

第一章 緒論

1.1 研究動機

在公共衛生的研究中，長期二元資料時常被觀察。長期二元資料，觀察得到的結果只有兩種狀況，舉例來說，觀察一個病人發病的『是』或『否』，或者經濟上，股市曲線『升』或『降』，在分析方面，習慣以符號 1 或 0 表示兩個狀態。在這邊討論的狀態都不是吸收狀態，就是兩個狀態都會反覆出現，以 1 或 0 來記錄。一般統計期刊所提出的分析長期二元資料時，會有兩點特徵。第一、觀察到的現象是兩個，需要 link function。第二、每一個個體會被不斷地觀察，不同個體之間是獨立的。

在分析這種資料結構，現在最常見的方法是 Generalized Estimating Equations (GEE) 和 Generalized Linear Mixed Model (GLMM)。^[1]

以遞移模型來說，除了知道解釋變數之值，還可以觀察到每個時間點的出象。^{[2][3]}在詹文耀博士與其學生的研究中，談及用卜瓦松過程來建構遞移模型，我們進一步延伸，讓遞移模型更具彈性。因此，在此篇論文中，對於長期二元資料，藉由連續時間馬可夫鍊的轉移機率，提出一個新遞移模型，進一步地，估計此模型的參數。

1.2 研究方法

這篇論文主要是提出一個新的遞移模型來分析長期二元資料。假設資料型態為 $Y(t) = (-1)^{N(t)}$ ，這是我們在 t 時間所觀察到的現象 $Y(t)$ ，即觀察到 -1 或 1，當，而 $Y(t)$ 完全被一個無法觀測到的隱藏隨機變數 $N(t)$ 所決定，假設 $N(t)$ 為齊一性卜瓦松過程，其發生率為 λ 。舉例來說，如圖：



(不失一般性，令 $t_0 = 0$)

在上圖中，有四個觀察時間點，分別是 t_0 、 t_1 、 t_2 、 t_3 ，而在這四個時間點分別是觀察到 1、-1、-1、1，如果用 $Y(t)$ 來表示的話，就是 $Y(t_0) = 1$ 、 $Y(t_1) = -1$ 、 $Y(t_2) = -1$ 、 $Y(t_3) = 1$ ，然而在每個時間點之間，可能有數次變化，也有可能沒有變化，例如：

情況一、時間點 t_0 到 t_1 之間，我們觀察到的現象是由 1 變成 -1。這段過程中，可能

變化有 $1 \rightarrow -1$ 一次變化； $1 \rightarrow -1 \rightarrow 1 \rightarrow -1$ 三次變化；

$1 \rightarrow -1 \rightarrow 1 \rightarrow -1 \rightarrow 1 \rightarrow -1$ 五次變化；... 等等，雖然中間的變化無法得知，但是可以發現到變化的次數是奇數次。

情況二、時間點 t_1 到 t_2 之間，我們觀察到的現象是由 -1 維持 -1。這段過程中，可

能沒有變化 -1 維持 -1，或者可能有其他變化 $-1 \rightarrow 1 \rightarrow -1$ 兩次變化；

$-1 \rightarrow 1 \rightarrow -1 \rightarrow 1 \rightarrow -1$ 四次變化；... 等等，雖然中間的變化依舊無法確定，但是可以發現到變化的次數是偶數次

為了方便討論，令 $X(t) = \frac{Y(t)+1}{2}$ ，則當 $Y(t)=1$ 時，此時 $X(t)=1$ ，而當 $Y(t)=-1$ 時，此時 $X(t)=0$ ，我們將觀察到的狀態轉成 0/1。

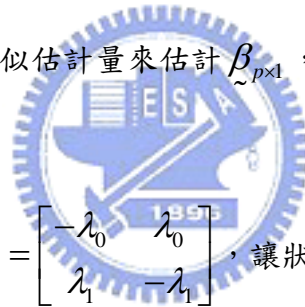
介紹完資料的特性後，因為卜瓦松過程就是最簡單的連續時間馬可夫鍊，因此用連續時間馬可夫鍊來建構新的遞移模型，以便推廣，加以應用。

首先，從最簡單的模型著手討論，由前述 $Y(t) = (-1)^{N(t)}$ ， $N(t)$ 為卜瓦松過程，其發生率為 λ ，經過轉換後，可以得到齊一性馬可夫鍊，其狀態是經過轉換後的 $X(t)$ ，

為 0/1，其微矩陣為 $Q(t) = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$ ，這裡考量令 $\lambda = \exp\{\tilde{x}'\tilde{\beta}\}$ ，其中 $\tilde{x}_{p \times 1}$ 為影響

發生率的因子，是可以測量到的值，而 $\tilde{\beta}_{p \times 1}$ 是每一個因子的係數向量，我們無法測量

到 $\tilde{\beta}_{p \times 1}$ 值，但是透過最大概似估計量來估計 $\tilde{\beta}_{p \times 1}$ ，獲得其估計值。



再來是考慮微矩陣 $Q(t) = \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix}$ ，讓狀態 0 跳到狀態 1 的發生率及狀態 1

跳到狀態 0 的發生率不同，此時可以考量令 $\lambda_0 = \exp\{\tilde{x}_0'\tilde{\beta}_0\}$ ，其中 \tilde{x}_0 為影響發生率 λ_0

的因子，為 $p \times 1$ 向量， \tilde{x}_0 是可以測量到的值，而 $\tilde{\beta}_0$ 是每一個因子的係數向量，令

$\lambda_1 = \exp\{\tilde{x}_1'\tilde{\beta}_1\}$ ，其中 \tilde{x}_1 為影響發生率 λ_1 的因子，為 $q \times 1$ 向量， \tilde{x}_1 也是可以測量到的

值，而 $\tilde{\beta}_1$ 是每一個因子的係數向量，無法直接測量到 $\tilde{\beta}_0$ 與 $\tilde{\beta}_1$ 之值，但透過最大概似估

計量來估計 $\tilde{\beta}_0$ 與 $\tilde{\beta}_1$ 之值，得到其估計值。

接著討論 $N(t)$ 為非齊性卜瓦松過程，其密度函數為 $\{\lambda(t), t \geq 0\}$ ，得到更有彈性的模型，但是數學計算方面也複雜許多，前述兩種模型都是齊一性連續時間馬可夫鍊，而假設 $N(t)$ 為非齊性卜瓦松過程，密度函數為 $\{\lambda(t), t \geq 0\}$ 後，可得到非齊性連續時間

馬可夫鍊，其狀態依舊為 0/1，微矩陣為 $Q(t) = \begin{bmatrix} -\lambda(t) & \lambda(t) \\ \lambda(t) & -\lambda(t) \end{bmatrix}$ ，這裡變成考慮

$\lambda(t) = f(t, \tilde{x}'\beta)$ ，仍然利用最大概似估計量來估計 $\beta_{p \times 1}$ ，得其估計值 $\hat{\beta}_{p \times 1}$ 。

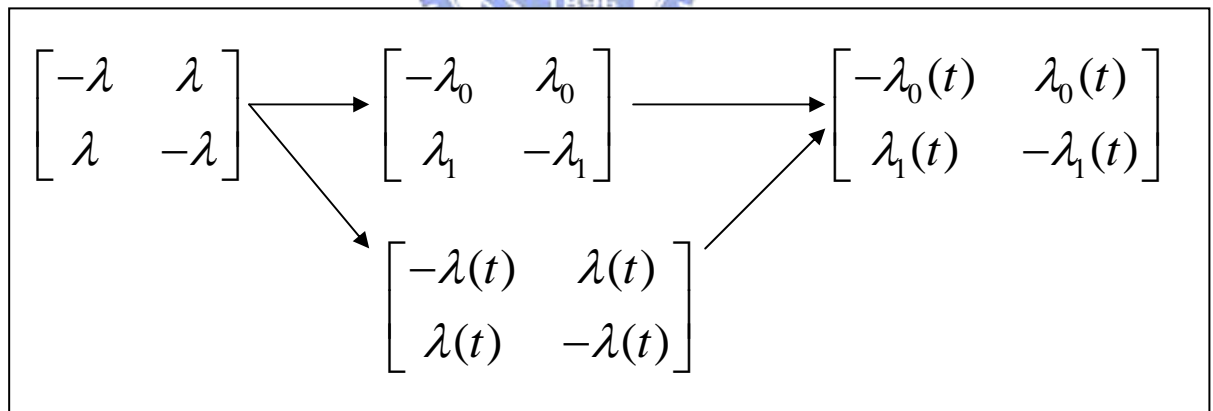
最後討論，最複雜的模型，我們將模型延伸到非齊性連續時間馬可夫鍊，當微矩

陣 $Q(t) = \begin{bmatrix} -\lambda_0(t) & \lambda_0(t) \\ \lambda_1(t) & -\lambda_1(t) \end{bmatrix}$ ，再令 $\lambda_0(t) = f(t, \tilde{x}_0'\beta_0)$ 及 $\lambda_1(t) = f(t, \tilde{x}_1'\beta_1)$ ，用最大

