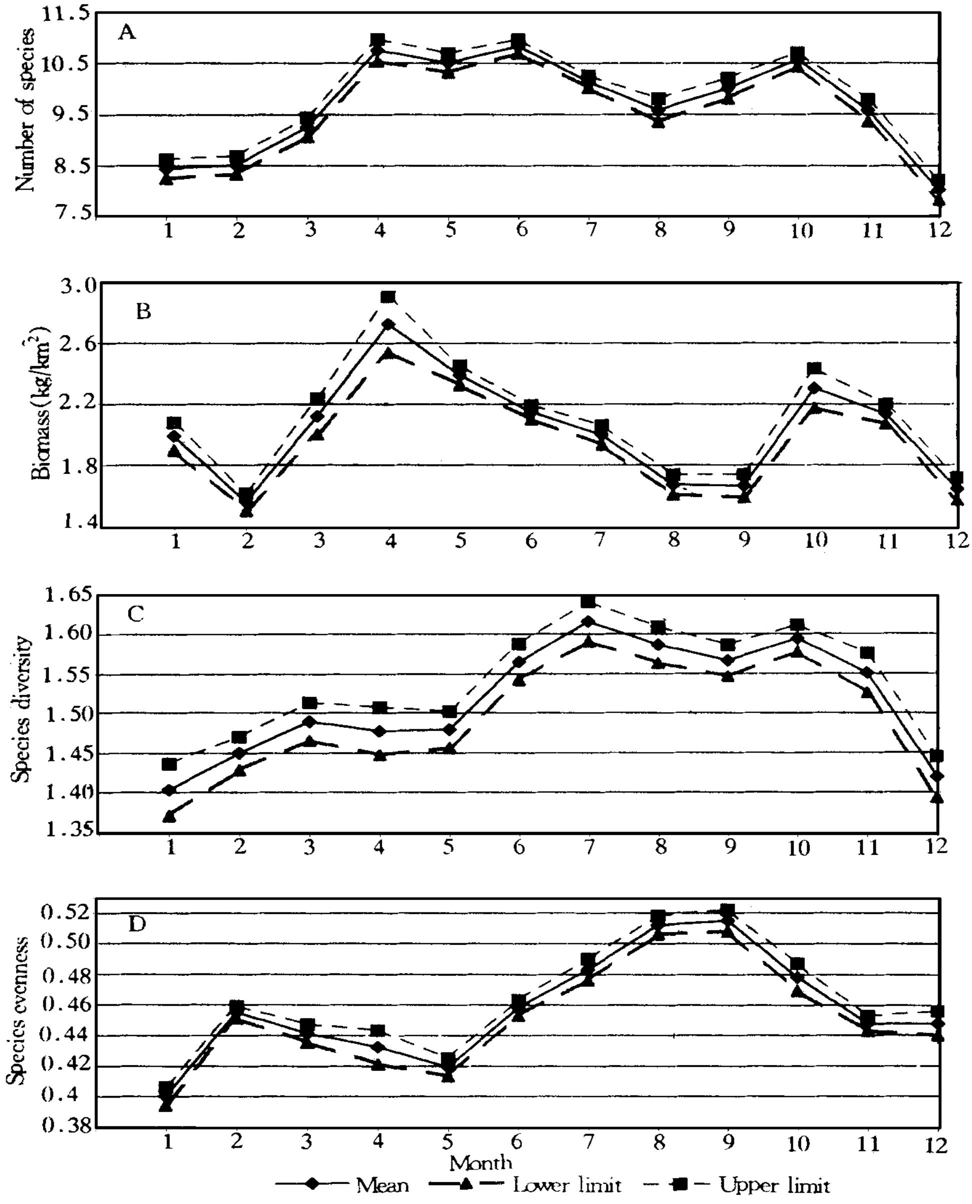

Fig. 3 The month-to-month series of means, upper and lower limits of 95% confidence intervals of 4 simulated community variables

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

maximum in June, while it is less than 8 from December through the next January and February. Just like the number of species, biomass gets its maximum in April, but it is less in February, August, September and December. Species diversity is greater from June through November than in other months, and gets its maximum in July, while the minimum in January. Species evenness shows almost the same pattern as that of species diversity. It is greater from June through October than in other months. It gets the maximum in September, while the minimum in January.

