

金融風險管理季刊
民 96，第三卷，第二期，1-30

信用卡使用者之違約風險研究—存活分析模型之應用^{*}

Credit Analysis of Credit Card Holders The Application of Survival Model

葉玫惠^{**}

Yeh, Clover M

台灣大學財務金融研究所博士候選人
Ph.D. Candidate, Graduate Institute of Finance,
National Taiwan University

張靖宜

Chang, Ching-Yi

正新橡膠工業股份有限公司財務專員
Financial Specialist, Cheng Shin Rubber Ind.
Co., Ltd.

廖咸興

Liao, Hsien-Hsing

台灣大學財務金融研究所副教授
Associate Professor, Department and Graduate
Institute of Finance, National Taiwan University

周國端

Jou, David G

宏泰人壽董事長暨台灣大學財務金融研究所
兼任副教授
Chairman, Hontai Life Insurance Co., Ltd. and
Adjunct Associate Professor, Department and
Graduate Institute of Finance, National Taiwan
University

摘要

目前，國內外關於信用卡使用者之違約風險評估，主要皆在固定時點判定未來一段特定期間內是否會發生違約（如區別分析）或發生違約的機率（如 Logit 模型以及 Probit 模型），無法提供持卡人在未來不同時點的違約機率（或存活率）。本研究應用在醫學及精算領域廣為使用的存活分析（Survival Analysis），透過與信用卡使用者違約相關的可能因素，建立預警模型及或存活率表，使銀行能以客觀的方式預估客戶未來各時點發生違約的機率，降低處理違約的後續成本。台灣信用卡資料的實證結果顯示，存活分析法於其具有實務上之應用性。

關鍵字：存活分析、多期信用風險模型、信用卡違約風險、存活率表

JEL 分類代號：G21

^{*} 感謝三位匿名審稿人寶貴的建議，使本文的內容更為充實。

^{**} 聯繫作者：葉玫惠，台灣大學財務金融研究所，E-mail: r91723060@ntu.edu.tw

Abstract

Most current methods on credit analysis of credit card holders are only able to predict the default (such as the discriminant analysis) or default probability (such as the Logit or the Probit model) in a period time of future time at a specific point of time. These methods cannot provide the information on the default probability in different time points in the future. Employing the survival analysis that is extensively used in medical and actuarial science, a pre-warning model and a predictable survival table for different groups of cardholders are established. Banks can use the pre-warning model or the survival table to estimate the default probability of a cardholder in future time points and therefore can reduce the costs for handling defaulted cardholders. The empirical analyses of Taiwan credit card data provide evidence supporting that the survival analysis is quite applicable.

Key Words: Survival Analysis, Multi-period Credit Risk Model, Credit Card Default Risk, Survival Table

JEL Classification: G21

1. 導論

近年來利率持續維持低檔，銀行業的傳統業務利潤微小，使得新種業務嶄露頭角，其中又以信用卡市場表現更為亮眼。在台灣使用信用卡消費的習慣已經非常普遍，流通在外的信用卡張數及刷卡金額也是逐年攀升，前者數據由民國八十五年底的546.7萬張急速竄升至民國九十四年底的4,549.4萬張，短短十年內就成長了將近九倍；在同期間內後者數據也從2,724億元增加至1兆4,210億元，成長率超過五倍。然而，持續膨脹的信用卡市場卻為鄰國南韓帶來一場可怕的金融災難，韓國最大信用卡發卡公司－LG信用卡公司，曾經三度發生現金週轉困難而中斷現金提取之服務，

幾近破產邊緣¹。國內信用卡市場的情況，自民國九十四年九月開始，卡債問題逐漸引起注意，而銀行如何控制相關風險，建立一套有效的信用卡違約風險之預警系統，為做好風險管理的重要基礎，也受到極大的重視。

過去國內外關於信用卡違約風險評估的相關文獻，以區別分析法 (Discriminant Analysis)、Logit 或 Probit 模型、類神經網路 (Neural Networks) 及資料探勘 (Data Mining) 為主，其中最為廣泛運用者為區別分析法及 Logit 模型。以區別分析作為分析工具的文獻有 Hayhoe et al. (1999) 和曾俊堯 (1995)；以 Logit 模型分析信用卡違約風險的文獻有 David (2001)、Thomas et al. (2002)、Noh, Roh and Han (2005) 和龔昶元

¹ 陳志強 (2004)，亞洲金融的轉型發展與風險－韓國信用卡公司財務危機探討，台灣經濟研究月刊第27卷第3期。

(1998)；以類神經網路作為分析工具的有 David (2001)、Noh et al. (2005)；而採用資料探勘方法的有 David and Gordon (2001)。Noh et al. (2005) 首度將信用卡違約風險與存活分析法 (Survival Analysis, Cox (1972)) 作結合，實證比較 Logit 模型、類神經網路及存活分析法運用於信用卡違約風險評估上，發現相較於 Logit 模型與類神經網路模型，存活分析模型不但有較小的型一錯誤，更包含較多的顯著變數。

雖然過去很少文獻將存活分析法與信用卡議題結合，但存活分析法在其他信用風險議題上均被廣泛運用，Lane et al. (1986) 運用存活分析法預測銀行違約機率；Keasey et al. (1990) 及 Luoma and Laitinen (1991) 運用存活分析法預測公司違約機率；Lando (1994) 則運用存活分析法預測債券違約機率；Lee and Urrutia (1996) 運用存活分析法預測產險公司違約機率；郭志安 (1997) 以台灣地區股票上市公司之財務資料分別建立危機發生前一年及前二年之預警模式；鄭婷月 (2003) 將存活分析法運用於汽車貸款客戶之風險研究；鍾岳昌 (2004) 則以比例危險模型估計房貸借款人提前清償及違約風險。

關於討論信用卡違約風險的研究，過去主要皆在固定時點判定未來一段特定期間內是否會發生違約（如區別分析法）或發生違約的機率（如 Logit 模型及 Probit 模型）。區別分析法假設變數為常態分配，其結果看不出程度上的差別（只有違約或不違約）；而 Logit 模型以及 Probit 模型之信

用評分方法，改進了區別分析法對於處理名目變數和分配假設上的缺點，但是仍無法提供持卡人在未來不同時點的違約機率（或存活率）。針對此問題，本文以醫學領域及精算領域廣為使用的存活分析法為基礎，建立一完整之信用卡違約風險預測模型，並對我國信用卡資料進行研究，確認其於實務上的應用性。另外，為加強實務上之可操作性，本研究除了建立信用卡之存活分析模型外，亦發展一套建立存活率參照表 (Survival Table) 及使用存活率表的方法，使得原本很複雜的存活分析法得以簡易的應用於金融機構的信用卡風險預測的操作上。此存活率表的使用方式類似保險上廣為運用的生命表，金融機構只要將信用卡使用者依其分數（由顯著變數值加總而得）來找出其風險組別，再參照存活率表，即可知此客戶在未來各時點的違約率（或存活率）。

與其他模型相較，存活分析的主要優點在於可以客觀的預測客戶未來各時點發生逾期或呆帳的機率，因此可以在適當時機做出處置。其他模型（如 Logit 模型）雖然亦能預測未來一段時間內發生違約的機率，但由於無法預測接近發生違約的時點，因此容易過早將尚未有違約風險的客戶列為危險客戶，如此容易損失客戶；相較而言，存活分析可以預測出接近違約的時點，因此，在預測出某客戶在未來的某個時點有高違約風險的情況下，可將之列為觀察對象，在接下來的時間以更新的資料來重新評估，做出適當處置。由於存活

分析可以預估客戶在未來各時點發生違約的機率，因此可以知道有多少時間來重新評估，既不會過早處置尚未有違約風險的客戶，也不會因為花不適當的時間重新評估而錯失處置時機。

本研究利用財團法人金融聯合徵信中心（簡稱金融聯合徵信中心）所蒐集之資料庫作為分析對象，透過與信用卡違約風險相關的可能因素建立存活率預警模型並與常用之Logit模型效力作比較，期能展示本研究模型之實務上之應用性與優越性，並對銀行信用風險管理模式提供建議。

本研究分為三部分，首先，討論相關研究方法，主要介紹Cox存活分析模型與說明所選擇的參數，並說明用來簡化預測變數數量的主成分分析法(Principal component analysis)；次之，為實證分析，運用國內信用卡資料建構可操作模型以及產出存活率表，並進行模型檢測；最後為結論與建議。

2. 研究方法

2.1 Cox 存活分析模型

由於違約的情況通常發生在信用卡使用者持有一段時間之後才會發生，運用存活分析法正好可以預測違約發生時點的機率。相較於過去相關文獻常採用的區別分析法、Logit或Probit模型，此法能提供未

來不同時點之存活機率預測，並可估算發生事件的機率值與估計未來存活時間，另外，根據Noh et al. (2005)的研究指出²，相對於其他模型，此法有較低的型一錯誤。存活分析法又稱作「危險模型」(Hazard Model)，亦稱作「存續期間模型」(Duration Model)，或稱作「Cox模型」。Cox模型的應用，以加速失敗時間模型(Accelerated Failure Time Model)及比率危險模型(Proportional Hazard Model, PHM)較為常見，前者強調的是一個停留狀態下的時間與存活函數的關係；而後者指的是個體間的危險函數呈比率關係，此模型具有「半參數」估計的特性，故較常於文獻中使用，本文亦採用此模型。以下為Cox存活分析模型之比率危險模型的說明。

自觀測樣本開始至樣本發生違約為止，這段期間稱為存活時間(Survival Time)。若樣本資料在研究區間內發生違約，我們稱該樣本之為完整資料(Completed Data)，這種資料的存活時間的始末都能被清楚地觀察；反之，若樣本資料存活時間的起點未知(左設限資料 / Left Censored Data)或者是終點未知(右設限資料 / Right Censored Data)，則都被視為是不完整的資料(Uncompleted Data)或是稱為設限資料(Censored Data)。若信用卡持有人的存活時間 T 超過某一時點 t ，則其存活率以存活函數(Survivor Function)表示為

