第二章 文獻探討

汽車保險費率釐訂中，可概分為損失金額的分配函數為已知或未知。但是由於損失金額往往會群聚在某些數值，且一般為右偏的資料型式，在統計檢定理論的限制下，要求得合適的分配模式往往非常困難，因此本章將著重在分配函數為未知的情況下，探討自由分配(distribution free)方法在費率釐訂的相關文獻內容。另外、近幾年發展的類神經網路(neural network)方法亦不需假設理論分配，可從以下引用的文獻中，了解相關的應用以做為本文研究方法的參考。

2.1 最小偏差估計法

2.1.1 理論介紹

在不假設理論分配的情況下，Bailey & Simon (1960)提出的最小偏誤(minimum Bias)估計法，可算是最早由資料的角度計算車險費率的方法。為了使整體費率達到收支均衡，保費收入與理賠支出的誤差應該愈小愈好，借由降低別個分類因素間的誤差來達到整體費率的不偏性，因為條件式永遠會比分類因素多一項，因此以疊代的方式求取模型的解析解(close form)，做為加減費係數的估計值。

舉例而言：假設汽車保險的被保險人可分為i種性別與j種年齡，x_i為第i種性別被保險人之估計費率係數，y_j為第j種年齡被保險人之估計費率係數，n_{ij}為第i種性別、第j種年齡被保險人的滿期車數，r_{ij}為第i種性別、第j種年齡被保險人的純保費，以疊代(iterative)的方式求出在給定的差異比率(percentage of difference)下x_i與y_j的估計值。

$$\sum n_{ij} \cdot r_{ij} = \sum n_{ij} \cdot \Phi(x_i, y_j) \qquad (2\text{-}1\text{-}1)$$

若 $\Phi(x_i,y_j)=x_i\times y_j$ 被稱為乘型模式；$\Phi(x_i,y_j)=x_i+y_j$ 則被稱為加型模式，加型模式假設被保險人(i, j)適用之費率為 $(BR_a)+x_i+y_j$，其中 BR_a 為基本保費，而第 i 種性別之平均誤差定義為：

$$AD_i=\frac{\sum_j n_{ij}(r_{ij}-x_i-y_j)}{\sum_j n_{ij}r_{ij}} \tag{2-1-2}$$

依貝里氏提出的限制條件[1]：就每一性別之平均誤差需等於零，每一年齡之平均誤差需等於零，令公式(2-1-2)為零可求出第 i 種性別之估計費率係數如公式(2-1-3)，以相同方式可求得第 j 種年齡之估計費率係數如公式(2-1-4)

$$x_i=\frac{\sum_j n_{ij}(r_{ij}-y_j)}{\sum_j n_{ij}}\text{；(2-1-3)} \qquad y_j=\frac{\sum_i n_{ij}(r_{ij}-x_i)}{\sum_i n_{ij}}\text{；(2-1-4)}$$

如上所述，利用疊代法求各費率係數。首先任取 x_i y_j 中的任一個，假設是 y_j，設定任一起值，由公式(2-1-3)求出 x_i，再將結果代入公式(2-1-4)求出 y_j，反覆此步驟至(x_i,y_j)收斂或誤差小於給定的範圍為止。

2.1.2 相關文獻

關於分類因素與平均保費之間的關係，除 Almer(1957)提出的乘法模型，以及 Bailey(1960)提出的加法模型外，為了提高分類因素的配適度，DuMouchel(1983)[2]則提出了混合模式(hybrid model)，最後 Brown(1988)以函數型態將分類因素與平均保費的關係一般化(Generalized)，在假設損失金額的分配函數已知的情況下，以最大概似法(maximum likelihood)及最小平方法(least squares estimate)計算各分類係數，文中提到以損失率為應變數(dependent variable)估計的結果優於損失金額，梁正德(民 82)以國

---

1 中文參考文獻[6]
2 中文參考文獻[7]

內強制汽車第三人責任保險比較貝里氏(Bailey)與布朗式(Brown)提出的估計方法，結論建議布朗式估計法可供國內在分類費率釐訂時的參考，而布朗式估計法也較普遍為世界各國所採用(Aderson,etc.al.2003)。

