

目錄

1	緒論	1
1.1	信用評分簡史	2
1.2	信用評分概述	2
2	文獻回顧	7
2.1	資料收集與分析	7
2.2	違約定義	11
2.3	變數篩選	11
2.4	信用評分模型	12
2.5	評分模型的驗證	13
2.6	樣本選擇	14
3	信用評分模型的建置	15
3.1	資料收集與分析	15
3.1.1	缺漏值樣本	16
3.1.2	案控樣本	16
3.2	變數序別化	17
3.2.1	ROC 線分析	19
3.2.2	組數的決定與界點的選擇	23
3.2.3	自體抽樣與平滑化	25
3.3	變數篩選	25
3.3.1	排除不穩定的解釋變數	25
3.3.2	變數縮減	26
3.3.3	控制變數	26

3.3.4	篩選統計量	27
3.3.5	篩選方式	27
3.3.6	評等壞率的單調性	28
3.4	Logit 評分模型	29
3.4.1	Logit 模型	29
3.4.2	評分模型的轉換	30
3.4.3	證據權重與 Logit 模型的關係	32
3.4.4	案控樣本對評分模型的影響	32
3.4.5	案控樣本對證據權重的影響	33
3.5	截斷點的選擇	33
4	評分模型的驗證	35
4.1	驗證統計量	35
4.1.1	信息值	35
4.1.2	PSI 值	36
4.1.3	離散統計量	37
4.1.4	HL 值	37
4.1.5	校準度分析	38
4.2	控制變數對模型驗證的影響	39
5	實證	41
6	結論與建議	55

第 1 章

緒論

在競爭激烈的借貸市場中，銀行業者為求迅速做出適當的授信決策必須藉助信用評分機制，於此「信用評分」已逐漸成為銀行業者授信決策的重要依據。信用評分機制固然方便，但當過度依賴時就必須慎選「信用評分模型」（以下簡稱評分模型）才能有效降低信用風險。對於貸放機構而言，若有能力自行建置評分模型，將能掌控信用評分的切確來源並做出適當的授信決策調整。

國內外雖有許多關於評分模型的相關文獻，但對於解釋變數序別化卻很少提及。變數序別化的主要功能是在區分貸款戶的相對優劣，提昇信用評分的辨識能力，進而增加銀行收益與降低信用風險。在本文第 3 章中將會針對解釋變數序別化做精闢的說明。

評分模型為了要能適用於各種經濟環境，必須加入控制變數以控制外在環境，本文認為評分模型一旦加入控制變數後，除了要計算信用評分的驗證統計量以辨識模型的優劣外，還必須將控制變數控制在給定的條件下進行驗證，才能確保信分模型的穩定性與實際運用時的可靠性。此外，由於違約借款戶案例相對於正常繳款的借款戶來得少很多，在違約借款戶過少的情況下所製作的評分將會很不穩定，因此在製作評分模型的過程中會採用「案控樣本」(case-control) 的方式建置評分模型以降低壞率過低影響評分模型之穩定性。使用案控樣本雖可提高樣本的整體違約率，但會造成樣本的信用評分低估。因此一旦使用案控樣本建置評分模型，則必須對評分進行調整。控制變數與案控樣本的問題在本文的 3.4 節將會做詳細的說明。

由於評分模型是建立在借款者未來的違約行為與過去相似的假設之上，難保借款者未來的借款行為不會改變，職是之故必須時常監控 (monitoring) 評分模型的可用性以確保評分模型能有效運作，故於本文的第 4 章有提供一些判別評分模型優劣的驗證方法，並於本文第 5 章實際建置一個預測企業財務危機的評分模型。

1.1 信用評分簡史

在 1930 年代郵購公司為了解決不同授信人員的授信決策不一致以及方便沒有經驗的人做決策，因而將評估信用的準則引入數值評分系統。1941 年 Durand 將 Fisher (1936) 所提出的區別分析方法 (Discriminant Analysis) 應用在區別借款申請者的好壞被喻為信用評分應用之開端。

在 1950 年代，數學家 Bill Fair 及工程師 Earl Isaac 在舊金山 (San Francisco) 成立第一家信用評分顧問公司，該公司將統計的分類技術與信用決策的自動化結合，不過當時的科技不足，故評分模型的發展受限。到了 1960 年代信用卡的出現，為了支應大量的申請案件，授信決策的自動化變得相對重要，隨著科技的快速演進，使得評分模型得以蓬勃發展。現今信用評分已可依據不同的需求，建置各式的評分模型以及相關的運用，諸如帳戶管理 (account management)、催繳欠款 (collection)、損失回收 (recovery)、以及風險定價 (risk-based pricing) 等，

1.2 信用評分概述

「信用貸款」是一種僅需透過借貸雙方對於貸款與償還行為之契約協議，而不需提供物資保證的借貸行為。不過借款者也有可能發生不履行借貸契約的違約行為，因此在授信之前銀行都必須先對借款申請者進行「徵信」，儘可能的避免放款給可能違約之借款者。

徵信即分析借款申請者外顯的性質 去預測其未來會違約的可能性。早期還沒有信用評分的觀念前，授信的決策過程都是以授信人員過去的審核經驗，此法往往會造成授信決策的不一致，因而延伸出信用評分的概念。評分模型的主要功能就是把可能和借款者違約行為相關的性質使用某一準則將其數量化為信用評分，且分數的高低可反映借款者的信用風險高低，銀行便可依此做出適當的授信決策和防範措施以降低風險進而提高經營績效。

不過端看授信人員的經驗所建置的評分模型其可靠信有待商榷，因此便陸續引入一些統計基礎加持，再藉由軟體及演算法的輔助使得評分模型能更為精確，且透過自動化的處理，銀行亦可迅速且準確的完成信用評估，一改過去冗長的徵信程序與減少商機流失。

信用評分模型的建置基礎

評分模型的建置基礎可分法則基礎 (rule-based) 和統計基礎 (statistically-based) 兩大類。法則基礎是使用授信人員過去經驗中認為會影響借款者違約行為的重要解釋變數所建置的評分模型，此類評分模型又被稱做判定基礎 (judgmental-based) 的評分模型。統計基礎則是採多變量分析來決定具統

計顯著性的解釋變數所建置的評分模型。使用統計基礎的評分模型相較於法則基礎的優點即是可以清楚得知影響信用風險的解釋變數及其貢獻性，以減少遺漏重要的解釋變數與避免包含無關緊要的解釋變數。

專屬與非專屬信用評分之比較

信用評分依資料來源可分專屬信用評分 (custom scores) 與非專屬信用評分 (generic scores) 兩大類。專屬信用評分是銀行使用本身的信貸資料所建立，非專屬信用評分則是徵信資料中心使用數家銀行與信貸業者的信貸資料所建立，銀行或借款申請者則可透過付費的方式向徵信資料中心取得信用評分。非專屬信用評分相較於專屬信用評分的優點是：(1) 沒有製作評分的問題；(2) 使用豐富的徵信資料，適用範圍較廣。非專屬信用評分相較於專屬信用評分的缺點是：(1) 由於非專屬信用評分並非針對銀行本身的申貸情況所設計，且徵信資料中心較缺乏借款者的正面訊息，因此專屬信用評分可能較不精確。(2) 不知徵信資料中心的評分細節，因此銀行無法彈性的擬定婉拒理由。(3) 當貸款數多時，非專屬信用評分的取得也是不可小覷的成本。

信用評分的優點與爭議

使用信用評分至少有以下 6 個優點：

1. 提高效率

信用評分可使信貸批准過程更有效率。根據 Mester (1997) 的觀察，過去企業的信貸處理需要花上半天至兩星期，信用評分則可縮減到一個小時以內。

2. 決策一致性

傳統依賴授信人員過去經驗的授信決策者會因不同的授信人員而產生決策不一致的情況，而信用評分則可以保證對每位申請者都是使用相同的標準，以減少主觀的認定過程。

3. 決策的精確性

信用評分相較於人為決定，其允許信貸的決策過程中，考慮到較多關於借款者不履約的因素。

4. 強化風險控管

預測個別借款者或中小型企業貸款的違約概率將有助於風險的控管，一改過去消極的避險行為，轉而追求利益極大。

5. 增加收益

Frame et al. (2001) 使用亞特蘭大聯邦準備銀行 1998 年的美國前 200 大的銀行中找出 99 家銀行的信貸資料, 研究使用信用評分對中小企業貸款的影響, 結果顯示銀行使用信用評分後對中小企業貸款數增加 0.084 倍 (將近 40 億美元)。

6. 降低成本

國內金融業者雖然已有信用評分的概念, 但多半是採委外的方式。評分模型每隔一段時間就必須重新驗證模型的可用性, 信用評分的執行若是委外管理, 需要花費數千美元到 5 萬美元不等, 長期下來將會使得銀行分析成本大幅上升, 若是委外建置評分卡最少也要 1 萬美元起跳。銀行若能雇有自己的專業統計人員, 自行建置專屬評分卡 (custom scores), 不但可立即且有效的維護及更新評分卡, 亦可降低成本。

Padhi et al. (1998) 針對不易取得銀行授信的中低收益型中小企業, 進行信用評分與放款數量之研究顯示: (1) 沒有信用評分機制的銀行, 對中小企業的放款數量明顯較有評分機制者少。(2) 使用信用評分的銀行對中小企業放款的數量與對其他企業的放款數量並無明顯差異。(3) 使用信用評分的銀行, 其分行數的多寡並不影響對中小企業放款的數量。可知信用評分機制對中低收益型中小企業的授信取得是有利的。

由於歐美國家的銀行在放款決策上對於信用評分的漸增依賴, 使得信用評分的取得管道、一致性、成本、以及精確性也引起一些爭議:

- 取得管道

借款申請者目前雖可透過許多管道取得自己的信用評分, 但索取的評分通常迥異於銀行所使用的評分。且不同信用報告機構的評分模型不盡相同, 這使得借款申請者無法比較各機構發佈的信用評分。銀行雖有義務對借款申請者揭露被婉拒的原因, 但通常無法得知評分模型中的變數及其相對重要性, 沒有這些資料, 借款申請者將無從得知諸如結清信用帳戶等金融交易對於其信用評分的影響為何, 除非借款申請者能目睹放款決策所使用的實際評分。

- 一致性與成本

不同信用報告機構所計算的信用評分之間也經常發生不一致的問題。此種不一致性的主要來源在於各信用報告機構自銀行取得的資訊可能不完整, 如遺漏正面的信用資訊。因此若使用單一信用報告機構的評分將有可能使借款申請者處於不利之地位, 故借款申請者往往需要申請多個信用評分使得申請成本增加。

- 精確性

有關信用評分的精確性之方法論的爭議，主要在於建置評分模型所使用的程序。經由全國樣本所建置的評分模型，可能忽略與地方和區域經濟條件相關的重要變數，使用未調整地方與區域經濟狀況之樣本來計算信用評分，可能造成信用評分的正確性。例如對於部份的少數族群，可能因代表性不足而在評分模型建置的選樣過程中遭到剔除。基於此種樣本推導的評分模型，對此一族群的借款申請者可能產生不適當的信用評分。

關於信用報告涵蓋資訊之正確性長久以來備受關注，Neal Walters and Sharon Hermanson (2001) 檢視信用報告的研究發現：在所有被檢視的信用報告中，超過半數以上含有錯誤的資訊。第二項研究發現：在接受調查的信用報告中，有 70% 含有不正確的資訊；其中的 29% 的錯誤足以對借款申請者的信用產生嚴重的不良後果。含有錯誤負面資訊的信用報告，將會降低借款申請者的信用評分，並可能造成其信用被婉拒，或被迫接受較為不利的信用條件。

第 2 章

文獻回顧

評分模型的建置過程大致可粗分為資料的收集和整理、違約意義的界定、變數篩選、評分模型的建置、以及評分模型的驗證等五個步驟，本章將依據這五個步驟以及樣本選擇問題進行相關的文獻回顧。

2.1 資料收集與分析

資料收集需視評分模型的貸款對象而有所不同，主要可分為「消費者貸款」與「企業貸款」兩類。消費者貸款的資料搜集方向主要是以個人的基本資料、財力證明、以及信用往來等記錄做為授信的依據。企業貸款的資料搜集方向則著重於公司營運健全與否，最具代表性的訊息即公司的財務報表。對於財報比率變數的使用，必須特別注意財報的時點問題，Ohlson (1980) 認為使用公司年底的財報來預測公司的破產行為是不合理的。因為財報需會計師簽證，因此財報會有延遲的現象，若是使用當年底的財報形同將已揭露公司破產行為的財報來做預測。此外由於不同規模模型的公司其財務報表的絕對值差異甚大，因此對於企業的財報分析多半是使用財務比率資料。

消費者貸款

陳宗豪 (1999) 參考某商業銀行的分類模式，將消費者小額信用貸款信用要素歸類成基本屬性、償債能力、信用往來、以及其它因素等四大項。其中基本屬性包括繳款情形、性別、年齡、職業、職位、工作年資、住宅所有權狀況、居所、婚姻狀況、以及房地產貸款金額；償債能力包括平均所得；信用往來包括行庫借款種類、信用貸款家數及金額、票性、信用卡張數與金額、信用卡使用額度、申貸金額、核貸金額、貸款期間、以及最近銀行查詢家數；其它因素包括借款用途、連保人、是否願意為家人所知、提供薪資種類、在職證明種類、以及是否為本行主管介紹。

何貴清 (2001) 延用陳宗豪 (1999) 年的四大項歸類方式 (但歸類方式有差異) 並外加財產狀況與

貸款條件兩項共分六大類。其中基本屬性包括性別、年齡、職業、職位、教育程度、工作年資、以及婚姻狀況；償還能力包括負債收入比、平均月所得、年收入、貸款金額、房租或房貸利息支出、以及銀行存款；信用往來包括擁有支票戶數、過去信用付款的記錄、與借款行往來關係、擁有信用卡張數、以及信用卡額度；財產狀況包括住宅所有權或其他不動產和是否為借款行房貸戶及貸款金額；貸款條件包括貸款期間、借款用途、以及保證人數與借款人關係；其他因素包括在公部門或私部門、扶養人口數、是否有人壽保險、最近銀行查詢家數、借款是否願為家人所知、以及現址居住年數。

戴堅 (2004) 將變數分成借款戶屬性、還款財源、借款戶展望、以及聯徵查詢資訊等四大類。其中借款戶屬性包含性別、婚姻狀況、教育程度、以及年齡；還款財源包含年所得、以及負債所得比率；借款戶展望包含服務機構與服務年資；聯徵查詢資訊包含信用卡張數、現金卡張數、是否使用循環利息、以及近三個月是否有他行查詢。

企業貸款

李明峰 (2001) 依財務比率的結構分償債能力、財務結構、經營效能、獲利能力、倍數分析、以及現金流量分析六類。其中償債能力包括流動比率、速動比率、以及營運資金百分比；財務結構包括固定長期適合率、債本比、以及負債比率；經營效能包括應收帳款週轉率、存貨週轉率、總資產週轉率、以及淨值週轉率；獲利能力包括稅前淨利率與淨值純益率；倍數分析包括財務費用佔營收比率與利息保障倍數；現金流量分析包括現金流量率與本期純益加折舊對總負債比率。

林妙宜 (2002) 將財務比率歸類為獲利能力、現金流量、成長率、償債能力、以及經營能力等五項指標。其中獲利能力指標包括資產報酬率、淨值報酬率 (稅後)、淨值報酬率 (常續利益)、營業毛利率、營業利益率、稅前淨利率、稅後淨利率、業外收支率、常續性利益率 (稅後)、每股盈餘、每股淨值、常續性 EPS、每股營業額、每股營業利益、以及每股稅前淨利；現金流量指標包括現金流量比率、每股現金流量、現金流量允當比率、現金再投資比率、現金流量對負債比、現金流量對總資產比、現金流量對銷貨收入比、以及現金流量對稅後淨利比；成長率指標包括營業收入成長率、營業毛利成長率、營業利益成長率、稅前淨利成長率、稅後淨利成長率、經常淨利成長率、總資產成長率、淨值成長率、固定資產報酬成長率、稅後常續性利益成長率、以及盈餘成長率；償債能力指標包括流動比率、速動比率、利息支出率、總負債與總淨值比、負債比率、淨值與總資產比、長期資金適合率、借款依存度、或有負債比率、利息保障倍數、以及現金比率；經營能力指標包括營業利益與實收資本比、稅前純益與實收資本比、存貨及應收帳款與淨值比、總資產週轉率、應收帳款週轉率、平均收帳天數、存貨週轉率、平均售貨天數、固定資產週轉率、淨值週轉率、應付帳款付現天數、淨營業周期、股東權益週轉率、營運資金週轉率、以及流動負債週轉率。

企業信用評估要素

國際上對影響企業信用的要素主要有 5C、6C、3F、6A、以及 5P 等說法。以下部份摘錄自蒲小雷與韓家平 (2001)、石新武 (2002)、以及孫國志與張炎培 (2005)。

5C 的說法早在 1910 年美國銀行家 William Post 於 “*The Fore Big c’s Factor in Extending Credit*” 一文中首先提出「品格」(Character)、「能力」(Capacity)、「資本」(Capital)、「擔保品」(Collateral) 等四項構成企業信用要素引起各方注意，之後銀行家 Edward F.Gee 主張加上「企業環境」(Condition of Business) 合稱 5C 信用要素。其中品格、能力、以及資本為信用評估的「內部因素」，擔保品與環境則為「外部因素」。

