七、 統計與機率.....
7.1 基本統計方法與意義.....2
7.2 機率概論*.....3
機率測度.....5
事件關係形式.....7
條件機率與獨立事件.....9
貝氏定理.....13

七、　統計與機率

7.1 基本統計方法與意義

1. 算術平均

$$\overline{x} = \tfrac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + ... + x_n}{n}$$

2. 中位數

中位數（又稱中值）代表一個樣本、種群或機率分佈中的一個數值，其可將數值集合劃分爲相等的上下兩部分。對於有限的數集，可以通過把所有觀察值高低排序後找出正中間的一個作爲中值。如果觀察值有偶數個，則中值不唯一，通常取最中間的兩個數值的平均數作爲中值。

一個數集中最多有一半的數值小於中值，也最多有一半的數值大於中值。如果大於和小於中值的數值個數均少於一半，那麼數集中必有若干值等同於中值。

設連續隨機變數 X 的分佈函數為 F(X)，那麼滿足條件 P(X≤m)=F(m)=1/2 的數稱為 X 或分佈 F 的中位數。

對於一組有限個數的數據來說，它們的中位數是這樣的一種數：這群數據里的一半的數據比它大，而另外一半數據比它小。 計算有限個數的數據的中位數的方法是：把所有的同類數據按照大小的順序排列。如果數據的個數是奇數，則中間那個數據就是這群數據的中位數；如果數據的個數是偶數，則中間那 2 個數據的算術平均值就是這群數據的中位數。

原始資料依大小排列 $x_1 \le x_2 \le x_3 ..... \le x_n$

若 n 為奇數： $Me = x_{\frac{n+1}{2}}$

若 n 為偶數： $Me = \tfrac{1}{2}(x_k + x_{k+1})$

3. 變異數

在統計中，一個隨機變數的「方差」描述的是它的發散程度，也就是該變數離其期望值的距離。即求各項數值與 M(算術平均數)之差的平方之平均數，一般以 S^2 表示之。

4. 標準差

假設有一組數值 x_1, ..., x_n （皆為實數），其平均值為：

$$\overline{x} = \tfrac{1}{n}\sum_{i=1}^{n} x_i$$

此組數值的標準差為：

$$\sigma = \sqrt{\tfrac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

從一大組數值當中取出一樣本數值組合 $x_1, ..., x_n$ ，常定義其*樣本標準差*：

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

6. 變異係數 C.V.

$$CV = \frac{S}{M} \times 100\%$$

7. 相關係數

$$r = \frac{\sum_{i=1}^{n}(X_i - M_x)(Y_i - M_y)}{n \cdot S_x \cdot S_y}$$

相關係數 r 之值介於 1~-1

a. $r = 1$ =>完全正相關。

b. $|r|$ 越趨近於 1 表高度相關，反之則否。

c. $r = 0$ 表零相關。

7.2 機率概論*

統計學研究的範圍可分為兩大部份：

敘述統計（descriptive statistics）：將其所得到的資料，歸納整理之後，進而單單描述並分析它們的特性。

推論統計（statistical inference）：以抽樣方法獲得樣本，並根據樣本實際資料的分析來推論或判斷母體的狀態，進而預測未來，做出決策。

EX. 擲一不公正的骰子(有瑕疵)，出現 1 點的機率未知，可以投擲骰子 N 次，若出現 n 次 1 點，則以 $\frac{n}{N}$ 作為出現 1 點的機率之估計值。

機率學：以已知的母體來測知樣本特徵。(和推論統計恰好相反)

EX.擲一公正的骰子一次，則出現 1 點的機率為 $\frac{1}{6}$ ，可以知道投擲骰子次數過多時，出現 1 點的可能性會接近 $\frac{1}{6}$ 。

【定義】機率即為衡量某一事件在未來發生的可能性，並將此可能性數量化。

【定義】一機率實驗所可能產生的結果，稱為此實驗的某一樣本點（sample point）。所有樣本點所構成的集合，稱為此實驗的樣本空間（sample space），通常以大寫英文 S 表之。

【定義】事件（event）即樣本空間 S 的子集合。只包含一個樣本點的事件稱為單一事件（simple event）。含有兩個樣本點以上的事件稱為複合事件（compound event）。

