或然率與常識

蕭文強

法國數學家拉普拉斯（LAPLACE ,1749-1827）曾經說過：「或然率理論不外是常識而已，只不過以計算來表達詮釋。」本文打算環繞這點舉一些例子，對活潑課堂教學應能起作用吧？

首先讓我提出幾個日常生活看到的現象，讀者能解釋嗎？(1)拋擲一枚銅板，正反兩面出現的次數相若。(2)在一杯清水裏放一滴墨水，然後攪拌，墨水化開於清水中，清水變色。再攪拌，還是一杯有色的水，不會回復原狀。(3)沒有可能把一枚曲別針扳成一根完全筆直的鐵絲。(4)張三往飯館吃飯，看見生意興隆，便向老板說句「生意眞好」，老板卻回答說生意並不好，是湊巧張三來到才看見那麼多顧客吧。(5)李先生在校務大樓九樓財務部工作，不時要到二樓學生事務處送交文件，每次他候電梯時，總覺得不走運；要上去時，來的電梯往下走；要下去時，來的電梯卻往上走。(6)擲一顆骰子，通常擲六次才出現某定點數。(7)網球員發球，往往先來一記“猛球”，失了便補上一記“普通球”。

以下我們將試用數學計算解釋這些現象和另一些看來並沒有那麼明顯的現象，需要用到的數學便是或然率理論了。已經有不少書本和文章介紹或然率的基本知識，本文不再重覆，下面需要用到的也只是幾個最基本的公式，爲了讀者方便，綜述如下：

(1)若A和B是互不相容的事件（不能同時發生），則P（A或B）$=P(A)+P(B)$。$P(X)$是事件X發生的或然率，是一個在0和1之間的數，$P(X)=0$表示X不可能發生，$P(X)=1$表示X必然發生。

(2)若A和B是互相獨立的事件（兩兩不影響），則P（A與B）$=P(A)P(B)$。更普遍的公式是P（A與B）$=P(A)P(B\mid A)$，這兒的$P(B\mid A)$是在A已經發生的條件底下B發生的或然率。

(3)期望值的計算。

* * * * * * * * * * *

1. 網球發球問題

以r表“猛球”，s表“普通球”，p_x表發x種類球（x是r或s）時發球成功的或然率，q_x表發球成功後終於取勝的或然率，所以發x種類球而取勝的或然率是$w_x=p_xq_x$。可以假定$0<p_r<p_s<1$和$1>q_r>q_s>0$，因而w_r可以大於w_s，w_r也可以小於w_s。以xy表先發x後發y這個策略，w_{xy}表用這個策略取勝的或然率，便有$w_{xy}=w_x+($1−

p_x) $w_y = w_x + w_y - p_x w_y$。試比較四種策略的取勝或然率，即是 w_{rr}、w_{rs}、w_{sr}、w_{ss}。由於

$$w_{rr} - w_{rs} = (1 - p_r)(w_r - w_s)$$

$$w_{sr} - w_{ss} = (1 - p_s)(w_r - w_s)$$

$$w_{rs} - w_{ss} = w_r - w_s[1 - (p_s - p_r)]$$

$$w_{rr} - w_{sr} = w_r[1 + (p_s - p_r)] - w_s = [1 + (p_s - p_r)]\{w_r - \frac{w_s}{1 + (p_s - p_r)}\}$$

$$w_{rs} - w_{sr} = p_r p_s (q_r - q_s) > 0$$

可知：

當 $0 < w_r \leqslant w_s[1 - (p_s - p_r)]$ 時，$w_{rr} < w_{sr} < w_{rs} \leqslant w_{ss}$；

當 $w_s[1 - (p_s - p_r)] < w_r \leqslant w_s / [1 + (p_s - p_r)]$ 時，$w_{rr} \leqslant w_{sr} < w_{ss} < w_{rs}$；

當 $w_s / [1 + (p_s - p_r)] < w_r \leqslant w_s$ 時；

$w_{sr} < w_{rr} \leqslant w_{ss} < w_{rs}$ 或

$w_{sr} \leqslant w_{ss} < w_{rr} \leqslant w_{rs}$；

當 $w_s < w_r < 1$ 時，$w_{ss} < w_{sr} < w_{rs} < w_{rr}$。

結論就是：若發“猛球”取勝很有把握，應採用 *rr*；若發“猛球”取勝的機會過低（相對於發“普通球”而言），應採用 *ss*，其餘情況，應採用 *rs*。無論如何，採用 *rs* 是穩當的策略，因爲即使它不是最佳，也是次佳。切勿採用 *sr*，因爲它總是最差或次差的策略。這解釋了爲何網球員多數先發“猛球”，如果失手便補發“普通球”，卻絕少反其道而行之。讀者會說：「這是常識吧，動用這麼多計算豈非多此一舉！」對的，我們起首不是說過或然率理論是常識嗎？是否“多此一舉”事屬見仁見智，但至少這說明了建立數學模型是怎樣的一回事，以下我們將經常做這一番工作，先來看一個簡單的情況是有益處的。以後要碰到的例子，不一定像這樣憑直覺已知答案的，到時讀者自然會感覺到數學的美妙了！