- 品格係指企業和管理者在經營活動中的行為和作風，如同人的品性一樣決定著企業信用的好壞。
- 能力係指經營者的管理、資金營運和信用調度的「經營能力」；企業營運、獲利和償債的「企業能力」；以及企業借款的「運用能力」。
- 資本一般係指企業的資本結構、資本安全性、流動性、以及獲利能力等財務狀況，企業資本來源有限或資本架構比例失調將會提高該企業的違約風險。
- 擔保品可降低銀行放款的潛在風險，儘管如此，擔保品並不會改善借款者的信用狀況，因此擔保品的作用僅是促使信用交易的完成而非授信的必要條件，故被列為信用評估的外部因素。
- 環境狀況又稱經濟要素，舉凡可能影響企業經營活動的外在因素都歸為環境狀況，如景氣循環、政經局勢、金融市場等變化。此一要素是企業所無法掌控的，有別於前四個要素。

6C 的說法即是原來的 5C 外加「保險」(Coverage insurance) 要素，凡是牽涉到債權保障方面的業務皆統稱為保險。保險的重要性是在現代商業活動中逐漸體現出來的，其功能與擔保品一樣惟主要差異在保險是透過第三者的保證而取得信用。

3F 的說法是建立在 5C 的基礎上，羅勃特摩里斯協會 (Robert Moris Associates) 將品格與能力合稱為「個人要素」(Personal Factor)，資本與擔保品合稱為「財務要素」(Financial Factor)，企業環境改為「經濟要素」(Economic Factor)。爾後 Milton Drake 又把個人要素改為「管理要素」(Management Factor) 形成目前企業信用 3F 要素。

美國國際復興銀行 (International Bank for Reconstruction and Development, IBRD) 將企業要素歸納為經濟原素 (Economic Aspects)、技術原素 (Technical Aspects)、管理原素 (Managerial

Aspects)、組織原素 (Organizational Aspects)、商業原素 (Commercial Aspects)、以及財務原素 (Financial Aspects) 合稱 6A。

- 經濟原素係指市場經濟大環境是否有利於企業的經營和發展。
- 技術原素係指企業在科技、生產能力、以及獲利能力上是否有利於還款。
- 管理原素係指企業內部各項管理措施是否完善與管理者經營作風和信譽狀況的優劣。
- 組織原素係指企業內部組織與結構是否健全。
- 商業原素係指企業原材料、動力、勞力、設備是否充分、產品銷售市場和價格競爭力等方向是否占有優勢。
- 財務原素係指企業的資金運用、資本結構、償債能力、流動性、以及獲利能力等財務狀況。

5P 是指申請者 (Personal)、借款目的 (Purpose)、還款來源 (Payment)、債權保障 (Protection)、以及授信展望 (Perspective)。

- 申請者係評估借款戶的責任感、經營成效、以及與銀行往來之情形。
- 借款目的可分短期周轉或長期投資，健全的授信業務應重視借款資金用途是否正當，並在貸放後加以追蹤是否依原定之資金運用，企業不當的使用資金，會威脅到銀行的債權。
- 企業能否按期還款取決於信用額度到期時企業即時的財產狀況與還款來源，短期借款的還款資金來源應來自企業短期的經營活動，中長期借款常被用於固定資產的購置，其還款資金應來自企業長期的累積利潤和投資情況。
- 債權保債可分為內部保障和外部保障，內部保障相當於擔保品，需要分析擔保品的合法性、可靠性、市場價格、以及銷售難易度等。外部保障是第三人承擔債務人信用責任的保障形式，採用模式有保證、背書、或第三人提供擔保品等。
- 銀行承作一授信案件之利益不外乎利息收入、手續費、保證費、以及承兌費等收入；或者借款戶所能為銀行增加收入之相關業務，如存款與外匯等之往來，銀行可以藉此評估授信展望。

2.2 違約定義

在界定樣本的好壞前必須先明確定義何謂違約行為，違約定義可視貸款類型與銀行本身對風險的承受度定之，不過多半的情況都是以超過繳款日 30、60 或 90 天做為違約定義，有時則是受限於違約樣本的多寡而去選擇違約定義。

此外在相關文獻中尚有以下定義：(1) 將列為催收款項的借款者視為違約戶，如呂美慧 (1999) 和李明峰 (2001)；(2) 將破產視為違約，如 Altman (1968)；(3) 以債信不足或債權到期無法償還本息者視為違約戶，如 Scott (1981) 及 Bahnson and Bartley (1992)。

違約定義除了好壞的二分法之外，亦有文獻以非二分法的方式定義違約的類型，如 Goldstein (1988) 分管理失敗、財務失敗、以及法律失敗三類。陳肇榮 (1983) 分財務危機、財務失調、以及破產倒閉三類。Hambrick and D veni (1988) 以營運績效分績效不好、持續惡化、損益兩平、以及破產邊緣四個等級。Lau (1987) 將違約狀況分成財務穩定、未發股息或減少股息支付、技術性違約與貸款償還違約、申請破產處分者、以及破產與清算等五種狀況。

台灣經濟新報將下述情況分成財務危機與準財務危機兩類：(1) 財務危機：跳票擠兌、倒閉破產、繼續經營疑慮、紓困、重整、接管、全額下市、財務吃緊停工、以及淨值為負；(2) 準財務危機包括掏空挪用、暫停交易、董事長跳票、銀行緊縮、大虧、景氣不佳停工、以及價值減損。

2.3 變數篩選

在多變量分析中，變數篩選的方法會因估計模型的不同而有所差異，以下是以往文獻所使用的變數篩選方法。

Hand (1997) 將變數篩選方式分成下述三類：(1) 使用專家知識、經驗以及資料在統計操作上的直覺。不過只靠專家知識與經驗有可能包含到對預測沒有幫助的變數，因此還需要一些統計上的佐證才比較可靠。(2) 逐步增加變數，挑選出能增加預測力的變數。不過使用此法虛擬變數很容易會被挑選出。(3) 找出能很好區分違約與未違約分配的變數。

呂美慧 (1999) 與李明峰 (2001) 使用檢定母體平均數差的方式篩選變數，其步驟為先行使用 Kolmogorov-Smirnov 檢定對各變數進行常態性檢定，符合常態的變數使用 t 檢定，不符合的變數則使用 Mann-Whitney U 檢定來篩選變數，並將上述篩選出的變數再使用逐步 logistic 迴歸法 (Step-wise Logistic Regression) 進行第二次的篩選。陳宗豪 (1999) 使用 Kruskal-Wallis 獨立樣本檢定法來檢驗各解釋變數對是否與正常繳款有顯著性差異存在。林建州 (2001) 使用因素分析法 (Factor

analysis) 和逐步 logistic 迴歸法來篩選變數。林妙宜 (2002) 使用母體平均數差檢定來篩選出具有顯著差異的變數。

Mays (2004) 將變數篩選過程分「變數分析」(variable analysis) 與「變數縮減」(variable reduction) 兩個步驟。第一步驟變數分析主要是根據單一解釋變數與應變數進行複迴歸的卡方值、Spearman 等級相關係數、以及信息值 (Information Value, IV) 三個統計量篩選出具統計顯著的變數。第二步驟變數縮減則是使用第一步驟所挑出的變數, 再進行群集分析法 (Cluster analysis), 將同質的變數刪去。

2.4 信用評分模型

多元區別分析 (Multiple Discriminant Analysis, MDA)

Altman (1968) 使用多元區別分析建立 Z-Score 評分模型, 其估計方法是假設模型是線性多元迴歸, 並以極大化組間變異與組內變異的比值來推估各變數的權重值, 分數即是這些權重值乘上解釋變數值的加總。

類神經網路 (Artificial neural network)

Coats and Fant (1993) 以類神經網路的方法製作財務預測模型。類神經網路是一種模擬人腦神經系統所發展出來的資訊處理系統, 其計算方式是將違約與否的應變數做為輸出層, 解釋變數做為輸入層, 再另行假設解釋變數間可能存在數層交互影響的隱藏層, 類神經網路處理系統便會依據樣本值推估各交互作用對輸出層的影響權重。

倒傳遞類神經網路是使用估出的權重值進行模擬, 計算從輸入層到輸出層的預測值後再使用最陡坡降法 (steepest descent) 修正權重值來極小化預測誤差, 不斷地反覆模擬即可推估出具有最小預測誤差的權重值, 再進一步使用各解釋變數與應變數的權重值設計評分方式。

江世傑 (2001) 使用模糊類神經網路估計模型。模糊類神經網路則是架構在類神經網路估計的概念上, 允許事先給定網路系統一些額外的訊息 (規則)。

Logit

Ohlson (1980) 使用 Logistic 迴歸模型進行公司破產預測。Wiginton (1980) 第一個發表 Logistic 迴歸模型與區別分析法應用在評分上的比較。由於 Logit 模型估計亦為本文建置評分模型的一個中

間過程, 故 Logit 評分模型在本文 3.4 節將有詳細的說明。

其它評分模型

除了上述常見的評分模型外, 尚有遞迴分割法 (recursive partitioning) (或決策樹法 (decision tree))、專家系統 (expert systems)、平滑無母數法 (smoothing onparametric methods)、以及時間變動法 (time varying models)。Makowki (1985) 有描述遞迴分割法 (或決策樹法) 在信用評分上的應用。專家系統的細節相對於其他估計模型來得稀少, 可以參閱 Zocco (1985)、Davis (1987)、以及 Leonard (1993b,c)。無母數中以鄰近法最常被應用在信用評分上, 如 Chatterjee and Barcun (1970)、Hand (1986)、以及 Henley and Hand (1996) 等文獻。時間變動法如 Bierman and Hausman (1970)、Dirickx and Wakeman (1976)、以及 Srinivasan and Kim (1987b) 使用 Bayesian 的基礎來提供企業貸款隨著時間經過的違約訊息, 並結合利潤基礎的方法來決定是否同意給予貸款。

近來也有文獻將馬可夫鏈模型 (Markov chain model) 與生存分析法模型 (survival model) 引入評分模型。馬可夫鏈模型可估算預測借款者從一特定狀態轉至另一狀態的轉換概率矩陣, 故可依此轉換概率矩陣去預測借款者的行為, 狀態也可以設定成多個狀態, 如遲繳、違約、及清償等三個狀態。生存分析法模型可用來預測違約事件將發生的時間, 此模型的優點是可以放時間變數 (time-varying variables)。

2.5 評分模型的驗證

建置完評分模型後的首要工作即是驗證評分模型的可用性。辨別模型的優劣可透過驗證統計量判別, Mays (2004) 將驗證模型的統計量歸類成分離統計量 (Separation statistics)、評等統計量 (Ranking Statistics) 與預測誤差統計量 (Prediction Error Statistics) 三類。其中分離統計量有離散統計量 (Divergence Statistics) 與 KS (Kolmogorov-Smirnov) 統計量; 評等統計量即 C 統計量 (又稱 AUC 值); 預測誤差統計量則為 HL 值 (Hosmer-Lemeshow)。

鍾經樊等 (2005) 將驗證統計量分成辨識度、序別相關、信息理論、以及違約配適度四類。其中辨識度包括 AUC 值、AR 值、KS 值、以及 Sym 值; 序別相關包括 Spearman's ρ_s 、Kendall's τ_a 、以及 Kendall's τ_b ; 信息理論包括信息值、離散統計量、條件熵比例、以及條件變異比例; 違約配適度包括 Brier 分數、HL 配適統計量、以及校準度檢定。

國內有一些文獻對於評分模型的驗證, 是以特定截斷點 (cutoff-point) 下的正確預測力做為驗證評分模型的優劣, 例如將 Logit 模型估出的預測違約率大於 0.5 的樣本視為違約樣本其餘視為未違

約樣本,並以評分模型的預測力來驗證模型的好壞。本文必須指出只看單一截斷點下的預測力來驗證評分模型的好壞是會受截斷點的選取所影響。換言之評分模型可經由截斷點的選擇來調整模型的預測力,若直接以單一截斷點下的預測力來驗證模型的優劣實為不妥。截斷點的選擇可以更有意義而不應只是用來控制模型的預測力¹。

2.6 樣本選擇

建置評分模型時若未包含拒絕授信的樣本,則建置模型的樣本分配將會是不完整的,此種資料遭截斷的情況將會產生樣本選擇的問題。針對樣本選擇問題有以下兩種常見的處理方式:(1)先使用現有的建置樣本先建立評分卡並使用此評分卡所建立的截斷點來取得拒絕樣本,再將這些拒絕樣本與原建置樣本重新製作評分卡。(2)另外建立樣本選擇模型進行估計。Ross (2000)指出若拒絕申請者貸款所使用的解釋變數與評分模型的解釋變數一樣,則樣本選擇偏誤的影響有限。Banasik et al. (2001)指出即使考慮了拒絕樣本的問題進行拒絕推論,通常對模型的改善有限。

關於樣本選擇的處理不宜從徵信公司取得拒絕樣本在其他銀行的表現來建置評方模型,使用此法同等於是假設二種貸款所帶來的違約行為是一樣的實屬不合理,此法可行的前提是,除非之後的貸款和先前在別家的貸款是接續下去而且是相同的貸款項目。此外並非所有拒絕樣本皆可在其他銀行找到借款記錄,就執行而言亦是個問題。

¹關於截斷點的選擇請參見本文 3.5 節。

第 3 章

信用評分模型的建置

3.1 資料收集與分析

資料收集的首要步驟就是先確定樣本的收集期間與違約的定義。由於評分模型是使用過去借款者的違約行為去預測未來借款者的違約行為，因此用來建置評分模型的樣本要和未來借款申請者有相當程度的類似性，以保證新舊借款者能有相近的借款行為，故樣本收集期間不宜過長。樣本收集期間也不宜過短，期間若太短將無法反映借款者的違約行為，例如設定樣本收集期間為 18 個月，現若一個為期三年的借款者在這 18 個月後才違約，則在建置樣本中將被視為好客戶，故樣本收集期間至少要足以涵蓋違約事件發生的顛峰時期才不會低估借款者的違約概率。Mays (2004) 指出違約顛峰時期雖會因貸款種類不同而有所差異，但一般違約事件發生的顛峰時期通常為 3 至 4 年。

收集期間一旦確定即可著手收集建置模型所需的樣本與變數。樣本收集的標準做法是產生兩群樣本，分別是建置樣本 (development sample) 與驗證樣本 (validation sample)。前者是實際用來建置評分模型的樣本，後者則是用來驗證模型的可用性。為了要確保評分模型在不同時期仍具有韌性 (robust)，驗證樣本的收集期間應不同於建置樣本，若沒有如此的樣本則可從建置樣本中隨機抽取部份樣本做為驗證樣本。至於變數的收集應儘可能收集與違約行為有關的解釋變數以減少遺漏重要的解釋變數，如收集相關理論之變數，以及授信專業人員經驗上認為重要的變數，理論上遺漏重要的解釋變數對模型不利的影響，較包含到不相關的解釋變數來得嚴重。

資料收集完緊接著就是清除不能用的資料，諸如貸款類型是舊型的未來不會再推出此類型的貸款；借款者是在貸款推出的測試期間所取得的，與未來取得貸款方式不同；違約原因是因為死亡或發生意外事故等樣本應刪除。變數值與理論值不符合的樣本則應視樣本數多寡將不合理樣本刪除或轉為缺漏值。

3.1.1 缺漏值樣本

當樣本所對應的解釋變數中有缺漏值則該筆樣本是無法進行 Logit 模型估計，因此缺漏值若不妥善處理將會使樣本大幅減少，亦會一併降低其他解釋變數所帶來的訊息。

本文將解釋變數缺漏值發生的原因分成「正常缺漏值」與「異常缺漏值」兩種。所謂正常缺漏值是指未來的資料仍會持續發生且合理的情況，例如財務比率變數是由數個會計科目運算而得，因此只要在運算的過程中使得分母為零將會得到缺漏值，或者是有些較細項的會計科目不見得各企業每次都會有記錄，而使得有些樣本的財務比率無法計算；亦或是由於個人隱私問題或者其他原因，得以讓申請者選擇性的提供資料所造成的缺漏值。異常缺漏值則是指在未來的資料收集過程中不太可能發生收集不到的情況。

關於缺漏值樣本所帶來的問題可以採序別化的方式將缺漏值組成一組，並以缺漏值組的證據權重做為缺漏值的序別化值以保留缺漏值樣本所帶來的訊息¹。不過要組成缺漏值組的前提是解釋變數的缺漏值要夠多，否則缺漏值樣本將不具代表性。若缺漏值樣本數過少且刪除缺漏值樣本後並不會使建置樣本的壞樣本大幅減少，則應刪除缺漏值樣本。若刪除缺漏值樣本會使壞樣本大幅減少，可考慮使用「缺漏值填補 (impute) 法」。

缺漏值填補法是先找出與解釋變數高度線性相關的變數，以解釋變數做應變數並與高度線性相關的變數進行線性迴歸，即可以配適值來替代缺漏值。此法雖沒有缺漏值樣本過少的問題，但未來的樣本若有缺漏值將無法評分。此外缺漏值填補法允許缺漏值歸屬到其他各組去，若這些缺漏值推估不當，將會破壞其他組的訊息，因此若未找到高度相關的變數不宜貿然使用缺漏值填補法。本文認為缺漏值填補法的適用時機為：(1) 刪除缺漏值後會造成壞樣本大量流失；(2) 缺漏值樣本數不多且為解釋變數的最高分或最低分組。直覺上缺漏值的問題是因訊息不確定，若給缺漏值組最高分或最低分，似乎就是判定缺漏值全為好樣本或全為壞樣本。當然，若為正常缺漏值樣本且樣本數夠多，僅管缺漏值組是最高分或最低分，獨立成組尚可接受，但若缺漏值樣本數介於中間地帶，則很難令人信服缺漏值所帶來的極端訊息。

