



•技术与方法•

样本量不一致时的 β 多样性计算

邹怡^{ID*}

西交利物浦大学, 江苏苏州 215123

摘要: 度量样方间物种组成的差异, 即 β 多样性, 是生态学研究中的常用手段。在开展生态学研究的过程中, 不同样方获取的样本量通常不同。使用物种稀疏曲线可以计算不同样本量的 α 多样性, 但常用的 β 多样性指数的计算却没有考虑样本量的差异。本文主要介绍了从稀疏曲线演化而来的可以计算不同样本量的 β 多样性指数——预期共享物种数(expected species shared, ESS)及其标准化后的指数, 其中详细介绍了弦标准化的预期共享物种数(chord-normalized expected species shared, CNESS)。利用真实采集的数据集, 本文演示了在不同样本参数 m 下, CNESS经过主坐标分析(principal coordinates analysis, PCoA)的二维排序结果, 并比较了样本量变化后, CNESS与基于多度的Chao-Jaccard相异性指数之间的差异。模拟结果表明, CNESS指数与Chao-Jaccard指数的PCoA结果具有相关性, 该相关性不随 m 值的变化而变化。CNESS指数较Chao-Jaccard指数具有更多优势, 通过调节样本参数 m , CNESS的结果可以分析优势种或者稀有种的物种组成差异, 同时CNESS指数对样本量不敏感。ESS系列相异指数是基于物种多度的计算, 适用于样本量不一致时的 β 多样性研究, 建议在开展昆虫等无脊椎动物的生态学研究中使用此指数。为了更加准确地获得样方之间的物种组成差异, 在数据分析的过程中应选取不同大小的 m 值计算CNESS。然而, 由于样本量小于特定 m 值的样方会在计算中被剔除, 因此, 在实际的取样工作中, 每个样方都应该尽量采集到足够多的个体, 才能保证在 m 值足够大的时候也不丢失样方信息。

关键词: β 多样性; CNESS; 主坐标分析(PCoA); 相异矩阵; 昆虫; 群落生态学

邹怡 (2021) 样本量不一致时的 β 多样性计算. 生物多样性, 29, 790–797. doi: 10.17520/biods.2021011.

Zou Y (2021) The calculation of β -diversity for different sample sizes. Biodiversity Science, 29, 790–797. doi: 10.17520/biods.2021011.

The calculation of β -diversity for different sample sizes

Yi Zou^{ID*}

Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123

ABSTRACT

Aims: Measuring differences in species composition between plots, i.e., β -diversity, is a common approach in ecological studies. In empirical studies, sample sizes between plots are often inconsistent. While species rarefaction curves can be used to calculate α -diversity for different sample sizes, commonly-used β -diversity indices do not take sample sizes into account. To overcome the limitation, this study introduced the species rarefaction-extended β -diversity index—the expected species shared (ESS) and its normalized format, with particular emphasis on the chord-normalized expected species shared (CNESS) index.

Methods: Based on empirical data, this study demonstrated the application of CNESS using principal coordinates analysis (PCoA) under different sample size parameter (m), and compared results between the CNESS and a commonly used abundance-based index, the Chao-Jaccard index.

Results: Simulation results showed that the PCoA results of the CNESS index and the Chao-Jaccard index were generally positively correlated and that the correlation was largely independent of m . By adjusting m , results of the CNESS can be tuned to focus on the species composition of both dominant and rare species, whereas the Chao-Jaccard cannot represent the relevant information. The CNESS index was not sensitive to the sample size, which offers advantages compared to the Chao-Jaccard index.

Conclusions: ESS-based dissimilarity indices are abundance-based and are suitable for the calculation of β -diversity

收稿日期: 2021-01-08; 接受日期: 2021-04-26

基金项目: 国家自然科学基金(31700363)

* 通讯作者 Author for correspondence. E-mail: yi.zou@xjtlu.edu.cn

when sample sizes vary among plots, which is especially important when studying insects and other invertebrates that commonly have vast differences in the number of samples among plots. In order to comprehensively understand species composition differences between plots, calculation of CNESS results under different m values is recommended. As plots with a sample size smaller than m will be excluded in the calculation, in practice, a sufficient number of individuals are required for each plot to ensure the integrity of plot information under a large m .

Key words: β-diversity; CNESS; principal coordinates analysis (PCoA); dissimilarity matrix; insects; community ecology

生物多样性的测度在生态学的研究中至关重要。Whittaker (1960)提出了生物多样性的3个概念: α多样性、β多样性和γ多样性。其中α多样性强调生境内的物种数, β多样性强调不同环境梯度下的物种组成差异, γ多样性强调区域内的总体物种丰富度。此后, 不同学者提出了一系列计算生物多样性的指数(马克平, 1994), 这些指数基本可归为两类: 一类是基于物种频度分布(presence/absence)的数据进行计算, 另一类是基于物种多度(abundance)的数据进行计算。

β多样性的算法大致可以分为两类: 一类是与γ多样性相关的倍性(γ/α)或加性($\gamma - \alpha$)分配法, 另一类是测量样方之间的相似性或相异性。度量样方之间的相似性(或相异性)被广泛用于测定时间或者空间维度上群落组成的变化(陈圣宾等, 2010; 斯幸峰等, 2017), 其计算结果通常是获得一个样方 × 样

