

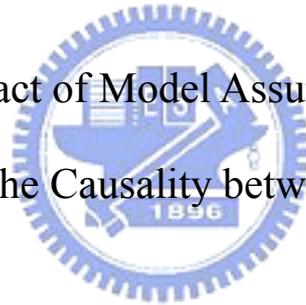
國立交通大學

統計學研究所

碩士論文

模型假設對分析變數間因果關係的影響

The Impact of Model Assumption for
Analyzing the Causality between Variables



研究生：戴如美

指導老師：洪慧念 教授

中華民國九十四年六月

模型假設對分析變數間因果關係的影響

The Impact of Model Assumption for Analyzing the Causality between Variables

研究生：戴如美

Student : Ru-Mei Dai

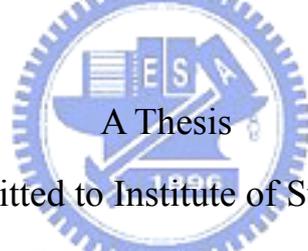
指導教授：洪慧念 博士

Advisor : Dr. Hui-Nien Hung

國立交通大學理學院

統計學研究所

碩士論文



Submitted to Institute of Statistics

College of Science

National Chiao Tung University

In Partial Fulfillment of the Requirements

For the Degree of

Master

In

Statistics

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

模型假設對分析變數間因果關係的影響

研究生：戴如美

指導教授：洪慧念 博士

國立交通大學統計學研究所

摘要

在日常生活或生物基因方面，變數之間總會有影響與被影響的情形發生；在這篇文章中，最主要的目標是希望將變數間有因果關係的資料，經由統計的方法，判斷出因果圖形結構之方向性，及探討模型假設對其結果之影響。

在這個領域之中，有很多科學家及統計學家提出了一些方法如：線性模型(D'haeseleer et al, 1999)、非線性模型(Weaver et al, 1999)、布爾數學邏輯網路模型 (Kauffman 1993, Somogyi and Sniegoski, 1996)等；隨後，Murphy and Mian (1998) 和 Friedman(1999)等人提出貝氏網路模型 (Bayesian Network)。在這篇文章中，試著結合貝氏的概念並探討使用概似函數的方法來分析變數間的因果關係，並希望能判斷出變數間圖形之方向性。

The Impact of Model Assumption for Analyzing the Causality between Variables

Student : Ru-Mei Dai

Advisor : Dr. Hui-Nien Hung

Institute of Statistics

National Chiao Tung University



In daily life or biological gene, there always exists variables which inference other variables or are influenced by some others. In this paper, our main goal is to analyze the variables which have causality by using statistical methods, to determine the directionality among causality graphs, and to find out the impact of model assumption on those methods.

In this field, lots of scientists and statisticians have proposed several methods, for example, linear models (D'haeseleer et al., 1999), nonlinear models (Weaver et al., 1999), and Boolean networks (Kauffman 1993, Somogyi and Sniegoski, 1996), etc. However, there should be some assumptions in all of these models. Murphy and Mian (1998) and Friedman et al. (1999) have suggested that use Bayesian network models of gene expression networks to get some improvements. In this paper, we discuss the concept of Bayesian's method and using likelihood functions to analyze the causality between two and three variables. We hope to determine the direction between variables.

誌 謝

兩年的研究生生涯將進入尾聲，即將面臨人生另一個旅程。回想起新竹的六年生活，很慶幸可以進入統計所這個大家庭，有著認真和藹的老師、親切助人的學長姐、互助互信同學及熱心可親的所辦助理，使這兩年生活充滿溫馨愉悅的氣氛。

最先要感謝我的指導教授—洪慧念老師，謝謝老師這兩年來的支持鼓勵與信任，一直很有耐心及細心的啟發教導，一步步地指引論文的研究，以期能更深入、嚴謹地探討；老師也教我很多生活中應有的態度及理念，給我很多的幫助，希望能不斷成長。感謝王維菁老師、謝淑蓉老師和黃信誠老師的細心教導與建議，使這篇論文能更完整。

最後要感謝我的父母：戴建福先生及吳寶琳女士，從小到大的栽培及支持；使我能專心的求學及生活。還要感謝育仕、淑儀、秀仁、揚波、雅靜及秀雯，不僅使我遠離寂寞與低潮，切磋討論時亦使我獲益良多。最後將這篇文章和大家分享，由衷感謝曾經幫助過我的人。

戴如美 謹誌于

國立交通大學統計學研究所

中華民國九十四年六月

目 錄

中文摘要.....	i
英文摘要.....	ii
中文誌謝.....	iii
目錄.....	iv
第一章 緒論	1
第二章 文獻回顧	4
2-1 貝氏網路法及其優缺點	4
2-2 文獻中提及的方法	13
2-2.1 PC Algorithm.....	13
2-2.2 Score Function	15
2-2.3 Ideal Interventions	16
2-2.4 Indicator Function	17
第三章 判斷變數間之方向性	19
3-1 檢定獨立的方法—Kendall's Tau	19
3-2 貝氏概念與概似函數.....	23
3-3 使用貝氏概念及概似函數判斷兩個變數間之方向性	24
3-4 考慮 (X,Y) 間不同的機率模型	26

3-4.1	(X, Y) 間機率模型為Normal分配	27
3-4.2	模擬五種不同分配，使用之公式資料服從Normal分配 ...	28
3-4.3	(X, Y) 間機率模型為Exponential分配	30
3-4.4	模擬四種不同分配，使用之公式資料服從Exponential.	31
3-5.1	考慮 (X, Y) 間的機率模型為線性關係之Normal	33
3-5.2	考慮 (X, Y) 間的機率模型為二次關係之Normal	35
3-5.3	模擬五種分配資料，以二次關係之Normal公式做判斷 ...	36
3-6	直接使用概似函數判斷三個變數間之方向性	41
3-6.1	考慮 (X, Y, Z) 間的機率模型為線性關係之Normal	44
3-6.2	模擬五種分配資料，以線性關係之Normal公式做判斷 ...	50
3-6.3	考慮 (X, Y, Z) 間的機率模型為二次關係之Normal	51
3-6.4	模擬五種分配資料，以二次關係之Normal公式做判斷 ...	56
第四章 結論		58
參考文獻		59
附 錄		62

第一章 緒論

在資訊發達及人類對於生物科技如此研發的現今，我們取得資料或生物資訊比以往容易了些，無論是所需的時間或是經費也都比以往經濟了許多；在這樣大量的資料中，應如何分析才能充份使用到每個觀測到的資訊且不遺漏任何資訊，並將資料充分應用到分析上，且使用最有效的方法得到欲解決問題之答案，正是目前統計學家都會面臨到的問題。

而在分析資料的過程中，通常會先討論變數與變數彼此之間的相關性，更進一步會討論變數與變數之間是否有所謂的因果關係，也有學者將「是否有因果關係」闡述為「是否有像親代與子代之間的關係」？

究竟因果關係是什麼，底下舉生活上的例子來說明：身高與體重之間是否有因果關係？一般而言，若把身高與體重視為兩個變數，經過簡單的統計分析，結果會顯示出這兩個變數是有相關性的；而這兩個變數彼此之間有沒有因果關係呢？不失一般性可以說，從身高這個變數的角度來看，身高增加時體重自然而然會隨著身高的增加而增加，而若從體重這個變數的角度來看，加重體重時身高並不一定會因體重的增加而增加。

底下我們分成兩種情況來說明：(i) 如果有兩個人，其中一

人身高較高，則此人體重也會傾向比另一人還重；(ii) 如果是同一人，我們想辦法將其身高變高，則此人在身高變高的同時，體重也應會變重；反過來，如果我們設法將此人體重變重，此時身高同時變高的可能性並不會很大。由上述，在(i) 中我們只能看出身高與體重有關係；在(ii) 中可發現某一變數受另一變數的影響，我們說這兩變數具有因果關係。在此例中，身高可稱之為體重的「因」或「親代」，另一方面，體重可稱之為身高的「果」或「子代」。從上述我們可以發現，如果每人只有一筆資料，我們只能看出關係，而無法看出因果關係。

若以圖形的角度來看，要如何表示變數之間的因果關係呢？底下先將上述身高體重的例子用變數命名之以方便用圖形說明：變數A表示身高，變數B表示體重，變數C代表一隱藏變數 (latent variable)；藉由這三個變數可以表示出 12 種以上的因果關係[†](註一)，而在上例的因果關係則可能有下列幾種情況：

(i) $A \rightarrow B$ ：身高對於體重是直接影響的關係，且是身高影響體重而不是體重影響身高；

(ii) $A \rightarrow C \rightarrow B$ ：身高對於體重是間接影響的關係，亦即身高

[†]註一：三個變數的圖形結構可以有底下 12 種可能：

$A \rightarrow B \rightarrow C$ 、 $A \rightarrow C \rightarrow B$ 、 $B \rightarrow A \rightarrow C$ 、 $B \rightarrow C \rightarrow A$ 、 $C \rightarrow A \rightarrow B$ 、 $C \rightarrow B \rightarrow A$
 $A \rightarrow B \leftarrow C$ 、 $A \rightarrow C \leftarrow B$ 、 $B \rightarrow A \leftarrow C$
 $B \leftarrow A \rightarrow C$ 、 $A \leftarrow B \rightarrow C$ 、 $A \leftarrow C \rightarrow B$

和體重這兩個變數之間隱藏了某些變數，而身高這個變數影響了某一隱藏變數，且這個隱藏變數直接影響了體重這個變數；也因此我們才能從資料發現身高這個變數對於體重這個變數是有影響的；但事實上，從資料中我們未必可以發現是否有一個隱藏變數存在。

在日常生活中或是生物基因方面，當分析出變數與變數之間有相關性時，並不代表變數之間有著因果關係，是需再更進一步的分析才能確定是否有所謂的因果關係。在不知道哪一個變數才是「因」或「親代」的情況時，要如何做才能得知變數之間正確的因果關係是當前很多學者想探討的問題；而本文主要的目標是希望將變數間有因果關係的資料，分析出其正確的因果關係及判斷出圖形的方向性。

第二章 文獻回顧

本文主要的目標是希望分析出資料的因果關係及判斷出圖形的方向性，在文獻上有很多統計學家使用各種不同的方法來描述基因網路模型，如：線性模型 (D'haeseleer et al, 1999)、非線性模型 (Weaver et al, 1999)、布爾數學邏輯網路模型 (Boolean Networks) (Kauffman 1993, Somogyi and Sniegoski, 1996) 等；但使用這些方法需要在底下兩個限制條件下才適用：

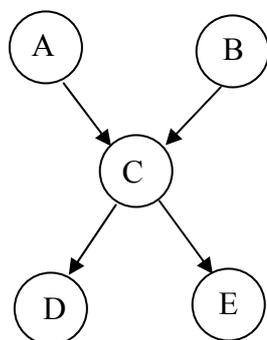
- (i) 每個變數都是可觀測的，亦即不能有所謂的隱藏變數；
- (ii) 所有在每個變數之間的關係需是決定性的 (deterministic)。

因此 Murphy and Mian (1998) 和 Friedman (1999) 等人建議使用貝氏網路模型 (Bayesian Network) 來設法改善其不足之處，貝氏網路在處理因果關係的領域上是很普遍卻也很重要的一個方法。

2-1 貝氏網路法及其優缺點

什麼是貝氏網路法呢？Heckerman 在 1995 年指出貝氏網路是一個根據先驗資料 (prior) 所架構出圖形模式 (graphical model) 的方法；簡單來說就是使用圖形結構分析隨機變數之間的因果關係。底下我們分成兩個部分來說明，一開始我們先說

明是使用何種圖形結構：貝氏網路法是使用一個有方向性且沒有循環的圖形結構 (Directed Acyclic Graph, 簡稱 DAG) (Peral and Verma, 1991), 由一些頂點 (node) 和連接頂點的邊 (edge) 所組成, 圖形中的每一個頂點都代表一個隨機變數, 而每個位於點和點之間的邊再加上有方向性的單箭號 (directed edge), 則代表這兩個相鄰的點之間的因果關係, 若是 $A \rightarrow B$ 則表示 A 直接影響 B; 以圖一來說明, 圖中共有 5 個頂點、4 條有向邊 (directed edge), 也就是有 A、B、C、D、E 五個隨機變數, 及 \overrightarrow{AC} 、 \overrightarrow{BC} 、 \overrightarrow{CD} 、 \overrightarrow{CE} 四條有向邊, 從圖一我們可以看出 A 和 B 是 C 的親代, C 是 D 和 E 的親代, 亦即 C 是 A 和 B 的子代, D 和 E 是 C 的子代; 在因果關係中, 有時將親代稱為因、子代稱為果。從圖一也可以看出, A 和 C、A 和 D、A 和 E、B 和 C、B 和 D、B 和 E、C 和 D、C 和 E 及 D 和 E 皆是相關的。



圖一：五個變數之 Bayesian Network 圖形

接著要繼續探討隨機變數間的因果關係，我們以條件機率的觀點做為出發點，想從條件機率的角度去看整個模型的機率密度函數。在列出整個模型的機率密度函數之前，我們先描述在條件機率方面，若要探討因果關係之前需先討論的兩種關係：

(i) 條件獨立：意指給定 X 、 Y 、 Z 三個隨機變數當中的隨機變數 Z 時，若此時隨機變數 Y 所得之訊息對於隨機變數 X 並沒有任何額外的訊息時，我們說在給定隨機變數 Z ，隨機變數 X 和 Y 會條件獨立；以圖一的關係來看，給定 C 時， D 和 E 會條件獨立，另外給定 C 時， A 和 D 及 B 和 E 亦會條件獨立。

(ii) 邊際相關：意指無論是在 causal chains $i \rightarrow m \rightarrow j$ 或 causal forks $i \leftarrow m \rightarrow j$ 兩種情況下，只要變數 i 和變數 j 相關，而若給定變數 m 時，變數 i 和變數 j 會條件獨立時，此時我們稱變數 i 及變數 j 是邊際相關；以圖一中的關係來看，將變數 C 視為變數 m 、變數 D 、 E 分別視為變數 i 、 j ，由於 D 和 E 是相關的，而若給定 C 時， D 和 E 會條件獨立，因此可以發現 D 和 E 是邊際相關，同理亦可以得知 A 和 D 、 B 和 E 是邊際相關。

值得一提的是：在一有方向性且沒有循環的圖形結構 G

中，對於任何隨機變數若給定其親代，則此隨機變數會與親代的親代還有親代的其它子代有條件獨立的因果關係。在這樣的 DAG 圖形結構中，還有一個值得留意的性質：沒有子代的變數或有子代但沒有給定子代的親代，彼此之間是獨立的；有了子代，且給定子代後的親代是相關的。以生活上的方式來說，沒有子孫的配偶或有子孫但沒有給定子孫的配偶，彼此之間是獨立的；有了子孫，且給定子孫後的配偶就是相關的了。

而貝氏網路究竟有什麼優點呢？Spirtes 等人在 2000 年指出下列五項：

- 
- (i) 能夠很明確的把 DAG 圖形結構中的因果關係與統計上的假設相關聯起來； <they explicitly relate the directed acyclic graph model of the causal relations among the gene expression levels to a statistical hypothesis.>
 - (ii) 包含了先前所提及的模型，如：線性模型、非線性模型、布爾數學邏輯網路模型（Boolean Networks）等，而且把 Hidden Markov Models 視為其特例； <the include all of the aforementioned models, and Hidden Markov Models, as special cases.>
 - (iii) 在觀測到的資料中，已開發了很多演算法是適合貝氏網路的； <there are already well developed algorithms for

searching for Bayesian networks from observational data >

- (iv) 允許隨機的元素及隱藏變數的存在； < they allow for the introduction of a stochastic element and hidden variables >
- (v) 允許資料被搜集的過程有明確的模型； < they allow explicit modeling of the process by which the data are gathered. >

