機率論為何要建立在機率空間上面?

蔡聰明

一. 一些問題的提出

機率論是研究隨機現象的學問。它要對隨機現象建立機率模型, 並且求算有趣事件的機率及推演出各種極限定理、機率法則, 研究隨機過程等等。

機率模型長得什麼模樣? 它的理型 (ideas and forms) 是什麼?

今日大家都會說, 機率論的出發點是機率空間 $(\Omega, \mathcal{F}, P)$, 其中 Ω 爲樣本空間 (sample space), 代表隨機實驗的所有可能發生結果; $\mathcal{F}$ 爲事件 (events) 所構成的 σ 代數 (σ-algebra); P 爲一個機率測度 (probability measure), 用來度量事件發生的「確定程度」(the degree of certainty)。「probability」一詞日文翻譯成「確率」, 我國早先的譯名是「或然率」、「概率」, 但目前通行的譯名是更佳的「機率」。

機率空間的理型是怎麼來的呢? 它當然是長久演化與試誤 (trial and error) 的產物, 經過人爲的選擇與創造才得到的。從長期累積的直觀經驗知識 (intuitions) 出發, 提煉出概念 (concepts) 與公理, 最後結晶爲理論 (theory)。事實上, 任何概念或理論都是這樣產生的, 用來承載或組織豐富的、既有的經驗知識, 進一步又可當作開拓新知的根據地。

機率論的公理化醞釀了約 30 年, 從 1900年到1993年, 正值現代數學公理化思潮的高峰期。

機率的概念起源於賭局 (games of chance), 從盤古開天地以來就出現, 長久不斷地累積經驗知識, 一直到 1713 年 J.Bernoulli 提出弱大數法則, 標誌著數學機率論的誕生。接著又有 De Moivre(1718年) 與 Laplace(1801年) 的中央極限定理, Gauss的誤差之正規分佈律 (1809年) 以及 Poisson 極限定理 (1832年) 等各種機率法則 (the laws of chance) 之獲得, 其中也出現了種種詭論 (paradoxes)。到了 1900 年, 公理化的時機成熟, Hilbert提出著名的23個問題, 其中第6個就是關於物理學與機率論的公理化問題。於是許多數學家開始嘗試解決機率論的公理化問題, 一方面繼續開拓機率論的領域, 一方面作各種試驗。結果在 1933

年才由俄國偉大數學家 Kolmogorov(1903-1987年) 完成機率論的公理化 (幾乎跟 Von Neumann 的量子力學公理化同時完成), 提出機率空間的理型作爲機率論的基礎。從此機率論作爲一個數學理論完全確立。Kolmogorov的機率空間模型, 普遍被數學家所接受。在這麼單薄的假設下, 居然開展出那麼豐富的結果, 這是機率論的榮耀。

機率論爲何要採用「機率空間」的理型? 這樣做有什麼優點? 不這樣做又有什麼缺點? 還有沒有其它選擇? 取捨的標準何在?

這些問題都是初學公理化機率論的人最易產生的困惑。如果回答說, 這樣取的「機率空間」是全測度爲1的特殊測度空間, 使得測度與積分論的工具都可以搬過來用, 這只是一個平凡的 (ordinary) 答案, 還是不能解惑。

Lebesgue在 1902-1903 年就創立測度與積分論, 爲什麼要等到 1933 年 (相差 30 年) 機率論才完成公理化呢? 因此,「機率空間」的出現必有更深刻的道理存在, 不論是內在的或外在的理由。1900年到1933年之間正是數學家找尋機率模型的試誤與演化階段。

機率空間的出現, 我們分成四個階段來討論: (i) 隨機實驗與命題演算算, (ii) 翻譯成集合論的語言得到初等機率空間之理型, (iii) 用各式各樣的例子來試驗初等機率空間的理型是否合用, (iv) 最後抽取出 Kolmogorov 的機率論公理系統。

本文我們要用一些具體的例子來闡明: 機率空間是公理化機率論的最佳選擇。

二. 初等機率空間的出現

甲. 隨機實驗與命題演算

所謂 「隨機實驗」(random experiment) 是指其出現結果有種種可能之實驗, 事先完全無法預料。例如丟一個銅板 100 次就是一個隨機實驗。

對於一個隨機實驗, 我們想知道各種有趣的事象發生的機會大小。這些事象通常以「命題」(propositions) 或 「敍述」 (statements) 的方式來出現, 例如令 p 表示「正面的次數介於 45 次到 58 次之間」, 這就是關於丟 100 次銅板的隨機實驗之一個事象。我們稱 p 爲一個隨機命題。

古典邏輯所研究的「敍述」都具有明確的眞假值, 不能時眞時假或沒有眞假值, 但是在機率論中的敍述則不然。作隨機試驗下來, 如果丟出 53 次正面, 則 p 成立; 如果丟出 40 次正面, 則 p 不成立。因此, 每一個事象都「說不準」(uncertainty)。衡量一個事象眞確的程度就叫做該事象的機率。

假設 $p, q, r, \cdots$ 爲一族命題, 那麼我們就可以談論命題的三個運算:

名稱	記號	意義
否定	$\sim$	$\sim p$ 表示「p 的否定」之命題
或	$\vee$	$p \vee q$ 表示「p 或 q」, 即 p、q 至少有一個成立之命題
且	$\wedge$	$p \wedge q$ 表示「p 且 q」, 即 p、q同時成立之命題