概似估計量來估計 β_0 及 β_1 。

我們從最簡單的齊一性卜瓦松過程，將模型延伸到非齊性連續時間馬可夫鍊，是本
篇論文最主要的貢獻。

用簡表說明流程：



第二章 介紹卜瓦松過程、馬可夫鏈與模擬


在這章節中，我們先介紹什麼是非齊性卜瓦松過程與非齊性連續時間馬可夫鏈，及如何去模擬非齊性卜瓦松過程與非齊性連續時間馬可夫鏈。

2.1 非齊性卜瓦松過程與非齊性連續時間馬可夫鏈

非齊性卜瓦松過程是最簡單的非齊性連續時間馬可夫鏈。在這一節當中，我們介紹非齊性卜瓦松過程與非齊性連續時間馬可夫鏈，因為經由瞭解非齊性卜瓦松過程，可以讓比較容易學習非齊性連續時間馬可夫鏈。

在介紹之前，先由多數人熟知的齊一性卜瓦松過程與齊一性連續時間馬可夫鏈開始。假設齊一性卜瓦松過程的發生率為 λ ，即當我們觀察發生次數時，在一段時間內，第0次到第1次的發生率為 λ ，第1次到第2次的發生率為 λ ，第2次到第3次的發生率為 λ ， \dots ，第 i 次到第 $i+1$ 次的發生率為 λ ，依此類推，這是發生率為 λ 的齊一性卜

瓦松過程，當我們將其以矩陣來表示時，此矩陣為


$$\begin{bmatrix} -\lambda & \lambda & 0 & \dots & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & & & \end{bmatrix}, \text{ 就很容}$$

易發現其實這是齊一性連續時間馬可夫鏈的一個例子。

說明完齊一性卜瓦松過程與齊一性連續時間馬可夫鏈的關係後，接著來談非齊性卜瓦松過程。

如果有一個計數過程 $N = \{N(t), t \geq 0\}$ 被稱為『具有密度函數為 $\{\lambda(t), t \geq 0\}$ 的非齊性卜瓦松過程』。有四個條件必須成立，第一點、在時間0的發生次數必須為0次。第二點、具備獨立增量性質，即沒有交集的時間區段，各別的發生次數是互相獨立的。

第三點、在很短的時間 h ($h \rightarrow 0$) 內，發生 1 次的機率為 $\lambda(t)h$ ，即

$P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$ 。第四點、在很短的時間 h ($h \rightarrow 0$) 內，發生 2 次以上的機率幾乎是零，可以用 $P\{N(t+h) - N(t) \geq 2\} = o(h)$ 來表示。

與齊一性卜瓦松過程比較，我們知道缺少穩定增量性質，因為原本的發生率是常數，非齊性卜瓦松過程變成跟時間有關的密度函數。因此，非齊性卜瓦松過程又稱為非穩定(nonstationary)卜瓦松過程。

在 2.1 節最後考慮一個連續時間隨機過程 $X = \{X(t), t \geq 0\}$ ，其狀態空間為 $S = \{0, 1, 2, \dots\}$ 。如果隨機過程 X 是一個非齊性連續時間馬可夫鍊，則此轉移機率 $P(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n)$ 是 $t_n, i_n, t_{n+1}, i_{n+1}$ 的函數，其中 $0 \leq t_n < t_{n+1}$ ， $i_n, i_{n+1} \in S$ 。

在此以符號 $p_{i_n, i_{n+1}}(t_n, t_{n+1})$ 來表示 $P(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n)$ 表示，即

$p_{i_n, i_{n+1}}(t_n, t_{n+1}) = P(X(t_{n+1}) = i_{n+1} | X(t_n) = i_n)$ 。我們也可以用矩陣來表示，對於任何 $0 \leq s < t$ ，轉移矩陣 $P(s, t) = (p_{ij}(s, t))_{i, j \in S}$ ，此矩陣中的元素 $p_{ij}(s, t)$ 就是前述的

$p_{i_n, i_{n+1}}(t_n, t_{n+1})$ 。

根據[4]Chapman-Kolmogorov equation，在一段時間 $[s, t]$ 內，我們可以任意找一點 u ，然後變成兩個區間，分別是 $[s, u]$ 和 $[u, t]$ ，得到 $p_{ik}(s, u)$ 和 $p_{kj}(u, t)$ ， k 是狀態空間中的一個狀態，先把 $p_{ik}(s, u)$ 和 $p_{kj}(u, t)$ 乘起來，再把所有 k 可能出現的狀態加起來，即是 $p_{ij}(s, t) = \sum_{k \in S} p_{ik}(s, u) p_{kj}(u, t)$ ，其中 $0 \leq s < u < t$ ， $i, j \in S$ 。同時也可以用矩陣 $P(s, t) = P(s, u)P(u, t)$ 來表示，於是 $P(s, s) = I$ ，對於 $0 \leq s \leq u \leq t$ 皆成立。

令 $p_{ij}(s, s) = q_{ij}$ ， $s \geq 0$ ， $i, j \in S$ 。

考慮微矩陣 $Q(t) = (q_{ij}(t))_{i,j \in S}$ ，有三點要注意，第一點，非對角線元素必須是非負的。第二點、對角線元素必須是負的。第三點、每一列元素的總和必須是零。

再假設 $\int_0^t q_{ii}(u)du < \infty$ ，對於所有 $t \geq 0$ ， $i \in S$ 。在這些假設下，我們可以找到唯一的轉移機率矩陣 P 使得

$$\lim_{\substack{h+h' \rightarrow 0^+ \\ (h=0 \text{ or } h'=0)}} \frac{P(t-h, t+h') - I}{h+h'} = Q(t)$$

進一步，得到

$$(1) P(s, t) = I + \int_s^t Q(x)P(x, t)dx \quad \text{或} \quad P(s, t) = I + \int_s^t P(s, x)Q(x)dx$$

然後也可以得到 backward 和 forward Kolmogorov equations

$$\begin{aligned} \frac{\partial P(s, t)}{\partial s} &= -Q(s)P(s, t) & 0 \leq s \leq t; \\ \frac{\partial P(s, t)}{\partial t} &= P(s, t)Q(t) & 0 \leq s \leq t. \end{aligned}$$

因為是非齊性連續時間馬可夫鏈，所以在計算轉移機率的時候，要小心時間點，時間的起始點或結束點不同的話，其轉移機率就會不同。

2.2 模擬非齊性卜瓦松過程的到達時間與二狀態馬可夫鏈

先介紹如何模擬生成非齊性卜瓦松過程的到達時間，再藉由非齊性卜瓦松過程的到達時間來模擬生成二狀態非齊性連續時間馬可夫鏈。談模擬生成非齊性卜瓦松過程的到達時間之前，必須先會生成齊一性卜瓦松過程的到達時間，再用[4]thinning方法得到非齊性卜瓦松過程的到達時間。這邊介紹了兩個方法來生成非齊性卜瓦松過程的到達時間，最後才談如何模擬生成二狀態非齊性連續時間馬可夫鏈。

為方便起見，除了介紹如何模擬生成卜瓦松過程的到達時間與二狀態馬可夫鏈，我們將演算法整理出來，列在下方。

2.2.1 模擬卜瓦松過程的到達時間

假設我們想要生成一段時間 T 內的卜瓦松過程的到達時間，此卜瓦松過程的發生率為 λ 。模擬的方法很簡單，先令 $S_0 = 0$ ，兩個發生點（假設為 S_i 與 S_{i+1} ）中間的時間長度（ $S_{i+1} - S_i$ ）為指數分配，期望值為 $\frac{1}{\lambda}$ ，就是表示 $S_{i+1} - S_i \sim Exp(\lambda)$ ，然後確定每段時間都是獨立的指數分配，期望值皆為 $\frac{1}{\lambda}$ 。收集直到 $S_{n+1} > T$ ，如此一來，就可以模擬時間 T 內的卜瓦松過程的到達時間數列 $\{S_i; i = 1, 2, \dots, n\}$ 。同時也可以得到時間 T 內，卜瓦松過程的到達次數為 n 。

模擬卜瓦松過程的到達時間演算法：

步驟 1、 $t = 0$ ， $i = 0$ 。

步驟 2、生成一個隨機變數 U ， $U \sim Uniform(0,1)$ 。

步驟 3、 $t = t - \frac{1}{\lambda} \log U$ 。如果 $t > T$ ，我們就停止模擬，如果 $t < T$ ，就進入步驟 4。

