利用相同方法分別算出其餘15組$l\{(\alpha_i, G_j) \mid \boldsymbol{Y}\}$的估計值，配合(3.2)式相同權重的先驗分配，便可得到(3.4)式離散型後驗分配的估計。有了此離散型後驗分配，我們對第2節的吉氏取樣稍做調整，在執行步驟2.之前，先依照(3.4)式$(\alpha_0, G_0)$的後驗機率分配抽取一組$(\alpha_0, G_0)$，在此組$(\alpha_0, G_0)$值下，再進行步驟2.的過程，最後得到一組$(\lambda_1, \ldots, \lambda_n)$。

直接由觀察值$\boldsymbol{Y}$將$\boldsymbol{\lambda}$分群並不容易，只給$G_0$唯一型式也非易事，之前我們想找一組最佳的$(\alpha_0, G_0)$來執行吉氏取樣，顯得不太可行。$(\alpha_0, G_0)$的離散型後驗分配讓此兩個重要參數在決定上有更多選擇，且幫我們對決定的參數做一檢視，當某一組$(\alpha_0, G_0)$的後驗分配機率太小，甚至接近零，我們可猜想那是不適合的。

# 4 無母數經驗貝氏估計法和母數經驗貝氏估計法

本節將以往估計卜瓦松均數$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$的方法，做一整理介紹，在第5節的模擬研究，我們會拿這些估計方法和本文提出的Dirichlet過程做比較。

## 4.1 Robbins的估計法

Robbin's (1955)提供了一個在$G$不知道的情況下估計卜瓦松均數的方法，他將(1.2)式中的$m(Y_i)$、$m(Y_i+1)$，以$\boldsymbol{Y}$中等同於$Y_i$值及$Y_i+1$值的個數來做估計，即

$$\hat{\lambda}_i = (Y_i + 1)\frac{\sharp\{Y_k : Y_k = Y_i + 1, k = 1, \ldots, n\}}{\sharp\{Y_k : Y_k = Y_i, k = 1, \ldots, n\}}\text{。} \tag{4.1}$$

並證明，當$n \to \infty$，此估計式的貝氏風險會趨近$G$爲已知型式下的貝氏風險。然而，在下一節的模擬，我們可以看出，此估計法的估計並不理想。

## 4.2 無母數最大概似估計法

在(1.1)卜瓦松模型

$$Y_i \mid \lambda_i \sim Poisson(\lambda_i)\text{，} \qquad i = 1, \ldots, n\text{，}$$
$$\lambda_i \sim G\text{。}$$

之下，觀察值$\boldsymbol{Y}$的邊際分配爲

$$m(\boldsymbol{Y}) = \prod_{i=1}^{k} m_i(Y_i)\text{，} \tag{4.2}$$

其中$m_i(Y_i) = \int f(Y_i \mid \lambda_i) dG(\lambda_i)$。

Laird (1978)證明使(4.2)有最大值的$G$可用少於或等於$n$個退化點分配的混合型式來表現，即使$G$是一個連續的分配。以下介紹無母數最大概似估計法（NPML）估計$G$的方法。

步驟1. 設定$\boldsymbol{\lambda}^{(0)} = (\lambda_1^{(0)}, \ldots, \lambda_n^{(0)}) = (Y_1, \ldots, Y_n)$和其對應機率$\boldsymbol{\pi}^{(0)} = (\pi_1^{(0)}, \ldots, \pi_n^{(0)})$ $= (\frac{1}{n}, \ldots, \frac{1}{n})$，當成估計先驗分配$\hat{G}$的初始值。

步驟2. 對於每一個$i$，$i = 1, \ldots, n$，

$$w_{ij}^{(1)} \propto \pi_j^{(0)} f(Y_i \mid \lambda_j^{(0)}) = \pi_j^{(0)} e^{-\lambda_j^{(0)}} \frac{\lambda_j^{Y_i}}{Y_i!}\text{，} \qquad j = 1, \ldots, n\text{，}$$

