什麼是機率與機率法則?

蔡聰明

機率是機率論中的一個核心概念, 深奧而難纏。丟一個公正銅板, 出現正面的機率是 1/2: 這絕不是說丟兩次銅板必出現一次正面, 也不是丟 10 次銅板必出現 5 次正面。甚至也不是說: 當 n 越來越大時, 丟 n 次銅板必會出現 $n/2$ 次正面。大家都聽過一則笑話: 醫生 對患有癌症的病人說,「我替病人開刀成功的機率是 1/10, 但是在你之前已有 9 位病人被我刀到命除, 你是幸運的第 10 位, 因此你開刀必然成功」。這是對機率作錯誤的解釋。

為了說明方便起見, 我們用隨機變數

$$\xi = \begin{cases} 1 & \text{正面, 機率 } \frac{1}{2} \\ 0 & \text{反面, 機率 } \frac{1}{2} \end{cases}$$

來表現丟一個公正銅板的實驗。丟出正面就報 1, 丟出反面就報 0, 此地 0 與 1 完全是為了點算正面的次數而設定的。今將此銅板獨立地一次接著一次丟下去, 令 ξ_n 表示第 n 次的可能結果, 於是就得到銅板序列 (ξ_n)。這是獨立且同佈 (i.i.d.) 的一列隨機變數, 它們都是 ξ 的抄本 (copies), 定義在某個機率空間 $(\Omega, \mathcal{F}, \mathcal{P})$ 上。這裡有兩個基本的問題:

問題1: 銅板序列 (ξ_n) 存在嗎? 說得更具體一點, 能否建構一個機率空間 $(\Omega, \mathcal{F}, \mathcal{P})$ 並且在其上定義一列獨立、同佈的隨機變數 (ξ_n), 使得滿足

$$P(\xi_n = 1) = \frac{1}{2} = P(\xi_n = 0), \forall n \in N?$$

問題2: 機率論為何要採用「機率空間」$(\Omega, \mathcal{F}, P)$ 之理型? 即為何 $\mathcal{F}$ 要取為 σ 代數 (σ-algebra), 不取為較接近常識的代數 (algebra), 而 P 要取為具有可列加性 (σ-additivity), 不取為具有有窮加性 (finite-additivity)?

銅板序列的存在性是沒有問題的, 而第二個問題更深刻。本文我們不預備討論它們。

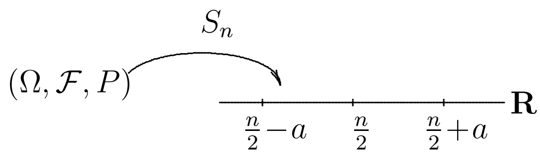
令 $S_n = \sum_{k=1}^n \xi_k$, 這是一個隨機變數, 表示丟 n 次銅板中出現正面的次數。它具有二項分佈

$$P(S_n = k) = {}_nC_k (\frac{1}{2})^k (\frac{1}{2})^{n-k}.$$

在「談 Stirling 公式」一文中 [5], 我們證明過: 將「機率 1/2」解釋為「在丟 n 次銅板中出現正面的次數“差不多”佔有一半, 即使當 $n \to \infty$ 時也不成, 因為對任意實數 $a > 0$, 恆有

$$\lim_{n\to\infty} P(-a \le S_n - \frac{n}{2} \le a) = 0$$

即以區間 $[\frac{n}{2}-a, \frac{n}{2}+a]$ 來網接 S_n 散佈過來的機率, 當 $n\to\infty$ 時, 仍然為0。



上述都是對於「機率1/2」作負面的斷言。不過我們應注意對偶的作用: 當我們知道夠多的負面斷言後, 對於正面的斷言就逐漸可以掌握。另外, 知道負面的斷言, 往往就已經很有用, 例如熱力學的定律皆斷言某些事情辦不到, 存在五次方程式沒有根式解, 幾何三大難題無法尺規作圖, 否定的解決也是解決。

本文我們要來追尋「什麼是機率」的正面答案, 乃至進一步的「機率法則」這個看似矛盾的概念。機率現象是說不準的, 居然有機率法則可循, 這不是既矛盾又神奇嗎?

一.Bernoulli 的弱大數法則

理想彼岸的「機率1/2」(Plato 理念世界) 與現實此岸的正面次數之「相對頻率 $\frac{S_n}{n}$」(經驗世界), 我們有很強的直覺相信它們具有密切的關係, 其間的橋樑是什麼呢? 這就是 James Bernoulli (1654-1705) 要解決的問題。他積20年的辛苦工作 (比美於 Tycho Brahe 20年的天文觀測, 提供數據讓 Kepler 發現行星運動三大定律), 終於有了突破性的發現, 得到機率論的第一個極限定理, 今日叫做 Bernoulli 的弱大數法則。

這個重要發現是如何得到的呢? 讓我們解說於下。

根據上述, 既然用任何有限長度的區間 $[-a,a]$ 在 $n\to\infty$ 時都網不到 $S_n-\frac{n}{2}$ 的任何機率, 自然就想到改用跟 n 有關的變動長度之區間。我們的目標是要網到所有的機率1。這樣的區間是什麼形式呢? n 的等級 (order) 是什麼?

首先觀察到下列事件都是等價的鐵定事件 (sure event):

$$
\begin{aligned}
& 0 \le S_n \le n \\
\Leftrightarrow\ & \mid S_n-\frac{n}{2} \mid \le \frac{n}{2} \\
\Leftrightarrow\ & \mid \frac{S_n}{n}-\frac{1}{2} \mid \le \frac{1}{2}
\end{aligned}
$$

即它們的機率皆為1。由此看出, 只要取 $l\ge\frac{1}{2}$, 則區間 $[-ln, ln]$ 就可網接 $S_n-\frac{n}{2}$ 的所有機率, 即 $P(|\frac{S_n}{n}-\frac{1}{2}|\le l)=1$, 此時根本不必用到 $n\to\infty$ (即不必訴諸無窮)。不過這只是一句大而無當的空話 (tautology), 相當於說「白馬是馬」。

顯然 l 越小越好。配合訴諸 $n\to\infty$, 可否將「$l\ge\frac{1}{2}$」作精進呢? 這是整個問題的關鍵, 也是困難的所在。

從對二項分佈的澈底研究中, Bernoulli 發現「$l\ge\frac{1}{2}$」可精進為「$l>0$」。這雖只是一小步, 但是在機率論史上卻是一大步! 一個偉大的時刻 (a great moment)!

