

## 第四章 實證分析

### 第一節 實證分析-利用 Logit 模型

本研究為了建立信用卡信用風險之授信評量模型，研究樣本包括了正常案件及違約案件共 12061，並且採用 9 解釋變數來建立研究模型。在研究變數操作上，首先在原始模型 56 個解釋變數，利用逐步模型選擇法，將不符合要求顯著水準之自變數一一刪除，則剩下較顯著的 9 的解釋變數進行logit模型分析違約風險，違約的定義即信用卡持卡人欠款M1 stage或大於M1，視為持卡人違約風險(0 代表非違約族群;1 代表違約族群)。

故下表分析是針對全部樣本(非違約族群與違約族群)進行 t 檢定以及簡單敘述統計。

【表 4-1】利用 Logit 模型求算出基本統計量

Variable	Coeff	T value
Constant	-1.684	-10.495
年收入(Income)	3.93E-07	5.133
職級(Job Level)	0.083	7.214
婚姻(Marital)	-0.027	-0.434
每月限額(Limit Amt)	-4.73E-07	-1.482
6 個月交易金額(6M Tran Amt)	-3E-05	-14.878
過去 6 個月平均繳款金額(AvgM Pay Amt)	-0.0001	-23.771
6 個月平均期初應繳金額(AvgM Bill Amt)	2.716E-05	23.549
過去6個月是否有循環息(revolving Interest)	1.600	15.151
過去 6 個月繳款金額/應繳金額(Payment Ratio)	-0.004	-0.752

由 **logit** 的實證分析以上 9 個變數

- 【1】 **年收入**：當只有考慮年收入因素時，當年收入愈高愈不容易違約。在其他的變數考慮進去時，由其是考慮了「過去 6 個月平均繳款金額」下，故相對會影響持卡人違約機率提高的因素，所以在這變數的影響的方向，當年收入愈高，其違約風險較大。

【表 4-2】利用 Logit 模型求算出年收入單獨影響持卡人違約機率

Variable	Coeff	T-Stat
截距	-1.383326869	-26.33318
年收入	-0.000000794	-8.66081

【表 4-3】利用 Logit 模型求算出年收入與過去 6 個月平均繳款金額影響持卡人違約機率

Variable	Coeff	T-Stat
截距	-0.0428	-0.80446
年收入	8.32E-08	1.45293
過去 6 個月平均繳款金額	-3.27E-04	-31.39191

【2】**職級**(01-自營,02-高級主管,03-一般主管,04-職員,05-專業技術人員,06-公務人員,07-教職,08-軍職,09-其它): 持卡人的職位與違不違約呈正相關, 持卡人是領固定薪水且職等不是很高時, 違約的機率較高。

【3】**婚姻狀況**(1-未婚,2-已婚,3-喪偶,4-離婚,9-其他): 有家庭者與單身者的差別在於, 承擔的責任與壓力不同, 所以婚姻狀況可能會影響個人信用情形。由上表顯示「婚姻」解釋變數不顯著, 現今社會的人會因為沒有錢而不敢結婚, 而有些人不想受到結婚後的金錢責任壓力, 及使是單身也有可能違約, 所以有沒有結婚對於信用卡違不違約是不構成影響因素持卡人的婚姻與違不違約呈正相關, 但由  $p\text{-value}=0.6639>0.1$  可知, 其婚姻狀況與否與持卡人違不違約其實是不相關的, 但實際有可能的情況當已結婚、離婚、喪偶都須扶養孩子的情況下, 則違約機率比單身未婚還高。

【4】**每月限額**: 持卡人每月限額與違不違約呈負相關, 可知當持卡人的每月限額愈大時, 則違約機率愈小。由  $p\text{-value}=0.1383>0.1$ , 可以發現每月限額不是顯著的變數。有些持卡人雖然是額度很小的卡戶, 但還是會利用以卡養卡的行為, 而造成信用卡違約機率很高。

**【5】6 個月交易金額/6 個月平均交易金額(一般消費、預借現金)：**6 個月交易金額與持卡人違不違約呈負相關，可知當持卡人 6 個月交易金額愈大，則違約機率愈小。

**【6】過去 6 個月平均繳款金額：**持卡人在過去 6 個月平均繳款金額與持卡人違不違約呈負相關，由  $p\text{-value}<0.1$ ，可知當持卡人在過去 6 個月平均繳款金額愈多，則違約機率愈小。

**【7】6 個月平均期初應繳金額：**持卡人在 6 個月平均期初應繳金額與持卡人違不違約呈正相關，由  $p\text{-value}<0.1$ ，持卡人在 6 個月平均期初應繳金額愈少，則違約機率愈小。

**【8】過去 6 個月是否有循環息：**持卡人在過去 6 個月是否有循環息與持卡人違不違約呈正相關，由  $p\text{-value}<0.1$ ，持卡人在過去 6 個月有循環息時，則違約機率愈高。

**【9】過去 6 個月繳款金額/應繳金額：**持卡人在過去 6 個月繳款金額/應繳金額的比率與持卡人違不違約呈相關，由  $p\text{-value}<0.1$ ，持卡人在「過去 6 個月繳款金額/應繳金額」比率愈小(應繳金額愈多或過去 6 個月繳款金額愈少)，則違約機率愈高。

