

# 中文摘要

應用資料採礦技術於貝氏網路架構之研究

線性代數與圖論之結合

研究生：黃意策

指導教授：王偉華 博士

東海大學工業工程與經營資訊研究所

## 摘要

科技的快速發展，使網際網路成為全球最重要的通訊媒介。資訊取得的便利，導致資訊量呈爆炸性地成長。想從繁雜的資料中有效率地過濾出有用、有價值的資訊加以處理運用，需要資訊處理技術的協助，造成資料採礦的廣泛應用。但現存的資料採礦技術大部分僅能於資料架構已知下，處理變數的問題。若能從資料中找出變數間的關係，建立資料的架構，將能提高資訊處理技術的效率及效能。

現存資料採礦技術中機率圖形相依模式 貝氏網路 的建立都是採用搜尋式的演算法。這類的演算法需要大量的資料才能保證網路架構的可靠性；變數多時，變得沒有效率，甚至無法計算，難以應用於實際情況。

本研究透過矩陣相似性轉換，找到相似矩陣，再結合圖論畫出貝氏網路架構。期望藉由數學明確性的特質，結合線性代數與圖論，提供一個非搜尋式的方法，找出適合資料的貝氏網路架構。

# 英文摘要

A Study of Finding Bayesian Network Structure

-- Combining Linear Algebra and Graph Theory

Student: Yi-Tse Huang

Advisor: Dr. Wei-Hua Andrew Wang

Institute of Industrial Engineering and Enterprise Information

Tunghai University

## Abstract

Due to the advanced technology innovation, internet becomes the most important communication media in these days. The convenience in accessing to the internet makes the information abound exponentially. If we want to find out the message behind the mass data, the fast processing technology is a must and the data mining technique is a good analytical method. But, in data mining area, most of the existed method can only cope with the known system structure. In this research, we are to develop a method to find out the relationships among data and build the structure accordingly.

In the existed methods of data mining, Bayesian network is a probabilistic graphical dependency model. Basically, Bayesian network is a search-based algorithm. Given the known structure, this algorithm needs mass data to guarantee the parameters setting of the network. And when the variables increase, the algorithms become inefficient and hard to apply in the real world.

In this paper, we have developed the similarity matrix by using the similarity transformation of the matrices. And, combined with graph theory to gain the structure of Bayesian network. The goal of this paper is that provide a non-search-based approach that is combined with linear algebra and graph theory to find out the structure of Bayesian network that is corresponded with data.

## 誌謝

兩年的時光過得很快，隨著這篇論文的完成求學路程也將告一段落。回想這兩年的風風雨雨、酸甜苦辣，首先要感謝指導教授王偉華博士的指導，無論是處事態度或是課業都給予許多的鼓勵與指導，相信這些對於未來的職場生涯有相當大的幫助。此外，也謝謝老師於求學期間給予諸多的寬容與體諒。

感謝陳麗霞老師、丁兆平老師、王文清老師及黃欽印老師於口試期間給予的諸多協助、建議和指正，使這篇論文能更加的完整。

感謝陪我走過這兩年研究所生涯的各位夥伴，俊和、立霆、淑芬、俊維、慕賢、啟信、怡文和玫媛。還有我親愛的室友們玉純、懿慧和宛璇，陪我走過低潮期。

最終，也最想要感謝的是，一直默默陪著我、支持我的家人，在我遇到低潮時，他們默默地支持我，給我最大的鼓勵及支持。沒有你們，我想我無法走完如此漫長的求學生涯，千言萬語只能化成一句謝謝。

僅以此研究成果獻給所有我關心的人和關心我的人。

黃意策 謹誌於

東海大學工業工程與經營資訊研究所

中華民國九十一年七月

# 目錄

中文摘要 .....	i
英文摘要 .....	ii
誌謝 .....	iii
目錄 .....	iv
表目錄 .....	vi
圖目錄 .....	ix
第一章 緒論 .....	1
1.1 研究動機.....	1
1.2 研究目的.....	4
1.3 研究方法.....	5
1.4 研究範圍與假設.....	8
1.5 論文架構.....	9
第二章 文獻探討 .....	10
2.1 貝氏網路架構的建立.....	10
2.1.1 K2 演算法.....	10
2.1.2 CB 演算法 .....	12
2.2 矩陣分解.....	14
2.2.1 區塊矩陣(block matrix).....	15
2.2.2 Jordan 標準型(Jordan canonical form) .....	16
2.3 矩陣與圖論.....	18
第三章 線性代數與圖論之結合 .....	22
3.1 資料產生.....	23
3.2 資料轉換.....	23
3.2.1 矩陣型式的設計.....	23
3.2.2 矩陣的元素值.....	25
3.3 矩陣相似性轉換.....	26

3.4 矩陣與圖形轉換.....	28
3.5 定性與定量轉換.....	29
3.6 相似性分析.....	29
<b>第四章 實證研究 .....</b>	<b>32</b>
4.1 範例一：Judea Pearl 警鈴(Alarm)貝氏網路 .....	32
4.1.1 本研究方法建立之警鈴貝氏網路架構.....	32
4.1.2 K2 演算法建立之警鈴貝氏網路架構.....	40
4.2 範例二：L. D. Phillips 和 W. Edward 文具袋(Book Bags)貝氏網路.....	42
4.2.1 本研究方法建立之文具袋貝氏網路架構.....	42
4.2.2 K2 演算法建立之文具袋貝氏網路架構.....	47
4.3 範例三：Steffen L. Lauritzen 和 David J. Spiegelhalter 胸腔門診(Chest Clinic) 貝氏網路.....	50
4.3.1 本研究方法建立之胸腔門診貝氏網路架構.....	50
4.3.2 K2 演算法建立之胸腔門診貝氏網路架構.....	62
4.4 本章小結.....	66
<b>第五章 結論與未來研究方向 .....</b>	<b>67</b>
5.1 結論.....	67
5.2 未來研究方向.....	68
<b>參考文獻 .....</b>	<b>70</b>
<b>附錄一 Judea Pearl 警鈴貝氏網路之相似性分析統計表.....</b>	<b>73</b>
<b>附錄二 L. D. Phillips 和 W. Edward 文具袋貝氏網路之相似性分析統計表.....</b>	<b>75</b>
<b>附錄三 Steffen L. Lauritzen 和 David J. Spiegelhalter 胸腔門診貝氏網路之相似性分析統計表 .....</b>	<b>79</b>
<b>附錄四 口試相關資料 .....</b>	<b>93</b>

## 表目錄

表 3.1 矩陣的型式 .....	25
表 3.2 資料型態 .....	26
表 4.1 警鈴貝氏網路之矩陣型式 .....	33
表 4.2 警鈴貝氏網路之 Jordan 標準型 .....	33
表 4.3 警鈴貝氏網路之相似矩陣 .....	35
表 4.4 警鈴貝氏網路之子矩陣 $M_{mja}$ .....	36
表 4.5 警鈴貝氏網路之子矩陣 $M_{mja}$ 的 Jordan 標準型 .....	36
表 4.6 警鈴貝氏網路之子矩陣 $M_{mja}$ 的相似矩陣 .....	37
表 4.7 文具袋貝氏網路之矩陣格式 .....	44
表 4.8 文具袋貝氏網路之 Jordan 標準型 .....	44
表 4.9 文具袋貝氏網路之相似矩陣 .....	44
表 4.10 胸腔門診貝氏網路之矩陣格式 .....	52
表 4.11 胸腔門診貝氏網路之 Jordan 標準型 .....	52
表 4.12 胸腔門診貝氏網路之相似矩陣 .....	53
表 4.13 胸腔門診貝氏網路之子矩陣格式 $M_{ctsv}$ .....	56
表 4.14 子矩陣 $M_{ctsv}$ 的 Jordan 標準型 $J_{ctsv}$ .....	56
表 4.15 胸腔門診貝氏網路之子矩陣 $M_{ctsv}$ 的相似矩陣 .....	57
表 4.16 胸腔門診貝氏網路之子矩陣格式 $M_{dxtb}$ .....	58
表 4.17 胸腔門診貝氏網路之子矩陣 $M_{dxtb}$ 的 Jordan 標準型 .....	58
表 4.18 胸腔門診貝氏網路之子矩陣 $M_{dxtb}$ 的相似矩陣 .....	58
表 4.19 架構相似度比較表 .....	66
表 4.20 資料相似度比較表 (p 值) .....	66
表 A-1.1 本研究方法的警鈴貝氏網路：新資料與原始資料之變數組合 統計表 .....	73

表 A-1.2 K2 演算法的警鈴貝氏網路：新資料與原始資料之變數組合統計表 .....	74
表 A-2.1 本研究方法的文具袋貝氏網路：新資料與原始資料之變數組合統計表 .....	75
表 A-2.2 本研究方法的文具袋貝氏網路：新資料與原始資料之變數組合統計表（續） .....	76
表 A-2.3 K2 演算法的文具袋貝氏網路：新資料與原始資料之變數組合統計表 .....	77
表 A-2.4 K2 演算法的文具袋貝氏網路：新資料與原始資料之變數組合統計表（續） .....	78
表 A-3.1 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表 .....	79
表 A-3.2 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 1） .....	80
表 A-3.3 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 2） .....	81
表 A-3.4 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 3） .....	82
表 A-3.5 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 4） .....	83
表 A-3.6 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 5） .....	84
表 A-3.7 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 6） .....	85
表 A-3.8 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表 .....	86
表 A-3.9 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 1） .....	87

表 A-3.10 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數 組合統計表（續 2） .....	88
表 A-3.11 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數 組合統計表（續 3） .....	89
表 A-3.12 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數 組合統計表（續 4） .....	90
表 A-3.13 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數 組合統計表（續 5） .....	91
表 A-3.14 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數 組合統計表（續 6） .....	92



# 圖目錄

圖 1.1 貝氏網路的建構流程 .....	3
圖 1.2 建構貝氏網路的方法 .....	6
圖 2.1 K2 演算法的演算過程 .....	11
圖 3.1 本研究之研究方法 .....	22
圖 3.2 資料分割示意圖 .....	24
圖 3.3 矩陣型式設計之概念圖 .....	24
圖 3.4 圖形示意圖 .....	28
圖 3.5 舉列說明之架構圖 $S$ 和 $T$ .....	30
圖 4.1 Judea Pearl 的警鈴貝氏網路.....	32
圖 4.2 警鈴貝氏網路之非循環方向圖 .....	35
圖 4.3 警鈴貝氏網路之子非循環方向圖 .....	37
圖 4.4 警鈴貝氏網路之非循環方向圖 .....	37
圖 4.5 本研究方法建立之警鈴貝氏網路 .....	37
圖 4.6 本研究方法的警鈴貝氏網路：魏克森符號等級和檢定 .....	39
圖 4.7 K2 演算法建立的警鈴貝氏網路之非循環方向圖 .....	40
圖 4.8 K2 演算法建立之警鈴貝氏網路架構 .....	40
圖 4.9 K2 演算法的警鈴貝氏網路：魏克森符號等級和檢定 .....	42
圖 4.10 L. D. Phillips and W. Edward 的文具袋貝氏網路.....	43
圖 4.11 文具袋貝氏網路之非循環方向圖 .....	45
圖 4.12 本研究方法建立之文具袋貝氏網路 .....	45
圖 4.13 本研究方法的文具袋貝氏網路：魏克森符號等級和檢定.....	47
圖 4.14 K2 演算法建立的文具袋貝氏網路之非循環方向圖 .....	48
圖 4.15 K2 演算法建立之文具袋貝氏網路架構 .....	48
圖 4.16 K2 演算法的文具袋貝氏網路：魏克森符號等級和檢定 .....	50

圖 4.17 Lauritzen 和 Spiegelhalter 的胸腔門診貝氏網路.....	51
圖 4.18 胸腔門診貝氏網路之子非循環方向圖 .....	54
圖 4.19 胸腔門診貝氏網路之子非循環方向圖 .....	57
圖 4.20 胸腔門診貝氏網路之子非循環方向圖 .....	57
圖 4.21 胸腔門診貝氏網路之子非循環方向圖 .....	59
圖 4.22 胸腔門診貝氏網路之非循環方向圖 .....	59
圖 4.23 胸腔門診貝氏網路之非循環方向圖 .....	60
圖 4.24 本研究方法的胸腔門診貝氏網路：魏克森符號等級和檢定 .	62
圖 4.25 K2 演算法建立的胸腔門診貝氏網路之非循環方向圖 .....	63
圖 4.26 K2 演算法建立之胸腔門診貝氏網路架構 .....	63
圖 4.27 K2 演算法的胸腔門診貝氏網路：魏克森符號等級和檢定 ....	65

# 第一章 緒論

## 1.1 研究動機

二十世紀以來，許多產業隨著科技的快速發展，都有了急遽的轉變。舉凡電子商務，網際網路等等皆成為二十世紀的新寵兒，尤其是網際網路，其傳播資訊的速度更是不可同日而語。

二十一世紀的新時代，網際網路滲入全球每個家庭中的速度更加地快速，成為全球最重要的通訊媒介。資訊的取得更加便利，資訊量呈爆炸性地成長。要想從龐大的資訊中有效率地過濾出有用、有價值的資訊加以處理運用，已不是人工所能完全負荷的。除此之外，各種資料處理所增加的成本，例如：人力成本、時間成本、彌補人為錯誤的修復成本等也快速增加。為了加快資訊的處理，節省人力成本、時間成本，資訊處理技術大量地取代人工來過濾和處理資訊。因此，形成資料採礦(Data mining)的廣泛應用。但現存的資料採礦技術大部分僅能於資料架構已知下，處理變數的問題。若能由資料中找出變數之間的關係，建立資料的架構，將能提高資訊處理技術的效率及效能。

目前已有幾個較成熟、較容易被接受的資料採礦技術，包含下列模式(Adriaans & Zantinge, 1996; Agrawal, *et al.*, 1996; Fayyad, *et al.*, 1996; Kennedy, *et al.*, 1998)：

1. 決策樹和決策規則(decision tree and rules)。
2. 線性模式(Linear models)。
3. 非線性模式(Non-linear models)。例如：類神經網路(Artificial neural network)。
4. 案例式方法(Example-based methods)。例如：最近鄰法(Nearest-neighbor methods)和案例式推論法(Case-based reasoning methods)。
5. 機率圖形相依模式(Probabilistic graphical dependency models)。例如：貝氏網路(Bayesian networks)。
6. 關聯屬性模式(Relational attribute models)。

近十年來，機率圖形相依模式中的貝氏網路已經成為專家系統中不確定專家知識的表示法[21]。貝氏網路是一種結合統計方法的圖形模式，表示變數間的機率性關係。其優點如下[21]：

1. 此模型表示所有變數間的相依性，所以容易處理不完整的資料集合，即 missing data 的問題。若有資料未被觀察到，大部分的模型都產生不精確的預測，因這些模型沒有表示變數間的關係。貝氏網路提供一個自然的方法表示此相依性。
2. 貝氏網路可以用於學習因果關係，可以用於了解問題領域和預測干擾的出現及影響。
3. 結合貝氏統計方法的貝氏網路促進領域知識 (domain knowledge) 與資料 (data) 的結合。
4. 結合貝氏統計方法的貝氏網路提供一個有效，有規則地避免資料過度配適 (overfitting) 的方法，即不需要堅持將某些資料用來作測試。使用貝氏方法的模型亦適用將所有可用資料用來訓練的方法。

建構貝氏網路的方法有兩種，一為透過領域專業人士的主觀認知勾勒出貝氏網路；另一則是透過客觀的資料分析取得[8]。

貝氏網路的建構流程，首先，選擇問題領域，定義問題。定義問題後，決定問題的變數或變數群，此過程有下列幾項主要作業：(1) 正確地確認問題目的，(2) 確認關於問題的可能觀測資料 (observations)，(3) 決定有用觀測資料的子集合，(4) 將觀測資料轉成具有互斥<sup>註1</sup>和互無遺漏<sup>註2</sup>狀態 (mutually exclusive and collectively exhaustive state) 的變數。接著，建立宣告條件獨立 (conditional independence) 的非循環方向圖 (Directed Acyclic Graph, DAG)，即貝氏網路基本架構，此架構通常是已知的。最後，評估貝氏網路架構的局部機率分配 (local probability distribution)，即參數的學習。貝氏網路的建構流程，如圖 1.1。

---

註1 互斥，是指不會有兩個狀態同時發生。

註2 互無遺漏，是指至少會有一個狀態發生。

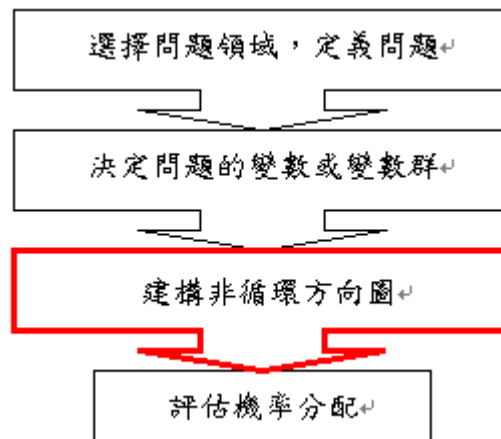


圖 1.1 貝氏網路的建構流程

前述貝氏網路的建構過程中，最主要的部分為建構非循環方向圖。為了決定貝氏網路架構，必須(1) 將變數排序，(2) 決定滿足條件獨立的變數集合。Herskovits 和 Cooper[16][24]提出一個貪婪演算法 (greedy algorithm)，即 K2 演算法。

根據 K2 演算法的基本假設，此演算法需要有一組節點(nodes)的排序才可以進行演算。但網路架構的品質對節點排序非常敏感[30]，節點的排序不好，可能導致無法顯示許多變數間的條件獨立性。所以要進行 K2 演算法時，須先給定一組適合的排序。除了 K2 演算法外，尚有另一個不需要節點排序來建構貝氏網路的方法，此方法有兩個假設必須存在，(1) 變數間的因果關係已知，(2) 因果關係要符合條件相依。因果關係必須仰賴專家知識或專家的協助，但想要從專家身上取得這些知識並不容易，因為專家可能相當地忙碌，或者專家無法將其內隱的知識外顯。此外，變數的數目太多時專家很難精確地描述出所有的關係。因此，藉由資料來建構貝氏網路是重要的。

Singh 和 Valtorta 提出一個整合條件獨立測試法(Conditional Independent tests, CI tests)與提供節點排序的演算法，為 CB 演算法<sup>註</sup>，解決節點排序的問題。CB 演算法只從案例(cases)資料庫中產生節點排序，不需要任何外部資訊，故使用者可以不需要提供變數排序[30]。

<sup>註</sup> CB 演算法，因先用條件獨立測試法提供節點排序，再執行貝氏演算法，故以此命名。

雖然 CB 演算法解決了建構貝氏網路時節點排序的問題，但 CB 演算法仍然存在著一些問題[30]：

1. CB 演算法是用條件獨立測試為基礎(CI-test-based)的演算法規劃所有節點排序，但節點排序隨著節點數增加呈指數成長。
2. 從案例資料庫中建構單一個「最好」的模型，視為真實模型。此點假設若案例資料庫中最可能網路的機率比第二可能網路的機率高，但此假設不適用於許多的真實情況。
3. 網路架構的品質對 CB 演算法第一部分所決定的排序非常敏感，需要找更好，更有效的探索法(heuristics)。

## 1.2 研究目的

根據條件獨立測試所設計的演算法，有兩個主要的缺點：[30]

1. 需要大量的資料才能保證這種方法的可靠性。

Verma 和 Pearl[34]指出，一般所有獨立關係的集合隨變數個數的增加呈指數性成長。因此，變數個數增加時，條件獨立為基礎的方法會變得難以計算。

2. Cooper 和 Herskovits[16]提出的貝氏非條件獨立測試為基礎的方法(Bayesian non-CI-test-based method)，即貝氏網路的貝氏學習(Bayesian learning of belief networks, BLN)，有四項假設必須成立：

- (1) 變數為離散型變數。
- (2) 案例間互不影響。
- (3) 所有變數均可完整地觀察。
- (4) 參數學習之前，貝氏網路架構的機率值不重要。一般以均勻分配處理學習前的機率值。

除上述貝氏學習的四項假設外，K2 演算法還有兩項假設[30]：

1. 一組可用的變數排序(ordering)。
2. 所有架構有相同地可能性(equally likely)。

如前所述，條件獨立測試所設計的演算法，需要大量的資料才能保證這種方法的可靠性，變數個數增加時，條件獨立為基礎的方法會變得無法計算。CB 演算法雖可以排除使用者必須提供變數排序的需求，但 CB 演算法節點數目增加時，節點排序之組合數呈指數性成長。如此，限制了此演算法的應用，不容易廣泛地應用於實際案例及特定領域，如：化學工業及醫療產業等等。此外，貝氏網路的貝氏學習所使用的變數侷限於離散型變數，對於建構連續型變數的貝氏網路是一項門檻。

綜合以上所述，機率圖形相依模式於執行時使用搜尋式的方法，如成對式地比較兩個節點間的因果關係及搜尋每個節點的父代集合。因此，節點數龐大時，此演算法變得相當耗時，甚至無法計算。基於這些因素，本研究期望能透過數學明確性的特質，提供一個非搜尋式的方法尋找問題領域的貝氏網路。

### 1.3 研究方法

建構貝氏網路有下列幾種可能的方法：

1. 根據資料，用排序架構(ordering structure)的方式，如 Singh 和 Valtorta 的 CB 和 K2 演算法，得到非循環方向圖。再根據得到的非循環方向圖，結合取得的案例，建構貝氏網路，最後執行參數的學習，修正貝氏網路。(圖 1.2 ①)
2. 根據資料，將資料直接以三角矩陣(Triangular matrix)的方式呈現，透過矩陣(matrix)與圖形的轉換，畫出非循環方向圖。再根據得到的非循環方向圖，結合取得的案例，建構貝氏網路，最後執行參數的學習，修正貝氏網路。(圖 1.2 ②)
3. 根據資料，將資料以矩陣的方式呈現，透過矩陣與圖形的轉換，畫出循環方向圖(Directed Cyclic Graph, DCG)。再將循環方向圖轉換成非循環方向圖，結合取得的案例，建構貝氏網路，最後執行參數的學習，修正貝氏網路。(圖 1.2 ③)

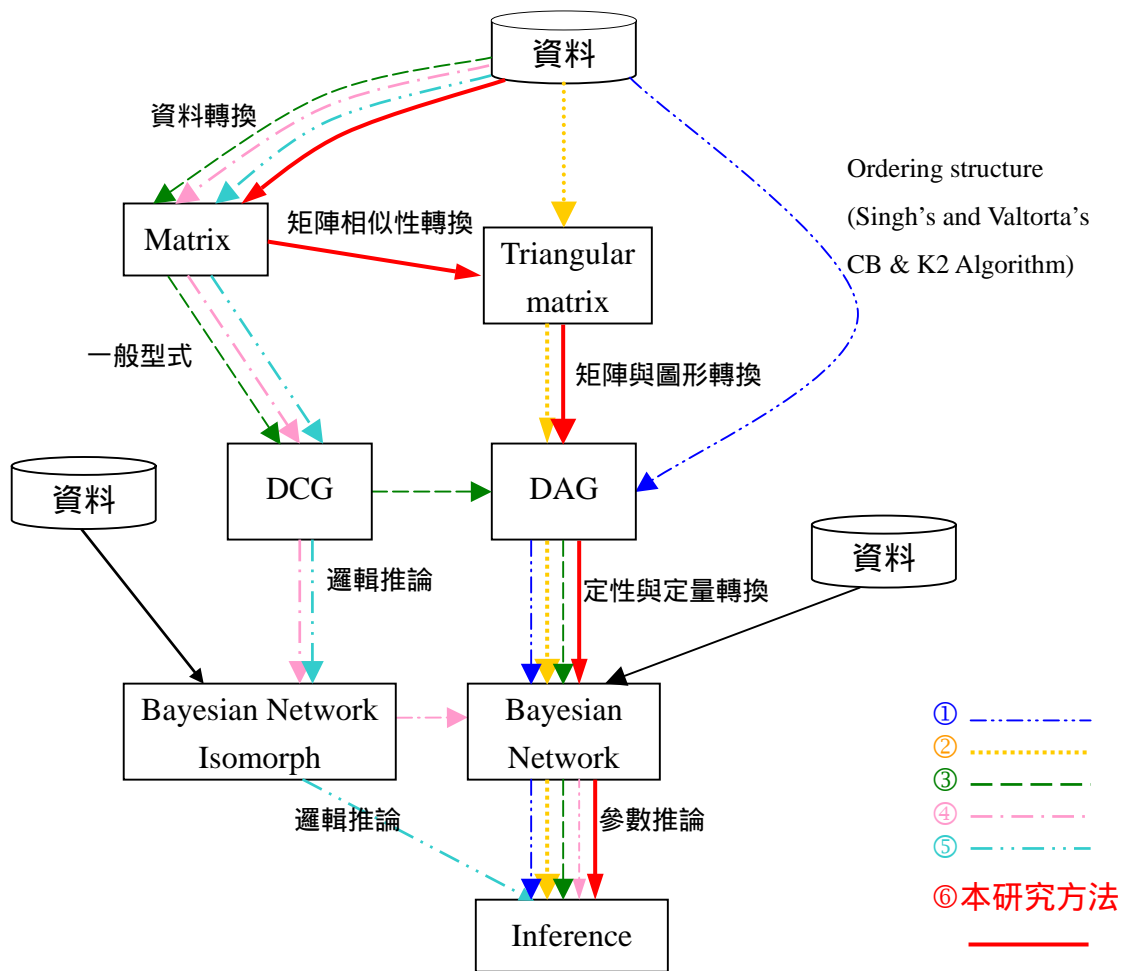


圖 1.2 建構貝氏網路的方法



4. 根據資料，將資料以矩陣的方式呈現，畫出循環方向圖，建構貝氏網路同形體(Bayesian Network Isomorph, BNI)。再將貝氏網路同形體轉換成貝氏網路，結合取得的案例，最後執行參數的學習，修正貝氏網路。(圖 1.2 ④)
5. 根據資料，將資料以矩陣的方式呈現，畫出循環方向圖，建構貝氏網路同形體，結合取得的案例，最後執行參數的學習，修正貝氏網路。(圖 1.2 ⑤)

上述五種建構貝氏網路可能的方法，除了圖 1.2 ①，用排序架構的方法，是目前建構貝氏網路的主流外，其它的方法尚未有完整的理論發展。排序架構方法，如 CB 和 K2 演算法，所發展的演算法已相當成熟，但此演算法受限於需要大量資料、節點排序及離散型變數。其中，節點數增加時，節點排序的可能性呈指數性成長，要從中選取適當的排序，需要有良好的探索法(heuristics)。因此使用此演算法時有節點數的限制。由於目前尚未發展出好的探索法，為了避免探索節點的排序，本研究期望藉由數學唯一性的特質，從數學的觀點，結合圖論(graph theory)，找出建構貝氏網路時必備的非循環方向圖。

除此之外，排序架構的方法不適合真實世界中一些問題領域的特質，例如：醫療產業短期內所能搜集到的案例很少，若採用排序架構的方法，得到的結果可信度很低；化學工業所能搜集到的案例多為連續性變數，依排序架構方法的假設，無法應用於這個問題領域。

