

探討壽險業優良顧客留存之實證研究

黃建丰

國立高雄應用科技大學企業管理系碩士在職專班研究生

f547945@yahoo.com.tw

葉惠忠

國立高雄應用科技大學企業管理系副教授

hcye@cc.kuas.edu.tw

摘要

人壽保險的特質是以「人」為本的事業，人是無價之寶，無法以金錢來衡量其價值；壽險業若無顧客，則其資產毫無價值，因此，壽險業的主要任務在吸引顧客與留下顧客，藉由具有競爭力的優勢商品或具差異化的服務吸引新顧客，並透過滿足顧客需求，提供大於期望的貼心服務，以有效提升顧客忠誠度與留存率。

展望二十一世紀顧客的型態已改變，市場經營也改變在市場激烈競爭的情況下，保險業對市場經營導向也產生重大變革，由商品導向演變為顧客導向。據此本研究以某樣本保險公司之 1122 位保戶，做為研究對象，在不同年齡、性別、婚姻、學歷、投保年資、職業、家庭背景等基本資料，保單狀態及售後服務等行為資料，透過資料庫行銷、資料探勘及顧客關係管理，來探討顧客留存率，其中運用區別分析及決策樹檢定法等統計方法，來了解顧客關係維護與優良顧客留存率之間的相關性作實證探討，以為日後業界此項方針的依據。

關鍵字：資料庫行銷、資料探勘、顧客關係管理、顧客留存、區別分析、決策樹

Keywords : Database marketing, Data mining, Customer relationship management, Customer retention, Discriminant, Decis tree

壹、緒論

一、研究動機

人類邁入 21 世紀，臺灣人身保險業，從民國 51 年政府准許民營保險公司成立，已經走過 40 年，在這新紀元時代人身保險業已然產生劇烈變化。隨著「金控法」之訂定實施，許多銀行為爭取更多的業務紛紛成立銀行保險經紀人/代理人部門，從事保險招攬業務的拓展，所以銀行保險業務有了蓬勃發展的現象（截至 97 年 12 月銀行保險首年度保費收入占率為 47.84%），人身保險業受到前所未有的衝擊。更由於保險市場國際化與自由化，21 世紀的保險市場的競爭與挑戰已趨向白熱化，故保險業者除不斷在銷售通路與商品開發積極改變及創新外，在顧客留存的議題下，如何防止顧客流失減少營運成本，增加公司的獲利方面，更是不遺餘力。根據統計資料顯示獲取一位新顧客的成本是維繫一位老顧客 6 倍，甚至 Geller (1997) 說是 5 倍或是 10 倍。因此，保險業者應該要特別注意老顧客的留存，期能使經營績效提高，以防止舊客戶的流失，造成經營成本上的大量損失。

近年來人身保險業者均面臨吸引新客戶成本提高及舊客戶嚴重流失的雙重難題，不論是從成本或獲利的角度來剖析，或是客戶因服務的不滿意及人情包袱而流失，都可清楚了解降低客戶流失率及維繫既有客戶，不僅是各家業者永續經營的關鍵，也是提昇競爭力及長期獲利的策略方向。再者，服務產業尤其是金融保險業在臺灣的市場，未來的挑戰不只是價格與獲利，而是服務品質，且是認知價值大於商品價值的服務。因此，本研究將以某保險公司 1122 位保戶做為研究對象，利用區別分析 (Discriminant)、決策樹 (Decision Tree)，探討顧客保留與年齡、性別、婚姻、學歷、年收入、職業、家庭背景等基本資料，保單狀態、售後服務、顧客忠誠及顧客滿意等行為資料，來了解顧客關係維護與優良顧客保留之間的相關性，作為實證探討。

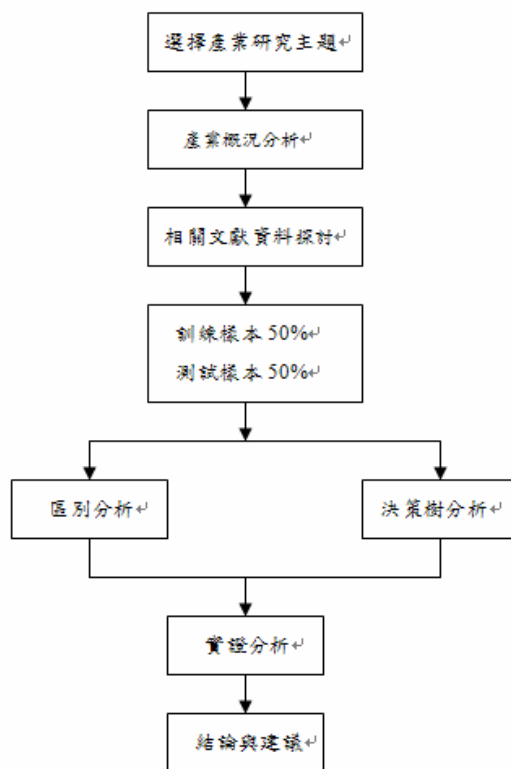
二、研究目的

由於可知，顧客保留是企業能否持續成長的重要因素，亦是能否永續發展的關鍵。因此，發掘並了解影響顧客保留的因素，對企業而言，是非常重要的議題，這也是本研究的目的。

本研究的主要目的如下：

1. 藉由保戶投保資料庫的資料，透過資料探勘、資料庫行銷及顧客關係管理，然後產生顧客忠誠及顧客留存進行實證分析，亦即結合人口統計基本資料變數與行為資料變數建立顧客留存的評估模式，期能找出可預測優良客戶的特徵。
2. 運用區別分析與決策樹分析等研究方法，建立適合的顧客留存的評估模式，供保險公司找出高留存度的優良顧客，以便利保險業者對優良客戶的管理，並藉此找到目標客戶群，長期用心經營，以增加利潤降低風險。

三、研究流程



一、我國人身保險事業概述

(一)、人身保險的功能

帶保險制度乃人類所創建的一種「社會經濟制度」，且是國家經濟發展的根本，人身保險業的主要精神所在就是幫助千千萬萬的個人，在發生意外或特定事故的時候，做為經濟生活的保障，保險業對個人所提供的保障，包括生命的保障、財產身家的保障，這也可以說是保險的功能。我們如果進一步探究的話，便會發現人身保險不僅對個人有好處，對社會、國家也有相當的功能；不僅是一種經濟制度，也是一種社會安全保障制度。對國家社會的助益除了穩定經濟制度安定政治的貢獻，保險業所累積「責任準備金」，可提供國家各項經濟建設大量資金，是動經濟成長，增加國富的因子。

(二)、臺灣人身保險事業之概況

由表 2-1 得知，目前臺灣人身保險事業之概況。

表 2-1 民國 99 年 9 月前核准成立並營業之壽險公司

公司名稱	備註欄
臺銀人壽	原中央信託局(人壽保險處)，民國 96 年 7 月與臺灣銀行合併，以臺灣銀行為存續公司，民國 97 年 1 月該人壽保險處改設立為股份有限