在介紹完貝氏網路的優點後，我們回過頭來談機率模型的部分，接著描述出其機率密度函數；在貝氏網路的結構中我們仍舊以條件機率的角度為出發點，由於事件有前因後果，因此會先有親代才有子代，所以我們利用 Pa_i^G 來表示在一有方向性且沒有循環的圖形結構 G 中隨機變數 X_i 的親代，接著以條件機率分配函數 $P(X_i | Pa_i^G)$ 表示隨機變數 X_i 在給定親代 Pa_i^G 後的條件機率分配，再利用隨機變數之間的條件獨立關係及機率論中的一些性質，就可以把所有欲考慮的隨機變數其機率密度函數分解 (factorized) 成數個給定親代 Pa_i^G 後的條件機率分配函數之乘積；以一般式來表示的話，隨機變數 (X_1, X_2, \dots, X_n) 的聯合機率密度函數可寫成底下的形式：

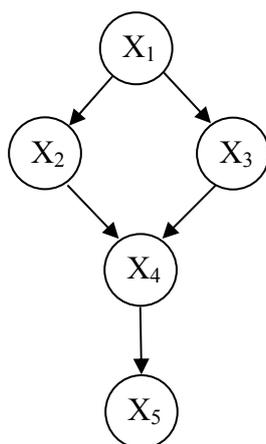
$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa_i^G)$$

以圖二為例，在貝氏網路結構中，隨機變數 $(X_1, X_2, X_3, X_4, X_5)$ 之聯合機率密度函數若利用機率論中的性質，可以寫成底下這種形式：

$$\begin{aligned} & P(X_1, X_2, X_3, X_4, X_5) \\ &= P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)P(X_4|X_1, X_2, X_3)P(X_5|X_1, X_2, X_3, X_4) \end{aligned}$$

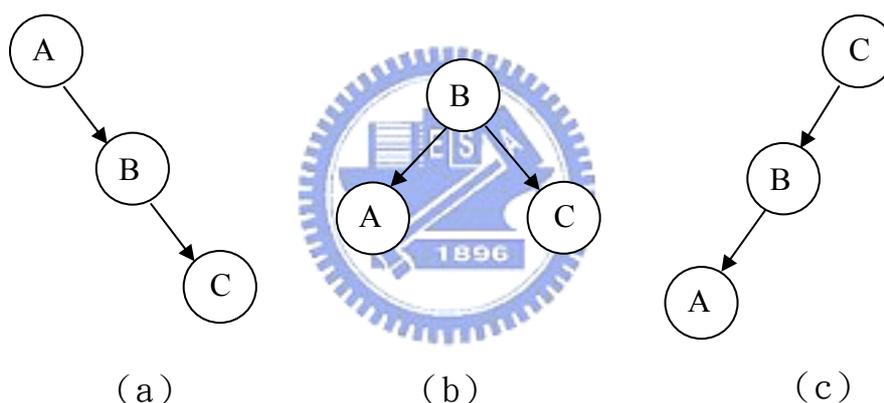
而從圖二中可以發現，在給定 X_1 時， X_2 和 X_3 會條件獨立，因此 $P(X_3|X_1, X_2) = P(X_3|X_1)$ ；同理給定 X_2 或 X_3 時， X_1 會和 X_4 條件獨立，即 $P(X_4|X_1, X_2, X_3) = P(X_4|X_2, X_3)$ ；而 $P(X_5|X_1, X_2, X_3, X_4) = P(X_5|X_4)$ 也是相同的道理；由上面的式子，我們可以將聯合機率密度函數簡化成底下的形式，使得在計算時可以簡單很多：

$$P(X_1, X_2, X_3, X_4, X_5) = P(X_1)P(X_2|X_1)P(X_3|X_1)P(X_4|X_2, X_3)P(X_5|X_4)$$



圖二：五個相關變數之 Bayesian Network 圖形

經由圖形之間的因果關係及機率論的性質，我們對於隨機變數 (X_1, X_2, \dots, X_n) 可以寫出一個唯一的分配 (distribution)；但有時，我們所寫出來的聯合機率密度函數和圖形結構之間並不是一對一 (one-to-one) 的對應，也就是說，雖然一個圖形結構只能簡化出一個唯一的分配，但有可能會有一個以上的圖形結構對應到同一個分配函數；比如下面三個圖形結構，在 DAG 圖形結構中所寫出來的聯合機率密度函數會是相同的。



圖三：三個密度函數相同的 DAG 圖形結構

由於上面三個 DAG 圖形結構有相同的「骨架」，亦即在不考慮方向性時，其圖形結構相同皆為 $A-B-C$ ，且變數間的因果關係有相同的獨立關係，此時我們說這兩個圖形是等價的 (Equivalent)。這三個等價的圖形結構之因果關係皆為如下： A 和 B 相關、 B 和 C 相關及給定 B ， A 和 C 會是條件獨立；也因此我們無法直接從資料中得知，上述三個圖形結構中的哪個圖

才是資料的真正圖形結構。從另一角度來看，上述三個圖形的模型分別可表示成如下：

$$\text{Model (a): } P(A), P(B|A), \text{ and } P(C|B)$$

$$\text{Model (b): } P(B), P(A|B), \text{ and } P(C|B)$$

$$\text{Model (c): } P(C), P(B|C), \text{ and } P(A|B)$$

經由簡單的計算可以驗證上述三個模型之密度函數皆等同於 $P(AB)$ 。

所以在貝氏網路法中，要分析隨機變數之間的因果關係時，需先將資料中所有因果關係之任何有可能的 DAG 圖形結構都畫出來，接著考慮前人的經驗累積、各個專家的知識經驗及任何可以提供我們分析圖形結構的資訊，也就是根據所謂的先驗分配以及資料，經過程式運算，再看哪一個圖形結構所得的隨機變數之聯合機率密度函數值的後驗分配會是最大的，經由這個方法使得有可能的 DAG 圖形結構變少，也使得我們有更大的機會選到正確的圖形結構。

由於在等價的圖形中我們並無法確定何者才是正確的圖形結構，但我們一定可以將所有等價的圖形結構表示成一個唯一的「部分有方向性且無循環」的圖形結構（Partially Directed Acyclic Graph，簡稱 PDAG），在 PDAG 的圖形結構中， $A \rightarrow B$

表示所有等價的圖形都包含 $A \rightarrow B$ 這個因果關係； $A - B$ 則表示在等價的圖形中， A 和 B 之間的因果關係有的為 $A \rightarrow B$ ，有的則為 $A \leftarrow B$ ，而我們還無法確認出方向性，因此記為 $A - B$ 。在 PDAG 的圖形結構中，有時我們可以確認出當中某些隨機變數之間的關係，但並不是全然都可以的。

事實上真正的因果圖形結構是無法完全確認出的，由於我們無法確認實際上會有多少個隨機變數，因此在我們設法畫出可能的圖形結構時，可能會有更多不同的圖形結構產生，而在每個可能的圖形結構中，還可能會有各種不同的隱藏變數存在，如此一來，會組合成更多不同的圖形結構，而不同的圖形結構會導致不同的因果關係，不同的因果關係會形成不同的聯合機率密度函數，再加上 DAG 圖形結構的可能組合是與我們所觀測到的隨機變數個數之間呈現指數倍關係，層層累積下來，可能的組合個數會大到根本無法繼續做因果關係的分析；所以我們在做因果關係的分析時，為了能找到適當的 DAG 圖形結構，必需做適當的假設，在一些條件限制下，才能找出適當的 DAG 圖形結構，底下列出較常見的假設：

(i) Causally Sufficient Assumption：在所有可觀測的變數

集合中，任何兩變數都不會被集合以外的變數所影響。

- (ii) Causal Markov Assumption：在給定親代後，任何變數都會與其非子代（non-descendants）之變數條件獨立。
- (iii) Causal Faithfulness Assumption：當中所有的條件獨立關係是繼承 Causal Markov Assumption；換個說法：沒有多餘的獨立關係了，因為所包含的獨立關係已經足夠了。

2-2 文獻中提及的方法

在上面的三個假設條件限制下，再加上前人的經驗累積、專家的知識經驗等等，可以使所有有可能的圖形結構之範圍縮小。接著介紹一些文獻中提及的用來分析貝氏網路的方法。

2-2.1 PC Algorithm

將我們欲所觀測的資料中，隨機變數之間的因果關係，利用 PC Algorithm（Verma and Pearl, 1990）來分析變數之間的條件獨立關係，將所得的關係依底下三個步驟來建構與資料最接近的 DAG 圖形結構：

Step1：在變數集合 V 中，對於任何一對變數 A 和變數 B ，尋找是否存在集合 S_{AB} ，使得給定集合 S_{AB} ，變數 A 和變數 B 會條件獨立；當不存在這樣的集合 S_{AB} 時，我們在變數 A

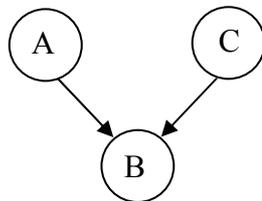
和變數 B 之間畫一條沒有方向性的邊，且這兩事件之間是為若且唯若的關係 (if and only if)。

Step2：對於每對不相鄰的變數 A 和變數 B，一一確認兩變數之間所有的共同相鄰變數 (common neighbor) 是否屬於集合 S_{AB} 。若屬於集合 S_{AB} 時則繼續確認的動作；反之，若兩變數之共同相鄰變數 C 不屬於集合 S_{AB} 時，則將兩箭號指向變數 C，亦即 $A \rightarrow C \leftarrow B$ 。

Step3：在我們設法確認出變數之間的因果關係方向性時，需遵守底下兩個條件：

(i) 在此步驟我們畫新的箭號時不能增加新的 v -結構；

我們簡單利用圖形 $A \rightarrow B \leftarrow C$ 解釋何謂 v -結構。若將圖形畫成底下圖四的格式時，可發現圖形呈現一個 v 字型，且兩變數皆指向另一變數；而因果關係為：A 和 C 在沒有給定 B 之前是獨立的，在給定 B 之後則為相關。



圖四： v -結構 (v -structure)

(ii) 在此步驟我們畫新的箭號時不能產生循環結構；

圖形 $A \rightarrow B \rightarrow C \rightarrow A$ 是為一 directed cycle，而圖形 $A \rightarrow A$ 則否，我們稱之為 self-loops，。

步驟 3 似乎有些抽象，Meek 在 1995 年提出只要依照底下四個規則去做就足夠滿足步驟 3：

R_1 ：當 $A \rightarrow B$ 存在且 A 和 C 不相鄰時，將 $B - C$ 畫成 $B \rightarrow C$ ；

R_2 ：當 $A \rightarrow C \rightarrow B$ 存在時，將 $A - B$ 畫成 $A \rightarrow B$ ；

R_3 ：當 $A - C \rightarrow B$ 和 $A - D \rightarrow B$ 皆存在且 C 和 D 不相鄰時，將 $A - B$ 畫成 $A \rightarrow B$ ；

R_4 ：當 $A - C \rightarrow D$ 和 $C \rightarrow D \rightarrow B$ 皆存在， C 和 B 不相鄰，且 A 和 D 相鄰時，將 $A - B$ 畫成 $A \rightarrow B$ 。

前面有說明過，即使同樣的條件獨立關係下也有可能建構出不唯一的圖形結構，需再利用其它資訊，如：事件發生的先後順序、專家的專業知識等等資訊，才能使我們再次縮減範圍，而儘可能得到唯一的一個 DAG 圖形結構。

2-2.2 Score Function

Heckerman 等人在 1995 年提出一個廣為流傳的貝氏方法，在給定資料 D 之下，在每個圖形 G 中利用所得的後驗分配求出一個分數函數 (Score function) $S(G:D)$ ，而我們所定義的這個

分數為後驗機率函數的對數值，會和 $P(G|D)$ 成正比，因此當後驗機率函數值愈高，也就是在給定的資料 D 下，此圖形 G 的可能性愈高時， $S(G:D)$ 的分數亦會愈高：

$$S(G:D) = \log P(G|D) = \log P(D|G) + \log P(G) + c$$

式子中的 c 是與圖形 G 無關的一個常數，而

$$P(D|G) = \int_{\Theta} P(D|G, \theta) P(\theta|G) d\theta,$$

由此可發現，圖形 G 之先驗機率函數 $P(G)$ 與 $P(\theta|G)$ 的選取是會直接影響圖形 G 的分數函數 $S(G:D)$ 。

取得分數後，更進一步使用靴環法（Bootstrap）來瞭解基因之間交互作用的潛在模型。接著使用此結果來推論統計信心指數（statistically confident feature），並使用這個準則來分析相關因果影響的分析。

2-2.3 Ideal Interventions

前面提及了圖形等價及分數函數，接著我們可以設法將方法整合使得能找出更多因果關係的方向性；此法我們稱之為 Ideal Interventions (Pearl, 2001)。在等價的圖形中，我們設法固定當中某一個隨機變數的值，如此一來會使原先等價圖形的因果關係產生變化，我們試著從這些變化確認出更多的方向性，

漸漸地會有愈來愈多因果關係的方向性被確定出來，使 PDAG 中不確定的關係可以逐漸被確認出來。底下利用一個簡單的例子實際說明這個方法，圖形 $G_1 : A \rightarrow B \rightarrow C$ 及圖形 $G_2 : A \leftarrow B \rightarrow C$ 兩個等價的 DAG 圖形結構，有著相同的因果關係及獨立關係，若我們控制變數 B 使之為一固定值，此時圖形在 G_1 中，變數 A 無法影響變數 B 了，而在圖形 G_2 中，變數 B 仍舊保留影響變數 A 的關係；則兩個原本等價的 DAG 圖形結構將不再是等價關係，此時再使用分數函數做比較，選擇分數比較高的那個圖形結構。如此一來，PDAG 圖形結構中，因果關係的方向性又多確定出一個了，使有可能的圖形結構之範圍縮小。



2-2.4 Indicator Function

把焦點回到條件獨立，如果要檢定在給定變數 Z 之下，變數 X 和變數 Y 是否條件獨立，我們必需檢驗在已知變數 Z 的情況下， X 和 Y 之間並沒有路徑可以傳遞訊息。我們利用一個指標函數 (Indicator Function) 來轉化這些抽象的訊息，將圖形結構的可能性稱為圖形結構的一個特徵 (feature)，如：圖形結構中存在著 $X \rightarrow Y$ 的因果關係、或給定變數 Z ，則變數 X 和變數 Y 會條件獨立...等特徵；指標函數設為如下：

$$f(G) = \begin{cases} 1, & \text{when } G \text{ satisfies the feature} \\ 0, & \text{otherwise} \end{cases}$$

