實驗概率和理論概率的關係
——中學數學教學如何受益於大學數學

蕭文強
香港大學數學系
黃毅英
香港中文大學課程與教學學系

楔子

這恐怕是常見的一堂概率課：老師著令每位學生擲毫 10 次，得出數百個數據後，發覺得正面（head）的概率與 $\frac{1}{2}$ 相去一段距離。老師說：這是因爲我們擲的次數不夠多，再多擲一些，實驗結果就會很貼近 $\frac{1}{2}$ 了。

另一些老師用電腦模擬擲骰，啓動了 Excel® 程式擲骰 100,000 次。很可惜，得出的結果亦與均勻分佈（uniform distribution）相去一段距離。老師也唯有說，擲的次數再大一點，擲出各點的概率就會趨於 $\frac{1}{6}$ 了。

這是真的嗎？

本文其中一位作者在一篇文章[1]裏亦談到這個問題。其實該問題是由另一位作者收到的一封電郵所引發。內容是這樣的：

> 今天我和我的碩士班同學討論功課的時候，談論到理論概率和實驗概率的關係。他們認為理論概率是理論層面，實驗概率是現實層面。但當我們重複做相關實驗無數次，我們得出的實驗概率的值便會趨向理論概率的值。如當我們重複做該實驗無限次時，實驗概率的值便相等於理論概率的值了。
>
> 但我記得你曾經解釋過，(希望我沒有記錯吧？！) 上述的論述是

1 蕭文強 (2007) 。回到未來——從大學講堂到中小學課室。《基礎教育學報》16 卷 1 期，97 – 114。

很錯誤的，因為縱使重複做相關實驗無限次，我們都不能把實驗概率和理論視為相同。我現在的問題是：

1. 我有沒有記錯你的解釋呢？

2. 如我真的沒有記錯，那我可如何解釋給他們明白，縱使做實驗無限次，兩者其實並不相同呢？

3. 他們認為以一個中三程度的學生，如他真的這樣理解實驗概率和理論概率的話，其實是可以接受的。這又是否真的可接受呢？

要更深入探討這個問題，我們首先要了解何謂實驗概率？何謂理論概率？兩者之間有何關係？

實驗概率和理論概率

顧名思義，實驗概率就是從大量實驗數據中總結出某個事件發生的可能性。譬如說，擲毫 100 次，有 48 次是正面，我們便說投擲該毫出現正面的實驗概率是 0.48；如果擲了 1000 次，有 532 次是正面，實驗概率便是 0.532。由此可見，實驗概率隨實驗次數多少會轉變，並沒有一個確定值。理論概率是按照已知的條件，根據合理的推論得出某個事件發生的可能性。譬如說，如果投擲一枚勻稱硬幣（fair coin），沒有理由相信出現正面或反面的機會不相同，因此投擲該毫出現正面的理論概率是 0.5。反過來說，如果硬幣並不勻稱，擲到正面的機會大於擲到反面的機會，那麼投擲該毫出現正面的理論概率便不再是 0.5，而是大於 0.5 了。至於它是什麼，以下我們再回去討論。

實驗概率趨於理論概率？

文章開首提到的說法：「實驗次數愈多，實驗概率會趨向於理論概率」，在不少教科書都以各種形式出現。有些說「實驗次數愈多，實驗概率便愈接近理論概率」，另一些說「當實驗次數足夠多時，實驗概率 ≈ 理論概率」。對於「趨向於」、「接近」或「≈」這些用語的精確意思，不同的作者或有不同的闡釋，但從字面上讀去，這樣的說法是有問題的。我們只可以說，實驗次數愈多，實驗概率**愈大可能**接近理論概率。

這種誤解，會否和課程文件有關呢？今天的中學數學課程以《1985 年中學數學科課程綱要》爲藍本。當時概率在初中和高中兩個階段分別出現。中四及中五「單元 8．概率和統計」之學習目的爲「加法和乘法定律」（頁 100）。而中三「單元 9．簡易概率的概念」之學習目的爲：(1) 明白「概率」的意義及欣賞它的應用價值；(2) 了解理論概率及實驗概率的分別。有關 (2) 的教學建議爲「進行實驗的目的在於驗算理論概率所求得的結果，而使學生同時了解到在現實生活中所接觸到的概率如失事率，犯罪率等都是實驗概率。」（頁 77）。其中提到以實驗去驗算理論概率，這個提法有待商榷。

