

第二章 文獻回顧

第一節 消費者信用貸款

壹、我國消費者貸款的定義¹

民國 74 年	消費者貸款，係指對於房屋購置及修繕、耐久性消費品（含汽車）、子女教育及其他個人小額貸款
民國 91 年 修正	房屋修繕、耐久性消費品（含汽車）、支付學費及其他個人之小額貸款（機關團體職工福利貸款、其他個人消費性貸款），及信用卡循環信用
民國 92 年 10 月 30 日 修正	所謂消費者貸款，謂會員以協助個人置產、投資、理財週轉、消費及其他支出為目的，而辦理之融資業務。消費者貸款係寄望以借款人之薪資、利息、租賃、投資或其他所得扣除生活支出後所餘之資金，作為其還款來源

貳、我國銀行業目前承作消費性金融業務如下

- 1.房屋貸款:提供自有或他人之不動產，向金融機構申請短、中長期貸款，以滿足其購屋、修繕或其他資金需求。
- 2.汽車貸款:消費者提供汽車供擔保向金融機構或分期付款付款公司申請購車融資。
- 3.小額消費性貸款:消費者為個人投資理財之需或資金週轉之目的，向金融機構申貸之擔保或無擔保小額融資。
- 4.信用卡業務:發卡機構依持卡人信用評等給予信用額度，在此信用額度範圍內，持卡人可以「先消費、後付款」，並可避免持有大量現金之風險。

¹ 我國消費者貸款的定義由財政部公告

5. 資產管理帳戶:將客戶現金存放、投資及消費性支付等資金往來。
6. 其他如行政機關公民營事業職工福利貸款、個人繳納綜合所得稅貸款、連鎖加盟店創業貸款、個人有價證券質押貸款等。

參、消費金融的特性

1. 貸款模式方面：消費金融之件數多，每筆金額小，此係因為消費金融面對的是消費大眾。
2. 核准流程方面：消費金融因貸款件數多，金額小，核貸時間短，必須由總行先行規劃消費金融產品，參考本身授信政策，市場狀況，訂定消費金融核貸條件、核貸程序暨作業細節，再交予各營業單位辦理，只要符合規定，營業單位主管即可核貸，一般稱為「公式化貸款」
3. 產品模式：消費金融則是標準化產品，產品藉大量製造過程，以強力行銷，激發客戶需求，進而向銀行購買。
4. 在客戶財務資料取得方面：消費金融之個人所得資料則不完全可得，即申貸者無法提供足夠資料證明其財力及還款來源，且消費金融往往不以擔保品之有無為唯一之考量，故徵信人員必須在有限的資料下，蒐集研判申貸人之信用狀況。
5. 貸款利率：因消費金融每筆金額小，貸款承做單位成本高，利率反映成本故消費者 貸款利率普遍較企業金融之利率高。
6. 風險哲學：依據客戶不同的風險程度，設計不同的款產品。

【表 2-1】 本國銀行辦理消費者貸款餘額統計表⁶

年月底	消費者貸款						
	小計	購置住宅貸款	房屋修繕貸款	汽車貸款	機關團體職工福利貸款	其他個人消費性貸款	信用卡循環信用餘額 ³
7712	840,875	415,201	236,062	7,742	62,779	119,091	
7812	975,375	490,867	283,357	7,581	58,517	135,053	
7912	1,070,806	536,653	345,663	6,300	47,436	134,754	
8012	1,246,127	667,269	398,835	5,414	42,918	131,691	
8112	1,528,810	888,471	457,423	11,613	39,872	131,431	
8212	1,742,423	1,032,875	529,349	17,823	38,108	124,268	
8312	2,177,078	1,373,730	591,004	26,505	37,232	148,607	
8412	2,559,158	1,696,043	634,031	25,807	37,970	165,307	
8512	3,022,425	2,002,590	683,108	30,109	47,637	258,981	
8612	3,598,047	2,289,903	765,514	33,158	58,303	370,185	80,984
8712	3,873,602	2,420,346	819,222	40,956	59,092	417,886	116,100
8812	4,050,223	2,539,707	799,768	46,431	68,727	444,514	151,076
8912	4,296,486	2,651,772	849,621	51,607	76,737	480,480	186,269
9012	4,270,124	2,623,276	780,717	63,520	73,354	495,968	233,289
9112	4,442,484	2,768,188	695,283	69,299	90,908	527,988	290,818
9212	4,969,421	3,038,881	690,016	80,452	172,455	613,140	374,477
9312	5,861,529	3,482,662	710,373	109,471	181,936	952,118	424,969
9412	6,627,973	3,972,402	725,369	140,374	171,361	1,152,875	465,592
9504	6,605,756	4,080,522	735,043	137,293	168,375	1,073,829	410,694

第二節 銀行授信評估原則之探討

由於信用卡的出現，改變了國人的消費習慣，也帶動消費性金商品的發展。各發卡銀行為提高發卡量或達到經濟規模，莫不以寬鬆的信用條件來審核申請人的各項財務狀況。如此一來，嚴重影響授信品質，導致信用卡催收帳款增加，呆帳比率提高。而授信審核是防範呆帳發生的第一道防線，完善的授信審核制度能客觀地判定申請人的信用狀況，以減少逾期帳款的發生。

