

仰觀宇宙之大，俯察品類之盛：如何量化生物多樣性

趙蓮菊 邱春火 王怡婷 謝宗震 馬光輝

國立清華大學統計研究所

摘要

生物多樣性係指所有生命個體的變異及其存在多樣生態系統的總稱，並且延伸出相似性與分化的相關概念。生物多樣性至少有三個層面：基因（遺傳）多樣性，物種（分類）多樣性和生態系統（功能）的多樣性。本文主要回顧物種（或分類）多樣性，其中涵蓋以下主題：(1) 根據 Hill 數值的傳統單一群落多樣性測度。這一類的度量只考慮物種數和物種的豐富度。(2) 依據 Hill 數值意義所建構的單一群落系統演化多樣性測度，此類測度除了考慮物種數和物種豐富度，也加入物種間演化/系統親緣距離的信息。(3) 傳統多樣性的分解及其相關的多群落間相似性（或分化）的測度。(4) 系統演化多樣性的分解與其相關的多群落間相似性（或分化）的測度。另外，我們舉一個實際的例子來說明統計模型在估計生物多樣性所扮演的角色。最後，對於未來一些相關的主題也一併進行討論。

關鍵詞：生物多樣性，多樣性的分解，系統演化多樣性，倍增原則，分類多樣性。

JEL classification: C13.

1. 緒論

1.1 何謂“生物多樣性”

生物多樣性 (biodiversity) 一詞係 biological diversity 的簡稱，最早是 E. O. Wilson 在 1986 年美國華盛頓區舉行之“生物多樣性國家論壇” (National Forum on Biological Diversity) 會議之論文中提出。生物多樣性泛指地球上所有生物的多樣化與變化性，它涵蓋了所有

從基因、個體、族群、群落、生態系到地景等各種層次的生命型式。生物多樣性越豐富，代表其組成越複雜，不易受到單一的破壞而產生滅絕。

1.2 生物多樣性的層次

生物多樣性層次包含基因 (遺傳) 多樣性、物種 (分類) 多樣性以及生態系 (功能性) 多樣性，這三個層級包含了地球所有生命現象的總和且彼此相互依存，缺一不可，其簡單的介紹如下：

基因(遺傳) 多樣性 Gene (genetic) diversity

基因的多樣性係指同一物種個體基因的變異與不同。基因多樣性也稱為遺傳多樣性，即物種基因隨時間，在不同的環境下，經基因漂變 (random drift) 或突變 (mutation) 等內在因素，及天擇 (natural selection) 或性擇 (sexual selection) 等外在因素，而使個體基因漸漸產生變異。當遺傳多樣性越高時，越容易適應環境，有利於族群的生存及演化。

基因的多樣性廣泛使用在農業與品種的改良，例如稻米因為基因的不同，從過去大家熟知的在來米、蓬萊米、糯米、長秈米，至今日基因改良的益全香米 (台農 11 號)、驚為天米 (台南 11 號)、壽司米 (台梗 9 號)、晶鑽米 (高雄 145 號)、醜美人 (高雄 139 號)，在市場上琳瑯滿目，眼花瞭亂。由於基因的不同，產品因而有不同的特質與優劣。例如，益全香米有特別的芋頭香味，驚為天米晶瑩剔透，一看就與眾不同，醜美人米雖不晶瑩好看但口感極佳。基因與品種的改良造就了所謂的優良品種，例如產量高、外觀佳的稻種。但這些優良品種在另一方面很可能失去了抵抗蟲害的能力。例如 70 年代末期東亞地區的稻米遭到草狀矮化病的侵襲，改良品種沒有抵抗病原的能力，人們賴以維生的稻米面臨重大的危機。最後科學家在印度的山谷找到具有抗病基因的三株野生種稻米才取得抗病的基因，因而挽回了稻米生產的危機。因此“天生物種必有其用”，而保育遺傳多樣性也是生物多樣性保育重要的一環。

物種(分類) 多樣性 Species (taxonomic) diversity

物種多樣性係指某一地區中物種種類的變異與不同。當我們在清華校園中觀賞鳥類時，城市三劍客的麻雀、綠繡眼、白頭翁總是一眼就看到，湖邊不時有小白鷺、翠鳥、夜鷺，而金背鳩、紅鳩、樹鵲、喜鵲、黃頭鷺、黑冠麻鷺也經常出現，運氣好的時候可以看到較稀有的黑枕藍鶺鴒，另外春天來臨時可以見到家燕、洋燕和小雨燕在天空遨翔，多姿多彩的鳥類豐富了校園的生命與生活。清華校園內大約有多少種的鳥類呢？曾經觀察到的大約 6、70 種。當我們看新竹地區或甚至擴大整個臺灣地區的時候，鳥的種類就更豐富多樣了。根據臺灣鳥類誌 (劉小如等人) 所撰，臺灣的鳥類大約有 543 種，留鳥中就有 15% 的臺灣特有種，是世界少有的“寶島”。當我們

放眼全世界時，又有多少鳥種呢？根據 Avibase 資料庫 (<http://avibase.bsc-eoc.org/>) 大約有一萬種鳥種。如果我們看全世界所有的物種，大約有多少呢？根據最近 Mora et al. (2011) 及 Costello (2013) 論文，全世界大約有 150 萬種已知且被命名，但仍有許多未被發現的物種，估計全球的物種大約在 200 萬到 800 萬之間。換句話說，絕大多數的物種都尚未被發現。因此可以知道欲估計全球未被發現的物種不是一個簡單的估計問題。本文主要探討“物種多樣性”這個層級，定義及討論量化多樣性的指標，見第二節的簡介與後續的詳細介紹。

生態系(功能性) 多樣性 Ecosystem (functional) diversity

生態系是一個動態的複合體，一個地區的生態系統由動物、植物、微生物及無生命的環境(陽光、空氣、水、礦物質、養份) 交織而成的生態環境。因地區的不同而有各種生態系統如森林生態系(熱帶雨林、溫帶闊葉林、寒帶針葉林)、海洋生態系、沙漠生態系、草原生態系、河流生態系、濕地生態系等。生態系或功能性多樣性即是指各物種在一個生態系中功能的差異與不同。各物種的特質(trait) 決定了此種差異。例如研究鳥類一般研究者會記錄鳥類的覓食行為包括鳥類的體型、體重、食物的種類、覓食區的平均高度與與高度範圍等。由這些特質得以量化物種之間的差別，再進一步探討生態系中各種物種形成的不同功能群組，以及探討物種之間的合作競爭與維持生態系的永續。此種生態系多樣性與生存在其中的物種多寡和物種系統演化的歷史也息息相關，因此功能性多樣性與物種多樣性也密切關連。

生物多樣性的重要性

生物多樣性的重要性，在於它可以提供一個穩定生態系與自然環境，是人類食衣住行育樂和藥物的來源，更是人類的健康與經濟所繫。生物多樣性是人類最珍貴的資產，後代子孫最不能原諒我們的就是生物物種的滅絕(Wilson, 1992)，為了挽救基因消失、物種滅絕、生態體系劣化，1992年6月巴西里約熱內盧舉行的聯合國環境與開發大會期間，全世界150國的政治領袖簽署了一份生物多樣性保育公約(Convention on Biological Diversity; CBD, 2012)，聯合國訂2010年為生物多樣性年(其標語為 Biodiversity is life, biodiversity is our life)，以期達到保育生物多樣性、永續利用、與公平合理分享生物多樣性資源所產生的利益。然而環境與物種因國而異，各國都應致力其本土的生物多樣性的保育，同時教育每一個人對自然的尊重與對生命的關懷才是最重要的關鍵所在。

2. 物種 (分類) 多樣性的量化簡介

當我們了解生物多樣性的意義與重要性後，接下來面對的問題是如何量化多樣性？如何在有限資源利用下，保護更多的生命多樣性，變成一極迫切的議題。本文主要探討“物種多樣性”這個層級，討論如何量化生物多樣性的指標。本文的主題與結果亦可應用在不同層級多樣性的研究。

古典生物多樣性指標(第三節)

雖然物種數在生態保育上是最直觀且應用最廣的指標，但由於抽樣的限制，往往無法調查到所有的物種，而估計在樣本中沒有調查的物種是一個極困難的統計問題 (Chao, 2005; Gotelli and Chao, 2012; Chao and Chiu, 2013)。另一方面，物種數並無考量物種豐富度的不同，(亦即有些物種個體數目多，有些物種個體數目少)。在第二節中我們將討論如何利用物種數與物種豐富程度來建立各種指標以量化生物多樣性，此種指標稱為古典多樣性指標。

系統演化多樣性指標(第四節)

古典多樣性指標都只考慮群落的物種數和其豐富度，沒有考慮物種間的差異(或相似)，即假設物種間的差異(或相似)都是一樣的，然而物種間存在著差異(或相似)，當物種以界、門、綱、目、科、屬、種分類時，可以分類來代表物種間的差異性(或相似性)。當我們考量此種差異時，古典多樣性指標必須推展至“類別多樣性”(taxonomic diversity)。同時由於分子生物技術的進步，各物種何時分化已很容易判定，因此物種多樣性亦可考量系統演化的歷史，而推展至“系統演化多樣性”(phylogenetic diversity)。本文將在第四節中討論“系統演化多樣性”。

比較多群落或地區的古典生物多樣性指標(第五節) 與

比較多群落或地區的系統演化多樣性指標 (第六節)

物種多寡的分布因空間與時間而有所不同。例如在熱帶的樹種種類繁多，而在寒帶的樹種稀少，在平地螞蟻物種多於平地，古時的蘭花種類沒有現在多。如何量化不同地區的差異性與相似性？一般多採用多樣性分解的方法。Whittaker (1960, 1972, 1977) 認為一個區域的生物多樣性可區分為 α 、 β 與 γ 三種， γ 多樣性即整個區域的多樣性， α 多樣性為各個群落多樣性的平均，即群落內之多樣性， β 多樣性即群落間之多樣性以描述群落間的不同與變化。因此 β 多樣性可用以架構多群落間的相似度或分化度。此種相似度或分化度進而可以用來比較多個區域(每一地區各有不同的群落數)。本文將在第五節中討論多群落或地區的古典生物多樣性指

標, 在第六節中討論多群落或地區的系統演化多樣性指標。

統計在生物多樣性研究的角色(第七節)

生物多樣性的研究由於牽涉學門甚廣(包括生物、生態、保育、地理資訊、生物資訊、計算機科學、數學、統計) 需要各領域之學者共同合作, 相互作用。本文將以一個實例來說明統計在生物多樣性估計中所扮演的角色, 以及相關軟體開發的重要性。

討論與展望(第八節)

在最後一節中, 我們將做一個綜合討論目前仍待解決的統計問題作為此研究方向的展望。

3. 單一群落的古典多樣性指標

生物多樣性的量化研究的發展可分為三部份 (Magurran, 2004; Magurran and McGill, 2010): 曲線配適法 (curve fitting)、參數法 (parametric approach) 與無參數法 (non-parametric approach), 分述於下三節, 本文主要專注於無參數法。

3.1 曲線配適法 (curve fitting)

生物學家早期量化的工作起始於描述物種多寡與調查之個體數目或面積的關係。例如物種累積曲線 (species accumulation curve) 描繪當調查之個體數目增加時, 物種的數目如何隨之增加, 或物種-面積曲線 (species-area curve) 描繪當調查面積增加時, 物種數目如何隨面積增加。此兩種曲線為一遞增曲線並且當個體數或面積夠大時平滑趨近一固定值, 此趨近值即為全部的物種數。

Arrhenins (1921) 最早提出指數曲線, 即 $S_t = \alpha\beta^t$ (S_t 代表 t 單位之個數或面積時所調查到之物種數目) 或 Gleason (1922) 提出之半對數曲線 $S_t = \alpha + \beta \log t$, 其中 α, β 為參數。此種曲線因為不會趨近平滑, 都僅應用在全部調查地區面積已知的情況或假設有限的調查個數可以完全看到所有的物種。

其後學者提出許多平滑且會趨近一固定值的曲線用以描述物種累積曲線或物種面積曲線, 最常用的有 (Flather, 1996):(下列之 α, β, μ 均為參數, S 為當 t 趨近無窮大時之收斂值)。

- (1) 廣義指數曲線: $S_t = S[1 - \exp(-\alpha t)]^\beta$ 及 $S_t = S\{1 - \exp[-\alpha(t - \beta)^\mu]\}$ (Weibull 曲線)。

(2) Michaelis-Menten 曲線 $S_t = St/(\beta + t)$ 及其推廣 $S_t = (\alpha + St)/(\beta + t)$ 及 $S_t = St^\alpha/(\beta + t^\alpha)$ (logistic 曲線)。

(3) 其它曲線 $S_t = S(1 - \alpha\beta^t)$ 以及 $S_t = S\{1 - [1 + (t/\alpha)^\beta]^{-\mu}\}$ 。

此法的好處是只須物種出現與否的記錄, 不須物種豐富度記錄, 因此不須描述物種豐富度分布, 並且可應用至所有的調查型態。但它的缺點是 (1) 描繪一條曲線必須要夠多的資料 (2) 許多不同的曲線可能都配適很好, 但其物種估計卻相差甚遠 (3) 好的配適並不一定得到好的物種估計 (4) 一般曲線要看取樣順序, 不同順序可能有不同估計值 (5) 因為沒有使用統計抽樣的模型, 所以無法探知抽樣誤差因而曲線的選擇沒有客觀的理論。

3.2 參數法 (parametric approach)

Fisher et al. (1943) 在 Journal of Animal Ecology 的論文可以說是生物多樣性統計理論之先趨奠基者。在此篇論文中 Fisher et al. 首度採用了統計抽樣的模型, 即假設第 i 個物種在抽樣時間內被看到的過程是一個卜瓦松過程具有參數 λ_i , $i = 1, 2, \dots, S$ 。這些參數代表了物種的豐富度 (λ_i 愈大表示被看到的期望次數愈多, 因而豐富度愈高)。Fisher 假定 $(\lambda_1, \lambda_2, \dots, \lambda_S)$ 為一個來自伽偶分布具有密度函數 $f(\lambda) = \lambda^{a-1}e^{-\lambda/b}/[\Gamma(a)b]^a$, 則任一物種出現 k 次的機率 $P_{a,b}(k)$ 為

$$\begin{aligned} P_{a,b}(k) &= \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} \frac{\lambda^{a-1}e^{-\lambda/b}}{\Gamma(a)b^a} d\lambda \\ &= \frac{\Gamma(a+k)}{k!\Gamma(a)} \left(\frac{1}{1+b}\right)^a \left(\frac{b}{1+b}\right)^k, \quad k = 0, 1, 2, \dots, \end{aligned}$$