且 $\sum_{j=1}^{n} w_{ij} = 1$。

步驟3. 更新$\boldsymbol{\lambda}$及$\boldsymbol{\pi}$。

$$\pi_j^{(1)} = w_{+j}^{(1)}\text{，} \qquad j = 1, \ldots, n\text{，}$$
$$\lambda_j^{(1)} = \frac{\sum_{i=1}^{n} w_{+j}^{(1)} Y_i}{\sum_{i=1}^{n} w_{ij}^{(1)}}\text{，} \qquad j = 1, \ldots, n\text{。}$$

重複$m$次步驟2、步驟3後，我們得到$\boldsymbol{\lambda}^{(m)} = (\lambda_1^{(m)}, \ldots, \lambda_n^{(m)})$及$\boldsymbol{\pi}^{(m)} = (\pi_1^{(m)}, \ldots, \pi_n^{(m)})$。當$m \to \infty$，$\boldsymbol{\lambda}^{(m)} = (\lambda_1^{(m)}, \ldots, \lambda_n^{(m)})$的值會收斂到幾個特定的值，記爲$\lambda_1^*, \ldots, \lambda_k^*$，$(k \leq n)$，且將其相對應的機率收集起來，記爲$\pi_1^*, \ldots, \pi_k^*$，以此當作先驗分配$G$的估

計，即

$$\hat{g}(\lambda_i) = \begin{cases} \pi_j^*\text{，} & \lambda_i = \lambda_j^*\text{，}j = 1\ldots,k\text{，} \\ 0\text{，} & \text{otherwise。} \end{cases}$$

有了估計先驗分配$\hat{G}$，$\lambda_i$的後驗分配爲

$$h(\lambda_i \mid Y_i) \propto l(\lambda_i \mid Y_i) \cdot \hat{g}(\lambda_j) = \begin{cases} e^{-\lambda_i}\frac{\lambda_i^{Y_i}}{Y_i!} \cdot \pi_j^*\text{，} & \lambda_i = \lambda_j^*\text{，}j = 1\ldots,k\text{，} \\ 0\text{，} & \text{otherwise。} \end{cases} \tag{4.3}$$

在損失函數$L(\hat{\lambda_i}, \lambda_i) = (\hat{\lambda_i} - \lambda_i)^2$之下，$\lambda_i$的貝氏估計量爲後驗分配(4.3)式的期望值

$$\hat{\lambda}_i^{NPEB} = \frac{1}{A}\sum_{j=1}^{k} \lambda_j^* \left[ e^{-\lambda_j^*}\frac{\lambda_j^{*Y_i}}{Y_i!} \cdot \pi_j^* \right]\text{，} \tag{4.4}$$

其中$A = \sum_{j=1}^{k} e^{-\lambda_j^*}\frac{\lambda_j^{*Y_i}}{Y_i!} \cdot \pi_j^*$。

## 4.3 母數經驗貝氏估計法

如前所述，假設分配$G$的型式爲已知，只是參數未定，以觀察值$\boldsymbol{Y}$來估計未知參數進而得到$G$的估計分配，稱爲母數經驗貝氏估計法（PEB）。在卜瓦松的模型下，爲了計算的可行性，我們往往假設$G$爲卜瓦松的共軛分配，亦即伽瑪分配，在此假設之下，(1.1)式的卜瓦松模型有下列型式：

$$\begin{aligned} Y_i \mid \lambda_i &\sim Poisson(\lambda_i)\text{，} \quad i = 1, \ldots, n\text{，} \\ \lambda_i &\sim \Gamma(a, b)\text{，} \end{aligned} \tag{4.5}$$

運用條件期望值和條件變異數的觀念，我們可以找到伽瑪分配參數$a$、$b$和隨機變數$Y_i$的關係，即

$$\begin{aligned} &E(Y_i) = E[E(Y_i \mid \lambda_i)] = E[\lambda_i] = ab\text{，} \\ &Var(Y_i) = Var[E(Y_i \mid \lambda_i)] + E[Var(Y_i \mid \lambda_i)] = Var[\lambda_i] + E[\lambda_i] = ab^2 + ab\text{。} \end{aligned} \tag{4.6}$$