本研究綜合前述可能建構方法中的 2 和 3，提出另一種可能的方法：

6. 根據資料，將資料轉換成矩陣型式。透過相似性轉換，先將矩陣轉換成相似的三角矩陣。透過矩陣與圖形的轉換，畫出非循環方向圖。再根據得到的非循環方向圖，結合取得的案例，將定性的非循環方向圖轉換成定量的貝氏網路。最後執行參數的學習，修正貝氏網路。(圖 1.2 ⑥)

## 1.4 研究範圍與假設

本研究主要目的是為了建立貝氏網路的基本架構，故本研究的研究範圍不包含貝氏網路的參數學習，這部分可以參考貝氏網路之參數學習的相關研究。

由於本研究的研究方法是根據線性代數所發展，因此，以線性代數的基本假設為本研究的基礎假設。

假設一：為了確保貝氏網路唯一。本研究方法之矩陣相似性轉換中，轉換矩陣(transformation matrix)不為奇異方陣(singular matrix)。

假設二：本研究以具有二元狀態的範例做為實證研究對象。

假設三：由於區塊矩陣的運算方式尚未有較完整的方法論，但基本概念與一般矩陣相似。一般矩陣相當於是一個具有  $1 \times 1$  子矩陣的區塊矩陣。故本研究以一般矩陣運算方式執行本研究方法。

假設四：矩陣之列與行的順序是已知的。此假設為本研究方法的強烈假設條件。若節點與矩陣的特徵值(eigenvalues)有相對應關係，則此假設不存在。

假設五：此項假設為本研究方法的設計假設。

### 1. 資料轉換

- 節點僅具有一個特質，使節點與節點間具有不相交性。
- 矩陣的元素值，以條件機率方式計算。

### 2. 矩陣相似性轉換

特徵值是否有 0 出現，判斷矩陣是否可以簡化節點間的關係。

### 3. 矩陣與圖形轉換

- 子矩陣是否為零的判斷方式，是以矩陣中的最大值為基準。若子矩陣中所有的相對值很小或趨近於零者，視此子矩陣為零矩陣。

- 因圖論定義之關係為矩陣之列項目到矩陣之行項目，故資料轉換時，採條件機率的方式，將對應關係設計為矩陣之列項目對應到矩陣之行項目。若矩陣設計是，採條件機率的方式，將對應關係設計為矩陣之行項目對應到矩陣之列項目。則相似矩陣於轉換圖形時，其對應關係須修正為矩陣之行對應到矩陣之列。

## 1.5 論文架構

本論文主要分為五章：第一章說明本研究之研究動機、研究目的、研究方法及研究範圍與假設等相關內容；第二章則是探討本研究內容涉及之相關文獻。包括貝氏網路架構的建立、矩陣分解、矩陣與圖論等相關文獻；第三章根據本研究之動機、目的與文獻探討的啟發，針對本研究之研究問題即結合線性代數與圖論建立貝氏網路架構，詳細描述本研究之研究方法的執行步驟；第四章為本研究方法的實證研究；第五章則根據本研究之結果，說明本研究的成果與未來發展方向。

## 第二章 文獻探討

### 2.1 貝氏網路架構的建立

從資料學習機率網路架構的方法可分成三群[30]：

1. 以線性和常態為基礎的假設。
2. 較普遍，但需要大量獨立關係的測試。
3. 貝氏方法(Bayesian approach)。

貝氏網路就是一種機率網路架構，是可以容易建構和解釋不確定知識的一種圖形表示法，可以藉由結合領域專家知識和統計資料學習新知識。是一種結合統計方法的圖形模型，表示變數間的機率性關係[20]。

通常貝氏網路的建構需要下列幾個步驟[6]：

1. 選擇問題領域，定義問題，找出與問題領域有關的變數集合，例如：一問題領域，有  $n$  個變數， $x_1, x_2, x_3, \dots, x_n$ ，則此問題領域的變數集合為  $X = \{x_1, x_2, x_3, \dots, x_n\}$ 。
2. 將變數集合中的變數做適當的排列。
3. 處理完變數集合後：
  - (1) 從變數集合中指定變數  $x_i$  做為貝氏網路中對應的節點。
  - (2) 於已建立的網路中針對  $x_i$ ，尋找一個最小的集合： $Parents(x_i)$  為  $x_i$  的『因』，使網路中條件獨立(Conditional Independence)存在。
4. 定義  $x_i$  的條件機率表。

#### 2.1.1 K2 演算法

Cooper 和 Herskovits[16]根據 (1) 變數為離散型變數，(2) 案例間互不影響，(3) 每個變數皆有對應值，(4) 觀察資料庫之前，貝氏網路架構的機率不重要，這四項假設推論出下列的結果：

考慮一組  $n$  個離散變數的集合  $Z$ 。每個變數  $x_i$  為  $Z$  的元素 (elements), 有  $r_i$  個可能的對應值  $(v_{i1}, \dots, v_{ir_i})$ 。假設  $D$  為有  $m$  個完整案例的資料庫, 即每個案例包含一個對應於  $Z$  中每個變數的值。設  $B_S$  代表一個貝氏網路的架構, 包含  $Z$  的變數。  $B_S$  中的每個變數  $x_i$  有一組父代集合  $\pi_i$ 。設  $w_{ij}$  代表  $\pi_i$  關於  $D$  的第  $j$  個單一對應 (unique instantiation), 假設  $\pi_i$  有  $q_i$  個這樣的單一對應。設  $N_{ijk}$  為案例個數, 其中案例是指  $D$  中  $x_i$  對應  $w_{ik}$ ,  $\pi_i$  對應  $w_{ij}$  的案例。設  $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$ , 則

$$P(B_S, D)^{\text{註1}} = P(B_S) \prod_{i=1}^n g(i, \pi_i) \quad (2.1.1)$$

其中,  $g(i, \pi_i)$  藉由下列的式子給定

$$g(i, \pi_i)^{\text{註2}} = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \quad (2.1.2)$$

這個結果是用來尋找資料庫中最有可能的網路架構。Herskovits 和 Cooper[16][24] 提出一個貪婪演算法, 藉由尋找每個變數使  $g(i, \pi_i)$  最大的父代集合, 讓  $P(B_S, D)$  最大的演算法, 稱為 K2 演算法。此演算法的演算過程如圖 2.1 :

```

for each node  $i, 1 \leq i \leq n$ , find  $\pi_i$  as follows :
 $\pi_i \leftarrow \phi$ 
 $P_{\text{old}} \leftarrow g(i, \pi_i)$ 
NotDone  $\leftarrow$  True
while NotDone do
   $\forall l \prec i (l \notin \pi_i), g_l \leftarrow g(i, \pi_i \cup \{l\})$ 
   $P_{\text{new}} \leftarrow \max_l g_l(i, \pi_i \cup \{l\});$ 
  Let  $z$  be the node which maximizes  $g_l$  above
  if  $P_{\text{new}} > P_{\text{old}}$ , then
     $P_{\text{old}} \leftarrow P_{\text{new}};$ 
     $\pi_i \leftarrow \pi_i \cup \{z\}$ 
  else NotDone  $\leftarrow$  false;
end {while};

```

圖 2.1 K2 演算法的演算過程

<sup>註1</sup>  $P(B_S, D)$  為貝氏網路架構  $B_S$  屬於資料庫  $D$  的機率。

<sup>註2</sup>  $g(i, \pi_i)$  為一組父代集合  $\pi_i$  屬於變數  $i$  的估計值。

## 2.1.2 CB 演算法

從資料建立貝氏網路架構的方法，不是高度依賴條件獨立測試，就是需要提供節點排序[30]。由 2.1.1 節可知，用 K2 演算法建立貝氏網路架構時，有一項假設就是變數排序是給定的。為了解決 K2 演算法的這項假設，Singh 和 Valtorta 提出 CB 演算法，應用條件獨立為基礎的演算法提供 K2 演算法一個所有節點的排序。

CB 演算法整合上述兩種方法：條件獨立測試，用於產生節點排序，再用非條件獨立基礎的方法，如 K2 演算法，建立基本貝氏網路架構[30]。這兩種方法的整合，解決過度依賴條件獨立測試和需要節點排序的問題，結果可以用來找所給資料庫中最可能的網路架構[30]。

CB 演算法的演算法如下：

$S_{ab}$ ： $a$  和  $b$  間的網路架構。

$A_{G_1}ab$ ：圖形  $G$  中鄰近  $a$  或  $b$  的點之集合，不包含  $a$  和  $b$ 。

$u$ ：一界限，受限於步驟 2 所產生之非方向圖的  $degree$ 。

$ord$ ：CI relations 的排序。

$\pi_i$ ：節點  $i$  的父代集合， $i$  介於 1 到  $n$  之間。

1. 以節點集合  $Z$  的完整圖(complete graph)  $G_1$  開始。

初始化  $ord$  為 0，對所有  $i$  介於 1 到  $n$  之間， $old\_pi_i$  為空集合， $old\_Prob$  為 0。

2. 修正  $G_1$ 。

- 每對  $G_1$  中相鄰的兩個點  $a$  和  $b$ ，若  $A_{G_1}ab$  有重要性大於或等於  $ord$ ，及  $I(a, S_{ab}, b)$ <sup>註</sup>，其中重要性  $ord$  的  $S_{ab}$  包含於  $A_{G_1}ab$ ，則移除  $a - b$  的連結。
- 儲存  $S_{ab}$ 。
- 若所有  $G_1$  中相鄰節點  $a$  和  $b$ ， $A_{G_1}ab$  重要性小於  $ord$ ，至步驟 10。
- 若  $G_1$  的  $degree$  大於  $u$ ，則  $ord+1$  後取代原來的  $ord$ ，重新開始步驟 2。

---

<sup>註</sup>符號  $I(S_1, S_2, S_3)$  表示  $S_2$  的存在使得  $S_1$  和  $S_3$  獨立。

3. 複製  $G_1$  , 令其為  $G$ 。

$G$  中每對不相鄰的變數  $a$  和  $b$  , 若有一個不在  $S_{ab}$  中的節點  $c$  , 及相鄰於  $a$  和  $b$  , 則給連結方向 , 如  $a \rightarrow c$  和  $b \rightarrow c$  。除了這樣的方向導致圖形中出現方向性循環(directed cycle)外 , 若連結已是反向 , 則給此連結成為雙向。

4. 若圖形中不會產生方向性循環 , 則嚐試指定方向給  $G$  中仍無方向的連結。

應用規則如下 :

規則 1 : 若  $a \rightarrow b$  和  $b - c$  , 及  $a$  和  $c$  不相鄰 , 則方向為  $b \rightarrow c$  。

規則 2 : 若  $a \rightarrow b$  ,  $b \rightarrow c$  , 和  $a - c$  , 則方向為  $a \rightarrow c$  。

規則 3 : 若  $a - b$  ,  $b - c$  ,  $b - d$  ,  $a \rightarrow d$  和  $c \rightarrow d$  , 則方向為  $b \rightarrow d$  。

規則 4 : 若  $a - b$  ,  $b - c$  ,  $a - c$  ,  $c - d$  和  $d \rightarrow a$  , 則方向為  $a \rightarrow b$  和  $c \rightarrow b$  。

此外 , 若  $a \rightarrow b$  ,  $b \rightarrow c$  , 和  $a \leftrightarrow c$  , 則方向為  $a \rightarrow c$  。

5. 令對所有  $i$  介於 1 到  $n$  之間 ,  $\pi_i$  為空集合。

對每個節點  $i$  , 將節點  $j$  的集合加入  $\pi_i$  , 使得局部非循環方向圖(Partially Directed Acyclic Graph, PDAG)  $G$  中的每個  $j$  有一個連結  $j \rightarrow i$  。

6. 局部非循環方向圖  $G$  中每個非方向或雙向的連結 , 選擇一個方向 , 如下 :

若非方向連結  $i - j$  ,  $\pi_i$  和  $G$  中相對應的父代集合  $\pi_j$  , 則計算下列的乘積 :

$$\begin{aligned} i_{\text{val}} &= g(i, \pi_i) \times g(j, \pi_j \cup \{i\})^{\text{註}} \\ j_{\text{val}} &= g(j, \pi_j) \times g(i, \pi_i \cup \{j\}) \end{aligned}$$

若  $i_{\text{val}}$  大於  $j_{\text{val}}$  , 則  $\pi_j \cup \{i\}$  取代原來的  $\pi_j$  。除連結的增加外 , 即  $i \rightarrow j$  , 導致局部非循環方向圖中的一個循環 , 選擇反向及改變  $\pi_i$  ( 取代  $\pi_j$  ) 。  $j_{\text{val}}$  大於  $i_{\text{val}}$  時 , 執行步驟一樣。

---

<sup>註</sup>符號  $g(i, \pi_i)$  , 見第 11 頁。

7. 步驟 6 所得之集合  $\pi_i$  , 對  $i$  介於 1 到  $n$  之間 , 定義非循環方向圖。因為每個節點  $i$  , 其  $\pi_i$  由方向性連結指向節點  $i$  的節點所組成。從非循環方向圖中執行拓撲式排序 , 產生一個節點的排序。
8. 應用 K2 演算法來找步驟 7 之排序中每個節點的父亲集合。設對所有  $i$  介於 1 到  $n$  之間 ,  $\pi_i$  表示節點  $i$  的父亲集合。  
 設  $new\_Prob = \prod_{i=1}^n g(i, \pi_i)$ 。
9. 若  $new\_Prob$  大於  $old\_Prob$  , 則  
 $new\_Prob$  取代  $old\_Prob$ 。  
 $ord+1$  取代原來的  $ord$ 。  
 對所有  $i$  介於 1 到  $n$  之間 ,  $\pi_i$  取代  $old\_pi_i$ 。  
 放棄  $G$ 。  
 回到步驟 2。  
 否則 , 跳到步驟 10。
10. 輸出  $old\_pi_i$  , 對所有  $i$  介於 1 到  $n$  之間。  
 輸出  $old\_Prob$ 。

## 2.2 矩陣分解

矩陣分解(Matrix decomposition or matrix factorization)是將矩陣拆解成數個三角矩陣, 依使用目的不同, 可分為三種矩陣分解法: (1) 三角分解法(Triangular Decomposition), (2) QR 分解法(QR Decomposition), (3) 奇異值分解法(Singular Value Decomposition)[4]。

### 1. 三角分解法(Triangular Decomposition)

三角分解法, 又稱為 LU 分解法, 是將一方陣分解成一個上三角形矩陣(upper triangular matrix)和一個下三角形矩陣(lower triangular matrix), 其用途主要是簡化一個大矩陣的行列式值的計算過程, 求反矩陣和求解聯立線性方程組。但要注意這種分解法所得到的上下三角形矩陣並非唯一, 還可找到數對不同的上下三角



形矩陣，上、下兩三角形矩陣相乘皆會得到原矩陣。舉以下一方陣  $\mathbf{A}$  為例：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -2 & -5 & 3 \\ -1 & -3 & 0 \end{bmatrix}$$

利用三角分解法可將  $\mathbf{A}$  矩陣分別拆解為上下三角形矩陣的乘積

$$\mathbf{A} = [\mathbf{L}][\mathbf{U}] = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix}。$$

## 2. QR 分解法(QR Decomposition)

QR 分解法是將矩陣分解成一個正規正交矩陣(orthonormal matrix)與上三角形矩陣，其中  $\mathbf{Q}$  代表正規正交矩陣， $\mathbf{R}$  代表上三角矩陣。如果矩陣大小為  $m \times n$ ，則  $\mathbf{Q}$  的大小為  $m \times m$ ， $\mathbf{R}$  的大小為  $m \times n$ 。例如：一  $3 \times 2$  的矩陣  $\mathbf{A}$ ，其正規正交矩陣  $\mathbf{Q}$  的大小為  $3 \times 3$ ，上三角矩陣  $\mathbf{R}$  的大小為  $3 \times 2$ 。

## 3. 奇異值分解法(Singular Value Decomposition)

奇異值分解(singular value decomposition, SVD)是另一種正交矩陣分解法，將矩陣拆解成兩個相互正交的矩陣及一個對角矩陣。用來解最小平方誤差法和數據壓縮。SVD 是最可靠的分解法，但是比 QR 分解法要花上近十倍的計算時間和 QR 分解法相同的地方為原矩陣亦沒有方陣的限制。

### 2.2.1 區塊矩陣(block matrix)

區塊矩陣是一種用較小矩陣來定義的矩陣，即矩陣被切割(partitioned)後所形成的矩陣。設一  $m \times n$  區塊矩陣  $\mathbf{A}$ ，其型式為

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \mathbf{A}_{m2} & \cdots & \mathbf{A}_{mn} \end{bmatrix} \quad (2.4.1)$$

其中， $\mathbf{A}_{ij}$  是  $\mathbf{A}$  的子矩陣(submatrix)[18]。

若兩個區塊矩陣有相同的型式，對角區塊是方陣，則區塊矩陣之加法(addition)與乘法(multiplication)與一般矩陣之加法與乘法相似。設  $\mathbf{A}$  和  $\mathbf{B}$  均為  $2 \times 2$  的區塊矩陣，則兩矩陣相加的結果如下[17]：

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \mathbf{A}_{12} + \mathbf{B}_{12} \\ \mathbf{A}_{21} + \mathbf{B}_{21} & \mathbf{A}_{22} + \mathbf{B}_{22} \end{bmatrix} \quad (2.4.2)$$

相乘的結果如下：

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix} \quad (2.4.3)$$

### 2.2.2 Jordan 標準型(Jordan canonical form)

Jordan 區塊(block)是一個除了對角線和上對角線的元素外，其它元素的值均為 0 的矩陣。此類矩陣稱為標準方格矩陣(canonical box matrices)，其對角線上的元素由單一特徵值所組成，上對角線的元素則由 1 組成[35]。其模型為

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \ddots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \quad (2.4.4)$$



## 2.3 矩陣與圖論

根據[22]。

### 定義 2.3.1

設  $\mathbf{A}=[a_{ij}]_{n \times n}$  為一  $n \times n$  矩陣，定義  $\mathbf{A}$  的圖形  $G(\mathbf{A})$  為含有點  $1, \dots, n$  的圖形，若  $a_{ij} \neq 0$ ，則從  $i$  到  $j$  存在一個弧線；反之亦然。

### 定義 2.3.2

設  $\mathbf{A}$  是一個  $q \times q$  區塊矩陣，其型式為

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1q} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{q1} & \mathbf{A}_{q2} & \cdots & \mathbf{A}_{qq} \end{bmatrix} \quad (2.5.1)$$

區塊矩陣  $\mathbf{A}$  的縮約圖形  $R(\mathbf{A})$  為含有點  $1, \dots, q$  的圖形，若  $\mathbf{A}_{ij} \neq \mathbf{O}$ ，則從  $i$  到  $j$  存在一個弧線；反之亦然。

### 定義 2.3.3

- (1) 設  $\mathbf{A}$  為一個  $q \times q$  區塊，其對角線的元素為方陣  $\mathbf{A}_{11}, \dots, \mathbf{A}_{qq}$ 。若矩陣  $\mathbf{A}_{ii}$  為奇異(singular)矩陣，則縮約圖形  $R(\mathbf{A})$  中的點  $i$  為奇異點(singular vertex)。
- (2) 縮約圖形  $R(\mathbf{A})$  中，奇異點  $i$  的階層(level)為終止(terminate)於  $i$  的路徑中奇異點的最大數量。
- (3) 設  $m$  為縮約圖形  $R(\mathbf{A})$  中一個奇異點的最大階層。 $\mathbf{A}$  的階層特徵值(level characteristic)  $\chi(\mathbf{A})$ ，定義為  $(\chi_1, \dots, \chi_m)$  序列，其中  $\chi_k$  是階層  $k$  之奇異點的個數。

### 定義 2.3.4

- (1) 設一個  $n \times n$  矩陣  $\mathbf{A}$ ，若圖形  $G(\mathbf{A})$  有強的連結，則  $\mathbf{A}$  是不可約的(irreducible)。即不論  $n=1$  或對所有  $i, j \in \{1, \dots, n\}$ ， $G(\mathbf{A})$  中存在一條從  $i$  到  $j$  的路徑。
- (2) 每個  $n \times n$  矩陣  $\mathbf{A}$  可以藉由置換相似於一個(下)區塊三角型式(triangular form)

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p1} & \mathbf{A}_{p2} & \cdots & \mathbf{A}_{pp} \end{bmatrix} \quad (2.5.2)$$

其中對角區塊  $\mathbf{A}_{11}, \dots, \mathbf{A}_{pp}$  是不可約的方陣。型式(2.5.2)稱為矩陣  $\mathbf{A}$  的 Frobenius 標準型(normal form)。

### 註釋 2.3.5

對應於矩陣  $\mathbf{A}$  之 Frobenius 標準型的縮約圖形依點的重新標註 (vertex relabeling) 被唯一決定 (uniquely determined)。

### 定義 2.3.6

- (1) 設  $\mathbf{A}$  為一個方陣,  $n_1, \dots, n_t$  為對應於  $\mathbf{A}$  之 0 特徵值的 Jordan 區塊大小。定義  $\mathbf{A}$  的奇異 Jordan 圖形 (singular Jordan graph)  $SJ(\mathbf{A})$  為包含不循環點 (loopless vertices), 長度  $n_1, \dots, n_t$  的  $t$  個不相交路徑 (disjoint paths), 稱為奇異路徑 (singular paths) 及  $rank(\mathbf{A})$  個有循環點的圖形。
- (2) 設  $\mathbf{A}$  為一個  $q \times q$  區塊矩陣, 其對角線的元素為方陣  $\mathbf{A}_{11}, \dots, \mathbf{A}_{qq}$ 。定義圖形  $GSJ(\mathbf{A})$  如下:

視  $SJ(\mathbf{A}_{11}), \dots, SJ(\mathbf{A}_{qq})$  為  $q$  個不相交圖形 (disjoint graph)。若  $\mathbf{A}_{ij} = \mathbf{O}$ , 對所有  $i > j$ , 則從  $SJ(\mathbf{A}_{ii})$  的每個點到  $SJ(\mathbf{A}_{jj})$  的每個點有一弧線。

根據[2]。

### 定義 2.3.7

一方陣  $\mathbf{A}$ , 若存在一個矩陣  $\mathbf{P}$ , 使得  $\mathbf{P}^t \mathbf{A} \mathbf{P}$  是一個上 (或下) 三角 (嚴格三角) 矩陣, 則  $\mathbf{A}$  為本質三角 (嚴格三角) 矩陣。

### 定義 2.3.8

一方陣  $\mathbf{A}$  是本質三角矩陣 (本質嚴格下三角矩陣) 若且唯若  $G(\mathbf{A})$  是非循環的方向圖。

### 定義 2.3.9

令  $\mathbf{A}$  是一個  $n \times n$  矩陣，設  $G(\mathbf{A})$  有  $r$  個強元素，命名為  $\alpha_1, \dots, \alpha_r$ 。 $\mathbf{A}$  的縮約圖形(reduced graph)，記作  $R(\mathbf{A})$ ，即含頂點(vertices)集合  $\{\alpha_1, \dots, \alpha_r\}$  的方向圖，使得存在一個弧線(arcs)  $(\alpha_i, \alpha_j)$  若且唯若  $\mathbf{A}$  之子矩陣  $A_{\alpha_i, \alpha_j}$  是非零的。 $R(\mathbf{A})$  亦是非循環方向圖。

### 定義 2.3.10

令  $\mathbf{A}$  是一個  $n \times n$  矩陣。 $\alpha$  和  $\beta$  是  $R(\mathbf{A})$  的兩個頂點。若  $\alpha = \beta$  或  $G(\mathbf{A})$  中存在一個鏈從  $\alpha$  中之某個頂點到  $\beta$  中之某個頂點，則  $\alpha$  有到達  $\beta$  (access to)。若  $i$  是  $G(\mathbf{A})$  的一個頂點， $\alpha$  是  $R(\mathbf{A})$  的一個頂點，則  $i$  有到達  $\alpha$ 。

根據[12]。

### 定義 2.3.11

若對應矩陣  $\mathbf{A}$  的方向圖  $D(\mathbf{A})$  中沒有弧線指向點  $i$ ，則點  $i$  為起始點(initial vertex)。

若對應矩陣  $\mathbf{A}$  的方向圖  $D(\mathbf{A})$  中沒有弧線離開點  $j$ ，則點  $j$  為最終點(final vertex)。

根據[23]。

### 定義 2.3.12

設一矩陣  $\mathbf{A}$ ，圖形  $G(\mathbf{A})$  定義為具有點集合  $\langle n \rangle$  的圖形，使得

若  $a_{ij} > 0$ ，則從  $i$  到  $j$  存在一個弧線。

假設  $\mathbf{A}$  是被分割的(partitioned)下三角  $r \times r$  區塊，其對角線部分是方陣。

### 定義 2.3.13

設一矩陣  $\mathbf{A}$ ，縮約圖形  $R(\mathbf{A})$  定義為具有點集合  $\langle r \rangle$  的圖形，使得若  $A_{ij} > 0$ ，則從  $i$  到  $j$  存在一個弧線。 $\mathbf{A}$  是下三角區塊矩陣，因此， $R(\mathbf{A})$  是三角圖形。

#### 定義 2.3.14

假設  $\mathbf{A}$  是被分割的(partitioned)下三角  $r \times r$  區塊，若  $\mathbf{A}_{ii}$  是奇異的(singular)，則縮約圖形  $R(\mathbf{A})$  中的點  $i$  為奇異點。 $R(\mathbf{A})$  中所有奇異點的集合，以  $S$  表示。

#### 輔助定理 2.3.15

設  $\mathbf{A}$  是一個  $n \times n$  的下三角矩陣。若  $\mathbf{A}$  有一個  $k \times k$  的非奇異(non-singular)子矩陣，則  $\mathbf{A}$  的奇異點形成一個奇異( $n-k$ )路徑(path)。

#### 註釋 2.3.16

設  $D_0(\mathbf{A})$  為包含點集合  $S$  的圖形，使得  $D_0(\mathbf{A})$  的弧線從點  $i$  到點  $j$  與  $G(\mathbf{A})$  的路徑從  $R(\mathbf{A})$  中的  $i$  到  $j$ ，這些路徑的中間點不屬於  $S$ 。即  $G(\mathbf{A})$  中每個奇異  $t$ -路徑為  $D_0(\mathbf{A})$  中  $(S, t)$ -路徑。(點集合  $S$ ，為  $R(\mathbf{A})$  中所有奇異點的集合)

根據[28]。

#### 註釋 2.3.17

若排列 Jordan 基底，使得每個特徵值所對應之 Jordan 區塊沿著對角線遞減，即將特徵值排序，則 Jordan 型式是唯一。

根據[37]。

#### 定義 2.3.18

設兩矩陣  $\mathbf{A}$  和  $\mathbf{B}$ ，若存在一個非奇異方陣  $\mathbf{X}$  使得  $\mathbf{A} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1}$ ，則  $\mathbf{A}$  和  $\mathbf{B}$  相似。映射(mapping)  $\mathbf{B} \rightarrow \mathbf{A}$ ，稱為相似性轉換(similarity transformation)。本研究將非奇異方陣  $\mathbf{X}$  稱為轉換矩陣。

### 第三章 線性代數與圖論之結合

為了能驗證最終所產生的貝氏網路架構是否合適，本研究的驗證方法採用實驗組與對照組方式進行。實驗組為本研究方法所建立之貝氏網路架構，對照組則為已知的貝氏網路架構。用對照組產生本研究的資料來源，執行本研究方法建立貝氏網路。最後進行實驗組與對照組的相似性分析(similarity analysis)。本章將說明資料的產生方法，本研究所提出的研究方法及實驗組與對照組之間的分析，如圖 3.1。