（這一節的內容，取材自 L. GILLMAN, MISSING MORE SERVES MAY WIN MORE POINTS, MATH. MAGAZINE, VOL. 58（1985），PP 222-224。）

2. 飯館的生意

先來看一個性質相同但較易令人明白的問題：數學系開設的 24 門課，每門課的學生人數如下表所示：

理化科數學	149	微分方程	79	代數㈠	90
初等分析	100	古典力學	36	代數㈡	34
線性代數	159	數值計算	47	離散數學	46
應用數學	109	作業研究㈠	75	泛函分析	63
社會科學數學	52	作業研究㈡	39	微分幾何	53
分析㈠	78	或然率論	38	幾何拓樸	19
分析㈡	42	數學發展史	21	分析特論	22
複變函數論	51	物理數學方法	34	代數特論	20

總學生課數是 149 + 100 +…+ 22 + 20 = 1456 人，所以平均每一門課有學生 1456/24 = 60.7 人。但如果隨意抽一位學生問平均每門他上的課有多少人，答案是否一樣呢？可以這樣考慮，發給每一門課上的每一名學生一張問卷，請他填上該門課的學生人數，把問卷收集後計算平均人數，就是答案了。應該是［149 × 149 + 100 × 100 +…+ 22 × 22 + 20 × 20］/ 1456 = 121844/1456 = 83.7 人。爲什麼數學系聲稱平均每門課有 61 人左右，但學生卻認爲平均每門課有 84 人左右呢？原因是多數的學生是在人數衆多的課上，所以從學生的角度看，他覺得每門課的人數比較大。

飯館情況是同樣道理，多數的顧客是在生意興隆的時候來，在他們眼中，生意當然十分好，但在飯館老板眼中，卻可能並不一樣。舉一個較極端的例子，一家飯館從早上九時至晚上九時營業，中飯時段很多顧客，中午十二時至下午一時有 300 人，下午一時至二時也有 300 人，其餘時間每一小時只有 1 人，於是老板計算得到的每小時平均人數是（300 + 300 + 10）/ 12 = 50.8 人，顧客卻認爲平均每小

時有(300×300＋300×300＋10)/610 ＝295.1人！

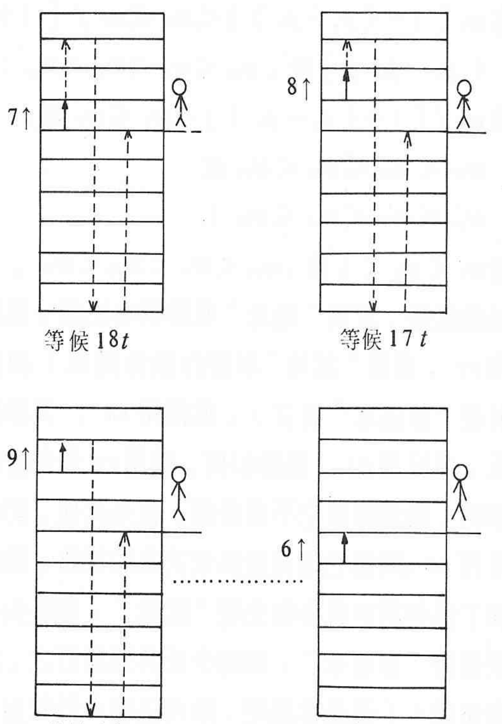
一般而言，前一個平均值總較後一個小了。用數學語言描述，把M寫成 $x_1+\cdots\cdots x_N$，前一個平均值是 $\overline{x}=M/N$，後一個平均值是 $x^*=(x_1^2+\cdots\cdots+x_N^2)/M$。讀者可試證明 $x^*\geqslant\overline{x}$，等式成立當且僅當 $x_1=\cdots=x_N=M/N$。即是說只有在最完全一致的情況下，兩者觀測得到的平均值才會相同。

(這一節的內容，取材自D. HEMENWAY，WHY YOUR CLASSES ARE LARGER THAN 'AVERAGE', MATH. MAGAZINE, VOL. 55(1982), PP. 162-164)