## 第二節 實證分析-利用 binary quantile regression 模型

【表 4-4】利用 Binary Quantile Regression 求算出各分量影響持卡人違約機率的基本

統計量<sup>1</sup>

分量	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	係數	係數	係數	係數	係數	係數	係數	係數	係數
截距	3.9182 (2.6688)	4.9596 (2.7526)	1.374 (3.0359)	3.3525 (2.9796)	-1.4968 (2.8530)	3.2288 (2.8463)	-4.3765 (2.6740)	3.0363 (2.8372)	1.2331 (3.1500)
職級	2.8261 (2.6802)	2.9422 (2.6406)	-2.3891 (2.8106)	4.2195 (3.2310)	1.5917 (2.9574)	0.2303 (2.3687)	-4.7157 (1.5740)	4.8940 (2.2477)	-2.1718 (3.6217)
婚姻	4.8948 (3.1356)	2.3327 (2.3182)	3.8293 (1.8475)	-4.4948 (3.3604)	-4.2153 (3.1123)	2.5338 (2.7829)	1.8991 (2.6488)	3.0286 (3.1834)	4.9246 (3.4849)
年收入	0.0046* (0.0020)	0.0117* (0.0048)	0.0824* (0.0110)	0.0187 (0.0097)	0.0289* (0.0222)	0.0074* (0.0010)	0.0414* (0.0012)	0.0420* (0.0031)	0.0554* (0.0078)
過去 6 個月平均繳款金額	-4.8667* (0.1389)	-4.9078* (0.1197)	-4.9470* (0.0833)	-4.9859* (0.0526)	-4.9732* (0.1846)	-4.9559* (0.3635)	-4.9984* (0.7848)	-5.0000* (0.6085)	-4.9379* (0.4649)
6 個月平均期初應繳金額	3.2096* (0.1330)	3.1244* (0.2581)	3.7574* (0.8827)	3.2597* (1.0222)	4.0117* (1.4045)	0.2935* (0.0224)	0.6524* (0.0453)	0.4534* (0.0351)	0.4860* (0.0457)
6 個月交易金額	-1.6337* (0.2262)	-1.7544* (0.3701)	-3.7990* (0.7477)	-3.0910* (1.2960)	-4.9947* (1.9251)	-0.1347* (0.0248)	-2.1646* (0.0222)	-0.7848* (0.0296)	-0.4080* (0.0764)
過去 6 個月是否有循環利息	-3.0837 (3.0694)	3.4471 (2.6519)	-2.4279 (2.9904)	-2.3137 (2.8031)	0.4099 (3.2923)	4.5410 (2.9631)	-0.4801 (3.1881)	4.3524 (3.3534)	1.4367 (3.4436)
過去 6 個月繳款金額/應繳金額	1.7619 (2.4937)	0.5320 (2.5590)	2.5167 (2.4523)	3.6270 (3.5830)	-3.8575 (3.0471)	-0.2587 (3.0740)	-2.2909 (3.0691)	3.8218 (3.5265)	-4.8469 (2.6684)
每月限額	-2.5677* (0.1134)	-2.4971* (0.2220)	-3.1630* (0.7477)	-2.2670* (0.7935)	-2.6249 (1.1107)	-0.0110 (0.0035)	-0.0844 (0.0026)	-0.0674 (0.0046)	-0.1110 (0.0207)

<sup>1</sup> 1.\*為顯著水準為 0.05，2. 括弧中為標準差

【表 4-5】9 個解釋變數的共變異數

	年收入	職級	婚姻	每月限額	6 個月交易金額	過去 6 個月平均繳款金額	6 個月平均期初應繳金額	過去 6 個月是否有循環息	過去 6 個月繳款金額/應繳金額
年收入	1.000								
職級	-0.163	1.000							
婚姻	0.175	-0.135	1.000						
每月限額	0.235	-0.246	0.223	1.000					
6 個月交易金額	0.196	-0.154	0.108	0.446	1.000				
過去 6 個月平均繳款金額	0.198	-0.158	0.120	0.470	0.862	1.000			
6 個月平均期初應繳金額	0.034	-0.133	0.054	0.340	0.369	0.394	1.000		
過去 6 個月是否有循環息	-0.140	0.019	-0.118	-0.111	-0.096	-0.088	0.424	1.000	
過去 6 個月繳款金額/應繳金額	0.068	-0.010	0.020	0.004	-0.001	-0.018	-0.075	-0.112	1.000