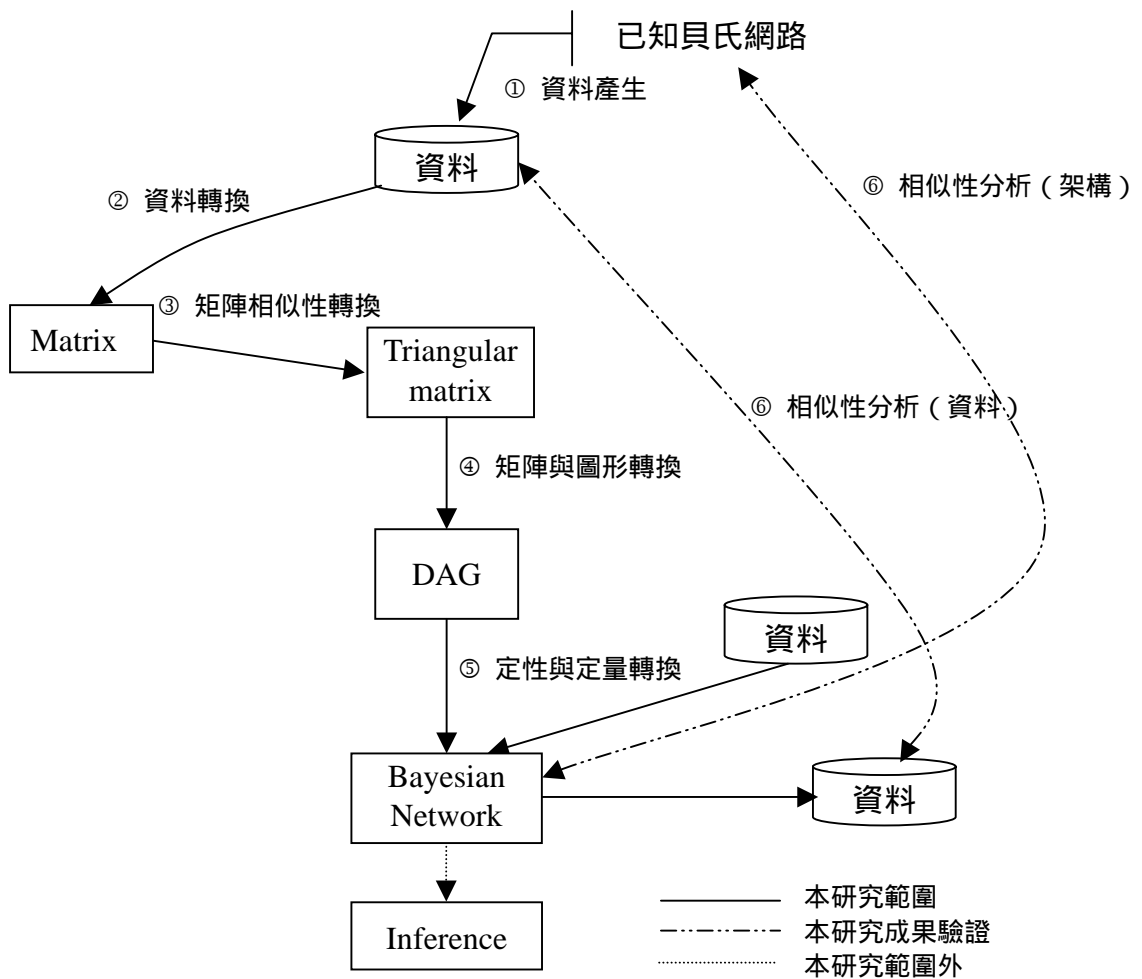


圖 3.1 本研究之研究方法



### 3.1 資料產生

本研究方法的步驟一，如圖 3.1 ①。以已知的貝氏網路架構為範例，利用貝氏網路應用軟體，產生一組資料，做為本研究的資料來源。為避免研究的結果因使用全部的資料量建立貝氏網路架構，再使用相同的資料做驗證，使驗證的結果失真。所以使用部分的資料做為建立貝氏網路架構的資料來源，剩下的資料則做為將非循環方向圖轉換為貝氏網路時所需的觀測資料。本研究擬以一般常用的資料分割方式，將三分之二的資料量做為建立架構的資料來源，剩下的三分之一資料量做為觀測資料<sup>註</sup>。圖 3.2 為資料分割的示意圖。

### 3.2 資料轉換

本研究方法的第二步驟，如圖 3.1 ②。以步驟一中產生的資料，將資料轉換成矩陣型式。

#### 3.2.1 矩陣型式的設計

每個問題領域中，存在與問題領域有關的節點（變數）集合。節點集合中，每個節點有屬於自己的狀態（屬性）集合。若將節點視為一個物件<sup>註</sup>，則節點所屬的狀態集合視為物件的屬性群。（如圖 3.3）

區塊矩陣是一種用較小矩陣來定義的矩陣，即矩陣被切割後所形成的矩陣。設一  $n \times n$  區塊方陣  $\mathbf{A}$ ，其型式為

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{n1} & \mathbf{A}_{n2} & \cdots & \mathbf{A}_{nn} \end{bmatrix} \quad (3.2.1)$$

其中， $\mathbf{A}_{ij}$  是  $\mathbf{A}$  的子矩陣。

---

<sup>註</sup>使用多少比例的資料量做為建構資料，多少比例資料量做為驗證資料，沒有一定的標準，視所擁有的資料量而定，例如：75%為建構資料，25%為驗證資料。

<sup>註</sup>物件，任何可以描述或有能力被描述的東西。例如：現實生活中有關人、事、物等任何實體的物體，或抽象的物體。

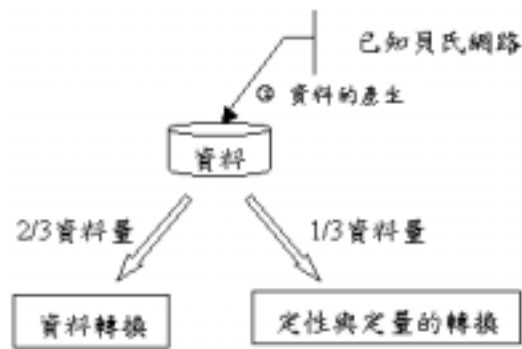


圖 3.2 資料分割示意圖

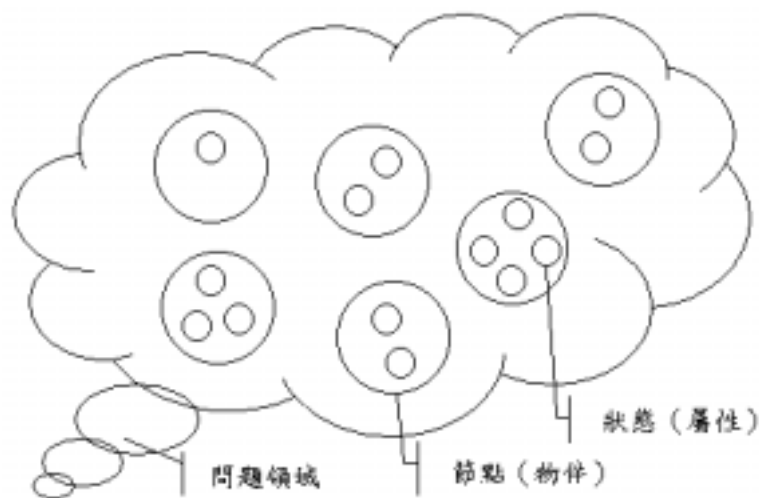


圖 3.3 矩陣型式設計之概念圖

將矩陣中每個對角線上的子矩陣視為由一個物件的屬性群所形成的矩陣型式，對應問題領域中的一個節點。因物件本身相當於一個個體，使物件與物件之間具有不相交性，所以相對應的節點亦具有不相交性。其餘子矩陣則為物件與物件之間的對應關係，這些子矩陣中的各項元素值則為任一物件內部屬性與另一物件內部屬性相對應的關係。

依上所述，本研究的矩陣設計方式如下：

假設取得的資料中有五個節點（物件），分別為 A、B、C、D 和 E，每個節點各有兩個狀態（屬性），分別為 Y 和 N，其中每個節點的狀態可以不同。將每個節點與個別對應的狀態視為矩陣之列與矩陣之行的項目。矩陣型式的設計如表 3.1 所示。

表 3.1 矩陣的型式

		A		B		C		D		E	
		Y	N	Y	N	Y	N	Y	N	Y	N
A	Y										
	N										
B	Y										
	N										
C	Y										
	N										
D	Y										
	N										
E	Y										
	N										

由表 3.1 可看出所設計的矩陣為一個 10×10 的矩陣，以  $M$  表示此矩陣。矩陣的元素是由列之節點的各個狀態對應行之節點的各個狀態所組成，即每個不同狀態之間的對應，共 10×10 個元素。

此外，這個矩陣亦可以視為一個 5×5 的區塊矩陣，由矩陣之列節點與矩陣之行節點之間的對應所形成。如矩陣之列項目 C 與矩陣之行項目 A 之間所形成的矩陣，為區塊矩陣的第三列第一行，相當於由一般矩陣  $M$  的第五列第一行、第五列第二行、第六列第一行及第六列第二行所形成的子矩陣（表 3.1 中粗實框線的部分）。

### 3.2.2 矩陣的元素值

上一節中已說明矩陣元素的組成方式。此一節中，將針對該組成方式做更詳細的說明。

貝氏網路具有因果關係的特質，以及利用事前機率(prior probability)與變數間兩者關係的條件機率來推論貝氏網路的參數值。有關條件機率的定義如下：

設  $A$  和  $B$  為樣本空間中的兩個事件， $B$  的機率  $P(B)$  大於 0。給定事件  $B$  發生的情況下， $A$  的條件機率  $P(A|B)$  則定義為

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (3.1)$$

為了使由資料轉換之矩陣內的元素亦具有因果關係的特質，所以本研究以矩陣之列項目做為條件機率中的給定對象，如方程式(3.1)中的事件  $B$ ；矩陣之行項目則做為欲求的對象，如方程式(3.1)中的事件  $A$ ；給定矩陣之列項目，則矩陣之行項目的條件機率，如方程式(3.1)中的條件機率  $P(A|B)$ 。

因此本研究方法矩陣元素值的形成方式如下：

以矩陣之列項目為給定的條件，計算出矩陣之行項目的條件機率。例如：矩陣第三列第一行（表 3.1 中粗虛框線的部分）的元素為節點  $B$  之狀態  $Y$  對應節點  $A$  之狀態  $Y$  所組成。此元素值為資料中節點  $B$  之狀態  $Y$  為給定的條件，計算節點  $A$  之狀態  $Y$  的條件機率。

假設今有一組資料  $D$ ，含有 8 筆案例，如表 3.2 所示。因此表 3.1 中，節點  $B$  之狀態為  $Y$  的機率是  $5/8 = 0.625$ 。若給定節點  $B$  之狀態為  $Y$ ，節點  $A$  之狀態為  $Y$  的條件機率是  $\frac{3/5}{8/8} = \frac{3}{5} = 0.6$ 。因此，表 3.1 中第三列第一行的元素值為 0.6。

### 3.3 矩陣相似性轉換

本研究方法的第三步驟，如圖 3.1 ③。透過相似性轉換，將矩陣轉換成相似的三角矩陣。

表 3.2 資料型態

編號	A	B	C	D	E
1	Y	N	N	N	Y
2	Y	Y	Y	N	N
3	N	N	N	N	N
4	N	Y	Y	Y	Y
5	Y	Y	N	Y	Y
6	N	Y	Y	Y	Y
7	N	N	Y	N	Y
8	Y	Y	Y	Y	Y

一組問題領域的節點間，若任意兩個節點之間可能的關係均為連結，則稱此為完全連結(fully connect)。但並非所有這些可能的連結均實際存在。為了避免多餘的節點關係，增加問題的複雜度或造成錯誤的判斷，所以需要簡化節點間的關係。例如：貝氏網路藉由條件獨立的方式，移除不存在的連結關係或微弱的連結關係，以簡化網路架構。

為了在不影響問題領域本質的情況下，簡化節點間的連結關係，本研究方法透過矩陣的相似性轉換，於保留問題領域本質的情形下，來簡化節點間的連結關係。

根據定義 2.3.18，設兩個方陣  $A$  和  $B$ ，若存在一個非奇異方陣  $X$  使得  $A=XBX^{-1}$ ，則  $A$  和  $B$  相似。映射(mapping)  $B \rightarrow A$ ，稱為相似性轉換(similarity transformation)。即一個矩陣轉換時所需的轉換矩陣存在，可透過相似性轉換求得相似矩陣。

資料轉換中設計的矩陣與相似矩陣之間具有相同的關係 保留矩陣之列到矩陣之行的關係，相似矩陣也保留了原矩陣的特徵值，故一個矩陣的相似矩陣足以代替原矩陣。符合上述條件的轉換矩陣不只一個，為了避免相似性轉換的過程中，因轉換矩陣的影響將變數組合成新的變數，本研究以最簡化的矩陣型式做為相似性轉換時的轉換矩陣。對角矩陣即是最簡化的矩陣型式。

Jordan 標準型是由一個矩陣之特徵值及每個特徵值重複的個數所形成的一個標準型，近似對角矩陣，保有原矩陣的本質。若該矩陣的特徵向量線性獨立，則 Jordan 標準型為一個對角矩陣。此外，Jordan 標準型亦具有最簡化及唯一性的特質。

由於 Jordan 標準型是由特徵值所形成的一種矩陣型式，若 Jordan 標準型中有特徵值 0 出現，則節點間有些關係可能不存在或是非常微弱，表示該矩陣可以簡化節點間的關係。因此本研究以 Jordan 標準型做為本研究相似性轉換時的轉換矩陣。

因此，進行相似性轉換之前，須先找出欲執行相似性轉換之矩陣  $M$  的 Jordan 標準型  $J$ 。再以求得的 Jordan 標準型做為相似性轉換時所需要的轉換矩陣，使得  $M=J T J^{-1}$ ，求矩陣  $M$  的相似型。最後，以  $J^{-1} M J$  求得相似於矩陣  $M$  的相似矩陣(Similarity matrix)  $T$ 。

### 3.4 矩陣與圖形轉換

本研究方法的第四步驟，如圖 3.1 ④。透過矩陣與圖形的轉換，畫出非循環方向圖。

假設有一組有限節點的集合  $V=\{V_1, V_2, \dots, V_n\}$  和一組弧線的集合  $E$ ，則由集合  $V$  和集合  $E$  所組成的圖形為  $G=(V, E)$ 。圖 3.4 為一個圖形的示意圖，其節點集合為  $V=\{A, B, C, D, E, F\}$ ，弧線集合為  $E=\{(A, B), (A, C), (B, D), (C, E), (C, F)\}$ 。

根據圖論定義 2.3.6， $5 \times 5$  區塊矩陣  $M$  的對角線元素  $M_{11}, \dots, M_{55}$ ，均為  $2 \times 2$  的正方子矩陣。每個對角線元素均對應問題領域中的一個節點，節點之間不會有相交的情況，故可視為 5 個不相交圖形。若區塊矩陣  $M$  的元素  $M_{ij} = 0$ ，對所有  $i \neq j$ ，則從點  $i$  到點  $j$  有一個弧線，即表示點  $i$  為點  $j$  的父節點(parent node)。

本研究於資料轉換的步驟中，以條件機率的方式計算矩陣元素值，其矩陣之列項目與矩陣之行項目的對應關係為矩陣之列對應到矩陣之行。因圖論定義之關係為矩陣之列項目到矩陣之行項目，故資料轉換時，採條件機率的方式，將對應關係設計為矩陣之列項目對應到矩陣之行項目。若矩陣設計是，採條件機率的方式，將對應關係設計為矩陣之行項目對應到矩陣之列項目。則相似矩陣於轉換圖形時，其對應關係須修正為矩陣之行對應到矩陣之列。即區塊矩陣  $M$  的元素  $M_{ij} = 0$  時，對所有  $i \neq j$ ，則從點  $j$  到點  $i$  增加一個弧線，即表示點  $j$  為點  $i$  的父節點。

根據定義 2.3.11，若對應區塊矩陣  $M$  的方向圖  $D(M)$  中沒有弧線指向點  $i$ ，則點  $i$  為起始點；若沒有弧線離開點  $j$ ，則點  $j$  為最終點。以此判斷那些節點是起始點，那些節點是最終點。

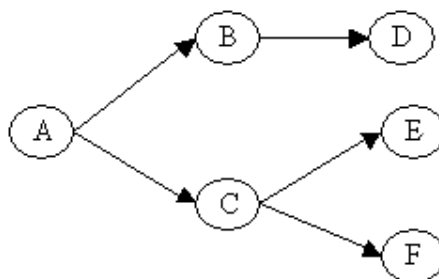


圖 3.4 圖形示意圖

此外，本研究對於中間節點的判斷方式有以下兩項規則：

規則一：根據註釋 2.3.16，本研究假設區塊矩陣  $M$  的對角線元素  $M_{ii} = 0$  時，則點  $i$  視為路徑的中間節點。

規則二：若點  $i$  到點  $j$ ，點  $j$  到點  $k$  及點  $i$  到點  $k$ ，則點  $j$  為點  $i$  和點  $k$  的中間點。

本研究矩陣與圖形的轉換，以矩陣相似性轉換求得之相似矩陣  $T$ ，根據上述的定義畫出非循環方向圖。

### 3.5 定性與定量轉換

本研究方法的第五步驟，如圖 3.1 ⑤。根據上一步驟得到的非循環方向圖和觀測資料，建立貝氏網路架構。此步驟的主要目的是將定性的非循環方向圖轉成定量的貝氏網路架構<sup>註</sup>。

定性的非循環方向圖中，只能描述節點間的連結關係。若將定性的非循環方向圖，透過訓練網路內部的參數（節點狀態的機率值），轉成定量的貝氏網路架構，則此網路架構可以用來做為診斷及預測工具。

本研究於資料產生時保留了三分之一的資料量做為此步驟的觀測資料。這些觀測資料即為貝氏網路參數訓練的來源，用以初始化步驟四產生之非循環方向圖的參數值。最終得到一個貝氏網路的雛型架構。

本研究方法的最終目的在於找出資料的貝氏網路架構，所以本研究將以貝氏網路應用軟體計算貝氏網路雛型的機率值，將定性的非方向循環圖，轉換成定量的貝氏網路架構。

### 3.6 相似性分析

此步驟為本研究方法產生之網路架構與範例間的相似性分析，如圖 3.1 ⑥。本研究以兩方面進行相似性分析，(1) 架構的相似性分析，(2) 資料結構的相似性。

---

<sup>註</sup>非循環方向圖是定性的，隨機的；貝氏網路是定量的，由資料的分配來決定。

## 架構的相似性分析

此部分是針對本研究方法產生之網路架構與範例間架構的相似性評估。本研究擬以 Jaccard 係數為架構相似性的評量法[33][36]。

假設  $S$  為一個架構圖， $i$  和  $j$  為  $S$  中的任兩個節點。此兩個節點間有一個方向性  $i \rightarrow j$  的弧線存在，以  $(i, j)$  表示。

若存在兩個架構圖  $S$  和  $T$ 。 $S$  和  $T$  中所有存在的弧線分別形成集合  $X$  和  $Y$ ，集中的弧線個數分別為  $x$  和  $y$ 。兩個集合共同弧線的個數為  $m$ ，則此兩個集合之間相似性的 Jaccard 係數為  $m/(x+y-m)$ 。

以下舉列說明之：

假設架構圖  $S$  和  $T$  中有節點  $A, B, C$  和  $D$ ，如圖 3.5 所示。 $S$  中所有存在之弧線形成的集合  $X = \{(A, B), (B, C), (A, C), (A, D)\}$ ，弧線個數  $x$  有 4 個； $T$  中所有存在之弧線形成的集合  $Y = \{(B, A), (B, C), (A, D), (B, D)\}$ ，弧線個數  $y$  有 4 個；兩個集合共同弧線的集合  $X \cap Y = \{(B, C), (A, D)\}$ ，弧線個數  $m$  有 2 個。則此兩個集合之間相似性的 Jaccard 係數為  $m/(x+y-m) = 2/(4+4-2) = 0.3333$ 。

## 資料的相似性分析

利用本研究方法建立之網路架構產生一組新資料，執行此組資料與原始資料間的相似性評估。

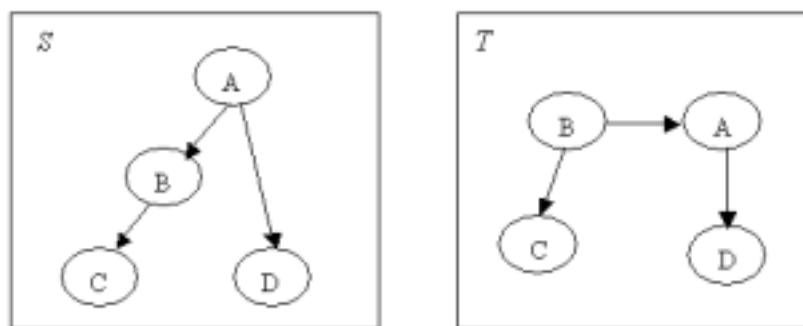


圖 3.5 舉列說明之架構圖  $S$  和  $T$



由於欲執行評估的(1) 目標是比較兩個母體，(2) 資料是量化的，不是常態分配，(3) 樣本是成對的。依據這樣的環境，本研究採用統計之無母數檢定法中成對比較的魏克森符號等級和檢定(Wilcoxon signed rank sum test)來執行資料的相似分析。

此相似性分析所檢定的是新資料與原始資料是否相同，故執行雙尾檢定。其兩個假設條件為

$H_0$ ：新資料與原始資料相同，即實驗組足以描述對照組

$H_1$ ：新資料與原始資料不相同，即實驗組不能描述對照組

假設兩個母體  $P_1$  和  $P_2$ ，其觀測值分別為  $x_1$  和  $x_2$ 。魏克森符號等級和檢定首先計算每對觀測值的差  $D = x_1 - x_2$ 。執行符號檢定時，消去  $D$  為 0 的值。接著，將  $D$  的絕對值由小到大排列，計算正差的等級和  $T^+$  及負差的等級和  $T^-$ 。任意選擇  $T^+$  來做檢定統計量，標記為  $T$ 。由於本研究欲檢定的總樣本數大於 30，故  $T$  近似常態分配。

其平均值為

$$E(T) = \frac{n(n+1)}{4} \quad (3.6.1)$$

標準差為

$$\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} \quad (3.6.2)$$

則其標準化的檢定統計量為

$$z = \frac{T - E(T)}{\sigma_T} \quad (3.6.3)$$

本研究以統計上一般常用的顯著水準 0.05 做為判定標準，執行雙尾的魏克森符號等級和檢定。若執行檢定後，求出的 p 值小於 0.025，則拒絕虛無假設  $H_0$ ，即實驗組不能描述對照組；反之，不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

## 第四章 實證研究

根據本研究的第二個研究假設，本研究以具有二元狀態的範例做為實證研究對象。所以本研究所舉範例之每個節點都擁有二元的狀態。此一章節中，除了以範例來說明本研究方法外，以搜尋式的演算法 K2 演算法 建立貝氏網路，再產生一組新資料，執行架構與資料的相似性分析。最後進行本研究方法與 K2 演算法的分析比較。

### 4.1 範例一：Judea Pearl 警鈴(Alarm)貝氏網路

#### 4.1.1 本研究方法建立之警鈴貝氏網路架構

假設你新買一個警鈴裝在家裡，它可以有效地偵測竊盜，但小地震發生時，警鈴也會產生反應。你有 John 和 Mary 兩個鄰居，他們都答應你如果他們有聽到鈴聲，會通知上班中的你。John 聽到鈴聲一定會通知你，但有時他會誤把電話聲當鈴聲，通知你。Mary 則是喜歡將音樂開得很大聲，因此有時會完全沒聽見鈴聲。

此範例的貝氏網路如圖 4.1。

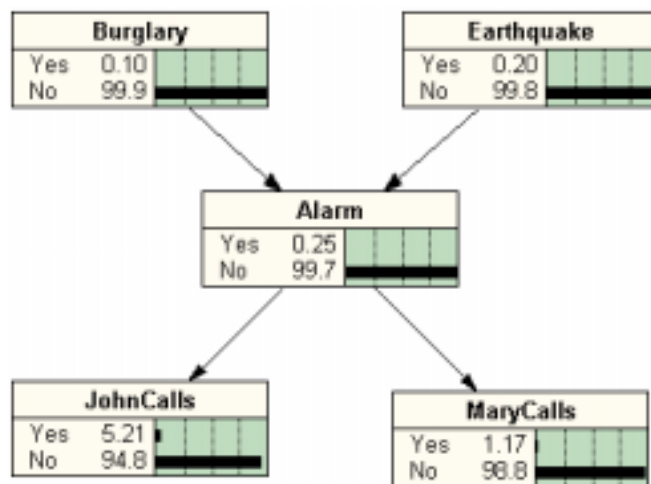


圖 4.1 Judea Pearl 的警鈴貝氏網路

Judea Pearl 的警鈴貝氏網路中，有五個節點 MaryCalls、JohnCalls、Alarm、Earthquake 和 Burglary，每個節點各有兩個狀態，分別為 Yes 和 No。

此範例的執行過程如下：

- (1) 利用 Netica 產生 1050 筆的資料<sup>註</sup>，做為此範例的資料來源。
- (2) 視每個節點與狀態為各列與各行的項目，透過 Microsoft Excel 處理(1)產生的資料之前三分之二 ( 700 筆 )，將資料轉成矩陣型式，可得一個 5×5 的區塊矩陣  $M$ ，如表 4.1 所示。其中矩陣之列與行的順序為 MaryCalls、JohnCalls、Alarm、Earthquake、Burglary。
- (3) 透過 Paul Godfrey[11]求 Jordan 矩陣分解的 MATLAB 程式找出區塊矩陣  $M$  的 Jordan 標準型  $J$ ，如表 4.2 所示。

表 4.1 警鈴貝氏網路之矩陣型式

		MaryCalls		JohnCalls		Alarm		Earthquake		Burglary	
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
MaryCalls	No	1	0	0.940751	0.059249	0.99711	0.00289	0.99711	0.00289	0.998555	0.001445
MaryCalls	Yes	0	1	0.625	0.375	0.625	0.375	1	0	0.875	0.125
JohnCalls	No	0.992378	0.007622	1	0	0.998476	0.001524	0.996951	0.003049	1	0
JohnCalls	Yes	0.931818	0.068182	0	1	0.909091	0.090909	1	0	0.954545	0.045455
Alarm	No	0.992806	0.007194	0.942446	0.057554	1	0	0.997122	0.002878	1	0
Alarm	Yes	0.4	0.6	0.2	0.8	0	1	1	0	0.6	0.4
Earthquake	No	0.988539	0.011461	0.936963	0.063037	0.992837	0.007163	1	0	0.997135	0.002865
Earthquake	Yes	1	0	1	0	1	0	0	1	1	0
Burglary	No	0.989971	0.010029	0.939828	0.060172	0.995702	0.004298	0.997135	0.002865	1	0
Burglary	Yes	0.5	0.5	0	1	0	1	1	0	0	1

