瘋狂大樂透

黃文璋
國立高雄大學應用數學系

1. 前言

台北銀行(以下簡稱北銀)發行的大樂透, 自第56至61期, 首度連續六期頭獎摃龜(摃龜通常表無人中頭獎), 投注因此大幅增溫。93年8月5日開獎的第62期, 總投注金額高達26.42億元, 創下大樂透上市以來的最高。頭獎獎金突破12.15億元, 由中獎二人均分。這是台灣彩券史上最高的獎金。第62期開獎前(見93年8月4日聯合報A5版), 北銀協理楊瑞東表示：

> 依統計來看, 大樂透"連七摃"的機率不到百分之二。換言之, 如本期投注量達十三億元以上, 且彩迷選號較平均分佈, 使得全部投注的號碼組合"覆蓋率"接近百分之百的話, 明晚開出頭獎的機率相當大。
>
> 大樂透"連六摃"機率低於百分之五, 結果上期還是出現"連六摃"。因此只要大樂透一千三百九十八萬多種號碼組合沒有全部被彩迷投注, 即使機率非常低, 也不能說完全沒有機會出現"連七摃"。

此講法引起公平會某位委員的關切, 認為北銀過度促銷。

北銀似乎是以銷售量來估計連七摃的機率。依其說法, 在開獎前二日(開獎前一日報上的新聞, 所以是開獎前二日講的), 他們就預測有九成八的機率, 第62期會開出頭獎。所依據的總投注金額為13億, 只有實際總投注金額的一半。宣稱有九成八的機率會開出頭獎, 當然會更鼓舞彩迷的投注。事實上, 由於各期開出的號碼為相互獨立(這是樂透彩吸引人的主要原因之一, 下一期那些號碼較易開出無法預知), 過去6期的開獎結果, 並不影響下一期會開出的頭獎號碼。而每期頭獎能否開出的機率, 與該期總投注數, 及彩迷的簽注行為有關, 北銀的講法是有問題的。不過如果連預估的總投注金額都只有實際的一半, 北銀各種預測之正確性, 是很令人存疑的。為了釐清一些概念, 並使一般人對於隨機性有較正確的了解, 本文將對彩券開獎的若干相關問題做探討。

2. 何謂大樂透

自民國九十一年一月起, 北銀發行公益彩券。其中的樂透彩, 由於可由購買者自選號碼, 也可由電腦代為選號, 有自主性, 也有方便性, 再加上高額獎金, 遂引起一陣風潮。當新鮮感逐漸失去後, 想再度引起購買熱潮, 自民國九十三年一月三日起, 北銀開始發行大樂透。為了區隔, 原先的樂透彩我們改以小樂透稱之。小樂透是由1至42的號碼中, 任選6碼; 大樂透則是由1至49的號碼中, 任選6碼。選號皆為不重複。

小樂透可能出現的號碼組合共有

$$\binom{42}{6} = \frac{42\cdot 41\cdot 40\cdot 39\cdot 38\cdot 37}{6!} = 5{,}245{,}786(\text{種})。 \tag{1}$$

大樂透可能出現的號碼組合共有

$$\binom{49}{6} = \frac{49\cdot 48\cdot 47\cdot 46\cdot 45\cdot 44}{6!} = 13{,}983{,}816(\text{種})。 \tag{2}$$

小樂透的中頭獎機率

$$\frac{1}{5{,}245{,}786}$$

已經很小了, 大樂透的中頭獎機率

$$\frac{1}{13,983,816}$$

又更小。小樂透我們曾經做了詳細的探討, 並對樂透彩開出的號碼, 及彩迷簽注的號碼, 是否符合隨機性, 給出統計分析, 可參考黃文璋(2003)第十九章"樂透彩開出號碼隨機性之檢定"一文, 及黃文璋與洪宛頻(2003)二文。底下便只介紹及討論大樂透。

大樂透每期除了由開獎機隨機地開出6個頭獎號碼, 還開出一個特別號。中獎方式見表1。

表1 大樂透之中獎方式

獎項	中獎方式
頭獎	6碼完全相同
貳獎	中5碼及特別號
參獎	中5碼
肆獎	中4碼及特別號
伍獎	中4碼
陸獎	中3碼及特別號
普獎	中3碼

表1中, 中頭獎表簽注的6個號碼, 與開出的6個頭獎號碼皆相同, 不計開出的順序; 中貳獎表簽注的6個號碼中, 有5個與頭獎號碼中的任5碼相同, 另一碼與特別號相同; 中參獎表簽注的6個號碼中, 有5個與頭獎號碼中的任5碼相同, 另一碼為剩餘的$42(=49-6-1)$個號碼中的任一個; 餘類推。

以第62期為例。當期頭獎號碼依序開出30, 21, 24, 14, 27, 13, 特別號20。由於順序無關緊要, 所以將頭獎號碼按小至大排列, 得13, 14, 21, 24, 27, 30。如果簽13, 20, 21, 24, 27, 30, 就中貳獎; 簽7, 14, 21, 24, 27, 30, 中參獎; 簽13, 14, 20, 27, 30, 47, 中肆獎; 簽5, 21, 24, 27, 30, 39, 中伍獎; 簽13, 20, 24, 28, 30, 45中陸獎; 簽14, 24, 30, 35, 42, 46, 中普獎。

