

NCAA籃球賽的預測

黃文璋

國立高雄大學應用數學系

2014年美國NCAA(National Collegiate Athletic Association)的男子籃球季後終於結束了。決賽於美國東部時間4月7日晚上9點開打,由康乃迪克大學(University of Connecticut)贏得年度總冠軍。這是該校隊史上,第四度登上NCAA男籃總冠軍寶座,上一次是2011年。

每年依例行賽的戰績全美排名,從幾百所大學裡,挑出68支勁旅,參與角逐這項長達3週的季後賽。按照排名,依南、西、中西、東的順序,大致以S型的方式,將球隊分入4區中,使各區間的球隊實力,差異不會太大。先淘汰4隊,剩64隊。每區的16支球隊,從第1排到第16種子。捉對廝殺,比賽採1戰單淘汰制。又依慣例,保障高種子球隊,即在每區裡,排名第1跟第16打,第2跟第15打,⋯,第8跟第9打。第1輪賽完,每區獲勝的8隊晉級,又是捉對廝殺。4輪過後,產生分區冠軍。4強再度捉對廝殺,最後產生總冠軍。過關斬將,連贏6場,就風光成為年度總冠軍。

不論整個例行賽的成績如何威震八方,季後賽只要輸1場就得打道回府,明年再來。像是例行賽34戰全勝,有史以來第6支不敗隊伍,中西區第1種子的威奇塔州立大學(Wichita State University),在第一輪還能談笑用兵,以64比37輕鬆擊敗對手,但第二輪便戰況激烈,最後76比78,以兩分之差,輸給該區第8種子,讓一堆人滿地找眼鏡。所以63場季後賽,每場都是殊死戰,不容絲毫閃失,各隊均相當拼命。不像NBA職籃,季

後賽皆採7戰4勝淘汰制，因此有時即使輸一、兩場，振奮一番後，仍可扳回。NCAA參賽的都是大學生，得顧慮到課業，無法打太多場，這可能是季後賽採1戰單淘汰賽制的原因之一。也導致季後賽常是高潮迭起，精彩萬分。

來個腦筋急轉彎，整個季後賽，64隊共須賽幾場？63場，因除了總冠軍外，每隊都恰好輸1場。不區分3、4名嗎？不必了。只有第1，連誰是第2都沒人在乎了，何況是第3？

每年NCAA男子籃球季後賽是從3月中旬開始(今年是3月18日)，至4月初結束(今年是4月7日)。由於備受全美眾多籃球愛好者之矚目，素來有“三月瘋”(March Madness)之稱。被一般人封為股神的巴菲特(Warren Edward Buffett, 1930-)，向來不甘寂寞，驚人之舉從來不少，他也為今年的三月瘋加把火。1月22日他宣佈，自3月3日開始，至3月16日截止，任何人都可以上網填寫，預測全部63場比賽的結果。他會邀請將決賽前那62場比賽的輸贏都猜中者，一起到現場觀看決賽。今年決賽，球場湧進7萬9千多名觀眾，可說一票難求。當然以巴菲特之能，要拿到什麼位子的票都有辦法。神算子若連決賽誰勝都猜對，便可獲得巴菲特提供的獎金10億美元。如果有不只一位63場全猜中者，獎金就均分。10億美元，約300億台幣，看起來極度誘人。天下還有比這更快速的致富方式嗎？又毫無損失，怎可不參加？但立刻有專家潑你冷水，巴菲特並非你想像的那麼樂善好施。他所設的NCAA賭盤，中獎機率不到900萬兆分之1。

900萬兆分之1是怎麼算出來的？且此值究竟是多小？

由於有63場比賽，若隨機猜，每場猜對的機率以1/2計，且各場比賽的結果假設相互獨立，則63場皆猜對的機率為 $(1/2)^{63}$ ，即

$$\frac{1}{9,223,372,036,854,775,808}。$$

簡單一點講，約為 $9.22 \cdot 10^{18}$ 分之1。1兆是10的12次方，而10的18次方比1兆多6個0，即為100萬兆。由於全猜對的機率，約為922萬兆分之1，因此才說中獎機率不到900萬兆分之1。相對於49取6的大樂透，要中頭獎已經覺得幾乎不可能了，因機率為1/13,983,816，約 $1.398 \cdot 10^7$ 分之1，要得到巴菲