一個好的圖解往往勝過千言萬語, 直指本心地反應出事物的本質。讓我們來觀察隨機變數 S_n 與 $\frac{S_n}{n}$ 散佈機率的圖像 (picture):

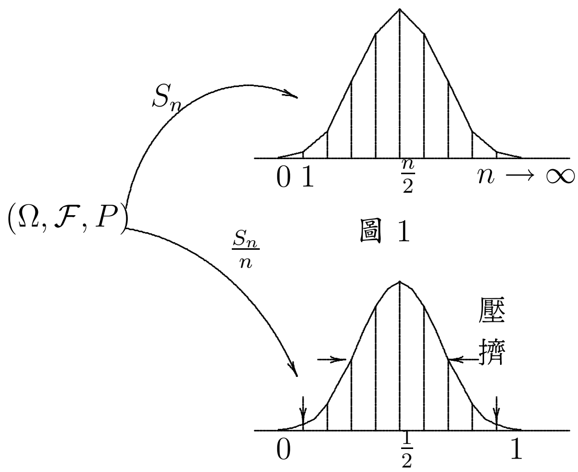


圖 1

圖 2

我們知道 S_n 與 $\frac{S_n}{n}$ 都具有二項分佈, 其圖形是左右對稱的, 中心點分別爲 $n/2$ 與 $1/2$, 而且最大項的等級爲 $\frac{1}{\sqrt{\pi n}} \to 0$, 當 $n \to \infty$ 時。

從 S_n 的眼光來看, 它將全部的機率 1, 散佈得越來越寬廣。當 $n \to \infty$ 時, 流佈於 $N = \{0, 1, 2, \ldots\}$ 之上, 但是每一點 $n \in N$ 。所分得的機率越來越小, 而趨近於0, 此圖不易看出機率逐漸聚集於哪一個範圍的趨勢。

改從調整過尺度的 $\frac{S_n}{n}$ 之眼光來看, $\frac{S_n}{n}$ 將全部機率 1 散佈在 $[0, 1]$ 之中, 當 $n \to \infty$ 時有越來越密集於 $1/2$ 點附近的趨勢, 而分在兩側的機率越來越被擠壓而消失掉 (參見圖 2)。

在這個觀察基礎下, Bernoulli 猜測到上述「 $l \geq \frac{1}{2}$ 」可以改成「 $l > 0$ 」: 即對任意 $\varepsilon > 0$, 當 $n \to \infty$ 時, 區間 $[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]$ 可以網接 $\frac{S_n}{n}$ 散佈過來的所有機率。更精鍊地說:

$$\text{對任意 } \varepsilon > 0, \lim_{n\to\infty} P(|\frac{S_n}{n} - \frac{1}{2}| \leq \varepsilon) = 1.$$

換言之, 無窮 $(n \to \infty)$ 像一根魔杖, 透過 $\frac{S_n}{n}$ 將機率全部趕入網羅 $[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon]$ 之中。值得注意的是, 微積分的求面積與求切線兩難題也是訴諸無窮才解開謎底的。

如果當初的銅板, 出現正面與反面的機率各爲 p 與 q , 其中 $0 < p < 1$ 且 $q = 1 - p$, 那麼上述的猜測可稍作推廣成爲:

$$\text{對任意} \varepsilon > 0, \lim_{n\to\infty} P(|\frac{S_n}{n} - p| \leq \varepsilon) = 1. \tag{1}$$

如何證明這個猜測呢? 當初 Bernoulli 是對二項分佈作估計, 而證出 (1) 式。這個證明, 不論是從歷史或數學或方法論的觀點來看, 都深具興味。

我們只需證明

$$\lim_{n\to\infty} P(|\frac{S_n}{n} - p| > \varepsilon) = 0 \tag{2}$$

就好了。這又等價於兩個「尾巴部份」(tail parts) 之證明:

$$\lim_{n\to\infty} P(S_n > np + n\varepsilon) = 0 \tag{3}$$

$$\text{與} \lim_{n\to\infty} P(S_n < np - n\varepsilon) = 0 \tag{4}$$

我們只需證明 (3) 式, 因爲將 p 與 q 的角色對調, (4) 式就同理可證。

現在證明 (3) 式:

$$\begin{aligned} & P(S_n > np + n\varepsilon) & (5) \\ = & \sum_{k > np + n\varepsilon} P_k = P_{k_0} + P_{k_0+1} + \cdots + P_n \end{aligned}$$

其中 $P_k = {}_nC_k P^k q^{n-k}$, 並且 k_0 爲第一個滿足 $k > np + n\varepsilon$ 之自然數 k 。我們要來估計 (5) 式。顯然 $P_{k_0}, P_{k_0+1}, \ldots, P_n$, 不是一個等比數列, 但是可以用一個等比數列來控制。爲此, 考慮比值:

$$\begin{aligned}\frac{P_k}{P_{k-1}} &= \frac{{}_nC_k p^k q^{n-k}}{{}_nC_{k-1}p^{k-1}q^{n-k+1}} \\ &= 1+\frac{(n+1)p-k}{kq} \qquad (6)\end{aligned}$$

$$\frac{P_{k+1}}{P_k} = 1+\frac{(n+1)p-(k+1)}{(k+1)q} \qquad (7)$$

由此看出, 當 $k>(n+1)p$ 時,

$$1>\frac{P_k}{P_{k-1}}>\frac{P_{k+1}}{P_k}>\cdots>\frac{P_n}{p_{n-1}}>0 \qquad (8)$$

因爲我們考慮的是 $n\to\infty$ 的情形, 故可取足夠大的 n , 使得 $np+n\varepsilon>(n+1)p$ 。因此, 當 $k>np+n\varepsilon$ 時, 也滿足 $k>(n+1)p$ 。參見下面圖3。

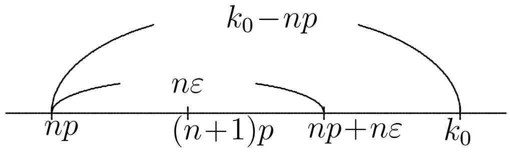


圖 3

特別地

$$1>\frac{P_{k_0}}{P_{k_0-1}}>\frac{P_{k_0+1}}{P_{k_0}}>\cdots>\frac{P_n}{p_{n-1}}>0 \qquad (9)$$

令 $\alpha=1+\frac{np-k_0}{k_0q}$, 則 $0<\alpha<1$ 且

$$\alpha\geq 1+\frac{(n+1)p-(k_0+1)}{(k_0+1)q}=\frac{P_{k_0+1}}{P_{k_0}} \qquad (10)$$

由 (9) 式知

$$\begin{aligned}&P_{k_0+1}<\alpha P_{k_0}\\ &P_{k_0+2}<\alpha P_{k_0+1}<\alpha^2 P_{k_0}\\ &\vdots\\ &P_n<\alpha P_{n-1}<\cdots<\alpha^{n-k_0}P_{k_0}\end{aligned}$$

因此, 由 (5) 式得

$$\begin{aligned}&P(S_n>np+n\varepsilon)\\ \leq\ & P_{k_0}(1+\alpha+\alpha^2+\ldots+\alpha^{n-k_0})\\ <\ & P_{k_0}\cdot\frac{1}{1-\alpha}=P_{k_0}\frac{k_0q}{k_0-np} \qquad (11)\end{aligned}$$

