Enabling A+ Decisions®
DALab Proprietary

CH7簡單貝氏分類法與貝氏網路

授課老師：簡禎富 講座教授

資料挖礦與大數據分析
Data Mining & Big Data Analytics

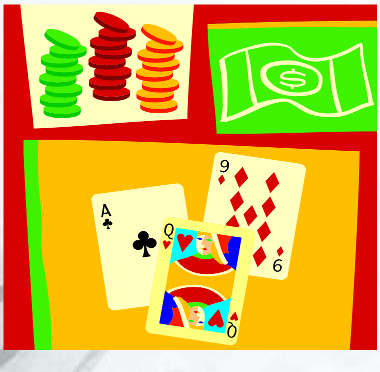
Enabling A+ Decisions®
DALab Proprietary

大綱

- 貝氏定理
- 簡單貝氏分類法
- 貝式網路
- 應用實例──台電饋線事故定位系統
- 結論

討論

- 賭場裡常有21點(Black Jack)，若莊家的牌面是8點，而自己是16點，是否會選擇補牌？
- 若你是補牌順序比較後面的閒家，前面已出現的補牌會怎麼影響你的決策？



Enabling A+ Decisions®
DALab Proprietary

貝氏分類(Bayesian Classifier)

- 藉由資料中分析屬性與反應變數之間的機率模型，根據貝氏定理(Bayes' Theorem)來更新資訊以推理判斷樣本資料歸屬的類別，作為分類和推論的依據
- 常用方法
  - 簡單貝氏分類法 (Naïve Bayesian Classifier)
  - 貝氏網路分類法 (Bayesian Network Classifier，簡稱貝氏網路)
- 面對沒有經驗、可參考的資訊過少或者沒有頻率機率存在的情況，貝氏網路採用主觀機率(subjective probability)，將認為該事件是否會發生的信心程度(degree of belief)的主觀判斷轉為主觀機率

貝氏定理(1/2)

- 根據新的資訊將事前機率(prior probability)修正為事後機率(posterior probability)的過程
  - 事前機率：尚未取得樣本資訊前，對事件原始可能發生的機率
  - 事後機率：根據取得樣本資訊後，修改事件可能發生的機率
  - 條件機率 (conditional probability)：根據某一事件發生的情況下，估計另一事件發生的機率
- 主要概念：一開始不知道目標事件 $\tilde{\theta}$ 的真實狀態，但知道 $\tilde{\theta}$ 服從機率分布 $P(\tilde{\theta})$，稱為事前機率。當得到新的樣本資訊或證據 ***E*** 後，可以根據貝氏定理，更新事後機率 $P(\tilde{\theta}|E)$

$$P(\tilde{\theta}=\theta_j|E)=P(\theta_j|E)=\frac{P(\theta_j\cap E)}{P(E)}=\frac{P(E|\theta_j)\times P(\theta_j)}{\sum_{j=1}^{m}P(E|\theta_j)\times P(\theta_j)}$$

$$P(\theta_j\cap E)=P(\theta_j|E)\times P(E)=P(E|\theta_j)\times P(\theta_j)$$

Enabling A+ Decisions®
DALab Proprietary

貝氏定理(2/2)

- 若$\theta_1,\theta_2,\ldots,\theta_m$ 為假設 $\tilde{\theta}$ 在樣本空間**S**中的一個分割，且事件 $E\subset S$ $P(\theta_j)\neq 0,\ j=1,\cdots,m$，則根據總和機率定理

$$\begin{aligned}P(E)&=P(E|\theta_1)\cdot P(\theta_1)+P(E|\theta_2)\cdot P(\theta_2)+\cdots+P(E|\theta_m)\cdot P(\theta_m)\\&=\sum_{j=1}^{m}P(E|\theta_j)\cdot P(\theta_j)\end{aligned}$$

- 在取得的新資訊事件**E**下，貝氏定理可修正假設 $\tilde{\theta}=\theta_j$的事前機率為事後機率 $P(\theta_j|E)$

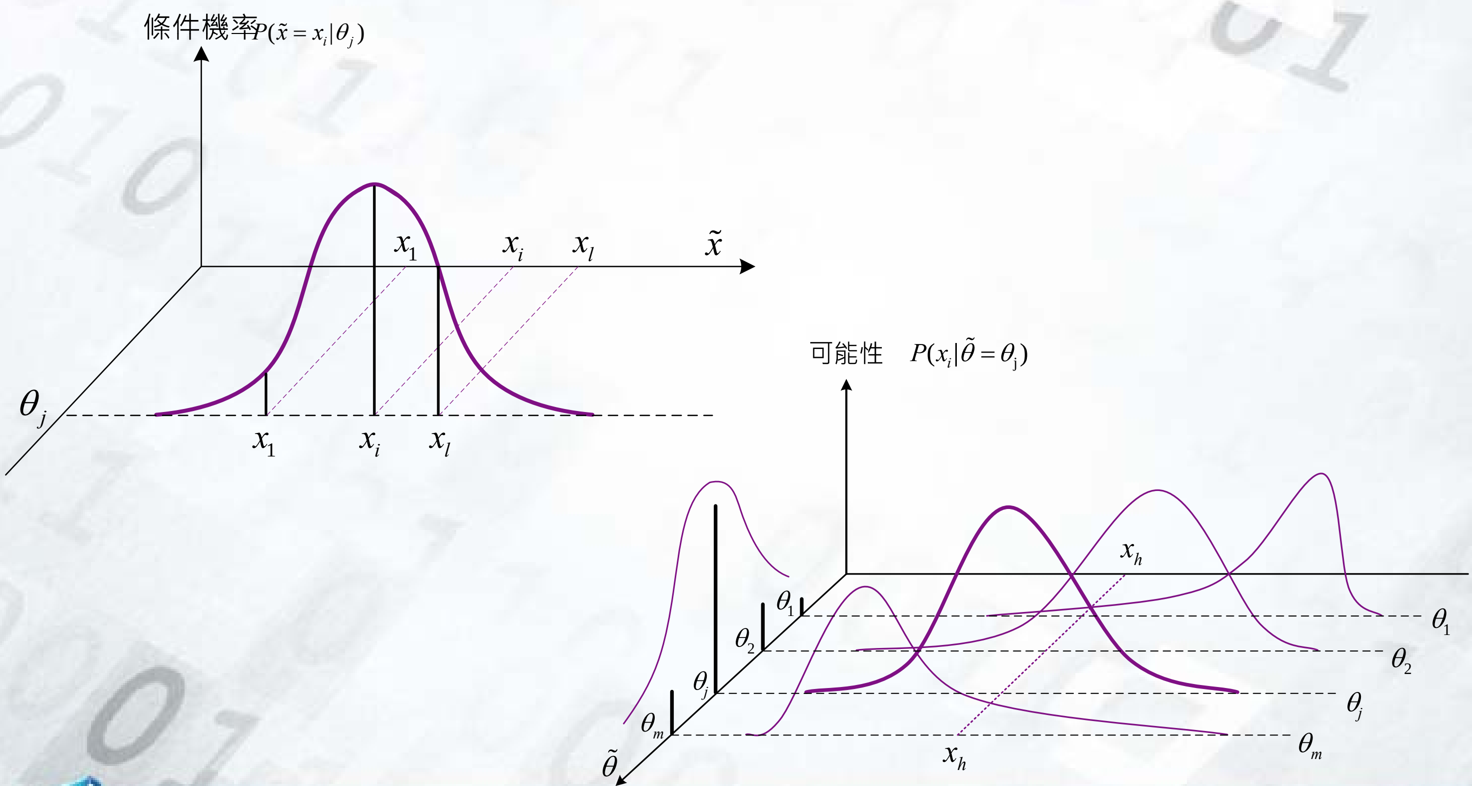
$$P(\theta_j|E)=\frac{P(E|\theta_j)\cdot P(\theta_j)}{\sum_{j=1}^{m}P(E|\theta_j)\cdot P(\theta_j)}$$

