

多様度と類似度、分類学的新指標

大垣俊一

いわゆる「群集生態学」には、2種類あると私は思っている。一つは、生物集団内の食物連鎖や住み込み現象など、種間関係を強調するもので、この立場から見た群集を英語でいえば **community** であろう。**community** としての群集を、関係の全体像と言った人もいる。これに対し、群集を単に種の集合体とみなし、種構成や分類群構成を基礎に、その時間的変化や、様々な空間スケールにおける変化を追及するという行き方もある。こちらの視点からの群集は、しばしば「生物相」と表現され、英語でいえば **assemblage** となる。生物相は、「その場の生物種の全体、ただし種間関係を考慮しない」というように定義されている（八杉ほか 1996）。この二つの立場をめぐっては互いに批判がありうるが、それは本稿のテーマではない。つきつめれば好みの問題だろう。

後者の群集研究においては、各地域の生物相の多様性や地域間の類似度を測る指標が重要な意味を持つ。多様度とは、あるサンプルの中に含まれる種がどれほど多様な構成になっているか、類似度とは、二つのサンプルが、種組成としてどれほど似ているかを測る尺度である。たとえばある勾配に沿って群集の多様性がどう変化するかを、各地点の種-個体数リストを並べて見てもよくわからない。類似度の場合、これらの地点群とあれたら地点群が、種組成の面でそれぞれ異なるクラスターを形成するというを示すのに、何百種類の分布図を出すわけにもいかない。「代表的な種」を選んで示すのも、恣意的な選択をしたのではないかと受取られるかもしれない。何らかの客観的な指標を選び、データを要約することが必要である。そのためこれまでいろいろな多様度、類似度の指数が考え出され、使われてきた。主な多様度指数として Simpson の λ 、Shannon-Wiener の H' 、類似度指数では Jaccard、Bray-Curtis、木元の C_{II} などがある。これらは 1960 年代ごろまでには出そろって、以後ほぼ 50 年にわたって使われ続けてきた。

この状況に変化が生まれたのは 1990 年代に入ってからで、イギリス、プリマス海洋研究所の研究者らが、海洋ベントスの研究に新しい多様度指数を導入した。その最大の特徴は、多様度の評価に分類の情報を組み込んだことにある。これまでの指数は、類似度も含めてそのすべてが、群集内の種を平等ないし均質に扱っていた。巻貝でも甲殻類でも同じ 1 種の扱いである。極端な場合、同属の巻貝が 1 個体ずつ 5 種いても、イソギンチャク、貝、カニ、ウニ、ホヤが 1 個体ずつ 5 種でも同じ多様度になるが、後者の多様度が高いと見るべきではないのか。少なくとも、そのような評価のしかたがあってよい。すべての種を同等に扱う従来型の指数は、ある意味では原データの持つ情報を生かし切れていないともいえる。環境政策の面で多様性保存の機運が高まるにつれ、このことはますます意味を持つようになった。こうして考え出された

指数が、のちに紹介する Warwick-Clarke の Δ^+ 、(デルタプラス)、 Δ^* (デルタスター) などのいわゆる ‘taxonomic distinctness index’ (分類学的多様度指数) である。分類組み込みの考え方は、その後類似度にも取り入れられたが、このことは、多数地点を類別する多変量解析的手法において、分類の情報を加えた分析を可能にした。たとえば MDS では、サンプル間の距離の尺度として Jaccard 指数や Bray-Curtis によっていたが、これを後述する分類学的非類似度 TD に置きかえることができる。

本稿では、多様度、類似度それぞれについて、これまで使われてきた指数の概略を紹介した上、新たに登場した分類学的指数を取り上げる。なお解説に当り、従来型の指数については主に木元 (1976) によった。

1. 多様度

① 従来 of 指数

種数 :

群集の多様性を測る尺度として最も単純なものは、調査時に記録された総種数である。実際に調査を行ってみれば、一見して生物相が豊富なところでは短時間で多くの種がリストされるのに対し、貧弱なところではいくら探しても少数の種しか見つからないということはしばしば起こる。したがって種数は群集の多様度を測る一つの有力な指標である。しかし種数で多様度を代表させることには二つの問題がある。一つは種ごとの個体数の情報が含まれていないということ、もう一つは、調査する人数、時間などの調査努力に影響されることである。このうち前者の問題を解決し、個体数を含めて多様度を評価することができるように考案されたのが、たとえば次に紹介する Simpson の λ 、Shannon-Wiener の H' である。

Simpson の λ :