Table 1 The means of 7 statistics of the number of species, biomass, species diversity and evenness

Statistics	Mean	Standard deviation	Coefficient of variation	Skewness	Kurtosis	Maximum	Minimum
Number of species	8.91	2.12	0.24	- 0.60	3.63	12.81	2.37
Biomass (kg/hm^2)	1.544	0.785	0.51	2.12	6.32	5.463	0.423
Species diversity	1.58	0.27	0.17	- 0.60	- 0.028	2.08	0.81
Evenness	0.73	0.08	0.10	- 0.47	- 0.17	0.88	0.53

Table 2 The 95% confidence intervals of 7 statistics of the variables characterizing the community

Statistics	Mean	Standard deviation	Coefficient of variation	Skewness	Kurtosis	Maximum	Minimum
Number of species	8.76, 9.07	1.90, 2.23	0.22, 0.25	- 0.78, - 0.41	- 0.50, 0.21	12, 13	2, 4
Biomass (kg/hm^2)	1.510, 1.577	0.742, 0.831	0.47, 0.54	1.5, 2.5	3.8, 9.7	4.97, 5.97	0.275, 0.554
Species diversity	1.53, 1.59	0.25, 0.31	0.16, 0.19	- 0.90, - 0.35	- 0.70, 1.30	1.96, 2.18	0.60, 0.95
Evenness	0.58, 0.62	0.06, 0.08	0.09, 0.12	- 0.90, 0.10	- 0.70, 0.55	0.83, 0.95	0.43, 0.59

Note: The numbers before and behind the comma in each cell represent, respectively, the lower and upper limit

2.2 Statistics of 4 simulated community variables from 10 000 samples

Table 1 shows the means of 7 statistics of the number of species, biomass, species diversity and evenness generated by simulation, which indicates the most possible values of 4 simulated community variables.

On the other hand, we are more concerned about the confidence intervals of 7 statistics of these variables, because they give us an overall view on the statistical characteristics of community parameters charac-

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

terizing the rodent community. Here we list the 95 % confidence intervals of 7 statistics of the 4 simulated community variables of 9 500 samples (Table 2).

The number of species is one of the basic parameters that show the organization of a community. Of 10 000 samples, the maximum is 13 (1 877 samples) or 12 (8 123 samples). The minimum number of species is 2 (6 403 samples), 3 (3 475 samples) or 4 (122 samples). It means all of the species never occurred simultaneously in the community during the same trapping period. The moderate coefficients of variation of the number of species (Table 2) show that this community variable is relatively stable comparing with biomass.

The mean biomass is 1.544 kg/hm^2. This variable shows the greatest fluctuating comparing with all other 3 variables. The maximum is between 4.338 and 6.642 kg/hm^2, and the minimum between 0.145 and 0.645 kg/hm^2. Species diversity showed less variation. Its maximum is between 1.92 and 2.29, while minimum between 0.20 and 1.05. Evenness fluctuates the least in all of 4 variables, which suggests that it is the most stable. Its maximum is between 0.825 and 1, while the minimum between 0.25 and 0.61.

2.3 Frequency distribution of the simulated community variables

When the null hypothesis was that a variable had a normal probability distribution, we conducted X^2 test. We pooled the data into 5 groups. Because there are 2 parameters for calculating the theoretical frequency of a normal distribution, the degree of freedom is 3. At confidence level of 0.90, 0.95, 0.99 and 0.999, we counted the number of samples in which variables are significantly normally distributed (Table 3).

Table 3 The cumulative number of the samples in which variables are significantly normally distributed at 0.9, 0.95, 0.99 and 0.999

Confidence level	0.90	0.95	0.99	0.999	0.999 9
X^2	< 6.251	< 7.815	< 11.345	< 16.266	> 16.266
Number of species	438	15	2 907	6 127	10 000
Biomass	56	109	291	510	10 000
Species diversity	4 355	5 529	8 627	8 895	10 000
Evenness	5 157	6 654	8 753	9 753	10 000

Both skewness and kurtosis of the normal probability distribution are equal to 0. In what extent the values of these two statistics are near to 0 can be a criterion for judging if a random variable has a normal distribution or not. On the other hand, that the values of these two vari-

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

ables are greater than 0 or less than 0 tells us the distribution is centralized and even or not.

The skewness of the number of species is always less than 0, as means its frequency distribution skews to the larger number of species. The kurtosis of 95% of samples is between -0.50 and 0.21. That is near the kurtosis of the normal distribution, which means that frequency distribution shows moderate even like the normal distribution. The results of simulation show that the number of species in only one tenth of samples has a normal distribution at the confidence of 95% and in three tenth of samples at the confidence level of 99%.

Comparing to the number of species, biomass has a stronger trend that the values of the variable tend to be less than the mean and more peaked. The frequency distribution of this variable in only 291 out of 10 000 samples is normal at the confidence level of 99%. Frequency distributions of species diversity, evenness and biomass show rather the same pattern. Both skewness and kurtosis are near 0. At the confidence level of 0.99 more than 85% of the samples have the normal distribution.

