## 相關文獻回顧

### 出現型數據

在大多數的物種數估計的研究中的理論假設皆為隨機且獨立的抽樣，但在通常情況下，個體抽樣的方式較難符合該假設，而相較於個體抽樣，區塊抽樣較能符合該理論假設。區塊且抽樣單位通常為陷阱、區塊與固定時間調查。在出現型數據的抽樣中，大多數的方法是將其中的研究區域劃分為多個面積大致相同的區塊，並從中隨機選擇區塊做為抽樣樣本進行調查。對於不同類型的物種，準確計算每個抽樣區塊中出現的個體數往往是一件相對困難的。因此在多數情況下，調查時僅記錄該物種抽樣區塊中的物種出現與否。

出現型數據由*t*個抽樣區塊的樣本所組成，其中每個抽樣區塊只紀錄物種的出現或是未出現，以形成一個以物種為列區塊為行的矩陣的矩陣。其中，若是*i*物種出現在*j*區塊中，則計為1；反之若尚未觀測到該物種則計為0。

以第一群落為例，被定義為落樣本中第物種出現的區塊數量，，表示在該樣本中第*i*物種出現的總區塊數量，故 。

並且，可令表示在出現頻率向量中出現*k*次的物種數，，且。故 為在該樣本中僅出現在一個區塊的物種數，為在該樣本中出現在兩個區塊的物種數，並依此類推。除此之外，為在該樣本並未被觀測到的物種數。而真實的物種數，應為被觀測到的物種數與未被觀測到的物種數之和。

### 取後放回之抽樣方式

在生態調查的研究中，物種數或稱物種豐富度是最直接呈現多樣性的指標之一。但普查物種數量往往需要消耗大量的人力、經費與時間等成本，這使得在抽樣的結果中能看見所有物種皆出現之狀況的機率大幅降低，因此在大部分的生態調查結果中，皆可能存在部分未被觀測到的物種。故需針對這些未被觀測到的物種進行估計，以獲取到更接近於真實物種豐富度的結果。

對於出現型數據所開發之物種豐富度估計的模型多數皆是依據捉放法 (capture-recapture) 的理論架構為基礎所建立。一般而言，傳統的捉放法是藉由該物種在樣本中所佔比例估計個體數，而在物種豐富度的估計中，可將捉放法「個體數」對應至「物種數」，以估計群落中的物種數作為物種豐富度的指標所使用。

在物種豐富度的調查結果中，可將物種大致分為豐富物種與稀有物種。樣本中稀有物種對於未被觀測到的物種提供了更為豐富的資訊。相較於皆為豐富物種的樣本，在某樣本中含有大量的稀有物種時，該樣本所抽樣地區應存在更多尚未被觀測到的物種。因此在過去許多研究中，皆是藉由稀有物種對物種豐富度的估計進行物種數估計。

依據2.2.1所述，並以第一群落為例，出現型矩陣可整理為，應為服從伯努力分佈 (Bernoulli distribution) 的隨機變數，且當時機率為 ，而時機率為。則發生率矩陣的機率分佈為：

又 服從二項分佈 (Binomial distribution)，則其機率分佈為：

#### 單群落物種數估計

Chao (1987) 針對出現型數據建立物種豐富度的無母數估計模型 *Chao2*。無母數的物種豐富度估計是一個基本且直接的觀念，所謂無母數估計意旨在該估計方法中，不對物種豐富度或者物種出現機率的分布進行假設。*Chao2*根據柯西-施瓦茨不等式 (Cauchy-Schwarz inequality)，利用稀有物種中所含的資訊以估計真實的物種豐富度。最終求得*Chao2*估計式為：

針對*Chao2* 估計式，Chiu 等人 (2014) 應用Good-Turing 頻率公式，並加入 與的資訊對其進行修正，得到更為準確的下界估計式 *iChao2* ：

#### 兩群落的共同物種數估計

在生態統計中，群落之間的共同物種在樣本中所佔的比例，可以表示兩群落間的物種多樣性，同時也能表現兩群落間的相似性 (Chao, et al. 2000)。在兩群落的抽樣樣本中除了共同種之外，也會分別存在只出現於其中單一群落的特有物種。為此，當在比較兩群落之間的物種豐富度時，並非僅考慮單一群落的物種豐富度，也必須針對群落間的共同物種數進行估計。與單群落的物種數估計相似，在大多數情況下，抽樣樣本無法觀測到所有存在的共同種。因此需針對未被抽樣觀測到的共同種進行估計，並加上已存在於樣本中的共同物種數，作為修正的共同物種數所使用。Pan等人(2009) 將*Chao2*的方法推廣至兩群落，建立一估計兩群落間存在的共同物種數之估計式*Pan*：