美夢成眞的機率有多少

樂透登場全民譁，競求財神助我家。
渾沌攪球稱隨機，亂中藏序是機率。
七期連五真神奇？字尾參差惹猜疑。
高額頭彩頻呼喚，期望回收剩一半。
中獎人數玄機露，死棋突困說策謀。
少量投注最上策，專心正事闔家樂！

劉應興

樂透登場全民譁，競求財神助我家

今年1月16日臺灣地區「樂透」彩券開跑，很快就成為社會上最熱門的話題，不僅新聞每天報導，專題網站亦應時而生，電子布告欄無數個版面都在討論相關的主題。民眾更是投注熱烈，風雨無阻，大排長龍，無非希望能一夕致富。

至2月26日，不到一個半月的時間，累計投注金額高達新台幣一百四十億元，雖為國庫增加了五十多億的收入，但是消費市場、投資市場卻同時受到資金排擠效應。前十一期的樂透彩，造就了三十餘位超級富豪，卻也造成了幾起違法案件及家破人亡的事例，甚至引發「社會土石流」的議論。投注站為爭取生意而服務至校園，一些BBS站甚至成立樂透專版，引起教育界的質疑。這些因彩券所引起的經濟、政治、社會及教育問題，點出了彩券雖小而問題大，值得多方探討。

樂透是一種機會遊戲，許多民眾卻因教育不足而迷信明牌，甚或誤信不法之徒的謬論，以為花點小錢即可致富。還有許多人求神問卜，期望神祇庇祐賜福；或追逐孩童、瘋子、野狗，以為藉此能參透天機。更有許多曾受「高等教育」者，竟也迷信「科學」明牌，以為藉由數學，利用電腦，即可算出中獎號碼。二月初某報導指出，有位高中學生以遞迴數列及渾沌現象預測明牌；另一報導則說有人能測天機，以

中若與開出的六個獎號有三個以上相同，即算是中獎，但中三個號碼（普獎）只得新台幣二百元。投注時可利用投注單劃記，交給經銷商放入投注機判讀；也可以口頭告知經銷商輸入；或由經銷商按投注機觸控螢幕之「快選」鈕由電腦隨機選號。投注號碼輸入後，由熱感式彩券列印機印出彩券。

開獎是將標有號碼的彩球一組四十二個放進開獎機。現行方式採抽球決定由01至42遞增，或由42至01遞降順序放入彩球。彩球先依序排列於滾球艙右側置球架上，便於觀眾檢視。然後開啟閘門讓四十二個彩球進入滾球箱，並吸入空氣產生氣流擾動四十二個彩球。在擾動中有一個球可能衝出上方活門，其他彩球繼續被箱內氣流擾動、衝出。如此開出七個彩球後停止，前六個彩球的編號就是當期獎號，第七個彩球的號碼稱為「特別號」，是區分二、三獎之用。

樂透開獎機用氣流將彩球攪亂，會開出的號碼事先不能預期。這就像丟一個十元硬幣，硬幣落地時會

一簡單公式計算明牌，連中數期。

其實只要略懂「機率」，即不致於被這些謬誤所騙。但亦有對機率、統計一知半解，卻以錯誤的方法解讀觀察結果，而衍生許多無謂的猜疑。本文探討樂透彩券的一些機率問題，並從統計學的觀點，分析幾個曾被認為奇特，甚至被認為有作弊嫌疑的現象。

渾沌擾球稱隨機，亂中藏序是機率

臺灣地區現行的「樂透」遊戲設計，投注者從01至42共四十二個號碼中，任選不重複的六個數字，其

（左）投注站終端機。
（右上）彩券印表機。
（右下）列印情形。

出現的結果事先並不知道，數學上稱這類實驗是「隨機實驗」。

「隨機」這兩個字沒人能給予完整的定義，或許可以這樣說：一個現象或一個實驗的結果被稱「隨機」，意指它的結果是毫無控制地亂，亂到觀測者沒有辦法去預測；但說它亂，它卻又能呈現出一種規律性。丟一個十元硬幣，觀察者無法預測下一次的結果是人像或文字（「拾圓」）；但丟了很多次，會發現出現人像的次數大約是丟擲次數的一半，而且丟越多次，出現人像的次數比例越接近二分之一。這種現象在數學（機率學）上稱為「大數法則」，而二分之一就是丟一個十元硬幣出現正面的機率。