	公司
台灣人壽	民國 87 年 7 月改制民營。
國泰人壽	
中國人壽	原為華僑人壽，民國 73 年更名；民國 96 年 10 月概括承受瑞士商環球瑞泰人壽；98 年併購保誠人壽台灣主要資產與營運業務。
南山人壽	
國華人壽	
新光人壽	
國寶人壽	
三商美邦人壽	原為三商人壽，民國 90 年更名。
朝陽人壽	原興農人壽 99 年 9 月變更。
幸福人壽	
遠雄人壽	原為中興人壽，民國 89 年更名；民國 93 年概括承受瑞士商蘇黎世人壽。
宏泰人壽	原為宏福人壽，民國 88 年更名。
安聯人壽	原為統一人壽，民國 88 年更名為統一安聯人壽，民國 96 年 7 月更名為安聯人壽。
保德信國際人壽	原為美商保德信人壽台灣分公司，民國 90 年改制為子公司。
全球人壽	原為美商投資者人壽，民國 83 年更名為美商全球人壽，87 年由荷商亞太全球人壽概括承受，民國 90 年概括承受美商全球人壽及澳商安盛國衛人壽，並改制為子公司。
國際紐約人壽	原為美商紐約人壽台灣分公司，民國 91 年改制為子公司。
大都會國際人壽	原為大都會人壽台灣分公司，民國 92 年改制為子公司
中華郵政(壽險處)	民國 96 年 2 月更名為臺灣郵政，民國 97 年 8 月回復為原中華郵政。
富邦人壽	原為美商美國安泰人壽台灣分公司，民國 90 年與美商喬治亞人壽合併稱 ING 安泰人壽台灣分公司，民國 95 年 3 月改制為子公司，民國 98 年 6 月與富邦人壽合併並更名為富邦人壽。
第一金人壽	原為第一英傑華人壽，99 年 01 月變更
康健人壽台灣分公司	原為美商康健人壽，95 年 10 月改由紐西蘭商康健人壽概括承包。
友邦人壽台灣分公司	原為美商美國人壽台灣分公司，民國 98 年 6 月移轉更名為英屬百慕達商友邦人壽台灣分公司。
宏利人壽台灣分公司	原為美商宏利人壽，民國 94 年 1 月改隸屬英屬百慕達商友邦人壽台灣分公司。
法國巴黎人壽台灣分公司	原為佳迪福人壽，民國 96 年 3 月改由法商法國巴黎人壽概括承受。
安達保險台灣分公司	
中泰人壽台灣分公司	
匯豐人壽台灣分公司	
蘇黎世國際人壽台灣分公司	
合作金庫人壽	民國 99 年 4 月成立
保誠人壽	

(資料來源：中華民國人壽保險商業同業公會)

二、顧客保留之相關理論

市場佔有率可以視為顧客保留，Gerpott, Rams, and Schindler, (2001) 認為顧客保留包含「行為」與「態度」兩種立場，所以將顧客保留定義為維持顧客與供應商的關係，主要經由兩種方式達成，其一為後續的購買，或是與供應商建立長期的契約關係，這是屬於行為的表現；另一則是未來向廠商再購買的意願，或是放契約關係的想法，這則是態度的表現。(Berne Mugicaand and Yague, 2001) 也認顧客保留包括「行為」與「態度」兩種因素，說明顧客留住的行為以及比較競爭者並加以選擇的態度。且 (Aspinall, Clive and Merlin 2001) 利用電話訪談調查顧客保留的定義。學者一致認為顧客保留可分為行為和態度兩方面，但以行為面為多數。其調查結果如表 2-7 所示：

表 2-2 顧客保留的定義

定義的項目	%
與顧客保持良好關係	23%
再購或更新	11%
回應企業活動	6%
滿意的態度	17%

資料來源：Aspinall, Clive and Merlin (2001)

(一)、顧客保留之定義

顧客流失對企業的經營成本將造成嚴重的損失，以服務產業的獲利來說，顧客保留較規模經濟、市場佔有率、單位成本等競爭優勢更有效果，透過成本的節省、購買量的增加溢價與口碑效果，當減少 5% 的顧客流失率，將可增加 25% 到 85% 獲利。(Reichheld and Sasser, 1990)

三、顧客關係管理

(一)、顧客關係管理的定義

何謂「顧客關係管理」(CRM)? 學者 R. S. Swift (2000) 對於 CRM 的詮釋為 CRM 是企業藉由與顧客充分的互動，來了解及影響顧客的行為，以提昇顧客的贏取率，顧客的留住率(Customer Retention)，顧客的忠誠度及顧客獲利率的一種經營模式。Tiwana (2000) 認為 CRM 是企業從各種不同的角度來了解及區別顧客，以發展出適合顧客個別需要的產品服務的一種企業程序與資訊科技的組合模式。其目的在用以管理與老顧客的關係，以使他們達到最高的忠誠度、留住率與利潤貢獻度，並同時有效率選擇性的吸引好的新顧客。CRM 也被解釋為持續性的關係行銷；吸引好的新顧客，當一個新顧客進來後就保持持續性的互動與服務維持良好的關係，將新客戶變成舊客戶則此客戶對企業的價值就大為增加了；其實也是顧客關係管理的一環。

四、資料探勘

(一)、資料探勘定義與模式

資料探勘 (Data Mining) 就是從資料中發掘資訊或知識，資料探勘泛指各式各樣從大量資料裡歸納出規則的方法，有人稱為「資料考古學」(Data Archaeology)、「資料樣型分析」(Data Pattern Analysis)、或是「功能相依分析」(Functional Dependency Analysis)。資料探勘的主要目標是從大量的資料中找出有用的隱藏資訊，這些資訊事先並不知道，但可以提供有用的決策參考。

五、資料庫行銷

資料庫行銷的真正意涵在於藉由行銷研究的統計模型，來分析與處理大量的顧客資料，以作為行銷決策之參考依據，不同於傳統的行銷活動以「交易」為核心，資料庫行銷是以「顧客」為核心所發展出來的行銷策略。資料庫行銷不只是企業對顧客的單向溝通管道，而是企業與顧客之間形成持續性互動的學習性關係。當該顧客與企業產生互動時，企業的窗口便可以立即地得到之前與該名顧客互動的所有記錄，並提出更符合顧客需求的解決方案。我們利用資料庫行銷可以鎖定每一位顧客，針對特定的顧客在特定的時間提供所需的產品，提高行銷活動的達成率，降低不必要行銷成本的浪費，因而提高公司的獲利。

(一)、資料倉儲(Data Warehouse)

Data Warehouse 是將顧客的所有的資料，包括年齡、性別、興趣、過去交易資料、貢獻價值等作整合。因此，Frawley and Thearling (1999) 將資料倉儲定義為為了達成精確性，資料倉儲把許多經由操作和外部得來的資料聯合匯整，統一了顧客和企業觀點的資料寶庫。

(二)、資料採礦 (Data Mining)

顧客關係管理藉由互動關係來蒐集顧客資料。一般而言，顧客關係管理蒐集之顧客資料包括行銷活動的顧客反應度、運銷和產品供應之相關數據資料、銷售與購買之資料、客戶資料、相關服務之數據、產品市場資料及網路銷售數據等。由於各項資料均為顧客與企業間之互動而產生，因此藉由分析顧客之各項資料，可有效掌握顧客特性。而分析資料的主要工具之一即為資料採礦。