EX.樣本空間，樣本點，及事件。

投擲一公平的骰子，其樣本空間為

S＝{1,2,3,4,5,6}

定義 A，B 兩事件為：

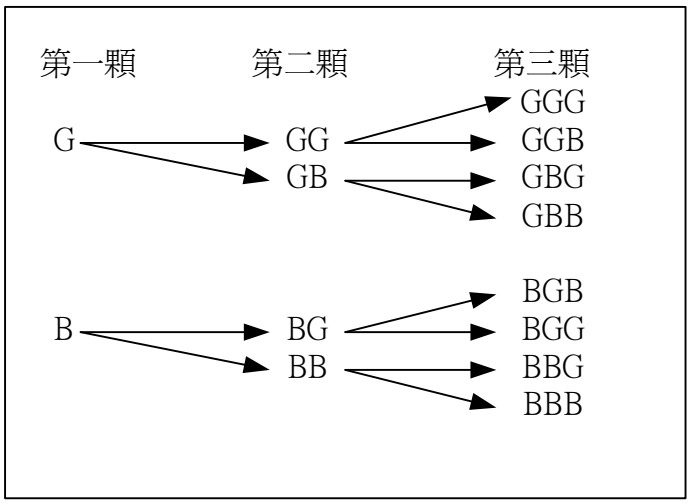
A：面朝上的點數，正好出現 3 點的事件。

B：面朝上的點數，出現超過 3 點的事件。

A＝{3}，為一單一事件。

B＝{4,5,6}，為一複合事件。

隨機抽取 3 顆電燈泡，我們就以良好（G），損壞（B）為符號，試著以樹狀圖列出此實驗所有可能的樣本點。



樣本空間為：

S＝{GGG，GGB，GBG，GBB，BGG，BGB，BBG，BBB}

定義 A，B 兩事件為：

A：抽取之 3 顆燈泡都為良好的事件。

B：抽取之 3 顆燈泡中，至少有兩顆損壞的事件。

則 A＝{GGG}，為一單一事件

B＝{GBB，BGB，BBG，BBB}，為一複合事件。

■ 機率測度

【定義】若A為某一機率事件，A 可能發生的機率以 P(A)表示。

若 P(A)=0：表示 A 不可能發生。

若 P(A)=1：表示 A 必然發生。

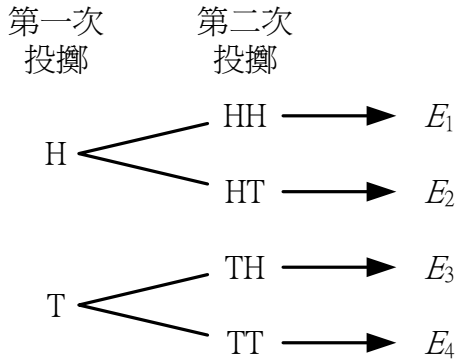
一般而言，P(A)會介於 0 與 1 之間。

即 0≦P(A)≦1。

【定義】計算事件所可能發生的機率，可先算出此事件所包含的每一樣本點，其個別所可能發生的機率，再將其加總即是。

EX.投擲一公平的硬幣兩次，試算出其剛好出現正面一次的機率。

解：



E2：第一次出現正面第二次出現反面的事件。

E3：第一次出現反面第二次出現正面的事件。

$P(E2)=\frac{1}{4}$　　$P(E3)=\frac{1}{4}$

故 $P(E)=P(E2)+P(E3)=\frac{1}{2}$

【定義】(古典法)：

當某一隨機實驗的樣本空間是有限的，且每一樣本點所可能發生的機率是相等的，則一事件A所可能發生的機率 $P(A)=\frac{N(A)}{N(S)}$ 。

其中，*N(S)*為樣本空間所包含的樣本點個數，*N(A)*為事件A本身所包含的樣本點個數。

EX.投擲一公平的硬幣兩次，試算出其剛好出現正面一次的機率。

解：S={HH，HT，TH。TT}

每個樣本點出現的機率相同，A 表剛好出現正面一次的事件，則 $P(A)=\frac{N(A)}{N(S)}=\frac{2}{4}=\frac{1}{2}$

【定義】(相對次數法)：

在相同的情況下，一直重覆某實驗 N 次，而事件 A 出現的次數為 N(A)，則 $\frac{N(A)}{N}$ 為事件 A 發生機率 P(A)的一個很好的估計值。