渾沌現象也如同隨機現象一般，難以預測下一刻會發生什麼結果，但彙總其結果又呈現某種規律性。不過，渾沌現象是由一個固定的機制產生的，隨機現象卻不是。前面提到有高中生利用渾沌遞迴序列預測樂透獎號，其實應該說是利用渾沌序列模擬樂透開獎結果。這樣的模擬，不僅沒有預測號碼的功能，並且因為還沒有足夠數據，也無法有效估計何種號碼比較容易出現。樂透開獎機內部的運作或許不能說是隨機的，但它產生的氣流應能造成紊流，也就是一種渾沌現象。而不同期開獎，又攙雜了許多不可測量的因素，使結果呈現了隨機性。像這種開獎裝置，在出廠前應做適當的測試，以保證各彩球出現機會均等。

由於有「神奇的39號」在前七期中五連莊，以及前十一期39,29兩號碼出現頻率偏高，不免令人懷疑是否彩球重量、形狀，以及機器設計、製造及開獎操作有些不當，致使結果不隨機或各號碼出現機會不均？對於這些問題後文將從機率學討論其發生的可能性，並以統計學觀點進行檢驗。但在檢驗之前，需要先把「機率」究竟是什麼弄清楚。

機率是介於0和1之間的一個小數，如前述丟硬幣實驗的二分之一，用來衡量「丟出人像」發生的機會。但實際去丟硬幣的結果不會剛好一半人像一半文字（如果這麼規律就稱不上「隨機」了）。那麼我們又如何能肯定地說，出現人像的機率是二分之一，而不是其它數值？

樂　透　模　擬

頭獎號碼						特別號
06	13	15	19	20	37	38
06	09	19	21	27	38	39
01	20	22	24	36	38	42
03	04	12	16	27	30	40
03	24	26	28	34	35	41
01	11	22	32	34	36	41
17	21	22	30	34	39	42
02	12	21	26	30	33	39
04	07	14	16	22	25	39
02	07	18	20	21	28	33

樂透彩中獎機率分析

- **頭獎：獎號六個數字全符合手中彩券**
  機率：$1/C(42,6)=1/5245786=.000000190629$
- **貳獎：獎號六個數字中之五個符合手中彩券，特別號開出彩券另一數字**
  機率：$[C(6,5)\ C(36,1)/C(42,6)]\times(1/36)=6/5245786=.000001143775$
- **參獎：獎號六個數字中之五個符合手中彩券，特別號不符彩券上的數字**
  機率：$[C(6,5)\ C(36,1)/C42,6)]\times(35/36)=210/5245786=.000040032132$
- **肆獎：獎號六個數字中之四個符合手中彩券**
  機率：$C(6,4)\ C(36,2)/C(42,6)=9450/5245786=.001801446$
- **普獎：獎號六個數字中之三個符合手中彩券**
  機率：$C(6,3)\ C(36,3)/C(42,6)=142800/5245786=.02722185$

機率學的研究往前可以推至十六、七世紀卡答諾（G. Cardano）、伽利略（Galileo）、巴斯卡（B. Pascal）和費瑪特（P. de Fermat），當時所處理的主要是有關骰子、撲克牌一類的機率問題。他們假設：公正的骰子每面出現的機會一樣，從充分洗牌後的撲克牌抽牌也一樣。伯努利（J. Bernoulli）最早給予機會均等假設一個基礎，稱為「理由不充分原理」；凱恩斯（M. Keynes）將它修訂為「無差別原理」：

> 如果沒有已知的理由可以使我們預見實驗結果中某一個比另一個更可能出現，並且這些可能結果對已知證據而言是對稱的，那麼它們就都具有相等的機率。

丟硬幣出現人像的機率為二分之一，就是依這個原理而設定的。它不一定對，但可經由實驗並利用統計方法檢驗。例如：假設丟了100次，結果人像有56次，這樣是否可說這枚硬幣不「公正」？統計人員假設出現人像的機率是P，拿$P=1/2$和P的不同值比較，來決定何種實驗結果可判定該硬幣不公正。就此例來說是看人像實際出現次數與理論值（50次）的差距。因$P=1/2$時兩者相差6以上的機率大約0.23，並不算罕見，表示差距不算太大，因此不能認為有問題。