3.1.2 案控樣本

因為壞樣本往往相對於好樣本少很多，組壞率將很受壞樣本影響，故在總壞率很低的情形下所製作的評分模型將很不穩定。為解決此問題，在建立評分模型時必須使用「案控 (case-control) 樣本」的方

¹序別化的方式請參閱本文 3.2 節。

式。所謂案控樣本是指實際用於建置模型的好、壞樣本，是從原建置樣本的好、壞樣本中抽取部份比例而得。通常案控的方式是壞樣本全數使用，而好樣本則是從建置樣本的好樣本中抽取部份比例的案控方式以調高全體總壞率。不過使用案控樣本所建置的評分模型其估計壞率將會是給定案控樣本下的條件概率，不同於原意要估計的無條件概率，若不修正評分模型將會造成整體壞率高估，亦即總評分偏低²。

3.2 變數序別化

變數序別化在評分模型中很常使用，如連續變數或序別變數會依其定義域分成數個區間，每個區間製成一組，如年薪分 30 萬以下、30 至 40 萬、40 至 60 萬、60 至 80 萬、以及 80 萬以上；學歷分高中職以下、高中職、專科、大學、以及博士/碩士。類別變數則是將某些類別重新歸類，如工作性質可分運輸、營建、資訊、電子、服務、觀光、百貨、金融、以及投信投顧等，可重新歸類為一般產業、資訊電子業與金融產業等。

變數序別化是評分模型的一個重要過程，不過一般文獻有在探討的並不多，本節將會說明本文認為合理的序別化分界點。由於變數序別化的方式即是對解釋變數進行「證據權重」(Weight Of Evidence, WOE) 轉換，且組數的決定與界點的選擇是以「ROC」(Receiver-Operating-Characteristic) 分析為基礎，因此在決定組數與界點前本文會先對證據權重與 ROC 分析做一詳細說明。

序別化的目的

評分模型的主要目的是在預測借款者的違約概率，因此重點是在抓住違約行為的主要趨勢而不是在配適樣本。變數序別化簡單的說就是將原解釋變數值進行轉換，以一個較具代表性的訊息來取代原解釋變數值所帶來的訊息，且此一訊息的大小可以反映出樣本的優劣。此外解釋變數進行序別化後，使得解釋變數得以依序別化值重新排序，使得序別化後的解釋變數能與應變數呈單調關係，如此便可使得原先與應變數呈非線性關係的解釋變數有機會做為預測變數³，而且使用序別化後的解釋變數製作評分模型可降低原變數值的影響進而提昇模型的辨識度。

可以轉換的理由是評分模型要看的是配適值，亦即變數值與係數估計值的乘積，變數一經轉換後原解釋變數值雖然會改變，但係數估計值也會跟著改變，使得轉換前後的配適值會近似。換言之變數

²關於案控樣本對評分模型的影響與解決方式，在本文的 3.4 節將有詳細說明。

³在迴歸模型裡，一般而言若沒有明顯的線性關係，很難有機會做為預測變數。

序別化的方式若正確，並不會改變模型應有的辨識能力，甚至可以透過序別化來提高模型的辨識力⁴。

序別化的步驟

先決定序別化的組數 J ，依組數來決定分界點 $c_i, i = 1, 2, \dots, J - 1$ 。原變數值若大於 c_{i-1} 小於等於 $c_i, i \neq J - 1$ 屬於第 i 組，大於 c_{J-1} 屬於第 J 組，小於等於 c_1 屬於第 1 組。同一組的樣本以該組的證據權重替代，如此便可將原變數的訊息濃縮成 J 個具代表性的數值。

序別化值 — 證據權重

本文使用證據權重做為序別化值，證據權重的計算公式為：

$$\text{WOE} \equiv \ln \left[\frac{f_0(x)}{f_1(x)} \right], \quad (3.1)$$

其中 $f_0(x)$ 與 $f_1(x)$ 分別為解釋變數 x 在給定好樣本與壞樣本的條件概率密度函數。從式 (3.1) 可知證據權重為 $f_0(x)/f_1(x)$ 的概似比轉換再取對數值。由於實際在進行序別化時都是以間斷型的公式計算，故本節將以間斷型的公式做說明。

證據權重的間斷型公式為：

$$\text{WOE}_j = \ln \left[\frac{g_j/G}{b_j/B} \right], \quad j = 1, 2, \dots, J, \quad (3.2)$$

j 為解釋變數值所對應的組數， g_j 表示第 j 組的好樣本數， b_j 表示第 j 組的壞樣本數， G 和 B 分別表示好樣本和壞樣本總數。我們可以進一步將式 (3.2) 改寫如下：

$$\text{WOE}(p_j|p_o) = \ln \left[\frac{(1-p_j)/(1-p_o)}{p_j/p_o} \right] = \ln \left[\frac{(1-p_j)}{p_j} \right] - \ln \left[\frac{(1-p_o)}{p_o} \right], \quad (3.3)$$

其中 p_o 與 p_j 分別表示建置樣本的總壞率與解釋變數在第 j 組的組壞率。從式 (3.3) 可知，一旦總壞率確定後，證據權重值相當於是在看各組的好壞率比⁵ 與總好壞率比之差，各組的好壞率比愈不同於總好壞率比證據權重的絕對值就愈大。其意涵為序別化後的好壞率比若與總好壞率比的差異愈大，表示對該解釋變數進行序別化有助於區分好樣本與壞樣本。此外若組好壞率比相對於總好壞率比高，

⁴請參見本文第 22 頁「變數概似比轉換對 ROC 線的影響」。

⁵本文為便於解說，將「1-組壞率」與組壞率的比值簡稱「好壞率比」，「1-總壞率」與總壞率的比值簡稱「總好壞率比」，好樣本與壞樣本簡稱「好壞樣本」。

表示該組傾向為好樣本，證據權重值愈大。反之，若組好壞率比相對於總好壞率比低表示該組傾向為壞樣本，證據權重值會愈小。此外，概似比取對數的功能是使好壞樣本在相同的概率密度函數值下有相同的絕對值效果，亦即：

$$f_1(x_1) = f_0(x_2), \quad f_0(x_1) = f_1(x_2)$$

$$\left| \ln \frac{f_1(x_1)}{f_0(x_1)} \right| = \left| \ln \frac{f_1(x_2)}{f_0(x_2)} \right|,$$

若以間斷型的公式而言，則是指好壞率比等於壞好率比時會有相同的絕對值效果。換言之，從證據權重值的正負與絕對值之大小，即可得知各組的好壞程度，因此以證據權重來表示各組的訊息再適合不過。不過當組壞率（或總壞率）為 0 或 1 時證據權重是無法計算的，此時可使用一個接近 0 或 1 的數值去替代。

由於「組好率（或總好率）」等於「1－組壞率（或總壞率）」，故證據權重值的大小亦可解釋成組壞率與總壞率的差異。組壞率愈不同於總壞率，表示序別化有助於區分好壞樣本，且組壞率愈大於總壞率表示該組傾向為壞組，證據權重值愈低。反之，組壞率愈小於總壞率表示該組傾向為好組，證據權重值愈高。

3.2.1 ROC 線分析

在說明「ROC」線前必須先定義 0、1 應變數⁶ (y) 與解釋變數 (x) 的正負相關以及「正確預測率」(True Positive Fractions, TPF) 與「錯誤預測率」(False Positive Fractions, FPF)。

由於應變數為二元變數，因此不應使用一般線性相關係數的方式計算解釋變數與應變數的相關程度，而必須改用 Spearman 等級相關係數。在計算 Spearman 等級相關係數時必須先用平均等級 (average rank) 的方式分別產生應變數與解釋變數的等級變數 (最小值為等級 1)，再用等級變數計算線性相關係數即為 Spearman 等級相關係數。令 r_y 和 r_x 分別為應變數與解釋變數的等級變數，則 Spearman 等級相關係數的計算公式為：

$$\hat{\rho}_s = \frac{\sum_{i=1}^N (r_{yi} - \bar{r}_y)(r_{xi} - \bar{r}_x)}{\sqrt{\sum_{i=1}^N (r_{yi} - \bar{r}_y)^2 \sum_{i=1}^N (r_{xi} - \bar{r}_x)^2}},$$

其中 \bar{r}_y 和 \bar{r}_x 分別為 r_{yi} 與 r_{xi} 的平均數， N 為總樣本數。當 Spearman 等級相關係數值為正，表示隨著解釋變數值愈大應變數愈傾向為壞樣本，本文定義此現象為解釋變數與應變數或壞樣本呈「正相

⁶0、1 應變數即應變數值只有 0 或 1 的二元應變數。

關」。反之，當 Spearman 等級相關係數值為負，隨著解釋變數值愈小應變數愈傾向為好樣本，本文定義則定義為解釋變數與應變數或壞樣本呈「負相關」。⁷

假設解釋變數與壞樣本呈正相關，令解釋變數大於某特定值 \bar{x} 後為壞樣本，則正確預測率為給定真實壞樣本下解釋變數大於 \bar{x} 的概率；錯誤預測率則為給定真實好樣本下解釋變數大於 \bar{x} 的概率。

$$\begin{aligned} \text{TPF}(\bar{x}) &= p(x > \bar{x} | y = 1), \\ \text{FPF}(\bar{x}) &= p(x > \bar{x} | y = 0). \end{aligned} \quad (3.4)$$

反之，假設解釋變數與壞樣本呈負相關，令解釋變數小於等於某特定值 \bar{x} 後為壞樣本，則正確預測率與錯誤預測率為：

$$\begin{aligned} \text{TPF}(\bar{x}) &= p(x \leq \bar{x} | y = 1), \\ \text{FPF}(\bar{x}) &= p(x \leq \bar{x} | y = 0). \end{aligned} \quad (3.5)$$

若將正確預測率擺縱軸錯誤預測率擺橫軸，當 \bar{x} 從負無限大到正無限大，正確預測率與錯誤預測率的連線即為 ROC 線。⁸

令 F_1 與 F_0 分別為壞樣本的累積分配與好樣本的累積分配，則：

$$\begin{aligned} F_1(\bar{x}) &= p(x \leq \bar{x} | y = 1), \\ F_0(\bar{x}) &= p(x \leq \bar{x} | y = 0), \end{aligned} \quad (3.6)$$

可知 ROC 線也可使用條件累積分配來表示。當解釋變數與壞樣本呈正相關， $\text{TPF}(\bar{x}) = 1 - F_1(\bar{x})$ ， $\text{FPF}(\bar{x}) = 1 - F_0(\bar{x})$ 。反之，解釋變數與壞樣本呈負相關， $\text{TPF}(\bar{x}) = F_1(\bar{x})$ ， $\text{FPF}(\bar{x}) = F_0(\bar{x})$ 。由累積分配介於 0、1 之間的特性可知 ROC 線會在邊長為 1 的正方形裡，若以正方形的左下角為原點 (0,0)，則 ROC 線會和原點 (0,0) 與右上角 (1,1) 相連。當解釋變數與壞樣本呈正相關時，依式 (3.4) 的定義可知 ROC 線會從右上角開始出發，先往左方走，再慢慢的往原點連線。反之，若解釋變數與壞樣本負相關，依式 (3.5) 的定義可知 ROC 線會從原點出發，先往上走再慢慢的往右上角。將正方形的左下角 (0,0) 與右上角 (1,1) 連線可以得到一條 45 度線，當解釋變數與壞樣本呈單調關係時 ROC 線皆會在 45 度線以上。ROC 線在 45 度線以上表示隨著錯誤預測率的增加正確預測率總是大於錯誤預測率，且 ROC 線愈接近左上角 (0,1) 表示正確預測率愈大於錯誤預測率，亦即解釋變數辨識好

⁷為方便說明，在本文中解釋變數與應變數的正負相關性一律指 Spearman 等級相關係數值的正負，若為線性相關係數會特別說明。

⁸由於「正確預測率」與「1-錯誤預測率」又可稱為「特質度」(specificity) 與「敏感度」(sensitivity)，所以 ROC 線也可用特質度與敏感度來解釋。

壞樣本的能力愈佳。ROC 線最好的極端情況是從原點 (0,0) 延著正方型的邊長連線至右上角 (1,1)。當 ROC 線延著 45 度線走, 表示無論 \bar{x} 值為何, 正確預測率與錯誤預測率總是很接近, 亦即該變數毫無辨識度可言。

現若不論解釋變數與壞樣本的關係為何, 正確預測力與錯誤預測力皆定義如下:

$$\text{TPF}(\bar{x}) = p(x \leq \bar{x} | y = 1) = F_1(\bar{x}),$$

$$\text{FPF}(\bar{x}) = p(x \leq \bar{x} | y = 0) = F_0(\bar{x}),$$

隨著 \bar{x} 從負無限大到正無限大或者說解釋變數值由小至大累積, 若解釋變數與壞樣本呈負相關, ROC 線會在 45 度線之上; 若解釋變數與壞樣本呈正相關, 則 ROC 線會在 45 度線之下。ROC 線會在 45 度線下是因解釋變數未依壞樣本的關係來變更正確預測率與錯誤預測率的定義, 使得一開始錯誤預測率一直大於正確預測率所造成的現象而非解釋變數不好, 此情況只需將縱軸與橫軸同時反轉 (正方形的右上角為原點 (0,0) 左下角為 (1,1)), 轉軸後的 ROC 線即是變更定義後的 ROC 線。換言之, 即使沒有依據解釋變數與壞樣本的關係來變更預測率的定義也不會影響 ROC 線的分析, 因此只要 ROC 線愈往左上方 (0,1) 或右下方 (1,0), 即表示解釋變數辨識好壞樣本的能力愈佳⁹。

若解釋變數與壞樣本呈二次型式的關係且極端值壞樣本較多的話, 則隨著解釋變數由小至大, ROC 線的走勢會近似 'N' 型, 只差在 ROC 線的中間部份會是往右或右上走; 若極端值好樣本較多的話, 隨著解釋變數由小至大, ROC 線的走勢會近似 'S' 型, 只差在 ROC 線的中間部份會是往上或右上走, 亦即解釋變數若與壞樣本呈二次形式的話, 則會有兩個明顯的大轉折點, 且無論 ROC 線是否穿過 45 度線。

AUC 值

由前述已知 ROC 線愈遠離 45 度線辨識度愈高, 因此使用 ROC 線下的面積「AUC」(Area Under roc Curve) 值比較變數間的辨識度高低是一個不錯的主意。令 $h = F_0(x)$, $v = F_1(x)$, 則 ROC 線與 AUC 的關係為:

$$v = \text{ROC}(h) = F_1[F_0^{-1}(h)] \quad (3.7)$$

$$\text{AUC} \equiv \int_{-\infty}^{\infty} F_1[F_0^{-1}(h)] dh.$$

⁹為方便說明, 本文往後的章節將「正確預測率」定義為壞樣本的累積分配 F_1 , 「錯誤預測率」定義為好樣本的累積分配 F_0 。

不過 AUC 值只適合用來與壞樣本呈單調關係的解釋變數，若是與壞樣本呈非單調關係的解釋變數則須視視 ROC 線的圖形做為最終的決定。此外兩條不同的 ROC 線是可以有相同的 AUC 值時，應以 ROC 線頂點較接近左上方或右下方者辨識力較高。

由前述概念可知，若解釋變數與壞樣本呈正相關則 AUC 值會小於 0.5，且愈小於 0.5 表示 ROC 線愈遠離 45 度線，解釋變數的辨識度愈高。若 ROC 線與壞樣本呈負相關則 AUC 值會大於 0.5，且愈大於 0.5 表示 ROC 線愈遠離 45 度線，解釋變數的辨識度愈高。若與壞樣本呈單調關係的解釋變數且 AUC 值愈接近 0.5，表示該解釋變數愈接近 45 度線，其辨識度愈低。此外，ROC 線在 45 度線以下是因固定預測率的定義所致，只要轉軸即可得到變更定義後的 ROC 線，故「轉軸後的 AUC 值」會等於「1 - 轉軸前的 AUC 值」。值得一提的是，Spearman 等級相關係數與 AUC 值雖沒有完全線性相關，但 Spearman 等級相關係數大於等於零時 AUC 值會小於等於 0.5，反之當 Spearman 等級相關係數係數小於零時 AUC 值會大於 0.5。

變數概似比轉換與 ROC 線的關係

再進一步說明概似比轉換對 ROC 線的影響前必須先知道 ROC 線上的切線斜率與概似比的關係。從式 (3.7) 可以得知 ROC 線的切線斜率為：

$$\frac{dROC(h)}{dh} = \frac{f_1[F_0^{-1}(h)]}{f_0[F_0^{-1}(h)]} = \frac{f_1(x)}{f_0(x)} \quad (3.8)$$

由式 (3.8) 可知 ROC 線的切線斜率為 $f_1(x)/f_0(x)$ 的概似比轉換。換言之，若將解釋變數進行 $f_1(x)/f_0(x)$ 的概似比轉換相當於是在將解釋變數值轉成 ROC 線上的切點斜率值。

$f_1(x)/f_0(x)$ 概似比所代表的意涵為壞樣本與好樣本的概率密度函數值比，若以間斷型公式表示則為：

$$\frac{f_1(x)}{f_0(x)} = \frac{F_1'(x)}{F_0'(x)} = \frac{\Delta F_1(x)/\Delta x}{\Delta F_0(x)/\Delta x} = \frac{\Delta F_1(x)}{\Delta F_0(x)} = \frac{b_j/B}{g_j/G}, \quad (3.9)$$