步驟 4、 $i = i + 1$ ， $S_i = t$ 。

步驟 5、回到步驟 2，繼續做下去。

2.2.2 模擬非齊性卜瓦松過程的到達時間

方法一、

若要模擬具有密度函數 $\lambda(t)$ 的非齊性卜瓦松過程，必須先找到 $\lambda(t)$ 的一個上界 λ ，使得 $\lambda(t) \leq \lambda$ 對於所有 $t \geq 0$ ，即先找到的 λ 值必須在每個時間點都要大於密度函數 $\lambda(t)$ 。然後接著生成發生率為 λ 的卜瓦松過程，收集發生時間點的數列 $\{S_i\}$ 。然後對於 $\{S_i\}$ ，每一個時間點取不取的方式用機率 $\lambda(S_i)/\lambda$ 來決定，將收集到的時間點用 $\{S'_i\}$ 來表示，則 $\{S'_i\}$ 即為具有密度函數 $\lambda(t)$ 的非齊性卜瓦松過程的到達時間。

模擬非齊性卜瓦松過程的到達時間演算法-方法一：

步驟 1、 $t = 0$ ， $i = 0$ 。

步驟 2、生成一個隨機變數 U_1 ， $U_1 \sim Uniform(0,1)$ 。

步驟 3、 $t = t - \frac{1}{\lambda} \log U_1$ 。如果 $t > T$ ，我們就停止模擬，如果 $t < T$ ，就進入步驟 4。

步驟 4、生成一個隨機變數 U_2 ， $U_2 \sim Uniform(0,1)$ ， U_1 與 U_2 互相獨立。

步驟 5、如果 $U_2 \leq \frac{\lambda(t)}{\lambda}$ ，則令 $i = i + 1$ ， $S'_i = t$ 。

步驟 6、回到步驟 2，繼續做下去。

方法二、

假設我們已經有密度函數 $\lambda(t)$ 的非齊性卜瓦松過程的到達時間點 $S_1 = s_1, \dots,$

$S_n = s_n$ 總共 n 個時間點，想要模擬生成下一個到達時間 S_{n+1} ，令 $\tau = S_{n+1} - S_n$ ，而 τ 的分配為

$$F_\tau(x | S_n = t) = 1 - \exp\{-[m(t+x) - m(t)]\} \quad x > 0$$

其中 $m(t) = \int_0^t \lambda(u) du$ 。

利用上式得到 τ 之後，就可推得 S_{n+1} 。

模擬非齊性卜瓦松過程的到達時間演算法-方法二：

步驟 1、生成第一個非齊性卜瓦松過程的到達時間點 $S_1 = s_1$ ，可以利用

方法一得到 s_1 。

步驟 2、 $S_1 = s_1, i = 1$ 。

步驟 3、生成一個隨機變數 $U, U \sim Uniform(0,1)$ ，然後令 $U = F_\tau(x | S_i = s_i)$ 。

求出 $x = F_\tau^{-1}(U | S_i = s_i)$ 。

步驟 4、 $t = s_i + x$ 。如果 $t > T$ ，我們就停止模擬，如果 $t < T$ ，就進入步驟 5。

步驟 5、 $i = i + 1, S_i = t$ 。

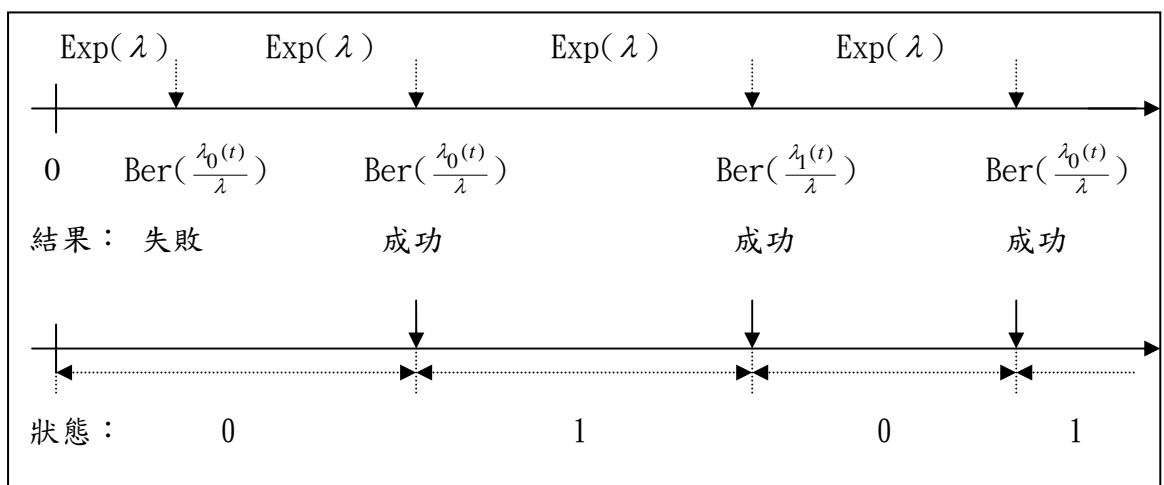
步驟 6、回到步驟 3，繼續做下去。

最後介紹如何用非齊性卜瓦松過程的到達時間來模擬生成二狀態非齊性連續時間馬可夫鏈。

2.2.3 模擬二狀態馬可夫鏈

模擬二狀態非齊性連續時間馬可夫鍊，利用非齊性卜瓦松過程的到達時間，先假定第一次的狀態為狀態 0 或狀態 1，如果是狀態 0 的話，就以狀態 0 的密度函數 $\lambda_0(t)$ 來模擬非齊性卜瓦松過程的到達時間，得到第一個狀態 0 的發生時間點，這時馬上跳到狀態 1，再以狀態 1 的密度函數 $\lambda_1(t)$ 來模擬非齊性卜瓦松過程的到達時間，相同地在第一個狀態 1 的發生時間點，又跳回狀態 0，依此步驟反覆做下去，就可以模擬生成二狀態非齊性連續時間馬可夫鏈。

程式模擬流程說明圖：



第三章 二狀態連續時間馬可夫鍊與遞移模型

先介紹非齊性連續時間馬可夫鍊的轉移機率如何求得，然後我們所需的是二狀態非齊性連續時間馬可夫鍊的轉移機率，因為必須有轉移機率，才能列出概似函數，進一步找出最大概似估計量。

3.1 轉移機率[6]

由 2.1 得到的 (1) 開始推導：

$$\begin{aligned} P(s,t) &= I + \int_s^t Q(x)P(x,t)dx \quad \text{其中，} I \text{ 是單位矩陣。} \\ &= I + \int_s^t Q(x)[I + \int_x^t Q(y)P(y,t)dy]dx \\ &= I + \int_s^t Q(x)dx + \int_s^t \int_x^t Q(x)Q(y)P(y,t)dydx \end{aligned}$$

持續上述過程，可以得到

$$\begin{aligned} (2) \quad P(s,t) &= I + \int_s^t Q(x)dx + \int_s^t \int_x^t Q(x)Q(y)dydx + \int_s^t \int_x^t \int_y^t Q(x)Q(y)Q(z)dzdydx + \dots \\ &= I + \int_s^t Q(x)dx + \int_s^t \int_s^x Q(y)Q(x)dydx + \int_s^t \int_s^x \int_s^y Q(z)Q(y)Q(x)dzdydx + \dots \end{aligned}$$

命題 1、

如果 $Q(t)$'s 具有可交換性， $Q(x)Q(y) = Q(y)Q(x)$ ，對於所有 x 和 y ，

這樣我們可以得到 $P(s,t) = \exp\{\int_s^t Q(x)dx\}$

如果 $Q(t)$'s 不具有可交換性，我們可以試著用下列方法，找出個別的轉移機率。

對於有限 n 個狀態馬可夫鍊，我們可以在 (2) 的兩邊同時乘上

向量 $(0, \dots, 1, \dots, 0)_{n \times 1}$ (第 i 個分量為 1) 和矩陣 A ，其中

$$A = \begin{bmatrix} 1 & \cdots & 1 & \cdots & 0 \\ 0 & \ddots & \vdots & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & \vdots & \ddots & 0 \\ 0 & \cdots & 1 & \cdots & 1 \end{bmatrix}_{n \times n}, \quad (\text{對角線及第 } i \text{ 行每個元素皆為 } 1), \text{ 計算後得}$$

$$A^{-1} = \begin{bmatrix} 1 & \cdots & -1 & \cdots & 0 \\ 0 & \ddots & \vdots & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & \vdots & \ddots & 0 \\ 0 & \cdots & -1 & \cdots & 1 \end{bmatrix}_{n \times n}, \quad (\text{對角線每個元素皆為 } 1 \text{ 及第 } i \text{ 行每個元素為 } -1, \text{ 對}$$