假設彩球、開獎機及操作程序沒問題，即可假設各號碼出現的機會均等，而可算出得頭獎機率是1/5245786。因為從42個號碼中取出6個而不考慮前後順序，則總共有5245786種不同的組合。就二獎而言，手中彩券上須恰有五個號碼與開出的獎號相同，這就有$C(6,5)\ C(36,1)$種情形，即：從手上6個號碼取5個，再從其他36個號碼之中取1個。但這包含二、三獎，必須特別號符合彩券上剩下（和獎號不符）的那個號碼，也就是開特別號時36個球剛好抓到手上剩下的一個號碼，就這樣算出二獎機率等於6/5245786。其他獎項算法類似，就不細述了。

七期連五真神奇？字尾參差惹猜疑

現行樂透中獎機率才2.906%，其中主要還是二百元的普獎（2.722%）撐場面。也就是說：花五十元投注，幾乎確定損龜，只有極少數特別幸運的人能中彩。就因這個緣故，從開辦以來，不公正、作弊等疑問不斷被提出。例如第三期開出後，即有人質疑「從17　27總共有連續的十一個號碼都沒有出現？」難道開獎機認得號碼嗎？為什麼單看「連續」號碼？如果不管號碼本身的意義，當時有23個號碼在三期中都沒出現，比起只看十一個號碼會覺得更極端。第六期後有人注意到39號連續出現四期而懷疑是否球重不同。為此第七期開獎時特別當眾檢驗了兩盒彩球的重量。

雖然結果顯示彩球都合乎規格，但誰又能證實些微的差距不會導致機會不均等？而該期又再度開出39號，更增添幾分神奇。究竟出現這種情形的機率有多大？有許多人是這樣算的：令A_{39}表示一期中開出39號，A_{39}^{5}表示特定五期連續開出39號，則因

$$P(A_{39})=7/42=1/6,\ 故P(A_{39}^{5})=(1/6)^{5}=.0001286$$

機率相當低！如果把兩期沒出現39號球的機率也算進去，結果更低！從機率的觀點，只要有正的機率，就有發生的可能。如前文分析，中頭獎的機率是數百萬分之一，但你可能只投一注就中頭獎。

然而，以統計學觀點來看，上述計算是不對的！這個算法必須事先指定特定的五期，然而這五期以及39號球卻都是事後觀察的結果。同樣情形我們可能看到13號球出現於1　5期，然後我們會懷疑：為什麼13號球能連續出現五期？應該是在一開始就指定號碼，然後看往後該號碼出現是否異常。以39號球而言，在第六期後被懷疑有問題，往後看，發現在第7期出現，然後第8　11期都沒出現，這樣的機率是：

$$P(E)=(1/6).\ (5/6)^{4}=0.0804$$

機率不算高，是因還有眾多可能情形；若39號球開出機率是其他球的兩倍，則同一事件的機率只有0.0716。這表示在第六期懷疑39號球較易開出而觀察往後情形時，不能只看一個特殊情況的機率，而應看他是否繼續更常出現。在機會均等時，39號球在7　11期至少出現一次的機率是0.598。

統計人員並不總是等著日後再驗證，39號球的問題可用另一種方式處理。我們不能因為看到39號球在第3　7期都出現，就直接算這個事件的機率。但我們可以問：在七期中有某個號碼連續出現至少五期的機率是多少？如果指定一特殊號碼，例如39號球，可算出它出現在第1　5期、第2　6期或第3　7期的機率是0.0003429。但這種事在每一個號碼都可能，39號可能只是偶然罷了！因此，利用機率學中的「取捨原理」

$$P(A_1\cup\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{i<j}P(A_i\cap A_j)+\cdots$$

可以計算出七期中至少一個號碼在連續五期中出現的機率是0.0144，而不是像先前列的那麼誇張。若考慮的不是「連續五期」而是「至少五期」，則39號球的現象更不是那麼稀罕。

無獨有偶地，29號球也是較常出現。在1　11期中，29號和39號球都各出現五次。某一個特定號碼在11期中出現五次並不稀奇，但特定兩個號碼都出現五次以上就比較引人遐思。因此，有觀察者做了陳述：「尾數是9的出現機會較大。」雖然這不像另一種說法：「29, 39兩號碼出現機率特高」那麼極端，然而，這個陳述還是犯了「事後解釋」的錯誤！以觀察到的資料建立假說，又用同樣的資料來檢驗這個假說，會產生嚴重誤導。