概似函數(likelihood function)

- **概似函數**定義狀態空間與樣本空間的關係
- **決策法則**(decision rule)定義樣本空間與行動空間的關係
- **損失函數**(loss function)定義行動空間與狀態空間不同組合下的收益或損失

- 概似函數 $P_{\tilde{\theta}}(x)$ 代表觀察到 x_i 時有多少可能性是來自於隨機變數 $\tilde{\theta} = \theta_{\mathrm{j}}$ 的情況
- $P_{\tilde{\theta}}(x)$ 越高則決策者觀察到樣本 x_i 後，對真實狀態為 θ_j 的信心(belief)越高

條件機率 vs 概似函數



Enabling A+ Decisions®
DALab Proprietary

[範例7.1]貝氏定理計算(1/2)

- 若某品牌手機主要由A、B兩家工廠生產，而工廠A的生產量為工廠B的4倍，且已知工廠A的良率為15/16 ，工廠B的良率為3/4

$$P(\text{良品}\mid\text{工廠A所生產}) = \frac{750}{800}$$

$$P(\text{良品}\mid\text{工廠B所生產}) = \frac{150}{200}$$

- 試求當檢驗結果為不良品時，該不良品來自於工廠A的可能性？來自工廠B的可能性？

Enabling A+ Decisions®
DALab Proprietary

[範例7.1]貝氏定理計算(2/2)

$$P(E=\text{工廠A所生產}\mid\tilde{\theta}=\text{不良品}) = \frac{P(\text{不良品且工廠A})}{P(\text{不良品且工廠A})+P(\text{不良品且工廠B})}$$

$$= \frac{P(\text{不良品}|\text{工廠A})P(\text{工廠A})}{P(\text{不良品}|\text{工廠A})P(\text{工廠A})+P(\text{不良品}|\text{工廠B})P(\text{工廠B})}$$

$$= \frac{\frac{50}{800}\times\frac{800}{1000}}{\frac{50}{800}\times\frac{800}{1000}+\frac{50}{200}\times\frac{200}{1000}} = 0.5$$

$$P(E=\text{工廠B所生產}\mid\tilde{\theta}=\text{不良品}) = \frac{P(\text{不良品且工廠B})}{P(\text{不良品且工廠B})+P(\text{不良品且工廠A})}$$

$$= \frac{P(\text{不良品}|\text{工廠B})P(\text{工廠B})}{P(\text{不良品}|\text{工廠B})P(\text{工廠B})+P(\text{不良品}|\text{工廠A})P(\text{工廠A})}$$

$$= \frac{\frac{50}{200}\times\frac{200}{1000}}{\frac{50}{200}\times\frac{200}{1000}+\frac{50}{800}\times\frac{800}{1000}} = 0.5$$

→ 當檢驗結果為不良品時，該不良品來自於工廠A的可能性有0.5，來自工廠B的可能性有0.5

簡單貝氏分類法 (Naïve Bayesian classification)

Enabling A+ Decisions®
DALab Proprietary

- 又稱單純貝氏分類法，有兩項基本假設：
  1. 已知各類別的事前機率，常依據專家意見、歷史資料或訓練資料設定
  2. 給定任一類別下，屬性資料相互獨立，即屬性資料條件獨立
- 當預測資料集不包含屬性資料時，只能依據事前機率預測觀察值屬於何種類別
- 當預測資料集包含屬性資料時，可建立各分類之條件機率模型，再利用屬性資料與貝式定理，算出每筆屬性資料屬於各分類的事後機率
- 能進行高維度資料分類，並快速建構可用於分類和預測的資料挖礦模型

簡單貝氏分類的步驟(1/2)

Enabling A+ Decisions®
DALab Proprietary

- 假設一訓練資料集包含***n***筆資料，$i=1,2,\cdots,n$，其中有***m***個類別 $\tilde{\theta}=\{\theta_1,\theta_2,\cdots,\theta_m\}$，其對應事前機率為 $P(\tilde{\theta}=\theta_j)$，$j=1,2,\cdots,m$

  定義第***i***筆資料中***k***個屬性的觀察值為 $\mathbf{E}_i=\{E_{i1},E_{i2},\cdots,E_{ik}\}$

  令 $\mathbf{E}=\{\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n\}^T$ 代表訓練資料集的所有屬性資料
- 簡單貝氏分類法利用最大化各類別的條件機率分布 $P(\mathbf{E}\,|\,\tilde{\theta}=\theta_j)$ 再利用資料集的屬性資料 $\mathbf{E}^*=(\mathbf{E}_1^*,\mathbf{E}_2^*,\cdots,\mathbf{E}_k^*)$ 與貝式定理算出各分類的事後機率

$$P(\theta_j\,|\,\mathbf{E}^*)=\frac{P(\mathbf{E}^*\,|\,\theta_j)\times P(\theta_j)}{P(\mathbf{E}^*)},\quad j=1,2,\cdots,m$$

簡單貝氏分類的步驟(2/2)

- 由假設(1)得到 $P(\theta_j \mid \mathbf{E}^*) > P(\theta_s \mid \mathbf{E}^*),\ j = 1,2,\cdots,m,\ j \neq s$
- 由假設(2)的條件獨立得

$$P(\mathbf{E}^* \mid \theta_j) = P(E_1^*, E_2^*,, \cdots, E_k^* \mid \theta_j)$$
$$= P(E_1^* \mid \theta_j) \times P(E_2^* \mid \theta_j) \times \cdots \times P(E_k^* \mid \theta_j)$$
$$= \prod_{l=1}^{k} P(E_l^* \mid \theta_j)$$

- 事後機率

$$P(\theta_j \mid \mathbf{E}^*) = \frac{\prod_{l=1}^{k} P(E_l^* \mid \theta_j) \times P(\theta_j)}{\sum_{j=1}^{m} \prod_{l=1}^{k} P(E_l^* \mid \theta_j) \times P(\theta_j)}$$

- 相依變數類別

$$P(E_l \mid \theta_j) = \frac{r_{lj}}{m_j}$$

簡單貝氏分類計算範例(1/5)

- 不動產公司蒐集了10筆顧客資料，包括三個類別屬性，目標變數為是否有購買不動產；假設 θ_1 表示有購買不動產，θ_2 代表沒有購買不動產。若想預測一位新客戶有無購買不動產

$$\mathbf{E}^* = (\text{婚姻} E_1^* = \text{已婚、年齡層 } E_2^* = \text{中年、收入 } E_3^* = \text{高})$$

ID (i)	婚姻 E_1^*	年齡層 E_2^*	收入 E_3^*	購買不動產決策變數 $\tilde{\theta}$
001	已婚	青年	低	有
002	已婚	中年	高	無
003	單身	中年	高	無
004	單身	青年	高	有
005	已婚	中年	中	有
006	單身	中年	低	有
007	單身	青年	高	無
008	已婚	青年	高	無
009	已婚	中年	高	有
010	已婚	青年	高	有

