

國立屏東大學應用數學系碩士班  
碩士論文

Department of Applied Mathematics

National Pingtung University

Master's Thesis

職棒賽事勝負之預測—以中華職棒二十八年賽  
季為例

Predict the Results of Professional Baseball Games—Take the  
28th season of Chinese Professional Baseball League for

The logo of National Pingtung University (NPTU) is centered on the page. It features a stylized mountain range in blue and green, with a sunburst or circular pattern in orange and yellow above it. Below the graphic, the letters 'NPTU' are written in a bold, blue, sans-serif font, and the word 'Example' is written in a smaller, black, sans-serif font directly underneath.

Example

指導教授：蔡典龍 博士

Advisor: Dr. Tien-Lung Tsai

研究生：林清榆

Student: Cing-Yu Lin

中華民國 106 年 7 月

July 2018

國立屏東大學應用數學系碩士班  
碩士論文

研究生：林清榆

職棒賽事勝負之預測——以中華職棒二十八年賽季為例

本論文業經審查及口試合格特此證明

碩士學位考試委員會 主席

蔣永川

委員

蔡典龍

委員

鄭品源

指導教授：蔡典龍 博士

蔡典龍

系主任：吳進通 博士

吳進通

中華民國 107 年 7 月 4 日

## 誌謝

本論文可以順利完成，特別感謝指導教授蔡典龍老師，研究所的兩年中在各方面指導我，在撰寫論文時提供我許多的協助與建議，除了學業上的幫助外，也提供了對未來工作及生活上的些許建議。這邊也感謝口試委員的鄭昌源老師與蔡永川老師，特地撥空進行論文審閱，並且在口試期間提供許多建議，使得論文更加的完善。

另外感謝研究所的同學旻學、翊誠、翊華、宜臻，在就學期間互相幫忙及關心，讓我能夠順利完成碩士學位。最後感謝我的父母，在他們支持下無後顧之憂地完成學業。

林清榆 謹致於

國立屏東大學應用數學系

中華民國 107 年 7 月

## 摘要

自從棒球確認回歸奧運會後，再次掀起了一股棒球風潮，尤其在 2017 年有許多重大的國際賽事，提到職業運動，除了欣賞刺激比賽外，運動彩券是個可以預測比賽勝負，額外獲得彩金的方法。預測職棒比賽結果的對象為中華職棒大聯盟 2017 年賽季(中華職棒二十八年)的四支球隊，將採用三種解釋變數處理方式，是為打擊成績與投手成績、打擊成績比與投手成績比、打擊成績差與投手成績差，在分別對 4 個時間點(賽季進行 20%、賽季進行 40%、賽季進行 70%、賽季進行 90%)建立邏輯斯迴歸模型，再對其後 10 場比賽以 6 種預測方式進行預測驗證。選擇解釋變數的方式，採用 Wald 統計量逐步選取。另外將賽季 70%、90%分別對下半季 20%、40%進行比較，也針對全賽季前一場與前面全賽季的三種解釋變數處理方式對照，判別其中差異。

研究結果三種解釋變數處理方式對 6 種預測方式的最高平均預測成功率分別是 70%、72.5%、67.5%，其中全賽季 70%、90%略優於下半季 20%、40%，而前一場的處理方式預測成功率也未必比這三種解釋變數處理方式好；預測方式則是以兩隊上次交手數據做為解釋變數時，預測成功率在三個解釋變數處理方式中都是最好的；選入變數方面，包含下半季及前一場共選入了 22 種變數，其中最高的三個為每打數平均打點、防禦率、三振四壞比。

關鍵詞：邏輯斯迴歸、職業棒球、預測

## Abstract

Since the baseball confirmed returns the Olympic Games, it has raised a baseball unrest once more, especially has many important international sports events in 2017, and mentioned the professional movement, besides the appreciation stimulation competition, the movement lottery ticket is may forecast the competition victory and defeat, extra obtains the lottery prize money the method.

Forecasts the Professional baseball competition result the object for the Chinese Professional Baseball League in 2017 the season (CPBL 28 years) four teams, we will use three explanation variable processing mode, that is batting score and pitching score, the division of batting score and pitching score, and the subtraction of batting score and pitching score, and establish the Logistic regression model for 4 times (20% for the season, 40% for the season, 70% for the season, 90% for the season), then predict and verify the last 10 games by 6 methods of prediction.

Use Wald test to choose the explanation variable, and compare the 70% and 90% for the season with 20% and 40% for the bottom season, and also control the 3 explanation variable for the previous game to all season and the previous season.

The research result is that the highest average predicted success rate is 70%,72.5% and 67.5% for 3 explanation variable to predict the 6 methods of prediction, in which 70% and 90% for all season is better than 20% and 40% for bottom season, but the preceding processing mode forecast success ratio not necessarily is also better than these three explanation variable processing mode; Forecast the way is previous time fights by two rows when the data does for the explanation variable, the forecast success ratio in three explanation variable processing mode all is best; selects into the variable aspect, contained next half Ji Jiqian altogether to select into 22 kind of variables, in which highest three has RBI,ERA and SO per bat.

Key words : Logistic regression, professional baseball, predict.

# 目錄

誌謝.....	I
摘要.....	II
Abstact .....	III
目錄.....	V
圖目錄.....	VIII
表目錄.....	IX
<b>第一章 緒論 .....</b>	<b>1</b>
<b>第一節 研究背景與動機 .....</b>	<b>1</b>
<b>第二節 研究目的 .....</b>	<b>4</b>
<b>第三節 名詞解釋 .....</b>	<b>4</b>
<b>第四節 研究流程 .....</b>	<b>9</b>
<b>第二章 文獻探討 .....</b>	<b>10</b>
<b>第一節 文獻探討 .....</b>	<b>10</b>
<b>第二節 邏輯斯迴歸 .....</b>	<b>14</b>
<b>第三章 研究方法 .....</b>	<b>16</b>
<b>第一節 研究對象 .....</b>	<b>16</b>
<b>第二節 資料來源 .....</b>	<b>16</b>
<b>第三節 資料整理 .....</b>	<b>16</b>
<b>第四節 統計分析 .....</b>	<b>19</b>
<b>第四章 研究內容 .....</b>	<b>21</b>
<b>第一節 敘述分析 .....</b>	<b>21</b>
<b>第二節 以主客隊各項成績建立模型與預測 .....</b>	<b>26</b>
<b>一、比賽進行 20%(賽季前 48 場).....</b>	<b>26</b>
<b>二、比賽進行 40%(賽季前 96 場).....</b>	<b>30</b>

三、比賽進行 70%(賽季前 168 場).....	34
四、比賽進行 90%(賽季前 216 場).....	38
五、比賽進行下半季 20%(賽季第 121-168 場).....	43
六、比賽進行下半季 40%(賽季第 121-216 場).....	44
七、前一場模型建立.....	45
八、驗證分析與比較.....	58
<b>第三節 以主客隊各項成績比建立模型與預測 .....</b>	<b>61</b>
一、比賽進行 20%(賽季前 48 場).....	61
二、比賽進行 40%(賽季前 96 場).....	63
三、比賽進行 70%(賽季前 168 場).....	65
四、比賽進行 90%(賽季前 216 場).....	67
五、比賽進行下半季 20%(賽季第 121-168 場).....	70
六、比賽進行下半季 40%(賽季第 121-216 場).....	70
七、前一場模型建立.....	71
八、驗證分析與比較.....	79
<b>第四節 以主客隊各項成績差建立模型與預測 .....</b>	<b>82</b>
一、比賽進行 20%(賽季前 48 場).....	82
二、比賽進行 40%(賽季前 96 場).....	84
三、比賽進行 70%(賽季前 168 場).....	86
四、比賽進行 90%(賽季前 216 場).....	89
五、比賽進行下半季 20%(賽季第 121-168 場).....	91
六、比賽進行下半季 40%(賽季第 121-216 場).....	92
七、前一場模型建立.....	93
八、驗證分析與比較.....	100
<b>第五節 選入的解釋變數統計 .....</b>	<b>103</b>
<b>第五章 結論與建議 .....</b>	<b>107</b>

<b>第一節 結論</b> .....	107
一、模型建立與預測間的差異.....	107
二、預測場次的差別.....	107
<b>第二節 討論與建議</b> .....	108
一、選入變數之討論.....	108
二、解釋變數處理與場次的建議.....	109
三、其餘建議.....	110
<b>參考文獻</b> .....	112
一、中文部分 .....	112
二、外文部分 .....	113
三、網站部分 .....	113

