

# 隨機隨想

黃文璋教授

國立高雄大學應用數學系

## 1 隨機世界

隨機究竟是什麼呢?你可能聽過隨機應變。字典的解釋(見台灣中華書局1986年出版的辭海)為“隨事機之變化靈活應付”。這是語出三國演義第五十七回,孫權問龐統“公平生所學,以何為主?”龐統回答“不必拘執,隨機應變。”翻閱字典,還有“隨機教學”,其內容(見文化圖書公司1980年出版,陸師成所主編的辭彙)為“不用固定的課本,不規定教學的時間,而由教師隨時利用機會施行教學。民國三十七年教育部所公佈的小學算術課程標準規定:一、二學年不特設教學時間,不用課本,而在日常生活、常識、工作、唱遊等科,以及自由遊戲中,隨機教學算術”。佛教聖地山西五臺山,有個著名的顯通寺。其中的大文殊殿有幅對聯,上聯為“法身無去無來住寂光而不動”,下聯為“德相非空非有應隨機以恆周”。此處“隨機”與“隨機應變”中的隨機意思近似。要以隨機的態度處事,才能恆久。

隨機就是有不確定性。由於這是個不確定性到處可見的世界,對隨機的內涵有些基本的了解,當有助於我們在這隨機世界活得更好。統計學家Smith曾說(見Smith(1984))“在我看來,任何尋求以單一答案解決複雜的不確定性狀態的科學推論方法,都是基於威權心態,對理性學習過程的拙劣模仿”(見張定綺譯(1998)p.147)。的確是如此,在隨機世界裡,多半狀態是不確定的,對事務的分析判斷,往往不能只尋求單一答案。市井小民如此,醫生如此,擔負國家重任者也該如此。以為凡事都該有個明確的選擇者,如果不是威權,就是愚昧了。

從前有個舉人進京準備應試,有天至一算命攤,算命者鐵口直斷他必高中一甲狀元進士。他聽了大喜,開始花天酒地。考前兩天經過那算命

攤,算命者一見到他,便說他印堂發黑,精神恍惚,必落榜無疑。此舉人大驚之下,收心勤奮讀書,總算勉強考上。

事實上,算命者第一次應是認為,以那舉人當時的情況,狀元及第的機會很大,但該舉人自己將其狀態改變了,結果當然有變了(這是所謂條件機率)。

## 2 隨機變數

從事一隨機實驗,或觀察一隨機現象,其所有可能的結果之集合,稱為一樣本空間(sample space),為一非空集合,通常以  $\Omega$  表之。 $\Omega$  中的每一元素,便稱為一樣本,通常以  $\omega$  表一樣本。令  $\mathcal{F}$  表包含  $\Omega$  之所有子集合(subset)之集合,每一  $A \in \mathcal{F}$  便稱為一事件(event)。對每一事件  $A$ , 給定一實數  $P(A)$ , 稱為  $A$  之機率( $P$  所須滿足的條件稍後再說明)。 $(\Omega, \mathcal{F}, P)$  就構成一機率空間(probability space)。

若  $\Omega$  為一包含  $n$  個元素之有限集合,則  $\Omega$  共有  $2^n$  個子集合。例如,取  $n = 20$ , 則  $2^{20} = 1,048,576$ 。隨著  $n$  的增大,子集合之個數成長快速。 $\Omega$  也可以是一無限集合,或一連續集合。例如,丟一銅板,直至出現一正面才停止,觀察總共之投擲數。則  $\Omega = \{1, 2, \dots\}$  為正整數之集合。又若自區間  $[0, 1]$  中任意挑選一數,則  $\Omega = [0, 1]$  為一連續集合。當然,  $\Omega$  還可以是各種奇怪的集合。

在較高等的機率論的書裡,  $\mathcal{F}$  往往允許取成只是  $\Omega$  之一些子集合之集合(但須滿足某些特定條件),而不需是包含  $\Omega$  之所有子集合之集合(見黃文璋(1994)第一章)。這樣放寬對  $\mathcal{F}$  之限制,不但使有些  $\Omega$  之子集合不是事件,也讓一些後續的討論,如隨機變數(random variable),更富彈性與變化。

樣本空間裡之  $P$ , 稱為機率測度(probability measure), 簡稱機率, 它是定義在  $\mathcal{F}$  上之一實函數, 且滿足

(i)  $\forall A \in \mathcal{F}, P(A) \geq 0$ ;

(ii) 若  $A_n, n \geq 1$ , 為互斥(mutually exclusive)事件, 則

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n);$$

(iii)  $P(\Omega) = 1$ 。

令  $\phi$  表空集合。二事件  $A, B$ , 若  $A \cap B = \phi$ , 便稱為互斥。條件(i)及(iii)指出, 機率值沒有負的, 且最大值為1, 最小值為0。至於條件(ii)則是要對“可數的無限”(countably infinite)個互斥事件之聯集的機率, 等於那些事件機率之和。在此一集合稱為可數的(countable), 若其為有限集合, 或可數的無限集合, 後者指一無限集合而與自然數的集合有1對1(one to one)且映成(onto)的對應。