在國內商業最普遍用的是傳統授信風險分析方法²而去延申作分析。然而隨著統

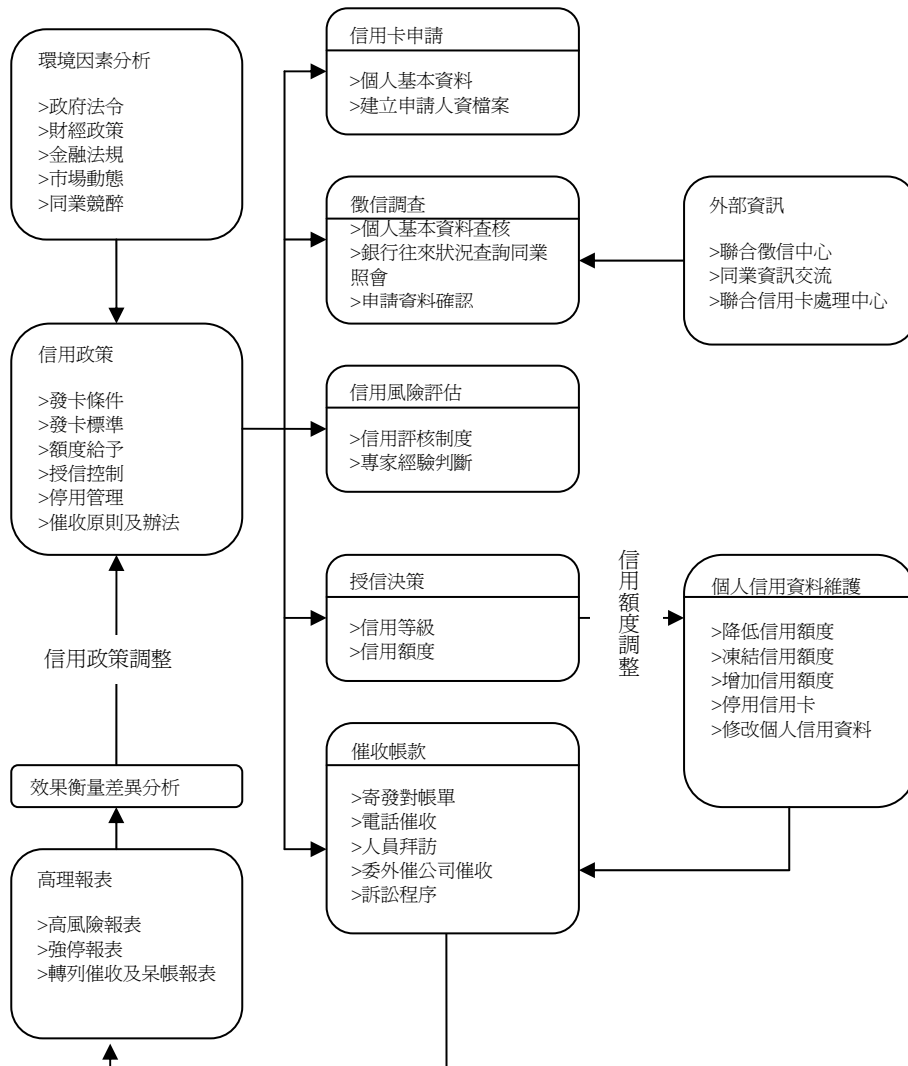
² 信用卡授信五個評估構面：客戶債信能力、客戶自有資產能力、客戶工作性質、客戶償債能力、客戶誠信度。

計學、計量工具的發展，各種衡量信用風險的模型model被架構推出，世界知名

銀行亦投注人力物力發展計量分析為主的風險管理部門，建立授信風險量化指

標。

【圖 2-1】 信用卡風險管理考量變數及作業流程³



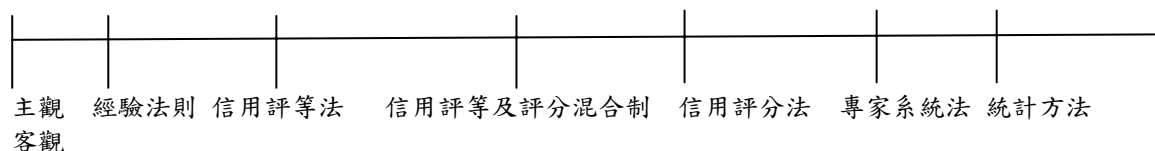
³ 「信用卡風險管理考量變數及作業流程」資料來源：張文生(2001)

壹、現行金融機構常用的授信評量方法

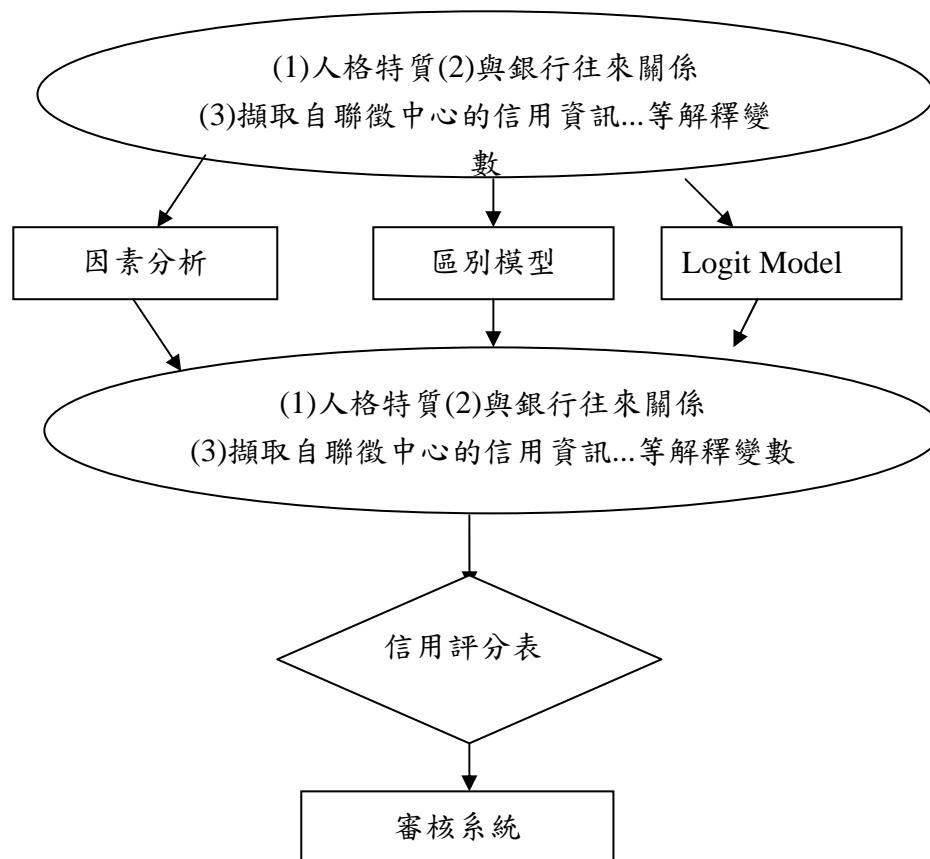
【表2-2】現行金融機構授信評量方法⁴

	授信評量方法	定義	優點	缺點
主觀的方法	經驗法則	此為傳統上金融機構用以評斷風險的方法，係指依據授信及回收經驗在試驗過程後，列出拒絕授信對象，或經由徵授信人員個人的經驗，以主觀判斷作成授信決定。	* 簡單易行 * 缺乏客觀標準	* 制度不健全，易生弊端
客觀的方法	(Rating) 信用評等法	將客戶信用品質細分為若干項目，然後有系統的對每個個別項目分別考評，並賦予適當的等級，用以代表該顧客的綜合信用評價。	* 徵信成本隨評等項目多寡而增減	* 評等不易 * 須有一定之評判標準 * 評等項目選擇不易
客觀的方法	(Ranking) 信用評分法	信用評分制度乃是將顧客信用資料透過信用評分表分析，把各考評項目之得分加總而得到一個信用分數，分數愈高表示顧客償債能力愈強。主要在改善信用評等所產生的問題，藉由將評估項目由評等改為實際量化資料，明確訂定各種變數實際分數，使得授信人員可以較客觀的評分。	* 具補償性 * 授信準則清楚 * 評分方式客觀	* 評分項目權重無絕對之客觀標準 * 評分項目選擇不易 * 評分項目細項區隔不易
客觀的方法	評分混合制	此乃綜合上述兩種方法而來，先利用信用評分表計算顧客的信用分數，按照其分數編入應屬等級，相同等級者給予相同之授信條件。	* 結合信用評等、評分制度的優點	* 複雜度高
客觀的方法	統計方法	國內外學者在研究銀行授信評估時，大多採用如區別分析Probit或Logistic Regression等統計方法。	* 較具客觀性 (較客觀，但不一定偵察較好)	* 變數選取不易 * 複雜度高
客觀的方法	專家系統法	利用資訊科技建立一套自動評核系統，用以輔助或取代授信專業人員的審核工作。其中包括傳統知識庫系統、範例學習法、類神經網路等。	* 具客觀性	* 實行不易