現在估計 P_{k_0} : 在區問 $[np,k_0]$ 內的自然數之個數小於等於 k_0-np 個, 其上的機率皆大於等 P_{k_0} 故

$$1\geq P(S_n\in[np,k_0])\geq(k_0-np)P_{k_0}$$

從而

$$P_{k_0}\leq\frac{1}{k_0-np} \qquad (12)$$

代入 (11) 式得

$$P(S_n>np+n\varepsilon)\leq\frac{k_0q}{[k_0-np]^2}<\frac{n}{[k_0-np]^2} \qquad (13)$$

今因 $k_0-np>n\varepsilon$ (參見圖3), 故得

$$\begin{aligned}&P(S_n>np+n\varepsilon) \qquad (14)\\ &<\frac{n}{(n\varepsilon)^2}=\frac{1}{n\varepsilon^2}\to 0,\ 當\ n\to\infty。\end{aligned}$$

這樣我們就證明了 (1) 式。M. Kac 對 Bernoulli 的這個證明評論說:「它並不難, 但也不頂容易。」(It is not hard but not very easy either!)

定理1 (Bernoulli 的弱大數法則, 1713年): 對任意 $\varepsilon>0,\lim_{n\to\infty}P(|\frac{S_n}{n}-p|\leq\varepsilon)=1$

對於這個重要結果, 我們再介紹另一種簡潔而利於推廣的證法: 首先我們注意到, 期

望值 $E(S_n) = np$, 另外由 (ξ_n) 之獨立性可得變異數, $\mathrm{Var}\,(S_n) = npq$ 。因此

$$
\begin{aligned}
& P(|\frac{S_n}{n} - p| > \varepsilon) = P(|S_n - np| > n\varepsilon) \\
= & \sum_{k:|k-np|>n\varepsilon} P(S_n = k) \\
\leq & \frac{1}{n^2\varepsilon^2} \sum_{k:|k-np|>n\varepsilon} (k-np)^2 \cdot P(S_n = k) \\
= & \frac{1}{n^2\varepsilon^2} \mathrm{Var}\,(S_n) = \frac{pq}{n\varepsilon^2} \to 0,\ \text{當}\ n \to \infty
\end{aligned}
$$

上述證明中的不等式就是所謂的 Chebyshev 不等式。比較起來, Bernoulli的證法美妙是美妙, 但只適用於二項分佈, 而無法推展到一般的分佈; Chebyshev 不等式的證法, 不但簡單明快, 而且可以推展到一般分佈的情形。A. Weil 說:「更一般與更簡潔是結伴同行的」(Greater generality and greater simplicity go hand in hand.) 這裡是一個例證。

經濟學家 Keynes 在1921年出版的「A treatise on probability」一書裡稱讚 Berno- ulli 爲「數學機率論的眞正創立者」(The real founder of mathematical probability)。 俄國在1913年於聖彼得堡 (St. Petersburg, 後來改爲列寧格勒, 現在又改回原名) 慶祝 Bernoulli 定理誕生200周年紀念, 由 J.V. Uspensky 將 Bernoulli 的原著: Ars conjectandi(1713) (即「猜測的藝術」, The art of conjecturing) 從拉丁文翻譯成俄文。一條重要的數學定理比人間任何帝國的基業還要天長地久, 今日蘇聯已崩解! Einstein 說:「政治是短暫的, 方程式是永恆的。」

Bernoulli稱他所發現的定理爲「Golden Theorem」(黃金定理), 可見其珍惜與寶貴。利用這個定理來解釋「丟一個公正銅板出現正面的機率爲1/2」就是: 對任意 $\varepsilon > 0$, 丟 n 次銅板出現正面的相對頻率 $\frac{S_n}{n}$ 會落在1/2左右 ε 範圍內的機率 , 當 $n \to \infty$ 時, 會趨近於1。這才是機率的正解之一。Bernoulli 認爲機率代表一個事件的「明確度」(degree of certainty), 這恰是日文的「確率」之來源。

Bernoulli說: 「經過了20多年的深思, 才得到這個定理, 我決定把它發表。它不但新奇、具有大用, 而且難度高。它的份量與價值超過全書其它所有章節。它對我深具意義, 遠比方圓問題 (幾何三大難題之一) 的解決更有意義, 因爲即使解決方圓問題, 也是沒有什麼用處的」。

他又說: 「對某件事作猜測就相當於度量它的機率。因此猜測的藝術或隨機的藝術可以定義爲儘可能準確地度量事件的機率, 使得我們作判斷與行動的時候, 可以採取較佳的策略。讓我們更深思熟慮, 而得到較安全與滿意的結果。只有在此中, 我們才見出哲學家的智慧與政治家的謀略。即使是最愚笨的人, 根據良知良能也知道, 觀測越多, 偏離眞理的危險越小。如果對任何事件都可作永恆長觀, 那麼機率就化成明確, 並且我們可以發現世界上所有事情皆按一定的原因與規律來發生。」

Belnoulli定理標誌著機運 (chance) 初步被馴服, 這個意義非凡, 因爲如此善變而說不準的機運都可被馴服, 還有什麼不可馴服的呢? 更重要的是它打開了機運的大門,

讓其後的數學家發現了美麗的機率天地 (花園)。

二.De Moiver - Laplace 的中央極限定理

對於公正銅板的情形, Bernoulli 定理可以改述為: 對任意 $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(S_n \in [\frac{n}{2}-\varepsilon n, \frac{n}{2}+\varepsilon n]\right) = 1$$

這表示以 $n/2$ 為中心, 左右 εn 的區間 $[\frac{n}{2}-\varepsilon n\ ,\ \frac{n}{2}+\varepsilon n]$, 在 $n\to\infty$ 時, 可以捕捉到 S_n 散佈的所有機率1。由於 Bernoulli 所用的區間是以 n 的等級來伸展, 非常快速, 故只要作粗略的估計就可證得它網到所有的機率。

二項分佈最大項的等級 (漸近相等式) 為 $\frac{1}{\sqrt{\pi n}}$, 因此以 n 為等級的區間可以捕捉到所有的機率, 這是順理成章的事。De Moivre (1667-1754) 的偉大貢獻是將 n 改為小一點的 $\sqrt{n}$, 即用區間 $[\frac{n}{2}-\varepsilon\sqrt{n}\ ,\ \frac{n}{2}+\varepsilon\sqrt{n}]$ 來捕捉 S_n 所散佈的機率。這「看似尋常」, 其實是「最奇絕」, De Moivre 計算出了, 在 $n\to\infty$ 時, 此區間所網到的機率。他說:「這是機運所能提出的最艱難的問題。」因此 De Moivre 的工作真的是「成如容易卻艱辛。」他採用 $\sqrt{n}$ 而不採用 $n^{1/3}$ 或 $n^{2/3}$ 等, 這是關鍵要害。