而可反應我們對某一特徵的信賴程度，亦即後驗機率函數則可表示為如下：

$$P(f(G)|D) = \sum_G f(G)P(G|D)$$

若我們完全沒篩選，而將所有 score 高的圖形結構列舉出來會很沒有效率，因為圖形結構的可能組合與隨機變數的個數是呈現指數倍成長的，所以我們必需把有代表性的圖形結構篩選出來，並把我們有興趣的特徵列出。Friedman 等人在 1999 年提出了無母數的靴環法 (Non-Parametric Bootstrap)，Friedman and Koller 在 2001 年又提出了一個更精確但花費也更多的 MCMC simulation。雖然 Friedman 等人在 2000 年利用靴環法模擬出來的資料和我們所觀測到的模型是吻合的，但也提出此法的錯誤比例是很高的；因為即使一個有高信賴程度的特徵沒有被偵測出來，並不代表這個特徵不存在於我們的圖形結構，而是因我們觀測到的這些資料沒有辦法強力支持這個特徵的存在。

第三章 判斷變數間之方向性

統計學家及生物學家在分析因果關係有使用各種不同的方法，前面提及的大多是以貝氏為主的方法；貝氏方法一開始先主觀地相信參數 θ 是來自於某一分配，再客觀地根據搜集的資料對於模型做修正；亦即需給定先驗分配 (prior distribution)，接著再根據資料修正參數 θ 的模型，而求得後驗分配 (posterior distribution)。由於先驗資料是很主觀的，所以我們的方法是設法著重在概似函數 (Likelihood function) 上面，而概似函數需要知道資料的模型才能做，因此在底下的方法中會探討若模型假設錯誤會造成什麼影響，並以數個不同的模型去做模擬，而誤差函數 (Error function) 亦會以數個不同的型式做模擬。

在分析因果關係的方向性之前，我們仍需先判斷資料中有哪些變數彼此之間是獨立的，或具有條件獨立的關係，在分析之前，先介紹在本文中所使用的檢定獨立方法。

3-1 檢定獨立的方法－Kendall's Tau

在文獻中有關檢定獨立的方法不只 Kendall's Tau 法，還有其它方法，如：Spearman Rank-Order Correlation Coefficient, Chi-Square Test, Pearson Product Moment Correlation (簡稱

Pearson's r), ...等，但由於 Kendall's Tau 法能檢定出的範圍比較廣，不像有些方法只能檢定出是否為“線性”相關；再加上其檢定之結果不會受到資料轉換的影響，尤其是在單調 (monotone) 的資料下；更有學者提及：當資料不是來自鐘形分配且樣本個數小於 20 的情況下，此法是最適合的，底下網頁有更進一步的說明。

<http://canis.tamu.edu/wfscCourses/WFSC620/exercises/ExerciG.htm>

基於上述之原因，我們採用 Kendall's Tau 法，底下對此法稍做說明；另外在附錄中，列出了一些不同的模型種類，其樣本數和型一誤差的關係表：



Kendall's Tau 的統計量依樣本數不同會有底下兩種表示法：

(i) $N < 40$ ：此時統計量可表示為 $\tau = \frac{C-D}{\binom{N}{2}}$ ；在觀測的資料中，對於所有數對 (x_i, y_i) 和 (x_j, y_j) ，將 $(x_i - x_j)(y_i - y_j) > 0$ 的個數記為 C ，而 $(x_i - x_j)(y_i - y_j) < 0$ 的個數記為 D ；在 $(x_i - x_j)(y_i - y_j) = 0$ 的部分，將數對中 $x_i = x_j$ 的個數記為 N_x ，而 $y_i = y_j$ 的個數記為 N_y ，當 $(x_i - x_j)(y_i - y_j) = 0$ 的個數很大時，Kendall's Tau 的統計量則需修正成如下：

$$\tau = \frac{C-D}{\sqrt{\left[\binom{N}{2} - N_x\right] \times \left[\binom{N}{2} - N_y\right]}}$$

而在文獻 Encyclopedia of Statistical Sciences Volume IV (Samuel Kotz & Normal L. Johnson) 中有提及當樣本個數 $N > 10$ 時，Kendall's Tau 即會服從常態分配，在程式的模擬驗證中，確實發現在樣本數大於 10 時是會服從常態分配。

(ii) $N \geq 40$ ：此時需再對所求得之統計量做如下的轉換，

且統計量： $t = \frac{\tau}{\sqrt{\frac{2 \times (2 \times N + 5)}{9 \times N \times (N - 1)}}}$ 仍舊服從常態分配。

我們利用 Kendall's Tau 來檢驗因果關係圖形結構中，隨機變數是否獨立，及是否有條件獨立的關係存在。

變數之間是否獨立只需利用上述提及的方法，將欲檢測的兩個變數作分析而得一統計量，再利用常態分配加以檢定即可得知；而變數之間是否有條件獨立的關係則需再做進一步分析。

底下欲探討的問題是：給定變數 Z ，變數 X 和變數 Y 之間是否有條件獨立的關係？先前我們在判斷變數間是否有獨立的關係，皆是看“兩個”變數之間是否獨立；若要看條件獨立的關係，我們可以使用這個觀點做為出發點，將三個變數之間的條件獨立關係設法轉化成只需藉由看兩個變數的獨立關係即可得知。將變數 X 和變數 Y 對於變數 Z 做分組，再分別去看在變數 Z 的每組值中 ($Z = c_j$) 變數 X 和變數 Y 是否獨立；其中 j 的範圍會依樣本個數及分組個數而改變，在變數 Z 的每一組值

中，變數 X 和變數 Y 的個數皆不能太小（一般會使之大於 5），以免影響分析。我們使用 Rank 離散化變數 Z 的方式將變數 X 和變數 Y 對於變數 Z 做分組，依序列為底下數個步驟：

Step 0：原始數對為 $(Z_i, X_i, Y_i) \quad i=1, \dots, n$ 。

Step 1：將資料 Z_1, \dots, Z_n 取 Rank，稱為 R_1, \dots, R_n 。

Step 2：使用 $\text{Int}\left[\frac{R_i}{n/k}\right]$ 將原始的連續資料離散化成 k 組，記為

$R_1^1, R_2^1, \dots, R_{i_1}^1, R_1^2, R_2^2, \dots, R_{i_2}^2, \dots, R_1^k, \dots, R_{i_k}^k$ ， R_i^j 之上標下標分別

表示第 j 組中第 i 個，而每組中分別有 i_j 個， $j=1, \dots, k$ 。

Step 3：檢查每個 i_j 是否都夠大足以分析。

Step 4：將每組之 R_i^j 值原對應的數對 $(X_i, Y_i)^j$ 分別收集起來。

Step 5：對於每組之 $(X_i, Y_i)^j$ 分別做 Kendall's Tau，來檢驗當

$Z = j$ 時之 $(X_i, Y_i)^j$ 是否獨立。

若當每組之變數 X 和變數 Y 皆獨立時，則我們認定在給定變數 Z 的條件下，變數 X 和變數 Y 有條件獨立的關係；反之，只要有一組變數 Z 的值會使變數 X 和變數 Y 相關，則我們認為在給定變數 Z 的條件下，變數 X 和變數 Y 不符合條件獨立的關係。而更進一步的判斷方式及更詳細的分析請見林育仕(2005)。

3-2 貝氏概念與概似函數

若以數學式來表示貝氏的概念，一般會將先驗分配記為 $\pi(\theta)$ ，而資料的分配記為 $f(x|\theta)$ ，則後驗分配可寫成如下式子：

$$\pi(\theta|x) = f(x|\theta)\pi(\theta)/m(x) \quad (3-2-1)$$

當中的 $m(x)$ 是資料的邊際分配： $m(x) = \int f(x|\theta)\pi(\theta)d\theta$

底下的分析接著上述再進一步的假設參數 θ 裡仍存在一個參數 λ ，亦即可寫成如下的型式：資料 X 服從函數 f 分配， $X \sim f(x|\theta)$ ，參數 θ 的先驗分配為 $\theta \sim \pi(\theta|\lambda)$ ，則 (X, θ) 的聯合分配為 $(X, \theta) \sim f(x|\theta)\pi(\theta|\lambda)$ ，接著對參數 θ 做積分可得資料 X 的邊際分配 $X \sim h(x|\lambda)$ ：

$$h(x|\lambda) = \int f(x|\theta)\pi(\theta|\lambda)d\theta \quad (3-2-2)$$

由於上式中只有參數 λ 是未知，因此利用已知的資料將函數最大化，以求得參數 λ 的 MLE (Maximum Likelihood Estimator)：

$$\max_{\lambda} h(x|\lambda) \Rightarrow \hat{\lambda}$$

此時我們的資料仍舊服從函數 f 分配， $X \sim f(x|\theta)$ ，而參數 θ 的先驗分配則可寫成 $\theta \sim \pi(\theta|\hat{\lambda})$ ，當中的 $\hat{\lambda}$ 是函數 h 之 MLE，所以藉由上述兩個式子 (3-2-1) 及 (3-2-2) 可以得知參數 θ 的後驗分配是和 $f(x|\theta)\pi(\theta|\hat{\lambda})$ 成正比。

3-3 使用貝氏概念及概似函數判斷兩個變數間之方向性

實際在計算的過程，會發現函數 h ，即式子 (3-2-2) 當中的積分並不是那麼容易求得，因此我們藉由參數 θ 的後驗分配會與函數 $f(x|\theta)\pi(\theta|\hat{\lambda})$ 成正比的關係為出發點，再利用分配的性質進而求得 $\hat{\lambda}$ 。一開始先從只有兩個變數的圖形結構開始，欲分析變數 X 和變數 Y 之間的因果關係方向性究竟為 $X \rightarrow Y$ 或 $X \leftarrow Y$ ？

Step 1：求出代表圖形結構 $X \rightarrow Y$ 的概似函數。

首先假設資料服從 $X \sim f_X(x|\theta_X)$ ， $Y|X \sim f_Y(y|x, \theta_Y)$ ，且參數的先驗分配中仍存在一個參數 λ ， $\theta_X \sim \pi(\theta_X|\lambda_X)$ ， $\theta_Y \sim \pi(\theta_Y|\lambda_Y)$ ，由以上假設我們可以得到底下的式子：

$$f(x, y|\theta_X, \theta_Y) = f_X(x|\theta_X)f_Y(y|x, \theta_Y) \quad (3-3-1)$$

$$\pi(\theta_X, \theta_Y|\lambda_X, \lambda_Y) = \pi(\theta_X|\lambda_X)\pi(\theta_Y|\lambda_Y) \quad (3-3-2)$$

從上小節我們知道參數 θ 的後驗分配會與函數 $f(x|\theta)\pi(\theta|\hat{\lambda})$ 成正比，因此得到： $\pi(\theta_X, \theta_Y|x, y) \propto f(x, y|\theta_X, \theta_Y)\pi(\theta_X, \theta_Y|\hat{\lambda}_X, \hat{\lambda}_Y)$ ，在這樣一個比例的式子下，只需再利用任何分配的機率密度函數之積分值恆為 1 的性質，即可把參數 (θ_X, θ_Y) 的後驗分配寫出來；接著使用上述的式子將資料 X 、 Y 的概似函數寫成如下：

$$f_1(x, y|\lambda_X, \lambda_Y) = \frac{f(x, y|\theta_X, \theta_Y)\pi(\theta_X, \theta_Y|\lambda_X, \lambda_Y)}{\pi(\theta_X, \theta_Y|x, y)} \quad (3-3-3)$$

使用上式 (3-3-3) 求出參數 (λ_X, λ_Y) 的 MLE: $(\hat{\lambda}_X, \hat{\lambda}_Y)$ ，再將 $(\hat{\lambda}_X, \hat{\lambda}_Y)$ 代回上式而求得圖形結構 $X \rightarrow Y$ 之概似函數 $f_1(\underline{x}, \underline{y} | \hat{\lambda}_X, \hat{\lambda}_Y)$ 。

Step 2: 求出代表圖形結構 $Y \rightarrow X$ 的概似函數。

如同 Step 1 的方法，假設資料的分配服從 $\underline{Y} \sim f_Y(\underline{y} | \theta_Y)$ ，而 $\underline{X} | \underline{Y} \sim f_X(\underline{x} | \underline{y}, \theta_X)$ ，且參數的先驗分配中仍存在一個參數 λ ，其分配為 $\theta_Y \sim \pi(\theta_Y | \lambda_Y)$ ， $\theta_X \sim \pi(\theta_X | \lambda_X)$ ，由以上假設我們可以得到底下的式子：

$$f(\underline{y}, \underline{x} | \theta_Y, \theta_X) = f_Y(\underline{y} | \theta_Y) f_X(\underline{x} | \underline{y}, \theta_X) \quad (3-3-4)$$

$$\pi(\theta_Y, \theta_X | \lambda_Y, \lambda_X) = \pi(\theta_Y | \lambda_Y) \pi(\theta_X | \lambda_X) \quad (3-3-5)$$

同理我們可以得到： $\pi(\theta_Y, \theta_X | \underline{y}, \underline{x}) \propto f(\underline{y}, \underline{x} | \theta_Y, \theta_X) \pi(\theta_Y, \theta_X | \hat{\lambda}_Y, \hat{\lambda}_X)$ ，再利用分配的機率密度函數之積分值恆為 1 的性質，可寫出參數 (θ_Y, θ_X) 的後驗分配；再將資料 X 、 Y 的概似函數寫成如下：

$$f_2(\underline{y}, \underline{x} | \lambda_Y, \lambda_X) = \frac{f(\underline{y}, \underline{x} | \theta_Y, \theta_X) \pi(\theta_Y, \theta_X | \lambda_Y, \lambda_X)}{\pi(\theta_Y, \theta_X | \underline{y}, \underline{x})} \quad (3-3-6)$$

使用上式 (3-3-6) 求出參數 (λ_X, λ_Y) 的 MLE: $(\hat{\lambda}_Y, \hat{\lambda}_X)$ ，將 $(\hat{\lambda}_Y, \hat{\lambda}_X)$ 代回上式而求得圖形結構 $Y \rightarrow X$ 之概似函數 $f_2(\underline{y}, \underline{x} | \hat{\lambda}_Y, \hat{\lambda}_X)$ 。

Step 3：判斷圖形結構何者的可能性最大？

經由 Step 1 及 Step 2 可得兩圖形結構之概似函數，欲知何者的可能性最大可利用一簡單的式子得知：

$$k = \frac{f_1(x, y | \hat{\lambda}_x, \hat{\lambda}_y)}{f_1(x, y | \hat{\lambda}_x, \hat{\lambda}_y) + f_2(y, x | \hat{\lambda}_y, \hat{\lambda}_x)} \quad (3-3-7)$$

當 $k > \frac{1}{2}$ 時，判定資料 X、Y 兩者之間的圖形結構為 $X \rightarrow Y$ ；反之當 $k < \frac{1}{2}$ 時，則判定資料 X、Y 兩者之間的圖形結構為 $Y \rightarrow X$ ；若當 $k = \frac{1}{2}$ 時，我們則需再尋找額外的資訊加以判斷資料之因果關係方向性。

若當對於資料 X、Y 之間的因果關係方向性之認知不是等價時，亦即認為圖形結構 $X \rightarrow Y$ 及 $Y \rightarrow X$ 兩者之間的可能性並非相同時，則可使用更一般性的式子，當 Prior 設為 p 時：

$$k = \frac{pf_1(x, y | \lambda_x, \lambda_y)}{pf_1(x, y | \lambda_x, \lambda_y) + qf_2(y, x | \lambda_y, \lambda_x)} \quad \text{where } p + q = 1 \quad (3-3-8)$$

3-4 考慮 (X, Y) 間不同的機率模型

在上一小節中，我們使用一般式來推導出判斷圖形結構之概似函數；下面將列出 (1) Normal 及 (2) Exponential 兩種機率模型的概似函數公式，並模擬 Normal、Exponential、Cauchy、Double Exponential、Gamma 等五種分配的資料，分別測試若模型假設