此即為一個負二項分布。Fisher et al. 證明了當 $a \rightarrow 0$ 時

$$P(\text{任一物種出現}k \text{ 次}|\text{該物種出現}) \rightarrow \frac{\alpha\theta^k}{k}, \quad k = 1, 2, \dots,$$

此地 $\alpha = -1/\log[b/(1-b)]$, $\theta = b/(1+b)$ 。以上序列即為對數級數 (log series), Fisher et al. 認為此地 α 可以用來表示一個群落之多樣性的特徵值, 此即為 Fisher “alpha” 的由來。

在 Fisher et al. (1943) 的導證中若伽偶分布中之參數 $a = 1$, 則任一物種出現 k 次的機率 $P_{a,b}(k)$ 成爲一個幾何序列, 此即為 MacArthur (1957) 之 broken-stick 模型。此模型相當於 $\left(\frac{\lambda_1}{\sum_{i=1}^S \lambda_i}, \frac{\lambda_2}{\sum_{i=1}^S \lambda_i}, \dots, \frac{\lambda_S}{\sum_{i=1}^S \lambda_i}\right)$ 爲一 Dirichlet $(1, 1, \dots, 1)$ 之分布。當我們從 $[0, 1]$ 的長度中任選 $S-1$ 個點而形成 S 線段, 此 S 線段即為 Dirichlet $(1, 1, \dots, 1)$ 分布。此如同一個單位長棍子, 若隨機斷裂成 S 片段, 則每一段的長度可以用來作為 S 物種豐富度之模型, 此即為 broken-stick 模型之理論意義。

除了 Fisher et al. (1943) 對物種豐富度所採用之伽偶模型, 及其特例指數模型外, 另外 Preston (1948) 提出對數常態模型也經常使用, 後來學者提出反高斯模型 (inverse Gaussian) 及一般化的反高斯模型 (generalized inverse Gaussian) 等。在參數模型中以某一參數或參數之函數做為量度多樣性的指標。

基於統計抽樣的參數模型開啓了生物多樣性統計理論研究。然而, 參數模型如同曲線配適法有其限制, 除有時很難找到配適良好的參數模型外, 不同的參數模型可能都配適很好, 但其物種數或多樣性的估計卻相差甚遠, 配適完美的參數模型並不一定得到好的物種數或多樣性估計值。另一最重要的限制是若兩群落配適不同的參數模型, 參數無法有統一的意義因而解釋亦不同, 將無法客觀比較群落多樣性, 此為本文專注在無參數的多樣性指標的主要原因。

3.3 無參數法 (non-parametric approach)

無參數法不須對豐富程度架構一個參數模型, 基本上利用物種的豐富程度來定義多樣性指標。假定一群落有 S 個物種, 任取一個個體有 p_i 之機會為第 i 個物種, 則無母數法直接利用 (p_1, p_2, \dots, p_S) 來定義多樣性指標。在早期此種指標最廣為使用的即為 Shannon (1948) 之熵指標, 其形式如下

$$H = - \sum_{i=1}^S p_i \log(p_i),$$

此原為熱力學中量度系統的不確定性指標, 應用於多樣性的意義則為任取一個體其所屬物種之不確定性。假想一個地區幾乎全部都是麻雀, 則任抽一個體, 我們可以知道有很大的機會是麻雀; 若一地區種類愈繁多, 我們就很難確定任抽一個個體的種類。因此量度不確定性的熵早期便廣泛使用以量度生物多樣性。

另一廣為使用者為 Simpson (1949) 年之 Simpson 指標, 其形式為

$$\lambda = \sum_{i=1}^S p_i^2,$$

以抽樣概念出發的 Simpson 指標, 其直觀解釋為「以隨機抽樣的方式從群落中取出兩個個體, 而此兩個個體為同一個物種的機率」, 此機率越高, 則多樣性越低。由於 Simpson 指標和多樣性成反比, 故一般採用與多樣性成正比之 Gini-Simpson 指標:

$$1 - \lambda = 1 - \sum_{i=1}^S p_i^2 = \sum_{i=1}^S p_i(1 - p_i)。$$

這兩種指標雖然在其它學門有其使用價值，但對量度生物多樣性而言，它們卻不滿足多樣性指標最基本的要求，即倍增原則或性質 (replication principle or doubling property)，如下所述。

倍增原則：多樣性指標的基本要求

生物多樣性指標必須具備倍增性質：Hill (1973) 及 Jost (2006, 2007) 認為當兩個具有相同物種豐富度結構的群落 (p_1, p_2, \dots, p_S) 相同，且完全互異 (completely distinct 即沒有共同種)，將兩個群落以等權數合併，則混合群落的多樣性即為原本個別群落的多樣性的兩倍。此倍增性可推廣到多個群落，即當具有相同物種豐富度結構的 N 個完全互異群落以等權數合併，混合群落的多樣性會是單一群落多樣性的 N 倍。

熵指標與 Gini-Simpson 指標都不滿足此基本的要求，例如在兩個地區各有 4 個豐富度均等的物種，且互相沒有共同種，則各自的熵為 $\log 4 = 1.39$ ，各自的 Gini-Simpson 指標為 $1 - 4(1/4)^2 = 0.75$ 。但當二個地區以等權數混合後，則此混合地區的熵為 $\log 8 = 2.08$ ，Gini-Simpson 指標為 0.875。可以很清楚說明物種數滿足倍增性，但熵及 Gini-Simpson 並不滿足。MacArthur (1965) 注意到若採用熵的指數則各自熵的指數為 4，混合地區為 8，即滿足倍增原則。而 Gini-Simpson 指標必須轉換成 Simpson 指標的倒數。在上例中各自 Gini-Simpson 倒數為 4，混合地區 Gini-Simpson 倒數為 8，亦滿足倍增原則。當我們使用熵的指數或 Simpson 的倒數時，它們的表現即如一般物種數的表現會滿足倍增原則。當群落的物種豐富度不同時，若其熵為 2.4，則 MacArthur 稱熵的指數 $\exp(2.4) = 11$ 為 11 個“有效物種” (effective number of species) 或“等價物種數” (number of equivalent species)，亦即此群落和另一個相等豐富度之群落 (有 11 個相同豐富度物種) 有相同的多樣性。又若一個群落其 Gini-Simpson 指標為 0.99，則 Simpson 倒數 $1/(1 - 0.99) = 100$ 之意義為此群落之多樣性和另一個群落 (有 100 個相同豐富度物種) 有相同的多樣性。此“有效物種數”即為 100。

Hill 指標族

Hill (1973) 將物種數、熵的指數及 Simpson 指標的倒數統合在一個多樣性指標族內。Hill (1973) 指標族的直觀作法如下：例如若物種平均豐富度為 $1/5$ ，則大約有 5 個物種，平均豐富度為 $1/200$ ，則大約有 200 個物種。所以 Hill (1973) 首先根據物種相對豐富度 (p_1, p_2, \dots, p_S) 來定義“平均豐富度”，再定義生物多樣性為“平均豐富度的倒數”。但物種相對豐富度 (p_1, p_2, \dots, p_S) 廣義的“加權平均”概念可以用下列表示 (加權數為 $\{w_1, w_2, \dots, w_S\}$ ，此地 w_i 代表對第 i 物

種的加權數)

$$h^{-1}\left(\sum_{i=1}^S w_i h(p_i)\right),$$

其中 h^{-1} 為函數 h 之反函數。若 $h(x) = h^{-1}(x) = x$ 則為一般的加權算術平均; 若 $h(x) = h^{-1}(x) = 1/x$ 則為一般的加權調和平均; 若 $h(x) = \log x$, $h^{-1}(x) = e^x$ 則為一般的加權幾何平均。

Hill 對第 i 個物種 (豐富度 p_i) 採用權數即為 $w_i = p_i$, 且 $h(x) = x^{q-1}$, $h^{-1}(x) = x^{1-q}$, 則平均豐富度為

$$h^{-1}\left(\sum_{i=1}^S w_i h(p_i)\right) = \left(\sum_{i=1}^S p_i^q\right)^{1-q},$$

Hill 定義位階數 (order) 為 q 之 Hill 多樣性指標, qD , 即階數為 q 之多樣性為此平均豐富度之倒數:

$${}^qD = \left(\sum_{i=1}^S p_i^q\right)^{\frac{1}{1-q}}, \quad q \in R,$$

由於 q 等於 1 時此函數不存在, 當 q 趨近於 1 可取極限並定義 1D 為

$${}^1D = \lim_{q \rightarrow 1} {}^qD = \exp\left[-\sum_{i=1}^S p_i \log(p_i)\right].$$

Hill 數值指標將常用的多樣性指標做一個統合, 當 $q = 0$, 則 0D 為物種數 S ; 當 $q = 1$ (逼近於 1), 則 1D 為 Shannon 指標之指數 $\exp(H)$; 當 $q = 2$, 2D 為 Simpson 指標之倒數 $1/\lambda$ 。

由物種相對豐富度 (p_1, p_2, \dots, p_S) 廣義的“加權平均”概念可知 $1/{}^2D$ 為相對豐富度之加權算術平均數; $1/{}^1D$ 為相對豐富度之加權幾何平均數; $1/{}^0D$ 為相對豐富度之加權調和平均數。由於調和平均數 \leq 幾何平均數 \leq 算術平均數, 所以 ${}^0D \geq {}^1D \geq {}^2D$ 。在任何相對豐富度結構下, 隨著 q 越大, 則 qD 會越小, 可由 Beck 與 Schögl (1993) 理論證明此遞減性質。我們可以 qD 對 q 作圖, 得到一個多樣性的剖面 (profile), 在圖 1 中針對不同的豐富度結構畫出其剖面圖。若豐富度為均等的群落, 則剖面圖為一直線, 其高度即為 ${}^qD = S$ 。若 (p_1, p_2, \dots, p_S) 變化不大, 則其剖面圖下降緩慢, 若 (p_1, p_2, \dots, p_S) 變化大 (表示有一些物種很豐富, 而又同時有很稀有的物種), 則剖面圖下降快速, 因此多樣性剖面圖可以完全刻畫一個群落的多樣性。豐富度 (p_1, p_2, \dots, p_S) 之變化程度可以其變化係數 (CV, coefficient of variation) $CV = \text{標準差}/\text{平均值}$ 來量度。若 $CV = 0$ 表示豐富度均等, 沒有變化; 若 CV 愈大, 則變化程度愈高。

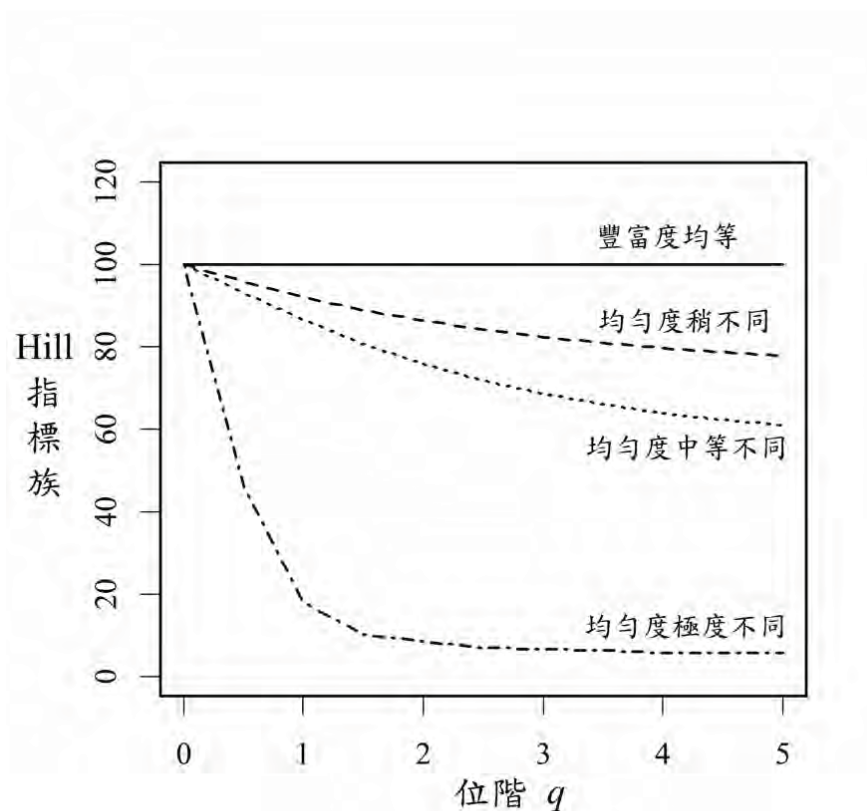


圖 1 (取自 Gotelli and Chao, 2012): 四個不同的豐富度結構的群落之多樣性的剖面 (profile), $0 \leq q \leq 5$ 。每一群落均有 100 物種, 500 個體。自上至下:(1) 豐富度為均等的群落 ($CV = 0$), 則剖面圖為一直線, 其高度即為 ${}^qD = S$; (2) 均勻度稍不同的群落 ($CV = 0.402$), 50 物種各有 7 個個體, 另 50 物種各有 3 個個體, 此結構記為 $\{50 \times 7, 50 \times 3\}$, 剖面下降緩慢; (3) 豐富度結構為 $\{22 \times 10, 28 \times 5, 40 \times 3, 10 \times 2\}$; 均勻度中等不同的群落 ($CV = 0.569$), 剖面下降較前為快; (4) 豐富度結構為 $\{1 \times 120, 1 \times 80, 1 \times 70, 1 \times 50, 3 \times 20, 3 \times 10, 90 \times 1\}$, 均勻度極度不同的群落 ($CV = 3.321$), 剖面下降最快。

Hill 在計算“平均豐富度”中採用 $w_i = p_i$ ，因此從另一角度來看， $\sum_{i=1}^S w_i h(p_i) = \sum_{i=1}^S p_i h(p_i)$ 此式可以視為對第 i 個物種（豐富度 p_i ）中每一個體給予權數 $h(p_i)$ 。當 $q > 1$ 時， $h(x) = x^{q-1}$ 是 x 遞增函數，愈大相對豐富度物種有愈高之權數，故 $q > 1$ 時，此平均豐富度及其 Hill 指標偏重豐富種。而當 $q < 1$ 時則偏重稀有種。當 $q \rightarrow 1$ 時，物種中每一個體其權數相同，不偏重豐富種或稀有種。當 q 越大，則豐富物種將越具有支配性，也就是說當 $q \rightarrow \infty$ ， ${}^\infty D$ 即為最豐富種之相對豐富度的倒數；相反地當 $q \rightarrow -\infty$ ， ${}^{-\infty} D$ 即義為最稀有種之相對豐富度的倒數。雖然 Hill 指標族可以定義在任一實數 q ，但對 $q < 0$ ，由於偏重稀有種，在統計上很難由樣本資料來估計，故一般只考慮 $q \geq 0$ 之 Hill 指標族。

