國 立 政 治 大 學 應 用 數 學 系

碩 士 學 位 論 文

逆轉勝機率之計算及其在台灣運動彩券投注策略上的應用

A computation on the probability of turning defeat into victory and its application to the betting strategy on Taiwan Sports Lottery

碩士班學生：楊顗錞　撰

指導教授：宋傳欽　博士

姜志銘　博士

中 華 民 國 107 年 7 月 5 日

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

致謝

在轉眼之間研究所生涯即將邁入尾聲，感謝這三年鼓勵我及在身邊陪伴我走過來的每位老師、朋友及同學們。首先要感謝我的指導教授宋傳欽老師及姜志銘老師，非常感謝老師們耐心的教導及督促，在論文撰寫期間也逐字逐句地給予我不同的建議，讓我在這段時間學習到許多不同的知識及研究的態度。在論文口試期間內，感謝陳政輝老師給予寶貴的建議及指正，讓此篇碩士論文可以更加完善。

在研究所求學過程中，研究生們之間的友情也是珍貴的回憶，感謝我的同門泰霖、沁如及張群，總是在我遇到困難或挫折時，幫助我解決問題或是替我加油打氣，感謝我的好朋友力夫、永昌，在課業上一起努力的讀書甚至不厭其煩的教導我，而在生活上遇到不開心的事情，總是聽我訴說讓我能有一個宣洩心情的地方，甚至帶來不同的樂趣就像開心果一樣。感謝承孝、昱翔、侑穎每個禮拜都抽空一起運動、比賽讓我的研究所生活中不只是求學，也留下了許多不同的回憶。

最後感謝生命中最重要的家人和于倢與她的愛犬奇米默默的在背後支持我，在我求學時不斷關心、鼓勵我，也在我最忙碌的時候不斷地幫助我處理瑣碎的事情，讓我可以無憂無慮的去完成我研究所的學業。感謝家人、研究所中的老師及同學們，很高興在我的研究所生活中遇到的各位，我會永遠的記住並珍惜這段美好的回憶。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

中文摘要

本研究旨在探討逆轉勝問題及其擴展應用，依據美國職業籃球 (NBA) 比賽歷史資料顯示，於比賽前三節結束時落後之球隊，最終仍有機會贏得勝利。Gill (2000) 假定了主、客隊的前三節得分和它們在第四節的分數皆具有常態分配，而且假設主、客隊在前三節落後之情況下，第四節得以逆轉勝的分數之條件分布為常態時，推導出逆轉勝機率之計算公式。

然而依常理，前三節得分分差的大小對逆轉勝機率應有不同程度的影響，為更貼近真實賽情，本研究擴展了Gill (2000) 模型，引入二元常態模型探討逆轉勝機率之計算公式。同時利用 NBA 歷年賽事資料對二元常態模型與Gill (2000) 模型進行預測逆轉勝機率之比較，我們發覺二元常態模型在預測逆轉勝機率較Gill (2000) 模型更精準。

最後本研究將上述結果進一步應用至台灣運動彩券分析上，利用二元常態模型計算出投注主隊或客隊獲利金額之期望值，以作為台灣運動彩券讓分盤是否投注之決策依據。然因資料可能不是非常符合二元常態模型，以致實際平均獲利在精準度上未達理想，為保守起見，因此下注 1 元時的獲利金額之期望值至少 0.5 元時，始建議下注。在前述決策方式下，我們發現下注客隊時，可獲得平均 0.1079 元的利潤。

關鍵字：逆轉勝機率、二元常態模型、運動彩券、美國職業籃球

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

Abstract

This study aims to explore the problem of turning defeat into victory and its extended application. According to the historical data of National Basketball Association, the team that falls behind on the first three quarters of the game will still have a chance to win. Gill (2000) assumes that the home team and away team scores at the end of the third quarter and during the fourth quarter are normally distributed. To give the formulas of the conditional probabilities of turning defeat into victory, Gill (2000) further assumes that each of these two corresponding conditional distributions follows a normal distribution.

In practice, the difference of scores during the fourth quarter depends not only on which teams falls behind but also on the difference of scores of these two teams at the end of the third quarter. Therefore, we extend Gill (2000) model to introduce a bivariate normal distribution model. We then give turning defeat into probability formulas by using this new model. Using NBA historical data, we compare our bivariate normal model with Gill (2000) model and find that the bivariate normal model is better than that of Gill (2000) .

This study also applies our bivariate normal model to analyze Taiwan sports lottery. We calculate the expected profit of betting a team to decide whether to bet on this team through Taiwan sports lottery. Although data may not fit the bivariate normal model well and prediction is not accurate enough, we still suggest people interested in gaming could bet on the teams having the expected profit at least 0.5

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

NTD per 1 NTD wager. Under such circumstances, it is found that our average profit per 1 NTD betting is still 0.1079 NTD on the away team.

Keywords: probability of turning defeat into victory, bivariate normal model, sports lottery, National Basketball Association

國立政治大學
National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

目錄

致謝 i
中文摘要 ii
Abstract iii
目錄 v
表目錄 vii
圖目錄 viii
第一章 緒論 1
第一節 研究背景與動機 1
第二節 研究目的 2
第三節 論文架構 3
第二章 文獻回顧 4
第一節 逆轉勝問題分析 4
第二節 運動彩券概述 7
第三節 台灣運動彩券玩法 8
第三章 研究方法 10
第一節 研究步驟 10
第二節 研究問題與假設 11
第三節 資料來源及整理 11
第四節 模型建立與公式推導 11
第五節 模型應用到運動彩券 19
第四章 實證結果 (以 NBA 為例) 22
第一節 球賽逆轉勝問題分析 22
第二節 運彩讓分盤預測結果 24

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第五章 結論與建議 28
附錄 A 不同年份資料預測讓分盤之實際投注平均獲利 29
附錄 B 程式碼 37
參考文獻 51

國立政治大學
National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表目錄

2.1 第四節分差彙整表 7

4.1 逆轉勝機率及其預測誤差之比較 23
4.2 以過去十年資料預測讓分盤之實際投注平均獲利 24
4.3 每場最少獲利期望值為 0.5 重抽樣法之結果 26
4.4 實際投注平均獲利之摘要統計值 (表 4.3) 27
4.5 最少獲利期望值為 0.5 元之場次，其實際平均獲利 95% 信賴區間 27

A.1 以過去九年資料預測讓分盤之實際投注平均獲利 29
A.2 以過去八年資料預測讓分盤之實際投注平均獲利 30
A.3 以過去七年資料預測讓分盤之實際投注平均獲利 31
A.4 以過去六年資料預測讓分盤之實際投注平均獲利 32
A.5 以過去五年資料預測讓分盤之實際投注平均獲利 33
A.6 以過去四年資料預測讓分盤之實際投注平均獲利 34
A.7 以過去三年資料預測讓分盤之實際投注平均獲利 35
A.8 以過去二年資料預測讓分盤之實際投注平均獲利 36

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

圖目錄

1.1 論文架構圖 3

3.1 研究流程圖 10
3.2 x_1, x_2積分區域 14
3.3 z_1, z_2積分區域 14
3.4 x_1, x_2積分區域 15
3.5 z_1, z_2積分區域 15
3.6 x_1, x_2積分區域 17
3.7 z_1, z_2積分區域 17
3.8 x_1, x_2積分區域 18
3.9 z_1, z_2積分區域 18
3.10 建議投注讓分盤流程圖 21

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第一章

緒論

第一節　研究背景與動機

比賽者無非勝負之爭，而球賽亦莫過於此，每當賽事展開，無論身在現場的群眾抑或關注電視的球迷，莫不為場上的情況所牽引、或加油、或祈禱，為得分振聲喝采、為失分扼腕長嘆，冀望著所支持球隊的精彩表現。然而賽事總是難以捉摸，其賽況或平淡、或曲折、或勾人心弦、或食之無味，而其中最為波折，最具有戲劇性，也最讓觀眾們熱血振奮的，無非是看見那一次次奇蹟般的逆轉勝。

因此在過去幾年中，亦有不少針對逆轉勝賽情的相關研究，如 Cooper et al. (1992) 對逆轉勝賽情的綜合分析，與 Gill (2000) 所提出之以常態模型預測逆轉勝機率，皆擴展了逆轉勝問題，為預測逆轉勝這一難以捉摸的問題作出相當的貢獻。

Gill (2000) 探討逆轉勝問題，並假設主、客場球隊前三節總分及第四節得分都為常態分布。定義前三節得分分差及第四節得分分差分別為主場球隊分數減去客場球隊分數， Gill (2000) 更進一步利用在已知主、客隊前三節落後之情況下，最後贏球之常態條件分布的假設下，用以預測逆轉勝的機率。

而本研究的另一項動機為運彩彩券的應用，在 19 世紀末時，發展出投注賽馬的比賽，也奠定了運動彩券的基礎。隨著全球運動盛行，運動的發展越來越多樣化，也因此運動彩券的種類也開始推陳出新，全球目前有八十多個國家 (包括義大利、法國、德國、日本、中國大陸、韓國…等) 發行運動彩券 (周錫洋，2005)，運動彩券也開始成為

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

世界的潮流。

台灣運動彩於 2008 年 5 月 2 日正式公開發行，一開始投注標的賽事只有棒球、籃球、足球，大多都是以國際運動賽事為主，近年來慢慢開始增加許多不同賽事 (包括羽球、網球……等)，甚至在 2014 年加入了台灣的體育賽事，讓投注者可以有更多的投注選擇，其中以新玩法「場中投注」[1]最受歡迎 (台灣運動彩券，2008a)。近來預測職業運動賽事的人數益發增長，讓投注者可以藉由此類分析結果作為投注的依據，以提高『過關』[2] 的可能。

第二節　研究目的

在實際的籃球比賽中，前三節兩球隊的分差大小會對第四節兩球隊的表現有不同程度的影響，Gill (2000) 的一維常態分布模型，並無法彰顯這種現象，為了解決這種問題，我們擴展了 Gill (2000) 模型，提出二元常態模型來描述逆轉勝問題，以期能更精準地預測逆轉勝之結果。

除了上述研究籃球逆轉勝問題之外，也因為運動彩券的盛行，我們更以此二元常態模型應用到台灣運動彩券上做分析，以讓分玩法為例，希望能讓投注者可以藉此分析來提高投注獲利金額之期望值，甚至為正期望值。

本研究使用美國職業籃球聯賽 (National Basketball Association，以下簡稱 NBA)2006-2007 年賽季至 2017-2018 年賽季的 12 年的例行賽資料在逆轉勝問題進行二元常態模型與 Gill (2000) 模型之比較，再利用 2007-2008 年賽季至 2016-2017 例行賽資料作為數據，以台灣運動彩券網站上每場比賽最終所開出讓分盤之分數為基準，預測 2017-2018 年賽季在讓分盤玩法下之投注獲利金額之期望值。

除此之外也利用 Python 程式提供一個計算單場讓分盤的程式，投注者只要輸入對戰隊伍的主、客場球隊名稱、台灣運動彩券所開出主隊讓分盤之分數以及投注主、客隊之賠率即可快速得到投注 1 元時主、客隊獲利金額之期望值以及是否為建議投注之比賽。

[1]Live Betting，俗稱走地，即可於賽事進行中，完結前隨時進行投注
[2]過關意指此賽事與投注者預測結果相符

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第三節　論文架構

本文將分為五個章節，第一章為緒論，了解本研究背景、動機與目的。第二章文獻回顧介紹逆轉勝問題、運動彩券的起源與現況，並將所參考之模型做簡單介紹，第三章研究方法介紹研究流程與驗證問題所需要的假設，最後提出模型理論及提供理論之推導過程，第四章為實證結果以 NBA 為例，陳述最後研究的結果及分析，第五章為結論與建議。