至於現行的課程文件，並沒有這個說法。1999 年的《中學課程綱要——數學科》只提出「比較實驗概率和理論概率」。「學與教資源套」[2] 則有列出教學建議。其中第三學習階段「概率的簡單概念」單元中有「比較實驗概率和理論概率」一項，提出「學生可作出相同的試驗，討論，並比較試驗的結果，從而留意不同的試驗，通常會得出不同的概率。除了使用實物的學習活動外，也可用電腦或計算機來模擬大量的試驗，讓學生明白概率是相對頻數的趨向值。教師應指導學生明白，有些日常生活的事件，例如交通意外率、犯罪率等，都是實驗概率」（頁 5）。其中提到概率是相對頻數的趨向值，這個提法也有待商榷。

在 1960 年代「新數學運動」把概率引進中學數學課程，當時風行英國的 SMP (The School Mathematics Project) 課本便清楚提到「實驗概率」（experimental probability）和「預期概率」（expected probability）。新數學時期最有代表性的本地中學教科書應算是半群學社在 1966 年出版之《新數學》，書中用很豐富的隨機試驗開始，讓學生以圖表方式紀錄實驗結果，例如擲毫若干次，由此引入實驗概率。書中也提到很有意思的一點，指出只有用多次試驗的方法才可以發現一顆骰子是不是勻稱的。綜觀當時幾本教科書，由實驗概率作爲動機引入理論概率之進路是一致的，由活生生的實驗出發不失爲有效引起動機之方式，當然這和當年強調「動手」和「探索」的潮流有關。無論如何，這個進路亦符合「數學化」（mathematisation）[3] 之歷程。雖說學習進路比較自然，實驗概率與理論概率的關係仍是要澄清的。

---

2 教育署數學組（2002）。《中一至中五數學科學與教資源套 6：數據處理範疇（第三學習階段）》。香港：作者。

3 黃毅英（2007）。數學化過程與數學理解。《數學教育》25 期，2 – 18。

現時有些教科書的進路改為先定義理論概率為：

$$\text{概率} = \frac{\text{符合事件結果的數量}}{\text{可能結果的總量}} \quad \text{（假設各可能結果發生的機會相等）}$$

然後討論實驗概率，容易導致文章開首述及的混亂，以為實驗之目的就是為了驗算或者求出理論概率，隱含「實驗概率趨於理論概率」這個有待商榷的想法。

實驗概率和理論概率的關係

有些人或者記得有條「大數定律」（Law of Large Numbers），是雅各布·伯努利（Jacob Bernoulli，1654–1705）在十八世紀初提出來的，不就是說明理論概率可以從大量實驗數據中求出來嗎？讓我們看看「大數定律」怎樣說。其實「大數定律」共有兩條：「弱大數定律」和「強大數定律」；和我們的討論有較密切關係的是「弱大數定律」；「強大數定律」涉及的數學背景知識超越了中學的認知範疇，這兒就不提了。

「弱大數定律」是這樣說的：

> 若 X 是一隨機變量，μ 為 X 之平均值，$X_1, \ldots, X_n$ 為 X 之隨機抽樣，而 $\overline{X_n} = \dfrac{1}{n}(X_1 + \ldots + X_n)$，則對任意 $\varepsilon > 0$，
> 有 $\lim\limits_{n\to\infty} P(\left|\overline{X_n} - \mu\right| \geq \varepsilon) = 0$。

〔這一節用上不少基礎概率論的技術名詞，在一個入門的概率論大學課程必會讀到。而且，該定理的證明亦不是特別困難，看透了，不外是把各項技術名詞的定義弄清楚後借助一道以切比雪夫（P. L. Chebyshev，1821 – 1894）命名的不等式寫出結果吧[4]。〕