【表 4-4】來看，在持卡人信用卡違約風險估計上，本研究是採用 Korads (2002) 所提出的 Binary Quantile Regression 探討在給定不同比例的分量下，2004 年 4 月台灣某家銀行的信用卡客戶資料來探討違約風險的因素。上表是為對年收入、職級、婚姻、每月限額、6 個月交易金額、過去 6 個月平均繳款金額、6 個月平均期初應繳金額、過去 6 個月是否有循環利息、過去 6 個月繳款金額/應繳金額等九個變數，分別作 0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9 分量迴歸所得到的係數估計值。

在年收入與持卡人信貸違約率呈正相關，在我們的觀念裡會想說當持卡人年收入愈高其信用卡違約機率會很小，但我們因為考慮了其他變數進去，則會得到當持卡人年收入愈高其信用卡違約機率會比較大。尤其是在分量 0.3，0.8 和 0.9 的情況下。

「過去 6 個月平均繳款金額」此變數與持卡人是否違約呈負相關，則當過去

6 個月平均繳款金額愈高，其違約機率較低。在一定的金額在同一個時間去償還，而償還的金額比較多，則持卡人相對違約機率會少很多，而銀行也不會收到較多呆帳情況。

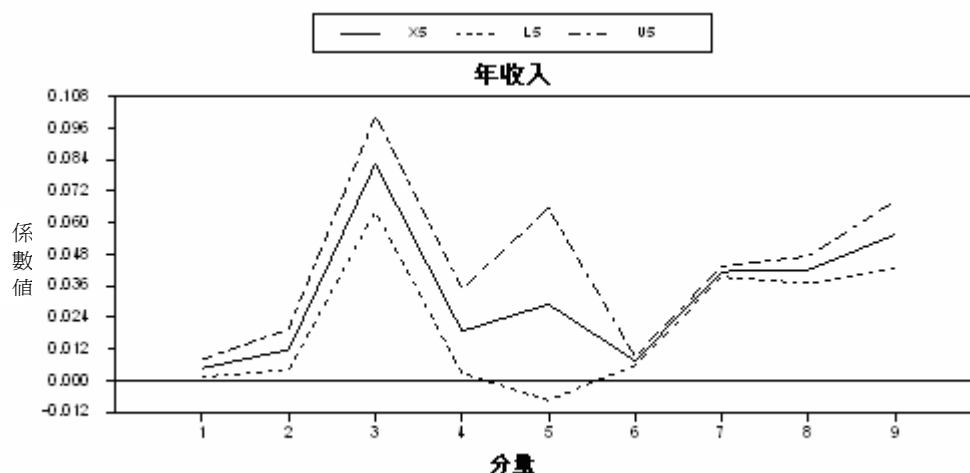
「6 個月平均期初應繳金額」此變數與持卡人是否違約呈正相關，則當 6 個月平均期初應繳金額須繳的愈多，持卡人的負債也愈多的情況下，則較容易發生違約的情況。更在每個分量的情況下，此變數為顯著影響持卡人信貸違約的機率。

「6 個月交易金額」此變數與持卡人是否違約呈負相關，則當 6 個月交易金額愈多，持卡人違約機率會愈低。一般人會覺得當你交易金額愈多反而應該違約機率會愈高。但因為加入了「過去 6 個月平均繳款金額」和「6 個月交易金額」有高度的正相關。當 6 個月平均繳款金額愈多且 6 個月底易金額也愈多，則違約機率愈小。更在每個分量的情況下，此變數為顯著影響持卡人信貸違約的機率。