## 圖目錄

圖 1.4.1 本研究之研究架構圖.....	9
------------------------	---

## 表目錄

表 1.3.1 比賽與球隊資訊.....	5
表 1.3.2 野手資訊.....	5
表 1.3.3 投手資訊.....	8
表 2.1.1 解釋變數處理方式與預測準確率統計.....	14
表 3.3.1 新增解釋變數資訊.....	17
表 3.3.2 解釋變數說明.....	18
表 3.3.3 預測說明.....	19
表 3.3.4 前一場變數處理方式說明.....	19
表 4.1.1 各項成績的敘述分析.....	21
表 4.1.2 各項成績比的敘述分析.....	23
表 4.1.3 各項成績差的敘述分析.....	25
表 4.2.1 比賽進行 20%時各項成績的單變數邏輯斯迴歸.....	27
表 4.2.2 比賽結果之參數估計與配適度.....	30
表 4.2.3 比賽進行 40%時各項成績的單變數邏輯斯迴歸.....	31
表 4.2.4 比賽結果之參數估計與配適度.....	34
表 4.2.5 比賽進行 70%時各項成績的單變數邏輯斯迴歸.....	35
表 4.2.6 比賽結果之參數估計與配適度.....	38
表 4.2.7 比賽進行 90%時各項成績的單變數邏輯斯迴歸.....	39
表 4.2.8 比賽結果之參數估計與配適度.....	42
表 4.2.9 比賽結果之參數估計與配適度.....	43
表 4.2.10 比賽結果之參數估計與配適度.....	45
表 4.2.11 比賽進行 20%時前一場各項成績的單變數邏輯斯迴歸.....	45
表 4.2.12 使用前一場建立模型-前 48 場.....	48
表 4.2.13 比賽進行 40%時前一場各項成績的單變數邏輯斯迴歸.....	49

表 4.2.14	使用前一場建立模型-前 96 場.....	51
表 4.2.15	比賽進行 70%時前一場各項成績的單變數邏輯斯迴歸 .....	52
表 4.2.16	使用前一場建立模型-前 168 場.....	55
表 4.2.17	比賽進行 90%時前一場各項成績的單變數邏輯斯迴歸 .....	55
表 4.2.18	使用前一場建立模型-前 216 場.....	58
表 4.2.19	主客隊各項成績的分析整合表.....	58
表 4.2.20	主客隊各項成績—前一場的分析整合表.....	60
表 4.3.1	比賽進行 20%時各項成績比的單變數邏輯斯迴歸 .....	61
表 4.3.2	比賽結果之參數估計與配適度.....	63
表 4.3.3	比賽進行 40%時各項成績比的單變數邏輯斯迴歸 .....	63
表 4.3.4	比賽結果之參數估計與配適度.....	65
表 4.3.5	比賽進行 70%時各項成績比的單變數邏輯斯迴歸 .....	65
表 4.3.6	比賽結果之參數估計與配適度.....	67
表 4.3.7	比賽進行 90%時各項成績比的單變數邏輯斯迴歸 .....	68
表 4.3.8	比賽結果之參數估計與配適度.....	70
表 4.3.9	比賽結果之參數估計與配適度.....	70
表 4.3.10	比賽結果之參數估計與配適度.....	71
表 4.3.11	比賽進行 20%時各項成績比前一場的單變數邏輯斯迴歸 .....	71
表 4.3.12	使用前一場建立模型-前 48 場.....	73
表 4.3.13	比賽進行 40%時各項成績比前一場的單變數邏輯斯迴歸 .....	73
表 4.3.14	使用前一場建立模型-前 96 場.....	75
表 4.3.15	比賽進行 70%時各項成績比前一場的單變數邏輯斯迴歸 .....	75
表 4.3.16	使用前一場建立模型-前 168 場.....	77
表 4.3.17	比賽進行 90%時各項成績比前一場的單變數邏輯斯迴歸 .....	77
表 4.3.18	使用前一場建立模型-前 216 場.....	79
表 4.3.19	主客隊各項成績比的分析整合表.....	79

表 4.3.20 主客隊各項成績比-前一場的分析整合表.....	81
表 4.4.1 比賽進行 20%時各項成績差的單變數邏輯斯迴歸.....	82
表 4.4.2 比賽結果之參數估計與配適度.....	83
表 4.4.3 比賽進行 40%時各項成績差的單變數邏輯斯迴歸.....	84
表 4.4.4 比賽結果之參數估計與配適度.....	86
表 4.4.5 比賽進行 70%時各項成績差的單變數邏輯斯迴歸.....	86
表 4.4.6 比賽結果之參數估計與配適度.....	88
表 4.4.7 比賽進行 90%時各項成績差的單變數邏輯斯迴歸.....	89
表 4.4.8 比賽結果之參數估計與配適度.....	91
表 4.4.9 比賽結果之參數估計與配適度.....	92
表 4.4.10 比賽結果之參數估計與配適度.....	92
表 4.4.11 比賽進行 20%時各項成績差前一場的單變數邏輯斯迴歸.....	93
表 4.4.12 使用前一場建立模型-前 48 場.....	94
表 4.4.13 比賽進行 40%時各項成績差前一場的單變數邏輯斯迴歸.....	95
表 4.4.14 使用前一場建立模型-前 96 場.....	96
表 4.4.15 比賽進行 70%時各項成績差前一場的單變數邏輯斯迴歸.....	96
表 4.4.16 使用前一場建立模型-前 168 場.....	98
表 4.4.17 比賽進行 90%時各項成績差前一場的單變數邏輯斯迴歸.....	98
表 4.4.18 使用前一場建立模型-前 216 場.....	100
表 4.4.19 主客隊各項成績差的分析整合表.....	100
表 4.4.20 主客隊各項成績差-前一場的分析整合表.....	102
表 4.5.1 主客隊各項成績選入的解釋變數統計.....	103
表 4.5.2 主客隊各項成績選入的解釋變數統計次數.....	104
表 4.5.3 主客隊各項成績比選入的解釋變數統計.....	104
表 4.5.4 主客隊各項成績比選入的解釋變數統計次數.....	105
表 4.5.5 主客隊各項成績差選入的解釋變數統計.....	105

表 4.5.6 主客隊各項成績差選入的解釋變數統計次數.....	106
表 4.5.7 選入解釋變數的統計總次數.....	106

# 第一章 緒論

## 第一節 研究背景與動機

棒球是一種團體球類運動，分進攻守兩方比賽，防守方必須於一個扇形球場上分配九人進行防守，由投手將球往本壘板方向進行投球，使比賽開始進行，如果打擊者將球擊出後，進行接球守備，傳向壘包造成出局數，達到三出局後攻守交換。而進攻方必須手持球棒，站在本壘板旁進行打擊，將球擊出後必須分別踩過一壘、二壘、三壘與本壘，自此可以得到一分。比賽會進行九局（少棒為 6 局、青少棒 7 局），每局分為上下半局，由兩隊分別進攻，九局打完之後得分最高者獲勝。若正規九局打完後雙方得分仍相同，則進入 12 局制延長賽(中華職棒，若 12 局打完仍平手則判定為和局)或依其他規定認定為和局(美國職業棒球大聯盟採無限局數延長賽，無和局制度)(竹板凳，2016)。

棒球的起源有幾個說法，一般認為棒球是源自於板球運動，板球除了是足球與棒球的始祖外，更是日後許多團體運動的典範。板球始源於中世紀英國小孩玩的一種用球棒或球拍進行擊球的遊戲，後來發展至英國貴族的紳士運動。棒球傳言是在十九世紀美國陸軍軍官 A. Doubleday 發明的，但我們所認識的棒球其實源自紐約市燈籠褲棒球隊(Knickerbocker Baseball Club)，1842 年該隊在卡萊特(Alexander Joy Cartwrigly)的主導下創立，首度編訂棒球規則，並且在 1846 年舉辦了第一場比賽，由 A. Doubleday 擔任裁判。南北戰爭後將棒球推廣至美國南方，使得更多人得知這項運動，讓更多的球隊成立。1869 年第一支職業棒球隊成立，隊名是辛辛那提紅長襪（即是現在的辛辛那提紅人隊），成為棒球職業化的先驅(張美惠譯，2008)。

台灣棒球的起源可以追溯到日據時期，中日甲午戰爭清朝戰敗，割讓台灣給日本，因此 1895 年日本開始了對台灣的統治，而同一年也是日本將美國傳入的 Baseball 統一譯為「野球」，稱為「野球元年」。最初日本人無意推動現代體育於台灣民眾，台灣人接受度也不高，但體育活動卻在最自然情形下進入台灣，棒球