在數學上我們很容易證明對所有的 q ，Hill 指數族均滿足 Hill (1973) 以及 Jost (2006, 2007) 之倍增原則。在 Chao et al. (2010) 更證明了 Hill 指數族更一般化的倍增原則。亦即 N 個群落並不一定要有相同的豐富度結構，只需有相同的 Hill 指標，則倍增原則仍然成立。Hill 指標族的優點除了滿足倍增原則外， ${}^q D$ 對所有 q 均可解釋為“有效物種數”，它們的單位均為“有效物種數”，因此可以互相比較。另外 Hill 指標族亦滿足 Patil and Taillie (1992) 與 Weikard et al. (2006) 之其它對多樣性指標的基本要求：

- (1) Hill's number 對任何 $q > 0$ ，都具有蕭爾凹函數 (Schur-concavity) 的性質 (Tong, 1972; Marshall and Olkin, 1979)。
- (2) “弱單調性” (weak monotonicity) 性質：當群落增加一新物種且是群落最稀有物種，此時群落的生物多樣性應該增加。此性質由蕭爾凹函數即可得。早期的“單調性” (monotonicity) 性質要求當群落增加一新物種，群落的生物多樣性應該增加。然而由於之前討論的指標都只強調物種數或視所有物種的豐富度都一致，並未考慮物種豐富度的差異性，因此 Hill 指標族除物種數外不會滿足此單調性。例如熵指標當群落增加極度豐富新物種時，物種不確定性會減少，故不滿足單調性。
- (3) “轉移原則” (principle of transfer)：若一個較豐富的物種轉移個體至較稀有的物種使得均勻度增加時，群落多樣性應增加。此為在經濟學上 (Dalton, 1920) 的“Robin Hood principle”。意即轉移富人財產到窮人使得收入較為平均時的理論。

Hill 指標族的另一優點是：一些常用的指標可以經由簡單的轉換至 Hill 指標族。例如常被採用的 Rényi (1961) 熵指標族及 Tsallis (1988) 熵指標族均可以經由簡單的轉換至 Hill 指標族。見表一所列。

表 1 常見指標與 Hill 數值指標的轉換關係

多樣性指標	指標形式	轉換形式	Hill 數值指標
物種數	$S = \sum_{i=1}^S p_i^0$	S	$q=0 D$
Shannon 熵指標	$H = -\sum_{i=1}^S p_i \log(p_i)$	$\exp(H)$	$q=1 D$
Simpson 指標	$\lambda = \sum_{i=1}^S p_i^2$	$1/\lambda$	$q=2 D$
Gini-Simpson 指標	$1 - \lambda = 1 - \sum_{i=1}^S p_i^2$	$1/(1 - \lambda)$	$q=2 D$
Rényi 指標族	$H_q = \log \left(\sum_{i=1}^S p_i^q \right) / (1 - q)$	$\exp(H_q)$	$q D$
Tsallis 指標族	$T_q = \left(\sum_{i=1}^S p_i^q - 1 \right) / (1 - q)$	$((1 - q)T_q + 1)^{1/(1-q)}$	$q D$

4. 單一群落系統演化多樣性指標

在第三節中，所談論的指標都只考慮群落的物種數和其豐富度，沒有考慮物種間的差異，即假設物種間的差異都是一樣的，此類指標雖然已具備完善的理論基礎且獲得廣泛的應用。但仍有生物學家提出這樣的疑慮：假設有兩群落各有三個物種，物種的豐富度都一致，其中一群落的物種是：老鷹、喜鵲、濱鵲，另一群落的物種為：喜鵲、樹鵲、藍鵲。傳統的所有指標皆會認定兩群落的多樣性一樣。但直觀上，卻會認為群落一的生物多樣性較高，因為其物種的差異較大，在生態系統上具有不同的生態區位或不同的生態功能。此類似觀點，最早是由 Pielou (1975) 提出，認為一群落的物種若涵蓋較多的屬別或科別，其多樣性應該大於只涵蓋較少屬別或科別的群落。因此生物學家普遍認為，多樣性指標若能考慮物種間的差異，應該更能反應群落生物多樣性的真義。

4.1 系統演化多樣性指標簡介

近幾年考慮物種差異的相關指標如雨後春筍地被提出，而根據物種系統演化樹 (phylogenetic tree) 資訊的系統演化多樣性 (phylogenetic diversity) 指標是最被廣為討論的指標，其中包含

- (1) PD (phylogenetic diversity; Faith, 1992): $PD = \sum_i L_i$, 其中 L_i 為系統演化樹中，第 i 支脈的長度。因此， PD 即群落中物種親緣關係樹所有支脈長度的總和。當以傳統林奈分類來取代物種親緣關係樹時，且定義各鄰近階層的距離都為 1 時 (即，同屬不同種的物種距離為一，同科不同屬的物種其差異為二，以此類推)，則 PD 等同於演化分類樹中各節點的總數，此多樣性的衡量方法，是由 Vane-Wright et al. (1991) 首先提出。而後和 PD 相關的研究 (Crozier, 1992, 1997; Warwick and Clarke, 1995) 相繼被發

表，因為此指標滿足“倍增性質”(doubling property)，具有生態直觀的解釋意義，因此成為最常用的指標之一。

(2) 多樣性函數 (diversity function; Weitzman, 1992, 1993, 1998)

此指標是經濟學上透過兩兩物種的距離矩陣，根據其遞迴演算條件得到，因此並沒有一個固定的數學形式。此指標的特性是不限制物種兩兩的距離需滿足何種條件，即使無法得到親緣關係樹，仍可透過其遞迴演算得到其多樣性指標值。d'Arnoldi et al. (1998) 指出，對於任何距離矩陣，透過此遞迴演算，可以得到一具有超度量 (ultrametric) 性質的距離矩陣。同時，當距離矩陣滿足超度量的性質，則此多樣性函數指標值等同於 PD 減去親緣關係樹的高度。

上述兩指標，並未考慮物種豐富度的資訊，也許在實際保育應用上已經足夠，或在一些資料收集、處理上有其實用性，如我們很難去量測微生物或叢聚樹種的族群量。但以統計的角度，大部份建構在物種豐富度的指標，即使小樣本都可以透統計估計的方法得到不錯的估計量。且以生態的角度而言，考慮物種間個體的交互作用和其演化歷史，可以讓我們瞭解群落的演化過程 (Hubbell, 2001)。有鑑於此，同時考慮物種豐富度和物種間差異的指標仍是生物學家探討的重點 (Allen et al., 2009; Cadotte et al., 2009 ; Clarke and Warwick, 1998; Graham and Fine, 2008; Helmus et al., 2007; Hardy and Senterre, 2007; Izsák and Papp, 1995, 2000; Pavoine et al., 2009; Rao, 1982; Ricotta and Szeidl 2006; Solow and Polasky, 1994; Vellend et al. 2010; Webb, 2000; Weikard et al., 2006)，而最常被討論與使用的指標為下列指標。

- (3) QD (quadratic entropy; Rao, 1982): $QD = \sum_{i,j} d_{ij} p_i p_j$ ，其中 d_{ij} 是第 i 和 j 物種的演化距離。因此，指標是根據兩兩物種的距離矩陣和物種相對豐富度來量測多樣性的大小，可以定義為以物種相對豐富度為權重的平均物種差異。且可以滿足直觀抽樣的解釋意義：從群落抽取兩個體的平均遺傳差異。另外，此指標具有幾個特性：(a) 當不同物種的距離 (d_{ij}) 都等於 1 時，則 QD 等同於 Gini-Simpson 指標。(b) 當物種的相對豐富度都一樣 (即 $p_i = 1/S$)，或只考慮物種的出現與否，則 $QD = J$ 指標 (Izsak and Papp, 2000)， J 指標可以解釋為：從群落抽取兩物種的平均遺傳差異。(c) 當考慮傳統分類樹且視各層級的差異都一樣時， QD 等同於各層級 Gini-Simpson 指標的和 (Shmatani 2001)。(d) 對於任何特性的距離矩陣都仍適用。因此， QD 成為系統演化多樣性指標中，最被廣為應用的指標之一，且其他大部分指標 (Barker, 2002; Helmus et al., 2007; Hardy and Senterre, 2007; Izsák and Papp, 1995, 2000; Solow et al., 1993; Solow

and Polasky, 1994; Warwick and Clarke, 1995; Webb, 2000) 都可視為 QD 指標的修正與應用。

- (4) PE (phylogenetic entropy; Allen et al., 2009): $H_p = -\sum_i L_i a_i \log a_i$, 其中 a_i 是演化樹中第 i 支脈的豐富度。此指標是根據群落物種親緣關係樹的支脈長度 (branch length) 和其相對應的豐富度, 來討論各支脈的熵值。因為此指標可視為 Shannon 熵指標的推廣, 因此指標的值不會特別偏重群落中豐富或稀有物種。

- (5) I_q (Pavoine et al., 2009): $I_q = \frac{\sum_i L_i a_i^q - T}{1-q}$, $q \geq 0$,

此指標族是考慮群落物種親緣關係樹為超度量 (ultrametric) 的情況下, 結合 Tsallis (1988) 熵指標族, 提出一系統演化多樣性指標族。其的最大優點是, 指標值不會隨演化樹的根 (root) 選取的差異而變, 且整合上述指標, 當 $q = 0$, I_0 等於 PD (Faith, 1992) 減去物種親緣關係樹的高度。當 $q = 1$, I_1 等於 PE (Allen et al., 2009)。當 $q = 2$, I_2 等於 QD (Rao, 1982), 因此 I_q 具備上述指標的性質。

4.2 結合 Hill 數值指標特性的系統演化多樣性指標

上述指標雖然相較傳統指標涵蓋更多的訊息 (物種差異的訊息), 實際上卻無法應用在生態保育上, 因為這些指標皆不滿足指標最基本的要求倍增性質。因此 Chao et al. (2010) 根據 Hill 數值的有效物種觀念, 結合系統演化樹的資訊, 建構 ${}^q\bar{D}(T)$ (phylogenetic mean diversity; Chao et al., 2010):

$${}^q\bar{D}(T) = \frac{1}{T} \left\{ \sum_{i \in B_T} L_i \left(\frac{a_i}{T} \right)^q \right\}^{1/(1-q)}, \quad q \geq 0,$$

其中, T 是從現今至過去的物種演化時間, T 可以取為 $[0^+, \infty)$ 的任何正數, 一般設 T 為群落物種的共同祖先第一次發生種化的時間點 (即演化樹的根)。另外, B_T 是指演化時間 $[-T, 0]$ 內, 演化樹中所有支脈的集合。 ${}^q\bar{D}(T)$ 可以解釋為“物種差異為 T 的有效物種數”, 也可以解釋為演化時間 $[-T, 0]$ 的“平均有效支脈或物種數” (mean effective number of lineages or species)。此指標具備以下的性質,

- (1) “尺度不變性” (scale invariant)

因為 ${}^q\bar{D}(T)$ 代表在演化時間 $[-T, 0]$ 中群落的平均有效物種數。所以無論時間的單位為何, ${}^q\bar{D}(T)$ 不受其影響。因此, 所謂的“尺度不變性”是指 ${}^q\bar{D}(T)$ 不受親緣系統樹的支脈長度的尺度的單位變化而變。

(2) “倍增性質” (doubling property)

因為考慮物種間差異的訊息，這裡將古典物種多樣性的倍增性質稍做修正，即當兩群落有相同的多樣性且群落沒有共同種，且來自不同群落的物種差異都達最大值 T ，則混合群落的多樣性為原個別群落多樣性的兩倍。另外，此性質並不侷限於兩群落的討論，對於 N 個群落，若各群落滿足上述的條件，則合併群落的 ${}^q\bar{D}(T)$ 為任何群落 ${}^q\bar{D}(T)$ 的 N 倍。

(3) “弱單調性” (weak monotonicity)

對之前討論之“單調性” (monotonicity) 即當一群落增加一新物種，則群落的多樣性應增加 (Izsák and Papp, 2000; Solow et. al., 1993; Solow and Polasky, 1994)。一般生物學家認為系統演化多樣性指標也需滿足此特性。如前所述，之前討論的指標都只強調物種的出現與否或視所有物種的豐富度都一致，並未考慮物種豐富度的差異性，因此常見的 QD 並不滿足此特性 (Izsák and Papp, 2000)。而在考慮 Hill 數值指標族時，也只有 $q = 0$ 時滿足此特性。因此，Weikard et al. (2006) 認為，當考慮物種的豐富度時，“單調性” (monotonicity) 的條件需做修正，即：當群落增加一最稀有新物種時，多樣性指標應增加，稱之為“弱單調性” (weak monotonicity)。如此，對於任何 $q > 0$ ，Hill 數值指標都滿足此性質，原因在於 Hill's number 對任何 $q > 0$ ，都具有蕭爾凹函數 (Schur-concavity) 的性質 (Tong, 1972; Marshall and Olkin, 1979)。當考慮物種的差異性時， ${}^q\bar{D}(T)$ 滿足“弱單調性”的條件為，當群落增加一最稀有新物種，且與群落其他物種的差異最大時，系統演化多樣性應增加。

(4) 當物種的差異是根據傳統林奈分類時，可將各層級的差異定義為 1，如相同屬的種間差異為 1；相同科不同屬的種間差異為 2，以此類推。因此，傳統分類多樣性可定義為

$${}^q\bar{D}(L) = \frac{1}{L} \left\{ \sum_{i \in B_L} \left(\frac{a_i}{L} \right)^q \right\}^{1/(1-q)}, \quad q \geq 0,$$

其中 B_L 為分類層級為 L 的分類樹，其所有支脈的集合。此指標可解釋為“分類差異為 L 的有效物種數”。

(5) 可以透過簡單的轉換和其他的指標作連結

(a) 與 Hill 數值指標的關係

當 $T = 0^+$ 或兩兩物種間的差異皆為 T 時， ${}^q\bar{D}(T) = \text{Hill's number}$ ，因此也可以將 ${}^q\bar{D}(T)$ 視為 Hill 數值指標，考慮物種間差異資訊的推廣指標。

(b) 與 Faith's PD 的關係

當 $q = 0$ 時, ${}^0PD(T) = PD$ (Faith, 1992), ${}^0D(T) = PD/T$ 。因此, ${}^qPD(T)$ 和 ${}^q\bar{D}(T)$ 可以視為 Faith (1992) 的 PD 考慮物種豐富度訊息後的推廣。