² Noh, H. J., Roh, T. H., and Han I., 2005, Prognostic Personal Credit Risk Model Considering Censored Information. Expert Systems with Applications, 28, 753-762

$$S(t) = P(T \geq t) = 1 - P(T < t) = 1 - F(t) \quad (1)$$

其中，存活函數 $S(t)$ 為存活時間超過時間 t 的機率。而單位時間的違約機率則以風險函數(Hazard Function)作為衡量，表示在某一時點 t 依然存活的情況下，該樣本在下一刻 $(t + \Delta t)$ 違約的單位時間之條件機率：

$$h(t) = \lim_{\Delta t \rightarrow 0^+} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} \quad (2)$$

存活函數與風險函數的關係為：

$$S(t) = \lim_{\Delta \tau_k \rightarrow 0} \prod_{k=0}^{r-1} [1 - h(\tau_k)(\tau_{k+1} - \tau_k)] \quad (3)$$

其中 $0 = \tau_0 < \tau_1 < \dots < \tau_{r-1} < \tau_r = t$

$$\begin{aligned} \Rightarrow S(t) &= e^{-\int_0^t h(u) du} \\ \Rightarrow h(t) &= \frac{-S'(t)}{S(t)} \end{aligned}$$

Cox 模型將風險函數 $h(t; \mathbf{x})$ 設為：

$$h(t; \mathbf{x}) = h_0(t) \exp(\beta' \mathbf{x}(t)) \quad (4)$$

其中， $\mathbf{x}(t)$ 為在時間 t 所選取對風險發生有影響的解釋變數， β 為各變數項的係數， $h_0(t)$ 表 $\mathbf{x}(t) = 0$ 時的基準風險函數 (Baseline Hazard Function)，若解釋變數不是時間的函數，則風險函數會改寫成：

$$h(t; \mathbf{x}) = h_0(t) \exp(\beta' \mathbf{x}) \quad (5)$$

由於在信用卡違約風險分析中，各項變數應為時間相依，因此本研究採用 (4) 式。

Cox 模型採取最大概似法來估計參數 β 。如果觀察到一位信用卡使用者 i 在時間 t_i 違約，則此樣本所包含的訊息佔整個風險集合所包含訊息的比例為：

$$\frac{\exp\{\beta' x_i(t_i)\}}{\sum_{l \in R\{t_i\}} \exp\{\beta' x_l(t_i)\}} \quad (6)$$

其中， $R\{t_i\}$ 為在 t_i 的整個風險集合， $x_i(t_i)$ 則是信用卡使用者 i 在時間 t_i 的解釋變數值。假設在觀察期間共觀察到 n 個信用卡持有者違約，違約時間依序為 t_1, t_2, \dots, t_n ，將風險訊息取對數後加總，可得概似函數(Log-Likelihood Function)為

$$L(\beta) = \sum_{i=1}^n \left(\beta' x_i(t_i) - \log \left[\sum_{l \in R\{t_i\}} \exp\{\beta' x_l(t_i)\} \right] \right) \quad (7)$$

解其一階及二階微分式，即可得各 β 值。

Cox 模型的適合度檢定是以 LR-test 做檢定：

$$\chi_{LR}^2 = -2 \log \frac{(\text{max. likelihood without the variable})}{(\text{max. likelihood with the variable})} \quad (8)$$

若 $\chi_{LR}^2 > \chi_{\alpha}^2(\nu)$ ，則拒絕 $H_0: \beta = 0$ ，其中 ν 為自由度。另外，此檢定方法可適用於整個模型以及個別係數 β 的檢定。

2.2 主成份分析法

在 Cox 模型的投入的解釋變數間若具有線性相依(linear dependent)的特性，便無法進一步作存活分析。利用主成份分析法可轉換原始投入變數使之成為一些互相獨立的主成分，而這些主成分為原始變數之線性組合。雖然大多數情況，最後選取的主成分數目較原始變數少，但仍保有原始變數大多數的資訊。詳言之，主成份分析法是考慮將資料中原有的 N 個變數作線性組合得到 M 個相互獨立的新變數（即主成分），為了達到資料簡化的目的，通常 M 比 N 小很多。而如何選取 M ，則必須考慮這 M 個新變數對原有的 N 個變數之「變異數—共變數結構」能有多少解釋能力而定。³ 當原來的變數轉為主成份後，便可將各變數的原始分數轉為主成份分數(Score)，以供進一步的統計分析。

假設原有 N 個數值變數，則可計算出 M 個主成份，每一個主成份是原始變數的線性組合：

$$Y_j = a_{j1}X_1 + a_{j2}X_2 + \dots + a_{jN}X_N \quad (9)$$

$$j = 1, 2, \dots, M$$

其中 Y_j 為第 j 個主成份， X 為原始變數， a_{ji} 則是在第 j 個主成份中第 i 個原始變數所佔的權重。由於在主成份分析中，有多少個原始變數便有多少個主成份，所以共同性(Community)會等於 1，亦即沒有誤差項 ϵ_j 。給定權重的方式可以是主觀的人為訂定，也可以是客觀地透過資料本身所傳達的訊息而訂定權重。主成份分析就是要將重要的變數適當地給予較大的權重，不重要的變數給予較小的權重。

使組合 $y = a_1x_1 + a_2x_2 + \dots + a_Nx_N$ 在 $a'a=1$ 下 $Var(y) = a' \Sigma a$ 最大的解 a 是矩陣 Σ 的最大特徵值(Eigenvalue) λ_1 所對應的特徵向量 (Eigenvector)。特徵向量通常取單位長度，且將主成份依照特徵值由大到小排列。其中 Σ 是隨機向量 $x = (x_1, x_2, \dots, x_p)$ 的共變異數矩陣，即 a 為矩陣 Σ 的特徵向量 λ 。在實務上， Σ 通常是未知的，此時可用抽樣資料求出樣本共變異數矩陣 S 來代替 Σ ；若變數的單位不同，可將資料先標準化後再作主成份分析，此時，可用相關矩陣 R 取代樣本共變異數矩陣 S 。模型建立所保留的成份之選擇，依 Kaiser 的建議保留特徵值大於 1 之成份。

經主成份萃取後可得到每個主成份下每個變數的權重，如此就可得到每筆資料經過主成份轉換後的主成份分數。主成份

³ 亦即將幾個變數予以線性組合，使經由線性組合而得到的成份之變異數為最大，使觀測值在這些成份上顯示出最大的個別差異來。為使變異數達到最大，通常主成份分析是不加以轉軸的。此外，主成份分析還可用來建構多種具有不同衡量單位變數之綜合指標，主成份亦可根據其負荷量來對主成份命名，決定多個變數的加權數值，依此訂出總指標。

分數的計算公式如下：

$$y_j = a_{j1}(x_{1i} - \bar{x}_1) + a_{j2}(x_{2i} - \bar{x}_2) + \dots + a_{jN}(x_{Ni} - \bar{x}_N) \quad j=1,2,\dots,M \quad (10)$$

即第 j 筆資料的第 j 個主成份，其中 \bar{x} 為 x 變數的平均數。

2.3 研究資料

本文之研究對象為金融聯合徵信中心所蒐集之信用卡資料隨機抽樣十萬筆匿名樣本⁴(其中八萬筆為訓練組 (Training Set)，兩萬筆為測試組 (Test Set))，以觀測基準點 (M) 前十二個月(績效表現期(Performance Period)：M-12~M-1) 的信用卡消費、繳款及其它紀錄作為參數⁵，以客觀、量化的方式建立模型，對未來十二個月(預測期 (Outcome Period)：M+1~M+12)內的違約

機率與違約時點做估計。圖 1 為資料觀測期間與預測時間的示意圖。由於受限於資料的可獲得性，本研究觀測基準點僅選取 2004 年 12 月 31 日、2005 年 01 月 31 日、2005 年 02 月 28 日、2005 年 03 月 31 日、2005 年 04 月 30 日、2005 年 05 月 31 日等六個觀測基準時點的抽樣資料進行分析；另外，對於樣本外的觀測時點則是選取 2005 年 12 月 31 日以及 2006 年 03 月 31 日兩個時點各兩萬筆樣本，分別依 2004 年 12 月 31 日及 2005 年 03 月 31 日建立之 Cox 模型進行驗證，其績效表現期一樣是觀測基準點前十二個月的期間，而預測期則是從觀測基準時點起算截至 2006 年 06 月 30 日為止分別有六個月以及三個月的期間。

由於個人信用狀態可能會隨時間變動，因此，金融機構需每隔一段時間即更新一次存活分析模型與存活率表，以確保預測之準確性。

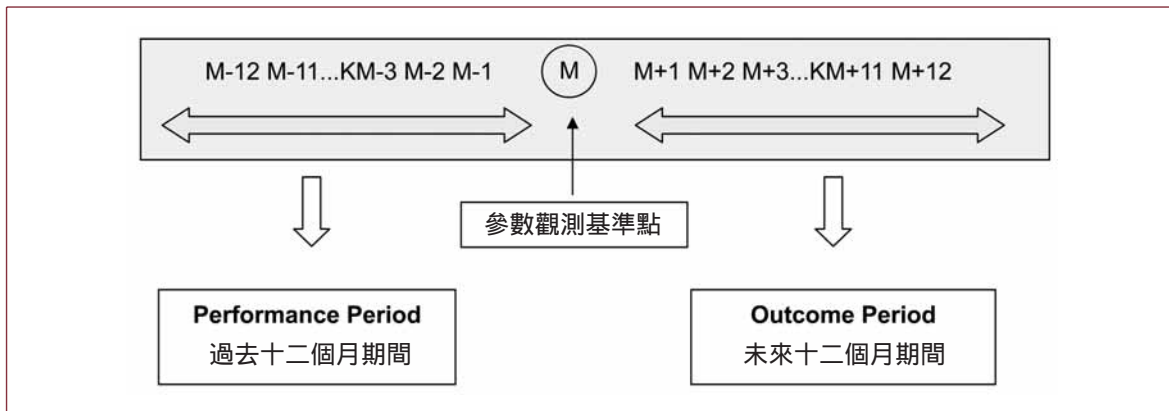


圖 1 資料觀測與預測期間示意圖

⁴ 金融聯合徵信中心資料是以一位信用卡持有人的資料作為一個樣本，因此所有會員金融機構皆可利用信用卡持有人在所有銀行的綜合消費及繳款狀況來對其信用風險做評估。

⁵ 個人其他信用資料不納入本研究當中，例如個人貸款資料、票信資料、銀行存款、資產狀況等。

2.4 投入變數

根據過去相關研究以及資料的可獲得性，本研究初步投入的變數分為人口資料變數、持卡人歷史資料變數及持卡人行為變數，其中，持卡人行為變數包括時間及統計量兩個維度：時間維度包括當月份資料、短期資料（最近三個月）、短中期資料（最近六個月）、中期資料（最近十二個月）；而統計量則包括了總次數、平均數、最大值與標準差。另外，本研究增加持卡人行為變化參數(增加量)，利用過去三個月的統計量與過去六個月的統計量之差作為短期與短中期行為差異的描述；過去六個月的統計量與過去十二個月的統計量