⁴資料來源：參考林建州（2001），銀行個人消費信用貸款授信風險評估模式之研究，中山大學財務管理研究所碩士論文（由上圖可以表示出主、客觀之程度）



貳、消費金融信用風險模型之實證研究架構



參、消費金融信用風險模型之 PD（違約機率）的計算

step1:搜集個人小額信貸的資料

step2:信用評分模型研究

step3:違約定義研究

step4:樣本選取-利用認為最適的模型選出最有效的解釋變數

step5:違約機率之預測

肆、國內外消費信用授信評量相關文獻之重要變數

國內、國外有關消費信用授信評量的文獻十分豐富，包含信用卡、信用貸款、房屋貸款等，諸多論文找出信用風險的因素，有些文獻三個構面來分

1.以持卡人之人格特性⁵-性別、年齡、職位、年收入等申請書的填寫項目作為研究因素

2.持卡人與發卡銀行的往來關係面⁶-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。

3.來自金融聯合徵信中心的信用資訊構面⁷-持卡人曾被金融機構查詢的次數、持卡人於金融聯合徵信中心遲繳或未繳的紀錄、持卡人過去是否有不良或異常記錄。

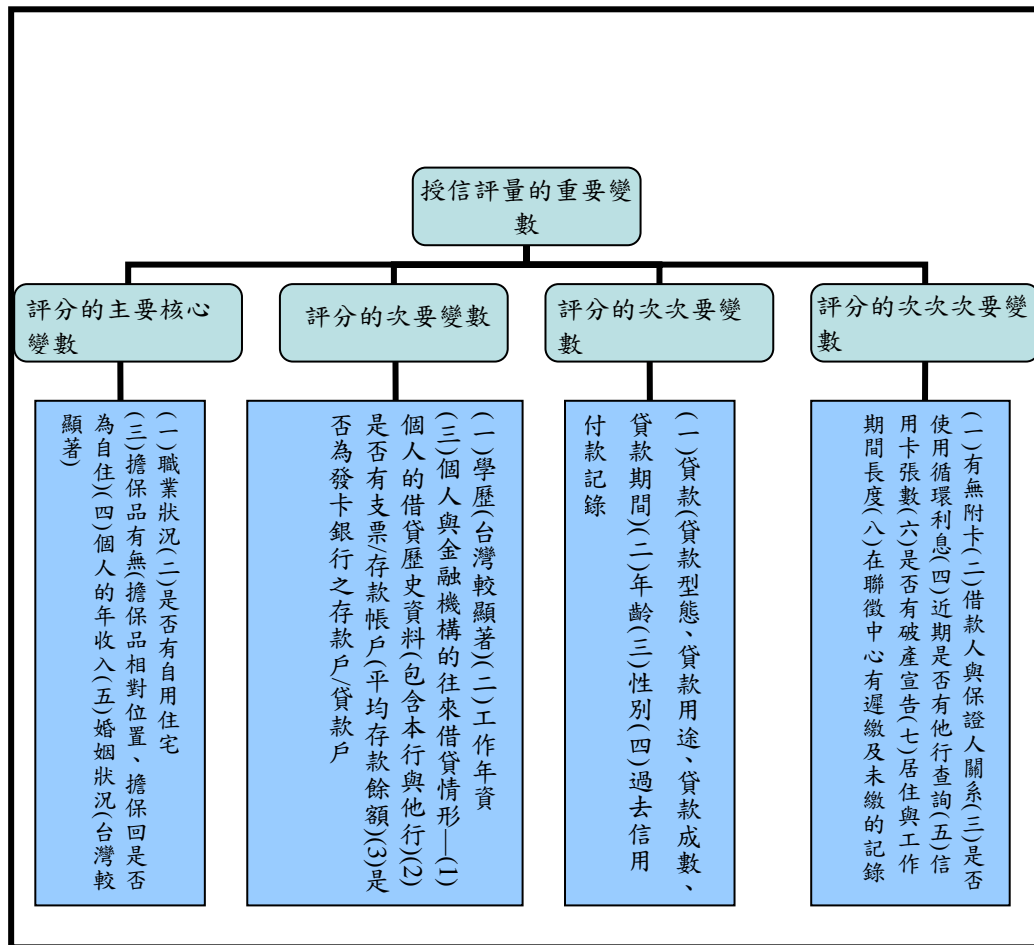
相對於現在信用卡持有者的行為模式較不易看出他們的違約機率。故本文對於研究變數，將部分重要之研究文獻整理說明如下：

⁵江淑娟(2003)，李美笑(2002)，李明謙(2002)與張文生(2002)

⁶江淑娟(2003)

⁷林重光(2004),林旭青(2004),施夢隆(1999)

【圖2-2】國內外授信評量相之重要變數⁸



⁸國內外授信評量相之重要變數是經由作者整理出來且由指導教授沈老師中華給的意見衡量出以上變數的分類。

第三節 信用卡市場的文獻探討

陳瑞雲(2004)在此研究中總共有 568 件樣本，其中使用循環利息的持卡人共 117 戶，共篩選出持卡人 21 個變數進行研究分析，分析結果發現有四個變數年資、年齡、年所得、有無附卡等四個變數與循環利息成顯著相關。

江淑娟(2003)以某家銀行之信用評分，對該銀行現有之信用卡持卡人作信用卡信用額度評等表中之各項信用評等因素與持卡人逾期違約可能機率之間的關係，以提供銀行信用卡核單位，作為信用卡徵審過程中，所使用之各項評等因素權重之依據。