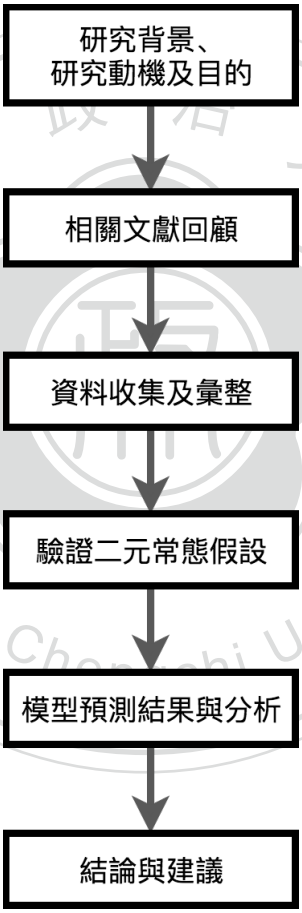

圖. 1.1. 論文架構圖

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第二章

文獻回顧

第一節　逆轉勝問題分析

在比賽前三節結束後落後之球隊，Cooper et al(1992) 認為約有 20% 比率會在最後贏得比賽，亦即有 20% 逆轉勝可能。

Gill (2000) 假設 (X,Y) 及 (U,V) 皆服從二元常態分配，其中 X,Y 分別為主場及客場球隊在前三節得到之總分，U,V 分別為主場及客場球隊第四節得到之分數。再令 Z=X-Y 及 W=U-V，則 Z,W 分別表示主場球隊在前三節領先之分數及第四節領先之分數。在已知前三節為輸或贏之情況下，Gill (2000) 進一步假定第四節得以逆轉勝的分數之條件分布為常態。在實務上這種限制條件不見得合理，因此本文將探討一般的二元常態模型下之逆轉勝問題。以下先重述Gill (2000) 模型的推導過程。

為了瞭解在第四節逆轉比賽的可能，必須先考慮前三節比賽結束時不是平手的情況，並計算前三節比賽結束時不是平手的機率，計算方式為利用機率總和減去平手之機率，故推導出平手的機率公式如下：

$$\Pr(\text{Tie}) = \Pr(-0.5 < Z < 0.5) = \Phi\left\{\frac{0.5-\mu}{\sigma}\right\} - \Phi\left\{\frac{-0.5-\mu}{\sigma}\right\} \tag{1.1}$$

其中 $\Phi(\cdot)$ 為常態分布累積分配函數

接下來推導出在前三節比賽結束時領先 n 分之公式，假定 H_n 和 A_n 分別為主場球隊和客場球隊在前三節比賽結束時領先 n 分

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

$$\Pr(H_n|\text{no Tie}) = \frac{\Pr\left(n-0.5 < Z < n+0.5\right)}{1-\Pr(\text{Tie})}$$

$$= \frac{\Phi\left\{\dfrac{n+0.5-\mu}{\sigma}\right\} - \Phi\left\{\dfrac{n-0.5-\mu}{\sigma}\right\}}{1-\Phi\left\{\dfrac{0.5-\mu}{\sigma}\right\} + \Phi\left\{\dfrac{-0.5-\mu}{\sigma}\right\}} \tag{1.2}$$

$$\Pr(A_n|\text{no Tie}) = \frac{\Pr\left(-n-0.5 < Z < -n+0.5\right)}{1-\Pr(\text{Tie})}$$

$$= \frac{\Phi\left\{\dfrac{-n+0.5-\mu}{\sigma}\right\} - \Phi\left\{\dfrac{-n-0.5-\mu}{\sigma}\right\}}{1-\Phi\left\{\dfrac{0.5-\mu}{\sigma}\right\} + \Phi\left\{\dfrac{-0.5-\mu}{\sigma}\right\}} \tag{1.3}$$

在考慮上述兩種不同情況下，將推導出前三節領先 n 分的球隊最後贏得比賽之公式

1. 在主場球隊前三節比賽結束領先情況下，假定 W 服從條件常態分配 $\mathbf{N}(\delta_1, \nu_1^2)$ ，主場球隊在前三節比賽結束時領先 n(>0) 分，則最後主場球隊贏得比賽的公式如下：

Pr(主場球隊贏得比賽｜主場球隊在前三節結束領先 n 分)

$$= \Pr(W > -n) = \Pr(W \geq -n - \frac{1}{2}) = 1 - \Phi\left\{\frac{-n-\frac{1}{2}-\delta_1}{\nu_1}\right\} \tag{1.4}$$

2. 在客場球隊前三節比賽結束領先情況下，假定 W 服從條件常態分配 $\mathbf{N}(\delta_2, \nu_2^2)$ ，客場球隊在前三節比賽結束時領先 n(>0) 分，則最後客場球隊贏得比賽的公式如

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

下：

$$\text{Pr(客場球隊贏得比賽｜客場球隊在前三節結束領先 n 分)}$$

$$= \Pr(W < n) = \Pr(W \leq n + \frac{1}{2}) = \Phi\left\{\frac{(n+\frac{1}{2}) - \delta_2}{\nu_2}\right\} \tag{1.5}$$

最後推導出在前三節領先之球隊，最後贏得比賽機率公式如下：

$$\text{Pr(主場球隊贏得比賽｜主場球隊在前三節結束領先)}$$

$$= \frac{\sum_{n=1}^{\infty}(1 - \Phi\left\{\frac{-n-\frac{1}{2}-\delta_1}{\nu_1}\right\}) \times \Pr(H_n|\text{no Tie})}{\sum_{n=1}^{\infty}\Pr(H_n|\text{no Tie})} \tag{1.6}$$

$$\text{Pr(客場球隊贏得比賽｜客場球隊在前三節結束領先)}$$

$$= \frac{\sum_{n=1}^{\infty}\Phi\left\{\frac{(n+\frac{1}{2})-\delta_2}{\nu_2}\right\} \times \Pr(A_n|\text{no Tie})}{\sum_{n=1}^{\infty}\Pr(A_n|\text{no Tie})} \tag{1.7}$$

接著Gill (2000) 利用 NBA1997-1998 年賽季資料來當作實證分析，在賽季資料中可得知 $\mu = 2.63$，$\sigma = 12.27$，利用式 (1.1) 計算平手機率為 0.032，以實際資料計算平手機率為 0.036

依式 (1.2) 及 (1.3) 計算主場球隊及客場球隊在前三節比賽結束時領先機率，$\Pr(H|\text{no Tie}) = \sum_{n=1}^{\infty}\Pr(H_n|\text{no Tie}) = 0.586$ 及 $\Pr(A|\text{no Tie}) = \sum_{n=1}^{\infty}\Pr(A_n|\text{no Tie}) = 0.414$，以實際資料計算，主場球隊及客場球隊在前三節比賽結束時領先機率分別為 0.576 及 0.424[1]。

[1]比賽場次共有 1145 場 (已排除在前三節比賽結束時平手場次)，主場球隊及客場球隊在前三節比賽結束時領先場次分別為 660 場、485 場

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. 2.1. 第四節分差彙整表

已知第三節狀況 / 統計量	主場球隊領先	客場球隊領先	總體狀況
平均	-0.27	1.19	0.34
標準差	7.15	7.32	7.28

表 2.1為第四節分差彙整表，分差即為主場球隊減去客場球隊之分數，以實際資料計算時，在只看第四節比賽得分分差下，主場球隊在第四節平均分差為 0.34 分，但在前三節結束主場球隊領先的情況下，主場球隊在第四節的平均分差為-0.27 分，如果前三節結束客場球隊領先的情況下，主場球隊第四節的平均分差為 1.19 分。

表 2.1顯示在第四節前已知主場球隊領先的情形下，平均來說客隊在第四節的表現比主隊來的好，反之在第四節前已知客場球隊領先的情形下，平均來說主隊在第四節表現比客隊來的好，這種的結果，再次確認前述 Cooper et al. (1992) 所提的存在相當的逆轉勝可能性。

在主場或客場領先時，使用表 2.1的參數 $\delta_1 = -0.27, \nu_1 = 7.15, \delta_2 = 1.19, \nu_2 = 7.32$，利用式 (1.4) 及 (1.5)，計算前三節領先 n 分之球隊，最後贏得比賽的機率。

最後用式 (1.6) 及 (1.7) 計算出在前三節領先之球隊，最後贏得比賽的機率，即為不被逆轉之機率，在主場球隊不被逆轉機率為 0.852，在客場球隊不被逆轉機率為 0.788，在利用前面所得出之結果，不論任何比賽，前三節領先之球隊不被逆轉之機率為 0.586×0.852+0.414×0.788=0.825，而逆轉勝機率為 1-0.825=0.175，以實際資料計算逆轉勝之機率為 0.163[2]

第二節　運動彩券概述

運動彩券起源於十九世紀末，由法國人皮耶為英國的賽馬設計一套彩池均分制，可以有效控制莊家風險以及提高彩金金額，比賽結果出爐後分派彩金，在當時也廣受歡迎，但隨著運動的蓬勃發展，運動彩券迷有了更完整的資訊來作為依據，利用這些資訊預測比賽結果，也從原來的彩池均分制轉變為現在運動彩券的固定賠率制 (林桓，

[2]比賽場次共有 1145 場 (以排除在前三節比賽結束時平手場次，逆轉場次為 186 場)

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

2008)。在二十世紀初，英國成立首家博彩公司，運動彩券也就此開始，由於運動彩券的盈餘，因此提供了體育發展經費，因此廣遍歐陸，很多的運動項目也逐漸被加入，但主要還是以職業運動為標的。

運動彩券在美洲地區，因 1992 年美國國會通過「職業及業餘運動員保護法」，禁止彩券以競技賽事為目標 (周錫洋，2005)，但實際上還是有不少並非以運動彩券方式存在，雖然當時未發行運動彩券，但運動賭博是合法的，以合法運動投注市場最大的美國拉斯維加斯為例，投注標的包括四大職業運動美式足球、職業籃球、職業棒球、冰上曲棍球及網球…. 等。

台灣運動彩券自 2008 年 5 月 2 日正式發行，初步開放的投注標的，種類只有美國職棒 (MLB)、美國職籃 (NBA) 以及英格蘭超級聯賽 (EPL)，玩法分為讓分、大小、勝分差、主客和、正確得分及冠軍等 6 種玩法，近年來因運動的蓬勃發展，台灣運動彩券的投注站林立，提供運動彩迷更多賽事、更多玩法、更多選擇，2014 年首度將台灣籃球聯賽 (SBL) 季後賽納入投注標的，以實際行動支持國內運動賽事。甚至提供「場中投注」，俗稱「走地」，隨著不同熱門運動或賽事適時推出供消費者投注，尤其是美國職籃的場中投注，消費者可以在比賽進行時，購買當節或是最終比賽結果，讓消費者可以邊看比賽邊投注，大幅提升比賽的緊張刺激程度。目前台灣運動彩券提供賽事包括棒球、籃球、足球、網球、羽球、冰球、美式足球、賽車、高爾夫、撞球等超過 10 種運動項目、20 種玩法提供消費者投注。

第三節　台灣運動彩券玩法

依現行台灣運動彩券發行機構公布所示，目前投注標的有棒球、籃球、足球、網球..... 等，其中籃球的玩法共有十三種，各種玩法規則分述如下：(台灣運動彩券，2008b)

1. 不讓分：預測一場籃球賽事在法定比賽時間結束時獲勝方球隊。

2. 讓分：依據過去的比賽紀錄以及目前球隊的狀況，將比賽中的兩支球隊分為「較強隊伍」與「較弱隊伍」，設定「較強隊伍」需讓分給「較弱隊伍」，其中讓出的分數由「台灣運動彩券」自訂。假設客隊聖安東尼奧馬刺隊 (以下簡稱馬刺隊) 出戰

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

主隊金州勇士隊 (以下簡稱勇士隊)，且勇士隊讓分設定為 12.5 分。若馬刺隊對勇士隊比賽最終比分為 107：125，勇士隊讓分之分數則為 125 分減 12.5 分，讓分後結果馬刺隊對勇士隊為 107：112.5，仍為勇士隊贏；若馬刺隊對勇士隊比賽最終比數為 107：110，讓分後結果馬刺隊對勇士隊則為 107：97.5，轉為馬刺隊勝出。

3. 大小：預測一場籃球賽事在法定比賽時間結束時，二隊得分之總和大於或小於台灣運動彩券所訂定的總得分。總得分會依每場不同的賽事來設定，且可能依狀況而調整。

4. 勝分差：預測一場籃球賽事在法定比賽時間結束時，主隊或客隊勝出的分數差距，比賽雙方的分差被劃分為 1-5、6-10、11-15、16-20、21-25、26+，等 6 個區間，主勝和客勝兩種賽果，總共有 12 個選項。

