

爲了DP法的使用，我們將原始資料（表5）整理成(1.1)卜瓦松模型的型式，令 $Y_1, \dots, Y_{57}$ 的值爲0， $Y_{58}, \dots, Y_{63}$ 的值爲1，以此類推， $Y_{99} = 9$ ， $Y_{100} = 10$ ；而相對應的估計值 $\hat{\lambda}_1, \dots, \hat{\lambda}_{57}$ 爲0.12， $\hat{\lambda}_{58}, \dots, \hat{\lambda}_{63}$ 爲2.56，以此類推， $\hat{\lambda}_{99} = 5.82$ ， $\hat{\lambda}_{100} = 6.42$ 。

有了每個樣區的估計均數，我們利用混和卜瓦松分配估計區域內海石花的個數分配，型式如下

$$\hat{f}(x) = \sum_{i=1}^{10} w_i \cdot e^{-\lambda_i^*} \frac{\lambda_i^{*x}}{x!}, \quad x = 0, 1, 2, \dots, \quad (5.1)$$

其中， $\mathbf{w} = (0.57, 0.06, 0.12, 0.05, 0.05, 0.05, 0.07, 0.01, 0.01, 0.01)$ ，

$\boldsymbol{\lambda}^* = (0.12, 2.56, 2.98, 3.38, 3.76, 4.12, 4.46, 4.83, 5.82, 6.42)$ 。

利用(5.1)式，重複抽取1000組樣本，每組100個觀察值，次數平均結果如表5的「估計分配的理論次數」。在虛無假設 $H_0$ 表示觀察值服從(5.1)式分配下，卡方配適度檢定(goodness-of-fit test)統計量爲

$$\begin{aligned} \psi &= \left[ \frac{57^2}{52.04} + \frac{6^2}{10.95} + \dots + \frac{1^2}{0.18} \right] - 100 \\ &= 15.122 < 19.675 = \chi_{0.95}^2(11), \end{aligned} \quad (5.2)$$

若顯著水準爲95%，我們不拒絕虛無假設，從而推論區域內海石花的個數分配爲(5.1)之型式。

## 6 結論

1. 從模擬結果可知，不論 $G$ 爲何種型態，DP估計法總是介於NPML及PEB之間，並趨向其中較好的估計法。
2. 採用Dirichlet過程當成 $G$ 的先驗分配，爲一種特別的NPEB法則，不僅使 $\boldsymbol{\lambda}$ 的先驗分配不受限於某種分配型式，當 $G$ 偏向某種分配時，還可以藉由 $G_0$ 的選擇，融入此資訊。

3. 本文在模擬研究時，假定我們不知 $\lambda$ 的型式，對於Dirichlet過程參數 $G_0$ ，採用均勻分配，但若我們知道 $\lambda$ 為伽瑪型式的混合分配，參數 $G_0$ 可採用伽瑪型式，也就是先利用PEB估計 $G_0$ ，再進行DP估計。
4. 如本文的實例探討，Dirichlet過程提供了一個 $\lambda$ 混和分配的估計法。

## 參考書目

- Antoniak, C. E. (1974). "Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems," *The Annals of Statistics*, **2**, 1152-1174.
- Archibald, E. E. A. (1948). "Plant Populations I. A New Application of Neyman's Contagious Distribution," *Annals of Botany*, **12**, 221-235.
- Blackwell, D., and MacQueen, J.B. (1973). "Ferguson Distribution via Polya Urn Schemes," *The Annals of Statistics*, **1**, 353-355.
- Escobar, M. D. (1994). "Estimating Normal Means with a Dirichlet Process Prior," *Journal of the American Statistical Association*, **89**, 268-277.
- Escobar, M. D. (1995). "Nonparametric Bayesian Methods in Hierarchical Models," *Journal of Statistical Planning and Inference*, **43**, 97-106.
- Ferguson, T. S. (1973). "A Bayesian Analysis of Some Nonparametric Problems," *The Annals of Statistics*, **1**, 209-230.
- Gelfand, A. E., and Smith, A. F. M. (1990). "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, **85**, 398-409.
- Laird, N. M. (1978). "Nonparametric Maximum Likelihood Estimation of a Mixing Distribution," *Journal of the American Statistical Association*, **73**, 805-811.