Simpson 指数は、確率の考え方に基づいて定義されている。今、S 群に色分けされた全数 N 個の球を箱に入れてよく混ぜ、任意の 1 個を取り出して色を確かめて箱に戻し、再び球を取り出して色を調べるという試行を考える。群集に即していえば、球は個体、色は種、S は種数、N は全個体数、箱は群集である。このとき、2 階の試行で色が一致する (= 同じ種になる) 確率は、i 番目の群 (種) の球の数 (個体数) を n_i として、

$(n_i/N) \times (n_i/N) = (n_i/N)^2$ これをすべての色 (種) について足し合せ、
 $\lambda = \sum_i P_i^2$ ($P_i = n_i/N$) となる。

実際に数値を入れてみる。著しく集中した個体数配分、たとえば 1 種のみが 5 個体いる、という場合、

$$\lambda = (5/5)^2 = 1$$

逆に、個体が各種に均等に配分されている場合、たとえば 5 種がそれぞれ 1 個体ずつなら、

$\lambda = (1/5)^2 \times 5 = 1/5$ で、1 よりかなり小さくなる。集中していれば多様度は低く、均等であれば高いと考えるべきであるから、多様度の指数としては λ を 1 から引き、

$$1 - \lambda = 1 - \sum_{i=1}^S P_i^2 = 1 - \sum_{i=1}^S \left(\frac{ni}{N} \right)^2 \quad 0 \leq 1 - \lambda < 1$$

S, 種数; ni , i 番目の種の個体数; N, 全個体数

これが Simpson の多様度指数である。この値は 0~1 の範囲にあり、先の例からもわかるように、個体数が完全に 1 種に集中していれば 0、均等に配分されて、しかも種数が多いと 1 に近づく (1 にはならない)。

Simpson 指数と種数の関係を見よう。5 種が 1 個体ずついる先の例では、

$$1 - \lambda = 1 - (1 / 5) = 0.8$$

10 種が 1 個体ずつなら、

$$1 - \lambda = 1 - (1 / 10)^2 \times 10 = 1 - (1 / 10) = 0.9$$

となり、発見種数 (と個体数) が多いほど、多様度が高くなる。

直観的には Simpson の多様度とは、フィールドを歩いてたまたま見かけた 2 個体が、同じ種に属する確率、つまり、たいてい同じ種を見かけるか、そのつど違った種に行き当たるか、そのどちらに偏っているかの程度と考えることができる。

Shannon-Wiener の H' :

Shannon-Wiener 関数は、Island Biogeography で有名な McArthur らが 1960 年代に使い始め、現在でも広く用いられている。これは情報理論に基づく指数である。情報量 (エントロピー) とは、「ある出来事がどれほど起こりにくいかを示す尺度」であるといわれる。つまり、起こる確率が小さいほど、その事象の情報量は大きい。一見理解しにくいのが、ありふれたことにはニュース性がなく、めったに起こらないことは、いったん起こると情報としての価値が高い、というたとえで説明できるかもしれない。つまり、そういう社会的背景を持つ概念なのである。数式的には、情報量は対数を用いて、

$$I = \log (1 / P) = -\log P \quad (P \text{ はそのことの起こる確率})$$

で定義される。対数を用いるのは数学的に処理しやすいかららしい。対数の底としては e、2、10 などがあり、どれを選ぼうと互いに定数倍で変換できるから本質的な違いはないが、群集分析では 2 をとることが多い。この場合、ある事象の起こる確率が $(1/2)^n$ であれば、上式により情報量は n となる。たとえば $P = (1/2)^{10}$ というような、きわめて稀にしか起こらない事象の情報量は 10、 $P = 1/2$ なら 1、必ずそれが起こるならば ($P = 1$)、0 となる。つまり、底 2 を取ると、コインの裏表のように「一か八か」「フィフティ・フィフティ」的な事象について、 I が整数値で与えられるという利点がある。

このとき、 P を群集内でのある種の存在確率 ni/N によみかえ、その種の持つ情報量 $-\log (ni/N)$ に、それが起こる確率をかけて積算したもの、つまり情報量の平

均値を求める。

$$H' = -\sum_{i=1}^s P_i \log_2 P_i = -\sum_{i=1}^s \frac{n_i}{N} \log_2 \frac{n_i}{N} \quad 0 \leq H'$$

S, 種数; n_i , i 番目の種の個体数; N, 全個体数

これが Shannon-Wiener の多様度指数と呼ばれるものである。

H' は、1 種のみ 5 個体いる、というように著しく個体数配分が偏っている場合、 $H' = -(5/5) \times \log_2(5/5) = 0$ のように小さくなる。