錯誤，我們能否判斷出正確之圖形結構。一開始先導出分配為 Normal 之概似函數公式。

3-4.1 (X, Y) 間機率模型為 Normal 分配

假設資料分配服從 $X \sim N(\theta_X, \sigma_X^2)$ ， $Y|X \sim N(X\theta_Y, \sigma_Y^2)$ ，且參數的先驗分配中仍存在一個參數 λ ， $\theta_X \sim N(\lambda_X, 1)$ ， $\theta_Y \sim N(\lambda_Y, 1)$ ，記為 (模型 3-4.1)；藉由上節之步驟 1 可得圖形結構 $X \rightarrow Y$ 之概似函數 $f_1(x, y | \hat{\lambda}_X, \hat{\lambda}_Y)$ ：

$$\frac{1}{2\pi\sqrt{(n+\sigma_X^2)(\sum x_i^2 + \sigma_Y^2)}} \exp\left\{-\frac{\sum x_i^2 - n\bar{X}^2}{2\sigma_X^2} - \frac{\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2}}{2\sigma_Y^2}\right\}$$

同理可以求得圖形結構 $Y \rightarrow X$ 之概似函數 $f_2(y, x | \hat{\lambda}_Y, \hat{\lambda}_X)$ ：

$$\frac{1}{2\pi\sqrt{(n+\sigma_Y^2)(\sum y_i^2 + \sigma_X^2)}} \exp\left\{-\frac{\sum y_i^2 - n\bar{Y}^2}{2\sigma_Y^2} - \frac{\sum X_i^2 - \frac{(\sum y_i X_i)^2}{\sum y_i^2}}{2\sigma_X^2}\right\}$$

經由上述兩個式子再利用上節之步驟 3 我們可以導出一判斷圖形結構公式其資料服從 Normal 分配：

$$k = \frac{p \frac{1}{2\pi\sqrt{(n+\sigma_X^2)(\sum x_i^2 + \sigma_Y^2)}} \exp\left\{-\frac{\sum x_i^2 - n\bar{X}^2}{2\sigma_X^2} - \frac{\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2}}{2\sigma_Y^2}\right\}}{p \frac{1}{2\pi\sqrt{(n+\sigma_X^2)(\sum x_i^2 + \sigma_Y^2)}} \exp\left\{-\frac{\sum x_i^2 - n\bar{X}^2}{2\sigma_X^2} - \frac{\sum y_i^2 - \frac{(\sum x_i y_i)^2}{\sum x_i^2}}{2\sigma_Y^2}\right\} + (1-p) \frac{1}{2\pi\sqrt{(n+\sigma_Y^2)(\sum y_i^2 + \sigma_X^2)}} \exp\left\{-\frac{\sum y_i^2 - n\bar{Y}^2}{2\sigma_Y^2} - \frac{\sum X_i^2 - \frac{(\sum y_i X_i)^2}{\sum y_i^2}}{2\sigma_X^2}\right\}}$$

3-4.2 模擬五種不同分配，使用之公式資料服從 Normal 分配

底下我們除了模擬出 Normal 的資料驗證其合理性外，並模擬來自不同分配的資料；在假設（模型 3-4.1）是對的情況下，測試若分配假設錯誤的情況下能否判斷出正確之圖形結構；設定樣本數為 100，重覆 1000 次，列出其在不同的分配及不同的先驗分配選擇下，判斷正確的次數，此時的先驗分配如同上述，假設 Prior 為 p 的情況。

在列出表格前，對於這篇文章中資料的模型分配代號先做個註記，接著列出部分模擬結果：

N : Normal Distribution、Exp : Exponential Distribution、
DE : Double Exponential Distribution、G : Gamma、C : Cauchy

X \ Y p	$N(0.5, 2.0)$	$N(2, 3.0)$	$Exp(0.5)$	$Exp(0.5)$
	$N(3.0x, 5.0)$	$N(2x^2 + 5x - 1, 2)$	$Exp(\frac{0.02}{x})$	$2x^2 + x + 1 + Exp(1)$
0.1	56	1000	1000	1000
0.2	633	1000	1000	1000
0.3	968	1000	1000	1000
0.4	995	1000	1000	1000
0.5	1000	1000	1000	1000
0.6	1000	1000	1000	1000
0.7	1000	1000	1000	1000
0.8	1000	1000	1000	1000
0.9	1000	1000	1000	1000

X \ Y p	$DE(1.0,2.4)$	$DE(1.0,2.4)$	$Cauchy(5.0,1.5)$
	$2.3x + DE(0,4)$	$x^2 + 2x - 5 + DE(0,4)$	$3x^2 + x + 5 + Cauchy(0,5)$
0.1	233	994	1000
0.2	696	998	1000
0.3	960	1000	1000
0.4	996	1000	1000
0.5	997	1000	1000
0.6	1000	1000	1000
0.7	1000	1000	1000
0.8	1000	1000	1000
0.9	1000	1000	1000

X \ Y p	$Cauchy(5.0,1.5)$	$Gamma(1,1)$	$Gamma(1,1)$
	$1.5x + Cauchy(0,5)$	$3x + Gamma(1,5)$	$2x^2 - 3x + 7 + Gamma(1,5)$
0.1	610	0	2
0.2	685	8	3
0.3	847	29	4
0.4	998	51	4
0.5	999	84	5
0.6	999	121	7
0.7	1000	187	8
0.8	1000	288	12
0.9	1000	419	17

從上面的表格中，發現除了 Gamma 分配並無法使用此法直接有效的判斷出圖形結構外，在其餘的模型下皆可以判斷出正確的圖形結構；仔細觀察 Gamma 分配的結果可以發現 Prior 的選取會大大地影響結果，於是我們希望能避免使用 Prior，以免

直接影響結果；而這當中也可能是因為 Normal 分配較有不易分辨出圖形結構的情況產生，於是我們先試著用不同的資料模型再做分析，底下將模型由 Normal 分配換成 Exponential 分配。

3-4.3 (X,Y)間機率模型為 Exponential 分配

假設資料 $X \sim Exp(\theta_x)$ [‡] (註二)， $Y|X \sim Exp\left(\frac{\theta_y}{X}\right)$ ，且參數的先驗分配中仍存在一個參數 λ ， $\theta_x \sim Exp(\lambda_x)$ ， $\theta_y \sim Exp(\lambda_y)$ ，記為(模型 3-4.3)；藉由上述方法可得圖形結構 $X \rightarrow Y$ 之概似函數 $f_1(x, y | \hat{\lambda}_x, \hat{\lambda}_y)$ ：

$$\frac{1}{n} [\Gamma(n+1)]^2 \bar{X} \sum \frac{y_i}{x_i} (\Pi x_i)^{-1} \left[\frac{1}{n} (\frac{1}{n} + 1)^2 \sum x_i \sum \frac{y_i}{x_i} \right]^{-(n+1)}$$

同理可以求得圖形結構 $Y \rightarrow X$ 之概似函數 $f_2(y, x | \hat{\lambda}_y, \hat{\lambda}_x)$ ：

$$\frac{1}{n} [\Gamma(n+1)]^2 \bar{Y} \sum \frac{x_i}{y_i} (\Pi y_i)^{-1} \left[\frac{1}{n} (\frac{1}{n} + 1)^2 \sum y_i \sum \frac{x_i}{y_i} \right]^{-(n+1)}$$

經由上述兩個式子，及假設 Prior 為 p 時，可將判斷圖形結構之公式整理成如下：

$$k = \frac{p(\Pi y_i) \left(\sum y_i \sum \frac{x_i}{y_i} \right)^n}{p(\Pi y_i) \left(\sum y_i \sum \frac{x_i}{y_i} \right)^n + (1-p)(\Pi x_i) \left(\sum x_i \sum \frac{y_i}{x_i} \right)^n}$$

上式即為資料服從 Exponential 分配 (模型 3-4.3) 的公式。

[‡]註二： $EX = \frac{1}{\theta_x}$

3-4.4 模擬四種不同分配，使用之公式資料服從 Exponential

本節採用的判斷公式中，資料是服從 Exponential 分配，我們亦想知道，在假設（模型 3-4.3）是對的情況下，測試若分配假設錯誤能否判斷出正確之圖形結構；仍舊設定樣本數為 100，重覆 1000 次，列出其在不同的分配及不同的先驗分配選擇下，判斷正確的次數；此時之先驗分配亦如同上述，假設 Prior 為 p 的情況。

在下頁的兩個表格中，由於 Exponential 分配皆為正數，所以捨棄了 Cauchy 分配，只採用了四種不同機率模型，而變數間以線性及二次兩種關係去模擬；在 Normal、Exponential、Gamma、Double Exponential 四種模型中，除了與判斷公式同模型的 Exponential 分配，無論變數間是線性或二次關係，皆可以判斷出正確的圖形結構；而其餘三種模型中，若變數間為線性關係時，除非 $p > 0.5$ ，否則無法判斷出正確的圖形結構；反之，變數間為二次關係時，只要 $p > 0.1$ 時，則可將正確的圖形結構判斷出。接著我們想再更進一步的使準確度提高，尤其是在變數之間為線性關係的情況下。

X \ Y Prior	X	$N(8,2)$	$N(10,2)$	$Gamma(3,2)$	$Gamma(3,2)$
	Y	$N(6x,1)$	$N(2x^2 + 5x,2)$	$5x + Gamma(1,5)$	$x^2 + x + Gamma(1,5)$
0.1	0	0	575	0	1000
0.2	0	0	1000	0	1000
0.3	0	0	1000	0	1000
0.4	0	0	1000	0	1000
0.5	724	0	1000	2	1000
0.6	1000	0	1000	34	1000
0.7	1000	0	1000	233	1000
0.8	1000	0	1000	661	1000
0.9	1000	0	1000	950	1000

X \ Y Prior	X	$Exp(0.5)$	$Exp(0.2)$	$DE(12,1.2)$	$DE(12,1.2)$
	Y	$Exp(\frac{0.02}{x})$	$2x^2 + x + 1 + Exp(1)$	$10x + DE(0,2)$	$2x^2 + 4 + DE(0,2)$
0.1	1000	1000	966	0	353
0.2	1000	1000	967	0	948
0.3	1000	1000	967	0	1000
0.4	1000	1000	967	0	1000
0.5	1000	1000	967	699	1000
0.6	1000	1000	968	999	1000
0.7	1000	1000	968	1000	1000
0.8	1000	1000	968	1000	1000
0.9	1000	1000	969	1000	1000

3-5 直接使用概似函數判斷兩個變數間之方向性

若只有兩個變數，則圖形方向性只有兩種可能：① $X \rightarrow Y$ 及
 ② $X \leftarrow Y$ ，底下先設法求出代表其中一個圖形結構的概似函數，

另一個圖形結構則可直接利用其相對的變數關係，即可求得。

首先假設資料的分配為 $X \sim f_X(x|\theta_X)$ ， $Y|X \sim f_{Y|X}(y|x, \theta_Y)$ ，因此我們可以得到圖形結構 $X \rightarrow Y$ 之概似函數：

$$L(x, y|\theta_X, \theta_Y) = f_X(x|\theta_X) f_{Y|X}(y|x, \theta_Y)$$

為了求出上式中參數的 MLE，一般會先取自然對數，接著將求出的 MLE $(\hat{\theta}_X, \hat{\theta}_Y)$ 代回上式而求得圖形結構 $X \rightarrow Y$ 之概似函數：

$$L(x, y|\hat{\theta}_X, \hat{\theta}_Y) = f_X(x|\hat{\theta}_X) f_{Y|X}(y|x, \hat{\theta}_Y) \quad (3-5-1)$$

同理可得圖形結構 $X \leftarrow Y$ 之概似函數為：

$$L(y, x|\hat{\theta}_Y, \hat{\theta}_X) = f_Y(y|\hat{\theta}_Y) f_{X|Y}(x|y, \hat{\theta}_X) \quad (3-5-2)$$

最後我們只要比較由式子 (3-5-1) 及 (3-5-2) 所得的兩個概似函數值之大小即可判斷出實際資料之圖形結構為何者。

3-5.1 考慮 (X, Y) 間的機率模型為線性關係之 Normal

假設資料中 X 、 Y 兩個變數都來自常態分配，而變數間是為線性關係，並採用來自常態分配的誤差，依照上小節之方法設法判定出有最大概似函數值的圖形結構。

首先假設資料的分配為 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX + b, \sigma_Y^2)$ ，記為 (模型 3-5.1)，由此模型假設我們可以得此圖形結構 $X \rightarrow Y$

之概似函數：

$$\begin{aligned} L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x_i - \mu_X)^2}{2\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(y_i - ax_i - b)^2}{2\sigma_Y^2}} \\ &= (2\pi\sigma_X^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} (2\pi\sigma_Y^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2}} \end{aligned}$$

$$\Rightarrow \ln L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2)$$

$$= -\frac{n}{2} \ln(2\pi\sigma_X^2) - \frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2} - \frac{n}{2} \ln(2\pi\sigma_Y^2) - \frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2}$$

將 $\ln L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2)$ 函數對參數作偏微，可得到參數之 MLE：

$$\begin{aligned} \hat{\mu}_X &= \bar{X}, \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2, \\ \hat{a} &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}, \quad \hat{b} = \bar{Y} - \hat{a} \bar{X} \\ \hat{\sigma}_Y^2 &= \frac{1}{n} \sum_{i=1}^n \left[y_i - \bar{Y} - \hat{a}(x_i - \bar{X}) \right]^2, \end{aligned}$$

將 $(\hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2)$ 代回函數 $L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2)$ 而求得圖形結構 $X \rightarrow Y$ 之

概似函數：

$$L(\underline{x}, \underline{y} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2) = \left(\frac{2\pi e}{n} \right)^{-n} \left\{ \sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2 - \left[\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \right]^2 \right\}^{-\frac{n}{2}}$$

同理我們可以求得圖形結構 $X \leftarrow Y$ 之概似函數：

$$L(\underline{y}, \underline{x} | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_X^2) = \left(\frac{2\pi e}{n} \right)^{-n} \left\{ \sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (X_i - \bar{X})^2 - \left[\sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X}) \right]^2 \right\}^{-\frac{n}{2}}$$

由上述兩步驟中求得的兩個函數，我們可以發現，其實兩個函數之值永遠會相等，亦即在兩變數之間為線性關係時，我們無法直接由概似函數判斷出正確之圖形結構；無論資料的模型為何，所得之函數值都相同。底下將線性關係改為二次關係，用同樣的方法，看結果如何？

3-5.2 考慮 (X, Y) 間的機率模型為二次關係之 Normal

在二次關係中，仍舊以 X 、 Y 兩個變數都來自常態分配為例，而誤差來自常態分配，依照上述的步驟設法判定出有最大概似函數值的圖形結構。

首先假設資料服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX^2 + bX + c, \sigma_Y^2)$ ，由以上假設我們可得圖形結構 $X \rightarrow Y$ 之概似函數：

$$L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2) = \left[(2\pi)^2 \sigma_X^2 \sigma_Y^2 \right]^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2}{2\sigma_Y^2}}$$