參、研究方法

一、研究對象

(一)、樣本資料來源

本研究係以國內某一大型保險公司地區性，特定業務員成交之被保險人為研究對象，母體資料中之保戶居住地區，以台南縣、台南市、高雄縣、高雄市、屏東縣佔大部分，小部分分佔台灣省各縣市。個案保險公司從81年8月所生效之保單契約開始，截至99年10月底止，總計有效契約數為1,122張，經樣本抽取篩選，符合有效樣本條者共計1,122件，本研究區分為訓練樣本（總樣本數50%，計556件）、策試樣本（總樣本數50%，計566件），全數作為本研究樣本，以供本研究作實證探討使用。

上述樣本為自81年8月1日至99年10月31日止，採礦於公司資料庫並經長期觀察擷取的保單記錄，截至99年10月31日止，保單生效日起，續期保費經催告期61天以後終止繳費，保單繼續率未達持續繳費5年以上為不正常契約，持續繳納保險費者為正常契約。

(二)、研究變數說明

本研究之變數選取是以基本資料變數及行為資料變數為主，基本資料變數是依據顧客填寫要保書所記載各項人口統計資料，行為資料變數則是保單狀態，如繳費方式、繳費能力、再購頻率、客服資料（如理賠經驗）、信用資料、購買需求產品種類及偏好，等消費行為資料。

(三) 變數操作性定義：

表 3-1 變數操作性定義

序號	變數名稱	變數定義
1	保單狀態(Y)	離散型資料:1-有效, 2-失效
2	性別(X1)	離散型資料:1-男, 2-女
3	年齡(X2)	離散型資料:1-20歲以下, 2-21~30歲, 3-31~40歲, 4-41~50歲, 5-51歲以上
4	教育程度(X3)	離散型資料:1-高中職(含)以下, 2-大專(學), 3-研究所(含)以上
5	婚姻狀態(X4)	離散型資料:1-已婚, 2-單身
6	職業(X5)	離散型資料:1-軍公教, 2-工, 3-服務業 4-商, 5-其他
7	年收入(X6)	連續型資料:1-30萬(含)以下, 2-31~50萬, 3-51~80萬, 4-81~100萬, 5-101萬以上
8	子女人數(X7)	離散型資料:1-0人, 2-1人, 3-2人, 4-3人, 5-4人,
9	與顧客認識多久(X8)	離散型資料:1-1年以內, 2-3年以內, 3-5年以內, 4-10年以內, 5-10年以上
10	居住地區(X9)	離散型資料:1-都會區, 2-鄉村地區, 3-偏遠地區
11	房屋貸款(X10)	離散型資料:1-是, 2-否
12	購買多家保單(X11)	離散型資料:1-是, 2-否
13	理賠經驗(X12)	離散型資料: 1-滿意, 2-抱怨
14	再購記錄(X13)	離散型資料:1-是, 2-否
15	推薦親朋好友(X14)	離散型資料:1-是, 2-否:
16	繳費方式(X15)	離散型資料:1-月繳, 2-季繳, 3-半年繳, 4-年繳, 5-躉繳
17	年繳化保費(X16)	連續型資料:1-1~3萬, 2-3.1~6萬 3-6.1~8萬, 4-8.1~10萬, 5-10.1萬以上
18	繳費管道(X17)	離散型資料:1-收費員, 2-銀行轉帳, 3-信用卡, 4-自行劃撥

19	保單貸款(X18)	離散型資料:1-是,2-否
----	-----------	---------------

上表中人口統計基本資料變數字有 7 項，行為變數有 11 項。

二、研究設計

本研究旨在建立壽險業優良顧客留存評估模型，希望達到挑選優良顧客的目標，以提供壽險業未來在經營策略上，針對顧客進行價值分析，迅速對該顧客作出判斷是否為目標客群。本研究有效樣本數為 11,22 件（訓練樣本 50%；測試樣本 50%）。步驟一為利用交叉分析對全體樣本進行樣本結構分析；步驟二利用訓練樣本以人口統計基本資料變數來建立留存評估模式；步驟三為利用測試樣本加入所有研究變數套入步驟二所建立的模型中作驗證，除檢驗模型的預測能力，也避免在選取樣本時可能發生的偏差或忽略可用性的重要因子，而造成模型在實際應用時產生落差。

本研究以優良顧客留存為目標，研究方法以區別分析（Discriminate）與決策樹（Decision Tree）兩種統計方法進行。所以在模型建構之初會先進行資料整理，並經過資料轉換再進行分析，現在就將兩種研究方法作一簡單介紹。

（一）區別分析

區別分析主要目的是了解群體的差異，它先利用區別變數建立區別規則（區別函數），然後再由區別規則對個體做分類，預測每個個體屬於各群組的可能機率。

1.區別分析的基本概念

區別分析是一種判別和分類的技術。是由一個分類變數（grouping variable）當因變數，與多個計量的區別變數（discriminate variable）當自變數的技術。此技術會建立此分類變數與其他多個計量變動間的對應關係從而建立判別函數（discriminate function）。其基本模型線性函數關係式如下：

$$Y=b_0+b_1x_1+b_2x_2+\dots+b_ix_i+\dots+b_kx_k$$

其中 Y 是判別函數； b_0 為常數； x_i 為是各區別變數； b_i 為相對應的判別函

（二）、決策樹（Decision Tree）

決策樹為資料探勘技術的方法之一，主要的作用在於從已知且分類的資料中，利用樹狀分枝的架構來產生規則。決策樹是一種決策過程的圖形，在決策過程中由許多不同的行動方案和不同的發生狀況或事件交互配合而成，其結果如樹狀分析圖，故稱決策樹。由文獻中得知常用的決策樹演算法為 CHAID、Exhaustive CHAID CRT 等。

其中 CHAID 是一個可用來建立“非二元樹”的一個演算法，因此，適用於分類超過 2 個以上的個案資料。CHAID 對較大型資料庫亦相當適用，尤其適合用於作市場行銷或市場區隔等研究分析用途。故本研究選擇建立決策樹分析中的 CHAID 建立選擇優良顧客留存的。

肆、實證分析結果

一、樣本結構交叉分析

1.性別

表 4-1 與性是否為有效保單別之交叉分析表

性別 是否為有效保單		男性	女性	總和
		有效保單	個數	351
	百分比	47.6%	52.4%	100.0%
失效保單	個數	182	202	384
	百分比	47.4%	52.6%	100.0%

總和	個數	533	589	1122
	百分比	47.5%	52.5%	100%
卡方值=0.003 P=0.958				

2.年齡

分佈表 4-2 是否為有效保單與年齡交叉分析表

是否為有效保單		年齡					總和
		20歲以下	21-30歲	31-40歲	41-50歲	51歲以上	
有效保單	個數	119	121	135	174	189	738
	百分比	16.1%	16.4%	18.8%	23.6%	25.1%	100.0%
失效保單	個數	33	35	103	116	97	384
	百分比	8.6%	9.1%	26.8%	30.2%	25.3%	100.0%
總和	個數	152	156	238	290	286	1122
	百分比	13.5%	13.9%	21.2%	25.9%	25.5%	100.0%
卡方值=33.178 P=0.000***							