方的距离矩阵(dissimilarity/distance matrix)。基于物种频度分布的β多样性指数有Jaccard指数、Sørensen指数、Cody指数、Whittaker指数等(Koleff et al, 2003); 基于物种多度的β多样性指数包括改进后的Jaccard与Sørensen指数(这里分别称为Chao-Jaccard指数和Chao-Sørensen指数)、Bray-Curtis指数(Chao et al, 2005; Schroeder & Jenkins, 2018), 或者直接采用物种多度的欧式距离(Euclidean distance)等(Legendre & Legendre, 2012)。为了更直观地表达样方之间的差异性, 距离矩阵又可以通过降维排序(ordination)得到一个二维的排序图(Legendre & Gallagher, 2001; Tuomisto, 2010) (图1)。排序的方法包括主坐标分析(principal coordinates analysis, PCoA, 等同于经典多维尺度变换 classical multidimensional scaling, MDS)、非度量多维尺度分析(non-metric multidimensional scaling, NMDS)等。

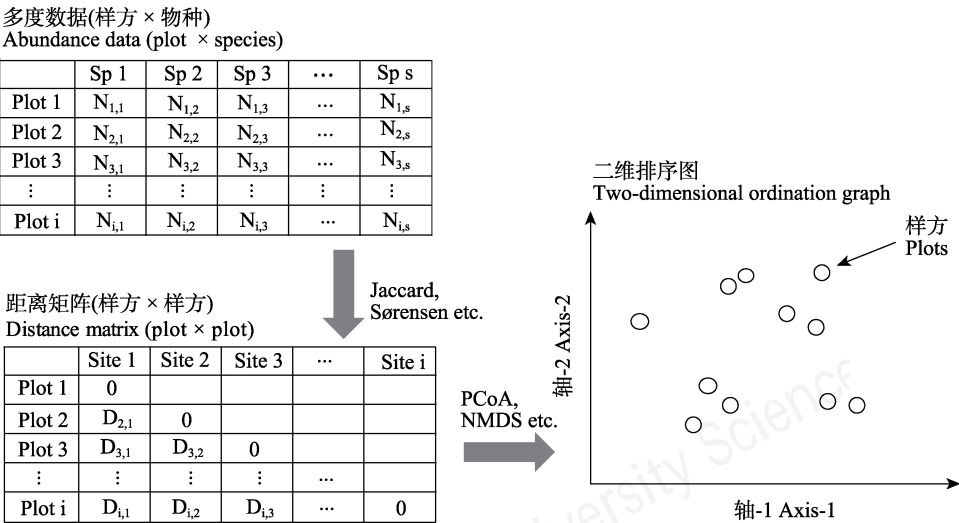


图1 从样方 × 物种的多度数据转换成二维排序图的流程图。原始的多度数据通过计算相异指数可以得到一个两两样方配对的距离矩阵; 该距离矩阵可以通过合适的排序方法得到二维的降序图; 二维图中的一个点代表一个样方, 点与点之间的距离代表样方之间的物种组成是相异(距离比较远)还是相似(距离比较近)。

Fig. 1 Flow chart for converting plot × species multi-dimensional data into a two-dimensional ordination graph. The raw data can be calculated into a pair-wised distance matrix. The distance matrix can be descended into a two-dimensional ordination graph. One point in the ordination graph represents a sampling plot, while the distance between the points represents that the species composition between plot is dissimilar (distance is far) or similar (distance is relatively close).

获取区域内的所有物种后可以对区域生物多样性进行无偏计算,但一般情况下这是很难实现的。样方中采集到的物种数与样本量(sample size)关系很大,而样本量往往是无法控制的,这对于种类繁多的昆虫等无脊椎动物尤其如此。例如,一个研究区域设置了A、B、C 3个样方,通过陷阱采样法比较3个样方地表甲虫多样性的差异。这3个样方中分别得到5、10和12个物种,以及20、100和200个个体。在 α 多样性的计算中,如果直接比较3个样方之间的物种数差异,会得出样方C拥有最高多样性的结论。但这一结论是不准确的,因为无法知道其他两个样方在取样达到200个个体时会得到多少个物种。这种情况下,我们可以采取稀疏标准化(Simberloff, 1972)的方法,通过随机抽样方法计算“种-个体”稀疏曲线(rarefaction curve)来估计出每个样方随机抽取一定个体后的物种数(如上述例子中3个样方都随机抽取20个个体),从而比较样本量标准化后3个样方之间的物种数。同理, β 多样性的计算中,如直接采用物种数而不考虑多度,其计算结果也会有偏差。由于不同样方采集的个体数不一致,因此采集到的物种数量本身会受到物种分布(即优势种或者稀有种的比例)的影响(Legendre & Gallagher, 2001)。而基于多度的 β 多样性指数如Chao-Jaccard、Chao-Sørensen、Bray-Curtis等相异性指数中,计算时只考虑采集到的个体数,并没有一个标准化的样本参数可供选择。