這三個運算顯然滿足下列運算律:

1. 交換律:

$$p \vee q = q \vee p,\ p \wedge q = q \wedge p$$

2. 結合律:

$$(p \vee q) \vee r \;=\; p \vee (q \vee r)$$
$$(p \wedge q) \wedge r \;=\; p \wedge (q \wedge r)$$

3. 分配律:

$$p \vee (q \wedge r) \;=\; (p \vee q) \wedge (p \vee r)$$
$$p \wedge (q \vee r) \;=\; (p \wedge q) \vee (p \wedge r)$$

4. 冪等律:

$$p \vee p = p, \quad p \wedge p = p$$

5. De Morgan 定律:

$$\sim (p \vee q) \;=\; (\sim p) \wedge (\sim q)$$
$$\sim (p \wedge q) \;=\; (\sim p) \vee (\sim q)$$

6.雙否定律:

$$\sim (\sim p) = p$$

7. 兩值律:

$$p \vee (\sim p) = 1, \quad p \wedge (\sim p) = 0$$

其中1表恆眞命題,0表恆假命題。

8. 0與1的特性:

$$p \vee 0 \;=\; p, \quad p \wedge 0 = 0$$
$$p \vee 1 \;=\; 1, \quad p \wedge 1 = p$$

定義:一族命題 $\mathcal{A}$ 滿足上述八條運算律就叫做一個 Boole 代數 (Boolean algebra), 記為 $\langle \mathcal{A}, \sim, \vee, \wedge, 0, 1 \rangle$。

註: Boole(1815-1864) 是自學成功的典範, 他在1854年出版 “An investigation of the laws of thought”一書, 引入命題演算, 嘗試將邏輯代數化。B. Russell 稱譽說:「純數學是 Boole 發明的。」

乙. 翻譯成集合論的語言: 初等機率空間

G. Cantor(1845-1918) 創立集合論, 堪稱為 「數學的法國大革命」。這是他追究「無窮」的產物。現代數學的特色之一, 就是一切都要化成集合論的語言來表達。

對於任意非空集 Ω, 令 2^{Ω} 表示 Ω 的所有子集全體, 將 $\sim$、$\vee$、$\wedge$ 分別解釋爲取補集 c、取聯集 $\cup$、取交集 $\cap$ 之操作, 再取 $0 = \emptyset$, $1 = \Omega$, 則 $\langle 2^{\Omega}, c, \cup, \cap, \emptyset, \Omega \rangle$ 形成一個 Boole 代數。反過來, M.H. Stone (1963年) 證明: 對於任何 Boole 代數 $\mathcal{A}$, 必可找到一個集合 Ω, 使得 $\mathcal{A}$ 中的命題表現爲 Ω 的子集, 並且 $\sim$、$\vee$、$\wedge$ 表現爲集合運算 c、$\cup$、$\cap$, 而成爲 2^{Ω} 的一個 Boole 子代數。

對於一個隨機實驗的一族隨機命題所成的 Boole 代數 $\mathcal{A}$, 令 Ω 表示實驗的所有可能結果之集合, 將每個命題 p 表現成 Ω 的子集 A , 亦即使 p 成立的元素所成之集, 叫做 p 的眞值集 (truth set), 如下列的對照表:

命題	眞値集
1	Ω
0	$\emptyset$
p	A
q	B
$\sim p$	A^c
$p \vee q$	$A \cup B$
$p \wedge q$	$A \cap B$
$\mathcal{A}$	$\mathcal{A}$

顯然, $\mathcal{A} \subset 2^{\Omega}$ 並且滿足下列性質:

(i) $\Omega \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(iii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

亦即 $\mathcal{A}$ 包含有宇集 Ω, 並且在取補集與聯集下具有封閉性。我們稱這樣的 $\mathcal{A}$ 爲一代數 (algebra), 其元素叫做事件 (events)。如果 $A \cap B = \emptyset$, 則稱 A、B 爲互斥 (disjoint) 事件。

接著是對每一個事件 $A \in \mathcal{A}$ 引入機率 P(A), 我們自然要求它滿足:

(i) 正性: $0 \le P(A) \le 1, \forall A \in \mathcal{A}$

(ii) 歸一性: $P(\Omega) = 1$

(iii) 加性: 若 A,$B \in \mathcal{A}$ 爲互斥事件, 則

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

我們稱 P 爲一個初等機率測度 (an elementary probability measure)。

定義: 一個初等機率空間 (an elementary probability space) 是指三合一空間 $(\Omega, \mathcal{A}, P)$, 其中

(i) Ω 爲一個非空集, 叫做樣本空間;

(ii) $\mathcal{A} \subset 2^{\Omega}$ 爲一個代數, 表事件全體;

(iii) P 爲一個初等機率測度。

註: 此地的初等機率測度與初等機率空間之用語, 我們根據的是 Itô [13]的書。改採「擬機率測度」與「擬機率空間」, 也是很好的用語。

初等機率空間是對一個隨機實驗去蕪存菁後的初步總結, 是一個很自然的機率模型。這個理型是否足以承載豐富的機率結果呢? 在既知的結果中, 最重要的是 Bernoulli 的弱大數法則、De Moivre-Laplace 的中央極限定理以及 Borel 的強大數法則。這些都可以生存在初等機率空間上面嗎?