Enabling A+ Decisions®
DALab Proprietary

簡單貝氏分類計算範例(2/5)

- 事前機率 $P(\tilde{\theta})$

$$P(\theta_1)=P(\text{有購買不動產})=6/10=0.60$$
$$P(\theta_2)=P(\text{無購買不動產})=4/10=0.40$$

→無任何其他資訊時，可合理猜測來訪的顧客，有購買不動產的機率為**0.6**

- 若加上屬性的訊息，可得 $P(\mathbf{E}^*|\theta_j)$的條件機率，以下先考慮僅有婚姻屬性預測該顧客是否已經購買不動產：

$$P(\text{購買不動產}=\text{有}|\text{婚姻}=\text{已婚})$$
$$=\frac{P(\text{已婚}|\text{有購買})P(\text{有購買})}{P(\text{已婚}|\text{有購買})P(\text{有購買})+P(\text{已婚}|\text{無購買})P(\text{無購買})}$$
$$=\left(\frac{4}{6}\times\frac{6}{10}\right)\bigg/\left(\frac{4}{6}\times\frac{6}{10}+\frac{2}{4}\times\frac{4}{10}\right)=0.67$$

$$P(\text{購買不動產}=\text{無}|\text{婚姻}=\text{已婚})$$
$$=\frac{P(\text{已婚}|\text{無購買})P(\text{無購買})}{P(\text{已婚}|\text{有購買})P(\text{有購買})+P(\text{已婚}|\text{無購買})P(\text{無購買})}$$
$$=\left(\frac{2}{4}\times\frac{4}{10}\right)\bigg/\left(\frac{4}{6}\times\frac{6}{10}+\frac{2}{4}\times\frac{4}{10}\right)=0.33$$

→發現考慮該顧客已經結婚下，推測可能已有購買不動產($0.67>0.33$)

Enabling A+ Decisions®
DALab Proprietary

簡單貝氏分類計算範例(3/5)

- 考慮加入其他顧客的屬性資料（婚姻、年齡層、收入），再以簡單貝氏分類法計算其是否已經購買不動產

$$P(\mathbf{E}^*|\tilde{\theta}=\theta_1)=\frac{P(\text{婚姻}=\text{已婚、年齡}=\text{中年、收入}=\text{高})}{P(\text{購買不動產}=\text{有})}=\frac{1/10}{6/10}=\frac{1}{6}$$

$$P(\mathbf{E}^*|\tilde{\theta}=\theta_2)=\frac{P(\text{婚姻}=\text{已婚、年齡}=\text{中年、收入}=\text{高})}{P(\text{購買不動產}=\text{無})}=\frac{1/10}{4/10}=\frac{1}{4}$$

- 由表可推估以下機率

$$P(E_1^*=\text{已婚}|\tilde{\theta}=\theta_1)=P(\text{婚姻}=\text{已婚}|\text{購買不動產}=\text{有})=4/6$$
$$P(E_1^*=\text{已婚}|\tilde{\theta}=\theta_2)=P(\text{婚姻}=\text{已婚}|\text{購買不動產}=\text{無})=2/4$$
$$P(E_2^*=\text{中年}|\tilde{\theta}=\theta_1)=P(\text{年齡層}=\text{中年}|\text{購買不動產}=\text{有})=3/6$$
$$P(E_2^*=\text{中年}|\tilde{\theta}=\theta_2)=P(\text{年齡層}=\text{中年}|\text{購買不動產}=\text{無})=2/4$$
$$P(E_3^*=\text{高}|\tilde{\theta}=\theta_1)=P(\text{收入}=\text{高}|\text{購買不動產}=\text{有})=3/6$$
$$P(E_3^*=\text{高}|\tilde{\theta}=\theta_2)=P(\text{收入}=\text{高}|\text{購買不動產}=\text{無})=4/4$$

簡單貝氏分類計算範例(4/5)

- 若假設三個屬性間為條件獨立，根據以上的條件機率，可預測該顧客是否有購買不動產的計算結果如下：

$$
\begin{aligned}
&P(\mathbf{E}^* \mid \tilde{\theta}=\theta_1) = P(\mathbf{E}^* \mid \text{購買不動產}=\text{有}) \\
&\propto P(\text{婚姻}=\text{已婚} \mid \text{購買不動產}=\text{有}) \times \\
&\quad P(\text{年齡}=\text{中年} \mid \text{購買不動產}=\text{有}) \times \\
&\quad P(\text{收入}=\text{高} \mid \text{購買不動產}=\text{有}) \\
&\quad = \frac{4}{6} \times \frac{3}{6} \times \frac{3}{6} = \frac{1}{6}
\end{aligned}
$$

$$
\begin{aligned}
&P(\mathbf{E}^* \mid \tilde{\theta}=\theta_2) = P(\mathbf{E}^* \mid \text{購買不動產}=\text{無}) \\
&\propto P(\text{婚姻}=\text{已婚} \mid \text{購買不動產}=\text{無}) \times \\
&\quad P(\text{年齡}=\text{中年} \mid \text{購買不動產}=\text{無}) \times \\
&\quad P(\text{收入}=\text{高} \mid \text{購買不動產}=\text{無}) \\
&\quad = \frac{2}{4} \times \frac{2}{4} \times \frac{4}{4} = \frac{1}{4}
\end{aligned}
$$