表2給出各獎中獎機率, 顯然連要中個普獎都不是很容易。

表2 大樂透中獎機率

獎項	中獎機率
頭獎	$\frac{1}{\binom{49}{6}}=\frac{1}{13{,}983{,}816}$
貳獎	$\frac{\binom{6}{5}}{\binom{49}{6}}=\frac{6}{13{,}983{,}816}=\frac{1}{2{,}330{,}636}$
參獎	$\frac{\binom{6}{5}\cdot\binom{42}{1}}{\binom{49}{6}}=\frac{252}{13{,}983{,}816}\doteq\frac{1}{55{,}491.3}$
肆獎	$\frac{\binom{6}{4}\cdot\binom{42}{1}}{\binom{49}{6}}=\frac{630}{13{,}983{,}816}\doteq\frac{1}{22{,}196.5}$
伍獎	$\frac{\binom{6}{4}\cdot\binom{42}{2}}{\binom{49}{6}}=\frac{12{,}915}{13{,}983{,}816}\doteq\frac{1}{1{,}082.8}$
陸獎	$\frac{\binom{6}{3}\cdot\binom{42}{2}}{\binom{49}{6}}=\frac{17{,}220}{13{,}983{,}816}\doteq\frac{1}{812.1}$
普獎	$\frac{\binom{6}{3}\cdot\binom{42}{3}}{\binom{49}{6}}=\frac{229{,}600}{13{,}983{,}816}\doteq\frac{1}{60.9}$

在表2裡, 我們將中獎機率皆化為分子為1的數。因一般人對那些介於0與1之間的機率值, 常不易理解其大小。以分子為1來表示, 就可知平均每簽多少注可中

獎。例如, 貳獎為平均簽2,330,636注可中一次。但要明白隨機性的涵意, 並非簽了2,330,636注就必能中貳獎。

大樂透每注50元, 每週一、四各開獎一次, 獎金由每期總投注金額中支付。既然名之為公益彩券, 當然要提撥一些收入做些公益, 再加上還有管銷及盈餘等支出, 總獎金只佔每期總投注金額之57%, 比小樂透的56%略高。陸獎每注固定獎金1,000元, 普獎每注固定獎金400元(在某特殊情況下, 普獎獎金可能少於400元)。頭獎至伍獎, 係將總獎金扣除陸獎及普獎之獎金總額後, 依表3之比率分配。若有數注同中某獎項(陸獎與普獎除外), 則由所有中獎者, 均分該獎項之獎金。

表3 大樂透頭獎至伍獎獎金分配比率

獎項	頭獎	貳獎	參獎	肆獎	伍獎
分配比率	56%	9%	10%	5%	20%

由於頭獎中獎機率很低, 常有無人中獎的情況發生(1至62期共有43期無人中頭獎), 這時就是"摃龜"。連續六期無人中頭獎便稱連六摃。北銀採累積獎金的方式。頭獎至伍獎, 若有某獎無人中, 則該獎之獎金累積納入次期該獎獎金。北銀大樂透的"特別約定"中說:

> 惟為降低賭風, 當頭獎獎金連續八期無人獲得時, 其累積之獎金將併入次期之總獎金中, 供所有獲獎者依獎金分配所述之方式, 由中獎人分享之(陸獎及普獎仍為原固定金額)。

北銀在商言商, 不斷設法刺激買氣, 有時還有意無意的講些不太正確的話, 當銷售量直直成長(就是賭風旺的時候), 北銀其實很高興, 不過還是要義正辭嚴的來這一套"惟為降低賭風, ···"。我們引一段報導(92年3月29日中國時報14版)大家便明白:

> 採用電腦選號可適度提升中獎率, 北銀評估暫停電腦選號主要是為了增加摃龜機會, 頭彩可以累積, 買氣自然上升。
>
> 楊瑞東進一步指出, 目前電腦選號比重約占六成, 六億元的銷售量等於有三億六千萬元採電腦選號。換算每二億六千三百萬元的銷售額就能開出一個頭獎, 與最近每期頭獎得主一到二名的實際情況相比, 就能證明電腦選號果然保證每期都能開出頭獎, 北銀樂透彩的銷售量就欲高不易。

前述報導中還指出, 因小樂透銷售額下降, 北銀曾考慮取消電腦選號, 後來因顧慮影響層面過大而未實施。另外, 報導中那句“就能證明電腦選號果然保證每期都能開出頭獎”也很有意思。看來隨機概念的缺乏是很普遍的。第4節我們會給出電腦選號損龜的機率(所以不“保證”每期都能開出頭獎)。

雖理論上, 每期各獎均可能無人中, 實際上, 不論大樂透或小樂透, 到目前為止, 只有頭獎及貳獎才曾無人中。當連續多期無人中頭獎, 此時頭獎累積的獎金愈來愈高, 民眾被挑起的期盼之心, 可以想象。就如第62期, 真是全民大瘋狂。開獎前, 當日下午6至7時, 全國各投注站, 一小時內的投注金額便達3.1億。當天傍晚, 若看到某店面門口擠滿人, 那大約就是投注站了。

因要先扣掉陸獎及普獎的獎金, 但每期各獎之中獎人數並不一定, 所以若只知當期總投注金額, 並無法獲知頭獎至伍獎各獎之獎金總額。由於投注使用電腦連線, 開獎後北銀在短時間內, 經電腦統計出陸獎及普獎之中獎注數, 就可分配頭獎至伍獎之獎金總額。然後依各獎簽中之注數, 便能決定每注之獎金。不過由表2之各獎中獎機率, 給定總投注數a, 可求出各獎中獎注數之期望值。即

$$\text{某獎中獎注數之期望值} = a \times \text{某獎中獎機率} = \frac{b}{50} \times \text{某獎中獎機率}, \tag{3}$$