之差則是作為短中期與中期的行為差異描述。詳細項目如下方表 1 所示，由表中的 47 個項目以及各種時間維度、統計量和增加量的組合可以獲得 484 個變數。由表中的 50 個項目以及各種時間維度、統計量和增加量的組合可以獲得 487 個變數。由表中的 50 個項目以及各種時間維度、統計量和增加量的組合可以獲得 487 個變數。

上述變數中，有 4 項名目變數，分別為編號 1 的性別、編號 3 的是否有自用住宅、編號 5 的教育程度代號、以及編號 33 的本月信用卡繳款中狀況最差代號。各項名目類別分列於附錄之附表 1 至附表 4。

表 1 變數參照表

項目	Label	Duration	Statistic Value	Increment
1	性別 N		N	N
2	年齡 N		N	N
3	是否有自用住宅 N		N	N
4	年薪_千元	N	N	N
5	教育程度代號 N		N	N
6	最近一次強制停卡距資料時點之月數	,6,12個月	Mean	增加量
7	本月有效主卡新增張數	,6,12個月	Mean	增加量
8	本月有效主卡新增家數	,6,12個月	Mean	增加量
9	本月非強停之有效主卡停卡張數	,6,12個月	Mean	增加量
10	本月非強停之有效主卡停卡家數	3,6,12個月	Mean	增加量
11	本月一般停卡之有效主卡停卡張數	3,6,12個月	Mean	增加量
12	本月一般停卡之有效主卡停卡家數	3,6,12個月	Mean	增加量
13	本月掛失停卡之有效主卡停卡張數	3,6,12個月	Mean	增加量
14	本月掛失停卡之有效主卡停卡家數	3,6,12個月	Mean	增加量
15	本月有效主卡家數	3,6,12個月	Mean	增加量
16	本月有效主卡張數	3,6,12個月	Mean	增加量
17	本月有效頂級卡主卡張數	3,6,12個月	Mean	增加量

表 1 變數參照表 (續)

項目	Label	Duration	Statistic Value	Increment
18	本月有效白金卡主卡張數	3,6,12個月	Mean	增加量
19	本月有效金卡主卡張數	3,6,12個月	Mean	增加量
20	本月有效普卡主卡張數	3,6,12個月	Mean	增加量
21	本月信用卡原始信用核定額度	3,6,12個月	Mean	增加量
22	本月信用卡繳款紀錄之信用核定額度	3,6,12個月	Mean, Max, Std.	增加量
23	本月信用卡應付帳款金額	3,6,12個月	Mean, Max, Std.	增加量
24	本月信用卡循環信用金額	3,6,12個月	Mean, Max, Std.	增加量
25	本月信用卡有繳款紀錄的家數	3,6,12個月	Mean, Max, Std.	增加量
26	本月信用卡應繳金額大於零的家數	3,6,12個月	Mean, Max, Std.	增加量
27	本月信用卡無繳款金額的家數(X)	3,6,12個月	Mean, Max, Std.	增加量
28	本月信用卡全額繳清的家數(AB)	3,6,12個月	Mean, Max, Std.	增加量
29	本月信用卡使用循環信用的家數(CD)	3,6,12個月	Mean, Max, Std.	增加量
30	本月信用卡無繳足最低金額或全額逾期末繳的家數(EF)	3,6,12個月	Mean, Max, Std.	增加量
31	本月信用卡繳款有延遲的家數(BDEF)	3,6,12個月	Mean, Max, Std.	增加量
32	本月信用卡預借現金餘額家數	3,6,12個月	Mean, Max, Std.	增加量
33	本月信用卡繳款中狀況最差代號 (9<A<B<C<D<E<F)	N	N	N
34	本月信用卡全額繳清之應付賬款金額(A OR B)	3,6,12個月	Mean, Max, Std.	增加量
35	本月信用卡循環信用之應付賬款金額(C OR D)	N	N	N
36	本月信用卡未繳足最低金額或全額逾期末繳之應付賬款金額(E OR F)	3,6,12個月	Mean, Max, Std.	增加量
37	本月信用卡繳款有延遲金額(B、D、E OR F)	3,6,12個月	Mean, Max, Std.	增加量
38	本月被查詢總家數	3,6,12個月	Total, Mean, Max, Std.	增加量
39	本月被查詢總次數	3,6,12個月	Total, Mean, Max, Std.	增加量
40	本月被查詢總產品類別數	3,6,12個月	Total, Mean, Max, Std.	增加量
41	本月被查詢總產品次數	3,6,12個月	Total, Mean, Max, Std.	增加量
42	本月現金卡授信金額	N	N	N
43	本月現金卡張數	3,6,12個月	Mean, Max, Std.	增加量
44	本月現金卡借款金額	3,6,12個月	Mean, Max, Std.	增加量
45	本月信用額度與所得比	3,6,12個月	Mean, Max, Std.	增加量
46	本月循環信用與應付帳款比	3,6,12個月	Mean, Max, Std.	增加量
47	本月有效附卡張數	3,6,12個月	Mean, Max, Std.	增加量

3. 模型應用之實證分析

3.1 主成份分析

為了處理初始投入變數的共線性問題，本研究採用主成份分析法萃取特徵值大於1的主成份。由於每一觀測基準時點的

樣本特性有些許相同，所萃取出來的主成份也不全然一樣，茲將各期主成份個數與累積總變異量之結果整理成表2。惟值得注意者，最後萃取出來之主成分的解釋力很小(見表3)，因此各期變異程度不大。

表2 各期主成份個數與累積總變異量

觀測基準時點	2004.12.31	2005.01.31	2005.02.28	2005.03.31	2005.04.30	2005.05.31
主成份個數	83	80	82	80	80	82
累積總變異量(%)	90.305%	90.162%	90.652%	90.428%	90.116%	90.478%

表3 2004.12.31 主成份分析結果

成份	初始特徵值		
	總和	變異數的%	累積
1	87.576	18.360	18.360
2	35.858	7.517	25.877
3	27.386	5.741	31.618
4	17.174	3.600	35.219
5	12.666	2.655	37.874
6	11.620	2.436	40.310
7	10.672	2.237	42.547
		⋮	
82	1.028	0.215	90.092
83	1.016	0.213	90.305

表3為2004.12.31樣本的投入變數主成份分析結果摘要，由表中可知，主成份分析將原始480個變數（扣除4個名目變數⁶）萃取出83個主成份，其解釋總變異量超過90%，而前11個主成份所累積的解釋總變異量就將近50%。末端未呈現出來的第84個成份到第298個成份累積了剩下接近10%的總變異量，第299個成份之後的累積總變異量幾乎微不足道。

主成份決定後，便將訓練組與測試組的原始變數值轉換成主成份分數，加上原來4個名目變數，投入Cox模型中進行存活

分析。另外，為了便於彙整各個樣本的訊息，本研究以樣本的主成分分數為基礎，依每個主成份對風險函數的影響，由低至高排序均勻分成十組給予0~9分⁷，待Cox模型篩選出顯著變數後，加總其顯著變數之分數，即為各樣本的分數。由於本研究的各期樣本其總分多數落在100分~300分之間，因此，將10分訂為組距來做分組；將100分以下的樣本歸為第一組，101分~110分歸為第二組，以此類推，至301分以上的樣本將歸為第二十二組。表4為樣本分數與分組對照表：

表4 樣本分數與分組對照表

分組	樣本總分	分組	樣本總分
第1組	≤ 100	第12組	201~210
第2組	101~110	第13組	211~220
第3組	111~120	第14組	221~230
第4組	121~130	第15組	231~240
第5組	131~140	第16組	241~250
第6組	141~150	第17組	251~260
第7組	151~160	第18組	261~270
第8組	161~170	第19組	271~280
第9組	171~180	第20組	281~290
第10組	181~190	第21組	291~300
第11組	191~200	第22組	≥ 301

⁶ 由於主成份分析不適用於名目變數，因此4個名目變數不納入主成份分析中。

⁷ 名目變數仍依原始分數做為分數，不另做變動。

3.2 存活分析 - Cox 模型

在建立模型之前，樣本必須先區分為發生事件者或被設限的資料兩種。本研究

所定義的事件為信用卡使用者發生違約，遭銀行強制停卡者；被設限的資料表示樣本在研究期間結束時存活期依然持續，無法觀察到真正的存活時間者。舉例來說，

表 5 各期 Cox 模型觀察值處理摘要

觀測基準時點	2004.12.31	2005.01.30	2005.02.28	2005.03.31	2005.04.30	2005.05.31
違約樣本數	1,788	2,031	2,179	2,339	2,545	2,939
(百分比%)	(2.2%)	(2.5%)	(2.7%)	(2.9%)	(3.2%)	(3.7%)
被設限的樣本數	78,200	77,955	77,783	77,638	77,436	77,041
(百分比%)	(97.8%)	(97.5%)	(97.3%)	(97.1%)	(96.8%)	(96.3%)
總和	79,988	79,986	79,962	79,977	79,981	79,980
(百分比%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)