3. 電梯問題

讓我們看看可憐的李先生是否眞的不走運。假定校務大樓共有10層，只有一部電梯。以 $j\uparrow$(或 $j\downarrow$)表示電梯正在第 j 層，將要上行(或下行)。對於站在九樓候電梯下去的李先生而言，有利的情況只有10↓和9↓，不利的情況卻有8↓、7↓、……、3↓、2↓、1↑、2↑、……、8↑、9↑，所以他等到下行電梯的或然率是2/18＝1/9，等到上行電梯的或然率是16/18＝8/9。(當然，我們有個假設，是電梯在任何一層的可能性是同樣大。)反過來，李先生站在二樓候電梯上去的時候，他等到上行電梯的或然率是1/9，等到下行電梯的或然率是8/9，難怪他常自嘆倒楣了！其實並非是他運氣不濟，是或然率作怪吧。

一般來說，如果有N層，在第k層等到上行電梯的或然率是 $P_K(\uparrow)=(k-1)/(N-1)$，在第$k$層等到下行電梯的或然率是 $P_K(\downarrow)=(N-k)/(N-1)$。(這些公式只對$k$不是1或$N$才適用，顯然 $P_1(\uparrow)=P_N(\downarrow)=1$，$P_1(\downarrow)=P_N(\uparrow)=0$，否則你可惹上麻煩了！)。結論是：越是位於高層，越小機會等到下行電梯；越是位於低層，越小機會等到上行電梯。這樣說，位於高層要下來的或者位於低層要上去的，豈不是極不方便嗎？但世事卻往往很公平，如果我們改看等候時間需多久，情況使人寬心多了。假定 t 是從一層到上一層(或下一層)所需的時間，也假定每次去到候電梯時它正好關上門(爲了計算方便)，從下面的圖中不難算得等候時間的期望值，是 $(18t)(1/18)+(17t)(1/18)+\cdots+(2t)(1/18)+(t)(1/18)=19t/2=9.5t$。請你注意，這個答案與你在那一層等候無關，也與你想等上行電梯抑下行電梯無關，即是說就等候時間而言，人人都期望等同樣久。(讀者可試算N層的一般情況。)



等候 18t　等候 17t　等候 16t　等候 t

不過，有時雖然在高層候電梯，卻是要上去；在低層候電梯，卻是要下去。量度"事與願違"的程度，是應該計算某種平均值的。說得清楚一些，設 f_K 是在第k層候電梯的"事與願違度"，則 $f_K=P_K(U)R_K(\downarrow)+P_K(D)P_K(\uparrow)$，這兒的 $P_K(U)$是在第k層候上行電梯的或然率，$P_K(D)$是在第k層候下行電梯的或然率。如果假設沒有人對某層有偏愛的話，$P_K(U)$是$(N-k)/(N-1)$，$P_K(D)=(k-1)/(N-1$

）（ $k\neq 1$ ，N ），$P_1(U)=P_N(D)=1$，$P_N(U)=P_1(D)=0$ 。所以 $f_0=f_1=0$ ，$f_k=(N-k)^2/(N-1)^2+(k-1)^2/(N-1)^2=1-2(N-k)(k-1)/(N-1)^2(k\neq 1,N)$ 。對整幢大樓而言，可以計算“氣結度”$f=f_1/N+f_2/N+\cdots+f_N/N=(N-2)(2N-3)/3N(N-1)$，這個值隨 N 增大而增大，並且趨於極限 2/3 。下班時份情況是否更糟呢？唯一與剛才的計算不同者，是 $P_K(U)$ 和 $P_K(D)$ 變更了，人人都想下去，所以 $P_K(U)=0,P_K(D)=1$ 。新計算得到 $f_K=(k-1)/(N-1)$，新計算得到“氣結度”$f=(N-2)/2N$，這個值也隨 N 增大而增大，並且趨於極限½。試比較平時與下班時分的“氣結度”，如下表所示：

N	平時的 f	下班時分的 f
10	0.50	0.40
20	0.58	0.45
30	0.61	0.46
40	0.62	0.47
50	0.63	0.48
100	0.65	0.49

結論就是下班時分情況不會更糟，反而改善了！（這一節的內容，取材自 A. WUFFLE, THE PURE THEORY OF ELEVATORS, MATH. MAGAZINE, VOL. 55(1982), PP. 30-37 。）

