瞻前顧後

黃文璋教授

國立高雄大學應用數學系

1　八方看看

副總統連戰於應中國廣播公司錄製“智慧的話”節目時,提出在他人生中,印象很深刻的一句話為“看看左右”(88年2月24日中國時報第4版,記者李建榮)、隨後在我見我思專欄,作者夏珍消遣行政院長蕭萬長的座右銘大概是“看上看下”(88年2月25日中國時報第15版)、政治人物可能是常需要左右看看競爭者有那些,以及上下看看君意及民意、

任賢齊有一首歌“對面的大孩看過來”,其中有句“我左看右看上看下看,原來每個大孩都不簡單”、仍舊是左右上下看,似乎前看後看連歌手都不重視、左看右看不知何者較重要?當一個人碰到不想回答的問題,往往顧左右而言他,看來左右是對等的、而有人重視上看,有人在乎下看、本文一方面介紹前看及後看,一方面要指出,通常做判斷時,常為人所輕忽的後看,其實是極重要的、能左右、上下、前後,八方看看自然是很好的,否則至少瞻前顧後,前後都要看看、

2　因果難分

前後看看之重要性,之所以為許多人所疏忽,乃源自於一般人邏輯不是很好,常常因果分不清、例如,在二月河(1997)p.56-57,施琅拿出一百枚康熙銅錢,向著台灣海域圖,然後說“倘若我軍出師順利,當有九十五枚以上的字面朝上!”結果一擲之下,有九十九枚是康熙字面朝上,將士大喜、

我們實在不知將士為何大喜?大家都知道,“若 A 則 B”成立不一定導致“若 B 則 A”成立、有時某一重要職務出缺,許多人汲汲營營地爭取,這時往往傳出所謂條件說、例如,要當行政院長,要具備下述五個條件:$\cdots$

、這些條件一出來,熱中於此職務的人低頭一看,自己都有,不禁高興萬分、可惜最後卻不是他、又如"李總統表示,他在'台灣的主張'一書中,強調公元兩千年總統的必備條件是'以新台灣人為基礎,開創新時代的人'、'言行一致'是未來擔任總統的必備條件"(88年6月19日聯合報第4版,記者何振中、邱貴雄)、要知這些條件,就是所謂必要條件、在"若 A 則 B"此一命題中,A 是使此命題成立的充分條件,B 是必要條件、也就是說,如果 A 成立了,B 一定成立、至於 B 成立時,A 是否必成立呢?那可不一定、我們只知道如果 B 不成立,則 A 必不成立、因此 B 才稱為此命題之必要條件、此道理看似簡單合理,但還是常會令人搞不清、對邏輯想進一步了解的讀者,可翻閱黃文璋(1999)第八章、

現依照施琅所述,當九十五枚以上的字面朝上,並不導致出師順利、而只知若沒有九十五枚以上的字面朝上,便會出師不順利、因此若施琅堅信他的銅板理論,便不應丟銅板,否則只會影響軍心、至於將士大喜,實在是白歡喜、二月河、施琅及將士,此三方必有一方是邏輯不好、

"因"導致"果",但不同的因卻可能導致相同的果、光看到果,是無法知道其原因為何? 可惜有人分不出因果關係,以為果也導致因,夾纏不清常因此產生、與這種人談話是很辛苦的、我們倒不冀求談笑有鴻儒,往來無白丁、但遇一邏輯不通,因果不分者,雖其不一定是白丁,只不過即使與之往來,也很難談笑、

由於因果不分,前後不知有異,也就並非太令人驚訝了、於是瞻前顧後之必要性,遂為許多人所疏忽、

3 條件機率

假設生男生女的機率皆為1/2、你看到有某孕婦,則猜她會生男的機率為1/2是合理的、某日有人按你家門鈴, 在開門前要不要猜按鈴者是男或是女的機率?當然都是1/2、對一有兩個小孩的家庭,若是一男孩及一女孩, 一般認為這是不錯的組合,剛好是中文裡的"好"字、你可能認為不太

稀奇,因有男大及大男兩種可能, 而其機率各為 $1/2 \cdot 1/2 = 1/4$,因此有一半的機率兩個小孩為一男及一大、這一切都是基於一個人會是男或大的機率各為1/2、眞是這樣嗎?其然乎,豈其然乎?

某日你至友人家拜訪,你早知他有二小孩,按門鈴後有一大孩替你開門、坐定後你順口問友人他是否有一男孩? 他給你肯定的答案、於是你誇獎他命好,一男一大最完美、又一回你另一朋友也有二小孩,其中有一男孩,你又問他是否另一個為大孩?又是肯定的答案、於是你又瞎捧一番、現在不少家庭都是兩個小孩,每回若你知道其中一個是男孩,便猜另一為大孩;而若知其中一個為大孩,便猜另一個是男孩、本來只是為了讓對方高興, 長期下來,居然猜中的比例不小,似乎是一半以上、於是你有點納悶,不要說一男一大的家庭是半數以上為不合理,一小孩為男或大的機率似乎都不是1/2了,這究竟是怎麼回事?

一家庭二小孩的性別有4種可能,即男男、男大、大男、大大,每一種可能之機率皆為1/4、若已看到一大孩,便知必為男男之外的三種可能之一,而其中有兩種可能會有男孩(男大及大男), 因此另一小孩為男孩之機率為2/3、若已看到一男孩之情況類似,另一小孩為大孩之機率亦為2/3、原來你以為自己經常猜對,其實並不是巧合,乃是有依據的、而道理也就是這麼淺顯不過、

等等!你覺得不制止這種匪夷所思的結論是不行的、任選一人為男或大的機率為何?你反問、1/2,答案很乾脆、那為何此家庭另一小孩為男或大的機率不再是1/2?你理直氣壯地反問、因為我們看到其中一個的性別了,理由還是這麼薄弱、你難道沒聽過若連生10個男孩,下一胎會是男或大的機率還是1/2嗎?你幾乎要受不了了、