(c) 與 Allen et al.'s HP 的關係

當 q 趨近於 1 時, ${}^1\bar{D}(T) \exp \left[- \sum_{i \in B_T} \frac{L_i}{T} a_i \log a_i \right]$, 且 T 為群落第一次發生種化的時間點, 則 ${}^1\bar{D}(T) = \exp \{H_p/T\}$ or $\log({}^1\bar{D}(T)) = H_p/T$ 。

(d) 與 Rao's QD 的關係

當 $q = 2$ 時, ${}^2\bar{D}(T) = \left\{ \sum_{i \in B_T} \frac{L_i}{T} a_i^2 \right\}^{-1}$, 且 T 為群落物種第一次發生種化的時間點, 則 ${}^2\bar{D}(T) = \frac{T}{T-QD} = \frac{1}{1-QD/T}$, 其中 QD 中物種間的距離, 在群落親緣關係樹的定義為兩物種間支脈長度的平均。

(e) 當 q 趨近無窮大時, ${}^\infty\bar{D}(T) = \frac{1}{\max\{a_1, a_2, \dots, a_{|B|}\}}$, 其中 T 為任意正數。

因此, 透過上述的關係, 可以將常見的多樣性指標, 轉換成 ${}^q\bar{D}(T)$, 使其具有直觀解釋意義的指標單位, 並且具備“倍增性質”。以下我們舉例說明“倍增性質”的重要性。

假設一個地區有 1000 個物種, 其物種間遺傳差異都是 7 且物種豐富度都一樣。表示每個物種對該地區多樣性的貢獻應該都是一樣的。假設有一開發案, 評估可能會使該地區 700 個物種消失, 則這個開發案對當地生物多樣性的影響為何? 若以 Rao's QD 作為生物多樣性指標, 則 1000 物種的 QD 值是 6.993, 而 300 個物種的 QD 值為 6.977。因此, 此數據可以得到 $6.977/6.993 = 0.998$ 。表示: 此開發案存活下來的物種, 其多樣性幾乎和原地區的多樣性一樣, 所以得到的結論是: 許可開發案進行。但另一派的觀點為, 因開發案導致 700 物種消失的 QD 值為 6.99, 因此, 得到消失物種的多樣性佔原地區多樣性的比率為 $6.99/6.993 = 0.999$, 表示開發案幾乎斷送該地區全部的多樣性, 所以得到的結論是: 不准開發案通過。為什麼同樣一件事, 卻有兩個極端的結果? 這是因為 QD 不滿足“倍增性質”, 造成消失 700 物種的多樣性加存活下來 300 物種的多樣性不會等同於原 1000 物種的多樣性。當 QD 轉換成 ${}^q\bar{D}(T)$, 則原 1000 物種的多樣性為 $1 / (1 - 6.993/7) = 1000$ “平均有效物種”; 因開發案而消失的 700 物種, 其多樣性為 $1 / (1 - 6.99/7) = 700$ “平均有效物種”, 而存活下來的 300 物種的多樣性為 $1 / (1 - 6.977/7) = 300$ “平均有效物種”。因此消失的多樣性佔原多樣性的 $700/1000 = 0.7$, 存活下來的多樣性是原多樣性的 $300/1000 = 0.3$ 。這樣的結論不會得到上述的謬論, 且符合我們直觀的想法, 這是因為 ${}^q\bar{D}(T)$ 滿足“倍增性質”。相同的問題一樣會出現在 Allen et al. 的 H_p 指標上, 但經過轉換 ${}^1\bar{D}(T) = \exp \{H_p/T\}$, 即可得到合理且具直觀的結果。

因此，我們認為無論古典指標或系統演化多樣性指標，“倍增性質”是多樣性指標最基本的要求，而 Chao et al. (2010) 所建構的 ${}^q\bar{D}(T)$ 指標，正可以解決以往指標在解釋推論的缺點。

因為系統演化多樣性是根據傳統 Hill 數值指標和群落內各物種的演化樹資訊所建構的指標。建議應該根據不同位階 q 和不同演化時間 T 所得到的剖面 (profile) 圖來進行分析，如此才能客觀、完整地描述整個群落不同演化歷史下的多樣性。例如，Chao et al. (2010) 分析人造林中，人工疏伐和未疏伐兩地區植被的系統演化多樣性，作者根據兩植被群落的豐富資料和系統演化樹資訊 (Shimatani, 2001)，如圖 2-1 所示。分別依不同的位階 $q = 0, 1, 2$ ，分析描述兩種類型的植被群落在演化時間 $[-T, 0]$ 的系統演化多樣性的關係，如圖 2-2 所示。資料分析顯示，當只考慮物種而不考慮其豐富度時 (位階 $q = 0$)，疏伐的植被群落在演化時間的任意時段中，都有較高的平均系統演化多樣性。而當考慮物種的豐富度時，無論強調豐富物種 (位階 $q = 2$) 的多樣性或視各物種等權重 (位階 $q = 1$) 的多樣性，未疏伐的植被群落在所有討論的演化歷史中，幾乎都有較高的平均系統演化多樣性。

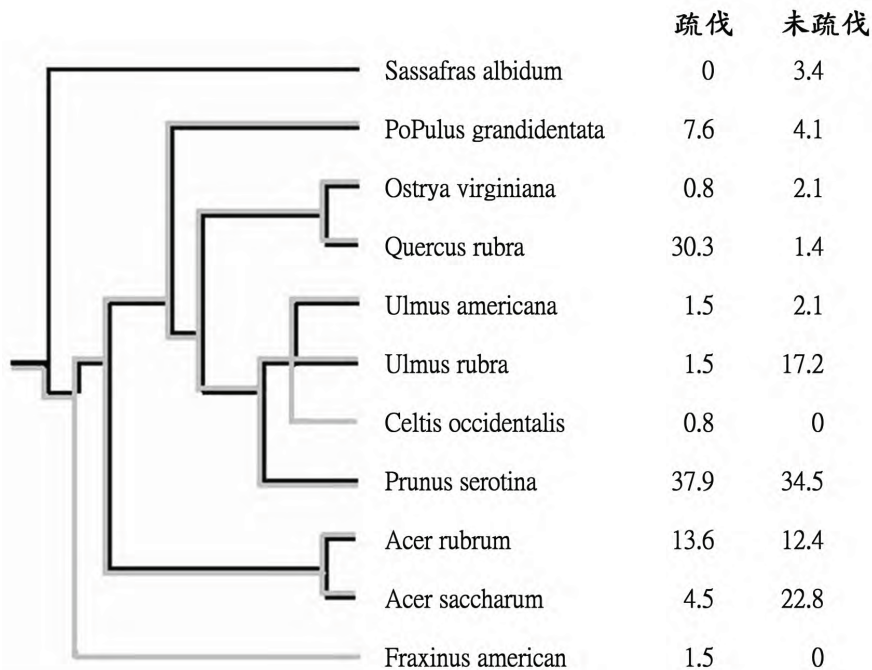


圖 2-1 (翻譯自 Chao et al. 2010): 疏伐與未疏伐植被族群中各物種的相對豐富度及各物種的系統演化樹

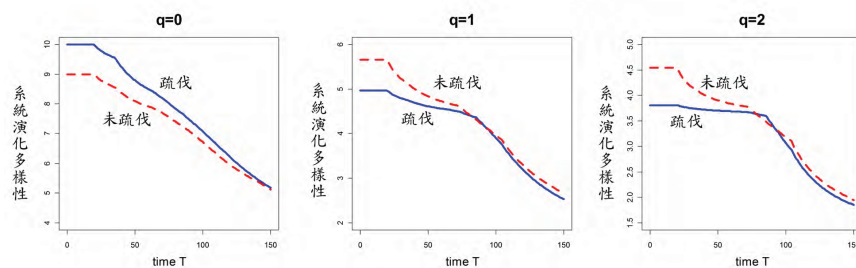


圖 2-2 (翻譯自 Chao et al. 2010): 兩種植被族群在演化時間為 $[-T, 0]$ 的平均系統演化多樣性

5. 多群落的生物多樣性

當討論多個群落的多樣性時，無論保育學家或生態學家所研究與探討的是：在不同地方（空間）或不同時間下，不同群落物種組成的差異有多大？進而探討是什麼因素造成其差異？因此如何正確量化群落間的差異性（differentiation）、轉變性（turnover），相似性（similarity）或重複性（overlap），變成此議題的重點。

關於多群落生物多樣性指標的發展已有一段歷史，Whittaker (1960, 1972, 1977) 以空間的概念將多群落的生物多樣性指標分為三個層次，即所謂的 α 、 β 與 γ 多樣性指標。其中 α 多樣性指標為量測單一群落的平均變異性； β 多樣性指標量測兩個以上不同群落的差異； γ 多樣性指標量測整體區域的變異性。

前人在探討多群落多樣性的議題時，大致可以分為兩個方向，其一是直接討論群落與群落間的相似或是分化程度，經由不同的解釋觀點或是物種優勢程度的偏好陸續定義出不同的群落相似或是分化指標。其二是 α 、 β 與 γ 多樣性的分解，依據三種指標的解釋意義，Whittaker 認為 γ 多樣性必須完全被 α 與 β 多樣性完全決定，因而各種 α 、 β 與 γ 多樣性指標的分解概念分別被提出。近年來興起將群落相似或分化性質以及多樣性分解這兩種方式整合的概念，Jost (2006, 2007) 認為所謂的相似或分化指標可以經由 β 多樣性的線性或是非線性的轉換而得。本節首先介紹傳統上生態學家常用的相似度指標，接著再引入多群落的多樣性分解，介紹 Jost 的 α 、 β 與 γ 多樣性指標，並說明 Jost 的 β 多樣性指標可以經由幾種簡單的轉換得到相似與分化指標族，而目前生態學家常用的指標恰屬於指標族中的特例。

5.1 相似度指標

文獻上對於 β 多樣性的討論可由兩群落的物種相似度 (similarity) 談起。群落的相似度與 β 多樣性乃是一相對的概念 (MacArthur, 1965), 一個生態系統中的多樣性高則其相似度低, 反之亦然。在比較兩個群落物種結構的相似程度時, 最基本的相似度指標是由群落一與群落二的出現的物種數 (分別記為 S_1 與 S_2), 以及兩群落的共同物種數 (記為 S_{12}) 來進行定義, 即所謂沒有考慮物種豐富度的出現型相似度指標 (incidence-based similarity index)。文獻上, 使用超過百年的出現型指標為 Jaccard (1908) 指標, 定義為 $I_J = S_{12}/(S_1 + S_2 - S_{12})$ 。因為 $S = S_1 + S_2 - S_{12}$ 為兩群落的混和之後的總物種數, 所以直觀上 Jaccard 指標可以解釋為: 兩群落共同物種數與總物種數之比值, 表示共同種在混和群落所佔的比例。另一種被大家熟知的出現型指標為 Sørensen (1948) 指標, 表示為 $I_S = S_{12}/[(S_1 + S_2)/2]$ 。直觀而言, 該指標表示: 兩群落共同物種數與兩群落的平均物種數之比值, 即共同種在平均群落所佔的比例。

一般而言, 群落中的物種豐富度並非完全一致, 有些物種可能較常見、有些物種可能較稀有, 僅使用出現型指標來判別兩群落的相似程度會有過於偏好稀有物種的情況。因此, 考量物種豐富度的相似指標 (abundance-based similarity index) 隨之而生。令 $(p_{11}, p_{21}, \dots, p_{S_1})$ 與 $(p_{12}, p_{22}, \dots, p_{S_2})$ 分別表示群落一與二的物種相對豐富度 Morisita (1959) 定義了 Morisita-Horn 相似指標為

$$C_M = \frac{\sum_{i=1}^S p_{i1} p_{i2}}{\frac{1}{2} \left(\sum_{i=1}^S p_{i1}^2 + \sum_{i=1}^S p_{i2}^2 \right)},$$

其中分子部分可以解釋為: 在兩個群落中分別隨機抽取一個個體為相同物種的機率; 而分母為: 在個別群落中隨機抽取兩個個體為相同物種機率之平均。Wolda (1983) 以及 Magurran (2004) 進一步指出, Morisita-Horn 指標乃是架構於 Simpson 指標之下, 因此會受到群落中優勢物種的影響。當優勢種為特有種時, 以 Morisita-Horn 指標來衡量相似度值會很低; 反之, 當優勢種為共同種時, 相似度會很高。換句話說, 若 Morisita-Horn 指標很高, 除了兩群落的物種分布幾乎一致之外, 也有可能是因為群落裡有少數共同種為優勢種, 或是許多特有種為稀有種。

Horn (1966) 將 Shannon 指標應用於 Morisita-Horn 指標, 提出 Horn 相似指標

$$C_H = \frac{\left(-\sum_{i=1}^S \bar{p}_i \log \bar{p}_i \right) - \left(-\frac{1}{2} \sum_{j=1}^2 \sum_{i=1}^S p_{ij} \log p_{ij} \right)}{\log 2},$$

其中 $\bar{p}_i = \sum_{j=1}^2 p_{ij}/2$ 。此指標可解釋為: 兩個群落物種相對豐富度結構的重複性 (overlap), 所謂重複性係指共同種佔平均群落物種之比例, 重複性越高即相似性越高。Horn 指標具有 Shan-

non 指標的性質，即同時考量優勢種與稀有種的影響，相對於 Sørensen 出現指標與 Morisita-Horn 指標，Horn 指標不強調群落間稀有或優勢物種的相似程度。

上述的相似度指標僅定義在兩個群落的比較上，然而現今生態統計所遇到的資料往往是多個群落，假若僅侷限於群落間兩兩比較，將喪失多群落共同物種的訊息。Chao et al. 在 2008 年的文章利用兩階段抽樣機率的概念，將 Morisita-Horn 相似指標推廣到多個群落 C_{qN} 指標族

$$C_{qN} = \frac{\frac{1}{N^q - N} \left(\sum_{i=1}^S \left(\sum_{j=1}^N p_{ij} \right)^q - \sum_{i=1}^S \sum_{j=1}^N p_{ij}^q \right)}{\frac{1}{N} \sum_{i=1}^S \sum_{j=1}^N p_{ij}^q}$$

該指標族具有重複性的解釋意義，即共同種佔平均群落物種數比例，其中足標 N 表示群落數、 q 表示對於相對豐富度的偏好程度越高越偏好豐富種，當 $N = 2$ 且 $q = 0, 1, 2$ 時可以分別對應到 Sørensen、Horn 以及 Morisita 指標。Chao et al. 進一步闡述，僅透過單一相似性指標來描述群落的相似程度，是極不客觀的。因為任何一種相似指標僅是描述群落間相似性的某一種面向，客觀的相似性比較應盡量採用指標族來作為衡量標準。