其中

$$b = 50a, \tag{4}$$

表當期之總投注金額。利用(3)式, 便可求出陸獎及普獎獎金之期望值, 再利用表3, 各獎之期望獎金便可求出了, 見表4。

表4　大樂透每期各獎期望獎金總額之近似值

獎項	期望獎金總額
頭獎	$11.592551229a \doteq 0.231851025b$
貳獎	$1.863088590a \doteq 0.037261772b$
參獎	$2.070098434a \doteq 0.041401969b$
肆獎	$1.035049217a \doteq 0.020700984b$
伍獎	$4.140196867a \doteq 0.082803937b$
陸獎	$1.231423526a \doteq 0.024628471b$
普獎	$6.567592137a \doteq 0.131351843b$

表4是如何求出? 以$C_1,\cdots,C_7$, 分別表頭獎至普獎之期望獎金總額。又令

$$A = 1,000 \times \frac{17{,}220}{13{,}983{,}816},$$
$$B = 400 \times \frac{229{,}600}{13{,}983{,}816}。$$

A,B分別表每一注會中陸獎及普獎之期望獎金。則

$$C_6 = Aa,$$
$$C_7 = Ba,$$

且

$$C_1 = 0.56a(28.5 - A - B),$$
$$C_2 = 0.09a(28.5 - A - B),$$
$$C_3 = 0.10a(28.5 - A - B),$$
$$C_4 = 0.05a(28.5 - A - B),$$
$$C_5 = 0.20a(28.5 - A - B)。$$

其中$28.5 = 50 \times 0.57$, 乃一注大樂透之期望獎金; $0.56,\cdots,0.20$, 乃頭獎至伍獎各獎獎金分配比率, 見表3。將上述$C_1,\cdots,C_7$, 分別除以50, 便得表4中, 以總投注金額b來表示之各獎期望獎金總額。

將表4中以a表示之各獎期望獎金總額, 分別除以當期總獎金$28.5a$, 便得表5每期各獎期望獎金總額, 在當期總獎金$28.5a$中, 所佔比例。

由頭獎期望獎金總額最高, 佔總獎金之40%以上, 及普獎期望獎金總額次高, 佔總獎金之23%以上, 可看出設計者之用心。一方面希望以高獎金吸引人, 一方面又使購買者偶而可中個400元的小獎, 因而能持續沈浸在購買彩券的樂趣中。

我們以第1期為例。該期大樂透總投注金額

$$b = 290{,}517{,}700(\text{元}),$$

總投注數

$$a = \frac{b}{50} = 5{,}810{,}354(\text{注})。$$

表5 大樂透每期各獎期望獎金總額佔總獎金之比例

獎項	期望獎金總額佔總獎金之比例
頭獎	40.6756%
貳獎	6.5372%
參獎	7.2635%
肆獎	3.6318%
伍獎	14.5270%
陸獎	4.3208%
普獎	23.0442%

利用北銀公布的資料、表2及表4, 表6給出各獎期望與實際情況的資料。由於本期陸獎及普獎實際中獎注數較預期的多, 頭獎至伍獎之實際獎金總額較預期的少。又期望獎金總額與實際獎金總額略有出入, 乃因採小數以下四捨五入造成的。

表6 第1期大樂透期望與實際情況的資料

獎項	期望中獎注數	實際中獎注數	期望獎金總額	實際獎金總額
頭獎	0.4	0	67,356,826	58,146,978
貳獎	2.5	7	10,825,204	9,345,050
參獎	104.7	178	12,028,005	10,383,389
肆獎	261.8	510	6,014,002	5,191,694
伍獎	5,366.3	8,119	24,056,009	20,766,778
陸獎	7,155.0	11,342	7,155,007	11,342,000
普獎	95,400.1	126,048	38,160,035	50,419,200
合計	108,290.8	146,204	165,595,088	165,595,089

表7　第62期大樂透期望與實際情況的資料

獎項	期望中獎注數	實際中獎注數	期望獎金總額	實際獎金總額
頭獎	3.8	2	612,714,615	1,215,450,076*
貳獎	22.7	22	98,471,992	100,474,965
參獎	952.5	796	109,413,324	111,638,850
肆獎	2,381.2	2,172	54,706,662	55,819,425
伍獎	48,814.4	43,885	218,826,648	223,277,701
陸獎	65,085.9	58,934	65,085,862	58,934,000
普獎	867,811.5	827,553	347,124,599	331,021,200
合計	985,072	933,364	1,506,343,702	2,096,616,217

∗ =625,177,561元(當期分配獎金)+590,272,515元(前6期累積獎金)

再以這引人注目的第62期為例。該期總投注金額

$$b = 2{,}642{,}708{,}250(\text{元}),$$

總投注數

$$a = \frac{b}{50} = 52{,}854{,}165(\text{注})。$$

表7給出本期期望與實際情況的資料。由各獎實際中獎注數均較期望中獎注數低, 可見對這一全台矚目, 人人想要一搏的第62期, 民眾雖千挑萬選, 求神問卜, 結果卻不比自己隨機選號, 或交給電腦選號, 更易中獎。要注意的是, 若只簽一注, 不論電腦選號或自選號碼, 中獎機率都是一樣的。但若簽多注, 或者多人簽注, 電腦選號或自選號, 中獎的機率就不一樣了。此因不論一人簽多注, 或多人簽注, 號碼往往有集中的現象, 此點稍後我們會再說明。

將表2中, 各獎中獎機率相加(就是所謂“中獎機率”), 得

$$\frac{260{,}624}{13{,}983{,}816} \doteq 1.864\%, \tag{5}$$

我們又定義

$$\text{中獎率} = \frac{\text{實際中獎注數}}{\text{該期總投注數}}。 \tag{6}$$

第62期的中獎率為

$$\frac{933{,}364}{52{,}854{,}165} \doteq 1.766\% < \text{中獎機率}。$$

當陸獎及普獎之實際中獎注數較期望值小, 頭獎至伍獎各獎能分配的獎金, 就比期望值大。再因有前6期損龜累積的獎金, 使第62期的頭獎獎金高達1,215,450,076元。而實際中獎人數2人, 又小於期望中獎人數3.8人, 所以此二位幸運的頭獎得主, 每人分得比預期多的607,725,038元。扣掉20%的稅金, 每人實際所得4.8618億元。