三. 初等機率空間的試驗

Bernoulli的弱大數法則起源於要澄清:「丟一個公正銅板, 出現正面的機率是1/2」這句話的意思。

爲此, 考慮在相同狀況下, 獨立地丟一個公正銅板 n 次之隨機實驗。我們爲其建立初等機率空間 $(\Omega_n, \mathcal{A}_n, P_n)$ 如下:

$$\Omega_n = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) : \omega_k = H \text{ 或 } T\}$$

其中 H 表正面 (head),T 表反面 (tail);

$$\mathcal{A}_n = 2^{\Omega_n}$$

由於每一樣本點的機會均等, 故對每一個 $\omega \in \Omega_n$, 指定

$$P_n(\omega) = \frac{1}{2^n}$$

從而, 對 $A \in \mathcal{A}_n$ 指定

$$P_n(A) = \frac{\#A}{2^n}$$

其中 $\#A$ 表示事件 A 所含的樣本點個數。

定義「隨機變數」(random variables)

$$\xi_k : \Omega_n \to \{0,1\}$$

如下: 對於 $k = 1, 2, \ldots, n$, 令

$$\xi_k(\omega) = \begin{cases} 1, & \text{當 } \omega_k = H \text{ 時} \\ 0, & \text{當 } \omega_k = T \text{ 時} \end{cases}$$

我們稱 $\xi_1, \xi_2, \ldots, \xi_n$ 爲一個 (有限的) 銅板序列。再令

$$S_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

則容易證得下面兩個定理:(參見 [2])

定理1: (Bernoulli弱大數法則) 對於任意 $\varepsilon > 0$, 恆有

$$\lim_{n\to\infty} P_n(|\frac{S_n}{n} - \frac{1}{2}| \leq \varepsilon) = 1 \qquad (1)$$

定理2: (De Moivre-Laplace 中央極限定理) 對任意 $x \in R$, 恆有

$$\begin{aligned} &\lim_{n\to\infty} P_n(\frac{S_n - n/2}{\sqrt{n}/2} \leq x) \\ =\ & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \qquad (2) \end{aligned}$$

因此, 初等機率空間的理型可以承載 Bernoulli 弱大數法則與 De Moivre-Laplace 中央極限定理。初步看起來, 似乎很不錯。這會是我們所要追尋的理型嗎?

再試驗 Borel 的強大數法則, 這必須涉及無窮多次丟銅板, 即銅板序列 (coin-tossing sequence)。

定義: 設 $(\Omega, \mathcal{A}, P)$ 爲一個初等機率空間。如果 $X_1, X_2, X_3 \ldots$ 爲定義在 $(\Omega, \mathcal{A}, P)$ 上, 取值 $\{0,1\}$ 的一列隨機變數, 滿足下列兩個性質:

(i) 同佈性: 對任意自然數n,

$$P(X_n = 0) = \frac{1}{2} = P(X_n = 1)$$

(ii) 獨立性: 對任意有限多個自然數 $i_1 < i_2 < \cdots < i_m$ 與由 $0, 1$ 所成的有限數列 $j_1, j_2, \ldots, j_m$

$$\begin{aligned} &P(X_{i1} = j_1, \ldots, X_{im} = j_m) \\ =\ & P(X_{i1} = j_1) \ldots P(X_{im} = j_m) \end{aligned}$$

則稱 (X_n) 爲一個銅板序列。

我們建造「丟無窮多次銅板」的初等機率空間及銅板序列如下:

取樣本空間

$$\Omega_\infty = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_k = H \text{ 或 } T\}$$

對於 $A \subset \Omega_n$, 令

$$C_A = \{\omega \in \Omega_\infty : (\omega_1, \omega_2, \ldots, \omega_n) \in A\}$$

表示以 A 爲底在 Ω_∞ 中的「柱集」(cylinder set)。令

$$\mathcal{A}_\infty = \{C_A : A \in \mathcal{A}, n \in \mathbf{N}\}$$

表示柱集的全體, 這是一個代數。對 $C_A \in \mathcal{A}_\infty$, 定義

$$P_\infty(C_A) = P_n(A) = \frac{\#A}{2^n}\text{。}$$

初等機率空間 $(\Omega_\infty, \mathcal{A}_\infty, P_\infty)$ 之建構只不過是將有窮多次丟銅板的一連串初等機率空間 $(\Omega_n, \mathcal{A}_n, P_n)$, $n \in \mathbf{N}$ 嵌入而已。

接著, 在 Ω_∞ 上定義無窮多個隨機變數:

$$\xi_k(\omega) = \begin{cases} 1, & \text{若 } \omega_k = H \ , \\ 0, & \text{若 } \omega_k = T \ , \end{cases}$$

$k = 1, 2, \ldots$。容易驗知 (ξ_n) 爲一個銅板序列。令

$$S_n = \xi_1 + \xi_2 + \cdots + \xi_n$$

所謂 Borel 的強大數法則是指

$$P_\infty(\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}) = 1 \tag{3}$$

這成立嗎? 讓我們作分析。令事件

$$B = \{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\}$$

則

$$\begin{aligned} B &= \bigcap_{\varepsilon>0} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \Big\{\Big|\frac{S_n}{n} - \frac{1}{2}\Big| < \varepsilon\Big\} \\ &= \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{|\frac{S_n}{n} - \frac{1}{2}| < \frac{1}{m}\} \end{aligned}$$