在每月限額的係數為負，則持卡人的每月限額愈高，其違約機率愈小。而尤其在分量 0.1，0.2，0.3，0.4 的情況下較顯著。

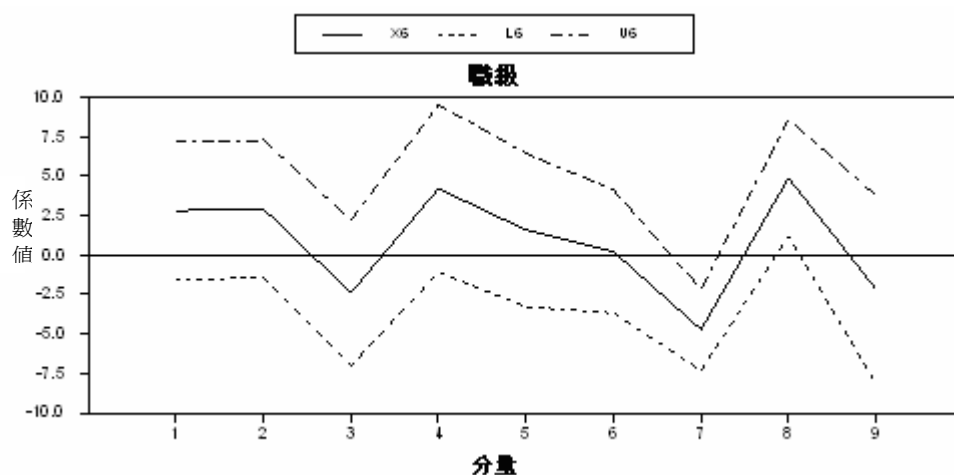
由【圖 4-1】可以看出年收入對於信用卡違約率的影響，在各分量的估計值皆為正的。而且 0.3 分量估計值和 0.6 分量估計值差距最大。

【圖 4-1】在各分量時年收入與信用卡預測違約情況



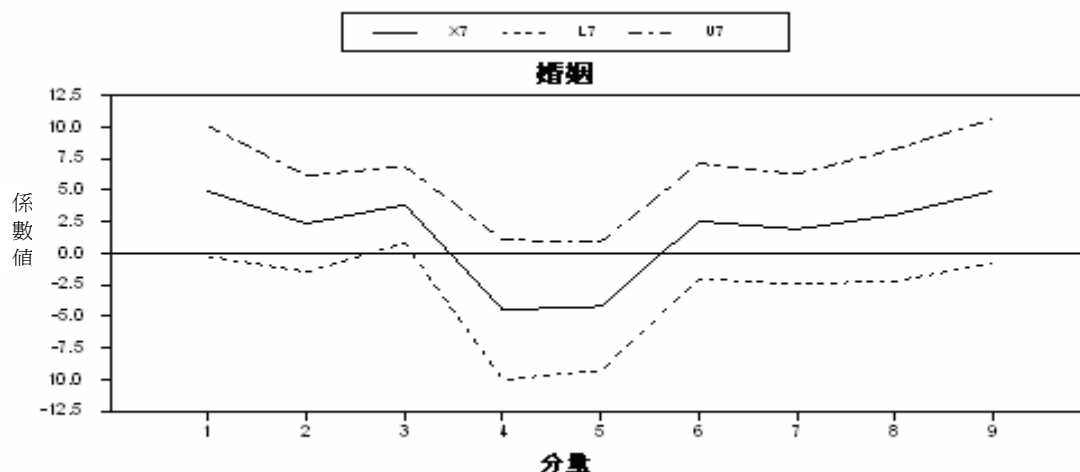
由【圖 4-2】可以看出在分量 0.3 與 07 時其估計值為負，其他分量為正。而且 0.7 分量估計值和 0.8 分量估計值差距最大。職級變數對持卡人是否為違約因素較不顯著。

【圖 4-2】在各分量時職級與信用卡預測違約情況



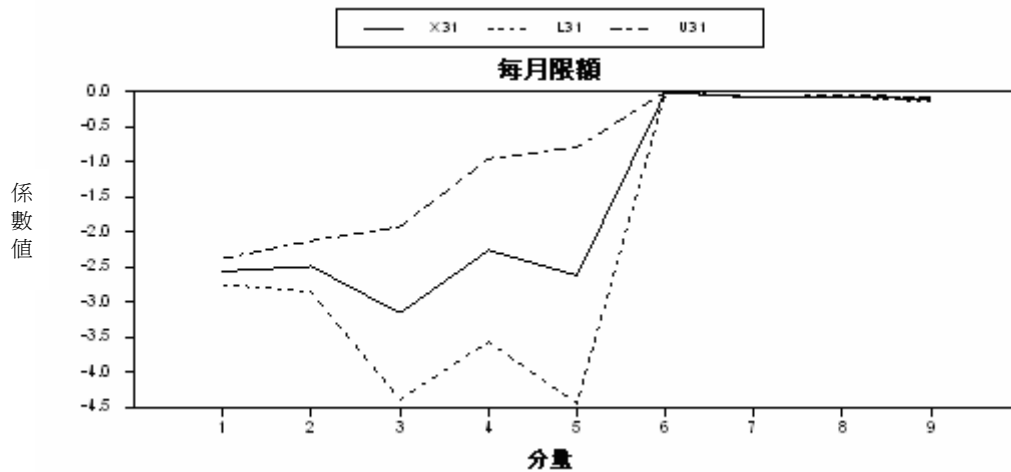
由【圖 4-3】可以看出在分量 0.4 與 0.5 時其估計值為負，其他分量為正。而且 0.4 分量估計值和 0.9 分量估計值差距最大。婚姻變數對持卡人是否為違約因素較不顯著。