**例1.** 丟一銅板  $n$  次, 觀測其所得之正面數。則  $\Omega = \{0, 1, 2, \dots, n\}$ ,  $\mathcal{F}$  為  $\Omega$  之所有子集合之集合, 共有  $2^{n+1}$  個元素。再令  $P(\{i\}) = p_i, i = 0, 1, \dots, n$ , 其中  $\sum_{i=0}^n p_i = 1$ , 且對  $\forall A \in \mathcal{F}$ , 令  $P(A) = \sum_{i \in A} p_i$ , 則  $(\Omega, \mathcal{F}, P)$  便構成一機率空間。

**例2.** 投擲一公正的銅板, 直至出現一正面才停止, 並記載總共之投擲數。則  $\Omega = \{1, 2, 3, \dots\}$  為一無限集合, 因此  $\mathcal{F}$  中有無限多個元素, 包含所有正整數之子集合。又設第一個正面在第  $k$  次投擲才出現之機率為  $2^{-k}, k \geq 1$ 。令  $f(k) = 2^{-k}$ , 則  $\sum_{k=1}^{\infty} f(k) = \sum_{k=1}^{\infty} 2^{-k} = 1$ 。對  $\forall A \in \mathcal{F}$ , 令

$$P(A) = \sum_{k \in A} 2^{-k}。$$

則  $(\Omega, \mathcal{F}, P)$  構成一機率空間。例如, 若令  $A$  表投擲數為偶數之事件, 即  $A = \{2, 4, \dots\}$ , 則

$$P(A) = \sum_{k=1}^{\infty} 2^{-2k} = \frac{1}{3}。$$

一般而言, 有下述例3及例4的結果。

**例3.** 令  $\Omega = \{\omega_1, \omega_2, \dots\}$  為一可數的集合,  $f$  為一定義在  $\Omega$  上之實函數, 且滿足

$$f(\omega_i) \geq 0, \forall i \geq 1, \text{ 且 } \sum_{i=1}^{\infty} f(\omega_i) = 1,$$

$\mathcal{F}$  為  $\Omega$  之所有子集合之集合, 再定義一  $\mathcal{F}$  上之函數  $P$  如下:

$$P(A) = \sum_{\omega \in A} f(\omega), A \in \mathcal{F},$$

則  $P$  為一機率測度。 $(\Omega, \mathcal{F}, P)$  即構成一離散(discrete)型的機率空間。

例4. 投擲一公正的骰子, 觀測所得點數, 則對應上例,

$$f(\omega_i) = \frac{1}{6}, \forall \omega_i \in \Omega = \{1, 2, 3, 4, 5, 6\}。$$

例5. 設  $\Omega$  為實數  $R$  上之一區間(有限或無限均可),  $f$  為  $\Omega$  上之一實函數, 且滿足

$$f(\omega) \geq 0, \forall \omega \in \Omega, \text{ 且 } \int_{\Omega} f(\omega) d\omega = 1。$$

則由下式可定義出一機率測度

$$P([a, b]) = \int_a^b f(\omega) d\omega, \forall a, b \in \Omega, \text{ 且 } a < b。$$

抽象思考的能力是很重要的, 凡事不能皆要眼見才為實。早期機率是處理一些實際遇到的問題, 如自5個人中任取3人組一球隊, 有幾種可能? 一開始的作法是將各種可能性寫出來, 後來便發展出排列組合的技巧。這便已是一大進步了, 否則當數字很大時, 要寫出所有可能的變化也是很辛苦。由寫出所有可能至排列組合, 抽象思考的影子已隱然若現。

但這仍然不夠。當一門學問發展到某一階段, 就必須更抽象化, 把這門學問的精髓取出, 而建立一思想體系。在這體系中, 當然可解決原來所遇到的實際問題, 也可處理更一般的問題。西元1929年, 年輕的Kolmogorov(1903-1987), 發表了一篇極重要的論文“General Theory of Measure and Probability Theory”。在此論文中, 他首度基於測度論(measure theory), 而以公理化的方式來描述機率理論。這位被認為是二十世紀最重要的俄國數學家, 至此為現代機率論奠定了基礎。機率論也因此成為一令人探索不盡的學科。

我們曾強調創意思考之不易。Kolmogorov 可說是極具創意的代表, 武俠小說裡, 高手雖多, 但也大約只有下述人物可與其相比(見金庸(1996)p.74):

他得覺遠傳授甚久, 於這部九陽真經已記了十之五六, 十餘年間竟然內力大進, 其後多讀道藏, 於道家鍊氣之術更深有心得。某一日在山間閒

遊,仰望浮雲,俯視流水,張君寶若有所悟,在洞中苦思七日七夜,猛地裏豁然貫通,領會了武功中以柔克剛的至理,忍不住仰天長笑。

這一番大笑,竟笑出了一位承先啓後,繼往開來的大宗師。他以自悟的拳理、道家沖虛圓通之道和九陽真經中所載的內功相發明,創出了輝映後世、照耀千古的武當一派武功。

後來北遊寶鳴,見到三峰挺秀,卓立雲海,於武學又有所悟,乃自號三丰,那便是中國武學史上不世出的奇人張三丰。

Kolmogorov 在其理論中,引進了機率空間。有時我們說“給我一些空間”。至此機率論有了活動的空間。若只是求一些簡單問題的機率,那時往往不太在意機率空間究竟是什麼。但在處理一些細膩的問題時,所涉及的機率空間究竟是什麼,就可能很重要了。