以 2004 年 12 月 31 日為觀測基準時點所隨機選取的 8 萬筆樣本中，在 2005 年 01 月~2005 年 12 月(Outcome Period)發生違約的樣本佔全部樣本的 2.2%，其餘皆為被設限樣本。茲將樣本分類及其所佔比重整理如表 5⁸。

Cox 模型篩選變數的方法為「向前逐步迴歸分析法（條件的 LR）」，應變數為存活時間，即持卡人截至發生違約或者研究期間結束時的持卡總月數；自變數為萃取出來的主成份以及名目變數等。由表 6 的結果

表 6 各期 Cox 模型檢定統計量

觀測基準時點	2004.12.31	2005.01.30	2005.02.28	2005.03.31	2005.04.30	2005.05.31
初始-2 對數概似數函數	38,367.21	43,534.54	46,531.97	49,921.54	54,291.94	62,536.00
-2對數概似數函數	31,036.95	35,082.29	37,591.99	40,642.75	44,047.07	50,614.3
卡方值	11,821.97	19,657.36	21,557.82	22,011.17	27,229.37	28,466.31
自由度(v)	60	58	55	58	62	69
顯著性	0	0	0	0	0	0

⁸ 由於抽樣時有部分樣本的變數有遺漏值，因此需刪除部分樣本，因此每期實際樣本數不同。

可看出各個觀測基準時點之 Cox 模型的適合度檢定量均為顯著：各期 -2 對數概似數函數均小於初始 -2 對數概似數函數，表示最後模型適合度最佳；同時，表示各期模型中的係數不全為 0。

在 Cox 模型中，顯著變數是以 Wald 檢定量來篩選，若 $\chi_w^2 = Z^2 = \left(\frac{\hat{\beta}}{SE(\hat{\beta})}\right)^2 > \chi_a^2(v)$

則拒絕虛無假設⁹ $H_0: \beta_i = 0$ 。附錄之附表 5 為觀測基準時點在 2004 年 12 月 31 日之 Cox 模型配適結果。研究結果顯示，篩選顯著變數時，Wald 檢定量與 p-value 所選出的結果有一致性。除了顯著變數之外，在 Cox 模型中的基準風險函數 $h_0(t)$ 也是存活模型的重要因子。由於 $h_0(t)$ 與參數無關，可

被假設為任意的形式且沒有分配上的限制，因此本研究以樣本來加以估計。有了基準風險函數 $h_0(t)$ 與 $\exp(\beta'x)$ 導出信用卡使用者在未來各時點的存活率 $S(t) = \left[e^{-h_0(t)}\right]^{\exp(\beta'x)}$ ，其中 x 為投入模型中的顯著變數（主成分或名目變數）。附錄之附表 6 為觀測基準時點在 2004 年 12 月 31 日以 Cox 模型所建立之之存活表。

本研究中，各期樣本所呈現的存活率幾乎都集中在 0.98~1.00 之間。圖 2 為以 2004 年 12 月 31 日為觀測基準時點的樣本外測試組資料，其中，有七成的樣本存活率是界於 0.99~1.00 之間，而將近二成的樣本其存活率是界於 0.98~0.99 之間，顯示整體的樣本存活率極高。

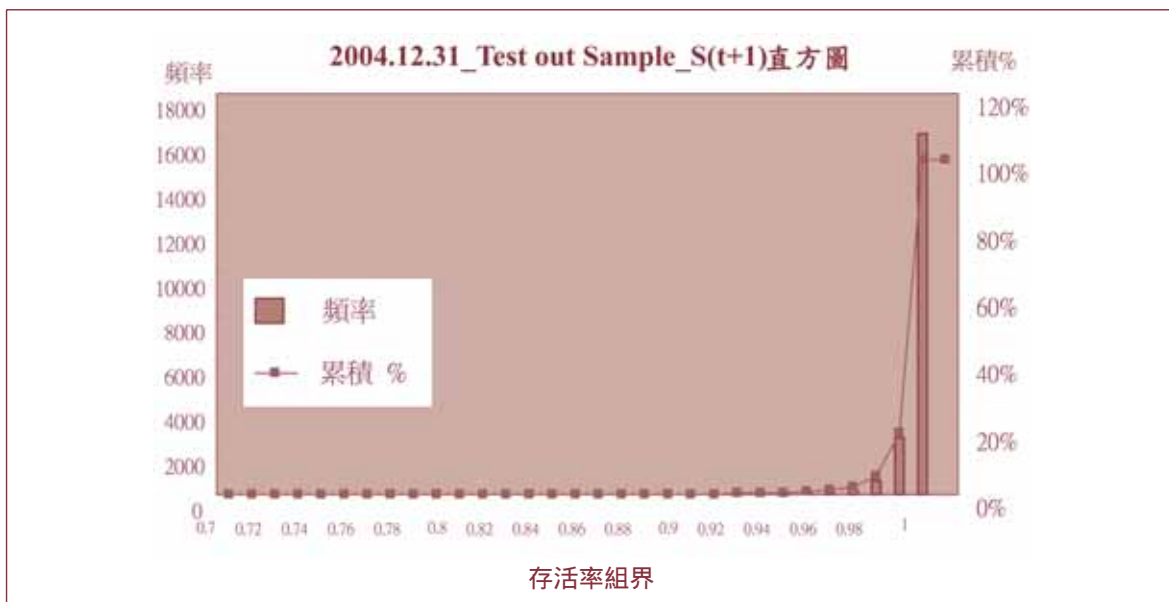


圖 2 2004.12.31 樣本外測試組之存活率分布圖

⁹ 本研究取 $\alpha = 0.05$ 之自由度為 1(或 6)的卡方值為 3.841(12.592)以及 $\alpha = 0.01$ 之自由度為 1(或 6)的卡方值為 6.635(16.812)來進行篩選。

3.3 存活率表之運用

如前所述，為了簡化 Cox 模型的使用，本研究發展一個建立存活率表的方法。而存活率表的使用方法，可由下述範例了解。假設某一信用卡使用者的分數為 163 分，且在觀測基準時點之持卡期間已有 14 個月，查表 4 可知該樣本所在組別為第八組。此時欲得知該樣本未來六個月存活率之變化，可查存活率表中 (如下表 7 所示) 第八組 (S(T)_8) 與存活時間 (T) 中第 15 個月至第 20 個月所對照出來的數值得之，其未來六個月的存活機率將由 99.9989% 下降至 99.9975%。一般而言，存活率與存活時間以及組別成反比，亦即存活率會隨著存活時間的增加而遞減；同時，組別排序越後面者存活率也較低，這是因為在本研究中總分越高者風險也越高之故。

除了提供個別存活率之數值，使用者亦可利用長期資料找出發生違約事件的存活率臨界值，進而判斷樣本會發生違約的時點。舉例來說，假設某金融機構設定 98% 為其存活率臨界值，假設觀察到一位信用卡持有者屬於第 21 組，且目前已經持有信用卡 15 個月了，由於第 21 組的第 21 個月的存活機率 (97.8028%) 小於 98%，因此可知其可能在未來第 6 個月左右成為高風險持有者，此金融機構應立即將其列為觀察對象，在接下來幾個月密集的檢視其消費與繳款情況，並在適當時機作出處置，以降低損失。

存活分析的好處在於可以預測違約接近的時點，雖然其他模型亦可能預測出未來一段時間內的違約機率 (如 Logit 模型)，但由於不能預測接近違約的時點，因此不

表 7 2004.12.31 存活率表(部分)

T	14	15	16	17	18	19	20	21	...
S(T)-3	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...
S(T)-7	0.999996	0.999995	0.999994	0.999993	0.999991	0.999990	0.999989	0.999988	...
S(T)-8	0.999991	0.999989	0.999986	0.999983	0.999980	0.999978	0.999975	0.999972	...
S(T)-9	0.999982	0.999979	0.999973	0.999968	0.999962	0.999958	0.999952	0.999946	...
S(T)-10	0.999966	0.999960	0.999949	0.999939	0.999928	0.999921	0.999910	0.999899	...
S(T)_11	0.999937	0.999926	0.999906	0.999887	0.999867	0.999853	0.999834	0.999813	...
S(T)_12	0.999897	0.999879	0.999846	0.999815	0.999783	0.999760	0.999728	0.999694	...
S(T)_13	0.999846	0.999819	0.999769	0.999724	0.999676	0.999641	0.999593	0.999541	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...
S(T)_21	0.992562	0.991277	0.988869	0.986708	0.984404	0.982751	0.980462	0.978028	...