鄭廳宜(1999)研究如何建立一套較適當的審核信用模式，以降低逾期放款比率，並提昇發卡機構的經營績效。其研究結果顯示：

- 1、信用授信評估以因素分析法可建立成五個評估構面，包括：客戶工作性質、客戶自有資產能力、客戶債信能力、客戶償債能力及客戶的誠信度。
- 2、信用卡授信五個評估構面中以客戶債信能力權重最高占全體25.8%權重，其次依序為客戶自有資產能力占24.8%權重，客戶工作性質占20.8% 權重，客戶償債能力占15.7%權重，最低權重為客戶的誠信度僅占全體12.9% 權重。
- 3、在評估項目方面，客戶的職業種類、客戶家庭收入高低、客戶與其他銀行往來狀況、客戶收入高低、客戶票信是否良好，其權重均占40%以上，顯示在信用卡評估上相當重視兩要素：1.所得：職業種類如醫生、律師、會計師等相對所得

較高；2.個人信用：此文研究發現「曾有退票紀錄」或「有不良信評」者，在信用卡申請上除非有其他非常有利之因素，否則獲得發卡機會將非常小。相對地只要有正當職業且無不良信評或退票紀錄，則很容易獲得發卡機構的發卡。

4、此研究所建立的AHP 授信評估模式與目前所使用的評分表進行實證比較分析，結果發現：(1)應用此研究AHP模式對逾期戶是呈現降低評估分數與發卡額度，對正常戶則是呈現提高評估分數與發卡額度(2)由此研究AHP方法所建立評分表與傳統授信評分表相比較可以顯著性地降低逾放比率與提高發卡機構的利潤。

第四節 信用風險模型的文獻探討

利用最小平方估計法(OLS)去作迴歸分析時，有下列三點是無法解決的⁹：

1. OLS 藉由極小化殘差平方和求得一組參數估計值，僅能用以描述解釋變數對被解釋變數的平均影響程度。
2. OLS 對極端值相當敏，意即一個太大或太小的樣本觀測值，就可能使估計的結果相當不準確。
3. OLS 估計式只有當誤差項為常態分配時，才有較好的統計效率。

因此研究者面臨上述問題時，往往必須仰賴其他的迴歸分析工具。

2.4.1 因素分析

1.模型建立：因素分析假設樣本單位在某一變數上的反應是由二個部份所組成(1)各變數共有的部份稱為「共同因素」

(2)各變數所獨有部份稱為「獨特因素」

獨特因素與共同因素和其他因素的獨特因素均無關聯。

因素分析依其目的可大致分為二類

(1)探索性因素分析:目的在於從一組雜亂無章的變數中找出共同的屬性，以建來新的假設或發展新的理論架構

(2)驗證性因素分析:目的在於驗證研究人員已有的理論架構

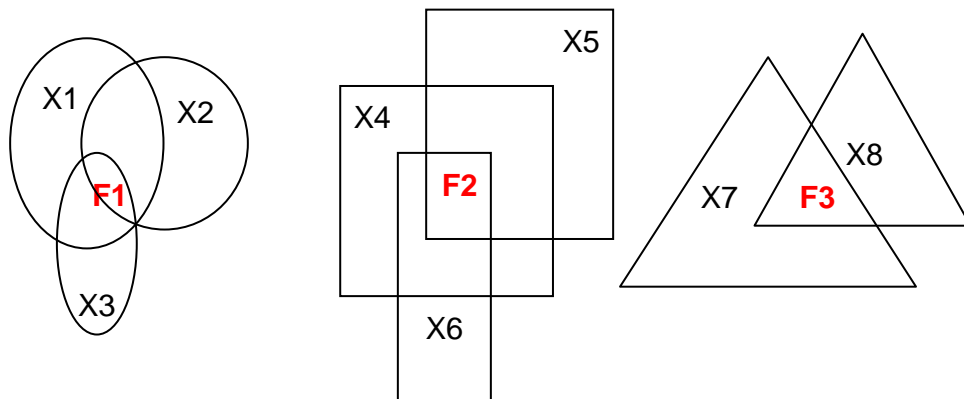
⁹ 引用王心妙(2005)，影響新生兒體重的因素分析：分量迴歸的應用，國立政治大學經濟研究所碩士論文。

決定共同因素之數目

(1)可採用 H.Kaiser:以變異數 1.0 為分野,變異數小於 1 之因素即不予選取,但此法會產生若變數少於 20,取出的變數偏少,而變數多於 50,取出的變數偏多的缺點。

(2)可採用 R.Cattel 發展之陡階檢定來決定,其方法是將每一個因素所能解釋之變異數畫在同一個圖形中,將各點連線,把陡降後曲線走勢趨於平坦的因素捨棄

【圖 2-3】因素分析之幾何圖形



令 Z 、 X 、 E 均為標準化之分數(即平均數為 0,標準差為 1)

因素分析模為 $Z_{ji} = \alpha_{j1}F_{1i} + \alpha_{j2}F_{2i} + \dots + \alpha_{jk}F_{ki}$

Z_{ji} = 第 i 個樣本單位在第 j 個變數上的值

F_{ki} = 第 i 個樣本單位在第 P 個共同因素上的值

α_{jp} = 因素權重,用以表示第 P 個共同因素對第 j 個變數之變異的貢獻

E_{ji} = 第 i 個樣本單位在第 j 個獨特因素上的值

β_j = 第 i 個變數之獨特因素的權重

k = 共同因素的數目

在進行因素分析之前，應先確定各變數觀察值間具有共同變異之存在，如此才值得作因素分析。

在決定選取變數時，採用每一個因素其因素負荷量大於 0.6 或小於 -0.6，作為投入的變數。依此準則選取代表的變數如下：

F1: 個人概況：1. 年齡 2. 職業狀況 3. 家庭狀況 4. 教育程度 5. 性別

F2: 個人所得：1. 年收入(含配偶) 2. 現任職業年資

F3: 個人負債：1. 票信記錄 2. 向金融機構總借款佔家庭收入每月

F4: 個人信用 1. 信用卡往來情形 2. 金融機構往來情況

把上面 F1.F2.F3.F4 代入 OLS 裡

2. 區別函數形式如下(假設每個資料有 P 種特性)

$$Z = W_0 + W_1 X_1 + W_2 X_2 + \dots + W_p X_p$$

Z 為區別函數，且對每一類資料而言， W_1 為 X_1 的係數， W_0 為常數項，

區別分析的目的是要辨識此一區別函數的各個權數(W_i)

即極大化之公式為
$$\text{Wilks' Lambda} = \frac{SS_b(\text{組間離均差平方和})}{SS_i(\text{組內離均差平方和})}$$

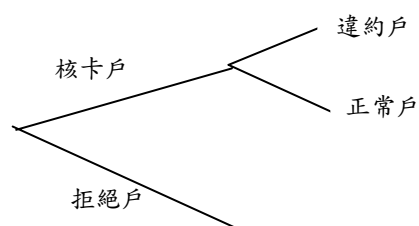
線性區別函數為資料觀察值在不同特性上的加權線性組合，加權的權數(W_i)則代表各種特性(X_i)的相對重要程度或影響力。