事实上,由 α 多样性“稀疏标准化”方法延伸而来的 β 多样性计算方法——预期共享物种数(expected species shared, ESS)可以选择样本量参数,然而该指数长期未被广泛应用。本文介绍了ESS的来源、数学表达式及其标准化方法,其中重点介绍了弦标准化的预期共享物种数(chord-normalized expected species shared, CNESS)。此外,利用真值数据,本文对比分析了CNESS指数与Chao-Jaccard指数之间的差异,并讨论了ESS系列指数的适用条件以及使用中的注意事项。

1 ESS的来源、算法及其标准化方法

稀疏曲线的计算来源于Hurlbert (1971)提出的预期物种数(expected species, ES),该指数基于超几何概率分布模型设置样本大小参数为 m ,其含义是指在一个拥有 S 个物种和 N 个有限个体的集合中,随

机不放回地抽取 m 个个体所得到的预期物种数。ES的数学表达式为:

$$ES_m = \sum_{k=1}^S \left[1 - \frac{\binom{N - N_k}{m}}{\binom{N}{m}} \right] \quad (1)$$

其中, S 代表观测到的物种总数, N 代表样本中个体总数, N_k 代表同一样本中物种 k 的个体总数。后来,Smith和Grassle (1977)提出,如果群体是无限大的,则抽取 m 个个体所得到的预期物种数符合多项分布(这里记为 ES'),其数学表达式如下:

$$ES'_m = \sum_{i=1}^S \left[1 - (1 - \pi_i)^m \right] \quad (2)$$

其中, π_i 是物种 i 的个体在群体中的概率。

为了解决样本量不一致时 β 多样性的计算,Grassle和Smith (1976)提出预期共享物种数(ESS)。顾名思义,ESS是由ES变换而来,是指两个样方间任意提取 m 个个体所得到的预期共有的物种数(ESS_{ijm}),其数学表达式如下:

$$ESS_{ijm} = \sum_{k=1}^S \left[1 - \frac{\binom{N_{i*} - N_{ik}}{m}}{\binom{N_{i*}}{m}} \right] \times \left[1 - \frac{\binom{N_{j*} - N_{jk}}{m}}{\binom{N_{j*}}{m}} \right] \quad (3)$$

其中, N_{i*} 与 N_{j*} 分别代表样方 i 和样方 j 的个体总数, N_{ik} 与 N_{jk} 代表物种 k 在样方 i 和样方 j 的个体数。Grassle和Smith (1976)也提出了根据 ES' 变换而来的ESS计算方法,但几乎未被使用。

从公式(3)可以看出,在 m 值很大的情况下, $\binom{N_{i*}}{m}$ 部分可能会超过计算机的计算范围,不过该式稍加推导即可解决该问题(推导方式见Zou & Axmacher, 2020):

$$ESS_{ijm} = \sum_{k=1}^S \left[1 - \prod_{n=0}^{m-1} \frac{(N_{i*} - N_{ik} - n)}{N_{i*} - n} \right] \times \left[1 - \prod_{n=0}^{m-1} \frac{(N_{j*} - N_{jk} - n)}{N_{j*} - n} \right] \quad (4)$$

与ES类似,ESS的计算结果大于0,代表了共享物种数的多少。然而,研究物种组成的相异性时,我们通常用值域0 (100%相同)至1 (100%相异)来表示,ESS直接计算的结果显然不够直观。为了解决这

一问题, Grassle和Smith (1976)进一步将ESS值标准化, 使其符合值域在0–1之间的相异指数, 称之为标准化的预期共享物种数(normalized expected species shared, NESS):

$$\text{NESS}_{ij|m} = \frac{2 \times \text{ESS}_{ij|m}}{\text{ESS}_{ii|m} + \text{ESS}_{jj|m}} \quad (5)$$

当 $m = 1$ 时, NESS的(相异性)值与Horn-Morisita指数的(相似性)值一致(即 $\text{Horn-Morisita} = 1 - \text{NESS}_{m=1}$), 因此可以认为NESS是Horn-Morisita指数的延伸(Grassle & Smith, 1976)。

NESS针对ESS的标准化使用的是算术平均值, 这易受极值的影响。Trueblood等(1994)在研究波士顿港泥滩生物时, 使用几何平均值来对ESS进行标准化, 称为弦标准化的预期共享物种数(CNESS), 可以弱化极值的影响。然而, Trueblood等(1994)的CNESS指数值域在 $0-\sqrt{2}$ 之间, 无法像传统的值域在0–1之间的 β 多样性指数那样直观地表达样方之间的相异性。近期Zou和Axmacher (2020)对Trueblood等(1994)提出的CNESS指数略加修改, 使其值域在0–1之间, 表达式如下:

$$\text{CNESS}_{ij|m} = \sqrt{1 - \frac{\text{ESS}_{ij|m}}{\sqrt{\text{ESS}_{ii|m} \times \text{ESS}_{jj|m}}}} \quad (6)$$