知道尚有多少時間能以更新的資料重新評估，只能立即就做處置，如此，可能由於過早處置而損失客戶。倘若被列為危險客戶的信用卡持有人在近期內改善其信用狀況，則過早處置會造成永遠損失這名客戶，這將是金融機構所不樂見的。

3.4 模型驗證

本模型除 LR test 外，亦作其他相關驗證。首先，針對六組模型(2004.12.31、2005.01.31、2005.02.28、2005.03.31、2005.04.30、2005.05.31)都作一樣本內測試，所謂樣本內測試，是取跟模型建立同期的其他樣本(各 20000 筆)，將其資料帶入建立好的模型中，驗證其存活率預測是否與實際存活率接近。

另外，針對 2004.12.31 及 2005.03.31 兩組，本研究亦作樣本外測試，又分為樣本外模型測試組與樣本外查表測試組。交叉比對這兩組樣本外測試組的結果，可以得出在現行的機制下，六個月建模檢驗一次與三個月建模檢驗一次的優劣¹⁰。

2004.12.31 組的樣本外模型測試組是隨機抽取二萬筆在 2005 年 12 月 31 日仍存活的樣本，以其 2005 年 1 月 1 日至 2005 年 12 月 31 日的變數帶入 2004.12.31 的模型中，估計出其未來六個月(2006 年 1 月 1 日至 2006 年 6 月 30 日)的理論存活率區間，再比較其未

來六個月的實際存活率，驗證其預測能力是否合乎預期。而 2005.03.31 的樣本外模型測試組是隨機抽取二萬筆在 2006 年 3 月 31 日仍存活的樣本，以其 2005 年 4 月 1 日至 2006 年 3 月 31 日的變數帶入 2005.03.31 的模型中，計算出其未來三個月(2006 年 4 月 1 日至 2006 年 6 月 30 日)的理論存活率區間，再比較其未來三個月的實際存活率，驗證其預測能力是否合乎預期。而樣本外查表測試組則是以相同的測試樣本，不帶入模型，而分別查 2004.12.31 與 2005.03.31 建立的存活表，再比較其未來六個月(三個月)的實際存活率，驗證其預測能力是否合乎預期。以下為模型驗證結果說明。

3.4.1 模型驗證結果

下圖 3、4、5 及 6 為 2004 年 12 月 31 日為觀測基準點的實驗結果。

圖 3 為以 2004 年 01 月至 2004 年 12 月期間的訓練組變數建立存活分析模型的結果，其中，Number 代表各組人數，Theoretical S(T) 為各組平均行存活率，Realized S(T) 為實際上各組在未來一年(2005 年 01 月至 2005 年 12 月)的存活率(即各組存活人數/總人數)。而圖 4 為樣本內測試組的模型分析結果，觀察期間與訓練組同為 2004 年 01 月至 2004 年 12 月，預測期間為 2005 年 01 月至 2005 年 12 月，測試樣本數

¹⁰ Basel II 推行後，相關主管機關規定，金融機構至少需每六個月對所有信用卡持有者作一次評估。雖然評估工作越密集越佳，但考量到每個月評估需耗費相當大的人力與物力，執行的可能性不高，因此，本研究樣本外測試採用採用每六個月及每三個月建模一次，比較其優劣，供金融機構參考。

為 20000 筆。其驗證方法為：依每一個樣本的存活時間，查同期訓練組算出的基準風險函數表找出其下一個月 (2005 年 01 月) 和其後第十二個月 (2005 年 12 月) 的基準風險函數，再配合其觀察期間的各主成份數

值，代入訓練組模型，即為其下一個月與其後第十二個月的預測存活率，將各組內樣本預測存活率取平均數，即可得各組在未來一年內的預測存活率區間 ($S(T+12)$, $S(T+1)$)¹¹。

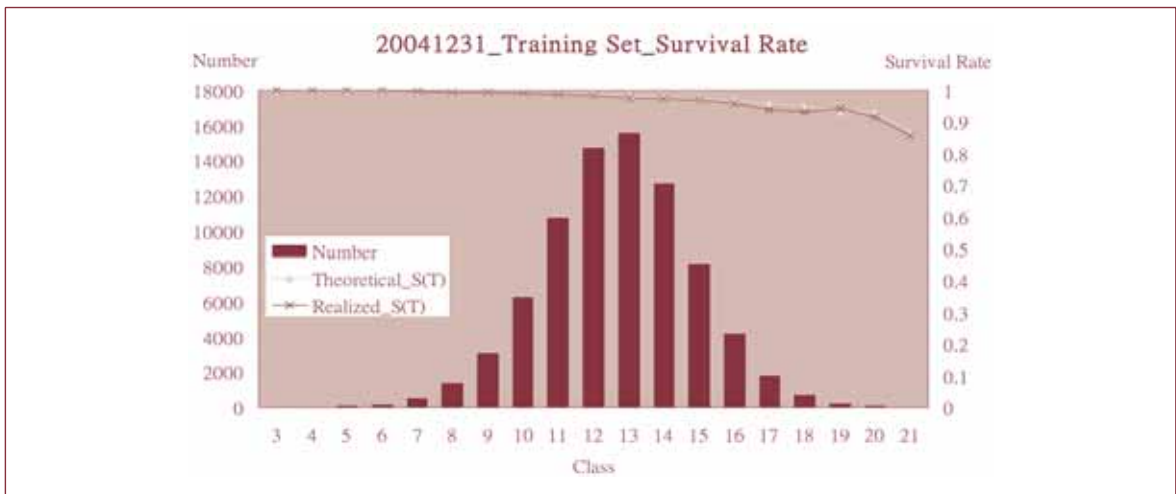


圖 3 2004.12.31 訓練組樣本之理論存活率與實際存活率

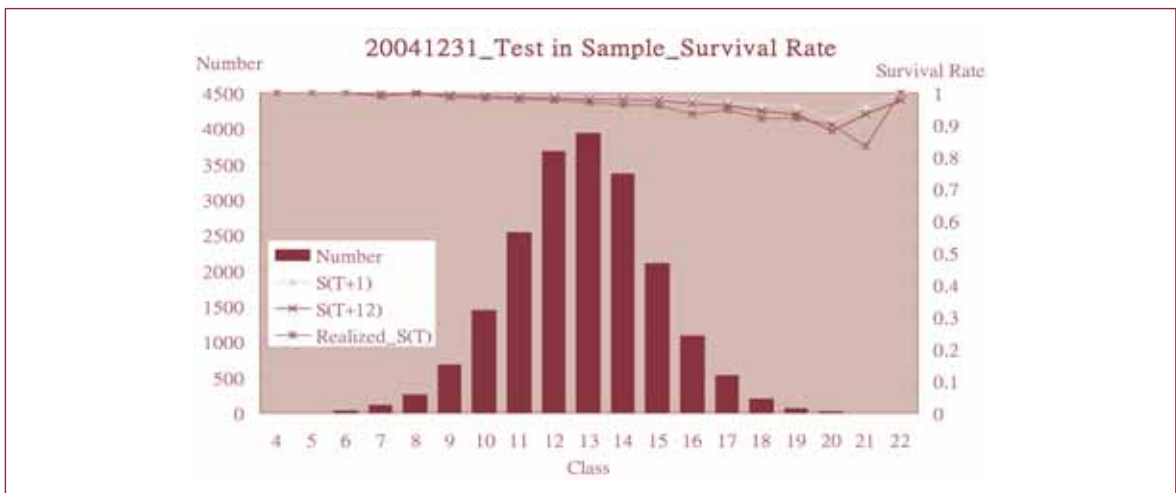


圖 4 2004.12.31 樣本內測試組之理論存活率與實際存活率

¹¹ 由於距離觀察日越久，其存活率越低，因此 $S(T+12)$ 小於 $S(T+1)$ 。

圖5為樣本外測試組的模型分析結果，觀察期間為2005年01月至2005年12月，預測期間為2006年01月至2006年06月，測試樣本數為20000筆。其中，Realized $S(T)$ 為實際上各組在未來六個月(2006年01月至2006年06月)的存活率。其驗證方法與樣本內測試組驗證方法類似。而圖6為樣本外測試組的查表分析結果。其查表驗證

方法為：依每一個樣本的分數，查訓練組所建立的存活率表(Survival Table)，找出其所屬組別，得其下一個月(2006年01月)和其後第六個月(2006年06月)的查表存活率，將各組內樣本查表存活率取平均值，即為各組未來六個月內的預測存活率區間($S(T+6)$, $S(T+1)$)。