2.4.2 Bivariate probit 模型

使用 Boyes et al.(1989), Jacobson-Roszbach(2003)發展出來的。主要是由兩條 Probit 迴歸式組成

(1)式是以發卡銀行是否核卡為分類標準，被解釋變數 Y1 為分類變數：意指拒絕戶為 0、核卡戶為 1

(2)式是以核卡戶是否違約作為分類標準，被解釋變數 Y2 同樣為分類變數：假設違約戶為 0、正常戶為 1



$$(1)Y_{1i} = \alpha_{1i} + \beta_1 * X_{1i} + \varepsilon_{1i}$$

$$(2)Y_{2i} = \alpha_{2i} + \beta_2 * X_{2i} + \varepsilon_{2i}$$

$$(1)Y_{1i} = \alpha_{1i} + \beta_1 * X_{1i} + \varepsilon_{1i}$$

$$(2)Y_{2i} = \alpha_{2i} + \beta_2 * X_{2i} + \varepsilon_{2i}$$

$i=1,2,\dots,N$ (申請者總數) $s=1,2,\dots,W$ (持卡戶總數)

α_j ：截距； β_j ：迴歸參數； ε_j ：殘差項； $j=1,2$

X_{1i} 為 $1 * K_i$ 向量的解釋變數， $K=1,2,\dots$ 為解釋變數的個數

X_{2i} 為 $1 * R_s$ 向量的解釋變數， $R=1,2,\dots$ 為解釋變數的個數

並假設模型符合均數為 0，標準差為 1 的二元標準常態分配，且相

關係數為 ρ

此模型可分為拒絕戶、正常戶、違約戶

$$l = \prod pr(\text{拒絕戶}) * \prod pr(\text{正常戶}) * \prod pr(\text{違約戶})$$

$$\begin{aligned} \ln l = & \sum_{i=1}^N (1 - y_{1i}) \ln[1 - \Phi(x_{1i}, \alpha_1)] \\ & + \sum_{i=1}^N y_{1i}(1 - y_{1i}) \ln\{\Phi(x_{1i}, \alpha_1) - \Phi_2(x_{1i}, \alpha_1, x_{2i}, \alpha_2; \rho)\} \\ & + \sum_{i=1}^N y_{1i}y_{2i} \ln \Phi_2(x_{1i}, \alpha_1, x_{2i}, \alpha_2; \rho) \end{aligned}$$

其中 $\Phi(X_{1i}, \alpha_1)$ 和 $\Phi_2(X_{1i}, \alpha_1, X_{2i}, \alpha_2; \rho)$ 分別代表單一元與二元標準常態累積

分配函數，後者的相關係數為 ρ

2.4.3 CART MODEL

這些我們用來尋找特徵的已分類資料可能是來自我們的現有的歷史性資料，

或是將一個完整資料庫做部份取樣，再經由實際的運作來測試；譬如利用一

個大的郵寄對象資料庫的部份取樣來建立一個 **Classification Model**，以後再

利用這個 **Model** 來對資料庫的其他資料或是新的資料作預測。

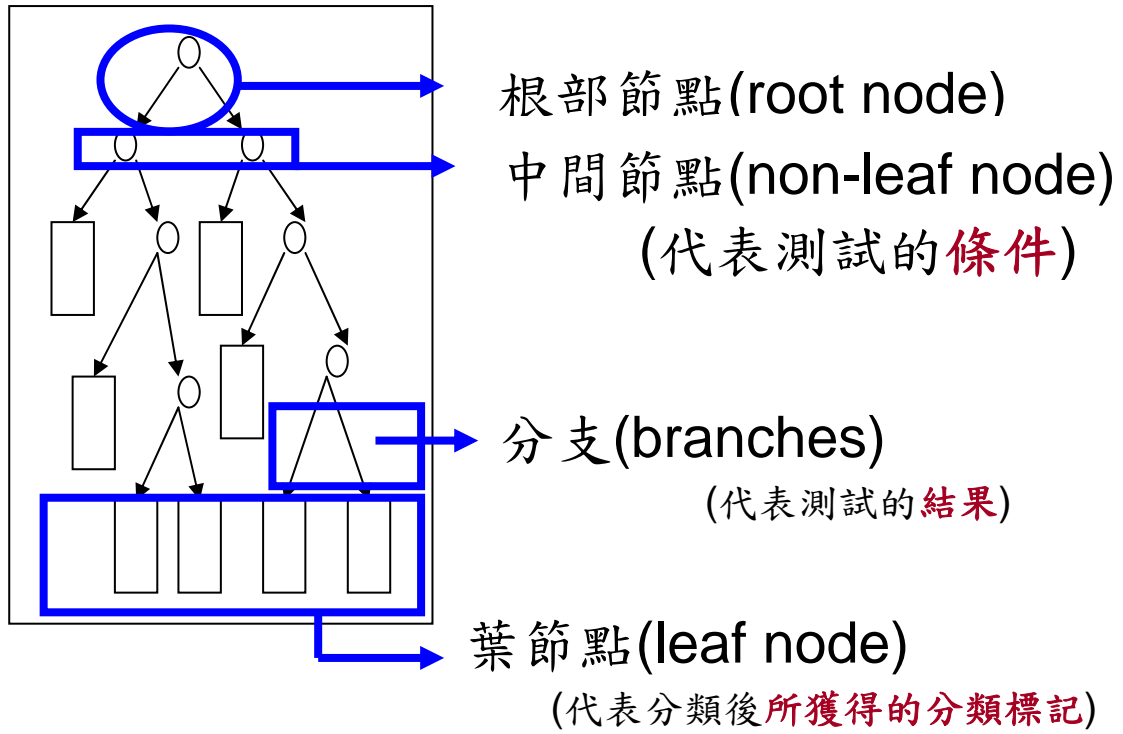
Classification 通常會牽涉到兩種統計方法：**Logistic Regression** 以及