魏長賢(民 83)以國內汽車保險業的經驗資料，以貝里氏法釐訂從人因素係數，提出以計點方式並考慮被保險人無肇事年度，做為加減費係數的基礎，由於觀察到純保費與賠款記錄點數呈線性關係，故以純保費的線性估計值代替分類中被保險人的純保費，彌補因為分類增加，造成同一分類中的經驗資料過少的問題。另費率代號係數表依貝里氏法以發照年份與費係代號計算各組的最適純保費(陳強,民 84)，陳建龍(民 84)指出在從車因素中，若其中含有從人因素影響的成份，則兩項因素將造成雙重懲罰的情形，會發生這樣的現象，導因於最小誤差法的獨立性假設，因此在考慮分類因素間的交叉影響下，逐漸有學者提出新的方法來計算費率（Guo, 2003）(Dugas et al., 2003)，例如決策樹(decision trees)、支向機(support vector machines)、以及類神經網路。

2.2 類神經網路(Artificial Neural Network)

心理學家 McCulloc(1943)與數學家 Pitts 共同提出了神經元最早的數學模式，Rosenblatt(1958)首先引用感知機（percepton）觀念來模擬大腦的感知與學習兩大能力，Minsky(1960)證明感知機無法處理互斥或(exclusive OR;XOR)問題，直至 Rumelhart(1986)借由加入隱藏層補足了感知機缺點，再度喚起研究風潮。

2.2.1 神經元（Neuron）[3]

欲了解類神經網路，必須先了解生物神經網路是如何運做？（圖 2-2-1）當外界資訊（如：聲、光、電、熱）等刺激，透過感覺器官轉換成電的訊號後，便會經由突觸間內部的電位變化，透過樹突傳送至神經

[3] 中文參考文獻[10] P2-9

細胞，再由軸突傳送至樹突，成為下一個神經元的輸入訊號，而早期近以神經元的數學模型（圖 2-2-1），X1，X2 稱為輸入值，W1，W2 為輸入值連絡至 Σ 節點（node）中的權重，以表示輸入訊號的強度，Σ 表示把所有輸送至此的輸入相加，b 為偏壓，因此：

$$n = X_1 \cdot W_1 \cdot + X_2 \cdot W_2 + b \qquad (2\text{-}2\text{-}1)$$

若以 $n = 0$ 為一界限，$n \geq 0$ 視為輸入值激發神經元，反之則輸入值被壓抑（未被激發），故改寫 $n = X_1 \cdot W_1 \cdot + X_2 \cdot W_2 + b - TH$，$TH$ 稱為閥值（threshold）或臨界點，意即外界的刺激需大於臨界點，神經元才會被激發，而 $f(n)$ 稱為激發（activation）或轉移（transfer）函數，功能在於將 n 轉化為一輸出值 a 以使大腦接收。