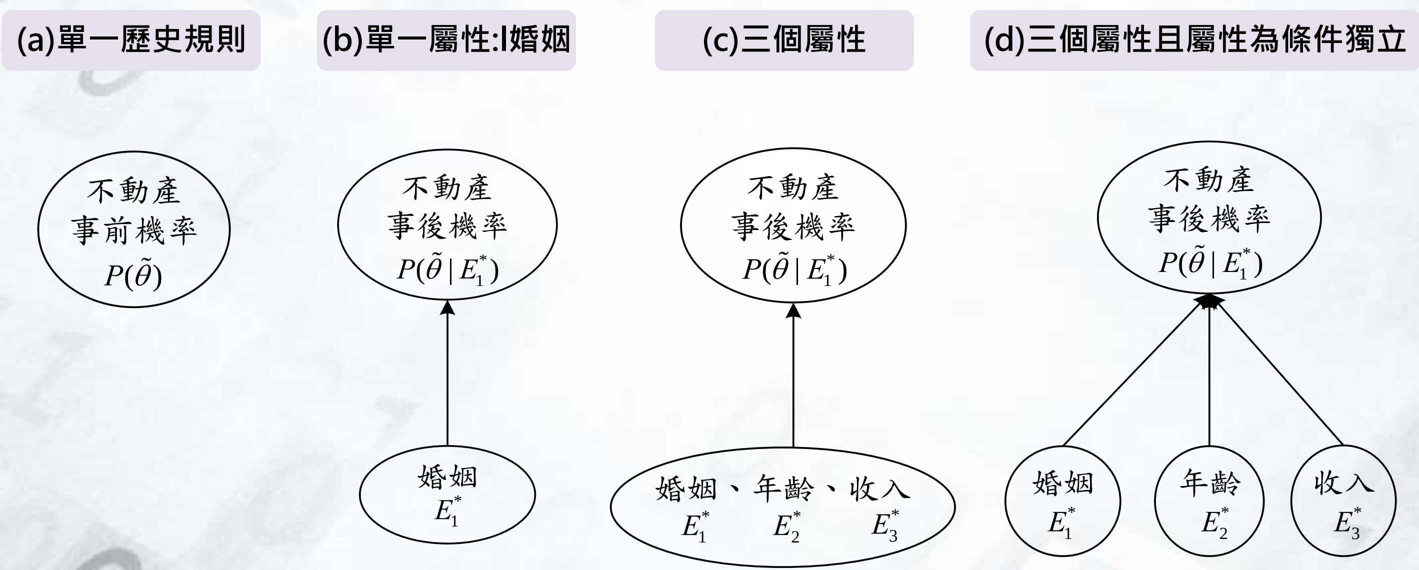
→比較上述結果，可檢驗在給定有無購買不動產下，三個屬性間為條件獨立

- 事後機率

$$P(\tilde{\theta}=\theta_1 \mid \mathbf{E}^*) = \frac{P(\mathbf{E}^* \mid \tilde{\theta}=\theta_1)P(\tilde{\theta}=\theta_1)}{P(\mathbf{E}^* \mid \tilde{\theta}=\theta_1)P(\tilde{\theta}=\theta_1) + P(\mathbf{E}^* \mid \tilde{\theta}=\theta_2)P(\tilde{\theta}=\theta_2)}$$
$$= 0.167 \times 0.60 / (0.167 \times 0.60 + 0.25 \times 0.40) = 0.5$$

$$P(\tilde{\theta}=\theta_2 \mid \mathbf{E}^*) = \frac{P(\mathbf{E}^* \mid \tilde{\theta}=\theta_2)P(\tilde{\theta}=\theta_2)}{P(\mathbf{E}^* \mid \tilde{\theta}=\theta_1)P(\tilde{\theta}=\theta_1) + P(\mathbf{E}^* \mid \tilde{\theta}=\theta_2)P(\tilde{\theta}=\theta_2)}$$
$$= 0.25 \times 0.40 / (0.167 \times 0.60 + 0.25 \times 0.40) = 0.5$$

簡單貝氏分類計算範例(5/5)

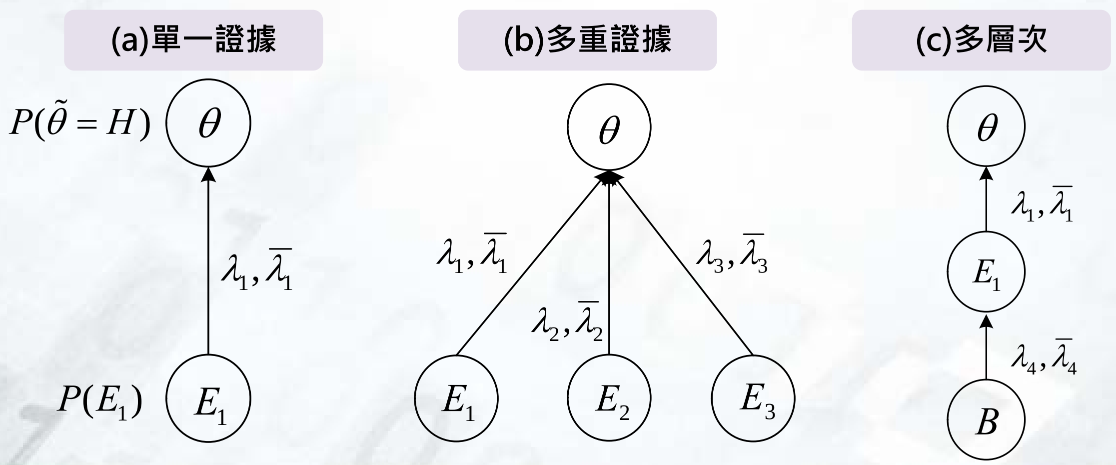


貝氏網路(Bayesian networks)

- **簡單貝氏分類**：假設屬性間互為條件獨立，但實務上屬性間往往存在相依關係，亦或一個目標事件的推理通常需要多個證據
- **貝氏網路**：以圖形呈現的統計推理模型，將多個不確定事件利用一組隨機變數以及變數間的影響關係來分析，能隨時根據新資訊或證據，修正相關的不確定事件之事後機率
  - 將複雜之不確定性判斷，解析為多個簡單有影響關係的不確定事件，每個不確定事件與目標假設的推論關係都是一個簡單判斷
  - 以網路來表達簡單節點間的因果推論關係，決策者對目標假設的評估，可由最底層節點觀察到的證據或樣本資訊逐層推演更新

貝氏網路之理論基礎

- 貝氏網路是用來處理複雜的推論關係，為「有向性非循環圖形」
- 貝式推理網路除網路圖外，還包含每個節點的事前機率，與每一個推論法則的強度，即證據或樣本資訊的概似函數或概似比
- 節點間的連結關係依照證據與目標事件的推理關係可分為：



單一證據推論

- 統計推論的最基本型態。以 $\tilde{\theta} = H$ 表示一個決策者有興趣的目標假設(通常是一個不確定事件），以機率 $P(\tilde{\theta} = H)$ 來表示事前機率，以*E*表示一個有關的證據

$$\text{If } E \text{ then } \tilde{\theta} = H$$

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E \mid H)P(H) + P(E \mid \overline{H})P(\overline{H})}$$

  - 例如，未進行檢查前，醫生只能經由一般數據判斷，有**1%**的國人會罹患肝硬化，即 **=1%**
  - 當醫生發現該病患是**B**型肝炎帶原者的新資訊時，根據「若**B**肝帶原，則罹患肝硬化」的推論，醫生會修正他認為該病患罹患肝癌的機率為 **(**肝硬化 **| B**肝帶原**)=25%**

概似函數(likelihood function)

- $P(E \mid H)$ 代表證據為**E**時，$\tilde{\theta} = H$ 的可能性，亦即當證據**E**出現隨著給定不同的 $\tilde{\theta}$ 條件而變化
- 事前機率與事後機率的正比關係

$$P(\tilde{\theta} \mid E) \propto P(\tilde{\theta}) \cdot P(E \mid \tilde{\theta})$$

- 事前機率取得方式：
  1. 大量的事前資訊，如歷史資料，可利用資料分析或資料挖礦方法計算機率
  2. 含糊的事前知識，可由專家判斷或決策者估計主觀機率
  3. 無先前資料提供任何資訊，則假設各種狀態機率相等

單一證據推理過程(1/2)

- 比率關係表示H發生和H不發生的比率為勝算(odds)

$$O(H)=\frac{P(\tilde{\theta}=H)}{P(\tilde{\theta}=\bar{H})}=\frac{P(H)}{1-P(H)}$$

- 確認*E*成立後，*H* 之事後勝算

$$O(\tilde{\theta}=H\,|\,E)=\frac{P(\tilde{\theta}=H\,|\,E)}{P(\tilde{\theta}=\bar{H}\,|\,E)}=\frac{P(E\,|\,H)}{P(E\,|\,\bar{H})}\times\frac{P(H)}{P(\bar{H})}$$

- 將*H* 發生條件下與*H* 不發生條件下*E*的概似函數以比率方式表達，稱為概似比 λ

$$\lambda=\frac{P(E\,|\,\tilde{\theta}=H)}{P(E\,|\,\tilde{\theta}=\bar{H})}$$

單一證據推理過程(2/2)