【圖 4-3】在各分量時婚姻狀況與信用卡預測違約情況



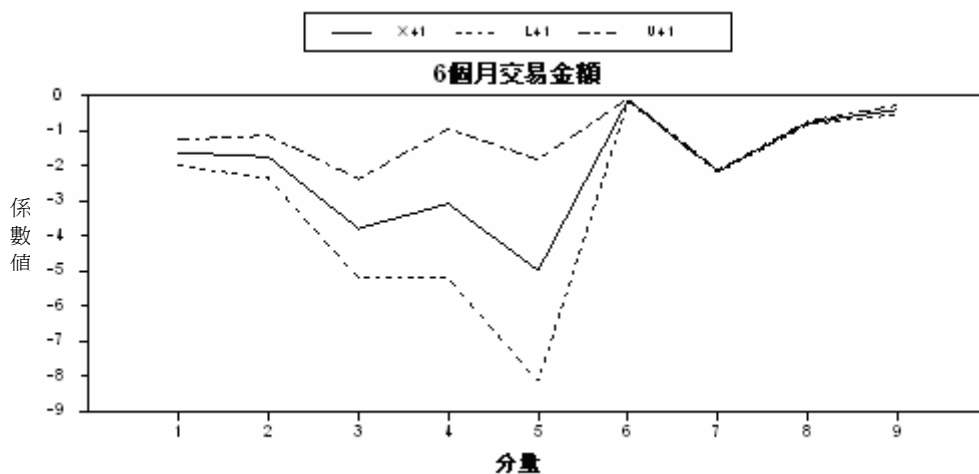
由【圖 4-4】可以看出在全部分量其估計值為負。而且 0.3 分量估計值和 0.6 分量估計值差距最大。而每月限額這變數對持卡人是否為違約因素較顯著。