---

4 任何一本基礎概率論的大學課本都會詳細討論這些內容，例如 Larson, H. J. (1969/1974/1982). *Introduction to probability theory and statistical inference* (1st/2nd/3rd edition). New York: John Wiley。如果讀者不要求知悉詳細技術內容，只求一個概括的了解，不妨看看一本數普讀物：蕭文強、林建（1982）。《概率萬花筒》。香港：廣角鏡出版社。重印（2007）：香港統計學會。

這條定律並沒有說當 n 愈大，樣本平均值 $\overline{X_n}$ 愈接近 μ，只是說當 n 愈大，$\overline{X_n}$ 偏離 μ 的可能性愈接近零。也就是說，當 n 足夠大時，很少機會 $\overline{X_n}$ 偏離 μ。用更普及的語言，「大數定律」告訴我們，在大量實驗中某事件發生的相對頻率（就是實驗概率），**通常**是非常接近事件發生的理論概率。我們可以這樣理解，以擲毫爲例（假設得正反兩面的概率均等），擲毫這事件形成的隨機變量稱爲 X，$X = 0$（正面）或 $X = 1$（反面）。我們作出抽樣 $X_1, \ldots, X_n$，假如得出

0　1　1　0　0　0　1　0　1　0　0　1　1　1　1　0　⋯

當 n 愈來愈大時，得出不同的 $\overline{X_n}$，如下表所示：

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	⋯	⋯
X_n	0	1	1	0	0	0	1	0	1	0	0	1	1	1	1	0	⋯	⋯
$\overline{X_n}$	0	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{4}$	$\frac{2}{5}$	$\frac{2}{6}$	$\frac{3}{7}$	$\frac{3}{8}$	$\frac{4}{9}$	$\frac{4}{10}$	$\frac{4}{11}$	$\frac{5}{12}$	$\frac{6}{13}$	$\frac{7}{14}$	$\frac{8}{15}$	$\frac{8}{16}$	⋯	⋯

在這特別的情況，$\overline{X_n}$ 確是愈來愈接近 $\frac{1}{2}$，但不表示這是必然的（這就是與上述提及的誤解之最大分別）。確有可能出現「特異」的情況，例如

1　1　0　0　0　0　0　0　0　0　0　0　0　0　0　0　⋯ 或

1　1　1　1　1　1　0　1　1　1　1　1　1　1　1　1　⋯

只是出現這些情況的可能性是非常低。

理論概率怎麼求？

怎樣界定和尋求某個事件發生的理論概率呢？有一種想法，認爲重複實驗許多次，該事件的相對頻率應該是頗穩定的，不如就把無限次實驗的相對頻率的極限叫做該事件發生的（理論）概率。用數學語言說，設 n 次實驗中事件 A 發生 $M(n)$ 次（M 是 n 的函數），則事件 A 發生的概率 $P(A)$ 等於 $\lim_{n\to\infty}\frac{M(n)}{n}$。

「大數定律」似乎爲這種想法提供了一些理論註腳，但想深一層又十分別扭！如果根據「大數定律」去**界定**概率，定律的內容卻又涉及概率，

豈不像是說「在無限次實驗中事件 A 發生的相對頻率與 A 的概率相差是某個任意量這個事件的概率趨於零」？既像拗口令，也不容易明白！

這樣界定概率，還有幾點令人不安。一者，$M(n)$ 隨 n 而變，除了 $M(n)$ 是 n 的非遞減函數以外，$M(n)$ 與 n 可沒有什麼關係，並無理據支持 $\frac{M(n)}{n}$ 有極限。二者，沒有可能實驗無限次，如此界定只是理想情況而已。三者（此乃「切膚之痛」！）有不少事件並不能採用重複實驗的方法去界定概率，例如「明天下雨的概率是 0.7」，或者「有八成機會『陳九秒』在今年奧運奪取一百米金牌」。

數學家採用了另一個「釜底抽薪」的方法去界定概率，索性把概率看作是定義在樣本空間（sample space）S 的子集上的實數值函數 P，滿足以下的公理：

(P1) $P(S) = 1$；

(P2) 若 A 是 S 的子集，則 $0 \le P(A) \le 1$；

(P3) 若 A_1、A_2、... 是 S 的兩兩不相交的子集，則 $P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$。