情報量に即していえば、この場合はどれをとってもその種になることが決まっているから、情報量はゼロである。逆に、5 種それぞれ 1 個体ずつというように均等に配分されていると、

$$H' = -(1/5) \log_2(1/5) \times 5 = \log_2 5 = 2.32$$

と、それなりの値をとる。何が出てくるかわからないという未知要素が含まれる分、情報量が多いということであろう。

H' は 0 以上の値をとり、上限はない。個体数との関係では、たとえば 5 種が 1 個体ずつでも 10 個体ずつでも、各 n_i/N は $1/5 = 10/50$ で変らないから、種間の個体数のばらつきと種数が同じなら、個体数の影響は受けない。一方、個体数は変わらず、それがより多くの種に分配されていると、個々の n_i/N は平均として小さくなるから H' としては大きい値を取ることになる。たとえば 10 個体が 2 個体ずつ 5 種に配分されていると、

$$H' = -(2/5) \log_2(2/5) \times 5 = 0.53$$

10 個体が 1 個体ずつ 10 種に配分されていれば、

$$H' = -(1/10) \log_2(1/10) \times 10 = \log_2 10 = 3.32$$

このように、種数が多いほど H' は大となり、個体数のばらつきの情報を取り込みつつ、種数を反映するという H' の特徴が現れている。

H' は対数の入った人工的な指数であるから、Simpson 指数のように、フィードに即した直観的解釈を与えることはむずかしい。情報理論の言葉で表現してもよくわからない。あえて言えば、ある約束のもとに測定した各種の稀少性の平均値を表しているということになるだろう。

② 分類学的指数

以上の指数は種数と個体数のみに基づき、出現種の分類学的組成が考慮されていない。ここで紹介する指数はそれを解決するために Warwick & Clarke (2001) によって考案された。この手法の核心は、指数のベースとなる各種間の「分類学的距離」の定量化である。分類学的多様度、taxonomic distinctness においては、i 番種と j 番種の分類学的距離 ω_{ij} を、次のように定義する。

i 番種と j 番種が、同じ種なら 0

〃 異なる種だが同じ属であるときは 1

〃 異なる属だが同じ科であるときは 2

以下同様に考え、まとめると、

ω_{ij} : 異種同属 1, 異属同科 2, 異科同目 3, 異目同綱 4, 異綱同門 5, 異門同界 6、である。

図 1 に、 ω の決定方法のイメージを示した。樹形図は A 種と B 種が同属異種、A と C は同科、A と D は同目であることを示している。このとき A-B、A-C、A-D の分類学的距離は、それぞれ点線を辿って 1, 2, 3 となる。

なお、分類学的多様度の場合、英語の表記で **diversity** ではなく **distinctness** になっている。これは最終的な指数を導く過程で、暫定的な指数に **diversity** を用いたため、それと区別しているもので、実際には **diversity** と同じ意味と考えて差し支えない。

最も単純な、種の存・不存のみのデータ (= 種リスト) では、 ω_{ij} をすべての種間で求めて足し合わせ、それを全組み合わせの数で割ることによって、次の多様度指数 Δ^+ (average taxonomic distinctness for presence / absence data) が求まる。

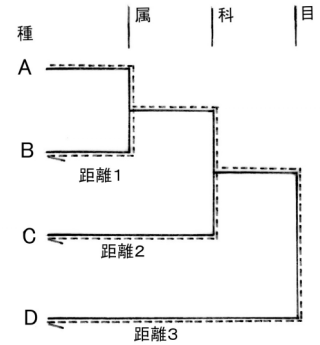


図 1.

$$\Delta^+ = \frac{\sum_{i=1}^S \sum_{j=1}^S \omega_{ij}}{\frac{S(S-1)}{2}} \quad i \geq j, \quad 1 \leq \Delta^+ \leq L-1$$

ω_{ij} , i 番種と j 番種の分類学的距離 (上記); S, サンプル内全種数
L, 使用した分類階層数

ここで分母の $S(S-1)/2$ は、種リストを縦軸と横軸に配置して総当りで ω_{ij} を求めたときの、右上か左下半分のマス目の数で、分子の方ではこのマス分の ω_{ij} を合計し、同じ組合せを 2 回数えることはしない。L は、 ω を求めるとき使った階層が、種のみなら 1、属までなら 2、科までなら 3 … などとなる。 Δ^* は全組み合わせ数で割っているから平均値であり、最大でも $L-1$ を越えない。たとえばサンプル構成種が科の違いまでしかなかったとすれば、目までの階層で足るから、L は種、属、科、目で 4、つまり $L-1=3$ 。一方分類学的距離はすべてのケースで 3 (同科異目) を越えないから、その平均である Δ^* は必ず 3 以下となる。