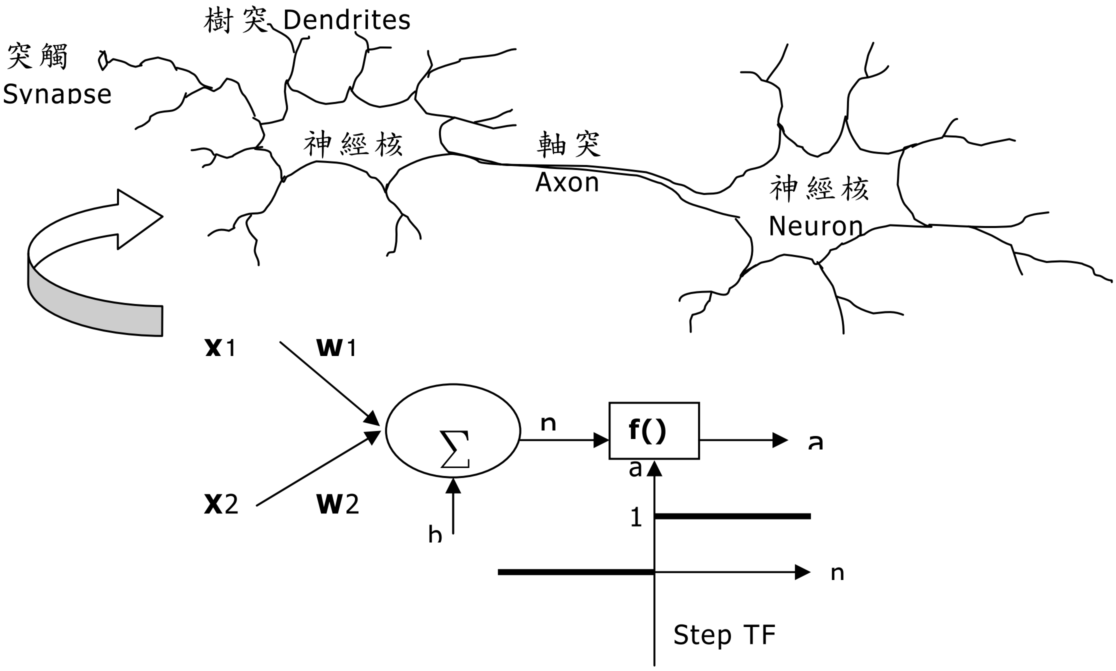

圖 2-2-1: 生物神經元數學模型

2.2.2 模型介紹

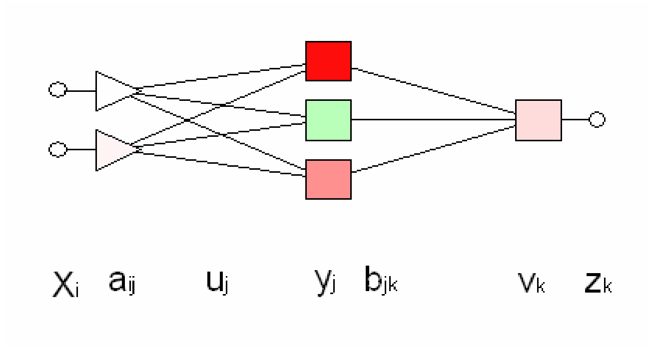

圖 2-2-2: 類神經網路例圖

按 2.1.1 節的例子說明，被保險人適用性別與年齡分類因素，在此 x_i 定義為分類因素的個數，$i=1,\ldots,I$ 依此例 $i=1,2$，意即網路有兩個輸入值，取一層非線性隱藏層，結點個數 $j=1,\ldots,J$(例 $J=3$)，轉換函數[4]取羅吉斯函數[5]，將我們所關心的 r_{ij} 視為目標輸出值 t_k，此處僅一個輸出值 $k=1$，網路輸出值設為 z、隱藏層輸出值為 y_j，以下導論類神經網路的計算過程:

(1) 前回饋過程(Feed Forward Pass)[6]

步驟 1: $u_j = a_{0j} + \sum_{i=1}^{2} a_{ij} x_i$ 計算輸入值的線性組合

步驟 2: $y_j = g(u_j) \equiv \dfrac{1}{1+e^{-u_j}}$ 透過轉換函數使隱藏層接收

步驟 3: $v = b_0 + \sum_{j=1}^{3} b_j y_j$ 為隱藏層的輸出值的線性組合

步驟 4: $z = g(v) \equiv \dfrac{1}{1+e^{-v}}$ 透過轉換函數成為輸出值

步驟 5: $Error = -\dfrac{1}{2}(z-t)^2$ 比較輸出值與目標值之間的誤差距離

(2) 誤差倒傳遞過程(Error Back Propagation Pass)

[4] 常用的轉換函數詳中文參考文獻[11],P2-11
[5] 或稱雙彎曲函數（sigmoid function）
[6] 英文參考文獻[15],P210

整個計算過程即是名為「倒傳遞」的由來，因為此過程描述一種遞迴關係，即從第 *m+1* 層的靈敏度來計算第 *m* 層的靈敏度。

步驟 1: $\frac{\partial Error}{\partial z}=(z-t)$ 衡量誤差與網路輸出的敏感度

步驟 2: $\frac{\partial z}{\partial v}=z(1-z)$ 衡量網路輸出與隱藏層輸出的敏感度

步驟 3: $\frac{\partial v}{\partial b_j}=\begin{cases}1 & j=0\\ y_j & j=1,\ldots,J\end{cases}$ 衡量隱藏層權重的敏感度