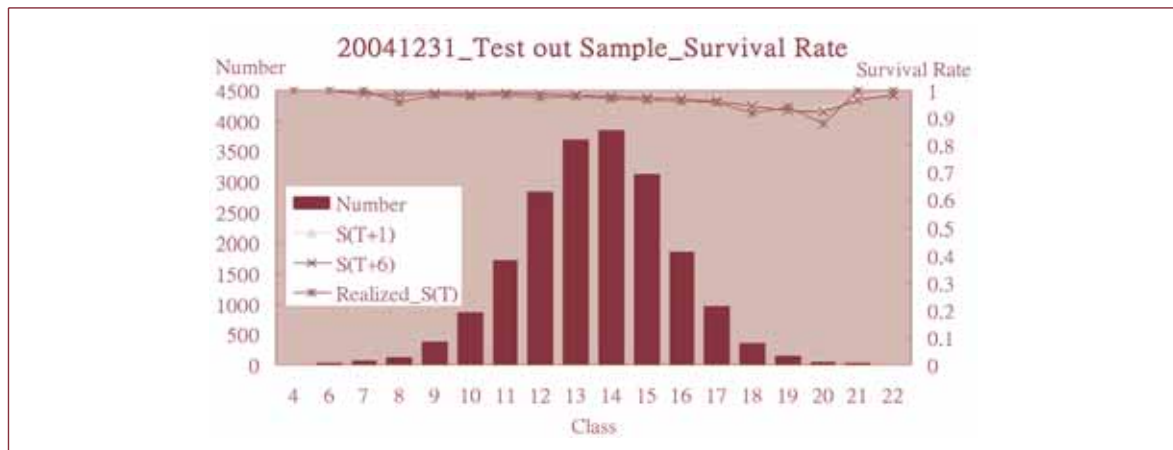


圖5 2004.12.31 樣本外測試組之理論存活率與實際存活率

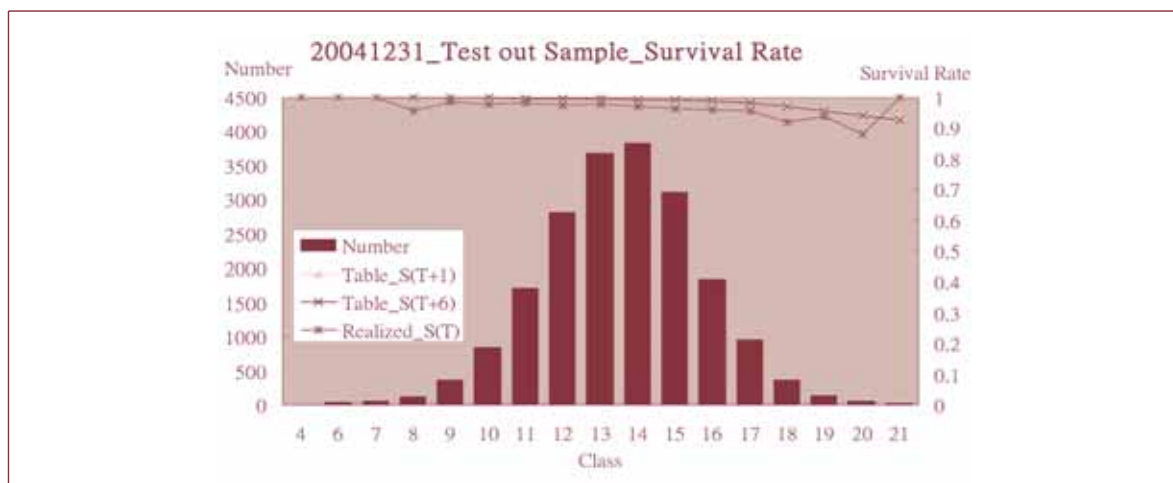


圖6 2004.12.31 樣本外測試組之實際存活率與存活率查表值

由圖3至6可知，無論訓練組、樣本內測試組、樣本外測試組均符合一共同之趨勢：風險越高的組（即總分越高），其平均理論存活率與實際存活率皆越低。由於先前分組時，是依各主成份與名目項目對風險函數的影響，由低至高給0至9分，亦即，總分越高者，其風險值越大，而其存活率則隨之越小。因此，圖中顯示之趨勢證明本研究提供的分組方式合理。圖4-6亦顯示出，無論是模型預測值或查表預測值，皆為 $S(T+12)$ 小於 $S(T+1)$ 或 $S(T+6)$ 小於 $S(T+1)$ ，此結果表示，模型建立之存活率會依存活時間而下降，與實際現象符合。另外，訓練組的理論存活率與實際存活率接近；樣本內測試組的實際存活率接近接下來十二個月的模型預測存活率區間；而樣本外測試組的實際存活率亦接近接下來六個月的模型預測存活率區間與查表預測存活率區間，這些結果皆顯示本研究適用於金融機構對信用卡持有者之信用風險評估上。

3.4.2 各組誤受與誤拒之機率

在討論模型誤拒與誤受機率時，本應找出一適用於每個時點的臨界值來作風險分類¹²，但受限於本研究樣本觀察點只有六期模型，很難找到各時點模型皆適用的臨界值，故在此以每一個時點的實際存活率為基準，將樣本分為兩群：低風險群與高

風險群，模型預測存活率高於實際存活率者歸於低風險群；模型預測存活率低於實際存活率者歸於高風險群。由於低風險群為模型結果認為較不具風險的持卡人，其授信機率較高，因此，其實際違約率即為誤受的機率；反之，高風險群為模型結果認為較具風險的持卡人，其授信機率因此降低，因此，其實際存活率即為誤拒的機率。誤受機率代表金融機構若採用Cox模型，會因為錯誤授信造成損失的機率；而誤拒機率則代表金融機構採用此模型時，會因為對被列為高風險持有者作處置（如停卡或降低信用額度）而造成損失的機率。

表8與表9為各期Cox模型於樣本內測試組與樣本外測試組的誤拒及誤受機率。在本研究中，模型誤受機率約在0.19%~1.82%，查表誤受機率稍高，約為1.73%~2.91%；而模型誤拒機率約在7.64%~24.64%，查表誤受機率稍低，約為2.16%~5.38%。觀察樣本外的預測，就誤受機率而言，以模型預測較查表預測準確。此現象是基於存活表為各組之理論存活率的平均值，是一個概觀值，對於每一個持卡人的預測自然有誤差。惟存活表有助於簡化操作，是故若誤差不大時，仍有其價值。另外，由表9可知，無論是模型預測能力或是查表預測能力，皆是2005.03組表現較佳，此結果顯示，每三個月建模一次的效果較每六個月建模一次佳。

¹² 本研究建議，金融機構未來在採用此模型時，應使用本身信用卡的長期資料，分析出適當的存活率臨界值，作為驗證模型之用。

表 8 各期樣本內測試組誤受機率及誤拒機率

單位: %	2004.12.31	2005.01.31	2005.02.28	2005.03.31	2005.04.30	2005.05.31
存活率 臨界值	97.61	97.31	96.12	96.90	96.61	96.12
誤受 機率	0.78~1.22	0.71~1.40	1.06~1.30	1.17~1.74	0.87~1.62	1.14~1.82
誤拒 機率	14.83~16.52	14.64~15.04	7.64~17.53	16.98~18.18	20.15~21.11	22.63~24.64

表 9 樣本外測試組誤受機率及誤拒機率

單位: %		2004.12.31	2005.03.31
模型	存活率臨界值	97.04	98.24
	誤受機率	0.63~0.81	0.19~0.24
	誤拒機率	18.38~18.95	7.95~8.08
查表	誤受機率	2.90~2.91	1.73~1.74
	誤拒機率	4.99~5.19	2.14~2.17

3.4.3 與 Logit 模型之比較

為與其他模型之預測能力比較，本研究特選擇過去文獻與目前金融機構最常使用的 Logit 模型來做比較。Logit 模型的各期樣本內、外測試組誤受機率及誤拒機率列於下表 10 中，而 Cox 與 Logit 模型各期顯著變數個數則列於表 11 中。

比較表 8、9 與表 10 可知，當以每一個時點的實際存活率為基準，將樣本分為低風險群與高風險群時，Logit 模型的誤受機率皆高於 Cox 模型的模型誤受機率；而其誤拒機率則有部分低於 Cox 模型的模型誤拒機率，此結果與 Noh et al. (2005) 的主要結果“Cox 模型具有較低的型一錯誤”相

符。另外，由表 11 可知，每一期 Cox 模型包含的顯著變數個數皆高於 Logit 模型篩選出的顯著變數個數，此結果亦支持 Noh et al. (2005) 的另一個結果：Cox 模型包含較多的顯著變數。由於降低型一錯誤可以減少錯誤授信對金融機構造成的損失；而較多的顯著變數亦可以提供更精細的資訊來做預測，因此，雖然相對於現行金融機構採用的模型，本模型較為複雜，使用成本可能因此較高，但本文仍建議金融機構未來除了原來的模型外，在各種成本效益符合的型況下，可考慮採用存活分析模型，透過與信用卡違約風險相關的可能因素建立預警模型，以降低處理違約的後續成

本，也能減少社會問題。

雖然比較表 8、9 與表 10 可知，相較於 Logit 模型，Cox 模型得查表誤受機率略高，但誠如先前所言，此現象是基於存活表為各組之理論存活率的平均值，是一個概觀值，對於每一個持卡人的預測自然有誤差，惟存活表有助於簡化操作，是故若誤差不大時，仍有其價值。另外，存活分析最主要的好處在於可以預測違約接近的時點，雖然 Logit 模型亦可預測出未來一段

時間內的違約機率，但由於不能預測接近違約的時點，因此不知道尚有多少時間能以更新的資料重新評估，只能立即就做處置，倘若被列為危險客戶的信用卡持有人在近期內改善其信用狀況，則過早處置會造成永遠損失這名客戶。反觀 Cox 模型，由於能預估期接近違約的時點，因此知道有多少時間藉由更新的資料來做評估，如此則可以在不平白損失客戶的情形下，降低金融機構的信用風險。

表 10 Logit 模型各期樣本內、外測試組誤受機率及誤拒機率

單位: %	樣本內						樣本外	
時間	2004.12	2005.01	2005.02	2005.03	2005.04	2005.05	2004.12	2005.03
存活率 臨界值	97.61	97.31	96.12	96.90	96.61	96.12	97.04	98.24
誤受 機率	1.04	1.03	1.08	1.24	1.10	1.18	0.80	0.33
誤拒 機率	11.21	13.12	15.89	16.15	19.27	21.60	14.99	6.84

表 11 Cox 與 Logit 模型各期顯著變數個數

	Cox	Logit		Cox	Logit
2004.12	50	29	2005.03	48	37
2005.01	48	24	2005.04	52	33
2005.02	45	31	2005.05	59	31

4. 結論與建議

國內常用於信用卡使用者之違約風險評估的方法，主要為區別分析及 Logit 模型

或 Probit 模型。這些模型皆在固定時點判定未來一段特定期間內是否會發生違約或發生違約的機率，無法提供持卡人在未來不同時點的違約機率。本研究應用在生物醫

學及精算領域廣為使用的存活分析法，透過與信用卡使用者違約相關的可能因素，進一步以主成份分析法萃取出主成份投入Cox模型，從中決定顯著的共變量，進而建立預測模型及或存活率表，使銀行能以客觀的預估客戶未來各時點發生違約的機率，這項能力是存活分析模型與其他模型最大的差異處，這項能力使得採用者可以更加掌握信用卡使用者的未來資訊，增加風險控管的積極性。應用存活分析法於台灣信用卡資料的實證結果顯示，所得之結果與Noh et al. (2005) 結果一致，存活分析模型不但有較小的型一錯誤，更包含較多的顯著變數；而且無論樣本內測試組或樣本外測試組，實際存活率皆接近模型預測存活率區間，表示本模型預測能力合乎預期，也顯示存活分析模型具有實務上之應用性。

本研究的獨特性是能對信用卡使用者在未來各時點發生違約的機率做預測，並建立一操作簡單之存活率表，因此，未來金融機構可應用此模型的預警功能對新信用卡申請人及現有信用卡持有人進行風險控管。由於目前多數金融機構皆與金融聯合徵信中心合作，只要信用卡申請人擁有任何一張信用卡，新信用卡申請銀行便可將其過去之信用卡紀錄參照存活率表，查出其未來發生違約的機率估計值，以判斷其是否為安全的申請者。而現有信用卡持有人之所有銀行信用卡紀錄，亦可提供金融機構估計其未來每個時點的違約機率，一旦違約機率高於金融機構認定之臨界