通过选择参数 m 值, NESS/CNESS可以比较两个样本量相同的样方之间的物种组成相异性。显然, 不同的 m 取值对于NESS/CNESS的计算结果是有影响的。 m 的取值越小, 其结果越受到优势种的影响; 反之, m 越大则结果越受到稀有种或者总体的物种组成的影响(Zou & Axmacher, 2020)。

NESS/CNESS适用于比较不同采样强度(sampling completeness)的样方间的物种组成差异, 被用于昆虫等无脊椎动物 β 多样性的研究(Beck & Vun Khen, 2007; Liu et al, 2010; 段美春等, 2012; Zou et al, 2014; Bonifácio et al, 2020)。然而, 相比于其他常见的相异性指数, NESS/CNESS并没有得到广泛应用, 原因之一是没有合适的软件。在此前的研究中, ESS系列指数相关的计算都是在基于MS-DOS系统开发的Compah96软件上完成(Gallagher, 1998), 该软件的链接目前已经无法打开。即使在使用Compah96软件时, 进行大量且重复的计算也十分困难。相比之下, 其他的 β 多样性指数的计算均可在开源软件R语言(R Core Team, 2020)

平台中实现。近期Zou和Axmacher (2020)以及Menot (2019)独立开发了用于计算ESS系列指数(包括ESS、NESS和CNESS)的R程序, 解决了软件计算的问题。此外, Zou和Axmacher (2020)运用模拟的数据, 对比了在不同样本量下CNESS指数与Bray-Curtis指数、Chao-Sørensen指数以及欧式距离所得到的结果稳定性的差异, 结果表明CNESS在不同的取样场景下表现都比较稳定。尽管如此, 针对真实的数据, 使用降维排序后, CNESS对比传统的 β 多样性在不同样本下的表现如何, 目前并没有相关的研究。

2 演示与模拟

本文利用真实采集的数据, 对比了在不同的样本参数 m 下的CNESS指数、Chao-Jaccard指数经过PCoA排序后的结果。此外, 对比了不同 m 值下CNESS与Chao-Jaccard的二维排序轴之间的关系。最后, 通过模拟样本量的变化, 评估成对样方间CNESS与Chao-Jaccard相异性平均值的稳定性。

本文的数据来自vegan包(Oksanen et al, 2018)中的“mite”数据集, 该数据集来自加拿大魁北克省蒙特利尔大学劳伦蒂斯生物研究站的甲螨数据(蜱螨目: 甲螨亚目), 包含70个土壤样品的35个形态种(Borcard & Legendre, 1994)。本研究将每个土壤样品视为一个样方, 该数据集共拥有70个样方, 编号1–70。共有9,800个个体, 其中样本量最小的样方包含8个个体, 最大的包含781个个体, 平均每样方包含140个个体。数据集中物种的多度分布见附录1。

首先, 计算CNESS与Chao-Jaccard相异指数, 得到样方之间配对的距离矩阵(pairwise distance matrix), 然后查看不同指数对群落物种结构相异性的变化。CNESS函数有两个参数: x 和 m , 其中 x 是样方 \times 物种(行 \times 列)的矩阵, m 是两个样地之间共有的样本量参数。理论上, m 值可为 ≥ 1 的正整数, 但是当 m 值大于样方内总体的样本数时, 该样方就会在计算中被剔除。本次计算中, CNESS分别采用了 $m = 1$ 、8和140这3个值, 分别代表最小 m 取值、样本量最小的样方所包含的个体数、所有样方的平均个体数。得到相异矩阵后, 用PCoA将相异矩阵二维化, 得到一个拥有两个轴(Axis-1和Axis-2)的二维图。

其次, 对基于两个不同指数得到的两个轴分别进行线性回归, 检验CNESS指数与Chao-Jaccard之间二维化之后所表达信息的关系。通过计算不同 m

值(1–140)的CNESS, 得到不同的回归系数, 查看CNESS指数与Chao-Jaccard指数得到的两个轴的相关关系是否受到 m 值的影响。

最后, 采取随机移除数据集中的个体的方式分别计算新的样本的CNESS与Chao-Jaccard的平均值, 移除的个体数从100依次递增至2,000 (约占总样本数的20%)。通过计算样本量变化后的CNESS ($m = 1, 8$ 与140)与Chao-Jaccard的样方间平均相异性的变化率来查看不同指数对于样本量的稳定性, 计算方法为: 变化率 = (变化后样本平均相异性 – 原样本平均相异性) / 原样本平均相异性 $\times 100\%$ 。

本文涉及的计算均在R软件(R Core Team, 2020)中完成, 使用vegan包(Oksanen et al, 2018)计算PCoA, 使用CommEcol包(Melo, 2014)计算Chao-Jaccard指数。CNESS指数的计算采用Zou和

Axmacher (2020)开发的ESS()函数。模拟与计算过程的R语言代码见附录2。

3 结果

从PCoA的排序图中可以看出, 不同 m 值的排序结果并不相同。数据集中采集个体数最少的样方包含8个个体, 因此当 $m = 8$ 时, 没有样方在计算中被剔除。当 $m > 8$ 时, 一定数量的样方被剔除; $m = 140$ 时, 排序图中只拥有24个样方(44个样方因少于140个个体而被剔除计算)(图2)。