4.擲骰子問題和擲銅板問題

期望擲一顆骰子多少次，才出現某定點數？即是計算這樣的期望值：假定擲得該定點數的或然率是 p ，而 $q=1-p$ ，則擲 1 次便成功的或然率是 p ；擲 2 次才成功的或然率是 qp ；擲 3 次才成功的或然率是 q^2p ；擲 4 次才成功的或然率是 q^3p ；……。要計算的期望值是

$$1\times p+2\times qp+3\times q^2p+4\times q^3p+\cdots$$
$$=p[1+2q+3q^2+4q^3+\cdots\cdots]$$
$$=p/(1-q)^2=1/p 。$$

如果骰子沒有乾坤，p 是⅙，所以通常擲 6 次便出現某定點數。因爲骰子有 6 面，直覺上也會認爲這是答案，是常識嗎？

擲銅板問題，計算是類似的，以 H 和 T 分別表正面和反面，連擲 100 次，期望出現多少次 H 呢？如果出現 H 的或然率是 p ，出現 T 的或然率是 $q=1-p$ ，則 100次裏沒有一次是 H 的或然率是 q^{100}；100 次裏只有 1 次是 H 的或然率是 $\binom{100}{1}pq^{99}$；100次裏只有 2 次是 H 的或然率是 $\binom{100}{2}p^2q^{98}$；100次裏只有 3 次是 H 的或然率是 $\binom{100}{3}p^3q^{97}$；……。要計算的期望值是

$$0\times q^{100}+1\times\binom{100}{1}pq^{99}+2\times\binom{100}{2}p^2q^{98}+\cdots+99\times\binom{100}{99}p^{99}q+100\times p^{100}$$
$$=100p(p+q)=100p 。$$

對一枚均勻銅板而言，$p=½$，所以期望出現 50次 H 50次 T 。請讀者注意，雖然 H 和 T 各佔一半是最可能出現的情況，這個機會還是不大的。事實上，連擲 100 次，剛好出現 50 次 H 和 50 次 T 這個事件並不常見，它發生的或然率連 0.1 也不到呢！

連擲 100次全部是 H 的或然率就更小了，只有 7.88×10^{-31} 。有個這樣的原理：一個事件發生的或然率如果是很小，在一次試驗中它是幾乎不會發生的。所以連擲 100次全部是 H ，幾乎肯定不會發生。（但不是不能發生，事實上，如果你把這個連擲 100次銅板的試驗重覆 $1/7.88\times 10^{-31}\sim 1.27\times 10^{30}$ 次，平均應有一次全部是 H 的。但即使你能每秒擲一次，也得花上 4×10^{24} 年才試畢，而宇宙年歲亦只是 1.3×10^{10} 年吧！）這個“小或然率原理”有一個應用；或者可以稱爲偶然世界的反證法。舉一個例子，有人把一枚銅板連擲 100次都是 H ，問下一次也擲得 H 的機會有多大？一位不懂或然率理論的賭徒說：「既然出現那麼多

次H，下次該輪到出現T了，賭它出現T的勝數一定很高了。」一位“吃古不化”的數學教師說：「按照獨立事件的定義，下次出現H的機會還是$\frac{1}{2}$。」一位活學活用的人說：「如果這是一枚均勻銅板，連擲100次也是H的或然率微乎其微，幾乎肯定不會發生。現在這卻發生了，莫非是銅板有乾坤？且慢，請你把銅板給我看看。」果然，銅板兩面都是H！當然，如果銅板眞的是均勻，下一次出現H的機會還是$\frac{1}{2}$，賭徒的想法是沒有科學根據的。

5.曲別針問題

把一根筆直的鐵絲扳成曲的，有很多很多不同的款式，其中一個就是曲別針。所以要把它還原，成功的或然率極小，幾乎肯定它不會發生。事實上，即使不要求還它本來面目，只要求把它扳成某種特定的款式，也是同樣困難的，只是筆直款式較其它是更有趣而已。說到這裏，不妨區分三回事：⑴小或然率事件，⑵有趣的事件，⑶使人驚奇的事件。最容易描述的是⑴，事件發生的或然率很小就是了。固然，何謂小何謂大，仍是沒有什麼標準可循，必須視乎情況。一批鉛筆中廢品或然率是0.1，說它小是無傷大雅，換了是一批針藥廢品或然率是0.1，卻可不算小了！⑵是個主觀的說法，你說它是有趣便有趣了。或許把⑵和⑶聯起來談更有意思，至少使人驚奇的事件該是有趣吧？而常常發生的事件也不會使人驚奇，所以使人驚奇的事件一定是小或然率事件。反過來卻不一定對，並非全部小或然率事件也使人驚奇，因爲有好些並不是有趣的。擧一個例子，如果玩橋牌時你得着一手這樣的牌：