值，便可以藉由提前調降其信用額度或其他處置來降低金融機構未來的損失；另一方面亦可在信用狀況恢復的短時間內恢復其信用額度，以避免流失良好客戶。由於本模型為一動態分析模型，其所建立之存活率表在未來每個時點皆有不同之估計值，金融機構可因此在適當時機對於未來有危險疑慮或是已脫離危險疑慮之信用卡持有者做適當處置，不會過早或過晚，相較於現行的信用評分方式（每一段固定時間調整一次），更具彈性。

關於後續研究方面，在研究信用卡使用者的違約風險時，任何持卡人的使用情況都是建立模型的重要資訊。不過本次研究中的資料已是經過各家銀行整理後再上傳給金融聯合徵信中心，過程中難免有遺漏值或者喪失部份資訊，樣本資料若儘可能地包含所有使用信用卡的訊息，則建立的模型將更加完備和準確。全國性的個人信用風險資料庫正慢慢建立完備的各種資訊，其中除了信用卡的訊息外，還有現金卡與授信資料。若能將這些關於個人信用的資訊都囊括進來，便可從信用卡違約研究的層面擴展至個人信用風險研究。

附錄

附表 1 性別代號表

性別	M(男)	F	(女)
代號	1	0	

資料來源：金融聯合徵信中心

附表 2 是否有自用住宅代號表

	有自用住宅者	無自用住宅或遺漏值
代號	1	0

資料來源：金融聯合徵信中心

附表 3 教育程度代號表

性別	博士	碩士	大學	專科	高中高職	其他
代號	1	2	3	4	5	6

資料來源：金融聯合徵信中心

附表 4 信用卡戶繳款狀況代號表

代號	說明	分數
X	未消費	0
A	全額繳清無延遲	1
B	全額繳清有延遲	2
C	循環信用無延遲	3
D	循環信用有延遲	4
E	未繳足最低金額	5
F	全額逾期末繳	6

資料來源：金融聯合徵信中心

附表 5 2004.12.31 Cox 模型配適結果

方程式中的變數	β	$SE(\beta)$	Wald	自由度	顯著性	$EXP(\beta)$
CP_WORST_PAYCODE		5	11.1832	6	0.0000	
CP_WORST_PAYCODE(1)	-0.5101	0.1579	10.4428	1	0.0012	0.6004
CP_WORST_PAYCODE(2)	-0.7491	1.0075	0.5529	1	0.4571	0.4728
CP_WORST_PAYCODE(3)	1.2745	0.1267	101.1413	1	0.0000	3.5769
CP_WORST_PAYCODE(4)	1.6232	0.1579	105.6256	1	0.0000	5.0695
CP_WORST_PAYCODE(5)	2.2241	0.1664	178.7556	1	0.0000	9.2454
CP_WORST_PAYCODE(6)	2.1888	0.1412	240.2786	1	0.0000	8.9242
H_CODE	-0.2255	0.0762	8.7452	1	0.0031	0.7981
EDU_CODE			65.1474	6	0.0000	
EDU_CODE(1)	-0.0907	0.2431	0.1391	1	0.7092	0.9133
EDU_CODE(2)	-0.5247	0.2692	3.7998	1	0.0513	0.5917
EDU_CODE(3)	-0.1122	0.1496	0.5623	1	0.4533	0.8939
EDU_CODE(4)	-0.0892	0.1509	0.3497	1	0.5543	0.9146
EDU_CODE(5)	0.3324	0.1355	6.0126	1	0.0142	1.3943
EDU_CODE(6)	0.3287	0.1391	5.5793	1	0.0182	1.3891
SEX	0.2679	0.0508	27.8132	1	0.0000	1.3072
FAC2_1	-1.1389	0.0296	1475.6380	1	0.0000	0.3202
FAC3_1	-0.5167	0.0136	1435.2143	1	0.0000	0.5965
FAC4_1	-0.0422	0.0159	6.9987	1	0.0082	0.9587
FAC5_1	0.2770	0.0208	177.9002	1	0.0000	1.3192
FAC6_1	-0.2377	0.0221	115.3771	1	0.0000	0.7884
FAC10_1	-0.1722	0.0202	72.8632	1	0.0000	0.8418
FAC14_1	-0.0658	0.0210	9.8248	1	0.0017	0.9363
FAC15_1	-0.2759	0.0324	72.2806	1	0.0000	0.7589
FAC17_1	-0.0515	0.0147	12.3250	1	0.0004	0.9498
FAC18_1	-0.0378	0.0176	4.6274	1	0.0315	0.9629
FAC19_1	-0.1090	0.0206	27.9438	1	0.0000	0.8967
FAC22_1	0.0847	0.0243	12.1641	1	0.0005	1.0884
FAC23_1	-0.0521	0.0223	5.4283	1	0.0198	0.9493
FAC27_1	0.0845	0.0242	12.1489	1	0.0005	1.0882
FAC28_1	0.1905	0.0171	124.2286	1	0.0000	1.2098
FAC29_1	0.0691	0.0172	16.1466	1	0.0001	1.0716

附表5 2004.12.31 Cox 模型配適結果 (續)

方程式中的變數	β	$SE(\beta)$	Wald	自由度	顯著性	$EXP(\beta)$
FAC32_1	-0.0808	0.0352	5.2808	1	0.0216	0.9223
FAC33_1	0.1091	0.0217	25.3830	1	0.0000	1.1153
FAC34_1	0.1130	0.0256	19.4610	1	0.0000	1.1196
FAC37_1	-0.1388	0.0170	66.8761	1	0.0000	0.8704
FAC38_1	0.0816	0.0192	17.9740	1	0.0000	1.0850
FAC40_1	-0.0994	0.0195	25.9887	1	0.0000	0.9054
FAC43_1	-0.1058	0.0213	24.5796	1	0.0000	0.8996
FAC44_1	0.0866	0.0161	28.8964	1	0.0000	1.0905
FAC45_1	0.0702	0.0172	16.6486	1	0.0000	1.0727
FAC46_1	0.0659	0.0169	15.2298	1	0.0001	1.0681
FAC47_1	-0.1124	0.0176	40.9376	1	0.0000	0.8937
FAC48_1	0.0831	0.0202	16.9841	1	0.0000	1.0866
FAC50_1	-0.0475	0.0201	5.5673	1	0.0183	0.9537
FAC52_1	0.0362	0.0159	5.2113	1	0.0224	1.0369
FAC53_1	-0.0887	0.0215	17.0716	1	0.0000	0.9151
FAC54_1	-0.0910	0.0213	18.3210	1	0.0000	0.9130
FAC63_1	-0.0963	0.0191	25.5242	1	0.0000	0.9082
FAC66_1	0.0588	0.0250	5.5323	1	0.0187	1.0606
FAC67_1	-0.0606	0.0161	14.2002	1	0.0002	0.9412
FAC68_1	-0.1177	0.0166	50.5011	1	0.0000	0.8889
FAC69_1	-0.0729	0.0167	19.1571	1	0.0000	0.9297
FAC71_1	0.0429	0.0137	9.8093	1	0.0017	1.0438
FAC72_1	-0.0611	0.0152	16.1324	1	0.0001	0.9407
FAC73_1	0.0572	0.0174	10.8438	1	0.0010	1.0588
FAC74_1	-0.0395	0.0150	6.9363	1	0.0084	0.9613
FAC75_1	0.1384	0.0337	16.8658	1	0.0000	1.1484
FAC76_1	0.0806	0.0274	8.6441	1	0.0033	1.0840
FAC79_1	-0.0891	0.0142	39.5720	1	0.0000	0.9147
FAC82_1	0.1213	0.0157	59.7006	1	0.0000	1.1289
FAC83_1	-0.1132	0.0274	17.0191	1	0.0000	0.8930

註：篩選條件為向前逐步迴歸分析法 (條件的LR)

附表 6 2004.12.31 Cox 模型之存活表

時間	基準線累積風險 $h_0(t)$	共變量平均值		
		存活機率	標準差	累積風險
5	0.00000293	0.999996	0.000002	0.000004
6	0.00000590	0.999992	0.000003	0.000008
7	0.00001194	0.999984	0.000005	0.000016
8	0.00002002	0.999973	0.000006	0.000027
9	0.00002918	0.999961	0.000008	0.000039
10	0.00003636	0.999952	0.000009	0.000048
11	0.00005593	0.999926	0.000011	0.000074
12	0.00007045	0.999907	0.000013	0.000093
13	0.00008925	0.999882	0.000015	0.000118
14	0.00012300	0.999837	0.000018	0.000163
15	0.00014433	0.999809	0.000021	0.000192
16	0.00018440	0.999755	0.000024	0.000245
17	0.00022045	0.999708	0.000028	0.000293
18	0.00025896	0.999656	0.000032	0.000344
19	0.00028665	0.999620	0.000034	0.000380
20	0.00032506	0.999569	0.000038	0.000431
21	0.00036602	0.999514	0.000041	0.000486
22	0.00041606	0.999448	0.000046	0.000552
23	0.00045576	0.999395	0.000050	0.000605
24	0.00050400	0.999331	0.000054	0.000669
25	0.00055638	0.999262	0.000058	0.000738
26	0.00062262	0.999174	0.000064	0.000826
27	0.00068326	0.999094	0.000070	0.000907
28	0.00072037	0.999045	0.000073	0.000956
29	0.00076534	0.998985	0.000077	0.001016
30	0.00081304	0.998922	0.000081	0.001079
31	0.00086428	0.998854	0.000085	0.001147
32	0.00092881	0.998768	0.000091	0.001232
33	0.00097267	0.998710	0.000095	0.001291
34	0.00102021	0.998647	0.000099	0.001354
35	0.00107532	0.998574	0.000103	0.001427

附表 6 2004.12.31 Cox 模型之存活表 (續)