各種の個体数も含むデータの場合は、まずすべての個体について、他の個体との分類学的距離を求めて足し合わせる。

$$\sum_i \sum_j \omega_{ij} n_i n_j$$

n_i は i 番目の、 n_j は j 番目の種の個体数であり、 ω_{ij} の距離を持つ組合せは $n_i \times n_j$ 組あるので、上式により、サンプル内部の全種、全個体間の分類距離の総和が求まる。しかしこのままでは種数や個体数が多くなるほど値が大きくなるから、何らかの標準化が必要である。そこで、サンプル内の全種が同属であった場合を仮定し、その時 ω_{ij} が常に 1 となることから、上式は $\sum_i \sum_j n_i n_j$ 、これで割ると、「サンプルと同種数、同個体数分布で全種同属の仮定的最単純群集」からの隔たりが求まり、次の Δ^* (taxonomic distinctness for abundance data) が定義される。

$$\Delta^* = \sum_{i=1}^S \sum_{j=1}^S \omega_{ij} \cdot n_i \cdot n_j / \sum_{i=1}^S \sum_{j=1}^S n_i \cdot n_j \quad i \geq j, \quad 1 \leq \Delta^* \leq L-1$$

ω_{ij} , i 番種と j 番種の分類学的距離; n_i , i 番種個体数; n_j , j 番種個体数
L, 使用した分類階層数

なお参考までに、海洋生態学で **abundance** といえば個体数 (**number of individuals**) のことであり、**biomass** は乾重量、湿重量などの生体重量のことである。

分類学的多様度指数の直観的意味は、すべての種間の分類学的距離を、すべての組合せ数、ないし全種同属と仮定した場合からの隔たりで表したものの、つまり標準化されたサンプル内の分類学的バラエティといえよう。

2. 類似度

① 従来の手法

Jaccard 共通係数 :

古くからある単純な指数で、今でも使われる。Jaccard 指数は存・不遜データ (種リスト) に対し、次の式によって与えられる。

$$J = \frac{c}{a+b-c} \quad 0 \leq J \leq 1$$

a, b, サンプル A, B に含まれる種数; c, A と B に共通に含まれる種数

要するに共通種数の非共通種数に対する割合であり、直観的にもわかりやすい。しかし個体数の情報を含んでいないという問題は、多様度に種数を用いた場合と同じである。

Bray-Curtis 指数 :

木元 (1976) において、Odum の差分百分率として紹介されているが、最近では MDS におけるサンプル間距離の尺度として多用され、そこでは Bray-Curtis 非類似度 (**dissimilarity index**) と呼ばれている (Field et al. 1982)。そのため海洋生態分野ではこの方が通りがよい。この指数を求めるには、比較する 2 サンプルに含まれる全種を縦軸にリストし、2 サンプルの個体数を対応させた表を作る。そして各種について、順番に個体数の差を求めて絶対値を取り、総和を求める。絶対値を取るのは、プラスマイナスが打消し合わないようにするためである。

	サンプル A	サンプル B	差の絶対値
種 1	n_{A1}	n_{B1}	$ n_{A1} - n_{B1} $
種 2	n_{A2}	n_{B2}	$ n_{A2} - n_{B2} $
:	:	:	:
種 i	n_{Ai}	n_{Bi}	$ n_{Ai} - n_{Bi} $
:	:	:	:
種 S	n_{AS}	n_{BS}	$ n_{AS} - n_{BS} $
和	N_A	N_B	$\sum_i n_{Ai} - n_{Bi} $

A にしか存在しない種があれば、 $n_{Bi} = 0$ とする。そして $\sum_i |n_{Ai} - n_{Bi}|$ を、A と B の種数の和で割り、次の Bray-Curtis 指数を得る。

$$\delta_{AB} = \frac{\sum_{i=1}^S |n_{Ai} - n_{Bi}|}{N_A + N_B} \quad 0 \leq \delta_{AB} \leq 1$$

n_{Ai} , サンプル A の i 番目の種の個体数; N_A , サンプル A の全個体数
(n_{Bi} , N_B も同様); S , 全種数

δ_{AB} は、A と B の種一個体数構成が全く同じであれば分子の絶対値記号内がすべて 0 だから 0 となり、全く重なりがなければ、

$$\text{分子} = \sum_i n_{Ai} + \sum_i n_{Bi} = N_A + N_B$$

となって 1 をとる。つまり正確に言えば非類似度であり、MDS などサンプル間の距離として使うときはこのままでよいが、類似度を評価するときは 1 から δ_{AB} を引くことになる。以上より Bray-Curtis 指数の意味は、個々の種についてサンプル間の個体数の差を求めてそれを平均したものである。