♠ J 9
♥ A J 5 3
♦ A Q 9 4
♣ 5 7 9

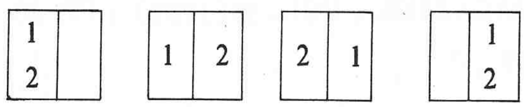
你大概不會到處向人家說：「噢，多使人驚奇呀！我得到黑桃J、9；紅桃A、J、5、3；……。」換了是一手這樣的牌：

♠ A K Q J 10 9 8 7 6 5 4 3 2

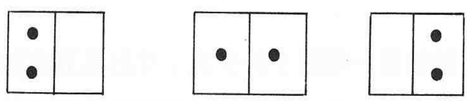
你的反應是截然不同吧？但其實得到前一手牌和得到後一手牌的或然率是一般小，都是$1/\binom{52}{13}\sim 1.6\times 10^{-12}$。

6.墨水問題

大自然現象，好像都從有序到無序。物理學家給的解釋，說後者發生的機會較前者是大得多，所以從有序到無序易，從無序到有序難。放一滴墨水在清水裏，起初清水和墨水分子的排列是有序，攪拌後便變成無序。清水與墨水各歸本位，並非不能發生，只是它發生的或然率極小，幾乎肯定不會發生吧。奧地利物理學家波爾茨曼（BOLTZMANN 1844-1906）便是用這種想法解釋下面的情況：兩個容器以管相連，當中有一開關掣，起初左面一個滿盛氣體，另一個不盛氣體；把掣打開，氣體從左面容器逃到右面，漸漸達至平衡狀態，二者氣壓相同。爲簡化敍述，假定只有兩個分子，稱其一爲1，另一爲2。如果察看“微觀狀態”，應有4個可能，如下圖所示：

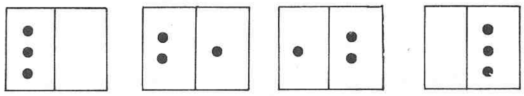


然而對一外在觀察者而言，何者是1何者是2，是分不開的，所以只有3種“宏觀狀態”，如下圖所示：



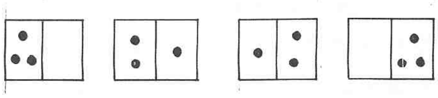
各佔或然率$\frac{1}{4}$、$\frac{2}{4}$、$\frac{1}{4}$，以每容器各含一個分子這情況是最可能發生。如果有三個分子，共有8個“微觀狀態”，歸納爲4種“宏觀狀態

”，如下圖所示：

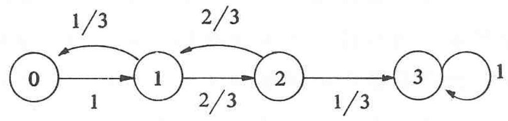


各佔或然率⅛、⅜、⅜、⅛。若有N個分子，得來的“宏觀狀態”共有$N+1$種，各佔或然率$\binom{N}{0}/2^N$、$\binom{N}{1}/2^N$、……、$\binom{N}{N}/2^N$。實驗觀測到一個體系總是自發地向一種特定的狀態發展，就是因爲該種狀態的微觀分佈方式數目最多，對應的（宏觀）狀態出現的可能性最大。但這僅是可能性最大，卻不排除體系向相反方向發展的可能。中國成語有句「守株待兔」，原意是叫人不要白費心機呆候，但其實這樣做不一定沒收穫的，只要等的時間足夠長，一定有兔呢！荷蘭物理學家愛蘭費斯（EHRENFEST）在1907年設計了這樣一個模型，把分子看成是標以號碼的球，每次隨意抽一個號碼，把標以該號碼的球從一個容器移往另一個。起初全部球都在左面的容器，試驗初期，左面容器的球的數目下跌得很快，趨於兩個容器的球的數目相若的情況，在其間擺動。使人驚奇者，是這個模型也告訴我們，全部球回到原來左面容器的或然率是1。不過，平均可要抽上2^N次（N是球的總數），即使$N=200$，2^N已經差不多是1.6×10^{60}，假定每秒移球一萬次，也得移上5×10^{48}年！

讓我們看一個類似的有趣問題：三隻蒼蠅在房間裏飛進飛出，使人不勝其煩，期望等多久才等到全部蒼蠅飛了出外，關上窗門，安享太平？我們用上面的語言，把蒼蠅看作球，房間看作（左面）容器，外面看作另一個（右面）容器。共有4個狀態，如下圖所示，標作0、1、2、3：