就是其中之一。最早的棒球活動是由日本人在台灣的銀行職員開始，起初只是用球進行傳接的運動，直到八、九年後經由日本人的引進才有正式的棒球隊與棒球比賽的產生。第二次世界大戰結束之後，台灣脫離日本，許多城市殘破不堪，但由於光復台灣，在「張燈結綵喜洋洋，勝利歌兒大家唱」的歡愉氣氛中，各行各業競相組隊，棒球賽成為點燃台灣社會熱情的發動機。台灣歷史最悠久的台電與合庫棒球隊，起源就是員工將棒球作為休閒活動，沒有太多外加的意義，他們並非專業選手，球技也不精湛，僅是當作一種純粹的運動，所以不好的球具以及有缺陷的球場等等各種缺陷都可以被包容(中華民國棒球協會，2006)。

「以竹為棒，以石為球」這句出自於 70 年代的紅葉少棒隊。當時台東紅葉國小克服許多困難，北上拿下第二十屆學童盃冠軍。1968 年八月與來自日本的關西聯隊進行交手(當初被認為是拿下威廉波特少棒賽冠軍的球隊，實際上拿到冠軍的是和歌山隊)，這場比賽被電視台轉播出去，擊敗了當時被認為是世界冠軍的日本隊，國人再次被點燃了對棒球熱情。然而第一個拿下世界冠軍的是金龍少棒隊。當時選出各地菁英組成的金龍隊，赴美參加威廉波特少棒賽，成功一路過關斬將晉級冠軍戰。冠軍賽開打之際，不管男女老幼都通宵達旦地守在電視機或收音機前，關注這歷史性的一戰。最後在球員努力下成功拿下了比賽勝利，全台為之沸騰，也開啟台灣三級棒球輝煌的戰績(中華民國棒球協會，2006)。

棒球在台灣帶起一波熱潮，國際賽也是中華隊所著重的。1980 年代的亞錦賽、世錦賽、洲際盃等等，甚至是被列為示範賽的 1984 年洛杉磯奧運、1988 年漢城奧運，中華隊將士用命，就為了取得好成績。棒球儼然成為了台灣的國球，這時也促使了台灣棒球職業化的可能(中華民國棒球協會，2006)。

在兄弟飯店負責人洪騰勝四處奔波努力下，「中華職棒聯盟」總算在 1989 年 10 月 23 日成立，台灣繼日本、韓國之後，正式成為亞洲第三個成立職業棒球運動的國家。元老創始球隊有兄弟象隊、味全龍隊、統一獅隊及三商虎隊，中華職棒也在 1990 年 3 月 17 日開始了第一個球季。起初各界人士對於職棒的誕生有三個疑問：一是人員問題，台灣棒球球員人手嚴重不足。二是場地問題，當時還有

許多業餘比賽要打，並沒有那麼多的球場可以消化這些比賽。三則是經費問題，職業隊伍每年支出上千萬，是否能平衡支出是最大的問題。但在各界關注以及民眾熱情支持下，這些問題暫時被消除了。在中華職棒開打之後也成功增加了兩支球隊，分別為時報鷹及俊國熊。原本在之後也要加入中華職棒的聲寶公司，則是因為中華職棒的某些協議而被拒絕於門外，同一時間因為電視轉播權問題的年代公司，也被終止與聯盟的合作。這兩家公司因此憤而決定另組聯盟，成立了那魯灣公司經營新的聯盟與旗下四支球隊。1996 年台灣第二個職業棒球聯盟「台灣職業棒球大聯盟」(簡稱為台灣大聯盟)成立，宣布旗下四支球隊分別為台北太陽隊、嘉南勇士隊、高屏雷公隊和台中金剛隊。1997 年 2 月 28 日台灣大聯盟第一個球季正式開始。中華職棒聯盟與台灣大聯盟經過幾年的惡鬥，以及各自產生的問題，促使兩聯盟開始決議合併一事。2003 年 1 月 13 日在陳水扁總統見證之下，兩聯盟順利整合，並更名為「中華職業棒球大聯盟」(英文全稱為 Chinese Professional Baseball League，以下簡稱中職或 CPBL)，球隊包含原中華職棒聯盟四隊，兄弟象、統一獅、興農牛、中信鯨，以及台灣大聯盟合併後的兩隊，誠泰太陽隊、第一金剛隊。同年 3 月 1 日為新聯盟開幕戰，中華職棒進入新的時代(中華民國棒球協會，2006)。

經過二十幾年的風風雨雨，有過球團的擴編與減少，甚至爆發過假球案，但在球團努力的運作以及球迷的支持之下，也努力維持了 28 個賽季。近年來亞洲各個國家也增加了許多彼此的交流，像是曾經的亞洲職棒大賽、季前跨聯盟的交流賽或熱身賽等等。自從棒球確認回歸奧運會後，再次掀起了一股棒球風潮，每年除了 3-10 月的職業棒球賽季之外，還有許多國際賽事或各國職棒交流賽會在季前或季後進行，尤其在 2017 年有許多的重大國際賽事，2017 世界棒球經典賽、台北世界大學運動會、U-18 世界盃棒球賽、亞洲棒球錦標賽，由台日韓三國年輕職棒選手為主的亞洲職棒冠軍爭霸賽以及由中華職棒主辦的亞洲冬季棒球聯盟。所以一整年都有滿滿的棒球賽事可以欣賞，而球迷們除了到場邊為喜愛的球隊、球員加油之外，在球場外進行些許博弈，也為比賽增加了一點觀賞性。

台灣運動彩券是一種與運動相關的彩券，自從 2008 年 5 月發行至今也已有 10 年的歷史，提供國人，一個合法、健康的投注管道。運動彩券的發行，除了負有打擊非法賭博的使命，更有發掘、培訓及照顧運動人才的長遠目標，運動彩券的銷售收入都有固定的比例提撥至國家運動發展基金，是政府發掘、培訓及照顧運動人才不可或缺的資金來源（台灣運彩，2008）。

我們球迷觀賞比賽之餘，若事先對比賽做功課，不但可以為自己喜愛的球隊加油，也可以為自己的荷包增加點厚度。我們若是想要透過運動彩券獲取些許利益，事前功課當然不能少，必須透過賽前的資訊進行觀察，再來對比賽進行預測。許璉方(2010)指出比賽紀錄對於預測下一場比賽結果是有助益的，李俊德(2012)利用 logistic 迴歸進行預測，可以達到 70%的準確率。雖然進行預測不可能達到百分之百的準確，但我們有效地透過賽前資料分析來精確判斷比賽的結果，將可以大大的提升投注報酬。我們在賽前分析比賽情勢，順便透過預測進行下注，除了欣賞比賽並且替喜愛的球隊加油外，我們投注運動彩券，可以透過政府培訓更多運動選手，進而幫助改善整個運動圈。另外透過運彩獲得些彩金，因此可能更有意願再次買票進到球場觀賞比賽，無疑也是替我們所喜愛的中職這個環境提供小小幫助。

## 第二節 研究目的

基於上述的研究背景與動機，本研究目的如下：

- 一、探討球員表現對中華職棒賽是勝負的影響。
- 二、建立影響中華職棒賽是勝負的統計模型。
- 三、探討與比較不同統計模型的預測能力

## 第三節 名詞解釋

表 1.3.1 中場次為賽季開始前已排定比賽順序，但實際比賽會因場地或下雨等天候問題而延賽，所以新增實際比賽場次來解釋真正比賽舉行之順序，若進行

延賽導致一日兩戰之情況，則以場次編號較小者為優先。本研究探討的為主場球隊之勝負，所以將主場球隊的比賽結果是為反應變數。

表 1.3.1 比賽與球隊資訊

數據種類	說明
場次	中華職棒大聯盟編排賽程的比賽編號。
實際比賽場次	實際比賽進行的順序。
球隊	中信兄弟、統一 7-ELEVEn 獅、LAMIGO 桃猿與富邦悍將。
球隊(客 VS 主)	兩支比賽球隊先攻者為客隊，後攻者為主隊。
主隊勝負	主場球隊以 1 表示獲勝，以 0 表示敗場。

表 1.3.2 為中職官網所記載野手方面的資訊有打數、得分、安打、打點、二壘安打、三壘安打、全壘打、雙殺打、被四壞球、被四死球、被三振、犧牲短打、犧牲飛球、盜壘成功、盜壘刺殺、失誤，本研究討論是團隊數據，故新增打擊率、壘打數、上壘率、長打率與攻擊指數。自中職官網上可得知打擊率為安打數/打數，而本研究使用的打擊率為團隊總安打數/總打數。

表 1.3.2 野手資訊

數據種類	說明
打數	打者上場打擊的次數扣掉被四死球、犧牲短打、犧牲飛球、妨礙打擊的次數。
得分	成功上壘後，回到本壘得分的次數。
安打	打者打擊出去後，球落在場內，不靠防守方失誤或野手選擇，安全上壘的次數。
打點	打者打擊出去後，跑者(包含打者)回到本壘所得到的分數。