12億元究竟有多少, 並不太容易想象。12億元相當於120萬張千元大鈔。一束1,000張千元大鈔(即一百萬元)約有10公分高, 120萬張就有

$$\frac{1{,}200{,}000}{1{,}000} = 1{,}200(\text{個})$$

10公分, 即120公尺。若一層樓高3公尺, 這筆巨款疊起來便有40層樓那麼高。眞是令人仰之彌高的獎金。

雖然有人質疑, 樂透彩打著公益的名號, 但購買者中有很高比例是低收入族, 屢買不中, 可能會製造一些社會問題。衛道之士更覺政府不該帶頭做莊, 鼓勵賭博。但社會上有些人具有賭徒性格並不是太壞的事, 見黃文璋(2003)第七章“賭國風雲”一文。另一方面, 有什麼方法可以用很少的錢, 而有機會在短時間內成為億萬富翁? 簽注樂透彩是一個人人都可以嘗試的簡單方法。花50元, 就能享受到一注在手, 希望無窮的樂趣。

北銀公益彩券的網址(http://www.roclotto.com.tw/)中說:

大樂透是使用電腦連線投注之樂透型機率遊戲。

既然是機率遊戲, 那就要相信機率理論。機率理論告訴我們, 在賭局裡, 想跟機率拼的人, 口袋裡的錢將隨風飄去, 很快便可贏得公益名。人定勝天之想法不對, 挑戰機率也非明智之舉, 這方面的一些說明可見黃文璋(2003)第十五章“挑戰機率”一文。

簽大樂透, 每花100元, 平均只得回57元, 43元不見了。所以不能抱著投資的心理簽注, 否則損失慘重。例如, 投資100,000元, 一期後, 平均只剩57,000元。若將獎金全部繼續簽注, 如此一個月(8期)後, 平均只餘

$$100{,}000 \times 0.57^8 \doteq 1{,}114.29(\text{元}),$$

下降速度驚人。

虧本的生意為何有人要做? 連續六次摃龜為何讓全台彩迷瘋透? 當然不是為了1,000元或400元的陸獎及普獎。而是想中大獎, 特別是那累積七期的頭獎獎金。

3. 贏的策略

上節不是才說大樂透總獎金僅佔總投注金額的57%, 且告訴大家不要跟機率拼, 那為何還會有贏的策略呢?

首先要說明, 大樂透之總獎金只有總投注金額之57%, 加上還要扣稅, 所以對購買者而言, 當然是一"不公正的賭局"。但假若你就是想試試運氣, 底下給出幾個宜遵守的投注原則。

簽大樂透, 雖人人想中頭獎, 但中的機率卻很小, 每注不中頭獎的機率則很大:

$$1-\frac{1}{13{,}983{,}816}, \tag{7}$$

此值當然很接近1。那簽愈多, 不中頭獎之機率是否就愈小? 也就是中頭獎之機率是否就愈大? 也不一定, 如果簽100注同樣的號碼, 則中與不中頭獎的機率, 仍維持不變。為了提高中頭獎的機率, 簽注的第一個原則是, 如果簽多注, 要簽不一樣組合的號碼。而且簽注中儘量不要有超過三個號碼相同。此因大樂透除了陸獎及普獎外, 都是由簽中者均分該獎之獎金。簽中4碼以上, 便會得伍獎以上, 自己跟自己分獎金有什麼意思呢?

其次, 如果中大獎當然希望與你同中的人愈少愈好。試想第62期, 如果有1,000人同中頭獎, 每人只分得1,215,450元。一百二十多萬元自然仍是一筆不少的錢, 但與六億多元相比, 是天壤之別。所以簽注的第二個原則是要簽冷門的號碼, 使得如果中大獎, 最好是獨得, 或愈少人與你均分獎金。第62期頭獎期望注數為3.8, 而實際只有2注中獎。簽那組頭獎號碼的兩個人, 感覺當然很甜美。

北銀發行公益彩券兩年多來, 彩迷們每週要上幾次機率課。似乎彩迷的機率概念有些提高了, 不再對一些其實很容易發生的事件感到訝異, 也少有人仍懷疑開獎的公平性。但由於巨額獎金的吸引力, 而且這麼多組號碼, 也真不知該選那一組。於是媒體上, 仍常有人在報明牌。難免就有些人抱著姑且相信的心理簽注。

有人會以近期發生的重大事件(如颱風日期等)中的數字簽注。求明牌的方式雖有異, 但都屬不智之舉。這些"明牌"的結局, 當然大部分是沒開不出來。只是若開出來, 可能會讓很多人欲哭無淚。德國的樂透彩亦是49取6。1993年10月16日那一期, 總銷售注數為6,803,090, Henze(1997)一文列出當期最熱門的20組號碼(亦見黃文璋(2003)第八章"隨機與密碼"一文), 其中7, 13, 19, 25, 31, 37那組, 共賣

出4,004注。不論獎金再高, 要4,004人均分, 大概都不會是件太高興的事。大家是否注意到7, 13, 19, 25, 31, 37, 形成一組等差數列, 第一個數字還是幸運的7。有人以為號碼既然是隨機地產生, 所以選號應均勻地分開。可惜有此想法的人不少, 造成等差數列往往多人簽注。看來等差數列是不能簽的。有人以為1, 2, 3, 4, 5, 6這組應少有人想到吧! 只是人同此心, 這是排名第四的一組, 有3,249人簽。

讀者大約可以了解避免簽注熱門號碼的道理了, 媒體上出現的明牌更是萬萬不能簽: 你看到別人也看得到; 你會相信, 也會有一些人跟你一樣。只是熱門號碼容易知道, 而那些是冷門號碼呢?