從一個狀態至另一個狀態的或然率不難算出（假定蒼蠅隨意飛進飛出），可以下面箭頭示意圖表達：



要抵達狀態3，必先經過狀態2；而要抵達狀態2，必先經過狀態1。從狀態0出發，必先到狀態1。因此，如果從狀態0到狀態3需走N步，過程必是從狀態0到狀態1，再經$N-3$步後回到狀態1，然後到狀態2，再到狀態3。從狀態1到狀態1（不准到狀態3），只能通過狀態0或2，所以$N-3=2m$，即是$N=3+2m$（$m=0,1,2,3,\cdots\cdots$）。因此，從狀態0通過$3+2m$步抵達狀態3的或然率是

$$\left(\frac{2}{9}\right)\left[\left(\frac{1}{3}\right)^m+\binom{m}{1}\left(\frac{1}{3}\right)^{m-1}\left(\frac{4}{9}\right)+\cdots\cdots\binom{m}{m-1}\left(\frac{1}{3}\right)\left(\frac{4}{9}\right)^{m-1}+\left(\frac{4}{9}\right)^m\right]$$

$$=\left(\frac{2}{9}\right)\left[\left(\frac{1}{3}+\frac{4}{9}\right)^m\right]=\left(\frac{2}{9}\right)\left(\frac{7}{9}\right)^m$$

（爲什麼？），

其餘是零，所以期望時間是

$$3\times\left(\frac{2}{9}\right)+5\times\left(\frac{2}{9}\right)\times\left(\frac{7}{9}\right)+7\times\left(\frac{2}{9}\right)\times\left(\frac{7}{9}\right)^2+\cdots\cdots$$

$$=\left(\frac{2}{9}\right)\left[3+5\left(\frac{7}{9}\right)+7\left(\frac{7}{9}\right)^2+\cdots\cdots\right]$$

$$=\left(\frac{2}{9}\right)\left(\frac{90}{2}\right)=10\text{。}$$

這個答案並不是明顯的常識了。更一般的情況還是可以處理的，但那已踏入馬爾可夫鏈（MARKOV CHAIN）的範圍，不述。

7.或然率與判斷

很多時，或然率雖屬常識，卻需小心運用。最常見的錯誤運用，是對事件賦予不正確的或然率，這往往是因情況資料不足，與數學無干，也不是單憑數學能解決的。譬如你見到前面有一個人來到三岔路口，兩條通往山上，一條通到山谷下面，你猜他走那條？既然有三條路，是否有⅓機會走通到山谷的路？但他或是上山或是下山，那豈不是只有½機會走通到山谷的路！也說不定那人擲骰決定，出現六點才往山谷，所以他有⅙機會走通到山谷的路。這種種皆與數學無干，是憑額外資料才解決得來。

第二點要小心的是不能只看“大局”。請看下面的例子（取自蕭文强、林建，概率萬花筒，香港廣角鏡出版社，1982，第 85-88頁）：兩所醫院的手術成功率，可由下面數據計算出來，醫院甲有916人動手術後康復，84人動手術後死亡，成功率是0.916；醫院乙有1880人動手術後康復，120人動手術後死亡，成功率是0.94。是否醫院乙水準較高呢？但如果作深入分析，卻發現是這樣，原來醫院甲動大手術較多，醫院乙動小手術較多，記錄如上表所示。從這些數據，可知醫院甲的大手術成功率是0.9，小手術成功率是0.98；醫院乙的大手術成功率是0.88，小手術成功率是0.96。即是說，不論大手術或小手術，醫院甲的表現都較佳，怎能說醫院乙水準更高呢？總的計算醫院乙成功率較高，只因為它的病人病情較輕而已，只有¼需動大手術，反觀醫院甲，卻有八成需動大手術。

醫院甲	大手術	小手術
康復人數	720	196
死亡人數	80	4

醫院乙	大手術	小手術
康復人數	440	1440
死亡人數	60	60

第三點要小心的是不要過份迷信數學的功用，拉普拉斯的故事是一個歷史教訓。拉普拉斯雖然對或然率理論有極大貢獻，但因爲他曾熱衷於把它不適當地用於某些社會現象，過份估計了數學的功用，被當時其他人形容爲“數學囈語”，連帶使或然率理論也蒙上汚名，使它的理論發展在西歐與當時俄國學派的工作比較，有一段相對停滯的時期。例如拉普拉斯認爲每位法官達致正確判決的或然率是一個常數，而且每位法官作的結論是彼此獨立的，所以只要審判時有很多法官參加，以大多數表決行事，審判便會非常公正了，因爲判決錯誤的可能性將減至極低。