表 1.3.2 野手資訊(續)

數據種類	說明
二壘安打	擊出安打後不透過野手失誤或野手選擇成功踏上二壘的次數。
三壘安打	擊出安打後不透過野手失誤或野手選擇成功踏上三壘的次數。
全壘打	將球擊出後，球落在界內區的全壘打牆外的次數。
雙殺打	一次打數造成兩人出局的次數。
被四壞球	對方投手投出四個壞球保送上壘的次數。
被四死球	包含四壞球及觸身球保送上壘的次數。
被三振	對方投手投出三個好球數導致出局的出局數。
犧牲短打	利用觸擊的方式使自己出局，達到跑者前進一個壘包的次數。
犧牲飛球	擊出高飛球使自己出局，讓跑者成功回到本壘得分的次數。
盜壘成功	利用投手投球的動作先行起跑至下一個壘包，並成功上壘的次數。
盜壘刺殺	利用投手投球的動作先行起跑至下一個壘包，但被刺殺的次數。
失誤	防守時守備失誤的次數。
打擊率	打擊率為打者上場打擊時，擊出安打的機率。即 $\text{打擊率} = \frac{\text{安打}}{\text{打數}}$
壘打數	打者依靠自己擊出安打後，所佔有的壘包數。

表 1.3.2 野手資訊(續)

數據種類	說明
上壘率	<p>打者上場打擊時，能夠站上壘包的機率。即</p> $\text{上壘率} = \frac{\text{安打} + \text{四死球}}{\text{打數} + \text{四死球} + \text{犧牲飛球}}$
長打率	<p>打者上場打擊時，平均能夠佔有的壘包數。即</p> $\text{長打率} = \frac{\text{壘打數}}{\text{打數}}$
攻擊指數	<p>攻擊指數為 1.000，則表示該名球員攻佔壘包的數目剛好等於他製造的出局數。數據越高代表球員越不容易出局，攻擊威脅性越大。即</p> $\text{攻擊指數} = \frac{\text{壘打數} + \text{盜壘} + \text{四死球}}{\text{打數} - \text{安打} + \text{盜壘刺殺} + \text{雙殺打}}$

表 1.3.3 中投手方面依照上場順序，第一位上場的為先發投手，其餘為後援投手，而本研究所討論的為團隊數據，故採用為加總或平均後的團隊成績，所以投手方面的資訊包含局數、面對打擊數、投球數、好球數、被安打、被全壘打、四壞球、四死球、奪三振、暴投、投手犯規、失分、自責分、防禦率、每局被上壘率。官網上所記載局數中，呈現 5.1 局或 7.2 局，分別代表的意思為投手投完 5 局又 1 個人次及投完 7 局又 2 個人次，為了方便計算後續數據，便先將其改為 5+1/3 及 7+2/3，再轉化為小數即 5.33 局或 7.67 局(四捨五入至第二位)。防禦率在中職官網記載為投手自責分×9/投球局數，本研究所使用防禦率則為團隊自責分×9/總投球局數。每局被上壘率(WHIP 值)在中職官網記載為(敵對安打數+四壞球)/投球局數，故此使用每局被上壘率為(敵對總安打數+總四壞球)/總投球局數。

表 1.3.3 投手資訊

數據種類	說明
局數	投手上場所投局數，造成三人次出局為投滿一局。
面對打擊數	投手面對打者的次數。
投球數	投手投出球的次數。
好球數	投手投出好球的次數。
被安打	投手被對方打者擊出安打的次數。
被全壘打	投手被對方打者擊出全壘打的次數。
四壞球	投手投出四壞球的保送次數。
四死球	投手投出四壞球及觸身球的保送次數。
奪三振	投手投出三個好球數，造成打者出局的出局數。
暴投	投手投出的球，令捕手無法正常守備的球數。
投手犯規	壘上有人時，投手有投手犯規違例之行為，使壘上跑者推進一個壘包的次數。
失分	投手在場上投球時所失去的所有分數。
自責分	投手所應負責的失分。
防禦率	又稱投手自責分率，為投手每投完九局平均會掉多少分數。即 $\text{防禦率} = \frac{\text{自責分} \times 9}{\text{投球局數}}$
每局被上壘率	扣除失誤（包含扣除觸身球、故意四壞、打者因妨礙打擊或妨礙跑壘而上壘的次數）外，投手平均每局讓打者上壘數。即 $\text{每局被上壘率} = \frac{\text{安打} + \text{四壞球}}{\text{投球局數}}$

## 第四節 研究流程

本研究流程依據研究背景與動機延伸出研究目的，再依目的進行文獻探討，確認研究對象與建立研究方法，針對研究對象建立資料以及進行資料處理，最後透過統計分析以得到研究結果。

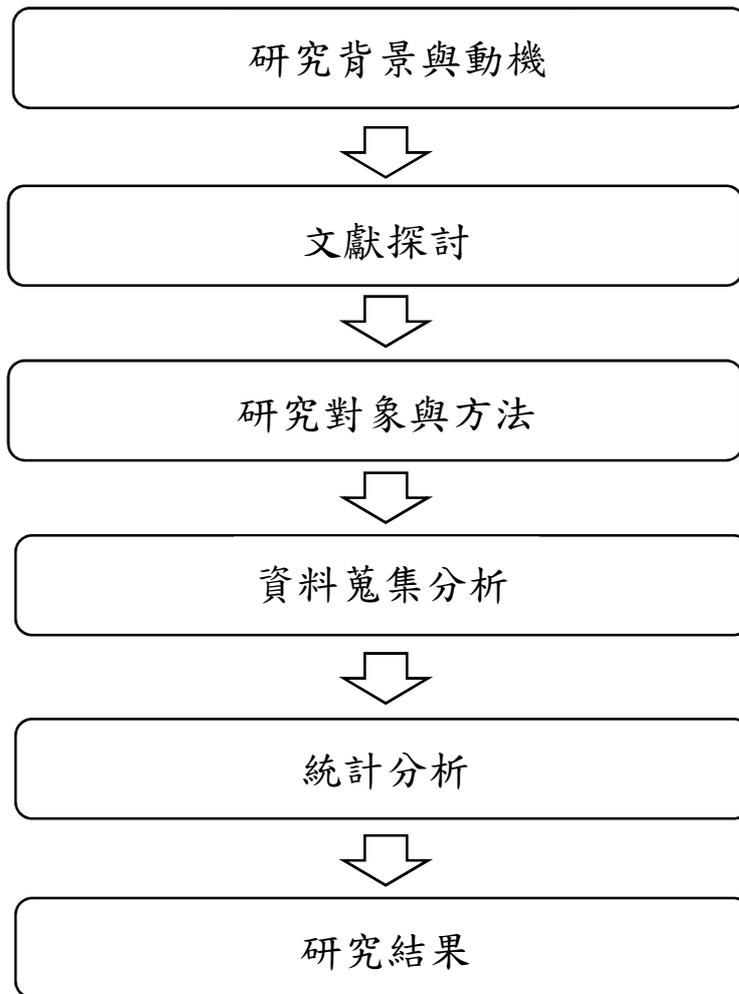


圖 1.4.1 本研究之研究架構圖

## 第二章 文獻探討

### 第一節 文獻探討

棒球簡單來說就是投、打、跑、傳、接。詳細的解釋棒球是由兩隊各 9 人，雙方進行攻守，在九局比賽結束後獲得最多分的那隊獲勝，所以棒球比賽在最後一局 3 出局以前，比賽的勝負都還很難說。球場上進行守備的人分別為投手與捕手各一名，內野手四名及外野手三名。

吳仁(1996)提到守備優於攻擊的隊伍獲得冠軍的機率比較大。棒球的守備，最講究團隊合作，投手固然是中心人物，但仍有賴於其他隊友的協助配合，若彼此不能互相支援，個人技巧再好，球隊也不可能有好成績。瘦菊子(2003)說到投手不論善投多少種變化球，球速有多快，沒有夠水準的控球力，鐵定無法成為好投手。林文斌與鄧元湘(2000)表示從防禦率可以看出一個投手的好壞，因為防禦率越低代表的就是越不容易失分，如果防禦率較低的投手沒有得到相對較多的勝場數，代表隊友打擊火力的供輸及守備方面的幫忙就不是那麼理想。

吳仁(1996)講到全壘打是球迷最熱烈擁護的，但是對球隊貢獻最大的，應該是適時安打，所以打點多及得分多的球員應該得到更多的掌聲。而他也提到棒球比賽可以說是上壘的累積，最後目的便是本壘。打者與跑者一有機會，都會想多進一壘，這是應有的態度。