EX. 欲估計某一不公平的骰子出現 1 點的機率。重複擲骰子 N 次，發現 1 點出現 N(A)次，則出現 1 點的機率之估計值為 $\frac{N(A)}{N}$ 。

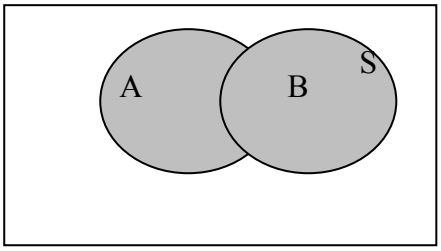
【定義】(主觀法)：

單由個人的經驗，觀點，相當主觀的推算事件發生的機率。

■ 事件關係形式

$A \cup B$ 表A發生或B發生的事件。

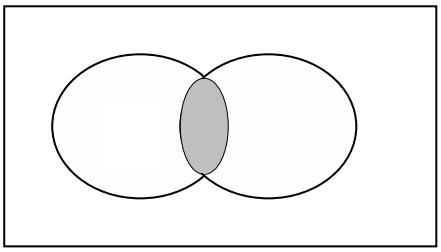
$A \cup B$ =(A發生B不發生)$\cup$ (A不發生B發生) $\cup$ (A發生B發生)

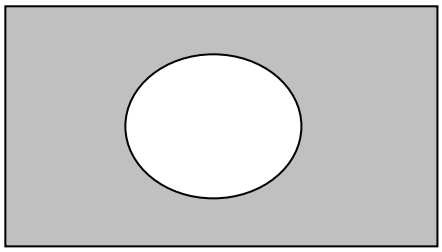


S 即整個大長方形，其代表著整個實驗的樣本空間。而兩圓形，則代表著A，B兩事件。

$A \cap B$ 表A發生且B發生的事件。

A^c 表A不發生的事件。

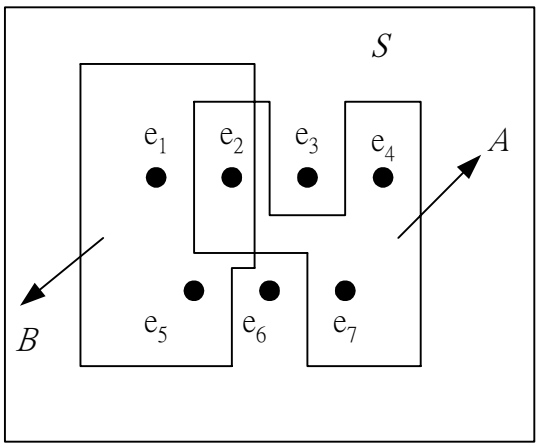




EX. 假設某一機率實驗其樣本空間包括 7 個樣本點：$S=\{e_1,e_2,e_3,e_4,e_5,e_6,e_7\}$定義兩事件A、B為；$A=\{e_2,e_4,e_7\}$ ，$B=\{e_1,e_2,e_5\}$； (a) 以作圖方式將A、B兩事件表現出來。

(b) 求 $A \cap B$ ，B^c ，$A \cap B^c$ ，$A \cup B$ 。

解：(a)



(b) $A \cap B = \{e2\}$ $B^c = \{e_3, e_4, e_6, e_7\}$

$A \cap B^c = \{e_4, e_7\} = A - B$

$A \cup B = \{e_1, e_2, e_4, e_5, e_7\}$

EX. 投擲三枚公平的硬幣(視同投擲一枚公平的硬幣 3 次)並觀察其正反面情形。定義兩事件 A，B 為

A：至少有一正面的事件。

B：至少有一反面的事件。

則

$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ A=S-{TTT}，B=S-{HHH}，

$A \cap B = S - \{TTT, HHH\}$

$A \cup B = S$ =S；因此，

$P(A) = 1 - \frac{1}{8} = \frac{7}{8}$，$P(B) = 1 - \frac{1}{8} = \frac{7}{8}$，$P(A \cap B) = 1 - \frac{2}{8} = \frac{6}{8}$，$P(A \cup B) = 1$。