3 DISCCUSION

When we are sampling a community, we get only a single value of community variables. If we estimate the corresponding parameters of these variables, the error is so large that we get a wrong impression of a community sometimes. On the other hand, sampling a community is a time-and-money consuming job, which limits the sample size. But we can simulate many times of sampling a community by computer to obtain larger-size samples. From a larger-size sample we can know more statistical characteristics of the simulated community variables.

Comparing the real community variables and the simulated community variables, we can see: 1) All of the values of the real community variables (see Table 2)[6] are between the upper and lower limits of the 95% confidence intervals of our simulated community variables (Table 3); 2) Both the number of species and species diversity showed conspicuous increasing trends.

The month-to-month mean time series of 4 simulated community variables show obviously similar interannual and monthly fluctuation patterns to the changes of physical environment in the Chihuahuan Desert[7]. In spring and fall, the temperature is moderate and it is the end of the rain season. Plants grow very well. The number of species and biomass are

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

greater than in other seasons. While more individuals of more species occurred in the community in summer, species diversity and evenness is greater than in other seasons.

The study of frequency distribution of rodent community variables in the Chihuahuan Desert shows that biomass, species diversity and evenness had a normal distribution[1], which is different from the results of simulation. The simulation results show that the probability of normally distributed samples of biomass is quite small.

But Tables 3 reveals that no one of 4 community variables shows normally distributed at the confidence level of 95%. We got normally distributed samples of them just by chance in the field. Even though we can obtain a point estimate of a community parameter characterizing a community. But it is more difficult to estimate the confidence interval of a variable, because we do not know its exact frequency distribution. When sample size is small, we cannot estimate the confience intervals of 4 community parameters by mean ±1.96 s, here s is the standard deviation of a variable. In order to estimate the confidence interval of a variable, a larger sample size is needed (e.g. $n > 50$).

REFERENCE

[1] Zeng Z, Yang Y, Song Z, Luo M. Dynamics of the rodent community in the Chihuahuan Desert of North America III. Frequency distributions and prediction of variables [J]. Acta Theriologica Sinica, 1995, **15** (4): 289～297.

[2] Sokal R R, Rohlf F J. Biometry [M]. San Francisco: W. H. Freeman, 1981.

[3] Zeng Z, Brown J H. Population ecology of a desert rodent: Dipodomys merriami in the Chihuahuan Desert [J]. Ecology, 1987, **68** (5): 1328～1340.

[4] Zeng Z, Yang Y, Song Z. Dynamics of the rodent community in the Chihuahuan Desert of North America II. Seasonality and periodicity [J]. Acta Theriologica Sinica, 1994, **14** (2): 100～207.

[5] Keen R E, Spain J D. Computer simulation in biology. A basic introduction [M]. New York: Wileyy-Lis Inc., 1992. 279～305.

[6] Zeng Z. Dynamics of the rodent community in the Chihuahuan Desert of North America I. Interannual fluctuations and trends [J]. Acta Theriologica Sinica, 1994, **14** (1): 24～34.

[7] Brown J H, Zeng Z. Comparative population ecology of eleven species of rodents in the Chihuahuan Desert [J]. Ecology, 1989, 70: 1507～1525.

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.

中 文 摘 要

北美 CHIHUAHUAN 荒漠啮齿动物群落动态 IV. 群落变量的模拟

曾宗永
(四川大学生物系, 成都, 610064)

利用 Monte Carlo 方法模拟了北美 CHIHUAHUAN 荒漠啮齿动物群落变量的总和、月间和年间变动情况及统计学特征, 这些变量是: 物种数、生物量、物种多样性和均匀性。模拟产生的 10 000 个样本显示这 4 个变量的平均值是 8.91, 1.544 kg/hm^2, 1.58 和 0.73。真实群落 4 个变量的所有值均在模拟群落变量的 95 %置信区间内。模拟群落物种数、生物量和物种多样性有相同的年间变动规律。物种数和生物量有相同的季节变动规律。当零假设为这些变量均服从正态分布时, X^2 检验及偏度和峭度说明, 在 95 %的显著性水平上, 物种数有 962 个样本, 生物量有 109 个样本, 物种多样性有 5 529 个样本, 均匀性有 6 654 个样本服从正态分布, 即在 95 %的显著性水平上, 没有一个变量是服从正态分布的。结果说明, 当样本含量较小时, 我们不能用 (平均数 + 1.96 ×标准误) 的公式来计算。

关键词: 计算机模拟; 生物量; 群落动态; 物种数; 概率分布; Monte Carlo 法

© 1995-2005 Tsinghua Tongfang Optical Disc Co., Ltd. All rights reserved.