行政院國家科學委員會專題研究計畫　成果報告

個人化馬可夫鏈模型在顧客價值遷徙路徑之分析

計畫類別：個別型計畫
計畫編號：NSC93-2416-H-002-008-
執行期間：93 年 08 月 01 日至 95 年 01 月 31 日
執行單位：國立臺灣大學國際企業學系暨研究所

計畫主持人：任立中

計畫參與人員：陳靜怡

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 95 年 2 月 23 日

顧客價值遷移路徑分析：馬可夫鏈模型

Customer Value Migration Analysis: Markov Chain Model

任立中
Lichung Jen
國立台灣大學國際企業學系副教授
台北市 106 羅斯福路四段 1 號
TEL: (02) 3366-4983
E-MAIL: lichung@management.ntu.edu.tw

陳靜怡
Ching-I Chen
國立暨南大學國際企業學系助理教授
南投縣 545 埔里鎮大學路 1 號
TEL: (049) 291-0960 EXT. 4911
E-MAIL: chingichen@ncnu.edu.tw

本研究感謝國科會編號 NSC93-2416-H-002-008 研究案之資助及於 2004 電子商務與數位生活研討會發表時學者的評論。

摘要

資料庫行銷是落實關係行銷與一對一行銷不可或缺的利器，廠商透過資料庫的分析得以衡量個別顧客之價值。傳統的顧客價值衡量方法多偏重於顧客於特定時點的平均價值，於廠商成長並無太大的助益。本研究認為廠商應從動態的角度觀察顧客價值成長或萎縮的過程，事前預判顧客價值的遷移路徑，方能達到開發顧客價值潛力以及預防顧客價值流失的效果。本研究以馬可夫鏈模型為基礎，根據顧客過去之價值狀態遷移路徑，搭配層級貝氏方法，如貝氏多項式模型、層級貝氏 Probit 模型以及層級貝氏 Logit 模型，建立個人化的移轉機率矩陣，進而預測顧客的價值遷移路徑，據此採取適當的顧客關係行銷之策略。

關鍵字：顧客價值、遷移模型、馬可夫鏈模型、層級貝氏方法

Abstract

Database Marketing is a requisite and efficient weapon for firms to put one-to-one relationship marketing in practice. The major task of this practice is to evaluate or calculate individual customer value through proper analyses of customers' purchase history data. Conventional approaches to measuring customer value are focused on the average value from past to present, which is less helpful to firm's growth strategy. What firms need to do is to be able to capture the migration patterns of a customer's value in advance, in order to explore potential customers and prevent inactive customers. This paper employs Markov chain model and hierarchical Bayesian approach to construct individual customer's transition probability matrix for forecasting the customer value migration process.

Keyword: Customer Value, Migration Patterns, Markov Chain Model, Hierarchical Bayesian Approach

一、緒論

廠商在資源有限的情況下，如何提高行銷效率以創造更大的利潤，一直是實務界所關心的議題。為了提升效率，行銷管理者致力於找出具有潛力的高價值顧客群，施以適當的行銷策略以維持良好的關係，從而提高或維持優質顧客的忠誠度。近年來由於資料庫的興起與發展，廠商得以鉅細靡遺的紀錄顧客的實際購買行為。然而，面對如此龐大繁瑣的資料庫，如何有效管理，進而得以精確衡量顧客價值並掌握其行為趨勢，方是關係行銷成為致勝武器的重要關鍵。

資料庫行銷經常使用最近購買期間（Recency）、購買頻率（Frequency）、購買金額（Monetary）等行為指標衡量顧客之獲利性（Hughes 1994; Stone 1995; Schijns & Schroder 1996）。過去這類的研究多偏重於靜態分析，將顧客過去的購買行為摘要成一具代表性的指標（如平均值或最近值），藉此衡量個別顧客之平均價值，作為顧客關係管理的依據。然而，靜態分析僅能呈現已經實現（亦即過去的、或截至目前為止的）的顧客價值，導致顧客關係管理的效果只是事後的錦上添花（如優惠並留住高價值顧客）或亡羊補牢（激勵或維持低價值顧客），這對廠商的成長並無太大的助益。

除了已實現的購買行為之外，廠商更重視的是顧客未來價值的演變（Dwyer 1989; Sewell & Brown 1990; Peppers & Rogers 1993; Hughes 1994; Kotler 1996）。這方面的研究大致分成兩個方向，一是衡量顧客的終生價值（customer lifetime value），有助於目標顧客的界定及行銷資源的配置；另是追蹤顧客價值的演變，在不同的價值時期施以適當的行銷策略。傳統的資料庫行銷所建立的終生價值模型，多屬於總合層次的價值折現模型，包括各期的銷售量、行銷成本、年折現率、顧客維持率等元素（Berger & Nasr 1998; Blatterberg & Deighton 1996），無法有效達到一對一行銷的目的。顧客生命週期方面的理論雖對顧客價值的演變多有著墨（Raphel & Raphel 1995; Gronroos 2000; Berry & Linoff 2000），但都似乎缺乏一套有效的追蹤方法。

Pfeifer & Carraway（2000）以個人化的移轉機率矩陣，在馬可夫鏈模型（Markov chain model）假設之下，捕捉顧客在不同購買決策狀態之間的移轉，有效落實顧客未來價值之衡量與追蹤。馬可夫鏈模型的最大優勢是彈性（flexibility），可廣泛涵蓋顧客於各種價值狀態之間的移轉狀況，形成常見的顧客遷移模型（customer migration model）或顧客維持模型（customer retention model）。過去在資料庫行銷中備受重視的顧客維持率（retention rates）、平均顧客利潤貢獻（average profit）等概念，皆屬於總合的摘要統計指標，僅能衡量一群而非單一顧客的終生價值。相對之下，由於馬可夫鏈模型本身係一機率模型，

廠商除可透過機率概念衡量顧客關係的不確定性之外，還可透過機率和期望值的概念預測公司與個別顧客的未來關係，可有效協助一對一行銷之落實。美中不足的是，該篇論文的討論僅止於移轉機率矩陣之應用，未曾說明如何建立個人化的移轉機率矩陣，也就無法真正達到一對一顧客關係行銷之目的。

本研究嘗試以顧客的過去購買紀錄為依據，透過貝氏統計模型估計個人化移轉機率矩陣，以追蹤每一個顧客的價值遷移路徑。接下來的章節安排如下：首先是馬可夫鏈模型之說明，在穩定性（stationary）的假設之下，透過移轉機率矩陣的運算，我們可以預測顧客未來行為路徑及最終行為落點。其次是移轉機率矩陣之估計，本研究使用貝氏統計模型，包括貝氏多項式模型、層級貝氏 Probit 模型、層級貝氏 Logit 模型等，推估個人化的移轉機率矩陣，藉此衡量顧客在不同狀態間移轉的可能性。然後根據樣本外預測的結果，我們挑選出擊中率較佳的模型，並以該模型的估計結果衡量顧客未來價值的演變。最後是顧客價值遷移路徑分析，確認不同路徑的顧客特質，以有助於廠商掌握現有顧客的價值及預測潛在顧客的行為。

二、模型

（一）馬可夫鏈模型

顧客消費行為的隨機變化，常使廠商不易掌握其後續發展。但馬可夫鏈模型的預測對象正是一個隨機變化的動態系統。馬可夫鏈模型的理論基礎為馬可夫過程（Markov Process），假設在已知目前狀態的條件，未來的狀態演變只與目前狀態有關，與過去的狀態變化無直接關係。茲定義狀態空間（state space）共存在 S 種狀態，描述狀態之間移轉可能性的移轉機率矩陣 P 可列示如下：

式（1）

$$P \equiv (t) \begin{bmatrix} p(1|1) & p(2|1) & \cdots & p(S|1) \\ p(1|2) & p(2|2) & \cdots & p(S|2) \\ \vdots & \vdots & \ddots & \vdots \\ p(1|S) & p(2|S) & \cdots & p(S|S) \end{bmatrix} \quad (t+1)$$

矩陣 P 稱為一階移轉機率矩陣（one-step transition probability matrix），矩陣內的元素 p(k|j)定義為已知本期（第 t 期）為狀態 j 的條件下，下期（第 t+1 期）狀態為 k 的條件機率，且 $\sum_{k=1}^{S} p(k \mid j) = 1$。矩陣 P 可再擴充為 m 階移轉機率矩陣（m-step transition probability matrix），即矩陣 P 自乘 m 次，如下所示：