3.教育程度

表 4-3 是否為有效保單與教育程度交叉分析表

是否為有效保單		教育程度			總和
		高中(職)以下	大專(學)	研究所以上	
有效保單	個數	332	346	60	738
	百分比	45.0%	46.9%	8.1%	100.0%
失效保單	個數	206	166	12	384
	百分比	53.6%	43.2%	3.1%	100.0%
總和	個數	538	512	72	1122
	百分比	48%	45.6%	6.4%	100%
卡方值=14.549 P=0.001**					

4.婚姻狀況

表 4-4 是否為有效保單與婚姻狀況交叉分析表

是否為有效保單		婚姻狀況		總和
		已婚	單身	
有效保單	個數	449	289	738
	百分比	60.8%	39.2%	100.0%
失效保單	個數	285	99	384
	百分比	74.2%	25.8%	100.0%
總和	個數	734	388	1122
	百分比	65.4%	34.6%	100%
卡方值=19.984 P=0.000***				

5.職業

表 4-5 是否為有效保單與職業交叉分析表

是否為有效保單		職業					總和
		軍公教	工	服務業	商	其他	
有效保單	個數	189	54	114	101	280	738
	百分比	25.6%	7.3%	15.4%	13.7%	37.9%	100.0%
失效保單	個數	70	53	65	48	148	384
	百分比	18.2%	13.8%	16.9%	12.5%	38.5%	100.0%
總和	個數	259	107	179	149	428	1122
	百分比	23.1%	9.5%	16%	13.3%	38.1%	100.0%

卡方值=17.737 P=0.001**

6.年收入

表 4-6 是否為有效保單與年收入交叉分析表

年收入		30 萬	31-50 萬	51-80 萬	81-100 萬	101 萬以上	總和
		是否為有效保單					
有效保單	個數	0	24	196	317	201	738
	百分比	.0%	3.3%	26.6%	43.0%	27.2%	100.0%
失效保單	個數	1	23	152	163	45	384
	百分比	0.3%	6.0%	39.6%	42.4%	11.7%	100.0%
總和	個數	1	47	348	480	246	1122
	百分比	0.1%	4.2%	31.0%	42.8%	21.9%	100.0%

卡方值=48.009 P=0.000***

7.子女人數

表 4-7 是否為有效保單與子女人數交叉分析表

子女人數		0 人	1 人	2 人	3 人	3 人以上	總和
		是否為有效保單					
有效保單	個數	282	79	258	107	12	738
	百分比	38.2%	10.7%	35.0%	14.5%	1.6%	100.0%
失效保單	個數	103	53	174	50	4	384
	百分比	26.8%	13.8%	45.3%	13.0%	1.0%	100.0%
總和	個數	385	132	432	157	16	1122
	百分比	34.3%	11.8%	38.5%	14.0%	1.4%	100.0%

卡方值=19.637 P=0.001**

8.與顧客認識多久

表 4-8 是否為有效保單與與顧客認識多久交叉分析表

認識多久		1 年以內	3 年以內	5 年以內	10 年以內	10 年以上	總和
		是否為有效保單					
有效保單	個數	0	8	23	241	466	738
	百分比	.0%	1.1%	3.1%	32.7%	63.1%	100.0%
失效保單	個數	26	38	53	142	125	384
	百分比	6.8%	9.9%	13.8%	37.0%	32.6%	100.0%
總和	個數	26	46	76	383	591	1122

9.居住地區

表 4-9 是否為有效保單與居住地區交叉分析表

居住地區		都會地區	鄉村地區	偏遠地區	總和
		是否為有效保單			
有效保單	個數	317	415	6	738
	百分比	43.0%	56.2%	0.8%	100.0%
失效保單	個數	145	238	1	384
	百分比	37.8%	62.0%	0.3%	100.0%
總和	個數	462	653	7	1122
	百分比	41.2%	58.2%	0.6%	100.0%

卡方值=4.324 P=0.115

10.房屋貸款

表 4-10 是否為有效保單與房屋貸款交叉分析表

房屋貸款		是否為有效保單		總和
		是	否	
有效保單	個數	174	564	738
	百分比	23.6%	76.4%	100.0%
失效保單	個數	195	189	384
	百分比	50.8%	49.2%	100.0%
總數	個數	369	753	1122
	百分比	32.9%	67.1%	100.0%
卡方值=84.689 P=0.000***				

11.購買多家保單

表 4-11 是否為有效保單與購買多家保單交叉分析表

購買多家保單		是否為有效保單		總和
		是	否	
有效保單	個數	156	582	738
	百分比	21.1%	78.9%	100.0%
失效保單	個數	92	292	384
	百分比	24.0%	76.0%	100.0%
總和	個數	248	874	1122
	百分比	22.1%	77.9%	100.0%
卡方值=1,167 P=0.280				

12.理賠經驗

表 4-12 是否為有效保單與理賠經驗交叉分析表

理賠經驗		是否為有效保單		總和
		滿意	抱怨	
有效保單	個數	730	8	738
	百分比	98.9%	1.1%	100.0%
失效保單	個數	250	134	384
	百分比	65.1%	34.9%	100.0%
總和	個數	980	142	1122
	百分比	87.3%	12.7%	100%
卡方值=261.218 P=0.000***				

13.再購記錄

表 4-13 是否為有效保單與再購記錄交叉分析表

再購記錄		是否為有效保單		總和
		是	否	
有效保單	個數	471	267	738
	百分比	63.8%	36.2%	100.0%
失效保單	個數	17	367	384
	百分比	4.4%	95.6%	100.0%
總和	個數	488	634	1122
	百分比	43.5%	56.5%	100.0%
卡方值=362.541 P=0.000***				

14.推薦親朋好友

表 4-14 是否為有效保單與是否推薦親朋好友交叉分析表

是否為有效保單		推薦親朋好友		總和
		是	否	
有效保單	個數	328	410	738
	百分比	44.4%	55.6%	100.0%
失效保單	個數	5	379	384
	百分比	1.3%	98.7%	100.0%
總和	個數	333	789	1122
	百分比	29.7%	70.3%	100.0%

卡方值=225.251 P=0.000***

15.繳費方式

表 4-15 是否為有效保單與繳費方式交叉分析表

是否為有效保單		繳費方式					總和
		月繳	季繳	半年繳	年繳	躉繳	
有效保單	個數	122	36	40	533	7	738
	百分比	16.5%	4.9%	5.4%	72.2%	0.9%	100.0%
失效保單	個數	118	19	14	232	1	384
	百分比	30.7%	4.9%	3.6%	60.4%	0.3%	100.0%
總和	個數	240	55	54	765	8	1122
	百分比	21.4%	4.9%	4.8%	68.2%	0.7%	100.0%

卡方值=32.298 P=0.000***

16.年繳化保費

表 4-16 是否為有效保單與年繳化保費交叉分析表

否為有效保單		年繳化保費					總和
		1-3 萬	3.1-6 萬	6.1-8 萬	8.1-10 萬	10.1 萬以上	
有效保單	個數	324	290	78	20	26	738
	百分比	43.9%	39.3%	10.6%	2.7%	3.5%	100.0%
失效保單	個數	234	117	18	8	7	384
	百分比	60.9%	30.5%	4.7%	2.1%	1.8%	100.0%
總和	個數	558	407	96	28	33	1122
	百分比	49.7%	36.3%	8.6%	2.5%	2.9%	100.0%