讀者諸君,這就是條件機率(conditional probability)、任選一人為男或大的機率,由過去的經驗, 可設為1/2,此即所謂事前機率(prior probability)、但若給了一些條件,則在這些條件下,其機率是有可能改變的、此時之機率,便稱為條件機率,是一種事後機率(posterior probability)、譬如說,你若到高雄大中任選一人,還會覺得挑中的是男生的機率為1/2嗎?有人來敲你門,你聽到像高跟鞋的走路聲, 難道不猜此大約是大

生嗎?給你一些條件,或說知道某些資訊,對事務的判斷將更精準,本來就很合理、福爾摩斯(Holmes)查案,不就一向如此?福爾摩斯由一些蛛絲馬跡做出正確的推測,你對他很佩服、而我們由一明確的資訊(看到一家庭中的一小孩為男或大),對另一看不到的事物(另一小孩的性別),只是有較精確些的推測(機率由1/2變為2/3),從邏輯上而言,有何不可?我們常以“事後諸葛”取笑人有後見之明、大部分的人當然連事後諸葛都不如,他們甚至連事後機率都無法理解、

其實一般人都知道,即使數據指出,任選一20歲的男生,其體重為60公斤以下的機率為0.5,但對某一特定的20歲男生,若知其身高為185公分,則可能僅會以很低的機率(譬如說0.05),猜其體重不到60公斤、這是因我們知道體重與身高並不“獨立”(independent)、但在一家庭的兩個小孩中,某一小孩的性別與其排序也不獨立,這就不是那麼明顯了、

底下我們給條件機率及獨立之定義、令 $P(A)$ 表事件(event)A 之機率、在此一實驗之某一可能的結果,稱為一事件、而一實驗之所有可能的結果之集合,稱為樣本空間(sample space),通常以 Ω 表之、

定義1. 設有 A, B 二事件,且 $P(B) > 0$、則在給定 B 發生之下,A 之條件機率,以 $P(A|B)$ 表之,且定義為

(1) $$P(A|B) = P(A \cap B)/P(B)、$$

又若

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

則稱 A 與 B 為獨立事件,或稱 A 與 B 為相互獨立(mutually independent)、

若 $P(B) > 0$,由(1)及(2)看出,當 A, B 獨立時,$P(A|B) = P(A)$、也就是在給定 B 發生之下,A 發生之可能性,並未受到影響、而二者同時發生之機率,就只是為二者分別發生機率之乘積、在機率統計裡的這種獨立的概念,又稱為統計獨立(statistically independent),或隨機獨立(stochastically independent)、其意義與一般所說的主權獨立、經濟獨立中的“獨立”意義並不相同、

若 $P(B)=0$,為了方便,我們定義 $P(A|B)=P(A)$ 、又由(1)得

$$P(A\cap B)=P(A|B)P(B) \text{ 、} \tag{3}$$

上式不論 $P(B)$ 是否為正都成立 、

我們亦有 n 個事件獨立之定義 、

定義2.設有事件 $A_1, A_2, \cdots, A_n, n \geq 3$ 、若對 $\forall k \geq 1$ 及$1 \leq i_1 < i_2 < \cdots < i_k \leq n$,

$$P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}), \tag{4}$$

則稱 $A_1, A_2, \cdots, A_n$ 為獨立事件 、

不難看出,若欲 n 個事件為獨立,要驗證

$$\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{n}=(1+1)^n-\binom{n}{0}-\binom{n}{1}=2^n-n-1$$

個條件,而不只是驗證

$$P(A_1\cap A_2\cap\cdots\cap A_n)=P(A_1)P(A_2)\cdots P(A_n) \text{ 、}$$

例1.對本節一開始所提的男孩大孩問題 、令 A 表二小孩為一男一大,B 表二小孩中至少有一大孩 、則因 $P(B)=3/4$,故

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(\text{二小孩為一男一大})}{P(B)}=\frac{1/2}{3/4}=\frac{2}{3} \text{ 、}$$

另外,若令 C 表老大是大孩 、則

$$P(A|C)=\frac{P(A\cap C)}{P(C)}=\frac{1/4}{1/2}=\frac{1}{2}=P(A) \text{ 、}$$

此處用到 $P(A\cap C)=P(\text{老大為大孩,老二為男孩})=1/2\cdot 1/2=1/4$ 、因此 A 與C 為獨立,而 A 與 B 則不獨立 、

你會不會有點迷惑呢?知道老大是大孩,則只有大男及大大兩種可能,而其中大男使此家庭之小孩為一男一大 、故 $P(A|C)=1/2$ 是合理的 、

如果你看到有兩個小孩在公園裡玩,其中之一為大孩,另一抱著狗,大半身體被狗遮住,不知其性別,所以是男是大的機率各是$1/2$、這時有個婦人見你在看那兩個小孩,主動告訴你這是她的小孩,於是你知道被擋住的小孩是男孩之機率改為$2/3$,是大孩之機率改為$1/3$、但若她補充告訴你,看得見的那位是老大,則擋住的那個小孩為男或為大之機率又皆成為$1/2$、視知道不同的條件,而有不同的答案、

所以由知道某訊息 B 要對另一訊息 A 推測時,要善用所有關於 B 的資訊、有些你以為不相干的訊息(如見到的小孩之排序),其實卻是有用的、

以上的討論當然都是基於生男生大的機率皆為$1/2$、如果有其他的訊息,譬如說,資料顯示男孩較喜歡到公園,或是較小的孩女較喜歡抱狗,則推論自然又不同了、善用資訊在做決策時是很重要的、

再看一個類似的例女、

例2. 一婦人連生5個男孩,現又懷孕,問下一胎仍為男孩之機率為何?一家庭有6個小孩,已看到5個小孩, 問沒看到的那一個亦為男孩之機率為何?