式（2）

$$P^m = \underbrace{P \cdot P \cdots P}_{\text{共m次}} = (t)\begin{bmatrix} p^{(m)}(1|1) & p^{(m)}(2|1) & \cdots & p^{(m)}(S|1) \\ p^{(m)}(1|2) & p^{(m)}(2|2) & \cdots & p^{(m)}(S|2) \\ \vdots & \vdots & \ddots & \vdots \\ p^{(m)}(1|S) & p^{(m)}(2|S) & \cdots & p^{(m)}(S|S) \end{bmatrix}^{(t+m)}$$

同理，矩陣 P^m 內的元素 $p^{(m)}(k|j)$定義為已知本期（第 t 期）為狀態 j 的條件下，預測未來第 m 期(第 t+m 期)狀態為 k 的條件機率。當 m 趨於無窮大之後$(m\to\infty)$，各種狀態的機率將趨於穩定，亦即 $p^{(m)}(k|1)= p^{(m)}(k|2)=\ldots=p^{(m)}(k|S)=\pi_k$，$k=1,2,\ldots,S$；$\pi=\{\pi_1,\pi_2,\ldots, \pi_S\}$稱為此馬可夫鏈之極限狀態機率（steady-state probability）向量。因此，只要得知顧客的移轉機率矩陣，就可根據其目前之起始狀態，以預測他們於未來各期的購買行為狀態。

廠商可同時根據數個購買行為指標（如 RFM 指標），將購買紀錄定義為數種型態，如圖 1 所示。圖 1（a）顯示顧客 i 有 L 期的購買紀錄，經操作性定義為 S 種不同狀態。除了當期狀態（j）之外，接連的下期狀態（k）亦是馬可夫鏈模型關心的重點。根據移轉機率矩陣之定義（式 1），矩陣元素是給定當期狀態（j）之條件下，下期發生狀態（k）之條件機率，可藉由連續兩期狀態之聯合發生機率推算而得。圖 1（b）之兩期狀態次數分配是計算聯合發生機率之依據，來自於圖 1（a）的計數（count）結果。例如，圖 1（a）顯示顧客 i 於第 1 期（本期）的狀態為 1，第 2 期（下期）的狀態為 3，形成該期觀察值(j,k)=(1,3)，對應至圖 1（b）則是於第 1 列第 3 欄記上一筆。

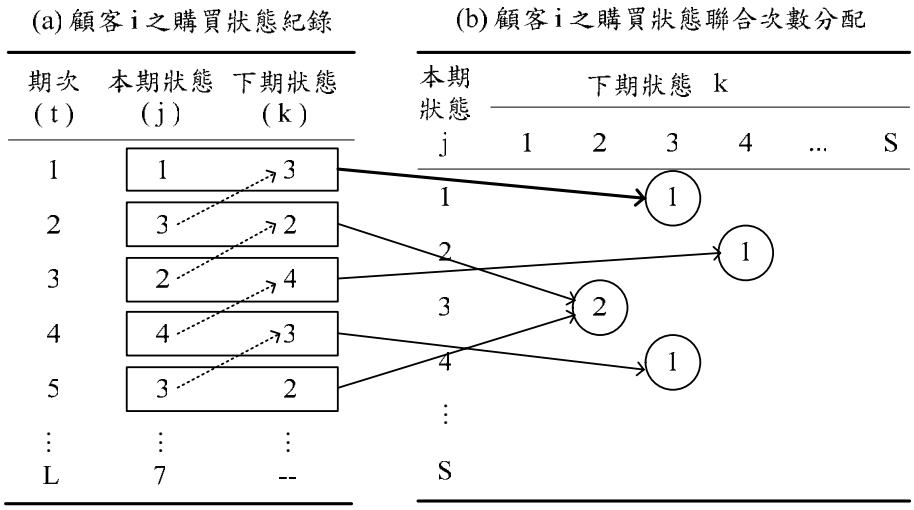


圖 1 購買狀態之聯合次數分配

狀態移轉機率屬於條件機率，相當於方格次數（F_{ijk}）相對於列次數和（$F_{ij\bullet}$）的比值，如表 1 所示。表中，F_{ijk} 為顧客 i 接連兩期的狀態是（j,k）的發生次數，亦即圖 1（b）的方格次數；$F_{ij\bullet}$是第 j 列的方格次數加總，$F_{i\bullet\bullet}$則是所有方格的次數加總。由圖 1（a）可知，我們無法觀察到顧客於最後一期（t=L）的下期狀態（k）而形成缺失值，故總次數 $F_{i\bullet\bullet}=\sum_{j=1}^{S}\sum_{k=1}^{S}F_{ijk}$ =L–1。顧客 i 接連兩期發生狀態（j,k）之聯合機率相當於方格次數對總次數之比值，即 $p_{ijk}= F_{ijk}/F_{i\bullet\bullet}$；顧客 i 於本期發生狀態（j）之邊際機率則相當於列次數和對總次數之比值，即 $p_{ij\bullet}= F_{ij\bullet}/F_{i\bullet\bullet}$。據此，狀態移轉機率 $p_i(k|j)$的計算公式如下所示：

式（3）　　$p_i(k|j)=F_{ijk}/F_{ij\bullet}= p_{ijk}/p_{ij\bullet}$　　$F_{ij\bullet}\neq0$或$p_{ij\bullet}\neq0$

式中，限制式 $F_{ij\bullet}\neq0$ 必須成立，這相當於要求每位顧客在觀察期間（第 1 期至第 L–1 期），每個購買狀態至少要有發生一次的紀錄，但個人的購買紀錄通常無法達到此項要求。為了解決資訊不足造成個人化參數難以估計的問題，貝氏統計分析（Bayesian Statistical Analysis）整合總體資訊和個體資訊以取得個人化參數估計的方法，近年來已被廣泛地運用在各種不同的行銷研究問題上（Allenby, Leone & Jen 1999; Allenby & Rossi 2000; Jen & Wang 1998）。

表 1 顧客 i 連續兩期狀態（j,k）之次數分布與聯合機率分配

狀態		下一期(k)				
		1	2	…	S	列總和
本期(j)	1	$F_{i11}(p_{i11})$	$F_{i12}(p_{i12})$	…	$F_{i1S}(p_{i1S})$	$F_{i1\bullet}(p_{i1\bullet})$
	2	$F_{i21}(p_{i21})$	$F_{i22}(p_{i22})$	…	$F_{i2S}(p_{i2S})$	$F_{i2\bullet}(p_{i2\bullet})$
	⋮	⋮	⋮	⋮	⋮	⋮
	S	$F_{iS1}(p_{iS1})$	$F_{iS2}(p_{iS2})$	…	$F_{iSS}(p_{iSS})$	$F_{iS\bullet}(p_{iS\bullet})$
行總和		$F_{i\bullet1}(p_{i\bullet1})$	$F_{i\bullet2}(p_{i\bullet2})$	…	$F_{i\bullet S}(p_{i\bullet S})$	$F_{i\bullet\bullet}$=L–1 ($p_{i\bullet\bullet}$= 1)

（二）貝氏多項式模型

本研究先採用貝氏多項式（Bayesian Multinomial, BMN）模型推估個別顧客連續兩期狀態之聯合機率矩陣，再轉換為個人化移轉機率矩陣。本研究應用貝氏統計理論將總體資訊處理為先驗知識（prior knowledge），個人資訊處理為樣本知識（sample knowledge）；再將這兩種資訊所形成的先驗分配和概似函數做比例組合，推導出參數之後驗分配（posterior distribution），其期望值即貝氏統計之參數點估計量。假設表 1 之方格次數 F_{ijk} 遵循多項式分配（Multinomial Distribution, MND），構成樣本概似函數，如下所示：

式（4）　$[F_{ijk}|p_{ijk},L^*] = MND(F_{ijk}|p_{ijk},L^*) = \frac{L^*!}{\prod_{j=1}^{S}\prod_{k=1}^{S} F_{ijk}!}\prod_{j=1}^{S}\prod_{k=1}^{S} p_{ijk}^{F_{ijk}}$ ，其中

$$\sum_j\sum_k F_{ijk} = F_{i\bullet\bullet} = L^* \text{，} \sum_j\sum_k p_{ijk} = p_{i\bullet\bullet} = 1 \text{，} 0 \le p_{ijk} \le 1$$

式中，$L^*=L-1$，即個別顧客之觀察期數減 1，亦相當於總方格次數（$F_{i\bullet\bullet}$）；p_{ijk} 是 MND 的個人化參數，其最大概似估計值 $\hat{p}_{ijk}=F_{ijk}/L^*$。由於個人資訊（樣本知識）的不足，$\hat{p}_{ijk}$ 值可能多半為 0，但這僅代表該顧客過去尚未發生過連續狀態為（j,k）的情形，並不代表未來發生的可能性為 0。因此，本研究以 BMN 模型整合總體資訊（先驗知識）及樣本資訊（樣本知識），藉此修正個人化參數估計。