De Moivre 做的是 $p=\frac{1}{2}$ 的特例, 即公正銅板, 後來 Laplace 考慮一般 p 的情形, $0<p<1$ 。下面我們要帶著現代的眼光重走一趟 De Moivre 與 Laplace 的探險之旅。

先從公正銅板的特例 $(p=\frac{1}{2})$ 思考起。此時 S_n 的期望值與變異數分別為

$$E(S_n)=\frac{n}{2} \text{ 與 } \mathrm{Var}\,(S_n)=\frac{n}{4},$$

而標準偏差為

$$\sqrt{\mathrm{Var}\,(S_n)}=\frac{\sqrt{n}}{2}$$

我們注意到, 事件

$$\frac{n}{2}-\varepsilon\sqrt{n}\le S_n\le\frac{n}{2}+\varepsilon\sqrt{n}$$

與事件

$$-2\varepsilon\le\frac{S_n-\frac{n}{2}}{\sqrt{n}/2}\le 2\varepsilon$$

是相同的。令 $S_n^*=\frac{S_n-\frac{n}{2}}{\sqrt{n}/2}$, 其中將 S_n 減去期望值 $\frac{n}{2}$ 的操作叫做無偏化 (即平移, 此時期望值變成0, 但變異數不變), 再除以標準偏差 $\frac{\sqrt{n}}{2}$ 得到 S_n^* 叫做規範化 (即尺度伸縮, 此時 S_n^* 的期望值為0, 變異數為1)。將一個隨機變數作無偏化再作規範化的操作叫做標準化。換言之, 我們將 S_n 標準化為 S_n^* 。

我們的偉大目標就是要計算

$$\lim_{n\to\infty} P(-2\varepsilon\le S_n^*\le 2\varepsilon)=?$$

或更一般的

$$\lim_{n\to\infty} P(a\le S_n^*\le b)=?$$

為此, 我們觀察下圖:

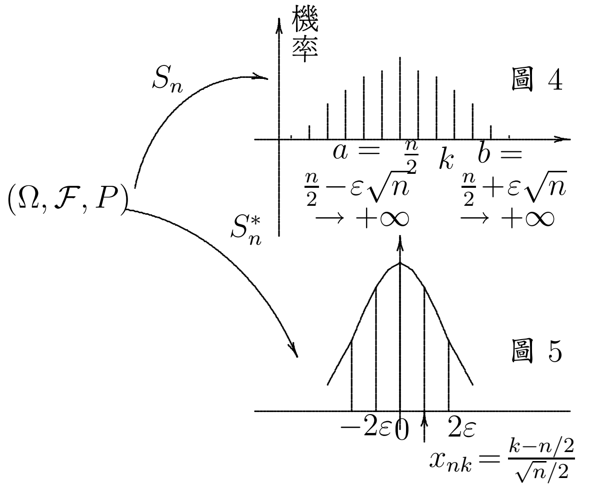


圖 4

圖 5

在圖4中, 區間 $[\frac{n}{2}-\varepsilon\sqrt{n},\frac{n}{2}+\varepsilon\sqrt{n}]$ 按 $\sqrt{n}$ 的等級來張開以捕捉 S_n 所散佈的機率。此時不易感覺到終究 $(n\to\infty)$ 捕捉到多少機率以及這個機率有何規律可尋。但是我們改觀察圖5時, 情形就改變了。令 $x_{nk}=\frac{k-\frac{n}{2}}{\sqrt{n/2}}$, 當 n 越來越大時, S_n^* 在固定區間 $[-2\varepsilon,2\varepsilon]$ 上所下的「機率雨」, 其間隔 $\Delta X_{nk}=\frac{1}{\sqrt{n/2}}$ 漸小, 即機率雨越下越細密, 而「雨點」越來越小。有如「大雨稀疏, 細雨密密。」我們簡直可以感覺得到, 在 $n\to\infty$ 時, 機率雨的減少與增加終究會達於適度的平衡: 機率不全落在 $[-2\varepsilon,2\varepsilon]$ 之中, 也不完全流失掉。有了這麼美妙的感覺, $[-2\varepsilon,2\varepsilon]$ 終究捕捉到多少機率, 其計算我們相信應該是有規律可尋的。

我們再用著名的 Galton 盤 (或叫機率盤) 來模擬 S_n^* 所下的「機率雨」之分佈情形 (參見圖 6)。從漏斗中輸入許多小球進到直立的盤箱中。盤中的黑點表示釘子, 球在落下的過程中, 會碰到釘子 (或球與球之間相撞) 而偏向左邊或右邊。在底部有許多相同大小的隔間, 用來收集落球。中間的地方收集得較多, 越向兩側收集得越少。從這個實驗中, 也反應出「機率雨」似乎是按某個美妙的規律來降落的。

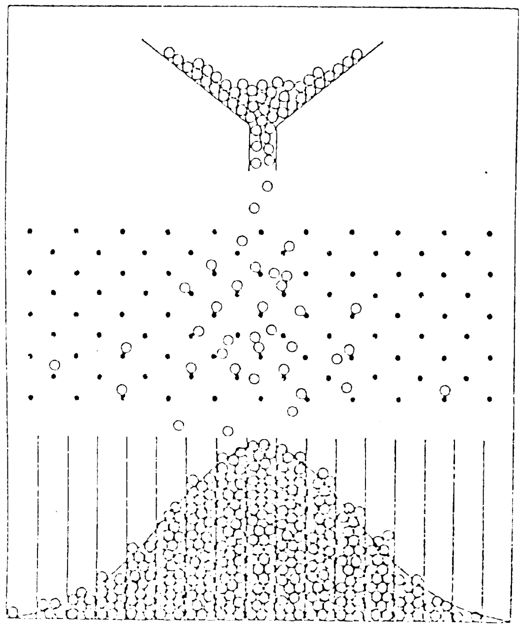


圖 6

有了規律感 (sense of order) 之後, 接著就是找尋規律。

現在開始做苦工。我們考慮一般 $p(0<p<1)$ 的情形。此時 $q=1-p$, 期望值 $E(S_n)=np$, 變異數 $\mathrm{Var}\,(S_n)=npq$, 標準偏差 $\sqrt{\mathrm{Var}\,(S_n)}=\sqrt{npq}$, 而 S_n 的標準化爲

$$S_n^*=\frac{S_n-np}{\sqrt{npq}}$$

我們要對任意有限區間 [a,b], 計算

$$\lim_{n\to\infty}P(a\le S_n^*\le b)$$

令 $X_{nk}=\dfrac{k-np}{\sqrt{npq}},\quad k=0,1,\ldots,n$ (15)

我們要考慮 n 越來越大 $(n\to\infty)$ ，但是 x_{nk} 滿足 $a\le x_{nk}\le b$。因此 k 亦隨 n 而變, 由 (15) 式知

$$k=np+\sqrt{npq}x_{nk} \text{ 且 } n-k=nq-\sqrt{npq}x_{nk} \tag{16}$$

所以當 $n\to\infty$ 時

$$k\sim np \text{ 且 } n-k\sim nq \tag{17}$$

其中記號 $a_n\sim b_n$ 表示 $\lim_{n\to\infty}\frac{a_n}{b_n}=1$, 即 (a_n) 與 (b_n) 漸近地相等。因此當 $n\to\infty$ 時, k 與 $n-k$ 也都趨近於無窮大。