再令

$$\begin{aligned} B_{lk}^m &= \bigcup_{n=k}^{l} \Big\{\Big|\frac{S_n}{n} - \frac{1}{2}\Big| < \frac{1}{m}\Big\} \\ B_n^m &= \Big\{\Big|\frac{S_n}{n} - \frac{1}{2}\Big| < \frac{1}{m}\Big\} \end{aligned}$$

則顯然有

$$\begin{aligned} B_{lk}^m \overset{l\to\infty}{\downarrow} \bigcap_{n=k}^{\infty} B_n^m \overset{k\to\infty}{\uparrow} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} B_n^m \\ \overset{m\to\infty}{\downarrow} B \end{aligned} \tag{4}$$

其中只有 $B_{lk}^m \in \mathcal{A}_\infty$, 其餘的都經過了取單調極限, 故都不屬於 $\mathcal{A}_\infty$。因此 (3) 式的左項沒有意義, 亦即在 $(\Omega_\infty, \mathcal{A}_\infty, P_\infty)$ 上無法承載 Borel 強大數法則。這顯示初等機率空間的理型似乎不夠用, 尤其是涉及樣本空間爲無窮集的情形。

這並沒有十足的說服力, 因爲對於一個銅板序列, 初等機率空間的造法可以有各種方式, 也許換個造法就可使 Borel 的強大數法則成立。

下面我們舉一個奇妙的例子, 這是從 [3]一文中取出來的。考慮從自然數集 $\mathbf{N}$ 中任取出一數的隨機實驗。我們可以造一個初等機率空間, 並且在其上定義出各種銅板序列, 其中有的使 Borel 強大數法則不成立, 也有的使 Borel 強大數法則成立。

從 $\mathbf{N}$ 中任取一數的隨機實驗, 原則上, 每一數被取到的機會均等, 但是由於 $\#\mathbf{N} = \infty$, 故無法對每一點指定相等的機率。我們退而求其次, 造一個機率模型保住大家都接受的常識就好了:

從 $\mathbf{N}$ 中任取一數, 那麼

(i) 取到偶數的機率是 $\frac{1}{2}$, 取到奇數的機率也是 $\frac{1}{2}$;
(ii) 取到 5 的倍數之機率是 $1/5$;
(iii) 取到 99 的機率是 0;
(iv) 取到的數小於 1000 的機率是 0;
(v) 取到的數落在等餘類 (congruence class)

$$A(n, r) = \{\omega \in \mathbf{N} : \omega \text{ 除以 } n \text{ 餘 } r\}$$

之中的機率是 $\frac{1}{n}$, 其中 $n \in \mathbf{N}$, $r = 0, 1, \ldots, n-1$。

我們先造一個保住常識 (v) 的初等機率空間:

取樣本空間 $\Omega = \mathbf{N}$, , 事件全體 $\mathcal{R}$ 爲有窮多個等餘類之聯集全體。容易驗知 $\mathcal{R}$ 爲一個代數, 並且存在唯一的初等機率測度 P, 定義在 $\mathcal{R}$ 上, 使得

$$P(A(n,r)) = \frac{1}{n}$$

因此, $(\Omega, \mathcal{R}, P)$ 爲一個初等機率空間。

其次, 定義銅板序列: 對於 $\omega \in \mathbf{N}$, 用二進位法唯一展開成

$$\omega = \alpha_1 \cdot 2^0 + \alpha_2 \cdot 2^1 + \alpha_3 \cdot 2^2 + \cdots$$

上式右邊只含有限項, 其中 $\alpha_1, \alpha_2, \ldots$ 皆爲 0或 1, 於是 ω 表成二進位法爲

$$\omega = \cdots \alpha_3\alpha_2\alpha_1$$

定義隨機變數

$$X_n(\omega) = \alpha_n, n \in \mathbf{N}$$

設 $i_1, i_2, \ldots, i_n$ 爲 0 或 1 所成的序列, 令

$$i_n i_{n-1} \cdots i_1 = i_1 \cdot 2^0 + i_2 \cdot 2^1 + \cdots + i_n \cdot 2^{n-1}$$

則易驗知

$$\begin{aligned} \{ &\ \omega \in \mathbf{N} : X_1(\omega) = i_1, \ldots, X_n(\omega) = i_n\} \\ = &\ A(2^n, i_n i_{n-1} \ldots i_1) \end{aligned}$$

從而 (X_n) 爲定義在 $(\Omega, \mathcal{R}, P)$ 上的一個公正的銅板序列。

對於任意 $\omega \in \mathbf{N}$, 存在自然數 k_0, 使得對所有 $k \geq k_0$。恒有

$$\alpha_k = X_k(\omega) = 0。$$

於是

$$P(\lim_{n\to\infty} \frac{S_n}{n} = 0) = P(\Omega) = 1$$

從而 Borel 強大數法則不成立, 事實上 $\{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\} = \emptyset$。

更有趣的是, 在同樣的初等機率空間 $(\Omega, \mathcal{R}, P)$ 上, 定義隨機變數

$$Y_n = \begin{cases} X_n, & \text{當 } n \text{ 爲偶數時}, \\ 1 - X_n, & \text{當 } n \text{ 爲奇數時}, \end{cases}$$