角線上的元素除外)。

$$(p_{i0}(s,t), p_{i1}(s,t), \dots, p_{i(i-1)}(s,t), 1, p_{i(i+1)}(s,t), \dots, p_{i(n-1)}(s,t)) \quad (\text{第 } i \text{ 個分量為 } 1)$$

$$= (0, \dots, 1, \dots, 0)P(s,t)A$$

$$= (0, \dots, 1, \dots, 0) \left[I + \int_s^t Q(x)dx + \int_s^t \int_x^t Q(x)Q(y)dydx + \dots \right] A$$

$$= (0, \dots, 1, \dots, 0) \left[I + \int_s^t Q(x)dx + \int_s^t \int_x^t Q(x)AA^{-1}Q(y)dydx + \dots \right] A$$

$$= (0, \dots, 1, \dots, 0) + \int_s^t (q_{i1}, q_{i2}, \dots, 0, \dots, q_{in})(x) \left[I + \int_x^t B(y)dy + \int_x^t \int_y^t B(y)B(z)dzdy + \dots \right] dx$$

$$\text{其中 } B(x) = A^{-1}Q(x)A = \begin{bmatrix} -q_{i0} + q_{00} & \cdots & 0 & \cdots & -q_{i(n-1)} + q_{01} \\ \vdots & & \vdots & & \vdots \\ q_{i0} & \cdots & 0 & \cdots & q_{i(n-1)} \\ \vdots & & \vdots & & \vdots \\ -q_{i0} + q_{(n-1)0} & \cdots & 0 & \cdots & -q_{i(n-1)} + q_{(n-1)(n-1)} \end{bmatrix} (x), \quad B(x)$$

的第 i 行每個元素皆為 0，接著定義 $B(x)$ 的子矩陣 $C_i(x)$ 如下：

$$C_i(x) = \begin{bmatrix} -q_{i0} + q_{00} & \cdots & -q_{i(i-1)} + q_{0(i-1)} & -q_{i(i+1)} + q_{0(i+1)} & \cdots & -q_{i(n-1)} + q_{0(n-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -q_{i0} + q_{(i-1)0} & \cdots & -q_{i(i-1)} + q_{(i-1)(i-1)} & -q_{i(i+1)} + q_{(i-1)(i+1)} & \cdots & -q_{i(n-1)} + q_{(i-1)(n-1)} \\ -q_{i0} + q_{(i+1)0} & \cdots & -q_{i(i-1)} + q_{(i+1)(i-1)} & -q_{i(i+1)} + q_{(i+1)(i+1)} & \cdots & -q_{i(n-1)} + q_{(i+1)(n-1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -q_{i0} + q_{(n-1)0} & \cdots & -q_{i(i-1)} + q_{(n-1)(i-1)} & -q_{i(i+1)} + q_{(n-1)(i+1)} & \cdots & -q_{i(n-1)} + q_{(n-1)(n-1)} \end{bmatrix} (x)$$

如果 $C_i(x)$ 滿足命題 1 的條件，則我們可得

$$\begin{aligned} & (p_{i0}(s,t), p_{i1}(s,t), \cdots, p_{i(i-1)}(s,t), p_{i(i+1)}(s,t), \cdots, p_{i(n-1)}(s,t)) \\ &= \int_s^t (q_{i0}(x), q_{i1}(x), \cdots, q_{i(i-1)}(x), q_{i(i+1)}(x), \cdots, q_{i(n-1)}(x)) \exp\left\{\int_x^t C_i(u) du\right\} dx \end{aligned}$$

$$\text{而 } p_{ii}(s,t) = 1 - \sum_{i \neq j} p_{ij}(s,t) \circ$$

3.1.1 二狀態馬可夫鍊的轉移機率

在這邊，我們需要用到一個特別的例子，二狀態馬可夫鍊 $Q(x) = \begin{bmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{bmatrix} (x)$ ，其中

$$q_{00} = -q_{01}, \quad q_{11} = -q_{10}, \quad C_0(x) = (-q_{01} + q_{11})(x) = -(q_{01} + q_{10})(x), \quad \text{而}$$

$$(3) \quad p_{01}(s,t) = \int_s^t q_{01}(x) \exp\left\{-\int_x^t (q_{01} + q_{10})(y) dy\right\} dx; \quad p_{00}(s,t) = 1 - p_{01}(s,t);$$

利用對稱性，可得

$$p_{10}(s,t) = \int_s^t q_{10}(x) \exp\left\{-\int_x^t (q_{10} + q_{01})(y) dy\right\} dx; \quad p_{11}(s,t) = 1 - p_{10}(s,t) \circ$$

3.2 用二狀態馬可夫鍊來建構新的遞移模型

我們利用二狀態馬可夫鍊來建構新的遞移模型，因為可以透過微矩陣 $Q(t)$ 來控制馬可夫鍊的行為，利用選取不同的 $Q(t)$ 來決定模型，而不同的 $Q(t)$ 就是其中的元素 q_{ij} 不同，選取不同的 q_{ij} 函數，再針對其函數中的參數 β 作估計。估計的方法，由 3.1 節求得的轉移機率 $p_{00}(s,t)$ 、 $p_{01}(s,t)$ 、 $p_{10}(s,t)$ 、 $p_{11}(s,t)$ ，寫出概似函數，然後求出最大概似估計量 $\hat{\beta}$ 。過程如下：

觀察得 $(n+1)$ 個樣本， $X(t_0), X(t_1), X(t_2), \dots, X(t_{n-1}), X(t_n)$ ，

$Q(t | \beta) = (q_{ij}(t | \beta))_{n \times n}$ ，在這邊我們用最大概似估計量來估計 β

$$L(\beta | X(t_0), X(t_1), X(t_2), \dots, X(t_{n-1}), X(t_n))$$

$$= p_{X(t_0), X(t_1)}(t_0, t_1) \times p_{X(t_1), X(t_2)}(t_1, t_2) \times \dots \times p_{X(t_{n-1}), X(t_n)}(t_{n-1}, t_n)$$

$$= \prod_{i=1}^n p_{x(t_{i-1}), x(t_i)}(t_{i-1}, t_i)$$

最大概似估計量 $\hat{\beta}$ 滿足 $\max L(\beta | X(t_0), X(t_1), X(t_2), \dots, X(t_{n-1}), X(t_n))$ 。

其中 $p_{X(t_{i-1}), X(t_i)}(t_{i-1}, t_i)$ 可用 (3) 去計算。

3.2.1 二狀態齊一性連續時間馬可夫鍊-1

先討論最基本的遞移模型，讓 $q_{01} = q_{10} = \lambda$ ，得到 $Q(t) = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$ ，這樣的微

矩陣 $Q(t)$ 代表，不管哪一個時間點，從狀態 0 跳到狀態 1 的發生率與從狀態 1 跳到狀態 0 的發生率是一樣的，可以透過新的遞移模型來估計 λ 。

$$\text{先算出 } p_{01}(s,t) = p_{10}(s,t) = \frac{1}{2}(1 - \exp\{-2\lambda(t-s)\})$$

$$p_{00}(s,t) = p_{11}(s,t) = \frac{1}{2}(1 + \exp\{-2\lambda(t-s)\})$$

發現 $p_{01}(s, t) = p_{10}(s, t)$ 、 $p_{00}(s, t) = p_{11}(s, t)$ ，於是做一個轉換，定義：

$\delta(\Delta t_i) = \delta(Y(t_i) - Y(t_{i-1})) \equiv Y(t_i) + Y(t_{i-1}) \pmod{2}$ ，則 $\delta(\Delta t_i)$ 為一伯努利分配，成

功的機率為 $p(\Delta t_i) = \frac{1}{2}(1 - \exp\{-2\lambda\Delta t_i\})$ ，其中 $\Delta t_i = t_i - t_{i-1}$ ：

$$\delta(\Delta t_i) = \begin{cases} 1 & \text{w.p. } p(\Delta t_i) = \frac{1}{2}(1 - \exp\{-2\lambda\Delta t_i\}); \\ 0 & \text{w.p. } 1 - p(\Delta t_i); \end{cases}$$

經由這個轉換後，可以很輕鬆地將概似函數寫出來

$$L = \prod_{i=1}^n \left[p(\Delta t_i)^{\delta(\Delta t_i)} (1 - p(\Delta t_i))^{1 - \delta(\Delta t_i)} \right]$$

對上式取對數函數， $l = \log L$ ，

$$\begin{aligned} l &= \sum_{i=1}^n \left[\delta(\Delta t_i) \log p(\Delta t_i) + (1 - \delta(\Delta t_i)) \log(1 - p(\Delta t_i)) \right] \\ &= \sum_{i=1}^n \left[\delta(\Delta t_i) \log \left[\frac{1}{2}(1 - \exp\{-2\lambda\Delta t_i\}) \right] + (1 - \delta(\Delta t_i)) \log \left[\frac{1}{2}(1 + \exp\{-2\lambda\Delta t_i\}) \right] \right] \end{aligned}$$

此時，如果假設 $\lambda = \exp\{\tilde{x}'\tilde{\beta}\}$ ，其中 $\tilde{x}_{p \times 1}$ 是會影響發生率 λ 的因子，帶入上式，可以得到

$$(4) \quad l = \sum_{i=1}^n \left[\delta(\Delta t_i) \log \left[\frac{1}{2}(1 - \exp\{-2e^{\tilde{x}'\tilde{\beta}}\Delta t_i\}) \right] + (1 - \delta(\Delta t_i)) \log \left[\frac{1}{2}(1 + \exp\{-2e^{\tilde{x}'\tilde{\beta}}\Delta t_i\}) \right] \right]$$