時間	基準線累積風險 $h_0(t)$	共變量平均值		
		存活機率	標準差	累積風險
36	0.00114848	0.998477	0.000110	0.001524
37	0.00118753	0.998425	0.000113	0.001576
38	0.00123257	0.998366	0.000117	0.001636
39	0.00129869	0.998278	0.000122	0.001723
40	0.00134301	0.998220	0.000126	0.001782
41	0.00139883	0.998146	0.000131	0.001856
42	0.00144567	0.998084	0.000135	0.001918
43	0.00150701	0.998002	0.000140	0.002000
44	0.00156473	0.997926	0.000145	0.002076
45	0.00161272	0.997862	0.000149	0.002140
46	0.00165891	0.997801	0.000153	0.002201
47	0.00171227	0.997731	0.000157	0.002272
48	0.00181154	0.997599	0.000165	0.002404
49	0.00188624	0.997500	0.000172	0.002503
50	0.00194505	0.997422	0.000177	0.002581
51	0.00198219	0.997373	0.000180	0.002630
52	0.00206241	0.997267	0.000186	0.002737
53	0.00211223	0.997201	0.000191	0.002803
54	0.00215628	0.997143	0.000194	0.002861
55	0.00221665	0.997063	0.000199	0.002941
56	0.00227457	0.996986	0.000204	0.003018
57	0.00232644	0.996918	0.000209	0.003087
58	0.00240431	0.996815	0.000215	0.003190
59	0.00245509	0.996748	0.000219	0.003258
60	0.00251140	0.996673	0.000224	0.003332
61	0.00258214	0.996580	0.000230	0.003426
62	0.00265456	0.996484	0.000236	0.003522
63	0.00273859	0.996373	0.000243	0.003634
64	0.00282673	0.996256	0.000251	0.003751
65	0.00288228	0.996183	0.000255	0.003825
66	0.00294520	0.996100	0.000260	0.003908

附表 6 2004.12.31 Cox 模型之存活表 (續)

時間	基準線累積風險 $h_0(t)$	共變量平均值		
		存活機率	標準差	累積風險
67	0.00298854	0.996042	0.000264	0.003966
68	0.00302219	0.995998	0.000267	0.004010
69	0.00308596	0.995914	0.000272	0.004095
70	0.00319263	0.995773	0.000281	0.004236
71	0.00326069	0.995683	0.000287	0.004327
72	0.00333037	0.995591	0.000293	0.004419
73	0.00338927	0.995513	0.000298	0.004497
74	0.00348387	0.995388	0.000306	0.004623
75	0.00358088	0.995260	0.000315	0.004752
76	0.00363876	0.995183	0.000320	0.004828
77	0.00368315	0.995125	0.000323	0.004887
78	0.00377392	0.995005	0.000331	0.005008
79	0.00389804	0.994841	0.000342	0.005172
80	0.00396155	0.994757	0.000347	0.005257
81	0.00400217	0.994704	0.000351	0.005311
82	0.00406964	0.994614	0.000357	0.005400
83	0.00415686	0.994499	0.000364	0.005516
84	0.00422819	0.994405	0.000371	0.005610
85	0.00434754	0.994248	0.000381	0.005769
86	0.00444294	0.994122	0.000389	0.005895
87	0.00454135	0.993992	0.000398	0.006026
88	0.00464328	0.993858	0.000407	0.006161
89	0.00470727	0.993773	0.000412	0.006246
90	0.00478540	0.993670	0.000419	0.006350
91	0.00486624	0.993564	0.000427	0.006457
92	0.00498603	0.993406	0.000437	0.006616
93	0.00516531	0.993169	0.000453	0.006854
94	0.00530146	0.992990	0.000466	0.007035
95	0.00541497	0.992840	0.000476	0.007185
96	0.00550288	0.992725	0.000484	0.007302
97	0.00560859	0.992585	0.000494	0.007442

附表 6 2004.12.31 Cox 模型之存活表 (續)

時間	基準線累積風險 $h_0(t)$	共變量平均值		
		存活機率	標準差	累積風險
98	0.00571897	0.992440	0.000504	0.007589
99	0.00591691	0.992179	0.000522	0.007851
100	0.00603742	0.992021	0.000533	0.008011
101	0.00612600	0.991904	0.000542	0.008129
102	0.00618144	0.991831	0.000547	0.008202
103	0.00623831	0.991756	0.000552	0.008278
104	0.00631704	0.991653	0.000560	0.008382
105	0.00642005	0.991517	0.000570	0.008519
106	0.00652823	0.991375	0.000580	0.008662
107	0.00662013	0.991254	0.000589	0.008784
108	0.00669201	0.991160	0.000596	0.008880
109	0.00681806	0.990994	0.000609	0.009047
110	0.00687039	0.990925	0.000614	0.009116
111	0.00695202	0.990818	0.000623	0.009225
112	0.00700901	0.990743	0.000629	0.009300
113	0.00724925	0.990427	0.000654	0.009619
114	0.00743866	0.990178	0.000674	0.009871
115	0.00750421	0.990092	0.000681	0.009958
116	0.00757310	0.990001	0.000688	0.010049
117	0.00771831	0.989811	0.000704	0.010242
118	0.00775596	0.989761	0.000708	0.010292
119	0.00783421	0.989658	0.000717	0.010395
120	0.00807516	0.989342	0.000744	0.010715
121	0.00820122	0.989177	0.000757	0.010882
122	0.00842521	0.988883	0.000782	0.011180
123	0.00851951	0.988759	0.000793	0.011305
124	0.00866455	0.988569	0.000810	0.011497
125	0.00876404	0.988438	0.000821	0.011629
126	0.00881563	0.988370	0.000827	0.011698
127	0.00892249	0.988230	0.000840	0.011839
128	0.00903572	0.988082	0.000854	0.011990

附表 6 2004.12.31 Cox 模型之存活表 (續)

時間	基準線累積風險 $h_0(t)$	共變量平均值		
		存活機率	標準差	累積風險
130	0.00909830	0.988000	0.000861	0.012073
131	0.00923246	0.987824	0.000878	0.012251
132	0.00930179	0.987733	0.000887	0.012343
133	0.00944629	0.987544	0.000906	0.012535
134	0.00967911	0.987239	0.000935	0.012843
136	0.00976517	0.987126	0.000947	0.012958
137	0.00985404	0.987010	0.000959	0.013076
138	0.01013054	0.986647	0.000998	0.013442
142	0.01023733	0.986508	0.001013	0.013584
144	0.01057832	0.986061	0.001064	0.014037
145	0.01069728	0.985906	0.001082	0.014194
146	0.01082130	0.985744	0.001101	0.014359
147	0.01094837	0.985577	0.001121	0.014528
149	0.01108442	0.985399	0.001142	0.014708
150	0.01122305	0.985218	0.001164	0.014892
151	0.01136436	0.985033	0.001187	0.015080
153	0.01152612	0.984822	0.001211	0.015294
155	0.01172289	0.984565	0.001248	0.015555
156	0.01192608	0.984300	0.001286	0.015825
162	0.01218297	0.983964	0.001338	0.016166
166	0.01248529	0.983569	0.001408	0.016567
173	0.01289141	0.983040	0.001520	0.017106
177	0.01336286	0.982425	0.001658	0.017731
241	0.01454380	0.980887	0.002293	0.019299
249	0.01580713	0.979244	0.002838	0.020975
250	0.01711017	0.977552	0.003316	0.022704
254	0.01847778	0.975780	0.003787	0.024519

參考文獻

1. 陳志強(2004), 亞洲金融的轉型發展與風險—韓國信用卡公司財務危機探討, 台灣經濟研究月刊第27卷第3期。
2. 郭志安(1997), 以Cox模型建立財務危機預警模式, 逢甲大學統計與精算研究所碩士論文。
3. 曾俊堯(1995), 信用卡信用風險評估模式之研究, 中州學報, 第八期P300-314。
4. 鄭婷月(2003), 汽車貸款客戶之風險研究, 成功大學統計所碩士論文。
5. 鍾岳昌(2004), 以比例危險模型估計房貸借款人提前清償及違約風險, 政治大學財管所碩士論文。
6. 龔昶元(1998), Logistic Regression 模式應用於信用卡信用風險審核之研究 - 以國內某銀行信用卡中心為例, 台北銀行月刊, 第28卷第9期, P35-49。
7. Cox, D. R. (1972), Regression Models and Life-Tables, Journal of the Royal Statistical Society, Series B, 34(2), 187-220
8. David, J. H. and Gordon, B. (2001), Prospecting for Gems in Credit Card Data, Journal of Management Mathematics 12(2):173-200
9. David, J. H. (2001), Modelling Consumer Credit risk, Journal of Management Mathematics 12(2):139-155
10. Hayhoe, C.R., Leach, L. and Turner, P.R. (1999), Discriminating the Number of Credit Cards Held by College Students Using Credit and Money Attitudes, Journal of Economic Psychology, 20(6), pp. 643-656(14)
11. Keasey, K., McGuinness, P., and Short, H. (1990), Maulti-logit Approach to Predicting Corporate Failure—Further Analysis and the Issue of Signal Consistency, Omega, 18(1), 85-94
12. Lando, D. (1994), The Essays on Contingent Claims Pricing, PhD Thesis, Cornell University, Ithaca, NY
13. Lane, W. R., Looney, S.W., and Wansley, J. W. (1986), An Application of the Cox Proportional Hazards Model to Bank Failure, Journal of Banking Finance, 511-531
14. Lee, S. H. and Urrutia, J. L. (1996), Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry : A Comparison of Logit and Hazard Models, The Journal of Risk and Insurance, 63(1), 121-130
15. Luoma, M., and Laitinen, E. K. (1991), Survival Analysis as a Tool for Company Failure Prediction, OMEGA International Journal of Management Science, 19(6), 673-678
16. Noh, H. J., Roh, T. H., and Han. I. (2005), Prognostic Personal Credit Risk Model Considering Censored Information, Expert Systems with Applications, 28, 753-762
17. Thomas, L. C., Edelman, D. B. and Jonathan, N. C. (2002), Credit Scoring and Its Application, SIAM Monographs on Mathematical Modeling and Computation, Philadelphia.