5.2 多群落中的多樣性分解

Whittaker (1960, 1972, 1977) 認為一個地區的生物多樣性可區分為 α 、 β 與 γ 三種，並主張三者應符合如下關係：

$$\gamma = \alpha \times \beta,$$

即整個區域的多樣性 γ 等於各個群落多樣性的平均 α 和整個區域中群落間的多樣性 β 的乘積。此外， γ 多樣性與 α 多樣性只是施用範圍或地域大小的不同，兩者應具有相似的特徵或相同的單位，因此 β 多樣性指標也可詮釋為對整個區域而言，不同群落多樣性的比值或差異。此稱之多樣性的乘法分解。

除了上述的乘法分解之外，Lande (1996) 提出另一種類 ANOVA 分解，將 α 多樣性指標定義為各個群落群落內的多樣性指標之加權平均，並稱之為“組內變異”，在定義 β 多樣性指標為“組間變異”。而 γ 多樣性指標為“組內變異”加“組間變異”，稱為“整體變異”，即

$$\gamma = \alpha + \beta,$$

此稱為多樣性的加法分解。

基於單一群落物種的生物多樣性等同於其有效物種數的概念逐漸被生態學家所接受，許多學者 (Anderson et al., 2011; Baselga, 2010; Chao et al., 2012; Ellison, 2010; Jost, 2006,

2007; Jost et al., 2010; Lande, 1996; Veech and Crist, 2010; Veech and Crist, 2010; Whittaker, 1972) 相繼將其概念投入多群落多樣性的應用之上。其中較為受到重視的多群落多樣性分解方式即為前述的乘法與加法分解，將有效物種數應用至乘法分解，使得 γ 多樣性指標具有整體區域有效物種數的解釋意義， α 多樣性指標解釋為各群落有效物種數的平均，而 β 多樣性指標為有效群落數。又如加法分解， γ 性指標表示整體區域的有效物種數， α 指標表示群落的平均有效物種數， β 指標表示平均而言單看一個群落而忽略的有效物種數。由於乘法與加法多樣性的分解各具解釋意義，因此透過多樣性直觀的公設來比較這兩種分解方法亦是相當重要的議題。在這之中，最廣為學者們所討論的概念是 α 與 β 指標的獨立性質 (Wilson and Shmida, 1984; Jost et al. 2010)。Chao et al. (2012) 綜合前人的討論，透過理論證明及數值模擬的方式指出，在給定物種數及群落數為有限整數之下，加法分解 α 與 β 指標具有強烈的相關性，假設物種數大於群落數則乘法分解的 α 與 β 指標具有統計上的獨立性質。

5.3 Jost 的多樣性指標

Jost (2007) 集前人對於生物多樣性研究之所成，提出 α 、 β 與 γ 多樣性的五個重要的公設，在此一一對其說明。

(1) α 、 β 彼此獨立

在多個群落的比較中，Wilson and Shmida (1984) 認為一生態系統，不同單位的測量必須彼此獨立由於 α 與 β 多樣性對於一個區域所詮釋的意義並不相同，因此 α 多樣性的大小不應影響 β 多樣性的大小，反之亦然。這種性質亦當反應在對應的 α 與 β 多樣性指標上面，倘若多樣性指標違反這個公設，將被認為是不適當的。

(2) 不同區域的多樣性指標可以相互比較

無論討論的是 α 、 β 或是 γ 多樣性，不同區域所計算出的多樣性指標必須都能相互比較，且相同的多樣性指標值應描述相同程度的多樣性。

(3) α 多樣性為群落間多樣性的平均值

α 多樣性乃是組成整個區域各個群落的多樣性之某種形式上的平均。假若每一個群落內的多樣性指標一樣時，整體區域的 α 多樣性指標仍維持不變。

(4) γ 必須要完全被 α 、 β 所決定

在探討一個區域中，多個群落間的多樣性時， α 、 β 與 γ 多樣性必然存在某種形式的關係，譬如前一節所提到的乘法性分解與加法性分解即是一例，然而在此公設中，並不對此

關係做任何限制。

(5) α 多樣性不能比 γ 多樣性大

這是一個較為直觀的公設，一般而言 α 多樣性係指在一個群落內，各個物種之間的多樣性，而 γ 多樣性乃是指大範圍整個區域之中，各個物種間的多樣性，因為 γ 多樣性所討論的範圍更大，多樣性的訊息不應該比小範圍的 α 多樣性小。因此一個適當的多樣性指標應符合 γ 多樣性指標大於等於 α 多樣性指標的性質。

Jost (2007) 經由乘法分解及其提出的五項公設條件之下，將 Hill 指標族推廣而得的 ${}^qD_\gamma$ 、 ${}^qD_\alpha$ 及 ${}^qD_\beta$ 為目前生態學家相當重視的物種多樣性指標族，其中 p_{ij} 表示群落 j 中物種 i 的相對豐富度， $j = 1, 2, \dots, N$; $i = 1, 2, \dots, S$ 。Jost 的指標族整理如下：

$$\begin{aligned} {}^qD_\gamma &= \left[\sum_{i=1}^S \left(\frac{p_{i1} + p_{i2} + \dots + p_{iN}}{N} \right)^q \right]^{\frac{1}{1-q}} \\ {}^qD_\alpha &= \left[\frac{\sum_{i=1}^S p_{i1}^q + \sum_{i=1}^S p_{i2}^q + \dots + \sum_{i=1}^S p_{iN}^q}{N} \right]^{\frac{1}{1-q}} \\ {}^qD_\beta &= {}^qD_\gamma / {}^qD_\alpha \end{aligned}$$

根據 β 多樣性與群落相似度呈負相關，且與群落差異呈正相關，經由標準化後得到 Jost 的群落轉換互補、群落均質性以及平均群落重複性指標族。Jost et al. 在 2010 年的文章中更指出平均群落重複性指標族等同 Chao et al. (2008) 所提出的 C_{qN} 指標族。另一方面，Chiu et al. (2013) 亦提出「群落轉換互補」指標族。上述四種指標族在階次 $q = 0, 1, 2$ 時恰為目前生態學家廣為採用的指標，期結果整理於表 2 中。

表 2 多群落相似性指標族及其特例

	轉換互補指標族	均質性指標	平均群落重複性指標	區域重複性指標
	$\frac{N - {}^qD_\beta}{N-1}$	$\frac{1/{}^qD_\beta - 1/N}{1-1/N}$	$\frac{(1/{}^qD_\beta)^{q-1} - (1/N)^{q-1}}{1-(1/N)^{q-1}}$	$\frac{(1/{}^qD_\beta)^{1-q} - (1/N)^{1-q}}{1-(1/N)^{1-q}}$
	(Jost, 2007)	(Jost, 2007)	(Jost 2007; Chao et al., 2008)	(Chiu et al., 2013)
$q = 0$	Sørensen	Jaccard	Sørensen	Jaccard
$q = 1$			Horn overlap	Horn overlap
$q = 2$	Regional-overlap	Morisita-Horn	Morisita-Horn	Regional-overlap

6. 多群落的系統演化生物多樣性

誠如第五章節所述, 討論多個群落的多樣性時, 主要是探討在不同地方 (空間) 或不同時間下, 不同群落間物種組成的差異有多大? 然而, 關於考慮物種演化差異的群落間相似、分化指標的文獻幾乎完全闕如。其中, 微生物學家最常採用類似傳統 Jaccard 或 Sorenson 指標的方法 (Lozupone and Knight, 2005; Lozupone et al., 2007; Hamady et al. 2009; Faith and Baker, 2006; Bryant et al., 2008; Bacaro et al., 2007), 根據物種演化樹的資訊, 將兩群落物種共同的總支脈長度佔全部總支脈長度的比例 (系統演化 Jaccard 相似性指標), 或佔平均群落總支脈長度的比例 (系統演化 Sorenson 相似性指標), 來衡量兩群落的相似性, 雖然這類指標已廣泛應用至微生物學的研究, 但這些指標並未考量物種的豐富度, 因此不足以完整地描述物種多樣性的架構。所而生態學家最常使用的指標是透過類似 ANOVA 的分解, 將群落間的變異佔全部變異的比例, 視為分化指標。因此, Nei (1982) 根據 QD (Rao, 1982) 提出的 N_{ST} 指標,

$$N_{ST} = \frac{QD_{\gamma} - QD_{\alpha}}{QD_{\gamma}},$$

其中 QD_{γ} 為合併群落的 QD , QD_{α} 為平均群落的 QD 。

N_{ST} 指標可以解釋為: 在合併群落內, 平均物種的遺傳差異, 相較於相同群落內, 所增加的比率。解釋上, 此指標可以被使用在群落間分化的量測上, 因此, N_{ST} 指標最先是被應用在群落遺傳的分化研究上。但此分化指標值會受平均群落的 QD 而影響, 因此在實際生態保育應用上, 可能會出現錯誤的結論 (Hardy and Jost, 2008)。所以, N_{ST} 指標不能作為定量群落間分化 (差異) 的指標。而類似的分解方法, Pavoine et al. (2009) 和 Mouchet and Mouillet (2011) 分別根據 I_q 和 H_p 定義考量物種演化差異的分化指標。然而如同 N_{ST} 分化指標一樣, 這些指標都無法客觀量度群落間的分化 (差異) 程度。

由上一節得知, 傳統群落間的相似指標, 如 Jaccard、Sorenson、Morisita-Horn、Horn 或 C_{qN} 相似指標, 皆可透過多樣性分解下的 β 多樣性 (也就是群落間的差異), 經單調函數的轉換得到。因此, Chiu et al. (2013), 根據 Chao et al. (2010) 建構的系統演化多樣性指標, 在滿足 Jost (2007) 提出的五個公設條件下, 將整個區域的多樣性 (γ 多樣性) 透過乘法分解得到兩個獨立的多樣性組成份, 即: 各個群落多樣性的平均 (α 多樣性) 和整個區域中群落間的

多樣性 (β 多樣性)。因此, 得到系統演化多樣性 γ 、 α 和 β :

$${}^q\bar{D}_\gamma(T) = \frac{1}{T} \left\{ \sum_{i \in B_T} L_i \left(\frac{\alpha_i}{T} \right)^q \right\}^{1/(1-q)},$$

$${}^q\bar{D}_\alpha(T) = \frac{1}{T} \left\{ \frac{\left[\sum_{i \in B_T} L_{1i} \left(\frac{\alpha_{1i}}{T} \right)^q + \sum_{i \in B_T} L_{2i} \left(\frac{\alpha_{2i}}{T} \right)^q + \dots + \sum_{i \in B_T} L_{Ni} \left(\frac{\alpha_{Ni}}{T} \right)^q \right]}{N} \right\}^{1/(1-q)},$$

$${}^q\bar{D}_\beta(T) = \frac{{}^q\bar{D}_\gamma(T)}{{}^q\bar{D}_\alpha(T)},$$

其中 a_{ji} 為第 j 群落中, 第 i 支脈的豐富度。另外,

${}^q\bar{D}_\gamma(T)$ 解釋為: 地區內, 在演化時間 $[-T, 0]$ 的平均有效物種數。

${}^q\bar{D}_\alpha(T)$ 解釋為: 平均群落內, 在演化時間 $[-T, 0]$ 的平均有效物種數。

${}^q\bar{D}_\beta(T)$ 解釋為: 在演化時間 $[-T, 0]$ 的平均相異群落數。

因為 β 多樣性是描述群落間的物種組成的差異, 因此 β 多樣性和群落分化成正比關係, 和群落物種相似度成反比關係, 且和群落數相關。因此, 透過不同單調函數的標準化轉換可得到以下指標,

- (1) 系統演化平均群落重複性指標: $\frac{N^{1-q} - [{}^q\bar{D}_\beta(T)]^{1-q}}{N^{1-q} - 1}$
- (2) 系統演化區域重複性指標: $\frac{[1/{}^q\bar{D}_\beta(T)]^{1-q} - (1/N)^{1-q}}{1 - (1/N)^{1-q}}$
- (3) 系統演化均質性指標: $\frac{1/{}^q\bar{D}_\beta(T) - 1/N}{1 - 1/N}$
- (4) 系統演化轉換的互補指標: $\frac{N - {}^q\bar{D}_\beta(T)}{N - 1}$

另外, 因為 ${}^q\bar{D}_\beta(T)$ 是根據 Hill 數值指標, 加入系統演化資訊的推廣指標。因此, Chiu et al. (2013), 將上述指標可和傳統的相似指標 (表 2) 相對應可得到表 3。

上述的相似性指標和群落平均的多樣性 (α 多樣性) 的大小無關, 因此皆可以衡量群落間物種的相似程度, 而其互補指標可當成分化指標。以下我們透過一簡單的例子說明 N_{ST} (Nei, 1982) 不適用於量測群落間物種組成的差異, 而 Chiu et al. (2013) 所定義的分化指標可以客觀衡量群落間物種組成的差異。

舉例說明, 假設欲分析兩地區內群落間的分化程度, 其中地區 1: 有兩群落分別有 500 個物種, 其物種差異都是 7 且豐富度都一樣, 當兩群落沒有共同種, 因此直觀上會認為兩群落的

表 3 系統演化相似性指標族及其特例。其中 $L_\gamma(T)$ 和 $L_\alpha(T)$ 分別為 PD (Faith, 1992) 的 γ 和 α 多樣性; $H_{p,\gamma}$ 與 $H_{p,\alpha}$ 分別為 H_P (Allen et al., 2009) 的 γ 和 α 多樣性; QD_γ 與 QD_α 分別為 QD (Rao, 1982) 的 γ 和 α 多樣性。