【圖 4-4】在各分量時每月限額與信用卡預測違約情況



由【圖 4-5】可以看出在全部分量其估計值為負。而且 0.5 分量估計值和 0.6 分量估計值差距最大。而 6 個月交易金額這變數對持卡人是否為違約因素較顯著。

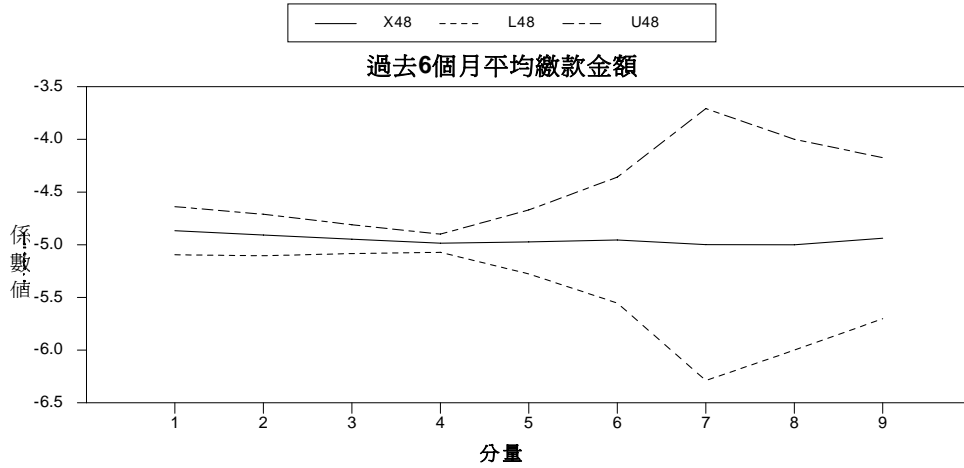
【圖 4-5】在各分量時過去六個月平均繳款金額與信用卡預測違約情況





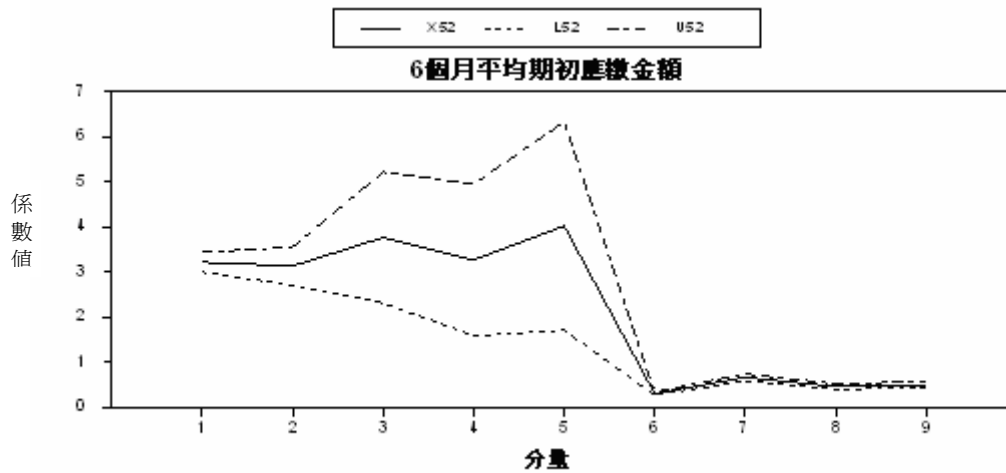
由【圖 4-6】可以看出在全部分量其估計值為負。而且 0.1 分量估計值和 0.8 分量估計值差距比較大。而過去 6 個月平均繳款金額這變數對持卡人是否為違約因素較顯著。

【圖 4-6】在各分量時 6 個月平均期初應繳金額與信用卡預測違約情況



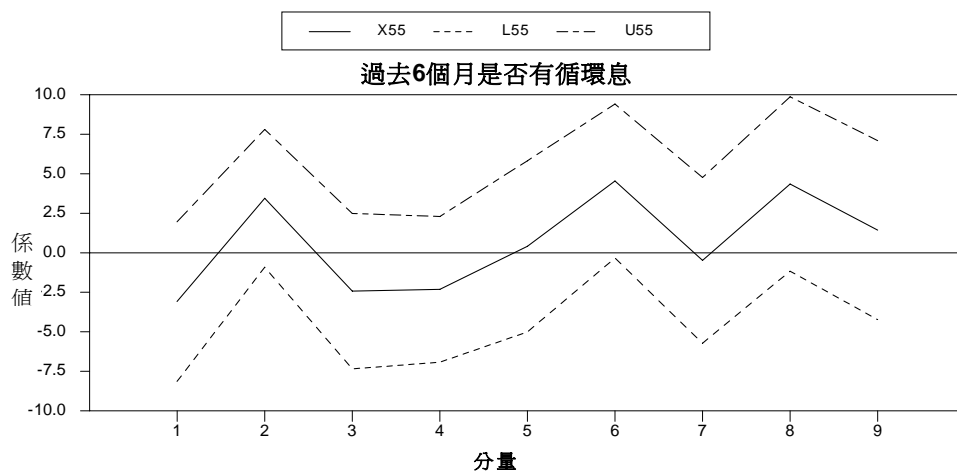
由【圖 4-7】可以看出在全部分量其估計值為負。而且 0.5 分量估計值和 0.6 分量估計值差距比較大。而 6 個月平均期初應繳款金額這變數對持卡人是否為違約因素較顯著。

【圖 4-7】在各分量時 6 個月平均期初應繳金額與信用卡預測違約情況



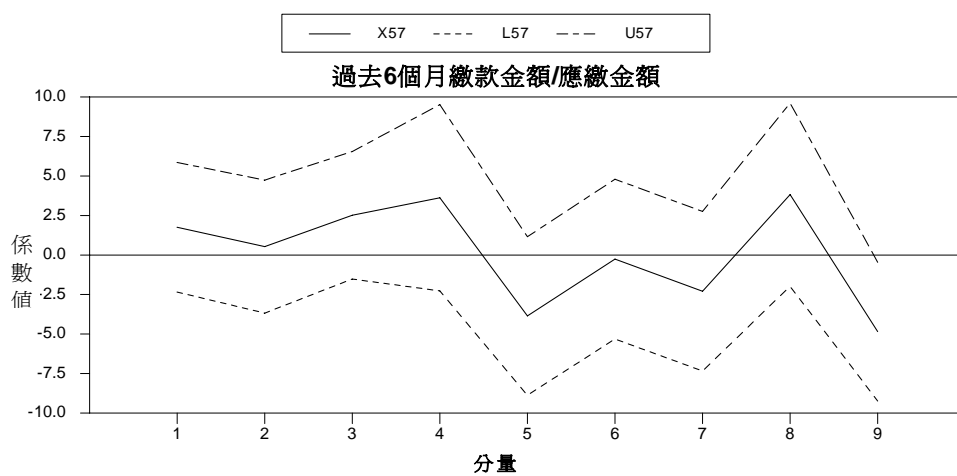
由【圖 4-8】可以看出在分量 0.1、0.3、0.4 與 0.7 時其估計值為負，其他分量為正。而且 0.1 分量估計值和 0.6 分量估計值差距最大。過去 6 個月是否有循環息變數對持卡人是否為違約因素較不顯著。