【定義】設 A，B 為兩事件且 A ∩ B ＝ ϕ，則稱 A，B 互斥（mutually exclusive）。

【公理】

1.對任一事件 A，0≦P(A)≦1。

2. S 為樣本空間，則 P(S)＝1。

3.對於兩兩彼此互斥的事件 A1，A2，…，An，

則 $P(A_1 \cup A_2 \cup ..... \cup A_n) = \sum_{i=1}^{n} P(A_i)$

EX. 假設某一機率實驗，其樣本空間包含了 5 個樣本點 $\{e_1, e_2, e_3, e_4, e_5\}$。

(a) 如果 $P(e_1) = P(e_2) = 0.15$，$P(e_3) = 0.4$ P 和 $P(e_4) = 2P(e_5)$，試求 $P(e_4)$ 及 $P\{e_5\}$。

(b) 如果 $P(e_1) = 3P(e_2) = 0.3$，且其餘樣本點機率皆相同，試求其餘樣本點之機率。

解：

(a) 假設 $P(e_5) = P$，則由題目可知 $P(e_4) = 2P$ P

而 $P(s) = 1 = P(e_1) + P(e_2) + P(e_3) + P(e_4) + P(e_5) = 0.15 + 0.15 + 0.4 + 2P + P$

因此 $3p = 0.3 \Rightarrow p = 0.1$ 故 $P(e_4) = 0.2$ $P(e_5) = 0.1$

(b) 假設 $P(e_3) = P(e_4) = P(e_5) = p$，則知

$1 = P(S) = 0.3 + 0.1 + 3p \Rightarrow p = 0.2$

故 $P(e_3) = P(e_4) = P(e_5) = 0.2$

■ 條件機率與獨立事件

1.聯合機率及邊際機率

◎考慮兩個以上事件同時發生的機率，稱為聯合機率。

◎考慮單一事件發生的機率，稱為邊際機率。

【例題】某班學生(樣本空間 S)，將 S 分割成男生(A)、女生(Ac)兩大塊。並將 S 分成統計學及格(B)與不及格(Bc)。

聯合機率表：

分數 性別	及格(B)	不及格(Bc)	合計
男生(A)	P(A∩B)	P(A∩Bc)	P(A)＝P(A∩B)＋P(A∩Bc)
女生(Ac)	P(Ac∩B)	P(Ac∩Bc)	P(Ac)＝P(Ac∩B)＋P(Ac∩Bc)
合計	P(B)＝P(A∩B)＋P(Ac∩B)	P(Bc)＝P(A∩Bc)＋P(Ac∩Bc)	100%

2.條件機率

【定義】給定一事件 B 已經發生，則事件 A 的條件機率（conditional probability）記為 P(A|B)，並定義為

$$\mathrm{P(A\mid B)}=\frac{P(A\cap B)}{P(B)}\text{，P(B)>0}$$

【例 4.9】假設有兩事件 A，B 且 P(A)＝0.5 ，P(B)＝0.3，P(A∩B)＝0.1。
試求(a)P(A|B) ， (b) P(B|A) 。

解：

(a) $\mathrm{P(A\mid B)}=\dfrac{P(A\cap B)}{P(B)}=\dfrac{0.1}{0.3}=\dfrac{1}{3}$

(b) $\mathrm{P(B\mid A)}=\dfrac{P(B\cap A)}{P(A)}=\dfrac{0.1}{0.5}=\dfrac{1}{5}$

EX 某一大學金融系大二班，男生有 30 人，女生 20 人。統計學期中考，60 分為及格。及格和不及格人數如下所述，括號中為其佔全班人數比例，今從此班任抽取一人：

分數 性別	及格 C	不及格 D	總和
男同學 A	10(0.2)	20(0.4)	30(0.6)
女同學 B	10(0.2)	10(0.2)	20(0.4)
總和	20(0.4)	30(0.6)	50(100%)