多群落系統演化相似性指標				
	$\frac{N-q\bar{D}_\beta(T)}{N-1}$	$\frac{1/q\bar{D}_\beta(T)-1/N}{2-1/N}$	$\frac{[1/q\bar{D}_\beta(T)]^{q-1}-(1/N)^{q-1}}{1-(1/N)^{q-1}}$	$\frac{[1/q\bar{D}_\beta(T)]^{1-q}-(1/N)^{1-q}}{1-(1/N)^{1-q}}$
$q = 0$	Phylo-Sørensen $\frac{N-L_\gamma(T)/L_\alpha(T)}{N-1}$	Phylo-Jaccard $\frac{L_\alpha(T)/L_\gamma(T)-1/N}{1-1/N}$	Phylo-Sørensen $\frac{N-L_\gamma(T)/L_\alpha(T)}{N-1}$	Phylo-Jaccard $\frac{L_\alpha(T)/L_\gamma(T)-1/N}{1-1/N}$
$q = 1$			Phylo-Horn overlap $1 - \frac{H_{p,\gamma}-H_{p,\alpha}}{T \log N}$	Phylo-Horn overlap $1 - \frac{H_{p,\gamma}-H_{p,\alpha}}{T \log N}$
$q = 2$	Phylo-regional-overlap $1 - \frac{QD_\gamma-QD_\alpha}{(N-1)(T-QD_\gamma)}$	Phylo-Morisita-Horn $1 - \frac{QD_\gamma-QD_\alpha}{(1-1/N)(T-QD_\alpha)}$	Phylo-Morisita-Horn $1 - \frac{QD_\gamma-QD_\alpha}{(1-1/N)(T-QD_\alpha)}$	Phylo-regional-overlap $1 - \frac{QD_\gamma-QD_\alpha}{(N-1)(T-QD_\gamma)}$

分化程度應該達最大。地區 2: 有兩群落分別有 10 個物種, 其物種差異都是 7 且豐富度都一樣, 當兩群落有共同種 9 種, 表示每一群落中, 只有一物種是特有種, 因此直觀上會認為此地區兩群落間的分化程度應該很小。然而地區 1, 假設欲分析兩地區內群落間的分化程度, 其中地區地區: 此地區間地區中, N_{ST} 值為 $(6.993 - 6.986) / 6.993 = 0.001$, 數值顯示兩群落幾乎完全一樣沒有分化。地區 2 的 N_{ST} 值為 $(6.335 - 6.3) / 6.335 = 0.0055$ 。若以 N_{ST} 值為分析群落分化的依據, 則會認為地區 2 的群落分化程度大於地區 1 的群落分化程度, 然而此結論卻和實際情況嚴重違背。這是因為 N_{ST} 和傳統的 G_{ST} 有相同的問題, 即 N_{ST} 值會受 QD_α (平均群落的 QD 值) 影響, 當 QD_α 很大時, 無論群落的分化程度如何, QD_S 都會很接近 QD_γ , 因而得到很小的 N_{ST} 值, 導致 N_{ST} 嚴重低估群落的分化程度 (如地區 1)。Jost (2008) 提出修正傳統 G_{ST} 的指標, 即 $D = G_{ST} / \max(G_{ST})$, 得到 $D = (1 - 1/2\bar{D}_\beta) / (1 - 1/N)$ 。Chiu et al. (2013) 根據 Jost 的概念得到 N_{ST} 的修正式: $N_{ST} / \max(N_{ST})$, 此修正式等同於 $\frac{N}{N-1} \frac{QD_\gamma - QD_\alpha}{T - QD_\alpha}$ (即表 3 中相似指標 Phylo-Morisita-Horn 的互補指標)。因此, 得到地區 1 的 $\frac{N}{N-1} \frac{QD_\gamma - QD_\alpha}{T - QD_\alpha} = 1$, 表示群落完全分化。地區 2: $\frac{N}{N-1} \frac{QD_\gamma - QD_\alpha}{T - QD_\alpha} = 0.1$, 表示群落幾乎沒有分化, 且指標值明確說明兩群落的分化情形: 因為在每個群落中, 10 物種中有 1 物種是特有種, 因此有 1/10 的分化程度。所以, $N_{ST} / \max(N_{ST})$ 的結果較符合實際群落的分化情況, 這是因為群落分化 (異質性) 的大小決定 $2\bar{D}_\beta(T)$ 值的大小, 而 $2\bar{D}_\beta(T)$ 的值不會受平均群落多樣性 $2\bar{D}_\alpha(T)$ (α 多樣性) 的影響。因此, $N_{ST} / \max(N_{ST}) = \frac{N}{N-1} \frac{QD_\gamma - QD_\alpha}{T - QD_\alpha}$ 能客觀地反應群落分化的真實情形。因此, 由上述簡單例子可以說明 Chiu et al. (2013) 所建構系統演化相似或分化指標相較傳統的 N_{ST} 更能客觀反應群落間物種組成的相似、分化程度。

如同單一群落的系統演化多樣性的分析，因為系統演化相似度（或分化）指標是根據傳統 β 多樣性和各物種的演化樹資訊所建構的指標。建議應該根據不同位階 q 和不同演化時間 T 所得到的系統演化相似度（或分化度）剖面（profile）圖來進行分析，如此才能客觀、完整地描述整個地區群落間在不同演化歷史下的相似（或分化）程度。

7. 統計在生物多樣性研究的角色

前面幾節的生物多樣性量化理論都是在假設物種豐富程度已知下所推導而成，亦及在參數層次（parameter-level）的討論。然而在實際的情況下，不可能對一個群落或地區物種及其豐富度做完整的普查。通常生物學家只能由收集的樣本資料來推論整個群落的多樣性，亦及在數據層次（data-level）的討論。但若以樣本中的觀察值做為群落中的參數值顯然是不合適。若豐富度變化很大，群落中有許多稀有種，通常會造成嚴重的低估。因此必須由數據層次推論至參數層次，統計模型用以連繫資料與群落扮演重要的角色。

我們以最簡單常用的研究為例，譬如要比較二個地區的生物多樣性，如果各取了一個樣本如何以數據來建立模型以比較多樣性呢？物種數的估計可以說是生物多樣性指標中最困難的估計，生物學家是如何比較二個地區的生物種數呢？又如何比較才合理？我們在下面以實例來說明。

7.1 比較兩地區物種數（標準化樣本數的稀釋-外插法）

此資料為 Janzen (1973 a, b) 所收集位於哥斯大黎加 (Costa Rica) 的奧薩 (Osa) 地區，原始林 (old growth) 與次森林 (second growth) 的甲蟲資料，其中在原始林收集了 237 隻甲蟲，共有 112 個物種；次森林收集了 976 隻甲蟲，共有 140 個物種。甲蟲在原始林與次森林資料的頻率分布見表 4 及表 5。

表 4 原始林甲蟲資料的頻率分布， S_{obs} 代表樣本中觀察到的物種數， n 代表樣本個體數， f_i 代表在樣本中出現 i 次物種數

i	1	2	3	4	5	6	7	8	14	42	S_{obs}	n
f_i	84	10	4	3	5	1	2	1	1	1	112	237

以觀察到的物種而言，原始林 (112 種) 較次森林 (140 種) 有較低的物種數。但原始林的樣本個體數只有 237，次森林的樣本個體數為 976。樣本數不同如何比較此二個地區的物種數？

表 5 次森林甲蟲資料, S_{obs} 代表樣本中觀察到的物種數, n 代表樣本個體數, f_i 代表在樣本中出現 i 次物種數

i	1	2	3	4	5	6	7	8	9	10	11	12	14	17	19
f_i	70	17	4	5	5	5	5	3	1	2	3	2	2	1	2
i	20	21	24	26	40	57	60	64	71	77		S_{obs}	n		
f_i	3	1	1	1	1	2	1	1	1	1		140	976		

傳統的“稀釋曲線”作法即為將次森林的樣本“稀釋”到 237。此種標準化生物多樣性樣本的作法最早由 Sanders (1968) 年提出, 其後經由生物學家 (Hurlbert, 1971; Simberloff, 1972; Gotelli and Colwell, 2001, 2011) 與統計學家 (Smith and Grassle, 1977) 的工作已有標準的“稀釋曲線”統計模型如下:

假設群落包含 S 種物種且物種 i 的相對豐富度表示為 p_i 。在歸還抽樣的方式下從一群落中抽取 n 個個體, 令 X_i 為第 i 個物種在樣本中出現的次數, 可假設隨機變數 (X_1, X_2, \dots, X_S) 服從多項式分配, 即

$$P(X_1 = x_1, X_2 = x_2, \dots, X_S = x_s) = \frac{n!}{x_1! \dots x_s!} p_1^{x_1} p_2^{x_2} \dots p_S^{x_s},$$

生態調查資料除了以物種出現次數 (X_1, X_2, \dots, X_S) 來呈現之外, 亦常使用物種出現頻率次數來表示, 令 $f_k = \sum_{i=1}^S I(X_i = k)$ 為樣本中恰出現 k 次的物種數, 其中 $k = 0, 1, \dots, n$ 。於此定義下, f_1 為樣本中恰出現一次的物種數, f_2 為樣本中恰出現兩次的物種數, f_0 為存在群落中之物種但其在樣本中的未被看到的物種數。此外, 樣本中看到的物種數等同於 $S_{obs} = \sum_{k \geq 1} f_k$, 樣本數 $n = \sum_{i=1}^S X_i = \sum_{k=1}^n k f_k$ 。

若令 $S(m)$ 為抽樣個數 m 時看到的物種數, 則其期望值 (Good, 1953) 為

$$E[S(m)] = S - \sum_{i=1}^S (1 - p_i)^m,$$

Smith and Grassle (1977) 證明 $S(m)$ 之最小變異數不偏估計為

$$\hat{S}(m) = S_{obs} - \sum_{X_i > 0} \frac{\binom{n-X_i}{m}}{\binom{n}{m}},$$

因此我們可利用樣本資訊將其套入上述估計式即可得到將個體數 n 稀釋至 m 的期望物種稀釋函數之估計, 進一步可以比較兩群落的物種數。而估計式 $\hat{S}(m)$ 之變異數估計式可採用 Chao and Jost (2012) 提出的拔靴方法 (bootstrap-method) 估計得到。進一步可以利用極限常態的特性得到 $E[S(m)]$ 信賴區間。

由以上 $\hat{S}(m)$ 的式子，可以求出對任意 $m < n$ 的物種數估計，此曲線稱為“稀釋曲線” (rarefaction curve) 基本上為一種內插 (interpolation) 的方法。見圖 3 中的實線部分即為原始林與次森林二地區的稀釋曲線，灰底陰影部份即為由拔靴方法得到之 95% 信賴區間。傳統生物學家的比較即是將次森林的樣本數由 976 稀釋到 237 而得到 70 之值 (見圖 3 中樣本數為 237 之曲線之值)，由此生物學家會下結論為：若樣本標準化到 $n = 237$ ，原始林有較多的物種數 (112 vs. 70)。

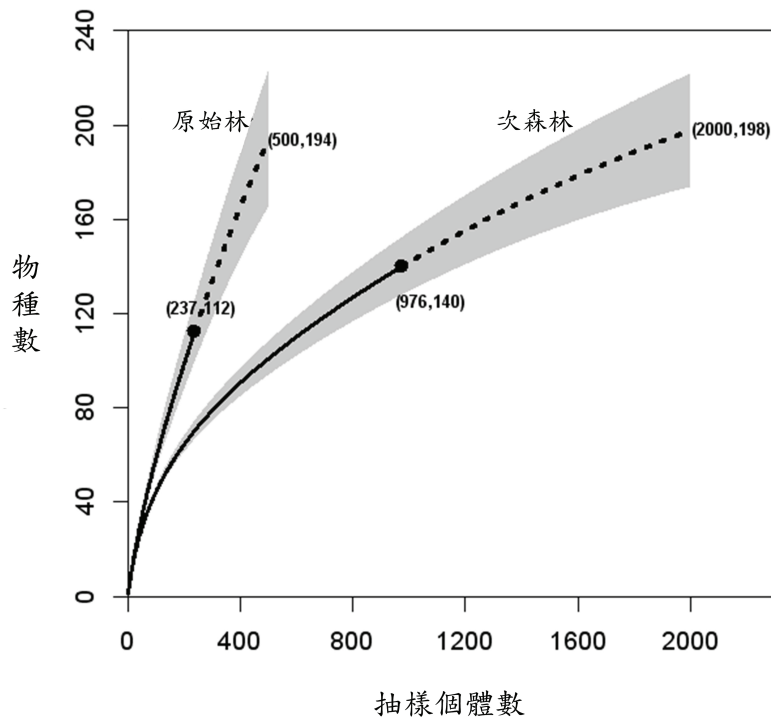


圖 3 甲蟲的物種稀釋與預測曲線對抽樣個體數作圖，其中，實心點為資料點，實線為物種稀釋部分，虛線為物種預測部分，灰底部分為物種數的 95% 信賴區間，括弧中的數字分別為個體數與物種數。

然而在以上的作法中，次森林的資料只有部分使用，浪費了大部分的資料。Colwell et al. (2012) 根據 Shen et al. (2003) 主張在稀釋曲線上，再加上“外插” (extrapolation) 曲線即為統計上的預測曲線 (prediction curve)。此部分的統計模型 (Shen et al., 2003) 如下：在以上建立稀釋曲線的多項式模型中，考慮再多收集 m^* 個個體，令 $S(n + m^*)$ 為在 $n + m^*$ 個

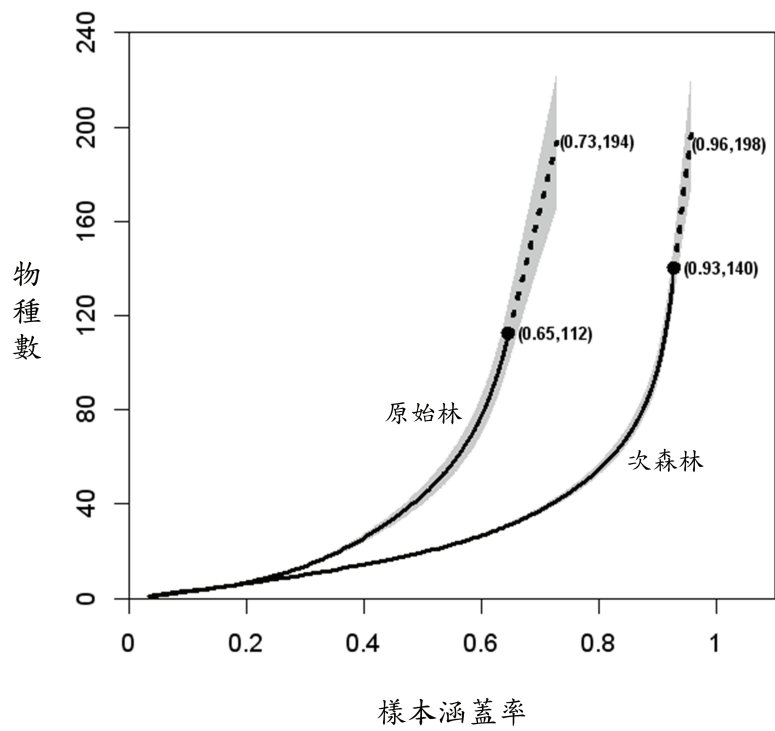


圖 4 甲蟲的物種稀釋與預測曲線對樣本涵蓋率作圖，其中，實心點為資料點，實線為物種稀釋部分，虛線為物種預測部分，灰底部分為物種數的 95% 信賴區間，括弧中的數字分別為樣本涵蓋率與物種數。