式 (3.9) 的 b_j 、 g_j 、 B 、 G 的定義與式 (3.2) 同，其意涵為組壞率與組好率比。組壞好率比小於 1，表示第 j 組的好樣本相對於壞樣本發生的機會愈高。反之，組壞好率比大於 1，表示第 j 組的好樣本相對於壞樣本發生的機會愈低。可知 $f_1(x)/f_0(x)$ 概似比轉換的功能是將樣本值依好壞樣本的概率密度函數關係重新定義，且轉換後的解釋變數值大小相當於樣本發生違約行為的可能性大小。故解釋變數經 $f_1(x)/f_0(x)$ 概似比轉換，無論原先解釋變數與壞樣本的關係為何，轉換後皆會與壞樣本呈正相關。換言之，解釋變數若經概似比轉換不但允許評分模型加入原先與壞樣本呈非線性關係的解釋變數，亦可增進模型的辨識力。

這裡值得一提的是,有些解釋變數雖與壞樣本呈二次型式,亦即解釋變數值愈極端,壞樣本會愈多或者是愈少的現象,此類與壞樣本呈二次型式的解釋變數,有些是有助於提昇評分模型的辨識度,若未經概似比轉換就將這些解釋變數直接放入 Logit 模型進行變數篩選,是不太可能被選入評分模型的,即使其極端值能很好的區分壞樣本或好樣本。

若解釋變數與壞樣本的關係呈二次型式以上本文則不建議引入評分模型,此類解釋變數若加入評分模型將會使得解釋變數與評分關係忽高忽低,很難令人接受。

證據權重與 ROC 線的關係

重新檢視證據權重可以發現證據權重也是 $f_1(x)/f_0(x)$ 的概似比轉換,只差在證據權重經概似比轉換後,還須做對數轉換並乘上一個負號。

$$\text{WOE} = -\ln \left[\frac{f_1(x)}{f_0(x)} \right],$$

又對數轉換為一單調遞增轉換¹⁰,乘上負號只會使證據權重與壞樣本呈負相關,故解釋變數做證據權重轉換,除了可用來反映各組的訊息外,還可提昇模型的整體辨識度。

3.2.2 組數的決定與界點的選擇

從前一小節裡可以得知一個重要的訊息,那就是在 ROC 線上具有相同斜率的樣本表示這些樣本的組好壞率比相似,亦即具有相同的違約行為。然而序別化的目的也是再將相似的樣本歸為一組,並以一個具代表性的數值取代。換言之,一個好的序別化方式應以配適解釋變數的 ROC 線為目標,將具有相同 ROC 線斜率的樣本歸為一組。

連續變數與序別變數

序別化組數的決定必須視解釋變數本身能夠將樣本分成幾群,亦即 ROC 線上有幾種明顯的斜率就可分成幾組,分界點即 ROC 線斜率改變時的轉折點。不過實際再畫 ROC 線時若解釋變數值是連續型的,其 ROC 線是階梯型式¹¹,雖可看出轉折點大致座落的區間,但使 ROC 線斜率轉折最大點並非都很明確,因此只能藉助「資料採礦」(data mining)的技術搜尋使 ROC 線斜率轉折最大的點。

¹⁰在第 ?? 頁已說明概似比取對數的功能,故在此不再贅述。

¹¹若好、壞樣本有相同的解釋變數值時,則當解釋變數為此樣本值時,ROC 線上會以斜線的方式呈現。

依 ROC 線的斜率進行序別化後，還必須考慮將某些組併掉。解釋變數與壞樣本呈單調關係時 ROC 線的斜率變化應呈同向遞增或遞減，若某些線段的斜率很不自然的夾在某些組前後，應考慮將其歸到組好壞率比相近的鄰近組去。若不將其併組，此種情況將會使得解釋變數與壞樣本的關係呈現「假二次型式」的現象，造成解釋變數錯誤評分。若解釋變數值與壞樣本呈二次型式，即解釋變數值愈極端壞樣本愈多，則 ROC 線斜率應先遞減後遞增；反之，若解釋變數值愈極端壞樣本愈少，則 ROC 線斜率應先遞增後遞減。這裡值得注意的是，若解釋變數與壞樣本是呈二次型式，卻硬是將其併成單調型式，也會造成解釋變數錯誤評分。解釋變數與壞樣本的關係若超過二次型式，則應考慮放棄該解釋變數。

序別變數的序別化方式與連續變數一樣，其主要差異在序別變數為間斷型且「單一值」(unique value) 個數少¹²，因此斜率轉折點會比連續變數來得較為明顯。

在進行序別化時必須兼顧變數的經濟意涵，若一個變數理論上應與壞樣本呈單調關係，但序別化後卻明顯不同，造成此種現象的可能原因為：(1) 資料處理錯誤。此種情形重新檢視資料以及定義的正確性即可解決。(2) 壞樣本過少。當壞樣本過少組壞率將很受壞樣本影響，此種情況應考慮併組。若非上述原因則應考慮將該變數捨去，否則必須交待清楚變數與理論不符的原因。

此外序別變數由於單一值少，故評分模型建置者常有一些主觀的序別化方式，如年齡在 20 至 65 歲間常以每間隔 5 歲分一組的方式進行序別化。連續變數也會有相同的問題，如本文實證資料中的借款依存度為 0 的樣本有 260 筆，足以自成一組，且直覺上有借款和沒借款的人性質應該有所差異。本文必須指出序別化的方式不應過於主觀，而應以資料本身所呈現的 ROC 線來進行序別化會較為適當。從本文的表 3 中可以清楚看出，借款依存度的第一個分界點為 20.32，亦即借款依存度在 20.32 以下所帶來的訊息是相似，因此實沒必要在借款依存度為 0 時去分一組，若任意進行序別化將會把不同違約特性的樣本組成一組，而降低模型的辨識力。

類別變數

類別變數與序別變數雖同為間斷型變數，但類別變數值只是代號並無大小區分。因此類別變數不能使用 ROC 線分析方式，只能依各類別各自組成一組，並使用類似 ROC 線分析的處理方式將組壞率相近的類別併成一組。

¹²同一解釋變數中不同樣本值的個數。

3.2.3 自體抽樣與平滑化

序別化的目的就是希望評分模型能抓住主要趨勢不要太受樣本值的影響，因此本文也不希望序別化太受抽樣影響。爲了不讓組界點太受單一建置樣本的影響，故本文建議使用「自體抽樣」(Bootstrap)，來取得一個較穩健的界點。自體抽樣即是將建置樣本視爲母體，使用「抽出放回」(with-replace) 的抽樣方式即可得出另一組建置樣本，並依此組建置樣本重新找各解釋變數的分界點，如此反覆進行自體抽樣，即可求得各分界點的分配，便可依自體抽樣後的平均數做爲最後的分界點。

此外由於壞樣本通常很少，若建置樣本的好壞樣本比差異大時，組壞率就會很受壞樣本的影響，進而影響證據權重。除了可使用案控樣本的方式處理外，亦可使用 Lowess¹³ 對組壞率進行平滑化以降低相鄰兩組的壞率差異過於極端。由於平滑化會改變各組所帶來的訊息，因此相鄰兩組壞率若無極端差異時不宜任意使用 Lowess 平滑法。此外本文認爲缺漏值組的序別化值所帶來的訊息不同於其他組別，故不建議缺漏值組與其他序別化值一起進行平滑化，否則缺漏值組的訊息將會改變鄰近組所帶來的訊息。

3.3 變數篩選

變數篩選的目的是從衆多的候選解釋變數¹⁴ 找出能幫助評分模型辨識好壞樣本的解釋變數。至於要如何從衆多的候選解釋變數中挑選，則需透過一些判斷的準則與統計量的輔助才得以執行，本節將會提供變數篩選的方法與步驟。

3.3.1 排除不穩定的解釋變數

篩選的首要步驟就是排除「不穩定」的候選解釋變數，這裡所謂的不穩定是指解釋變數所反映的好壞樣本的分配會隨著時間變動而有明顯的改變。¹⁵ 解釋變數若不穩定，即使在樣本內的預測力很高，但對樣本外的預測可能毫無幫助，甚至造成錯誤預測。不過要執行此篩選步驟，驗證樣本的收集期間必須不同於與建置樣本才能看出解釋變數是否具穩定性。

變數的穩定性可以看候選解釋變數在建置及驗證樣本中好壞樣本的概率密度函數，由於解釋變數

¹³Lowess 的計算方式請參閱 Royston, P. (1991)。

¹⁴候選解釋變數係指尚未選入評分模型前的待選變數。

¹⁵一般而言，解釋變數多少都會隨時間改變，只要解釋變數所反映的好、壞樣本分配不要差異過大即可接收。

中好壞樣本的真實概率密度函數是不知道，因此只能使用核密度估計法 (kernel density estimation) 去推估其概率密度函數，故初步篩選只能排除掉概率密度函數明顯不同的變數。這裡值得一提的是，概率密度函數位移的候選解釋變數不可以排除，因為概率密度函數位移可能是因為外在環境所影響，使得候選解釋變數值整體偏高或偏低，一旦加入控制變數後便可補捉此種效果。除了看概率密度函數外，也可以看好壞樣本在建置與驗證的累積分配，如 ROC 線與 AUC 值，若解釋變數在建置與驗證樣本的 ROC 圖與 AUC 值很明顯不同也應排除。

3.3.2 變數縮減

變數縮減的目的就是將候選解釋變數依線性相關係數進行分群，將線性相關程度高達 0.9 以上的候選解釋變數歸成同一群，屬於同一群的變數只需找出較具辨識力與顯著性的變數投入篩選程序即可。

高度線性相關的候選解釋變數因彼此提供的訊息相近，若一起放入篩選程序中，往往會造成某些解釋變數變得毫無貢獻甚或是係數符號不正確等問題。因此若將一堆訊息極為相似的候選解釋變數放入篩選程序，只會徒增篩選成本以及混淆變數所提供的訊息。

這裡是看線性相關而不是 Spearman 等級相關的原因是，一般而言變數間彼此若有高度線性相關亦會存在高度等級相關，序別化之後，各組的樣本數與壞率皆相似，故序別化後仍會有高度的相關性。反之，只具有高度等級相關的變數則不一定會有高度線性線相關，因此序別化後的相關性則有高有低。

將解釋變數分群後，同一群的候選解釋變數則依穩定度¹⁶、AUC 值、以及單變量 Logit 模型的 t 統計量，挑出最好的候選預測變數做代表即可。其餘無法分類的候選解釋變數則全數保留至篩選程序中。

3.3.3 控制變數

在進行經過初步排除不穩定的解釋變數以及變數縮減後，接下來才算是真的在篩選出能做為評分模型的解釋變數，不過影響借款者違約與否的因素除了借款者本身可控制的內在因素外，還包括非借款者所能掌控的外在因素，如景氣好壞與否。評分模型為了要能適用於各種經濟環境，必須加入控制變數以控制外在環境才能確保評分模型實際運用時的穩定性與可靠性。因此在篩選預測變數之前必須先決定控制變數才能挑出真正對評分模型有貢獻的解釋變數。

¹⁶這裡指的穩定度係指序別化後，建置樣本與驗證樣本的 PSI 值。PSI 值請參見第 36 頁。

一般而言控制變數都是使用總體變數,如經濟成長率、通貨膨脹率、利率等。不過總體變數通常都有高度線性相關的現象,因此也不宜過多。使用的控制變數也應具有一定的顯著性,否則如同沒有控制變數。

3.3.4 篩選統計量

篩選變數時必須要先設定篩選的依據,由於評分模型著重的是在預測好壞樣本,因此本文建議以解釋變數的辨識度高低做為篩選變數的依據,最具代表性的統計量即 AUC 值,其他用以判別辨識度高低的統計量多半都與 AUC 有關,諸如 AR 值、KS 值與 Sym 值等¹⁷。此外也可以看解釋變數在模型中的貢獻度,貢獻度低的解釋變數對評分無太大助益可考慮將其刪除。評分的貢獻度有以下兩種計算方式:

$$\text{貢獻度1} = \frac{1}{N} \sum_{i=1}^N \frac{|\Delta s_{ki}|}{\sum_{k=1}^K \sum_{i=1}^N |\Delta s_{ki}| / N} = \frac{\sum_{i=1}^N |\Delta s_{ki}|}{\sum_{k=1}^K \sum_{i=1}^N |\Delta s_{ki}|},$$

$$\text{貢獻度2} = \frac{1}{N} \sum_{i=1}^N \frac{|\Delta s_{ki}|}{\sum_{k=1}^K |\Delta s_{ki}|},$$

其中 $|\Delta s_{ki}| = |s_{ki} - \bar{s}_k|$ 為第 i 個樣本在第 k 個解釋變數的評分離均差絕對值。兩貢獻度主要差異在前者是在比較解釋變數的整體評分表現,後者則是在比較解釋變數對各樣本的評分表現。

3.3.5 篩選方式

評分模型為多變量模型,不宜以單變量的篩選統計量高低做為篩選變數方式,因為有些解釋變數雖然在單變量評分模型中整體辨識力不高,但對某些區間的辨識度很高,使得在多變量的情況下可以彌補整體辨識力較強的解釋變數中相對較弱的區間,進而提昇評分模型的整體辨識度。若是採用單變量的方式篩選變數,將有可能刪除此類必須搭配相對較強的解釋變數才能顯示其效果的變數。

多變量的篩選方式主要有「正向選入法」(forward selection)與「反向淘汰法」(backward elimination)。不過本文的篩選基準為 AUC 值,與一般以 p 值為主的篩選方式有所出入。本文的正向選入法是要選入模型的解釋變數,必須符合以下條件:

1. 係數符號要正確

¹⁷請參閱鍾經樊等 (2005)。

若解釋變數相互條件獨立，則經證據權重轉換的解釋變數係數值會是負數¹⁸。若係數值為正表示解釋變數間存在高度相關性，則應考慮刪除係數為正的解釋變數，或者是刪除與該解釋變數呈高度相關的解釋變數。若一個解釋變數使用數個線性相關的變數取代，就評分模型的穩定性而言，應使用多個具有相關性的解釋變數可降低單一解釋變數不穩定對評分模型所造成的不利影響。故就解釋變數係數符號的正確性而言只要係數不為正即可接受。

2. 必須具有一定的 t 統計量

評分的貢獻度與 t 統計量雖沒有正比關係，不過 t 統計量絕對值很小的解釋變數其評分貢獻度通常也不高，且評分的貢獻度會受解釋變數的個數影響，t 統計量則較無此問題。故使用正向選入法可以看 t 統計量絕對值是否有超過一特定值，本文是以 1 做為門檻。

3. 能使總評分的 AUC 值相對提高較多且相對較穩定者

最後選入模型的解釋變數則是從符合上述兩個條件的候選解釋變數中，能使總評分的 AUC 值相對提高較多者。此外應以變數穩定性為輔，若一味的挑選使評分模型的 AUC 值最大的解釋變數進入模型中而不顧及變數穩定性，一旦不慎挑選到不穩定的解釋變數將會大幅消滅評分模型的預測力。在變篩選過程中的穩定性則是指解釋變數在建置樣本裡的 AUC 值排名，與驗證樣本裡的 AUC 值排名不可相差過大。

反向淘汰法與正向選入法的差異是先把候選解釋變數全數放入模型中，再依序從係數符號為正、t 統計量絕對值小於 1 的候選解釋變數中，排除使模型 AUC 值下降最小且最不穩定者。

本文建議以正向選入法來篩選變數，正向選入法除了能確實掌控選入模型的解釋變數具有相當的辨識度外，也可再加入變數後從係數的估計值中得知解釋變數間的相關性。若使用反向淘汰法，最後選入模型中的變數只是因為刪除較無辨識度後所留下來的變數，對於解釋變數間的複雜關係全然毫無概念。

3.3.6 評等壞率的單調性

各評等對應的「評等壞率」(同一評等的樣本預測壞率平均值) 單調性應做為篩選變數的最後關卡，若評等壞率不呈單調性則必須放寬篩選條件，選入使評等呈單調性的預測變數。評等若不呈單調表示較低高等級的評等壞率比較低等級的評等壞率來得高，故此種評等將無法用來判別樣本與壞率的關係。

¹⁸在本文第 32 頁證據權重與 Logit 模型的關係將會說明。

3.4 Logit 評分模型

在建置評分模型的衆多方法中, 本文選用 Logit 模型製作評分模型的主因是 Logit 模型在評分模型的應用上有許多便利的操作方式與解釋。Logit 模型估計只是建置評分模型的一個中間過程, 不應將評分模型視為一般的 Logit 模型, 因為評分模型著重的是辨識好壞樣本, 且估計前必須先對解釋變數進行轉換, 估計後還必須對配適值做評分轉換, 與原迴歸模型著重於迴歸係數的估計與解釋是截然不同。