5. 客隊 (或主隊) 大小：預測一場籃球賽事在法定比賽時間完成時，客隊 (或主隊) 的總得分大於或小於台灣運動彩券所訂定的客隊 (或主隊) 的總分。

6. 單節大小：在一場籃球賽事中，投注者依據自己指定之節數，來預測該節兩隊的總得分大於或小於台灣運動彩券所訂定的該節總分。

7. 半場大小：在一場籃球賽事中，投注者依據自己指定之半場 (上半場或下半場)，來預測該半場兩隊的總得分大於或小於台灣運動彩券所訂定的該半場總分。

8. 單節不讓分：在一場籃球賽事中，投注者依據自己指定之節數，來預測該節比賽結果為客隊勝、和局或主隊勝。

9. 單節讓分：在一場籃球賽事中，投注者依據自己指定之節數，來預測該節比賽經讓分數調整正式比數的結果。

10. 半場不讓分：在一場籃球賽事中，投注者依據自己指定之半場 (上半場或下半場)，來預測該半場比賽結果為客隊勝、和局或主隊勝。

11. 半場讓分：在一場籃球賽事中，投注者依據自己指定之半場 (上半場或下半場)，來預測該半場比賽經讓分數調整正式比數的結果。

12. 單雙：預測一場籃球賽事在法定時間完成時，選擇兩隊得分總數為單數 (如 1,3,5,⋯) 或雙數 (如 0,2,4,⋯)。

13. 冠軍：預測比賽指定的總冠軍、小組或分區冠軍球隊。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第三章

研究方法

第一節　研究步驟

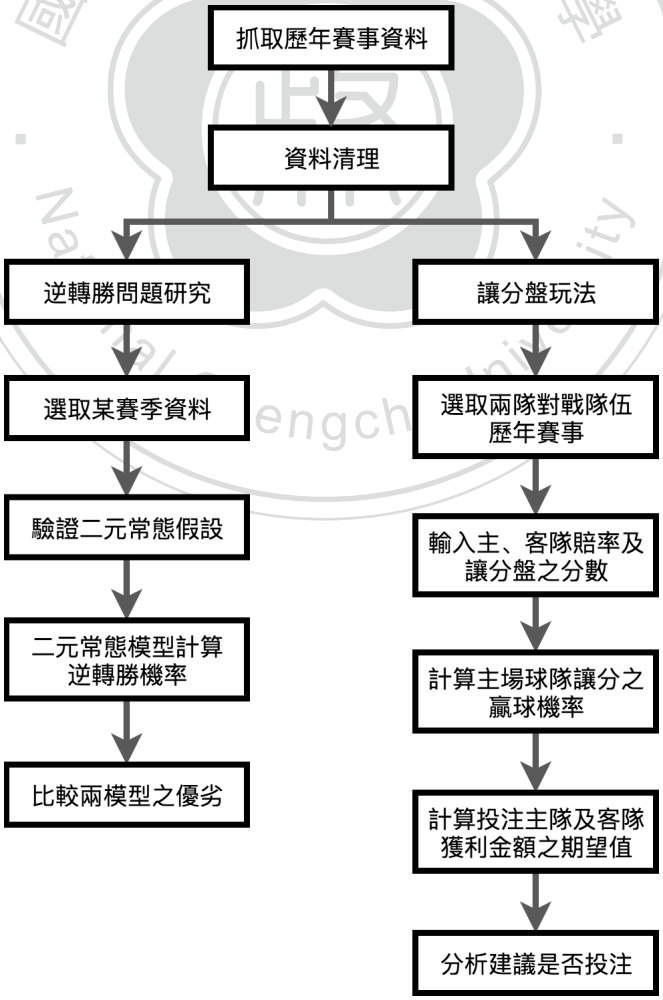

圖. 3.1. 研究流程圖

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第二節　研究問題與假設

由於在研究逆轉勝問題分析時，假定了兩變數 (此為前面所提到的前三節得分分差與第四節得分分差) 服從二元常態分配，故在利用二元常態模型計算之前，需先驗證兩個變數是否為二元常態分配。

有很多統計軟體可驗證二元常態分配，本研究以 R 語言的 MVN(Multivariate Normality) 套件，檢驗各年資料是否滿足二元常態分配假設。

第三節　資料來源及整理

本研究資料來源為「BASKETBALL REFERENCE」[1]的網站，並使用 R 語言的網路爬蟲抓取歷年美國職籃各球隊例行賽中每場比賽的總得分及每一節得分的分數，將資料匯出成 Excel 檔之後，再把各年資料的主場球隊及客場球隊在比賽中的每一節的得分整理成前三節得分分差與第四節得分分差用以分析逆轉勝問題。

將本研究應用到台灣運動彩券的讓分盤比賽，資料來源為「玩預測」的網站[2]，使用 Python 程式取得台灣運動彩券在每一場比賽最終所開出讓分盤之分數，所以抓取 2007-2008 年賽季至 2016-2017 年賽季的 10 年的例行賽資料作為數據，以預測 2017-2018 賽季比賽。

第四節　模型建立與公式推導

本節將介紹二元常態模型與其性質，並說明本研究如何推導二元常態模型的公式，利用此公式算出逆轉勝問題機率。

若 $\mathbf{X} = (X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，其聯合機率密度函數可表示成:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp(-\frac{q(x_1, x_2)}{2(1-\rho^2)}) \tag{4.1}$$

[1]BASKETBALL REFERENCE 網址：https://www.basketball-reference.com/
[2]玩預測網址：https://www.playsport.cc/index.php

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

其中 $q(x_1,x_2)=(\frac{x_1-\mu_1}{\sigma_1})^2-2\rho(\frac{x_1-\mu_1}{\sigma_1})(\frac{x_2-\mu_2}{\sigma_2})+(\frac{x_2-\mu_2}{\sigma_2})^2$，$-1\le\rho\le 1$

再來介紹二元常態分布的條件機率密度函數，若 $\det(\Sigma)>0$，在給定 X_1 條件下，則 X_2 的條件分布為常態分布，且條件期望值、條件變異數為：

$$E(X_2|X_1=x_1)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\qquad Var(X_2|X_1=x_1)=\sigma_2^2(1-\rho^2)$$

其中 $\mu_1=E(X_1)$，$\mu_2=E(X_2)$，$\sigma_1=\sqrt{Var(X_1)}$，$\sigma_2=\sqrt{Var(X_2)}$，ρ 為 X_1 和 X_2 的相關係數，則

$$X_2|X_1=x_1\sim N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\ ,\ \sigma_2^2(1-\rho^2))$$

推導逆轉勝問題分析時，利用二元常態模型來做雙重積分計算出逆轉之機率，計算的時候將式 (4.1) 改寫成向量的形式

$$f_{\mathbf{X}}(\mathbf{x})=\frac{1}{(2\pi)(\det(\Sigma))^{\frac{1}{2}}}\exp(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}))\tag{4.2}$$

其中 $\mathbf{X}=\begin{pmatrix}X_1\\X_2\end{pmatrix}$，$\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix}$，$\Sigma=\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\\sigma_{21} & \sigma_2^2\end{pmatrix}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$

接下來首先假設 Σ 的第 i 對的特徵值、特徵向量分別為 λ_i、$\mathbf{e}_i$，則可將 Σ 對角化使得其中的第 i 行為 $\mathbf{e}_i$ 且對角矩陣 Λ 的第 i 個對角元素為 λ_i，因此 $\Sigma^{-1}=\Gamma\Lambda^{-1}\Gamma^T=\Gamma\Lambda^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}\Gamma^T=(\Gamma\Lambda^{-\frac{1}{2}})(\Gamma\Lambda^{-\frac{1}{2}})^T$，(4.2) 可進一步表示成：

$$f_{\mathbf{X}}(\mathbf{x})=\frac{1}{2\pi(\det(\Sigma))^{\frac{1}{2}}}\exp(-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T\Gamma\Lambda^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}\Gamma^T(\mathbf{X}-\boldsymbol{\mu}))\tag{4.3}$$

最後再將二元常態隨機向量 $\mathbf{X}=\begin{pmatrix}X_1\\X_2\end{pmatrix}$ 標準化，亦即 $\mathbf{Z}=\begin{pmatrix}Z_1\\Z_2\end{pmatrix}$，其 x_1,x_2 和 z_1,z_2 的關係如下：

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \Sigma^{-\frac{1}{2}} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} = \Lambda^{-\frac{1}{2}} \Gamma^T \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \Gamma \Lambda^{\frac{1}{2}} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \tag{4.4}$$

此時 **Z** 為二元標準常態分配，其聯合機率密度函數可表示為：

$$f_{Z_1,Z_2}(z_1, z_2) = \frac{1}{2\pi} \exp(-\frac{1}{2}(z_1^2 + z_2^2)) \tag{4.5}$$

最後可利用式 (4.5) 做雙重積分，以計算逆轉勝之機率。接著將推導逆轉勝機率之一般公式：

本文所探討的機率問題 (包含逆轉勝的機率) 的區域僅有 $x_1 > 0$ (或 $x_1 < 0$) 及 $x_1 + x_2 > 0$ (或 $x_1 + x_2 < 0$)，因此它們的積分區域經由式 (4.4) 關係式轉換對應於 z_1 及 z_2 的積分區域，亦即：

$$\text{若} x_1 > 0, \text{則} \begin{cases} z_2 > \dfrac{-\mu_1 - az_1}{b} & , \text{if } b > 0 \\ z_2 < \dfrac{-\mu_1 - az_1}{b} & , \text{if } b < 0 \\ az_1 > \mu_1 & , \text{if } b = 0 \end{cases}$$

$$\text{若} x_1 + x_2 > 0, \text{則} \begin{cases} z_2 > \dfrac{-(\mu_1 + \mu_2) - (a+c)z_1}{b+d} & , \text{if } b + d > 0 \\ z_2 < \dfrac{-(\mu_1 + \mu_2) - (a+c)z_1}{b+d} & , \text{if } b + d < 0 \\ (a+c)z_1 > (\mu_1 + \mu_2) & , \text{if } b + d = 0 \end{cases}$$

其中 a, b, c, d 為 $\Gamma \Lambda^{\frac{1}{2}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 中之元素

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

1. 在主場球隊前三節比賽結束領先情況下且最後主場球隊贏得比賽機率之一般公式：

$$\Pr(X_1+X_2>0, X_1>0)$$

$$=\int_{-\infty}^{s_1}\int_{-\infty}^{\frac{-\mu_1-az_1}{b}}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1+$$

$$\int_{s_1}^{\infty}\int_{-\infty}^{\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d}}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1$$

$$=\int_{-\infty}^{s_1}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{-\infty}^{\frac{-\mu_1-az_1}{b}}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1+$$

$$\int_{s_1}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{-\infty}^{\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d}}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1$$

$$=\int_{-\infty}^{s_1}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\Phi(\frac{-\mu_1-az_1}{b})dz_1+$$

$$\int_{s_1}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\Phi(\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d})dz_1$$

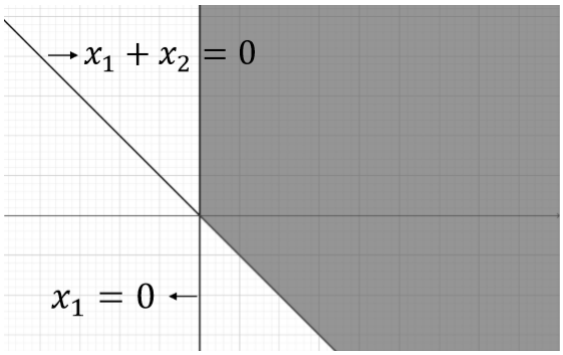

圖. 3.2. x_1, x_2積分區域

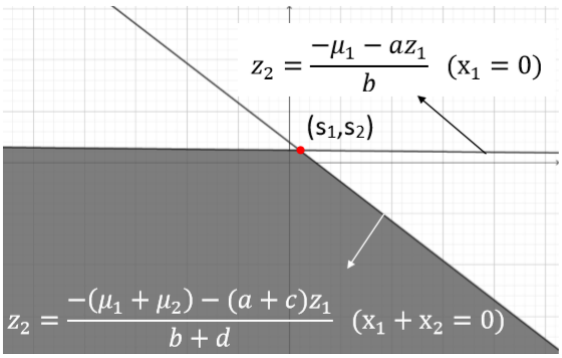

圖. 3.3. z_1, z_2積分區域

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

2. 在客場球隊前三節比賽結束領先情況下且最後客場球隊贏得比賽機率之一般公式：

$$\Pr(X_1+X_2<0, X_1<0)$$

$$=\int_{-\infty}^{s_1}\int_{\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d}}^{\infty}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1+$$

$$\int_{s_1}^{\infty}\int_{\frac{-\mu_1-az_1}{b}}^{\infty}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1$$

$$=\int_{-\infty}^{s_1}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1+$$

$$\int_{s_1}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{\frac{-\mu_1-az_1}{b}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1$$

$$=\int_{-\infty}^{s_1}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})(1-\Phi(\frac{-(\mu_1+\mu_2)-(a+c)z_1}{b+d}))dz_1+$$

$$\int_{s_1}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})(1-\Phi(\frac{-\mu_1-az_1}{b}))dz_1$$