BMN 模型通常假設個人化參數 p_{ijk} 的先驗分配為 Dirichlet 分配，即 MND 之自然共軛機率分配（natural conjugate distribution）（Berger 1985），有助於推導 p_{ijk} 之後驗分配。Dirichlet 分配之機率密度函數如下（Johnson & Kotz 1972）：

式（5）　$[p_{ijk}|\alpha_{jk}] = Dirichlet(p_{ijk}|\alpha_{jk}) = \frac{\Gamma(\alpha_0)}{\prod_{j=1}^{S}\prod_{k=1}^{S}\Gamma(\alpha_{jk})}\prod_{j=1}^{S}\prod_{k=1}^{S} p_{ijk}^{(\alpha_{jk}-1)}$ ，其中

$$\sum_j\sum_k \alpha_{jk} = \alpha_0 \text{，} 0 \le \alpha_{jk} \le \alpha_0 \text{，} \sum_j\sum_k p_{ijk} = 1 \text{，} 0 \le p_{ijk} \le 1$$

式中，α_{jk} 與 α_0 皆是 Dirichlet 分配之參數，又稱為純先驗參數（pure priors），由研究者自行設定，與先驗期望值之關係為 $E(p_{ijk}|\alpha_{jk})=\alpha_{jk}/\alpha_0$；$\alpha_0$ 是所有 α_{jk} 參數之總和。結合 MND（概似函數）與 Dirichlet 分配（先驗分配），個人化參數 p_{ijk} 的後驗分配之推導如下所示：

式（6）　$[p_{ijk}|\alpha_{jk}, F_{ijk}] \propto MND(F_{ijk}|p_{ijk}, L^*)\times Dirichlet(p_{ijk}|\alpha_{jk})$

$$\propto \prod_{j=1}^{S}\prod_{k=1}^{S} p_{ijk}^{F_{ijk}} \cdot \prod_{j=1}^{S}\prod_{k=1}^{S} p_{ijk}^{(\alpha_{jk}-1)} = \prod_{j=1}^{S}\prod_{k=1}^{S} p_{ijk}^{(F_{ijk}+\alpha_{jk}-1)} \propto Dirichlet(p_{ijk}|F_{ijk}+\alpha_{jk})$$

如推導結果所示，p_{ijk} 的後驗分配正比於 Dirichlet 分配，與先驗分配相同，但參數調整為（$F_{ijk}+\alpha_{jk}$）。後驗分配之期望值即為貝氏統計之參數估計量（p^*_{ijk}）：

式（7）　$$p^*_{ijk} = E(p_{ijk}|F_{ijk}+\alpha_{jk}) = \frac{F_{ijk}+\alpha_{jk}}{L^*+\alpha_0} = \frac{F_{ijk}}{L^*+\alpha_0}\cdot\frac{L^*}{L^*} + \frac{\alpha_{jk}}{L^*+\alpha_0}\cdot\frac{\alpha_0}{\alpha_0}$$

$$= \frac{L^*}{L^*+\alpha_0}\left(\frac{F_{ijk}}{L^*}\right) + \frac{\alpha_0}{L^*+\alpha_0}\left(\frac{\alpha_{jk}}{\alpha_0}\right) = w_1\hat{p}_{ijk} + w_2\gamma_{jk}$$

如上式所示，貝氏統計量（p^*_{ijk}）是樣本估計值（$\hat{p}_{ijk}$）與先驗期望值（γ_{jk}）之加權和，權數（w_1,w_2）取決於樣本參數 L^*（即個人觀察期數）與先驗參數α_0（由研究者自行設定）之相對大小。因此，先驗參數α_0之設定不宜與L^*相差太多，否則後驗估計值可能會完全偏向先驗期望值（γ_{ij}）或樣本估計值（$\hat{p}_{ijk}$），無法同時反映總體資訊及展現個人的異質性。因此，本研究對於先驗參數之設定如下所示：

（1） $\gamma_{ij} \equiv \alpha_{jk}/\alpha_0 = \sum_{i=1}^{N} F_{ijk} \Big/ \sum_{i=1}^{N}\sum_{j=1}^{S}\sum_{k=1}^{S} F_{ijk} = \sum_{i=1}^{N} F_{ijk} \Big/ NL^*$ ；

（2） $w_1 \equiv L^*/(L^*+\alpha_0) = 0.8, 0.5, 0.2$。

首先，先驗期望值γ_{ij}，即參數α_{jk}相對於α_0之比值，被設定為所有 N 位顧客連續兩期處於狀態（j,k）之總期數相對於總觀察期數的比值。其次，樣本資訊的權數 w_1 被設定為 0.8, 0.5, 0.2 等三種數值，使後驗估計值分別呈現偏向樣本資訊、二者平均、偏向先驗資訊等結果。後驗估計值（p^*_{ijk}）為計算個人化移轉機率之依據，即 $p_i(k|j) = p^*_{ijk} / \sum_{k=1}^{S} p^*_{ijk}$ 。

然而，BMN 模型較適用於具購買紀錄的現有顧客的行為預測，對於尚無購買紀錄之潛在顧客則難以預測。相較之下，層級貝氏 Probit（Hierarchical Bayesian Probit, HB Probit）模型與層級貝氏 Logit（Hierarchical Bayesian Logit, HB Logit）模型可將顧客特質變數納入機率模型之中（Rossi & McCulloch 1994），有助於預測潛在顧客的未來購買狀態，更具管理意義。Lenk（2001）曾提出建立 HB Probit 模型與 HB Logit 模型之完整過程，是本研究之主要參考依據。

（三）HB Probit 模型與 HB Logit 模型

Probit 模型與 Logit 模型是探討分類性變數行為的統計方法，皆可應用於估計移轉機率矩陣。不過，這兩個模型皆不適合處理組數太多的次數資料，如本研究之連續狀態組數就高達 S^2 組。為了減少組數，本研究的作法是先給定本期狀態為 j 之條件下，以 Probit 模型或 Logit 模型（泛稱為條件狀態選擇模型）直接估計下期狀態 k 之條件發生機率，k=1,2,…,S，構成移轉機率矩陣（如式 1）之第 j 列數值。由於移轉機率矩陣共有 S 列，即 j=1,2,…,S，故本研究須建立 S 個條件狀態選擇模型，方能逐列補滿移轉機率矩陣之參數估計。為能整合全體資訊及個人資訊推導個人化參數估計，本研究以 HB 方法建立個人化條件狀態選擇模型。

然而，若個人的購買紀錄並未包含狀態 j，則就欠缺以本期狀態 j 為條件之下期狀態資料，以致無法建立對應的個人化條件狀態選擇模型。如圖 2 所示，

在建立第 j 個模型之前（j=1,2,...,S），投入的顧客購買紀錄必須加以篩選。圖 2（a）是顧客 i 的購買紀錄，以 S 種狀態表示之，其中只有「本期狀態為 j」的下期狀態選擇資料，才適合用以建立第 j 個模型，如圖 2（b）所示。例如，若欲建立顧客 i 的第 3 個 HB 條件狀態選擇模型，則投入資料必須是本期狀態為 3（j=3）的下期狀態選擇（k），如第 2 期與第 5 期這兩筆資料，k=2,6，方適合用以配適第 3 個模型。因此，在估計第 j 個模型的時候，顧客 i 僅有 $L_{i|j}$ 筆下期狀態選擇資料可用以配適該模型，$L_{i|j} \le L^*$；而僅有 N_j 位顧客具有符合條件的資料可以推論個人化移轉機率矩陣的第 j 列移轉機率向量，（$N-N_j$）位顧客則因缺乏資料而必須依靠總體資訊估計之，詳細說明於後。