8. 法庭上的或然率

說到審判，的確有不少事例是與或然率的計算有關，西方法律書籍把“有關的證據”界定爲“TENDENCY TO MAKE THE EXISTENCE OF ANY FACT THAT IS OF CONSEQUENCE TO THE DETERMINATION OF THE ACTION MORE PROBABLE OR LESS PROBABLE THAN IT WOULD BE WITHOUT THE EVIDENCE”。裁定被控者有罪與否，應按照證據是否確鑿，所謂“無可置疑”（BEYOND REASONABLE DOUBT）。固然，怎樣才算無可置疑，是大有文章可做的。下面是幾樁實例：

(一) 1962 年瑞典警方有一次控告某人在只許停車若干鐘頭的地方過時停車，警方使用的證據是這樣的，警察記錄了停車前後輪胎上打氣活瓣的位置（以時鐘時針的12個位置作記號），如果逾時發現記號位置沒更改，便顯示車子停在該處過久了。被告者反駁說，在此期間他曾駕車離去，後來回去剛好把車子停在同一地點，而且兩個輪胎又剛好停在先前的位置！法庭計算這個事件發生的或然率，是 $1/12 \times 1/12 = 0.007$，認爲還沒有足夠小至使人難於

相信，結果判被告無罪。法庭還說，如果警察事前記錄全部四個輪胎的位置，這個事件發生的或然率是 $1/12\times1/12\times1/12\times1/12=0.00005$，才算小至使人難以相信它會發生，所以證據才算“無可置疑”。讀者們，你認爲法庭的計算合理嗎？

(二)在二十年代美國德州有位老人萊斯（RICE）忽然逝世，遺下財產六百萬美元。有位名叫栢德力（PATRICK）的律師拿出一份據說是萊斯立下的遺囑，承受人就是他本人。這自然引起萊斯親友的猜疑，請法庭判別眞僞。遺囑有四頁紙，每頁寫上萊斯的簽名。字跡專家奧斯本（OSBORN）細心檢查，發現那四個簽名完全一模一樣，但通常的簽名，每次總有少許出入，極難完全一模一樣的。兩個簽名完全一模一樣的或然率極小，何況是四個簽名一模一樣，簽名的人更是年邁老翁呢？法庭根據這一點判決遺囑無效，因爲它是揑造的。萊斯的財產結果用以建立了德州有名的萊斯學院（RICE INSTITUTE）。

(三)1915年紐約律師瑞斯李（RISLEY）被控塗改法庭文件，控方使用的證據是這樣的。更改的六個字裏，顯示了打字機的11處特徵，而瑞斯李辦公室的打字機正好有這些特徵。一位數學教授作證，說這種吻合的或然率只有40億份之一，既然這樣小，使人難以相信它是巧合。但法庭終判瑞斯李無罪，理由是(1)數學教授並非打字機專家，(2)他的答案不知憑何計算而得，(3)他企圖以數字混亂陪審團。

(四)六十年代在美國洛杉磯有一老婦被刼掠，有人看見一對男女從現場逃走，男的是個黑人，蓄小鬍子，也蓄大鬍子；女的是個白人，金髮束成馬尾狀；兩人駕駛一部黃色汽車。後來警方逮捕了哥連士夫婦（COLLINS），男的是個黑人，曾蓄小鬍子和大鬍子（被捕時已剃掉）；女的是個白人，金髮，平日喜束馬尾狀；兩人擁有一部黃色汽車。控方作如下的或然率計算：

駕黃汽車者	1/10
蓄小鬍子男子	1/4
束馬尾狀女子	1/10
金髮女子	1/3
蓄大鬍子黑人	1/10
黑白夫婦	1/1000

把各值相乘，得一千二百萬份之一。控方就憑這一點計算，認爲洛杉磯居民中符合上述特徵者可謂“萬中無一”，現在既然捕捉了一對這樣的男女，不是犯罪者還有誰？就是這樣，哥連士夫婦被判有罪入獄。過了幾年後，他們上訴得直，才給釋放。上訴得直的理由是(1)控方提供的那些或然率不知怎樣得來，(2)控方企圖以數字混亂陪審團，(3)法庭重新計算在捕捉了一對符合上述特徵的男女條件底下，還有另一對符合上述特徵的男女的或然率，是 0.42，並不算小，又怎能一口咬定哥連士夫婦就是犯罪者呢？法庭的計算是這樣子，首先假定符合特徵的男女的數目，服從普阿松（POISSON）分佈，即是說，有m對男女符合特徵的或然率是 $\lambda^m e^{-\lambda}/m!$，$\lambda$ 是平均有多少對男女符合特徵。即使以先前控方的估計，$\lambda=1$，要計算的或然率是

$$P(\text{有多於一對男女符合特徵})/P(\text{有至少一對男女符合特徵})$$
$$=[1-P(\text{有0對})-P(\text{有一對})]/[1-P(\text{有0對})]$$
$$=[1-e^{-1}-e^{-1}]/[1-e^{-1}]$$
$$=0.42$$

讀者們，你對此案的或然率計算有什麼意見呢?