- *H*的事後或然率等於*E* 對*H* 的概似比乘以*H* 的事前勝算

$$O(H\,|\,E)=\lambda\times O(H)$$

- $\bar{E}$ 的概似比 $\bar{\lambda}=\frac{P(\bar{E}\,|\,\tilde{\theta}=H)}{P(\bar{E}\,|\,\tilde{\theta}=\bar{H})}=\frac{1-P(E\,|\,H)}{1-P(E\,|\,\bar{H})}$

- 建立當證據*E*不成立時*H*之事前或然率與事後或然率之修正關係

$$O(H\,|\,\bar{E})=\bar{\lambda}\times O(H)$$

- 事前機率與或然率的轉換關係式 $P(H)=\frac{O(H)}{1+O(H)}$

- 事後機率與事後或然率的關係式 $P(H\,|\,E)=\frac{O(H\,|\,E)}{1+O(H\,|\,E)}$

- $\bar{\lambda}$ 與 λ 的關係 $\bar{\lambda}=\frac{1-P(E\,|\,H)}{1-P(E\,|\,\bar{H})}=\frac{1-\lambda\cdot P(E\,|\,\bar{H})}{1-P(E\,|\,\bar{H})}$

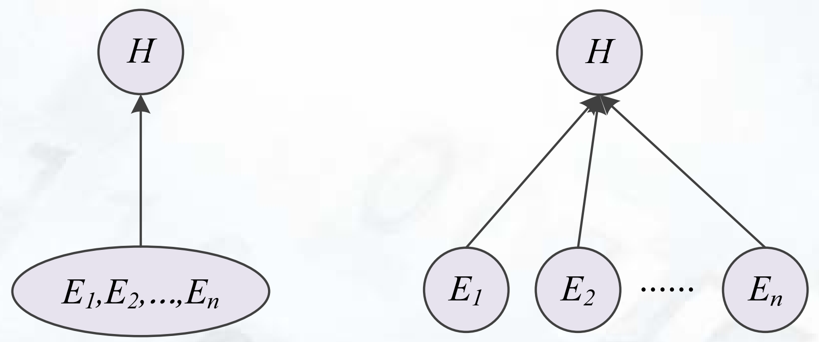
多重證據推論(1/3)

- 推論時參考的資訊或觀察的證據不只一種

$$\text{If } E_1 \text{ and } E_2 \text{ and} \ldots \text{and } E_n, \text{ then } \tilde{\theta} = H$$

***H* 的事後機率**

$$P(H \mid E_1, E_2, \cdots, E_n) = \frac{P(E_1, E_2, \cdots, E_n \mid H) \times P(H)}{P(E_1, E_2, \cdots, E_n)}$$



(a) 證據E_1, E_2, …, E_n聯合　(b) 證據E_1, E_2, …, E_n為條件獨立

多重證據推論(2/3)

- 若證據$E_1, E_2, \ldots, E_n$在給定 $\tilde{\theta} = H$時為條件獨立，亦即每個證據 E_i對$\tilde{\theta} = H$之可能性或概似函數 ，均不受其他證據和 $\tilde{\theta} = H$的推理關係影響，則$E_1, E_2, \ldots, E_n$ 對$\tilde{\theta} = H$ 的聯合概似函數 $P(E_1, E_2, \cdots, E_n \mid \tilde{\theta} = H)$ 為個別概似函數 $P(E_i \mid \tilde{\theta} = H)$ 的乘積：

$$P(E_1, E_2, \cdots, E_n \mid \tilde{\theta} = H) = \prod_{i=1}^{n} P(E_i \mid \tilde{\theta} = H)$$

$$P(\tilde{\theta} = H \mid E_1, E_2, \cdots, E_n) = \frac{\prod_{i=1}^{n} P(E_i \mid \tilde{\theta} = H) \times P(\tilde{\theta} = H)}{P(E_1, E_2, \cdots, E_n)}$$

- $\tilde{\theta} = \bar{H}$ 的事後機率

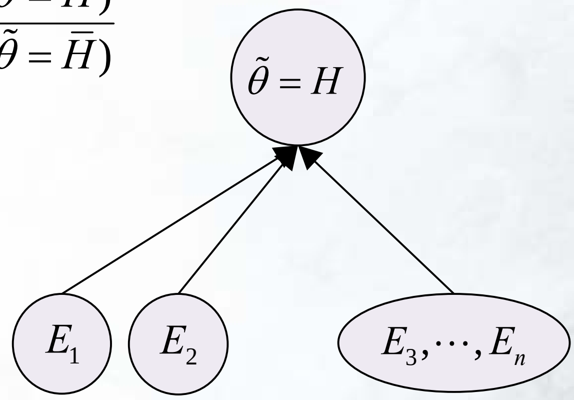
$$P(\tilde{\theta} = \bar{H} \mid E_1, E_2, \cdots, E_n) = \frac{\prod_{i=1}^{n} P(E_i \mid \tilde{\theta} = \bar{H}) \times P(\tilde{\theta} = \bar{H})}{P(E_1, E_2, \cdots, E_n)}$$

- 事後勝算　$O(\tilde{\theta} = H \mid E_1, E_2, \cdots, E_n) = O(\tilde{\theta} = H) \times \prod_{i=1}^{n} \lambda_i$

Enabling A+ Decisions®
DALab Proprietary

多重證據推論(3/3)

- λ_i 為證據 E_i 成立的概似比 $\lambda_i = \dfrac{P(E_i \mid \tilde{\theta} = H)}{P(E_i \mid \tilde{\theta} = \bar{H})}$
- $\bar{\lambda}_i$ 為證據 E_i 不成立的概似比 $\bar{\lambda}_i = \dfrac{P(\bar{E}_i \mid \tilde{\theta} = H)}{P(\bar{E}_i \mid \tilde{\theta} = \bar{H})}$
- 在單一證據與多重證據的貝氏網路推論中，每一個推論關係都具有證據成立的概似比 ，與證據不成立的概似比 ，分別代表成立或不成立時對假設的修正及其強度



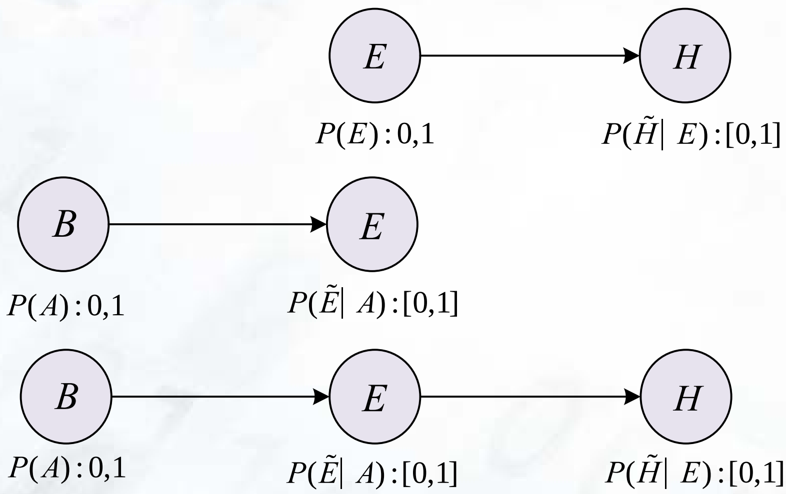
$$P(E_1, E_2, \cdots, E_n \mid \tilde{\theta} = H)$$
$$= P(E_1 \mid \tilde{\theta} = H) \cdot P(E_2 \mid \tilde{\theta} = H) \cdot P(E_3, E_4, \cdots, E_n \mid \tilde{\theta} = H)$$

