條件機率的應用

譚克平
國立台灣師範大學科學教育研究所

在一般的數學教科書中，關於機率的介紹，常採用丟硬幣、擲骰子或抽撲克牌等作爲例子，雖然可以方便教導機率各有關概念，但也常讓學生認爲機率只是在做賭博遊戲時才派上用場的錯覺。其實機率可以十分實用，它不但可以應用到日常推理思考的層面，而且還可以應用到科學推理之上。本文假設讀者對機率這概念已有基本的認識，首先會從介紹條件機率的定義入手，進而介紹一些實際應用的例子，最後還會從 Popper 的論著中，介紹機率在科學哲學中可以應用的例子。

一、基本概念

首先，讓我們以下列的列聯表來介紹機率與條件機率的基本意義。假設某廠商爲新產品做了一系列的廣告推銷之後，接著想了解一下該廣告系列的成效，故在某百貨公司屬下的攤位進行一次調查，結果在受訪的一百人之中，發現當中有 60 位是因爲看到廣告之後而來的，另外 40 位則未曾看到該廣告系列。在受訪者中，共有 40 位顧客購買了該新產品，而其中有 30 位表示是看過廣告之後才來的。

根據以上的資料，我們可總結出以下的列聯表以做進一步的討論。

表一

	有看過廣告	沒看過廣告	共
有購買產品	30	10	40
沒購買產品	30	30	60
共	60	40	100

倘若在這群受訪者中，隨機抽出一位，則此人有購買新產品的機率是多少？

按照古典的機率定義　，亦有人稱之爲分析性的定義（analytical definition），某事件 A 發生的機率等於事件 A 可能發生的次數，除以此事件所屬的樣本空間 S 的樣本數。如果用符號表示，設 P(A)代表事件 A 之機

率，n(A)代表事件 A 可能發生的次數，n(S)代表樣本空間的樣本數，則

$$P(A) = \frac{n(A)}{n(S)} \qquad \text{............}\ (1)$$

在式（1）中，重要的是弄清楚事件 A 所對應的樣本空間 S 的性質。在這個例子中，整體的樣本空間 S 是指所有的受訪者；而事件 A 是指當中符合「有購買新產品」這種情況的受訪者，故 $P(A) = \frac{40}{100} = 0.4$。

同理，如果事件 B 是指「沒看過廣告但卻有購買新產品」這種情況，則按表一中，共有 10 人符合上述的規定，故

$$P(B) = \frac{n(\text{未看廣告且購買產品})}{n(S)} = \frac{10}{100} = 0.1$$

進一步來說，如果在這些受訪者中隨機抽出一位，而且已經得知對方沒有看過廣告。在這預先設定的條件之下，要估計此人有購買產品的機率，可用下述之條件機率（conditional probability）這個概念來運算。

設已發生的事件爲 D，現欲求在知道條件 D 發生了的情況下，事件 C 所發生的機率，我們可用符號 P(C｜D)表示此一機率，式中的直線代表「已經知道」之意。此一情況其實相當於要求事件 C 與 D 同時發生的機率，但所相對的樣本空間卻需先調整爲符合事件 D 的樣本集合，這是因爲我們已經知道事件 D 已發生這個先決條件，既然有更多的資料，故此對應的樣本空間便縮小了。即

$$P(C \mid D) = \frac{n(C\text{且}D)}{n(D)} \text{............}(2)$$

針對目前的例子，D 指已知對方「沒看過廣告」，C 指「有購買產品」，則

$$P(\text{有購買產品} \mid \text{沒看過廣告}) = \frac{\text{沒看過廣告並有購買產品的人數}}{\text{沒看過廣告的人數}} \text{.........}(3)$$

經運算後得　　$P(\text{有購買產品} \mid \text{沒看過廣告}) = \frac{10}{40} = 0.25$

由此可見，P(有購買產品｜沒看過廣告) 並不等於 P(沒看廣告卻有購買產品)。

式（2）是以符合事件的樣本數來求得條件機率，除此之外，也可用事件發生的機率來求得條件機率。在式（3）的右端，若將分子及分母同時除以受訪者總人數，得

$$P(\text{有購買產品} \mid \text{沒看過廣告}) = \frac{10/100}{40/100}$$

因事件「沒看過廣告」所對應的總體樣本空間是所有受訪者，故上式的分母等於 P(沒看過廣告)， 同理，上式的分子相等於 P(「沒看過廣告」且「有購買產品」），即

$$P(\text{有購買產品} \mid \text{沒看過廣告}) = \frac{P(\text{沒看過廣告且有購買產品})}{P(\text{沒看過廣告})}$$

倘若我們將這結果一般化，即可得 $P(C \mid D) = \dfrac{P(C\text{ 且 }D)}{P(D)}$ ……………(4)

二、貝氏定理

根據式（4），得知 $P(D \mid C) = P(C\text{ 且 }D) / P(C)$。一般而言，$P(C \mid D)$ 並不等於 $P(D \mid C)$，那麼，我們會追問 $P(C \mid D)$ 與 $P(D \mid C)$ 之間有什麼關係？