則易驗知 (Y_n) 仍然爲一個銅板序列, 並且 Borel 強大數法則成立:

$$P(\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}) = P(\Omega) = 1$$

進一步, 如果取 $i_1, i_2, \ldots$ 爲 0 與 1 所組成的數列, 使得極限 $\lim_{n\to\infty} \frac{i_1+i_2+\cdots+i_n}{n}$ 不存在, 定義

$$Z_n = \begin{cases} X_n, & \text{若 } i_n = 0, \\ 1 - X_n, & \text{若 } i_n = 1, \end{cases}$$

則 (Zn) 也是一個銅板序列, 並且

$$P(\lim_{n\to\infty} \frac{S_n}{n} \text{ 不存在}) = P(\Omega) = 1$$

對於任何有理數 α, $0 \leq \alpha \leq 1$ 都可以在 $(\Omega, \mathcal{R}, P)$ 上造一個銅板序列, 使得

$$P(\lim_{n\to\infty} \frac{S_n}{n} = \alpha) = 1$$

初等機率空間 $(\mathbf{N}, \mathcal{R}, P)$ 雖然奇妙, 但還是有所不足。例如單身集等, $\{2\}$、$\{7\}$ 等, 以及 $\mathbf{N}$ 的任何有限子集都不屬於 $\mathcal{R}$ 。因此, 它並沒有保住所有的常識。

一個改進之道是擴大 $\mathcal{R}$, 並且延拓 P, 這就是下面要介紹的更重要的一個例子。

對於任意子集 $A \subset \mathbf{N}$ 及 $m \in \mathbf{N}$, 令

$$A[m] = \{k : 1 \leq k \leq m, k \in A\}$$

如果極限 $\lim_{m\to\infty}\frac{\#A[m]}{m}$ 存在, 則記此極限值爲 $P^*(A)$, 叫做 A 的算術密度 (arithmetical density)。令

$$\mathcal{D}=\{A\subset\mathbf{N}:\lim_{n\to\infty}\frac{\#A[m]}{m}\text{ 存在}\}$$

表示所有具有算術密度的子集全體。於是得到一個三合一空間 $(\mathbf{N},\mathcal{D},P^*)$, 這是機率式數論 (probabilistic number theory) 的出發點, 結果豐碩。

命題 1: (i) $\mathbf{N}$ 中的單身集與有窮子集皆屬於 $\mathcal{D}$ 且其機率爲 0。

(ii) 等餘類$A(n,r)\in\mathcal{D}$ 且

$$P^*(A(n,r))=\frac{1}{n}$$

因此,$(\mathbf{N},\mathcal{D},P^*)$ 是 $(\mathbf{N},\mathcal{R},P)$ 的延拓。

證明: (i) 是顯然的。

(ii) 令 $A=A(n,r)$, 則

$$\frac{m}{n}-1<\#A[m]=[\frac{m}{n}]\le\frac{m}{n}$$

$$\text{所以}\frac{1}{n}-\frac{1}{m}<\frac{\#A[m]}{m}\le\frac{1}{n}$$

令 $m\to\infty$, 得到 $P^*(A(n,r))=\frac{1}{n}$。因此,$A(n,r)\in\mathcal{D}$。

因此, 上述在 $(\mathbf{N},\mathcal{R},P)$ 上所定義的各式各樣的銅板序列都可搬到 $(\mathbf{N},\mathcal{D},P)$ 上來, 仍然具有相同的奇異性質。

命題 2:

(i) $\emptyset,N\in\mathcal{D}$ 且 $P^*(\emptyset)=0$,$p^*(N)=1$,

(ii) $A\in\mathcal{D}\Rightarrow A^c\in\mathcal{D}$, 且 $P^*(A^C)=1-P^*(A)$,

(iii) $A,B\in\mathcal{D}$ 且 $A\subset B\Rightarrow B\backslash A\in\mathcal{D}$,

(iv) $A_1,A_2,\ldots,A_n\in\mathcal{D}$ 且兩兩互斥 $\Rightarrow\bigcup_{k=1}^n A_k\in\mathcal{D}$。

證明:(i) 與 (iv) 是顯然的。

(ii)$A\in\mathcal{D}$, 由於 $\#A^c[m]=m-\#A[m]$, 故

$$\begin{aligned}\lim_{m\to\infty}\frac{\#A^c[m]}{m}&=\lim_{m\to\infty}\frac{m-\#A[m]}{m}\\&=1-\lim_{m\to\infty}\frac{\#A[m]}{m}\end{aligned}$$

今因 $\lim_{m\to\infty}\frac{\#A[m]}{m}$ 存在, 所以 $\lim_{m\to\infty}\frac{\#A^c[m]}{m}$ 也存在, 並且

$$P^*(A^c)=1-P(A)\text{。}$$

(iii) 設 $A,B\in\mathcal{D}$ 且 $A\subset B$ 則

$$\#(B\backslash A)[m]=\#B[m]-\#A[m]$$

$$\text{所以}\lim_{m\to\infty}\frac{\#(B\backslash A)[m]}{m}$$

$$=\lim_{m\to\infty}\frac{\#B[m]}{m}-\lim_{n\to\infty}\frac{\#A[m]}{m}\text{ 存在,}$$