特的10億美元，比中大樂透真是難太多了。附帶一提，因1京為1萬兆，所以922萬兆分之1，即922京分之1。

2002年1月，台北銀行開始發行42取6的公益彩券。為了想了解人的天性是否具有隨機性，經由幾個不同的途徑，我收集了一些人們填寫的樂透彩號碼，以做隨機性檢定。又因想使創建中的高雄大學師生，增加若干生活情趣，我對全校教職員及學生發行彩券。參加者完全免費，每人可從1至42，任意挑選6碼。活動截止後，以電風扇吹42顆彩球的方式公開產生中獎號碼，凡6碼全中者，可得獎金1萬元，當然是台幣。成立才第二年的高大，12個系共有18班，學生有八百多人，教職員有1百人左右。你可預料無人中獎，因中獎機率為 $1/5,245,786$ ，而在高大共僅收集到718組號碼。事實上，自始至終，我對口袋裡那1萬元，一直相當放心，不擔心有人會向它招手。巴菲特的重賞，吸引了約1,500萬人參加預測。相對於微乎其微的922京分之1的中獎機率，你仍可大膽判斷，無人能完美預測。而的確也是如此。

當各家媒體都指出巴菲特的遊戲，僅不到900萬兆分之1的機率能產生幸運者，我卻有些好奇，以 $(1/2)^{63}$ 當做幸運獲得10億美元獎金的機率，合理嗎？

兩隊賽球，贏者晉級。當終場的哨音響起，打成平手。經過延長賽後，又是平手。二度延長賽，仍分不出勝負。時間實在太晚了，裁判徵求兩隊同意後，以投擲銅板決定輸贏。雙方都能接受，因沒有其他資訊，只好假設銅板出現正、反面的機率皆為 $1/2$ 。離開座位一會兒，回來後發現桌上多了杯咖啡。誰送來的？並不難猜，應是隔壁辦公室A與B兩位善心小姐之一。究竟是那一位？是否與投擲銅板的情況一般，兩位的可能性相同，即機率各 $1/2$ ？不見得，看過“女人香”(Scent of a Woman, 1992)那部電影嗎？空氣中飄散著一股淡淡的茉莉花香，那是A小姐慣用的香水味，八成是她。

並非永遠以相同的可能性，來訂定事件發生的機率。當有可供參考的資訊時，人們對事件發生機率的研判，常隨之而變。

對於NCAA男子籃球季後賽的結果，如果是隨機猜，則第一輪的32場

比賽，皆猜對的機率為

$$\frac{1}{2^{32}} = \frac{1}{4,294,967,296}$$

約為 $4.29 \cdot 10^9$ 分之1，即小於40億分之1。因此不必全部63場，預測能闖過第一輪，便已經是高難度了。換句話說，參加預測的那1,500萬人，若都隨機填寫每場比賽的結果，則根本不必填那麼多，因僅是第一輪，可預期便將人人繳羽而歸。但一般人豈會以隨機的方式預測比賽結果？這假設根本是錯的。我們已指出每區的第一輪比賽，都是第1跟第16打，第2跟第15打，…。所以交手兩隊的強弱，皆有所事資訊，因此判斷輸贏，將不致於從頭到尾皆隨機地猜。由此可見，各家媒體報導中，均提及的那不到900萬兆分之1的中獎機率，聳動歸聳動，卻沒參考價值。

不隨機填，那是否就完全按例行賽的排名填寫輸贏？即預測排名在前的球隊將獲勝？球是圓的，輸贏不見得那麼毫無意外。如果比賽結果都與事先的排名吻合，這樣的球賽，恐怕吸引不了太多觀眾。季後賽在退無可退的壓力下，各隊均有可能激發出潛能，鹹魚翻身經常可見。當然強隊也有可能因壓力過大，而演出大走樣。畢竟大學球員，不像職業隊球員那麼經驗老到。因此過度迷信例行賽的排名，可能不是太好的主意。