なお、性質が既知の群集からのサンプリングシミュレーションの結果によれば、検討した 10 種の類似度指数の中で、Bray-Curtis は実際の群集間の性質差を最もよく反映する指数の一つと評価されている (Faith et al. 1987)。一方ユークリッド距離は適合性が悪く、このことが、ユークリッド距離を非類似度のベースとする PCA (主成分分析) が、Bray-Curtis ベースの MDS に比べ、生物データへの適用が不適当と判断される一つの根拠になっている (Clarke 1993)。

木元の C_{Π} (森下の C_{λ}):

森下の C_{λ} は、多様度における Simpson 指数の応用というべき、確率に基づく指数である。のち木元によって修正され、 C_{Π} 指数が提案された。ここでは C_{Π} を解説する。

今、サンプル A (全個体数 N_A) と B (全個体数 N_B) の全種 (S) に、同種であれ

ば同じ番号になるように共通の番号をつけておき、AとBから各1個体取り出したとき、それが同じ*i*番種である確率は、

$(n_{Ai} / NA) \times (n_{Bi} / NB)$ その総和は、

$$\sum_i \{(n_{Ai} / NA) \times (n_{Bi} / NB)\} = \sum_i (n_{Ai} \times n_{Bi}) / (NA \times NB) \dots \textcircled{1} \quad (i=1 \sim S)$$

これは異なる2群集から任意に1個体ずつ取り出したとき、それがたまたま、1～S番のうちのどれか同じ種になる確率だから、この値が大きいかほど両群集は似ており、小さいほど異なると考えることができる。この値は、2群集に全く重なりがない場合、掛算のどちらかは必ず0だから、総和も0となる。一方、2群集が全く同じ組成の場合でも1にはならない。たとえばA、Bともに3種から成り、1, 2, 3番種の個体数が2, 3, 5の場合、①式は、

$$(2^2 + 3^2 + 5^2) / 10 \times 10 = 38 / 100 = 0.38$$

そこで①を次の値で割る。

$$(\Pi_A + \Pi_B) / 2 \quad \text{ここで } \Pi_A = \sum_i (n_{Ai}^2 / NA^2) \quad (\Pi_B \text{ も同様})$$

これは、指数を0から1の範囲に収めるための標準化である。たとえば、上と同じ種組成の場合に

$$(\Pi_A + \Pi_B) / 2 = \{(2/10)^2 + (3/10)^2 + (5/10)^2\} \times 2 / 2 = 38 / 100 = 0.38$$

であるから、これで割れば1になる。以上より、木元の指数は次のようになる。

$$C_{\Pi} = \frac{\sum_{i=1}^S n_{Ai} n_{Bi}}{N_A N_B} \bigg/ \frac{\Pi_A + \Pi_B}{2} \quad \text{ただし } \Pi_A = \sum_{i=1}^S (n_{Ai}^2 / N_A^2), \quad 0 \leq C_{\Pi} \leq 1$$

NA, サンプルA全個体数; n_{Ai} , サンプルAの*n*番種個体数; S, 全種数
(Bも同様)

木元の C_{Π} の本質はその分子にあり、その直観的意味とは、2サンプルから1個体ずつ取り出して、それがたまたま同種に属するかどうかの程度であるということができる。

② 分類学的指数

類似度における分類組み込み型の指数は、多様度におけるのと同様の考え方で、今のところ存・不存データに対してのみ与えられている (Izsak & Price 2001)。個体数を含むデータでも、分類学的類似度を構成することはそれほどむずかしくないはずだが、今のところ発表されていない。

まず、種間の分類学的類似度 ω を、多様度の場合と同じように定義する。

同種…0, 異種同属…1, 異属同科…2, 異科同目…3, 以下同様

これに従い、2サンプル (A, B) 間で、種ごとに、相手方のサンプル中で最も近い分類学的距離を持つものを探して距離の最小値、 $\min \omega_i$ を求める。そしてその総和を、全種数で割った次の値が、サンプル間の分類学的距離 TD (taxonomic distance) である。

$$TD = \frac{\sum_{i=1}^{SA} \min \omega_{Ai} + \sum_{i=1}^{SB} \min \omega_{Bi}}{S_A + S_B} \quad 0 \leq TD \leq L - 1$$

$\min \omega_{Ai}$, サンプル A の i 番種からサンプル B 中のすべての種に対して総当りで分類学的距離を求めたときの最小値

S_A, S_B , サンプル A, B の種数; L , 使用分類階層

TD は、A 群と B 群が全く同じ種構成ならば、各種間の最小分類学的距離がすべて 0 になるから、0 である。TD は一種のサンプル間距離（非類似度）なので、MDS などにはそのまま使えるが、類似度にはなっていない。類似度として使う場合、標準化して 0~1 に収まるようにした上、1 から引き、次の taxonomic similarity、 Δs とする。

$$\Delta s = 1 - \frac{TD}{L - 1} \quad 0 \leq \Delta s \leq 1$$

L , 分析に用いた分類階層の数

L が 1 だと、第 2 項の分母が 0 になってまずいようだが、 $L = 1$ とはサンプル A も B もすべて同種であることを意味し、分子の TD も 0 になるから、この場合は $\Delta s = 1$ と考えるようである。