基于CNESS的PCoA排序结果与Chao-Jaccard的排序结果具相似之处, 但也有明显差异。如样地7, 13, 21和24在CNESS ($m = 1$ 和 $m = 8$)和Chao-Jaccard的排序中均聚集在一起, 而样地57和样地62与其他样地的差异均较大(图2)。线性回归结

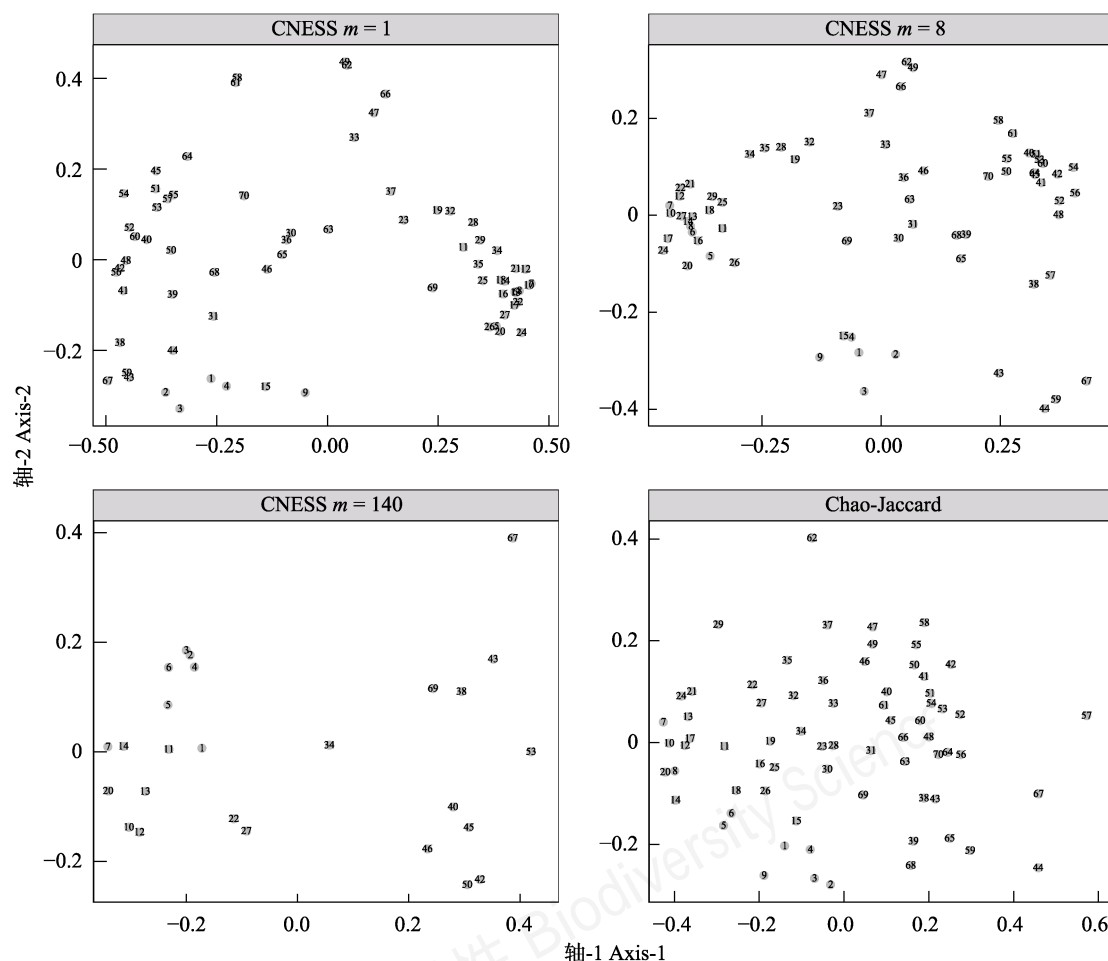


图2 不同取样参数大小($m = 1, 8$ 和140)下的基于弦标准化的预期共享物种数(CNESS)指数以及Chao-Jaccard指数的主坐标分析(PCoA)二维排序图

Fig. 2 Two-dimensional principal coordinates analysis (PCoA) plots based on chord-normalized expected species shared (CNESS) index with different sampling size parameter ($m = 1, 8, 140$) and based on the Chao-Jaccard index

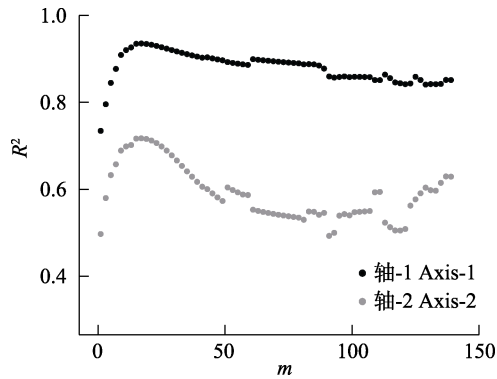


图3 基于弦标准化的预期共享物种数(CNESS)指数的主坐标分析(PCoA)排序产生的两个轴(Axis-1与Axis-2)与基于Chao-Jaccard的PCoA排序产生的两个轴分别进行线性回归后的回归系数(R^2)在不同 m 值下的变化