當我們得到觀察值$\boldsymbol{Y}$後，便可利用觀察值的均數$\overline{Y}$和變異數$S^2$，估計參數$a$、$b$。

令$\overline{Y} = E(Y_i)$、$S^2 = Var(Y_i)$，我們得到

$$\hat{b} = \frac{S^2 - \overline{Y}}{\overline{Y}} \quad \text{且} \quad \hat{a} = \frac{\overline{Y}}{\hat{b}}\text{。}$$

有了$\lambda_i$的估計先驗分配$\Gamma(\hat{a},\hat{b})$，接著再利用貝氏定理，我們得到$\lambda_i$的後驗分配

$$\begin{aligned} h(\lambda_i \mid Y_i) &\propto l(\lambda_i \mid Y_i)\cdot g(\lambda_i) \\ &\propto e^{-\lambda_i}\frac{\lambda_i^{Y_i}}{Y_i!}\cdot\frac{\lambda_i^{\hat{a}-1}e^{-\lambda_i/\hat{b}}}{\Gamma(\hat{a})\hat{b}^{\hat{a}}} \\ &\propto \lambda_i^{(\hat{a}+Y_i)-1}e^{-\lambda_i/(\frac{\hat{b}}{1+\hat{b}})} \\ &\equiv \Gamma\left(\hat{a}+Y_i,\tfrac{\hat{b}}{1+\hat{b}}\right)\text{。} \end{aligned} \tag{4.7}$$

同樣地，我們求(4.7)式的期望值，並將之當成$\lambda_i$的貝氏估計量

$$\hat{\lambda}_i^{PEB}=E(\lambda_i \mid Y_i)=\tfrac{(\hat{a}+Y_i)\hat{b}}{1+\hat{b}}=\tfrac{\hat{a}\hat{b}}{1+\hat{b}}+\tfrac{Y_i\hat{b}}{1+\hat{b}}=\tfrac{1}{1+\hat{b}}\overline{Y}+\tfrac{\hat{b}}{1+\hat{b}}Y_i\text{。} \tag{4.8}$$

由(4.8)式可知，$\lambda_i$的PEB估計量爲所有觀察值平均$\overline{Y}$和$Y_i$加權平均的結果，其權重分別爲$1/(1+\hat{b})$、$\hat{b}/(1+\hat{b})$，且和爲1。當資料$\boldsymbol{Y}$很集中時，變異數$S^2$小，$\hat{b}$跟著小，$1/(1+\hat{b})$相對較大，此時用較多其他觀察值的資訊來幫助估計$\lambda_i$；當資料$\boldsymbol{Y}$較分散時，$1/(1+\hat{b})$相對較小，此時相信$\lambda_i$應該和$Y_i$以外的觀察值關係不大。

# 5 模擬研究和實例探討

## 5.1 模擬研究

從已知分配$G$中抽取$\lambda_i$，再以此$\lambda_i$當成卜瓦松的參數，抽取$Y_i$值，其中$i=1,\ldots,n$。爲了不讓估計方法受限於特定的觀察值，我們取$n=15$、30、50三種樣本大小，且分別討論$G$爲「伽瑪分配」及「特定伯努力分配」的狀況，型式如下：

表2. $G$為伽瑪分配以及特定伯努力分配的機率密度函數

| $G$爲伽瑪分配 | $G$爲特定伯努力分配 |
|---|---|
| $f(\lambda_i)=\dfrac{\lambda_i^{1-1}e^{-\lambda_i/b}}{\Gamma(1)b^1}$，<br>其中，$b=1$、2、3.16。 | $P(\lambda_i)=\begin{cases}0.8\text{，} & \lambda_i=0.5\text{，}\\ 0.2\text{，} & \lambda_i=c\text{，}\\ 0\text{，} & \text{otherwise，}\end{cases}$<br>其中，$c=1.5$、5.5、10.5。 |