這兒所謂樣本空間就是某項實驗可能發生的所有結果的集合，它的子集便叫做事件（event）。例如投擲一顆骰子，$S = \{1, 2, 3, 4, 5, 6\}$ 有六個結果，數字 k 代表 k 點。$A = \{1, 3, 5\}$ 代表擲到奇數點這個事件。

注意上述的界定倒不明言如何設定那個函數 P，只要求它滿足 (P1)、(P2)、(P3)。靠著這三條公理，加上知道某部份事件——稱爲基本事件——的概率，我們便能計算任何事件的概率了。就以 $S = \{1, 2, 3, 4, 5, 6\}$ 爲例，如果骰子是勻稱的，合理的假設是 $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$，因此 $P(\{1, 3, 5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = 3 \times \frac{1}{6} = 0.5$，即是說，擲到奇數點與擲到偶數點的機會是相同的。要是你擲骰 100 次，發現只有 12 次擲到奇數點，再擲 100 次，發現只有 8 次擲到奇數點，你便會懷疑那顆骰子是否勻稱了。

概率公理化與建模過程

從上一節公理化處理的角度看，實驗概率與理論概率是兩種不同的觀念，在某個意義而言，它們沒有直接關係。但它們卻有另一層面的微妙關連，讓我們在這一節談談。

不妨把公理化處理看成是一個建模過程（modeling process），即是把實際問題譯成數學語言，建立一個數學模型。在這個過程中我們把實際情形理想化及簡化，以便用適當的數學方法進行分析求解。由於我們把情形理想化及簡化，得到的答案可能只是近似答案，是否合用還得回到實際觀測作檢驗，逐步調整模型去求更好的答案[5]。「大數定律」是基於 (P1)、(P2)、(P3) 這些公理，運用邏輯推論得來的結果，在實際觀測亦的確如是，印證了概率建模符合實際情理。

理論概率是如何設定的呢？在一些情況，如投擲一枚勻稱硬幣，我們按常理設定理論概率；既曰勻稱，按定義，$P(\text{正面}) = P(\text{反面}) = \frac{1}{2}$。但在另一些情況，實驗可以是一種參考點。譬如一台機器會生產不合規格的產品，這牽涉到保險和賠償等問題，我們必須評估生產不合規格產品的可能性，於是我們對這台機器進行大量的測試。比方經過 10 萬次測試，得出生產不合規格產品的可能性是 $\frac{1}{10000}$，這當然是實驗概率，不是理論概率；而且可能測試愈多，機器越磨損，生產不合規格產品的機會愈大！不過，我們爲了評估風險建立數學模型時，採取這個實驗概率作爲理論概率，亦合乎常理。一旦我們作了這個設定，其他風險評估等計算，就可依循數學方式進行。我們有理由相信這台機器生產不合規格產品的理論概率是客觀存在的，只是我們未必有足夠的資料和技術找出來吧。

讓我們多看一個運用計算幫助我們增進理解的例子。面對一堆電子元件，一半產自甲廠，另一半產自乙廠，從中抽取一件，問這一件產自甲廠的機會是多少；如果單單知道這麼多，我們只能說機會是 $\frac{1}{2}$。如果我們再知道這一件是壞件，而且知道甲廠生產壞件的可能性是乙廠的三倍，你

5 蕭文強（1978）。《為甚麼要學習數學》。香港：學生時代出版社。第二版（1992）：香港新一代文化協會。增訂本（1995），台灣：九章出版社。第四章。

認爲那機會是否仍然是 $\frac{1}{2}$ 呢？我們多了資料，可以運用一道以貝斯（Thomas Bayes，1702 – 1761）命名的公式去計算，得到的概率是 $\frac{3}{4}$。貝斯公式用來計算某個事件的條件概率（conditional probability），是概率論的基本知識，讀者可以在課本找到[4]，這兒不贅，要點是這個答案是從概率模型中計算得來的。固然，我們可以再回到實驗觀測，假如實驗結果並不接近 $\frac{3}{4}$，我們便會懷疑那堆電子元件是不是真的一半產自甲廠，另一半產自乙廠呢？〔如果從實驗觀測得來的答案幾乎就是 $\frac{1}{2}$，你能否再用一次貝斯公式計算那堆電子元件有多少產自甲廠？有多少產自乙廠呢？〕