職業棒球所求的即是勝利，比賽一場一場贏下來，累積夠多的勝場數後就能進入到季後賽，季後賽過關斬將之後就能獲得冠軍金盃，獲得冠軍的球隊同時也能得到較多球迷的支持，球迷最想看的就是支持的球隊贏下一場又一場的比赛。職棒球隊為了獲得勝利，將最好的比賽呈現給球迷欣賞，進行各方面的補強與訓練是必要的，而會影響比賽勝負的則有許多不同因素。

一場棒球比賽的開始從投手將球投出的那一刻開始，根據現今棒球規則，投手板距離本壘板 60 呎 6 吋(18.44 公尺)，假設投手以時速 144 公里/小時投出，換算成秒速為 40 公尺/秒，投手投球必須向前跨一步，約是 1.4 公尺，所以實際上

球出手到本壘板的距離為 17.04 公尺，時間等於距離/速度，所以  $17.04/40$  大約 0.426 秒左右就到達本壘板，而打者就必須在這 0.426 秒內做出反應，比賽往往就在這眨眼間便決定了勝負。投手為了獲得出局數，要不投出打者打不到的球，要不將球投到刁鑽的位置使打者打不好，所以投手除了加快球速，增加變化球，最重要就是好的控球力，畢竟球速就算 170 公里，可以投出各種變化球，投不進好球帶都是不行的。打者要思考的也有許多，除了投手投過來的球要怎麼打到外，也要思考壘包上有無跑者的各種情況，假設一壘有人，是否要採取犧牲短打或者打滾地球讓跑者推進。若是三壘有跑者，是否要打成高飛球，讓他回來得分……等，想盡辦法得到多分才能使獲勝的機率提高。球場上的其他野手也要隨時注意球的動向，在打者站上打擊區時，先思考著這位打者是怎樣的風格，若是強打者，是否要將位置站得遠一些，球飛過來時，判斷球的走向，趕到正確的位置進行接傳球，幫助投手將失分壓到最低。除了投手要有好的表現外，其他隊友的打擊和守備也要有表現，畢竟棒球是一個 9 人的團體運動。

所以對比賽進行預測的時候，不論是投手的防禦率、三振數、好壞球數...等等，或是打者的打擊率、安打數、失誤數...等等，只要有可能影響勝負的各項數據，都必須納入預測的解釋變數中。葛學文(2014)以 2010 年至 2013 年美國職棒大聯盟 30 支球隊，團隊年度打擊、投球及守備紀錄，運用資料探勘技術探討研究球隊績效，得知影響球隊進入季後賽最關鍵的因素為防禦率、攻擊指數和對戰同分區球隊之勝率。林家鳳(2016)以 2010 年至 2015 年中華職棒球員以打擊紀錄與球員身高、體重、年齡等身體質量變數為解釋變數，歸納影響球員次年是否留在一軍的因素，其中提到攻擊指數對球隊獲勝是很重要的因素，而二壘安打率、長打率與勝率皆有正向顯著相關。Bock(2015)提到教練可以利用歷史數據判斷對手的下一步行動，也就是可以透過預測的方式來增加勝率。

李俊德(2012)以中華職棒二十一年 4 支球隊的各項數據，在主場球隊選入的解釋變數中，得分、勝率、盜壘成功與犧牲短打次數增加會提高主場球隊獲勝的機率，被三振次數、投手面對被打席數、被全壘打數、先發投手壞球數、先發投

手投球數、投手每局被打擊率(WHIP 值)與先發投手失分增加會降低主場球隊比賽獲勝的機率。而客場球隊選入的解釋變數中，失誤、投手防禦率、投手失分(含先發投手)與先發投手安打數增加會提升主場球隊比賽獲勝的機率，得分、打數、長打率、二壘安打數、打點、投手投球局數、投手奪三振數(含先發投手)增加會降低主場球隊比賽獲勝的機率，預測準確率大約為 70%左右。

施致平、黃蕙娟、倪瑛蓮(2010)以中職 18 年對戰組合攻守紀錄為研究主軸，運用文件分析與德爾菲法，建立影響中華職棒比賽勝負之解釋變數，再用邏輯斯迴歸分析，建構中職比賽勝負之預測。其得到的結果為：一、影響中職比賽勝負之關鍵預測要素，攻擊成績指標包含打數、打點、得分、安打、二壘安打、三壘安打、雙殺打、犧牲短打、犧牲飛球、四壞球、四死球、盜壘成功、被三振、出賽數、打席、全壘打、一壘以上安打數、壘打數、盜壘成功率、打擊率、上壘率、長打率共 22 項、投手成績指標包含出賽數、完投、完封、無四死球、勝場、勝率、敗場、救援成功、中繼點、防禦率、面對打席、被安打、被全壘打、四壞球、救援成功、四死球、奪三振、暴投、被盜壘數、抓盜壘數、失分、自責分、WHIP、被長打率、被上壘率、被打擊率、平均每局投球數、三振四壞比、每九局四壞數、每九局三振數、每九局被安打數共 32 項，守備成績指標包含出局數、助殺、守備失誤、雙殺、被盜壘成功、盜壘阻殺、守備率共 7 項。二、影響勝負攻擊方面的解釋變數中，解釋力高低的解釋變數分別為得分、打數、打點、失誤、犧牲短打。再由勝算比分析發現，打數、失誤為負面解釋變數，其中又以失誤之勝算比最小；而得分、打點及犧牲短打為正向解釋變數。三、影響勝負投球方面的解釋變數中，解釋力高低之解釋變數分別為投球局數、面對打席、被安打、四壞球及自責分。再由勝算比分析發現，面對打席及自責分為負面解釋變數；而投球局數、被安打與四壞球為正向解釋變數。

馮瑞祥(2013)利用資料探勘的技術預測美國大聯盟職棒賽事之勝負與得分。首先他使用皮爾森積差相關分析去除掉較無相關之變數，挑選適當之投入變數，再利用類神經網路中的倒傳遞網路將挑選出之變數投入並建立模型。所提出的模

型可預測比賽勝負，但準確率約為五成上下，尚待加強。陳麒文、李天行、梁玉秋(2014)利用資料探勘中的人工類神經網路以及邏輯斯迴歸來建構美國職棒大聯盟洋基隊與紅襪隊的勝隊預測模式，並且以兩隊在2006年至2012年共127場例行賽，所有例行賽中的各項變數，包含主客場球隊、主客場先發投手之左右投狀況、主客場先發投手上一場之變數、主客場打擊群上一場之變數、賽事時間、比賽時之華氏溫度等共60個解釋變數。結果以人工類神經網路所建構出來的預測模式，其勝隊預測準確率高達72.22%，而邏輯斯迴歸的部分得到的預測準確率為61.11%。

鄧文揚(2013)使用逐步迴歸、邏輯斯迴歸、類神經網路和決策樹對中華職棒兩支球隊(La new 熊 VS 兄弟象隊)進行勝負預測模型建立，先找出影響比賽勝負之解釋變數，再利用篩選出的解釋變數建立2009年兩支球隊對戰勝負預測模型，最後再去預測兩支球隊2010年季賽的比賽勝負，其中選入的解釋變數先進行前置處理，而使用的解釋變數包含打擊方面(打數比、打點比、得分比、安打比、二壘打比、三壘打比、雙殺打比、犧短比、犧飛比、四壞比、四死比、盜壘成功比、盜壘刺比、失誤比、被三振比、打擊率比)以及投球方面(投球局數比、面對打席比、投球數比、好球數比、壞球數比、被安打比、被全壘打比、投手四壞比、投手四死比)，篩選來的重要解釋變數則有打點比、投球局數比、被安打比、失誤比、盜壘刺比、暴投比、投手三振比。

方敬硯(2017)利用資料探勘建立美國職業大聯盟小熊隊的預測模型，使用的方式是將原始資料分別投入決策樹、邏輯斯迴歸、倒傳遞類神經網路進行訓練測試，將其得到的重要變數分別投入倒傳遞類神經網路以及邏輯斯迴歸分別進行訓練及測試。此研究將60個變數透過共線性評估後，刪除了10個具有共線性的變數，最終投入的變數共50個。預測結果邏輯斯迴歸的部分，其準確率為88.89%。