Discriminant Analysis。然而因為 **Data Mining** 已漸普遍，所以 **Neural Nets**

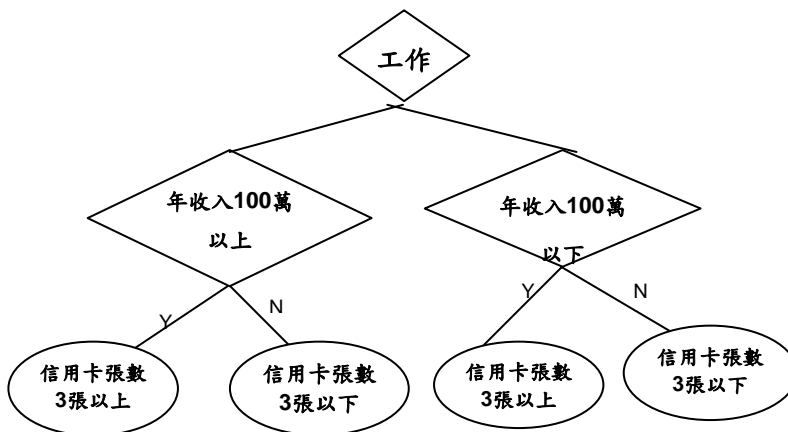
以及 **Decision Tree** 也漸漸受到採用。雖然這些統計方法本身都十分複雜，但

使用者並不會牽涉到這些繁雜的統計。

【圖 2-4】 CART MODEL 之 Decision trees



【圖 2-5】 CART MODEL 之 Decision trees 例子



2.4.4 Logistic Regression

Logit 模式：於累加機率分配函數 Logistic 函數；為屬質因變數迴歸模型的一種，應變數為分立。

$$\text{Logit 迴歸基本模式 } P_i = F(Y_i) = \frac{1}{1 + e^{-Y_i}}$$

$$Y_i = \alpha_i + \sum \eta_{ih} \text{Personal}_{ih} + \sum \rho_{im} \text{solvency}_{im} + \sum \theta_{in} \text{Bankcontact}_{in} + \varepsilon_i$$

$$P_i : \text{為事件 } i \text{ (違約或未違約) 發生的機率}$$

$$Y_i \text{ (實際觀察到的應變數)} = \begin{cases} y_i = 1 \text{ (不良戶) if } y^* > 0 \\ y_i = 0 \text{ (正常戶) if } y^* \leq 0 \end{cases}$$

$F(\square)$ 為 logit 的累積機率分配函數, $i=1,2,\dots,N$ (總樣本數)

$h,m,n=1,2,\dots$ (三大構面的變數個數)

α_i = 截距; $\eta_{ih}, \rho_{im}, \theta_{in}$: 迴歸參數; ε_i : 殘差項

Personal_{ih} = 持卡戶的人格特質面與變數(性別、申請時的年齡、學歷、婚姻、職位、工作年數)

solvency_{im} = 擷取自銀行的資訊(年收入、過去6個月平均繳款金額、過去6個月(3個月、2個月)繳款次數、6個月平均期初應繳金額、6個月平均每月消費款/每月限額、過去6個月繳款金額/應繳金額)

bankcontact_{in} = 擷取自銀行的資訊(擁有信用卡張數、擁有信用卡種類數、擁有Master卡數、擁有VISA卡數、擁有金卡數、擁有白金卡數、每月限額)

林建州（2001）與蔡明憲（2002）研究採用區別模型、Logit模型探討影響消費者信用貸款客戶可能產生正常戶及違約戶的關鍵因素，並利用信用評分方式建立一套消費者信用貸款審核模式，其實證結果Logit模型的準確率高於區別模型的準確率。

陳鴻文（2002）尋求建立一個適用於銀行個人小額信用貸款貸款評量模型，係以表內變數及表外變數為參考值。利用Logit模型為該論文研究方法，實證結果：顯著變數的重要性由大至小依序為是否使用循環利息，近期是否有它行查詢、性別、服務年資，最小為信用卡張數和本研究在只有表內變數時鑑別度為67%，再加入表外變數後鑑別度可提升至91%。結論：良好的授信模式有利銀行的經營，個人小額信用貸款尤其需良好的評分表及風險控管，發現表外變數的加入更可大幅提升模型的鑑別能力，而增加模型的可利用性。

2.4.5 Quantile Regression 與 Binary Quantile Regression

Quantile Regression 係最早由 Koenker and Bassett(1978)首先提出將條件平均的最小平方模型發展為條件分量函數的模型估計且極小化不對稱之絕對誤差的加權總和來估計不同分量下之係數。分量迴歸模型之估計係數，可用於探討解釋變數對於被解釋變數在給定「特定分位數之下的邊際效果」，提供不同分位數的估計

結果，可用於觀察被解釋變數的整個分配狀況。在實證上，分量迴歸模型不只可用來解釋平均的狀況，更常用來觀察分配尾端的情況。

理論模型詳述如下

假設一個隨機變數 Y 的分配為 F_Y ，而 Y 的 θ 分量記為 $Q_\theta(Y)$, $\theta \in (0,1)$

令 $Q_\theta(Y)$ 等於某個數值 q_θ ，則 $Q_\theta = F_Y^{-1}(\theta) = q_\theta$

(1)

式(1)表示該分配其中會有 θ 部分小於或等於 q_θ ，而 $(1-\theta)$ 部分大於或等於 q_θ ，故 q_θ 可經由下列條件求解得到

$$q_\theta = \arg \min [\theta \int_{y \geq q_\theta} |y - q_\theta| dF_Y(y) + (1-\theta) \int_{y < q_\theta} |y - q_\theta| dF_Y(y)]$$

(2)

令 X 與 Y 為兩個隨機變數，此時在給定 X 下，設定一條件分配為 $F_{Y|X}$ ，則 Y 的條件分量為 $Q_\theta(Y|X) = F_{Y|X}^{-1}(\theta)$

(3)

條件分量 $Q_\theta(Y|X)$ 是 X 的函數，令 $Q_\theta(Y|X) = q_\theta(\theta)$ ，在給定 $X = x$ 下， Y 小於 $q_\theta(\theta)$ 的機率等於 θ ，此時可看出 $X = x$ 時之條件分量，其亦可表示如下：

$$q_\theta = \arg \min [\theta \int_{y \geq q_\theta(x)} |y - q_\theta(x)| dF_{Y|X=x}(y) + (1-\theta) \int_{y < q_\theta(x)} |y - q_\theta(x)| dF_{Y|X=x}(y)]$$

(4)

若 $\{y_t, x_t\}_{t=1}^T$ 代表一組雙變量隨機變數， y_t 與 x_t 分別為從 Y 與 X 所得到的 T 個樣本，因此樣本條件分量 $q_\theta(y_t | x_t)$ 則為

$$q_{\theta}(y_t | x_t) = \arg \min[\theta \int_{y_t \geq q_{\theta}(x_t)} |y_t - q_{\theta}(x_t)| + (1-\theta) \int_{y_t < q_{\theta}(x_t)} |y_t - q_{\theta}(x_t)|]$$

(5)