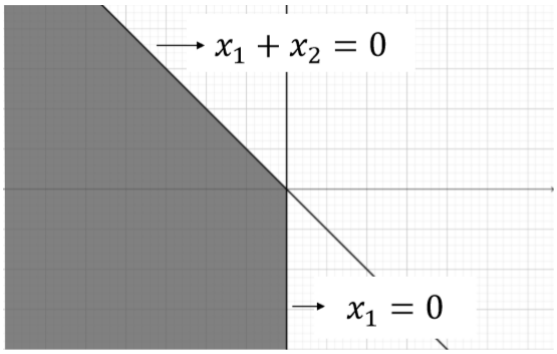

圖. 3.4. x_1, x_2積分區域

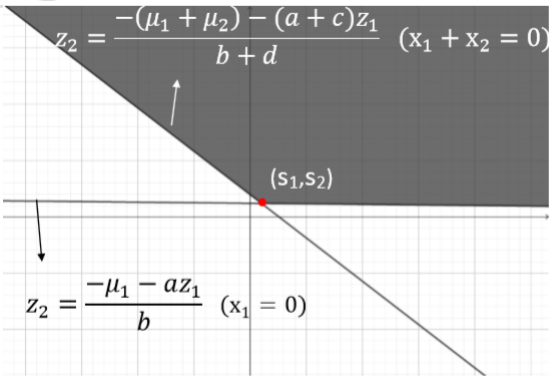

圖. 3.5. z_1, z_2積分區域

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

最後推導出來的積分式子將利用數值解的方法去算出機率值，而在利用二元常態模型計算逆轉勝之機率，利用 python 把一連串計算的過程寫成程式，在附錄中附上 python 的程式碼，以提供計算使用。

以下我們使用 2017-2018 賽季之資料作為例子以介紹上述的計算過程。在 2017-2018 賽季中，從賽季資料得到 $\mathbf{X} = (X_1, X_2) \sim N_2(\mu_1 = 1.5966, \mu_2 = 0.5376, \sigma_1^2 = 178.3203, \sigma_2^2 = 59.0681, \rho = -0.2227)$ 可將聯合機率密度函數表示成：

$$
\begin{aligned}
&f_{X_1,X_2}(x_1, x_2) \\
&= \frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} \exp\left(x_1 - 1.5966, x_2 - 0.5376\right)^T \begin{pmatrix} 178.3203 & -22.8582 \\ -22.8582 & 59.0681 \end{pmatrix} \begin{pmatrix} x_1 - 1.5966 \\ x_2 - 0.5376 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\Sigma &= \begin{pmatrix} 178.3203 & -22.8582 \\ -22.8582 & 59.0681 \end{pmatrix} \\
&= \begin{pmatrix} -0.1820 & -0.9833 \\ -0.9833 & 0.1820 \end{pmatrix} \begin{pmatrix} 54.8368 & 0 \\ 0 & 182.5516 \end{pmatrix} \begin{pmatrix} -0.1820 & -0.9833 \\ -0.9833 & 0.1820 \end{pmatrix}
\end{aligned}
$$

令 $\Gamma = \begin{pmatrix} -0.1820 & -0.9833 \\ -0.9833 & 0.1820 \end{pmatrix}$，$\Lambda = \begin{pmatrix} 54.8368 & 0 \\ 0 & 182.5516 \end{pmatrix}$

經由標準化後，可以得知 x_1, x_2 和 z_1, z_2 的關係如下：

$$
\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0.1350 & 0 \\ 0 & 0.0740 \end{pmatrix} \begin{pmatrix} -0.1820 & -0.9833 \\ -0.9833 & 0.1820 \end{pmatrix} \begin{pmatrix} x_1 - 1.5966 \\ x_2 - 0.5376 \end{pmatrix}
$$

$$
\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -0.0246 & -0.1328 \\ -0.0728 & 0.0135 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0.1106 \\ 0.1090 \end{pmatrix}
$$

$$
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1.3479 & -13.2855 \\ -7.2815 & 2.4593 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} 1.5966 \\ 0.5376 \end{pmatrix}
$$

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

1. 計算在主場球隊前三節比賽結束領先情況下且最後主場球隊贏得比賽機率之過程如下：

$$
\begin{aligned}
&\Pr(X_1+X_2>0\,,X_1>0)\\
&=\int_{-\infty}^{0.1106}\int_{-\infty}^{-0.1015z_1+0.1202}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1+\\
&\int_{0.1106}^{\infty}\int_{-\infty}^{-0.7971z_1+0.1971}\frac{1}{2\pi}\exp(\frac{-1}{2}(z_1^2+z_2^2))dz_2dz_1\\
&=\int_{-\infty}^{0.1106}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{-\infty}^{-0.1015z_1+0.1202}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1+\\
&\int_{0.1106}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\int_{-\infty}^{-0.7971z_1+0.1971}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_2^2}{2}))dz_2dz_1\\
&=\int_{-\infty}^{0.1106}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\Phi(-0.1015z_1+0.1202)dz_1+\\
&\int_{0.1106}^{\infty}\frac{1}{\sqrt{2\pi}}\exp(\frac{-z_1^2}{2})\Phi(-0.7971z_1+0.1971)dz_1\\
&=0.4641
\end{aligned}
$$

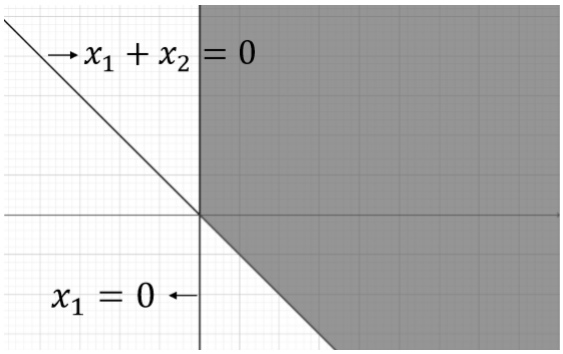

圖. 3.6. x_1, x_2積分區域

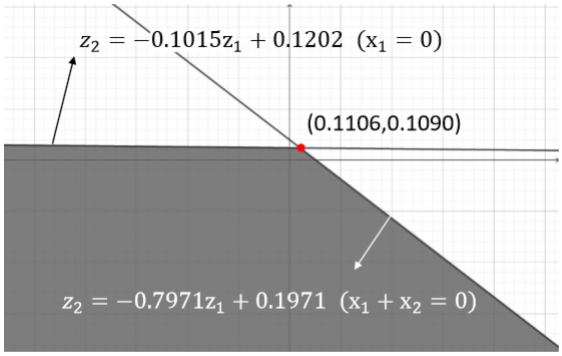

圖. 3.7. z_1, z_2積分區域

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

2. 計算在客場球隊前三節比賽結束領先情況下且計算最後客場球隊贏得比賽機率之過程如下：

$$
\begin{aligned}
&\Pr(X_1 + X_2 < 0, X_1 < 0) \\
&= \int_{-\infty}^{0.1106} \int_{-0.7971z_1+0.1971}^{\infty} \frac{1}{2\pi} \exp(\frac{-1}{2}(z_1^2 + z_2^2)) dz_2 dz_1 + \\
&\quad \int_{0.1106}^{\infty} \int_{-0.1015z_1+0.1202}^{\infty} \frac{1}{2\pi} \exp(\frac{-1}{2}(z_1^2 + z_2^2)) dz_2 dz_1 \\
&= \int_{-\infty}^{0.1106} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_1^2}{2}) \int_{-0.7971z_1+0.1971}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_2^2}{2})) dz_2 dz_1 + \\
&\quad \int_{0.1106}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_1^2}{2}) \int_{-0.1015z_1+0.1202}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_2^2}{2})) dz_2 dz_1 \\
&= \int_{-\infty}^{0.1106} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_1^2}{2})(1 - \Phi(-0.7971z_1 + 0.1971)) dz_1 + \\
&\quad \int_{0.1106}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(\frac{-z_1^2}{2})(1 - \Phi(-0.1015z_1 + 0.1202)) dz_1 \\
&= 0.3553
\end{aligned}
$$

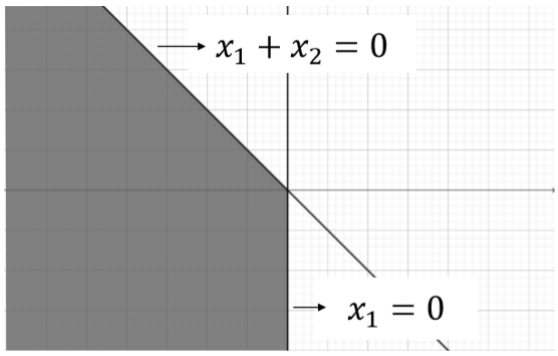

圖. 3.8. x_1, x_2積分區域

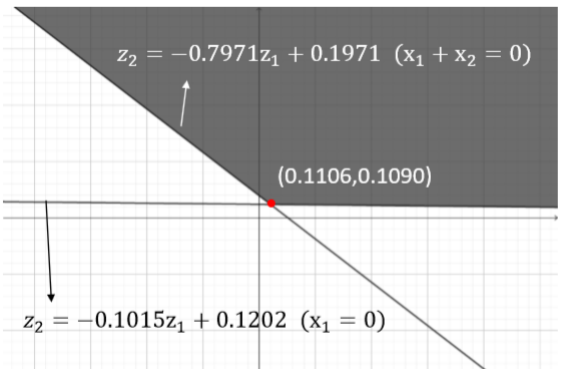

圖. 3.9. z_1, z_2積分區域

用上述二元常態模型計算出當前三節領先之球隊，最後又贏得比賽的機率，

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

即為不被逆轉之機率，利用前面所得出不論任何比賽球隊不被逆轉之機率為 0.4641+0.3553=0.8194，而逆轉之機率為 1-0.8194=0.1806。

第五節　模型應用到運動彩券

本節將利用上一節中模型應用到台灣運動彩券上，推導出讓分盤之主隊贏球機率，利用此機率及投注主、客隊賠率 (以台灣運彩為例大約在 1.65 至 1.85 間) 計算出投注 1 元主隊及客隊獲利金額之期望值，並建議是否下注此場比賽讓分盤。

首先將推導主場球隊讓分 (以 k 來表示，k>0 代表主場球隊受讓，k<0 代表主場球隊讓分) 後贏球機率之一般公式，亦即需計算前三節得分分差加第四節得分分差加主場球隊讓分之分數大於零之機率。

若 $\mathbf{X}=\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right)$，則 X_1+X_2 服從常態分配：

$$X_1+X_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$$

因為在讓分盤中，最終比賽會分出勝負，故不會發生平手，所以我們在計算主場球隊讓分後贏球機率時必須限制在非平手情況下即 X_1+X_2≠0。又 $2k$≠0 且 $2k \in$ Z，所以在計算下面 P_1 時，不需要做連續化修正。以下為主場球隊讓分 (k<0) 或受讓分 (k>0) 後贏球機率之一般公式，以 P 來表示：

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

$$
\begin{aligned}
\mathrm{P}_1 &= \Pr(X_1 + X_2 + k > 0) = \Pr(X_1 + X_2 > -k) \\
&= \Pr(Z > \frac{-k - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}) \\
&= 1 - \Phi(\frac{-k - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}})
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P}_2 &= \Pr(X_1 + X_2 = 0) \\
&= \Phi(\frac{0.5 - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}) - \Phi(\frac{-0.5 - (\mu_1 + \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}})
\end{aligned}
$$

$$
\mathrm{P} = \begin{cases} \dfrac{\mathrm{P}_1}{1 - \mathrm{P}_2} & \text{, if } k < 0 \\ \dfrac{\mathrm{P}_1 - \mathrm{P}_2}{1 - \mathrm{P}_2} & \text{, if } k > 0 \end{cases} \tag{5.1}
$$

其中 X_1，X_2 分別代表前三節及第四節得分分差。客場球隊讓分後贏球機率為1-P。

接下來將投注主隊和客隊賠率輸入程式及利用式 (5.1)，即可計算出每投注 1 元，主隊和客隊獲利金額之期望值。以下為投注主隊及客隊獲利金額之期望值之計算公式：

$$
\text{投注主隊獲利金額之期望值} = \text{投注主隊賠率} \times \mathrm{P} - 1
$$

$$
\text{投注客隊獲利金額之期望值} = \text{投注客隊賠率} \times (1 - \mathrm{P}) - 1
$$

再來觀察投注兩隊之獲利金額之期望值建議是否投注該場讓分比賽，當主隊獲利金額之期望值大於客隊獲利金額之期望值且大於零，則建議投注主隊，當客隊獲利金額之期望值大於主隊獲利金額之期望值且大於零，則建議投注客隊，當主隊及客隊獲利金額之期望值皆小於零，則不建議投注。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

