可能性理論淺介

楊敏生　劉曼君

模糊理論 (fuzzy theory) 的發展和應用, 一路行來已是 30 年。在這 30 年中, 模糊理論帶給工業界和工程界一番嶄新的氣象, 特別是近年來與神經網路結合, 更展現出一種革命性的面貌。由於根據馮紐曼 (J. von Neumann) 原理所設計的電腦在思維及學習等方面遭遇瓶頸, 神經網路已成爲電腦人腦化的最佳途徑, 特別是與處理自然語言、知識表示等的最佳工具 — 模糊理論結合後, Neuro & Fuzzy 將發展成爲新一代的電腦 (或許可稱之爲第六代電腦)。模糊集合不僅跳出了傳統集合的排中律限制, 可將人類思維和概念之過渡邊界以數學方式表達及運算; 另一方面也擴展了不確定性 (uncertainty) 的現象, 將隨機性所無法表達的另一種稱之爲模糊性 (fuzziness) 的不確定性呈現出來。隨機性可透過機率測度來衡量, 相對地, 模糊性卻透過所謂的可能性 (possibility) 測度來得到, 因此, 可能性理論就成爲與機率論平行前進的另一種處理不確定性的理論。到底什麼是可能性? 它和我們所熟悉的機率 (probability) 又有什麼不同呢? 在數學傳播十八卷一期的「模糊理論簡介」(參考文獻 [1]) 一文中曾對模糊理論作了介紹, 在本文中, 我們將再進一步探討可能性理論。

自從 Wiener 和 Shannon 利用機率統計的概念發展出通訊理論, 無論是編碼或是資料的傳輸和接收, 都與機率理論息息相關。時至今日, 我們不僅要求資料能正確傳輸, 更希望能解讀其中的涵義, 以將這些資料作更進一步的利用。因爲許多資料是以自然語言的方式表示並儲存, 而語言中正有其模糊性存在, 這可說是模糊理論發展的重要因素之一。而所謂的可能性理論是模糊測度的主要理論, 架構上是以模糊集爲基礎; 探其背後思想, 正與自然語言所傳達的信息有密切關係。

數學中的特徵函數 $I_A(x)$ 只有0和1兩個值, 當 $x \in A$ 時 $I_A(x) = 1$, 當 $x \notin A$ 時 $I_A(x) = 0$。將其一般化之後, $I_A(x)$ 擴展爲 [0,1] 區間連續函數 $\mu_A(x)$, 即若 $x \in X$, $\mu_A(x) \in [0,1]$, 函數 $\mu_A(x)$ 就是所謂的隸屬函數 (membership function)。通常我們將一種模糊的概念以隸屬函數來表示, 如: 年輕、美麗等, 而隸屬函數值的大小則表示隸屬程度的不同。模糊集合的概念最早由扎德 (L. A. Zadeh) 在1965年提出, 原始的想法便是在處理自然語言 (natural language) 上的模糊 (不確定) 性, 希望以嚴謹的數學方式描述模糊的現象。但爲什麼不能以機率來描述這些不確定性呢? 主要就在於機率測度上「可加性」(additivity) 的限制, 在自然語意的表達與處理上是不必要的。怎麼說呢? 例如, 在語言的表達上, 我們說這個人看起來還蠻年輕的, 或說這個人是年輕的程度爲0.8, 如

此並沒有意味著說他「不年輕」甚至「年老」的程度是0.2的意思。但如果把「年輕」這樣一個概念用機率來表達的話, 就只能說他「年輕」的「機率」是0.8,「不年輕」的「機率」必是0.2了。這和一般的語言習慣確實有出入, 可見以機率來表達語言上的模糊性是不恰當的。

事實上, 機率和可能性都是用來處理不確定性, 差別在於前者探討的是發生與否的不確定, 也就是隨機性; 而後者執掌的則是來自於模糊性的不確定。什麼是具模糊性的不確定呢? 就以下面兩種情況來描述。一是自然語言中的不確定 (例如: 年輕、美麗、熱等), 另一種是因測量儀器不夠精密而產生的不精確 (imprecision), 例如, 身高測量值爲170公分的人, 他「眞正」的高度可能介於169.5到170.5公分之間。這兩種情況都是典型具模糊性之不確定的類型。我們可以說, 可能性測度主要便是處理機率測度所不能測量的另一部分不確定性。

雖說可能性的本質和機率不同, 但是, 無可諱言的, 兩者的結構間卻有頗多共通處, 例如: 隸屬函數不免令人聯想起機率密度函數; 而隸屬程度和機率值同樣介於0和1之間, 等等。下面讓我們來看看機率空間 (Ω, F, P) 及可能性空間 (Γ, G, π) 的定義。