3.4.1 Logit 模型

假設 y 是 0、1 變數, 1 表示壞樣本, 0 表示好樣本, 故 y 的概率密度函數 $f(y)$ 會呈柏努利 (Bernoulli) 分配的形式:

$$y \sim f(y; p) = p^y(1 - p)^{1-y},$$

其中 p 為壞率。依據柏努利分配的定義可知 y 的期望值即壞率, 變異數則為壞率與好率的乘積。有變數的概率密度函數即可使用最大概似估計法 (Maximum Likelihood Estimation, MLE) 估計壞率。現若收集 N 個關於 y 的樣本, 這些樣本將呈 N 元的聯合柏努利分配, 概率密度函數將變得不易求得, 因此必須先對樣本做一些合理的簡化假設才有辦法寫出概似函數並使用最大概似估計法估計壞率。一般皆假設為樣本間彼此獨立表示樣本間的違約行為不會互相影響, 且假設樣本雖可以有不同的壞率但背後影響壞率的解釋變數相同, 亦即

$$E(y_i) = p_i = \mathbf{x}_i' \boldsymbol{\beta},$$

其中 \mathbf{x} 即影響違約行為的解釋變數, $\boldsymbol{\beta}$ 表示各解釋變數的係數值。不過如此假設無法保證壞率介於 0、1 之間, 因此一般都會假設成:

$$E(y_i) = p_i = F(\mathbf{x}_i' \boldsymbol{\beta}), \quad (3.10)$$

其中 F 是一個值域介於 0、1 之間的單調函數, 如此便能確保壞率介於 0、1 之間。若有加控制變數 \mathbf{z} , 則式 (3.10) 的壞率設定應改寫成:

$$p_i = F(\mathbf{z}_i' \boldsymbol{\gamma} + \mathbf{x}_i' \boldsymbol{\beta}),$$

其中 $\boldsymbol{\gamma}$ 為控制變數的係數值。上述轉換過程中, 若 F 是使用常態分配的累積分配函數稱為 Probit 模型, 若是使用 Logistic 分配的累積分配函數即為 Logit 模型。

有了上述假設後, 則 N 元的聯合伯努利概率密度函數便可簡化成:

$$\begin{aligned} f(y_1, y_2, \dots, y_N; p_1, p_2, \dots, p_N) &= \prod_{i=1}^N f_i(y_i; p_i) \\ &= \prod_{y_i=1} p(y_i = 1) \prod_{y_i=0} p(y_i = 0), \end{aligned}$$

令 $p_i = \exp(\mathbf{x}'_i \boldsymbol{\beta}) / (1 + \exp(\mathbf{x}'_i \boldsymbol{\beta}))$ 便可進一步使用最大概似法估計 Logit 模型的 $\boldsymbol{\beta}$ 值, 其目標函式如下:

$$L(\boldsymbol{\beta}) = \prod_{y_i=1} p(y_i = 1 | \mathbf{x}_i; \boldsymbol{\beta}) \prod_{y_i=0} p(y_i = 0 | \mathbf{x}_i; \boldsymbol{\beta}), \quad (3.11)$$

3.4.2 評分模型的轉換

當 Logit 模型估計完後, 接下來就是要將 Logit 模型轉成評分模型。Logit 模型可以線性迴歸的型式表示成:

$$\ln \frac{\hat{p}_i}{1 - \hat{p}_i} = \hat{\gamma}_1 \bar{z}_1 + \dots + \hat{\gamma}_M \bar{z}_M + \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_K x_{Ki}, \quad (3.12)$$

上式 Logit 模型的配適值本身就可視為一種評分, 但一般為了較能符合直覺上的解釋都會讓總評分愈高表示違約率愈低, 因此都會將 (3.12) 式取負號後才將其視為總評分。令轉換前的分數為:

$$\begin{aligned} \tilde{s}_i &= \ln \left(\frac{1 - \hat{p}_i}{\hat{p}_i} \right) = -(\hat{\gamma}_1 \bar{z}_1 + \dots + \hat{\gamma}_M \bar{z}_M) - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_K x_{Ki}) \\ &= \tilde{s}_0 + \tilde{s}_c + \sum_{k=1}^K \tilde{s}_{ki}, \end{aligned} \quad (3.13)$$

其中 $\tilde{s}_0 = -\hat{\beta}_0$ 為調整前底分, $\tilde{s}_c = -(\hat{\gamma}_1 \bar{z}_1 + \dots + \hat{\gamma}_M \bar{z}_M)$ 為調整前控制分, 其中 $\bar{z}_m, m = 1, \dots, M$, 是指第 m 個控制變數的控制值。 $\tilde{s}_{ki} = -\hat{\beta}_k x_{ki}, k = 1, \dots, K$, 為第 i 個樣本在第 k 個解釋變數的調整前評分, 由式 (3.13) 中可以注意到控制變數都是給定值不會隨樣本改變, 因為控制值必須是事先給定的, 本文中稱給定控制值的總評分為「基準分」, 若控制變數是使用實際值則稱為「配適分」¹⁹, 由式中可知控制分的功能是在提高或降低全體樣本的底分。

上述評分是直接使用 Logit 模型的配適值, 這些數值是樣本的預測好壞率比的對數值, 可想而知一定是有正有負, 且其絕對值也都不會太大。從這些評分的相對高低, 只能看出壞率的相對大小, 並

¹⁹在稍後 4.2 節還會有一種「調整分」。

無法提供額外的訊息。因此可考慮轉換總評分，使得轉換後的總評分相對大小較有意義，且不改變樣本原始分數的相對大小。總評分的轉換公式為：

$$\begin{aligned}
 S_i &= A + B\tilde{S}_i & (3.14) \\
 &= A + B\tilde{s}_0 + B\tilde{s}_c + B \sum_{k=1}^K \tilde{s}_{ki} \\
 &= s_0 + s_c + \sum_{k=1}^K s_{ki},
 \end{aligned}$$

其中 $s_0 = A + B\tilde{s}_0$, $s_c = B\tilde{s}_c$ 分別為調整後的底分與控制分, s_{ki} 則為第 i 個樣本在第 k 個解釋變數的調整後評分。

依據 A 、 B 的定義不同，使得評分有不同的轉換意義。對於 A 、 B 兩個轉換值的設定主要可分兩種型式：(1) 給定一預測壞率對應的分數，與設定分數的變化方式。(2) 是限制分數的最高分與最低分。第一種型式有以下兩種設定方法：

$$A = \bar{S} - B \cdot \ln(\bar{r}), \quad B = \Delta s / \ln \alpha; \quad (3.15)$$

$$A = \bar{S} - B \cdot \ln[(1 - \bar{p})/\bar{p}], \quad B = \Delta s / \ln \alpha. \quad (3.16)$$

已知 $\tilde{S}_i = \ln[(1 - \hat{p})/\hat{p}]$ ，若令 $B = \Delta s / \ln(\alpha)$ ，當險算比 $(1 - \hat{p})/\hat{p}$ 等於 α 時總評分會增加 Δs ，因此險算比每擴增 α 倍時總評分都將增加 Δs 。至於 A 的設定，(3.15) 式表示當險算比為 \bar{r} 時總評分等於 \bar{S} ，(3.16) 式表示當壞率為 \bar{p} 時總評分等於 \bar{S} 。第二種型式也有兩種設法如下：

$$A = S_{min} - B\tilde{S}_{min}, \quad B = \frac{(S_{max} - S_{min})}{(\tilde{S}_{max} - \tilde{S}_{min})}; \quad (3.17)$$

$$\begin{aligned}
 A &= S_{min} - B \ln\left(\frac{1 - p_{min}}{p_{min}}\right), \\
 B &= \frac{(S_{max} - S_{min})}{\ln[(1 - p_{max})/p_{max}] - \ln[(1 - p_{min})/p_{min}]}.
 \end{aligned} \quad (3.18)$$

(3.17) 式是設定轉換前的最低總評分 \tilde{S}_{min} 與最高總評分 \tilde{S}_{max} 轉換後要是最低分 S_{min} 與最高分 S_{max} 。(3.18) 式則是去限定預測壞率分別為 p_{max} 與 p_{min} 時總評分要是 S_{max} 與 S_{min} 。(3.17) 與 (3.18) 式主要差異在前者可把分數限制在 S_{min} 與 S_{max} 之間，後者的設定等同於是只要預測壞率低於 p_{max} 即認定是好客戶，預測壞率高於 p_{min} 即認定是壞客戶。

3.4.3 證據權重與 Logit 模型的關係

由第 23 頁的證據權重與 ROC 線的關係中已知證據權重與壞樣本呈負相關，且由式 (3.3) 得知證據權重可用組壞率與總壞率表示，組壞率愈高證據權重值愈小，可知證據權重與組壞率呈負相關。因此證據權重愈小組壞率愈高的組別，我們可以預期該組的樣本其預測違約率也會相對較高。從 Logit 模型可以用 (3.12) 式表示可知解釋變數若為證據權重，則係數值會是負的。

3.4.4 案控樣本對評分模型的影響

建置評分模型時若是採用案控樣本而非使用全部樣本時，樣本的壞率估計值將不同於無條件下所求得的壞率估計值，因此必須修正模型才能得到正確的估計值。案控樣本是指實際建置模型時並非使用收集期間裡的所有樣本，而是只從所收集的樣本中隨機抽出部分樣本進行建模，令 $\delta_i = 1$ 為被選用的樣本， $\delta_i = 0$ 則否，由於案控樣本存在另一隨機變數 δ ，因此式 (3.11) 的目標函數應改為：

$$L(\boldsymbol{\beta}) = \prod_{y_i=1} p(y_i = 1|\delta_i = 1, \mathbf{x}_i) \prod_{y_i=0} p(y_i = 0|\delta_i = 1, \mathbf{x}_i)。$$

換言之，壞率估計值不再是 $\hat{p}(y_i = 1|\mathbf{x}_i)$ 而是 $\hat{p}(y_i = 1|\delta_i = 1, \mathbf{x}_i)$ ，不過由貝氏定理可推得 $p(y_i = 1|\delta_i = 1, \mathbf{x}_i)$ 與 $p(y_i = 1|\mathbf{x}_i)$ 的關係如下：

$$\frac{p(y_i = 1|\mathbf{x}_i)p(\delta_i = 1|y_i = 1, \mathbf{x}_i)}{p(y_i = 0|\mathbf{x}_i)p(\delta_i = 1|y_i = 0, \mathbf{x}_i) + p(y_i = 1|\mathbf{x}_i)p(\delta_i = 1|y_i = 1, \mathbf{x}_i)}。 \quad (3.19)$$

令 $q_1 = p(\delta_i = 1|y_i = 1, \mathbf{x}_i)$ ， $q_0 = p(\delta_i = 1|y_i = 0, \mathbf{x}_i)$ ，由於案控樣本是隨機抽樣，因此可以進一步假設 δ_i 和 \mathbf{x}_i 獨立，即 $q_1 = p(\delta_i = 1|y_i = 1)$ $q_0 = p(\delta_i = 1|y_i = 0)$ 。將 q_0 、 q_1 代入 (3.19) 式即可推得 (3.20) 式：

$$\frac{p(y_i = 1|\mathbf{x}_i)}{1 - p(y_i = 1|\mathbf{x}_i)} = \left(\frac{q_0}{q_1}\right) \frac{p(y_i = 1|\delta_i = 1, \mathbf{x}_i)}{1 - p(y_i = 1|\delta_i = 1, \mathbf{x}_i)}。 \quad (3.20)$$

令 $S_i^* = \ln [1 - \hat{p}(y_i = 1|\delta_i = 1, \mathbf{x}_i)] / \hat{p}(y_i = 1|\delta_i = 1, \mathbf{x}_i) = -\mathbf{x}_i' \hat{\boldsymbol{\beta}}$ ，由 (3.20) 式可知：

$$\begin{aligned} \tilde{S}_i &= \ln \frac{1 - \hat{p}(y_i = 1|\mathbf{x}_i)}{\hat{p}(y_i = 1|\mathbf{x}_i)} = \ln \left(\frac{\hat{q}_1}{\hat{q}_0}\right) + \ln \frac{1 - \hat{p}(y_i = 1|\delta_i = 1, \mathbf{x}_i)}{\hat{p}(y_i = 1|\delta_i = 1, \mathbf{x}_i)} \\ &= \ln \left(\frac{\hat{q}_1}{\hat{q}_0}\right) + S_i^*, \end{aligned} \quad (3.21)$$

其中 \hat{q}_0 為從原建置樣本中抽出的好樣本佔原建置樣本中好樣本的實際比例， \hat{q}_1 為從原建置樣本中抽出的壞樣本佔原建置樣本中壞樣本的實際比例。從 (3.21) 式可知若使用 Logit 模型建置評分模型，

只需加上 $\ln(\hat{q}_1/\hat{q}_0)$ 調整即可解決案控樣本對評分模型的影響。一般而言 \hat{q}_1 都會大於 \hat{q}_0 ，若不調整總評分則總評分將會低估。直覺上的解釋是因為好樣本少放，使得整體壞率上升，總評分下降，因此必須靠 \hat{q}_0 、 \hat{q}_1 修正整體壞率，反映在評分模型上即是整體評分加一個常數即可。

3.4.5 案控樣本對證據權重的影響

由於證據权重也是使用案控樣本下的壞率進行轉換，因此與壞率有關的證據权重轉換也必須調整，從 (3.20) 式已知案控樣本對壞好率比的調整方式，將 (3.20) 式代入證據权重公式可以發現組壞率與總壞率的 q_1/q_0 調整效果正好相互抵消，故解釋變數若經證據权重轉換將不受案控樣本影響。

評等

當總評分轉換好後通常還會進一步對總評分進行分等，如此便可針對不同評等的借款者快速做出適當的授信決策。分等的方式也可同序別化的方式一樣看總評分能分出幾群樣本，即可分成幾等。不過一個好的總評分要能夠將壞樣本集中在低分群，好樣本集中在高分群。換言之，若模分模型夠好，則總評分的 ROC 線就應該要是一條很圓滑的曲線且遠離 45 度線。當 ROC 線是一條很圓滑的曲線時，只會有一個明顯的轉折點區分好壞樣本群，因此可以先將此點找出後，視銀行想要再從好壞樣本群中分成幾個等級，而採均分組的方式即可。

3.5 截斷點的選擇

截斷點 (cutoff point) 是用來拒絕借款申請者的臨界點，當總評分低於截斷點將無法取得借款，其目的就是把違約可能性較高的客戶排除在外。

直覺上截斷點應以盡可能減少好客戶流失的前提下排除傾向為壞客戶的借款者，因此截斷點應以最能區分好壞樣本的分數來做為截斷點，亦即使用 KS (Kolmogorov-Smirnov) 點做為截斷點：

$$KS = \text{Max}|F_1(S^*) - F_0(S^*)| \quad (3.22)$$

其中 F_1 與 F_0 分別為壞樣本與好樣本的累積分配，若單純以區分好壞樣本的話，當總評分為 S^* 時即為最適截斷點。由 KS 值的目標函式可求得一階條件為 $f_1(S^*) = f_0(S^*)$ ，可知 KS 值發生在 ROC 線的切線斜率為 1 的點。若 ROC 線是一條很圓滑的曲線且在 45 度線上，則當 ROC 線斜率要從大於 45 度變成 45 度時即為 KS 點，此點也將會是 ROC 線上到 45 度線垂直距離最大的點。

使用 KS 點來選擇截斷點雖能排除掉大多數的壞樣本，不過這是在不考慮客戶所帶來的收入與成本下所做的選擇，此點未必是利潤極大點。銀行在拒絕樣本的同時也拒絕掉了許多好客戶使得收入減少，換言之，銀行為了降低風險同時勢必要以減少收入做為代價，亦即降低風險與減少收入存在抵換關係 (trade-off)。因此可以考慮使預期利潤極大的總評分做為截斷點，其目標函式為：

$$E\pi = \eta F_0(S)(1 - P_0) - \lambda F_1(S)P_0 \quad (3.23)$$

其中 $E\pi$ 為預期利潤， S 為總評分， η 與 λ 則分別表示每增加一位好借款者的「單位收益率」與每增加一位壞借款者的「單位損失率」。由一階條件式可得：

$$\eta f_0(S)(1 - P_0) = \lambda f_1(S)P_0 \quad (3.24)$$

(3.24) 式的意涵即「邊際收益」等於「邊際損失」，移項後可得到：

$$\frac{f_1(S)}{f_0(S)} = \frac{\eta}{\lambda} \left(\frac{1 - P_0}{P_0} \right), \quad (3.25)$$

從 (3.25) 式可以清楚得知，使 ROC 線上的斜率等於單位收益率乘上好率與單位損失率乘上壞率的比值的總評分即利潤最大的截斷點。

第 4 章

評分模型的驗證

當評分模型建置完畢後必須驗證評分模型的可用性，一個好的評分模型其總評分應能很好的區分好壞樣本的概率密度函數以及累積分配。除此之外，也可以從評等指標進行配適度分析，故驗證統計量主要可以分成訊息理論、辨識度、以及評等配適度三大類。各類都有諸多相關的統計量，本節只說明本文有使用到的統計量。¹

4.1 驗證統計量

訊息理論本文使用息值 (Information Value, IV) 與離散統計量 (divergence statistic)，辨識度本文使用 AUC 值，評等配適度分析則為 HL 統計量 (Hosmer-Lemeshow, HL)、校準度 (calibration) 檢定與母體穩定度 (PSI)。由於本文的辨識度驗證統計量是使用 AUC 值，在 3.2.1 節已做說明故在此不多贅述。此外配適度分析中的 PSI 值與訊息理論中的信息值概念一樣，因此 PSI 值與信息值放在一起說明。