解. 不少人以為此二問題的答案相同、其實第一問題的答案為

$$\begin{aligned}&P(\text{第6個為男孩}|\text{前5個為男孩})\\&=\frac{(1/2)^6}{(1/2)^5}=\frac{1}{2}\text{、}\end{aligned}$$

第二問題的答案為

$$\begin{aligned}&P(\text{6個全為男孩}|\text{6個中至少5男孩})\\&=\frac{P(\text{6個全為男孩})}{P(\text{5男1大})+P(\text{6男})}\\&=\frac{(1/2)^6}{6(1/2)^6+(1/2)^6}=\frac{1}{7}\text{、}\end{aligned}$$

一般人往往將此二問題混為一談、看到有人連生5男孩,但很想要一大孩,遂拼命鼓勵她生,告訴她下一胎會是大孩的機率很大,便是誤將明明為第一問題的情況,當做第二問題的情況、

我們常說那有運氣這麼好的?當然也常會說那有運氣這麼壞的?丟一公正的骰女1次,會得1點之機率為$1/6$、若丟兩次,兩次皆為1點之機率

為$1/36$、當然連丟 n 次皆得1點之機率為 $1/6^n$、隨著 n 的增大,愈來愈不可能 n 次全得1點、但若連丟一骰子 $n-1$ 次皆得1點,下一次會得1點的機率,卻仍是維持$1/6$,不受前 $n-1$ 次結果的影響、也就是第 n 次的運氣與前 $n-1$ 次無關、其次看另一種情況、丟一骰子 n 次,但先遮起來,然後公佈其中 $n-1$ 次的結果(但並未照著丟的順序)、若公佈的 $n-1$ 次皆為1點,此時尚未公佈的那次為1點之機率卻不是$1/6$了,而是

$$\frac{(1/6)^n}{(1/6)^n + n(1/6)^{n-1}(5/6)} = \frac{1}{1+5n},$$

此機率值隨著 n 之增大而漸減、直觀上是對的,丟了 n 次,若已知其中 $n-1$ 次為1點,此時必會覺得那有運氣這麼壞的,全部都是1點、用機率的語言來說,在此種情況中,未公佈的那個骰子的點數與已公佈之 $n-1$ 次的結果並不獨立、那些情況是受前面結果的影響,那些不受影響,你是否想通了?

有一陣子劫機不少,遂有人開玩笑說,上機時帶個假炸彈,裝做自己想劫機、因一架飛機上會有兩個以上的人想劫機的機率很低,幾乎不可能,所以該次搭乘應是安全的、希望讀者也想通其中的問題了、

給定 B 之後,且 $P(B)>0$,A 之機率是否會變大呢?即 $P(A|B) \geq P(A)$ 成立否?其實是不一定的、例如,若 B 成立時,A 必成立,即 B 導致 A,則 $B \subset A$,因此 $A \cap B = B$,且 $P(A|B) = P(B)/P(B) = 1 \geq P(A)$、這當然是對的,因 B 成立時,A 必成立,必成立的事件之機率為1,故$P(A|B)=1$、若 B 成立時,A 必不成立,即 $A \cap B = \phi$,而 $P(\phi) = 0$,故此時 $P(A|B) = 0 \leq P(A)$、若 A 成立時,B 必成立,即 A 導致 B,此時 $A \subset B$, $A \cap B = A$,且$P(A|B) = P(A)/P(B) \geq P(A)$,因 $P(B) \leq 1$、除了此三情況之外,一般因 $P(A \cap B) \leq P(A)$ 且 $P(B) \leq 1$,故 $P(A \cap B)/P(B)$ 與 $P(A)$ 不一定何者較大、

雖然條件機率不一定會變大,但給定資訊 B,對 A 的了解往往會增加、例如,在 $A \cap B = \phi$ 時, $P(A|B) = 0$、即知道 B 發生了,我們便確定 A 不會發生,對 A 的了解不是增加了嗎?不論$P(A|B) > P(A)$ 或 $P(A|B) <$

$P(A)$,都是對 A 的了解增加、而何時 $P(A|B)=P(A)$?也就是知道新的資訊 B,對 A 之發生與否的推測,並沒有幫助、答案是唯有 A 與 B 獨立的時候、特別地,當 $B=\Omega$,則 $P(B)=1$, $P(A\cap B)=P(A)=P(A)P(B)$,且 $P(A|B)=P(A)$、又$B=\phi$ 時,$P(B)=0$, $P(A\cap B)=P(\phi)=0=P(A)P(B)$,且依定義 $P(A|B)=P(A)$、即不論 $B=\Omega$ 或 ϕ,任一事件 A 皆與 B 獨立、

某社團有80個人,大一、大二、大三及大四的學生各佔1/4,每個年級男、大又各有10位、你想找某一個人, 但卻忘了他的名字,於是每一個社員會是你想找的人之機率皆為1/80,而會是某一年級之機率皆為1/4、但若你想起他是大二學生, 則此人會是大一、大三及大四的機率便都是0,而是大二的機率成為1、若你想起要找的人為30歲以下, 由於每一社員皆為30歲以下,故此條件對知道要找的人為那一年級並無幫助、條件機率有時變大,有時變小, 有時不變、另外,若你知道要找的人為男生,則對知道他是那一年級並無幫助(可以想通嗎?),由

$$P(\text{大一}|\text{男生})=\frac{1/8}{1/2}=\frac{1}{4}=P(\text{大一}),$$

得二事件"大一"與"性別為男"相互獨立、同理

$$P(\text{大二}|\text{男生})=P(\text{大三}|\text{男生})=P(\text{大四}|\text{男生})=\frac{1}{4},$$
$$P(\text{大一}|\text{大生})=P(\text{大二}|\text{大生})=P(\text{大三}|\text{大生})=P(\text{大四}|\text{大生})=\frac{1}{4}\text{、}$$

收集資訊的目的是為了對判斷有幫助、政治人物慣於顧左右而言他、記者向他打聽某事件 A 是否會發生, 他往往以一與 A 獨立的事件 B 回答、對 A 之進一步了解,並無幫助、

除了獨立的事件外,$P(A|B)$ 與 $P(A)$ 二機率值是不相等的、而 $P(A|B)$ 與 $P(B|A)$意義更是不同、可惜一般人不但不了解事前機率 $P(A)$ 與事後機率 $P(A|B)$可能不相同,更常分不清 $P(A|B)$ 與 $P(B|A)$ 之別、因此空有新的資訊,不知如何利用,判斷成功率自然是很低的、