(1) 機率空間 (Ω, F, P)

Ω: 樣本空間

F: 所有事件所成的集合

P: 機率測度, 必須滿足以下三個性質:

(a) $P(\emptyset)=0$, $P(\Omega)=1$;

(b) $0 \leq P(A) \leq 1$ 其中 $A \in F$;

(c) 若 $A_1, A_2, \ldots \in F$ 且互斥, $\bigcup_{k=1}^{\infty} A_k \in F$, 則

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

(2) 可能性空間 (Γ, G, π)

Γ: 圖訊空間, 和樣本空間 Ω 一樣, 會隨實驗不同而變動, 例如在處理字元辨識問題時, Γ 便是所有可能符號所成的集合。

G: Γ 中所有子集所成的集合。

π: 可能性測度, 必須滿足以下兩個性質:

(a) $\pi(\emptyset)=0$, $\pi(\Gamma)=1$;

(b) 對集合 $A_\alpha \in G$ 所成的任意 (有限、可數或不可數) 集合,

$$\pi(\bigcup_{\alpha} A_\alpha) = \sup_{\alpha} \pi(A_\alpha);$$

由以上機率空間和可能性空間的比較中, 我們可以發現, 機率測度中最重要的可加性條件在可能性測度上不見了! 代之而起的是 sup, 限制條件明顯地放寬了。

機率有機率分配, 是經由機率空間透過隨機變數產生。相同地, 可能性亦有可能性分配 (possibility distribution), 它乃是由可能性空間透過模糊變數而產生。什麼是模糊變數 (fuzzy variable) 呢? 依 Zadeh[2]的觀點, 不妨以「模糊限制」(fuzzy restriction) 的角度來看。我們以最簡單的敍述形式「X 是 F」來說明, 並以「小明是年輕的」爲例。敍述中, X(如: 小明) 是變數, 在圖訊空間 Γ(如:[0,100], 表年紀) 中取值; F 就是對 X 的模糊限制 (如: 年輕)。於是,「X 是 F」就可以轉爲「$R(X) = F$」或「$R(A(X)) =$

F」的形式, 其中 $A(X)$ 表 X 的屬性; 而 $R(X)$ 和 $R(A(X))$ 意思相同, 都是對 X 的限制。可以這麼寫:

「小明是年輕的」→
「R(小明)=年輕」或「$R(A$(小明)) =年輕」

式中「年紀」就是小明的屬性,「年輕」則是模糊限制; 也就是說, 小明透過「年紀」這個屬性和「年輕」這個模糊限制產生關聯。可能性空間 (Γ, G, π) 經過特殊的模糊限制 F 後定義出模糊變數 X, X 是一個實値函數, $X: \Gamma \to \mathcal{R}$, 在上例中就是屬性「年紀」。如此便可定義出一個對模糊變數 X 的可能性分配。這和由機率空間 (Ω, F, P) 經過特定限制 (如: 擲 n 次銅板中出現正面的次數) 而定義出隨機變數 (random variable)X, 再由此決定 X 是何種形態之機率分配的概念是一樣的, 其間的差別只在於可能性空間中的限制具模糊特性 (如: 年輕), 而機率空間則否。

那麼, 可能性分配函數 (possibility distribution function) 又是怎麼得到的呢? 可能性分配函數和可能性測度間有什麼關係? 回想一下模糊限制 F, 它是一個模糊集, 那麼必可定義出一個 F 的隸屬函數 $\mu_F(x)$, 用來描述 F 這個模糊概念。我們將可能性分配函數 r_X 定義爲 μ_F, 也就是 $r_X = \mu_F$。爲什麼這麼定呢? 我們在上一段曾說明, 可能性分配是由模糊限制 F 透過模糊變數 X 而定出來的, 因此, 把 μ_F 定爲 X 的可能性分配函數, 似乎再自然不過了。而可能性分配函數和可能性測度間的關係, 就可寫爲 $\pi(\{x\}) = r_X(x)$。因此, 我們可以得到 $\pi(\{x\}) = r_X(x) = \mu_F(x)$ 這樣的關係。經過這樣一個過程, 對應於「小明是年輕的」這個敍述的可能性分配便確定了。

假設「年輕」這個模糊集的隸屬函數, 定義如下:(其中 x 表年齡)

$$\mu_{\text{young}}(x) = \begin{cases} 1, & 0 \le x \le 25 \\ \left[(1 + (\frac{x-25}{5})^2\right]^{-1} & 25 < x \le 100 \end{cases}$$