在不同的情況下,機率空間可以完全不同。有時必須講清楚所欲討論問題之機率空間。這就相當於讓人了解你的活動範圍及遊戲規則。

樣本空間中的元素,可以是很實在的事物,如球、蘋果、銅板及人等。它們各具備一些性質,其中有些是可以測量的。如紅球的個數,蘋果之重量,銅板之正面數。至於對一個人,除了他生理上的特徵,如年紀、身高、體重等,尚有一些數值資料伴隨他。如智商、兄弟姊妹之人數、收入,或過去三年發表之論文總篇數等。換句話說,在很多時候,除了實驗所呈現的表面的結果外,我們可能對此結果之某一數值上的特性更有興趣。甚至即使不是對任何數值上的特性感興趣,我們一向習慣將欲探討的事物數字化。如丟銅板,將正面當作1,反面當作0。而袋中若有紅、白、黑三色球,則可分別以1,2,3表之。

在科學上通常對處理數量較得心應手。除了學生有學號,國民有身分證字號,郵局有郵遞區號。爲了討論聰明才智,人們訂出智商;爲了說明冷熱,有了溫度;空氣污染之嚴重情況,也有指標。其他例子尚有比重、大氣壓力等。

由於以上這些原因,隨機變數因應而生。所謂隨機變數,就是一個定義在樣本空間 $\Omega$ 之實函數。若以 $X$ 表一隨機變數,則此函數之定義域

為  $\Omega$ , 對應域為實數  $R$ , 值域就是  $R$  之一子集合。換句話說, 對  $\forall \omega \in \Omega$ ,  $X(\omega) \in R$ 。隨機變數, 可能只是想將實驗結果數字化, 如

$$X(\text{正面}) = 1, X(\text{反面}) = 0,$$

或是對每一實驗結果, 給一數值資料, 如若  $\Omega =$  某校新生之集合, 則對  $\forall \omega \in \Omega$ , 即對每一新生, 可令

$$X(\omega) = \omega \text{ 之身高。}$$

如果  $\Omega$  本來就是一些數字的集合, 如任選一正整數, 則  $\Omega = \{1, 2, 3, \dots\}$ , 此時一致函數(identity function)  $X(\omega) = \omega$  亦為一隨機變數。其他諸如  $Y(\omega) = \omega^2$ ,  $Z(\omega) = e^\omega$ , 皆為隨機變數。隨機變數通常以英文大寫字母表示。

隨機變數其實是要滿足對  $\forall x \in R$ ,  $\{\omega | X(\omega) \leq x\} \in \mathcal{F}$ , 若將樣本空間  $(\Omega, \mathcal{F}, P)$  中的  $\mathcal{F}$  取得小些, 則有些  $\Omega$  上的實函數就不一定是隨機變數了。此時對一給定的  $\Omega$  上的實函數, 便要檢視其是否為隨機變數, 問題也就變化多端, 挑戰性因而增加。因我們將  $\mathcal{F}$  取成包含  $\Omega$  之所有子集合之集合, 因此每一  $\Omega$  上的實函數皆為隨機變數。

有些初學者對隨機變數常感不解, 不知為何要引進此概念。事實上, 只要記住它是個函數就易明白了。實驗的結果本來是  $\omega$ , 映至實數  $X(\omega)$ ,  $X(\omega)$  是個變數, 但因是個隨機實驗的結果, 所觀察到的  $\omega$ , 每次可能不同(不若一般函數  $f$ , 指定一  $x$ , 而給出函數值  $f(x)$ ), 所以稱為隨機變數。

對一隨機變數  $X$ , 我們定義其分佈函數(distribution function)  $F$  為

$$F(x) = P(X \leq x) = P(\{\omega | X(\omega) \leq x\}), \forall x \in R。$$

嚴格講  $P(X \leq x)$  應寫成  $P(\{X \leq x\})$ , 因  $\{X \leq x\}$  才是一集合。不過由於不致混淆, 往往只簡寫成  $P(X \leq x)$ , 而  $P(\{\omega | X(\omega) \leq x\})$  也就寫成  $P(\omega | X(\omega) \leq x)$ 。

若一隨機變數  $X$  之所有可能值的集合為可數的, 則此隨機變數稱為離散的。即存在一可數的集合  $C = \{x_1, x_2, \dots\}$ , 使得

$$\sum_{i \in C} f(x_i) = 1,$$

其中  $f(x_i) = P(X = x_i)$ 。而對任一實數集合  $A$ ,

$$P(A) = P(X \in A) = \sum_{x \in A \cap C} f(x),$$

且對  $\forall x \in R$ ,

$$F(x) = \sum_{y \leq x, y \in C} P(X = y) = \sum_{y \leq x, y \in C} f(y)。$$

由於在實際應用時,常會考慮型如  $\{X \leq x\}$  之事件的機率,所以才會定義分佈函數。在一機率空間上,我們可定義很多不同的隨機變數,也就是增加了其變化性。由隨機變數再轉入分佈函數,可說將其變化性減小。很多完全不同的隨機變數,卻可有相同的分佈函數(見底下例7及例8)。這一放一收之間,端視我們要探討問題到多細膩。