於是 $B\backslash A\in\mathcal{D}$。 #

註: 上述命題中的 (iv),「兩兩互斥」不能放棄, 有限聯集也不能改爲可列聯集。

命題 3: $\mathcal{D}$ 不是一個代數, 更不是一個 σ 代數。

我們只需舉出 $E,F\in\mathcal{D}$, 但是 $E\cap F\notin\mathcal{D}$ 就好了。

取 E 爲奇數集, 再取

$$\begin{aligned}F=\{&[2^{2n},2^{2n+1}]\text{ 中的奇數與}\\&[2^{2n+1},2^{2n+2}]\text{ 中的偶數},n\ge0\}\end{aligned}$$

則 $E = A(2,1) \in \mathcal{R} \subset \mathcal{D}$ 且
$P^*(E) = P(E) = \frac{1}{2}$。

其次證 $F \in \mathcal{D}$。首先觀察到, 對任意 $n \in \mathbf{N}$, 存在 $n \in \mathbf{N}$ 使得 $m \in [2^{2n}, 2^{2n+1}]$, 並且當 $m \to \infty$ 時,$n \to \infty$。

在 $[1,2]$ 中有一個奇數, 在 $[2^{2n}, 2^{2n+1}]$ 中有 2^{2n-1} 個奇數, $n = 1, 2, 3, \ldots$; 在 $[2^{2n}, 2^{2n+1}]$ 中有 $2^{2n}+1$ 個偶數, $n = 0, 1, 2, \ldots$。於是

$$
\begin{aligned}
&[1 + (2^0+1) + 2^1 + (2^2+1) + \cdots \\
&\quad +2^{2n-3} + (2^{2n-2}+1)]/(2^{2n}+1) \\
\geq\ & \frac{\#F[m]}{m} \\
\geq\ & [1 + (2^0+1) + 2^1 + (2^2+1) + \cdots \\
&\quad +2^{2n-1} + (2^{2n}+1)]/(2^{2n+2})
\end{aligned}
$$

亦即

$$
\begin{aligned}
&\frac{2^{2n-1}+n+2}{2^{2n}} \geq \frac{2^{2n-1}+n+2}{2^{2n}+1} \\
\geq\ & \frac{\#F[m]}{m} \geq \frac{2^{2n-1}+n+2}{2^{2n+2}} \\
&\frac{1}{2} + \frac{n+2}{2^{2n}} \geq \frac{\#F[m]}{m} \geq \frac{1}{2} + \frac{n+2}{2^{2n+2}}
\end{aligned}
$$

由夾擠原理知

$$\lim_{m\to\infty} \frac{\#F[m]}{m} = \frac{1}{2}$$

因此 $F \in \mathcal{D}$。

考慮 $E \cap F = \{[2^{2n}, 2^{2n+1}]\}$ 中的奇數, $n = 0, 1, \ldots\}$。我們要證明 $\lim\limits_{m\to\infty} \frac{\#(E\cap F)[m]}{m}$ 不存在。利用歸謬法, 如果該極限存在, 則取子序列也收斂到同一值。今取 $a_n = 2^{2n+1}$, $b_n = 2^{2n+2}$,

則

$$
\begin{aligned}
\#(E\cap F)[a_n] &= \#(E\cap F)[b_n] \\
&= 2^1 + 2^3 + \cdots + 2^{2n-1},
\end{aligned}
$$

所以

$$
\begin{aligned}
\frac{\#(E\cap F)[a_n]}{a_n} &= \frac{1}{3} - \frac{1}{3\cdot 2^{2n}} \\
\frac{\#(E\cap F)[b_n]}{b_n} &= \frac{1}{6} - \frac{1}{3\cdot 2^{2n+1}}
\end{aligned}
$$

從而

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{\#(E\cap F)[a_n]}{a_n} = \frac{1}{3} \neq \frac{1}{6} \\
=\ & \frac{\#(E\cap F)[b_n]}{b_n},\ \text{矛盾。}
\end{aligned}
$$

於是 $E \cap F \notin \mathcal{D}$。 #

因此, $(\mathbf{N}, \mathcal{D}, P^*)$ 既不是初等機率空間, 更不是機率空間。但是, 它依然可以承載許多數論上的結果。參見 Kac 的小書 [7]。

總結上述的例子:

甲、在初等機率空間 $(\Omega_n, \mathcal{A}_n, P_n), n = 1, 2, \ldots$, 上可承載 Bernoulli弱大數法則與 De Moivre-Laplace 中央極限定理 。

乙、在初等機率空間 $(\Omega_\infty, \mathcal{A}_\infty, P_\infty)$ 上

$$\{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\} \notin \mathcal{A}_\infty$$

丙、在初等機率空間 $(\mathbf{N}, \mathcal{R}, P)$ 或非初等機率空間 $(\mathbf{N}, \mathcal{D}, P^*)$ 上可以定義各式各樣的銅板序列, 使得

$$\{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\} = \emptyset; \{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\} = \Omega;$$

$$\{\lim_{n\to\infty}\frac{S_n}{n}=\alpha\}=\Omega, \forall\alpha \text{ 爲有理數},$$
$$0\le\alpha\le 1;$$
$$\{\lim_{n\to\infty}\frac{S_n}{n} \text{ 不存在}\}=\Omega.$$

Borel的強大數法則可以成立, 也可以不成立, 堪稱「百花齊放」, 結果繁多。

四. Lebesgue 空間上的機率論之試驗

自從 Lebesgue 創立測度與積分論後,Lebesgue 空間 $((0,1],\mathcal{L},\mu)$ 就一直被當作是開發機率論 (乃至分析學) 的絕佳空間, 其中 $\mathcal{L}$ 爲 $(0,1]$ 中 Lebesgue 可測集全體,μ 爲 Lebesgue 測度。