那麼, 在指定「小明是年輕的」的情況下, 關於它的可能性分配是如何呢? 由上一段可知, 可能性分配函數 $r_X = \mu_{\text{young}}$。經由計算可得: 若 $x = 20$, 則 $\mu_{\text{young}}(x) = 1$; 若 $x = 30$, 則 $\mu_{\text{young}}(x) = 0.5$; 若 $x = 40$, 則 $\mu_{\text{young}}(x) = 0.1$。這樣的結果有什麼意義呢? 可以這麼說: 如果我們已經知道小明是年輕的, 那麼小明是20歲的「可能性」是1, 也就是說「20歲」和「年輕」之間的一致程度 (compatibility) 是1; 30歲的可能性降到0.5,40歲的可能性就更低了, 只剩下0.1。年齡愈大, 求算出的可能性値就愈小, 表示愈大的年紀和年輕這個模糊概念的一致程度愈低。這樣的結果十分直觀, 似乎也很合理, 不是嗎? 在這個例子中, 如果將「可能性」三個字換成「機率」, 似乎就十分怪異, 這一點我們可以從兩方面來看。就觀念上而言, 這個例子並不是發生與否的問題, 而是一種因語意上的不明確所產生的問題; 若就定義而言, 可能性和機率不同, 並不具可加性。所以顯然此時「機率」二字並不恰當,「可能性」概念的引入就變得十分自然了。

下面我們再使用扎德所舉的一個十分生動且有趣的例子來說明, 也許更能幫助我們區分機率和可能性之間的差異。

(表一)

u	1	2	3	4	5	6	7	8
$r_X(u)$	1	1	1	1	0.8	0.6	0.4	0.2
$P_X(u)$	0.1	0.8	0.1	0	0	0	0	0

小明以 u 個雞蛋作爲早餐, 表一中的 u 值從1到8, $P_X(u)$ 表機率, $r_X(u)$ 表可能性, 例如, 小明早餐吃兩個蛋的機率是0.8, 吃6個的可能性是0.6。有時候, 我們可以將可能性視爲一種「執行上的難易程度」, 値愈大表愈容易。那麼, 從表中可知對小明而言, 吃1個蛋到4個蛋都很容易 (此時可能性皆是最大值1), 然後可能性値隨著蛋個數的增加而遞減, 表示蛋愈多、吃起來愈不容易, 十分合乎常理! 到9個時可能性爲0, 表示 — 小明根本不可能吃得下9個蛋! 而機率這列顯示出的訊息就不同了。我們可以發現小明早餐吃2個蛋的機率是0.8, 明顯大於1個和3個的0.1, 表示小明「通常」以2個蛋爲早餐, 這也許是基於健康上的考量或其他因素, 但至少和他的能力無關。値得注意的是, 機率仍謹守可加性, 也就是三個機率値的和爲1。可能性則無此限制。

以上的例子, 是否有助於你更進一步了解可能性呢? 不過, 有一點要說明的是, 上列可能性所指的「難易程度」不一定是指生理上執行的難易; 有時候, 所謂的難易程度只是一種象徵性的意義。從上面的例子, 我們很容易了解「對小明而言, 吃9個蛋根本是力不能及的事」; 但是, 在「若小明年輕, 則他是30歲的可能性爲0.5」這個例子中,「30歲」和「年輕」間的容易程度, 就得看爲是「一致程度」了。

電腦科技日新月異, 使各種應用科學得以憑藉它建立更複雜的模型, 處理更龐大繁複的資料。這些資料常囿於儀器設備或人爲因素而顯得不夠完整, 無法以精確的形態表之, 另方面也因爲系統日趨龐大, 分析日漸複雜, 可能性理論由是應運而生, 且恰如其份地在這股潮流中扮演舉足輕重的脚色。本文從自然語言的模糊性這個觀點切入, 說明可能性的重要; 再將其和大家所熟知的機率相比較, 使讀者能體會兩者間的差異, 並舉例說明。的確, 從圖形分類、控制過程到決策分析、專家系統中不確定性的管理, 無論在理念或實務上, 可能性理論在在顯示出其必要性。進一步關於可能性理論的了解, 可參考[3]、[4]、[5]、[6]。

参考文獻

1. 楊敏生, 模糊理論簡介, 數學傳播,Vol.18(1), 頁7-11,83年3月。
2. L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1, 3-28(1978).
3. M. L. Puri and D. Ralescu, Fuzzy random variable, J. of Math. Analysis and Applications 114, 409-422(1986).
4. D. Dubois and H. Prade, Possibility Theory, Plenum Press, New York (1988).
5 G. J. Klir and B. Parviz, Probability-possibility transformations: a comparison, Int. J. of General Systems 21(3), 291-310(1992).
5. R. R. Yager, On the completion of qualitative possibility measures, IEEE Trans. on Fuzzy Systems 3,184-194(1993).

—本文作者分別任教和畢業於中原大學數學研究所—