另外,若存在一非負函數  $f$ ,滿足

$$\int_{-\infty}^{\infty} f(x)dx = 1,$$

且對任意  $a < b$ ,

$$(1) \quad P(a < X \leq b) = \int_a^b f(x)dx,$$

則  $X$  稱為絕對連續(absolutely continuous,有時只簡單地說連續)的隨機變數,由(1)得

$$F(x) = \int_{-\infty}^x f(y)dy。$$

在前述離散及絕對連續的隨機變數中的  $f$ ,稱為該隨機變數之機率密度函數(probability density function,簡稱p.d.f.)。

**例6.** 在例1中,令  $X(\omega) = \omega$ ,表丟一銅板  $n$  次所得之正面數。則

$$F(x) = \begin{cases} 0 & , x < 0, \\ \sum_{i=1}^{[x]} p_i & , 0 \leq x < n, \\ 1 & , x \geq n, \end{cases}$$

其中  $p_i = P(X = i)$ ,  $[x]$  表不超過  $x$  之最大整數。

又若丟銅板每次出現的結果設為獨立,且每次出現正面的機率設為  $p$ ,則此時

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \dots, n。$$

具有這種 p.d.f. 之隨機變數,稱為有參數  $n$  及  $p$  之二項分佈(binomial distribution),以  $\mathcal{B}(n, p)$  表之,其中  $n$  為正整數,  $0 \leq p \leq 1$ 。

二隨機變數之分佈函數(或機率密度函數)若相同,便稱為有相同分佈(identically distributed)。分佈相同不表二隨機變數相等。見下二例。

**例7.** 設丟一公正的銅板  $n$  次,令  $X, Y$  分別表出現的正面數及反面數,則  $X, Y$  皆有  $\mathcal{B}(n, 1/2)$  分佈。但因  $X + Y = n, X = Y$  當然不對。

**例8.** 設  $X$  有標準常態分佈(standard normal distribution),以  $\mathcal{N}(0, 1)$  表之。令  $Y = -X$ ,則  $Y$  亦為  $\mathcal{N}(0, 1)$  分佈。即  $X$  與  $Y$  之分佈相同。但  $X = 3$  時,  $Y$  必為  $-3$ ,故  $X = Y$  當然不對。

簡單地講,由於我們討論的是隨機現象,二隨機變數之分佈相同,表其會有相同的隨機變化。在例6中,  $X = i$  之機率與  $Y = i$  之機率皆相同,  $i = 0, 1, \dots, n$ 。即會出現  $i$  個正面與會出現  $i$  個反面之機率相同。但此並不表出現的正面數要等於反面數。即使是同一個銅板,連丟數次,每次出現的面當然不一定要相同,雖每次出現正面的機率皆相等。這就是隨機現象迷人之處,你永遠不知你的下一個結果,即使你知道其分佈情況。

設有二隨機變數  $X, Y$ , 定義在同一機率空間。對任二實數  $x, y$ , 有時我們會想知道事件  $\{X \leq x\}$  與  $\{Y \leq y\}$  同時發生之機率。此機率以  $P(X \leq x, Y \leq y)$  表之,其意義為  $P(\omega | X(\omega) \leq x, \text{且 } Y(\omega) \leq y)$ 。我們便以  $F(x, y) = P(X \leq x, Y \leq y), x, y \in R$ , 當作  $X, Y$  之聯合分佈函數(joint distribution function)。

由  $F(x, y)$  可得  $X$  及  $Y$  之邊際分佈函數(marginal distribution func-



tion):

$$P(X \leq x) = \lim_{y \rightarrow \infty} F(x, y),$$
$$P(Y \leq y) = \lim_{x \rightarrow \infty} F(x, y)。$$

但反過來,若只知  $X$  及  $Y$  之邊際分佈,並無法得知  $X$  及  $Y$  之聯合分佈。例如,若  $X$  及  $Y$  皆有  $B(n, 1/2)$  分佈,是否可得到  $X, Y$  之聯合分佈? 答案是否定的。如對例6中的  $X, Y$ ,

$$P(X = i, Y = j) = \begin{cases} 0 & , j \neq n - i, \\ P(X = i) & , j = n - i。 \end{cases}$$

另外,若  $X$  有  $B(n, 1/2)$  分佈,而  $Y = X$ ,則  $Y$  當然亦有  $B(n, 1/2)$  分佈,此時

$$P(X = i, Y = j) = \begin{cases} 0 & , j \neq i, \\ P(X = i) & , j = i。 \end{cases}$$

甚至還可有其他可能。

這並非太奇怪的事。假設你了解某甲之行爲變化,也了解乙的,但二人放在一起會如何呢? 兩人可能互不影響,也可能互相干擾,除非有進一步資訊,否則會如何就不得而知了。

一二變數之非負函數  $f$ , 若滿足

$$P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv, \quad \forall x, y \in R,$$

便稱爲  $X, Y$  之聯合機率密度函數(此爲連續型)。而若存在一平面上的可數集合  $C$ , 及一非負函數  $f(x, y) = P(X = x, Y = y), \forall (x, y) \in C$ , 且

$$\sum_{(x, y) \in C} f(x, y) = 1,$$

則此爲離散型的情況。

二隨機變數  $X$  及  $Y$ , 若滿足

$$(2) \quad P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y), \quad \forall x, y \in R,$$