步驟 4: $\frac{\partial v}{\partial y_j}=b_j$ 計算隱藏層權重

$$\text{令}\quad p\equiv\frac{\partial E}{\partial z}\cdot\frac{\partial z}{\partial v}=(z-t)\cdot z\cdot(1-z)\quad ;\ E=Error$$

$$\text{依上述}\quad \frac{\partial E}{\partial b_j}=\begin{cases}p & j=0 & \text{誤差權重}\\ py_j \quad ; & j=1\cdots J & \text{隱藏層權重}\\ px_{j-J} & j=J+1\cdots J+I & \text{輸入層權重}\end{cases}$$

步驟 5: $\frac{\partial y_j}{\partial u_j}=y_j(1-y_j)$ 衡量隱藏層與輸入層的敏感度

步驟 6: $\frac{\partial u_j}{\partial a_{ij}}=\begin{cases}1 & i=0\\ x_i & i=1,\ldots,I\end{cases}$ 計算輸入層權重

$$\text{已知}\ \frac{\partial E}{\partial a}=\frac{\partial E}{\partial y}\cdot\frac{\partial y}{\partial u}\cdot\frac{\partial u}{\partial a}$$

$$\because \frac{\partial E}{\partial y_j}=\sum_{j=1}^{J}p\cdot\frac{\partial v}{\partial y_j}=\sum_{j=1}^{J}p\cdot b_j$$

$$\therefore \frac{\partial E}{\partial a_{ij}}=(\sum_{j=1}^{J}p\cdot b_j)\cdot y_j(1-y_j)$$

$$\text{令}\ q\equiv\frac{\partial E}{\partial y}\cdot\frac{\partial y}{\partial u}$$

$$\text{則}\ \frac{\partial E}{\partial a_{ij}}=\begin{cases}q_j & i=0 & \text{誤差權重}\\ q_jx_i & i=1\cdots I & \text{輸入層權重}\end{cases}\ ;$$

步驟 7: 將第一次疊代取得的權重值代入網路，再將第二組資料輸入，計算第二次疊代所得權重，直至誤差不在有明顯變動。

2.2.3 相關文獻

類神經網路於產險領域的應用，最初被用在預測產險公司是否失卻清償能力（Brockett et al., 1994），Speights（1999）則以類神經網路預估勞工補償保險的賠案處理時間，Cristina(2002)在個人汽車保險費率擬訂上，比較廣義線性模型、類神經網路、分類與迴歸樹(CART)[7]等方法，結論肯定類神經網路較不受離群值（Outlier）與散佚值（missing value）的影響，但是可能有過度配適（over fitting）的危險，存在不易了解與解釋的黑盒子（black box），在簡單又好用的標準下，建議選擇廣義線性模型。Dugas 等人(2003）探討統計學習理論在汽車保險費率釐訂上的應用，文中提出廣義線性模型、類神經網路、決策樹(decision trees)與支向機(support vector machine)等模型，結論類神經網路能反映因素間的複雜關係且具有預測能力。因此、Mano (表 2-2-1)與 Dugas (表 2-2-2)兩人對於類神經網路在汽車保險費率釐訂上的應用有著截然不同的看法。Mano 著重在模型的解釋能力；Dugas 則強調模型的預測能力。

表 2-2-1 模型比較表 (Mano & Rasa)

++ :優 / xx: 劣 / - - :無關	廣義 線性模型	類神經 網路	分類與 迴歸樹
參數選取	++	xx	--
過度配適的危險	xx	++	--
散佚值的處理能力	xx	++	++
異常值的處理能力	++	++	++
計算過程	--	xx	--
計算時間	++	xx	xx
可讀性與可了解程度	++	--	++
可行性	++	xx	xx
整體評估	++	xx	xx

資料來源: 27th International Congress of Actuaries

[7] CART: Classification and Regression Trees

表 2-2-2 模型比較表 (Dugas et al.)