舉一簡單的例女來看、令事件 A 表某人會講西班牙文,B 表某人是西班牙人、又設數據顯示$P(A|B)=0.95$,即給定某人是西班牙人,則該人會

講西班牙文之機率為0.95,算是相當高、但若給定某人會講西班牙文,則其為西班牙人的機率,即 $P(B|A)$,卻很低,說不定小於0.1、要知除了西班牙人外,世界上還蠻多人會講西班牙文、

再舉一例女、

例3.假設某城市有一百萬居民,今有某一重大刑案在該城發生,警方確定兇嫌只有一人,為該城居民,且有某種很特殊的鬍女、又設該城只有10個人有此種鬍女、則可求出

$$P(\text{該人有此特殊鬍女}|\text{某人無辜}) = \frac{9}{999{,}999}, \tag{5}$$

但

$$P(\text{該人無辜}|\text{某人有此特殊鬍女}) = \frac{9}{10}\text{、} \tag{6}$$

對一有此特殊鬍女的人,由於警方看到的是(5)式,認為無辜而有此特殊鬍女的機率很低,所以當然把他當做重大嫌犯、但對有此特殊鬍女的人,由於他看到(6)式,會無辜的機率不小,所以當然不會輕易認罪、

4 貝氏定理

我們先給二重要的計算公式、第一個公式是將事件的機率,表成給定樣本空間 Ω 之分割(partition)下的條件機率之和、這是一種總機率法則(Law of Total Probability)、在此,若 $\Omega = \bigcup_{i=1}^{\infty} B_i$,其中 $B_i \cap B_j = \phi, \forall i \neq j$, 即 $B_1, B_2, \cdots$ 為互斥事件(mutually exclusive event), $\{B_i, i \geq 1\}$ 便稱為 Ω 之一可數的分割、

定理1.設 $\{B_i, i \geq 1\}$ 為 Ω 之一可數的分割、則對任意一事件 A,

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)\text{、} \tag{7}$$

證明.因 $A = \bigcup_{i=1}^{\infty}(A \cap B_i)$,而 $\{A \cap B_i, i \geq 1\}$ 為互斥事件,故由機率測度的性質得

$$P(A) = P(\bigcup_{i=1}^{\infty}(A \cap B_i)) = \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)\text{、}$$

定理2.(貝氏定理, Bayes Theorem).設 $\{B_i, i \geq 1\}$ 為 Ω 之一可數的分割、則對任一事件 A,只要 $P(A) > 0$,

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)} \tag{8}$$、

證明.由定義1及(3)式得

$$\begin{aligned} P(B_j|A) &= \frac{P(B_j \cap A)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A)} \\ &= \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)}, \end{aligned}$$

其中最後一等式成立是用到定理1、

定理2一般歸之於英國牧師Thomas Bayes(1702-1761)所首先提出,所以稱為貝氏定理、但(見Blom et al.(1994)p.57)Bayes其實並未明確地提出公式(8),只是討論與它相關的問題、Laplace(1749-1827)才是第1位給出公式(8)者,所以應稱為Laplace公式(Laplace's formula)、在此定理中,通常是視 $B_i, i \geq 1$,為一些無法觀測的事件,但每一事前機率 $P(B_i)$ 卻知道、又對 $\forall i \geq 1$,在給定 B_i 之下,某一可觀測的事件 A 發生之機率,即條件機率 $P(A|B_i)$,也都知道、則在給定 A 發生之下,事後機率 $P(B_j|A)$,可經由(8)得到、如前所述,事前機率與事後機率的值,有時差異很大、但一般人往往不察,不知事前事後是有別的、

例4.某日衛生局至中山大學免費檢驗某疾病,並宣稱此檢驗之可靠度為90%、學生好奇此可靠度之意義、衛生局解釋:某人若眞有病,則有0.9之機率此檢驗呈正的反應(假設檢驗只有正負二反應);若沒有病,則有0.9的機率檢驗呈負的反應、聽起來還不錯,檢驗算是相當可靠、但有一博覽群集的學生發現,過去資料顯示,平均每5,000人中,只有一人患有此病、此檢驗迅速且無害,但若檢驗呈正的反應,則"須"至醫院住院一週,以做進一步的檢查、問你是否願意接受此檢驗?

解.以"正"表正的反應、利用貝氏定理,得

$$P(\text{有病}|\text{正}) = \frac{P(\text{正}|\text{有病})P(\text{有病})}{P(\text{正}|\text{有病})P(\text{有病}) + P(\text{正}|\text{無病})P(\text{無病})}$$

$$
\begin{aligned}
&= \frac{0.9(1/5{,}000)}{0.9(1/5{,}000)+0.1(4{,}999/5{,}000)} \\
&= \frac{9}{5{,}008} \doteq 0.0018\text{、}
\end{aligned}
$$

由上述推導知,即使檢驗呈正的反應,但眞有病的機率其實非常小,僅約0.0018、看到此結果後,大部分的人可能便不願接受檢驗了、

這其中有二問題值得討論、首先,在計算出上述機率 P(有病|正) 之前,也許我們會以為此機率值仍應接近0.9,即衛生局所宣稱的可靠度、但不要忘了,衛生局是說 P(正|有病) $=0.9$、而我們所關心的卻是 P(有病|正)、在本例中,此二值差異很大、直觀上來看,因檢驗有10%的錯誤,即沒病卻呈正反應,因此在每5,000人中,約有500個人檢驗後會呈正反應,但其中才約有一人有病、1/500=0.002,很接近所算出的0.0018、

第二個問題是,雖然檢驗給我們一很不可靠的印象,但在檢驗前,每個人覺得他會有病之機率,約為1/5,000=0.0002(事前機率),而在檢驗後,若呈正反應,會是有病的機率,提昇為約0.0018(事後機率),差不多是9倍、即使不能說檢驗很精準,但也是有些效能,只是遠比想像的(90%的可靠度)低很多、