北銀提出了"覆蓋率"一詞(見前言), 其定義應是

$$\text{覆蓋率} = \frac{\text{某期投注之總相異組合數}}{13{,}983{,}816}, \tag{8}$$

其中13,983,816是總共的號碼組合數。93年8月4日聯合報A5版指出, 第61期投注的覆蓋率是67%(93年8月6日中國時報A1版則說是66%)。北銀掌握每期無人投注的號碼組合, 因此冷門號碼亦為其所掌握。可惜這份在北銀內部中很可能不是秘密的致富號碼組合名單, 儘管北銀一再提到覆蓋率, 我們一般人卻無從取得。

另一簽注的原則是, 若前期摃龜, 下一期便是較好的簽注時機。連續摃龜愈多期, 時機便愈佳。各位看, 第62期由於有前6期之累積獎金, 因此

$$\frac{\text{實際總獎金}}{\text{總投注金額}} = \frac{2{,}096{,}616{,}217}{2{,}642{,}708{,}250} \doteq 79.34\%,$$

遠大於原設訂的57%。

第62期獎金雖高, 若有人異想天開, 包所有的牌, 即每一號碼組合各簽一注, 那會有什麼後果? 答案是會中不少獎。表8給出此人會中的各獎注數(由表2可得)。在其他一切投注情況不變下(所以此人之外的中獎人數仍如表7所給), 總投注數因此人的包牌而增加, 各獎總獎金也隨之提高。則該投注者之期望總所得(採北銀算法, 小數無條件捨去)為

$$\begin{aligned} & \frac{1{,}377{,}558{,}180}{3} + 6 \times \frac{126{,}528{,}053}{28} + 252 \times \frac{140{,}586{,}726}{1{,}048} + 630 \times \frac{70{,}293{,}363}{2{,}802} \\ & \quad + 12{,}915 \times \frac{281{,}173{,}452}{56{,}800} + 17{,}220 \times 1{,}000 + 229{,}600 \times 400 \\ \doteq & \quad 708{,}897{,}687(\text{元}), \end{aligned}$$

其中各獎中獎注數, 為原先的中獎注數加上此人之中獎注數。期望總所得略大於總投資的

$$50 \times 13{,}983{,}816 = 699{,}190{,}800(\text{元})。$$

表8 採全部包牌會中的各獎注數

獎項	頭獎	貳獎	參獎	肆獎	伍獎	陸獎	普獎
中獎注數	1	6	252	630	12,915	17,220	229,600

表面看是一好的策略, 但因還要繳稅, 就划不來了 。況且要在數天內簽一千三百多萬注, 耗費的人力也應不少 。

顯然在連六摃之下, 包全部的牌, 仍不是上策 。

現在假設在全部的13,983,816種組合中, 有1,000,000種組合名單, 極少有人簽注, 就假設沒人簽注, 而幸運的你擁有此名單 。則每簽一注會得多少獎金呢? 當然不一定, 是一隨機變數 。以第62期為例 。該期頭獎期望獎金總額為612,714,615元 。則每簽一注(由於這1,000,000組是極度冷門的號碼, 假設若中頭獎為你獨得), 光是頭獎獎金之期望值便有

$$612{,}714{,}615 \times \frac{1}{13{,}983{,}816} \doteq 43.816(\text{元})\text{ 。}$$

說不定還會中一些其他的獎, 期望所得可能會超過一注的50元 。

不過這種連七摃的機會並不是太多, 況且獎金還要扣稅 。所以以上討論, 只是強調如果你真的要簽注, 冷門號碼之重要性 。對此題材有興趣的讀者, 可參考Chernoff(1981)一文, 該文算是很早引進避免簽注熱門號碼, 而使期望淨所得為正之概念者 。

4. 摃龜機率

在93年8月4日聯合報A5版的一則新聞中提到:

> 北銀內部統計顯示, 大樂透“連五摃”機率低於百分之十, “連六摃”(機率)低於百分之四, “連七摃”機率低於百分之二, “連八摃”的機率更是微乎其微, 低於百分之一 。

我們不知北銀內部是如何統計出這些機率的, 新聞中並未說明 。1至61期連摃的統計見表9 。其中連1摃表只有一期摃龜, 下一期便有人中頭獎了 。數據如此少, 實在很難建立機率模型, 而得到連續幾摃之機率 。不過我們猜想北銀並不是經由建立機率模型, 而得到這些連續幾摃的機率 。他們可能有某種“方式”來統計, 只是可不可靠呢? 我們知道對同一事件, 不同單位公佈的民調結果常不太一致, 即使事後證實當初所預測太離譜, 只要以“民意如流水”一句帶過, 別人也無可奈何 。有個笑

表9 1至61期連摃統計

連1摃	連2摃	連3摃	連4摃	連5摃	連6摃
5	4	4	3	0	1

話, 某教授對氣象局宣稱明日降雨機率為6成, 或7成等, 常感到很納悶, 好奇地問他們是如何得到這些機率的? 氣象局局長說, 我們局裡共有10個人, 若有7人, 認為明天會下雨, 我們就宣佈明天降雨的機率為7成。