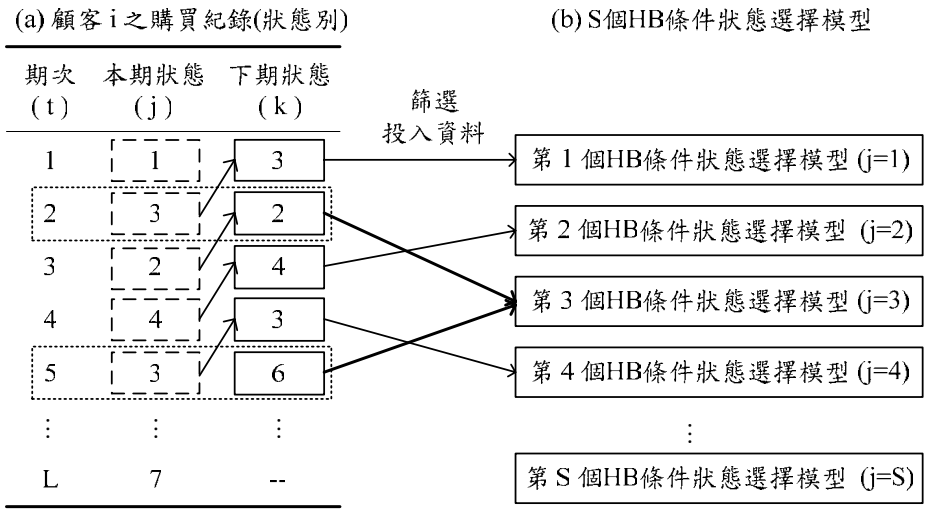


圖 2　HB 條件狀態選擇模型之投入資料篩選

本研究分別根據 Probit 模型及 Logit 模型之假設，建立兩個 HB 條件狀態模型。HB Probit 模型假設狀態選擇是狀態相對效用之比較結果，且假設所有狀態之相對效用為無法觀察之連續型潛伏資料（continuous latent data）向量，遵循多變量常態分配。以第 j 個 HB Probit 模型為例，假設經資料篩選後，顧客 i 僅有 $L_{i|j}$ 筆購買紀錄適合投入模型配適，則第一層模型之設定如下所示：

式（8）

$$y^*_{itm|j} = \mu_{im|j} + \varepsilon_{itm|j} \qquad m=1,2,\ldots,S-1$$

$$\begin{cases} y^*_{itk|j} \ge \max\left(0, y^*_{itm|j}\right), & 若 y_{it|j} = k \quad k=1,2,\ldots,S-1 \\ y^*_{itm|j} < 0, & 若 y_{it|j} = S \quad t=1,2,\ldots,L_{i|j} \end{cases}$$

式中

$y_{it|j}$=顧客 i 於時點 t 的條件狀態選擇，共分為 S 種；

$y^*_{itm|j}$ =顧客 i 於時點 t 的條件狀態選擇為 m 的潛伏相對效用（令狀態 S 為比較基礎），相對效用之大小排序與觀察值（$y_{it|j}$）具對應的關係；

$\mu_{im|j}$ =顧客 i 認為處於條件狀態 m 可得到之平均相對效用，

$\varepsilon_{itm|j}$ =相對效用之誤差項，（S–1）個誤差項構成的向量（$\varepsilon_{it1|j}, \varepsilon_{it2|j},\ldots, \varepsilon_{it,S-1|j}$）被假設遵循多變量常態分配，期望值為 0 向量，變異數矩陣為Σ_j，並限制 $Var(\varepsilon_{it,S-1|j})=\sigma^2_{S-1|j}=1$。

式（8）說明潛伏相對效用之大小排序與狀態選擇觀察值具有對應的關係。例如，若狀態選擇觀察值為 k，則意指狀態 k 之相對效用（$y^*_{itk|j}$）不僅是正值之外，也大於其他所有狀態之相對效用；若狀態選擇觀察值為 S（即基礎狀態），則代表（S–1）個狀態的相對效用皆為負值。HB 方法並設定參數（平均相對效用向量，$\mu_{i|j}$）之先驗分配做為第二層模型，如下所示：

式（9）　　$\mu_{i|j}=\Theta'_j z_i + \delta_{i|j}$　　$i=1,2,\ldots,N_j$

式中

$\mu_{i|j}$=平均相對效用向量，係一[(S–1)×1]向量；

z_i=顧客 i 之顧客特質向量，包括性別、年齡等，係一(P×1)向量，經資料篩選後，假設符合資格投入模型配適的顧客共有 N_j 位；

Θ_j =迴歸係數矩陣，衡量平均相對效用如何因顧客特質而異，係一[P×(S–1)]矩陣；

$\delta_{i|j}$ =誤差項向量，係一[(S–1)×1]向量，假設遵循多變量常態分配 $MN(0,\Lambda_j)$。

此一層次模型之參數估計結果，可應用於預測具相同特質之顧客處於各種狀態時得到的平均相對效用，進而預測「平均」的條件狀態選擇機率，有助於補足缺乏購買紀錄顧客（共 $N–N_j$ 位）之個人化移轉機率矩陣的第 j 列機率向量。然而，不論是參數估計或是機率預測，皆須經過複雜的高維度積分運算。為解決此一問題，HB 模型以馬可夫鏈蒙地卡羅（MCMC, Markov Chain Monte Carlo）方法中的吉布斯抽樣（Gibbs Sampling）方法逼近高維度積分結果，詳細過程請見附錄（一）。

HB Logit 模型與 HB Probit 模型類似，二者差異主要在於誤差項的機率分配設定。Logit 模型假設相對效用之誤差項遵循 Logistic 分配（Mcfadden 1973），可直接建立條件狀態選擇機率模型，即 HB Logit 模型的第一層模型，如下所示：

式（10）　$$p_i(k|j)=\frac{\exp(\mu_{ik|j})}{\exp(\mu_{i1|j})+\exp(\mu_{i2|j})+...+\exp(\mu_{i,S-1|j})+\exp(\mu_{i,S|j})}$$

式中

$p_i(k|j)$=顧客 i 給定本期是狀態 j 的條件下，下期狀態為 k 的條件選擇機率，即個人化移轉機率矩陣第 j 列第 k 欄元素；

$\mu_{im|j}$ =顧客 i 選擇條件下期狀態為 m 的平均相對效用，m=1,2,…,S；令 $\exp(\mu_{i,S|j})=\exp(0)=1$。

除了第一層模型設定之外，HB Logit 模型之其餘與參數估計有關的先驗分配設定皆與 HB Probit 模型相同，即第二層模型亦設定為平均相對效用對顧客特質之線性迴歸模型。然而，由於 HB Logit 模型的第一層模型（樣本分配）與第二層模型（先驗分配）無法結合成一具封閉型式（close form）的後驗分配函數，也就無法透過吉布斯抽樣方法逼近高維度積分結果，故本研究另以梅托普利斯-海斯丁運算法（Metropolis-Hastings Algorithm）（Metropolis et al. 1953; Hastings 1970）的作法逼近之，詳細過程請見附錄（二）。第二層模型之參數估計結果亦可應用於預測具相同特質之顧客處於各種狀態時之平均相對效用，利用式（10）即可預測「平均」的條件狀態選擇機率，有助於補足缺乏購買紀錄顧客之移轉機率矩陣估計值。

三、實證分析

（一）資料特性與狀態定義

本研究以國內某家電信業者之資料庫為實證研究對象，自資料庫隨機抽取出的 1,000 位顧客做為樣本。觀察期間設定從 2004 年 9 月起至 2005 年 4 月止，共 35 週，交易紀錄包括客戶編號、通話秒數、通話次數與通話金額。本研究將交易紀錄分成兩個部份，前 25 期（週）作為估計樣本，最後 10 期（週）作為預測樣本；預測樣本之配適結果將作為評估模型預測效度之依據。顧客特質資料包括性別、年齡、擁有手機數、使用費率方案、付費方式等。如表 2 所示，此次自資料庫隨機抽取之 1000 位顧客中，約 2/3 為男性，九成以上的年齡為 30 歲以上，約有一半擁有 2 隻以上手機，八成使用的費率方案為月租 B 型及 C 型（註[1]），六成顧客仍採用自行付費方式。

註[1]：月租方案分為 A、B、C 型及其他；A 型係指月租費抵通話費 88 型，B 型係指月租費抵通話費 188 型，C 型係指月租費抵通話費 288 型。

表 2 顧客特質之百分比分配

性別	(%)	年齡	(%)	手機數	(%)	費率方案	(%)	付費方式	(%)
女	37.9	<30	7.2	1 隻	55.1	A 型	4.7	自行付費	61.8
男	62.1	30~40	28.2	2 隻	25.2	B 型	50.7	銀行代繳	24.6
		40~50	39.4	≥3 隻	19.7	C 型	33.3	其他	13.6
		≥50	25.2			其他	11.3		

本研究以當週通話次數及平均通話秒數等兩個構面定義馬可夫鏈模型之狀態空間，共分為 16 種，如表 3 所示。表中，狀態 1 是通話次數與秒數皆為最少的狀態，狀態 16 則是二者皆為最多的狀態；狀態 4 為通話次數多、通話秒數少的狀態，狀態 13 則完全相反。根據狀態空間之定義，本研究將屬於連續型資料的通話紀錄，轉換為離散型的通話狀態資料。表 4 是估計樣本（共 25,000 筆交易紀錄）於 16 種狀態之次數分配，其中狀態 1 與狀態 16 之次數較高，狀態 4 與狀態 13 之次數較低，代表當週通話次數與平均通話秒數呈正向相關。