(五)最後一件案件是著名的德雷福斯（DREYFUS）賣國案，震撼了法國軍政界與民間。德雷福斯是一位任職於法國陸軍參謀總部的軍官，忽然在1894年10月15日被捕，控以賣國罪。原來在1894年9月在巴黎德國使館一位武官的字紙簍裏找到一封沒署名的信，寫信的人答應向德國提供軍方的情報，信的字跡酷肖德雷福斯的字跡。德雷福斯堅持自己清白，但當時軍方爲了政治需要，必須定他的罪來維持聲譽，便千方百計搜集證據。當時的國防

部長甚至指使情報組僞做秘密檔案以指證德雷福斯。結果德雷福斯被判有罪，終身監禁於魔鬼島。這件事引起民憤，很多有名望的知識份子組織起來營救德雷福斯。在1899年8月7日德雷福斯第二次被審，仍判有罪，但法庭含糊其詞，並且教他向總統請求特赦。終於民主力量得勝，在1906年法庭宣判以前的裁決無效，不只把德雷福斯復職，還頒予勳章。

控方當時的一項指證，說德雷福斯寫的信，字母出現頻率並不吻合統計得到的一般法國文字中字母出現的頻率，所以他一定是以密碼傳遞情報了。當時法國最負盛名的數學家龐卡萊（POINCARÉ）也被請上法庭作證，他正確地指出即使最可能發生的事件，眞的要它發生，或然率是很小的（記得連擲100次銅板要出現50次正面的或然率，十份之一也不到嗎？），所以這種指證不能成立。控方不只不接受，還譏笑龐卡萊！另一位法國數學家彭利維（PAINLÉVÉ）憤怒極了；他說：「給我RACINE（十七世紀法國著名劇作家）的作品，我將以你們的愚蠢辦法證明他是個賣國賊，因爲在他的作品裏，如同德雷福斯的信件一樣，也不吻合字母出現的最可能分佈！」

（這一節的內容，取材自E. B. MODE, PROBABILITY AND CRIMINALISTICS, J. AMER. STAT. ASSO., VOL. 58(1963), PP. 628-640；M.W. GRAY, STATISTICS AND THE LAW, MATH. MAGAZINE, VOL. 56(1983), PP. 67-81）

* * * * * * * * * * *

看過這衆多事例，讀者認爲或然率理論是否常識而已？也許不儘是，不過拉普拉斯說過另一句話，卻很貼切，他說：「生活上多數重要的問題，其實都是或然率的問題。」

—本文作者任教於香港大學—

質數十一講

近年來，數學上最轟動的事件之一是有名的黎曼假設（The Riemann Hypothesis）似已被證實（後來提不出證明來，只空歡喜一場），要談黎曼假設，就得先從黎曼 zeta－函數說起，對任意正整數 n 及複數 $z=\lambda+i\sigma$，由

$$|n^{-z}|=n^{-\lambda}$$

及比較審斂法知道無窮級數 $\sum_{n=1}^{\infty} n^{-z}$ 在 $Re z=\lambda>1$ 時是絕對收斂而定出一函數，這函數即是有名的黎曼 zeta 一函數，一般以 $\zeta(z)$ 來表示，亦即

$$\zeta(z)=\sum_{n=1}^{\infty}\frac{1}{n^z}, \quad Re z>1$$

對於函數 $\zeta(z)$，一般人應不會生疏，例如在判定級數的斂散性時，常跟這級數做比較，又如

$$\zeta(2)=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots=\frac{\pi^2}{6}$$

$$\zeta(4)=1+\frac{1}{2^4}+\frac{1}{3^4}+\cdots+\frac{1}{n^4}+\cdots=\frac{\pi^4}{90}$$

都是已知而有趣的事實；由正整數的分解定理（即每一正整數都可表成 $p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_n^{\alpha_n}$ 的質數乘積），亦可將 $\zeta(z)$ 表成無窮乘積形式

$$\zeta(z)=\sum_{n=1}^{\infty}\frac{1}{n^z}=\prod_{p\in P}(1-p^{-z})^{-1}, Re z>1$$