每一期摃龜之機率, 並無法預知。因不知該期究竟有多少總投注數, 每一期總投注數之變異可是很大的。即使知道總投注數, 如果不知其中有多少注是電腦選號, 有多少注是彩迷自選號碼(以及如何選號)也無法求出摃龜之機率。假設都是電腦"隨機選號", 或即使彩迷自選號碼亦符合隨機性。也就是每一簽注, 都有

$$\frac{1}{13{,}983{,}816}$$

的機率, 選中任一組號碼, 且各次選號為相互獨立。則只要知道總投注數, 便可求出當期摃龜機率。而將接連各期摃龜機率相乘, 就是那連續幾期摃龜的機率。此因各期之開獎及投注也假設相互獨立。表10給出自56期至62期, 大樂透之總投注數, 及各期摃龜機率。其中若某期之總投注數為m, 則因每一注不中頭獎之機率如(7)式所給, 故當期摃龜(即m注皆不中)之機率為

$$(1-\frac{1}{13{,}983{,}816})^m \text{。} \tag{9}$$

不論總投注數m多大, 上述摃龜機率永遠為正。這是隨機選號下的結果。當然m愈大, 摃龜機率愈小。

在知道總投注數及隨機選號的假設下, 56至60期皆摃龜(連5摃)之機率約為

$$0.7452 \times 0.7196 \times 0.6550 \times 0.6466 \times 0.5413 \doteq 0.1229,$$

56至61期皆摃龜(連6摃)之機率約為

$$0.03129,$$

56至62期皆摃龜(連7摃)之機率約為

$$0.0007 \text{。}$$

表10　第56期至62期各期大樂透之總投注數及摃龜機率

期別	總投注數	摃龜機率
56	4,112,586	0.7452
57	4,601,198	0.7196
58	5,917,595	0.6550
59	6,097,816	0.6466
60	8,582,570	0.5413
61	19,135,676	0.2545
62	52,854,165	0.0228

因連7摃實際並未發生, 所以我們就停止求連8摃之機率。這是大樂透發行以來, 第一次連6摃發生, 表10中各期總投注數逐漸增加, 尤其連6摃之後的第62期, 總投注數超過前面6期總投注數之和。如果第62期又摃龜, 有誰能預測第63期的總投注數呢? 恐怕是很難吧! 如果下一回又有連6摃發生, 你覺得民眾仍會如此瘋狂嗎? 可能會, 也可能不會。會與不會, 都有相當的理由支持。在數據不足的情況下, 我們實在不敢預測未來連續幾摃之機率。北銀宣稱的連續幾摃的機率, 既不知是指這回, 以後, 甚或任一次。無論如何, 他們對宣佈統計的結果, 顯然並不很嚴謹。

但這還不是北銀說法最可議之處。機率值會變, 是隨機現象裡的特性。簡單地講, 機率裡有所謂條件機率: 給定某條件後, 事件之機率可能會改變。大家知道一家庭有三個小孩, 皆是男孩之機率為1/8。這是基於生男生女之機率皆為1/2, 且各次生產為相互獨立之假設。如果某家庭頭兩胎已經都是男孩了, 則他們家頭三胎皆是男孩之機率仍為1/8嗎? 顯然不是, 應與一胎為男孩之機率相同, 即1/2。

連6摃已經發生了, 即使62期總投注金額達13億元(如前言新聞中的數據), 即26,000,000注, 則依(9)式, 第62期摃龜(如此連7摃就發生了)之機率為

$$(1 - \frac{1}{13{,}983{,}816})^{26{,}000{,}000} \doteq 0.1558 \text{。} \tag{10}$$

絕非如北銀所說的摃龜機率不到百分之二。至於他們所說

> 如$\cdots$, 且彩迷選號較平均分佈,$\cdots$, 明晚開出頭獎的機率相當大。

也是不明所以。彩迷又無法串通, 選號如何能“平均分佈”? 平均分佈又是什麼意思呢? 至於說“覆蓋率接近百分之百的話, 明晚開出頭獎的機率相當大”, 這句話就彷

佛“考滿分就是第一名”是白講的 。下一節我們將看到在隨機選號下, 選中的號碼是很難“平均分佈”的 。

5. 覆蓋率

每一期之覆蓋率如何計算? 投注再多, 也可能只集中在某幾組號碼, 使得覆蓋率很低 。每一期的覆蓋率(為一隨機變數)取的值可能為

$$\frac{1}{13{,}983{,}816}, \frac{2}{13{,}983{,}816}, \ldots, \frac{13{,}983{,}815}{13{,}983{,}816}, 1 \text{ 。}$$

假設是隨機選號(否則就無法計算覆蓋率了), 且給定總投注數, 則可求出覆蓋率之機率分佈 。只是那些繁瑣的排列組合, 你絕不會想動筆去算的 。

我們給一較簡單的例子 。

例1. 假設有2個箱子, 將2個球分別隨機地放進其中任一箱 。即每一個球均有1/2的機率放進任何一箱中 。每箱中各恰有一球的機率為

$$\frac{2!}{2^2} = \frac{2}{4} = \frac{1}{2} \text{ 。}$$

如果是10個球隨機地放進10個箱子中, 則每箱中各恰有一球的機率為

$$p = \frac{10!}{10^{10}} = \frac{3{,}628{,}800}{10^{10}} = 0.00036288 \text{ 。} \tag{11}$$

此機率可說是非常地小 。一般而言, n個球隨機地放進n個箱子中, 則每箱中各恰有一球的機率為$n!/n^n$, 此值隨著n之增大而漸減 。

上述現象, 可能違反一般人的直觀 。n個球隨機地放進n個箱子中, 每個箱子中各恰有一球(平均分佈?), 似乎看起來才較符合隨機性的, 結果卻是極不容易發生 。甚至由底下的表11及表12, 球集中在某幾個箱子中, 都遠較各箱中有一樣多的球容易發生 。