要瞭解「尾數是9」的出現機會是否異乎尋常，統計上可採用「卡方檢定」。結果發現：在各號碼球出現機會均等的假設下，11期結果會像實際上出現的這樣或更極端情況的機率大約1/4，並不稀奇。雖然“9”字尾出現頻率似乎偏高，不過，在十組字尾中出現這麼一個稍微偏高的離差值，也不算太奇怪。在筆者所做的模擬中，也曾遇到某一字尾出現13次而另一字尾只出現三次的情況，比實際資料還極端。隨機的現象就是這樣：要它乖一點，它卻亂跳不受控制；要它極端些，它偏要向中道靠攏。

神奇的39號球

- A (i) 代表在7期中 i 號球連續出現5期以上
- 令1代表出現第 i 號球，0表不出現，“*”不管出不出現，則A (i) 代表結果序列為 (11111**), (011111*), (*011111) 三種
- $P〔A(i)〕=\left(\frac{1}{6}\right)^{5}+\left(\frac{5}{6}\right)\cdot\left(\frac{1}{6}\right)^{5}+1\cdot\left(\frac{5}{6}\right)\cdot\left(\frac{1}{6}\right)^{5}$
  $=.00034293553$
- $P〔A(i)\ or\ A(j)〕=.51939666\times10^{-7}$
- $S1=\Sigma P〔A(i)〕,$
- $S2=\Sigma P〔A(i)\ and\ A(j)〕, i<j$
- 則$S1-S2\leq P$〔A (l) or ... or A (n) 發生〕$\leq S1$
- 結果P〔A (1) or ... or A (42) 發生〕=.0144

高額頭彩頻呼喚，期望回收剩一半

中獎機會不到3%的樂透造成風潮，固然為政府帶來不少的收入，但就投注民眾的觀點來說，能獲得多少回收才是重要的。高額頭獎彩金確實誘人，十二期彩券創造了數十位千萬級富翁讓人羨慕。然而，頭獎、二獎中獎機率合計才百萬分之一強，不到3%的三、四獎及普獎得獎人想必把獎金再投入也是一去不回。一家歡樂何只萬家愁？

具隨機性收益的活動如樂透彩的價值評估，第一個被考慮的是期望值。假設機會性遊戲只有兩種結果：一是得到新台幣1,000元，有1/5的機會；另一是沒有回收。根據前面提過的大數法則，如果玩很多次，平均每五次有一次得新台幣1,000元，另四次一毛錢也沒有，等於平均每次得新台幣200元。假設上述遊戲改成有1/10機會獲得1,200元，有1/5的機會獲得400元，其他情形沒有回收。這個遊戲的期望值就是：

$$1200\times(1/10)+400\times(1/5)+0\times(7/10)=200$$

假設上列遊戲是一種彩券，發行一萬張，其中1,000張每張可得獎金新台幣1,000元，2,000張每張可得獎金新台幣400元，其餘7,000張無獎。則買彩券的人每買一張的期望回收是新台幣200元，一萬張彩券合計新台幣200萬元；而發行彩券的公司支出的獎金也是200萬元（1200元×1000張＋400元×2000張），這提供計算彩券期望值的簡便算法：彩券期望值即是所有獎金平分給每張彩券的值。

現行樂透彩投注額的56%為獎金，因此如果不考慮稅捐，長期而言每注的平均回收是28元；若扣除普獎以外的所得稅和印花稅，只得23.4元。

對投注人而言，期望值卻不等於上述長期平均回收款，主要原因是可能某一類獎項無人中獎，或有前期累積獎金；另一個因素是投注人的選號並不是隨機的，這尤其使得期望值的計算成為不可能的任務。

如果所有投注者的選號是隨機的、互不相干的、沒有號碼偏好，那麼每注的期望值是一樣的。此時計算期望值，只需加上前期累積獎金及本期某獎項無人中獎而獎金移入下期累計的情形。臺灣地區樂透投注數除第一期1,552萬注最少以外，至少都在1,900萬注以上；即使假設以後逐漸退燒，每期應不致低於1,000萬注。二獎中獎機率6/5245786，平均中獎人（注）數11.44。在隨機選號等條件下這是屬於「二項實驗模型」，並且可用「卜瓦松分布」描述中獎人數的機率分布。但在平均數11.44之下無人中二獎的機率只有十萬分之一，對期望值的影響近於零，可忽略。