以下把上述方法透過流程圖來表示是否投注：

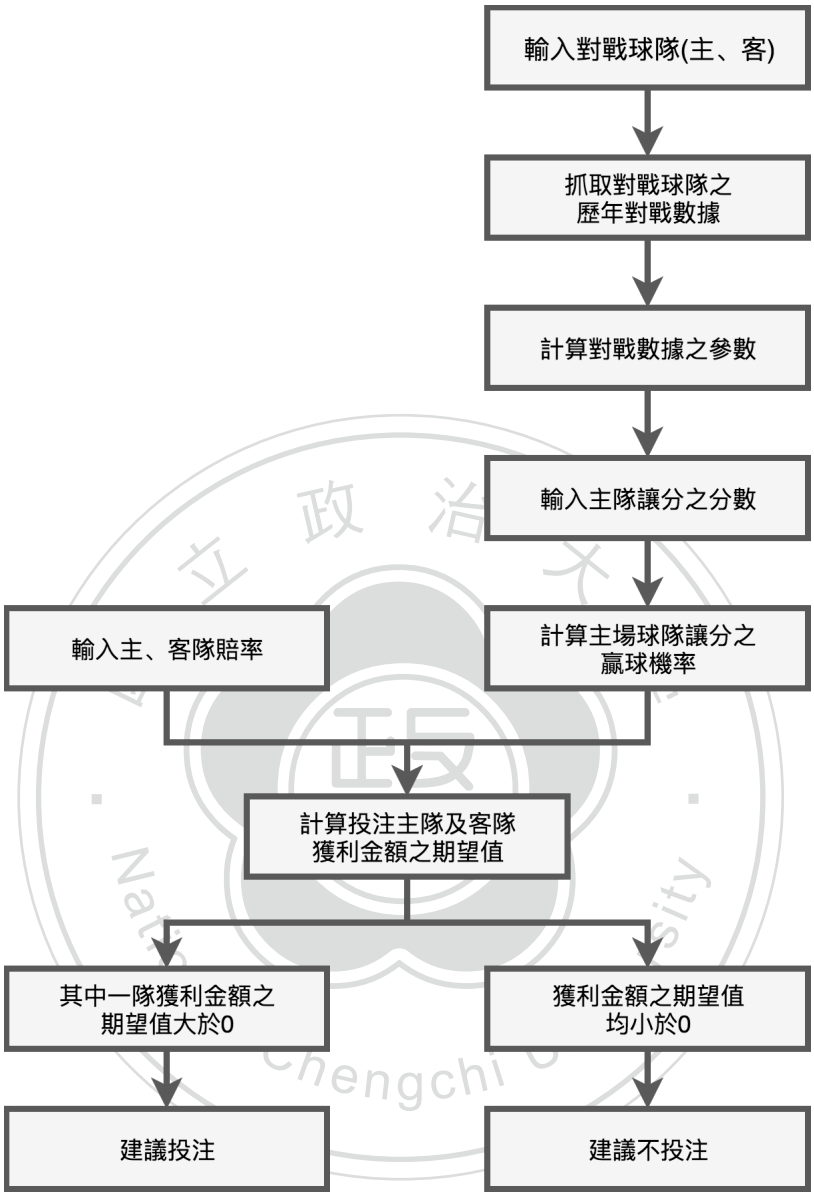

圖. 3.10. 建議投注讓分盤流程圖

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第四章

實證結果 (以 NBA 為例)

第一節　球賽逆轉勝問題分析

本節中使用 NBA 中 2006-2007 賽季至 2017-2018 年賽季共 12 年資料，每年都是利用所有球隊例行賽比賽資料作為數據，每年比賽場次大約 1230 場，本節分析逆轉勝問題，所以在計算上，我們必須先扣除每個賽季中前三節得分分差為平手的場次，再以當年賽季資料計算實際的觀測值，進行二元常態模型與 Gill (2000) 模型之比較。

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. 4.1. 逆轉勝機率及其預測誤差之比較

年份	Gill 模型		二元常態模型		實際觀察值
2006-2007	0.2072	(0.0166)	0.1876	**(0.0030)**	0.1906
2007-2008	0.1869	(0.0125)	0.1776	**(0.0032)**	0.1744
2008-2009	0.1962	(0.0119)	0.1818	**(0.0025)**	0.1843
2009-2010	0.1945	**(0.0076)**	0.1767	(0.0102)	0.1869
2010-2011	0.2002	**(0.0087)**	0.1819	(0.0096)	0.1915
2011-2012	0.1814	(0.0178)	0.1716	**(0.0080)**	0.1636
2012-2013	0.1919	(0.0179)	0.1775	**(0.0035)**	0.1740
2013-2014	0.1901	(0.0232)	0.1738	**(0.0069)**	0.1669
2014-2015	0.1896	(0.0259)	0.1763	**(0.0126)**	0.1637
2015-2016	0.1993	(0.0174)	0.1811	**(0.0008)**	0.1819
2016-2017	0.2011	(0.0190)	0.1799	**(0.0022)**	0.1821
2017-2018	0.2015	(0.0226)	0.1807	**(0.0018)**	0.1789

註：粗體部分為預測誤差較小

依表4.1所示，我們發現除了 2009-2010、2010-2011 兩個賽季利用 Gill (2000) 模型比較靠近實際觀察值之外，其他賽季中使用二元常態模型會比使用 Gill (2000) 模型更靠近實際觀察值。雖然在 2006-2007 賽季的資料不符合二元常態的假設，但預測的結果還是很不錯。所以在加入前三節得分分差及第四節得分分差相關性是有助於更精準預測逆轉勝，因此我們在第二節將利用二元常態模型來預測台灣運動彩券讓分盤結果，觀察是否可以提高我們投注的獲利。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第二節　運彩讓分盤預測結果

本節使用過去幾年的 NBA 資料來預測 2017-2018 例行賽讓分盤之結果。首先抓取對戰球隊不含 2017-2018 之歷史對戰資料作為訓練數據，再利用所算得的獲利金額之期望值，更進一步將投注後可獲得的利潤去建議是否投注。將投注情況分為三種：不分主客隊、主隊、客隊，在不同的每場最少獲利期望值下，計算每下注 1 元，可獲得的利潤 (以下稱為平均獲利)。以下為每投注 1 元實際平均獲利公式：

$$(\text{每下注 1 元})\ \text{實際平均獲利} = \frac{\text{正確投注場次之賠率和} - \text{投注場次數}}{\text{投注場次數}}$$

表. 4.2. 以過去十年資料預測讓分盤之實際投注平均獲利

每場最少	投注對象					
	不分主客隊		主隊		客隊	
獲利期望值	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0	934	-0.0745	498	-0.1084	436	-0.0358
0.05	820	-0.0595	436	-0.0827	384	-0.0331
0.1	705	-0.0456	379	-0.0704	326	-0.0167
0.15	612	-0.0355	327	-0.0575	285	-0.0102
0.2	526	-0.0284	287	-0.0801	239	0.0337
0.25	442	-0.0180	252	-0.0635	190	0.0424
0.35	290	-0.0053	177	-0.0537	113	0.0704
0.4	232	-0.0213	145	-0.0628	87	0.0477
0.45	166	0.0286	107	-0.0266	59	0.1288
0.5	111	0.0009	73	-0.0548	38	0.1079

在實證中，我們分別利用過去兩年資料至過去十年資料分別當作訓練數據，來預測 2017-2018 例行賽 (亦即以今年比賽結果為測試資料) 讓分盤時，在不同的每場最少獲利期望值下，比較實際 (測試) 平均獲利，發現利用資料最多的過去十年資料 (如表4.2) 預測較好。(其餘年份預測之讓分盤結果見附錄 A)

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

當我們提高每場最少獲利期望值時，實際平均獲利也逐漸上升，相對可投注的場次也逐漸變少。而在投注不分主客隊時，雖然將每場最少獲利期望值提高至 0.5，但在預測上得到的實際平均獲利並不是太高，接著我們進一步將投注分成主隊及客隊，在預測主隊時，實際平均獲利卻不盡理想，而在預測客隊時，當每場最少獲利期望值為 0.2 時，建議投注客隊的實際平均獲利方由負轉正，甚至在每場最少獲利期望值為 0.5 時，建議投注客隊的情況下，實際平均獲利可達到 0.1079 元。

在此我們利用重抽樣法 (resampling)，以過去十年資料共有 12059 場比賽，每次去除 500 場比賽 (1000 筆資料)，共分成 25 組資料，在每場最少獲利期望值為 0.5 時，去檢驗預測 2017-2018 年賽季實際平均獲利是否有在預期內，表4.3為每場最少獲利期望值為 0.5 時，重抽樣法之結果。最後將表4.3之結果總結成表4.4，以及利用表4.4建立 95% 信賴區間，如表4.5。

國立政治大學
National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. 4.3. 每場最少獲利期望值為 0.5 重抽樣法之結果

	不分主客隊場次	不分主客隊實際平均獲利	主隊場次	主隊實際平均獲利	客隊場次	客隊實際平均獲利
去除 0 到 1000	130	0.0300	86	-0.0134	44	0.1148
去除 1000 到 2000	120	-0.0892	81	-0.1481	39	0.0333
去除 2000 到 3000	120	0.0146	80	-0.0700	40	0.1838
去除 3000 到 4000	118	0.0017	74	-0.0662	44	0.1159
去除 4000 到 5000	111	0.0167	72	-0.0417	39	0.1244
去除 5000 到 6000	120	0.0300	78	-0.0468	42	0.1726
去除 6000 到 7000	114	-0.0886	73	-0.1534	41	0.0268
去除 7000 到 8000	126	-0.0214	84	-0.0958	42	0.1274
去除 8000 到 9000	114	0.0193	74	-0.0689	40	0.1825
去除 9000 到 10000	123	0.0037	80	-0.0494	43	0.1023
去除 10000 到 11000	118	0.0157	75	-0.0333	43	0.1012
去除 11000 到 12000	117	-0.0222	74	-0.0696	43	0.0593
去除 12000 到 13000	103	-0.0267	66	-0.1182	37	0.1365
去除 13000 到 14000	113	-0.0327	73	-0.0795	40	0.0525
去除 14000 到 15000	110	-0.0382	70	-0.0900	40	0.0525
去除 15000 到 16000	108	-0.0542	72	-0.0924	36	0.0222
去除 16000 到 17000	117	-0.0359	76	-0.1158	41	0.1122
去除 17000 到 18000	108	-0.0204	69	-0.0768	39	0.0795
去除 18000 到 19000	111	0.0320	73	-0.0315	38	0.1539
去除 19000 到 20000	117	-0.0368	77	-0.1052	40	0.0950
去除 20000 到 21000	113	0.0292	74	-0.0446	39	0.1692
去除 21000 到 22000	115	-0.0196	71	-0.0789	44	0.0761
去除 22000 到 23000	114	0.0364	74	-0.0196	40	0.1400
去除 23000 到 24000	120	-0.0154	81	-0.0605	39	0.0782
去除 24000 到 25000	114	0.0053	75	-0.0333	39	0.0795

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. 4.4. 實際投注平均獲利之摘要統計值 (表 4.3)

	實際平均獲利		
	不分主客隊	主隊	客隊
均數	-0.0107	-0.0721	0.1037
標準差	0.0353	0.0370	0.0480
最小值	-0.0892	-0.1534	0.0222
第一四分位數	-0.0327	-0.0924	0.0761
中位數	-0.0154	-0.0696	0.1023
第三四分位數	0.0167	-0.0446	0.1365
最大值	0.0364	-0.0134	0.1838

表. 4.5. 最少獲利期望值為 0.5 元之場次，其實際平均獲利 95% 信賴區間

	實際平均獲利 95% 信賴區間
不分主客隊	(-0.0245,0.0032)
主隊	(-0.0866, -0.0576)
客隊	(0.0849, 0.1225)

從表4.4及表4.5中，發現在建議投注主隊或不分主客隊實際平均獲利的信賴區間為負或有正有負之間，故在這種情況下投注並不是特別有利。但投注客隊的實際平均獲利信賴區間都大於 0，因此投注客隊仍然可以得到不錯的利潤。