首先是定義出銅板序列。對於任意 $\omega\in(0,1]$, 作二進位法展開。

$$\omega=0.\xi_1(\omega)\xi_2(\omega)\xi_3(\omega)\cdots$$

其中每個 $\xi_n(\omega)$ 都是 0 或 1。有的數有兩種展開法, 例如

$$\frac{1}{2}=0.1000\cdots=0.0111\cdots$$

我們就規定取後者。於是 (ξ_n) 就是定義在 Lebesgue 空間 $((0,1],\mathcal{L},\mu)$ 上的一個銅板序列。令 $S_n=\xi_1+\cdots+\xi_n$。

再令 $r_n(\omega)=2\xi_n(\omega)-1, n=1,2,\ldots$, 則 r_n 取值 -1 與 $+1$, 叫做 Rademacher 函數。

利用分析學工具, 我們就可以證得下列諸定理:

定理 B: (Borel強大數法則, 1909年)

$$\mu(\lim_{n\to\infty}\frac{S_n}{n}=\frac{1}{2})=1$$

對於 S_n 的漲落 (fluctuation) 之研究, 有一系列的結果, 最後導致重覆對數法則 (Law of the iterated logarithm) 的產生。爲方便起見, 下面令 $S_n=r_1+r_2+\cdots+r_n$, 則 Borel 的強大數法則可以改述成:

$$S_n=o(n)\quad \text{a.s.(殆遍)}$$

由此當然可得

$$S_n=O(n)\quad \text{a.s.}$$

進一步, 我們有:

定理 4: (Khinchine重覆對數法則,1924年)

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\ln\ln n}}=1,\ \text{a.s.}$$

註: 在1929年 Kolmogorov 再推廣到一般獨立隨機變數列的情形。

Steinhaus與 Wiener 在1922年提出一個問題: 在調和級數 $1+\frac{1}{2}+\frac{1}{3}+\cdots$ 每一項之前以丟銅板來加正、負號, 討論其斂散行爲。這個問題是往後獨立隨機變數和 (sum of independent ramdom variables) 及隨機積分 (Wiener 積分與 Itô 積分) 理論的發源地。

定理5: (Rademacher, 1922; Kolmogorov 與 Khinchine, 1924)

設 $\sum_{n=1}^{\infty}C_n$ 爲一個複數項之級數, 則

$$\sum_{n=1}^{\infty}C_nr_n \text{ 殆遍收斂} \Leftrightarrow \sum_{n=1}^{\infty}|C_n|^2<\infty。$$

復次,Wiener 在1923年將 (S_n) 作連續化 (考慮銅板越丟越快, 步幅越來越小), 在 Lebesgue 空間上建構出 Wiener 過程 $(W(t))t \geq 0$ (或叫做 Brown 運動)。這是最重要的一個隨機過程, 是整個現代隨機過程論的出發點。

結論是: 在 Lebesgue 空間上開展的機率論相當成功。

五. 回顧與前瞻

欲承載 Borel 的強大數法則, 是機率論公理化的著眼關鍵。Borel 當初建立這個結果是在 Lebesgue空間上做出來的, 而 Lebesgue 空間是典型的機率空間。

如果改用正準建構 (canonical construction) 的初等機率空間 $(\Omega_\infty, \mathcal{A}_\infty, P_\infty)$ 來看, 我們發現

$$B \equiv \{\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\} \notin \mathcal{A}_\infty$$

$\mathcal{A}_\infty$ 太小了, 事件 B 無處容身。顯然 B 生存在由 $\mathcal{A}_\infty$ 所生成的單調類 (monotone class) $m(\mathcal{A}_\infty)$ 之中。由單調類定理 (monotone class thesrem) 知, 代數 $\mathcal{A}_\infty$ 所生成的單調類 $m(\mathcal{A}_\infty)$ 就是其所生成的 σ 代數 $\sigma(\mathcal{A}_\infty)$。換言之, $B \in \sigma(\mathcal{A}_\infty)$。另一方面, 由測度論知, P_∞ 可以唯一延拓到 $\sigma(\mathcal{A}_\infty)$ 上, 成爲一個機率測度, 仍記爲 P_∞, 使得 $(\Omega_\infty, \sigma(\mathcal{A}_\infty), P_\infty)$ 爲一個機率空間, 並且在其上 Borel 強大數法則成立。

進一步, 對於在任何機率空間 $(\Omega, \mathcal{F}, P)$ 上的銅板序列, 同樣都可以證明 Borel 強大數法則成立。

但是對於初等機率空間 $(\mathbf{N}, \mathcal{R}, P)$, 卻可在其上定義各種的銅板序列, 使得 Borel 強大數法則有時成立, 有時不成立。成立與否跟隨機變數的函數本身有關, 而不只是依賴於隨機變數的機率分佈。這是初等機率空間的致命傷。

理想的機率論應該只依賴於隨機變數的機率分佈, 而跟隨機變數作爲函數本身的對應取值無關 (不妨叫做機率論的不變性理論觀點), 這在機率空間上發展出的機率論確實如此。