將上式函數取自然對數再對參數作偏微，可得到參數之 MLE：

$$\begin{aligned} \hat{\mu}_X &= \bar{X}, \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2, \\ \hat{a} &= \frac{S_{XX}S_{X^2Y} - S_{XY}S_{XX^2}}{S_{XX}S_{X^2X^2} - (S_{XX^2})^2}, \quad \hat{b} = \frac{S_{X^2X^2}S_{XY} - S_{X^2Y}S_{X^2X}}{S_{X^2X^2}S_{XX} - (S_{X^2X})^2}, \\ \hat{c} &= \bar{Y} - \hat{a} \frac{1}{n} \sum_{i=1}^n x_i^2 - \hat{b} \bar{X} \\ \hat{\sigma}_Y^2 &= \frac{1}{n} \sum_{i=1}^n \left[(y_i - \bar{Y}) - \hat{a} \left(x_i^2 - \frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \hat{b} (x_i - \bar{X}) \right]^2, \end{aligned}$$

當中之 $S_{AB} = \sum_{i=1}^n (A_i - \bar{A})(B_i - \bar{B})$ ；再將 $(\hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2)$ 代回概似函數 $L(\underline{x}, \underline{y} | \mu_X, \sigma_X^2, \sigma_Y^2)$ ，進一步求得圖形結構 $X \rightarrow Y$ 之概似函數：

$$L(\underline{x}, \underline{y} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2) = \left[(2\pi e)^2 \hat{\sigma}_X^2 \hat{\sigma}_Y^2 \right]^{-\frac{n}{2}}$$

$$= \left(\frac{2\pi e}{n} \right)^{-n} \left\{ S_{XX} \left[S_{YY} + \frac{2S_{XY}S_{XX^2}S_{X^2Y} - S_{XX}(S_{X^2Y})^2 - S_{X^2X^2}(S_{XY})^2}{S_{XX}S_{X^2X^2} - (S_{XX^2})^2} \right] \right\}^{-\frac{n}{2}}$$

同理圖形結構 $X \leftarrow Y$ 之概似函數可直接將上式之 X 、 Y 互換而得

$$L(\underline{y}, \underline{x} | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_X^2) = \left[(2\pi e)^2 \hat{\sigma}_Y^2 \hat{\sigma}_X^2 \right]^{-\frac{n}{2}}。$$

3-5.3 模擬五種分配資料，以二次關係之 Normal 公式做判斷

底下我們將模擬在不同的資料模型中，若仍沿用上小節中資料為 Normal 分配，而使用變數之間為二次關係所導出之概似函數，欲知是否可以判斷出正確的圖形結構？只要當實際資料的圖形結構之概似函數值大於等於另一圖形結構函數值即為判斷正確。底下使用 100、1000 兩種樣本數做模擬，皆重覆 1000 次，列出判斷正確的次數。

Case 1：資料為 Normal 分配，且變數之間的關係為線性關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
N(1,1)	N(-3X+7,1)	100	491	1000	487
N(1,3)	N(X+5,5)	100	489	1000	475
N(3,2)	N(6X-2,6)	100	532	1000	480
N(0,1)	N(-3X+1,3)	100	479	1000	530
N(0,5)	N(-X+10,2)	100	484	1000	495

從上表發現，在 Normal 分配下，若兩變數之間為線性關係時，不容易判斷圖形是為 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。此為合理結果，由於在線性關係方面，當 X 為 Y 的線性時亦可寫成 Y 為 X 的線性；在這個例子中，我們也可以將 X 為 Y 的二次寫成 Y 為 X 的二次，因此兩圖形被判斷正確的可能性皆為一半左右。

Case 2：資料為 Normal 分配，而變數之間的關係為二次關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
N(1,1)	$N(X^2-3X+7,1)$	100	1000	1000	1000
N(1,3)	$N(4X^2+X+5,5)$	100	1000	1000	1000
N(3,2)	$N(2X^2+6X-2,6)$	100	1000	1000	1000
N(0,1)	$N(5X^2-3X+1,3)$	100	1000	1000	1000
N(0,5)	$N(2X^2-X+10,2)$	100	1000	1000	1000

從上表發現，在 Normal 分配下，若兩變數之間為二次關係時，無論資料變異有多大，皆可判斷出正確的圖形結構 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 3：資料為 Exponential 分配，變數間的關係為線性關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
Exp(5)	$-X+10+\text{Exp}(5)$	100	0	1000	0
Exp(3)	$-3X+7+\text{Exp}(1)$	100	1	1000	0
Exp(2)	$-2X+6+\text{Exp}(5)$	100	104	1000	0
Exp(1)	$X+1+\text{Exp}(0.5)$	100	160	1000	6
Exp(1)	$2X+1+\text{Exp}(2)$	100	137	1000	7

從上表發現，在 Exponential 分配，若兩變數之間為線性關係時，不容易判斷圖形是為 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 4：資料為 Exponential 分配，變數間的關係為二次關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
Exp(5)	$2X^2-X+10+\text{Exp}(5)$	100	989	1000	1000
Exp(1)	$X^2+X+1+\text{Exp}(0.5)$	100	998	1000	1000
Exp(3)	$6X^2-3X+7+\text{Exp}(1)$	100	999	1000	1000
Exp(2)	$X^2-2X+6+\text{Exp}(5)$	100	998	1000	1000
Exp(1)	$X^2+2X+1+\text{Exp}(2)$	100	1000	1000	1000

從上表發現，在 Exponential 分配，若兩變數之間為二次關係時，無論資料變異有多大，幾乎皆可判斷出正確的圖形結構 $X \rightarrow Y$ 或 $X \leftarrow Y$ ；在這個模擬結果中，我們亦發現，誤差項的變動對於判斷結果似乎不是很重要。

Case 5：資料為 Double Exponential 分配，而變數間為線性關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
DE(5,1)	$-X+10+\text{DE}(0,1)$	100	303	1000	288
DE(1,1)	$X+1+\text{DE}(0,1)$	100	309	1000	278
DE(3,3)	$-3X+7+\text{DE}(0,3)$	100	347	1000	363
DE(2,4)	$2X+1+\text{DE}(0,2)$	100	369	1000	375
DE(1,6)	$-2X+6+\text{DE}(0,5)$	100	381	1000	357

從上表發現，在 Double Exponential 分配下，若兩變數之間為線性關係時，不容易判斷圖形是為 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 6：資料為 Double Exponential 分配，而變數間為二次關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
DE(5,1)	$2X^2-X+10+\text{DE}(0,1)$	100	1000	1000	1000
DE(1,1)	$X^2+X+1+\text{DE}(0,1)$	100	1000	1000	1000
DE(3,3)	$6X^2-3X+7+\text{DE}(0,3)$	100	1000	1000	1000
DE(2,4)	$X^2+2X+1+\text{DE}(0,2)$	100	1000	1000	1000
DE(1,6)	$X^2-2X+6+\text{DE}(0,5)$	100	1000	1000	1000

從上頁的表格發現，在 Double Exponential 分配下，若兩變數之間為二次關係時，無論資料變異有多大，皆可判斷出正確的圖形結構 $X \rightarrow Y$ 或 $X \leftarrow Y$ ；在這個模擬結果中，我們亦發現，誤差項的變動對於判斷結果似乎不是很重要。

Case 7：資料為 Cauchy 分配，而變數之間的關係為線性關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
C(5,1)	C(-X+10,1)	100	48	1000	4
C(1,1)	C(X+1,1)	100	46	1000	7
C(1,6)	C(-2X+6,5)	100	67	1000	13
C(2,4)	C(2X+1,2)	100	78	1000	15
C(3,3)	C(-3X+7,3)	100	90	1000	17

從上表發現，在 Cauchy 分配下，若兩變數之間為線性關係時，不容易判斷圖形是為 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 8：資料為 Cauchy 分配，而變數之間的關係為二次關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
C(5,1)	C(2X ² -X+10,1)	100	992	1000	998
C(1,1)	C(X ² +X+1,1)	100	986	1000	1000
C(3,3)	C(6X ² -3X+7,3)	100	999	1000	1000
C(2,4)	C(X ² +2X+1,2)	100	991	1000	1000
C(1,6)	C(X ² -2X+6,5)	100	994	1000	1000

從上表發現，在 Cauchy 分配下，若兩變數之間為二次關係時，無論資料變異有多大，皆可判斷出正確的圖形結構 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 9：資料為 Gamma 分配，而變數之間的關係為線性關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
G(1,6)	$-X+10+G(1,2)$	100	1	1000	0
G(1,1)	$-X+5+G(1,5)$	100	6	1000	0
G(5,1)	$-3X+7+G(1,1)$	100	320	1000	42
G(3,3)	$6X-2+G(1,6)$	100	318	1000	112
G(2,4)	$2X+1+G(1,3)$	100	231	1000	139

從上表發現，在 Gamma 分配下，若兩變數之間為線性關係時，不容易判斷圖形是為 $X \rightarrow Y$ 或 $X \leftarrow Y$ 。

Case 10：資料為 Gamma 分配，而變數之間的關係為二次關係。

資料 X	資料 Y	樣本數	正確次數	樣本數	正確次數
G(5,1)	$X^2-3X+7+G(1,1)$	100	1000	1000	1000
G(1,1)	$4X^2-X+5+G(1,5)$	100	1000	1000	1000
G(3,3)	$2X^2+6X-2+G(1,6)$	100	1000	1000	1000
G(2,4)	$5X^2+2X+1+G(1,3)$	100	1000	1000	1000
G(1,6)	$2X^2-X+10+G(1,2)$	100	1000	1000	1000

從上表發現，在 Gamma 分配下，若兩變數之間為二次關係時，無論資料變異有多大，皆可判斷出正確的圖形結構 $X \rightarrow Y$ 或 $X \leftarrow Y$ ；在這個模擬結果中，我們亦發現，誤差項的變動對於判斷結果似乎不是很重要。

由以上 10 個 Cases，我們在此做個小整理，若資料中的變數為二次關係時，無論是何種分配我們皆可判斷出正確的圖形結構；反之，資料中的變數為線性關係時，仍無法直接由（模型 3-5.2）判斷出正確的圖形結構，整理的表格如下頁所示。

資料模型	變數間為線性關係	變數間為二次關係
Normal	✗	✓
Exponential	✗	✓
Double Exponential	✗	✓
Cauchy	✗	✓
Gamma	✗	✓

3-6 直接使用概似函數判斷三個變數間之方向性

若有 X 、 Y 、 Z 三個變數時，圖形之因果關係方向性有底下三種類型：① $X \rightarrow Y \rightarrow Z$ 、② $X \rightarrow Z \leftarrow Y$ 、③ $Y \leftarrow X \rightarrow Z$ ，而圖形①中 X 、 Y 、 Z 三個變數的位置若置換順序有 6 種組合的可能，圖形②、③各有 3 種組合的可能；詳細的 12 種組合情形已在第一章時列出，底下即開始分析代表這三類圖形結構的概似函數，每一類型中， X 、 Y 、 Z 三個變數的順序僅以其中一個組合為代表，亦即當圖形類型相同，而三個變數的順序若有改變時，概似函數中只要變數做相對應的代換即可。

Step 1：求出代表圖形結構 $X \rightarrow Y \rightarrow Z$ 的概似函數

假設資料 $X \sim f_X(x|\theta_X)$ ， $Y|X \sim f_{Y|X}(y|x, \theta_Y)$ ， $Z|Y \sim f_{Z|Y}(z|y, \theta_Z)$ ，

由以上假設我們可以得到此圖形之概似函數：

$$L_1(x, y, z|\theta_X, \theta_Y, \theta_Z) = f_X(x|\theta_X) f_{Y|X}(y|x, \theta_Y) f_{Z|Y}(z|y, \theta_Z)$$

為了求出參數的 MLE，一般會對上式取自然對數，接著將求出的 MLE $(\hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z)$ 代回上式而求得圖形結構 $X \rightarrow Y \rightarrow Z$ 之概似函數 $L_1(x, y, z | \hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z)$ 。

而在 X, Y, Z 三變數的資料中，依照同樣的方法求得圖形結構 $X \rightarrow Z \rightarrow Y$ 之概似函數：

$$L_1(x, z, y | \hat{\theta}_X, \hat{\theta}_Z, \hat{\theta}_Y) = f_X(x | \hat{\theta}_X) f_{Z|X}(z | x, \hat{\theta}_Z) f_{Y|Z}(y | z, \hat{\theta}_Y)$$

同理，可求出圖形結構相對應關係為 $A \rightarrow B \rightarrow C$ 之概似函數：

$$\begin{aligned} Y \rightarrow X \rightarrow Z &: L_1(y, x, z | \hat{\theta}_Y, \hat{\theta}_X, \hat{\theta}_Z) \\ Y \rightarrow Z \rightarrow X &: L_1(y, z, x | \hat{\theta}_Y, \hat{\theta}_Z, \hat{\theta}_X) \\ Z \rightarrow X \rightarrow Y &: L_1(z, x, y | \hat{\theta}_Z, \hat{\theta}_X, \hat{\theta}_Y) \\ Z \rightarrow Y \rightarrow X &: L_1(z, y, x | \hat{\theta}_Z, \hat{\theta}_Y, \hat{\theta}_X) \end{aligned}$$

Step 2：求出代表圖形結構 $X \rightarrow Z \leftarrow Y$ 的概似函數

假設資料 $X \sim f_X(x | \theta_X)$ ， $Y \sim f_Y(y | \theta_Y)$ ， $Z | X, Y \sim f_{Z|X,Y}(z | x, y, \theta_Z)$ ，

where $X \perp Y$ ，由以上假設我們可以得到此圖形之概似函數：

$$L_2(x, z, y | \theta_X, \theta_Z, \theta_Y) = f_X(x | \theta_X) f_Y(y | \theta_Y) f_{Z|X,Y}(z | x, y, \theta_Z)$$

接著求出參數之 MLE $(\hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z)$ 並代回上式而求得圖形結構 $X \rightarrow$

$Z \leftarrow Y$ 之概似函數 $L_2(x, z, y | \hat{\theta}_X, \hat{\theta}_Z, \hat{\theta}_Y)$ 。

依照同樣的方法求得圖形結構 $X \rightarrow Y \leftarrow Z$ 之概似函數：

$$L_2(\underline{x}, \underline{y}, \underline{z} | \hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z) = f_X(\underline{x} | \hat{\theta}_X) f_Z(\underline{z} | \hat{\theta}_Z) f_{Y|X,Z}(\underline{y} | \underline{x}, \underline{z}, \hat{\theta}_Y)$$

同理，可求出圖形結構 $Y \rightarrow X \leftarrow Z$ 之概似函數：

$$L_2(\underline{y}, \underline{x}, \underline{z} | \hat{\theta}_Y, \hat{\theta}_X, \hat{\theta}_Z) = f_Y(\underline{y} | \hat{\theta}_Y) f_Z(\underline{z} | \hat{\theta}_Z) f_{X|Y,Z}(\underline{x} | \underline{y}, \underline{z}, \hat{\theta}_X)$$

Step 3：求出代表圖形結構 $Y \leftarrow X \rightarrow Z$ 的概似函數

假設資料 $X \sim f_X(\underline{x} | \theta_X)$ ， $Y|X \sim f_{Y|X}(\underline{y} | \underline{x}, \theta_Y)$ ， $Z|X \sim f_{Z|X}(\underline{z} | \underline{x}, \theta_Z)$ ，

where $Y \perp Z | X$ ，由以上假設我們可以得到此圖形之概似函數：

$$L_3(\underline{y}, \underline{x}, \underline{z} | \theta_Y, \theta_X, \theta_Z) = f_X(\underline{x} | \theta_X) f_{Y|X}(\underline{y} | \underline{x}, \theta_Y) f_{Z|X}(\underline{z} | \underline{x}, \theta_Z)$$