從式（4）得知 $P(C\text{ 且 }D) = P(C \mid D) \times P(D)$，同理也可以導出

$P(C\text{ 且 }D) = P(D \mid C) \times P(C)$，即 $P(C\text{ 且 }D) = P(C \mid D) \times P(D) = P(D \mid C) \times P(C)$

故此有

$$P(C \mid D) = \frac{P(D \mid C)P(C)}{P(D)} \quad \text{……………..(5)}$$

式（5）是貝氏定理其中一種呈現方式，在此式中，P(C) 可稱爲 C 的先驗機率（a prior probability），因爲 P(C) 是沒有參考資料 D 之前便以斷定下來的，而 $P(C \mid D)$ 則稱爲 C 的後設機率（posterior probability），因這是在得知資料 D 之後才斷定 C 發生的機率。

此外，還有一個很有用的全機率定理 (Theorem of total probability)，現在以前述例子作基本介紹。

從表一中，讓我們集中注意在那 40 位沒看過廣告者身上，他們只有兩種可能的情況，一就是他們有買新產品，不然就是沒有，他們僅僅可能有這兩種互相排斥的情況，所以，如果我們要計算受訪者沒有看過廣告的機率，這應相等於沒有看過廣告而有買產品的機率和沒有看廣告也沒有買產品的機率的總和，即

$$\begin{aligned} P(\text{沒看過廣告}) &= P(\text{沒看過廣告 且 有購買產品}) \\ &\quad + P(\text{沒看過廣告 且 沒購買產品}) \quad \text{------(6)} \end{aligned}$$

這點也可藉由數據直接得知：

$$P(\text{沒看過廣告 且 有購買產品}) + P(\text{沒看過廣告 且 沒購買產品})$$
$$= \frac{10}{100} + \frac{30}{100} = \frac{40}{100} = P(\text{沒看過廣告})$$

此外，如果直接應用式(4)的話，我們可得知 P(沒看過廣告 且 有購買產品) = P(沒看過廣告｜有購買產品) × P(有購買產品)；同樣的道理，P(沒看過廣告 且 沒購買產品) = P(沒看過廣告｜沒購買產品) × P(沒購買產品)，因此，式(6)可改用下式表示：

$$\begin{aligned} P(\text{沒看過廣告}) &= P(\text{沒看過廣告} \mid \text{有購買產品}) \times P(\text{有購買產品}) \\ &\quad + P(\text{沒看過廣告} \mid \text{沒購買產品}) \times P(\text{沒購買產品}) \end{aligned}$$

如果我們把這個觀察符號化，沒看廣告這事件以 A 作代表，用 B_1 代表有購買產品，而 B_2 代表沒有購買產品的話，則上式可改寫爲

$$P(A) = P(A \mid B_1) \times P(B_1) + P(A \mid B_2) \times P(B_2) \quad \text{- - - - - - - - - - - -} (7)$$

這就是全機率定理的一種表示方式。

再者，我們還可將式(7)更爲一般化，如果有一組事件 B_1 及 B_2 兩兩互斥，且 B_1 及 B_2 爲樣本空間的一個有限劃分，一則相對於任一屬於同樣本空間的事件 A，全機率定理的一般表示方式爲

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i) \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

將此方程式帶入貝氏定理中，我們可得到貝氏定理的另一種常見的表示方式：

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{i=1}^{n} P(A \mid B_i)P(B_i)}$$

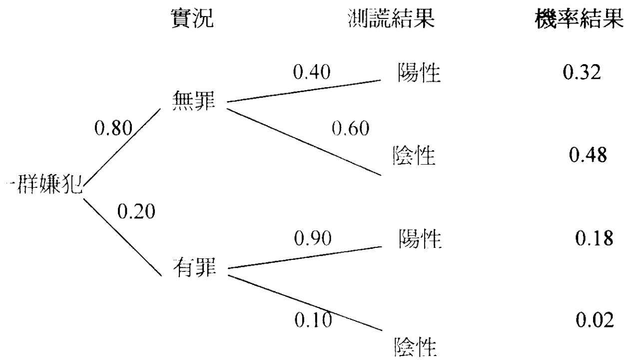
有了以上的基礎之後，接著讓我們來探討條件機率的一個實際應用例子。

三、測謊機的例子

在現代社會中，測謊機已是警方辦案時常用的工具，但是測謊機有多準確，現在讓我們以簡化的方式來研究一下。

假設按照過去的記錄，某受過訓練的測謊機操作員在使用該機器時，如果受試者真的有犯案的話，則測謊機呈現陽性反應的次數高達 90%，但如果對方沒有犯案，則呈陽性反應的次數還是達 40%。另外，假設我們擁有絕對的資料，得知某一群嫌犯中有 20%是有罪的，而其他 80%卻是無辜的。這假設在現實生活中當然是不可能的，但爲了討論之故，且暫時接受這個假設。