至於無人中頭獎機率，當投注數1,000萬時有0.1486；投注數2,000萬時也還有0.0221。在扣除前期累積獎金後，前11期平均每注之頭獎獎金最低是第3期8.09元，最高為第4期9.20元，故無人中獎機率不能忽略。於是得無累積獎金且已知總投注數為n時，每注之期望值是：

$$E[X\mid N=n, W=0]\quad 28-A\times P〔無人中頭獎〕$$
$$=(28-A)+A\times P〔有人中頭獎〕$$

其中A代表無累積獎金下平均每注頭獎獎金。普獎人（注）數以理論值0.02722185n計算，則A＝8.57114。有頭獎累積獎金w時，

$$E[X\mid N=n, W=w]$$
$$(28-A)+(A+w/n)\times P〔有人中頭獎〕$$

雖然有頭獎累積獎金時簽注數會大幅增加，但無論如何它確能提高期望值。只不過，即使提高，仍遠不及投注成本（50元）。前12期分別在第1,5,11期無人中頭獎，對次期期望值的貢獻依次為6.90元，4.45元，5.72元。有高額累積頭獎，必帶來高額投注量，可假設n足夠大，故無人中頭獎機率可忽略。在不計稅捐下，期望值要超過彩券成本，必須累積獎金被投注數均分後達22元，若考慮稅捐則需要33.4元。這代表連續未發出頭獎，獎金各期累計簽注數至少分別是本期的2.57或3.90倍，但這種情況發生的可能性微乎其微。

以上結果給投注者的啟示是：等候有累積獎金再來投注，或有累積獎金時增加投注，可以提高期望回收；但扣除投注成本後，幾乎可說期望值一定是負的。期望值是負值，而贏（中獎）機率接近零的機會性遊戲只能玩玩，卻不能太認真啊！機率理論（大數

法則）告訴我們：長期玩下去的結果，期望值不只是理論，而是事實！這就是「久賭必輸」的道理。至於單期大量投注，只是加速這個過程罷了！

中獎人數玄機露，死棋突困說策謀

以上關於期望值的結果依賴於假設「所有投注號碼是隨機的、互不相干的、無偏好的」，但事實上不論國外或國內資料，都證實投注號碼不滿足上述假設。臺灣地區1　11期樂透頭獎實際中頭獎人（注）數比理論值有高有低。比理論值高的，可計算在假設模型下比該實際中獎數高或至少相等的機率；比理論值低的，則計算比該數小或相等的機率　（這等於統計學上做雙邊假說檢定時P值的一半）結果第5、7、11三期機率都偏低。第5期及第11期無人中頭獎都算是較特殊的，但還比不上第7期一人獨得頭獎的稀奇；至於第六期十二人平分頭獎就一點也不奇怪了！

投注號碼並不隨機或不均勻是可預期的，若能取得詳細投注資料或統計，可以分析投注者的號碼偏好和變化。如果能及時知道各號碼組合的投注數，基於開獎機對號碼無偏的假設，選擇較少人簽注的號碼組合，則可以提高期望值。舉個簡單的例子：假想一種彩券只有六個號碼，大獎總獎金新台幣16,000元，上下附獎新台幣4,000元，各由得獎人均分。例如1　6號投注數依次為2,8,15,25,50,100時，若獎號是1號，則簽1號的投注者每人各得新台幣8,000元，而簽2號和6號的投注者108人均分新台幣4,000元。

在各獎號機會均等的假設下，選擇最冷門號碼（1）的期望值是選熱門號碼（6）的36倍；若投注數差異縮小。期望值差異也縮小；另外，投注數組成和給獎規則的交互作用也會影響期望值差異。北銀的樂透彩當然較此例複雜許多，但基本特性是一樣的（只是實際期望值差距多大因沒有投注號碼資料無從計算）。不過，如果所有投注者都認識到這一點而捨去「明牌」，則冷門號碼也可能變熱門號碼。

中頭獎之期望報酬

μ	中1注	中2注	中3注
0.5	0.78694	0.85225	0.88653
1.0	0.63212	0.73576	0.79272
1.5	0.51791	0.64278	0.71444
2.0	0.43233	0.56767	0.64850
2.5	0.36717	0.50627	0.59248
3.0	0.31674	0.45551	0.54449
3.5	0.27709	0.41309	0.50306
4.0	0.24542	0.37729	0.46703
4.5	0.21975	0.34678	0.43548
5.0	0.19865	0.32054	0.40768
5.5	0.18108	0.29779	0.38302
6.0	0.16625	0.27792	0.36104