因此，我們可以透過上式用最大概似估計量來估計，得 $\hat{\tilde{\beta}}$ 。

3.2.2 二狀態齊一性連續時間馬可夫鍊-2

再來討論，當 $Q(t) = \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix}$ 時，這時候的發生率不是相同的，狀態 0 跳到狀

態 1 的發生率為 λ_0 ，狀態 1 跳到狀態 0 的發生率為 λ_1 ，可以算出

$$p_{01}(s, t) = \frac{\lambda_0}{\lambda_0 + \lambda_1} (1 - \exp\{-(\lambda_0 + \lambda_1)(t - s)\}) ;$$

$$p_{00}(s, t) = 1 - p_{01}(s, t) ;$$

$$p_{10}(s, t) = \frac{\lambda_1}{\lambda_0 + \lambda_1} (1 - \exp\{-(\lambda_0 + \lambda_1)(t - s)\}) ;$$

$$p_{11}(s, t) = 1 - p_{10}(s, t) ;$$

因為當 $Q(t) = \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix}$ 的情況下， $p_{01}(s, t)$ 、 $p_{00}(s, t)$ 、 $p_{10}(s, t)$ 、 $p_{11}(s, t)$ 四個

機率皆不同，所以無法有類似前一種狀況 $Q(t) = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix}$ 的作法，轉換成伯努利分

佈。

寫出其概似函數

$$L = \prod_{i=1}^n [p_{X(t_{i-1})X(t_i)}(t_{i-1}, t_i)]$$

對上式取對數函數，

$$l = \sum_{i=1}^n [\log(p_{X(t_{i-1})X(t_i)}(t_{i-1}, t_i))]]$$

再來假設 $\lambda_0 = \exp\{x_0' \beta_0\}$ ， $\lambda_1 = \exp\{x_1' \beta_1\}$ ，其中 x_0 是會影響發生率 λ_0 的因子， x_1

是會影響發生率 λ_1 的因子， x_0 為 $p \times 1$ 向量， x_1 為 $q \times 1$ 向量，帶入上式，一樣用最大

概似估計量來估計，得 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 。

3.2.3 二狀態非齊性連續時間馬可夫鍊

最後討論 $Q(t) = \begin{bmatrix} -\lambda_0(t) & \lambda_0(t) \\ \lambda_1(t) & -\lambda_1(t) \end{bmatrix}$ ，此時，模型複雜許多，狀態一跳到狀態二的

密度函數與狀態二跳到狀態一的密度函數不僅不相同，而且還分別是時間的函數。因此，計算轉移機率會複雜許多，列出其轉移機率

$$p_{01}(s,t) = \int_s^t \lambda_0(x) \exp\left\{-\int_x^t (\lambda_0 + \lambda_1)(y) dy\right\} dx ;$$

$$p_{00}(s,t) = 1 - p_{01}(s,t) ;$$

$$p_{10}(s,t) = \int_s^t \lambda_1(x) \exp\left\{-\int_x^t (\lambda_0 + \lambda_1)(y) dy\right\} dx ;$$

$$p_{11}(s,t) = 1 - p_{10}(s,t) ;$$

寫出其概似函數

$$L = \prod_{i=1}^n \left[p_{X(t_{i-1})X(t_i)}(t_{i-1}, t_i) \right]$$

對上式取對數函數，

$$l = \sum_{i=1}^n \left[\log(p_{X(t_{i-1})X(t_i)}(t_{i-1}, t_i)) \right]$$

再來假設 $\lambda_0(t) = f(t, \underline{x}_0' \underline{\beta}_0)$ ， $\lambda_1(t) = f(t, \underline{x}_1' \underline{\beta}_1)$ ，其中 \underline{x}_0 是會影響密度函數 $\lambda_0(t)$ 的

因子， \underline{x}_1 是會影響密度函數 $\lambda_1(t)$ 的因子， \underline{x}_0 為 $p \times 1$ 向量， \underline{x}_1 為 $q \times 1$ 向量，帶入上式，

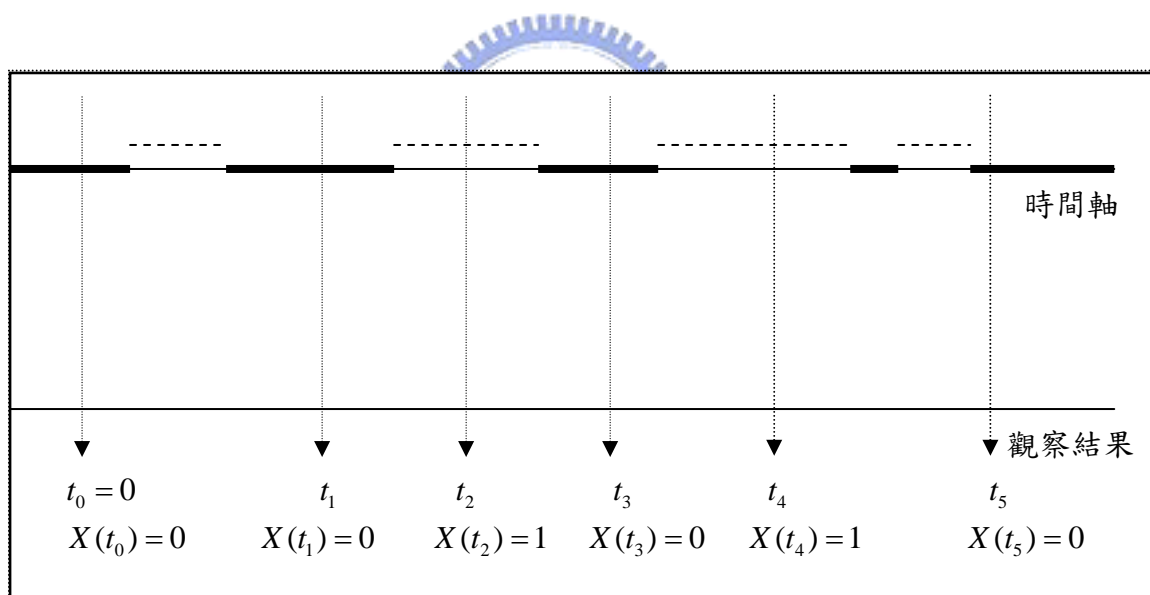
一樣用最大概似估計量來估計，得 $\hat{\underline{\beta}}_0$ 和 $\hat{\underline{\beta}}_1$ 。

第四章 模擬實驗

在這章節中，我們假設三種不同類型的微矩陣 $Q(t)$ ，透過電腦做些簡單的模擬實驗，先模擬實際的二元長期資料的狀態，然後觀察某些時間點 $\{t\}$ ，得到狀態樣本 $\{X(t)\}$ ，將時間點樣本與狀態樣本帶入欲估計參數 β 的遞移模型中，計算出欲估計參數 β 的最大概似估計量 $\hat{\beta}$ 。

請看舉例說明圖：

第一條線模擬實際的二元長期資料，粗體線表示狀態 0，虛線表示狀態 1，兩個狀態交替出現，箭頭部分即我們觀察所選取的時間點，可以得到狀態樣本。



4.1 微矩陣 $Q(t)$ 與時間無關

先模擬最簡單的例子，令 $q_{01} = q_{10} = \exp\{x\beta\}$ ，得到微矩陣 $Q_1(t)$

$$Q_1(t) = \begin{bmatrix} -\exp\{x\beta\} & \exp\{x\beta\} \\ \exp\{x\beta\} & -\exp\{x\beta\} \end{bmatrix}.$$

此微矩陣 $Q_1(t)$ 表示，不管哪一個時間點，從狀態 0 跳到狀態 1 的發生率與從狀態 1 跳到狀態 0 的發生率皆為 $\exp\{x\beta\}$ ，其發生率與時間無關，是常數，我們可以透過新的遞移模型來估計 β 。

在微矩陣 $Q_1(t)$ 情況下，由 (4) 得其對數概似函數為

$$l = \sum_{i=1}^n \left[\delta(\Delta t_i) \log \left[\frac{1}{2} (1 - \exp\{-2\exp(x\beta)\Delta t_i\}) \right] + (1 - \delta(\Delta t_i)) \log \left[\frac{1}{2} (1 + \exp\{-2\exp(x\beta)\Delta t_i\}) \right] \right]$$

對上式的 β 微分，得

$$\frac{d}{d\beta} l = \sum_{i=1}^n \left[\delta(\Delta t_i) \left(\frac{2x\Delta t_i \exp\{-2e^{x\beta}\Delta t_i + x\beta\}}{1 - \exp\{-2e^{x\beta}\Delta t_i\}} \right) + (1 - \delta(\Delta t_i)) \left(-\frac{2x\Delta t_i \exp\{-2e^{x\beta}\Delta t_i + x\beta\}}{1 + \exp\{-2e^{x\beta}\Delta t_i\}} \right) \right]$$