【圖 4-8】在各分量時過去六個月是否有循環息與信用卡預測違約情況



由【圖 4-9】可以看出在分量 0.5、0.6、0.7 與 0.9 時其估計值為負，其他分量為正。而且 0.4 分量估計值和 0.9 分量估計值差距最大。「過去 6 個月繳款金額/應繳金額」變數對持卡人是否為違約因素較不顯著。

【圖 4-9】在各分量時「過去六個月繳款金額/應繳金額」與信用卡預測違約情況

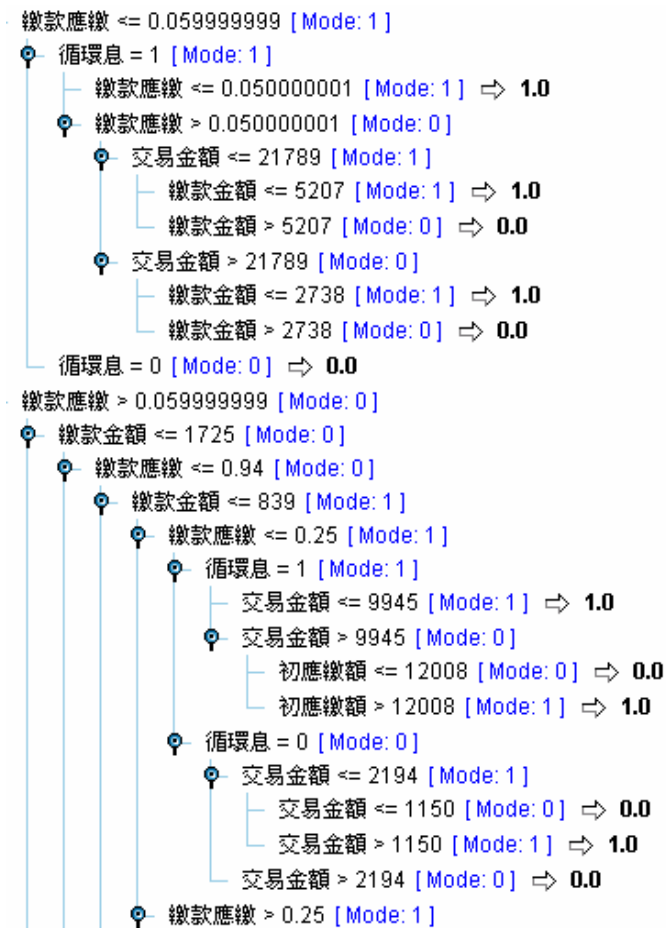


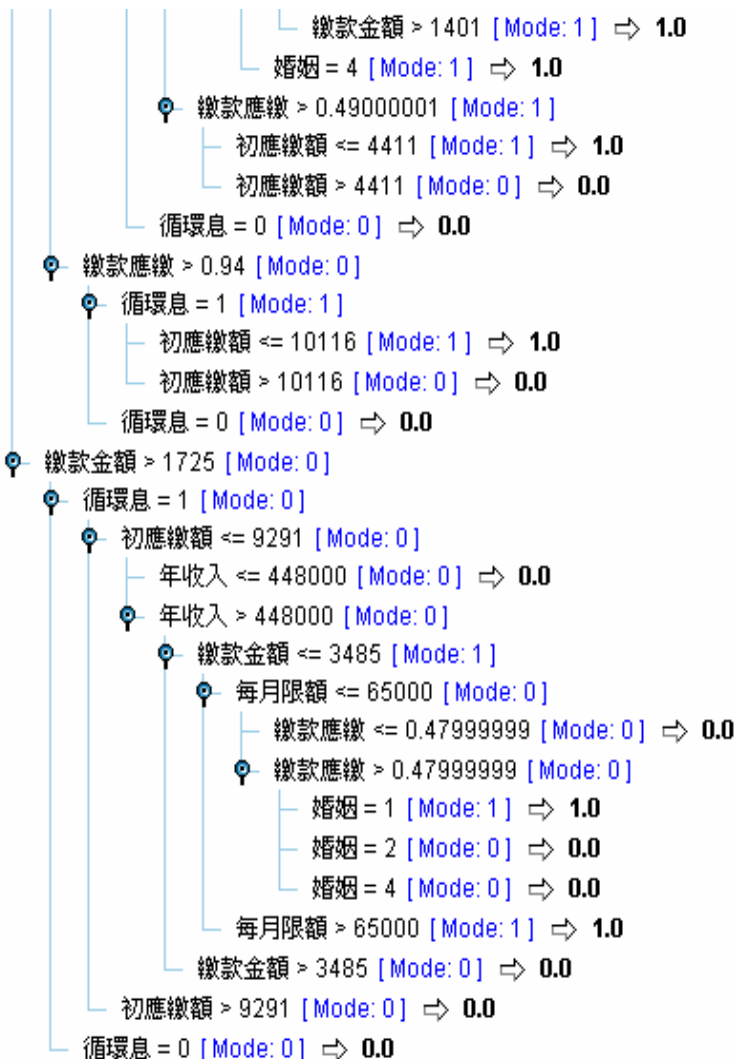
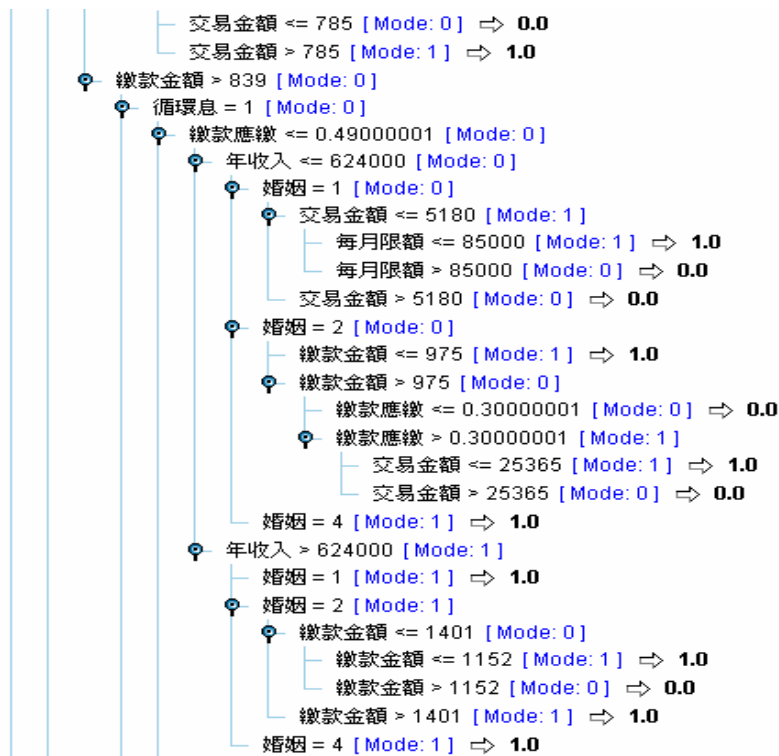
### 第三節 實證分析-利用 Data mining-Clementine(克萊門泰)

Data Mining 所用方法就是用數值方法與演算法的 Recursive(遞迴) 等的技巧，且除了程式設計以外，到最後是花更多的時間去做系統分析 (System Analysis) 與最佳化 (Optimization)。主要還是在於區隔樹、關聯分析與類神經網路這三種方法最為特殊。本文分析著重於「區隔樹」的方法

區隔樹方法的功能，會針對所需要區隔的多變量樣本特性，依據不同變數的解釋能力，使用該變數做出分層區隔的樹狀圖。本文利用九個解釋變來預測持卡人信用卡違約機率的現象且分別使用區隔樹方法來加以說明。

【圖 4-10】使用 Data mining 來預測信用卡違約情況