++:優 / xx:劣 / - - :無關	貝里氏 模型	廣義 線性模型	類神經 網路
簡單模型	++	++	xx
考慮因素間複雜的 交互關係	xx	xx	++
預測能力	xx	xx	++
穩定性	++	++	++
反映個體差異性	xx	++	++

資料來源: ApStar., CAS Seminar 2003

國內文獻方面，王健亞(民 88) 以類神經網路推估全民健保醫療費用，結論類神經網路具有動態精算功能，並可做為精算人員的輔助決策，其後則偏重於分類問題的探討，劉坤河(民 90)以類神經網路建構全民健保論人計酬風險調整模型，分類出高低風險的醫療成本，預測正確率達71.7%，翁永富(民 91)則以類神經網路選取強制汽車責任保險的分類因素，整體預測正確率 85%，林民宏(民 92)比較類神經網路與羅吉斯迴歸預測壽險保單早期失效的機率，結論類神經網路的誤判成本較低，許勝仁(民 92)以汽車車體損失險為例，以自組織映射網路(Self-organizing Map)[8]的聚類結果探擷影響風險程度的特徵，可藉此做為費率釐訂的參考方式。

2.3 綜合討論~對本文的啓示

依本文求取損失金額不偏估計量的研究目標，在滿足自由分配的假設下，貝里氏最小偏差估計即是最早符合的研究方法，並且為國內現行制度，而在新發展的資料萃取方法中，本文選擇類神經網路的原因，在

[8]學習規則分成兩大類：監督式學習（supervised learning）給定相對應的目標輸出值以供比較，而無監督式學習（unsupervised learning）則沒有目標輸出值可供比較，自組織映射網路即屬無監督式學習。

於其以最小平方法，求得觀察值與估計值的最佳映射函數(符合本文研究目標)、決策樹則以同一類別中佔大多數的機率特徵做為分類標準(不符合本文研究目標)、支向機旨在可線性分離各分類因素的維度空間中，找到可以區分各分類因素的迴歸式，雖然符合本文研究目標，但是需要冗長的運算時間，不若類神經網路的平行運算效率。綜上所述，本文採用類神經網路做為實驗組，以貝里氏(加型)模型及布朗式最小平方法做為對照組，進行模型比較。雖然加型與乘型均符合不偏性(unbiased)的要求，但乘型的有效性(efficiency)不如加型(Weisberg, 1982)，故採用貝里氏加型模型以做為比較[9]。另採用梁正德(民 82)的研究建議(貝里氏乘型、布朗式常態分配(乘型)、布朗式最小平方估計(乘型)為最佳)，以布朗式最小平方法與貝里氏模型加以比較。

2.3.1 模型差異性比較

依 Bailey 提出的理論，估計費率與實際損失間之平均誤差需為零，採用最小平方法借由反覆的疊代，找出收斂的參數估計值，而在不強調變異數齊一($Var(\varepsilon_i)=\delta^2$)的情況下，以最小平方法所求得之估計式，雖仍滿足不偏性(unbiased)，但已不再具有效性(efficiency)，意即變異數非為最小，使得參數變異數之估計產生偏誤的現象，迫使最小平方估計不再是最佳估計式。另外、迴歸分析的過程中，若是獨立性的假設無法成立，則可能導致錯誤的結論，殘差變異及各項迴歸係數之標準誤可能造成過度樂觀的估計，意即估計值低於實際發生值。另一方面，採用類神經網路計算加減費係數，則在於將估計費率與實際損失間的誤差，倒向分配予各網路輸入層權重與隱藏層權重，使得估計值與觀察值之間的誤差最小。

兩種方法的主要差異在於誤差的來源。貝里氏透過降低分類誤差的估計值來達到整體誤差的最小化；類神經網路則直接縮小個別資料的誤

[9] 在美國加型、乘型或混合式均有保險公司使用，歐洲則以加型居多(中文參考文獻[9]P.153。)

(將學習率 α 視為函數微分的倒數),借由隱藏層的加入,使類神經網路可表現輸入處理單元的交互影響,使得類神經網路不需預先假設觀測值與估計值的對映關係Φ。

模型	最小誤差估計	類神經網路模型
誤差函數	$AD=\sum n_{ij}(r_{ij}-\Phi(x_i,y_j))$	$SSE=E(e^2)=E[(t-a)'(t-a)]$
最適化過程	$\Delta x=x_{n+1}-x_n=-\dfrac{f(x)}{f'(x)}$	$w_{i,j}^m(k+1)=w_{i,j}^m(k)-\alpha\dfrac{\partial\hat{F}}{\partial w_{i,j}^m}$
方法	牛頓(Newton-Raphson)法	最陡坡降(gradient steepest descent)法

表 2-3-1 貝里氏與類神經網路演算式比較表