少量投注最上策，專心正事闔家樂

前面提到兩個投注策略有助於提高期望值，一是在有前期累積獎金的情況投注，二是選擇冷門號。然而，真實的冷門號碼難預知，對提高期望值有多少助益不得而知，何況大家搶冷門就變熱門了！因此，極不可能使期望值由負轉正。那麼，購買彩券的利基何在？大量簽注是否有利？

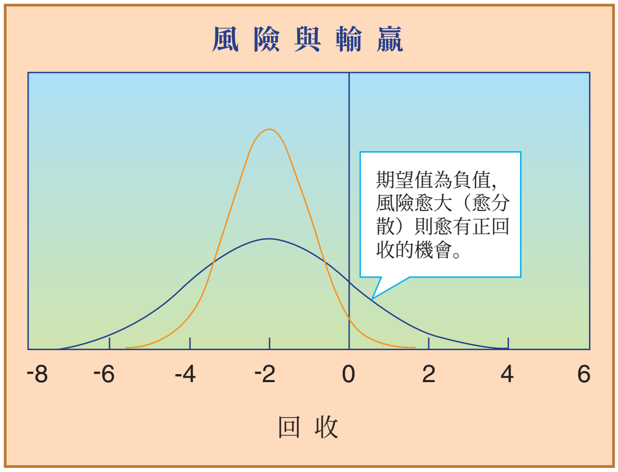
前面談期望值，結論是久賭必輸，大量投注則加速「必輸」的過程。細言之，則要依單期簽注策略來談。

如果每張彩券號碼都是任選，而且互不相干的，則結果等於一期當多期來玩。這好比丟硬幣，一個硬幣丟十次出現人像的次數，和一次丟十個硬幣而計算出現人像的個數，其機率上的性質是完全一樣的。在所有投注者都是隨機選號，而且沒有累積獎金的條件下，一期投注數萬元的，等於一次投注玩了數年，大數法則將發揮作用使這種簽注法輸的機會更高。只是簽一注輸的機率是97.1%，簽1,000注則至少有一個三獎或超過十個四獎以上才可能贏，而這樣的機會在千分之八以下！

由於樂透彩的獎金不是固定的，因此簽注模式會影響到期望值，結果是簽同一號碼組合則平均期

望值降低，甚至號碼不完全相同也會影響二獎以下的獎金。以頭獎而言，若本來中獎人數可用平均數μ的卜瓦松分布描述，則某人投注而中頭獎時期望分享頭獎獎金，簽二注與只簽一注相比，可能只增加分享8%（μ＝0.5時），也可能增加70%以上（μ＞6時）。但基於頭獎比二獎實際分享金額大了數十倍，因此除非μ很低，否則在下文考慮的情形下，還是划算。另一方面，顯然簽太多注是不划算的！

現在考慮的問題是一期要簽多注的策略問題。在期望值是負的情況下，風險愈大（投資學上的風險常以統計學上的「變異數」量度）愈有贏的機會，像前面所提隨機簽多注會降低贏的機率，就是在於風險降低。有些投注者專注於提高中獎機率，但中獎機率雖提高（甚至有人已找出只要110張樂透彩至少能中一張普獎的方法）並不等於提高回收！假設兩張樂透彩券號碼有k個相同，當k＝0時表示兩彩券號碼都不同，這時一張中獎則一張不中獎機會較大，這兩張彩券的回收將有負的關聯，但它們可提高中獎機率，降低風險。要提高風險，也就是兩張彩券要能有正的相關，那就要k愈大愈好，因為這時它們同時中大獎，或同時中獎，或同時不中獎的機會較高。雖然兩張彩券同號在一人獨得頭獎時變成浪費，但基於前段的分析知道：除非μ很小（投注數太少），否則多一張彩券即使是唯一頭獎號碼也可增加分享彩金成數，但投注太多則效益降低。總結來說，單期多注的策略是：二注同號，小規模包牌，過多無益。



最重要的是：不論如何用心，中獎機會低，贏錢機會更低。買張彩券有個夢無妨，切莫把老本放！ □

劉應興
成功大學統計學系