體之樣本中會看到的物種數，則統計相關的問題為如何估計。由上面的討論，Good (1953) 的式子對任何樣本數均成立，即

$$E[\hat{S}(n + m^*)] = S - \sum_{i=1}^S (1 - p_i)^{n+m^*},$$

但此時不再有不偏估計量。要估計以上的外插曲線必須仰賴對物種數的估計。而物種數的估計可以說是生物多樣性指標中最困難的估計。Colwell et al. (2012) 建議採用 Chao (1984, 1987) 之下界估計量，此估計量相當於利用出現 1 次 (f_1) 與 2 次 (f_2) 的物種來估計未在樣本中出現的物種 (f_0) 即

$$\hat{f}_0 = \frac{n-1}{n} \frac{f_1^2}{2f_2} \quad (f_2 > 0)$$

$$\text{或 } \hat{f}_0 = \frac{n-1}{n} \frac{f_1(f_1-1)}{2} \quad (f_2 > 0)。$$

由此，Shen et al. (2003) 及 Chao et al. (2009) 可得到 $D[\hat{S}(n + m^*)]$ 的一個估計量：

$$\hat{S}(n + m^*) = S_{obs} + \hat{f}_0 \left[1 - \left(1 - \frac{f_1}{n\hat{f}_0 + f_1} \right)^{m^*} \right]。$$

由於 \hat{f}_0 是一個下界估計量，因而外插的範圍就有所限制，Chao and Jost (2012) 建議 $m^* \leq n$ ，亦即外插的範圍最多為兩倍，即不能超過樣本數。若超過的話，可能產生較大的偏差。在圖 3 中我們加上了外插至兩倍的預測曲線。見圖 3 中的虛線部分即為原始林與次森林二地區的外插曲線，灰底陰影部份即為由拔靴方法得到之 95% 信賴區間。除樣本數極小外兩者的信賴區間都不重疊，因此二曲線及信賴區間可以知道對任何樣本小於 $237 \times 2 = 474$ 的比較，原始林有顯著較高的多樣性。

7.2 比較兩地區物種數 (標準化樣本涵蓋的稀釋-外插法)

在前一節中，生物學家一直以為若樣本數一致，則比較同一樣本數的物種數是一種公平的比較。然而對一個物種數稀少的寒帶地區，可能即可以調查到所有的物種，但對一個物種繁多的熱帶雨林，同樣可能只看到極少一部分的物種。因此對生物多樣性的樣本，往往不能以相同的樣本數作客觀公平的比較。Chao and Jost (2012) 提出以樣本的完整性 (Sample Completeness) 作為標準化的基準。長久以來，幾乎所有的生物學家都不認為樣本的完整性可以準確的估計。然而超乎想像的，樣本的完整性可以很精準以樣本涵蓋 (sample coverage) 來估計。

早在第二次世界大戰期間，計算機科學之父 Alan Turing 及 I. J. Good 針對破譯密碼的需要就已發展樣本涵蓋的觀念。Turing 的統計工作在戰後由 Good (1953, 1956, 2000) 發表。其中一個最重要的統計觀念即為樣本涵蓋，對一組樣本數為 n 的樣本涵蓋定義為

$$C(n) = \sum_{i=1}^S p_i I(X_i > 0),$$

$C(n)$ 即為所有在樣本中看到物種之機率的總和，或為所有在樣本中看到物種其在群落個體數佔全部群落個體數的比例。此樣本涵蓋可以作為樣本完整性的客觀量度。或許大家會認為此種完整性不容易由有限的樣本來估計，令人訝異的是 Turing 及 Good 發現此完整度可以僅以出現 1 次的物種數之比例就可精準的估計，其估計量為

$$\hat{C}(n) = 1 - \frac{f_1}{n}.$$

Chao and Jost (2012) 採用更精確的估計：

$$\hat{C}(n) = 1 - \frac{f_1}{n} \left[\frac{(n-1)f_1}{(n-1)f_1 + 2f_2} \right].$$

為了要標準化樣本涵蓋，對任一樣本數為 m 的樣本涵蓋 $C(m)$ 之期望值為

$$E[C(m)] = 1 - \sum_{i=1}^S p_i (1 - p_i)^m.$$

對稀釋的樣本 $m < n$ ，Chao and Jost (2012) 導出最小變異不偏估計量為

$$\hat{C}(m) = 1 - \sum_{X_i \geq 1} \frac{X_i}{n} \frac{\binom{n-X_i}{m}}{\binom{n-1}{m}}, \quad m < n.$$

對外插的樣本數 $n + m^*$ ，其期望樣本涵蓋 $E[C(n + m^*)]$ 沒有不偏估計量，且須要一個未出現物種的估計量才能得到。若如上節採用相同 \hat{f}_0 ，則可得到

$$\hat{C}(n + m^*) = 1 - \frac{f_1}{n} \left[\frac{(n-1)f_1}{(n-1)f_1 + 2f_2} \right]^{m^*+1}.$$

因此一個整合性的樣本涵蓋為基準的稀釋曲線即為 $\hat{S}(m)$ 對 $\hat{C}(m)$ 作圖，而外插曲線即為 $\hat{S}(n + m^*)$ 對 $\hat{C}(n + m^*)$ 作圖，兩部份會平滑相接於資料點 $(\hat{C}(n), S_{obs})$ 。

針對次森林與原始林的資料，其樣本涵蓋為標準化的稀釋-外插曲線如圖 4。根據採用對 f_1 與 f_2 來估計的樣本涵蓋，在原始林中為 65%；而在次森林中為 93%。如果樣本數加倍，即原始林樣本數由 237 增至 500，則樣本涵蓋由 65% 增至 73%；次森林樣本數由 976 增至 2000，

樣本涵蓋由 93% 增至 96%，增加很少。因此對樣本完整性為 73% 以下的任何值都可比較如圖 4，可得除樣本涵蓋極小外兩者的信賴區間都不重疊，因此對任何樣本涵蓋小於 73% 的標準化樣本，原始林都有顯著較高的多樣性。

上述之兩種稀釋-外插曲線估計，可以看出統計之工作者經由簡單的多項式模型進而架構稀釋-外插曲線，並且可以推論在稀釋部分為最小變異不偏估計，而在外插部分經由統計的模型可以作有限度的外插估計，在小於 2 倍的範圍內得到有用可靠的結果。雖然以上都只有比較物種數，相關的稀釋-外插理論最近已推廣至 Hill 指標族 (Chao et al., 2013)。最近的將來並將推廣至系統演化指標。

以上的研究都係由生物學家與統計學者共同研究而成。統計學者必須了解到生物數據背景與限制條件為何？所須要的分析為何？而生物學家也必須了解統計模型的優劣，才能共同合作完成研究。

7.3 相關軟體簡介

生物多樣性的量度與計算都仰賴相關的軟體，多年來我們發展了下列的軟體，簡述軟體的內容如下：

(1) SPADE (Species Prediction And Diversity Estimation)

本文第三節及第五節所討論的古典多樣性指標與多群落的相似性 (或分化性) 之所有相關統計估計。

(2) PhD (Phylogenetic Diversity)

第四節及第六節所討論的系統演化指標之統計估計。

(3) iNEXT (interpolation/Extrapolation)

第七節所討論之以樣本數或樣本涵蓋為標準化的稀釋 (內插) 與外插曲線之估計。

以上軟體皆可經由網頁 <http://chao.stat.nthu.edu.tw/softwareCE.html>，免費下載使用。

8. 討論與展望

因為多樣性的相關研究是近幾年才逐漸受到重視，因此與生態理論或保育政策相關的新生物多樣性問題也相繼被提出。而為了解決這些問題，首先需克服的是，如何將這些生態問題用嚴謹的數學來描述，並且定義相關的多樣性指標，使其指標值可以客觀地描述生態多樣性的相關

特性, 進而解決生態學或保育決策的問題。以下將介紹未來生物多樣性指標發展的幾個重點,

功能性多樣性指標

這是考量物種間生態功能差異的指標, 生態學家認為群落的穩定與否, 取決於該群落的生態功能是否健全, 因此生態功能多樣性指標, 是另一生態學家強調的議題, 雖然已經很多相關指標陸續被提出 (Bady et al. 2005; Botta, 2005; Diaz and Cabido, 2001; Heemsbergen et al., 2004; Mouchet et al., 2010; Petchey and Gaston, 2002, 2006, 2007; Podani, 1999; Ricotta, 2005; Ricotta and Burrascano, 2008; Ricotta et al., 2010, 2011, 2012; Schleuter et al., 2010; Swenson et al., 2011; Tilman, 2001; Villéger et al., 2008; Walker et al., 2008), 但這些指標大都無法滿足“倍增性質”, 或者沒有考慮物種的豐富度, 因此無法客觀衡量群落的功能性多樣性。所以如何透過物種生態特徵的差異來定量群落功能性多樣性是亟待解決的問題。

多層次多樣性指標

現今多樣性指標只能描述單一地區的多樣性(γ 多樣性) 或一個地區中不同群落間物種組成的差異(β 多樣性), 然而對於多樣性在多層次間的交互關係卻仍未被討論。如: 群落由物種組成, 而地區由群落組成, 地景由地區所構成, 在這樣層層的生態體系架構中, 如何透過多樣性的角度去衡量各層級間的關係是生態學家欲解決的問題。

時間與空間交互作用的多樣性指標

探討一地區物種在時間和空間的變化, 一直是生態學上一個最重要的研究主題之一, 因此如何透過物種多樣性的角度來回答這個問題, 變成一個解決這個問題的一個重要窗口。所以, 如何客觀定量時間與空間交互作用的多樣性是生態學家一直努力的方向。

多樣性指標之統計估計

雖然自 Fisher et al. (1943) 論文以來, 以統計理論為基礎的生物多樣性研究已近七十載, 然而由有限資料來推論多樣性的工作事實上仍有許多亟待解決的問題。即使以古典的 Hill 指標族而言, 自 Hill 在 1973 年提出後, 如何從有限的樣本中得到可靠的多樣性剖面都仍未有定論, 更遑論系統演化指標的剖面估計。而多群落的相似度與分化度的估計亦有待完成, 因此未來多樣性的估計是生態學家及統計學家可以共同合作發展的研究領域。

多樣性理論應用於其他學科

在本文所討論的多樣性測度量化理論多半都是近年的研究，由於此理論與遺傳統計有許多共同之處，因此如何應用於遺傳學將是極具挑戰的方向。預期在其他學科（例如生物資訊、訊息科學及經濟學）也會有類似的應用。

結語

量度一個地區的生物多樣性，如果不考慮物種間分類或系統演化的關係，則 Hill 指標族滿足倍增原則及一些生物學家所認為生物多樣性測度所應具有的性質，故 Hill 指標族的剖面圖（見圖 1）可以完全刻畫生物多樣性。同樣考量物種的分類或系統演化時，以 Hill 指標推廣的系統演化指標是吾人認為最理想的指標，因此根據不同位階 q ($q = 0, 1, 2$)，對不同演化時間做剖面圖（如圖 2-2），亦完全刻畫系統演化的多樣性。在比較多群落的多樣性時，多樣性的分解（無論考慮物種分類或系統演化與否），吾人亦建議採用上述之指標，並利用上述之指標所得到的相似（或相異）指標來比較多群落。

Nanny (2004, p.721) 曾言：“We are all blind men (and women) trying to describe a monstrous elephant of ecological and evolutionary diversity.” 同樣，對於生物多樣性的研究，我們亦如瞎子摸象，可能都只瞭解極小的一部份，但此方向帶給我們師生夙興夜寐、夙夜匪懈、孜孜不倦的研究主題以及和生物學家們的共同合作與互相學習之樂趣卻是無窮的。

誌謝詞

感謝主編陳玉英教授的邀請撰寫此篇回顧論文，趙蓮莉感謝國科會相關研究計畫資助，並感謝多年來一起做研究的學生們及共同合作學習的生物學家們。邱春火的研究由清華大學博士後人員計畫資助；王怡婷的研究由國科會博士人員計畫資助。謝宗震與馬光輝由國科會計畫下博士班獎助金資助。

參考文獻

劉小如、丁宗蘇、方偉宏、林文宏、蔡牧起、顏重威 (2010)。台灣鳥類誌 (上)、(中)、(下)。行政院農業委員會林務局。

Allen, B., Kon, M. and Bar-Yam, Y. (2009). A new phylogenetic diversity measure

- generalizing the Shannon index and its application to phyllostomid bats. *The American Naturalist*, **174**, 236–243.
- Anderson, M. J., Crist, T. O., Chase, J. M., et al. (2011). Navigating the multiple meanings of beta diversity: a roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.
- Arrhenius, O. (1921). Species and area. *Journal of Ecology*, **9**, 95–99.
- Bacaro, G., Ricotta, C. and Mazzoleni, S. (2007). Measuring beta-diversity from taxonomic similarity. *Journal of Vegetation Science*, **18**, 793–798.
- Bady, P., Dolédec, S., Fesl, C., Gayraud, S., Bacchi, M. and Schöll, F. (2004). Use of invertebrate traits for the biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. *Freshwater Biology*, **50**, 159–173.
- Barker, G. M. (2002). Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. *Biological Journal of the Linnean Society*, **76**, 165–194.
- Baselga, A. (2010). Multiplicative partition of true diversity yields independent alpha and beta components; additive partition does not. *Ecology*, **91**, 1974–1981.
- Beck, C. and Schögl, F. (1995). Thermodynamics of chaotic systems: an introduction (Vol. 4). Cambridge: Cambridge University Press.
- Botta-Dukát, Z (2005). Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, **16**, 533–540.
- Bryant, J. A., Lamanna, C., Morlon, H., et al. (2008). Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences USA*, **105**, 11505–11511.
- Cadotte, M. W., Jonathan Davies, T., Regetz, J., Kembel, S. W., Cleland, E. and Oakley, T. H. (2009). Phylogenetic diversity metrics for ecological communities:

- integrating species richness, abundance and evolutionary history. *Ecology Letters*, **13**, 96–105.
- CBD, 2012. Covention on Biological Diversity. <http://www.cbd.int/>.
- Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, **11**, 265–270.
- Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, **43**, 783–791.
- Chao, A. (2005). Species estimation and applications. Pages 7907-7916 in S. Kotz, N. Balakrishnan, C. B. Read, and B. Vidakovic, editors. *Encyclopedia of Statistical Sciences*, New York: Wiley.
- Chao, A., Chiu, C.-H. and Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 3599–3609.
- Chao, A., Chiu, C.-H. and Hsieh, T. C. (2012). Proposing a resolution to debates on diversity partitioning. *Ecology*, **93**, 2037–2051.
- Chao, A., Colwell, R. K., Lin, C. W. and Gotelli, N. J. (2009). Sufficient sampling for asymptotic minimum species richness estimators. *Ecology*, **90**, 1125–1133.
- Chao, A. and Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, **93**, 2533–2547.
- Chao, A., Jost, L., Chiang, S. C., Jiang, Y.-H. and Chazdon, R. L. (2008). A two-stage probabilistic approach to multiple-community similarity indices. *Biometrics*, **64**, 1178–1186.
- Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K. and Ellison, A. M. Rarefaction and extrapolation of Hill numbers: a unifying sampling framework for biodiversity studies. *Under revision, Ecological Monographs*.