4.1.1 信息值

信息值原意是用來測度一隨機變數 x 的分配不確定下，給定兩個可用來描述隨機變數 x 的概率密度函數，計算兩概率密度函數的「相對熵」(relative entropy) (或「Kullback Leibler 距離」) 的加總值。相對熵 $D(f|g) \equiv -E_f[\ln g(x)] - E_f[\ln f(x)]$ ，其中 f 是可以描述隨機變數 x 的 (真實) 概率密度函數， g 則是另一個不希望用來描述隨機變數 x 的 (錯誤) 概率密度函數。相對熵所代表的意義是用錯誤的概率密度函數來描述隨機變數 x 所產生的效率損失。 $-E_f[\ln f(x)]$ 為隨機變數 x 的「熵」(entropy)。由於 f 和 g 的相對熵並不對稱，因此信息值便將兩概率密度函數的相對熵加總，以求得

¹其它相關統計量請參閱鍾經樊等 (2005)。

一個對概率密度函數 f 和 g 而言皆是對稱的函數。在評分模型的應用上, f 和 g 則分別用來表示在給定總評分 S , 好樣本的條件分配 f_0 與壞樣本的條件分配 f_1 的相對熵和, 故信息值的數式可寫成:

$$\begin{aligned} \text{IV}(f_0, f_1) &= D(f_0|f_1) + D(f_1|f_0) = E_{f_0} \left[\ln \frac{f_0(S)}{f_1(S)} \right] + E_{f_1} \left[\ln \frac{f_1(S)}{f_0(S)} \right] \\ &= \int_{-\infty}^{\infty} [f_0(S) - f_1(S)] \cdot \ln \left[\frac{f_0(S)}{f_1(S)} \right] dx, \end{aligned} \quad (4.1)$$

從 (4.1) 式可以看出信息值可以分成 $f_0(S) - f_1(S)$ 與 $\ln[f_0(S)/f_1(S)]$ 兩部份, 前者的意思為在給定總評分 S 之下好壞樣本的概率密度函數的高度差; 後者則為概率密度函數的高度比的對數值。後者的功能除了讓高度差有如取絕對值的效果外, 由於高度比的對數值可以反映高度差座落的位置, 好壞樣本的概率密度函數若愈分離則高度差都可以得到較大的權重²。反之, 若概率密度函數愈接近, 即使高度差很大信息值也不會太大。可知信息值的大小可以反映出 f_0 與 f_1 的「分離程度」。實際在運用時只能以 (4.2) 式的信息值間斷型公式計算:

$$\begin{aligned} \text{IV} &= \sum_{j=1}^J \left(\frac{g_j}{G} - \frac{b_j}{B} \right) \cdot \ln \left[\frac{g_j/G}{b_j/B} \right] \\ &= \sum_{j=1}^J \left(\frac{g_j}{G} - \frac{b_j}{B} \right) \cdot \text{WOE}_j \end{aligned} \quad (4.2)$$

其中 g_j 、 b_j 、 G 、 B 的定義與式 (3.2) 一樣。故信息值可以用來看各評等是否能很好的區分好壞樣本的概率密度函數。³

4.1.2 PSI 值

信息值可以用來看概率密度函數的分離程度相反的也可以用來看概率密度函數的接近程度, PSI 值即是以此概念所延伸出來的。PSI 值可用來比較各評等的樣本率 (組樣本數除上總樣本數) 與壞樣本率 (壞樣本除上總壞樣本數) 在建置樣本與驗證樣本是否相似。

PSI 值的計算方式即信息值的計算公式, 只要將好樣本與壞樣本的概率密度函數改成建置樣本與驗證樣本在各評等的樣本率, 即可計算全樣本的 PSI 值。若改成建置樣本與驗證樣本各評等的壞樣本率, 即可求得壞樣本的 PSI 值。

²相同的概率密度函數高度差, 是可以有不同的概率密度函數高度比。

³由於信息值會受組數的影響, 除非解釋變數分組的方式與方組數皆一樣, 否則不宜用來篩選變數。

4.1.3 離散統計量

離散統計量與信息值的功能一樣也是用來測度分配的分離程度, 不過其計算方式較為直接, 只看兩概率密度函數的平均數與變異數, 其計算公式為:

$$DS = \frac{(\mu_0 - \mu_1)^2}{(\sigma_0^2 + \sigma_1^2)/2},$$

其中 μ_0 與 σ_0^2 分別表示好樣本的平均數與變異數, μ_1 與 σ_1^2 即壞樣本的平均數與變異數。由上式可知, 若兩分配的平均數差距愈大, 則離散統計量愈大。若好壞樣本為具有相同變異數的常態分配, 則離散統計量會正好等於信息值⁴。此法相較於信息值的優點是不受分組數的影響。

4.1.4 HL 值

HL 值是從預測誤差的概念為基礎, 計算各評等的預測壞率與實際壞率的誤差值, 其愈小表示預測誤差愈小各評等的預測力愈準確。HL 值的計算公式為⁵:

$$HL = \sum_{j=1}^J \frac{n_j(\hat{p}_j - p_j^*)^2}{p_j^*(1 - p_j^*)}, \quad j = 1, 2, \dots, J,$$

其中 p_j^* 表示第 j 個評等的評等壞率 (預測壞率), \hat{p}_j 表示第 j 個評等的實際壞率, J 為總分等數, n_j 為第 j 組的樣本數。HL 值會呈卡方分配, 自由度則須視樣本來源, 若為建置樣本則自由度為 $J - 2$,

⁴假設在給定總評分 S 下的好、壞樣本概率密度函數 $f_0(S)$ 與 $f_1(S)$ 呈常態分配 $f_0(S) \sim \mathcal{N}(\mu_0, \sigma)$, $f_1(S) \sim \mathcal{N}(\mu_1, \sigma)$, 則 IV 可寫成:

$$\begin{aligned} IV &= \int_{-\infty}^{\infty} [f_0(S) - f_1(S)] \ln \frac{f_0(S)}{f_1(S)} \\ &= \frac{1}{2\sigma^2} [(S - \mu_1)^2 - (S - \mu_0)^2] \frac{1}{\sqrt{\pi\sigma^2}} \left\{ \exp\left[-\frac{(S - \mu_0)}{2\sigma^2}\right] - \exp\left[-\frac{(S - \mu_1)}{2\sigma^2}\right] \right\} \\ &= \frac{1}{2\sigma^2} [E_0(S - \mu_1)^2 - E_0(S - \mu_0)^2 - E_1(S - \mu_1)^2 + E_1(S - \mu_0)^2] \\ &= \frac{(\mu_0 - \mu_1)^2}{\sigma^2}. \end{aligned}$$

⁵由於在驗證時本文是將評等壞率視為各驗證樣本樣的預測壞率值, 理論上應稱為 Pearson 值, HL 值與 Pearson 值兩者主要差別在前者的預測壞率是使用同一評等裡的各樣本預測壞率平均值, 而後者則是有自然的組預測壞率值, 為 HL 值的特例。本文為方便說明, 一律稱 HL 值。

若不為建置樣本則自由度為 J 。不過由於 HL 值假設樣本之間是獨立，而樣本之所以會被歸為同一評等且具有相同的預測壞率，必然是因為這些壞樣本或者好樣本是受到共同因素的影響使得這些樣本不太可能是彼此獨立的，因此使用 HL 值用來驗證評等預測壞率則較為嚴苛，往往會使得各評等的 HL 值在驗證上容易拒絕掉各評等的實際壞率與各評等的預測壞率相等的虛無假設，故可以考慮使用校準度分析。

4.1.5 校準度分析

校準度分析則是在考慮二元狀態變數之相關性所做的檢定，一個常見的設定是「單因子門檻模型」。假設樣本的好壞是根據某一潛在指標 s_i^* 與門檻值 c_i^* 所決定：

$$y_i = \begin{cases} 1, & \text{若 } s_i^* \leq c_i^*, \\ 0, & \text{若 } s_i^* \geq c_i^*, \end{cases}$$

假設潛在指標具有如下的單因子結構：

$$s_i^* = \sqrt{\rho}x + \sqrt{1 - \rho}\varepsilon_i, \quad (4.3)$$

其中 x 乃是 $s_i^*, i = 1, \dots, N$ 所共有的「系統風險因子」， ε_i 是各指標的「特有風險因子」，彼此互相獨立且與 x 皆呈常態分配。 ρ 為系統風險因子和特有風險因子之間的相對重要性，是一個介於 0, 1 之間的值。由於評分模型是假設同一評等的樣本具有相同壞率 $p_j = p(y_i = 1) = \Phi(c_i)$ ，因此 $c_i = \Phi^{-1}(p_j)$ 。在給定上述假設下，則 y_i 在給定系統風險因子的條件壞率為：

$$\begin{aligned} p(y_i = 1|x) &= p(s_i^* \leq c_i^*|x) = p\left[\sqrt{\rho}x + \sqrt{1 - \rho}\varepsilon_i < \Phi^{-1}(p_j) \middle| x\right] \\ &= p\left[\varepsilon_i < \frac{\Phi^{-1}(p_j) - \sqrt{\rho}x}{\sqrt{1 - \rho}} \middle| x\right] = \Phi\left[\frac{\Phi^{-1}(p_j) - \sqrt{\rho}x}{\sqrt{1 - \rho}} \middle| x\right], \end{aligned} \quad (4.4)$$

根據大數法則可以證明各評等的實際壞率會收斂到 (4.4) 式⁶。假設第 j 個評等壞率預測值為 q_j ，在虛無假設為各評等的真實壞率 (p_j) 等於評等壞率 (q_j) 且對立假設為各評等的真實壞率大於評等壞

⁶證明過程可見 Bluhm, et al. (2003)。

率的假設下，則：

$$\begin{aligned}
 p(p_j > q_j) &\approx p \left\{ \Phi \left[\frac{\Phi^{-1}(p_j) - \sqrt{\rho}x}{\sqrt{1-\rho}} \middle| x \right] > q_j \right\} \\
 &= p \left[x < \frac{\Phi^{-1}(p_j) - \sqrt{1-\rho}\Phi^{-1}(q_j)}{\sqrt{\rho}} \right] \\
 &= \Phi \left[\frac{\Phi^{-1}(p_j) - \sqrt{1-\rho}\Phi^{-1}(q_j)}{\sqrt{\rho}} \right]。
 \end{aligned}$$

若 $p(p_j > q_j)$ 小於顯著水準則拒絕虛無假設。

4.2 控制變數對模型驗證的影響

控制變數的功能是用來控制影響違約行為的外在因素，在實際評分時都必須事先將外在因素控制在特定情況之下，因此在驗證模型時除了要看模型配適與預測的表現外，也要驗證模型在給定控制值⁷之下的表現。

不過在給定控制值之下驗證模型是有問題的，其問題在於解釋變數值的基準不一致，例如某解釋變數與景氣好壞呈正相關，景氣愈好解釋變數值就大，若將樣本在景氣好時所得到的解釋變數值控制在景氣壞的時候，則該樣本會得到異常的高分。反之，若將樣本在景氣壞時所得到的解釋變數值控制在景氣好的時候，則該樣本將會得到異常的低分。因此要驗證給定控制值的評分模型表現必須先將解釋變數值調整成同樣的條件才可以進行驗證。

要將樣本的解釋變數值調整成與控制值的基準一致，必須先行估計各解釋變數值與控制變數的關係才有辦法進行調整，因此必須先以各解釋變數做為應變數，控制變數做為解釋變數進行線性迴歸，求得各解釋變數變動一單位對解釋變數的影響就可以進行調整。假設解釋變數與控制變數的關係如下：

$$x_{ki} = \hat{\lambda}_{k1}z_{1i} + \dots + \hat{\lambda}_{kM}z_{Mi} + e_{ki},$$

$\hat{\lambda}_{km}$ 為第 m 個控制變數 z_{mi} 對第 k 個控制變數 x_{ki} 的迴歸係數， e_{ki} 則為迴歸模型的殘差值，下標 i 表示第 i 個樣本。現若要將控制變數控制成 $\bar{z}_1, \dots, \bar{z}_m$ ，則解釋變數值應調整如下：

$$\bar{x}_{ki} = \hat{\lambda}_{k1}\bar{z}_{1i} + \dots + \hat{\lambda}_{kM}\bar{z}_{Mi} + e_{ki},$$

⁷本文的給定控制值是指給定控制變數特定值。

故建置樣本的轉換前的總評分 \tilde{S}_i 應修正成：

$$\tilde{\tilde{S}}_i = -(\hat{\gamma}_1 \bar{z}_{1i} + \dots + \hat{\gamma}_M \bar{z}_{Mi}) - (\hat{\beta}_1 \bar{x}_{1i} + \dots + \hat{\beta}_K \bar{x}_{Ki}), \quad (4.5)$$

本文將解釋變數依控制值調整過後的總評分的 (4.5) 式稱為「調整分」，由 (4.5) 式可推得調整前後的差異為：

$$\tilde{\tilde{S}}_i - \tilde{S}_i = \sum_{m=1}^M (\hat{\gamma}_m + \sum_{k=1}^K \hat{\beta}_k \hat{\lambda}_{km}) (z_{mi} - \bar{z}_m). \quad (4.6)$$

由 (4.6) 式可知當控制變數值與壞樣本成反比 $\hat{\gamma}_m$ 會傾向為負，又序別化解釋變數與應變數的關係為負，故 $\hat{\lambda}_{km}$ 傾向為正。若將外在環境控制在壞情況下則 z_{mi} 傾向大於 \bar{z}_m ，故調整前的分數偏高，若未調整分數進行評等驗證樣本將會集中在較壞的等級。反之，若將外在環境控制在好情況下驗證樣本將會集中在較高的等級。

若控制變數與壞樣本成正比 $\hat{\gamma}_m$ 會傾向為正，又序別化解釋變數與應變數的關係為負，故 $\hat{\lambda}_{km}$ 傾向為負。若將外在環境控制在壞情況下，則 z_{mi} 傾向小於 \bar{z}_m ，故調整前的分數偏高，若未調整分數進行評等驗證樣本將會集中在較壞的等級。反之，若將外在環境控制在好情況下驗證樣本將會集中在較好的等級。

第 5 章

實證

由於銀行對於顧客的存款、放款或匯款等有關資料有保守秘密的義務¹，因此消費者的借貸資料不易取得，故本文使用台灣經濟新報資料庫 (TEJ) 所提供的企業基本資料與財務比率資料建置企業信用評分模型，預測企業發生財務危機的行為。財報資料取自上市 (櫃) 或曾經上市 (櫃) 一般產業的資料庫，財務危機的定義則延用台灣經濟新報的財務危機與準財務危機定義 (請見 2.2 節)。為方便說明，本節將曾發生過財務危機的企業稱為危機企業或壞樣本，未發生過財務危機則稱正常企業或好樣本。

樣本收集期間

本文樣本收集期間為 1996 年至 2003 年，從 1996 年開始取的主因是 1998 和 1999 兩年有許多企業發生財務危機，為了要收集較多的壞樣本同時又為了解決 Ohlson (1980) 的財報選用問題，因此必須使用一至一年半前的財報來做預測。取自 2003 的原因是 2004 年的樣本必須觀察到 2006 底才能確認樣本的好壞，由於尚無 2006 底的財報資料，因此只能取至 2003 底。對於曾發生財務危機的企業財報選用規則如表 1 所示。對於曾發生財務危機的企業，儘管財務危機已解決，也不將這些企業其他年度的財報視為好樣本，其原因是企業會發生危機事件有可能已反映在更早期 (一年半以前) 的財報裡，因此不將壞樣本發生危機一年半前或更早期的財報視為正常企業。企業發生危機事件後，即使財務危機已解決，由於企業商譽曾經受損，多少會影響企業未來的營運狀況，因此亦不將危機處理後的財報視為正常企業。至於正常企業的財報理論上也要依表 1 的方式取樣，但因為正常企業的財報皆是用來預測未來仍是正常企業的訊息，因此只要在未來一年半裡沒有發生財務危機，正常企業有無往前取一年半的財報是沒有關係的。本文中的好樣本皆是觀察到 2006 年仍未發生危機的企業，將這些未曾發生危機的企業視為好樣本是沒有問題。

¹銀行法第 48 條第 2 項之規定。

表 1：財報選用時點

財務危機發生時點 (年/月)	財報選用時點 (年/月)
1997/12、1998/03、1998/06、1998/09	1996/12
1998/12、1999/03、1999/06、1999/09	1997/12
1999/12、2000/03、2000/06、2000/09	1998/12
2000/12、2001/03、2001/06、2001/09	1999/12
2001/12、2002/03、2002/06、2002/09	2000/12
2002/12、2003/03、2003/06、2003/09	2001/12
2003/12、2004/03、2004/06、2004/09	2002/12
2004/12、2005/03、2005/06、2005/09	2003/12

經上述方式處理資料並刪除不合理的樣本值後²，本文將 2001 至 2003 年的 4,202 筆樣本做為驗證樣本，其中好樣本有 4,113 筆，壞樣本有 89 筆。將 1996 至 2000 年視為建置樣本，共 6,022 筆，其中好樣本佔 5,888 筆，壞樣本佔 134 筆。之所以使用 2000 年來劃分建置樣本與驗證樣本的主因是考量壞樣本的多寡，好讓建置樣本可以有較多的壞樣本用於建置評分模型，且驗證評分模型時驗證樣本的壞樣本也不致於過少。

變數缺漏值處理

資料庫裡的財務比率變數雖有 78 個，不過營收變動率、營業利益變動率和稅後淨利變動率的資料全為缺漏值，股利殖利率、現金股利率、當季季底本益比 (PER)、當季季底市價淨值比 (PBR) 和當季季底市價營收比 (PSR) 的缺漏值佔總樣本比重皆為 64%，現金流量允當比率為 47%，或有負債/淨值有將近 68% 的樣本是零，本文認為上述 10 個變數本身所能提供的訊息不具代表性，故實際用來進行篩選的變數有 68 個 (請見附錄)。此外本文認為解釋變數的缺漏值樣本若不足 2%，則缺漏值組所帶來的訊息可能不具代表性，由於刪除缺漏值後只會使壞樣本減少 2 筆，因此本文將比率變數中缺漏值不足 2% 的樣本刪除。刪除後樣本數為 6,013 筆，其中壞樣本為 134 筆，其餘 5,879 筆皆為好樣本。

