## 取後不放回之抽樣方法的估計方式

依據2.2.3節所示，並以第一群落樣本為例的情況下，在目標區塊中檢視到物種*i*所存在的區塊數，其分布應服從一個二項分佈；且第*i*物種出現的區塊數所組成的出現次數，在給定 的情況下，應服從超幾何分佈。又與有關，來自於，因此可推導出：

其中，為指標函數，表示若發生時，則該式結果計為1，反之則該式結果計為0。

Shen 和 He (2008) 針對取後不放回的抽樣方式，開發有母數的估計方法，假設 遵循超幾何分佈，且() 為服從 的隨機變數，故機率密函數 可表示為：

其中 與 ，且。

並且 的機率分布可依據 與 獲得 (Shen & He, 2008)。設 為樣本中出現的區塊數正好為的平均機率，則：

令表示在個區塊中準確觀測到的物種數，而為在單群落樣本中出現個區塊數。故根據，可定義出 為：

依據上述式子可知、以及分別為：

並基於Good-Turing頻率公式與柯西不等式之概念，針對單一群落樣本的估計得出近似式：。從中可以得知，在物種估計時，採取出現較少次的物種，可以更多提供未出現物種的資訊，有助於減少物種數估計結果的偏差。

隨後依據單群落估計式的概念，將估計推廣至兩群落。則 與 皆分別遵循超幾何分佈，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，。

並又令 為在樣本中第一群落樣本出現*k*次且第二群落樣本出現*l*次的區塊數，則 為樣本中觀測到的共同物種數量，。藉此，可獲得、以及：

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

依據式 (1) 與式 (2) 成立以下近似值。並將 設定為1，且假設：

由上述式子可得出：

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

將並將其代入 ，可得：

同理 也依此證明，可得：

|  |  |
| --- | --- |
|  | (5) |

並將該結果帶入可求得：

同時，在兩群群落樣本同時為、以及分別為：

|  |  |
| --- | --- |
|  | (6) |
|  | (7) |
|  | (8) |

並依據式 (6) 與式 (7) 成立以下近似值：

將式 (4)、(5) 代入 後，並加入對估計式進行調整，最終得估計式：

其中、以及分別為：

|  |  |
| --- | --- |
|  | (9) |
|  | (10) |
|  | (11) |

在的基礎上，加入 的資訊對 的估計進行修正，依據式 (6) 與式 (7) 成立以下近似值：

並經由該式可推得出：

|  |  |
| --- | --- |
|  | (12) |

又經由可以得知，以及 分別為：

|  |  |
| --- | --- |
|  | (13) |
|  | (14) |

並依式 (13)與式 (14) 成立以下近似式：

並由上述式子推得：

|  |  |
| --- | --- |
|  | (15) |

又可從式 (12) = 式 (15) 得：

並依公式解，且皆大於0，故得：

|  |  |
| --- | --- |
|  | (16) |

最終得估計式為：

其中，將式 (16) 的與 的結果分別帶入 與 中可分別求得 與 的估計式；而使用式 (11) 估計。