- Chiu, C.-H., Jost, L. and Chao, A. (2013). Phylogenetic beta diversity, similarity, and differentiation measures. *Under revision, Ecological Monographs*.
- Clarke, K. R. and Warwick, R. M. (1998). A taxonomic distinctness index and its statistical properties. *Journal of Applied Ecology*, **35**, 523–531.
- Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S. Y., Mao, C. X., Chazdon, R. L. and Longino, J. T. (2012). Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology*, **5**, 3–21.
- Costello, M. J., May, R. M. and Stork, N. E. (2013). Can we name earth's species before they go extinct?. *Science*, **339**, 413–416.
- Crozier, R. H. (1992). Genetic diversity and the agony of choice. *Biological Conservation*, **61**, 11–15.
- Crozier, R. H. (1997). Preserving the information content of species: genetic diversity, phylogeny, and conservation Worth. *Annual Review of Ecology and Systematics*, **28**, 243–268.
- Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, **30**, 348–361.
- d'Arnoldi, C. T., Foulley, J. L. and Ollivier, L. (1998). An overview of the Weitzman approach to diversity. *Genetics Selection Evolution*, **30**, 149–161.
- Diaz, S. and Cabido, M. (2001). Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, **16**, 646–655.
- Ellison, A. M. (2010). Partitioning diversity. *Ecology*, **91**, 1962–1963.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.
- Faith, D. P. and Baker, A. M. (2006). Phylogenetic diversity (Pd) and biodiversity conservation: some bioinformatics challenges. *Evolutionary Bioinformatics* **2**, 121.

-
- Fisher, R. A., Corbet, A. S. and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, **12**, 42–58.
- Flather, C. H. (1996). Using landscape ecology to test hypotheses about large-scale abundance patterns in migratory birds. *Ecology*, **77**, 28–35.
- Gleason, H. A. (1922). On the relation between species and area. *Ecology*, **3**, 158–162.
- Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, **40**, 237–264.
- Good, I. J. (2000). Turing’s anticipation of empirical bayes in connection with the cryptanalysis of the naval enigma. *Journal of Statistical Computation and Simulation*, **66**, 101–111.
- Good, I. J. and Toulmin, G. (1956). The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika*, **43**, 45–63.
- Gotelli, N. J. and Chao, A. (2013). Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. Levin, SA, editor. *The Encyclopedia of Biodiversity*, New York: Elsevier.
- Gotelli, N. J. and Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.
- Gotelli, N. J., and Colwell, R. K. (2011). Estimating species richness. Pages 39–54 in A. E. Magurran and B. J. McGill, editors. *Biological diversity: frontiers in measurement and assessment*, New York: Oxford University Press.
- Graham, C. H. and Fine, P. V. A. (2008). Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, **11**, 1265–1277.

- Hamady, M., Lozupone, C. and Knight, R. (2009). Fast Unifrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and phylochip data. *The ISME Journal*, **4**, 17–27.
- Hardy, O. J. and Senterre, B. (2007). Characterizing the phylogenetic structure of communities by an Aaditive partitioning of phylogenetic diversity. *Journal of Ecology*, **95**, 493–506.
- Heemsbergen, D. A., Berg, M. P., Loreau, M., Van Hal, J. R., Faber, J. H. and Verhoef, H. A. (2004). Biodiversity effects on soil processes explained by interspecific functional dissimilarity. *Science*, **306**, 1019–1020.
- Helmus, M. R., Bland, T. J., Williams, C. K. and Ives, A. R. (2007). Phylogenetic measures of biodiversity. *The American Naturalist*, **169**, 68–83.
- Helmus, M. R., Savage, K., Diebel, M. W., Maxted, J. T. and Ives, A. R. (2007). Separating the determinants of phylogenetic community structure. *Ecology Letters*, **10**, 917–925.
- Hill, M. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, **54**, 427–432.
- Horn, H. S. (1966). Measurement of “overlap” in comparative ecological studies. *The American Naturalist*, 419–424.
- Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton: Princeton University Press.
- Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, **52**, 577–586.
- Izsák, J. and Papp, L. (1995). Application of the quadratic entropy indices for diversity studies of drosophilid assemblages. *Environmental and Ecological Statistics*, **2**, 213–224.
- Izsak, J. and Papp, L. (2000). A link between ecological diversity indices and measures of biodiversity. *Ecological Modelling*, **130**, 151–156.

-
- Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles*, **44**, 223-292.
- Janzen, D. H. (1973a). Sweep samples of tropical foliage insects: description of study sites, with data on species abundances and size distributions. *Ecology*, **54**, 659–686.
- Janzen, D. H. (1973b). Sweep samples of tropical foliage insects: effects of seasons, vegetation types, elevation, time of day, and insularity. *Ecology*, **54**, 687–708.
- Jost, L. (2006). Entropy and diversity. *Oikos*, **113**, 363–375.
- Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427–2439.
- Jost, L. (2010). Independence of alpha and beta diversities. *Ecology*, **91**, 1969–1974.
- Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos*, **76**, 5–13.
- Lozupone, C. and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.
- Lozupone, C. A., Hamady, M., Kelley, S. T. and Knight, R. (2007). Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, **73**, 1576–1585.
- MacArthur, R.H. (1957). On the Relative Abundance of Bird Species. *Proceedings of the National Academy of Sciences of the United States of America*, **43**, 293–295.
- MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews*, **40**, 510–533.
- Magurran, A. E. (2004). Measuring biological diversity. Oxford: Blackwell.

- Magurran, A. E. and McGill, B. J. (2010). Biological diversity. Oxford University Press.
- Marshall, A. W. and Olkin, I. (1979). Theory of Majorization and its Applications. *Academic*, New York.
- Mora, C., D. P. Tittensor, S. Adl, A. G. B. Simpson, and B. Worm. (2011). How many species are there on Earth and in the Ocean?. *PLoS Biology*, **9**, e1001127.
- Morisita, M. (1959). Measuring of interspecific association and similarity between communities. *Memoires of the Faculty of Science, Kyushu University, Series E*, **3**, 65–80.
- Mouchet, M. A. and Mouillot, D. (2011). Decomposing phylogenetic entropy into α , β and γ components. *Biology Letters*, **7**, 205–209.
- Mouchet, M. A., Villéger, S., Mason, N. W. H. and Mouillot, D. (2010). Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology*, **24**, 867–876.
- Nanney, D. L. (2004). No trivial pursuit. *Bioscience*, **54**, 720–721.
- Nei, M. (1982). Evolution of human races at the gene level. Human genetics, part A: the unfolding genome. *Alan R. Liss, New York*, 167–181.
- Patil, G. P. and Taillie, C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association*, **77**, 548–561.
- Pavoine, S., Love, M. S. and Bonsall, M. B. (2009). Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically structured species assemblages: application to rockfish (genus: *Sebastes*) in the Southern California Bight. *Ecology Letters*, **12**, 898–908.
- Petchey, O. L. and Gaston, K. J. (2002). Functional diversity (FD), species richness and community composition. *Ecology Letters*, **5**, 402–411.

-
- Petchey, O. L. and Gaston, K. J. (2006). Functional diversity: back to basics and looking forward. *Ecology Letters*, **9**, 741–758.
- Petchey, O. L. and Gaston, K. J. (2007). Dendrograms and measuring functional diversity. *Oikos*, **116**, 1422–1426.
- Pielou, E. C. (1975). *Ecological Diversity*. New York: Wiley.
- Podani, J. (1999). Extending Gower's general coefficient of similarity to ordinal characters. *Taxon*, **48**, 331–340.
- Preston, F. W. (1948). The commonness and rarity of species. *Ecology*, **29**, 254–283.
- Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, **21**, 24–43.
- Rényi, A. (1961). On measures of entropy and information. Proceedings of the 4th Berkeley Symposium on Mathematics. *Statistics and Probability*, **1**, 547–561.
- Ricotta, C. and Szeidl, L. (2006). Towards a unifying approach to diversity measures: bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*, **70**, 237–243.
- Sanders, H. L. (1968). Marine benthic diversity: a comparative study. *American Naturalist*, **102**, 243–282.
- Schleuter, D., Daufresne, M., Massol, F. and Argillier, C. (2010). A user's guide to functional diversity indices. *Ecological Monographs*, **80**, 469–484.
- Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.
- Shen, T.-J., Chao, A. and Lin, J.-F. (2003). Predicting the number of new species in further taxonomic sampling. *Ecology*, **84**, 798–804.
- Shimatani, K. (2001). On the measurement of species diversity incorporating species differences. *Oikos*, **93**, 135–147.

- Simpson, E. H. (1949). Measurement of diversity. *Nature*, **163**, 688.
- Smith, W. and Grassle, J. (1977). Sampling properties of a family of diversity measures. *Biometrics*, **33**, 283–292.
- Solow, A. R. and Polasky, S. (1994). Measuring biological diversity. *Environmental and Ecological Statistics*, **1**, 95–103.
- Solow, A., Polasky, S. and Broadus, J. (1993). On the measurement of biological diversity. *Journal of Environmental Economics and Management*, **24**, 60–68.
- Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske skrifter*, **5**, 1–34.
- Swenson, N. G., Anglada-Cordero, P. and Barone, J. A. (2011). Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient *Proceedings of the Royal Society B: Biological Sciences*, **278**, 877–884.
- Tilman, D. (2001). Functional diversity. *Encyclopedia of biodiversity*, **3**, 109–120.
- Tong, Y. L. (1972). Some distribution properties of the sample species-diversity indices and their applications. *Biometrics*, **39**, 999–1008.
- Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, **52**, 479–487.
- Vane-Wright, R. I., Humphries, C. J. and Williams, P. M. (1991). What to protect: systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.
- Veech, J. A. and Crist, T. O. (2010). Diversity partitioning without statistical independence of alpha and beta. *Ecology*, **91**, 1964–1969.
- Veech, J. A. and Crist, T. O. (2010). Toward a unified view of diversity partitioning. *Ecology*, **91**, 1988–1992.

-
- Vellend, M., Vornwell, W., Magnuson-Ford, K. and Mooers, A. O. (2010). Measuring phylogenetic biodiversity. *Biological diversity: Frontiers in measurement and assessment* Oxford: Oxford University Press.
- Villéger, S., Mason, N. W. H. and Moullot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, **89**, 2290–2301.
- Walker, S. C., Poos, M. S. and Jackson, D. A. (2008). Functional rarefaction: estimating functional diversity from field data. *Oikos*, **117**, 286–296.
- Warwick, R. and Clarke, K. (1995). New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology Progress Series*, **129**, 301–305.
- Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, **156**, 145–155.
- Weikard, H. P., Punt, M. and Wessler, J. (2006). Diversity measurement combining relative abundances and taxonomic distinctiveness of species. *Biology Letters*, **12**, 215–217.
- Weitzman, M. L. (1992). On diversity. *Quarterly Journal of Economics*, **107**, 363–406.
- Weitzman, M. L. (1993). What to preserve? An application of diversity theory to crane conservation. *Quarterly Journal of Economics*, **108**, 155–183.
- Weitzman, M. L. (1998). The Noah's Ark problem. *Econometrica*, **66**, 1279–1298.
- Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs*, **30**, 279–338.
- Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, **12**, 213–251.

- Whittaker, R. H. (1977). Evolution of species diversity in land communities. *Evolutionary Biology*, **10**, 1–67.
- Wilson, E. (1992). *The Diversity of Life*. Belknap Press/Harvard University Press. Cambridge Massachusetts, USA:.
- Wilson, M. and Shmida, A. (1984). Measuring beta diversity with presence-absence data. *The Journal of Ecology*, **72**, 1055–1064.
- Wolda, H. (1981). Similarity indices, sample size and diversity. *Oecologia*, **50**, 296–302.

[Received February 2013; accepted February 2013.]

趙蓮菊教授於 1977 年自美國威斯康辛大學獲得博士學位後，次年回台在國立清華大學（清大）統計研究所任教。其間獲國科會傑出獎（1989-1995）、教育部學術獎（2003），更榮任國家講座教授（2005-2008），自 2006 年起獲聘為清大特聘講座教授。趙教授於 1997 年獲選為數理統計學會會士，並且在 2001 年榮獲 SCI 的「經典引文」獎。除著力於研究外，趙教授也熱心學術服務，曾擔任 *Statistica Sinica* 的副主編，迄今仍擔任 *Biometrics*, *J. Agricultural, Biological, and Environmental Statistics* 及 *Australian & New Zealand J. Statistics* 三份國際期刊的副主編。

HOW TO QUANTIFY BIODIVERSITY?

Anne Chao, Chun-Huo Chiu, Y. T. Wang, T. C. Hsieh and K. H. Ma
Institute of Statistics, National Tsing Hua University

ABSTRACT

Biodiversity refers to the variety and variability among living organisms and the ecological complexes in which they occur, and related concepts of similarity and differentiation. Biodiversity has at least three aspects: gene (genetic) diversity, species (taxonomic) diversity and ecosystem (functional) diversity. This paper reviews species (or taxonomic) diversity and cover the following topics: (1) traditional diversity measures in a community based on Hill numbers. This class of measures only considers species richness and species abundances. (2) Phylogenetic diversity measures based on Hill numbers for an assemblage. This class of measures additionally incorporates information on evolutionary/phylogenetic distances between species. (3) Traditional diversity decomposition and related similarity (or differentiation) measures for multiple communities. (4) Phylogenetic diversity decomposition and related similarity (or differentiation) measures for multiple assemblages. We also use a real example to illustrate the role of statistical models in estimating biodiversity measures. Some future topics are also discussed.

Key words and phrases: biodiversity, diversity decomposition, phylogenetic diversity, replication principle, taxonomic diversity.

JEL classifications: C13.