因此，在給定的 x_t 之下，將有 θ 部分的 y_t 小於或等於 $q_{\theta}(x_t)$ ，而有 $(1-\theta)$ 部分的

y_t 大於或等於 $q_{\theta}(x_t)$ ，此可看出 Y 在不同的 θ 下，受 x 影響程度之差異，故分量

迴歸模型之估計模式將可詳述如下。假設一線性模型如下：

$$y_t = x_t' \beta + e_t \quad t = 1, \dots, T$$

(6)

其中 x_t 是 $K \times 1$ 的行向量，表示 K 個所有解釋變數(包含一個常數項)的第 t 個觀察值， β 是 $K \times 1$ 的行向量，包含了各個變數的迴歸係數， e_t 是誤差項，也可寫為 $e_t(\beta)$ 。

依(5)及(6)式，迴歸參數 β_{θ} 的估計式可表為

$$\hat{\beta}_{\theta} = \arg \min[\theta \int_{y_t \geq x_t' \beta} |y_t - x_t' \beta| + (1-\theta) \int_{y_t < x_t' \beta} |y_t - x_t' \beta|]$$

(7)

此為 Koenker and Bassett(1978)首先提出的分量迴歸，亦即只要給定正負絕對值誤差不同的權數，即可得出不同分量迴歸估計式。當 $\theta = 0.5$ 時，則上式乘以 2 即成為 $\sum \sum_{t=1}^T |y_t - x_t' \beta|$ ，此為 LAD 的估計式， $\theta = 0.5$ 時的分量迴歸也稱為中位數迴歸，而中位數迴歸只是分量迴歸中的一個特例。

且 Buchinsky(1994) 利用分量迴歸法來應用在「工資分配結構」上，其描述

整個條件工資分配，並探討在工資分配的不同樣本點下，工作技能報酬的改變，特別是受到教育和工作經驗影響方面。在國內許瑞宏(2003)以誤差修正模型之分量迴歸來估計「台灣貨幣需求函數」。

Korads (2002)運用Quantile Regression Model 以「應變數為二元型態¹⁰」的方式作延申且得到為binary quantile regression model，以binary quantile regression model運用在估計「銀行客戶的信用評分」，以德國銀行 1000 名貸款戶為樣本，將其樣本分為估計組(800 人)和測試組(200 人)，比較同質、異質的probit model 和binary quantile regression model 之違約機率，並將誤差機率分解成型I誤差(貸款違約戶沒被預測出違約)與型II誤差(正常戶被預測成違約)。在模型的正確率預測方面，估計組的結果顯示在所有的機率截點上，分量迴歸結果優於probit模型；但在樣本外的估計結果顯示，在 20%~60%的機率範圍內，probit模型的表現卻是優於分量迴歸模型的。

而二元分量迴歸與分量迴歸最大的不同，在於分量迴歸的被解釋變數為連續性的變數，而二元分量迴歸的被解釋變數為間斷的兩種屬質變數，如此的變數設定和logit及Probit模型的對於變數的設定相同，用於分析被解釋變數為二元的情形。關於二元分量迴歸，我們可以設定以下模型：

$$Q_{y_i^*}(\tau | x_i) \equiv F_{y_i^*}^{-1}(\tau | x_i) = x_i' \beta(\tau) \quad (7)$$

其中 $\tau \in (0,1)$ ， y^* 為真實違約與否之變數， $Q(\cdot)$ 為分量函數， $F(\cdot)$ 為 y^* 的分配函數，

¹⁰應變數為二元型態，應用在信用風險上其應變數為信用卡為違約與不違約的情況

而真正 y^* 無法藉由觀測 (7) 式而直接獲得。我們需要將此變數轉換成另一種可以觀測到的指標變數 $y_i = 1(y_i^* \geq 0)$ ，而可以透過單調轉換將此設定轉換成：

$$g\{Q_{y_i^*}(\tau | x_i)\} = Q_{g(y_i^*)}(\tau | x_i) \quad (8)$$

而既然 $g(y_i^*) = 1\{y_i^* \geq 0\}$ 也是單調轉換函數，我們可做如此設定：

$$1\{Q_{y_i^*}(\tau | x_i) > 0\} = Q_{1\{y_i^* > 0\}}(\tau | x_i) \equiv Q_{y_i}(\tau | x_i) \quad (9)$$

而根據 (9) 式則(7)式可改寫為：

$$Q_{y_i}(\tau | x_i) = 1\{x_i' \beta(\tau) \geq 0\} \quad (10)$$

(10)式之模型即是 Manski(1975,1985)的 Maximum Score Estimation. 而我們可以透

過極小化下列式求得欲估計之參數：

$$b_N^*(\tau) \in \arg \min_b N^{-1} \sum_{i=1}^N \rho_\tau(y_i - 1\{b' x_i \geq 0\}) \quad (11)$$

此處的 $\rho_\tau(v) \equiv [\tau - 1\{v < 0\}]v$ ，為 Koender 和 Bassett(1978)所提及的勾函數(check

function)，而 Manki(1985)證明了 Maximum Score Estimation 或二元分量迴歸模型

所估計的結果符合了一致性。

2.4.6 信用風險模型效力之驗證與比較-CAP&ROC

在 BASEL II 的內部模型法中，特別要求銀行自行研發的信用風險模型必須要

具有一定的效力。我們在信用模型間的預測效力比較上，使用林公韻(2005)中的

主要兩種方法：CAP 曲線 (Cumulative Accuracy Profiles) 、ROC (Relative or

Receiver Operation Characteristic)。

1. CAPs 曲線 (Cumulative Accuracy Profiles)

這個曲線適用於在信用卡的信用評等模型上，在考慮評分之下，我們可以知道低違約率的持卡戶有較高的評用分數，所有債款人第一次向銀行借款時，銀行會給予持卡人一個信用評等。

模型間的表現好壞，可以直接從 CAP 的圖形上來看。假設共有 (N+D) 個持卡人，其中有 D 為信用卡持卡人之違約戶（即應變數為 1），N 為正常戶（即應變數為 0）。經過模型配適後，每一持卡人都會有個模型所配適出的「違約機率」。