接著求出參數之 MLE $(\hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z)$ ，並將之代回上式而求得圖形結

構 $Y \leftarrow X \rightarrow Z$ 之概似函數 $L_3(\underline{y}, \underline{x}, \underline{z} | \hat{\theta}_Y, \hat{\theta}_X, \hat{\theta}_Z)$ 。

依照同樣的方法求得圖形結構 $X \leftarrow Y \rightarrow Z$ 之概似函數：

$$L_3(\underline{x}, \underline{y}, \underline{z} | \hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z) = f_Y(\underline{y} | \hat{\theta}_Y) f_{X|Y}(\underline{x} | \underline{y}, \hat{\theta}_X) f_{Z|Y}(\underline{z} | \underline{y}, \hat{\theta}_Z)$$

同理，可求出圖形結構 $X \leftarrow Z \rightarrow Y$ 之概似函數：

$$L_3(\underline{x}, \underline{z}, \underline{y} | \hat{\theta}_X, \hat{\theta}_Z, \hat{\theta}_Y) = f_Z(\underline{z} | \hat{\theta}_Z) f_{X|Z}(\underline{x} | \underline{z}, \hat{\theta}_X) f_{Y|Z}(\underline{y} | \underline{z}, \hat{\theta}_Y)$$

Step 4：判斷圖形結構何者的可能性最大？

經由上述三個步驟，我們可以得到 X 、 Y 、 Z 三個變數中所有可能圖形結構之概似函數，共 12 個；由這 12 個概似函數值

中，我們可以找出值為最大的。當最大值只有一個圖形結構時，我們判定此概似函數值所對應之圖形結構為我們所尋求之因果關係圖；反之，當最大值不只一個時，我們只能判斷出和資料之因果關係圖等價的圖形結構。

在本小節中，我們使用一般式來推導出判斷圖形結構之概似函數；下面將直接推導出 (X, Y, Z) 三變數皆來自 Normal 分配的概似函數公式，仍舊會分成線性關係及二次關係。並使用程式模擬 Normal、Exponential、Double Exponential、Cauchy、Gamma 等五種分配的資料，當中包含了變數間關係為線性及二次的資料，分別測試若假設錯誤，我們能否判斷出正確之圖形結構。



3-6.1 考慮 (X, Y, Z) 間的機率模型為線性關係之 Normal

假設 X 、 Y 、 Z 三個變數都來自常態分配，而變數間是為線性關係，並採用來自常態分配的誤差，依照上述四個步驟設法判定出有最大概似函數值的圖形結構。

Step 1：求出代表圖形結構 $X \rightarrow Y \rightarrow Z$ 的概似函數

首先假設資料分配服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX+b, \sigma_Y^2)$ ， $Z|Y \sim N(cY+d, \sigma_Z^2)$ ，由以上假設我們可以得此圖形結構之概似函數如下頁所示：

$$\begin{aligned}
& L_1(x, y, z | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2) \\
&= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x_i - \mu_X)^2}{2\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(y_i - ax_i - b)^2}{2\sigma_Y^2}} \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-\frac{(z_i - cy_i - d)^2}{2\sigma_Z^2}} \\
&= (2\pi\sigma_X^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} (2\pi\sigma_Y^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2}} (2\pi\sigma_Z^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (z_i - cy_i - d)^2}{2\sigma_Z^2}}
\end{aligned}$$

$$\begin{aligned}
& \Rightarrow \ln L_1(x, y, z | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2) \\
&= -\frac{n}{2} \ln(2\pi\sigma_X^2) - \frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2} - \frac{n}{2} \ln(2\pi\sigma_Y^2) - \frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2} - \frac{n}{2} \ln(2\pi\sigma_Z^2) - \frac{\sum_{i=1}^n (z_i - cy_i - d)^2}{2\sigma_Z^2}
\end{aligned}$$

將 $\ln L_1(x, y, z | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2)$ 函數對參數作偏微，可得到參數之 MLE：

$$\begin{aligned}
\hat{\mu}_X &= \bar{X} \quad , \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad , \\
\hat{a} &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad , \quad \hat{b} = \bar{Y} - \hat{a} \bar{X} \\
\hat{\sigma}_Y^2 &= \frac{1}{n} \sum_{i=1}^n \left[y_i - \bar{Y} - \hat{a} (x_i - \bar{X}) \right]^2 \quad , \\
\hat{c} &= \frac{n \sum_{i=1}^n y_i z_i - \sum_{i=1}^n y_i \sum_{i=1}^n z_i}{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2} \quad , \quad \hat{d} = \bar{Z} - \hat{c} \bar{Y} \\
\hat{\sigma}_Z^2 &= \frac{1}{n} \sum_{i=1}^n \left[z_i - \bar{Z} - \hat{c} (y_i - \bar{Y}) \right]^2
\end{aligned}$$

$(\hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$ 代回函數 $L_1(x, y, z | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2)$ 而求得圖形結構

$X \rightarrow Y \rightarrow Z$ 之概似函數 $L_1(x, y, z | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$ 。

而在 X, Y, Z 三變數的資料中，依照同樣的方法求得圖形

結構 $X \rightarrow Z \rightarrow Y$ 之概似函數： $L_1(x, z, y | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$

同理，可求出圖形結構相對應關係為 $A \rightarrow B \rightarrow C$ 之概似函數：

$$Y \rightarrow X \rightarrow Z : L_1(y, x, z | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$$

$$Y \rightarrow Z \rightarrow X : L_1(y, z, x | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2, \hat{\sigma}_X^2)$$

$$Z \rightarrow X \rightarrow Y : L_1(z, x, y | \hat{\mu}_Z, \hat{\sigma}_Z^2, \hat{\sigma}_X^2, \hat{\sigma}_Y^2)$$

$$Z \rightarrow Y \rightarrow X : L_1(z, y, x | \hat{\mu}_Z, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2, \hat{\sigma}_X^2)$$

Step 2：求出代表圖形結構 $X \rightarrow Z \leftarrow Y$ 的概似函數

假設資料分配為 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y \sim N(\mu_Y, \sigma_Y^2)$ ，where $X \perp Y$ ， $Z | X, Y \sim N(aX + bY + c, \sigma_Z^2)$ ，由以上假設我們可以得到此圖形之概似函數：



$$\begin{aligned} & L_2(x, z, y | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x_i - \mu_X)^2}{2\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(y_i - \mu_Y)^2}{2\sigma_Y^2}} \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-\frac{(z_i - ax_i - by_i - c)^2}{2\sigma_Z^2}} \\ &= (2\pi\sigma_X^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} (2\pi\sigma_Y^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i - \mu_Y)^2}{2\sigma_Y^2}} (2\pi\sigma_Z^2)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (z_i - ax_i - by_i - c)^2}{2\sigma_Z^2}} \\ &\Rightarrow \ln L_2(x, z, y | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2) \\ &= -\frac{n}{2} \ln(2\pi\sigma_X^2) - \frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2} - \frac{n}{2} \ln(2\pi\sigma_Y^2) - \frac{\sum_{i=1}^n (y_i - \mu_Y)^2}{2\sigma_Y^2} - \frac{n}{2} \ln(2\pi\sigma_Z^2) - \frac{\sum_{i=1}^n (z_i - ax_i - by_i - c)^2}{2\sigma_Z^2} \end{aligned}$$

將 $\ln L_2(x, z, y | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2)$ 函數對參數作偏微，得參數之 MLE：

$$\hat{\mu}_X = \bar{X} \quad , \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad ,$$

$$\hat{\mu}_Y = \bar{Y} \quad , \quad \hat{\sigma}_Y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{Y})^2 \quad ,$$

$$\hat{a} = \frac{\left(\bar{XZ} - \frac{1}{n} \sum_{i=1}^n x_i z_i\right) \hat{\sigma}_Y^2 - \left(\bar{YZ} - \frac{1}{n} \sum_{i=1}^n y_i z_i\right) \left(\bar{XY} - \frac{1}{n} \sum_{i=1}^n x_i y_i\right)}{\left(\bar{XY} - \frac{1}{n} \sum_{i=1}^n x_i y_i\right) - \hat{\sigma}_X^2 \hat{\sigma}_Y^2},$$

$$\hat{b} = \frac{\left[\hat{a} \left(\bar{XY} - \frac{1}{n} \sum_{i=1}^n x_i y_i\right) - \left(\bar{YZ} - \frac{1}{n} \sum_{i=1}^n y_i z_i\right)\right]}{\hat{\sigma}_Y^2}$$

$$\hat{\sigma}_Z^2 = \frac{1}{n} \sum_{i=1}^n \left[\left(z_i - \bar{Z} \right) - \hat{a} \left(x_i - \bar{X} \right) - \hat{b} \left(y_i - \bar{Y} \right) \right]^2$$

再將 $(\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$ 代回函數 $L_2(\underline{x}, \underline{z}, \underline{y} | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2)$ 而求得圖形結構 $X \rightarrow Z \leftarrow Y$ 之概似函數 $L_2(\underline{x}, \underline{z}, \underline{y} | \hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$ 。

而在 X, Y, Z 三變數的資料中，依照同樣的方法求得圖形結構 $X \rightarrow Y \leftarrow Z$ 之概似函數：

$$L_2(\underline{x}, \underline{y}, \underline{z} | \hat{\mu}_X, \hat{\mu}_Z, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$$

同理，可求出圖形結構 $Y \rightarrow X \leftarrow Z$ 之概似函數：

$$L_2(\underline{y}, \underline{x}, \underline{z} | \hat{\mu}_Y, \hat{\mu}_Z, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$$

Step 3：求出代表圖形結構 $Y \leftarrow X \rightarrow Z$ 的概似函數

假設資料的分配服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX + b, \sigma_Y^2)$ ， $Z|X \sim N(cX + d, \sigma_Z^2)$ ，where $Y \perp Z|X$ ，由以上假設我們可以得到此圖形之概似函數：

$$L_3(\underline{y}, \underline{x}, \underline{z} | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2)$$

$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x_i - \mu_X)^2}{2\sigma_X^2}} \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{(y_i - ax_i - b)^2}{2\sigma_Y^2}} \frac{1}{\sqrt{2\pi\sigma_Z^2}} e^{-\frac{(z_i - cx_i - d)^2}{2\sigma_Z^2}}$$

$$= (2\pi\sigma_X^2)^{\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} (2\pi\sigma_Y^2)^{\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2}} (2\pi\sigma_Z^2)^{\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (z_i - cx_i - d)^2}{2\sigma_Z^2}}$$

$$\begin{aligned} &\Rightarrow \ln L_3(\underline{y}, \underline{x}, \underline{z} | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2) \\ &= -\frac{n}{2} \ln(2\pi\sigma_X^2) - \frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2} - \frac{n}{2} \ln(2\pi\sigma_Y^2) - \frac{\sum_{i=1}^n (y_i - ax_i - b)^2}{2\sigma_Y^2} - \frac{n}{2} \ln(2\pi\sigma_Z^2) - \frac{\sum_{i=1}^n (z_i - cx_i - d)^2}{2\sigma_Z^2} \end{aligned}$$

將 $\ln L_3(\underline{y}, \underline{x}, \underline{z} | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2)$ 函數對參數作偏微，可得參數之 MLE：

$$\begin{aligned} \hat{\mu}_X &= \bar{X} \quad , \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad , \\ \hat{a} &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad , \quad \hat{b} = \bar{Y} - \hat{a} \bar{X} \\ \hat{\sigma}_Y^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - \hat{a} x_i - \hat{b})^2 \quad , \\ \hat{c} &= \frac{n \sum_{i=1}^n x_i z_i - \sum_{i=1}^n x_i \sum_{i=1}^n z_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad , \quad \hat{d} = \bar{Z} - \hat{c} \bar{X} \\ \hat{\sigma}_Z^2 &= \frac{1}{n} \sum_{i=1}^n (z_i - \hat{c} x_i - \hat{d})^2 \end{aligned}$$

再將 $(\hat{\mu}_X, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$ 代回函數 $L_3(\underline{y}, \underline{x}, \underline{z} | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2)$ ，進而求得圖形結構 $Y \leftarrow X \rightarrow Z$ 之概似函數 $L_3(\underline{y}, \underline{x}, \underline{z} | \hat{\mu}_X, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$ 。

而在 X, Y, Z 三變數的資料中，依照同樣的方法求得圖形結構 $X \leftarrow Y \rightarrow Z$ 之概似函數：

$$L_3(\underline{x}, \underline{y}, \underline{z} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$$

同理，可求出圖形結構 $X \leftarrow Z \rightarrow Y$ 之概似函數：

$$L_3(\underline{x}, \underline{z}, \underline{y} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$$

Step 4：判斷圖形結構何者的可能性最大？

由上述三個步驟中可得到 12 個概似函數，而每個概似函數在代入 MLE 估計值後形式都可表示成下式：

$$L_i(x, y, z | \hat{\theta}_{MLE}) = (2\pi\hat{\sigma}_X^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} (2\pi\hat{\sigma}_Y^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} (2\pi\hat{\sigma}_Z^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} \quad i=1,2,3$$

上式可整理成 $\left[(2\pi e)^3 (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 \hat{\sigma}_Z^2) \right]^{-\frac{n}{2}}$ ；由於取了 $-\frac{n}{2}$ 次方後，每一個概似函數值都變的相當小，而會影響我們判斷概似函數的大小。

由於函數 $f(x) = x^{-\frac{n}{2}}$ 為一遞減函數，因此為了找出概似函數

$\left[(2\pi e)^3 (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 \hat{\sigma}_Z^2) \right]^{-\frac{n}{2}}$ 的最大值，需找出 $\left[(2\pi e)^3 (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 \hat{\sigma}_Z^2) \right]$ 的最小值。

在程式模擬過程中，我們發現若三個變數間之關係若為線性，不可能判斷出有「一個」最大概似函數值的圖形結構，因為在線性關係中，那些等價的圖形結構之概似函數會相等，因此大部分都只能判斷出符合資料因果關係的等價圖形結構。比如資料的因果關係為： $X \rightarrow Y \rightarrow Z$ 時，我們模擬出來的概似函數中，底下三個圖形結構值會相同：

$$X \rightarrow Y \rightarrow Z, Z \rightarrow Y \rightarrow X, X \leftarrow Y \rightarrow Z$$

而上述三圖之等價關係為：給定 Y，X 和 Z 會條件獨立，記為類型①。同理，底下兩類圖形之等價關係，亦會使其概似函數值相同：

(i) 給定 Z ， X 和 Y 會條件獨立，記為類型②：

$$X \rightarrow Z \rightarrow Y, Y \rightarrow Z \rightarrow X, X \leftarrow Z \rightarrow Y$$

(ii) 給定 X ， Y 和 Z 會條件獨立，記為類型③：

$$Y \rightarrow X \rightarrow Z, Z \rightarrow X \rightarrow Y, Y \leftarrow X \rightarrow Z$$