²將各年資料合併後，刪除沒有財報資料的企業以及非單一企業的綜類財報，並依表 1 的方式取得壞樣本的資料後，共有 10,473 筆樣本 (1,631 家企業)，其中壞樣本佔 226 筆 (226 家企業)。將淨值、營收、營業成本小於等於零或缺漏值等 249 筆樣本刪除，刪除後樣本剩 10,224 筆，其中好樣本尚有 10,001 筆 (1,389 家企業)，壞樣本則有 223 筆 (223 家企業)。

案控樣本

本文建置樣本的好樣本將近壞樣本 45 倍之多，總壞率只有 0.0223。由於使用壞率過低的樣本建置評分模型會使得評分模型很不穩定，因此本文採案控樣本的方式從建置樣本的好樣本中隨機抽出 40% 的好樣本用來建置模型。故實際用來建模的樣本數為 2,486 筆，其中壞樣本有 134 筆，好樣本有 2,352 筆，總壞率為 0.0539。

控制變數

企業是否會發生財務危機除了從企業內部的營運狀況觀之，還應考慮企業外部的經濟環境因素，因此本文除了使用企業的財務比率資料外，還有使用總體變數做為模型的控制變數以掌控經濟環境對企業發生危機的外在因素，

本文從台灣經濟新報總體資料庫的景氣指標、就業統計、貨幣金融統計和國民所得統計四類中，挑出可能對企業有營運有影響的十個候選控制變數，景氣指標包括景氣對策信號綜分、躉售物價指數 (WPI%)(2001=100)、消費者物價指數 (CPI%)(2001=100)、和台灣股價指數 (月平均)；就業統計包括 15 歲以上民間人口數；貨幣金融統計包括 M2 年增率 (月底值)、擔保放款融通利率、短期融通利率、隔夜加權平均拆款利率；國民所得包括實質國民所得成長率 (2001=100)。不過十個控制變數中短期融通利率因為五年值都一樣無法使用；景氣對策信號綜分、台灣股價指數、十五歲以上民間人口、及擔保放款融通利率很不顯著；躉售物價指數與消費者物價指數雖然皆很顯著，但因彼此高度線性相關，一起放入模型後將會使得兩者皆不顯著，因此選用相對較為顯著之消費者物價指數。剩餘的控制變數中只有廣義貨幣 (M2) 年增率 (月底值) 與消費者物價指數一起放入模型中仍可維持各自的顯著性，不過在進行篩選解釋變數的過程中，加入變數後皆會造成 M2 年增率變得很不顯著，因此本文最後只使用 CPI 做為評分模型的控制變數。

變數序別化與變數篩選

篩選的首要步驟即排除不穩定的解釋變數，在 68 個候選解釋變數中營業費用率%、用人費用率%、研究發展費用率%、現金流量比率%、有息負債利率%、稅率 (A)%、稅率 (B)%、每股淨值 (B)(元)、每股營業額 (元)、每股淨值 (A)(元)、每股淨值 (C)(元)、稅後淨利成長率%、經常淨利成長率%、總資產成長率%、內部保留比率%、總資產週轉率 (次)、存貨週轉率 (次)、平均銷售天數 (天)、固定資產週轉率 (次)、淨值週轉率 (次)、應付帳款收現天數 (天)、淨營業週期 (天)、每人營收 (千元)、每人營業利益 (千元)、以及每人配備率等 25 個變數在建置樣本與驗證樣本的好壞樣本的概率密度函數或

AUC 值有明顯的不同；長期資金適合率 (A)%、存貨應收帳款/淨值、應收帳款週轉率 (次)、及應收帳款收現天數 (天) 等 4 個比率變數的 ROC 線沿著 45 度線走，本文認為這 4 個變數對模型沒有貢獻，因此將以上 29 個變數做初步排除。此外淨值比率%與負債比率%的呈完全線性相關，本文選用負債比率%，故最後本文使用 38 個候選解釋變數進行序別化並採用正向選入法進行篩選變數，篩選出借款依存度%、資產報酬成長率%、淨值報酬率 (常續性利益)%、資產報酬率 (稅前息前折舊前)%、每股現金流量 (元)、營業毛利率%、以及營業毛利成長率等 7 個財務比率變數做為評分模型之解釋變數，其序別化結果置於表 3，在序別化的過程中若組壞率為 0 或 1 時將無法做證據權重轉換，因此本文中若組壞率為 0 則以 0.0003 取代，若組壞率為 1 則以 0.9997 取代。

變數說明

借款依存度% 測度長短期借款占自有資本之比率，就債權保障觀點而言，本比率愈低愈佳。

計算公式：

$$\text{借款依存度} = \text{長短期借款} / \text{淨值} \times 100\%$$

長短期借款包括短期借款、應付商業本票、一年內到期負債與長期負債。

資產報酬成長率 (稅前息前折舊前)% 衡量公司報酬率增加的比率，能了解公司的獲利與營運成長狀況，就經營績效衡量觀點，成長率愈高愈好。

計算公式：

$$\begin{aligned} \text{資產報酬成長率 (稅前息前折舊前)\%} = \\ (\text{稅前息前折舊前淨利} - \text{前期稅前息前折舊前淨利}) / \text{期初資產總額} \times 100\% \end{aligned}$$

淨值報酬率 (常續性利益)% 測度企業自有資本之稅後獲利能力，就經營績效衡量觀點而言，本比率愈高愈佳。

計算公式：

$$\text{淨值報酬率 (常續性利益)\%} = \text{常續性利益} / [(\text{期初淨值} + \text{期末淨值}) / 2] \times 100\%$$

常續性利益依台灣經濟新報資料的定義如下：

$$\begin{aligned} \text{常續性利益} &= \text{稅後淨利} + (\text{處分投資損失} + \text{投資跌價損失} + \text{處分資產損失}) \\ &\quad - (\text{處分投資利得} + \text{投資跌價損失回轉} + \text{處分資產利得}) \\ &= \text{稅後淨利} + \text{處份固定資產及處分投資損失} \\ &\quad - \text{處份固定資產及處分投資利益} \end{aligned}$$

資產報酬率 (稅前息前折舊前)% 用來測度公司資產能夠創造出來的稅後息前淨利能力, 代表著公司所運用一元的資產能夠創造出多少的稅後息前淨利。此比率愈高代表公司資產之獲利愈高。

計算公式:

$$\text{資產報酬率 (稅前息前折舊前)\%} = \text{稅前息前折舊前之常續性利益} / [(\text{期初資產總額} + \text{期末資產總額}) / 2] \times 100\%$$

每股現金流量 從短期的觀點來看, 每股現金流量比每股盈餘更能顯示從事資本支出及支付股利的能力, 本比率愈高愈好。

計算公式:

$$\text{每股現金流量} = (\text{營業活動淨現金流量} - \text{特別股股利}) / \text{加權平均股本} \times 10。$$

$$\text{營業活動淨現金流量} = \text{流動資產} - \text{流動負債} - \text{滯銷的存貨}。$$

營業毛利率% 測度企業產銷效能, 就經營績效衡量觀點而言, 本比率愈高愈佳。

計算公式:

$$\text{營業毛利率\%} = \text{營業毛利} / \text{營業收入淨額} \times 100\%$$

營業毛利成長率% 營業毛利成長率通常所代表的是一個企業的創新能力, 本比率愈高愈佳。

計算公式³:

$$\text{營業毛利成長率} = \text{營業毛利增減額} / \text{abs(去年同期毛利)} \times 100\%$$

$$\text{營業毛利} = \text{營收淨額} - \text{營業成本},$$

實證結果

表 2 為 Logit 模型估計結果以及各變數的評分貢獻度與 AUC 值。從表 2 中可看出各解釋變數的係數皆為負, 乃是因變數進行證據權重轉換的關係所致。從資產報酬成長率的 AUC 值可以得知, 做為解釋變數不見得 AUC 值都要很高, 只要係數的符號正確, 且能對評分模型有相當程度的評分貢獻度, 仍然可選入評分模型。從表 3 中可得知貢獻度高的變數乃是因各組的壞率差異大, 表示貢獻度的高低可反映解釋變數的辨識能力。由控制變數的係數值為負值可知消費者物價指數與壞率呈負相關, 表示消費者物價指數愈低企業愈傾向於發生財務危機。

³式中的 abs 表示取絕對值之意。

表 2：Logit 模型估計結果與各解釋變數評分貢獻度及 AUC 值

解釋變量	係數估計值	<i>t</i> 值	貢獻度1	貢獻度2	AUC 值
借款依存度	-0.890	-7.90	43.43	45.29	0.783
資產報酬成長率	-0.373	-2.70	14.22	12.97	0.608
淨值報酬率	-0.345	-1.96	12.47	12.02	0.737
資產報酬率	-0.195	-1.01	9.23	9.75	0.752
每股現金流量	-0.289	-1.60	8.60	8.44	0.658
營業毛利率	-0.333	-2.35	6.51	7.20	0.695
營業毛利成長率	-0.282	-2.10	5.54	4.34	0.630
消費者物價指數	-0.267	-2.28			
截距項	-2.597	-13.98			

表 3 為本文所製作的評分卡，解釋變數名稱後括號裡的代號為變數在台灣經濟新報資料庫裡的代號。樣本率或壞率有加（調）的欄位表示原案控樣本已依案控比例調整過。PSI 欄中的“.”表示該組無法計算 PSI 值。本文使用 (3.18) 式進行總評分之轉換，設定壞率為 0.0003 與 0.9997 時總評分要分別為 1000 分與 0 分。經此設定後 (3.14) 式中的 A 為 500 分， B 為 61.64 分。由於本文只抽取 40% 的好樣本進行建置評分模型，依第 32 頁的案控樣本設定，應分別設定 $\hat{q}_0 = 0.4$ 與 $\hat{q}_1 = 1$ ，將轉換前的總評分調高 $\ln(1/0.4)$ ，否則轉換前的總評分將偏低。底分經上述調整後，轉換後的底分為 716.58。

由表 3 中可以清楚得知各變數評分與 44 頁的變數說明一致，比率愈高愈好的變數組壞率會隨比率增加而遞減，各組評分則會隨比率增加而遞增。反之，比率愈低愈好的變數組壞率會隨比率增加而遞增，各組評分則會隨比率增加而遞減。從建置樣本的組壞率與驗證樣本的組壞率亦可看出，在案控樣本之下建置樣本的組壞率相較於驗證樣本的組壞率皆有偏高的傾向。此外從 PSI 值可知各解釋變數若不考慮缺漏值的樣本率，則 PSI 值都相當低，可知解釋變數的全樣本分配與壞樣本分配在建置樣本與驗證樣本中皆很穩定。

表 3：信用評分

1. 借款依存度% (r509), 全樣本 PSI: 0.230, 壞樣本 PSI: 0.094									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$0 \leq x \leq 20.32$	137.00	0.355	0.392	0.004	0.345	0.392	857	1648
2	$20.32 < x \leq 50.50$	19.88	0.245	0.288	0.007	0.243	0.288	603	1209
3	$50.50 < x \leq 93.15$	-21.61	0.188	0.209	0.002	0.191	0.209	475	879
4	$93.15 < x < \infty$	-73.63	0.145	0.108	0.011	0.157	0.108	391	455
5	缺漏值	260.27	0.067	0.003	0.207	0.064	0.003	160	11
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$0 \leq x \leq 20.32$	137.00	0.002	0.006	0.087	0.005	0.006	4	9
2	$20.32 < x \leq 50.50$	19.88	0.016	0.011	0.004	0.038	0.011	23	13
3	$50.50 < x \leq 93.15$	-21.61	0.033	0.027	0.000	0.078	0.027	37	24
4	$93.15 < x < \infty$	-73.63	0.080	0.095	0.003	0.179	0.095	70	43
5	缺漏值	260.27	0.000	0.000	.	0	0	0	0
2. 資產報酬成長率% (r411), 全樣本 PSI: 0.314, 壞樣本 PSI: 0.023									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -4.32$	-23.66	0.100	0.152	0.022	0.105	0.152	262	638
2	$-4.32 < x \leq 3.76$	-4.14	0.426	0.514	0.017	0.428	0.514	1065	2159
3	$3.76 < x < \infty$	8.63	0.323	0.316	0.000	0.320	0.316	796	1326
4	缺漏值	109.19	0.151	0.019	0.275	0.146	0.019	363	79
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -4.32$	-23.66	0.060	0.047	0.016	0.137	0.047	36	30
2	$-4.32 < x \leq 3.76$	-4.14	0.027	0.020	0.003	0.064	0.020	68	42
3	$3.76 < x < \infty$	8.63	0.015	0.013	0.005	0.038	0.013	30	17
4	缺漏值	109.19	0.000	0.000	.	0.000	0.000	0	0

(續) 表 3：信用評分卡

3. 借淨值報酬率 (常續性利益)% (r104), 全樣本 PSI: 0.141, 壞樣本 PSI: 0.144									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -6.37$	-28.95	0.106	0.140	0.010	0.115	0.140	286	590
2	$-6.37 < x \leq 2.71$	-11.27	0.147	0.191	0.011	0.150	0.191	374	802
3	$2.71 < x \leq 11.14$	3.50	0.281	0.286	0.000	0.279	0.286	694	1201
4	$11.14 < x \leq 19.79$	14.40	0.218	0.192	0.003	0.214	0.192	533	808
5	$19.79 < x < \infty$	54.96	0.194	0.185	0.000	0.188	0.185	467	779
6	缺漏值	100.92	0.055	0.005	0.117	0.053	0.005	132	22
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -6.37$	-28.95	0.082	0.085	0.064	0.182	0.085	52	50
2	$-6.37 < x \leq 2.71$	-11.27	0.037	0.015	0.067	0.088	0.015	33	12
3	$2.71 < x \leq 11.14$	3.50	0.019	0.017	0.001	0.046	0.017	32	20
4	$11.14 < x \leq 19.79$	14.40	0.012	0.009	0.012	0.028	0.009	15	7
5	$19.79 < x < \infty$	54.96	0.002	0.000	.	0.004	0.000	2	0
6	缺漏值	100.92	0.000	0.000	.	0.000	0.000	0	0
4. 資產報酬率 (稅前息前折舊前)% (r100), 全樣本 PSI: 0.215, 壞樣本 PSI: 0.101									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -2.39$	-15.92	0.080	0.111	0.010	0.087	0.111	216	466
2	$-2.39 < x \leq 4.76$	-9.69	0.161	0.218	0.017	0.167	0.218	416	916
3	$4.76 < x \leq 12.27$	0.72	0.348	0.348	0.000	0.346	0.348	864	1464
4	$12.27 < x < \infty$	26.15	0.321	0.314	0.000	0.311	0.314	774	1319
5	缺漏值	57.16	0.090	0.009	0.188	0.087	0.009	216	37
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -2.39$	-15.92	0.079	0.079	0.051	0.176	0.079	38	37
2	$-2.39 < x \leq 4.76$	-9.69	0.049	0.033	0.001	0.113	0.033	47	30
3	$4.76 < x \leq 12.27$	0.72	0.021	0.013	0.050	0.051	0.013	44	19
4	$12.27 < x < \infty$	26.15	0.003	0.002	0.000	0.007	0.002	5	3
5	缺漏值	57.16	0.000	0.000	.	0.000	0.000	0	0

(續) 表 3：信用評分卡

5. 借每股現金流量 (元) (r303), 全樣本 PSI: 0.210, 壞樣本 PSI: 0.017									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -0.01$	-11.78	0.307	0.256	0.009	0.316	0.256	785	1075
2	$-0.01 < x \leq 2.48$	1.15	0.350	0.445	0.023	0.349	0.445	868	1870
3	$2.48 < x < \infty$	20.43	0.278	0.295	0.001	0.272	0.295	675	1241
4	缺漏值	84.54	0.066	0.004	0.176	0.064	0.004	158	16
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -0.01$	-11.78	0.042	0.052	0.004	0.099	0.052	78	56
2	$-0.01 < x \leq 2.48$	1.15	0.021	0.013	0.012	0.051	0.013	44	24
3	$2.48 < x < \infty$	20.43	0.007	0.007	0.001	0.018	0.007	12	9
4	缺漏值	84.54	0.000	0.000	.	0.000	0.000	0	0
6. 營業毛利率% (r105), 全樣本 PSI: 0.017, 壞樣本 PSI: 0.063									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq 10.71$	-19.92	0.197	0.244	0.010	0.206	0.244	513	1023
2	$10.71 < x \leq 14.63$	-3.52	0.120	0.130	0.001	0.121	0.130	300	546
3	$14.63 < x \leq 37.64$	10.51	0.543	0.487	0.006	0.536	0.487	1332	2048
4	$37.64 < x < \infty$	32.22	0.141	0.139	0.000	0.137	0.139	341	585
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq 10.71$	-19.92	0.057	0.051	0.013	0.131	0.051	67	52
2	$10.71 < x \leq 14.63$	-3.52	0.026	0.020	0.003	0.063	0.020	19	11
3	$14.63 < x \leq 37.64$	10.51	0.014	0.010	0.031	0.033	0.010	44	21
4	$37.64 < x < \infty$	32.22	0.005	0.009	0.017	0.012	0.009	4	5