令 $\frac{d}{d\beta} l = 0$ ，用 [7] Newton-Rapbson 法，解出 β 的數值解。

對上式 $\frac{d}{d\beta} l$ 的 β 微分，得

$$\frac{d^2}{d\beta^2} l = \sum_{i=1}^n \left[\delta(\Delta t_i) \left(-\frac{4x^2(\Delta t_i)^2 \exp\{-4e^{x\beta}\Delta t_i + 2x\beta\}}{(1 - \exp\{-2e^{x\beta}\Delta t_i\})^2} + \frac{2x\Delta t_i(x - 2x\Delta t_i e^{x\beta}) \exp\{-2e^{x\beta}\Delta t_i + x\beta\}}{1 - \exp\{-2e^{x\beta}\Delta t_i\}} \right) \right. \\ \left. + (1 - \delta(\Delta t_i)) \left(-\frac{4x^2(\Delta t_i)^2 \exp\{-4e^{x\beta}\Delta t_i + 2x\beta\}}{(1 + \exp\{-2e^{x\beta}\Delta t_i\})^2} - \frac{2x\Delta t_i(x - 2x\Delta t_i e^{x\beta}) \exp\{-2e^{x\beta}\Delta t_i + x\beta\}}{1 + \exp\{-2e^{x\beta}\Delta t_i\}} \right) \right]$$

給初始值 β_0 之後，

將觀察值 $\{t, X(t)\}$ 帶入下式：

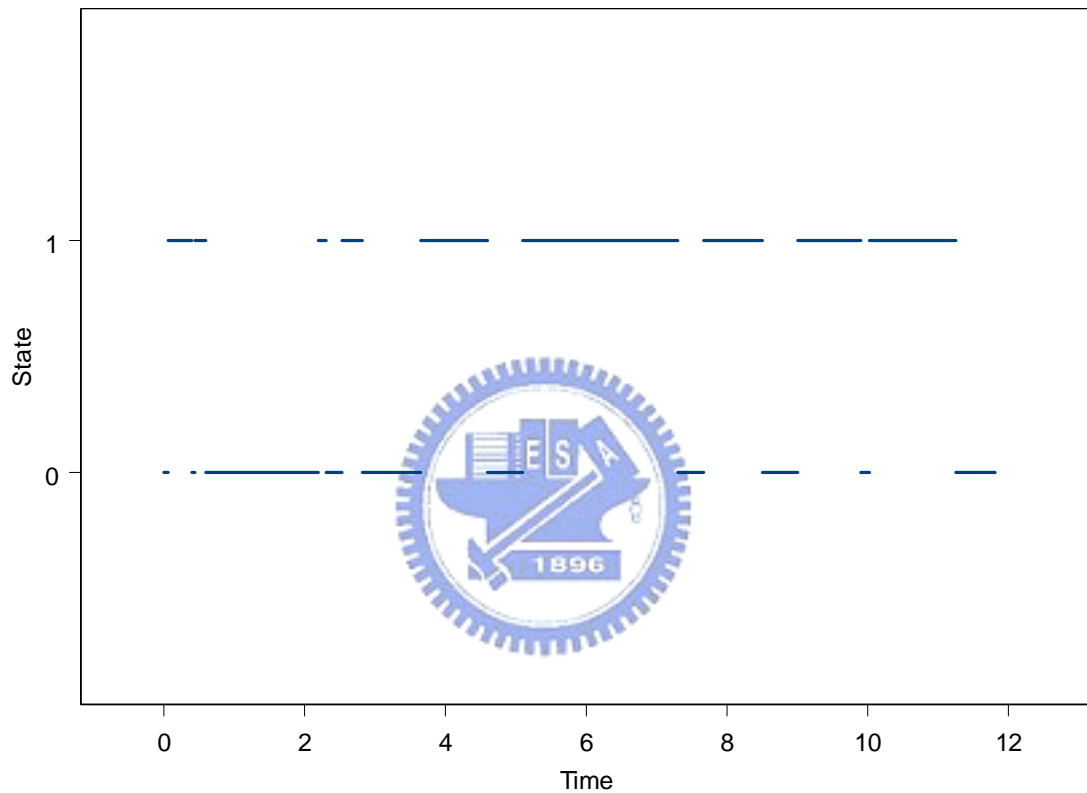
$$\beta_i = \beta_{i-1} - \left(\frac{d^2}{d\beta^2} l \right)^{-1} \left(\frac{d}{d\beta} l \right), \quad i = 1, 2, 3, \dots$$

透過不斷重複上述步驟，算出 β 的最大概似估計量的近似值 $\hat{\beta}$ 。

模擬結果：

模擬的設定為 $x = 3$ ， $\beta = 0.2$ ，則 $Q_1(t) = \begin{bmatrix} -e^{0.6} & e^{0.6} \\ e^{0.6} & -e^{0.6} \end{bmatrix}$ 。

模擬時間長度為 5000，總共模擬出 9194 次變化，為方便解釋，畫出前二十次變化來看，圖形如下：



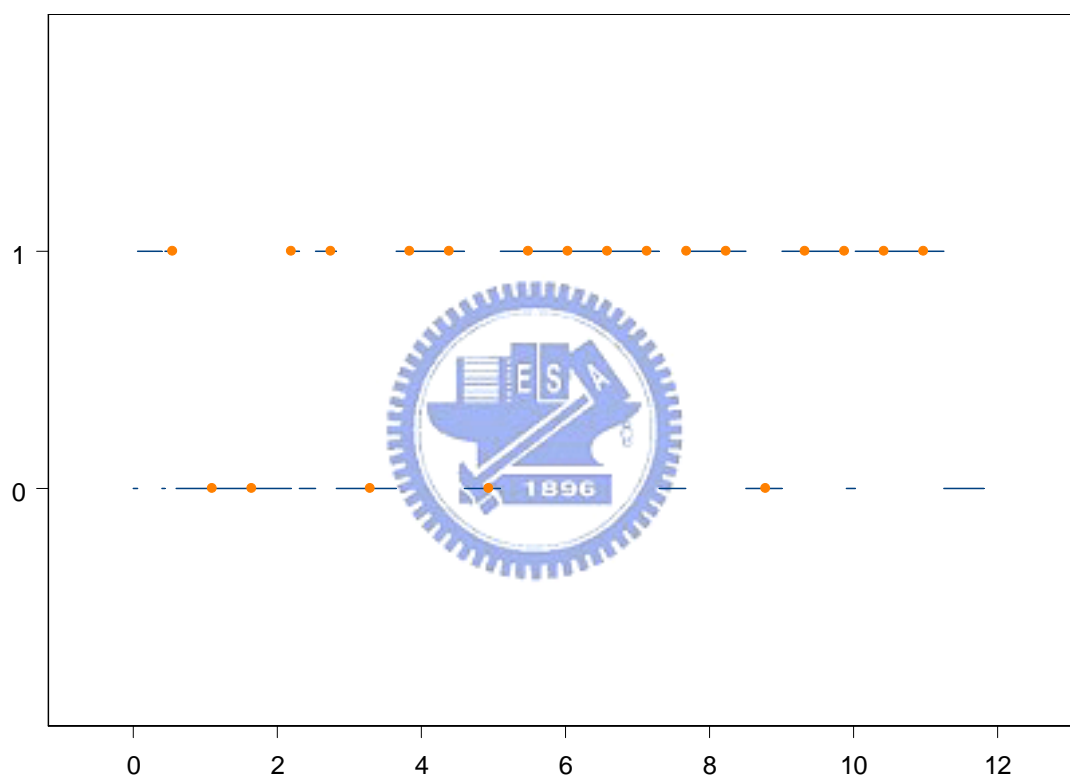
前二十次模擬狀態變化的時間如下：

時間	0.060	0.396	0.438	0.594	2.192	2.307	2.532	2.818	3.648	4.596
狀態	0	1	0	1	0	1	0	1	0	1
時間	5.096	7.299	7.669	8.504	9.007	9.902	10.026	11.254	11.811	12.386
狀態	0	1	0	1	0	1	0	1	0	1

觀察時間點，總共取了 9107 個時間點，列出前面二十個時間點的觀察時間 (t) 及觀察得到的狀態 ($X(t)$) 如下：

時間	0.548	1.097	1.646	2.195	2.744	3.292	3.841	4.390	4.939	5.488
狀態	1	0	0	1	1	0	1	1	0	1
時間	6.036	6.585	7.134	7.683	8.232	8.780	9.329	9.878	10.427	10.976
狀態	1	1	1	1	1	0	1	1	1	1

當我們將上述兩表得到的數據，畫在一起時，將更容易發現其變化：



結果利用 Matlab 軟體計算，估計 β 的最大似估計量為 0.2099。

第二個例子，令 $q_{01} = \exp\{x_0\beta_0\}$ ， $q_{10} = \exp\{x_1\beta_1\}$ ，得到微矩陣 $Q_2(t)$

$$Q_2(t) = \begin{bmatrix} -\exp\{x_0\beta_0\} & \exp\{x_0\beta_0\} \\ \exp\{x_1\beta_1\} & -\exp\{x_1\beta_1\} \end{bmatrix}。$$