以10個球10個箱子為例 。我們列出恰有i個空箱之機率在表11, 其中當然符合

$$\sum_{i=0}^{9} P(i\text{個空箱}) = 1 \text{ 。}$$

表11　10個球隨機地放進10個箱子中, 空箱數之機率分佈

空箱數	機率
0	p
1	$45p$
2	$375p = (315 + 60)p$
3	$980p = (35 + 420 + 525)p$
4	$(7609/8)p = (21/2 + 525/4 + 175/2 + 525 + 1575/8)p$
5	$(2835/8)p = (7/4 + 21 + 35 + 315/4 + 105 + 105 + 63/8)p$
6	$(6821/144)p = (1/6 + 7/4 + 7/2 + 35/16 + 21/4 + 35/2 + 35/9 + 35/8 + 35/4)p$
7	$(311/168)p = (1/112 + 1/14 + 1/6 + 1/4 + 1/8 + 1/2 + 5/16 + 5/12)p$
8	$(73/5760)p = (1/4032 + 1/896 + 1/336 + 1/192 + 1/320)p$
9	$p/9!$

我們再列出一些事件之機率在表12。表11及表12中之機率, 讀者不妨自行驗證。我們只給出1箱中有2球, 8箱中各有1球, 另1空箱之機率的算式如下:

$$\frac{\binom{10}{1}\binom{9}{8}\binom{10}{2}\cdot 8!}{10^{10}} = 45\cdot\frac{10!}{10^{10}} = 45p\ 。$$

表12 10個球隨機地放進10個箱子中, 一些事件之機率

事件	機率
1箱中有2球, 8箱中各有1球, 另1空箱	$45p$
1箱中有3球, 7箱中各有1球, 另2空箱	$60p$
2箱中各有3球, 2箱中各有2球, 另6空箱	$(35/4)p$
2箱中各有4球, 2箱中各有1球, 另6空箱	$(35/16)p$
1箱中有1球, 1箱中有2球, 1箱中有3球, 1箱中有4球, 另6空箱	$(35/2)p$
1箱中有6球, 1箱中有2球, 2箱中各有1球, 另6空箱	$(7/4)p$
5箱中各有2球, 另5空箱	$(63/8)p$

由表11, 即得10個球隨機地放進10個箱子中, 覆蓋率的機率分佈:

$$P(\text{覆蓋率} = \frac{i}{10}) = P(10 - i\text{空箱}), \quad i = 1, \cdots, 10 \text{。} \tag{12}$$

例如,

$$\begin{aligned} &P(\text{覆蓋率} = 0.7) = P(3\text{空箱}) = 980p \\ &= \frac{980 \cdot 10!}{10^{10}} \doteq 0.3556 \text{。} \end{aligned}$$

設有i空箱, 令

$$\text{空箱率} = \frac{i}{10} \text{。} \tag{13}$$

則

$$\text{覆蓋率} = 1 - \text{空箱率},$$

因此(將上式兩側取期望值)

$$E(\text{覆蓋率}) = 1 - E(\text{空箱率}) \text{。} \tag{14}$$

例1是以球為例, 不難看出, 自1至n的n個號碼中, 隨機選號m次, 是類似的問題。而只要將大樂透的所有組合號碼, 自1至13,983,816編號(如令{1, 2, 3, 4, 5, 6}為1, {1, 2, 3, 4, 5, 7}為2, 餘類推), 則當某期總銷售注數為m, 隨機選號, 就相當於自1至n =13,983,816個號碼, 隨機選號m次。

由例1, 可看出要計算覆蓋率之機率分佈是很辛苦的。但給出機率分佈似無多大意義, 期望值在這裡反而較有意義:

$$\begin{aligned} &E(\text{空箱率}) \\ &= \frac{1}{10}p(1 \cdot 45 + 2 \cdot 375 + 3 \cdot 980 + 4 \cdot \frac{7{,}609}{8} + 5 \cdot \frac{2{,}835}{8} + 6 \cdot \frac{6{,}821}{144} \\ &\quad + 7 \cdot \frac{311}{168} + 8 \cdot \frac{73}{5{,}760} + 9 \cdot \frac{1}{9!}) \\ &\doteq 0.348678, \end{aligned}$$

其次若自10個號碼中, 隨機抽出1個號碼, 就稱做頭獎號碼, 則損龜(即選號皆為其餘9個號碼)之機率為

$$(1 - \frac{1}{10})^{10} \doteq 0.348678,$$

期望空箱率與損龜機率驚人的吻合, 會是巧合嗎? 利用下例便可回答此一問題。

例2. 自1至n的n個號碼, 隨機選號m次。令N表沒被選中的號碼個數。則
(i) $E(N)=n(1-\frac{1}{n})^m$;
(ii) $E(N/n)=(1-\frac{1}{n})^m$。
證明. 首先N可表示為n個指示函數(indicator function)之和:

$$N=I_1+\cdots+I_n, \tag{15}$$

其中$I_i=1$, 若第i個號碼沒被選中, $I_i=0$, 若第i個被選中。利用期望值的線性性質, 得

$$E(N)=E(I_1+\cdots+I_n)=E(I_1)+\cdots+E(I_n)\text{ 。}$$

由於1至n中的第i個號碼沒被選中, 表m次選號皆只能選其餘的$n-1$個號碼。每次要選中其餘$n-1$個號碼之機率為$1-1/n$, 又各次選號為相互獨立。故

$$P(I_i=1)=P(\text{第}i\text{個號碼沒被選中})=(1-\frac{1}{n})^m, \tag{16}$$

且

$$P(I_i=0)=P(\text{第}i\text{個號碼被選中})=1-(1-\frac{1}{n})^m, \tag{17}$$

故

$$E(I_i)=1\cdot P(I_i=1)+0\cdot P(I_i=0)=P(I_i=1)=(1-\frac{1}{n})^m\text{ 。}$$

由此得

$$E(N)=n\cdot(1-\frac{1}{n})^m\text{ 。} \tag{18}$$

得證(i)。由於

$$E(\frac{N}{n})=\frac{E(N)}{n}=\frac{n(1-\frac{1}{n})^m}{n}=(1-\frac{1}{n})^m,$$