表 4.2 警鈴貝氏網路之 Jordan 標準型

5	4.4E-16	0	0	0	0	0	0	0	0	0
0	2.037741	-4.48E-16	0	0	0	0	0	0	0	0
0	0	1.000816	-3.88E-16	0	0	0	0	0	0	0
0	0	0	0.88667	-4.51E-17	0	0	0	0	0	0
0	0	0	0	0.758702	-1.11E-16	0	0	0	0	0
0	0	0	0	0	0.316071	-2.28E-15	0	0	0	0
0	0	0	0	0	0	-5.43E-17	-1.81E-16	0	0	0
0	0	0	0	0	0	0	3.88E-17	-1.68E-16	0	0
0	0	0	0	0	0	0	0	-1.09E-16	2.48E-16	0
0	0	0	0	0	0	0	0	0	0	-2.11E-17

<sup>註</sup>為了驗證資料量適合特殊問題領域，如：醫療產業，本研究的資料量設定為 1050 筆。

- (4) 用求得的  $J$ ，做為相似性轉換時所需要的轉換矩陣，求  $J^{-1}MJ$ ，可得相似於  $M$  的下三角矩陣  $T$ 。如表 4.3 所示，其中對角線上的粗虛線框表示該點為中間點，其餘粗實線框表示兩變數間有關係存在。
- (5) 根據(4)的結果  $T$  及圖論的定義畫出非循環方向圖。
- 區塊矩陣  $T$  中，子矩陣  $T_{41}$ 、 $T_{51}$ 、 $T_{42}$ 、 $T_{52}$ 、 $T_{43}$  及  $T_{53}$  不為零矩陣。
  - 由定義 2.3.2 及定義 2.3.6 可知，區塊矩陣  $T$  的縮約圖形中，存在弧線  $4 \rightarrow 1$ 、 $5 \rightarrow 1$ 、 $4 \rightarrow 2$ 、 $5 \rightarrow 2$ 、 $4 \rightarrow 3$  及  $5 \rightarrow 3$ 。即 Earthquake  $\rightarrow$  MaryCalls、Burglary  $\rightarrow$  MaryCalls、Earthquake  $\rightarrow$  JohnCalls、Burglary  $\rightarrow$  JohnCalls、Earthquake  $\rightarrow$  Alarm 及 Burglary  $\rightarrow$  Alarm。
- (6) 將(5)的結果以圖形的方式表示，得到圖 4.2。
- (7) 架構的相似性分析
- 此範例的貝氏網路中，所有存在之弧線形成的集合  $X = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls})\}$ 。  $X$  中的弧線個數  $x$  為 8 個。
  - 本研究的最終結果中，所有存在之弧線形成的集合  $Y = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls})\}$ 。  $Y$  中的弧線個數  $y$  為 6 個。
  - 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  $X \cap Y = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 6 個。
  - 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{6}{8 + 6 - 6} = 0.750$$

由此可知，兩個架構的相似程度為 75.0%。

表 4.3 警鈴貝氏網路之相似矩陣

		MaryCalls		JohnCalls		Alarm		Earthquake		Burglary	
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
MaryCalls	No	1	0	0.188304	0.010507	0.151302	0.000183	-1.21E-17	-3.6E-17	-2.18E-17	4.96E-17
MaryCalls	Yes	1.09E-15	1	0.306963	0.163172	0.232703	0.058166	-4.45E-16	-8.86E-17	-4.67E-17	1.05E-16
JohnCalls	No	4.957845	0.015519	1	-4.93E-32	0.756928	0.000481	-5.75E-17	-1.8E-16	-1.09E-16	2.48E-16
JohnCalls	Yes	5.254592	0.156695	2.88E-17	1	0.777887	0.032406	-2.95E-16	-2.04E-16	-1.17E-16	2.66E-16
Alarm	No	6.54279	0.019322	1.243195	0.067261	1	5.2E-13	-7.13E-17	-2.37E-16	-1.44E-16	3.27E-16
Alarm	Yes	29099.98	11475.67	156.8398	10014.81	153.2725	3555.654	-2.63E-11	-2.2E-12	8.08E-14	-1.85E-13
Earthquake	No	4.04E+18	1.59E+18	2.17E+16	1.39E+18	2.13E+16	4.94E+17	-3645.592	-305.8068	11.24635	-25.82242
Earthquake	Yes	-1.24E+18	-4.79E+17	-1.17E+16	-4.18E+17	-1.06E+16	-1.48E+17	1096.355	92.93382	-2.780761	6.392889
Burglary	No	-3.16E+17	-1.11E+17	-8.65E+15	-9.65E+16	-6.95E+15	-3.43E+16	253.0289	21.21751	1.004434	0
Burglary	Yes	-1.19E+17	-4.83E+16	10.61782	-4.21E+16	2.140026	-1.5E+16	110.5647	8.567789	0	1

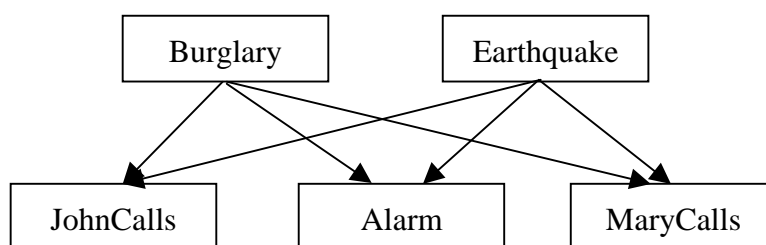


圖 4.2 警鈴貝氏網路之非循環方向圖

為了簡化圖形架構，去除多餘的變數關係。故本研究於此範例中，建立子代節點集合  $C_{mja} = \{ \text{MaryCalls}, \text{JohnCalls}, \text{Alarm} \}$  的子圖形。結合圖 4.2 和子圖形，形成此範例的非循環方向圖。

(8) 重複(2)到(6)相同的步驟，找出子代節點集合

$$C_{mja} = \{ \text{MaryCalls}, \text{JohnCalls}, \text{Alarm} \}.$$

(9) 由(8)和(2)可知子圖形有 MaryCalls、JohnCalls 和 Alarm，每個節點各有兩個狀態，分別為 Yes 和 No。可得一個  $3 \times 3$  的區塊矩陣  $M_{mja}$ ，如表 4.4 所示。

(10) 找出區塊矩陣  $M_{mja}$  的 Jordan 標準型  $J_{mja}$ ，如表 4.5 所示。

表 4.4 警鈴貝氏網路之子矩陣  $M_{mja}$

Amja		MaryCalls	MaryCalls	JohnCalls	JohnCalls	Alarm	Alarm
		No	Yes	No	Yes	No	Yes
MaryCalls	No	1	0	0.188304	0.010507	0.151302	0.000183
MaryCalls	Yes	1.09E-15	1	0.306963	0.163172	0.232703	0.058166
JohnCalls	No	4.957845	0.015519	1	-4.93E-32	0.756928	0.000481
JohnCalls	Yes	5.254592	0.156695	2.88E-17	1	0.777887	0.032406
Alarm	No	6.54279	0.019322	1.243195	0.067261	1	5.2E-13
Alarm	Yes	29099.98	11475.67	156.8398	10014.81	153.2725	3555.654

表 4.5 警鈴貝氏網路之子矩陣  $M_{mja}$  的 Jordan 標準型

3	0	0	0	0	0
0	1.6014	0	0	0	0
0	0	0.8858	0	0	0
0	0	0	0.5128	0	0
0	0	0	0	-2E-16	0
0	0	0	0	0	1E-16

- (11) 用求得的  $J_{mja}$ ，做為相似性轉換時所需要的轉換矩陣，求  $J_{mja}^{-1} M_{mja} J_{mja}$ ，可得相似於  $M_{mja}$  的下三角矩陣  $T_{mja}$ 。如表 4.6 所示，其中對角線上的粗虛線框表示該點為中間點，其餘粗實線框表示兩變數間有關係存在。
- (12) 根據(11)的結果  $T_{mja}$  及圖論的定義畫出非循環方向圖。
- 區塊矩陣  $T$  中，子矩陣  $T_{31}$  及  $T_{32}$  不為零矩陣。
  - 由定義 2.3.2 及定義 2.3.6 可知，區塊矩陣  $T$  的縮約圖形中，存在弧線  $3 \rightarrow 1$  及  $3 \rightarrow 2$ 。即 Alarm  $\rightarrow$  MaryCalls 及 Alarm  $\rightarrow$  JohnCalls。
- (13) 將(12)的結果以圖形的方式表示，得到圖 4.3。
- (14) 結合圖 4.2 和圖 4.3，可得圖 4.4。
- (15) 利用 Netica，結合(14)得到的非循環方向圖與(1)剩下的三分之一資料量（350 筆），建構貝氏網路，最終得到圖 4.5，貝氏網路架構。

表 4.6 警鈴貝氏網路之子矩陣  $M_{mja}$  的相似矩陣

		MaryCalls		JohnCalls		Alarm	
		No	Yes	No	Yes	No	Yes
MaryCalls	No	1	0	0.277775	0.010127	-6.61E-17	1.24E-19
MaryCalls	Yes	0	1	0.345711	0.120073	-7.76E-17	3.02E-17
JohnCalls	No	3.360922	0.013779	1	0	-2.24E-16	2.22E-19
JohnCalls	Yes	5.451713	0.212939	0	1	-3.52E-16	2.29E-17
Alarm	No	-1.5E+16	-5.8E+13	-4.2E+15	-1.48E+14	1	0
Alarm	Yes	9.3E+15	7.44E+15	1.37E+15	3.18E+15	0	1

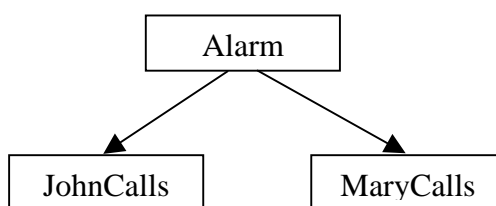


圖 4.3 警鈴貝氏網路之子非循環方向圖

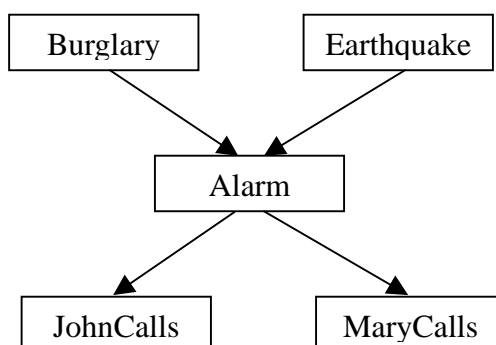


圖 4.4 警鈴貝氏網路之非循環方向圖

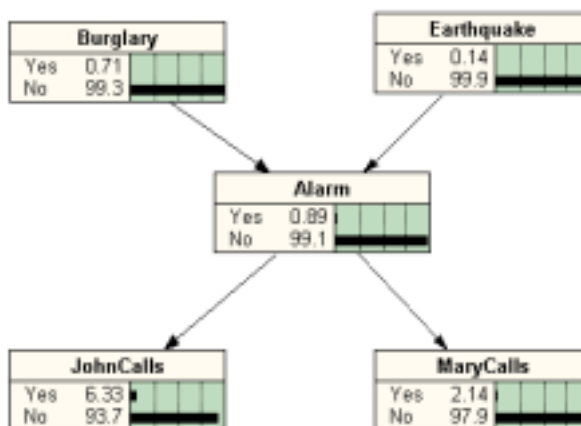


圖 4.5 本研究方法建立之警鈴貝氏網路

(16) 架構的相似性分析

- a. 此範例的貝氏網路中，所有存在之弧線形成的集合  $X = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls})\}$ 。  $X$  中的弧線個數  $x$  為 8 個。
- b. 本研究的最終結果中，所有存在之弧線形成的集合  $Y = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls})\}$ 。  $Y$  中的弧線個數  $y$  為 8 個。
- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  $X \cap Y = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 8 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{8}{8 + 8 - 8} = 1$$

由此可知，兩個架構的相似程度為 100%。

(17) 資料的相似性分析

- a. 顯著水準 0.05 之假設檢定的兩個假設條件為
- $H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組  
 $H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組
- b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。（本研究方法的警鈴貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄一的表 A-1.1。）
- c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的  $p$  值。計算結果如圖 4.6 所示，



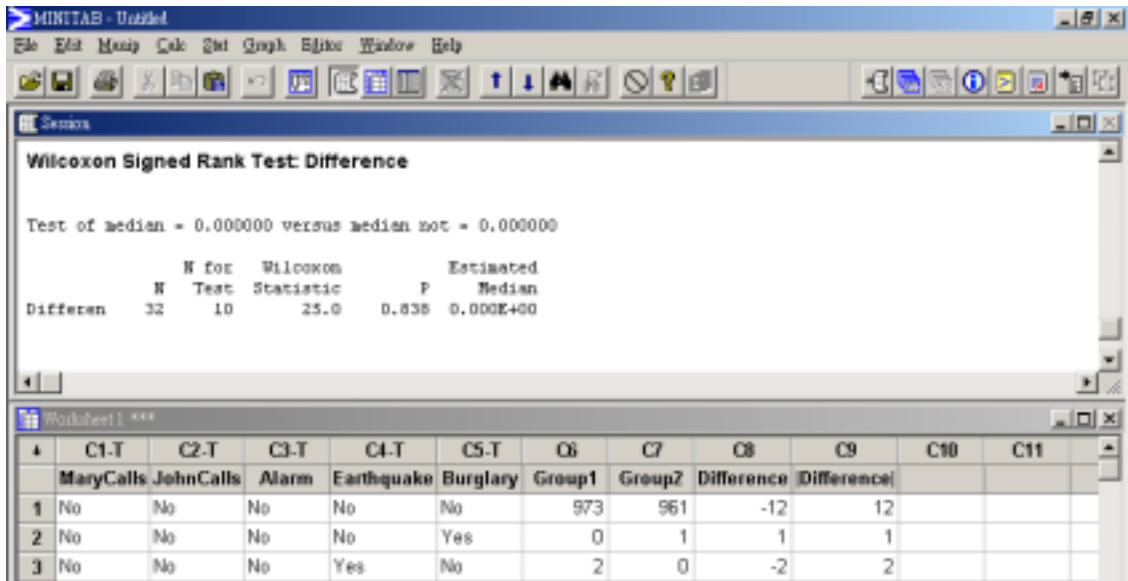


圖 4.6 本研究方法的警鈴貝氏網路：魏克森符號等級和檢定

p 值為 0.838 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

此範例由架構的相似性分析可知，本研究方法建立之此範例架構與已知架構間的相似度為 100%，表示本研究方法建立之此範例架構可以完全表達此範例變數間的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示本研究方法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

但是由(7)和(16)的架構相似性分析可知，本研究方法於未簡化圖形架構前，所得架構與已知架構間的相似度為 75.0%；簡化圖形架構後，所得架構與已知架構間的相似度為 100%。此現象說明，本研究方法於此範例中，簡化前的架構雖不是很理想，但以相同的方法簡化架構，架構會比簡化前更符合已知的貝氏網路架構。

## 4.1.2 K2 演算法建立之警鈴貝氏網路架構

此一節中，本研究擬以 K2 演算法產生此範例之架構，進行架構與資料的相似性評估。由於 K2 演算法執行前須給定一組節點的排序，因此給定一組與本研究相同的排序，則節點排序為 MaryCalls、JohnCalls、Alarm、Earthquake、Burglary。

透過貝氏網路應用軟體 Bayesware Discoverer 執行此一步驟，可得以下之結果：

- (1) K2 演算法產生之非循環方向圖，如圖 4.7。
- (2) 利用 Netica，結合(1)得到的非循環方向圖與 4.1.1 之(1)剩下的三分之一資料量，建構貝氏網路，最終得到圖 4.8，貝氏網路架構。

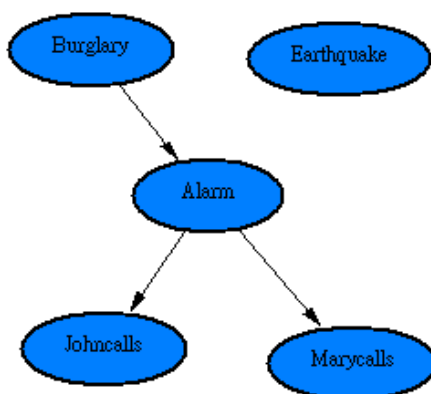


圖 4.7 K2 演算法建立的警鈴貝氏網路之非循環方向圖

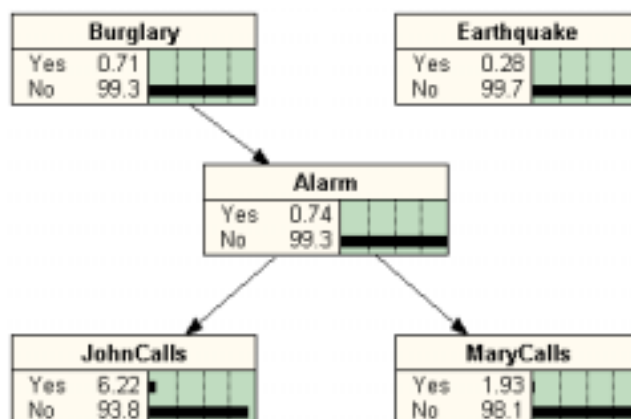


圖 4.8 K2 演算法建立之警鈴貝氏網路架構

### (3) 架構的相似性分析

- a. 此範例的貝氏網路中，所有存在之弧線形成的集合  
 $X = \{(\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Burglary}, \text{MaryCalls}), (\text{Earthquake}, \text{Alarm}), (\text{Earthquake}, \text{JohnCalls}), (\text{Earthquake}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls})\}$ 。  $X$  中的弧線個數  $x$  為 8 個。
- b. K2 演算法產生之貝氏網路架構中，所有存在之弧線形成的集合  $Y = \{(\text{Burglary}, \text{MaryCalls}), (\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls})\}$ 。  $Y$  中的弧線個數  $y$  為 5 個。
- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  
 $X \cap Y = \{(\text{Burglary}, \text{MaryCalls}), (\text{Burglary}, \text{Alarm}), (\text{Burglary}, \text{JohnCalls}), (\text{Alarm}, \text{MaryCalls}), (\text{Alarm}, \text{JohnCalls})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 5 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{5}{8 + 5 - 5} = 0.625$$

由此可知，兩個架構的相似程度為 62.5%。

### (4) 資料的相似性分析

- a. 顯著水準 0.05 之假設檢定的兩個假設條件為  
 $H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組  
 $H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組
- b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。( K2 演算法的警鈴貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄一的表 A-1.2。 )
- c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的 p 值。計算結果如所示，p 值為 0.308 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

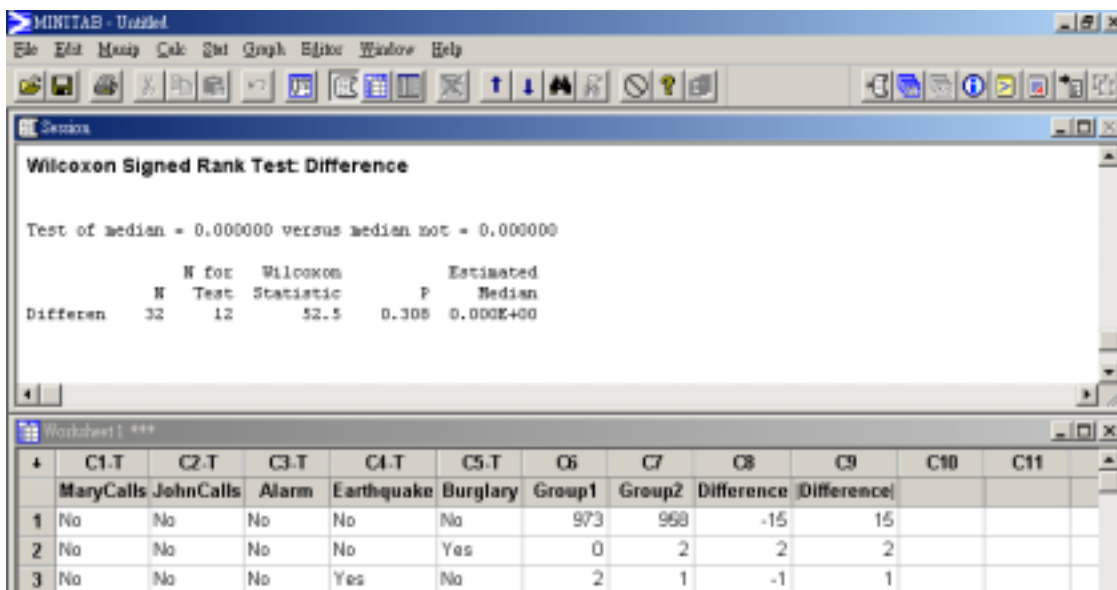


圖 4.9 K2 演算法的警鈴貝氏網路：魏克森符號等級和檢定

由架構的相似性分析可知，K2 演算法建立之此範例架構與已知架構間的相似度為 62.5%，表示 K2 演算法表達此範例變數間 62.5% 的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示 K2 演算法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

## 4.2 範例二：L. D. Phillips 和 W. Edward 文具袋(Book Bags) 貝氏網路

### 4.2.1 本研究方法建立之文具袋貝氏網路架構

有兩個文具袋，每個文具袋中裝有 10 個撲克籌碼。一個袋子中 7 個紅的和 3 個藍的，另一個袋子中有 3 個紅的和 7 個藍的。五個籌碼從其中一個袋中抽出，秀給實驗者檢查（一次一個籌碼，然後放回袋中），實驗者不知道籌碼是從那個袋子中抽出。每個袋子被抽到的機會相等。每一次抽籤後，實驗者回報他認為下一個籌碼會從那一個袋子抽出，機率有多少。

此範例的貝氏網路如圖 4.10。

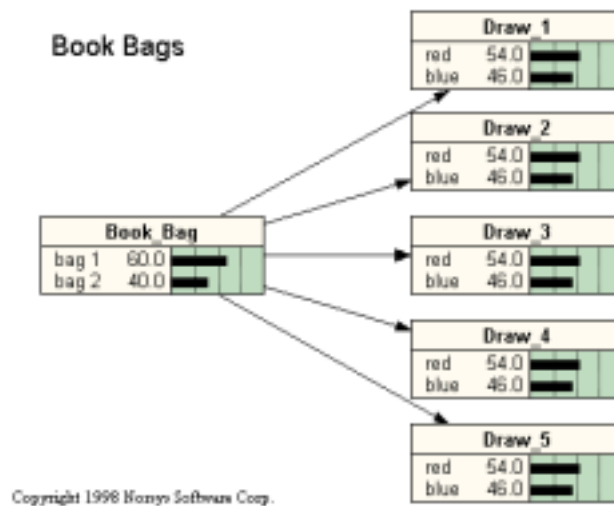


圖 4.10 L. D. Phillips and W. Edward 的文具袋貝氏網路

L. D. Phillips and W. Edward 的文具袋貝氏網路中，有六個節點 Book\_Bag、Draw\_1、Draw\_2、Draw\_3、Draw\_4 和 Draw\_5，每個節點除 Book\_Bag 為 bag\_1 及 bag\_2 外，其餘各項各有兩個狀態，分別為 red 和 blue。

此範例的執行過程如下：

- (1) 利用 Netica 產生 1050 筆的資料，做為此範例的資料來源。
- (2) 視每個節點與狀態為各列與各行的項目，透過 Microsoft Excel 處理(1)產生的資料之前三分之二 ( 700 筆 )，將資料轉成矩陣型式，可得一個  $6 \times 6$  的區塊矩陣  $M$ ，如表 4.7 所示。其中矩陣之列與行的順序為 Draw\_1、Draw\_2、Draw\_3、Draw\_4、Draw\_5、Book\_Bag。
- (3) 透過 Paul Godfrey[11]求 Jordan 矩陣分解的 MATLAB 程式找出區塊矩陣  $M$  的 Jordan 標準型  $J$ ，如表 4.8 所示。
- (4) 用求得的  $J$ ，做為相似性轉換時所需要的轉換矩陣，求  $J^{-1}MJ$ ，可得相似於  $M$  的下三角矩陣  $T$ 。如表 4.9 所示，其中粗實線框表示兩變數間有關係存在。
- (5) 根據(4)的結果  $T$  及圖論的定義畫出非循環方向圖。
  - a. 區塊矩陣  $T$  中，對角線元素  $T_{44}$  不為零矩陣，子矩陣  $T_{41}$ 、 $T_{51}$ 、 $T_{61}$ 、 $T_{42}$ 、 $T_{52}$ 、 $T_{62}$ 、 $T_{43}$ 、 $T_{53}$ 、 $T_{63}$ 、 $T_{54}$  及  $T_{64}$  亦不為零矩陣。

表 4.7 文具袋貝氏網路之矩陣格式

		Draw_1	Draw_1	Draw_2	Draw_2	Draw_3	Draw_3	Draw_4	Draw_4	Draw_5	Draw_5	Book_Bag	Book_Bag
		blue	red	red	blue	red	blue	blue	red	blue	red	bag_2	bag_1
Draw_1	blue	1	0	0.492447	0.507553	0.468278	0.531722	0.537764	0.462236	0.501511	0.498489	0.546828	0.453172
Draw_1	red	0	1	0.596206	0.403794	0.634146	0.365854	0.390244	0.609756	0.406504	0.593496	0.203252	0.796748
Draw_2	red	0.425587	0.574413	1	0	0.613577	0.386423	0.407311	0.592689	0.386423	0.613577	0.203655	0.796345
Draw_2	blue	0.529968	0.470032	0	1	0.485804	0.514196	0.523659	0.476341	0.529968	0.470032	0.561514	0.438486
Draw_3	red	0.398458	0.601542	0.604113	0.395887	1	0	0.385604	0.614396	0.398458	0.601542	0.197943	0.802057
Draw_3	blue	0.565916	0.434084	0.475884	0.524116	0	1	0.553055	0.446945	0.517685	0.482315	0.575563	0.424437
Draw_4	blue	0.552795	0.447205	0.484472	0.515528	0.465839	0.534161	1	0	0.524845	0.475155	0.571429	0.428571
Draw_4	red	0.404762	0.595238	0.600529	0.399471	0.632275	0.367725	0	1	0.388889	0.611111	0.190476	0.809524
Draw_5	blue	0.525316	0.474684	0.468354	0.531646	0.490506	0.509494	0.53481	0.46519	1	0	0.553797	0.446203
Draw_5	red	0.429688	0.570313	0.611979	0.388021	0.609375	0.390625	0.398438	0.601563	0	1	0.210938	0.789063
Book_Bag	bag_2	0.707031	0.292969	0.304688	0.695313	0.300781	0.699219	0.71875	0.28125	0.683594	0.316406	1	0
Book_Bag	bag_1	0.337838	0.662162	0.686937	0.313063	0.702703	0.297297	0.310811	0.689189	0.317568	0.682432	0	1