便稱為獨立(independent),否則稱為不獨立(dependent)。設有  $X, Y$  二隨機變數,機率密度函數分別為  $g$  及  $h$ ,聯合機率密度函數為  $f$ 。則  $X$  與  $Y$  獨立,若且唯若

$$(3) \quad f(x, y) = g(x)h(y), \quad \forall x, y \in R。$$

我們也可將以上聯合分佈及獨立的概念推廣到  $n$  個變數的情況,讀者不妨去翻閱一般機率論的書。這裡的獨立,又稱為隨機獨立(stochastically independent),是指二隨機變數機率上的“行為”互不受對方影響。與國家有自治能力,內政外交均不受他國干涉之“主權獨立”中的獨立,意義不盡相同。即使是同一個人,用同一銅板,兩次丟出之結果,也可以是獨立,只要(2)式成立。在線性代數中,所謂向量之線性獨立(linearly independent),意義也不一樣。其他中文裡尚有金雞獨立、遺世而獨立,以及獨立三邊靜及獨力揚新令(分別見唐詩三百首劉長卿的送李中丞歸漢陽別業及盧綸的塞下曲),都有站立的意思,與隨機獨立毫不相干。

對一事件  $A$ ,令  $I_A$  表  $A$  之指示函數(indicator function)。即若  $A$  發生,則  $I_A = 1$ ,否則  $I_A = 0$ 。 $I_A$  便是“指示”事件  $A$  是否發生,因此

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

$I_A$  為一隨機變數,只取0及1兩個可能的值,為一型式簡單的隨機變數。

**例9.** 考慮有三個小孩之家庭。令事件  $A$  表家庭中有男孩及女孩之事件, $B$  表最多有一女孩之事件。令  $X = I_A, Y = I_B$ 。 $(X, Y)$  之值只有4種可能,即(0,0), (0,1), (1,0)及(1,1)。因(假設生男生女之機率皆為1/2)

$$P(X = 1) = \frac{3}{4}, \quad P(X = 0) = \frac{1}{4}, \\ P(Y = 1) = \frac{1}{2}, \quad P(Y = 0) = \frac{1}{2},$$

且

$$P(X = 1, Y = 1) = P(\text{二男孩一女孩}) = \frac{3}{8} = P(X = 1)P(Y = 1),$$

$$\begin{aligned}
P(X = 1, Y = 0) &= \frac{3}{8} = P(X = 1)P(Y = 0), \\
P(X = 0, Y = 1) &= \frac{1}{8} = P(X = 0)P(Y = 1), \\
P(X = 0, Y = 0) &= \frac{1}{8} = P(X = 0)P(Y = 0),
\end{aligned}$$

故  $X$  與  $Y$  獨立。可驗證若改為家庭中有兩個(或四個)小孩,則  $X$  與  $Y$  便不獨立。本例顯示有時獨立性並不是那麼容易看出。

### 3 隨機過程

有時我們會觀測一系列之隨機變數。如依序丟一銅板,令  $X_1, X_2, \dots, X_n$  分別表  $n$  次結果。又如令  $X(t)$  表某校至時間  $t$  之畢業生總數。在這些例子裡,我們分別得到隨機過程(stochastic process,或稱random process) $\{X_i, 1 \leq i \leq n\}$ 及 $\{X(t), t \geq 0\}$ 。我們並將其中之指標(index) $i$ ,及  $t$  視為“時間”,可以是一離散集合,或一連續集合;可以是一有限集合,也可以是一無限集合。一般可表示為  $\{X(t), t \in T\}$ ,  $T$  為指標之集合(index set)。 $X(t)$  即表此過程在時間  $t$  之值,或說狀態(state)。

早期有所謂人口函數,如以  $f(t)$  表某地區在時間  $t$  之人口數。也就是對每一時間  $t$ ,以一定值  $f(t)$  表其人口。但在隨機世界裡,在任一時刻  $t$ ,其人口數應皆為一隨機變數。所以以一隨機過程做為其依時間而變之人口的模式,應是較合理的。由隨機變數至隨機過程,也就成為一自然的推廣。我們再略微說明如下。

大家都知道,由常數至函數,在概念上是一大突破。自小學起我們學數量,每個學生體重若干,汽車時速若干,都是一定值。在中學時,我們學了函數,學生體重可不相同了,第  $n$  個學生重  $a_n$ ;汽車也可不以等速行駛,以  $f(t)$  表在時間  $t$  之速度。學生不再是等重,汽車也不再是等速行駛。函數的概念當然比常數的概念合理且較一般,事實上常數可視為一種特別的函數。

然後我們進一步有隨機變數的概念。也就是任選的一學生,其體重不一定是定值;汽車在一特定時間,速度也可有不同的值。如果再考慮時間

的改變,就得到一雙重的“變”,隨機過程便產生了。由於其合理性及一般性,隨機過程遂成爲討論很多實際的問題中,所採用的模式。

能以隨機性來看事物之有所變,對我們做決策會有很大幫助。我們引一段莊子山木篇如下:

莊子行於山中,見大木,枝葉盛茂,伐木者止其旁而不取也。問其故,曰“無所可用。”莊子曰“此木以不材得終其天年。”夫子出於山,舍於故人之家。故人喜,命豎子斂雁而烹之。豎子請曰“其一能鳴,其一不能鳴,請悉斂?”主人曰“斂不能鳴者。”明日,弟子問於莊子曰“昨日山中之木,以不材得終其天年,今日之雁,以不材死,先生將何處?”