表2.1.1是各文獻中所使用的解釋變數處理方式、預測方法以及預測準確率的統計，解釋變數處理方式有兩種，其中邏輯斯迴歸預測的預測準確率大約都在60%以上。

表 2.1.1 解釋變數處理方式與預測準確率統計

論文作者	解釋變數處理方式	預測準確率(預測方式)
李俊德(2012)	主客隊各項成績	70%(邏輯斯迴歸)
馮瑞祥(2013)	主客隊各項成績	50%(類神經網路)
陳麒文等人 (2014)	主客隊各項成績	72.22%(類神經網路) 61.11%(邏輯斯迴歸)
方敬碩(2017)	主客隊各項成績	81.48%(決策樹) 88.89%(邏輯斯迴歸) 92.59%(類神經網路)
鄧文揚(2013)	主客隊各項成績比	85%(逐步迴歸) 100%(邏輯斯迴歸) 95.6%(類神經網路) 88.4%(決策樹)

## 第二節 邏輯斯迴歸

王濟川與郭志剛(2003)提到線性迴歸模型(Linear Regression Model)是最流行的統計分析方法，然而在許多情況下，線性迴歸會受到限制。在分析分類變數時，通常會採用的一種統計方法是對數線性模型(log-linear model)。當對數線性模型中的一個二元變數被當作應變數並定義一系列解釋變數的函數時，對數線性模型就變成了 logistic 迴歸模型。

$$P(y_i = 1|x_i) = \frac{1}{1 + e^{-(\alpha + \beta x_i)}}$$

將事件發生的條件機率標記原則定為  $P(y_i=1 | x_i)=p_i$ ，我們就能得到下列 logistic 迴歸模型

$$p_i = \frac{1}{1 + e^{-(\alpha + \beta x_i)}}$$

$$= \frac{e^{\alpha+\beta x_i}}{1 + e^{\alpha+\beta x_i}}$$

其中， $p_i$ 為第  $i$  個案例發生事件的機率，它是一個由解釋變數 $x_i$ 構成的非線性函數(non-linear function)。然而這個非線性函數可以被轉換為線性函數(linear function)。

定義不發生事件的條件機率為

$$1 - p_i = 1 - \left( \frac{e^{\alpha+\beta x_i}}{1 + e^{\alpha+\beta x_i}} \right) \\ = \frac{1}{1 + e^{\alpha+\beta x_i}}$$

事件發生機率與事件不發生機率之比為

$$\frac{p_i}{1 - p_i} = e^{\alpha+\beta x_i}$$

這個比稱之為事件的發生比(the odds of experiencing an event, odds)。Odds 一定為正值，因為  $0 < p_i < 1$ ，而且沒有上界。將 odds 取自然對數就能夠得到一個線性函數：

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \beta x_i$$

當有  $N$  個自變數時可以將 $p_i = e^{\alpha+\beta x_i} / 1 + e^{\alpha+\beta x_i}$ 擴展為：

$$p_i = \frac{e^{\alpha+\sum_{k=1}^N \beta_k x_{ki}}}{1 + e^{\alpha+\sum_{k=1}^N \beta_k x_{ki}}}$$

Logistic 迴歸模型則為下列形式：

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \sum_{k=1}^N \beta_k x_{ki}$$

其中 $p_i = P(y_i = 1 | x_{1i}, x_{2i}, \dots, x_{Ni})$ 為在給定自變數 $x_{1i}, x_{2i}, \dots, x_{Ni}$ 的值時的事件發生機率。

一旦擁有各個案例的觀測自變數 $x_1$ 至 $x_N$ 值構成的樣本，以及擁有其事件發生與否的測量值，我們就能夠使用這些訊息來分析和描述特定下事件的發生以及發生的機率。

# 第三章 研究方法

## 第一節 研究對象

研究對象為中職 4 隊(中信兄弟、統一 7-ELEVEn 獅、LAMIGO 桃猿、富邦悍將)於 2017 年例行賽(中華職棒二十八年)的上半季及下半季的表現，2017 年 3 月 25 日至 2017 年 10 月 11 日(扣除因雨延賽)共 236 場的比賽紀錄，此為公開性的資料。利用 236 場的比賽資料，分別於賽季進行 20%(前 48 場)、40%(前 96 場)、70%(前 168 場)、90%(前 216 場)還有下半季 20%(第 121 場到第 168 場)、下半季 40%(第 121 場到第 216 場)共六個時間點進行分析且建立模型，分別驗證其後 10 場比賽之結果。

## 第二節 資料來源

中華職棒二十八年賽季(2017)球隊對戰紀錄是從中華職棒大聯盟全球資訊網中成績看板取得，使用 Excel 2016 進行建檔統計，所有資料皆是由中職官方進行紀錄的，之後再以 SPSS 21 統計軟體進行資料分析(中華職棒大聯盟全球資訊網，2017)。

## 第三節 資料整理

中華職棒大聯盟全球資訊網上面記載的資料包含打擊成績及投手成績，打擊成績有打數、得分、安打、打點、二壘安打、三壘安打、全壘打、雙殺打、四壞球、死球、三振、犧牲短打、犧牲飛球、盜壘成功、盜壘刺殺、失誤和打擊率，投手成績有投球局數、面對打擊數、投球數、好球數、被安打、被全壘打、四壞球、死球、三振、暴投、投手犯規、失分、自責分、防禦率和每局被上壘率。其中打擊方面的得分、安打、全壘打、四壞球、死球、三振，與投手方面的失分、被安打、被全壘打、四壞球、死球、三振等變數，本研究認為這幾個變數是一體的兩面，舉例來說，進攻方的安打數，應該會等於防守方的被安打數，所以在這

裡只會取用單方面的資料成績。

表 3.3.1 則是參照陳麒文等人（2014）採用了主場先發投手上一場三振率、客場打擊群上一場三振四壞比、客場先發投手上一場三振率、客場打擊群上一場每打點平均打數，其中三振四壞比及每打點平均打數在中職官網中沒有紀錄，故在資料整理的部分新增，並且將每打點平均打數做調整為每打數平均打點，避免分母為零時產生的問題。

表 3.3.1 新增解釋變數資訊

數據種類	說明
每打數平均打點	每個打數可以獲得的打點。
三振四壞比	投手奪三振數與四壞球數的比值。

本研究認為在當場比賽選入得分、打點、失分、自責分建立模型是不恰當的，原因在棒球比賽結束時，得分較高之隊伍獲勝，所以我們若是選入這些變數，我們在進行預測前，已知這些變數，就可以推論出比賽之勝者，這樣進行預測就沒有意義，故不選入這些變數，而在前一場比賽選入則不影響。另外本研究也認為投球局數的部分選入模型也是不恰當，原因在於棒球比賽進行九局，但若主隊在九局上半結束時已領先客隊，此時九局下半是不需要進行的，客隊投手在此時的投球局數會停在八局，只有在九局上半平手或領先時會進入到九局下半，所以若是客隊投球局數比主隊投球局數少時，也可以判斷出比賽之勝負，故不選入此變數。

選擇解釋變數的方式採用 Wald 統計量逐步選取，將會採取三種解釋變數處理，第一種是將所有打擊及投手方面的資料，以主客場做依據，投入到邏輯斯迴歸模型中建立模型。第二種是將打擊與投手的各項成績，以客隊除上主隊，形成一個成績比的資料去建立模型。第三種則是將打擊與投手的各項成績，用客隊減去主隊，產生一個成績差的資料去建立模型。第二種解釋變數處理的部分，Steve(1995)提到在計數統計上遇到資料為零的狀況，若樣本數足夠大，這樣的問

題就會消失，但如果不夠，則會產生麻煩始得無法計算估計值，要讓此問題消失最簡單的方法，就是在單位格增加 0.5，以消除空單元格，產生估計值。其它詳細使用解釋變數及說明如表 3.3.2。

解釋變數總計有打數、每打數平均打點、二壘安打、三壘安打、雙殺打、犧牲短打、犧牲飛球、盜壘成功、盜壘刺殺、失誤、打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、好球數、被安打、被全壘打、四壞球、四死球、奪三振、三振四壞比、暴投、投手犯規、防禦率、每局被上壘率。各項成績中要再分成主客隊，例如客隊防禦率、主隊防禦率，各項成績比則是要將客隊除上主隊形成比值的形式，而各項成績差是要將客隊成績減掉主隊成績變成形成一個差的變數。

表 3.3.2 解釋變數說明

解釋變數處理方式	說明
打擊成績與投手成績(以下簡稱主客隊各項成績)	所有打擊與投手成績以主客隊的方式選入模型。
打擊成績比與投手成績比(以下簡稱主客隊各項成績比)	將打擊與投手成績以客隊除上主隊的方式，產生一個成績比(客隊÷主隊)。
打擊成績差與投手成績差(以下簡稱主客隊各項成績差)	將打擊與投手成績以客隊減去主隊的方式，產生一個成績比(客隊-主隊)。