表 4.8 文具袋貝氏網路之 Jordan 標準型

6	-5.6E-16	0	0	0	0	0	0	0	0	0	0	0	0
0	2.141409	1.2E-16	0	0	0	0	0	0	0	0	0	0	0
0	0	0.939385	9.55E-17	0	0	0	0	0	0	0	0	0	0
0	0	0	0.867861	-8.3E-18	0	0	0	0	0	0	0	0	0
0	0	0	0	0.837835	6.51E-17	0	0	0	0	0	0	0	0
0	0	0	0	0	0.823924	1.14E-16	0	0	0	0	0	0	0
0	0	0	0	0	0	0.389585	1.32E-15	0	0	0	0	0	0
0	0	0	0	0	0	0	-9.6E-17	-2E-16	0	0	0	0	0
0	0	0	0	0	0	0	0	-2.6E-16	3.28E-16	0	0	0	0
0	0	0	0	0	0	0	0	0	8.85E-17	-1E-16	0	0	0
0	0	0	0	0	0	0	0	0	0	-2.2E-16	1.03E-16	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1.04E-18	0

表 4.9 文具袋貝氏網路之相似矩陣

		Draw_1	Draw_1	Draw_2	Draw_2	Draw_3	Draw_3	Draw_4	Draw_4	Draw_5	Draw_5	Book_Bag	Book_Bag
		blue	red	red	blue	red	blue	blue	red	blue	red	bag_2	bag_1
Draw_1	blue	1	0	0.0771	0.073414	0.06539	0.073016	0.034917	1.11E-16	-3.7E-17	3.48E-17	-2.8E-17	9.45E-18
Draw_1	red	-1.5E-16	1	0.261541	0.163648	0.248112	0.140765	0.070997	2.13E-16	-1.1E-16	8.68E-17	-4.9E-17	1.01E-17
Draw_2	red	2.718295	1.309423	1	0	0.547248	0.338927	0.168921	5.11E-16	-2.3E-16	1.93E-16	-1.1E-16	2.32E-17
Draw_2	blue	3.663964	1.159783	7.13E-17	1	0.468997	0.488164	0.235072	7.43E-16	-2.7E-16	2.48E-16	-2E-16	6.7E-17
Draw_3	red	2.853478	1.537472	0.677334	0.410074	1	1.25E-29	0.179302	5.36E-16	-2.7E-16	2.19E-16	-1.2E-16	2.53E-17
Draw_3	blue	4.121129	1.128199	0.542572	0.552065	-3.9E-13	1	0.261507	8.33E-16	-2.7E-16	2.58E-16	-2.1E-16	7.23E-17
Draw_4	blue	9651.022	6818.497	3105.002	1290.737	2832.151	1163.629	570.2035	1.61E-12	-1E-12	7.73E-13	-3.4E-13	1.57E-15
Draw_4	red	-2.8E+18	-2E+18	-9.2E+17	-3.8E+17	-8.4E+17	-3.4E+17	-1.7E+17	-474.715	306.5849	-228.243	99.29779	-0.41997
Draw_5	blue	1.37E+18	9.68E+17	4.41E+17	1.83E+17	4.02E+17	1.65E+17	8.14E+16	230.0886	-146.779	109.4797	-47.5031	0.100289
Draw_5	red	1.09E+18	7.66E+17	3.49E+17	1.45E+17	3.18E+17	1.31E+17	6.48E+16	183.4974	-116.843	87.35226	-37.8328	0.254193
Book_Bag	bag_2	9.09E+17	6.47E+17	2.94E+17	1.22E+17	2.69E+17	1.1E+17	5.42E+16	153.1093	-99.2227	74.18688	-31.0611	5.55E-17
Book_Bag	bag_1	1.95E+18	1.36E+18	6.21E+17	2.61E+17	5.66E+17	2.36E+17	1.16E+17	330.3913	-210.57	158.1993	-67.6259	1

- b. 由定義 2.3.2 及定義 2.3.6 可知, 區塊矩陣  $T$  的縮約圖形中, 存在弧線  $4 \rightarrow 1$ 、 $5 \rightarrow 1$ 、 $6 \rightarrow 1$ 、 $4 \rightarrow 2$ 、 $5 \rightarrow 2$ 、 $6 \rightarrow 2$ 、 $4 \rightarrow 3$ 、 $5 \rightarrow 3$ 、 $6 \rightarrow 3$ 、 $5 \rightarrow 4$ 、及  $6 \rightarrow 4$ 。即  $\text{Draw}_4 \rightarrow \text{Draw}_1$ 、 $\text{Draw}_5 \rightarrow \text{Draw}_1$ 、 $\text{Book\_Bag} \rightarrow \text{Draw}_1$ 、 $\text{Draw}_4 \rightarrow \text{Draw}_2$ 、 $\text{Draw}_5 \rightarrow \text{Draw}_2$ 、 $\text{Book\_Bag} \rightarrow \text{Draw}_2$ 、 $\text{Draw}_4 \rightarrow \text{Draw}_3$ 、 $\text{Draw}_5 \rightarrow \text{Draw}_3$ 、 $\text{Book\_Bag} \rightarrow \text{Draw}_3$ 、 $\text{Draw}_4 \rightarrow \text{Draw}_4$ 、 $\text{Draw}_5 \rightarrow \text{Draw}_4$  及  $\text{Book\_Bag} \rightarrow \text{Draw}_4$ 。此外,  $T_{44}$  不為零矩陣, 表示 4 為中間點, 即  $\text{Draw}_4$  為圖形中的中間點。
- (6) 將(5)的結果以圖形的方式表示, 得到圖 4.11, 貝氏網路之非循環方向圖。
- (7) 利用 Netica, 結合(6)得到的非循環方向圖與(1)剩下的三分之一資料量, 建構貝氏網路。得到圖 4.12, 貝氏網路架構。

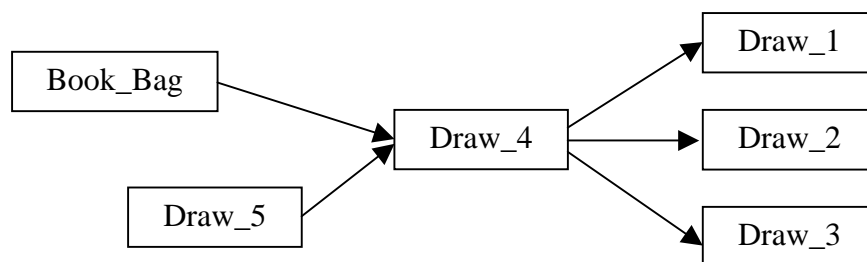


圖 4.11 文具袋貝氏網路之非循環方向圖

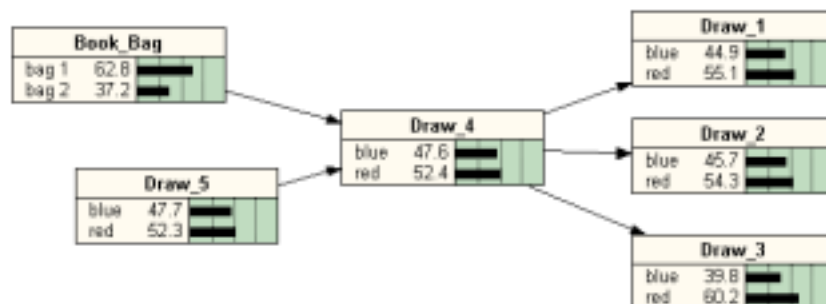


圖 4.12 本研究方法建立之文具袋貝氏網路

(8) 架構的相似性分析

- a. 此範例的貝氏網路中，所有存在之弧線形成的集合  
 $X = \{(\text{Book\_Bag}, \text{Draw\_1}), (\text{Book\_Bag}, \text{Draw\_2}), (\text{Book\_Bag}, \text{Draw\_3}), (\text{Book\_Bag}, \text{Draw\_4}), (\text{Book\_Bag}, \text{Draw\_5})\}$ 。  $X$  中的弧線個數  $x$  為 5 個。
- b. 本研究的最終結果中，所有存在之弧線形成的集合  
 $Y = \{(\text{Book\_Bag}, \text{Draw\_4}), (\text{Book\_Bag}, \text{Draw\_1}), (\text{Book\_Bag}, \text{Draw\_2}), (\text{Book\_Bag}, \text{Draw\_3}), (\text{Draw\_5}, \text{Draw\_4}), (\text{Draw\_5}, \text{Draw\_1}), (\text{Draw\_5}, \text{Draw\_2}), (\text{Draw\_5}, \text{Draw\_3}), (\text{Draw\_4}, \text{Draw\_1}), (\text{Draw\_4}, \text{Draw\_2}), (\text{Draw\_4}, \text{Draw\_3})\}$ 。  $Y$  中的弧線個數  $y$  為 11 個。
- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  
 $X \cap Y = \{(\text{Book\_Bag}, \text{Draw\_1}), (\text{Book\_Bag}, \text{Draw\_2}), (\text{Book\_Bag}, \text{Draw\_3}), (\text{Book\_Bag}, \text{Draw\_4})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 4 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{4}{5 + 11 - 4} = 0.3333$$

由此可知，兩個架構的相似程度為 33.33%。

(9) 資料的相似性分析

- a. 顯著水準 0.05 之假設檢定的兩個假設條件為
- $H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組  
 $H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組
- b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。（本研究方法的文具袋貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄二的表 A-2.1。）
- c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的  $p$  值。計算結果如圖 4.13 所示， $p$  值為 0.425 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。



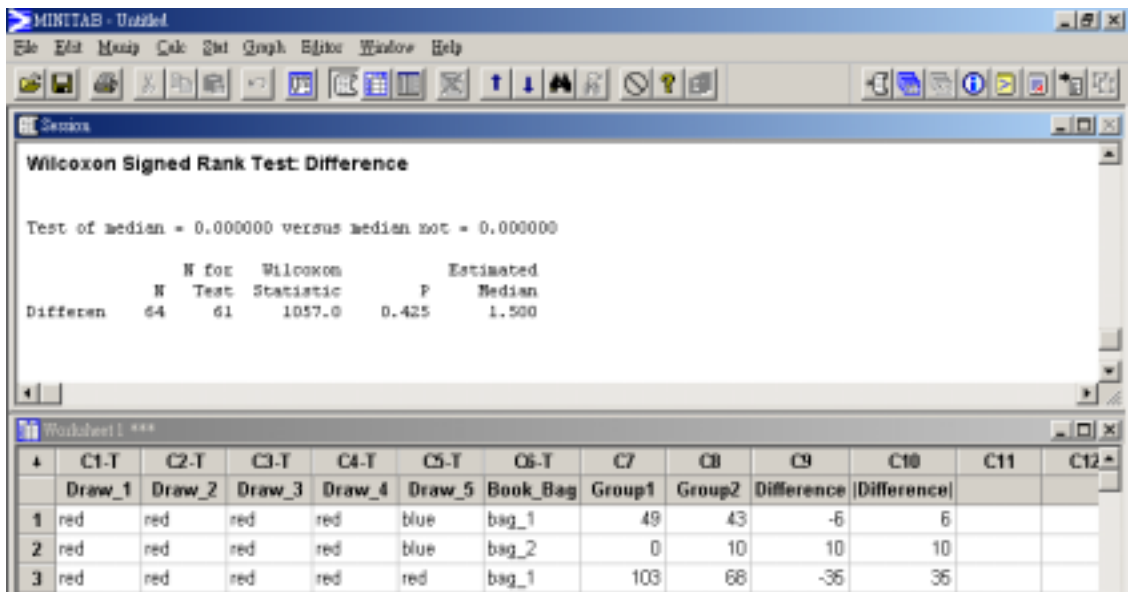


圖 4.13 本研究方法的文具袋貝氏網路：魏克森符號等級和檢定

此範例由架構的相似性分析可知，本研究方法建立之此範例架構與已知架構間的相似度為 33.33%，表示本研究方法只能表達此範例變數間 33.33%的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示本研究方法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

#### 4.2.2 K2 演算法建立之文具袋貝氏網路架構

此一節中，本研究擬以 K2 演算法產生此範例之架構，進行架構與資料的相似性評估。由於 K2 演算法執行前須給定一組節點的排序，因此給定一組與本研究相同的排序，則節點排序為 Draw\_1、Draw\_2、Draw\_3、Draw\_4、Draw\_5、Book\_Bag。

透過貝氏網路應用軟體 Bayesware Discoverer 執行此一步驟，可得以下之結果：

- (1) K2 演算法產生之非循環方向圖，如圖 4.14。
- (2) 利用 Netica，結合(1)得到的非循環方向圖與 4.2.1 之(1)剩下的三分之一資料量，建構貝氏網路，最終得到圖 4.15。

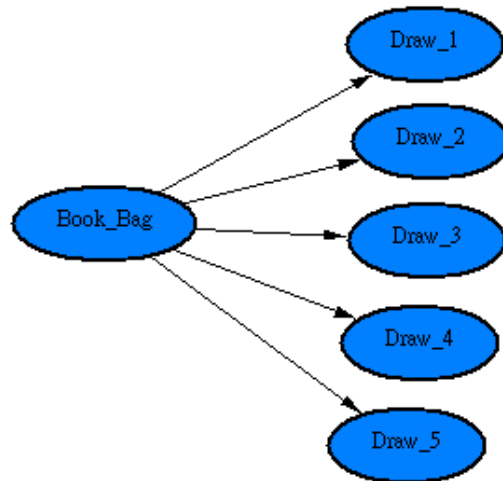


圖 4.14 K2 演算法建立的文具袋貝氏網路之非循環方向圖

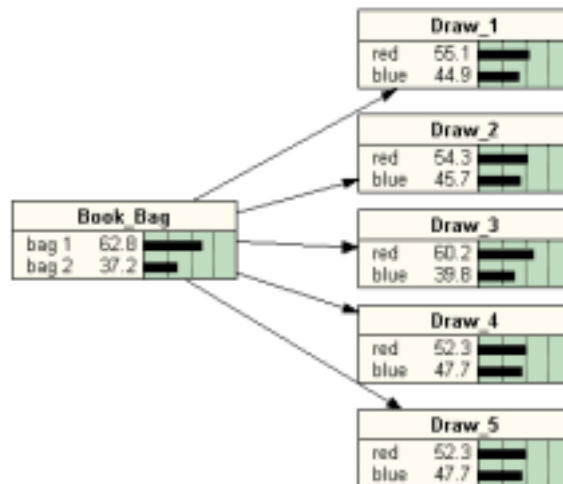


圖 4.15 K2 演算法建立之文具袋貝氏網路架構

(3) 架構的相似性分析

a. 此範例的貝氏網路中，所有存在之弧線形成的集合

$X = \{(Book\_Bag, Draw\_1), (Book\_Bag, Draw\_2), (Book\_Bag, Draw\_3), (Book\_Bag, Draw\_4), (Book\_Bag, Draw\_5)\}$ 。  $X$  中的弧線個數  $x$  為 5 個。

b. K2 演算法產生之貝氏網路架構中，所有存在之弧線形成的

集合  $Y = \{(Book\_Bag, Draw\_1), (Book\_Bag, Draw\_2), (Book\_Bag, Draw\_3), (Book\_Bag, Draw\_4), (Book\_Bag, Draw\_5)\}$ 。  $Y$  中的弧線個數  $y$  為 5 個。

- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  
 $X \cap Y = \{(\text{Book\_Bag}, \text{Draw\_1}), (\text{Book\_Bag}, \text{Draw\_2}), (\text{Book\_Bag}, \text{Draw\_3}), (\text{Book\_Bag}, \text{Draw\_4}), (\text{Book\_Bag}, \text{Draw\_5})\}$ 。 $X \cap Y$  中的弧線個數  $m$  為 5 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{5}{5 + 5 - 5} = 1$$

由此可知，兩個架構的相似程度為 100%。

#### (4) 資料的相似性分析

- a. 顯著水準 0.05 之假設檢定的兩個假設條件為

$H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組

$H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組

- b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。( K2 演算法的文具袋貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄二的表 A-2.3。 )
- c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的  $p$  值。計算結果如圖 4.16 所示， $p$  值為 0.988 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

由架構的相似性分析可知，K2 演算法建立之此範例架構與已知架構間的相似度為 100%，表示 K2 演算法完全表達此範例變數間的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示 K2 演算法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

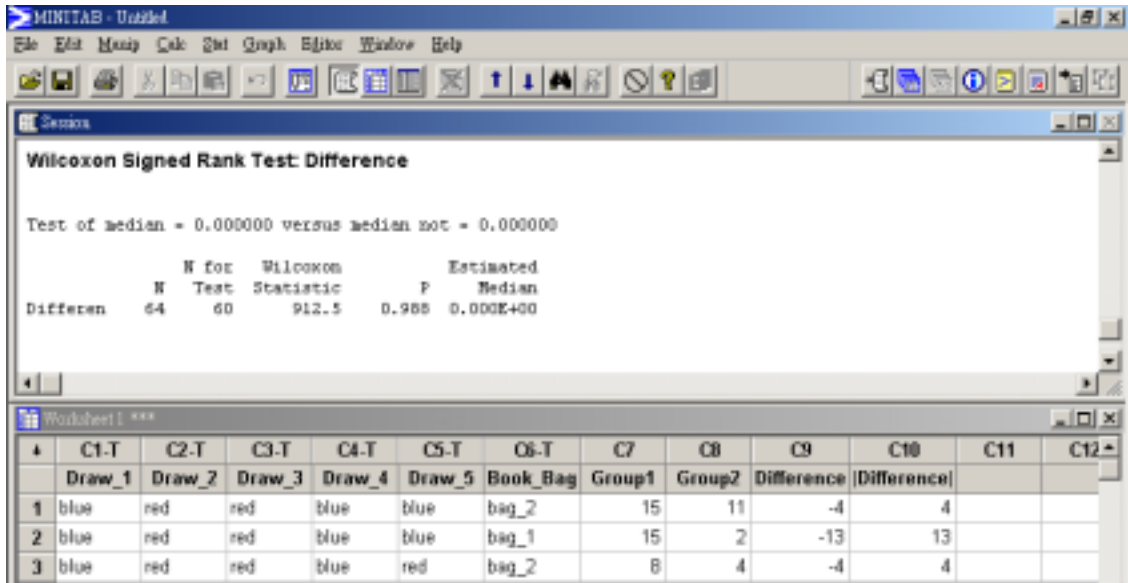


圖 4.16 K2 演算法的文具袋貝氏網路：魏克森符號等級和檢定

### 4.3 範例三：Steffen L. Lauritzen 和 David J. Spiegelhalter 胸腔門診(Chest Clinic)貝氏網路

#### 4.3.1 本研究方法建立之胸腔門診貝氏網路架構

醫生面臨胸腔門診的情況，主要工作是從癥狀 是否呼吸困難，觀察資料 胸腔 X 射線，兩個危險因子 是否抽煙和曾拜訪亞洲和三種疾病 癌症、肺結核和支氣管炎等這些情況做適當地推論。

呼吸困難可能是由肺結核、肺癌、支氣管炎或這些疾病的混合型所引起的癥狀，也有可能是與這些疾病無關的疾病。近期內曾拜訪過亞洲地區，會使感染肺結核的機會增加，抽煙是罹患肺癌和支氣管炎的危險因子。單純胸腔 X 射線的結果無法區別肺癌和肺結核，同樣地，無法判斷是否有呼吸困難的癥狀。

此範例的貝氏網路如圖 4.17。

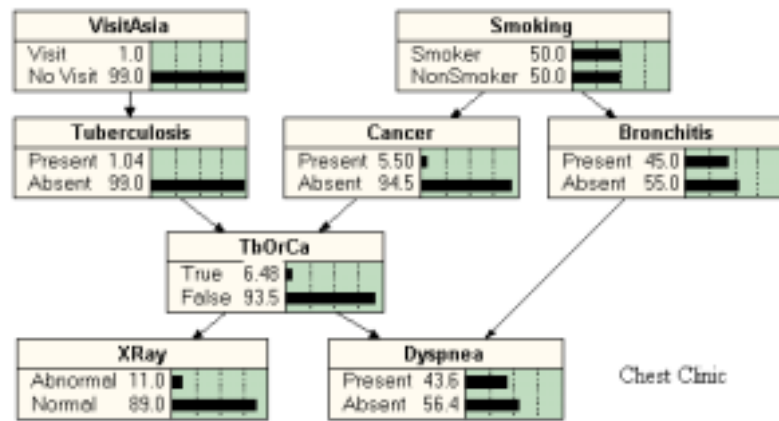


圖 4.17 Lauritzen 和 Spiegelhalter 的胸腔門診貝氏網路

Steffen L. Lauritzen 和 David J. Spiegelhalter 的胸腔門診(Chest Clinic)貝氏網路中，有八個節點 VisitAsia、Smoking、Tuberculosis、Cancer、Bronchitis、TbOrCa、XRay 和 Dyspnea，每個節點各有兩個狀態，VisitAsia 的狀態為 Visit 和 No\_Visit，Smoking 的狀態為 Smoker 和 NonSmoker，TbOrCa 的狀態為 True 和 False，Xray 的狀態為 Abnormal 和 Normal，剩下的 Tuberculosis、Cancer、Bronchitis 及 Dyspnea 的狀態均為 Present 和 Absent。

此範例的執行過程如下：

- (1) 利用 Netica 產生 1050 筆的資料，做為此範例的資料來源。
- (2) 視每個節點與狀態為各列與各行的項目，可得一個  $8 \times 8$  的區塊矩陣  $M$ ，如表 4.10 所示。其中矩陣之列與行的順序為 Dyspnea、XRay、TbOrCa、Bronchitis、Cancer、Tuberculosis、Smoking、VisitAsia。
- (3) 透過 Paul Godfrey[11]求 Jordan 矩陣分解的 MATLAB 程式找出區塊矩陣  $M$  的 Jordan 標準型  $J$ ，如表 4.11 所示。
- (4) 用求得的  $J$ ，做為相似性轉換時所需要的轉換矩陣，求  $J^{-1}MJ$ ，可得相似於  $M$  的下三角矩陣  $T$ 。如表 4.12 所示，其中對角線上的粗虛線框表示該點為中間點，其餘粗實線框表示兩變數間有關係存在。

表 4.10 胸腔門診貝氏網路之矩陣格式

		Dyspnea	Dyspnea	XRay	XRay	TbOrCa	TbOrCa	Bronchitis	Bronchitis	Cancer	Cancer	Tuberculosis	Tuberculosis	Smoking	Smoking
		Absent	Present	Normal	Abnormal	FALSE	TRUE	Absent	Present	Absent	Present	Absent	Present	Smoker	NonSmoker
Dyspnea	Absent	1	0	0.9490291	0.0509709	0.9781553	0.0218447	0.8543689	0.1456311	0.9805825	0.0194175	0.9975728	0.0024272	0.3980583	0.6019417
Dyspnea	Present	0	1	0.8472222	0.1527778	0.8715278	0.1284722	0.1527778	0.8472222	0.8819444	0.1180556	0.9895833	0.0104167	0.6145833	0.3854167
XRay	Normal	0.615748	0.384252	1	0	0.9984252	0.0015748	0.5685039	0.4314961	0.9984252	0.0015748	1	0	0.4551181	0.5448819
XRay	Abnormal	0.3230769	0.6769231	0	1	0.3076923	0.6923077	0.5384615	0.4615385	0.3692308	0.6307692	0.9384615	0.0615385	0.8	0.2
TbOrCa	FALSE	0.616208	0.383792	0.969419	0.030581	1	0	0.5749235	0.4250765	1	0	1	0	0.4571865	0.5428135
TbOrCa	TRUE	0.1956522	0.8043478	0.0217391	0.9782609	0	1	0.4347826	0.5652174	0.0869565	0.9130435	0.9130435	0.0869565	0.9130435	0.0869565
Bronchitis	Absent	0.8888889	0.1111111	0.9116162	0.0883838	0.9494949	0.0505051	1	0	0.9545455	0.0454545	0.9949495	0.0050505	0.3611111	0.6388889
Bronchitis	Present	0.1973684	0.8026316	0.9013158	0.0986842	0.9144737	0.0855263	0	1	0.9210526	0.0789474	0.9934211	0.0065789	0.6513158	0.3486842
Cancer	Absent	0.6139818	0.3860182	0.9635258	0.0364742	0.993921	0.006079	0.5744681	0.4255319	1	0	0.993921	0.006079	0.4574468	0.5425532
Cancer	Present	0.1904762	0.8095238	0.0238095	0.9761905	0	1	0.4285714	0.5714286	0	1	1	0	0.952381	0.047619
Tuberculosis	Absent	0.5905172	0.4094828	0.9123563	0.0876437	0.9396552	0.0603448	0.566092	0.433908	0.9396552	0.0603448	1	0	0.487069	0.512931
Tuberculosis	Present	0.25	0.75	0	1	0	1	0.5	0.5	1	0	0	1	0.5	0.5
Smoking	Smoker	0.4809384	0.5190616	0.8475073	0.1524927	0.8768328	0.1231672	0.4193548	0.5806452	0.8826979	0.1173021	0.9941349	0.0058651	1	0
Smoking	NonSmoker	0.6908078	0.3091922	0.9637883	0.0362117	0.9888579	0.0111421	0.7047354	0.2952646	0.994429	0.005571	0.994429	0.005571	0	1
VisitAsia	No_Visit	0.5892598	0.4107402	0.9071118	0.0928882	0.934688	0.065312	0.5660377	0.4339623	0.9404935	0.0595065	0.9941945	0.0058055	0.4876633	0.5123367
VisitAsia	Visit	0.5454545	0.4545455	0.9090909	0.0909091	0.9090909	0.0909091	0.5454545	0.4545455	0.9090909	0.0909091	1	0	0.4545455	0.5454545

表 4.11 胸腔門診貝氏網路之 Jordan 標準型

8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2.943007	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1.686388	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.813719	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1.028479	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0.998329	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0.285663	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.244413	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	-2.7E-13	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	2.73E-13	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	8.89E-16	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	-2.2E-16	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	-1.5E-16	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	-1.5E-16	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.95E-17	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.22E-16