(a) 若已知被抽取此人為一男同學，試求此男同學不及格的機率。

(b) 若已知被抽取此人為一女同學，試求此女同學及格的機率。

解：(a) $P(D\mid A)=\dfrac{P(D\cap A)}{P(A)}=\dfrac{0.4}{0.6}=\dfrac{2}{3}$

(b) $P(C\mid B)=\dfrac{P(C\cap B)}{P(B)}=\dfrac{0.2}{0.4}=\dfrac{1}{2}$

3.獨立事件

A，B 為任二事件，若有下列情形之一成立：

(1) P(A|B)＝P(A)

(2) P(B|A)＝P(B)

(3) P(A∩B)＝P(A)P(B)

則稱 A 與 B 兩事件獨立（independent），反之則稱 A 與 B 兩事件相依（dependent）。

註：兩事件獨立，意謂著其中一事件發生與否，都不影響另一事件發生的機率。

EX.

分數 性別	及格 C	不及格 D	總和
男同學 A	10(0.2)	20(0.4)	30(0.6)
女同學 B	10(0.2)	10(0.2)	20(0.4)
總和	20(0.4)	30(0.6)	50(100%)

試問 A，D 兩事件是否獨立？

解：

$$\mathrm{P(A\mid D)}=\frac{P(A\cap D)}{P(D)}=\frac{0.4}{0.6}=\frac{2}{3}$$

$\mathrm{P(A)}\times 0.6$

因 $\mathrm{P(A\mid D)}\neq P(\mathrm{A})$

所以此兩事件並不獨立，而是相依。

互斥（mutually exclusive）　　$P(A\cap B)=0$

獨立（independent）　　$P(A\cap B)=P(A)P(B)$

EX. 欲由編號 1 至 50 號的摸彩卷中抽出 1 張，抽中者可獲得黃天受老師簽名照 1 張。今某君持有 3 號及 28 號兩張摸彩卷，令 A 代表抽出為偶數的事件，B 事件={3，28}。則

$$P(A\cap B)=\{28\}\neq\phi$$

$$P(B\mid A)=\frac{P(B\cap A)}{P(A)}=\frac{1/50}{25/50}=\frac{1}{25}=P(B)$$

知 A 和 B 為獨立的，但 A 和 B 不為互斥。

4.機率法則

【定義】(加法法則)

$P(A\cup B)=P(A)+P(B)-P(A\cap B)$

若兩事件互斥，則此法則改寫為：

$P(A\cup B)=P(A)+P(B)$

【例 4.14】舉行期中考，已知班上同學有 50%數學不及格，有 30%英文不及格，而有 20%數學與英文均不及格。今從此班級中隨機抽取一人，試問此同學在這兩科中，至少有一科不及格的機率為何？

解：若令事件

A：抽取此同學，數學不及格的事件

B：抽取此同學，英文不及格的事件

$P(A)=50\%$，$P(B)=30\%$

數學與英文均不及格事件的機率為

$P(A\cap B)=20\%$

至少有一科不及格事件的機率為

$P(A\cup B)=P(A)+P(B)-P(A\cap B)$

$=50\%+30\%-20\%\ =60\%$

推廣：

任意三事件 A，B，C

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)-P(C\cap A)+P(A\cap B\cap C)$

若三事件互為互斥事件，則

$P(A\cup B\cup C)=P(A)+P(B)+P(C)$

【定義】(乘法法則)