爲了求算

$$P(a\le S_n^*\le b)=\sum_{k:a\le x_{nk}\le b}{}_nC_kp^kq^{n-k}$$

在 $n\to\infty$ 時之極限, 我們先找尋「局部」一項 ${}_nC_kp^kq^{n-k}$ 的漸近相等式。換言之, 我們要估算

$$\begin{aligned}P(S_n=k)=&\ P(S_n^*=x_{nk})\\ =&\ \tfrac{n!}{k!(n-k)!}p^kq^{n-k}\end{aligned}$$

由 Stirling 公式知

$$\begin{aligned}&\frac{n!}{k!(n-k)!}p^kq^{n-k}\\ =&\ \frac{\sqrt{2\pi n}n^ne^{-n}\quad p^kq^{n-k}}{\sqrt{2\pi k}k^ke^{-k}\sqrt{2\pi(n-k)}(n-k)^{n-k}e^{-(n-k)}}\\ =&\ \frac{1}{\sqrt{2\pi}}\frac{\sqrt{n}}{\sqrt{k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\\ \sim&\ \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\\ &\text{(由(17)式)}\end{aligned} \tag{18}$$

接著探尋 $\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}$ 之漸近相等式。由 (16) 式知

$$\frac{np}{k}=1-\frac{\sqrt{npq}}{k}x_{nk},\ \frac{nq}{n-k}=1+\frac{\sqrt{npq}}{n-k}x_{nk} \tag{19}$$

其次, 由 Taylor 展式

$$\begin{aligned}\ln(1+x)\ =&\ x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots\\ &+(-1)^{n+1}\frac{x^n}{n}+\cdots,\quad |x|<1,\end{aligned} \tag{20}$$

可知, 對任意 $0<A<1$, 當 $|x|<A$ 時, 恆有

$$\begin{aligned}&\left|\ln(1+x)-(x-\frac{x^2}{2})\right|\\ \le&\ \sum_{n=3}^{\infty}\frac{|x|^n}{n}\le\frac{1}{3}\sum_{n=3}^{\infty}|x|^n\\ =&\ \frac{1}{3}\frac{|x|^3}{1-|x|}\le\frac{1}{3}\frac{|x|^3}{1-A}\\ =&\ \frac{1}{3}K\cdot|x|^3\end{aligned} \tag{21}$$

其中 $K=\frac{1}{1-A}$, 並且

$$\begin{aligned}&\ln\left(\frac{np}{k}\right)^k=k\ln\left(1-\frac{\sqrt{npq}}{k}x_{nk}\right)\\ =&\ k\left(-\frac{\sqrt{npq}}{k}x_{nk}-\frac{npq}{2k^2}x_{nk}^2\right.\\ &\left.-\frac{(npq)^{3/2}}{3k^3}x_{nk}^3-\cdots\right)\end{aligned} \tag{22}$$

$$\begin{aligned}&\ln\left(\frac{nq}{n-k}\right)^{n-k}\\ =&\ (n-k)\ln\left(1+\frac{\sqrt{npq}}{n-k}x_{nk}\right)\\ =&\ (n-k)\left(\frac{\sqrt{npq}}{n-k}x_{nk}-\frac{npq}{2(n-k)^2}x_{nk}^2\right.\end{aligned}$$

$$+\frac{(npq)^{3/2}}{3(n-k)^3}x_{nk}^3-\cdots\Big) \tag{23}$$

注意到, 由 (17) 式及 $a \le x_{nk} \le b$ 可知, 當 n 夠大時, 可以使

$$\left|\frac{\sqrt{npq}}{k}x_{nk}\right|<1 \text{ 與 } \left|\frac{\sqrt{npq}}{n-k}x_{nk}\right|<1。$$

因此(22) 與 (23) 兩式之收斂不成問題。

(22)+(23) 得到

$$\begin{aligned}
&\ln\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\right]\\
=\;&\frac{-n^2pq}{2k(n-k)}x_{nk}^2\\
&+\left[\left(\frac{(npq)^{3/2}}{3(n-k)^2}\right.\right.\\
&\left.\left.-\frac{(npq)^{3/2}}{3k^2}\right)x_{nk}^3+\cdots\right]
\end{aligned} \tag{24}$$

令 $R(x)=\left(\frac{(npq)^{3/2}}{3(n-k)^2}-\frac{(npq)^{3/2}}{3k^2}\right)x_{nk}^3+\cdots$。只要 n 夠大就有

$$\begin{aligned}
&\left|\frac{\sqrt{npq}}{k}x_{nk}\right|<A<1\\
\text{且}\quad&\left|\frac{\sqrt{npq}}{n-k}x_{nk}\right|<A<1,
\end{aligned}$$

於是由 (21) 式得到

$$\begin{aligned}
|R(x)| &\le \frac{1}{3}K\left(\frac{(npq)^{3/2}}{(n-k)^2}+\frac{(npq)^{3/2}}{k^2}\right)|x_{nk}|^3\\
&\le \frac{1}{3}K\left(\frac{n^{3/2}}{(n-k)^2}+\frac{n^{3/2}}{k^2}\right)\cdot M^3
\end{aligned}$$

其中我們用到了$pq<1$且$M=\max(|a|,|b|)$, 再由 (17) 式得知, 當 $n\to\infty$ 時, 上式右項趨近於0。從而 (24) 式可以改寫成

$$\begin{aligned}
&\ln\left[\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\right]\\
\sim\;&\frac{-n^2pq}{2k(n-k)}x_{nk}^2\\
\sim\;&-\frac{1}{2}x_{nk}^2, (\text{由}(17)\text{式}),
\end{aligned} \tag{25}$$

亦即

$$\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}\sim e^{-\frac{1}{2}x_{nk}^2} \tag{26}$$

代回 (18) 式就得到

$$\begin{aligned}
P(S_n=k) &= P(S_n^*=x_{nk})\\
&\sim \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}e^{-\frac{1}{2}x_{nk}^2}
\end{aligned} \tag{27}$$

這眞是一個艱苦而漂亮的估計過程。上式對於使得 x_{nk} 落在 $[a,b]$ 中的所有 k 皆成立。換言之, (27) 式對 k 是均勻收斂的。

進一步, 我們注意到 $\Delta x_{nk}=\frac{1}{\sqrt{npq}}$, 故 (27) 式可改寫成

$$P(S_n^*=x_{nk})\sim\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x_{nk}^2}\cdot\Delta x_{nk}$$