我們來檢視過去5年，即從2009至2013年間，第一輪32場比賽的資料。排名在前若輸給排名在後者，我們稱為有一“逆轉”。表1給出各年逆轉的次數；表2給出5年間，各區的排名第1至第8，各被逆轉的總次數。逆轉總次數從7至10，但沒有8次，這應只是純屬巧合。過去5年，第一輪共賽了160場，其中有46次逆轉，佔了28.75%，不到3成，數據看起來合理。例行賽30幾場的成績，仍相當反映各隊的實力。因此是會有些逆轉，但不該太多。各區排名第1至第8，5年間各共有20場賽事，其中排名第5至第8，被逆轉的次數皆較多。這也合理，因其對手之排名分別是12、11、10，及9，交戰雙方排名較接近，實力差距不致於太大。至於各區排名第1者，5年間那20場，皆未曾爆出冷門。這仍是相當合理，因其對手是第16名，實力差了一截，想逆轉較困難。

現在來看2014年第一輪的預測。32場比賽，依據前5年的資料，不妨

表1 2009至2013年, 各年逆轉總次數

2009	2010	2011	2012	2013
10	9	7	10	10

表2 2009至2013年, 各排名被逆轉總次數

1	2	3	4	5	6	7	8
0	3	1	5	10	9	10	8

就大膽假設逆轉次數僅可能介於7至10次。從32場中, 分別挑出7、8、9, 及10場來逆轉, 則其可能性分別有 $\binom{32}{7}$ 、 $\binom{32}{8}$ 、 $\binom{32}{9}$, 及 $\binom{32}{10}$ 種。即分別有3,365,856、10,518,300、28,048,800, 及64,512,240種。4種不同的逆轉次數, 再各以1/4的機率加權, 得平均可能性有26,611,299種。如果4區的第1名對第16名, 你又大膽假設皆不會有逆轉, 則逆轉可能性分別有 $\binom{28}{7}$ 、 $\binom{28}{8}$ 、 $\binom{28}{9}$, 及 $\binom{28}{10}$ 種, 即分別有1,184,040、3,108,105、6,906,900, 及13,123,110種。再各以1/4的機率加權, 得平均可能性有6,080,538.75種, 較26,611,299種減少很多。而此二值, 均遠小於42.9億, 表示如果逆轉次數, 果真介於7至10間, 則會有人第一輪過關的機率將增大不少。

該來檢視2014年第一輪的實際結果了。表3給出第1至第8, 各排名被逆轉的次數。共被逆轉8次, 且排名在前2者, 皆未被逆轉。亦即若採二策略: (1)相信排名第1者不會被逆轉; (2)扣除有排名第1的4場, 分別自剩下的28場中, 隨機挑出7、8、9, 及10場來逆轉, 再各以1/4的機率加權, 則第一輪全部結果皆猜中之機率為1/6,080,538.75。將近600萬分之1的機率, 在約1,500萬人參與下, 便很可能會有人第一輪全命中了。

表3 2014年各排名被逆轉次數

1	2	3	4	5	6	7	8
0	0	1	0	3	2	1	1

表2顯示，各區排名第5至第8者，被逆轉的頻率，遠較排名第1至第4者高出很多。靈敏的讀者可能會連想到，在選出幾場比賽做為逆轉時，應也可以較大的加權機率，放在有排名第5至第8名的賽事。這是對的，只是便不像目前只要從32或28場中，隨機取出幾場那麼單純，計算將略複雜些。加權機率也要能較準確的估計才行。由於本文只是想以淺顯的方式，並以第一輪為例，讓讀者了解，若適當藉助統計，將可大幅提昇預測命中率。所以提供的方法，簡單易懂至上，並不追求最佳。事實上，若有更多資料，如不是僅5年，而是收集更長時間，以及收集更多面相的資料，像是各球隊的韌性、球員抗壓性，與兩兩球隊交戰的紀錄(如是否有某支球隊為另一支球隊的“天敵”)等，對研判比賽勝負的精準性，都可能有幫助。