如果我們讓這群嫌犯接受測謊機的試驗，那麼當一個嫌犯的測試結果呈陽性反應時，他真的有犯案的機率爲何？爲了要方便解答這個問題，我們可以建立如下的機率樹狀圖(tree diagram)，從中得到幫助。



圖一　測謊機機率樹狀圖

在圖一中，機率結果一欄中的數據，是採用條件機率規則而得到的結果，比如機率等於.32 是指無罪而測謊機呈陽性反應的機率，它的符號表示為 P(無罪 且 測試呈陽性反應)，由式(4)得知

$$\begin{aligned}&P(\text{無罪 且 測試呈陽性反應})\\&=P(\text{陽性} \mid \text{無罪})\times P(\text{無罪})=(0.40)(0.80)=0.32\end{aligned}$$

其餘機率亦可如此類推。

為了要考慮測謊機的準確性，我們需要計算出 P(陽性｜無罪)的值。首先，我們可以應用全機率定理來計算測驗呈陽性的機率：

$$\begin{aligned}&P(\text{陽性})=P(\text{陽性} \mid \text{無罪})\times P(\text{無罪})+P(\text{陽性} \mid \text{有罪})\times P(\text{有罪})\\&=(0.40)(0.80)+(0.20)(0.90)=0.50\end{aligned}$$

在推算出 P(陽性)的數據之後，我們便可應用貝氏定理來推算測謊機的準確性了，即 $P(\text{有罪} \mid \text{陽性})=\dfrac{P(\text{陽性}|\text{有罪})}{P(\text{陽性})}=\dfrac{(0.90)(0.20)}{0.50}=0.36$。這個機率是出奇的小。我們最需要注意的是在一般情況之下， P(陽性｜有罪) 並不等於 P(有罪｜陽性)，因此，雖然 P(陽性｜有罪) 高達 90%，但這並不等於測謊呈陽性時嫌犯有犯罪的機率也高達 90%。所以在司法的過程中，還需要有其它資料以做佐證，這一點是極為重要的。這類的推論也可應用到 AIDS 的檢驗中。若測謊機及愛滋病的檢驗系統還未達到絕對準確的地步的話，就需要留意到 $P(A \mid B)$及 $P(B \mid A)$ 很可能不相等的情況。

當然，上述的推算中也不無弱點，例如，我們如何確定在受過訓練的測謊機操作員手中，對於一個真的有犯罪者，在受測時會有 90%的機率呈現陽性結果，這些數據雖有一點可爭議性，但從上述的例子中，對於條件機率的應用及限制，應該可見一斑。

四、科哲中的應用

除此之外，條件機率在一些哲學論點上的推論中也可派上用場。現在讓我們參考一下波普(Popper)對科學的一個看法，看看條件機率在他的論點中佔有怎樣的一席之地。因為篇幅所限，以下只是一個扼要的介紹，有興趣的讀者敬請參閱他的原著。

波普的看法是認爲科學不是歸納的，他覺得要從觀察的結果中歸納出一個涵蓋觀察本身的概論(generalization)，無論我們如何選擇那些觀察，都不能夠增加那概論成立的機率。波普這個見解是建基於他對每一個的先驗機率(prior probability)必須爲零的假設。換句話來說，他覺得對一個廣泛適用的概論，在未蒐集任何證據之前，這概論爲真的機率爲零。若以符號 S 表示概論，則 $P(S) = 0$。無論我們是否同意波普這個看法，我們暫且接受這點，先看看波普接下來的論點。根據這個觀點，可應用貝氏定理來引證波普對無論擁有多少觀察到的證據，皆不能提高其歸納出來的概論爲真的機率。

現簡化其論證如下，因爲 $P(S \mid E) = \dfrac{P(E \mid S)P(S)}{P(E)}$，既然 E 是概論 S 的一個特例，故 $P(E \mid S) = 1$。另外假設證據 E 絕對爲真，即 $P(E) = 1$，因前述之 $P(S) = 0$，故明顯得知 $P(S \mid E) = 0$。

在這個推論中，雖然我們只提到一個證據 E，但同樣的推論也可以用在超過一個證據之上，其所得的結果將會相同，即無論證據有多少，均不能增加涵蓋這些證據的概論爲真的機率！這就是波普並不接受歸納法的其中一個原因。

對波普來說，科學的發展是靠提出假設以及反駁假設的過程而進展的，也就是說，科學是靠演繹而進展的。波普的論證自有其弱點，例如有些人士並不接受他對任何一個概論的先驗機率爲零的假設。有興趣的讀者，可閱讀參考文獻中所列舉的書籍，以作進一步的了解。

本文的目的在介紹條件機率的概念，並介紹其眾多應用中的一些例子。條件機率就好比一條小小幽徑，可以引領我們走得很遠。

參考文獻

[1] Popper, K. R. (1968). The logic of scientific discovery, London: Hutchinson.
[2] Newton - Smith, W. H. (1981). The rationality of science, London: Routledge ,Kegan Paul.
[3] Lakatos, I. (1968). The problem of inductive logic, North - Holland Publishing Company.