表 3 通話狀態空間之定義

平均通話秒數（秒/次）	通話次數（次/週）			
	0~6	7~12	13~18	>18
>60	狀態 13	狀態 14	狀態 15	狀態 16
>45 且≤60	狀態 9	狀態 10	狀態 11	狀態 12
>30 且≤45	狀態 5	狀態 6	狀態 7	狀態 8
>0 且≤30	狀態 1	狀態 2	狀態 3	狀態 4

表 4 通話狀態空間之次數分配

平均通話秒數（秒/次）	通話次數（次/週）				列總和
	0~6	7~12	13~18	>18	
>60	968	1,667	1,481	2,256	6,372
>45 且≤60	770	1,579	1,429	1,959	5,737
>30 且≤45	1,292	2,385	1,888	2,050	7,615
>0 且≤30	1,966	1,810	913	587	5,276
欄總和	4,996	7,441	5,711	6,852	25,000

（二）模型預測效度之比較

本研究首先根據估計樣本（第 1~25 週之通話狀態），透過 BMN 模型、HB Probit 模型、HB Logit 模型等，估計個人化移轉機率矩陣，係一（16×16）矩陣。然後，再以預測樣本（第 26~35 週之通話狀態）評估各個模型之預測效度。樣本外預測的作法是，以顧客於第 25 期（估計樣本之最後一期）的狀態為已知訊息，先根據一階移轉機率矩陣（P_i^1）預測顧客於下一期（即第 26 期）處於各種狀態的條件機率分布，並以機率最大者做為預測狀態；再根據二階移轉機率矩陣（P_i^2）預測下兩期（即第 27 期）的狀態，其他以此類推。本研究以逐期之擊中率（hit rate）做為評估模型預測效度之指標，即預測狀態與實際狀態相符之樣

本數除以總樣本數之比率。如圖 3 所示，三個 BMN 模型之擊中率皆明顯高於 HB Logit 模型與 HB Probit 模型，其中又以樣本資訊權重設為 0.5（w_1=0.5）的 BMN 模型表現較優；HB Logit 模型則比 HB Probit 模型略勝一籌。

造成 BMN 模型優於 HB 條件狀態選擇模型之可能原因有二，包括模型假設及資料特性等。在模型假設方面，HB Logit 與 Probit 模型皆假設外在的狀態選擇乃源自於各狀態之潛伏相對效用的比較結果，但這些被設定遵循 Logistic 分配或多變量常態分配的潛伏相對效用似乎無法確切反映觀察到的狀態選擇；BMN 模型直接以狀態選擇觀察值估計移轉機率矩陣，得到較好的預測結果。在資料特性方面，由於以週為單位的通話紀錄在本質上就缺乏外在因素資料，如通話時點是否為特價時段、通話對象是否為網內互打等，導致 HB Logit 與 Probit 模型可以整合外在因素解釋狀態選擇之優勢無法充份發揮，致使其預測效度不如 BMN 模型。在接下來的顧客價值路徑遷移分析，本研究將採用 w_1=0.5 的 BMN 模型追蹤顧客價值之未來演變。

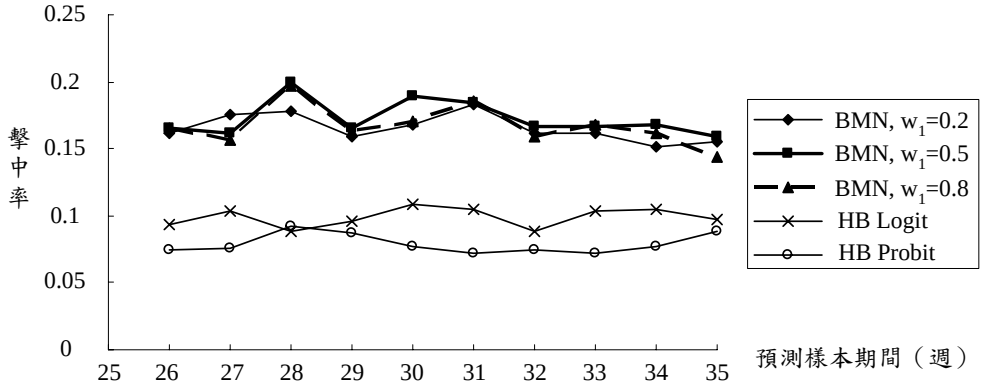
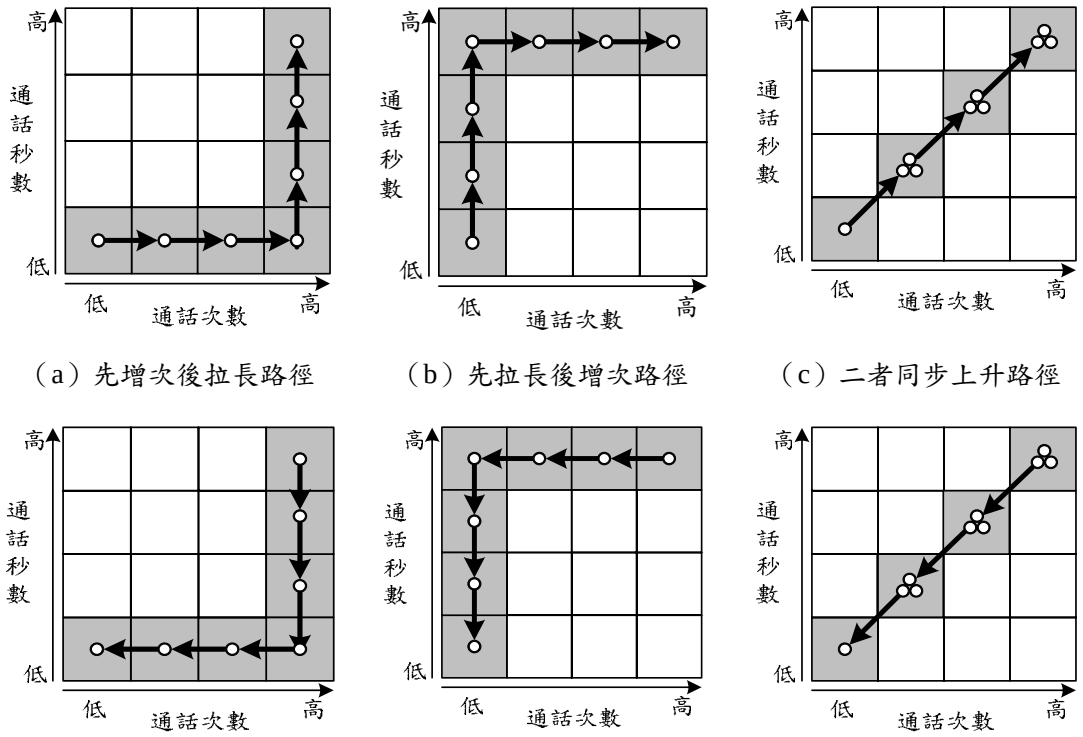


圖 3 不同模型之擊中率比較

（三）顧客價值路徑遷移分析

追蹤顧客價值的未來演變是做好顧客關係管理的重要關鍵。針對不同的顧客價值遷移路徑，廠商宜採取不同的顧客管理方法，以有效提升顧客價值或預防逐漸下降之趨勢。因此，本研究設定數種可能的顧客價值上升或下降路徑探討之，並透過個人化移轉機率矩陣衡量每位顧客遵循該價值遷移路徑的可能性，協助廠商進行有效的一對一顧客關係管理。如圖 4 所示，價值遷移路徑之設定分為上升和下降兩大類，然後根據通話次數、通話秒數及二者同步等再各自細分為三種，共計六種路徑，分別以灰色方格標明。每種路徑皆包含 13 個步驟，圓圈（◦）代表在原來狀態連續停留兩期，箭頭（→）代表由原來狀態移轉至另一個狀態。



（a）先增次後拉長路徑　（b）先拉長後增次路徑　（c）二者同步上升路徑

（d）先縮短後減次路徑　（e）先減次後縮短路徑　（f）二者同步下降路徑

圖 4 顧客價值遷移路徑

以圖 4（a）先增次後拉長路徑為例，其經過之價值狀態先是通話次數逐漸增加，然後才是平均通話秒數之逐漸上升，最後達到最高的顧客價值。本研究設定此一遷移路徑之狀態移轉步驟為：