$P(A\cap B)=P(B)P(A|B)=P(A)P(B|A)$

若兩事件獨立，則此乘法法則改寫為：

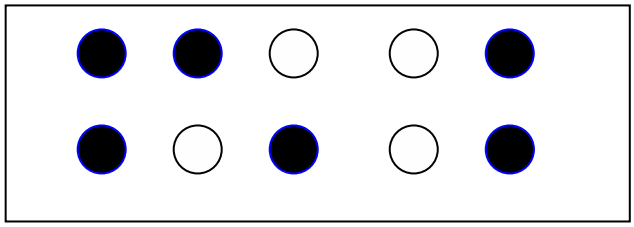
$P(A\cap B)=P(B)P(A)$

EX. 在一不透明的桶子中，裝有 4 顆白球(W)及 6 顆黑球(B)。試求抽取的兩顆球中，一顆為白球，一顆為黑球的機率。

解：

定義　Wi：抽取之第 i 顆球為白色的事件

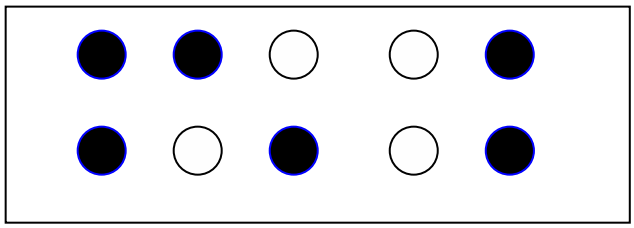
Bi：抽取之第 i 顆球為黑色的事件 i＝1,2



a.抽取不放回（without replacement）

P(一個白球，一個黑球)

$$P=(W_1\cap B_2)+P(B_1\cap W_2)$$
$$=P(W_1)P(B_2\mid W_1)+P(B_1)P(W_2\mid B_1)$$
$$=(4/10)(6/9)+(6/10)(4/9)=\frac{4}{15}+\frac{4}{15}=\frac{8}{15}$$



b.抽取放回（with replacement）

P(一個白球，一個黑球)

$$=P(W_1\cap B_2)+P(B_1\cap W_2)$$
$$=P(W_1)P(B_2\mid W_1)+P(B_1)P(W_2\mid B_1)$$
$$=(4/10)(6/10)+(6/10)(4/10)=\frac{6}{25}+\frac{6}{25}=\frac{12}{25}$$

5.餘集法則

【定義】(餘集法則)

對任一事件 A，則：

$P(A) = 1 - P(A_c)$ 或 $P(Ac) = 1 - P(A)$

EX. 有一神射手，其槍法十分準確，有 90%的機率可射擊命中。今對此目標連續射發十次，並假設十次擊發彼此皆為獨立事件。試問此神射手至少失誤一發的機率。

解：

A：至少失誤一發的事件

Ac：十發完全命中無失誤的事件

$P(A_c) = (0.9)10$

$P(A) = 1 - P(A_c) \times 1 - (0.9)10 = 0.6513$

■ 貝氏定理

【定理】(全機率法則)

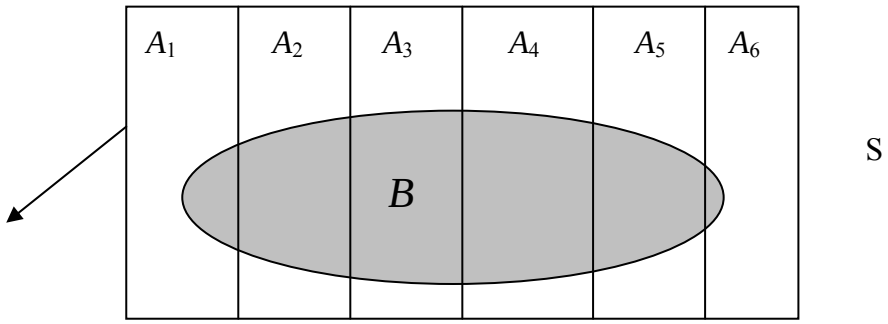
假設 S=A1∪A2∪……∪Ak；P(Ai)>0，i=1,2,……k ，且對 i≠k 而言，Ai∩Ak＝ϕ。

則對任意事件 B：

$$P(B) = \sum_{i=1}^{k} P(A_i)P(B \mid A_i)$$

註：$A_1, A_2, ...A_k$ 為 S 的一個分割(partition)。

舉例證明：

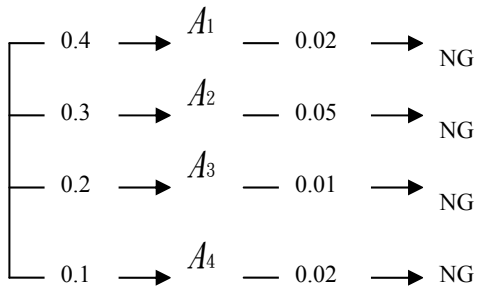


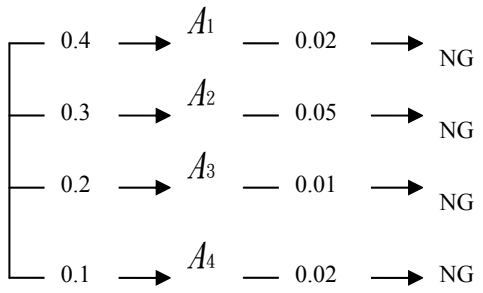
$P(B) = (A_1 \cap B) + P(A_2 \cap B) + ... + P(A_6 \mid B)$