最後我們可以發現雖然提高每場最少獲利期望值，但在投注讓分盤所得之實際平均獲利卻並未達到理想的目標。在理論上，當給定每場最少獲利期望值為 0 時，實際平均獲利應為正值，但實際結果為負值，直到每場最少獲利期望值增加至 0.2 時，客隊實際平均獲利方由負轉正。雖然在逆轉勝問題上，二元常態模型表現不錯，但在投注獲利問題上，每兩隊隊伍對戰利用二元常態模型可能不是非常適合。實際上在預測 2017-2018 年賽季時，比賽總場次有 1230 場，但在每場最少獲利期望值為 0 時，預測出來投注場次卻達到 934 場，在預測上或許有高估的情形，導致在實際平均獲利沒有太好。雖然在預測並未到達期望獲利，但在預測客隊時，當每場最少獲利期望值為 0.5 時，在符合建議投注客隊的情況，實際投注 1 元平均獲利可達到 0.1079 元，卻還是可以獲得不錯的利潤。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

第五章

結論與建議

本研究認為在實際的籃球比賽中，前三節兩球隊的分差大小會對第四節兩球隊表現有不同程度地影響。而Gill (2000) 假設主、客場球隊在已知前三節落後情況之下，第四節得以逆轉勝的分數之條件分布為常態，然而此假設未能完整描述比賽情形。為了更貼近實際賽況，我們擴展了Gill (2000) 模型，引入二元常態模型進行逆轉勝機率的預測，發現二元常態模型預測逆轉勝機率會比Gill (2000) 模型更貼近實際的觀察值。

接著將二元常態模型應用至預測台灣運動彩券讓分盤，但在預測讓分盤時，精準度未達理想。在理論上，當給定每場最少獲利期望值為 0 時，平均獲利應為正值，但實際結果為負值，直到每場最少獲利期望值增加至 0.2 時，客隊平均獲利方由負轉正。雖然在逆轉勝問題上，二元常態模型表現不錯，但在投注獲利問題上，每兩隊隊伍對戰利用二元常態模型可能不是非常適合，在預測上或許有高估的情形。雖然在符合建議投注主隊實際平均獲利都為負值，但在每場最少獲利期望值到 0.5 時，符合建議投注客隊只能投注 38 場，在實際每投注 1 元，仍可獲得 0.1079 元的利潤。

在未來研究上，對於某兩球隊對戰，預測主場球隊讓分後獲勝機率之精準度的改善，可透過先經由驗證歷史資料是否符合二元常態，若不符合的話，應先經過適當之資料轉換，再做計算其預測機率，甚至可以尋求更適當之模型。而近幾年來台灣運動彩券新增場中投注之玩法，由於這種玩法的賠率會隨著比賽時間的進行與進球與否，賠率會一直不斷的變動，在這種情況下如有記錄賠率的變動或是在未來有辦法抓取到賠率變動的資料，則可將此模型或新模型更進一步應用到場中投注。

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

附錄 A

不同年份資料預測讓分盤之實際投注平均獲利

表. A.1. 以過去九年資料預測讓分盤之實際投注平均獲利

每場最少獲利期望值	投注對象					
	不分主客隊		主隊		客隊	
	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	932	-0.0668	491	-0.1138	441	-0.0145
0.05	836	-0.0775	442	-0.1149	394	-0.0354
0.1	727	-0.0548	388	-0.0915	339	-0.0128
0.15	633	-0.0422	342	-0.0779	291	-0.0002
0.2	542	-0.0601	293	-0.1046	249	-0.0078
0.25	462	-0.0642	256	-0.1055	206	-0.0129
0.3	388	-0.0711	224	-0.1033	164	-0.0271
0.35	317	-0.0402	187	-0.0853	130	0.0246
0.4	251	-0.0120	155	-0.0671	96	0.0771
0.45	184	-0.0139	122	-0.0156	62	-0.0105
0.5	123	0.0028	81	-0.0179	42	0.0429

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.2. 以過去八年資料預測讓分盤之實際投注平均獲利

每場最少	投注對象					
	不分主客隊		主隊		客隊	
獲利期望值	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	932	-0.0837	488	-0.1265	444	-0.0367
0.05	838	-0.0671	436	-0.1110	402	-0.0194
0.1	740	-0.0648	391	-0.1175	349	-0.0057
0.15	638	-0.0641	339	-0.1119	299	-0.0099
0.2	558	-0.0875	301	-0.1347	257	-0.0323
0.25	473	-0.0868	267	-0.1300	206	-0.0308
0.3	402	-0.0432	235	-0.1013	167	0.0386
0.35	326	-0.0405	185	-0.1043	141	0.0433
0.4	265	-0.0513	161	-0.0693	104	-0.0236
0.45	198	-0.0235	126	-0.0905	72	0.0938
0.5	157	-0.0481	97	-0.1273	60	0.0800

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.3. 以過去七年資料預測讓分盤之實際投注平均獲利

每場最少獲利期望值	投注對象					
	不分主客隊		主隊		客隊	
	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	935	-0.0811	484	-0.1124	451	-0.0475
0.05	845	-0.0831	445	-0.1293	400	-0.0316
0.1	762	-0.0799	398	-0.1237	364	-0.0320
0.15	657	-0.0610	346	-0.1134	311	-0.0026
0.2	578	-0.0733	312	-0.1250	266	-0.0126
0.25	485	-0.0656	268	-0.1323	217	0.0168
0.3	407	-0.0205	228	-0.0886	179	0.0662
0.35	332	-0.0476	196	-0.1010	136	0.0294
0.4	272	-0.0377	163	-0.0816	109	0.0280
0.45	213	-0.0505	133	-0.0981	80	0.0288
0.5	161	-0.0168	101	-0.0569	60	0.0508

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.4. 以過去六年資料預測讓分盤之實際投注平均獲利

每場最少	投注對象					
	不分主客隊		主隊		客隊	
獲利期望值	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	956	-0.0832	499	-0.1147	457	-0.0487
0.05	860	-0.0827	449	-0.1293	411	-0.0319
0.1	758	-0.0705	398	-0.1152	360	-0.0211
0.15	665	-0.0589	350	-0.1236	315	0.0130
0.2	561	-0.0540	298	-0.1243	263	0.0257
0.25	479	-0.0391	263	-0.1017	216	0.0370
0.3	404	-0.0561	229	-0.0921	175	-0.0089
0.35	350	-0.0307	200	-0.0485	150	-0.0070
0.4	295	-0.0588	167	-0.0614	128	-0.0555
0.45	225	-0.0713	131	-0.0729	94	-0.0691
0.5	173	-0.0353	107	-0.0785	66	0.0348

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.5. 以過去五年資料預測讓分盤之實際投注平均獲利

每場最少獲利期望值	投注對象					
	不分主客隊		主隊		客隊	
	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	966	-0.0818	504	-0.1304	462	-0.0288
0.05	875	-0.0827	463	-0.1177	412	-0.0434
0.1	782	-0.0747	416	-0.1234	366	-0.0193
0.15	672	-0.0613	353	-0.1310	319	0.0158
0.2	570	-0.0475	307	-0.1101	263	0.0257
0.25	493	-0.0481	273	-0.0960	220	0.0114
0.3	422	-0.0714	236	-0.1123	186	-0.0196
0.35	352	-0.0663	195	-0.0792	157	-0.0503
0.4	289	-0.0642	163	-0.0721	126	-0.0540
0.45	231	-0.1264	132	-0.1208	99	-0.1338
0.5	183	-0.1270	110	-0.0723	73	-0.2096

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.6. 以過去四年資料預測讓分盤之實際投注平均獲利

每場最少獲利期望值	投注對象					
	不分主客隊		主隊		客隊	
	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	984	-0.0881	518	-0.1501	466	-0.0192
0.05	888	-0.0810	465	-0.1259	423	-0.0316
0.1	790	-0.0980	413	-0.1517	377	-0.0391
0.15	696	-0.0970	362	-0.1539	334	-0.0353
0.2	610	-0.0955	315	-0.1443	295	-0.0434
0.25	536	-0.0992	278	-0.1390	258	-0.0562
0.3	457	-0.0935	244	-0.1209	213	-0.0622
0.35	397	-0.0943	210	-0.0879	187	-0.1016
0.4	335	-0.0943	172	-0.0698	163	-0.1202
0.45	262	-0.0750	137	-0.0369	125	-0.1168
0.5	201	-0.0580	107	0.0173	94	-0.1436

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.7. 以過去三年資料預測讓分盤之實際投注平均獲利

每場最少獲利期望值	投注對象					
	不分主客隊		主隊		客隊	
	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	991	-0.1189	531	-0.1789	460	-0.0496
0.05	914	-0.1269	486	-0.1785	428	-0.0682
0.1	826	-0.1211	439	-0.1786	387	-0.0558
0.15	744	-0.1114	395	-0.1543	349	-0.0629
0.2	666	-0.1096	362	-0.1352	304	-0.0791
0.25	574	-0.1345	306	-0.1600	268	-0.1054
0.3	508	-0.1361	273	-0.1557	235	-0.1134
0.35	437	-0.1240	233	-0.1232	204	-0.1250
0.4	375	-0.1205	201	-0.1080	174	-0.1351
0.45	316	-0.1062	168	-0.0693	148	-0.1480
0.5	270	-0.0769	142	-0.0345	128	-0.1238

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

表. A.8. 以過去二年資料預測讓分盤之實際投注平均獲利

每場最少	投注對象					
	不分主客隊		主隊		客隊	
獲利期望值	場次	實際平均獲利	場次	實際平均獲利	場次	實際平均獲利
0.0	1029	-0.1241	562	-0.1685	467	-0.0707
0.05	951	-0.1185	515	-0.1642	436	-0.0646
0.1	860	-0.1230	462	-0.1554	398	-0.0853
0.15	772	-0.1277	415	-0.1527	357	-0.0986
0.2	692	-0.1434	376	-0.1773	316	-0.1030
0.25	644	-0.1504	353	-0.1688	291	-0.1280
0.3	581	-0.1464	317	-0.1582	264	-0.1322
0.35	521	-0.1417	291	-0.1552	230	-0.1248
0.4	465	-0.1367	261	-0.1523	204	-0.1167
0.45	413	-0.1334	235	-0.1404	178	-0.1242
0.5	358	-0.1175	194	-0.1299	164	-0.1027

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

附錄 B

程式碼

爬蟲抓取 NBA 資料

```
#######################引入套件#######################
#install.packages("rvest")
#install.packages("plyr")
library(rvest)
library(plyr)
#################### 抓取資料執行 ##################
##建立網址
URL_part <- c("https://www.basketball-reference.com/boxscores/?month=",
              "&day=",
              "&year=")
path <-data.frame(DT = character(),URL = character())
for(y in 2017:2018){
  for(m in 1:12){
    for(d in 1:31){
      path <- rbind(path,
                    data.frame(DT =paste(y,"_",m,"_",d,sep=""),
                               URL =paste(URL_part[1],m,URL_part[2],d,URL_
                                   part[3],y,sep="")))
    }
  }
}
DATA <- list()
for(i in 1:length(path[[1]])){
  doc <- read_html(as.character(path[[2]][i]))
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
  if((length(doc %>% html_nodes(".teams+ table"))) != 0){
    DATA[[as.character(path[[1]][i])]] <- doc %>% html_nodes(".teams+ table") %>% html_table()
    for(j in 1:length(DATA[[as.character(path[[1]][i])]])){
      rownames(DATA[[as.character(path[[1]][i])]][[j]]) <- DATA[[as.character(path[[1]][i])]][[j]][,1]
      DATA[[as.character(path[[1]][i])]][[j]] <- DATA[[as.character(path[[1]][i])]][[j]][,-1]
      DATA[[as.character(path[[1]][i])]][[j]] <- cbind(DATA[[as.character(path[[1]][i])]][[j]],Total = c(sum(DATA[[as.character(path[[1]][i])]][[j]][1,]),sum(DATA[[as.character(path[[1]][i])]][[j]][2,])))
    }
  }else {
    DATA[[as.character(path[[1]][i])]] <-(paste(path[[1]][i],"Do not have any data."))
  }
  Sys.sleep(round(runif(1,min = 1 ,max = 5)))
}
#------------------------資料整理------------------------------------------------
#-------------X月-------------
MonthX <- data.frame('1' = character(),'2' = character(),'3' = character(),'4' = character(),'1OT' = character(),Total = character())
datalist <- list()
for(d in 1:31){
  if(is.list(DATA[[paste0("2017_X_",d)]])){
    for(b in 1:length(DATA[[paste0("2017_X_",d)]])){
      datalist <- c(datalist,DATA[[paste0("2017_X_",d)]][b])
    }
  }
}
MonthX <- rbind.fill(datalist)
MonthX$teamname <- unlist(lapply(datalist, row.names))
MonthX <- MonthX[c('teamname',1,2,3,4,'1OT','Total')]
write.csv(MonthX,"MonthX-2017.csv")
```