1◦1→2◦2→3◦3→4◦4→8◦8→12◦12→16◦16

此一路徑從初始狀態（設為狀態 1）開始，歷經 13 個的狀態移轉步驟，包括 7 個停留在原來狀態（◦）及 6 個移轉至不同狀態（→）。這些步驟皆是先增次後拉長路徑可能歷經之狀態移轉情形，本研究將之建構成路徑（a）；其他路徑亦以此類推。比較特殊的是，由於二者同步上升路徑（c）及二者同步下降路徑（f）所歷經的灰色方塊較少，亦即由原來狀態移轉至不同狀態的情況（→）較少，本研究對這兩條路徑刻意增加停留在原來狀態（◦）的情況，使其移轉步驟個數仍然維持 13 個，便於與其他路徑之發生可能性進行比較。

每位顧客遵循不同價值遷移路徑之可能性，乃根據個人化移轉機率矩陣加以計算而得。假設遷移路徑之狀態移轉步驟為 j→k◦k→m，則顧客 i 遵循該路徑之聯合機率的計算如下所示：

式（11）　　$p_i(j,k,k,m) = p_i(j) \times p_i(k|j) \times p_i(k|k) \times p_i(m|k)$

由上式可知，路徑發生機率乃一連串狀態移轉機率之乘積。由於狀態移轉機率本身是介於 0~1 的數值，若路徑設定的步驟愈多，則路徑發生機率將愈小。因此，為能公平比較不同路徑之發生機率，本研究設定所有路徑皆歷經 13 個狀態移轉步驟。而且，由於狀態移轉機率之相乘結果幾乎很小且容易逼近 0，本研究先對於路徑發生機率取自然對數（*ln*）後，才做進一步分析，如下所示：

式（12）　　$ln\,p_i(j,k,k,m) = ln\,p_i(j) + ln\,p_i(k|j) + ln\,p_i(k|k) + ln\,p_i(m|k)$

根據上式，個別顧客遵循六條價值遷移路徑之可能性即可加以計算，有助於廠商瞭解每位顧客傾向遵循何種路徑，並施以適當的顧客管理策略。

本研究以預測力較佳的 BMN（w_1=0.5）模型推估個別顧客之遵循六條路徑之可能性，並以可能性最大者預測顧客之遵循路徑，次數分配如表 5 所示。約有四成顧客被歸為價值上升群，遵循路徑（a）、（b）、（c）；約有六成顧客被歸為價值下降群，遵循路徑（d）、（e）、（f）。遵循路徑（c）者之顧客價值為最高，約佔顧客樣本的一成，廠商應設法留住這些顧客。然而，幾乎有一半的顧客被預測遵循路徑（d），有先減少通話次數再縮短通話時間之行為趨勢；廠商對於這些顧客應先設法預防通話次數的降低，如採用次數折扣之方式。值得注意的是，廠商只能就具有充足交易紀錄的顧客，透過 BMN 模型計算個人化的移轉機率矩陣，並進而預測其顧客價值遷移路徑。對於資料筆數尚少的潛在顧客，廠商僅能就其顧客特質預測其未來可能的通話狀態。因此，廠商有必要探討路徑發生可能性與顧客特質之間的關係，藉此管理潛在顧客。

表 5 顧客價值遷移路徑群

群別	顧客價值遷移路徑	人數	百分比	群別	顧客價值遷移路徑	人數	百分比
價值上升群	（a）先增次後拉長	256	25.6	價值下降群	（d）先減次後縮短	555	55.5
	（b）先拉長後增次	22	2.2		（e）先縮短後減次	3	0.3
	（c）二者同步上升	117	11.7		（f）二者同步下降	47	4.7

本研究以 SUR（Seemly Unrelated Regression）模型探討路徑可能性與顧客特質之間的關係，實證結果如表 6 所示。由表 6 可明顯的看出，路徑（a）（d）、（b）（e）、（c）（f）等三對組合之組內迴歸係數符號與顯著性幾乎完全相同，代表顧客具高度同質性。例如，傾向遵循路徑（a）與（d）的顧客顯著的皆具有男性、50 歲以下、只有一支手機、採用費率高於 C 型方案等特質。這可能是因為本研究在設定六種價值遷移路徑時（如圖 4 所示），路徑（a）（d）、（b）（e）、（c）（f）等三對組合之組內差異僅在於價值遷移方向為往上升或往下降，但路徑所經過的狀態（即圖 4 的灰色方塊）卻是完全相同。

表 6 遷移路徑之顧客特質分析－SUR 模型

遷移路徑 顧客特質	（a） 先增次 後拉長	（b） 先拉長 後增次	（c） 二者同 步上升	（d） 先縮短 後減次	（e） 先減次 後縮短	（f） 二者同 步下降
男性	0.259*	0.332*	0.090	0.197	0.310*	-0.067
<30歲	0.676**	0.191	0.118	0.857**	-0.151	0.043
30~40歲	0.381*	0.232	0.143	0.412**	0.261	0.010
40~50歲	0.348*	0.222	0.293	0.313*	0.378*	0.051
只有一隻手機	0.399**	-0.391*	0.189	0.200	-0.185	-0.042
擁有兩隻手機	0.081	-0.348	-0.363	-0.015	-0.081	-0.318
A型費率方案	-0.046	-1.067**	0.823	0.234	-1.038*	0.648
B型費率方案	-0.625**	-1.443**	-0.022	-0.708**	-1.619**	-0.290
C型費率方案	-0.571**	-1.053**	-0.753**	-0.709**	-1.075**	-0.905**
自行付費	-0.012	0.371	0.322	0.233	0.133	0.405
銀行代繳	0.069	-0.007	0.367	0.085	-0.191	0.456
租約年資（月）	-0.156	-0.077	-0.291	-0.188	-0.284*	-0.159
截距項	-28.333**	-32.324**	-30.583**	-27.913**	-32.122**	-31.455
F值	2.0794**	3.5344**	2.4879**	2.7795**	4.0157**	1.8675**
R^2	0.0247	0.0412	0.0294	0.0327	0.0466	0.0222

**：p 值<0.05，*：p 值<0.1

因此，本研究再以逐步迴歸（stepwise regression）（註[2]）探討配對路徑之可能性差異（註[3]）與顧客特質之間的關係，迴歸係數估計結果如表 7 所示。非自行付款、只擁有一隻手機以及使用非 A 型費率方案的顧客，傾向遵循先增加通話次數再延長通話時間的價值上升路徑，而非反方向的價值下降路徑。同理，未滿 30 歲的顧客傾向遵循先延長通話時間，再增加通話次數的價值上升路徑。只擁有一隻手機、使用 B 型費率方案以及 40~50 歲的顧客則是傾向遵循通話次數與通話時間同步增加的價值上升路徑。

表 7 配對遷移路徑之顧客特質分析－逐步迴歸模型

（括弧內為檢定 p 值）

顧客特質	路徑（a）/（d）	顧客特質	路徑（b）/（e）	顧客特質	路徑（c）/（f）
自行付費	-0.251 (0.020)	未滿 30 歲	0.331 (0.185)	一隻手機	0.259 (0.037)
一隻手機	0.146 (0.168)			B 型費率	0.180 (0.143)
A 型費率	-0.385 (0.120)			40~50 歲	0.170 (0.181)
截距項	-0.117 (0.249)	截距項	0.667 (0.001)	截距項	0.524 (0.001)
F 值	3.200 (0.023)	F 值	1.760 (0.185)	F 值	2.490 (0.059)

註[2]：由於普通迴歸模式之總檢定皆不顯著，本研究改以逐步迴歸衡量變數之間的關係，並設定篩選解釋變數之準則為係數檢定 p 值須小於 0.2。

註[3]：以（a）（d）路徑為例，二者可能性之差異被定義為 *ln*Pr(路徑 a)– *ln*Pr(路徑 d)=*ln*[Pr(路徑 a)/ Pr(路徑 d)]，即兩條路徑發生機率之比值再取對數值。

四、結論

資料庫行銷是在顧客存在「異質性」與「動態性」的前提下，以顧客交易行為為基礎，透過統計模型推估個別顧客的行為模型與特性，並剖析現有顧客和潛在顧客的輪廓，藉此進行一對一行銷並與顧客建立最佳的關係。本研究分別以貝氏多項式模型（BMN 模型）、層級貝氏 Probit 模型（HB Probit 模型）與層級貝氏 Logit 模型（HB Logit 模型）對於顧客交易狀態移轉行為加以探討。藉由模型的建立，我們得以估計資料庫中每位顧客之狀態移轉機率矩陣，並據此預測現有顧客之未來價值，研擬適當的顧客管理策略。根據樣本外預測的結果，BMN 模型（w_1=0.5）是所有模型中擊中率最佳之模型，該模型設定個人化機率的後驗估計值是樣本估計值與總體估計值各佔一半權重的結果。本研究進一步建立六條具行銷意義的顧客價值遷移路徑，用以衡量顧客價值的可能演變；根據個人化移轉機率矩陣，我們得以計算每位顧客遵循這些路徑的可能性，並藉此分析不同路徑的顧客特質，供廠商做顧客管理之參考。