看事情絕非簡單地分爲不材好或不材不好。更何況與時推移,我們所處的是一隨機過程的世界,更不應只以單一的答案看萬物。

#### 4 隨機取樣

在機率或統計的應用中,我們常會提到所謂“隨機”,底下解釋一些其主要的含意。

第一種含意爲隨機變數的隨機樣本。設  $X$  爲一隨機變數,且設  $X$  可重複觀察,而得到  $X_1, X_2, \dots$ 。例如,若  $X_i$  表第  $i$  個燈泡的壽命,  $\{X_i, i \geq 1\}$  設爲獨立且有共同分佈(independent and identically distributed, 簡稱 i.i.d.), 則  $X_1, X_2, \dots, X_n$  稱爲隨機變數  $X$  之  $n$  個隨機樣本(random sample of size  $n$ )。此處“隨機”樣本,表示這些樣本爲 i.i.d. 之隨機變數。

第二種含意是針對有  $\Omega$ -的母體。自一有  $\Omega$ -的母體中抽取一樣本,若每個樣本均有相同的機率(即樣本數之倒數)被抽中,則我們可以說自此母體中隨機地抽取一樣本。

第三種是描述下述情況:自區間  $(a, b)$  中隨機地取  $n$  個點,若以  $X_1, X_2, \dots, X_n$  表取中之點,則  $X_1, X_2, \dots, X_n$  爲 i.i.d.  $U(a, b)$  之隨機變數。在

此,若隨機變數  $X$  之機率密度函數為

$$f(x) = \begin{cases} (b-a)^{-1}, & x \in (a, b), \\ 0 & , \text{其他}, \end{cases}$$

則稱  $X$  在  $(a, b)$  上有均勻分佈(uniform distribution),以  $\mathcal{U}(a, b)$  表之。同理可定義  $\mathcal{U}[a, b)$ ,  $\mathcal{U}(a, b]$  及  $\mathcal{U}[a, b]$ 。

**例10.** 假設平面上有一圓,半徑為  $r$ 。在圓上隨機地畫一條弦,我們想求此弦長度大於圓之內接等邊三角形之邊長  $a$  的機率。令  $A$  表一隨機的弦其長度會大於  $a$  的事件。而一隨機的弦其長度由下述三者之一皆可決定:此弦與圓心之距離  $D$ ; 此弦的中點位置  $M$ ; 此弦與圓心所張開的角  $\theta$ 。底下分別假設  $D, M, \theta$  在其範圍內均勻分佈,而求出  $A$  之機率。

(1) 設  $D$  在  $[0, r]$  均勻分佈。此假設是基於若有一隻尺,自與圓相切處以等速平行地向圓心移動,隨時停止便與圓交出一條弦。由幾何性質知,事件  $A$  發生,若且唯若  $D < r/2$ , 故  $P(A) = 1/2$ 。

(2)  $M$  在整個圓內均勻分佈。此假設是基於由圓內任意取一點  $M$ , 當做弦的中點,自  $M$  做垂直此點與圓心的連線便得到一弦。易見此弦長度是否大於  $a$ , 就看  $M$  是否落在半徑為  $r/2$  之同心圓內。故  $P(A) = P(M \text{ 落在小圓內}) = \text{小圓面積}/\text{大圓面積} = 1/4$ 。

(3) 設  $\theta$  在  $[0, 2\pi]$  均勻分佈。此假設是基於自圓周上任取一點當做弦的一個端點,然後自此點沿著圓周順時針等速移動,隨時停止便得到弦的另一端點。易見事件  $A$  發生,若且唯若此弦二端點與圓心所張開之角  $\theta$  介於  $2\pi/3$  與  $4\pi/3$  之間,故  $P(A) = (4\pi/3 - 2\pi/3)/2\pi = 1/3$ 。

故在不同的假設下,本問題的答案可以是  $1/2, 1/4$ , 或  $1/3$ 。甚至在另外的假設下,可能得到其他答案。此問題即法國數學家 Bertrand (1822-1900), 在西元 1889 所提出的詭論(Bertrand paradox)。在對“隨機”有不同的解釋下,得到不同的答案,其實是不足為奇的。

靈敏的讀者也很可能已看出,在(1)、(2)及(3)中之機率空間皆不相同,甚至連樣本空間也不相同(分別為區間  $[0, r]$ , 整個圓及區間  $[0, 2\pi]$ )。機率空間是討論機率的根本,機率空間不同,就彷如雞同鴨講,不能期待有相

同的結論。

前述Bertrand 詭論,將幾何引進了機率。這類問題,就是所謂幾何機率(geometric probability)。其起源可溯自西元1777年提出的Buffon針問題(Buffon's needle problem)。Buffon (1707-1788) 為法國博物學家(naturalist),他自西元1733年起,開始對將機率應用在幾何上的問題感興趣。他利用積分巧妙地解出了針問題,雖然其解法中有一錯誤,而於他提出34年後,才由另一法國數學家Laplace (1749-1827)所更正。