從而

$$\begin{aligned}
P(a\le S_n^*\le b) &= \sum_{k:a\le x_{nk}\le b}P(S_n^*=x_{nk})\\
&\sim \frac{1}{\sqrt{2\pi}}\sum_{k:a\le x_{nk}\le b}e^{-x_{nk}^2/2}\cdot\Delta x_{nk}
\end{aligned}$$

這是定積分 $\int_a^b\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$ 的 Riemann 和, 故

$$\lim_{n\to\infty}P(a\le S_n^*\le b)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx \tag{28}$$

這樣我們就發現了 S_n^* 在任何有限區間 $[a,b]$ 中落下「機率雨」在 $n\to\infty$ 時之計算

規則。更令人驚奇的是這個規則可由世界上最重要的一條曲線

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2} \tag{29}$$

來給出, 它叫做標準常態分佈(standard normal distribution)。

總結上述, 我們得到一個美妙而重要的定理:

定理2 (De Moivre - Laplace 的中央極限定理): 設 $0<p<1, q=1-p$, 且 $S_n^*=\frac{S_n-np}{\sqrt{npq}}$,

$$x_{nk}=\frac{k-np}{\sqrt{nqp}},\quad k=0,1,\cdots,n。$$

(i)De Moivre 的局部極限定理 (1733年):

當 $n\to\infty$ 時, 對於使得 x_{nk} 落在任何有限區間 $[a,b]$ 中的 k 而言,

$$P(S_n^*=x_{nk})\sim\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}e^{-x_{nk}^2/2}$$

並且對 k 爲均勻收斂。

(ii)Laplace 的積分極限定理 (1812年):

$$\lim_{n\to\infty}P(a\le S_n^*\le b)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx$$

標準常態分佈 (又叫標準正規分佈), 它的圖形如下:

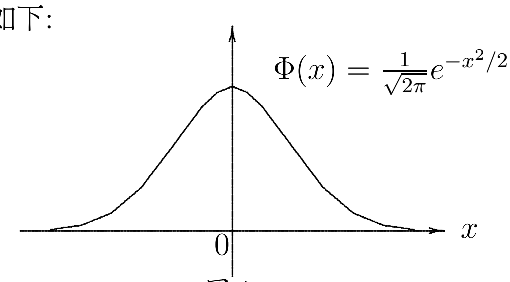
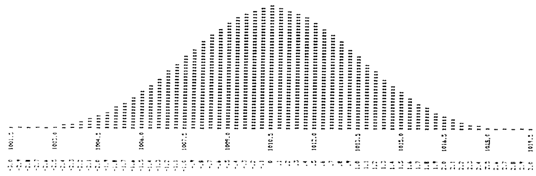


圖 7

我們也可以利用電腦作出更美麗的圖:



對於這條神奇而迷人的曲線, Laplace (1749-1827) 曾比喻說: 「正規分佈律就好像是天體力學中的萬有引力定律一樣, 佔有核心的地位。只要是由無窮多個無窮小的獨立隨機因子所湊成的一個隨機變量就會遵循正規分佈律。」一個有趣的對照: 積分 $\int_a^b f(x)dx$ 可以看成是由無窮多個無窮小的矩形之連續求和。「玉山不辭細土, 故能成其壯美。」

在上述中央極限定理中

$$\lim_{n\to\infty}P(a\le S_n^*\le b)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx \tag{30}$$

這個收斂很特別, 並不是隨機變數 S_n^* 的機率密度函數收斂到 $\Phi(x)$ 。事實上, 它是分析學中更根本而奧妙的"測度弱收斂"概念, 上式是這個概念的胚芽 (germ)。

如果當初我們用

$$\xi=\begin{cases}+1 & \text{正面, 機率 } \frac{1}{2}\\ -1 & \text{反面, 機率 } \frac{1}{2}\end{cases}$$

來表現丟一個公式銅板的實驗, 那麼 $E(S_n)=0, \mathrm{Var}(S_n)=n$, 於是 $S_n^*=\frac{S_n}{\sqrt{n}}$, 從而(30) 式變成

$$\lim_{n\to\infty}P\left(a\le\frac{S_n}{\sqrt{n}}\le b\right)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}dx \tag{31}$$

因此中央極限定理有時又叫做" n 的平方根定律"(the law of square root n)。

中央極限定理是一個典型例子, 在計算式中同時出現 π 、e 這兩個經常涉及宇宙奧密的偉大常數。有人問 H. Steimhaus(1887 - 1972) 說, 為何在機率論中到處會出現 π ? 他回答得很妙: 「機運是風水輪流轉的」(Fortune moves in circle)。如果是物理學家 R. Feynman, 他會問:「圓在哪裡? 」(Where is the circle?) 我們可以回答說:「 π 出自 Stirling 公式, Stirling 公式又出自 Wallis 公式, 而 Wallis 公式源自 Wallis 之求圓的面積。」

中央極限定理比 Bernoulli 的估計更精細, 因此可用來重新推導出 Bernoulli 定理, 這是順理成章的事:

$$\begin{aligned}
&\lim_{n\to\infty} P\left(\left|\frac{S_n}{n}-p\right|\leq\varepsilon\right)\\
=&\lim_{n\to\infty} P\left(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\leq S_n^*\leq\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)\\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-x^2/2}dx=1
\end{aligned}$$

此地我們用到了機率積分

$$\int_{-\infty}^{\infty}e^{-x^2/2}dx=\sqrt{2\pi} \qquad (32)$$

這是微積分的一個好習題。因此, 中央極限定理是兩面刃, 一面砍下 Bernoulli 定理, 另一面砍中了機率論中最重要的正規分佈, 將理想的 p 跟現實的 $\frac{S_n}{n}$ 連繫得更緊密、更好。

正規分佈的出現, 讓人首次深刻感受到, 在機運的混沌中出現了秩序 (order out of chaos), 如蓮之出污泥。這比自古以來從自然現象中看出規律 (laws), 更令人震撼。從此, 人們有了“機率法則” (probability law or the law of chance) 的概念, 這是機率要成為一門數學的先決條件。英國數學家兼哲學家 A.N. Whitehead(1861-1947, B. Russell 在劍橋大學三一學院的老師及合作者) 說得好 (參見 [6]): 「活生生的科學是無法產生的, 除非人們對於事物存在有規律, 特別地, 自然存在有規律具有普遍而近乎直覺的信仰。」(There can be no living science unless there is a widespread instinctive conviction in the existence of an Order of Things, and, in particular, of an Order of Nature.) 英國 Bedford 地方有一個中學, 校園內的標語寫著「自然含妙理, 等待著你去追尋。」劍橋大學 Cavendish 實驗室的座右銘更標舉出: 「造物者的工作是偉大的, 追尋出他的所有傑作令人樂在其中。」(The works of the Lord are great, sought out of all them that have pleasure there in.) 這些都非常振奮、鼓舞人心。

如何找尋、猜測到規律? 這是求知、思考活動的核心, 較深刻 (profound) 而有趣。當我們猜測到規律後, 要加以證明或否證 (proof or refutation), 通常是平凡 (ordinary) 而順理成章的。注意到, 只有數學才有證明! 偉大數學家 Poincaré(1854-1912) 說:「我們利用邏輯來證明, 但是我們透過直覺來發明。」(It is by logic we prove, but by intuition that we invent.) Leibniz 把規律與秩序看成是自然的一種先定和諧 (a pre-estabilished harmony)。「欲窺探

此先定和諧的願望, 是支持科學家工作的永不止息的耐力與恆心之泉源」, Einstein如是道出他的寶貴經驗。

三. 機率的各種解釋

什麼是機率? 長久以來數學家與哲學家爭論不休。各門各派的人對機率提出了各種不同的解釋, 這恰好反應出機率概念的深刻, 滑溜而不易把捉。比美於歐氏幾何中的點: 什麼是點? 點有多大?

