# 使用動差法估計共同物種數

## 取後放回之抽樣方法的估計方式

在單群落的情況下，以第一群落的樣本為例，假設在目標區域實際存在種物種， 其中為一個未知參數。且抽樣方法是從目標區域中針對其中的抽樣區塊進行隨機抽樣，並記錄每個區塊中的物種存在與否。若是該樣本總共包含個抽樣區塊，並且 表示第*i*物種在樣本中出現的區塊數量。則遵循總數為，機率為的二項分佈 (binomial distribution) 。在此，除了取決於第*i*物種的族群規模外，也與其多種的生物特徵因素相關。

Chiu (2022) 使用Beta二項式模型，建立一個新的針對單群落物種數的有母數估計方法。假設 遵循二項分佈，且() 為機率密度函數 服從 的隨機變數。假設 為樣本中出現的區塊數正好為的平均機率，最終獲得以下樣本之物種出現區塊數的機率分佈如下：

又令表示在樣本中 個區塊的物種數，而。並根據樣本中物種出現次數的機率分佈，可知可以表示為：

依據上述式子，可獲得、 以及 ：

並根據柯西-施瓦茨不等式 (Cauchy-Schwarz inequality) 之概念與Good-Turing頻率公式 (Good, 1953, 2000) 得出近似式：。由該近似式可以得知，出現於較少區塊的稀有物種可以為未被觀測到的物種數提供更多的估計資訊。

將上述概念推廣至兩群落。假設兩群落皆為隨機且取後放回的抽樣，則 與 皆分別遵循二項分佈以及。並假設，；，，且機率密度函數分別為服從 與的與。將定義為樣本中出現的區塊數正好分別為 和 的平均機率。則：

其中與。令 表示在第一群落的樣本中存在*k*個區塊，且在第二群落的樣本中存在*l*個區塊的物種數。且樣本中觀測到的共同物種數為。藉此，可獲得、以及：

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |
|  | (3) |

將 設定為1，可藉由式 (1)、式 (2) 與式 (3) 成立以下近似值：

並經由化簡 可得：

|  |  |
| --- | --- |
|  | (4) |

以及經由求得：

並由於 為適用於所有二項分佈的下界估計 (Pan et al., 2009; Chao 1987)，因此可依據 (4) 得知，在假設物種出現於群落的比例 () 服從Beta二項分佈時，則估計量 的偏差近似於 。

且由柯西-施瓦茨不等式可推得 ，並結合式 (4) 可得知 的偏差絕對幅度的下界近似為不等式：

故將 帶入式 (4)，且為了確保估計是的穩定與，將滿足 ，最終可獲得：

且在上述等式中皆表示為：若 時，則；而若 時，則表示為。同理，可經由上述相同方式推導出與：

最後加入對估計式進行修正，最終得估計式：

其中、以及分別為：

且在上述等式中皆表示為：若 時，則；而若 時，則表示為。