多重證據推論，只有部分證據滿足條件獨立之假設

Enabling A+ Decisions®
DALab Proprietary

多層推論(1/2)

- 多層次的貝氏網路中，節點間的因果關係較為複雜，網路中的某一節點，可能同時是其後續節點的因，也是前行節點的果



(a) $P(\tilde{H}| E):[0,1]$

(b) $P(\tilde{E}| A):[0,1]$

(c) $P(\tilde{H}| A):[0,1]$

- 當根部節點有新證據時，貝氏推理即往上逐層修正每一個節點的機率，求得每個事件的事後機率

多層推論(2/2)

- 經過觀測事件B後，僅能在某些程度上確認E是否成立，則將E成立的機率表示為 $P(E|B)$。根據機率理論，可將 $P(H|B)$轉換：

$$P(H\mid B)=P(H,E\mid B)+P(H,\bar{E}\mid B)$$
$$=P(H\mid E,B)P(E\mid B)+P(H\mid \bar{E},B)P(\bar{E}\mid B)$$

$$P(H\mid B)=P(H\mid E)P(E\mid B)+P(H\mid \bar{E})P(\bar{E}\mid B)$$

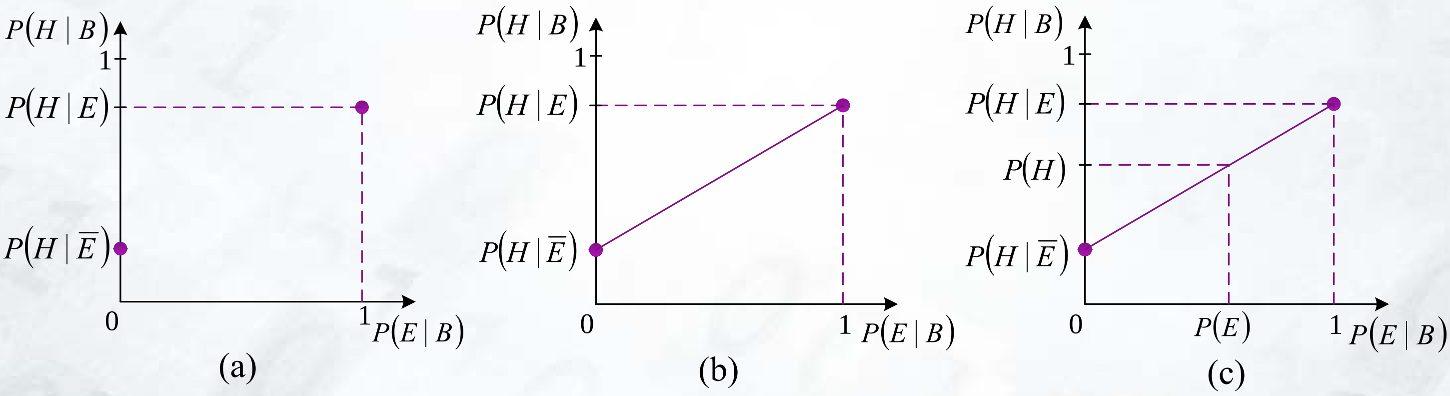
- 當證據*E*已發生與否已經確定時（即***P*(*E*) = 1**或**0**）
- 證據**B**直接對**H**的有效概似比 λ_B 為 $\lambda_B=\dfrac{P(B\mid H)}{P(B\mid \bar{H})}$

改寫為 $\lambda_B=\dfrac{P(B\mid H)}{P(B\mid \bar{H})}=\dfrac{P(H\mid B)}{P(\bar{H}\mid B)}\times\dfrac{P(\bar{H})}{P(H)}=\dfrac{O(H\mid B)}{O(H)}$，其中 $O(H\mid B)=\dfrac{P(H\mid B)}{P(\bar{H}\mid B)}=\dfrac{P(H\mid B)}{1-P(H\mid B)}$

Enabling A+ Decisions®
DALab Proprietary

貝氏網路的不一致性修正(1/2)

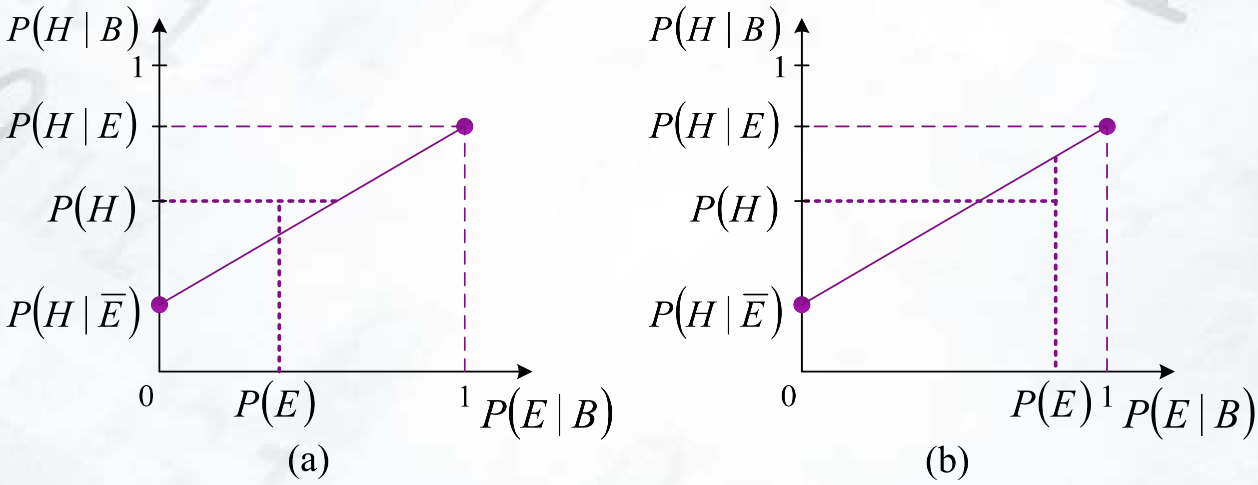
- 由於貝氏網路中，多層推論時的中間層節點是由其他機會節點推論而來，因此某節點的事前機率和由該節點的先行節點所推得的機率可能會產生不一致(inconsistent)



理想狀況下 $P(E\mid B)$ 對 $P(H\mid B)$ 之關係圖

Enabling A+ Decisions®
DALab Proprietary

貝氏網路的不一致性修正(2/2)