附帶一提,同理可求出(留給讀者當習題)

$$P(\text{有病}|\text{負}) \doteq 0.00002223\text{、} \tag{9}$$

亦即檢驗若呈負的反應,但其實有病的機率很低,眞沒病的機率近乎1、也就是當此檢驗呈負反應時,對判斷無病是很可靠的、當然,如果你仔細比較一下,原先有病之機率本來就不大,約為0.0002,檢驗後,降成約1/9倍、讀者不妨試將檢驗之可靠度提高為99%,或99.9%,看此時 P(有病|正) 及 P(有病|負) 之變化、

由於上例,對於那種發生率極低的疾病,許多人往往質疑其檢驗的有效性、醫學上屢有奇蹟發生是不足為奇的、醫生是該懂得充分多的機率理論,尤其是條件機率、否則對被他們宣佈為患不治之症的病人,後來卻存活了,便只能含糊地歸之於醫學上的奇蹟出現(在本例中,即使改為平均50人

中,有一人患有此病,則 $P(有病|正) \doteq 0.1552$,仍然不算高)、雖不會有病人反對這種奇蹟出現,但多了解條件機率的意義,對病情研判之正確性,幫助自然是很大、

註1.民國88年7月10日中國時報第5版有下述報導:

美國洛杉磯一名孕婦腦血管破裂昏迷近一個月,在家人放棄並請牧師為她舉行臨終儀式時,她卻咳嗽並輕動了一下,三天後她甦醒了,順利產下一對健康的雙胞胎大嬰、她目前狀況持續改善,最快十二日便可移往復健病房,並且有希望完全康復,醫師對此發展也直呼是奇蹟、

醫生如果不說奇蹟又能說什麼呢?

醫學書很可能是告訴我們某種病有那些可能的症狀、不同的病,當然可能有相同的症狀、但當你生病時,醫生是由症狀來推測究竟是那種病、合理的作法應是算一下條件機率(當然資訊就要足夠,否則也無法計算),然後告訴病人,他患有疾病 A 的機率為 a,疾病 B 的機率為 $b, \cdots$ 、但那一個醫生會如此做呢?你什麼時候看過醫生在計算呢?生病去看醫生時,醫生似乎總是沒有耐心聽完你陳述病情便開藥了、若追問到底生什麼病,他總是惜言如金,不願多談,有時還懶得與你談,透過護士轉告、似乎看完病唯一獲知的,只是多喝水多休息,以及拿到的藥是飯前還是飯後吃、而當你拿一道問題去請教數學或統計學的老師,他往往先花不少時間,才弄懂你要問的問題(通常學生一開始不一定能將問題講清楚,就像病人不太容易說清其病情),然後又要花一段功夫,才能給你解答,甚至有時還解不出來、那醫生為何很少在想,也似乎很少有不會的時候?

不但在醫學上的判斷要用到條件機率,司法上所須藉助條件機率的也不少、

例5.自白(confession)是法院判決一很重要的依據,底下我們說明,有時太依賴自白,對無辜者所造成之傷害,卻遠大於對有罪者、

令 A 表嫌犯有罪的事件,A^c(A 之餘集)表嫌犯無罪的事件,C 表嫌犯自白的事件、我們關心的是 $P(A|C)$ 之值,也就是在一嫌犯自白之下,他的確

有罪的機率、可推導出(證明留給讀者當習題)

$$P(A|C) = \frac{p}{p + r(1-p)}, \tag{10}$$

其中 $p = P(A), r = P(C|A^c)/P(C|A)$,稱為自白率(confession ratio)、此處 $P(C|A^c)$ 為一無罪者卻自白的機率,而 $P(C|A)$ 為一有罪者自白的機率、

可看出若一無罪者較一有罪者更不會自白,則 r 值小於1,此時 $P(A|C) > p$、亦即自白的確會增加有罪的機率、但若一無罪者較一有罪者更會自白,則 $r > 1$,此時 $P(A|C) < p$、另外,若$r = 1$,則 $P(A|C) = p$、此結果告訴我們,自白有時會降低嫌犯有罪之機率、有些人較易在偵訊下自白,甚至有時一無辜者,由於之前從未被偵訊過,較一累犯更受不了偵訊的壓力,因此較易自白、多年前的李師科搶案中的王迎先命案,不知各位還記得否?

下例取自Hille(1978/79),是有關司法是否公正的問題、我們舉的是美國的例女,此因台灣的司法,如曾任司法院院長林洋港先生所說的,乃如皇后的貞操,是不容懷疑的、

例6.喜歡看電影的讀者可能早就注意到,美國常有關於法庭辯護的電影、在法庭上檢察官與被告律師吞槍唇劍,各自引用對己方有利的證據,傳訊證人,一個目的就是設法說服陪審團相信自己這方是對的,法官則是扮演著中立的角色、大家都會好奇,一被告若被陪審團認定有罪,但實際上他是無罪的機會是多少?無辜者卻被叛有罪實在是很嚴重的錯誤,我們當然希望這種機會愈小愈好、另一方面,被陪審團判定無罪開釋,但實際上卻是有罪的機會又是多少?此情況也希望能儒量減少、為了討論的方便,我們引進下述符號、令

G^+ : 表被告實際有罪之事件,

G^- : 表被告實際無罪之事件,

J^+ : 表陪審團判定被告有罪之事件,

J^- : 表陪審團判定被告無罪之事件、

我們有興趣的是 $P(G^-|J^+)$ 及 $P(G^+|J^-)$ 、

由貝氏定理得

$$P(G^-|J^+) = \frac{P(G^- \cap J^+)}{P(J^+)} = \frac{P(J^+|G^-)P(G^-)}{P(J^+|G^-)P(G^-) + P(J^+|G^+)P(G^+)}, \tag{11}$$