Δs の意味を大ざっぱに言えば、サンプル内の各種が、相手サンプル内にどれだけ自分に近縁な種を見つけられるか、その程度を平均化したもの、ということになるだろう。

3. 分類組込み指数の評価

Warwick & Clarke (2001) は、分類学的多様度の総説の中で、従来型の指数、とくに Shannon-Wiener の H' は、サンプルサイズが増すと増大する欠点があると述べている。木元 (1976) も、「(H' 等の) 多様度指数で測定された一連の数値を種数で割ると、しばしば一定の数値になることがある。このため、多様度指数は単に種類の多少によるものであり、それ以外の何物でもないという議論もある」と述べる。本稿の具体例の検討でも、 H' は発見種数と共に増加する傾向を示した。一方、 $\Delta+$ や Δ^* などの分類学的指数は、種数や個体数が増えても増加することなく、一定値に収束することが、実際のサンプルからのシミュレーションによって示されている (図 3)。なおこの図では横軸が個体数なので、 H' が個体数の影響を受けているように見える。これは先の H' の項での説明と矛盾するようだが、ここでは個体数と同時に種数が増えていると思われるので、種数の影響と見るべきである。

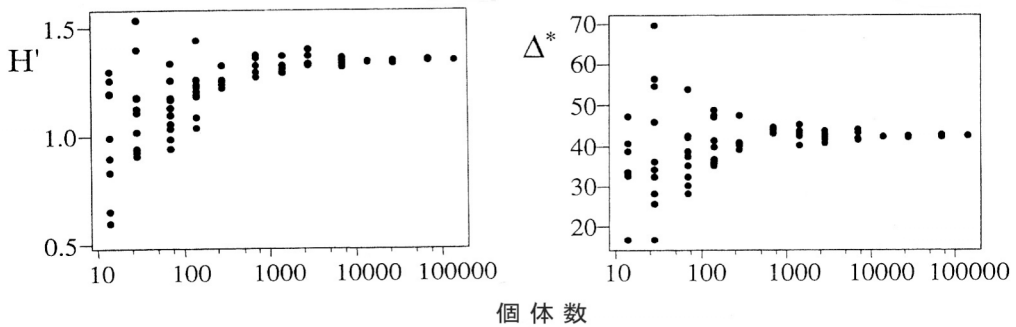


図3. 海洋マクロベントスの実データに基づく、サンプル個体数と多様度数値の関係のシミュレーション。Shannon-Wiener H' と分類学的多様度 Δ^* の比較。Warwick & Clarke (2001) による。

以上から Warwick & Clarke は、 H' は調査努力が変動して発見種数や個体数が増減するとその影響を受けて不安定であり、そのため調査努力の変動しやすい広域、長期の調査では、分類学的指数が有利であると強調している。しかし図3を見ると、個体数の少ないところでは両指標ともかなりのばらつきがあり、その意味では共に調査努力に影響される。唯一の違いは H' が種数の増大と共に全体的に高い値をとることだが、これは H' の数式的構成からの当然の帰結であろう。もともと H' や Simpson 指数は、種数に個体数分布を加えた指標を追求した結果生まれたのであり、むしろ種数を反映するところにその特徴がある。調査努力に依存することは一つの問題だが、それがコントロールされていれば、分類学的指数とは異なる情報を示しうる。

一方、類似度についての従来型指数と分類学的指数の比較について、提唱者の Izsak & Price (2001) は、同一母集団からの抽出シミュレーションの結果として、 Δs はより少ない種数で平衡状態に達し、変化にバラつきが少なくスムーズな挙動を示すとしている。これは両者の定義からも納得できる。Jaccard の場合、相手方サンプルに同一種があれば1、なければ0で、いわば 'all or nothing' であるのに対し、 Δs では、同一種が存在しなくても、分類学的距離を元に何らかの数値が与えられる。このことが、挙動の smoothness に反映するのだと思われる。そのことからかれらは、 Δs は調査努力の変動に対して抵抗性があると、多様度の分類学的指標と同様の結論を導いている。

4. 実例による検討

ここでは筆者が2004年に行った、田辺湾内外の海岸貝類分布調査の結果 (Ohgaki 2007a) をもとに、分類学的指数と従来型の指数をさらに比較、検討してみる。多様度、類似度指数の計算は、プリマスの Clarke らが開発した海洋生態学用多変量解析ソフト、Primer v6 によった。これは様々な指数をまとめて計算でき、PCA や MDS についてもオプションが豊富である。 H' や Bray-Curtis は自分で Excel などプログラムを組んでもできるが、分類学的指数は簡単にプログラムできない。今のところこれを使うしかないようだ。