$= P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + ... + P(A_6)P(B \mid A_6)$

$\sum_{i=1}^{6} P(A_i)P(B \mid A_i)$

EX. 有 4 部機器 (A_1, A_2, A_3, A_4) 生產同一產品，各機器出產產品數量各佔總產量之比為 0.4 ,0.3 ,0.2 ,0.1。再令產品為不良品的事件為 B。各部機器產品的不良率分別為 0.02,0.05,0.01,0.02 試問若隨機抽取一產品，其為不良品的機率為？

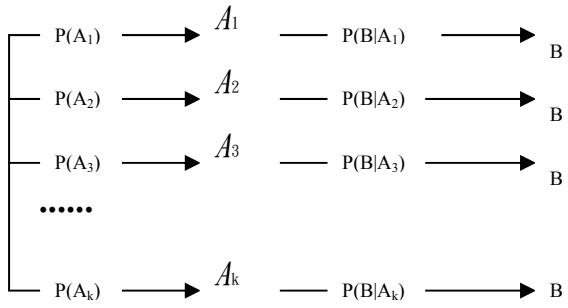




解:

$$\mathrm{P(NG)} = \sum_{i=1}^{4} P(A_i \cap NG) = P(A_1 \cap NG) + P(A_2 \cap NG)$$

$$= \mathrm{P(A_1)P(NG \mid A_1) + P(A_2)P(NG \mid A_2) + (A_3)P(NG \mid A_3) + P(A4)P(NG \mid A_4)}$$
$$= 0.4 \times 0.02 + 0.3 \times 0.05 + 0.2 \times 0.01 + 0.1 \times 0.02$$
$$= 0.027$$



【定理】(貝氏定理；Bayes' Theorem)

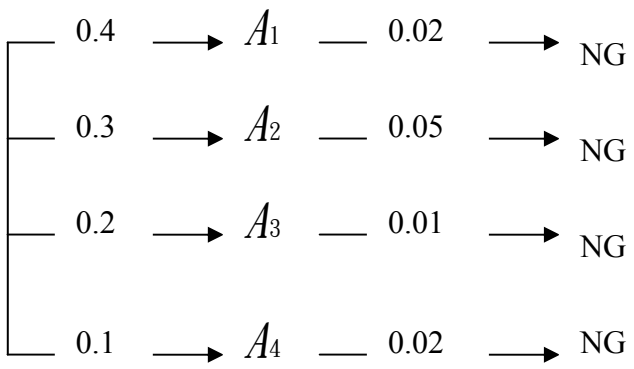
假設事件{A1，A2，…，Ak}為樣本空間 S 之一分割，P(Ai)>0，i=1,2,……k，則對其他任意事件 B 而言：

$$P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)}$$

$$= \frac{P(A_j)P(B \mid A_j)}{\sum_{i=1}^{k} P(A_i)P(B \mid A_i)}$$

EX. 前例中，若從工廠的成品中，隨機抽取一產品。假若已知此抽取之產品為不良品，試問它是由機器 A1 所生產的機率為何？

解：



$$P(A_1 \mid NG) = \frac{P(A_1 \cap NG)}{P(NG)}$$ 『條件機率』

$$= \frac{P(A_1)P(NG \mid A_1)}{P(NG)}$$ 『乘法法則』

$$= \frac{P(A_1)P(NG \mid A_1)}{\sum_{i=1}^{4} P(A_i)P(NG \mid A_i)}$$ 『全機率法則』

$$= \frac{(0.4)(0.02)}{0.027} = \frac{8}{27} = 0.296$$