Fig. 3 Linear regression value (R^2) between respective axes (Axis-1 and Axis-2) that generated by chord-normalized expected species shared (CNESS) index based and Chao-Jaccard based principal coordinates analysis (PCoA) plots at different m values

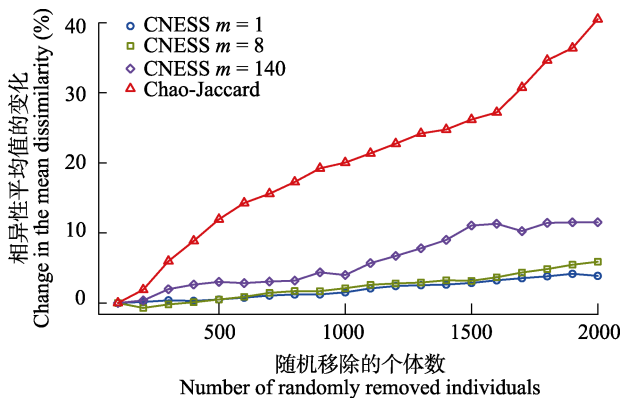


图4 弦标准化的预期共享物种数(CNESS) ($m = 1$ 、8和140)和Chao-Jaccard相异性的平均值的变化率与随机移除的个体数之间的关系

Fig. 4 Relationship between the changes of mean dissimilarity distance with the number of individuals randomly removed from the dataset for chord-normalized expected species shared (CNESS) index with different m value ($m = 1$, 8 and 140) and the Chao-Jaccard index

果表明, 基于CNESS与基于Chao-Jaccard排序产生的轴之间的相关程度比较高(对于1–140之间的任意 m 值, 轴1的相关系数 $R^2 > 0.7$, 轴2大多数 $R^2 > 0.4$, P 值均 < 0.001), 其相关系数并未随 m 值的变化发生显著变化(图3)。

对比随机移除一定数量个体后的新数据集与完整的数据集的结果, CNESS所得的样方间均值变化较小, 在移除2,000个个体时, $m = 1$ 、8和140的均值变化分别为4.9%、5.4%和8.0%。相应地, Chao-

Jaccard的变化率高达31% (图4)。

4 讨论

通过对来自mite数据集的计算, 本文演示了不同样本量的CNESS指数在二维排序后的结果, 并比较了CNESS与基于多度的Chao-Jaccard指数所得结果的差异。虽然本文只针对CNESS指数进行了模拟, 但由于ESS系列的其他相异指数(如NESS)与CNESS有类似的特点, 所得的结果也可以推广到其他ESS系列的相异性指数。

ESS系列相异指数的优势之一是可以自由选择样本参数 m 值的大小。 m 取值较小时, 其结果侧重于优势种, m 取值较大时, 其结果侧重于整体样本, 即随着 m 的增大, 稀有种的侧重度会逐渐增加(Zou & Axmacher, 2020)。PCoA排序结果表明, mite数据集在反映优势种和整体样本的物种组成差异上并不一致, 而这一结果在Chao-Jaccard指数上无法呈现。虽然依据不同 m 值计算出的CNESS指数不同, 但其PCoA的结果跟Chao-Jaccard指数之间的关系却是独立于样本参数的, 说明在任意 m 值下, CNESS的计算结果跟Chao-Jaccard等指数都具有可比性。

ESS系列相异指数的另一优势是其结果对于实际采集的样本量不敏感。从相异性平均值的变化率可以分析出, 即使在移除了2,000个个体后, CNESS的平均值仍然跟原始数据十分接近, 而Chao-Jaccard却产生了将近1/3的变化, 该结果跟Zou和Axmacher (2020)的模拟结果一致。Zou和Axmacher (2020)通过对比CNESS与Chao-Sørensen指数、Bray-Curtis指数以及基于个体百分比的欧氏距离发现, 即使取样完整度很低, 多次模拟重复采样后CNESS结果的变异系数也比其他指数低, 说明CNESS指数在采样不完整时计算结果具稳定性。在开展昆虫等无脊椎动物的生物多样性调查工作中, 经常采取以生物个体为单位的离散型采样(discrete sampling)方式(Colwell et al, 2012; Chao & Chiu, 2016), 这通常会导致不同样方之间的样本量具有差异。由于CNESS具有对样本量不敏感的特性, 其计算结果在不同样本量下有较好的稳定性, 推荐用于取样大小不一致时的 β 多样性计算。

CNESS (以及NESS)是由ESS标准化而来, 用以描述两个样方之间物种 β 多样性的差异。值得注

意的是, ESS指数本身所得到的结果是(样方之间共享的)物种数量。近期, Zou和Axmacher (2021)通过对ESS的外延, 开发了一个新的指数——总预期共享物种数(total expected species shared, TESS)。在采样不完全的情况下, 该指数可估计出两个样方之间拥有共享物种的数量。通过结合 α 多样性的估计值, 如ACE (abundance-based coverage estimator)、Chao-1指数等, TESS拥有对任一基于频度分布的 β 多样性指数(如传统的Jaccard和Sørensen指数)进行无偏估算的潜力, 因而拥有广泛的应用前景(Zou & Axmacher, 2021), 需开展更多的评估性研究。