卡方值=33.254 P=0.000***

17.繳費管道

表 4-17 是否為有效保單與繳費管道交叉分析表

繳費管道		是否為有效保單				總和
		收費員	銀行轉帳	信用卡	自行劃撥	
有效保單	個數	175	289	72	202	738
	百分比	23.7%	39.2%	9.8%	27.4%	100.0%
失效保單	個數	71	32	34	247	384
	百分比	18.5%	8.3%	8.9%	64.3%	100.0%
總和	個數	246	321	106	449	1122
	百分比	21.9%	28.6%	9.4%	40.0%	100.0%

卡方值=173.435 P=0.000***

18.保單貸款

表 4-18 是否為有效保單與保單貸款交叉分析表

保單貸款		是否為有效保單		
		是	否	總和
有效保單	個數	31	707	738
	百分比	4.2%	95.8%	100.0%
失效保單	個數	78	306	384
	百分比	20.3%	79.7%	100.0%
總和	個數	109	1013	1122
	百分比	9.7%	90.3%	100.0%

卡方值=74.755 P=0.000***

綜合以上分析，可以看出 7 個人口統計基本資料變數中，「年齡」、「教育程度」、「婚姻狀況」、「職業」、「年收入」、「子女人數」等六項具顯著性，而 11 個行為資料變數，以「與顧客認識多久」、「房屋貸款」、「理賠經驗」、「再購記錄」、「推薦親朋好友」、「繳費方式」、「年繳化保費」、「繳費管道」、「保單貸款」等 9 項較具顯性，所以初步分析結果，行為資料變數較人口統計基本資料變數有顯著性，就比率上及項目個數方面均以行為資料變數較佳。

二、實證結果

(一) 逐步區別分析模型的建立

1. 利用人口統計基本資料變數建立模型：

首先在訓練樣本資料中輸入 7 項人口統計基本資料變數，進行逐步區別分析，利用 F 值與 F 機率值「選入」、「刪除」條件，篩選新變數，對貢獻不大的變數逐一刪除，直到模型內變數都達到要求的條件為止，經以逐步區別分析統計找出顯著變數，依序為年收入、婚姻狀況、教育程度等三個區別能力較佳的變數

判別函數：判別是否為有效保單的判別函數：X3 代表教育程度，X4 代表婚姻狀態，X6 代表年收入。舉例如下：

例 1. 現有一被保險人甲，學歷為大專以上，婚姻狀態為單身年收入在 101 萬以上，則依 101 萬以上，則

$$Y = a + b_3 * X_3 + b_4 * X_4 + b_6 * X_6$$

其判別是否為有效保單的判別函數

$$Y_1 = -17.951 + 2.953 * X_3 + 6.906 * X_4 + 4.961 * X_6$$

$$Y_2 = -14.586 + 2.607 * X_3 + 5.967 * X_4 + 4.343 * X_6$$

$$Y_1 = -17.951 + 2.953 * 2 + 6.906 * 2 + 4.961 * 5 = 26.572$$

$$Y_2 = -14.586 + 2.607 * 2 + 5.967 * 2 + 4.343 * 5 = 24.277$$

因為 $Y_1 > Y_2$ ，故將此被保險人判為第一類是有效保單顧客，依此條件，保險公司可以根據上述方程式判別是否為有效保單之優良顧客的依據。

(1) 事前機率：在是否為有效保單作統計分類時，已知樣本中有效保單明顯多於失效保單(385/171)，故觀察值以「依據組別大小計算」，其值為有效保單(385/556=0.692)；失效保單(171/556=0.308)；與樣本中該保險公司的標準值(70%/30%)無顯著差異。

(2) 準確率：逐步區別分析訓練樣本556個，在是否為有效保單預測準確率如表4-19，總預測準確率為72.5%
 準確率 = (355 + 48) / 556 = 72.5%

型1錯誤(預測為有效保單，實則為失效保單) = 123/556 = 22.1%

型2錯誤(預測為失效保單，實則為有效保單) = 30/556 = 5.3%

表4-19 第一階段訓練樣本逐步區別分析模型判別結果分析表

是否為有效保單		預測的各組成員		總和
		有效保單	失效保單	
原始的	個數 有效保單	355	30	385
	失效保單	123	48	171
%	有效保單	92.2	7.8	100.0
	失效保單	71.9	28.1	100.0

a. 72.5% 個原始組別觀察值已正確分類。

上表為分類結果分析，其顯示有高達153個觀察值分類錯誤，由此可知「人口統計資料變數」對樣本的區別能力很弱，造成逐步區別的效果不佳，致產生過高錯誤率。

2. 利用人口統計基本資料變數和行為資料變數共同建立模型：

利用訓練樣本將所有18個變數套入上述分析所建立的模型中，同樣以逐步區別分析統計找出顯著變數，依序為「認識多久」、「保單貸款」、「再購記錄」、「房屋貸款」、「購買多家保單」、「理賠經驗」、「推薦親朋好友」、「年齡」、「年收入」、「繳費管道」10個項目。

(1) 判別函數：

判別是否為有效保單的判別函數：X2代表年齡，X6代表年收入，X8代表與顧客認識多久，X10代表房屋貸款，X12代表理賠經驗，X13代表再購記錄，X14代表推薦親朋好友，X17代表繳費管道，X18代表保單貸款。舉例如下：

例2. 現有一被保險人乙，其年齡為30歲，年收入在101萬以上，與服務人員交情10以上，沒有房貸與保單貸款，亦無友家保單，曾經申請理賠感到滿意，連續3年加購保單，又常推薦親朋好友，繳費管道為銀行轉帳，則

$$Y = a + b_2 * X_2 + b_6 * X_6 + b_8 * X_8 + b_{10} * X_{10} + b_{11} * X_{11} + b_{12} * X_{12} + b_{13} * X_{13} + b_{14} * X_{14} + b_{17} * X_{17} + b_{18} * X_{18}$$

其判別是否為有效保單的判別函數：

$$Y_1 = -95.195 + 2.048 * X_2 + 5.357 * X_6 + 8.133 * X_8 + 7.642 * X_{10} + 11.732 * X_{11} + 12.401 * X_{12} + 4.712 * X_{13} + 11.204 * X_{14} + 1.092 * X_{17} + 25.458 * X_{18}$$

$$Y_2 = -87.485 + 2.258 * X_2 + 4.926 * X_5 + 6.918 * X_6 + 5.770 * X_{10}$$

$$+ 11.076 * X_{11} + 14.978 * X_{12} + 6.413 * X_{13} + 11.986 * X_{14} + 1.387 * X_{17} + 21.966 * X_{18}$$

$$Y_1 = -95.195 + 2.048 * 2 + 5.357 * 5 + 8.133 * 5 + 7.642 * 2$$

$$+ 11.732 * 2 + 12.401 * 1 + 4.712 * 1 + 11.204 * 1 + 1.092 * 2$$

$$+ 25.458 * 2 = 96.516$$

$$Y_2 = -87.485 + 2.258 * 2 + 4.926 * 5 + 6.918 * 5 + 5.770 * 2 \\ + 11.076 * 2 + 14.978 * 1 + 6.413 * 1 + 11.986 * 1 + 1.387 * 2 \\ + 21.966 * 2 = 89.971$$