六. Kolmogorov 的公理系統

總結上述,Kolmogorov 在1933年提出機率空間的理型作爲機率論的基礎, 這是順理成章的最佳選擇。

Kolmogorov按兩個階段來建立機率論的公理系統:

甲、樣本空間 Ω 是有窮集的情形

此時初等機率空間 $(\Omega, \mathcal{A}, P)$ 的理型就夠用, 通常取 $\mathcal{A} = 2^\Omega$。這含蓋了所有 Laplace 型的古典機率空間。

乙、樣本空間 Ω 是無窮集的情形

仍然取初等機率空間 $(\Omega, \mathcal{A}, P)$ 作爲模型, 但是進一步要求 P 具有連續性:

$$(A_n) \subset \mathcal{A} \text{ 且 } (A_n) \downarrow \emptyset \text{ (遞降至空集)}$$
$$\Rightarrow \lim_{n\to\infty} P(A_n) = 0 \tag{5}$$

根據測度論,(5) 式等價於 P 具有可列加性 (σ-additrvity):

$$(A_n) \in \mathcal{A} \text{ 兩兩互斥且 } \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$$

$$\Rightarrow P(\bigcup_{n=1}^{\infty}) = \sum_{n=1}^{\infty} P(A_n) \qquad (6)$$

從而 P 可以唯一延拓到由 $\mathcal{A}$ 所生成的 σ 代數 $\sigma(\mathcal{A})$ 上, 仍記爲 P, 而得到機率空間 $(\Omega, \sigma(\mathcal{A}), P)$ 。整個機率論就在這個延拓的機率空間上順利地開展。

因此, 機率論的基礎就是機率空間 $(\Omega, \mathcal{F}, P)$ 定義如下:

定義: 設 Ω 爲一個非空集。如果 $\mathcal{F} \subset 2^{\Omega}$ 滿足:

(i) $\Omega \in \mathcal{F}$,

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,

(iii) $(A_n) \subset \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, 則稱 $\mathcal{F}$ 爲 Ω 上的一個 σ 代數。如果 $P : \mathcal{F} \to R_+ = [0, \infty)$ 滿足:

(iv) $0 \leq P(A) \leq 1, \forall A \in \mathcal{F}$,

(v) $P(\Omega) = 1$,

(vi) $(A_n) \subset \mathcal{F}$, 兩兩互斥 $\Rightarrow P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$,

則稱 P 爲 $\mathcal{F}$ 上的一個機率測度。如果 $\mathcal{F}$ 爲 Ω 上的一個 σ 代數, P 爲 $\mathcal{F}$ 上的一個機率測度, 那麼三合一空間 $(\Omega, \mathcal{F}, P)$ 就叫做一個機率空間。

七. 結語

Kolmogorov說:「在描述任何可觀測的隨機過程時, 我們只得到初等機率空間的資料, 而延拓的機率空間只是實際隨機過程的一種理想化。我們“任意地”(arbitrarily) 劃定滿足公理 (5) 的系統作爲機率模型。這個限制, 在各種研究中, 我們發現是適當的且方便的。」

在 Kolmogorov 的公理系統之下, 三合一空間 $(\mathbf{N}, \mathcal{R}, P)$ 與 $(\mathbf{N}, \mathcal{D}, P^*)$ 都是「化外之民」, 因爲顯然它們都不能延拓成爲機率空間。但是它們並不失其重要性。Kolmogorov 說他的公理系統是「任意地」割捨, 這很高明地保留一個活結, 使得非機率空間上的機率論仍然有研究的餘地。畢竟機率論是一種數學理論, 但是它跟經驗世界又具有密切關係。有一個時期, 機率論被看作是介於物理學 (或哲學) 與數學之間的一門學問。

放棄較自然與直觀的初等機率空間, 改採較局限的機率空間, 倒不是因爲初等機率空間不足以承載機率論, 而是承載的東西花樣繁多, 有的甚至違背了常識。作局限使得事情單純化, 而不失其豐富性。由代數 $\mathcal{A}$ 延拓爲 σ 代數 $\sigma(\mathcal{A})$, 相當於從有理數系 Q 延拓爲實數系 R, 多出了許多經驗世界接觸不到的事件, 但是對於澄清經驗世界卻大有用處。某種理想化、連續化、無窮化是「了解」的必要手段, 這是機率論給我們的啓示。

参考資料

1. 楊維哲: 機率論, 台北, 正中書局, 1979年。
2. 蔡聰明: 什麼是機率與機率法則? 數學傳播, 十九卷一期, 1995。
3. W.D. Sudderth and S. Ramakrishnan: A sequence of coin toss variables for

which the strong law fails. Amer. Math. Monthly, 939-941, 1988。

4. A. N. Kolmogorov: Foundations of the theory of probability, Chelsea, New York, 1956。
5. L. Breimam: Probability. Addison-Wesley, 1968。
6. K. Itô :Introduction to probability theory, Cambridge Univ. Press, 1984。
7. M. Kac: Statistical independence in probability, Analysis and number theory, Carus Math. Monographs 12, 1959。
8. J. Barone and A. Novikoff: A History of the axiomatic formulation of probability from Borel to Kolmogorov, Part I. Arch. Hist. Exact Sci. 18, 123-190, 1978。
9. H. Royden: Real Analysis, third Edition, McMillan, 1988。
10. L.E. Maistrov: Probability theory, A Historcal Sketch, Academic press, 1974.

—本文作者任教於台灣大學數學系—