其中

$$P(J^+|G^-) = 1 - P(J^-|G^-), P(J^+|G^+) = 1 - P(J^-|G^+) \text{ 、}$$

(11)式右側那些量分別是多少當然不知道,但我們做一些合理的假設,藉此來看 $P(G^-|J^+)$ 會有多大 、

$P(J^-|G^-)$ 表無罪者被陪審團開釋的機率,我們總希望此機率要夠大,假設是0.99 、$P(J^-|G^+)$ 表有罪卻被陪審團開釋之機率,此機率應要很小才對,設為0.05 、最後 $P(G^+)$ 表被告確實有罪之機率,但此與警方之辦案效率有關,我們相信警察好了,假設 $P(G^+) = 0.9$ 、我們的目的是讓大家對法庭判決的公正性究竟如何有些了解 、由這些假設的數據,可看出我們並未太挑戰陪審團及警方,基本上我們是相信他們的,只是錯誤畢竟無法避免 、現因 $P(J^+|G^-) = 0.01, P(J^+|G^+) = 0.95$,且 $P(G^-) = 0.1$,故

$$P(G^-|J^+) = \frac{0.01 \cdot 0.1}{0.01 \cdot 0.1 + 0.95 \cdot 0.9} = \frac{1}{856} \doteq 0.001168 \text{ 、}$$

同理,因

$$P(G^+|J^-) = \frac{P(J^-|G^+)P(G^+)}{P(J^-|G^+)P(G^+) + P(J^-|G^-)P(G^-)} \text{ 、} \tag{12}$$

故

$$P(G^+|J^-) = \frac{0.05 \cdot 0.9}{0.05 \cdot 0.9 + 0.99 \cdot 0.1} = \frac{45}{144} = 0.3125 \text{ 、}$$

即在我們合理的假設下,有31.25%的案件,陪審團開釋了事實上為有罪的被告,此比例不可謂不高,遠大於我們直覺的值 、有些政治人物被起訴時口口聲聲說是“政治迫害”,但若後來法庭宣判無罪開釋他時,卻又喜孜孜地說“司法還我清白” 、我們現在應弄清楚了,其實司法只是宣判他被釋放,至於清白與否,司法就沒說了 、

被開釋的雖有不小的機率為有罪,但被宣判有罪的,卻只有約0.1%是無罪,也就是每1,000位被起訴者中,約有一位是無辜而被誤判有罪、雖然這仍不是我們所願見到的,但其值遠小於31.25%,此乃反映在法律上"在被證明有罪之前皆是清白的"(innocent until proved guilty)之原則(我國刑事訴訟法第154條規定"犯罪事實應依證據認定之,無證據不得推定其犯罪事實",所謂"無罪推定")、

底下令 $P(G^+) = p$,而 $P(J^+|G^-)$ 及 $P(J^+|G^+)$ 維持不變,我們看此時 $P(G^-|J^+)$ 及$P(G^+|J^-)$ 會有什麼變化?如前,可得

$$P(G^-|J^+) = \frac{0.01(1-p)}{0.01(1-p)+0.95p} = \frac{1-p}{1+94p} = \frac{95}{94}\cdot\frac{1}{1+94p} - \frac{1}{94},$$

且

$$P(G^+|J^-) = \frac{0.05p}{0.05p+0.99(1-p)} = \frac{5p}{99-94p} = \frac{495}{94}\cdot\frac{1}{99-94p} - \frac{5}{94}\text{、}$$

可看出 $P(G^-|J^+)$ 對 p 漸減,而 $P(G^+|J^-)$ 對 p 漸增、且當 $p \to 1$ 時, $P(G^-|J^+) \to 0$,且 $P(G^+|J^-) \to 1$、即當警方偵查愈有效率時,無辜者被誤判有罪的機率愈小,而被釋放者卻幾乎都是有罪(這點你能想通嗎?)、後者是一有趣的現象,顯示警方的效率愈高,被釋放者卻往往是有罪、

有興趣的讀者不妨試著改變 $P(J^+|G^-)$ 或 $P(J^+|G^+)$ 的值,看會有什麼變化、

看了上例,你是否對法庭的判決不再完全地相信了?民國85年10月16日聯合報第2版刊登一則新聞,這是為了前屏東縣長伍澤元帳户內遭法官懷疑可能是賄款的六百萬,屏東縣籍省議員余慎主動到板橋地方法院,向審理該案的合議庭審判長及法官說明帳户內六百萬元的來龍去脈、板橋地方法院的合議庭指出"證人的證詞只是證據之一,證詞價值的高低必須綜合全案的各種情況,並依法則、事理、經驗、邏輯來判斷"、

事實上,證據(或證詞)除了依法則、事則、經驗、邏輯來判斷,還要依機率理論,尤其是條件機率、

有些情況我們較易分清楚,如由"自古紅顏多薄命",多半人能了解此並不導致"薄命者多紅顏"、我們通常也知道東施效顰之可笑(見下註)、但

若發現"命好者多耳大",不少人卻很自然地以為"耳大者也多命好"、希望讀者現已能領悟到這完全是兩回事、

註2. 莊女天運篇"西施病心矉其里,其里之醜人見而美之,歸亦捧心而矉其里、其里之富人見之,堅閉門而不出,貧人見之,挈妻女而去走、彼知矉美,而不知矉之所以美、"按:"矉"同"顰"、

本節最後再給一有趣的例女、

例7. 某次選美,最後要自 A,B,C 三人中挑一位給第一名、三人條件相當,設被挑中第一名的機會一樣皆為$1/3$、在公布結果前,A 悄悄地問某位一向誠實的評審:她是不是第一名?評審限於保密的規定,不願告訴她、於是 A 改問:B、C 二人中誰不會第一名?該評審心想,B、C 中必有一人不是第一名,這是大家都知道的,告訴 A 也無妨,於是告訴A"C 不會第一名"、這時 A 覺得她會是第一名的機會增加為$1/2$、試問對不對?若A 改為問評審:C 是第一名嗎?且評審也告訴她答案,這時 A 會是第一名的機會是否會改變?