在此模型中亦是 $Y_1 > Y_2$ ，故將被保險人乙判定為第一類為有效保單之顧客。從結果可以顯示，在這兩個個案條件下保險公司可以根據上述方程式判別被保險人是否留存且為優良顧客的依據。

綜合以上二階段實證結果，本研究將訓練樣本與測試樣本的預測能力結果呈現於表4-20，由結果可觀察出在準確率的評估中，第一階段逐步區別分析，訓練樣本對於能正確的預測有效保單優良顧客的能力約略優於測試樣本，但由於在第一階段只加入「人口統計基本資料」其區別能力很弱，分類錯誤率過高，造成分類效果不佳的現象。但於第二階段的實證分析時加入「行為資料」變數後，錯誤率明顯下降，準確率皆大幅提升。

表4-20 是否為有效保單優良顧客逐步區別預測能力評估表

評估指標	階段	訓練樣本	測試樣本
準確率	第一階段	72.5%	72.1%
	第二階段	87.8%	87.8%

(二) 決策樹之實證結果

決策數之實證分析說明如下：

a. 利用人口統計資料變數建立模型：

在7個分割變數中，篩選出4個分割屬性，即「年收入」、「職業」、「婚姻狀況」、「教育程度」。底下的樹形圖以「年收入」分割能力最強；其分割效果的CHAID檢定可知，卡方值為29.922，P值達0.000，產生下列規則，分析如圖4-1，說明如下：

規則一：若有一被保險人，為有效保單且「年收入」大於100萬，職業在變數定義中為其他者，則將被歸納為優良顧客的條件。

規則二：若有一被保險人，為有效保單其「年收入」小於或等於100萬，婚姻狀況為單身，將被歸納為優良顧客的條件。

規則三：若有一被保險人，為有效保單其「年收入」小於或等於100萬，婚姻狀況為已婚，教育程度為是大專以上，將被歸納為優良顧客的條件。

規則四：若有一被保險人，為有效保單其「年收入」小於或等於100萬，婚姻狀況為單身，教育程度為是高中(職)以下，將被歸納為一般顧客的條件。

分類結果：

表4-21 第一階段決策樹分類結果分析表

觀察值	預測值		
	優良顧客	一般顧客	百分比 (%)
優良顧客	341	44	88.6%
一般顧客	103	68	39.8%
概要百分比 (%)	79.9%	20.1%	73.6%

註：成長方法：CHAID 依變數：是否為優良顧客

上表4-32為分類結果分析，其顯示有147位觀察值分類錯誤，分類正確率為73.6%。分類效果不佳，其原因乃決策樹分類方式是依層次分割不純的群體，不似區別分析是依直線的判別函數分割群體，決策樹的分類效果會受到分割規則與樣本特性等因素影響所致。

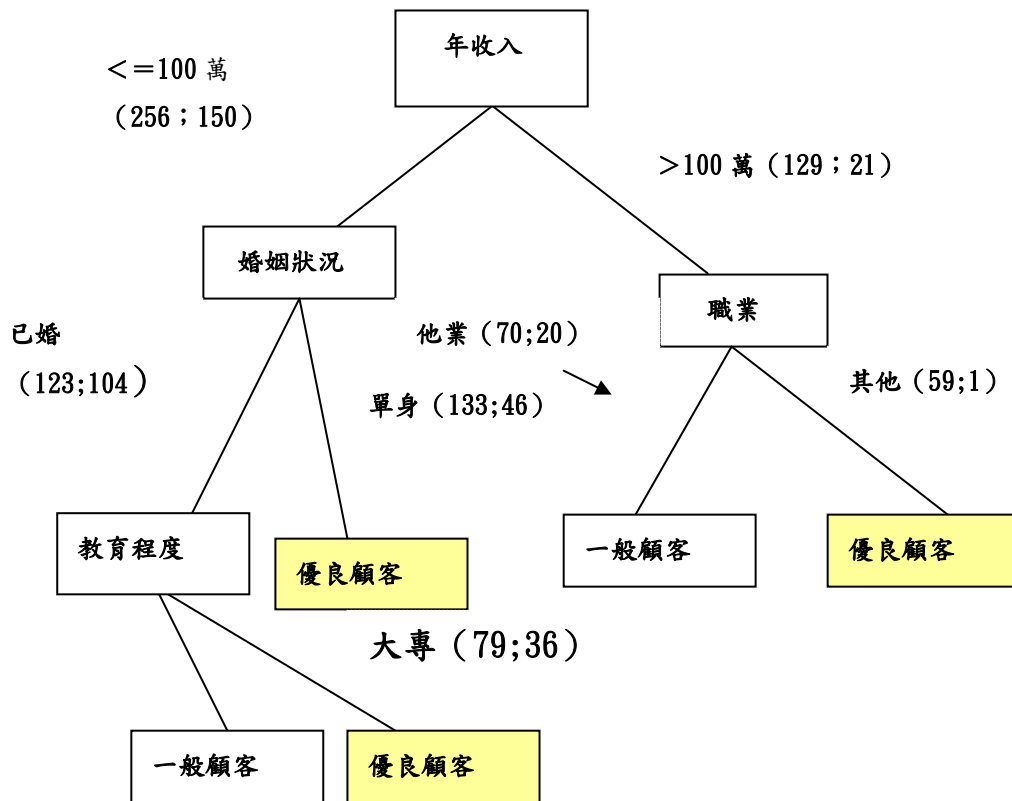


圖4-1第一階段決策樹分析圖

2. 利用人口統計變數與行為資料變數共同建立模型：

在18個分割變數中，篩選出5個分割屬性，即「再購記錄」、「房屋貸款」、「婚姻狀況」、「認識多久」、「保單貸款」。由圖4-1得知，以「再購記錄」分割能力最強，卡方檢定值146.491，P值為0.000，因此，產生下列規則，分析如圖4-2，說明如下：

規則一：若被保險人有「再購記錄」，但有「房屋貸款」，將被歸納為一般顧客。

規則二：若被保險人有「再購記錄」，無「房屋貸款」，「婚姻狀況」為已婚者，則被歸納為優良顧客。

規則三：若被保險人無「再購記錄」，認識年限等於或小於3年時，將被歸納為一般顧客。

規則四：若被保險人無「再購記錄」，認識年限大於3年時且無保單貸款，則被歸納為優良顧客。

分類結果：

表4-22第二階段決策樹分類結果分析表

觀察值	預測值		
	優良顧客	一般顧客	百分比 (%)
優良顧客	382	9	97.7%
一般顧客	74	101	57.7%
概要百分比 (%)	80.6%	19.4%	85.3%