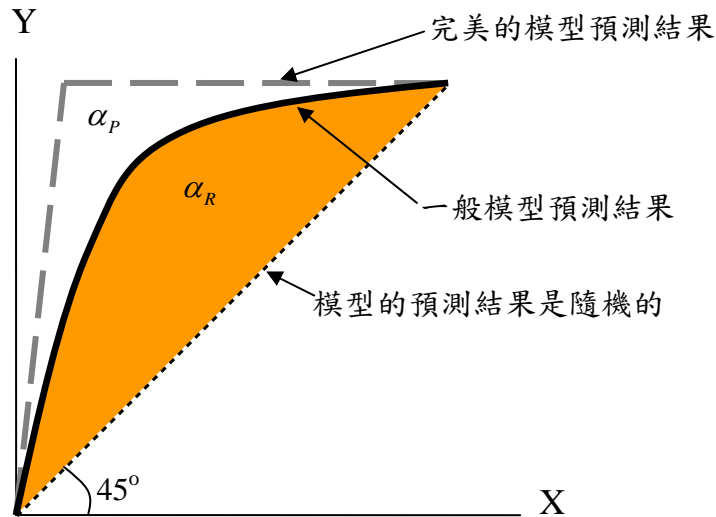
我們先把下表的預測值先由最可能違約排至最前面。

	實際值	預測值	
D 個違約戶	1	0.98	取前 x% (= (N+D)*x%) 的持卡戶，計算這前 x% 持卡戶中，模型所配適出違約機率為 (1-x)% 以上且實際上也是持卡人違約數目佔所有持卡人違約數目之百分比。令其為 y(x%)。以 x 為橫軸，y(x%) 為縱軸，則可繪製出 CAP 曲線。
	1	0.928	
	:	:	
	1	0.7	
N 個正常戶	0	0.4	
	:	:	
	0	0.0001	

舉例：假設現在有 10,000 持卡戶，其中有 1,000 人為違約戶，9,000 人為正常戶則由這例子可知 $N + D = 10,000$ ， $D = 1,000$ ， $N = 9,000$ 。經過模型配適後，這 10,000 持卡戶各有模型配適後的違約機率，將這些違約機率由大排到小之後，取前 5% 家公司（即 $x=5$ ，代表前 500 位的持卡人），發現模型所配適的違約機率為 95% 以上且實際上應變數為 1 的持卡戶（即實際上為違約戶）的數目總計有 80 個，這代表當模型以 95% 為截斷點（Cut Point）時，模型可以偵測出 8%

$\left(\frac{80}{1000} \times 100\% = 8\% \right)$ ，即 $y(5\%) = 8\%$ 的違約公司

【圖 2-6】CAP 曲線



【圖 2-6】中的實線為依模型配適結果所繪製的 CAP 曲線，45 度角的粗點

虛線表示這個模型預測結果幾乎是隨機的，也就是表示這個模型並不具有任何的偵測能力。最理想的模型應該對所有違約的持卡戶都能預測出來。也就是在 x 軸越前端的部分，y 值越大越好，也就是 CAP 曲線越往左上，CAP 曲線下的面積越大，則代表模型的偵測能力越好，【圖 2-6】中的灰色長條虛線表示完美的模型預測結果。

α_p 為完美模型預測結果下的面積， α_R 為一般模型預測結果模型預測結果下的面積

故 $AR = AR = \frac{\alpha_p}{\alpha_R}$ 值愈等於一，其預測信用卡違約風險效率愈高。

2. ROC 曲線 (Relative or Receiver Operation Characteristic)

ROC 曲線的概念與 CAP 曲線概念很像，在討論 ROC 曲線之前，要先定義【表 2-3】中的幾個名詞：

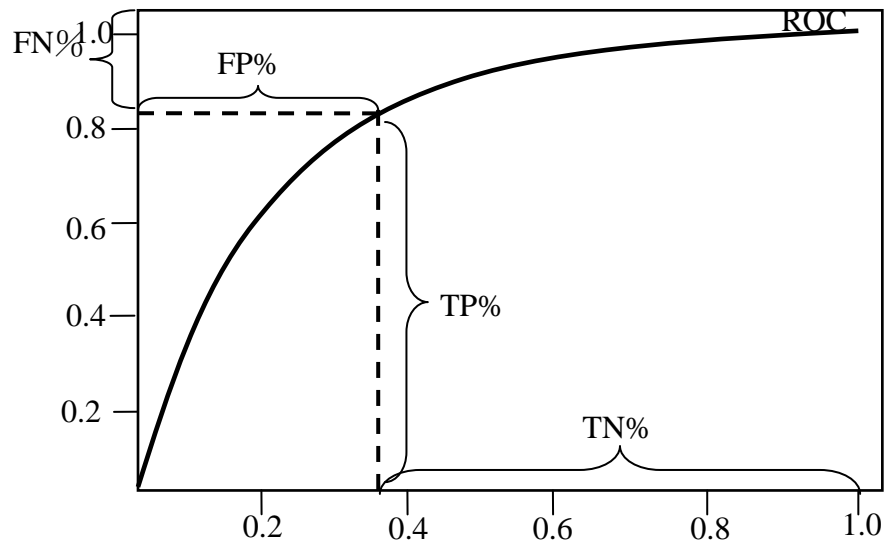
【表 2-3】 實際情況與模型判定交叉分類表

模型判定 \ 實際情況	違約戶 ($y_i = 1$)	正常戶 ($y_i = 0$)
違約戶 $\hat{y}_i = F(x_i \hat{\beta}^*) \geq \text{截斷點}$	TP% \times D	FP% \times N (Type II Error)
正常戶 $\hat{y}_i = F(x_i \hat{\beta}^*) < \text{截斷點}$	FN% \times D (Type I Error)	TN% \times N

依據不同的截斷點，可以得到不同的交叉分類表，而 ROC 曲線就是在繪製 FP% 與 TP% 之間的關係。舉例來說，同樣地，

假設現在有 10,000 持卡戶，其中有 1,000 人為違約戶，9,000 人為正常戶則由這例子可知 $N + D = 10,000$ ， $D = 1,000$ ， $N = 9,000$ 。經過模型配適後，這 10,000 持卡戶各有模型配適後的違約機率，將這些違約機率由大排到小之後，即以 $c\%$ 為截斷點時，在 9,000 正常戶中，有 450 ($9,000 \times 5\% = 450$) 持卡人被判定為違約戶，找到這個截斷點 $c\%$ 之後，再依據這個截斷點，計算 TP%。若在 1000 個違約戶中，以 $c\%$ 為截斷點時，該模型可以捕捉到 80 個違約戶（即觀察值為 1 且估計的違約機率在 $c\%$ 以上的樣本公司），則代表 TP% 為 8% ($80 \div 1000 = 8\%$)。接著進行重複的動作，即對應 10%、15%、...、100% 的型二誤差，找到不同的截斷點 $c\%$ ，再根據不同的截斷點找到不同的 TP₁%、TP₂%、...、TP₂₀%。以型二錯誤為橫軸，即 5%、10%、15%、...、100%，縱軸為 TP₁%、TP₂%、...、TP₂₀%，則可繪製出 ROC 曲線，如【圖 2-7】。

【圖 2-7】 ROC 曲線



由【圖2-7】 ROC曲線來看，在同一個TP%之下，FP%越小，則代表所犯的型二錯誤越小，代表模型的檢驗效力越好，因此，ROC曲線越往左上，ROC曲線下的面積越大，則代表模型越好。