表 4.12 胸腔門診貝氏網路之相似矩陣

		Dyspnea	Dyspnea	XRay	XRay	TbOrCa	TbOrCa	Bronchitis	Bronchitis	Cancer	Cancer	Tuberculosis	Tuberculosis	Smoking	Smoking	VisitAsia	VisitAsia
		Absent	Present	Normal	Abnormal	FALSE	TRUE	Absent	Present	Absent	Present	Absent	Present	Smoker	NonSmoker	No_Visit	Visit
Dyspnea	Absent	1	0	0.2001	0.0052	0.1258	0.0027	0.0305	0.0044	-3E-14	7E-16	1.1E-16	-7E-20	-7E-18	-1E-17	1E-17	4E-19
Dyspnea	Present	0	1	0.4855	0.0422	0.3046	0.0436	0.0148	0.0704	-8E-14	1E-14	3E-16	-8E-19	-3E-17	-2E-17	3E-17	1E-18
XRay	Normal	2.921	0.6706	1	0	0.6089	0.0009	0.0963	0.0625	-2E-13	3E-16	5.3E-16	0	-4E-17	-5E-17	5E-17	2E-18
XRay	Abnormal	3.1763	2.4483	0	1	0.3889	0.8494	0.189	0.1386	-1E-13	2E-13	1E-15	-2E-17	-1E-16	-4E-17	1E-16	4E-18
TbOrCa	FALSE	4.7932	1.0982	1.5895	0.0242	1	0	0.1597	0.101	-3E-13	0	8.6E-16	0	-6E-17	-8E-17	8E-17	3E-18
TbOrCa	TRUE	1.5678	2.3712	0.0367	0.7974	0	1	0.1244	0.1384	-2E-14	2E-13	8.1E-16	-2E-17	-1E-16	-1E-17	8E-17	5E-18
Bronchitis	Absent	24.893	1.1447	5.3816	0.2518	3.4185	0.1765	1	0	-9E-13	4E-14	3.1E-15	-4E-18	-2E-16	-3E-16	3E-16	1E-17
Bronchitis	Present	6.4601	9.6646	6.2188	0.3285	3.8481	0.3493	0	1	-1E-12	9E-14	3.6E-15	-6E-18	-4E-16	-2E-16	3E-16	1E-17
Cancer	Absent	-2E+13	-4E+12	-6E+12	-1E+11	-4E+12	-2E+10	-6E+11	-4E+11	1	0	-0.0032	4.9E-06	0.0002	0.0003	-3E-04	-1E-05
Cancer	Present	6E+12	9E+12	1E+11	3E+12	0	4E+12	4E+11	5E+11	0	1	0.00326	0	-5E-04	-3E-05	0.0003	2E-05
Tuberculosis	Absent	5E+15	1E+15	2E+15	8E+13	1E+15	7E+13	2E+14	1E+14	-287.9	18.515	1	0	-0.08	-0.0838	0.0881	0.004
Tuberculosis	Present	-9E+15	-1E+16	0	-4E+15	0	-5E+15	-6E+14	-6E+14	1232.3	0	0	1	0.3285	0.3285	-0.36	0
Smoking	Smoker	-1E+16	-5E+15	-5E+15	-4E+14	-3E+15	-4E+14	-4E+14	-5E+14	796.36	-106	-2.9273	0.00429	1	0	-0.26	-0.011
Smoking	NonSmoker	-2E+16	-3E+15	-5E+15	-1E+14	-3E+15	-4E+13	-7E+14	-2E+14	897.16	-5.033	-2.9281	0.00408	0	1	-0.259	-0.012
VisitAsia	No_Visit	6E+16	2E+16	2E+16	1E+15	1E+16	8E+14	2E+15	1E+15	-3221	204.06	11.1115	-0.0161	-0.89	-0.9354	1	0
VisitAsia	Visit	2E+16	6E+15	7E+15	3E+14	4E+15	4E+14	7E+14	5E+14	-1113	111.46	3.99601	0	-0.297	-0.3561	0	1

- (5) 根據(4)的結果  $T$  及圖論的定義畫出非循環方向圖。
- 區塊矩陣  $T$  中，子矩陣  $T_{51}$ 、 $T_{61}$ 、 $T_{71}$ 、 $T_{81}$ 、 $T_{52}$ 、 $T_{62}$ 、 $T_{72}$ 、 $T_{82}$ 、 $T_{53}$ 、 $T_{63}$ 、 $T_{73}$ 、 $T_{83}$ 、 $T_{54}$ 、 $T_{64}$ 、 $T_{74}$  及  $T_{84}$  均不為零矩陣。
  - 由定義 2.3.2 及定義 2.3.6 可知，區塊矩陣  $T$  的縮約圖形中，存在弧線  $5 \rightarrow 1$ 、 $6 \rightarrow 1$ 、 $7 \rightarrow 1$ 、 $8 \rightarrow 1$ 、 $5 \rightarrow 2$ 、 $6 \rightarrow 2$ 、 $7 \rightarrow 2$ 、 $8 \rightarrow 2$ 、 $5 \rightarrow 3$ 、 $6 \rightarrow 3$ 、 $7 \rightarrow 3$ 、 $8 \rightarrow 3$ 、 $5 \rightarrow 4$ 、 $6 \rightarrow 4$ 、 $7 \rightarrow 4$  及  $8 \rightarrow 4$ 。即  $\text{Cancer} \rightarrow \text{Dyspnea}$ 、 $\text{Tuberculosis} \rightarrow \text{Dyspnea}$ 、 $\text{Smoking} \rightarrow \text{Dyspnea}$ 、 $\text{VisitAsia} \rightarrow \text{Dyspnea}$ 、 $\text{Cancer} \rightarrow \text{XRay}$ 、 $\text{Tuberculosis} \rightarrow \text{XRay}$ 、 $\text{Smoking} \rightarrow \text{XRay}$ 、 $\text{VisitAsia} \rightarrow \text{XRay}$ 、 $\text{Cancer} \rightarrow \text{TbOrCa}$ 、 $\text{Tuberculosis} \rightarrow \text{TbOrCa}$ 、 $\text{Smoking} \rightarrow \text{TbOrCa}$ 、 $\text{VisitAsia} \rightarrow \text{TbOrCa}$ 、 $\text{Cancer} \rightarrow \text{Bronchitis}$ 、 $\text{Tuberculosis} \rightarrow \text{Bronchitis}$ 、 $\text{Smoking} \rightarrow \text{Bronchitis}$  及  $\text{VisitAsia} \rightarrow \text{Bronchitis}$ 。
- (6) 將(5)的結果以圖形的方式表示，得到圖 4.18。
- (7) 架構的相似性分析
- 此範例的貝氏網路中，所有存在之弧線形成的集合  $X = \{(\text{VisitAsia}, \text{Tuberculosis}), (\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{TbOrCa}, \text{XRay}), (\text{TbOrCa}, \text{Dyspnea}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Bronchitis}, \text{Dyspnea})\}$ 。  $X$  中的弧線個數  $x$  為 18 個。

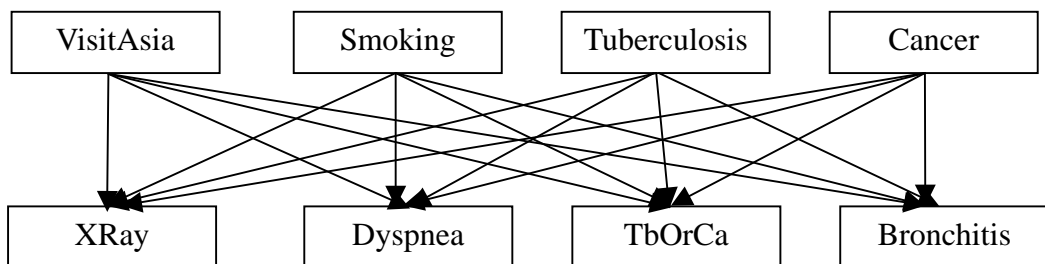


圖 4.18 胸腔門診貝氏網路之子非循環方向圖



- b. 本研究的最終結果中，所有存在之弧線形成的集合  
 $Y = \{(\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{VisitAsia}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Smoking}, \text{Bronchitis}), (\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{Tuberculosis}, \text{Bronchitis}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Cancer}, \text{Bronchitis})\}$ 。  $Y$  中的弧線個數  $y$  為 16 個。
- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  
 $X \cap Y = \{(\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Smoking}, \text{Bronchitis}), (\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 13 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{13}{18 + 16 - 13} = 0.6190$$

由此可知，兩個架構的相似程度為 61.90%。

為了簡化圖形架構，去除多餘的變數關係。故本研究於此範例中，將(6)的結果分成父代節點集合  $P_{ctsv} = \{\text{VisitAsia}, \text{Smoking}, \text{Tuberculosis}, \text{Cancer}\}$  與子代節點集合  $C_{dxtb} = \{\text{Xray}, \text{Dyspnea}, \text{TbOrCa}, \text{Bronchitis}\}$ ，分別建立兩集合的子圖形。結合此兩個子圖形，形成此範例的貝氏網路架構。

- (8) 重複(2)到(6)相同的步驟，找出父代節點集合  $P_{ctsv} = \{\text{VisitAsia}, \text{Smoking}, \text{Tuberculosis}, \text{Cancer}\}$  的子圖形。
- (9) 由(8)和(2)可知子圖形有節點 VisitAsia，狀態為 No\_Visit 和 Visit；節點 Smoking，狀態為 NonSmoker 和 Smoker；節點 Tuberculosis 和節點 Cancer，狀態均為 Absent 和 Present。可得一個  $4 \times 4$  的區塊矩陣  $M_{ctsv}$ ，如表 4.13 所示。
- (10) 找出區塊矩陣  $M_{ctsv}$  的 Jordan 標準型  $J_{ctsv}$ ，如表 4.14 所示。

表 4.13 胸腔門診貝氏網路之子矩陣格式  $M_{ctsv}$

		Cancer	Cancer	Tuberculosis	Tuberculosis	Smoking	Smoking	VisitAsia	VisitAsia
		Absent	Present	Absent	Present	Smoker	NonSmoker	No_Visit	Visit
Cancer	Absent	1	0	0.993921	0.006079	0.4574468	0.5425532	0.9848024	0.0151976
Cancer	Present	0	1	1	0	0.952381	0.047619	0.9761905	0.0238095
Tuberculosis	Absent	0.9396552	0.0603448	1	0	0.487069	0.512931	0.9841954	0.0158046
Tuberculosis	Present	1	0	0	1	0.5	0.5	1	0
Smoking	Smoker	0.8826979	0.1173021	0.9941349	0.0058651	1	0	0.9853372	0.0146628
Smoking	NonSmoker	0.994429	0.005571	0.994429	0.005571	0	1	0.9832869	0.0167131
VisitAsia	No_Visit	0.9404935	0.0595065	0.9941945	0.0058055	0.4876633	0.5123367	1	0
VisitAsia	Visit	0.9090909	0.0909091	1	0	0.4545455	0.5454545	0	1

表 4.14 子矩陣  $M_{ctsv}$  的 Jordan 標準型  $J_{ctsv}$

4	-3E-16	0	0	0	0	0	0	0
0	1.235957	-7.5E-16	0	0	0	0	0	0
0	0	1.010997	9.44E-16	0	0	0	0	0
0	0	0	0.990344	-5.6E-17	0	0	0	0
0	0	0	0	0.762702	-5.6E-16	0	0	0
0	0	0	0	0	-2.2E-17	-7.9E-17	0	0
0	0	0	0	0	0	2.52E-17	-6.7E-17	0
0	0	0	0	0	0	0	0	-6.7E-17

- (11) 用求得的  $J_{ctsv}$ ，做為相似性轉換時所需要的轉換矩陣，求  $J_{ctsv}^{-1} M_{ctsv} J_{ctsv}$ ，可得相似於  $M_{ctsv}$  的下三角矩陣  $T_{ctsv}$ 。如表 4.15 所示，其中對角線上的粗虛線框表示該點為中間點，其餘粗實線框表示兩變數間有關係存在。
- (12) 根據(11)的結果  $T_{ctsv}$  及圖論的定義畫出非循環方向圖。
- 區塊矩陣  $T_{ctsv}$  中，對角線元素  $T_{ctsv\ 33}$  不為零矩陣，子矩陣  $T_{ctsv\ 31}$ 、 $T_{ctsv\ 41}$ 、 $T_{ctsv\ 32}$ 、 $T_{ctsv\ 42}$  及  $T_{ctsv\ 43}$  亦不為零矩陣。
  - 由定義 2.3.2 及定義 2.3.6 可知，區塊矩陣  $T_{ctsv}$  的縮約圖形中，存在弧線  $3 \rightarrow 1$ 、 $4 \rightarrow 1$ 、 $3 \rightarrow 2$ 、 $4 \rightarrow 2$  及  $4 \rightarrow 3$ 。即 Smoking  $\rightarrow$  Cancer、VisitAsia  $\rightarrow$  Cancer、Smoking  $\rightarrow$  Tuberculosis、VisitAsia  $\rightarrow$  Tuberculosis、及 VisitAsia  $\rightarrow$  Smoking。此外， $T_{ctsv\ 33}$  不為零矩陣，表示 3 為中間點，即 Smoking 為子圖形中的中間點。
- (13) 將(12)的結果以圖形的方式表示，得到圖 4.19。
- (14) 結合圖 4.18 和圖 4.19，可得圖 4.20。
- (15) 重複(2)到(6)相同的步驟，找出  $C_{dxtb} = \{Xray, Dyspnea, TbOrCa, Bronchitis\}$  的子圖形。

表 4.15 胸腔門診貝氏網路之子矩陣  $M_{ctsv}$  的相似矩陣

		Cancer	Cancer	Tuberculosis	Tuberculosis	Smoking	Smoking	VisitAsia	VisitAsia
		Absent	Present	Absent	Present	Smoker	NonSmoker	No_Visit	Visit
Cancer	Absent	1	0	0.251213	0.0015051	0.087224	-6.7E-17	-4.6E-18	-1.7E-17
Cancer	Present	2.25E-15	1	0.817987	7.635E-16	0.587708	-4.3E-16	1.68E-17	-5.4E-17
Tuberculosis	Absent	3.717738	0.073772	1	-1.972E-31	0.367448	-2.8E-16	-1.6E-17	-6.7E-17
Tuberculosis	Present	4.038999	-8.4E-17	-1.7E-15	1	0.385069	-2.9E-16	-1.5E-17	-6.8E-17
Smoking	Smoker	-138.242	3.921309	-30.9558	-0.7611803	-1.66975	1.87E-15	-1E-15	2.1E-15
Smoking	NonSmoker	-2E+17	5.13E+15	-4.4E+16	-1.056E+15	-3.7E+15	3.568989	-1.46074	3.004487
VisitAsia	No_Visit	5.32E+15	-1.5E+15	-1.4E+14	2.285E+14	1.03E+15	-0.72471	1.100763	0
VisitAsia	Visit	-5.4E+16	-1.7E+15	-1.5E+16	-13.981885	-5.1E+15	3.919621	0.642142	1

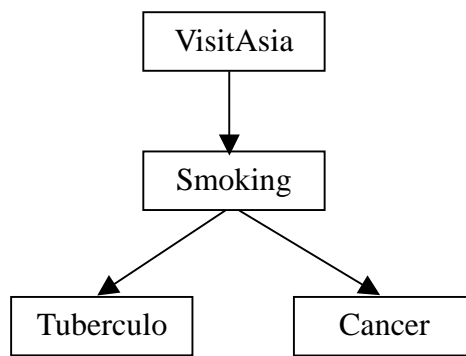


圖 4.19 胸腔門診貝氏網路之子非循環方向圖

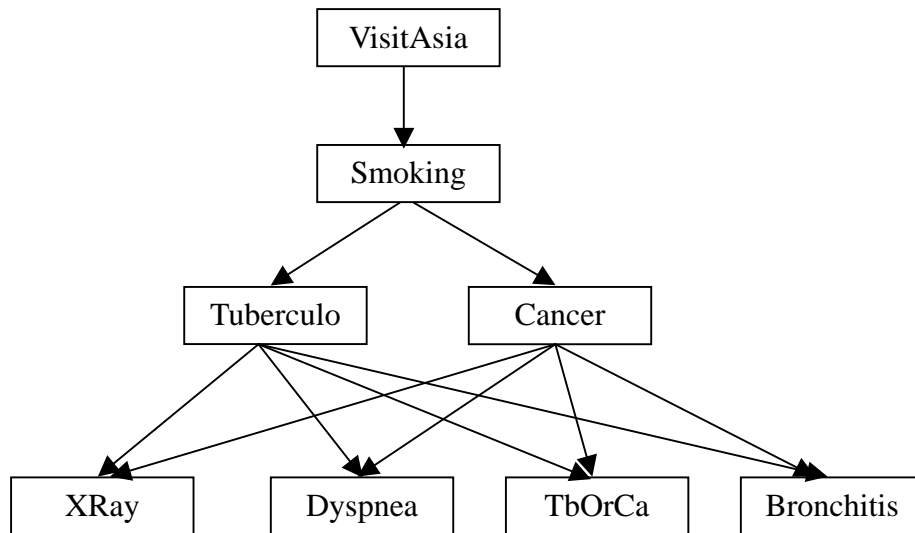


圖 4.20 胸腔門診貝氏網路之子非循環方向圖

- (16) 由(15)和(2)可知子圖形有節點 XRay，狀態為 Normal 和 Abnormal；節點 TbOrCa，狀態為 FALSE 和 TRUE；節點 Dyspnea 和節點 Bronchitis，狀態均為 Absent 和 Present。可得一個  $4 \times 4$  的區塊矩陣  $M_{dxtb}$ ，如表 4.16 所示。
- (17) 找出區塊矩陣  $M_{dxtb}$  的 Jordan 標準型，可得一個矩陣  $J_{dxtb}$ ，如表 4.17 所示。
- (18) 用求得的  $J_{dxtb}$ ，做為相似性轉換時所需要的轉換矩陣，求  $J_{dxtb}^{-1} M_{dxtb} J_{dxtb}$ ，可得相似於  $M_{dxtb}$  的下三角矩陣  $T_{dxtb}$ 。如表 4.18 所示，其中對角線上的粗虛線框表示該點為中間點，其餘粗實線框表示兩變數間有關係存在。

表 4.16 胸腔門診貝氏網路之子矩陣格式  $M_{dxtb}$

		Dyspnea	Dyspnea	XRay	XRay	TbOrCa	TbOrCa	Bronchitis	Bronchitis
		Absent	Present	Normal	Abnormal	FALSE	TRUE	Absent	Present
Dyspnea	Absent	1	0	0.9490291	0.0509709	0.9781553	0.0218447	0.8543689	0.1456311
Dyspnea	Present	0	1	0.8472222	0.1527778	0.8715278	0.1284722	0.1527778	0.8472222
XRay	Normal	0.615748	0.384252	1	0	0.9984252	0.0015748	0.5685039	0.4314961
XRay	Abnormal	0.3230769	0.6769231	0	1	0.3076923	0.6923077	0.5384615	0.4615385
TbOrCa	FALSE	0.616208	0.383792	0.969419	0.030581	1	0	0.5749235	0.4250765
TbOrCa	TRUE	0.1956522	0.8043478	0.0217391	0.9782609	0	1	0.4347826	0.5652174
Bronchitis	Absent	0.8888889	0.1111111	0.9116162	0.0883838	0.9494949	0.0505051	1	0
Bronchitis	Present	0.1973684	0.8026316	0.9013158	0.0986842	0.9144737	0.0855263	0	1

表 4.17 胸腔門診貝氏網路之子矩陣  $M_{dxtb}$  的 Jordan 標準型

4	-5.8E-16	0	0	0	0	0	0	0
0	2.003901	-1.1E-15	0	0	0	0	0	0
0	0	1.518007	3.33E-16	0	0	0	0	0
0	0	0	0.288944	-1.9E-17	0	0	0	0
0	0	0	0	0.189149	8.67E-18	0	0	0
0	0	0	0	0	-1.4E-16	8.03E-17	0	0
0	0	0	0	0	0	7.29E-17	2.51E-17	0
0	0	0	0	0	0	0	3.74E-16	0

表 4.18 胸腔門診貝氏網路之子矩陣  $M_{dxtb}$  的相似矩陣

		Dyspnea	Dyspnea	XRay	XRay	TbOrCa	TbOrCa	Bronchitis	Bronchitis
		Absent	Present	Normal	Abnormal	FALSE	TRUE	Absent	Present
Dyspnea	Absent	1	4.93E-32	0.360158	0.003682	0.0462543	1.37E-18	1.6E-17	1.9E-17
Dyspnea	Present	8.54E-16	1	0.641793	0.022029	0.0822638	-5E-18	1.07E-17	1.6E-16
XRay	Normal	1.622517	0.507246	1	-2.5E-32	0.1244073	5.56E-18	2.74E-17	1.16E-16
XRay	Abnormal	4.472527	4.694644	-2E-15	1	0.2014223	-3.2E-16	3.28E-16	6.43E-16
TbOrCa	FALSE	12.0004	4.563879	7.314089	0.132734	0.9379806	-5.8E-19	2.31E-16	9.9E-16
TbOrCa	TRUE	2.25E+16	-1.1E+16	1.02E+16	-1.9E+15	1.352E+15	1.012709	-0.20275	-1.62278
Bronchitis	Absent	4.8E+16	1.57E+15	1.77E+16	3.24E+14	2.304E+15	0.021653	1.04931	0
Bronchitis	Present	2.11E+15	4.31E+15	3.66E+15	7.63E+13	4.631E+14	-0.01009	0.018387	1

- (19) 根據(18)的結果  $T_{dxtb}$  及圖論的定義畫出非循環方向圖。
- 區塊矩陣  $T_{dxtb}$  中，對角線元素  $T_{dxtb 33}$  不為零矩陣，子矩陣  $T_{dxtb 31}$ 、 $T_{dxtb 41}$ 、 $T_{dxtb 32}$ 、 $T_{dxtb 42}$  及  $T_{dxtb 43}$  均不為零矩陣。
  - 由定義 2.3.2 及定義 2.3.6 可知，區塊矩陣  $T_{dxtb}$  的縮約圖形中，存在弧線  $3 \rightarrow 1$ 、 $4 \rightarrow 1$ 、 $3 \rightarrow 2$ 、 $4 \rightarrow 2$  及  $4 \rightarrow 3$ 。即  $TbOrCa \rightarrow Dyspnea$ 、 $Bronchitis \rightarrow Dyspnea$ 、 $TbOrCa \rightarrow XRay$ 、 $Bronchitis \rightarrow XRay$ 、及  $Bronchitis \rightarrow TbOrCa$ 。此外， $T_{dxtb 33}$  不為零矩陣，表示 3 為中間點，即  $TbOrCa$  為子圖形中的中間點。
- (20) 將(19)的結果以圖形的方式表示，可得到圖 4.21。
- (21) 結合圖 4.20 和圖 4.21，可得圖 4.22。

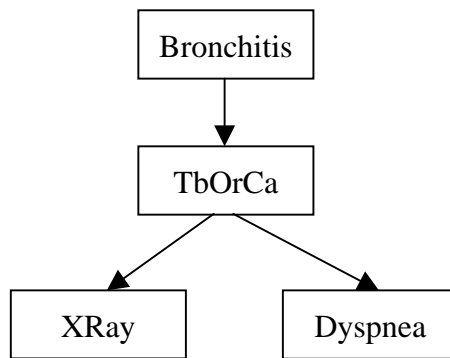


圖 4.21 胸腔門診貝氏網路之子非循環方向圖

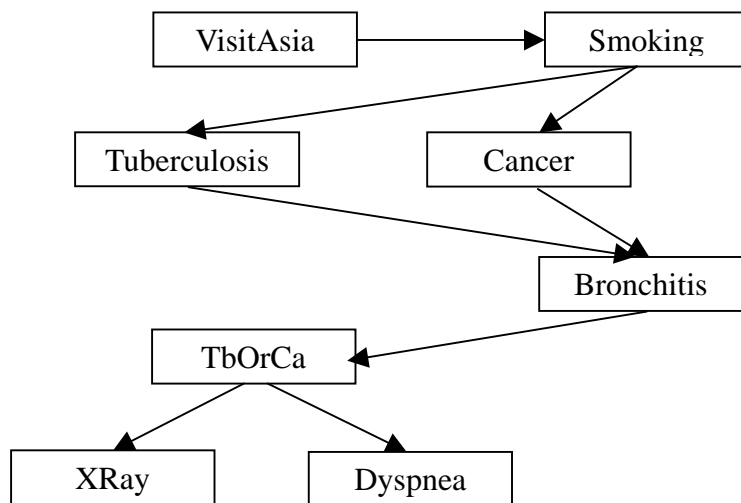


圖 4.22 胸腔門診貝氏網路之非循環方向圖

(22) 利用 Netica , 結合(21)得到的非循環方向圖與(1)剩下的三分之一資料量 ( 350 筆 ) , 建構貝氏網路。最終得到圖 4.23。

(23) 架構的相似性分析

a. 此範例的貝氏網路中 , 所有存在之弧線形成的集合

$X = \{(\text{VisitAsia}, \text{Tuberculosis}), (\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{TbOrCa}, \text{XRay}), (\text{TbOrCa}, \text{Dyspnea}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Bronchitis}, \text{Dyspnea})\}$ 。  $X$  中的弧線個數  $x$  為 18 個。

b. 本研究的最終結果中 , 所有存在之弧線形成的集合

$Y = \{(\text{VisitAsia}, \text{Smoking}), (\text{VisitAsia}, \text{Tuberculosis}), (\text{VisitAsia}, \text{Bronchitis}), (\text{VisitAsia}, \text{Cancer}), (\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{Smoking}, \text{Tuberculosis}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Tuberculosis}, \text{Bronchitis}), (\text{Tuberculosis}, \text{TbOrCa}),$

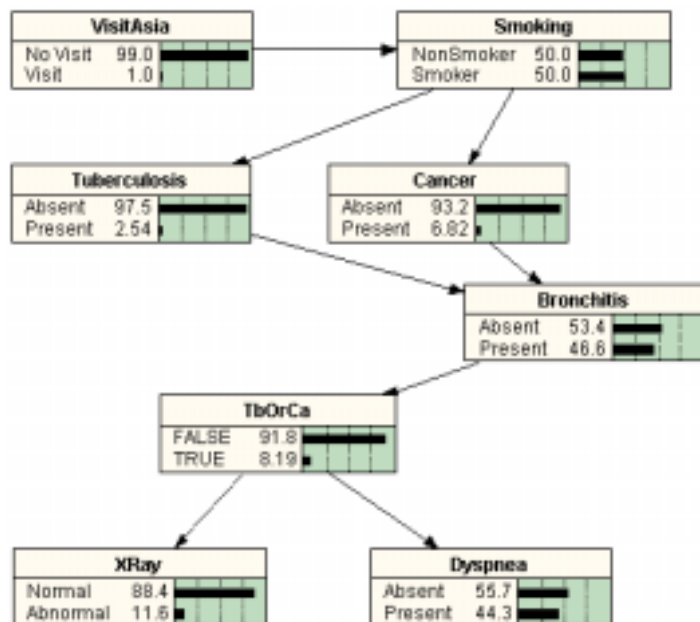


圖 4.23 胸腔門診貝氏網路之非循環方向圖

(Tuberculosis, XRay), (Tuberculosis, Dyspnea), (Cancer, Bronchitis), (Cancer, TbOrCa), (Cancer, XRay), (Cancer, Dyspnea), (Bronchitis, TbOrCa), (Bronchitis, XRay), (Bronchitis, Dyspnea), (TbOrCa, XRay), (TbOrCa, Dyspnea)}。Y 中的弧線個數  $y$  為 26 個。

- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  $X \cap Y = \{(VisitAsia, Tuberculosis), (VisitAsia, TbOrCa), (VisitAsia, XRay), (VisitAsia, Dyspnea), (Tuberculosis, TbOrCa), (Tuberculosis, XRay), (Tuberculosis, Dyspnea), (TbOrCa, XRay), (TbOrCa, Dyspnea), (Smoking, Cancer), (Smoking, Bronchitis), (Smoking, TbOrCa), (Smoking, XRay), (Smoking, Dyspnea), (Cancer, TbOrCa), (Cancer, XRay), (Cancer, Dyspnea), (Bronchitis, Dyspnea)\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 18 個。

- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{18}{18 + 26 - 18} = 0.6923$$

由此可知，兩個架構的相似程度為 69.23%。

#### (24) 資料的相似性分析

- a. 顯著水準 0.05 之假設檢定的兩個假設條件為

$H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組

$H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組

- b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。(本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄三的表 A-3.1。)
- c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的  $p$  值。計算結果如圖 4.24 所示， $p$  值為 0.027 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