由表4-22得知，在加入行為資料變數後，錯誤分類的數量雖明顯下降，但仍不盡理想，其總準確率雖也大幅提高，與逐步區別相較還是略遜一籌。

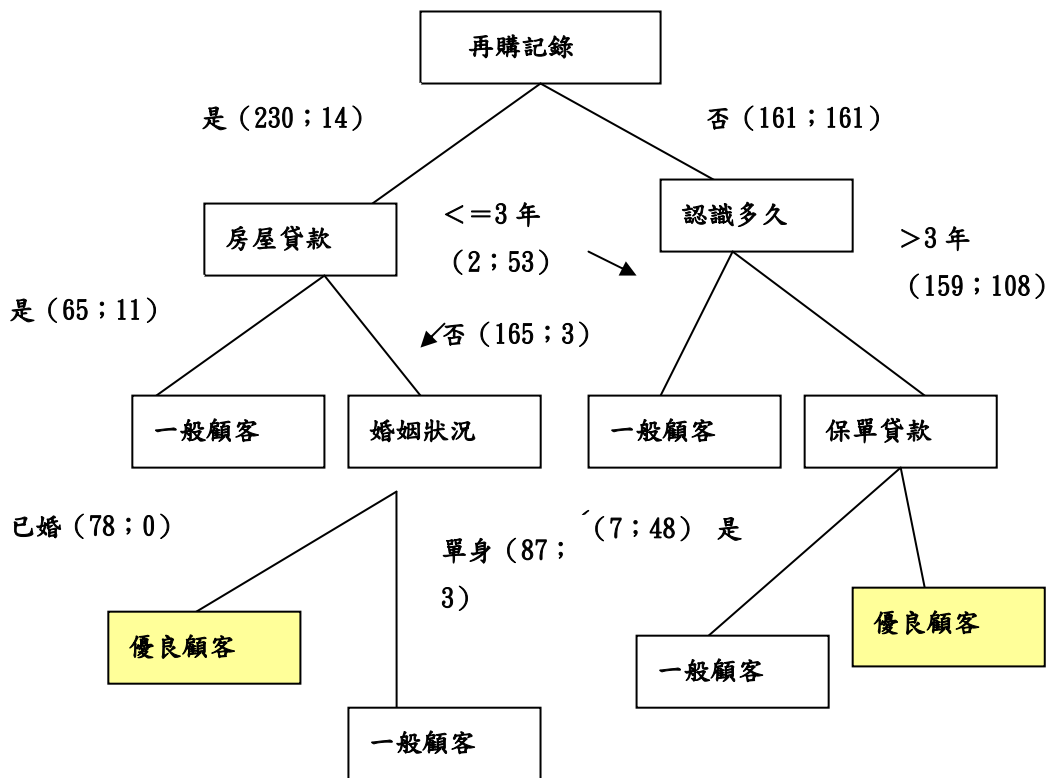


圖4-2第二階段決策樹分析圖

綜合以上二階段實證分析可以發現，決策樹在二階段的分析中，第一階段人口統計資料變數分析時，從7個預測變數裏產生了4個變數為分割屬性，而在第二階段加入行為資料變數時，所有18個變數產生5個變數為分割屬性，在第一階段中具有顯著性變數只有「婚姻狀況」變數仍然出現在第二階段為分割屬性中，這是由於「婚姻狀況」變數預測能力較佳所致。

決策樹方法的優點，在於可以生成可理解的規則；和一些資料探勘演算法比較之下，決策樹演算法產生的規則比較容易理解，而且決策樹模型建立的過程也很直觀。所以能根據決策樹產生的分類規則，作為預測保險業是否為有效保單優良顧客留存的依據。

茲將決策樹於訓練樣本與測試樣本在不同階段的測試評估結果製成表4-23，由結果可以觀察出準確率的評估中，第一階段決策樹在訓練樣本正確預測優良顧客的能力優於測試樣本，在第二階段則不分軒輊。

表 4-23 決策樹預測能力分析表

評估指標	階段	訓練樣本	測試樣本
準確率	第一階段	73.6%	73.5%
	第二階段	85.3%	85.3%

從以上兩種方法的實證分析結果得知，逐步區別分析準確率的評估與決策樹，在第一階段測試評估結果，訓練樣本略優於測試樣本，第二階段準確率則相等。就兩種方法的總準確率作比較，第一階段實證分析的總準確率，決策樹略優於逐步區別分析，但在第二階段的實證分析，則是逐步區別分析高於決策樹。因此，本研究最後決定以兩種方法交互運用，作為優良顧客留存的模式。最後結果也發現在第一階段作實證分析時，只有以人口統計資料作實證分析時，變數準確率不佳，第二階段加入行為資料變數後，不論是逐步區別分析或決策數分析，準確率都大。

伍、結論與建議

1. 一、結論本研究首先以交差分析進行樣本結構分析，得到結論為「年齡」、「教育程度」、「婚姻狀況」、「職業」、「年收入」、「子女人數」等6項人口統計資料變數，「與顧客認識多久」、「房屋貸款」、「理賠經驗」、「再購記錄」、「推薦親朋好友」、「繳費方式」、「年繳化保費」、「繳費管道」、「保單貸款」9項行為資料變數，共計15項變數具顯著性。
2. 以分析結果而論：兩種模式所得到的顯著變數抑或屬性，分別都涵蓋在人口統計資料變數與行為資料變數中，但行為資料變數顯著項較人口統計資料變數為多數，其中又以行為資料變數的加入使模型的預測能力更佳，與交叉分析的差異在於第一階段實證分析時，人口統計資料變數有「年收入」、「婚姻狀況」、「教育程度」3項具顯著性；而第二階段則是「與顧客認識多久」、「保單貸款」、「再購記錄」、「房屋貸款」4項，且兩種模式皆分別於第一階段與第二階段，都在這些變項中具顯著性。
3. 在變數的選取過程上，行為資料變數的預測能力明顯的比人口統計資料變數較佳。在研究方法的部分，逐步區別分析根據變數的重要性建立模型，作為決策依據；而決策樹有預測的優點，其預測的結果也以簡單的規則呈現，但在實證分析上，顯著變數的項目，也出現於逐步區別的分析中。由分類準確率可發現，決策樹通常與區別分析的準確率會約略相等，這表示同樣的資料，不論是用何種分類技術，所得到的分類效果差異並不是太大，本研究最後採取逐步區別分析與決策樹兩種方法交互運用，取擷兩者優點，將有益於實務的運作，對於保險公司業務經營及顧客篩選上皆有助益。
4. 本研究最終評估模式，採取逐步區別分析與決策樹交互運用，其有助於保險業者在進行優良顧客留存分類判別與準客戶篩選時，加速對顧客的屬性的判斷，對於A級優良顧客則施以更完備且有效的行銷策略，可達到減少成本增加利潤的最大功效，並將顧客分群管理，在日益競爭的壽險市場，以最短的時間，最快速的行動，最有效率的策略，運作於實務上業務拓展，及全力維護優良保單，便能奠定保險及從業人員良好的社會形象。經營效率提升。