3-6.2 模擬五種分配資料，以線性關係之 Normal 公式做判斷

底下我們亦將模擬在不同的資料模型中，若仍用上小節中資料為 Normal 分配且變數間為線性關係所導出之概似函數，是否可以判斷出正確的圖形結構？只要當實際資料的圖形結構之概似函數值大於或等於所有其它之圖形結構函數值即為判斷正確。在底下的模擬過程中，樣本數皆設定為 100，重覆試驗 1000 次，而原始資料之圖形結構為 $X \rightarrow Y \rightarrow Z$ ；表中第四欄為判斷出與正確圖形結構等價的圖形次數，亦即當圖形結構為類型①： $X \rightarrow Y \rightarrow Z$ 、 $Z \rightarrow Y \rightarrow X$ 及 $X \leftarrow Y \rightarrow Z$ ，此三個圖形結構函數值為所有圖形結構函數值中最大的次數。表中第五欄將判斷錯誤部分中次數最多的圖形結構類型及判斷次數列出，前②、③類如上小節所述，類型②：給定 Z ， X 和 Y 會條件獨立、類型③：給定 X ， Y 和 Z 會條件獨立；而類型④： $X \rightarrow Z \leftarrow Y$ 、類型⑤： $X \rightarrow Y \leftarrow Z$ 及類型⑥： $Y \rightarrow X \leftarrow Z$ 。

從下頁表發現，在上述之資料模型中，若資料中變數間的

關係和判斷公式中變數間的關係同為線性關係時，幾乎皆可判斷出和正確圖形結構 $X \rightarrow Y \rightarrow Z$ 等價的圖形結構；當資料中變數間的關係為二次而判斷公式中變數間的關係為線性時，能判斷出等價圖形結構的模型僅有 Gamma 分配，在其它模型分配下，則很難判斷出來。列出部分模擬結果後，我們接著考慮變數間的關係為二次的情況，先導出判斷公式後亦使用程式模擬之。

資料 X	資料 Y	資料 Z	等價	類型
N(1, 1)	N(X+0.5, 4)	N(2Y+6, 3)	983	②: 17
N(2, 5)	N(X ² -3X+7, 2)	N(4Y ² +2Y-1, 5)	410	②: 505
Exp(1)	2X+2+Exp(2)	4Y+6+Exp(3)	911	②: 89
Exp(2)	X ² +3X-5+Exp(2)	2Y ² -2Y+5+Exp(3)	372	③: 365
DE(2, 3)	-X+10+DE(0, 2)	2Y-5+DE(0, 1)	959	②: 41
DE(2, 5)	X ² -3X+7+DE(0, 2)	4Y ² +2Y-1+DE(0, 5)	304	②: 564
C(2, 3)	C(-3X+7, 2)	C(2Y-5, 1)	923	②: 69
C(2, 3)	C(X ² -3X+7, 2)	C(3Y ² +2Y-5, 1)	445	②: 521
G(3, 1)	-3X+7+G(1, 2)	2Y-5+G(1, 1)	985	②: 15
G(3, 1)	X ² -3X+7+G(1, 2)	3Y ² +2Y-5+G(1, 1)	1000	

3-6.3 考慮(X,Y,Z)間的機率模型為二次關係之 Normal

在二次關係中，仍舊以 X、Y、Z 三個變數都來自常態分配為例，而誤差來自常態分配，依照上述四個步驟設法判定出有最大概似函數值的圖形結構；因為方法相同，所以部分推論過程則不列出。

底下分成四個步驟說明之。

Step 1：求出代表圖形結構 $X \rightarrow Y \rightarrow Z$ 的概似函數

首先假設資料服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX^2 + bX + c, \sigma_Y^2)$ ，
 $Z|Y \sim N(dY^2 + eY + f, \sigma_Z^2)$ ，由以上假設我們可得圖形之概似函數：

$$L_1(\underline{x}, \underline{y}, \underline{z} | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2) \\ = \left[(2\pi)^3 \sigma_X^2 \sigma_Y^2 \sigma_Z^2 \right]^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2}{2\sigma_Y^2}} e^{-\frac{\sum_{i=1}^n (z_i - dy_i^2 - ey_i - f)^2}{2\sigma_Z^2}}$$

將上式函數取自然對數再對參數作偏微，可得到參數之 MLE：

$$\hat{\mu}_X = \bar{X}, \quad \hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2, \\ \hat{a} = \frac{S_{XX}S_{X^2Y} - S_{XY}S_{XX^2}}{S_{XX}S_{X^2X^2} - (S_{XX^2})^2}, \quad \hat{b} = \frac{S_{X^2X^2}S_{XY} - S_{X^2Y}S_{X^2X}}{S_{X^2X^2}S_{XX} - (S_{X^2X})^2} \\ \hat{c} = \bar{Y} - \hat{a} \frac{1}{n} \sum_{i=1}^n x_i^2 - \hat{b} \bar{X} \\ \hat{\sigma}_Y^2 = \frac{1}{n} \sum_{i=1}^n \left[(y_i - \bar{Y}) - \hat{a} \left(x_i^2 - \frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \hat{b} (x_i - \bar{X}) \right]^2, \\ \hat{d} = \frac{S_{YY}S_{Y^2Z} - S_{YZ}S_{YY^2}}{S_{YY}S_{Y^2Y^2} - (S_{YY^2})^2}, \quad \hat{e} = \frac{S_{Y^2Y^2}S_{YZ} - S_{Y^2Z}S_{Y^2Y}}{S_{Y^2Y^2}S_{YY} - (S_{Y^2Y})^2} \\ \hat{f} = \bar{Z} - \hat{d} \frac{1}{n} \sum_{i=1}^n y_i^2 - \hat{e} \bar{Y} \\ \hat{\sigma}_Z^2 = \frac{1}{n} \sum_{i=1}^n \left[(z_i - \bar{Z}) - \hat{d} \left(y_i^2 - \frac{1}{n} \sum_{i=1}^n y_i^2 \right) - \hat{e} (y_i - \bar{Y}) \right]^2$$

再將 $(\hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$ 代回函數 $L_1(\underline{x}, \underline{y}, \underline{z} | \mu_X, \sigma_X^2, \sigma_Y^2, \sigma_Z^2)$ 而求得圖形結構 $X \rightarrow Y \rightarrow Z$ 之概似函數 $L_1(\underline{x}, \underline{y}, \underline{z} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$ 。

同樣的方式可以得到下頁所列五個圖形結構之概似函數：

$$X \rightarrow Z \rightarrow Y : L_1(\underline{x}, \underline{z}, \underline{y} | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$$

$$Y \rightarrow X \rightarrow Z : L_1(\underline{y}, \underline{x}, \underline{z} | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$$

$$Y \rightarrow Z \rightarrow X : L_1(\underline{y}, \underline{z}, \underline{x} | \hat{\mu}_Y, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2, \hat{\sigma}_X^2)$$

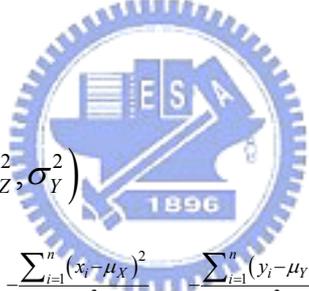
$$Z \rightarrow X \rightarrow Y : L_1(\underline{z}, \underline{x}, \underline{y} | \hat{\mu}_Z, \hat{\sigma}_Z^2, \hat{\sigma}_X^2, \hat{\sigma}_Y^2)$$

$$Z \rightarrow Y \rightarrow X : L_1(\underline{z}, \underline{y}, \underline{x} | \hat{\mu}_Z, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2, \hat{\sigma}_X^2)$$

Step 2：求出代表圖形結構 $X \rightarrow Z \leftarrow Y$ 的概似函數

假設資料服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y \sim N(\mu_Y, \sigma_Y^2)$ ，where $X \perp Y$ ，
 $Z | X, Y \sim N(aX^2 + bX + cY^2 + dY + e, \sigma_Z^2)$ ，由以上假設我們可以得到此

圖形之概似函數：



$$L_2(\underline{x}, \underline{z}, \underline{y} | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2) = \left[(2\pi)^3 \sigma_X^2 \sigma_Y^2 \sigma_Z^2 \right]^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} e^{-\frac{\sum_{i=1}^n (y_i - \mu_Y)^2}{2\sigma_Y^2}} e^{-\frac{\sum_{i=1}^n (z_i - ax_i^2 - bx_i - cy_i^2 - dy_i - e)^2}{2\sigma_Z^2}}$$

將上式函數取自然對數再對參數作偏微，可得到參數之 MLE：

$$\hat{\mu}_X = \bar{X} \quad , \quad \sigma_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad ,$$

$$\hat{\mu}_Y = \bar{Y} \quad , \quad \sigma_Y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{Y})^2 \quad ,$$

再由下列聯立方程組解出參數 $(\hat{a}, \hat{b}, \hat{c}, \hat{d})$ ：

$$\begin{cases} a \sum X_i^2 X_i^2 + b \sum X_i X_i^2 + c \sum Y_i^2 X_i^2 + d \sum Y_i X_i^2 = \sum Z_i X_i^2 \\ a \sum X_i^2 X_i + b \sum X_i X_i + c \sum Y_i^2 X_i + d \sum Y_i X_i = \sum Z_i X_i \\ a \sum X_i^2 Y_i^2 + b \sum X_i Y_i^2 + c \sum Y_i^2 Y_i^2 + d \sum Y_i Y_i^2 = \sum Z_i Y_i^2 \\ a \sum X_i^2 Y_i + b \sum X_i Y_i + c \sum Y_i^2 Y_i + d \sum Y_i Y_i = \sum Z_i Y_i \end{cases}$$

$$\hat{e} = \bar{Z} - \hat{a} \frac{1}{n} \sum_{i=1}^n x_i^2 - \hat{b} \frac{1}{n} \sum_{i=1}^n x_i - \hat{c} \frac{1}{n} \sum_{i=1}^n y_i^2 - \hat{d} \frac{1}{n} \sum_{i=1}^n y_i$$

$$\sigma_Z^2 = \frac{1}{n} \sum_{i=1}^n \left[(z_i - \bar{Z}) - \hat{d} \left(y_i^2 - \frac{1}{n} \sum_{i=1}^n y_i^2 \right) - \hat{e} (y_i - \bar{Y}) \right]^2$$

再將 $(\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$ 代回函數 $L_2(\underline{x}, \underline{z}, \underline{y} | \mu_X, \mu_Y, \sigma_X^2, \sigma_Z^2, \sigma_Y^2)$ 而求得圖形結構 $X \rightarrow Z \leftarrow Y$ 之概似函數 $L_2(\underline{x}, \underline{z}, \underline{y} | \hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$ 。

同樣的方式可以得到下列圖形結構之概似函數：

$$X \rightarrow Y \leftarrow Z : L_2(\underline{x}, \underline{y}, \underline{z} | \hat{\mu}_X, \hat{\mu}_Z, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$$

$$Y \rightarrow X \leftarrow Z : L_2(\underline{y}, \underline{x}, \underline{z} | \hat{\mu}_Y, \hat{\mu}_Z, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$$

Step 3：求出代表圖形結構 $Y \leftarrow X \rightarrow Z$ 的概似函數

假設資料分配服從 $X \sim N(\mu_X, \sigma_X^2)$ ， $Y|X \sim N(aX^2 + bX + c, \sigma_Y^2)$ ， $Z|X \sim N(dX^2 + eX + f, \sigma_Z^2)$ ，where $Y \perp Z|X$ ，由以上假設我們可得圖形之概似函數：

$$L_3(\underline{y}, \underline{x}, \underline{z} | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2)$$

$$= \left[(2\pi)^3 \sigma_X^2 \sigma_Y^2 \sigma_Z^2 \right]^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^n (x_i - \mu_X)^2}{2\sigma_X^2}} e^{-\frac{\sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2}{2\sigma_Y^2}} e^{-\frac{\sum_{i=1}^n (z_i - dx_i^2 - ex_i - f)^2}{2\sigma_Z^2}}$$

將上式函數取自然對數再對參數作偏微，可得到參數之 MLE：

$$\hat{\mu}_X = \bar{X} \quad , \quad \sigma_X^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad ,$$

$$\hat{a} = \frac{S_{XX} S_{X^2Y} - S_{XY} S_{XX^2}}{S_{XX} S_{X^2X^2} - (S_{XX^2})^2} \quad , \quad \hat{b} = \frac{S_{X^2X^2} S_{XY} - S_{X^2Y} S_{X^2X}}{S_{X^2X^2} S_{XX} - (S_{X^2X})^2}$$

$$\hat{c} = \bar{Y} - \hat{a} \frac{1}{n} \sum_{i=1}^n x_i^2 - \hat{b} \bar{X}$$

$$\sigma_Y^2 = \frac{1}{n} \sum_{i=1}^n \left[(y_i - \bar{Y}) - \hat{a} \left(x_i^2 - \frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \hat{b} (x_i - \bar{X}) \right]^2,$$

$$\hat{d} = \frac{S_{XX} S_{X^2Z} - S_{XZ} S_{XX^2}}{S_{XX} S_{X^2X^2} - (S_{XX^2})^2}, \quad \hat{e} = \frac{S_{X^2X^2} S_{XZ} - S_{X^2Z} S_{X^2X}}{S_{X^2X^2} S_{XX} - (S_{X^2X})^2}$$

$$\hat{f} = \bar{Z} - \hat{d} \frac{1}{n} \sum_{i=1}^n y_i^2 - \hat{e} \bar{Y}$$

$$\sigma_Z^2 = \frac{1}{n} \sum_{i=1}^n \left[(z_i - \bar{Z}) - \hat{d} \left(x_i^2 - \frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \hat{e} (x_i - \bar{X}) \right]^2,$$

再將 $(\hat{\mu}_X, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$ 代回函數 $L_3(y, x, z | \mu_X, \sigma_Y^2, \sigma_X^2, \sigma_Z^2)$ 而求得圖形結構 $Y \leftarrow X \rightarrow Z$ 之概似函數 $L_3(y, x, z | \hat{\mu}_X, \hat{\sigma}_Y^2, \hat{\sigma}_X^2, \hat{\sigma}_Z^2)$ 。

同樣的方式可以得到下列圖形結構之概似函數：

同樣的方式可以得到下列圖形結構之概似函數：

$$X \leftarrow Y \rightarrow Z : L_3(x, y, z | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$$

$$X \leftarrow Z \rightarrow Y : L_3(x, z, y | \hat{\mu}_X, \hat{\sigma}_X^2, \hat{\sigma}_Z^2, \hat{\sigma}_Y^2)$$

Step 4：判斷圖形結構何者的可能性最大？

上面三個步驟中，我們仍舊可以整理出式子：

$$L_i(x, y, z | \hat{\theta}_{MLE}) = (2\pi\hat{\sigma}_X^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} (2\pi\hat{\sigma}_Y^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} (2\pi\hat{\sigma}_Z^2)^{-\frac{n}{2}} e^{-\frac{n}{2}} \quad i=1,2,3$$

因此和線性的做法相同，為了找出概似函數 $\left[(2\pi e)^3 (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 \hat{\sigma}_Z^2) \right]^{-\frac{n}{2}}$ 的最大值，我們需找出 $\left[(2\pi e)^3 (\hat{\sigma}_X^2 \hat{\sigma}_Y^2 \hat{\sigma}_Z^2) \right]$ 的最小值。

在二次關係的程式模擬過程中，通常都可以將變數間為二次關係的因果關係之圖形結構判斷出來。比如二次關係的資料之因果關係為： $X \rightarrow Y \rightarrow Z$ 時，通常模擬出來的概似函數中，我

們可以很容易的發現圖形結構 $X \rightarrow Y \rightarrow Z$ 相對應的值會“最小”，亦即其概似函數的值會是最大的。即使我們使用不同的分配去模擬，結果在二次關係時仍舊是不錯的。