得證(ii)。

N/n表沒有被選中的號碼佔全部號碼之比例。上例給出其期望值

$$E(\frac{N}{n})=(1-\frac{1}{n})^m, \tag{19}$$

恰為某一特定號碼沒被選中之機率。利用指示函數, 沒有求出複雜的N之機率分佈, 我們巧妙地求出$E(N)$及$E(N/n)$, 因而得到覆蓋率

$$\frac{n-N}{n}=1-\frac{N}{n}$$

之期望值

$$E(\frac{n-N}{n})=1-E(\frac{N}{n})=1-(1-\frac{1}{n})^m \text{。} \tag{20}$$

即得證

$$E(\text{覆蓋率})=1-P(\text{摃龜}) \text{。} \tag{21}$$

或等價地

$$P(\text{摃龜})=1-E(\text{覆蓋率}) \text{。} \tag{22}$$

上式直觀上是對的。開出的頭獎號碼, 若未落在號碼之覆蓋處, 就是摃龜, 反之亦然。我們曾以第62期的總投注數52,854,165, 經由摸擬30次, 所得的覆蓋率, 到小數第3位, 都與

$$1-P(\text{摃龜})=1-(1-\frac{1}{13{,}983{,}816})^{52{,}854{,}165}\doteq 1-0.022830581\doteq 0.977169418$$

一致。

註1. N之變異數亦可求出, 也是利用(15)式對N的表示法, 其推導過程在此略去。

$$\mathrm{Var}(N)=n(1-\frac{1}{n})^m+(n^2-n)(1-\frac{2}{n})^m-n^2(1-\frac{1}{n})^{2m} \text{。} \tag{23}$$

註2. 依公式(21), 由表10, 在第61期

$$E(\text{覆蓋率})=1-P(\text{摃龜})\doteq 1-0.2545=0.7455 \text{。}$$

實際覆蓋率67%, 比此期望值低。這顯然是因自選號碼造成的號碼較集中簽注的現象。表面上看來, 74.55%與67%之差異似乎不大。不過在如此大的樣本下, 這種差異算是很大的(由覆蓋率之變異數極小亦可得知, 見習題)。統計學裡的假設檢定(testing hypothesis)裡, 有較科學的方法, 來判定此結果顯示彩迷選號不符合隨機性。不過由北銀強調

投注量達十三億以上, $\cdots$, 使得$\cdots$ "覆蓋率"接近百分之百$\cdots$,

顯然他們對覆蓋率之大小毫無概念。利用(10)及(21)式, 當總投注數達十三億, 期望覆蓋率約為

$$1 - 0.1558 = 0.8442,$$

在此情況下, 覆蓋率並不易接近100%。如果第61期的覆蓋率為75%, 說不定北銀仍會以為很低呢!

6. 結語

公益彩券上市以來, 帶給社會不小的衝擊, 彩券影響及牽動著不少人的生活。大樂透第62期雖有極高的投注金額, 第63期又回復到常態, 總投注金額僅有251,080,100元, 不到第62期的十分之一。放鬆心情下, 不過度熱中追求頭獎, 中獎率反而高達2.27%, 是自第38期以來最高。大部分的時候, 一般民眾似乎了解, 不必對大樂透太寄予厚望, 此態度是正確的。媒體上充滿著各種令人緊張或匪夷所思的消息, 報報明牌也不是太嚴肅的事, 就當做飯後茶餘。只是主事者即使有業績壓力, 也不宜提供不正確或似是而非的訊息。

習題

1. 大樂透第62期, 若某人包所有的牌, 在其他一切投注情況不變下, 試驗證此人之期望總所得確為708,897,687元。
2. 在第3節裡大樂透有1,000,000種組合名單無人簽注之例, 試說明對第62期, 可否求出每簽一注此份名單中的號碼之期望獎金。
3. 試以第62期大樂透的總投注數52,854,165注, 模擬30次覆蓋率。
4. 在例2裡,

   (i) 試求$\mathrm{Cov}(I_i, I_j)$;

   (ii) 設$m, n \to \infty$, 且$m/n \to a < \infty$, 試求此時之$\lim \mathrm{Cov}(I_i, I_j)$。

   (解. (i) $i = j$時, $\mathrm{Cov}(I_i, I_j) = (1-1/n)^m - (1-1/n)^{2m}$, $i \neq j$時, $\mathrm{Cov}(I_i, I_j) = (1-2/n)^m - (1-1/n)^{2m}$。(ii) $i = j$時, $\mathrm{Cov}(I_i, I_j) \to e^{-a} - e^{-2a}$, $i \neq j$時, $\mathrm{Cov}(I_i, I_j) \to 0$)

5. 試推導出(23)式。
6. 試利用(23)式, 對第61期大樂透之投注數, 給出覆蓋率之變異數及標準差(解. 變異數$\doteq 7.229435 \cdot 10^{-9}$, 標準差$\doteq 8.502609 \cdot 10^{-5}$, 此為計算到小數30位所得之近似值)。

參考文獻

1. 黃文璋(2003). 隨機思考論。華泰文化事業股份有限公司, 台北。
2. 黃文璋、洪宛頻(2003). 簽注之隨機性探討。中國統計學報, 41, 475-492。
3. Chernoff, H. (1981). How to beat the Massachusetts numbers game. *Mathematical Intelligence* **3**, 166-172.
4. Henze, N. (1997). A statistical and probabilistic analysis of popular lottery tickets. *Statistica Neerlandica*, Vol. **51**, No.2, 155-163.