從實驗數據企圖獲取某些參數用以建立更好的數學模型，或者從部份抽樣估計全局情況，是概率論的重要應用。這種應用叫做統計推斷（statistical inference），踏入一個現今在預科應用數學課程範圍內，將來在新高中數學課程延伸部份的一個單元，是大部份學生較少機會認識的課題，那可是另一篇文章的題材了。

中學數學與大學數學

中小學數學只是制度上的分段，兩者是一個連續體，小學的數學經驗可以變成連綿「數學化」的基礎。故此「數學化」並不是指把多些「硬」數學放進教學內容，而是提出一條相當長的、由具體到抽象、由歸納到演繹之路。同理，中學與大學亦是一個連續體。我們其中一位作者早已提出，不單是中學數學的學習延續到大學數學，從大學數學的高觀點更往往能啓迪中小學數學的教學[6]。問題不僅在於技術性內容上，而是整個數學理念、概念系統和知識結構上。本文所討論的便是其中活生生的一例，其他事例可謂不勝枚舉[7]。只要教學雙方不單滿足於「做妥數學題」（應試）而要真正弄明白真相，大學數學知識往往能提供教學上之依據。固然，我們不是

---

6 Siu, M.K. (1981). School Algebra? University Algebra? *Mathematics Bulletin, 2*, 5-6.

7 其他例子如 0.9 是否等於 1 之類，可參考黃毅英（2006）。「老師，用『A 簿』還是用『B 簿』？」。《數學教育》23 期，27 – 36。

說應該由大學數學指導中小學數學教學，因爲我們仍要考慮學生因素，如智性水平與動機等等問題。

我們好像將「數學建模」、「公理化」、「數學化」混作一談，事實正恰恰如是。學生學習往往依循著數學發展上的類似軌跡 [8]，從具體到抽象、從直觀到正規 [5]。1989 年美國的《學校數學課程與評核標準》也提出這種「數學建模」正正是一種數學化過程 [9]。數學教育名家 Hans Freudenthal [10] 也說過「數學化」這個觀念不是由他發明的，在數學界流傳已久，不過有時用「抽象化」、「理想化」不同的詞彙表達這個意思吧了 [3]。由此可見這三個觀念的共通性，也可見大學數學往往能對學校數學（不局限於某一個階段）的教學給出有用的啓示。

一篇介紹「數學學養教師」的文章 [11] 引用了清代文學家袁枚說的話：「學如弓弩，才如箭鏃，識以領之，才能中鵠」。在教師而言，我們也許同時需要數學學科知識（mathematics subject knowledge）和教學知識（pedagogical knowledge），當中包括所謂教學內容知識（pedagogical content knowledge）的引領，再加上對學生這個「學習主體」的了解和教學技巧，教與學才會提升到另一台階。

作者電郵： mathsiu@hkucc.hku.hk
nywong@cuhk.edu.hk

---

8 Siu, F. K., & Siu, M. K. (1979). History of mathematics and its relation to mathematical education. *International Journal of Mathematics Education for Science and Technology, 10*(4), 561-567.

9 National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

10 Freudenthal, H. (1991). *Revisiting mathematics education (China lectures)*. Dordrecht: Kluwer Academic Publishers.

11 陳鳳潔、黃毅英、蕭文強（1994）。「教（學）無止境：數學學養教師的成長」。載林智中、韓考述、何萬貫、文綺芬、施敏文（編）。《香港課程改革：新時代的需要研討會論文集》（頁 53 – 56）。香港：香港中文大學課程與教學學系。後載黃毅英（2005）（編）。《迎接新世紀：重新檢視香港數學教育──蕭文強教授榮休文集》（頁 38 – 45）。香港：香港數學教育學會。