由于ESS系列指数的计算是基于物种被采集到的概率随着物种多度而变化的超几何分布模型, 因此, 包括CNESS在内的ESS系列指数均不适用于基于物种数(频度分布)的数据。而大多数 β 多样性的计算均仅考虑了物种数而忽略物种多度分布(Koleff et al, 2003), 可能是由于其结果比较容易解释。目前ESS系列指数只针对以个体为单位的离散型采样方式, 而在植被等调查中, 多采用以面积为单位的连续型采样(continuous sampling)方式(Chao & Chiu, 2016), 针对后者ESS并没有相应的计算方式。相对而言, 在 α 多样性的计算中, 针对这两种采样方式有“种-个体”曲线和“种-面积”曲线可供选择(Simberloff, 1972)。虽然将ESS进一步开发成可以针对面积的计算方法或许并不困难, 但其应用场景可能比较有限, 主要由于植被调查中不同样方之间的“样本(面积)大小”通常是一致的, 而ESS恰可用于样本大小不一致的 β 多样性计算。即使样方面积大小不同, 设置了(能够用于ESS进行标准化)多个小样方以后, 该样方或已经包含了大部分植物物种(相对于种类繁多的昆虫等无脊椎动物而言)。这种情况下, 基于物种数的 β 多样性指数与使用较大 m 值的CNESS/NESS指数的结果虽然在绝对值上可能不同, 但在样方之间相对距离的结果将趋与一致。上述局限性或许是ESS系列指数没有受到广泛关注的原因之一。

综上所述, ESS系列相异指数适用于样方之间样本量不一致、样本量会影响观测到的物种的情况, 且目前该系列指数只可用于物种多度的计算。因此, ESS系列指数比较适合用于样本量很难统一且种类繁多的昆虫等无脊椎动物的研究, 也可用于与样本量关系密切的基于操作分类单元(operational

taxonomic units, OTU)的微生物群落的 β -多样性研究中。在实际研究工作中使用CNESS指数(或ESS系列的其他指数)时, 建议选取多个 m 值进行计算, 呈现的不同结果既可以关注优势种也可以关注整体的物种组成的差异。值得注意的是, m 的取值较大时虽然有利于估计整体物种组成的差异, 但是为了使样方之间具有可比性, 计算结果中样本量小于 m 的样方会被剔除。因此, 为了保证样方的信息能够得到表达, 在采样的过程中, 建议每一个样方都尽可能地采集足够数量的个体, 这对开展昆虫等无脊椎动物的物种组成研究具有重要意义。

致谢: 衷心感谢三位审稿专家以及编委对本文提出的宝贵意见和建议; 感谢太原师范学院孙晓杰博士在文稿校对方面提供的帮助。

ORCID

邹怡  <https://orcid.org/0000-0002-7082-9258>

参考文献

- Beck J, Vun Khen C (2007) Beta-diversity of geometrid moths from northern Borneo: Effects of habitat, time and space. *Journal of Animal Ecology*, 76, 230–237.
- Bonifácio P, Martínez Arbizu P, Menot L (2020) Alpha and beta diversity patterns of polychaete assemblages across the nodule province of the eastern Clarion-Clipperton Fracture Zone (equatorial Pacific). *Biogeosciences*, 17, 865–886.
- Borcard D, Legendre P (1994) Environmental control and spatial structure in ecological communities: An example using oribatid mites (Acari, Oribatei). *Environmental and Ecological Statistics*, 1, 37–61.
- Chao A, Chazdon RL, Colwell RK, Shen TJ (2005) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, 8, 148–159.
- Chao A, Chiu C (2016) Species richness: Estimation and comparison. In: *Wiley StatsRef: Statistics Reference Online* (eds Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL), pp. 1–26. John Wiley & Sons, Inc.
- Chen SB, Ouyang ZY, Xu WH, Xiao Y (2010) A review of beta diversity studies. *Biodiversity Science*, 18, 323–335. (in Chinese with English abstract) [陈圣宾, 欧阳志云, 徐卫华, 肖懿 (2010) Beta多样性研究进展. *生物多样性*, 18, 323–335.]
- Colwell RK, Chao A, Gotelli NJ, Lin SY, Mao CX, Chazdon RL, Longino JT (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of*