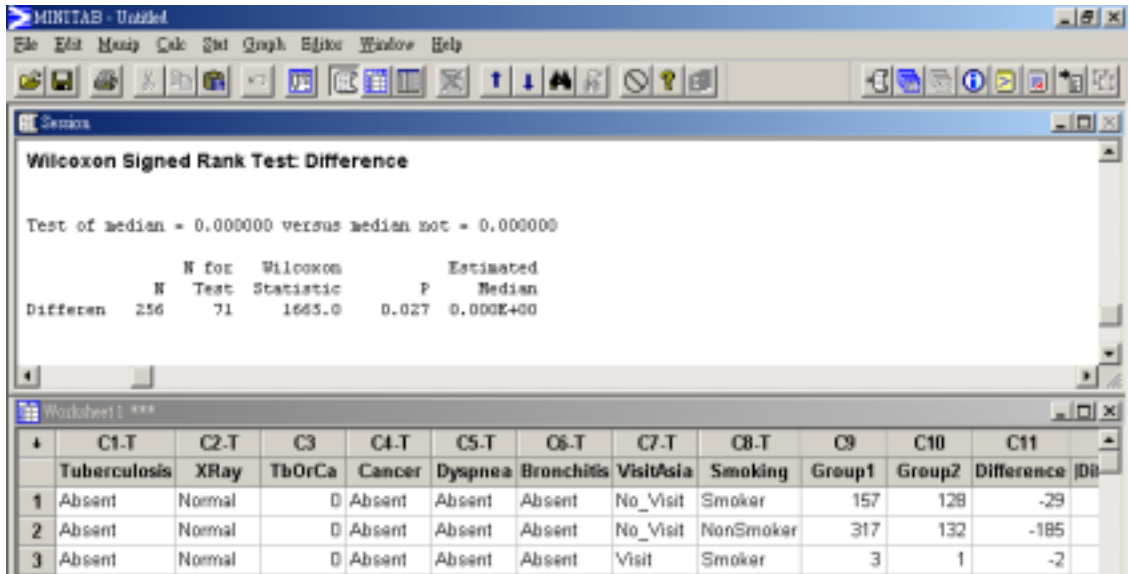


圖 4.24 本研究方法的胸腔門診貝氏網路：魏克森符號等級和檢定

此範例由架構的相似性分析可知，本研究方法建立之此範例架構與已知架構間的相似度為 69.23%，表示本研究方法只能表達此範例變數間 69.23%的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示本研究方法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

但是由(7)和(23)的架構相似性分析可知，本研究方法於未簡化圖形架構前，所得架構與已知架構間的相似度為 61.90%；簡化圖形架構後，所得架構與已知架構間的相似度為 69.23%。此現象說明，本研究方法於此範例中，簡化前的架構雖不是很理想，但以相同的方法簡化架構，架構會比簡化前更符合已知的貝氏網路架構。

#### 4.3.2 K2 演算法建立之胸腔門診貝氏網路架構

此一節中，本研究擬以 K2 演算法產生此範例之架構，進行架構的相似性評估。由於 K2 演算法執行前須給定一組節點的排序，因此給定一組與本研究相同的排序，則節點排序為 Dyspnea、XRy、TbOrCa、Bronchitis、Cancer、Tuberculosis、Smoking、VisitAsia。



透過貝氏網路應用軟體 Bayesware Discoverer 執行此一步驟，可得以下之結果：

- (1) K2 演算法產生之非循環方向圖，如圖 4.25。
- (2) 利用 Netica，結合(1)得到的非循環方向圖與 4.3.1 之(1)剩下的三分之一資料量，建構貝氏網路，最終得到圖 4.26。

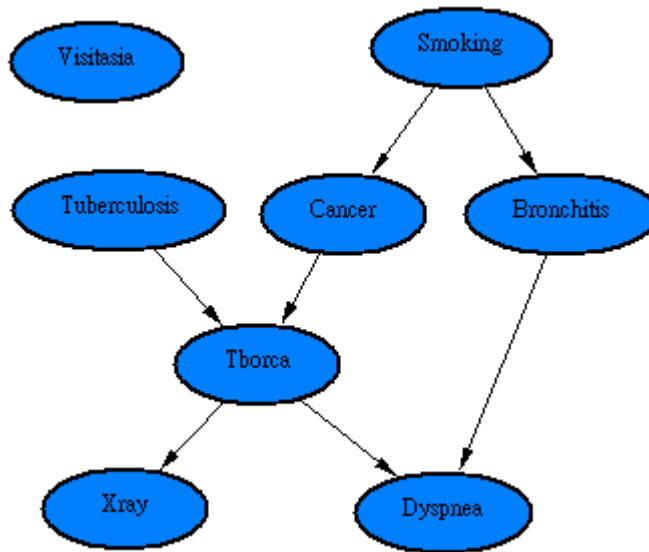


圖 4.25 K2 演算法建立的胸腔門診貝氏網路之非循環方向圖

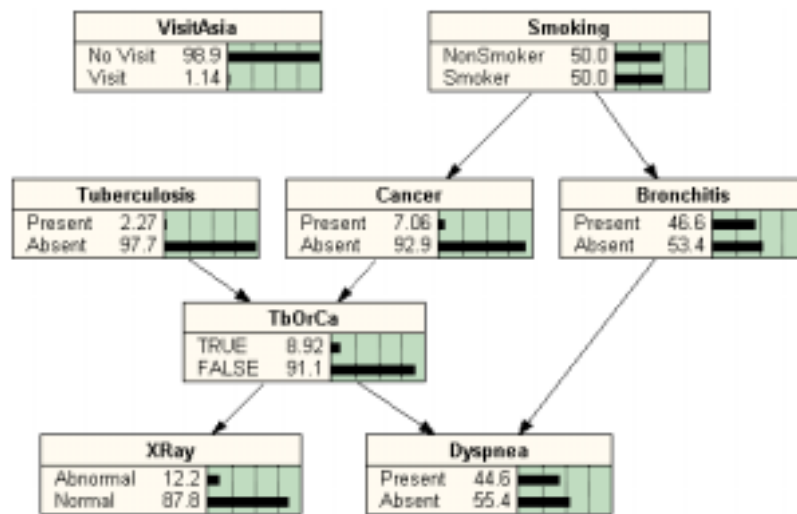


圖 4.26 K2 演算法建立之胸腔門診貝氏網路架構

(3) 架構的相似性分析

- a. 此範例的貝氏網路中，所有存在之弧線形成的集合  
 $X = \{(\text{VisitAsia}, \text{Tuberculosis}), (\text{VisitAsia}, \text{TbOrCa}), (\text{VisitAsia}, \text{XRay}), (\text{VisitAsia}, \text{Dyspnea}), (\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{TbOrCa}, \text{XRay}), (\text{TbOrCa}, \text{Dyspnea}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Bronchitis}, \text{Dyspnea})\}$ 。  $X$  中的弧線個數  $x$  為 18 個。
- b. K2 演算法產生之貝氏網路架構中，所有存在之弧線形成的集合  $Y = \{(\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{TbOrCa}, \text{XRay}), (\text{TbOrCa}, \text{Dyspnea}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Bronchitis}, \text{Dyspnea})\}$ 。  $Y$  中的弧線個數  $y$  為 14 個。
- c. 由 a 和 b 可知， $X$  和  $Y$  之共同弧線形成的集合為  
 $X \cap Y = \{(\text{Tuberculosis}, \text{TbOrCa}), (\text{Tuberculosis}, \text{XRay}), (\text{Tuberculosis}, \text{Dyspnea}), (\text{TbOrCa}, \text{XRay}), (\text{TbOrCa}, \text{Dyspnea}), (\text{Smoking}, \text{Cancer}), (\text{Smoking}, \text{Bronchitis}), (\text{Smoking}, \text{TbOrCa}), (\text{Smoking}, \text{XRay}), (\text{Smoking}, \text{Dyspnea}), (\text{Cancer}, \text{TbOrCa}), (\text{Cancer}, \text{XRay}), (\text{Cancer}, \text{Dyspnea}), (\text{Bronchitis}, \text{Dyspnea})\}$ 。  $X \cap Y$  中的弧線個數  $m$  為 14 個。
- d. 由 a，b 和 c 可計算集合  $X$  和  $Y$  之間相似性的 Jaccard 係數為

$$\frac{m}{x + y - m} = \frac{14}{18 + 14 - 14} = 0.7778$$

由此可知，兩個架構的相似程度為 77.78%。

(4) 資料的相似性分析

a. 顯著水準 0.05 之假設檢定的兩個假設條件為

$H_0$ : 新資料與原始資料相同，即實驗組足以描述對照組

$H_1$ : 新資料與原始資料不相同，即實驗組不能描述對照組

b. 統計新資料與原始資料之變數組合的觀測次數及兩個觀測次數的差。( K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表，請參閱附錄三的表 A-3.8。 )

c. 本研究使用統計應用軟體 MINITAB 計算雙尾的魏克森符號等級和檢定的  $p$  值。計算結果如圖 4.27 所示， $p$  值為 0.177 大於 0.025，故不否定虛無假設  $H_0$ ，即實驗組足以描述對照組。

由架構相似性分析可知，K2 演算法建立之此範例架構與已知架構間相似度為 77.78%，表示此法表達此範例 77.78% 變數間的關係。由資料的相似性分析可知，兩個架構各自產生的兩組資料間，變數結構是相同的。表示 K2 演算法建立之此範例架構，再經由參數值的初始化後，足以描述此範例變數的發生情形。

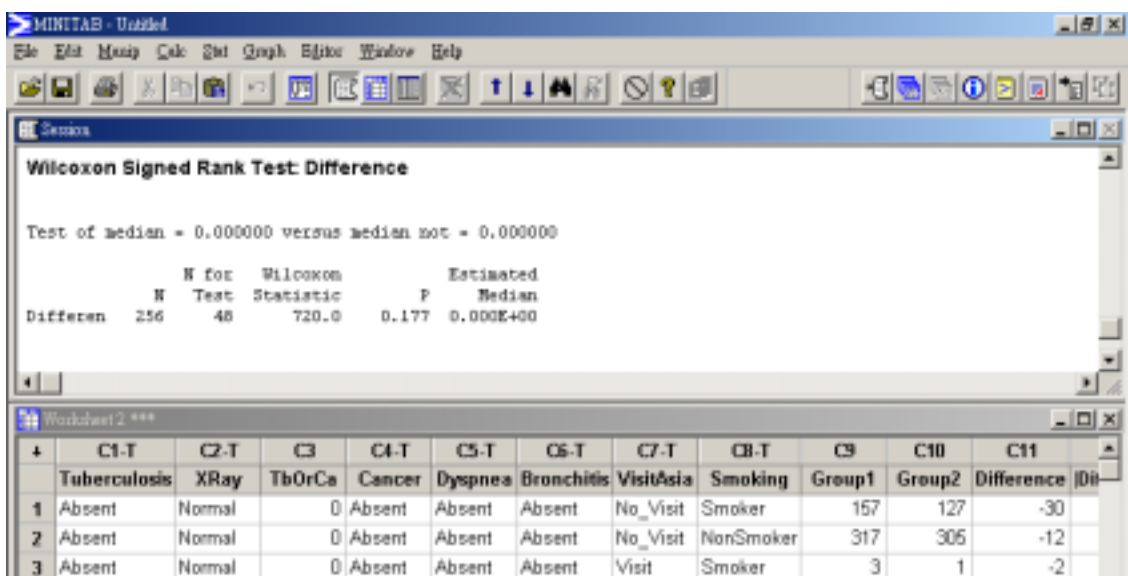


圖 4.27 K2 演算法的胸腔門診貝氏網路：魏克森符號等級和檢定

## 4.4 本章小結

根據前幾節的實證研究，可得表 4.19 和表 4.20 的結果。由實證研究的結果與這兩個表格，本研究得到以下之結論：

### 1. 警鈴貝氏網路

本研究方法建立之此範例架構較 K2 演算法所建立之架構好。資料部分兩者均足以描述問題，表示本研究方法於此範例建立之架構的描述能力不會比 K2 演算法差。

### 2. 文具袋貝氏網路

本研究方法建立之此範例架構較 K2 演算法所建立之架構差。資料部分兩者均足以描述問題，表示本研究方法於此範例建立之架構的描述能力不會比 K2 演算法差。

### 3. 胸腔門診貝氏網路

本研究方法建立之此範例架構較 K2 演算法所建立之架構差。資料部分兩者均足以描述問題，表示本研究方法於此範例建立之架構的描述能力不會比 K2 演算法差。

表 4.19 架構相似度比較表

貝氏網路 建立方法	警鈴	文具袋	胸腔門診
本研究方法	100%	33.33%	69.23%
K2 演算法	62.5%	100%	77.78%

表 4.20 資料相似度比較表 (p 值)

貝氏網路 建立方法	警鈴	文具袋	胸腔門診
本研究方法	0.838	0.425	0.027
K2 演算法	0.308	0.988	0.177

\*顯著水準 0.05 之雙尾魏克森符號等級和檢定

## 第五章 結論與未來研究方向

### 5.1 結論

本研究期望藉由數學明確性的特質，從數學的觀點，結合圖論，找出建構貝氏網路時必備的非循環方向圖。本研究透過矩陣的相似性轉換，找到相似矩陣，再結合圖論畫出貝氏網路架構。由於此方法並非搜尋法，所以資料的多寡不影響本研究方法的執行效率。

經過第三章之研究方法設計與第四章之三個範例的實證研究，本研究得到以下之結論：

警鈴、文具袋及胸腔門診三個範例的架構相似度分別為 100%、33.33% 及 69.23%。本研究方法建立之架構只有警鈴貝氏網路的表現較好，其餘兩個範例的表現並不理想。但資料相似度部分三個範例均足以描述問題，表示本研究方法於此範例建立之架構的描述能力不會比搜尋式演算法差。

根據以上之結論，本研究探討可能造成上述現象的原因：

1. 矩陣相似性轉換中，節點排序是一個給定的可能排序，這樣的排序對於單純的架構影響較小，如警鈴貝氏網路。但對於複雜及節點間因果關係不強的架構較為敏感，如胸腔門診貝氏網路及文具袋貝氏網路。
2. 矩陣與圖形轉換中，本研究僅以子矩陣是否為零矩陣來判定節點間的方向性。子矩陣內部元素的效應未納入考慮，例如：子矩陣內部元素值的大小，是否可視為節點間方向性的權重。還有元素值的正負是否會影響子矩陣方向性的判定結果。

除了上述可能的原因外，以下將探討本研究方法的設計假設是否可能影響結果：

1. 資料轉換中，節點與節點間具有不相交性。

節點僅具有一個特質，使節點與節點間具有不相交性。此項設計使矩陣型式為區塊矩陣，但本研究方法中將矩陣中的最小元素  $1 \times 1$  視為一個單位，以一般矩陣的處理方式執行運算。此一過程，可能使兩節點間的關係受到其它節點的影響。若未來能以區

塊矩陣的處理方式執行本研究方法，將能使運算過程更符合此項設計，避免其它節點的影響。

2. 資料轉換中，矩陣的元素值，以條件機率方式計算。

此項假設是為了符合貝氏網路具有因果關係的設計。此項設計使矩陣元素間的關係更密切，使矩陣的設計更能表現資料的因果關係。

3. 矩陣相似性轉換中，以 Jordan 標準型中特徵值 0 出現與否，判斷矩陣是否可以簡化節點間的關係。

本研究的矩陣元素順序是給定的，是一組較符合此項設計的順序。若能找出節點與 Jordan 標準型各個對角線元素的相對應關係，將使本研究方法的執行更有效率。

4. 矩陣與圖形轉換中，因資料轉換時，採條件機率的方式。若矩陣於執行過程曾經過轉置，則相似矩陣於轉換圖形時，其矩陣之列項目與矩陣之行項目的對應關係須修正為矩陣之行對應到矩陣之列。

本研究方法的實證研究中，矩陣均未經過轉置，故無法判定此項假設是否會影響結果。

## 5.2 未來研究方向

根據上一節的結論，本研究方法有以下幾項可能發展的方向：

1. 變數的排序與轉換矩陣之間的關係。例如：兩者之間是否有對應關係。
2. 子矩陣內部元素值是否會影響子矩陣方向性的判定結果。
3. 關係較複雜的圖形，是否可能藉由 Jordan 標準型，於不影響變數間關係的情況下，將變數切割成較小的集合。
4. 區塊矩陣方法論。

除了本研究方法外，本論文第一章曾提及，貝氏網路架構的建立方法，除本研究所提出的研究方法外，尚有五種可能的建立方法。目前這五種方法中，雖然排序架構的方式仍有尚未克服的問題點，但此方法已足夠穩定，被廣泛地應用。其餘四種方法則為概念，尚未有完整的理論發展。

因此，貝氏網路架構的建立有以下發展方向：

1. 排序架構方式最大的問題點在於節點的排序。若能發展出更好，更有效的探索法，解決節點的排序問題，則此方法所產生的結果將會有更好的效果。
2. 若能不影響資料間的相關性，直接將資料以三角矩陣的方式表示，透過圖論，畫出非循環方向圖，即可得到貝氏網路的基本架構。
3. 資料以矩陣的方式表示，可以依圖論畫出循環方向圖。若能發展出將循環方向圖轉成非循環方向圖的規則或方法論，如此亦是建立貝氏網路架構的方法。
4. 從資料中找到描述此資料的循環方向圖後，可以直接將此方向圖畫成貝氏網路同形體的方法論。此貝氏網路同形體可能存在貝氏網路中不允許存在的因果循環。
5. 若有一個貝氏網路同形體存在，將此同形體轉成貝氏網路的方法論。或是直接結合貝氏網路同形體與資料進行參數學習的方法論。

上述的方法中，2 和 3 的發展可以使貝氏網路架構的建立較為快速，因為 2 和 3 的發展是以數學為理論基礎。數學具有唯一性，可以避免因探索所造成的時間成本。讓貝氏網路的架構的學習更有效率。此外，貝氏網路最為垢病之處即是因果循環不存在的基本假設，但現實世界中，因果循環是存在的。貝氏網路同形體具有因果循環的特性，所以貝氏網路同形體的發展更可能成為貝氏網路的一大突破，應用的範圍也將更為廣泛，更切合實際情況。

## 參考文獻

- [1] 尼維爾和蓋利, 結構分析 統一古典矩陣求解法, 馬吉康(譯), 徐氏基金會, 臺北市, 民國 65 年。
- [2] 林祥良, 「矩陣的高度特徵數及階層特徵數之關係」, 碩士論文, 淡江大學數學研究所, 民國 80 年。
- [3] 林惠玲和陳正倉, 應用統計學, 雙葉書廊有限公司, 台北, 1999。
- [4] 吳俊謀, 計算機應用, <http://webclass.ncu.edu.tw/~junwu/index.html>
- [5] 呂靜芳, 「由網站行為歷程以貝式學習建立學習者模式之引導系統」, 碩士論文, 中央大學資訊工程研究所, 民國 88 年。
- [6] 廖嘉雄, 「以貝氏網路為基礎的知識處理機制之研究」, 碩士論文, 東海大學工業工程研究所, 民國 89 年。
- [7] 蔡易達, 「運用貝氏網路建立台灣地區高風險上市公司之偵測模型」, 碩士論文, 元智大學資訊研究所, 民國 89 年。
- [8] 潘俊帆, 「以貝氏網路為基礎的個人差異化知識處理機制」, 碩士論文, 東海大學工業工程研究所, 民國 90 年。
- [9] 鄭重山, 「貝氏網路之一致性修正比較及其實證研究」, 碩士論文, 清華大學工業工程與工程管理學系, 民國 89 年。
- [10] 蘇祐萱, 「貝氏網路於輔助盈餘預估分析之研究」, 碩士論文, 元智大學資訊研究所資訊管理組, 民國 89 年。
- [11] Abdo, George and Paul Godfrey, “Four MATLAB programs for the complex floating point Jordan matrix decomposition,” <http://winnie.fit.edu/~gabdo/jordan.txt>
- [12] Berman, Abraham and Mark Krupnik, “Spectrum Preserving Lower Triangular Completions -- The Nonnegative Nilpotent Case,” *The Electronic Journal of Linear Algebra*, Vol. 2, 9-16, July 1997.
- [13] Boyd, Stephen, “Lecture 10: Jordan canonical form,” Introduction to Linear Dynamical Systems, <http://www.stanford.edu/class/ee263/>, Jan. 2001.
- [14] Brualdi, Richard A. and Herbert J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, Cambridge, NY, 1991.
- [15] Clarke, Michael, “Implementing matrix algorithms on matrix coprocessors,” M.Sc. Thesis, Uni. of South Australia, Dec. 1997.



- [16]Cooper, Gregory F. and Edward Herskovits, “A Bayesian Method for the Induction of Probabilistic Networks from Data,” *Machine Learning*, 9, 309-347, 1992.
- [17]Dongarra, Jack J. and Victor Eijkhout, “Numerical linear algebra algorithms and software,” *Journal CAM (Numerical) Linear Algebra*, Vol. 31, Number 4, October 1999.
- [18]Gere, James M. and William Weaver, *Matrix Algebra for Engineers*, Van Nostrand, 1965.
- [19]Haddawy, P., “An Overview of Some Recent Developments in Bayesian Problem Solving Techniques”, *AI Magazine*, Spring, 1999.
- [20]Heckerman, David, “A Tutorial on Learning Bayesian Networks,” *Technical Report MSR-TR-95-06*, Microsoft Research, Microsoft Corporation, July 1995.
- [21]Heckerman, David, “A Tutorial on Learning With Bayesian Networks,” *Technical Report MSR-TR-95-06*, Microsoft Research, Microsoft Corporation, March 1995.
- [22]Hershkowitz, Daniel, “The combinatorial structure of generalized eigenspaces – from nonnegative matrices to general matrices,” Elsevier Science Inc, 1999.
- [23]Hershkowitz, Daniel and Hans Schneider, “Path Coverings of Graphs and Height Characteristics of Matrices,” *Journal of Combinatorial Theory*, Series B 59, 172-187, 1991.
- [24]Herskovits, E. H., “Computer-based probabilistic-network construction,” Doctoral dissertation, Medical Information Sciences, Stanford University, Stanford, CA, 1991.
- [25]Irony, Dror and Sivan Toledo, “Trading Replication for Communication in Parallel Distributed-Memory Dense Solvers,” *Parallel Processing Letters*, July 2001.
- [26]Keller, G. and B. Warrack, *Statistics for Management and Economics*, Fifth Edition, Duxbury Press, 594-600, 1999.
- [27]McGuire, *William and Richard H. Gallagher*, *Matrix structural analysis*, Wiley, New York, 374-390, 1978.

- [28]Quigg, John, “Uniqueness of Jordan Form,”  
<http://math.asu.edu/~quigg/teach/courses/442/2000f/lectures/>, 2001.
- [29]Searle, Shayle R., *Matrix Algebra Useful for Statistics*, Wiley, New York, 1985.
- [30]Singh, Moninder and Marco Valtorta, “Construction of Bayesian Network Structures From Data,” *International Journal of Approximate Reasoning*, 12:111-131, 1995.
- [31]Stoudt, Gary S., “Jordan Canonical Form,”  
<http://www.ma.iup.edu/projects/CalcDEMma/JCF/jcf.html>, 1994.
- [32]Strang, Gilbert, *Linear algebra and its applications*, Academic Press, New York, 1980.
- [33]Thseng, F.S., *et al.*, “Glycine formosana Hosokawa in Taiwan- pod morphology allozyme and DNA polymorphism,” *Bot. Bull. Acad. Sin.*, 40:251-257, 1999.
- [34]Verma, T. and J. Pearl, “An algorithm for deciding if a set of observed independencies has a causal explanation,” *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, 323-330, 1992.
- [35]Weisstein, Eric, *world of mathematics*, <http://mathworld.wolfram.com/>, site last updated: Apr 2002
- [36]Yang, Yau-Wen, Warren H.J. Kuo, and Teck Hiang Wong, “Genetic polymorphism of seven populations of *Capsella bursapastoris* based on RAPD markers,” *Bot. Bull. Acad. Sin.*, 39: 17-21, 1998.
- [37]Yousef, Saad, *Iterative Methods for Sparse Linear Systems*, <http://www.cs.umn.edu/~saad>, Second edition, Jan 2000.