樂透彩之開獎。由於每期頭獎號碼均隨機產生，所以知道再多過去的資料，對預測下一期，乃毫無幫助。每年NCAA籃球賽的結果，與過去有些關係，但並非太大。有些人試圖以統計裡的迴歸模型來預測，顯然將徒勞無功。要再度提醒一點，球賽裡有各種偶發事件，人算永不如天算。如雖自1985年起，各分區排名第1者，在第一輪比賽從未輸過，但誰曉得？若那一年出現這種大逆轉，並不必太訝異，該來的總是會來。一旦發生某區排名第1者破天荒地輸球，則採納我們提供的排名第1者不會被逆轉之預測策略的那些人，便將全部損龜。

在我們建議的策略(2)裡，扣除有排名第1的4場，分別自剩下的28場中，隨機挑出7、8、9，及10場來逆轉，再各以1/4的機率加權。如此當任何一支非排名第1的球隊被逆轉時，所有參與預測者中，均至少約有 $7/28 = 1/4 = 25\%$ 正確。2014年的實際情況如何？

當南區排名第6的俄亥俄州立大學(Ohio State University)，以1分之差，輸給排名第11的戴頓大學(University of Dayton)後，就使約86%的預測者扼腕。至於中西區排名第3的杜克大學(Duke University)，爆冷輸給排名第14的梅西大學(Mercer University)，有最多的約97.6%的參與者未料中。這兩場賽完，便僅剩約1.9%的參與者了。反之，如果依我們的策略，經兩場逆轉後，將至少約留下 $(1/4)^2 = 6.25\%$ 的預測者繼續角逐。當南區排名第10的史丹佛大學(Stanford University)，打敗排名第7的新墨西

哥大學(University of New Mexico)後,僅餘16人還能對10億美元懷抱希望。在中西區排名第11的田納西大學(University of Tennessee),贏了排名第6的麻塞諸塞大學(University of Massachusetts)後,剩下6人了。又在西區排名第8的岡薩加大學(Gonzaga University),勝過排名第9的奧克拉荷馬州立大學(Oklahoma State University)後,碩果僅存3人。這場其實並非逆轉。當東區排名第8的曼菲斯大學(University of Memphis),擊潰排名第9的喬治華盛頓大學(George Washington University)後,雖非逆轉,但僅餘的3人也笑不出來了。至此參與預測者,全軍覆沒。僅僅賽了25場,第一輪尚未結束,便讓1,500萬人10億美元的幻想,全破滅了。

由上述幾場逆轉或未逆轉的預測結果,我們猜想,大部分人的預測,大致是依例行賽之排名,而當兩隊排名接近時,便用上一些判斷。太仰賴排名當然不是太好,只是看起來眾人的判斷力,也不像太睿智,似比不上我們的兩個簡單策略。

我們的想法是,像這類預測,變數太多,與其苦思最佳策略而不可得,還不如大略觀察分析後,找出其中若干關鍵,給出一些簡單策略,結局反而可能更好。並非天性好賭,而是生活上常需做各種預測,如考完試選填志願,也是在進行預測。既然脫離不了預測,一旦遇到,便抽絲剝繭,了解問題中的邏輯,不必想太多,就訂出簡單預測策略,之後就靜觀其變了。

最後,今年總冠軍對決的兩隊,康乃迪克大學及肯塔基大學(University of Kentucky),前者是東區第7種子,全美排名第26,後者是中西區第8種子,全美排名第29。今年例行賽唯一不敗球隊威奇塔州立大學,是誰讓它在第二輪便止步?正是肯塔基大學。冠亞軍都是排名不算高的球隊,若不加進一些隨機性,全憑年度例行賽的戰績,則光是要預測中爭冠的兩支球隊,便非易事。