每個人對機率都有一些常識概念, 但是細思起來又常會出現漏洞或矛盾。正如 St. Augustine 所說的: 「什麼是時間? 如果沒有人問我, 我是淸楚明白的; 如果我要解釋給問我的人聽, 我就迷惑了。」

機率論具有廣泛而豐富的應用: 在研究大自然方面,「自然的眞正理路在於機率的演算」(The true logic of Nature is in the Calculus of Probability. J.C. Maxwell 之名言); 在日常事物方面,「機率更是生活的指南」(Probability is the very guide of life. J. Butler 之名言。) 爲了應用機率, 就迫切需要給予解釋, 以便跟直觀常識連結起來。歷來對機率的解釋不下十餘種, 我們只選擇五種重要而有代表性者來介紹。

甲. 機會均等說 (Equiprobability theory)

這一派又叫做古典的觀點或 Laplace 的觀點, 主要的提倡者有 Pascal、Fermat、D'Alembert、Laplace 等人。他們首次嘗試給機率定量化以及一些基本的演算規則。我們分成四項來敍述:

(I) **基本假設**

對於一個隨機實驗的所有可能出現結果, 在沒有理由來預期或偏好某一結果時, 每一個結果的地位都相同, 皆應視爲機會均等 (equi- probable)。在文獻上, 這通常叫做不充足理由原理 (Principle of insufficient reason) 或無差別原理 (Principle of indifference) 或對稱性原理。

(II) **機率的定義與演算規則**

我們採用現代集合記號。設 Ω 表示一個隨機實驗的所有可能結果, 叫做樣本空間 (Sample space)。$A \subset \Omega$ 表示一個事件 (event), $\#A$ 表示集合 A 的元素個數。

(1) 當 $\#\Omega < \infty$ 時, 我們定義事件 A 的機率爲

$$P(A) = \frac{\#A}{\#\Omega} \tag{33}$$

(2) 當 $\#\Omega = \infty$ 時, 若 Ω 可表現爲幾何空間的領域, 則定義

$$P(A) = \frac{m(A)}{m(\Omega)} \tag{34}$$

其中$m(A)$ 表示 A 的長度、面積或體積等等。(34) 式就是所謂的幾何機率 (geometric probability)。

至於機率的演算規則 (Calculus of probabilities), 這不外是四則運算 $(+, -, \times, \div)$:

(i) **加法**: 若A,B 兩事件互斥, 即 $A \cap B = \phi$

, 則 A 或 B 發生的機率爲

$$P(A\cup B)=P(A)+P(B) \tag{35}$$

(ii) **減法**: 若 $A^c=\Omega\setminus A$ 表示 A 的補事件, 則

$$P(A^c)=1-P(A) \tag{36}$$

(iii) **乘法**: 若A,B 互相獨立, 則 A 與 B 同時發生的機率爲

$$P(A\cap B)=P(A)\cdot P(B) \tag{37}$$

(iv) **除法**: 若 $P(B)>0$, 則在 B 發生下重估 A 發生的機率爲

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \tag{38}$$

這叫做 A 對 B 的條件機率 (conditional probability)。

(v) 0或1之機率: 不可能事件 ϕ 的機率爲0, 鐵定事件 Ω 的機率爲1

$$P(\phi)=0, P(\Omega)=1 \tag{39}$$

(III) **優點**

在機會均等的假設下所定義的機率, 其計算只是排列與組合之點算。因此它的優點有

(1) 直觀易懂、簡潔、實用。我們對機率的直觀常識皆起源於此。

(2) 適用於日常生活的賭局 (games of chance) 以及一些科學的領域。今日初等的「點算機率」仍然通行。

(IV) **缺點**

(1) 當 $\#\Omega=\infty$ 時, (33) 式的定義失效。

(2) 無法作爲統計學的基礎, 因爲統計學時時會涉及 $\#\Omega=\infty$ 的情形。

(3) 對於有些情況, 不易確認「機會均等」所要施展的樣本空間, 因而容易產生一些詭論 (paradoxes), 例如著名的 Bertrand 詭論: 在單位圓的內部任意劃一弦, 這弦的長度大於圓內接正三角形邊長的機率可以是 1/2,1/3與 1/4。又如在統計力學中, 我們考慮 r 個粒子置於 n 個袋子的問題, 其樣本空間有各種情形, 如下表:

	球可分辨	球不可分辨
袋子可重複置球	n^r Maxwell-Boltzmann統計	${}_nH_r={}_{n+r-1}C_r$ Bose-Einstein 統計
袋子不可重複置球	${}_nP_r,\ (r\le n)$	${}_nC_r,\ (r\le n)$ Fermi-Dirac統計

因此要對於哪個樣本空間施展「機會均等」的假設, 並不是自明的。理論與經驗之間的關係往往很微妙而複雜。

(4) 容易對於機率產生誤解, 例如 D'Al- embert 就說:「一件事情 A 只有發生與不發生, 故 $P(A)=\frac{1}{2}$ 。」又如將丟一個公正銅板出現正面的機率是1/2解釋爲丟 2次就出現1次正面等等。

古典的機率理論, 首次嘗試給機率建立嚴格的數學基礎。雖然不完美, 有許多缺點, 但是平心而論, 若沒有它, 就產生不出往後的更豐富而美麗的新理論。正如 Galileo 是在批判 (破)Aristotle 物理學的錯誤之下, 才逐漸立下新物理學。

乙. 公理化的觀點 (Axiomatic approach)

從純數學的眼光來看, 自從 1933 年 Kolmogorov 提出機率論的公理系統 (Hilbert在1900年提出著名的23個問題, 其中第6個問題就是機率論的公理化問題), 對一個隨機現象採用機率空間 $(\Omega,\mathcal{F},P)$ 的理型來描述, 那麼機率只不過是一個可列加的 (σ -additive) 測度 P , 總測度爲 $P(\Omega)=1$, 或者乾脆看作是公理系統中的一個無定義名詞 (undefined term)。這種不解釋機率是什麼, 只管機率的演算規則的辦法, 不失爲是一個高招, 相當於「無言」或「留白」的妙用, 「天何言哉, 四時行焉, 百物生焉。」比較起來, 歐氏 (Euclid) 反而是有點畫蛇添足, 他定義點只佔有位置而沒有大小, 但從來沒有用到它。因爲有困難: 點沒有長度, 如何累積出有長度的線段?