3-6.4 模擬五種分配資料，以二次關係之 Normal 公式做判斷

底下我們亦將模擬在不同的資料模型中，若仍用上小節中資料為 Normal 分配所導出之概似函數，是否可以判斷出正確的圖形結構？只要當實際資料的圖形結構之概似函數值比所有其它之圖形結構函數值大即為判斷正確。在底下的模擬過程中，樣本數皆設定為 100，重覆試驗 1000 次，而原始資料之圖形結構為 $X \rightarrow Y \rightarrow Z$ ，表中第四欄為判斷正確圖形結構的次數，第五欄列出判斷成其它圖形結構的次數僅列出次數最高的類型及判斷次數，共有底下幾種可能之類別；類型①： $Z \rightarrow Y \rightarrow X$ 、類型②： $X \leftarrow Y \rightarrow Z$ 、類型③： $X \rightarrow Z \rightarrow Y$ 、類型④： $Y \rightarrow Z \rightarrow X$ 、類型⑤： $Y \rightarrow Z \rightarrow X$ 、類型⑥： $Y \rightarrow X \rightarrow Z$ 、類型⑦： $Z \rightarrow X \rightarrow Y$ 、類型⑧： $Y \leftarrow X \rightarrow Z$ 及類型⑨： $X \rightarrow Y \leftarrow Z$ 。

從下頁表發現，在上述之資料模型中，若資料中變數間的關係與判斷公式中變數間的關係同為二次關係時，幾乎都可以判斷出正確的圖形結構 $X \rightarrow Y \rightarrow Z$ ，僅在模型結構為 Cauchy 時判斷正確率約為 0.988。但當資料中變數關係為線性而判斷公式中

變數間的關係為二次時，僅有在模型結構為 Exponential 時，判斷正確率為 0.931，其它的模型結構皆小於 0.5，判斷正確的比率仍舊不高。

資料 X	資料 Y	資料 Z	正確	類型
N(2, 5)	N(-3X+7, 2)	N(2Y-1, 5)	398	①:382
N(2, 5)	N(X ² -3X+7, 2)	N(4Y ² +2Y-1, 5)	1000	
Exp(3)	2X+2+Exp(5)	4Y+6+Exp(2)	931	①: 65
Exp(3)	2X ² +2X+2+Exp(5)	3Y ² +4Y+6+Exp(2)	1000	
DE(2, 3)	-X+10+DE(0, 2)	2Y-5+DE(0, 1)	246	①:419
DE(2, 3)	2X ² -X+10+DE(0, 2)	3Y ² +2Y-5+DE(0, 1)	1000	
C(2, 3)	C(-3X+7, 2)	C(2Y-5, 1)	41	①:709
C(2, 3)	C(X ² -3X+7, 2)	C(3Y ² +2Y-5, 1)	988	④: 5
G(3, 1)	-3X+7+G(1, 2)	2Y-5+G(1, 1)	80	①:565
G(3, 1)	X ² -3X+7+G(1, 2)	3Y ² +2Y-5+G(1, 1)	1000	

綜合以上，我們從模擬的結果可以發現，模型改變會影響我們的判斷結果；當資料與判斷公式有相同模型或變數間的關係相同，不論同為線性或同為二次時，我們幾乎都可以判斷出確切的結果；只是當變數間關係為線性的時候，我們有時需要再藉由一些其它的資訊幫助分析。當模型或變數間的關係皆不相同時，判斷正確的比率相對的減少很多。由此可見，在做分析之前，我們最好要做好 Model Checking 的動作，以期能得到更佳的分析結果。

第四章 結論

分析出資料的因果關係及判斷出圖形的方向性有很多方法，尤其以貝氏網路的方法最受矚目，但難度比較高。在此我們探討使用貝氏概念及概似函數的判斷準則，及模型假設對結果的影響；當資料中的變數關係與判斷公式中的變數關係一致時，我們幾乎都可以判斷出圖形的方向性；而資料中的變數若為線性關係時，需再藉由其它的資訊，如：專家的知識經驗、時間的先後順序等等才能更準確的判斷出其圖形結構。

在分析的過程中，我們可以發現，當誤差太大時，會混淆資料的資訊而影響我們判斷的準確度；因此，如何設法消除那些影響我們判斷的雜訊，或將雜訊所代表的意義分析出來，亦是一個值得探討的問題。

此外，在本文中，我們只有探討兩個變數及三個變數間的因果關係，如何延伸到很多個變數及複雜化變數間的關係，甚至我們是否可以將變數分類，使得每一類當中只有小數量的變數，再將各類變數做結合，進而推廣到大數量變數的因果關係，亦是一個可以深入研究的問題。

參考文獻

- Arjas, E., Parner, J. (2004). Causal reasoning form longitudinal data. *Scandinavian Journal of Statistics*, Vol. 31: pp. 171-187.
- D'haeseleer, P., Wen, X., Fuhman, S. and Somogyi. (1999). R. Linear modeling of mRNA expression levels during CNS development and injury. In *Proc. of the Pacific Symp. on Biocomputation*,. 4:41-52.
- Friedman, N., Goldszmidt, and Wyner, N. (1999). Data analysis with Bayesian networks: A bootstrap approach. In *UAI'99*, pp.206-215.
- Friedman, N. and Koller, D. (2001). Being Bayesian about Bayesian network structure: A Bayesian approach to structure discovery in Bayesian networks. *Machine Learning*. To appear. Earlier version appeared in UAI'00.
- Friedman, N., Lineal, M., Nachman, I., and Pe'er, D. (2000). Using Bayesian networks to analyze expression data. Accepted to *Journal of Computational Biology*, 7, pp.601-620.
- Heckerman, D. (1995). A Bayesian approach to learning causal networks. In *Proc. of Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 285-295. Also appears as Technical Report MSR-TR-95-04, Microsoft Advanced Technology Division, Microsoft Corporation, Seattle, Washington.
- Judea Pearl. (2001). *Causality: Models, Reasoning, and Inference*. Cambridge.
- Kauffman, S. (1993). *The origins of order. self-organization and selection in evolution*. Oxford University Press.
- Lauritzen, S., L. (2004). Discussion on causality. *Scandinavian Journal of Statistics*, Vol. 31: pp.189-201.
- Meek, C. (1995). Causal inference and causal explanation with background knowledge. In P. Bensnard and S. Hanks (Eds.), *Uncertainty in Artificial Intelligence*, Vol. 11, pp. 403-410. San Francisco: Morgan Kaufmann.

- Murphy, K. and Mian, S. (1999). Modeling gene expression data using dynamic Bayesian networks. Technical Report, University of California at Berkeley, Department of Computer Science.
- Pearl, J., and Verma, T. (1991). A theory of inferred causation. In J. A. Allen, R. Fikes, and E. Sandewall (Eds.), *Principles of Knowledge Representation and Reasoning: Proc. of the second International Conference*, pp. 441-52. San Mateo, CA: Morgan Kaufmann.
- Pe'er D., Regev A., Elidan G., Friedman N. (2001). Inferring subnetworks from perturbed expression profiles. *Bioinformatics* 17 Suppl. 1: S215-S224.
- Perrin, B., Ralaivola, L., Mazurie, A., Bottani, S. Mallet, J. & D'Alche-Buc, F. (2003). Gene networks inference using dynamic Bayesian networks. *Bioinformatics*, Vol. 19, *Suppl. 2*: ii138-ii148.
- Samuel Kotz and Norman L. Johnson. (1989). *Encyclopedia of statistical sciences*, Vol. 4. New York: Wiley.
- Somogyi, R. and Sniegowski, C. (1996). Modeling the complexity of genetic networks: understanding multigenetic and pleiotropic regulation. *Complexity*, 1, 45-63.
- Spirtes, P. and Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review* 9: 62~72.
- Spirtes, P., Glymour C., Scheines R. Kauffman, S., Aimale, V. and Wimberly, F. (2000). Constructing Bayesian network models of gene expression networks from Microarray data. *Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems and Technology*.
- Verma, T., and Pearl, J. (1990). Equivalence and synthesis of causal models. In *Proceedings of the 6th Conference on Uncertainty in Artificial Intelligence* (July, Cambridge, MA), pp.220-227. Reprinted in P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence*, Vol. 6, pp. 255-268. Amsterdam: Elsevier.
- Weaver, C., Workman, C., and Stormo, G. (1999). Modeling regulatory networks with weight matrices. In *Proc. of the Pacific Symp. on Biocomputing*, 4:112-123.

呂傑瀚(2003) 判斷兩變間的因果關係，交通大學統計所碩士論文

鄭佳瑜(2003) 用無母數迴歸法分析變數間之因果關係，交通大學統計所
碩士論文

鄭新慶(2004) 台灣醫療支出成長因素之探討－時間序列與有向圖之應用，
逢甲大學會計與財稅所碩士論文



附 錄

底下的數據是以 $X, error \sim N(0,1)$ where $X \perp error$ 為根據所模擬出來的結果，模型組合方式有很多種，一一列於表格上方，若依誤差的不同共分成五大類，列為如下：

- (I) 誤差用加的；
- (II) 誤差用乘的，變異數不變；
- (III) 誤差用乘的，變異數改變；
- (IV) 誤差為指數次方，變異數不變；
- (V) 誤差為指數次方，變異數改變；

當中每個表格共四欄，每欄的內容如下：

第一欄：型一誤差， $\alpha = 0.02, 0.05,$ 及 0.10 三種；

第二欄：當樣本數為表格中所列之大小時， p -value 的期望值會小於所設定之型一誤差；

第三欄：當樣本數為表格中所列之大小時，在所設定的型一誤差下，型二誤差（模型假設列於各表格上方）的期望值會小於所設定之 α 值；

第四欄：第二欄及第三欄所得之 Kendall's Tau 平均值。

(I) 誤差用加的；

(1) 當誤差為變數的十倍時 $Y = 0.1X + \text{error}$

0.02	700	2000	0.064
0.05	600	1500	0.064
0.10	400	1000	0.063

(2) 當誤差為變數的五倍時 $Y = 0.2X + \text{error}$

0.02	170	600	0.124
0.05	130	400	0.126
0.10	90	300	0.129

(3) 當誤差為變數的兩倍時 $Y = 0.5X + \text{error}$

0.02	30	100	0.297
0.05	30	70	0.296
0.10	20	50	0.301

(4) 當誤差和變數一樣大時 $Y = X + \text{error}$

0.02	15	35	0.502
0.05	10	25	0.519
0.10	10	15	0.521

(5) 當變數為誤差的二倍時 $Y = 2.0X + \text{error}$

0.02	10	15	0.706
0.05	10	15	0.694
0.10	5	10	0.726

(6) 當變數為誤差的三倍時 $Y = 3.0X + \text{error}$

0.02	10	15	0.795
0.05	10	10	0.789
0.10	5	10	0.814

(7) 當變數為誤差的四倍時 $Y = 4.0X + \text{error}$

0.02	10	10	0.850
0.05	10	10	0.830
0.10	5	10	0.861

(8) 當變數為誤差的五倍時 $Y = 5.0X + \text{error}$

0.02	10	10	0.876
0.05	5	5	0.866
0.10	5	5	0.891

(II) 誤差用乘的，僅平移，變異數不變；

(1) 一樣大，平移 0 $Y = X \times |\text{error}|$

0.02	10	15	0.760
0.05	10	15	0.752
0.10	5	10	0.749

(2) 一樣大，平移 0.2 $Y = X \times (0.2 + \text{error})$

0.02	150	600	0.139
0.05	90	400	0.144
0.10	70	300	0.150

(3) 一樣大，平移 0.5 $Y = X \times (0.5 + \text{error})$

0.02	25	100	0.344
0.05	20	70	0.359
0.10	15	40	0.366

(4) 一樣大，平移 1.0 $Y = X \times (1.0 + \text{error})$

0.02	10	30	0.607
0.05	10	20	0.603
0.10	10	15	0.615

(5) 一樣大，平移 2.0 $Y = X \times (2.0 + \text{error})$

0.02	10	10	0.825
0.05	5	10	0.818
0.10	5	10	0.822

(6) 一樣大，平移 3.0 $Y = X \times (3.0 + \text{error})$

0.02	10	10	0.889
0.05	5	10	0.888
0.10	5	10	0.884

(7) 一樣大，平移 4.0 $Y = X \times (4.0 + \text{error})$

0.02	10	10	0.914
0.05	5	10	0.914
0.10	5	5	0.918

(III) 誤差用乘的，不僅平移，變異數亦改變；

(1) 誤差 var=0.75，平移 0 $Y = X \times \sqrt{0.75}|\text{error}|$

0.02	10	10	0.725
0.05	10	10	0.753
0.10	5	10	0.769

(2) 誤差 var=0.75，平移 0.2 $Y = X \times (0.2 + \sqrt{0.75} \text{error})$

0.02	110	500	0.159
0.05	70	300	0.168
0.10	50	190	0.175

(3) 誤差 var=0.75，平移 0.5 $Y = X \times (0.5 + \sqrt{0.75} \text{error})$

0.02	20	90	0.383
0.05	15	50	0.403
0.10	10	30	0.425

(4) 誤差 var=0.75，平移 1.0 $Y = X \times (1.0 + \sqrt{0.75} \text{error})$

0.02	10	25	0.633
0.05	10	15	0.655
0.10	5	10	0.683

(5) 誤差 var=0.75，平移 2.0 $Y = X \times (2.0 + \sqrt{0.75} \text{error})$

0.02	10	15	0.822
0.05	5	10	0.842
0.10	5	10	0.835

(6) 誤差 var=0.75, 平移 3.0 $Y = X \times (3.0 + \sqrt{0.75} \text{ error})$

0.02	10	10	0.887
0.05	5	10	0.902
0.10	5	10	0.887

(IV) 誤差為指數次方, 變異數不變;

(1) 平移 0 $Y = X \times \exp^{\text{error}}$

0.02	10	10	0.723
0.05	10	10	0.743
0.10	5	10	0.718

(2) 平移 0.2 $Y = X \times (0.2 + \exp^{\text{error}})$

0.02	10	15	0.755
0.05	10	10	0.779
0.10	5	10	0.727

(3) 平移 0.5 $Y = X \times (0.5 + \exp^{\text{error}})$

0.02	10	15	0.799
0.05	5	10	0.815
0.10	5	10	0.766

(V) 誤差為指數次方, 變異數改變;

(1) 誤差 variance=0.4 $Y = X \times \exp^{(0.2 \times \text{error})}$

0.02	7	9	0.926
0.05	5	10	0.933
0.10	5	5	0.906

(2) 誤差 variance=2 $Y = X \times \exp(\sqrt{2} \times \text{error})$

0.02	9	15	0.688
0.05	10	15	0.703
0.10	7	9	0.663

(3) 誤差 variance=4 $Y = X \times \exp(2 \times \text{error})$

0.02	9	17	0.643
0.05	10	15	0.650
0.10	7	11	0.624

(4) 誤差 variance=9 $Y = X \times \exp(3 \times \text{error})$

0.02	11	19	0.599
0.05	10	15	0.598
0.10	5	5	0.585

(5) 誤差 variance=16 $Y = X \times \exp(4 \times \text{error})$

0.02	11	19	0.585
0.05	10	20	0.578
0.10	7	13	0.572

(6) 誤差 variance=25 $Y = X \times \exp(5 \times \text{error})$

0.02	11	23	0.566
0.05	10	20	0.559
0.10	7	13	0.559