(續) 表 3：信用評分卡

7. 營業毛利成長率% (r402), 全樣本 PSI: 0.170, 壞樣本 PSI: 0.031									
分組	組界點	組評分	樣本率 (調)		全樣本	樣本率		樣本數	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -24.90$	-21.55	0.098	0.173	0.043	0.105	0.173	261	728
2	$-24.90 < x \leq 15.45$	0.39	0.320	0.368	0.007	0.320	0.368	796	1545
3	$15.45 < x < \infty$	5.82	0.508	0.447	0.008	0.504	0.447	1252	1879
4	缺漏值	82.34	0.074	0.012	0.112	0.071	0.012	177	50
分組	組界點	組評分	壞率 (調)		壞樣本	壞率		壞樣本	
			建模	驗證	PSI	建模	驗證	建模	驗證
1	$-\infty < x \leq -24.90$	-21.55	0.073	0.050	0.019	0.165	0.050	43	36
2	$-24.90 < x \leq 15.45$	0.39	0.022	0.016	0.004	0.053	0.016	42	25
3	$15.45 < x < \infty$	5.82	0.016	0.015	0.008	0.039	0.015	49	28
4	缺漏值	82.34	0.000	0.000	.	0.000	0.000	0	0

評分模型的驗證結果

製作好表 3 的信用評分卡後，加總樣本在各解釋變數的分數以及底分與調整分後即可求得樣本的總評分。至於控制分的計算，本文是使用 2001 至 2003 年的消費者物價指數平均值 -0.16 做為控制變數的控制值，且總評分有依控制變數控制值進行 (4.5) 式的修正，總評分的分等方式則依 ROC 線的形式大致均分成六等，分等結果置於表 6。總評分在建置樣本與驗證樣本的驗證統計量置於表 4。驗證樣本除了使用 2001 至 2003 年所有樣本一起進行驗證外，亦有分年進行驗證，2001 至 2003 年的樣本數依序為 1,383、1,407、及 1,412 個，壞樣本數依序為 17、29、及 43 個。

就評分模型的辨識度而言，從表 4 的驗證統計量可看出除了 2002 年的驗證結果與較為偏低外，其餘的 AUC 值至少都有 0.84 以上，表示本評分模型具有高度的辨識能力，且信息值至少都有 1 以上，離散統計量也都有 2 以上，表示總評分能很好區好壞樣本的概率密度函數。從驗證樣本的 HL 值⁴偏大的情況可知建置樣本與驗證樣本的評等壞率有明顯的不同，不過由表 3 各解釋變數的 PSI 值可知，解釋變數的 PSI 值偏大的主因來自於缺漏值組，扣除缺漏值組後其實各解釋變數的 PSI 值都很低，可知 HL 值偏大的情況應是受這些缺漏值樣本的影響。此外建置樣本的期間屬於景氣較好的時期，而驗證樣本的期間則屬於景氣較壞的時期，本文只使用消費者物價指數做為控制變數，故所能捕捉到景氣變化對評分的影響有限，使得驗證樣本的組壞率與預測壞率有明顯的不同。儘管如此，本文的評等指標至少都有符合校準度檢定之要求，校準度檢定 p 值請參見表 5，本文校準度檢定的風險相

⁴在計算 HL 統計量時，建置樣本的實際壞率是使用經案控比例調整過後的實際壞率，驗證樣本則無案控的問題。

關係數值設為 0.25。

各年度的驗證結果差異主要是受解釋變數在各年度裡的穩定性影響，若解釋變數值在建置樣本與驗證樣本有顯著的差異，則模型的驗證結果亦會有明顯的不同。雖然 2002 年的驗證結果相較於其他年來得偏低，但比較 2002 年前後的驗證統計量，就整體的驗證結果而言，本文所製作的評分指標尚有一定程度的穩定性。

表 4：驗證統計量彙總⁵

驗證統計量	建置樣本	驗證樣本			
		2001-2003	2001	2002	2003
AUC 值	0.880	0.841	0.942	0.787	0.846
信息值	1.734	1.459	0.935	0.852	1.002
離散統計量	2.076	1.964	4.184	1.250	2.124
HL 統計量	0.720	32.029	28.764	15.983	5.526

表 5：校準度檢定統計量 p 值

評等	壞率	建模
1	0.000	.
2	0.001	0.208
3	0.005	0.242
4	0.012	0.272
5	0.024	0.299
6	0.103	0.367

驗證評分指標的優劣除了看表 4 與表 5 的驗證統計量外，還必須看評等表的 PSI 值以確保評等表的穩定度與可用性。檢驗評等的穩定度時必須看在給定控制值下各評等的全樣本 PSI 值與壞樣本 PSI 值，表 6 為基準分的評等分佈，驗證評等的穩定度時必須使用基準分的原因是從 1996 至 2003 年的消費者物價指數依序為 3.07、0.9、1.68、0.18、1.25、-0.01、-0.2、-0.28，若解釋變數不進行調整，依消費者物價指數與壞率呈反相關係可知建置樣本的解釋變數值會是屬於景氣相對較好下的結果，若解釋變數不進行調整，建置樣本的總評分會偏高。由全樣本 PSI 值與壞樣本 PSI 值皆小於 1 可知，本文的評等指標在不同時期下仍具有相當程度的預測率。

⁵其中信息值與 HL 統計量皆為間斷型的統計量，本文是使用配適分均分六等的方式計算而得。

表 6：評等表

全樣本 PSI: 0.028, 壞樣本 PSI: 0.052									
評等	評等組界	樣本率 (調)		全樣本	樣本率		樣本數		
		建模	驗證	PSI	建模	驗證	建置	驗證	
1	904 - ∞	0.173	0.142	0.006	0.167	0.142	416	595	
2	811 - 903	0.171	0.216	0.011	0.165	0.216	411	906	
3	738 - 810	0.172	0.168	0.000	0.167	0.168	416	705	
4	677 - 737	0.169	0.192	0.003	0.167	0.192	414	808	
5	613 - 676	0.167	0.165	0.000	0.167	0.165	415	695	
6	-∞ - 612	0.149	0.117	0.008	0.167	0.117	414	493	

(續) 表 6：評等表

評等	評等組界	評等	壞率 (調)		壞樣本	壞率	
		壞率	建模	驗證	PSI	建模	驗證
1	904 - ∞	0.000	0.000	0.002	.	0.000	0.002
2	811 - 903	0.001	0.001	0.002	0.017	0.002	0.002
3	738 - 810	0.005	0.005	0.007	0.008	0.012	0.007
4	677 - 737	0.012	0.012	0.012	0.005	0.029	0.012
5	613 - 676	0.024	0.024	0.027	0.006	0.058	0.027
6	-∞ - 612	0.103	0.103	0.106	0.017	0.222	0.106

與 TCRI 信用指標的比較

台灣經濟新報資料庫亦有對各企業進行信用評等，其評等指標命名為「TEJ 企業信用風險指標」(Taiwan Corporate Credit Risk Index, TCRI)，表 7 為本文的評等指標與 TCRI 在建置樣本、驗證樣本、2001 年、2002 年、以及 2003 年的驗證統計量比較，由於建置樣本與驗證樣本中有些樣本並無 TCRI 的評等資料，為比較基準一致，在表 7 中本文的評等指標⁶ 是使用與 TCRI 相同樣本下所計算而得並與 TCRI 一樣分 9 個評等，分等方式則以簡單的均分方式。由於有些評等沒有壞樣本使得信息值、HL 值無法計算，此外因為不知道 TCRI 對各評等的預測壞率為何，亦無法進行各評等的校準度分析，故表 7 中並無信息值、HL 值以及校準度分析等比較。

⁶這裡的評等指標是使用配適分。基本上除了計算評等指標的 PSI 值外，所有的驗證統計量都應看配適分。

表 7：TCRI 與本文評等指標的驗證統計量

	本文評等指標					TCRI				
	建置	驗證	2001	2002	2003	建置	驗證	2001	2002	2003
AUC 值	0.851	0.824	0.886	0.802	0.825	0.770	0.797	0.825	0.775	0.807
離散統計量	2.218	1.843	2.897	1.608	1.808	1.034	1.429	1.799	1.177	1.518

由表 7 可以清楚得知，無論是在建置樣本、驗證樣本或對各年度的評等比較，皆顯示本文所製作的評等指標的辨識度優於 TCRI。不過從對 TCRI 的驗證統計量可知，台灣經濟新報歷年來對 TCRI 的修正確實有提高 TCRI 的辨識度，不過還是有改進空間。此外從表 7 對 TCRI 的驗證結果可知 TCRI 與本文面臨相同的情況，亦即對 2002 年的驗證結果相較於其他年皆明顯不佳，可知 2002 年的企業發生財務危機之原因有明顯異於其他年。

第 6 章

結論與建議

本文將 ROC 分析的概念引入解釋變數的序別化過程，以增進解釋變數辨識度的方式進行序別化，不同於以往對解釋變數的序別化皆採行簡單的均分或主觀的分組方式。從實證結果顯示無論是在建置樣本、驗證樣本或對各年度的評等比較，皆顯示本文所製作的評等指標優於行之多年之 TCRI。此外從對 TCRI 的驗證結果可知 TCRI 與本文面臨相同的情況，亦即對 2002 年的驗證結果相較於其他年皆明顯不佳，可知 2002 年的企業發生財務危機之原因，有明顯異於其他年，換言之，除了財務比率等資料可以用來預測企業發生財務危機的可能性外，尚存在其他非財務比率等訊息會影響企業是否會危機。

由於配適分的驗證統計量高只能反應壞樣本集中在低分群，好樣本則集中在高分群，並無法反映樣本整體的評等分佈是否改變。本文除了看配適分的驗證統計量外，尚有提出加入控制變數後對評分模型的驗證方式，從實證結果顯示本文的評等表的全樣本與壞樣本 PSI 值只有 0.028 與 0.052，可知評分模型具有相當程度的穩定性。

就本文所製作的評分模型而言，本文認為尚有以下三點值得進一步研究之處，礙於篇幅問題並未做進一步探討。

1. 因為建立模型所使用的樣本只有曾獲得借款的申請者才可被觀察到，然而被拒絕申請的樣本則是無法觀察到的，此種資料遭截斷 (truncation) 的情形將會造成樣本選擇偏誤的問題。樣本選擇偏誤的大小，會隨著截斷點 (cutoff point) 的高低而改變，一個具有較高截斷點的評分卡，會有較嚴重的樣本選擇偏誤問題。不過由於借貸資料攸關資料洩密的問題不易取得，因此本文的實證資料是使用上市、上櫃的財報資料，此樣本並沒有樣本選擇的問題，所以無法進行樣本選擇的實證。
2. 產業間之差異性頗大，由於其特性不一，遭逢之風險也不一，直覺上產業變數應會對評分模型有所影響。不過本文直接將產業別的類別變數進行序別化，並將訊息相似的產業別併組後放入

模型中，對評分模型的預測並沒有作用。其原因可能是因解釋變數皆為財務比率變量，因而無法反映各會計科目的原始大小。

但也不宜將產業變數直接作為控制變數放入模型中，因為控制變數的作用是控制影響違約行為的外在因素，而產業別若具有效果應屬於內生因素，若將未來的借款者設定為特定產業別亦不合理，因此也不能將產業別做為控制變數。

對於此種情形本文認為必須依各產業分別建立各自的評分卡並比較各產業間的評分差異，才能反映哪些產業有顯著的不同。除了產業別外，舉凡直覺上認為會隨某種類別不同而有不同的違約行為，如地區別，都應各別製卡才能比較出各評分模型的評分方式是否有顯著差異。

3. 解釋變數進行序別化後雖可保有缺漏值組所帶的訊息，不過缺漏值組的代表性如何，是否應獨立成組，是否應併組，應如何併組等議題也都尚有討論空間。

附錄

候選解釋變量

候選解釋變量			
獲利能力	償債能力	經營能力	成本費用率
資產報酬率 (稅前息前折舊前)%	現金再投資%	存貨應收帳款/淨值	營業費用率%
資產報酬率 (稅後息前)%	流動比率%	總資產週轉率 (次)	用人費用率%
資產報酬率 (稅後息前折舊前)%	速動比率%	應收帳款週轉率 (次)	研究發展費用率%
淨值報酬率 (稅後)%	利息支出率%	應收帳款收現天數 (天)	呆帳費用率
淨值報酬率 (常續性利益)%	總負債/總淨值%	存貨週轉率 (次)	現金流量比率%
營業毛利率%	負債比率%	平均銷售天數 (天)	有息負債利率%
已實現銷貨毛利率%	淨值比率%	固定資產週轉率 (次)	稅率 (A)%
營業利益率%	長期資金適合率%	淨值週轉率 (次)	稅率 (B)%
稅前淨利率%	借款依存度%	應付帳款收現天數 (天)	
稅後淨利率%	利息保障倍數	淨營業週期 (天)	
業外收支率%	營業利益/實收資本		
常續性利益率 (稅後)%	稅前純益/實收資本		
成長率	每股比率	其它比率	法定比率
營收成長率%	每股淨值 (A)(元)	每人營收 (千元)	營運槓桿度
營業毛利成長率%	每股淨值 (B)(元)	每人營業利益 (千元)	財務槓桿度
已實現銷貨毛利成長率%	每股淨值 (C)(元)	每人配備率	
營業利益成長率%	常續性每股盈餘 (EPS)		
稅前淨利成長率%	每股現金流量 (元)		
稅後淨利成長率%	每股營業額 (元)		
經常淨利成長率%	每股營業利益 (元)		
常續性利益成長率 (稅後)	每股稅前淨利 (元)		
總資產成長率%			
淨值成長率%			
折舊性固定資產成長率%			
總資產報酬成長率%			
內部保留比率%			

參考文獻

- 呂美慧 (1999), 《銀行授信評等模式 – Logistic Regression 之應用》, 國立政治大學金融學系碩士論文。
- 陳宗豪 (1999), 《消費者小額信用貸款之信用風險研究 – 甄選的觀點》, 國立中山大學人力資源管理研究所碩士論文。
- 江世傑 (2001), 《模糊類神經網路在消費性貸款之應用》, 國立成功大學工業管理學系碩士論文。
- 李明峰 (2001), 《銀行業對企業授信『信用評等表』財務比率預警有效性之實證分析》, 國立中山大學財務管理學系研究所碩士論文。
- 何貴清 (2001), 《消費者小額信用貸款之信用風險研究 – 以一商業銀行客戶為例》, 國立中山大學人力資源管理研究所碩士論文。
- 林建州 (2001), 《銀行授信客戶違約機率之衡量》, 國立中山大學財務管理研究所碩士論文。
- 林妙宜 (2002), 《公司信用風之衡量》, 國立政治大學金融研究所碩士論文。
- 陳肇榮 (1983), 《運用財務比率預測財務危機之實證研究》, 政治大學企業管理研究所博士論文。
- 戴堅 (2004), 《個人消費性信用貸款授信評量模式之研究》, 國立中正大學國際經濟研究所碩士論文。
- 鍾經樊、黃嘉龍、謝有隆與黃博怡 (2005), 「台灣地區企業信用評分系統的建置、驗證、和比較」, 台灣經濟學會與北美華人經濟學會 2005 年聯合年會發表。
- 蒲小雷與韓家平 (2001), 《企業信用管理典範》, 中國對外經濟貿易出版社。
- 石新武 (2002), 《資信評估的理論和方法》, 經濟管理出版社。
- 孫國志與張炎培 (2005), 《企業信用調查實務》, 中國青年出版社。
- Altman, E. I. (1968), “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, *Journal of Finance*, 123, 4, 589-609.
- Banasik, J., J. Crook, and L. Thomas (2003), “Sample Selection Bias in Credit Scoring Models”, *Journal of the Operational Research Society*, 54, 822-832.
- Bluhm, C., L. Overbeck, and C. Wagner (2003), *An Introduction of credit Risk Modeling*, Boca Raton, FL: Chapman & Hall/CRC.

- Coats, P. and L. Fant (1993), "Recognizing Financial Distress Patterns Using a Neural Network Tool", *Financial Management*, 22, 3, 142-155.
- Frame, W. S., A. Srinivasan, and L. Woosley (2001), "The Effect of Credit Scoring on Small-Business Lending", *Journal of Money, Credit and Banking*, 33, 813-825.
- Hand, D. J. and W. E. Henley (1997), "Statistical classification Methods in Consumer Credit Scoring: A Review", *Journal of the Royal Statistical Society. Series A(Statistics in Society)*, 160, 523-541.
- Mays, E. (2004), *Credit scoring for risk managers: the handbook for lenders*, 1-155, Mason, Ohio: Thomson/South-Western.
- Mester, L. J. (1997), "What's the Point of Credit Scoring?", *Federal Reserve Bank of Philadelphia*, 3-16.
- Ross, S. L. (2000) "Mortgage Lending, Sample Selection and Default", *Real Estate Economics*, 28, 581-621.
- Neal Walters and Sharon Hermanson, (2001), "Credit Scores and Mortgage Lending", Research Report, AARP Public Policy Institute.
- Ohlson J. M., (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, 18, 1, 109-131.
- Padhi, M. s., L. W. Woosley, and A. Srinivasan (1998), "Credit Scoring and Small Business Lending in Low- and Moderate- Income communities", Federal Reserve Board of Atlanta.
- Royston, P. (1991), gr6: Lowess smoothing. *Stata Technical Bulletin* 3: 7-9. Reprinted in *Stata Technical Bulletin Reprints*, vol. 1, 41-44.