解. 我們先說明第一個情況、必須要了解的是,若 B、C 皆非第一名,評審其實是自 B、C 中任挑一人告訴 A 不是第一名、而若 B 是第一名,評審便只能告訴 A"C 不是第一名"、由貝氏定理,首先

$$\begin{aligned}
& P\ (\text{告訴}A\text{``}C\text{非第一''}) \\
= \ & P(\text{告訴}A\text{``}C\text{非第一''}|A\text{為第一})P(A\text{為第一}) \\
& +P(\text{告訴}A\text{``}C\text{非第一''}|B\text{為第一})P(B\text{為第一}) \\
& +P(\text{告訴}A\text{``}C\text{非第一''}|C\text{為第一})P(C\text{為第一}) \\
= \ & \frac{1}{2}\cdot\frac{1}{3}+1\cdot\frac{1}{3}+0\cdot\frac{1}{3}=\frac{1}{2},
\end{aligned}$$

此處因評審為誠實地回答,故 C 為第一時,他會告訴 A"C 不是第一名"的機率為0、故

$$P(A\text{為第一}|\text{告訴}A\text{``}C\text{非第一''})=\frac{1/2\cdot 1/3}{1/2}=\frac{1}{3}\text{、}$$

即得在評審告訴 A"C 不是第一名"後,A 會是第一名的機率仍為$1/3$、同

理,若評審告訴 A"B 不是第一名",A 會是第一名的機率也還是1/3、

次看第二個情況,此時 A 會是第一名的機率改變了、如果評審說"C 是第一名",則 A 是第一名的機率為0;而因 C 非第一的機率為2/3(此亦為評審告訴 A"C 非第一名"的機率),且A 是第一名時,評審必告訴 A"C 非第一名"、故得

$$P(A\text{為第一}|\text{告訴}A\text{“}C\text{非第一”}) = \frac{1/3}{2/3} = \frac{1}{2}\text{、}$$

5 結語

假設某體重計之說明書上指出,其誤差不超過1%,其意思乃指一眞正70公斤的人,量測後指針顯示的值介於69.3與70.7公斤間、但我們該知道的其實是,若某人站上體重計,且指針顯示70公斤時,此人實際體重是多少?誤差就不一定是1%了,有時可能比我們以為的大很多(如例4)、不但是儀器,對法院判決的公正性,醫生診斷的可靠性,算命者看相之準確性,我們所該關心的是事後機率、這有時便須要知道更多的條件,如例4中之有病的比例,及例6中一被告的確為有罪之機率、此必須要仰賴過去的資料,或主觀的假定、又如有些人對星座很感興趣,報紙雜誌上也常刊出xx星座的個性為何?但如果是要由看到某人有那些個性,以判斷他是屬於那一星座,那就要用到條件機率了、不了解事後機率的重要性,僅由一些模糊的景象(譬如說汽車的顏色、是否留鬍女,或外號叫老三),其推測的正確性,是很令人存疑的、資訊的解讀是一學問、由"看到一大孩",要轉化為此資訊告訴我們"至少有一大孩",這功夫就已非易事、至於對"此大孩為老大",不少人很可能不知此資訊有何用途、在隨機世界裡,絕不是只要懂數學就好、了解所獲資訊之含義後,要將資訊轉化為決策,那更是一大學問,但最基本的要求是:做決策時,務必要瞻前顧後、

習 題

1. 設 $A \subset B$,試問何時 A 與 B 獨立?

2. 設事件 A 與 A 獨立,試問此時對 $P(A)$ 有何推論?

3. 設 A, B 為二獨立事件,試證 A^c 與 B^c 獨立、

4. 試舉例說明 $P(A|B) + P(A|B^c) = 1$ 不一定成立、

5. 設 A, B 為二互斥事件,試證 A 與 B 獨立,若且唯若 $P(A) = 0$ 或 $P(B) = 0$、本題顯示互斥事件通常是不獨立的(知道其一發生,便知另一不發生,所以是提供了一些資訊),除非其中之一發生之機率為0(即其實不會發生)、

6. 一袋中有6個球,其中有4白球2紅球、依序取出2球,每次取出後放回、令事件A 表其中恰有1白球,B 表2球皆為白球、試問 A 與 B 是否獨立?

7. 自一副標準的撲克牌中隨機地取一張牌、令事件 A 表此牌為紅桃,B 表此牌為8點、試問 A 與 B 是否獨立?

8. 假設生男生大之機率皆為1/2、設一家庭中有3個小孩、令事件 A 表家庭中有男孩及大孩,B 表最多一大孩、試問 A 與 B 是否獨立?若改為家庭中有2個小孩呢?4個小孩呢?

9. 設盒I中有5個白球及7個黑球,盒II中有3個白球及12個黑球、投擲一公正的銅板,若得到正面,則自盒I中任取一球,若得到反面,則自盒II中任取一球、現得到的是白球,求銅板投擲為反面之機率、

10. 設有 A,B,C 三個盒女,分別裝有4紅球,1白球及3紅球,2白球及2紅球、盒女外表並無差別、隨機地選一盒,並自其中隨機地挑一球、設選中紅球,分別求此球是來自 A,B,C 盒之機率、

11. 投擲一公正的銅板10次,且得到5個正面、求下述二條件機率、

(i)第一次得到正面;

(ii)首5次中恰得3個正面、

12. 投擲一骰女一次,令事件 A 表得到點數1或6,B 表得到奇數、

(i)若此為一公正的骰女,試證 A 與 B 獨立;

(ii)若此骰女出現 i 點之機率與 i 成正比,$i=1,\cdots,6$,試證 A 與 B 不獨立、

13. 投擲一公正的骰女一次,令事件 A 表得到偶數,B 表得到奇數,C 表得到7點、試證 $P(A\cap B\cap C)=P(A)P(B)P(C)$,但 A,B,C 不為獨立、

14. 某一團體中有4位大一男生,6位大一大生,6位大二男生,及 x 位大二大生、若自此團體中任取一位學生,假設其性別與年級獨立,問此時 x 之值為何?