進行預測時，因為無法得知該場比賽的各項成績，只能透過已完成的賽事進行估計，所以在此將預測驗證的部分則是分為 6 種，分別是前一場數據(不分主客場)、兩隊上次對戰數據、前兩種的平均、前一場數據(分主客場)、第二種與第四種的平均、近五場比賽的平均，詳細說明請見表 3.3.3。

表 3.3.3 預測說明

預測代號	預測方法	說明
預測 1	前一場數據(不分主客場)	即 A 隊上一場比賽成績與 B 隊上一場比賽成績
預測 2	兩隊上次對戰數據	AB 兩隊上次對戰成績，若主客場不同需交換(A 隊這次比賽為主隊，上一次為客隊時需交換)
預測 3	預測 1 跟預測 2 平均	即前兩種的平均
預測 4	前一場數據(分主客場)	即 A 隊上一場比賽成績與 B 隊上一場比賽成績(若 A 隊這場為主隊，即上一場擔任主隊之成績)
預測 5	預測 2 與預測 4 平均	即預測 2 與預測 4 之平均
預測 6	近五場比賽的平均	兩隊最近五場比賽成績的平均值

陳麒文等人(2014)及李俊德(2012)在其研究中使用了前一場的成績進行建立模型及預測，故在此也將其納入預測驗證當中，在之後也將與其他部分進行比較，詳細說明請見表 3.3.4。

表 3.3.4 前一場變數處理方式說明

解釋變數處理方式	說明
各項成績前一場(以下簡稱前一場)	使用前一場的數據對當場的比賽結果進行建立模型。

## 第四節 統計分析

### 一、敘述分析

本研究選入各項成績的敘述分析包含最大值、最小值、平均數及標準差。

## 二、推論統計

1. 單變數邏輯斯迴歸模型，摘要訊息包含參數估計值、標準誤、Wald、自由度、P 值以及勝算比估計值。
2. 多變數邏輯斯迴歸模型透過 Wald 統計量逐步選取解釋變數，摘要訊息包含參數估計值、標準誤、Wald、自由度、P 值以及勝算比估計值。

# 第四章 研究內容

## 第一節 敘述分析

表 4.1.1 是主客場球隊各項成績的敘述分析，表中顯示的項目有最小值、最大值、平均數及標準差。

表 4.1.1 各項成績的敘述分析

變數	最小值	最大值	平均數	標準差
客隊打數	22	53	35.688	4.466
客隊每打數平均打點	0	0.48	0.145	0.088
客隊二壘安打	0	7	2.067	1.516
客隊三壘安打	0	2	0.154	0.405
客隊雙殺打	0	5	0.933	0.939
客隊犧牲短打	0	3	0.350	0.581
客隊犧牲飛球	0	3	0.354	0.596
客隊盜壘成功	0	5	0.708	0.993
客隊盜壘刺殺	0	2	0.425	0.642
客隊失誤	0	5	0.941	1.005
客隊打擊率	0.07	0.51	0.279	0.081
客隊壘打數	2	46	15.958	7.284
客隊上壘率	0.1	0.54	0.347	0.085
客隊長打率	0.07	0.98	0.434	0.155
客隊攻擊指數	0.14	1.89	0.778	0.322
客隊面對打擊數	21	54	38.925	4.880
客隊投球數	89	223	149.138	22.884
客隊好球數	54	135	92.733	12.384
客隊被安打	0	23	9.904	3.663

表 4.1.1 各項成績的敘述分析(續)

解釋變數	最小值	最大值	平均數	標準差
客隊被全壘打	0	4	0.938	0.973
客隊四壞球	0	10	3.442	2.069
客隊四死球	0	3	0.463	0.702
客隊奪三振	1	16	7.142	2.721
客隊三振四壞比	0.25	25	3.552	4.147
客隊暴投	0	3	0.529	0.754
客隊投手犯規	0	1	0.017	0.128
客隊防禦率	0	19.13	4.952	3.499
客隊每局被上壘率	0.33	3.13	1.562	0.553
主隊打數	17	51	34.279	3.998
主隊每打數平均打點	0	0.4	0.143	0.083
主隊二壘安打	0	8	1.871	1.361
主隊三壘安打	0	4	0.179	0.472
主隊雙殺打	0	4	0.838	0.888
主隊犧牲短打	0	4	0.392	0.670
主隊犧牲飛球	0	3	0.338	0.585
主隊盜壘成功	0	6	0.638	1.054
主隊盜壘刺殺	0	3	0.296	0.525
主隊失誤	0	4	0.796	0.826
主隊打擊率	0	0.5	0.938	0.973
主隊壘打數	0	43	3.442	2.069
主隊上壘率	0.11	0.53	0.463	0.702
主隊長打率	0	0.81	7.142	2.721

表 4.1.1 各項成績的敘述分析(續)

解釋變數	最小值	最大值	平均數	標準差
主隊攻擊指數	0	1.96	0.774	0.411
主隊面對打擊數	23	58	40.471	5.579
主隊投球數	104	243	153.204	23.648
主隊好球數	64	146	95.171	13.275
主隊被安打	2	27	10.217	4.010
主隊被全壘打	0	7	1.121	1.171
主隊四壞球	0	11	3.554	2.249
主隊四死球	0	5	0.521	0.808
主隊奪三振	1	20	7.288	2.777
主隊三振四壞比	0.27	27	3.511	3.888
主隊暴投	0	4	0.521	0.760
主隊投手犯規	0	1	0.004	0.065
主隊防禦率	0	20	5.090	3.613
主隊每局被上壘率	0.33	3.33	1.513	0.547

表 4.1.2 是主客場各項成績比的敘述分析，表中顯示的項目有最小值、最大值、平均數及標準差。

表 4.1.2 各項成績比的敘述分析

解釋變數	最小值	最大值	平均數	標準差
打數	0.65	1.50	1.048	0.130
每打數平均打點	0	10.88	1.417	1.610
二壘安打	0	13	1.789	2.197
三壘安打	0	5	1.127	0.871

表 4.1.2 各項成績比的敘述分析(續)

解釋變數	最小值	最大值	平均數	標準差
雙殺打	0	9	1.639	1.740
犧牲短打	0	5	1.250	1.132
犧牲飛球	0	7	1.277	1.146
盜壘成功	0	11	1.718	1.898
盜壘刺殺	0	5	1.479	1.356
失誤	0	11	1.661	1.770
打擊率	0.28	5.67	1.095	0.586
壘打數	0.13	61	2.188	5.601
上壘率	0.33	2.75	1.066	0.437
長打率	0.14	10	1.179	0.864
攻擊指數	0.18	15.12	1.512	1.745
面對打擊數	0.64	1.48	0.973	0.137
投球數	0.51	1.54	0.989	0.178
好球數	0.60	1.51	0.988	0.157
被安打	0	4.60	1.126	0.660
被全壘打	0	9	1.455	1.686
四壞球	0	11	1.556	1.767
四死球	0	7	1.341	1.384
奪三振	0.10	8	1.169	0.782
三振四壞比	0.03	18.2	1.947	2.759
暴投	0	7	1.448	1.494
投手犯規	0	3	1.029	0.265
防禦率	0	19	1.993	2.946

表 4.1.2 各項成績比的敘述分析(續)

解釋變數	最小值	最大值	平均數	標準差
每局被上壘率	0.2	4.13	1.194	0.667

表 4.1.3 是主客場各項成績差的敘述分析，表中顯示的項目有最小值、最大值、平均數及標準差。

表 4.1.3 各項成績差的敘述分析

解釋變數	最小值	最大值	平均數	標準差
打數	-16	14	1.408	4.377
每打數平均打點	-0.31	0.4	0.003	0.123
二壘安打	-5	6	0.196	1.910
三壘安打	-4	2	-0.025	0.613
雙殺打	-4	4	0.096	1.285
犧牲短打	-4	2	-0.042	0.862
犧牲飛球	-3	3	0.017	0.781
盜壘成功	-6	5	0.071	1.463
盜壘刺殺	-3	2	0.129	0.851
失誤	-4	5	0.146	1.316
打擊率	-0.33	0.33	-0.005	0.114
壘打數	-32	35	1.196	10.721
上壘率	-0.32	0.31	-0.005	0.122
長打率	-0.64	0.64	0.006	0.211
攻擊指數	-1.44	1.43	0.004	0.532
面對打擊數	-20	16	-1.546	5.761
投球數	-119	67	-4.067	27.450

表 4.1.3 各項成績差的敘述分析(續)

解釋變數	最小值	最大值	平均數	標準差
好球數	-58	37	-2.438	15.058
被安打	-15	18	-0.313	5.100
被全壘打	-7	4	-0.183	1.528
四壞球	-9	7	-0.113	2.864
四死球	-4	3	-0.058	1.013
奪三振	-15	10	-0.146	3.971
三振四壞比	-23	21.50	0.041	5.593
暴投	-3	3	0.008	1.075
投手犯規	-1	1	0.013	0.144
防禦率	-17	15	-0.138	5.095
每局被上壘率	-2.22	2.08	0.047	0.796