機率空間的理型是經過長期的試誤 (trial and error), 演化才創造出來的。在機率空間的模型下, 我們可以推導出大數法則, 用來連結一個事件的機率與可經驗的相對頻率。這是機率的頻率說的發源地。

數學之公理的、形式的終究本質 (ultimate reality) 由 Hilbert 完全揭露, 他發現在歐氏幾何中將點、線、面改爲椅子、桌子、茶杯並不影響整個體系。機率論當然也不例外。這種由公理出發的演繹體系, 完全是空架性的, 只是一種傳遞眞值 (truth value) 的邏輯網路: 如果由源頭輸入的是眞值, 那麼眞理就會流佈於整個系統。一切都非常淸澈與明確, 自我完足。但是, 數學所付出的代價就是羅素 (B.Russell) 所說的一句嚴肅的俏皮話: 「數學是這樣的一門學問, 我們永遠不知所云, 也不知道我們所說的是否爲眞。」(Mathematics is the subject in which we never know what we are talking about nor what we are talking about is true.) 因爲要對最初輸入的東西 (即公理) 作解釋, 跟自然世界連結, 並且確定其爲眞, 這些並不是數學份內的事, 而是應用數學 (包括理論物理學) 的工作。從哲學知識論的觀點來看, 這個工作並不容易, 是產生爭論與不確定之源。愛因斯坦 (Einstein) 說得好: 「當數學定律指涉到實體世界時, 它們就不是確定的; 當它們是確定時, 它們就不指涉到實體世界。」(As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.) 追求確定不移的知識是人類的渴望, 但是擁抱自然卻要付出得不到完全確定的代價。由於科學理論之「天地不全」, 故需不斷破舊立新, 而產生科學革命。

將機率視爲公理系統中的無定義名詞是公理化機率論的優點。如果要問什麼是機率, 那麼就透過大數法則, 由相對頻率來偷窺一下機率的影子。這可以說是大多數數學家對機率所採取的態度。

丙. 頻率說 (Frequency theory)

這又叫做機率的統計觀點或經驗觀點, 主要的提倡者有 Venn(1886), Von Mises(1928), Reichenbach(1935) 等人。

這一派的主要論點: 機率是自然世界客觀的一個屬性, 像質量、距離或其它物理一樣, 必須透過觀測與度量來揭露。機率由自然世界的特性唯一決定, 只能後驗地 (a posteriori) 得知; 在相同狀況下作大量的觀測, 經由相對頻率來反應。對一個事件作觀測, 當觀測越長期, 其次數越大時, 相對頻率漸趨穩定, 其極限值就是該事件的機率。機率論是描述自然的一部分, 或至少是物理理論架構的一部分。

頻率說的優點是注重經驗, 適合於壽險精算與統計學的用途。但是它的主要缺點有三:

(1) 對於不能重複作實驗觀測的事件無法賦予機率。不過我們仍然可以說: 明天下雨的機率是 $1/3$, 民主黨大選獲勝的機率是 $5/9$, 某生考上大學的機率是 $4/7$, 而頻率說無能爲力。

(2) 人生有涯, 我們的經驗是有窮的, 不是無窮的, 我們只能作有窮多次而無法作無窮多次的實驗觀測。因此純靠觀測經驗無法掌握機率。

(3) 相對頻率的極限值不見得存在。

丁. 邏輯關係說 (Logical relation theory)

這一派又叫做機率的先驗說 (A priori theory of probability), 主要的提倡者有 Key- nes(1921), Jeffreys(1948), Koopman(1940) 以及 Carnap(1950) 等人。

他們的主要論點如下: 機率是先驗決定的, 而不是靠純經驗得知的。機率是從定命事件 (或敍述) 的恆眞或恆假 (機率爲 1 或 0) 延拓成有時眞有時假的事件, 即在恆眞與恆假之間作連續插值。機率是敍述之間的邏輯關係, 永遠相對於給定的資訊來說的。因此, 機率就是度量給定資訊 (證據、知識) 對一個事件的邏輯支持程度。當給定的資訊改變時, 相應地一個事件的機率也跟著改變。例如, 當我們知道某人是英國人, 其餘一無所知, 在這個證據下, 他很可能右眼是藍色的, 但是當我們進一步知道他的左眼是藍色時, 原先的機率就完全改變。這使得歸納邏輯 (inductive logic) 以及從經驗 (或錯誤) 中學習變成可能。這一派企圖用機率手法來解決科學哲學中深刻的歸納問題。

戊. 機率的主觀說 (The suljectivistic theory of probability)

這一派的主要提倡者有 Ramsey (1931), De Finetti(1937) 以及 Savage(1954) 等人。他們的論點如下:

每一個人對任何事件都持有信仰度 (degree of belief), 機率就是一個人對一個事件信仰度的表白。欲知一個人對一個事件所持的機率最好是考察他的行爲, 尤其是他在打賭行爲中願冒的風險。世界上並無客觀的機率, 或至少客觀機率是較不重要的一種機率概念。一個事件並無唯一的機率, 每一個人在邏輯上都有自由選取自己的機率。合理的信仰度必跟機率的演算規則一致。這一派爲 Bayes 觀點的統計決策論提供了基礎。

四. 結語

以上對機率的五種解釋, 基本上已含蓋了我們平常對機率概念的各層面之理解。「無尙妙趣, 了悟之樂。」

不過, 我們還是沒有完全抓住機率, 它仍然在「忽隱忽現」之中, 我們有點像是瞎子摸象。每一次我們加以捕捉, 起先以爲捉住了, 後來又發現給溜了。事實上, 機率就像命運、機運一樣, 也許永遠說不清, 因爲它們都生存在無窮遠處的理想彼岸, 有如鏡花水月, 可望不可即, 但卻形成一個永恆的引力中心, 引人入勝。

雖然我們對機率無法求得甚解。但這並不影響我們對它的數學研究。機率論就是要把機運可用數學說出的部分說清楚, 如此而已。

參考文獻

1. J.M Keynes: A treatise on probability, MacMillan And Co. Ltd London, 1921.
2. R. Weatherford: Philosophical Foundations of probability theory, Routledge & Kegan Paul Ltd, 1982.
3. A. Hald: A History of probability and statistics and their applications before 1750, John Wiley and Sons, 1990.
4. Y.G. Sinai: Probability theory, An Introductory Course, Springer-Verlag,1991.
5. 蔡聰明: 談 Stirling 公式, 數學傳播, 十七卷二期, 1993.
6. A.N. Whitehead: Science and the modern world, MacMillan, 1925.
7. S.M. Stigler: The History of Statistics, The Measurement of Uncertainty before 1900, Harvard Univ. press, 1986.

—本文作者任教於台灣大學數學系—