- Plant Ecology, 5, 3–21.
- Duan MC, Liu YH, Wang CL, Axmacher JC, Li LT, Yu ZR (2012) Diversity of geometrid moth (Lepidoptera: Geometridae) in cropland and reforested semi-natural habitats at different altitudes of Bashang Plateau, Hebei Province of China. *Chinese Journal of Applied Ecology*, 23, 785–790. (in Chinese with English abstract) [段美春, 刘云慧, 王长柳, Axmacher JC, 李良涛, 宇振荣 (2012) 坝上地区不同海拔农田和恢复半自然生境下尺蛾多样性. *应用生态学报*, 23, 785–790.]
- Gallagher ED (1998) Compah96. <http://www.es.umb.edu/faculty/edg/files/pub/COMPAH.EXE> (accessed on 2013-09-01)
- Grassle JF, Smith W (1976) A similarity measure sensitive to the contribution of rare species and its use in investigation of variation in marine benthic communities. *Oecologia*, 25, 13–22.
- Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecology*, 52, 577–586.
- Koleff P, Gaston KJ, Lennon JJ (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, 72, 367–382.
- Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129, 271–280.
- Legendre P, Legendre L (2012) *Numerical Ecology*, 3rd edn. Elsevier, Amsterdam.
- Liu YH, Axmacher JC, Wang CL, Li LT, Yu ZR (2010) Ground beetles (Coleoptera: Carabidae) in the intensively cultivated agricultural landscape of Northern China—Implications for biodiversity conservation. *Insect Conservation and Diversity*, 3, 34–43.
- Ma KP (1994) Measurement of biotic community diversity. I. α diversity (Part 1). *Chinese Biodiversity*, 2, 162–168. (in Chinese) [马克平 (1994) 生物群落多样性的测度方法. I. α 多样性的测度方法(上). *生物多样性*, 2, 162–168.]
- Melo AS (2014) CommEcol: Community Ecology Analyses. R package version 1.6.5. <http://CRAN.R-project.org/package=CommEcol>. (accessed on 2021-02-01)
- Menot L (2019) ness: Hypergeometric standardization and probabilistic measures of similarity for community ecology. R package version 0.1.0, Zenodo. <https://doi.org/10.5281/zenodo.3254750>. (accessed on 2021-02-01)
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MRH, Wagner H (2018) vegan: Community Ecology Package. R package version 2.5-6. 10.1111/j.1654-1103.2003.tb02228.x. (accessed on 2021-02-01)
- R Core Team (2020) R: A Language and Environment for Statistical Computing. Version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria.
- Schroeder PJ, Jenkins DG (2018) How robust are popular beta diversity indices to sampling error? *Ecosphere*, 9, e02100.
- Si XF, Zhao YH, Chen CW, Ren P, Zeng D, Wu LB, Ding P (2017) Beta-diversity partitioning: Methods, applications and perspectives. *Biodiversity Science*, 25, 464–480. (in Chinese with English abstract) [斯幸峰, 赵郁豪, 陈传武, 任鹏, 曾嶙, 吴玲兵, 丁平 (2017) Beta多样性分解: 方法、应用与展望. *生物多样性*, 25, 464–480.]
- Simberloff D (1972) Properties of the rarefaction diversity measurement. *The American Naturalist*, 106, 414–418.
- Smith W, Grassle JF (1977) Sampling properties of a family of diversity measures. *Biometrics*, 33, 283–292.
- Trueblood DD, Gallagher ED, Gould DM (1994) Three stages of seasonal succession on the Savin Hill Cove mudflat, Boston Harbor. *Limnology and Oceanography*, 39, 1440–1454.
- Tuomisto H (2010) A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, 33, 23–45.
- Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30, 279–338.
- Zou Y, Axmacher JC (2020) The Chord-Normalized Expected Species Shared (CNESS)-distance represents a superior measure of species turnover patterns. *Methods in Ecology and Evolution*, 11, 273–280.
- Zou Y, Axmacher JC (2021) Estimating the number of species shared by incompletely sampled communities. *Ecography*, 44, ecog.05625.
- Zou Y, Sang WG, Zhou HC, Huang LY, Axmacher JC (2014) Altitudinal diversity patterns of ground beetles (Coleoptera: Carabidae) in the forests of Changbai Mountain, Northeast China. *Insect Conservation and Diversity*, 7, 161–171.

(责任编辑: 王少鹏 责任编辑: 黄祥忠)

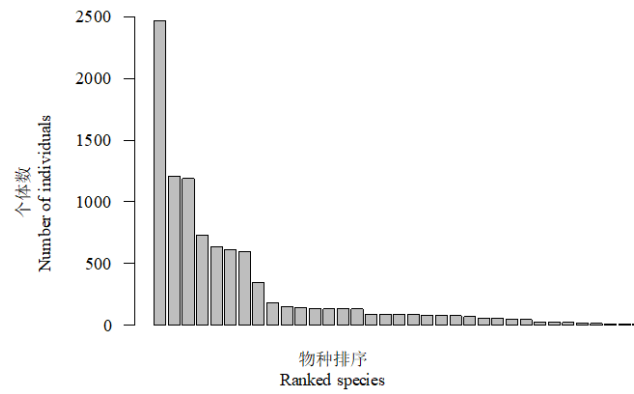
附录 Supplementary Material

附录1 “mite”数据集各物种的多度分布

Appendix 1 The abundance distribution for the species in the “mite” dataset
<https://www.biodiversity-science.net/fileup/PDF/2021011-1.pdf>

附录2 用于演示和模拟的R语言代码

Appendix 2 R scripts used for demonstration and simulation
<https://www.biodiversity-science.net/fileup/PDF/2021011-2.txt>



附录 1 “mite”数据集各物种的多度分布

Appendix 1 The abundance distribution for the species in the “mite” dataset