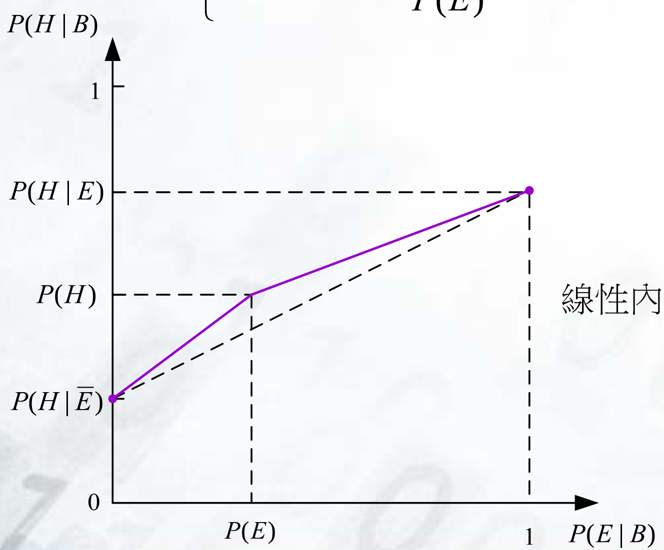


不一致狀況下 $P(E|B)$ 對 $P(H|B)$ 之關係圖

Enabling A+ Decisions®
DALab Proprietary

線性內插修正

$$P(H|B)=\begin{cases} P(H)+\dfrac{[P(E|B)-P(E)]}{1-P(E)}\cdot[P(H|E)-P(H)], & P(E|B)\ge P(E) \\ P(H)-\dfrac{[P(E|B)-P(E)]}{P(E)}\cdot[P(H|\bar{E})-P(H)], & P(E|B)<P(E) \end{cases}$$



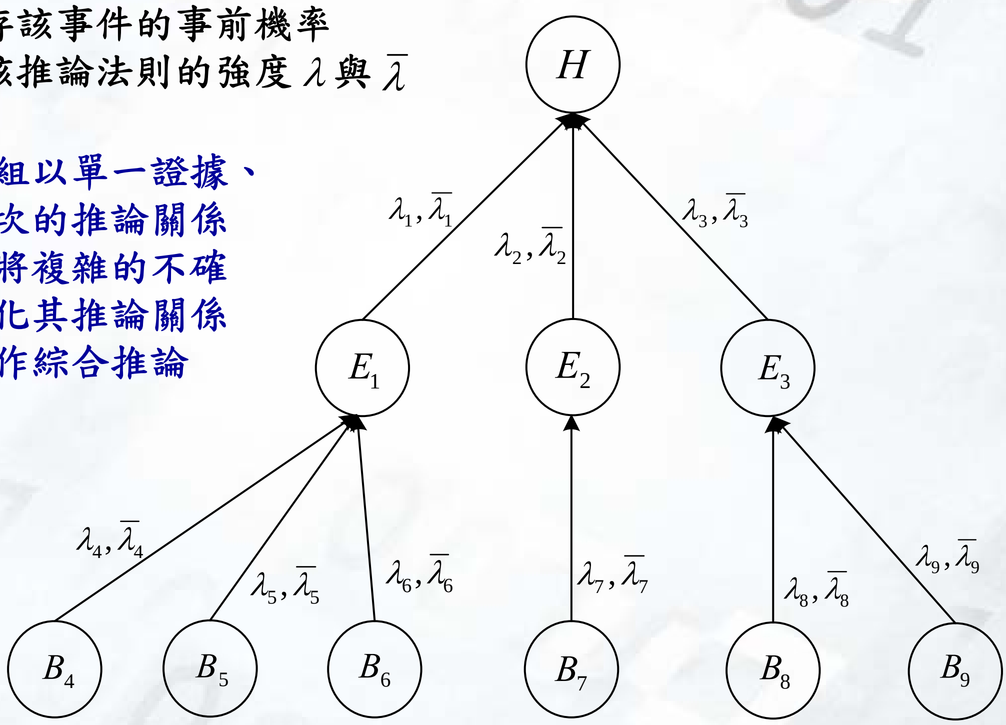
線性內插修正後 $P(E|B)$ 對 $P(H|B)$ 之關係圖

多層次、多重證據之貝氏網路圖

Enabling A+ Decisions®
DALab Proprietary

每一個節點必須儲存該事件的事前機率
每一個箭號須儲存該推論法則的強度 λ 與 $\bar{\lambda}$

- 貝氏網路包含一組以單一證據、多重證據與多層次的推論關係所連結之節點，將複雜的不確定事件分解並簡化其推論關係後，再整合起來作綜合推論



資料挖礦個案研究

Enabling A+ Decisions®
DALab Proprietary

台電饋線事故定位系統

PDF

Reference:

Chien, C.F., Chen, S. and Lin, Y. (2002), "Using Bayesian Network for Fault Location on Distribution Feeder of Electrical Power Delivery Systems," *IEEE Transactions on Power Delivery*, 17(13), 785-793.

Enabling A+ Decisions®
DALab Proprietary

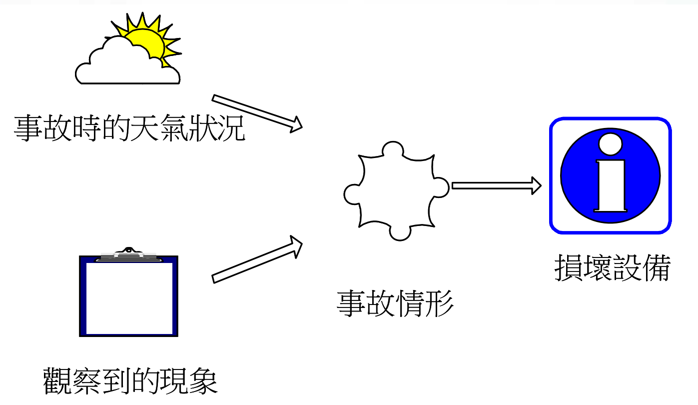
案例簡介

- 案例背景：目前台電的配電饋線多半尚未自動化，當配電饋線發生故障時，必須根據經驗判斷並立即趕赴事故現場，執行試送電，以確定故障位址，故障檢測相當耗時。即使已自動化的饋線，對於分歧線上的事故定位仍須仰賴經驗
- 主要目的：在饋線發生事故後，迅速檢出故障區以加速隔離並轉供其他電源，以縮短用戶的停電時間，減少經濟損失和社會成本
- 本案例(Chien, Chen & Lin, 2002)針對以貝氏網路為基礎之事故定位專家系統之發展與實證，本系統可以推論在不同的事故狀況下，各設備的相對損壞可能性

Enabling A+ Decisions®
DALab Proprietary

分析過程

1. 確定目標假設(即最頂層的節點)，以及與目標假設有關的隨機變數
2. 建立影響圖，以定義變數間的相依性。指向同一節點的所有先行節點須為條件獨立
3. 對每一個變數建立局部條件機率分布以進行模式評估與分析