抓取讓分、賠率程式碼

```
#引入套件
import pandas as pd
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
from urllib.request import quote
from urllib.request import urlopen
from urllib.request import Request
from bs4 import BeautifulSoup
import numpy as np
import time
from urllib.error import HTTPError, URLError
from datetime import date
from dateutil.rrule import rrule, DAILY
import time
import random
#引用網址
chaper_url_head = 'https://www.playsport.cc/predictgame.php?action=result&alliance
    id=3&gametime='
headers = {'User-Agent':'Mozilla/5.0 (Windows NT 6.1; WOW64; rv:23.0) Gecko
    /20100101 Firefox/23.0'}
#輸入賽季時間、並抓取每場讓分盤之讓分
DD = {}
a = date(2017, 10, 18)
b = date(2018, 4, 12)
for dt in rrule(DAILY, dtstart=a, until=b):
    time.sleep(2+random.random())
    req = Request(url=chaper_url_head+dt.strftime("%Y%m%d"), headers=
        headers)
    A = urlopen(req).read()
    data = pd.read_html(str(BeautifulSoup(A,"lxml").select('.gamedata-
        results')),encoding='utf-8')[0]
    if len(data.iloc[2,0]) != 5:
        D = pd.DataFrame(columns=["分數", "Team", "讓分"])
        for i in range(int((len(data)-2)/3)):
            try :
                D = pd.concat([D,pd.DataFrame([[dt.strftime("%Y%m%d"),data.
                    iloc[3+3*i,0].split()[0],data.iloc[2+3*i,3],data.iloc
                    [2+3*i,8]],[dt.strftime("%Y%m%d"),data.iloc[3+3*i,0].
                    split()[2],data.iloc[2+3*i,5],data.iloc[6+3*i,2]]],
                    columns=["日期", "分數", "Team", "讓分"])],axis=0,
                    ignore_index=True)
            except:
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
                        k = 0
        DD[dt.strftime("%Y%m%d")] = D
#賠率先塞空值
Data_coll_2017_2018['賠率']=np.nan
#抓取讓分盤之主、客隊賠率
for i in range(len(Data_coll_2017_2018.loc[:,'賠率'])):
    if Data_coll_2017_2018.loc[i, '讓分'] is not None:
        Data_coll_2017_2018.loc[i,'賠率'] = Data_coll_2017_2018.loc[i,'讓分'].split()[1]
#將抓取主隊讓分資的部分
for i in range(len(Data_coll_2017_2018.loc[:,'讓分'])):
    if Data_coll_2017_2018.loc[i,'讓分'] is not None:
        Data_coll_2017_2018.loc[i,'讓分'] = Data_coll_2017_2018.loc[i,'讓分'].split()[0][1:-1]
    if i % 2 == 0:
        Data_coll_2017_2018.loc[i,'讓分'] = np.nan
```

逆轉勝問題程式碼

```
#讀取套件
import numpy as np
from scipy.linalg import fractional_matrix_power
from scipy.stats import norm
import sympy
import sympy.stats
import scipy.integrate as integrate
import math
import matplotlib.pyplot as plt
import pandas as pd
get_ipython().magic('matplotlib inline')
#讀取當年賽季資料
data = pd.read_excel('2017-2018NBA資料.xlsx')
data = data.reset_index()
#去除當前三節得分分差為0的場次
data = data.dropna()[data.Z != 0]
#計算二元常態模型參數
mu_Z   = (data['Z']).mean()
mu_W   = (data['W']).mean()
var_Z  = (data['Z']).var()
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
var_W   = (data['W']).var()
cor     = data.loc[:,['Z','W']].corr().loc['Z','W']
cov_ZW  = cor*np.sqrt(var_Z)*np.sqrt(var_W)
M       = np.array([[var_Z, cov_ZW], [cov_ZW, var_W]])
data_mu = np.array([[mu_Z],[mu_W]])
#計算eigenvalue and eigenvector、最後得出之X1,x2與Z1,2關係式
def eigenvalue_and_eigenvector(M,data_mu):
    M_eigva,M_eigvt = np.linalg.eig(M)
    #Swapping th columns
    M_eigvt[:,[0, 1]] = -M_eigvt[:,[1, 0]]
    M_eigva[[0, 1]]   =  M_eigva[[1, 0]]
    #將eigenvalue轉為對角矩陣
    M_eigva2 = np.diag(M_eigva)
    #計算D^-1/2
    D_half_inverse = fractional_matrix_power(M_eigva2,-0.5)
    D_half_inverse_dot_M_eigvt = np.dot(D_half_inverse,M_eigvt)
    D_half_inverse_dot_neg_data_mu = np.dot(D_half_inverse_dot_M_eigvt,-data_mu)
    inverse_of_D_half_inverse_dot_M_eigvt = np.linalg.inv(D_half_inverse_dot_M_eigvt)
    inverse_of_D_half_inverse_dot_D_half_inverse_dot_M_eigvt_dot_data_mu = np.dot(inverse_of_D_half_inverse_dot_M_eigvt,-D_half_inverse_dot_neg_data_mu)
    return inverse_of_D_half_inverse_dot_M_eigvt, inverse_of_D_half_inverse_dot_D_half_inverse_dot_M_eigvt_dot_data_mu
#計算二元常態模型預測之逆轉勝機率
def integral(inverse_of_D_half_inverse_dot_M_eigvt, inverse_of_D_half_inverse_dot_D_half_inverse_dot_M_eigvt_dot_data_mu):
    x,y = sympy.symbols('x y')
    #將矩陣輸入sympy來做符號計算
    A = sympy.Matrix(inverse_of_D_half_inverse_dot_M_eigvt[[0],:])
    # where A is coefficient of Z=0
    B = sympy.Matrix(inverse_of_D_half_inverse_dot_M_eigvt[[1],:])
    c = sympy.Matrix(inverse_of_D_half_inverse_dot_D_half_inverse_dot_M_eigvt_dot_data_mu)
    st = v = sympy.Matrix([[x,y]]).T
    ### Home
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
    #計算兩不等式的交點，以求積分區域
    cross = sympy.solve([(A*st+c.row(0))[0] , (A*st+c.row(0))[0]-(-B*st-c.row(1))[0]],[x,y])
    #k是交點的x值
    k=cross[x]
    #建立參數式
    Z          = sympy.solve((A*st+c.row(0))[0] ,y)
    Z_plus_W = sympy.solve((A*st+c.row(0))[0]-(-B*st-c.row(1))[0],y)
    #計算積分
    f = Z[0]
    g = sympy.lambdify(x, f, "numpy")
    h = lambda x: 1/(np.sqrt(2*np.pi))*np.exp((-1*x**2)/2)*norm.cdf(g(x))
    sol1 = integrate.quad(h, -math.inf,k)
    f = Z_plus_W[0]
    g = sympy.lambdify(x, f, "numpy")
    h = lambda x: 1/(np.sqrt(2*np.pi))*np.exp((-1*x**2)/2)*norm.cdf(g(x))
    sol2 = integrate.quad(h,k,math.inf)
    H = sol1[0]+sol2[0]
    ### AWAY
    #建立參數式
    Z          = sympy.solve((A*st+c.row(0))[0] ,y)
    Z_plus_W = sympy.solve((A*st+c.row(0))[0]-(-B*st-c.row(1))[0],y)
    #計算積分
    f = Z_plus_W[0]
    g = sympy.lambdify(x, f, "numpy")
    h = lambda x: 1/(np.sqrt(2*np.pi))*np.exp((-1*x**2)/2)*(1-norm.cdf(g(x)))
    sol1 = integrate.quad(h, -math.inf,k)
    f = Z[0]
    g = sympy.lambdify(x, f, "numpy")
    h = lambda x: 1/(np.sqrt(2*np.pi))*np.exp((-1*x**2)/2)*(1-norm.cdf(g(x)))
    sol2 = integrate.quad(h,k,math.inf)
    A = sol1[0]+sol2[0]
    return H,A
#計算Gill(2000)模型預測之逆轉勝機率
def Pr_hn(data):
    PR_Hn = list()
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
    for n in list(range(1,int(max(data.Z)+1))):
        A =  norm.cdf((n + 0.5 - np.mean(data.Z))/np.std(data.Z))-norm.cdf
            ((n - 0.5 - np.mean(data.Z))/np.std(data.Z))
        B = 1 - norm.cdf((0.5 - np.mean(data.Z))/np.std(data.Z)) + norm.cdf
            ((-0.5 - np.mean(data.Z))/np.std(data.Z))
        PR_Hn.append(A / B)
    A = 0
    for n in list(range(0,len(PR_Hn))):
        if (n+1) in list(data.Z):
            A = A + PR_Hn[n]
    A
    return PR_Hn
def Pr_an(data):
    PR_An = list()
    for n in list(range(1,int(-min(data.Z)+1))):
        A =  norm.cdf((-n + 0.5 - np.mean(data.Z))/np.std(data.Z))-norm.cdf
            ((-n - 0.5 - np.mean(data.Z))/np.std(data.Z))
        B = 1 - norm.cdf((0.5 - np.mean(data.Z))/np.std(data.Z)) + norm.cdf
            ((-0.5 - np.mean(data.Z))/np.std(data.Z))
        PR_An.append(A / B)
    B = 0
    for n in list(range(0,len(PR_An))):
        if n+1 in list(-data.Z):
            B = B + PR_An[n]
    B
    return PR_An
def Pr_hw(data,PR_Hn,PR_An):
    Pr_HlHw = list()
    for n in list(range(1,int(max(data.Z)+1))):
        if n in list(data.Z):
            Pr_HlHw.append((1 - norm.cdf((-n - data[data.Z>0].W.mean()) /
                data[data.Z>0].W.std()))*PR_Hn[n-1])
    Pr_Hw_numerator = np.sum(Pr_HlHw)
    Pr_AlAw = list()
    for n in list(range(1,int(-min(data.Z)+1))):
        if n in list(-data.Z):
            Pr_AlAw.append((norm.cdf((n - data[data.Z<0].W.mean()) / data[
                data.Z<0].W.std()))*PR_An[n-1])
```

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
    Pr_Aw_numerator = np.sum(Pr_AlAw)
    return (Pr_Aw_numerator+Pr_Hw_numerator)
#二元常態模型計算不被逆轉機率
ml, m2 = eigenvalue_and_eigenvector(M, data_mu)
H,A = integral(ml,m2)
#Gill(2000)模型計算不被逆轉機率
PR_Hn = Pr_hn(data)
PR_An = Pr_an(data)
#觀察值不被逆轉機率
data[data.Z+data.W>0][data.Z>0]
#Z>0 & Z+W>0 叫出對應場次及 Z & W
Home =len(data[data.Z+data.W>0][data.Z>0])
#主隊前三節領先情況下，主隊獲勝場次數
Away =len(data[data.Z+data.W<0][data.Z<0])
#客隊前三節領先情況下，客隊獲勝場次數
Total=Home+Away
```

計算讓分盤程式碼