① 多様度

田辺湾内外 海岸貝類群集 多様度変化

図 1. 田辺湾外—湾奥 9 地点の、多様度指数の変化。各地点各種の出現区数に基づく。

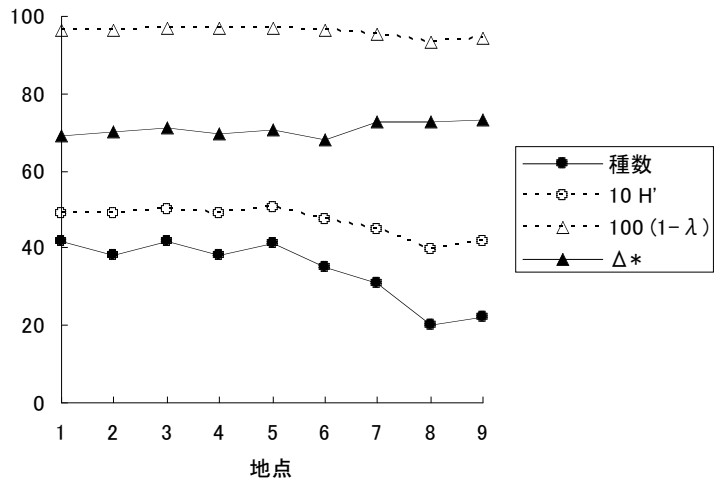


図 1 に、田辺湾外—湾奥 9 地点の種数、 H' 、 $1-\lambda$ と Δ^* の変化のパターンを示した。これらの地点は開放度の大きい順に並び、1 は湾外、2-6 は湾口、7-9 は湾奥にある。これを見ると、種数と H' は似た挙動を示し、 H' と Δ^* には差が見られる。つまり H' は湾奥で減少するのに対し、 Δ^* は逆に増加する。つまり湾奥では、種レベルで見た稀少性の総和は減少する（＝特定種のモノポリーの状態に近づく）一方、分類学的バラエティーはむしろ増大する。

このような違いの出る理由は、次のように考えられる。湾口部では出現種数が多く、分類学的多様度の計算において、種ごとに比較的低い階層でペアが見つかる。その結果、指数値は全体として低くなる。これに対し、湾奥部では種数が減って各種にとっての近縁種がいなくなり、ペアの相手との共通の分類階層が上がって、平均すると多様度が高く出のだろう。しかしこれが正しいとしても、あまり意味のある結果とはいえない。しかし、逆のパターンが出ると意味を持つことがある。石垣島の名蔵干潟では、1989 年と 1998 年の比較で、巻貝類の種数と H' が上がったが (Ohgaki & Kosuge 2005)、 Δ^* で評価すると、むしろ多様度は下がっている。このことは、種レベルでの多様度は上がったが、属以上のタクサで種組成が単調化したことを意味する。具体的には、底質の泥地化に伴い、それに適応した属が多様化する一方、他の属が衰退しつつあると考えられる。これはたとえば、有機汚濁が進んで多毛類の種数は増えたが軟体動物や棘皮動物が姿を消して、上位タクサの多様性が失われて行く場合などと、似た変化といえよう。

② 類似度

類似度についての分類学的指数は、今のところ存・不存データ用の Δs （または非類似度としての TD）しかないのを、また従来型指数として Jaccard ($1-J$)

を選び、MDS を使って比較する。MDS については以前の Newsletter でも紹介したが（大垣 1999）、要約すれば、主成分分析（PCA）のノンパラメトリクス版として、各サンプルをその種組成に基づき類別する手法である。直観的理解としては、 n 個のサンプル相互の距離を何らかの非類似度で測り、それを順位に換算した上、距離を保ったまま $n-1$ 次元空間に位置づけ、その像を、最も矛盾の少ない形で平面（2次元）に押しつけたものと考えて差しつかえない。平面上に位置づけるのにどれほど無理をしているかの指標が、MDS 図の中にある Stress 値で、これが 0.2 までなら許容範囲ということになっている。この時、サンプル間距離の測度（非類似度）としていろいろなものを選択できる。個体数データの含まれる通常のサンプルでは Bray-Curtis 指数が用いられることが多いが、ここではそれを Jaccard ($1-J$) と TD にするというわけである。