## 附錄一 Judea Pearl 警鈴貝氏網路之相似性分析 統計表

表 A-1.1 本研究方法的警鈴貝氏網路：新資料與原始資料之變數組合  
統計表

MaryCalls	JohnCalls	Alarm	Earthquake	Burglary	Group1	Group2	Difference	Difference
No	No	No	No	No	973	961	-12	12
No	No	No	No	Yes	0	1	1	1
No	No	No	Yes	No	2	0	-2	2
No	No	No	Yes	Yes	0	0	0	0
No	No	Yes	No	No	1	0	-1	1
No	No	Yes	No	Yes	0	0	0	0
No	No	Yes	Yes	No	0	0	0	0
No	No	Yes	Yes	Yes	0	0	0	0
No	Yes	No	No	No	58	57	-1	1
No	Yes	No	No	Yes	0	0	0	0
No	Yes	No	Yes	No	0	0	0	0
No	Yes	No	Yes	Yes	0	0	0	0
No	Yes	Yes	No	No	0	3	3	3
No	Yes	Yes	No	Yes	2	3	1	1
No	Yes	Yes	Yes	No	0	0	0	0
No	Yes	Yes	Yes	Yes	0	0	0	0
Yes	No	No	No	No	9	22	13	13
Yes	No	No	No	Yes	0	0	0	0
Yes	No	No	Yes	No	0	0	0	0
Yes	No	No	Yes	Yes	0	0	0	0
Yes	No	Yes	No	No	0	0	0	0
Yes	No	Yes	No	Yes	0	0	0	0
Yes	No	Yes	Yes	No	0	0	0	0
Yes	No	Yes	Yes	Yes	0	0	0	0
Yes	Yes	No	No	No	1	1	0	0
Yes	Yes	No	No	Yes	0	0	0	0
Yes	Yes	No	Yes	No	0	0	0	0
Yes	Yes	No	Yes	Yes	0	0	0	0
Yes	Yes	Yes	No	No	2	1	-1	1
Yes	Yes	Yes	No	Yes	2	1	-1	1
Yes	Yes	Yes	Yes	No	0	0	0	0
Yes	Yes	Yes	Yes	Yes	0	0	0	0

表 A-1.2 K2 演算法的警鈴貝氏網路：新資料與原始資料之變數組合統計表

MaryCalls	JohnCalls	Alarm	Earthquake	Burglary	Group1	Group2	Difference	Difference
No	No	No	No	No	973	958	-15	15
No	No	No	No	Yes	0	2	2	2
No	No	No	Yes	No	2	1	-1	1
No	No	No	Yes	Yes	0	0	0	0
No	No	Yes	No	No	1	1	0	0
No	No	Yes	No	Yes	0	2	2	2
No	No	Yes	Yes	No	0	0	0	0
No	No	Yes	Yes	Yes	0	0	0	0
No	Yes	No	No	No	58	56	-2	2
No	Yes	No	No	Yes	0	0	0	0
No	Yes	No	Yes	No	0	1	1	1
No	Yes	No	Yes	Yes	0	0	0	0
No	Yes	Yes	No	No	0	0	0	0
No	Yes	Yes	No	Yes	2	4	2	2
No	Yes	Yes	Yes	No	0	0	0	0
No	Yes	Yes	Yes	Yes	0	0	0	0
Yes	No	No	No	No	9	15	6	6
Yes	No	No	No	Yes	0	0	0	0
Yes	No	No	Yes	No	0	0	0	0
Yes	No	No	Yes	Yes	0	0	0	0
Yes	No	Yes	No	No	0	1	1	1
Yes	No	Yes	No	Yes	0	1	1	1
Yes	No	Yes	Yes	No	0	0	0	0
Yes	No	Yes	Yes	Yes	0	0	0	0
Yes	Yes	No	No	No	1	1	0	0
Yes	Yes	No	No	Yes	0	0	0	0
Yes	Yes	No	Yes	No	0	0	0	0
Yes	Yes	No	Yes	Yes	0	0	0	0
Yes	Yes	Yes	No	No	2	1	-1	1
Yes	Yes	Yes	No	Yes	2	6	4	4
Yes	Yes	Yes	Yes	No	0	0	0	0
Yes	Yes	Yes	Yes	Yes	0	0	0	0

## 附錄二 L. D. Phillips 和 W. Edward 文具袋貝氏網路之相似性分析統計表

表 A-2.1 本研究方法的文具袋貝氏網路：新資料與原始資料之變數組合統計表

Draw_1	Draw_2	Draw_3	Draw_4	Draw_5	Book_Bag	Group1	Group2	Difference	Difference
red	red	red	red	blue	bag_1	49	43	-6	6
red	red	red	red	blue	bag_2	0	10	10	10
red	red	red	red	red	bag_1	103	68	-35	35
red	red	red	red	red	bag_2	0	20	20	20
red	red	red	blue	blue	bag_1	18	12	-6	6
red	red	red	blue	blue	bag_2	7	7	0	0
red	red	red	blue	red	bag_1	47	14	-33	33
red	red	red	blue	red	bag_2	2	18	16	16
red	red	blue	red	blue	bag_1	15	27	12	12
red	red	blue	red	blue	bag_2	5	5	0	0
red	red	blue	red	red	bag_1	37	18	-19	19
red	red	blue	red	red	bag_2	2	5	3	3
red	red	blue	blue	blue	bag_1	19	9	-10	10
red	red	blue	blue	blue	bag_2	11	15	4	4
red	red	blue	blue	red	bag_1	21	12	-9	9
red	red	blue	blue	red	bag_2	3	23	20	20
red	blue	red	red	blue	bag_1	29	32	3	3
red	blue	red	red	blue	bag_2	7	15	8	8
red	blue	red	red	red	bag_1	52	31	-21	21
red	blue	red	red	red	bag_2	4	9	5	5
red	blue	red	blue	blue	bag_1	9	7	-2	2
red	blue	red	blue	blue	bag_2	9	18	9	9
red	blue	red	blue	red	bag_1	25	15	-10	10
red	blue	red	blue	red	bag_2	5	22	17	17
red	blue	blue	red	blue	bag_1	5	10	5	5
red	blue	blue	red	blue	bag_2	13	9	-4	4
red	blue	blue	red	red	bag_1	19	28	9	9
red	blue	blue	red	red	bag_2	3	3	0	0
red	blue	blue	blue	blue	bag_1	1	10	9	9
red	blue	blue	blue	blue	bag_2	25	12	-13	13
red	blue	blue	blue	red	bag_1	7	15	8	8
red	blue	blue	blue	red	bag_2	10	18	8	8
blue	red	red	red	blue	bag_1	20	22	2	2
blue	red	red	red	blue	bag_2	3	6	3	3
blue	red	red	red	red	bag_1	47	30	-17	17
blue	red	red	red	red	bag_2	1	9	8	8
blue	red	red	blue	blue	bag_1	15	10	-5	5
blue	red	red	blue	blue	bag_2	15	24	9	9
blue	red	red	blue	red	bag_1	25	18	-7	7

( 下頁繼續 )

表 A-2.2 本研究方法的文具袋貝氏網路：新資料與原始資料之變數組合統計表（續）

Draw_1	Draw_2	Draw_3	Draw_4	Draw_5	Book_Bag	Group1	Group2	Difference	Difference
blue	red	red	blue	red	bag_2	8	29	21	21
blue	red	blue	red	blue	bag_1	9	15	6	6
blue	red	blue	red	blue	bag_2	17	5	-12	12
blue	red	blue	red	red	bag_1	22	26	4	4
blue	red	blue	red	red	bag_2	5	6	1	1
blue	red	blue	blue	blue	bag_1	2	11	9	9
blue	red	blue	blue	blue	bag_2	26	18	-8	8
blue	red	blue	blue	red	bag_1	9	11	2	2
blue	red	blue	blue	red	bag_2	10	20	10	10
blue	blue	red	red	blue	bag_1	12	25	13	13
blue	blue	red	red	blue	bag_2	9	3	-6	6
blue	blue	red	red	red	bag_1	21	33	12	12
blue	blue	red	red	red	bag_2	3	11	8	8
blue	blue	red	blue	blue	bag_1	5	15	10	10
blue	blue	red	blue	blue	bag_2	29	27	-2	2
blue	blue	red	blue	red	bag_1	4	13	9	9
blue	blue	red	blue	red	bag_2	17	19	2	2
blue	blue	blue	red	blue	bag_1	3	14	11	11
blue	blue	blue	red	blue	bag_2	26	3	-23	23
blue	blue	blue	red	red	bag_1	8	14	6	6
blue	blue	blue	red	red	bag_2	12	0	-12	12
blue	blue	blue	blue	blue	bag_1	1	9	8	8
blue	blue	blue	blue	blue	bag_2	69	22	-47	47
blue	blue	blue	blue	red	bag_1	5	11	6	6
blue	blue	blue	blue	red	bag_2	30	11	-19	19

表 A-2.3 K2 演算法的文具袋貝氏網路：新資料與原始資料之變數組合  
統計表

Draw_1	Draw_2	Draw_3	Draw_4	Draw_5	Book_Bag	Group1	Group2	Difference	Difference
blue	red	red	blue	blue	bag_2	15	11	-4	4
blue	red	red	blue	blue	bag_1	15	2	-13	13
blue	red	red	blue	red	bag_2	8	4	-4	4
blue	red	red	blue	red	bag_1	25	21	-4	4
blue	red	red	red	blue	bag_2	3	4	1	1
blue	red	red	red	blue	bag_1	20	18	-2	2
blue	red	red	red	red	bag_2	1	2	1	1
blue	red	red	red	red	bag_1	47	49	2	2
blue	red	blue	blue	blue	bag_2	26	32	6	6
blue	red	blue	blue	blue	bag_1	2	1	-1	1
blue	red	blue	blue	red	bag_2	10	14	4	4
blue	red	blue	blue	red	bag_1	9	5	-4	4
blue	red	blue	red	blue	bag_2	17	13	-4	4
blue	red	blue	red	blue	bag_1	9	4	-5	5
blue	red	blue	red	red	bag_2	5	0	-5	5
blue	red	blue	red	red	bag_1	22	6	-16	16
blue	blue	red	blue	blue	bag_2	29	39	10	10
blue	blue	red	blue	blue	bag_1	5	11	6	6
blue	blue	red	blue	red	bag_2	17	9	-8	8
blue	blue	red	blue	red	bag_1	4	12	8	8
blue	blue	red	red	blue	bag_2	9	23	14	14
blue	blue	red	red	blue	bag_1	12	6	-6	6
blue	blue	red	red	red	bag_2	3	4	1	1
blue	blue	red	red	red	bag_1	21	16	-5	5
blue	blue	blue	blue	blue	bag_2	69	66	-3	3
blue	blue	blue	blue	blue	bag_1	1	2	1	1
blue	blue	blue	blue	red	bag_2	30	28	-2	2
blue	blue	blue	blue	red	bag_1	5	3	-2	2
blue	blue	blue	red	blue	bag_2	26	35	9	9
blue	blue	blue	red	blue	bag_1	3	3	0	0
blue	blue	blue	red	red	bag_2	12	17	5	5
blue	blue	blue	red	red	bag_1	8	6	-2	2
red	red	red	blue	blue	bag_2	7	3	-4	4
red	red	red	blue	blue	bag_1	18	23	5	5
red	red	red	blue	red	bag_2	2	0	-2	2
red	red	red	blue	red	bag_1	47	61	14	14
red	red	red	red	blue	bag_2	0	6	6	6
red	red	red	red	blue	bag_1	49	36	-13	13
red	red	red	red	red	bag_2	0	1	1	1
red	red	red	red	red	bag_1	103	118	15	15
red	red	blue	blue	blue	bag_2	11	12	1	1
red	red	blue	blue	blue	bag_1	19	5	-14	14
red	red	blue	blue	red	bag_2	3	5	2	2

(下頁繼續)

表 A-2.4 K2 演算法的文具袋貝氏網路：新資料與原始資料之變數組合  
統計表（續）

Draw_1	Draw_2	Draw_3	Draw_4	Draw_5	Book_Bag	Group1	Group2	Difference	Difference
red	red	blue	blue	red	bag_1	21	17	-4	4
red	red	blue	red	blue	bag_2	5	3	-2	2
red	red	blue	red	blue	bag_1	15	16	1	1
red	red	blue	red	red	bag_2	2	1	-1	1
red	red	blue	red	red	bag_1	37	36	-1	1
red	blue	red	blue	blue	bag_2	9	9	0	0
red	blue	red	blue	blue	bag_1	9	13	4	4
red	blue	red	blue	red	bag_2	5	5	0	0
red	blue	red	blue	red	bag_1	25	24	-1	1
red	blue	red	red	blue	bag_2	7	9	2	2
red	blue	red	red	blue	bag_1	29	33	4	4
red	blue	red	red	red	bag_2	4	3	-1	1
red	blue	red	red	red	bag_1	52	56	4	4
red	blue	blue	blue	blue	bag_2	25	30	5	5
red	blue	blue	blue	blue	bag_1	1	6	5	5
red	blue	blue	blue	red	bag_2	10	11	1	1
red	blue	blue	blue	red	bag_1	7	9	2	2
red	blue	blue	red	blue	bag_2	13	10	-3	3
red	blue	blue	red	blue	bag_1	5	7	2	2
red	blue	blue	red	red	bag_2	3	3	0	0
red	blue	blue	red	red	bag_1	19	13	-6	6



# 附錄三 Steffen L. Lauritzen 和 David J. Spiegelhalter 胸腔門診貝氏網路之相似性分析統計表

表 A-3.1 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Normal	FALSE	Absent	Absent	Absent	No_Visit	Smoker	157	128	-29	29
Absent	Normal	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	317	132	-185	185
Absent	Normal	FALSE	Absent	Absent	Absent	Visit	Smoker	3	1	-2	2
Absent	Normal	FALSE	Absent	Absent	Absent	Visit	NonSmoker	3	1	-2	2
Absent	Normal	FALSE	Absent	Absent	Present	No_Visit	Smoker	59	105	46	46
Absent	Normal	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	29	105	76	76
Absent	Normal	FALSE	Absent	Absent	Present	Visit	Smoker	1	1	0	0
Absent	Normal	FALSE	Absent	Absent	Present	Visit	NonSmoker	1	1	0	0
Absent	Normal	FALSE	Absent	Present	Absent	No_Visit	Smoker	9	103	94	94
Absent	Normal	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	37	101	64	64
Absent	Normal	FALSE	Absent	Present	Absent	Visit	Smoker	0	1	1	1
Absent	Normal	FALSE	Absent	Present	Absent	Visit	NonSmoker	1	1	0	0
Absent	Normal	FALSE	Absent	Present	Present	No_Visit	Smoker	209	69	-140	140
Absent	Normal	FALSE	Absent	Present	Present	No_Visit	NonSmoker	115	88	-27	27
Absent	Normal	FALSE	Absent	Present	Present	Visit	Smoker	1	0	-1	1
Absent	Normal	FALSE	Absent	Present	Present	Visit	NonSmoker	2	1	-1	1
Absent	Normal	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	15	15	15
Absent	Normal	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	2	2	2
Absent	Normal	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Present	No_Visit	Smoker	0	18	18	18
Absent	Normal	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	8	8	8
Absent	Normal	FALSE	Present	Absent	Present	Visit	Smoker	0	2	2	2
Absent	Normal	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	No_Visit	Smoker	0	9	9	9
Absent	Normal	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	1	1	1
Absent	Normal	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Present	No_Visit	Smoker	0	14	14	14
Absent	Normal	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	3	3	3
Absent	Normal	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.2 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 1）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Normal	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	1	1	1
Absent	Normal	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	1	1	1
Absent	Normal	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	No_Visit	Smoker	0	1	1	1
Absent	Normal	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	No_Visit	Smoker	1	0	-1	1
Absent	Normal	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	Smoker	9	5	-4	4
Absent	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	13	4	-9	9
Absent	Abnorma	FALSE	Absent	Absent	Absent	Visit	Smoker	0	1	1	1
Absent	Abnorma	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	4	4	4
Absent	Abnorma	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	4	4	4
Absent	Abnorma	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Absent	No_Visit	Smoker	1	5	4	4
Absent	Abnorma	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	1	6	5	5
Absent	Abnorma	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Absent	Visit	NonSmoker	1	0	-1	1

（下頁繼續）

表 A-3.3 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 2）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Abnorma	FALSE	Absent	Present	Present	No_Visit	Smoker	5	4	-1	1
Absent	Abnorma	FALSE	Absent	Present	Present	No_Visit	NonSmoker	3	3	0	0
Absent	Abnorma	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	1	1	1
Absent	Abnorma	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	7	7	7
Absent	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	3	3	3
Absent	Abnorma	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	3	3	3
Absent	Abnorma	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	6	6	6
Absent	Abnorma	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	8	8	8
Absent	Abnorma	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	8	8	8
Absent	Abnorma	TRUE	Absent	Present	Absent	Visit	Smoker	0	1	1	1
Absent	Abnorma	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Present	No_Visit	Smoker	0	15	15	15
Absent	Abnorma	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	17	17	17
Absent	Abnorma	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Present	Visit	NonSmoker	0	1	1	1
Absent	Abnorma	TRUE	Present	Absent	Absent	No_Visit	Smoker	6	0	-6	6
Absent	Abnorma	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	4	0	-4	4
Absent	Abnorma	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Present	No_Visit	Smoker	1	0	-1	1

（下頁繼續）

表 A-3.4 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 3）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Abnorma	TRUE	Present	Absent	Present	No_Visit	NonSmoker	2	0	-2	2
Absent	Abnorma	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Absent	No_Visit	Smoker	16	1	-15	15
Absent	Abnorma	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Present	No_Visit	Smoker	30	2	-28	28
Absent	Abnorma	TRUE	Present	Present	Present	No_Visit	NonSmoker	1	1	0	0
Absent	Abnorma	TRUE	Present	Present	Present	Visit	Smoker	1	0	-1	1
Absent	Abnorma	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Absent	No_Visit	Smoker	0	3	3	3
Present	Normal	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	0	2	2	2
Present	Normal	FALSE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	3	3	3
Present	Normal	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	2	2	2
Present	Normal	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	No_Visit	Smoker	0	4	4	4
Present	Normal	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	0	1	1	1
Present	Normal	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Present	No_Visit	Smoker	0	3	3	3
Present	Normal	FALSE	Absent	Present	Present	No_Visit	NonSmoker	0	4	4	4
Present	Normal	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	No_Visit	Smoker	0	1	1	1
Present	Normal	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.5 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 4）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Normal	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	No_Visit	Smoker	1	0	-1	1
Present	Normal	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0

（下頁繼續）

表 A-3.6 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 5）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Abnorma	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Present	No_Visit	Smoker	0	1	1	1
Present	Abnorma	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	1	1	1
Present	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	1	0	-1	1
Present	Abnorma	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Absent	No_Visit	Smoker	1	0	-1	1
Present	Abnorma	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	2	0	-2	2
Present	Abnorma	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Present	No_Visit	Smoker	3	1	-2	2
Present	Abnorma	TRUE	Absent	Present	Present	No_Visit	NonSmoker	1	1	0	0
Present	Abnorma	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.7 本研究方法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表（續 6）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchiti	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Abnorma	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	No_Visit	Smoker	1	0	-1	1
Present	Abnorma	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	No_Visit	Smoker	1	0	-1	1
Present	Abnorma	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0

表 A-3.8 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Normal	FALSE	Absent	Absent	Absent	No_Visit	Smoker	157	127	-30	30
Absent	Normal	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	317	305	-12	12
Absent	Normal	FALSE	Absent	Absent	Absent	Visit	Smoker	3	1	-2	2
Absent	Normal	FALSE	Absent	Absent	Absent	Visit	NonSmoker	3	5	2	2
Absent	Normal	FALSE	Absent	Absent	Present	No_Visit	Smoker	59	55	-4	4
Absent	Normal	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	29	30	1	1
Absent	Normal	FALSE	Absent	Absent	Present	Visit	Smoker	1	1	0	0
Absent	Normal	FALSE	Absent	Absent	Present	Visit	NonSmoker	1	0	-1	1
Absent	Normal	FALSE	Absent	Present	Absent	No_Visit	Smoker	9	13	4	4
Absent	Normal	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	37	31	-6	6
Absent	Normal	FALSE	Absent	Present	Absent	Visit	Smoker	0	1	1	1
Absent	Normal	FALSE	Absent	Present	Absent	Visit	NonSmoker	1	1	0	0
Absent	Normal	FALSE	Absent	Present	Present	No_Visit	Smoker	209	231	22	22
Absent	Normal	FALSE	Absent	Present	Present	No_Visit	NonSmoker	115	113	-2	2
Absent	Normal	FALSE	Absent	Present	Present	Visit	Smoker	1	4	3	3
Absent	Normal	FALSE	Absent	Present	Present	Visit	NonSmoker	2	2	0	0
Absent	Normal	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Present	No_Visit	Smoker	0	2	2	2
Absent	Normal	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)



表 A-3.9 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數組合統計表（續 1）

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Normal	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	No_Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	1	1	1
Absent	Normal	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	1	1	1
Absent	Normal	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	No_Visit	Smoker	1	0	-1	1
Absent	Normal	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	No_Visit	Smoker	0	1	1	1
Absent	Normal	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	No_Visit	Smoker	0	1	1	1
Absent	Normal	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Normal	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	Smoker	9	9	0	0
Absent	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	13	9	-4	4
Absent	Abnorma	FALSE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	5	5	5
Absent	Abnorma	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	1	1	1
Absent	Abnorma	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Absent	No_Visit	Smoker	1	2	1	1
Absent	Abnorma	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	1	1	0	0
Absent	Abnorma	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Absent	Visit	NonSmoker	1	1	0	0
Absent	Abnorma	FALSE	Absent	Present	Present	No_Visit	Smoker	5	5	0	0
Absent	Abnorma	FALSE	Absent	Present	Present	No_Visit	NonSmoker	3	2	-1	1
Absent	Abnorma	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.10 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表 (續 2)

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Abnorma	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Present	No_Visit	Smoker	0	1	1	1
Absent	Abnorma	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Absent	No_Visit	Smoker	6	6	0	0
Absent	Abnorma	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	4	2	-2	2
Absent	Abnorma	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Present	No_Visit	Smoker	1	6	5	5
Absent	Abnorma	TRUE	Present	Absent	Present	No_Visit	NonSmoker	2	1	-1	1
Absent	Abnorma	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Absent	No_Visit	Smoker	16	8	-8	8
Absent	Abnorma	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	7	7	7

(下頁繼續)

表 A-3.11 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表 (續 3)

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Absent	Abnorma	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Absent	Abnorma	TRUE	Present	Present	Present	No_Visit	Smoker	30	21	-9	9
Absent	Abnorma	TRUE	Present	Present	Present	No_Visit	NonSmoker	1	3	2	2
Absent	Abnorma	TRUE	Present	Present	Present	Visit	Smoker	1	2	1	1
Absent	Abnorma	TRUE	Present	Present	Present	Visit	NonSmoker	0	1	1	1
Present	Normal	FALSE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	0	2	2	2
Present	Normal	FALSE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Present	No_Visit	Smoker	0	1	1	1
Present	Normal	FALSE	Absent	Present	Present	No_Visit	NonSmoker	0	1	1	1
Present	Normal	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.12 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表 (續 4)

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Normal	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	0	1	1	1
Present	Normal	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	No_Visit	Smoker	0	1	1	1
Present	Normal	TRUE	Absent	Present	Present	No_Visit	NonSmoker	0	1	1	1
Present	Normal	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	No_Visit	Smoker	1	0	-1	1
Present	Normal	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Normal	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.13 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表 (續 5)

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Abnorma	FALSE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	No_Visit	Smoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	No_Visit	NonSmoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Visit	Smoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Visit	NonSmoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	No_Visit	Smoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	No_Visit	NonSmoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Visit	Smoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Visit	NonSmoker	0	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	FALSE	Present	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	Smoker	0	3	3	3
Present	Abnorma	TRUE	Absent	Absent	Absent	No_Visit	NonSmoker	1	3	2	2
Present	Abnorma	TRUE	Absent	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	No_Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Absent	No_Visit	Smoker	1	5	4	4
Present	Abnorma	TRUE	Absent	Present	Absent	No_Visit	NonSmoker	2	5	3	3
Present	Abnorma	TRUE	Absent	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Present	No_Visit	Smoker	3	3	0	0
Present	Abnorma	TRUE	Absent	Present	Present	No_Visit	NonSmoker	1	3	2	2
Present	Abnorma	TRUE	Absent	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Absent	Present	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	No_Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	No_Visit	NonSmoker	0	0	0	0

(下頁繼續)

表 A-3.14 K2 演算法的胸腔門診貝氏網路：新資料與原始資料之變數  
組合統計表 (續 6)

Tuberculosis	XRay	TbOrC	Cancer	Dyspne	Bronchitis	VisitAsia	Smoking	Group1	Group2	Difference	Difference
Present	Abnorma	TRUE	Present	Absent	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	No_Visit	Smoker	0	1	1	1
Present	Abnorma	TRUE	Present	Absent	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Absent	Present	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	No_Visit	Smoker	1	1	0	0
Present	Abnorma	TRUE	Present	Present	Absent	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Absent	Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	No_Visit	Smoker	1	0	-1	1
Present	Abnorma	TRUE	Present	Present	Present	No_Visit	NonSmoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	Visit	Smoker	0	0	0	0
Present	Abnorma	TRUE	Present	Present	Present	Visit	NonSmoker	0	0	0	0

## 附錄四 口試相關資料

黃欽印老師

- ◆ 建議

建議：建議使用更公平的評量法，做為結果的評量法。

修正：資料相似性評估，改以魏克森符號等級和檢定為評量法。

建議：建議將間接關係納入架構評量法的評估範圍。

修正：已納入評量中，請參見第 30 頁。

建議：做結果的診斷準確性評估，即診斷及預測的能力。例如：固定 6 個變數，視其它 2 個變數的結果是否相同。

修正：因修正資料相似性的檢定方法後，結果是可以被接受的，故未再修正此建議。

- ◆ 問題

問題：矩陣是否有循環的測試？

答：從矩陣到圖形的轉換，中間過程會先找出矩陣的相似三角矩陣，再以該相似三角矩陣畫出非循環方向圖。非對角線的部分，不會有循環的情形；對角線的部分，用以判斷中間點。

問題：矩陣轉換時，若有循環如何解決？

答：問題領域本質有循環，其解決方法請參見第 69 頁未來研究方向中貝氏網路同形體的發展。

- ◆ 其它

1.問題變數多時，用唯一的方法解決問題，此方法的使用效益較大。

王偉華老師

- ◆ 建議

建議：Michael Jordan 及 David Heckermann 的方式，說明其研究。與他們的相同及相異之處。

建議：比較 K2 演算法和本研究方法的資料相似性檢定。

修正：已修正，請參見第四章之第 4.1.2 節、第 4.2.2 節和第 4.3.2 節。

◆ 問題

問題：K2 演算法是用什麼搜尋法？

答：貪婪搜尋法和全域搜尋法。

丁兆平老師

◆ 建議

建議：建議採用無母數檢定法 符號檢定的魏克森符號等級和檢定 來檢定資料的相似性。

修正：將資料藉由 Microsoft Excel 處理完後，再透過 MINITAB 求出符號檢定的結果 p 值。

建議：與 K2 演算法的比較中，僅考慮架構相似性的比較，建議納入資料相似性的比較。

修正：均改採用魏克森符號等級和檢定。

◆ 問題

問題：排序問題，固定列找 Jordan form。若改變順序有何影響？是否不同？Program 是以 eigenvalues 大小排列或是以重覆來排列？以現階段的排序排列有何好處？

答：本研究方法中的 Jordan 標準型是採用 Paul Godfrey 的求解程式碼，其排列方式無法確定，故採用固定排序的方法執行本研究方法。節點的排序是本研究的未來研究方向，請參見第 68 頁未來研究方向中節點排序與轉換矩陣之間的關係。

問題：建構貝氏網路的可能方法中，②和③有何差異？為何需要轉成矩陣再轉成三角矩陣？是否可以直接做三角矩陣？

答：目前未看到相關文獻。此外，若問題領域本質可以直接轉換成三角矩陣，則可以採用②。但一般問題領域較少可以直接轉成三角矩陣，故透過③。



問題：子矩陣內部的值大小是否有影響？

答：子矩陣內部值的大小及正負號是否影響方向性的判定，現階段未納入考慮，列為未來的研究方向。

陳麗霞老師

◆ 問題

問題：貝氏網路的貝氏用於何處？

答：節點參數值的計算方式，即貝氏網路的計算方式。

問題：如何看出變數間有因果關係？條件機率看方向性，是否有因果關係？是否有文獻說明？

答：假設有兩個事件  $A$  和  $B$ ，分別計算兩者的條件機率  $P(A|B)$  和  $P(B|A)$ 。若  $P(A|B)$  大於  $P(B|A)$ ，則事件  $A$  為因，事件  $B$  為果。文獻說明可參考貝氏網路的相關文獻。

問題：何謂節點與節點具有不相交性？

答：請參見第 13 頁矩陣型式的設計。

問題：區塊分解方式與一般有什麼不同？

答：目前本研究方法是採用一般矩陣分解方式求矩陣的相似矩陣。若改為區塊矩陣分解方式，是否會使本研究的結果更好？此為未來的研究方向。

問題：何以要求非循環方向圖？

答：因為貝氏網路的基本條件為非循環方向圖。

問題：相似矩陣中，元素大小代表何種意義？

答：矩陣元素值的大小目前未納入考慮，列為未來研究方向。

問題：第一章中，何謂局部機率分配？

答：已修正，請參見第 2 頁。

問題：第一章中，變數排序是指重要性或是影響力？

答：是指變數的因果關係。

問題：CB 演算法，為何叫做 CB？

答：已修正，請參見第 3 頁 CB 演算法的註腳。

問題：第 11 頁， $P(B_s)$ 如何決定？

答：因為目前一般解貝氏網路問題，多是於架構已知的情形下進行。所以， $P(B_s)$ 為 1。

王文清老師

◆ 問題

問題： $n \times n$  區塊方陣， $n$  代表什麼？

答：代表節點的個數。

其它

◆ 建議

建議：區塊與區塊的乘法，區塊僅做為表示，不需特別做乘法。

建議：矩陣之行與列的對應關係，會因矩陣的設計而需要修正。並非是矩陣轉換過程中矩陣被轉置而需修正。

修正：已修正，請參見第 9 頁。

◆ 問題

問題：區塊矩陣之相似矩陣是否與一般矩陣作法相同？

答：目前未找到相關文獻，故無法確定是否相同，此為未來的研究方向。

問題：第 11 頁， $P$  指得是否為機率？ $g$  是否為機率值？

答：已加入註腳說明，請參見第 11 頁。