```
#讀取套件
import numpy as np
from scipy.linalg import fractional_matrix_power
from scipy.stats import norm
import sympy
import sympy.stats
import scipy.integrate as integrate
import math
import matplotlib.pyplot as plt
import pandas as pd
get_ipython().magic('matplotlib␣inline')
#讀取歷年NBA資料
data = pd.read_excel('近X年NBA資料.xlsx')
data = data.reset_index()
#讀取2017–2018賽事資料(包含每場讓分及賠率)
new_data= pd.read_excel('new_data.xlsx')
#抓取對戰歷年數據、計算主場球隊讓分後贏球機率
def prob(inA,inB,data,hometeam_for_point):
    mergeB = [True if ((data.iloc[i,1] == inA) & (data.iloc[i+1,1] ==
        inB) & (i%2==0))
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
                        or ((data.iloc[i,1] == inB) & (data.iloc[i-1,1] ==
                            inA) & (i%2!=0))
                else False for i in range(0, len(data)-1)]
    Home_team_data = data[0:len(data)-1][mergeB]
    mu_Z   = (Home_team_data['Z']).mean()
    mu_W   = (Home_team_data['W']).mean()
    var_Z  = (Home_team_data['Z']).var()
    var_W  = (Home_team_data['W']).var()
    cor    = Home_team_data.loc[:,['Z','W']].corr().loc['Z','W']
    cov_ZW = cor*np.sqrt(var_Z)*np.sqrt(var_W)
    M      = np.array([[var_Z, cov_ZW],
                       [cov_ZW, var_W]])
    data_mu = np.array([[mu_Z],[mu_W]])
    P1 = 1-norm.cdf((-hometeam_for_point-mu_Z-mu_W)/np.sqrt(M.sum()))
    P2 = norm.cdf((0.5-mu_Z-mu_W)/np.sqrt(M.sum()))-norm.cdf((-0.5-mu_Z-
        mu_W)/np.sqrt(M.sum()))
    if hometeam_for_point<0:
        hometeamwinpr=P1/(1-P2)
    else:
        hometeamwinpr=(P1-P2)/(1-P2)
    return hometeamwinpr
#win_or_lose為2017_2018賽季每一場球賽的真實資料(紀錄主隊or客隊贏球)
win_or_lost = []
for i in range(int(len(new_data)/2)):
    if (new_data.loc[1+2*i,'Total'] + new_data.loc[1+2*i,'讓分']) >
        new_data.loc[0+2*i,'Total']:
        win_or_lost += [1]
    else:
        win_or_lost += [0]
#2017_2018賽事在不同期望最小獲利值建議下注場次和期望獲利金額
Different_threshold = dict()
for j in np.arange(.1,0.55,0.05):
    pred = []
    for i in range(int(len(new_data)/2)):
        At=new_data.loc[0+2*i,'team0me']
        Bt=new_data.loc[1+2*i,'team0me']
        hometeam_for_point = new_data.loc[1+2*i, '讓分']
        try:
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
            pr_homewin = prob(At,Bt,data,hometeam_for_point)
            Away_expected =new_data.loc[0+2*i,'賠率']*(1-pr_homewin)-1
            Home_expected =new_data.loc[1+2*i,'賠率']*pr_homewin-1
            if (Home_expected>Away_expected) & (Home_expected>j):#1print('下主隊')
                pred += [1]
            elif (Away_expected>Home_expected) & (Away_expected>j):#0print('下客隊')
                pred += [0]
            else:
                pred += [np.nan]
        except:
            pred += [np.nan]

    不分主客隊場次 = len(np.array(pred)[pd.notnull(pred)])
    主隊場次 = len(np.array(pred)[np.array(pred)==1])
    客隊場次 = len(np.array(pred)[np.array(pred)==0])

    merge_0 = []
    for i in np.array(pred)==0:
        merge_0 = merge_0 + [i, False]
    odds_to_away = new_data[merge_0].loc[:,"賠率"].values
    oppsite_to_endorse = -(np.array(win_or_lost)[np.array(pred)==0]-1)
    客隊平均單位獲利 = (odds_to_away@oppsite_to_endorse-客隊場次)/客隊場次

    merge_1 = []
    for i in np.array(pred)==1:
        merge_1 = merge_1 + [False, i]
    odds_to_home = new_data[merge_1].loc[:,"賠率"].values
    oppsite_to_endorse = (np.array(win_or_lost)[np.array(pred)==1])
    主隊平均單位獲利 = (odds_to_home@oppsite_to_endorse-主隊場次)/主隊場次

    不分主客隊單位獲利 = (客隊平均單位獲利*客隊場次+主隊平均單位獲利*主隊場次)/不分主客隊場次
    Different_threshold["Threshold␣=␣" + str(j)] = pd.DataFrame([[不分主客隊場次,不分主客隊單位獲利,主隊場次,主隊平均單位獲利,客隊場次,客隊平均單位獲利]],columns = ['不分主客隊場次','不分主客隊單位獲利','主隊場次','主隊平均單位獲利','客隊場次','客隊平均單位獲利'])
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
#將不同期望最小獲利值所計算出的場次、期望獲利金額彙整
end = .55
collection = pd.DataFrame(columns = ['不分主客隊場次','不分主客隊單位獲利',
    '主隊場次','主隊平均單位獲利','客隊場次','客隊平均單位獲利'])
for j in np.arange(0,end,.05):
    collection = pd.concat([collection, Different_threshold["Threshold␣=␣"
        + str(j)]])
index = [ str(j) for j in np.arange(0,end,.05)]
collection.index = index
```

計算單場讓分盤程式碼

```
import numpy as np
from scipy.linalg import fractional_matrix_power
from scipy.stats import norm
import sympy
import sympy.stats
import scipy.integrate as integrate
import math
import matplotlib.pyplot as plt
import pandas as pd
get_ipython().magic('matplotlib␣inline')
#讀取歷年NBA資料
data = pd.read_excel('近十年NBA資料.xlsx')
data = data.reset_index()
#  # 球隊名稱
#波士頓塞爾蒂克 = Boston  布魯克林籃網 = Brooklyn  紐約尼克 = New York
#費城76人 = Philadelphia 多倫多暴龍 = Toronto  芝加哥公牛 = Chicago
#克里夫蘭騎士 =Cleveland  底特律活塞 = Detroit  印地安那溜馬 = India0
#密爾瓦基公鹿 =Milwaukee  亞特蘭大老鷹= Atlanta  夏洛特黃蜂 = Charlotte
#邁阿密熱火 =Miami  奧蘭多魔術 = Orlando  華盛頓巫師 = Washington
#達拉斯獨行俠 = Dallas  休士頓火箭 = Houston  曼菲斯灰熊 = Memphis
#紐澳良鵜鶘 =New Orleans 聖安東尼奧馬刺= San Antonio 丹佛金塊 = Denver
#明尼蘇達灰狼= Minnesota  波特蘭拓荒者 = Portland  猶他爵士 = Utah
#奧克拉荷馬雷霆 = Oklahoma City 金州勇士 = Golden State
#洛杉磯快艇 = LA Clippers  洛磯山湖人=LA Lakers
#鳳凰城太陽 = Phoenix  沙加緬度國王 = Sacramento
#請輸入任意兩隊隊伍名稱
客場球隊=input("客場球隊名稱：")
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
主場球隊=input("主場球隊名稱：")
#抓取任意 A & B 兩隊對戰資料
merge = [True  if  ((data.iloc[i,1] == 客場球隊) & (data.iloc[i+1,1] == 主場球隊) & (i%2==0))
                or ((data.iloc[i,1] == 主場球隊) & (data.iloc[i-1,1] == 客場球隊) & (i%2!=0))
                or ((data.iloc[i,1] == 主場球隊) & (data.iloc[i+1,1] == 客場球隊) & (i%2==0))
                or ((data.iloc[i,1] == 客場球隊) & (data.iloc[i-1,1] == 主場球隊) & (i%2!=0))
            else False for i in range(0, len(data)-1)]
data[0:len(data)-1][merge]
#抓取 'B隊' 為主場資料場次
mergeB = [True  if  ((data.iloc[i,1] == 客場球隊) & (data.iloc[i+1,1] == 主場球隊) & (i%2==0))
                or ((data.iloc[i,1] == 主場球隊) & (data.iloc[i-1,1] == 客場球隊) & (i%2!=0))
            else False for i in range(0, len(data)-1)]
Home_team_data = data[0:len(data)-1][mergeB]
Home_team_data
#資料平均數、變異數、相關係數
mu_Z  = (Home_team_data['Z']).mean()
mu_W  = (Home_team_data['W']).mean()
var_Z  = (Home_team_data['Z']).var()
var_W  = (Home_team_data['W']).var()
cor  = Home_team_data.loc[:,['Z','W']].corr().loc['Z','W']
cov_ZW = cor*np.sqrt(var_Z)*np.sqrt(var_W)
M  = np.array([[var_Z, cov_ZW], [cov_ZW, var_W]])
data_mu = np.array([[mu_Z],[mu_W]])
#抓取對戰歷年數據、計算主場球隊讓分後贏球機率
def prob(客場球隊,主場球隊,data,主場球隊讓分之分數):
    mergeB = [True  if  ((data.iloc[i,1] == 客場球隊) & (data.iloc[i+1,1] == 主場球隊) & (i%2==0))
                    or ((data.iloc[i,1] == 主場球隊) & (data.iloc[i-1,1] == 客場球隊) & (i%2!=0))
            else False for i in range(0, len(data)-1)]
    Home_team_data = data[0:len(data)-1][mergeB]
    mu_Z  = (Home_team_data['Z']).mean()
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
    mu_W    = (Home_team_data['W']).mean()
    var_Z   = (Home_team_data['Z']).var()
    var_W   = (Home_team_data['W']).var()
    cor     = Home_team_data.loc[:,['Z','W']].corr().loc['Z','W']
    cov_ZW  = cor*np.sqrt(var_Z)*np.sqrt(var_W)
    M       = np.array([[var_Z, cov_ZW],
                        [cov_ZW, var_W]])
    data_mu = np.array([[mu_Z],[mu_W]])
    P1 = 1-norm.cdf((-主場球隊讓分之分數-mu_Z-mu_W)/np.sqrt(M.sum()))
    P2 = norm.cdf((0.5-mu_Z-mu_W)/np.sqrt(M.sum()))-norm.cdf((-0.5-mu_Z-
        mu_W)/np.sqrt(M.sum()))
    if 主場球隊讓分之分數<0:
        hometeamwinpr=P1/(1-P2)
    else:
        hometeamwinpr=(P1-P2)/(1-P2)
    return hometeamwinpr
#輸入主場讓分，主場球隊讓分之分數 = 主場讓分(用 - 表示)、主場被讓(用 + 表示
    )
主場球隊讓分之分數=float(input("主場球隊讓分之分數:"))
#輸入下注主隊及客隊賠率
投注主隊之賠率=float(input('投注主隊之賠率='))
投注客隊之賠率=float(input('投注客隊之賠率='))
#主場球隊讓分後贏球機率
prob(客場球隊,主場球隊,data,主場球隊讓分之分數)
print('主場球隊贏球機率=',prob(客場球隊,主場球隊,data,主場球隊讓分之分數))
#下注主場球和客場隊期望獲利金值
Home_expected =投注主隊之賠率*prob(客場球隊,主場球隊,data,主場球隊讓分之分
    數)-1
print('投注主隊期望值=',Home_expected)
Away_expected =投注客隊之賠率*(1-prob(客場球隊,主場球隊,data,主場球隊讓分之
    分數))-1
print('投注客隊期望值=',Away_expected)
if (((1-prob(客場球隊,主場球隊,data,主場球隊讓分之分數))*投注客隊之賠率
    -1>0) and ((1-prob(客場球隊,主場球隊,data,主場球隊讓分之分數))*投注客隊
    之賠率-1<0.2)):
    print('可以投注客隊，但獲利機率不大，風險自行承擔')
elif(((1-prob(客場球隊,主場球隊,data,主場球隊讓分之分數))*投注客隊之賠率
    -1>0.2) and ((1-prob(客場球隊,主場球隊,data,主場球隊讓分之分數))*投注客
```


DOI:10.6814/THE.NCCU.MATH.002.2018.B01

```
    隊之賠率-1<0.5)):
        print("建議投注客隊，風險自行承擔")
elif(1-prob(客場球隊,主場球隊,data,主場球隊讓分之分數))*投注客隊之賠率
    -1>0.5:
        print("建議投注客隊，獲利機率高，風險自行承擔")
else:
        print('此比賽不建議投注')
```

國立政治大學
National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01

參考文獻

Cooper, H., DeNeve, K. M., and Mosteller, F. (1992). Predicting professional sports game outcomes from intermediate game scores. *CHANCE*.

Gill, P. S. (2000). Late-game reversals in professional basketball, football, and hockey. *The American Statistician*, 54(2):94–99.

台灣運動彩券 (2008a)。台灣運彩銷售破百億-半年挹注運動發展基金及公益 10 億；https://www.sportslottery.com.tw/zh/web/guest/news-release。

台灣運動彩券 (2008b)。運動彩券投注規範；https://www.sportslottery.com.tw/zh/web/guest/sports-regulation。

林桓 (2008)。運動特種公益彩券管理辦法相關法律問題。*行政院體育委員會研究報告 Sac-Res-096-02*。

周錫洋 (2005)。運動彩券之行銷策略研究。*台北大學企業管理學系碩士在職進修專班*。

國立政治大學 National Chengchi University

DOI:10.6814/THE.NCCU.MATH.002.2018.B01