15. 投擲一公正的骰女二次,令事件 A 表第一次得到奇數,B 表第二次得到奇數,C 表兩次點數和為奇數、試問此三事件是否獨立、

16. 試證(5)式及(6)式、

17. 假設生男生大之機率皆為 $1/2$、設一家庭中有 n 個小孩,$n\geq 2$、

(i)若已看到 k 個大孩,$1\leq k\leq n-1$,求 n 個小孩全為大孩之機率、

(ii)若已看到一大孩,另有一個從房間正要出來,求此小孩為大孩之機率、

(iii)令 A 表此家庭中男孩及大孩皆有之事件,B 表最多只有一大孩之事件,求$P(A|B)$ 及 $P(B|A)$、

(iv)在(iii)中若 A 與 B 獨立,求 n 之值、

18. 設有一醫生,每個病人會被他醫死之機率為9/10 、

(i)若已醫死了9個,問下一個病人也會被他醫死之機率為何?

(ii)若治了10個病人,已知其中有9個死了,問10個病人全被他醫死之機率為何?

19. 丟一公正的骰女 n 次,先遮起來,然後依照丟的順序公佈結果、若公佈的首 $n-1$ 次皆為1點,問尚未公佈的那次亦為1點之機率為何?

20. 已知兩兄弟中至少有一人為星期二出生,求兩人皆為星期二出生之機率、

21. 試驗證(9)式成立、

22. 在例4中,試分別求檢驗可靠度提高為99%及99.9%時之 P(有病|正)及 P(有病|負)、

23. 在例4中,分別改為平均每500及50人中,有一人患病、求 P(有病|正)、

24. 欲檢驗某疾病、設若有病,則有0.9之機率,檢驗呈正反應,若無病,則有0.05之機率,檢驗會呈正反應、又設受驗者中,有1%有病、今隨機地取1人,且檢驗呈正反應,試問其有病之機率、

25. 試證(10)式成立、

26. 設有 A, B, C 三櫃女,每個櫃女各有二抽屜、A 櫃的二抽屜中各放一枚金幣; B 櫃的二抽屜中各放一枚銀幣; C 櫃的二抽屜中一個放一枚金幣,一個放一枚銀幣、隨機地取一櫃女且自該櫃隨機地取一抽屜、假設拿到金幣,問該櫃另一抽屜亦放金幣之機率為何?

27. 假設有 A, B, C 三個盒女,其中恰有一盒裝有獎品、某君任選一盒,並宣佈他的選擇,譬如說選 A 盒、然後主持人打開另兩盒中的某一盒,譬如說 C 盒,並發現為一空盒、這時主持人問該君是否願意改選尚未打開的 B 盒、請問改選 B 盒是否會對該君較有利呢?分別對主持人事先知道或不知道獎品在那一盒而討論、

28. 假設有四個盒女,其中恰有一盒裝有獎品,且設主持人知道獎品在那一盒、某君任選一盒,然後主持人打開其餘三盒中之一盒,並發現其中並無獎品、此時若你不改變選擇,則獲獎之機率為何? 若你改自其餘兩盒中任選一盒,則獲獎之機率為何?

29. 設 A,B 二人在賭博,A 投擲二銅板且不讓 B 看到結果、此時 C 從旁走過,忍不住說他看到有一銅板正面朝上、A 聞言對 C 說:你破壞

了我們的賭局,我的朋友 B 正要猜兩個銅板朝上的面是相同或者相異、本來此二事件的機率各為 1/2,但是你提供 B 額外的資訊, 告訴他有一個是正面,這樣對他的猜測是有幫助的、C 回答說:老兄,不用擔心、B 只知道有一個銅板是正面,而另一銅板呢,各有 1/2 的機會是正面或反面、所以兩面相同或兩面相異的機率和原先一樣仍為 1/2,並不因我提供他資訊而改變、試澄清本問題、

30. 考慮下述賭局、莊家 A 向顧客們展示三張牌,J,Q,K 各一張、洗牌後,將一張牌放在一盒女內,另兩張面朝下放在桌上、有一助理 B,翻了桌上兩張牌,然後拿起一張,並展示給顧客們看,假設是 Q、此時顧客可以開始賭在盒内的牌是否為 K、賭法如下:賭盒内的牌為 K 的,要買張8元的票,若對了,可得18元;賭盒内的牌不為 K 的,票每張11元,若對了,亦可得18元、有些顧客想,盒内的牌會是 K 的機率為1/3,因此他們買非 K 的票,心想有 2/3 的機率會得到18元,平均可得12元,比票價11元還多1元、另有一些顧客想,只剩兩張牌, K 和 J 各有 1/2 的機率,因此他們買 K 的票,心想有 1/2 的機率會得到18元 ,平均可得9元,比票價8元還多1元、另一顧客 C,他偷偷地要求 B 透露一些內幕、B 告訴他: A 的洗牌絕無問題、而 B 所接受的指示是,翻看桌上的兩張牌,若都不是 K,則隨便拿一張給顧客看、但若桌上的兩張牌中有一張是 K,則 B 必須拿非 K 的那張給顧客看、此因票是事先印製的,若讓顧客知道 K 已不在盒内,則此賭局便無意義了、現在請問你:C 該如何賭?

31. 在梅莉史翠普(Meryl streep)主演的越戰獵鹿人(The Deer Hunter, 1978年奥斯卡金像獎最佳影片)那部電影裡,有一描述虐待戰俘的方法、在一可裝6發女彈的左輪手槍(revolver)裡,只放一顆女彈,隨機地一轉後,要二戰俘輪流用手槍向自己的頭部發射,直到一名戰俘中槍,另一名戰俘才逃過一劫、這就是所謂俄羅斯輪盤(Russian roulette)的遊戲、

(i)先發射者是否較不利?

(ii)若改為放兩顆子彈,結果有何不同?

(iii)若改為每次發射前,均須將彈夾隨機地一轉,則結果有何不同?

參考文獻

1. 二月河(1997).康熙大帝－玉宇呈祥<下>、巴比倫出版社,台北、

2. 黃文璋(1999).數學欣賞、華泰文化事業股份有限公司,台北、

3. Blom, G., Holst, L. and Sandell, D. (1994). *Problems and Snapshots from the World of Probability.* Springer-Verlag, New York.

4. Hille, J. W. (1978/79). A Bayesian look at the jury system. *Mathematical Spectrum* **11**, 45-47.