ベースは上記多様度と同じ、田辺湾内外の貝類分布データである。個体数、区数は無視して、種リストのみを input した。図 2 にその結果を示す。

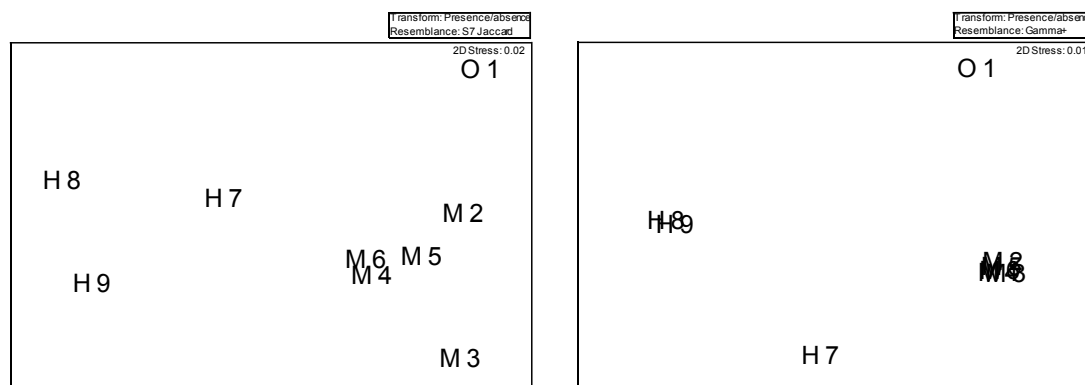


図 2. 田辺湾内外 9 地点の、異なる類似度指数に基づく MDS 分析。左は Jaccard 指数 ($1-J$)、右は分類学的非類似度 TD を用いた。O, 湾外; M, 湾口; H, 湾奥。

これによると、どちらの場合も最湾外 (O1)、湾外—湾口 (M2-M6)、湾奥 (H7-H9) に分類されることでは共通している。これは、分布区数を含めた PCA でも同じ傾向になる (Ohgaki 2007b)。この例では、Jaccard ベースと TD ベースで湾口群集と湾奥群集それぞれの内部各点の相対的接近度に差が見られるが、これにどういう意味があるのか、またどちらが現場の状態をよりの確に反映しているのかなどはわからない。それよりもむしろ、PCA の結果も含めた共通性、つまり、どれで分析しても、湾外／湾口／湾奥がそれぞれ異なるクラスターを構成するという事実を重視すべきである。つまり、成り立ちの異なる指標で評価してもなおかつ同じ結果が出るほど、この 3 域に差があるということである。

多様度、類似度共に、分類学的指数は、調査努力云々よりも、従来無視されてきた各種の分類学的位置という重要な情報を取り込んだ点に意味がある。たとえばこれを生物地理学に応用すれば、種数—個体数型指標の時代には断絶していた、生態学的群

集評価と系統分類学的視点がこれらの指標を介して結びつき、新しい視点を提供することも期待される。その一方、 H' などの従来型指数もそれと別の特徴を持ち、依然有効である。分類学的指数が多様度、類似度のすべてを包含して、従来型を無用のものにするということではない。むしろ分類組込み指標の登場によって、生態現象の分析により豊富な手段が提供されるに至ったことを、多とすべきであろう。

引用文献

- Clarke KR 1993. Non-parametric multivariate analyses of changes in community structure. *Aust.J.Ecol.* 18, 117-143.
- Faith DP, Minchin PR, Belbin L 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69, 57-68.
- Field JG, Clarke KR, Warwick RM 1982. A practical strategy for analysing mutispecies distribution patterns. *Mar.Ecol.Prog.Ser.* 8, 37-52.
- Izsak C, Price ARG 2001. Measuring β -diversity using a taxonomic similarity index, and its relation to spatial scale. *Mar.Biol.Prog.Ser.* 215, 69-77.
- 木元新作 1976. 動物群集研究法 I - 多様性と種類組成. 共立出版
- 大垣俊一 1999. 群集組成の多変量解析. *Argonauta* 1, 15-26.
- Ohgaki S, Kosuge T 2005. A circa-decadal change in the gastropod fauna on a tidal flat on an island mangrove estuary. *Zool.Sci.* 22, 49-56.
- Ohgaki S 2007a. List of intertidal molluscan species around Tanabe Bay, 2004. *Argonauta* 14-19.
- Ohgaki S 2007b. Molluscan assemblages along a bay coastline: gradients related to geographical distribution and feeding habits of species. *Plankton & Benthos Res.* 2, 28-37.
- Warwick RM, Clarke KR 2001. Practical measures of marine biodiversity based on relatedness of species. *Ocean.Mar.Biol.Ann.Rev.* 39, 207-231.
- 八杉龍一・小関治男・古谷雅樹・日高敏隆 1996. 生物学辞典, 第4版.