在六條顧客價值遷移路徑中，服從二者同步上升（路徑 c）的顧客最具發展潛力，廠商應設法維繫與這群顧客的長期關係，避免讓顧客有任何轉換品牌的想法（Elliott & Glynn 1998）。對於遵循先增次後拉長（路徑 a）或先縮短後減次（路徑 d）的顧客，廠商應著重於拉長秒數策略，如給予通話時間折扣，當通話時間超過一定秒數後即適用於較低的費率。另外，對於遵循先拉長後增次（路徑 b）或先減次後縮短（路徑 e）的顧客，廠商應著重於增加次數策略，如設計免費通話次數優惠，促使顧客養成多打電話的習慣。遵循二者同步下降（路徑 f）的顧客已漸漸失去顧客價值，廠商應先評估投資的成本效益予以篩選之，對於較具潛力的顧客可持續追蹤管理，至於其餘客戶則不須有特別的行銷活動。

在資料庫行銷的模型建立方面，本研究提供另一個不同的思考方向。綜觀國內外之文獻，關於顧客價值遷移路徑相關主題的文獻較少，缺乏深入的探討；而與馬可夫鏈理論有關的研究，更只有應用在財務、工業分析方面，應用在行銷領域也未獲詳細介紹。本研究嘗試融合馬可夫鏈理論與層級貝氏模型探討個別顧客之交易狀態移轉行為，並加上實際資料庫之分析，企圖透過理論探討與實證研究的結合，能夠對顧客價值遷移路徑分析有更進一步的瞭解與認知。然而，馬可夫鏈理論假設移轉機率矩陣不會因時而變，尚缺乏實證支持。因此，這方面的研究仍需後續研究者繼續投入發展。希望藉由本研究的拋磚引玉，提供後續研究者一個新的研究方向。

參考文獻

Allenby, G. M. & Rossi, P. E. 2000. Statistics and marketing. ***Journal of the American Statistical Association.*** 95: 635-638.

Allenby, G. M., Leone, R. P. & Jen, L. 1999. A dynamic model of purchase timing with application to direct marketing. ***Journal of the American Statistical Association.*** 94 (446): 365-373.

Berger, J. O. 1985. ***Statistical decision theory and bayesian analysis***. Soringer-Verlag, New York, NY: Spriger.

Berger, P. D. & Nasr, N. I. 1998. Customer lifetime value: marketing models and application. ***Journal of Interactive Marketing.*** 12: 17-30.

Berry, M. J. A. & Linoff, G. 2000. ***Mastering data mining: the art and science of customer relationship management.*** New York, NY: John Wiley & Sons.

Blatterberg, R. & Deighton, J. 1996. Manage marketing by the customer equity test. ***Harvard Business Review.*** July-August: 136-144.

Dwyer, F. R. 1989. Customer lifetime valuation to support marketing decision making. ***Journal of Direct Marketing.*** 3 (2): 8-15.

Elliott, G. & Glynn,W. 1998. Segmenting financial services markets for customer relationships: A portfolio-based approach. ***The Service Industries Journal.*** 18(3): 38-54.

Gelfand, A. & Smith A. 1990. Sampling-based approaches to calculating marginal densities. ***Journal of the American Statistical Association.*** 85: 398-409.

Geman, S. & Geman, D. J. 1984. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. ***IEEE Transactions on Pattern Analysis and Machine Intelligence***. 6: 721-741.

Gronroos, C. 2000. ***Service management & marketing: A customer relationship management approach*** (2nd ed.), New York, NY: John Wiley & Sons.

Hastings, W. K. 1970. Monte carlo sampling methods using markov chains and their applications. ***Biometrika***. 57: 97-109.

Hughes, A. M. 1994. ***Strategic database marketing***. Chicago, IL: Probus Publishing Company.

Jen, L. & Wang, S. J. 1998. Incorporating heterogeneity in customer valuation: An empirical study of health care direct marketing in Taiwan. ***International Journal of Operations Quantitative Management.*** 4 (3): 217-228.

Johnson, N. L. & Kotz, S. 1972. ***Distributions in statistics: continuous multivariate distributions.*** New York, NY: John Wiley & Sons.

Kotler, P. 1996. ***Marketing management: Analysis, planning, implementation & control*** (7th ed.), New Jersey, NJ: Prentice-Hall.

Lenk, P. 2001. ***Bayesian inference and markov chain monte carlo.*** Paper presented at the Bayesian Applications and Methods in Marketing Conference and Tutorial, Columbus.

Mcfadden, D. 1973. Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), ***Frontiers in Economics.*** New York, NY: Academic Press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. 1953. Equations of state calculations by fast computing machines. ***Journal of Chemical Physics.*** 21: 1087-1092.

Peppers, D. & Rogers, M. 1993. ***The one to one future: Building relationships one customer at a time***. New York, NY: Dubleday/Currency.

Pfeifer, P. E. & Carraway, R. L. 2000. Modeling customer relationships using markov chains. ***Journal of Interactive Marketing.*** 14 (2): 43-55.

Raphel, M. & Raphel, N. 1995. ***Up the loyalty ladder: Turning sometime customers into full-time advocates of yours business.*** New York, NY: Harper Business.

Rossi, P. E. & McCulloch, R. E. 1994. An exact likelihood analysis of the multinimial probit model. ***Journal of Econometrics***. 64: 207-240.

Rossi, P. E., Allenby, G. M. & McCulloch, R. E. 2005. ***Bayesian statistics and marketing.*** Hoboken, NJ: John Wiley & Sons.

Schijns, J. M. C. & Schroder, G. J. 1996. Segment selection by relationship strength. ***Journal of Direct Marketing.*** 10: 69-79.

Sewell, C. & Browm, P. 1990. ***Customer for lift: How to turn that one-time buyer into a lifetime customer.*** New York, NY: Doubledat/Currency.

Stone, B. 1995. ***The successful direct marketing.*** Lincolnwood, IL: NTC Business Books.

附錄

（一）層級貝氏 Probit 模型

層級貝氏統計方法對於參數後驗分配之推導分為兩大步驟：先是推導所有參數之完整條件後驗分配（full conditional posterior distribution）；然後以 MCMC 方法模擬產生可能的後驗估計值，再以模擬數值的相對次數分配逼近真正的後驗分配，平均數做為後驗的點估計值（Rossi, Allenby & McCulloch 2005）。以第 j 個 HB Probit 模型為例，假設顧客 i 有 $L_{i|j}$ 筆購買紀錄適合投入模型估計，且樣本模型之設定如式（8），參數模型之設定如式（9），則這兩個模型所有參數之完整條件後驗分配的推導說明如下。

1. 式（8）的參數：平均相對效用向量（$\mu_{i|j}$）

此一向量包含 S^* 個平均相對效用參數（令 $S^*= S-1$），其樣本概似函數由式（8）潛伏相對效用向量（$y^*_{it|j}$）之聯合常態分配（Multivariate Normal Distribution, MN）所構成，先驗分配來自於式（9）的設定，故條件後驗分配之推導如下所示：

$$[\mu_{i|j}|\, y^*_{it|j}, \text{rest}] \propto \prod_{t=1}^{L_{i|j}} MN_{S^*}\left(y^*_{it|j} \mid \mu_{i|j}, \Sigma_j\right) \times MN_{S^*}\left(\mu_{i|j} \mid \Theta'_j z_{i|j}, \Lambda_j\right)$$

$$\propto\ MN_{S^*}\left(\mu_{i|j} \mid u_{i|j}, V_{i|j}\right)$$

其中，$\mu_{i|j} = V_{i|j} \cdot \left(\sum_{t=1}^{L_{i|j}} \Sigma_j^{-1} y^*_{it|j} + \Lambda_j^{-1}\Theta'_j z_{i|j}\right)$；$V_{i|j} = \left(L_{i|j}\Sigma_j^{-1} + \Lambda_j^{-1}\right)^{-1}$，rest 代表其餘所有參數。