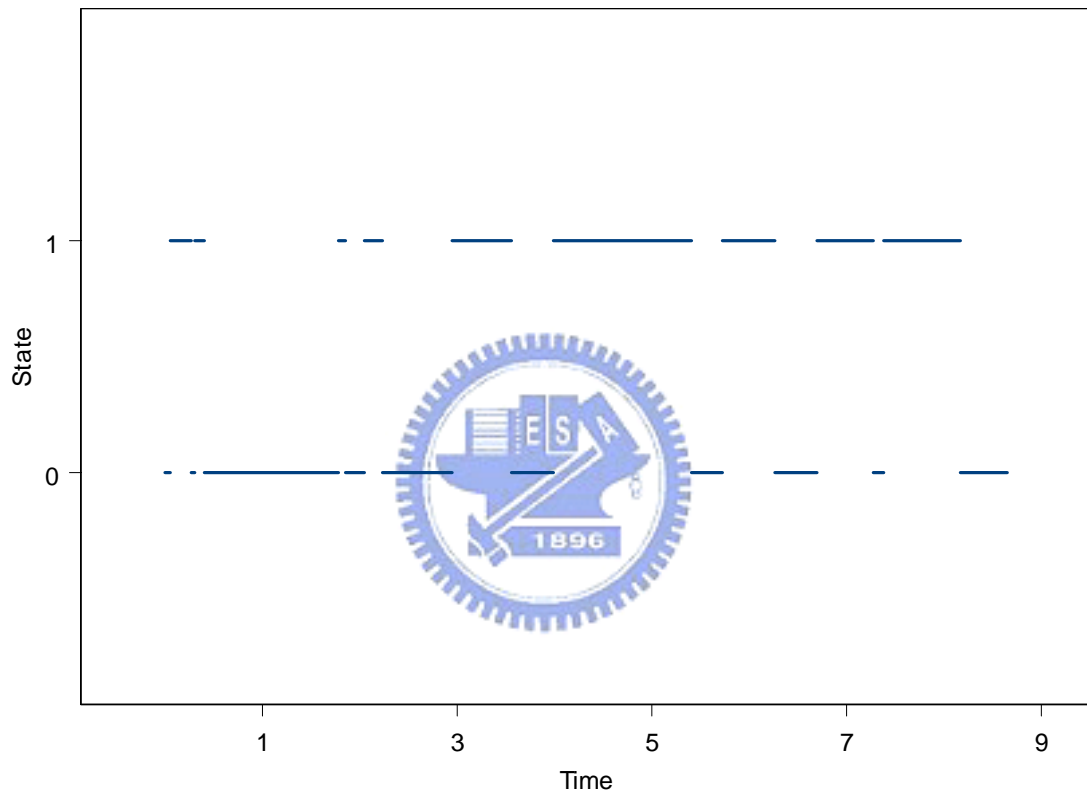
此微矩陣 $Q_2(t)$ 表示，不管哪一個時間點，從狀態 0 跳到狀態 1 的發生率為 $\exp\{x_0\beta_0\}$ ，與從狀態 1 跳到狀態 0 的發生率 $\exp\{x_1\beta_1\}$ 是不同的，但是這兩個發生率依

舊是常數。我們可以透過新的遞移模型分別來估計 β_0 與 β_1 。

模擬結果：

模擬的設定為 $x_0 = 0.5$ ， $\beta_0 = 2$ ， $x_1 = 0.8$ ， $\beta_1 = 3$ ，則 $Q_2(t) = \begin{bmatrix} -e & e \\ e^{2.4} & -e^{2.4} \end{bmatrix}$ 。

模擬時間長度為 10000，總共模擬出 24226 次變化，為方便解釋，畫出前二十次變化來看，圖形如下：



前二十次模擬狀態變化的時間如下：

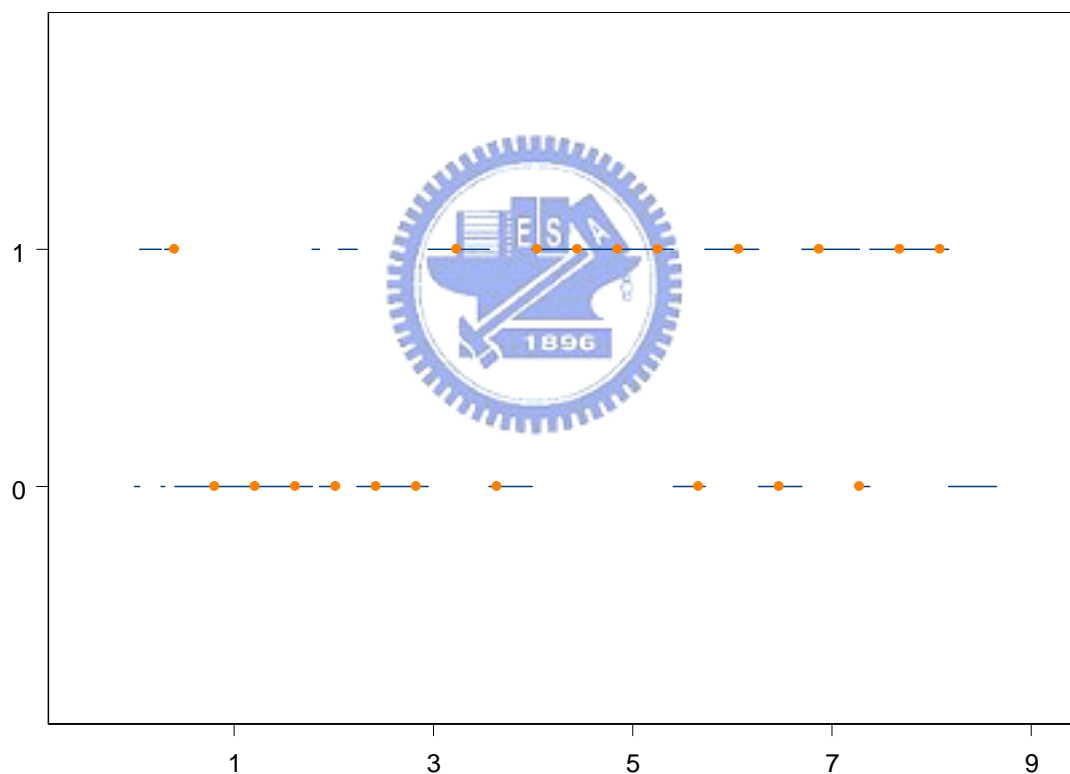
時間	0.052	0.268	0.305	0.405	1.780	1.854	2.048	2.232	2.947	3.557
狀態	0	1	0	1	0	1	0	1	0	1
時間	3.988	5.406	5.725	6.263	6.696	7.272	7.379	8.170	8.649	9.020
狀態	0	1	0	1	0	1	0	1	0	1

觀察時間點，總共取了 24722 個時間點，列出前面二十個時間點的觀察時間 (t)

及觀察得到的狀態 ($X(t)$) 如下：

時間	0.404	0.808	1.213	1.617	2.021	2.426	2.830	3.234	3.639	4.043
狀態	1	0	0	0	0	0	0	1	0	1
時間	4.447	4.852	5.256	5.660	6.065	6.469	6.873	7.278	7.682	8.086
狀態	1	1	1	0	1	0	1	0	1	1

當我們將上述兩表得到的數據，畫在一起時，將更容易發現其變化：



經由 Matlab 軟體計算後，估計 β_0 及 β_1 的最大概似估計量分別為 2.0578 及 3.0909。

4.2 微矩陣 $Q(t)$ 為時間的函數

介紹完與時間無關的微矩陣後，接下來，我們要將微矩陣 $Q(t)$ 延伸到與時間有關的函數， $Q(t) = \begin{bmatrix} -\lambda_0(t) & \lambda_0(t) \\ \lambda_1(t) & -\lambda_1(t) \end{bmatrix}$ ，而對於計算求參數的最大概似估計量而言，將會複雜許多。