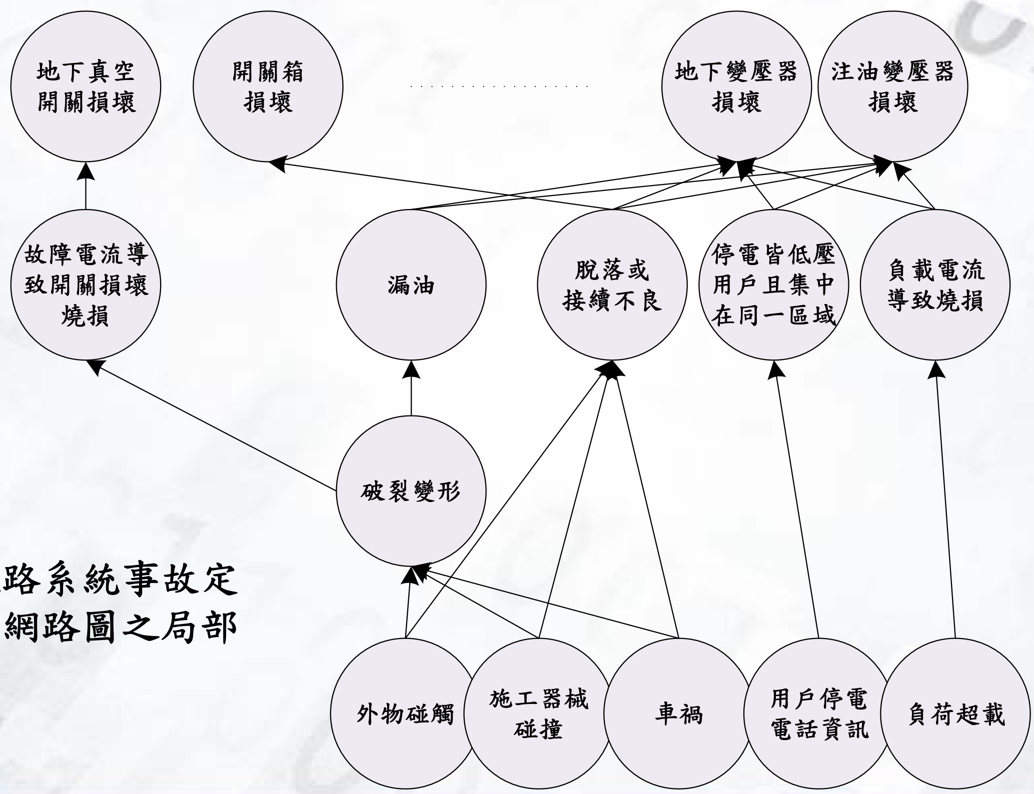
Enabling A+ Decisions®
DALab Proprietary

建構貝氏網路圖(1/2)

- 與配電調度領域的專家進行多次結構性訪談，來驗證研究小組建構的貝氏推理網路模型與專家的推理邏輯是否一致
  - 分析專家心智架構以建立貝氏網路模型
  - 擷取專家知識，對貝氏網路中每一個推理關係給定參數值
- 分析台電現行的配電系統事故停電統計資料，協助決定貝氏推理網路所需要的變數(節點)
  - 配電事故停電記錄表
    - ✓ 事故日期、時間、地點、發生事故的設備(如變壓器)、事故原因（如火災）
- 將分析頻率、相關性、事前機率與條件機率等相關項目整合為單一項目，以降低網路的節點數

Enabling A+ Decisions®
DALab Proprietary

建構貝氏網路圖(2/2)



地下線路系統事故定位貝氏網路圖之局部

給定貝氏推理網路之參數

$$\lambda_i = \frac{P(E_i \mid H)}{P(E_i \mid \overline{H})} = \frac{0.20}{0.10} = 2, \quad \overline{\lambda}_i = \frac{P(\overline{E}_i \mid H)}{P(\overline{E}_i \mid \overline{H})} = \frac{0.80}{0.90} = 0.89$$

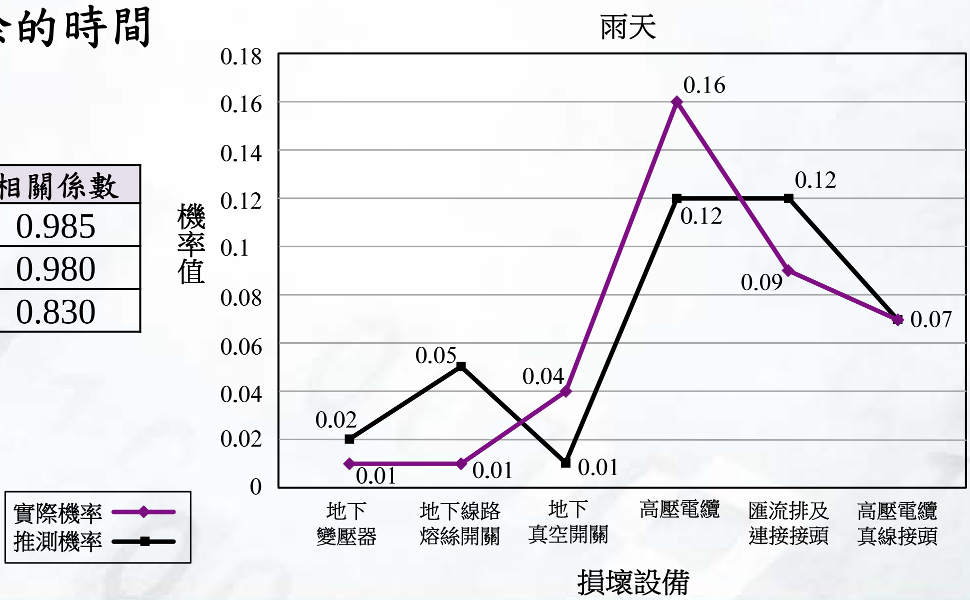
狀況	i.e.	相對機率
1.當注油變壓器發生故障，會觀察到漏油事件的機率為何	$P(E_i \mid H)$	20%
2.當其他設備發生故障（不包含注油變壓器），會觀察到漏油事件的機率為何	$P(E_i \mid \overline{H})$	10%
3.當注油變壓器發生故障,不會觀察到漏油事件的機率為何	$P(\overline{E_i} \mid H)$	80%
4.當其他設備發生故障，不會觀察到漏油的機率為何	$P(\overline{E_i} \mid \overline{H})$	90%

- 每個假設節點連結的先行節點(證據節點)須條件獨立，互不影響
  - 當一條推理法則改變時，其他指向同一假設的推理法則是否改變
  - 若專家認為會改變，就必須修正網路圖，直到所有的推論法則都滿足條件獨立假設

驗證貝氏推理網路

- 當發生停電事故時，工作人員可藉由貝氏網路，輸入其觀察到的現象，快速地將可能發生故障的設備鎖定在有限範圍內
- 藉由推論得到的機率值排序可能發生故障的設備，依序檢查，以降低故障排除的時間

情況	樣本數	相關係數
自然劣化	447	0.985
施工器械碰觸	111	0.980
雨天	209	0.830



案例小結

- 貝氏網路可藉由擷取大量的專家知識，模仿實際配電饋線中各種造成事故的因素之間的因果關係，達到多重目的

	發展	設計	評估	測試
階段1 架構與定義問題	定義問題 回顧文獻	架構問題 參數設定	評估專家的心智模式以建立定量架構	測試定量架構
階段2 發展模式	發展萃取專家知識的工具	調整貝氏網路的架構	評估貝氏網路的輸入項	測試輸入項的可靠度
階段 3 評估與分析模式	調整架構好的貝氏網路	設計驗證模式有效性的實驗	評估專家對真實案例的直接判斷	測試模式的可靠度與有效性

貝氏網路建構包含之活動列表

結論

- 貝氏網路可透過分析歷史資料、結合主觀機率與貝氏推論，以建立結合統計決策理論、實證資料和專家判斷之資料挖礦方法
- 優點：在抽樣資訊不足時，亦可利用事前機率來計算未來風險，不會因為資料不足而無法分析
- 貝氏推論架構可說是針對資料本質去選擇合適的事前機率與概似函數，使資料特性與貝氏推論模式相符合，並透過對事前機率分布之修正與驗證來獲得有效的決策模式