2. 式（8）的參數：誤差項之共變數矩陣（Σ_j）

此一矩陣係一（$S^* \times S^*$）之方陣，其樣本概似函數由式（8）之潛伏相對效用向量（$y^*_{it|j}$）之聯合常態分配所構成，先驗分配則通常設定為 Inverse Wishart 分配（簡稱 IW 分配），則條件後驗分配之推導如下所示：

$$[\Sigma_j|\, y^*_{it|j}, \text{rest}] \propto \prod_{i=1}^{N_j}\prod_{t=1}^{L_{i|j}} MN_{S^*}\left(y^*_{it|j} \mid \mu_{i|j}, \Sigma_j\right) \times IW_{S^*}\left(\Sigma_j \mid s_0, R_0^{-1}\right)$$

$$\propto IW_{S^*}\left(\Sigma_j \mid s_j, R_j^{-1}\right)$$

其中，$s_j = s_0 + \sum_{i=1}^{N_j} L_{i|j}$，$R_j^{-1} = R_0^{-1} + \sum_{i=1}^{N_j}\sum_{t=1}^{L_{i|j}} \left(y^*_{it|j} - \mu_{i|j}\right)\left(y^*_{it|j} - \mu_{i|j}\right)'$。另外，本研究設定純

先驗參數 $s_0= S^*+2$，$R_0^{-1}=I_{S^*}$（單位矩陣）。

3. 式（9）的參數：迴歸係數矩陣（Θ_j）

Θ_j 係一（$P\times S^*$）矩陣，衡量 P 個顧客特質變數（$z_{i|j}$）與第 j 個 HB 條件狀態模型之 S^*個平均相對效用（$\mu_{i|j}$）之間的關係。根據式（9），Θ_j 之樣本概似函數由平均相對效用向量（$\mu_{i|j}$）之聯合常態分配所構成，先驗分配則可搭配設定為多變量常態分配。為便於推導後驗分配，令 $\Theta_j^*=\text{vec}(\Theta_j)$，將迴歸係數矩陣予以向量化；並令 $\mu_j^{*\prime}=[\mu_{1|j}\ \ldots\ \mu_{Nj|j}]$，將所有 N_j 位顧客的平均效用向量推疊成一（$N_jS^*\times1$）向量，改寫樣本概似函數，則條件後驗分配之推導如下：

$$[\Theta_j^*|\mu_j^*,z_i,\text{rest}]\propto MN_{N_jS^*}\left(\mu_j^*|\left(Z_j\otimes I_{S^*}\right)\Theta_j^*,\ I_{N_j}\otimes\Lambda_j\right)\times MN_{PS^*}\left(\Theta_j^*|u_0,V_0\right)$$

$$\propto MN_{PS^*}\left(\Theta_j^*|u_j,V_j\right)$$

其中，$u_j= V_j\cdot\left[\left(Z_j'\otimes\Lambda_j^{-1}\right)\mu_j^*+V_0^{-1}u_0\right]^{-1}$，$V_j=\left(Z_j'Z_j\otimes\Lambda_j^{-1}+V_0^{-1}\right)^{-1}$，二者皆是後驗分配之參數推導結果。另外，本研究設定純先驗參數 $u_0= 0$，係一零向量；$V_0=100I_{Nj}$，二者之設定皆在加重樣本資訊的權重，降低純先驗分配對條件後驗分配的影響。

4. 式（9）的參數：誤差項之共變數矩陣（ Λ_j）

Λ_j 係一（$S^*\times S^*$）之方陣，其樣本概似函數由式（2）的平均相對效用向量（$\mu_{i|j}$）之聯合常態分配所構成，先驗分配同樣設定為 Inverse Wishart 分配，則條件後驗分配之推導如下所示：

$$[\Lambda_j|\mu_{i|j},\text{rest}]\propto\prod_{i=1}^{N_j}MN_{S^*}\left(\mu_{i|j}|\Theta_j'z_{i|j},\Lambda_j\right)\times IW_{S^*}\left(\Lambda_j|f_0,G_0^{-1}\right)$$

$$\propto IW_{S^*}\left(\Lambda_j|f_j,G_j^{-1}\right)$$

其中，$f_j=f_0+N_j$，$G_j^{-1}=G_0^{-1}+\sum_{i=1}^{N_j}\left(\mu_{i|j}-\Theta_j'z_{i|j}\right)\left(\mu_{i|j}-\Theta_j'z_{i|j}\right)'$。此處，本研究設定先驗分配之參數 $f_0= S^*+2$，$G_0^{-1}=I_{S^*}$。

5. MCMC 方法

在貝氏統計中，所有參數（註[4]）皆被視為隨機變數，故個別參數之條件後驗分配函數必須先對其餘參數（rest）進行多重積分，方能成為真正的後驗分配。

註[4]：不包含純先驗參數，如Σ_j之先驗分配參數 s_0和 G_0 由研究者自行設定，不被視為隨機變數。

MCMC 方法是最常被用來解決多重積分問題的統計工具。MCMC 方法是一種抽樣基礎（sampling-based）方法，根據馬可夫鏈性質模擬（simulate）出一組樣本，其近似分配（approximate distribution）即為邊際後驗分配，而不必直接對條件後驗分配進行多重積分。以 HB Probit 模型為例，一旦給定所有參數之初始值（initial value），即可根據式（8）隨機產生符合值域限制的潛伏相對效用數值，連同其餘參數初始值代入特定參數之條件後驗分配，就可隨機產生該參數之候選值（candidate），再代入其他條件後驗分配產生其他參數之候選值。經過反覆的參數代入及隨機抽樣足量的候選值之後，再產生的候選值將漸近遵循真正的後驗分配（註[5]），而這些候選值之平均即為貝氏統計之後驗點估計值。此種根據完整條件後驗分配之抽樣候選值逼近後驗分配的方法，又稱為吉布斯抽樣方法（Gibbs Sampling）（Geman & Geman 1984; Gelfand & Smith 1990）。

6. 狀態機率預測

HB Probit 模型之機率預測也涉及多重積分的問題，本研究亦以模擬方法解決此一問題。在給定個人化參數後驗估計值（$\mu_{i1|j},\ldots,\mu_{i,S-1|j}$）之下，即可根據式（8）隨機產生（S–1）個相對效用數值（$y^*_{i1|j},\ldots,y^*_{i,S-1|j}$）；再根據這些相對效用數值的相對大小，預測顧客之所處狀態。在重覆進行多次此一模擬過程之後，以各種狀態之相對次數，做為狀態機率之預測值（註[6]）。

（二）層級貝氏 Logit 模型

式（10）是 HB Logit 模型之第一層樣本模型；根據顧客的狀態選擇紀錄所建構之樣本概似函數如下所示：

$$L(\mu_{i|j})=\left(\frac{\exp\left(\mu_{i1|j}\right)}{\sum_{k=1}^{S}\exp\left(\mu_{ik|j}\right)}\right)^{F_{ij1}}\times\left(\frac{\exp\left(\mu_{i2|j}\right)}{\sum_{k=1}^{S}\exp\left(\mu_{ik|j}\right)}\right)^{F_{ij2}}\times\cdots\times\left(\frac{\exp\left(\mu_{iS|j}\right)}{\sum_{k=1}^{S}\exp\left(\mu_{ik|j}\right)}\right)^{F_{ijS}}$$

式中的 F_{ijk} 同表 1 之定義，即顧客 i 連續兩期狀態為（j,k）的發生次數。此一概似函數無法找到自然共軛的先驗分配，故二者之乘積無法被推導成一具封閉型式的條件後驗分配，吉布斯抽樣也就無法被用以產生後驗估計值。在此情況下，MCMC 法中的梅托普利斯-海斯丁運算法（Metropolis-Hastings Algorithm, M-H）是最常用以推導後驗分配之方法。M-H 法由 Metropolis 等人（1953）所發展，其後由 Hastings（1970）予以一般化。做法是先產生參數準候選值，再將該值代

註[5]：本研究設定先進行 5000 次的反覆抽樣之後，再進行 1000 次的反覆抽樣，並以這 1000 次的抽樣結果的平均做為參數之後驗點估計值。

註[6]：本研究設定重覆進行 1000 次的模擬。

入概似函數（如上式）與先驗分配（式 9），計算二者機率密度之乘積；若乘積高於門檻值，則該準候選值留下成為候選值，否則就予以淘汰並以上次候選值取代之。在經過足量的反覆選取準候選值的過程之後，再留下的參數候選值將逼近真正的後驗分配（註[7]）。詳細過程請參考 Lenk（2001）。除了第一層模型之外，HB Logit 模型的其餘部份之設定與推導，與 HB Probit 模型完全相同。

註[7]：本研究設定先進行 10000 次的反覆選取過程，再以後續 1000 次的選取結果的平均做為後驗點估計值。