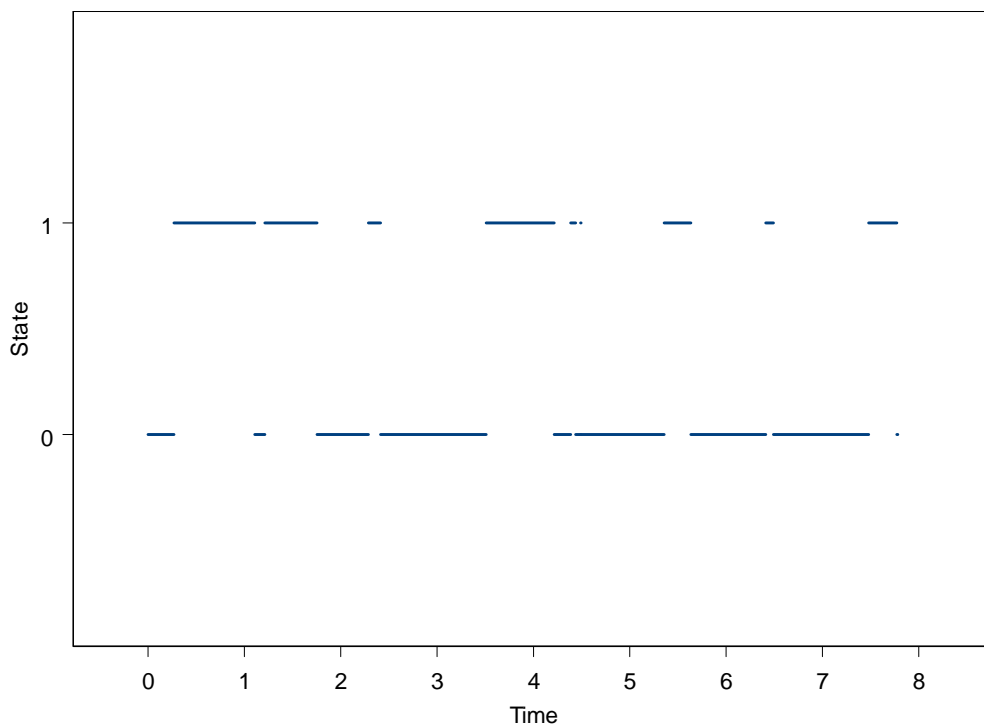
第三個例子，令 $q_{01}(t) = \exp(\sin(x_0\beta_0 t) + 1)$ ， $q_{10}(t) = \exp(\cos(x_1\beta_1 t) + 1)$ ，得到

$$Q_3(t) = \begin{bmatrix} -\exp((\sin(x_0\beta_0 t) + 1)) & \exp((\sin(x_0\beta_0 t) + 1)) \\ \exp((\cos(x_1\beta_1 t) + 1)) & -\exp((\cos(x_1\beta_1 t) + 1)) \end{bmatrix}$$

模擬結果：

模擬的設定為 $x_0 = 2$ ， $\beta_0 = 2$ ， $x_1 = 1.5$ ， $\beta_1 = 3$ ，則 $Q_3(t) = \begin{bmatrix} -e^{\sin 4t+1} & e^{\sin 4t+1} \\ e^{\cos 4.5t+1} & -e^{\cos 4.5t+1} \end{bmatrix}$ ，

模擬時間長度為 10000，總共模擬出 29879 次變化，取前二十次變化來看，圖形如下：



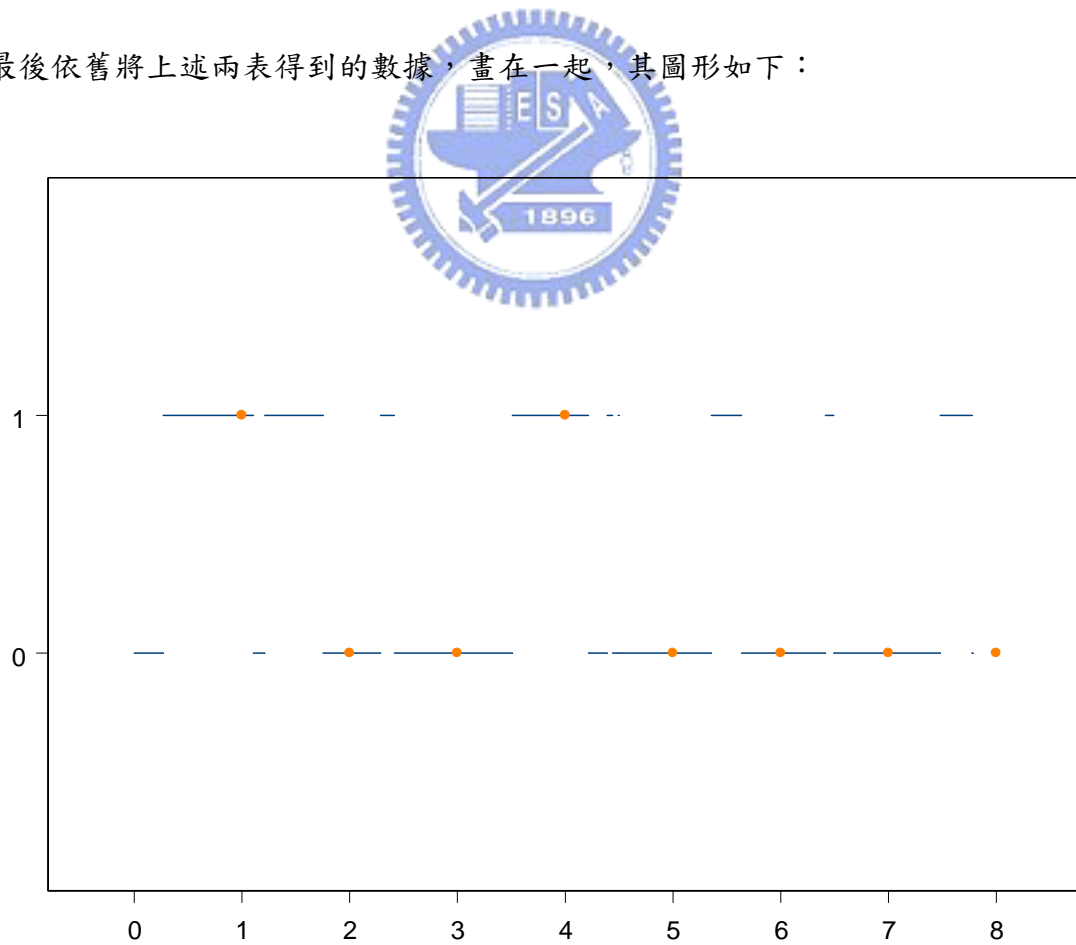
前二十次模擬狀態變化的時間如下：

時間	0.269	1.106	1.212	1.753	2.286	2.414	3.508	4.215	4.388	4.438
狀態	0	1	0	1	0	1	0	1	0	1
時間	4.491	4.496	5.356	5.635	6.413	6.492	7.480	7.773	7.783	7.968
狀態	0	1	0	1	0	1	0	1	0	1

觀察時間點，總共取了 9995 個時間點，列出前面二十個時間點的觀察時間 (t) 及觀察得到的狀態 ($X(t)$) 如下：

時間	1	2	3	4	5	6	7	8	9	10
狀態	1	0	0	1	0	0	0	0	0	1

最後依舊將上述兩表得到的數據，畫在一起，其圖形如下：



利用 Matlab 軟體計算，估計 β_0 及 β_1 的最大概似估計量分別為 2.0571 及 3.0429。

第五章 結論

這篇論文的主要目的是對於長期二元資料提供一個新的遞移模型，此遞移模型具有非常大的彈性，這是此遞移模型的優點，藉由非齊性連續時間馬可夫鍊，推導出其轉移機率，將密度函數受時間影響的因素考量進去，我們可以在任意時刻開始觀察，將第一次觀察到的現象當作起始點 t_0 ，不受觀察起始點的影響，而觀察的時間區間長度，可以取任意長度，甚至每次取的時間長度都不一樣也可以，這篇論文所舉的例子中，時間長度是固定的。對於微矩陣的選取，本篇論文中並沒有提出具體的方法，基本上，這需要研究者長期的經驗來決定各種函數是否適宜。

此遞移模型的缺點是計算繁雜，需要花費比較多的時間，因為我們考量的因素比較多，計算參數的最大概似估計量時，是利用 Newton-Rapbson 法，而不同的起始值會影響 Newton-Rapbson 法的收斂時間的長短，除了這個方法外，還可以利用 EM 演算法來幫助找出參數的最大概似估計量。

我們針對單一個體長時間的觀察後，估計出其參數的最大概似估計量，可以延伸到針對不同的獨立個體，假設其微矩陣後，進一步估計不同獨立個體參數的最大概似估計量，例如個體 1 的微矩陣為 $Q(t | \underline{x}_1' \underline{\beta})$ ，個體 2 的微矩陣為 $Q(t | \underline{x}_2' \underline{\beta})$ ， \dots ，個體 n 的微矩陣為 $Q(t | \underline{x}_n' \underline{\beta})$ ，參數 $\underline{\beta}$ 是相同的，因為假設獨立，因此可以很容易寫出其概似函數，估計出其參數 $\underline{\beta}$ 的最大概似估計量。

雖然此篇論文中，只有利用到二狀態連續時間馬可夫鍊的轉移機率，但是我們有提出 n 狀態非齊性連續時間馬可夫鍊的轉移機率求法，亦可利用 n 狀態非齊性連續時間馬可夫鍊的轉移機率針對 n 個不同出象的觀察過程提出類似此論文二狀態馬可夫鍊的遞移模型，相同地，估計有興趣因子參數的最大概似估計量。