二、建議

1. 本研究兩種實證方法，都是在加入行為資料變數後，其總準確率都大幅提升，行為資料變數，代表的是與顧客互動過程留下的服務記錄，而人口統計資料變數是訂定新契約時告知的基本資料，顧客提供給保險公司的資料，保險公司會經過核保過程的驗證，對於保單的持續率影響程度有限；但顧客的行為資料會因為保單效力行使期間而有所改變，這種與保險公司行為互動產生的記錄，在挑選良質保單上，對保險公司應該是最有用的資訊，經過逐步區別與決策樹的實證分析，今後在改善與控制良質保單持續率，在思維上與傳統作業習慣上，應有不同的作為，對於優良顧客的篩選，應該將重心著墨在行為資料變數，因為「優良顧客才能創造良好的利潤」。
2. 在本研究的兩種實證方法中，以「年收入」的預測能力最佳，實質上，「年收入決定顧客繳費能力」，這個變項值得業務人員的深思，因為一張良質保單，包括了保戶的付款能力，但這並不代表只在較高收入的市場，去尋找準顧客，在每個不同收入的階層中，必有良質保單的準顧客存在，本著「客製化」一對一量身訂做的行銷觀念，在進行市場區隔作顧客分類的時候，將「年收入」列入重要考量因素，身為業務人員不能只注重業績與配套商品的銷售，應該推銷適當的保單，來配合保戶的繳費能力。否則，保單在第二次保費交付之前，便已失效，對保險公司而言，是一種直接的回收損失。因此繳費能力也是決定保單持續力不可或缺的要素。不過也要避免銷售過低額的保單，首年度的保單銷售成本，往往高於實際保費的收入，利潤隨著保單大小而增減，
3. 上述建議1提及行為資料變數，才是優良顧客留存的重點，因為行為資料變數是保單效力行使期間，顧客與保險公司行為互動產生的記錄，其對保單的留存甚鉅。例如：「再購記錄」，當顧客與業務人員的互動過程產生服務滿意，一位滿意的顧客，不但，會自己重複購買，更會將好的經驗傳達給他人，高度滿意的顧客可為公司帶來許多好處，除了會將正面的意見提供別人，也會維持長期的忠誠，以留住具有貢獻度的顧客，減少開發新市場的成本。
4. 另外，「繳費方式」與「繳費管道」，在本研究雖未出現在顯著變數中，但其在「優良顧客留存」的實務上

扮演的重要性亦不容忽視，研究發現繳費方式為「月繳」的顧客，大都採取的繳費管道為「自行劃撥」佔多數，一般顧客容易因繁忙而忘記繳納保費，造成保單的失效，我們要補充說明的是，「月繳」保單，如果透過銀行或郵局轉帳，準時繳費，其持續率也可另人欣慰。

5. 在日益競爭的保險市場，保單持續率，除了對保險業者有息息相關的重大影響外，還能影響保險公司維持具有競爭性的費率、或分紅的能力。大量的保單失效，使得保險公司必須設定較高的保費，支付較低紅利，或給予較低現金價值。如此一來，勢必減弱公司的競爭力，徹底認識保單失效所產生的損失，並體認維持保單持續率與發展新業務同等重要。故在競爭的保險市場，除積極的開發新的市場，努力的維護現有顧客，更是當務之急。
6. 良質保單與否，關係著社會大眾的福祉。資料顯示，在美國及加拿大兩地，將近有值180億美元的失效保單，其原因是未能如期在保費催告期限內繳納續期保費。而接近200萬張保單，未能領取任何利益之前，保單就失效了，這群的保戶甚至於付出極大代價，最後卻一無所有。對社會投保大眾來說，意謂著願望未能達成，計劃遭到破壞，也代表失去保障，在大多數的情況下，便對保險及其從業人員失去應有的尊重，若能重視顧客權益。

參考文獻

一、中文部分：

1. 中華民國人壽保險同業公會編制（98）人身保險業務員資格測驗統一教材。
2. 江朝國,2003,保險業之資金運用,財團法人保險事業發展中心
3. 林震岩,2010,多變量分析:SPSS 的運用與操作
4. 陳順宇,多變量分析 華泰文化（四版）
5. 陳雅鈴，2004，顧客滿意、顧客忠誠與顧客保留之關係——一個實證上的檢驗，碩士論文。
6. 楊世瑩,2009,SPSS 統計分析.
7. 郭文德(譯),1994,壽險行銷管理.
8. 蕭振農，2003，探討壽險業顧客關係管理的策略與運作流程——以南山人壽為例，國立中山大學高階經營管理研究所，碩士論文。
9. 曾泰鋒，2005，現金卡優良客戶挑選之實證研究，國立高雄應用科技大學商務經營研究所，碩士論文。

二、英文部分：

1. Aspinall, Edward, Clive Nancarrow and Merlin Stone (2001), "The Meaning and Measurement of Customer Retention," *Journal of Targeting, Measurement and Analysis for Marketing*, 10 (1), 79-87.
2. Berné, Carmena, Múgica, José M. and Jesús Yagüe, M. (2001), "The Effect of Variety-Seeking on Customer Retention in Services," *Journal of Retailing and Consumer Services*, Vol. 8, Issue 6, pp.335-45.
3. Fornell, Claes and Birger Wernerfelt (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, Vol. 24, No. 3, pp.337-46
4. Geller, Lois (1997), "Customer Retention Begins with the Basics," *Direct Marketing*, Vol. 60, Issue 5, pp.58-62.
5. Gerpott, Torsten J., Wolfgang Rams and Andreas Schindler (2001), "Customer Retention, Loyalty, and Satisfaction in the German Mobile Cellular Telecommunications Market," *Telecommunications Policy*, 25, 249-69.
6. Geller, Lois (1997), "Customer Retention Begins with the Basics," *Direct Marketing*, Vol. 60, Issue 5, pp.58-62.
7. Hughes, Arthur M. (1994), "Strategic Database Marketing", Chicago : Probus Publishing.7. Kalakota , Ravi and Marcia Robinson (1999), "e-Business-Roadmap for success", Addison-Wesley.
8. Huberty, C. J. (1989). **Problems with stepwise methods-Better alternatives.**
In B. Thompson (Ed.), *Advances in Social Science Methodology*, (Vol. 1, pp.43-70). Greenwich: JAI Press Inc.

9. Innis, Daniel E. and Bernard J. La Londe (1994), "Customer Service: The Key to Customer Satisfaction, Customer Loyalty, and Market Share," *Journal of Business Logistics*, 15 (1), 1-27.
10. Payne, Adrian (2002), "The Value Creation Process in Customer Relationship Management," *Insight Interactive*, pp. 1-17
11. Reichheld, Frederick F. and W. Earl Sasser, Jr. (1990), "Zero Defections: Quality Comes to Services," *Harvard Business Review*, Vol. 14, Issue (March), pp.495-507.
12. Shani David, Sujana Chalasani (1992) "Exploiting Niches Using Relationship Marketing" *The Journal of Consumer Marketing*, Vol.9, pp.33-42
13. Swift, R., "Accelerating Customer Relationships" Prentice Hall, 2001
14. Weinstein, Art (2002), "Customer-Specific Strategies Customer retention: A Usage Segmentation and Customer Value Approach," *Journal of Targeting, Measurement and Analysis for Marketing*, 10 (3), 259-68.
15. Zeithaml, Leonard, L. Berry and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, Vol. 60, Issue 2, pp.31-46.
16. J. R. Quinlan, "Introduction of Decision Trees," *Machine Learning*, Vol. 1, pp. 81-106, 1986.