例11.設一平面上有無限條間距為 $l$ 之平行線。隨機地投擲一長度為 $a$ 之針至此平面上, $a < l$ ,求針會與其中某一條平行線相交之機率。

解. 將平行線的一個方向叫做左,另一方向叫做右,而與平行線垂直的方向一稱為上,一稱為下。令 $X$ 表針與右方向之夾角; $Y$ 表針下方那個端點與最接近的上方的一條平行線的距離。則所謂隨機地投擲一針至平面上,即表 $X, Y$ 分別有 $\mathcal{U}[0, \pi)$  及 $\mathcal{U}(0, l]$  分佈。若設 $X$ 與 $Y$ 獨立,則 $X, Y$ 之聯合p.d.f. 為

$$f(x, y) = \frac{1}{\pi l}, x \in [0, \pi), y \in (0, l]。$$

令 $A$ 表針會與某一平行線相交之事件,則 $A$ 發生,若且唯若 $Y \leq a \sin X$ 。

故

$$(4) \quad P(A) = P(Y \leq a \sin X) = \int_0^\pi \int_0^{a \sin x} \frac{1}{\pi l} dy dx = 2a/\pi l。$$

前面設 $X$ 與 $Y$ 獨立,此假設可經由實驗來檢定。事實上,若重複地投擲一針至一平面上 $n$ 次,且令 $n(A)$ 表其中與平行線相交之次數,則

$$\frac{n(A)}{n} \doteq \frac{2a}{\pi l}。$$

故 $n$ 很大時,  $(2an)/(ln(A))$ 應為圓周率 $\pi = 3.14\dots$ 之一很好的估計值,而實驗顯示結果很正確。關於此問題的討論可參考Uspensky (1937), p.113,及Gani(1980/81)。

例11中之步驟乍看之下是有些奇怪的,因我們是以一隨機的設計來估計一非隨機的量 $\pi$ 。事實上,這種想法就是蒙地卡羅法(Monte Carlo

method) 的基礎, 為一有用的技術。當直接計算有困難時, 可藉助計算機及機率的設計以解之。

底下亦為一有趣的例子。

**例12.** 自區間  $(0, 1)$  隨機地取兩數, 以  $X$  及  $Y$  表之。求最接近  $X/Y$  的整數恰為一偶數之機率。

**解.** 有對稱性概念的人, 會以為此機率應為  $1/2$ 。我們來看對不對。首先由題意知,  $X, Y$  為獨立的隨機變數, 且皆有  $U(0, 1)$  分佈。對任意  $X$  及  $Y$ , 最接近  $X/Y$  的整數恰為偶數, 若且唯若  $X/Y < 0.5$ , 或存在一正整數  $n$ , 使得  $2n - 0.5 < X/Y < 2n + 0.5$ 。

現因

$$P(X/Y < 0.5) = P(2X < Y) = \int_0^{0.5} \int_{2x}^1 dy dx = \frac{1}{4},$$

又

$$\begin{aligned} & P(2n - 0.5 < X/Y < 2n + 0.5) \\ &= P\left(\frac{2X}{4n+1} < Y < \frac{2X}{4n-1}\right) \\ &= \int_0^1 \left(\frac{2x}{4n-1} - \frac{2x}{4n+1}\right) dx \\ &= \frac{1}{4n-1} - \frac{1}{4n+1}, \end{aligned}$$

故

$$\begin{aligned} & P(\text{最接近 } X/Y \text{ 的整數恰為一偶數}) \\ &= \frac{1}{4} + \sum_{n=1}^{\infty} \left(\frac{1}{4n-1} - \frac{1}{4n+1}\right) \\ &= \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \frac{1}{9} + \dots\right) \\ &= \frac{1}{4} + 1 - \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots\right) \\ &= \frac{1}{4} + 1 - \arctan 1 \\ &= \frac{1}{4} + 1 - \frac{\pi}{4} = \frac{5-\pi}{4} < \frac{1}{2}. \end{aligned}$$

事實上我們所求即為  $X/Y \in S$  之機率,其中

$$S = (0, 0.5) \cup (1.5, 2] \cup (2, 2.5) \cup (3.5, 4] \cup (4, 4.5) \cup \dots,$$

包含每一整數區間  $(i, i+1), i \geq 0$ , 的一半部分。但此機率卻小於  $1/2$ , 即使  $X, Y$  原為 i.i.d. 之均勻分佈。

隨機取樣應用廣泛, 我們再給兩例。

**例13.** 某民意調查機構之一訪問員問某受訪者, 上次總統選舉時投給那一位後選人? 受訪者不願回答。訪問員靈機一動說“那請問一下, 你投的候選人後來當選了嗎?” 受訪者立即高興地說“當然當選了。”有時在調查一些較敏感性的問題, 受訪者可能不樂意回答, 或不誠實回答。譬如說涉及名譽、隱私等問題, 一般人大概很難誠實回答。底下我們提供一種常用的隨機回答的技巧(randomized response technique)。