我們可以由上先看到第一層是由持卡人的繳款率來判斷是否違約，當繳款率小於 0.059 時較會發生違約情況(Mode=1)且繳款率大於 0.059 時比較不會發生違約情況(Mode=0)，依序一層層作分類，找出持卡人在什麼情況下容易發生違約。

【表 4-6】實際違約值與預測違約值

		實際違約值	
		0	1
預測違約值	0	10920	495
	1	352	1355

Results for output field 違約值

Comparing \$C-違約值 with 違約值

<b>Correct</b>	12,275	93.55%
<b>Wrong</b>	847	6.45%
<b>Total</b>	13,122	

Confidence Values Report for \$CC-違約值

<b>Range</b>	0.521 - 0.966
<b>Mean Correct</b>	0.0
<b>Mean Incorrect</b>	0.793
<b>Always Correct Above</b>	0.966 (0% of cases)
<b>Always Incorrect Below</b>	1.0 (0% of cases)
<b>90% Accuracy Above</b>	Never reached requested level
<b>2.0 Fold Correct Above</b>	Never reached requested level

從【表 4-6】可以看出，利用 Data Mining 的區隔樹其預測出違約的機率高達 96.6% 且可以看出信用卡持卡人共同擁有何種特性會發生信用卡違約，利用九個解釋變數來找出一定的規律條件。