這種技巧就是調查機構先設計二問題, 一個是所欲問的敏感性的問題, 一個是一不相干的問題。經由一隨機的實驗, 如丟一銅板、丟一骰子、抽一張撲克牌, 或自袋中取一球, 以決定受訪者回答那一問題。但訪問員並不知道受訪者回答那一問題。

例如, 問題一為敏感問題: 你是否考試作弊過? 問題二為: 你是否在四月出生? 丟一公正銅板, 若出現正面則回答問題一, 否則回答問題二。

設  $p$  為受訪者曾作弊過之機率, 這是我們所想估計的, 而假設受訪者在四月出生之機率為  $1/12$ 。令  $\lambda$  表回答“是”之機率,  $H$  表正面出現之事件,  $H^c$  表反面出現之事件,  $Y$  表回答“是”之事件。則

$$\begin{aligned}\lambda &= P(Y) = P(Y|H)P(H) + P(Y|H^c)P(H^c) \\ &= \frac{1}{2}p + \frac{1}{2} \cdot \frac{1}{12},\end{aligned}$$

即得  $p = 2\lambda - 1/12$ 。由受訪者回答“是”之比例  $\hat{\lambda}$  來估計  $\lambda$ , 則可以  $2\hat{\lambda} - 1/12$  來估計  $p$ 。假設實際調查時, 得到  $\hat{\lambda} = 0.44$ , 則  $p$  之估計值為  $0.88 - 1/12$ 。

**例14.** 有時你會看到某種野生動物只剩多少頭的報導, 你是否會奇怪是如



何估計的,難道是全抓到而數一遍?事實上經由“記號法”,也就是捕捉後做記號放走,隔一段時間後看捉到的之中,有幾頭是有做記號的,由其間的比值,便可拿來做為該動物現存幾頭之估計值。

例如,自某魚池抓取1,000條魚,做記號後放回。數日後又抓了800條魚,並發現其中有100條有記號。設魚池中有  $N$  條魚,利用

$$\frac{100}{800} = \frac{1,000}{N},$$

解出  $N$  為8,000。即估計魚池中有8,000條魚。

上述做法當然要有一些假設,你能想到一些嗎?

隨機取樣用途廣泛,例如在民意調查中扮演著重要的角色。對隨機的內涵,有一最基本的了解是必要的。

## 習 題

1. 丟一公正的銅板3次,令  $X, Y$  分別表所得之正面數及反面數。試證  $X$  與  $Y$  之分佈相同,但對每一樣本  $\omega$ ,  $X(\omega)$  與  $Y(\omega)$  皆不相等。
2. 設有二人約會,時間為下午5時至6時間,兩人並約好最多只等對方10分鐘。若兩人之抵達時刻為獨立,且皆在下午5時至6時間均勻分佈,求兩人會碰面之機率。
3. 設某路線公車自上午6時起,每15分鐘自起站發一班車。 $A, B$  二人在6時至7時間,獨立且隨機地抵達起站。試求二人會搭上同一班車之機率。
4. 在例9中,針對一家庭有(i)兩個, (ii)四個小孩,分別驗證此時  $X$  與  $Y$  並不獨立。

5. 將  $2r$  個球隨機地放進  $r$  個盒子中,令  $X_i$  表第  $i$  個盒子中之球數。求
- (i)  $X_1, X_2, \dots, X_r$  之聯合p.d.f.;
  - (ii) 每個盒子中恰有二球之機率。
6. 隨機地將一竹子折成兩段,試求長段之長度超過短段之長度2倍之機率。
7. 在區間  $[0, 1]$  隨機地任取兩點  $X$  及  $Y$ 。以  $X, Y$  為分割點,將此區間分為三段。求此三段構成一三角形之機率。
8. 任取兩正整數,求其會互質之機率。(提示:利用  $\sum_{i=1}^{\infty} i^{-2} = \pi^2/6$ )。
9. 試以例10中之三個不同的假設,求在圓上隨機地畫一條弦,其長度會大於圓半徑的機率。
10. 設  $X, Y$  為二隨機變數。令  $U = X + Y, V = X - Y$ 。
- (i) 試舉一組獨立的  $X$  及  $Y$ ,且  $U$  與  $V$  不獨立;
  - (ii) 試舉一組不獨立的  $X$  及  $Y$ ,且  $U$  與  $V$  獨立。
11. 設有三隨機變數  $X, Y, Z$ 。試問若  $X + Y$  與  $X + Z$  之分佈相同,是否必導致  $Y$  與  $Z$  分佈相同。

## 參考文獻

1. 金庸(1996).倚天屠龍記,第三版。遠流出版社,台北。
2. 黃文璋(1994).機率論講義。全華科技圖書股份有限公司,台北。
3. 張定綺譯(1998).與天為敵—人類戰勝風險的傳奇故事。商業周刊出版股份有限公司,台北。
4. Gani, J. (1980/81). The problem of Buffon's needle. *Mathematical Spectrum* **13**, 14-18.

5. Smith, A. F. M. (1984). Present position and potential developments: some personal views of Bayesian statistics. *Journal of Royal Statistical Association* **147**, part3, 245-259.
6. Uspensky, J. V. (1937). *Introduction to Mathematical Probability*. McGraw-Hill Book Company, Inc., New York.