## 第二節 以主客隊各項成績建立模型與預測

### 一、比賽進行 20%(賽季前 48 場)

表 4.2.1 是主客場各項成績在比賽進行 20%時選入 56 個解釋變數的單變數邏輯斯迴歸，其中客隊每打數平均打點、客隊失誤、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊每打數平均打點、主隊團隊打擊率、主隊壘打數、主隊上壘率、主隊長打率、主隊攻擊指數、客隊被安打、客隊被全壘打、客隊防禦率、客隊每局被上壘率、主隊面對打擊數、主隊被安打、主隊四壞球、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.1 比賽進行 20%時各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S.E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.202	0.109	3.442	1	0.064	0.817
客隊每打數平均打點	-19.112	6.401	8.915	1	0.003	<0.001
客隊二壘安打	-0.081	0.205	0.157	1	0.692	0.922
客隊三壘安打	-1.550	1.087	2.035	1	0.154	0.212
客隊雙殺打	-0.185	0.346	0.285	1	0.594	0.831
客隊犧牲短打	-0.302	0.510	0.350	1	0.554	0.740
客隊犧牲飛球	0.414	0.488	0.721	1	0.396	1.514
客隊盜壘成功	-0.306	0.288	1.134	1	0.287	0.736
客隊盜壘刺殺	0.259	0.441	0.344	1	0.558	1.295
客隊失誤	1.023	0.419	5.967	1	0.015	2.782
客隊團隊打擊率	-14.564	6.358	5.247	1	0.022	<0.001
客隊壘打數	-0.162	0.066	5.996	1	0.014	0.851
客隊上壘率	-16.306	5.864	7.732	1	0.005	<0.001
客隊長打率	-6.641	2.776	5.724	1	0.017	0.001
客隊攻擊指數	-4.864	1.707	8.119	1	0.004	0.008
主隊打數	0.105	0.093	1.274	1	0.259	1.110
主隊每打數平均打點	19.020	6.443	8.715	1	0.003	$1.821 \times 10^8$
主隊二壘安打	0.128	0.264	0.235	1	0.628	1.136
主隊三壘安打	-0.597	0.964	0.383	1	0.536	0.551
主隊雙殺打	0.091	0.356	0.066	1	0.798	1.095
主隊犧牲短打	0.030	0.594	0.003	1	0.960	1.030
主隊犧牲飛球	0.693	0.778	0.794	1	0.373	2.000
主隊盜壘成功	0.369	0.344	1.151	1	0.283	1.446

表 4.2.1 比賽進行 20%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S.E.	Wals	df	P 值	勝算比估計值
主隊盜壘刺殺	-0.391	0.655	0.356	1	0.551	0.677
主隊失誤	-0.152	0.364	0.174	1	0.677	0.859
主隊團隊打擊率	13.977	5.251	7.084	1	0.008	1175701.010
主隊壘打數	0.149	0.056	7.006	1	0.008	1.161
主隊上壘率	14.102	5.092	7.670	1	0.006	1332207.026
主隊長打率	10.358	3.350	9.562	1	0.002	31515.643
主隊攻擊指數	3.685	1.274	8.366	1	0.004	39.846
客隊面對打擊數	0.102	0.071	2.057	1	0.152	1.107
客隊投球數	0.003	0.016	0.037	1	0.848	1.003
客隊好球數	-0.015	0.029	0.252	1	0.615	0.985
客隊被安打	0.294	0.120	5.972	1	0.015	1.342
客隊被全壘打	1.237	0.457	7.327	1	0.007	3.445
客隊四壞球	0.326	0.168	3.756	1	0.053	1.386
客隊四死球	-1.067	0.557	3.668	1	0.055	0.344
客隊奪三振	-0.242	0.135	3.217	1	0.073	0.785
客隊三振四壞比	-0.175	0.099	3.149	1	0.076	0.839
客隊暴投	-0.368	0.360	1.040	1	0.308	0.692
客隊投手犯規	21.158	28420.722	0.000	1	0.999	$1.545 \times 10^9$
客隊防禦率	0.465	0.155	8.983	1	0.003	1.593
客隊每局被上壘率	2.983	0.922	10.461	1	0.001	19.749
主隊面對打擊數	-0.212	0.084	6.312	1	0.012	0.809
主隊投球數	-0.026	0.015	2.948	1	0.086	0.974
主隊好球數	-0.012	0.024	0.228	1	0.633	0.988

表 4.2.1 比賽進行 20%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S.E.	Wals	df	P 值	勝算比估計值
客隊被安打	0.294	0.120	5.972	1	0.015	1.342
客隊被全壘打	1.237	0.457	7.327	1	0.007	3.445
客隊四壞球	0.326	0.168	3.756	1	0.053	1.386
客隊四死球	-1.067	0.557	3.668	1	0.055	0.344
客隊奪三振	-0.242	0.135	3.217	1	0.073	0.785
客隊三振四壞比	-0.175	0.099	3.149	1	0.076	0.839
客隊暴投	-0.368	0.360	1.040	1	0.308	0.692
客隊投手犯規	21.158	28420.722	0.000	1	0.999	1.545×10 <sup>9</sup>
客隊防禦率	0.465	0.155	8.983	1	0.003	1.593
客隊每局被上壘率	2.983	0.922	10.461	1	0.001	19.749
主隊面對打擊數	-0.212	0.084	6.312	1	0.012	0.809
主隊投球數	-0.026	0.015	2.948	1	0.086	0.974
主隊好球數	-0.012	0.024	0.228	1	0.633	0.988
主隊被安打	-0.330	0.138	5.678	1	0.017	0.719
主隊被全壘打	-0.543	0.286	3.614	1	0.057	0.581
主隊四壞球	-0.305	0.147	4.308	1	0.038	0.737
主隊四死球	-0.377	0.357	1.115	1	0.291	0.686
主隊奪三振	0.000	0.116	0.000	1	1.000	1.000
主隊三振四壞比	0.127	0.079	2.560	1	0.110	1.135
主隊暴投	-0.035	0.452	0.006	1	0.939	0.966
主隊投手犯規	X	X	X	X	X	X
主隊防禦率	-0.510	0.168	9.237	1	0.002	0.600
主隊每局被上壘率	-2.769	0.947	8.560	1	0.003	0.063

由表 4.2.2 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -8.962 - 3.616 \times \text{客隊暴投} + 13.975 \times \text{客隊每局被上壘率} - 1.82 \times \text{主隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.2 得到，客隊暴投的參數估計值為-3.616，勝算比估計值為 0.027，表示在其它條件固定下，客隊每多一次暴投，主隊獲勝的機率變為原本 0.027 倍，但這與常理判斷並不相同，客隊暴投增加時，理應主隊勝率會提升；客隊每局被上壘率的參數估計值為 13.975，勝算比估計值 1172871.497，表示在其它條件固定下，客隊每局每多讓一個人上壘，主隊獲勝的機率變為原本 1172871.497 倍；主隊防禦率的參數估計值為-1.82，勝算比估計值為 0.162，表示在其它條件固定下，主隊投手每九局每多失一自責分，主隊獲勝的機率變為原本的 0.162 倍。Hosmer-Lemeshow 配適度的統計量 0.733，P 值為 0.998 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.2 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-8.962	4.585	3.82	1	0.051	0.000
客隊暴投	-3.616	1.718	4.429	1	0.035	0.027
客隊每局被上壘率	13.975	5.713	5.985	1	0.014	1172871.497
主隊防禦率	-1.82	0.725	6.303	1	0.012	0.162
配適度	卡方	0.733	df	7	P 值	0.998

## 二、比賽進行 40%(賽季前 96 場)

表 4.2.3 是主客場各項成績在比賽進行 40%時選入 56 個解釋變數的單變數邏輯斯迴歸，其中客隊打數、客隊每打數平均打點、客隊失誤、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊每打數平均打點、主隊團隊打擊率、主隊壘打數、主隊上壘率、主隊長打率、主隊攻擊指

數、客隊被安打、客隊被全壘打、客隊四壞球、客隊三振四壞比、客隊防禦率、客隊每局被上壘率、主隊面對打擊數、主隊投球數、主隊被安打、主隊被全壘打、主隊四壞球、主隊四死球、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.3 比賽進行 40%時各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.171	0.060	8.005	1	0.005	0.843
客隊每打數平均打點	-20.266	4.424	20.982	1	0.000	<0.001
客隊二壘安打	-0.252	0.146	2.953	1	0.086	0.778
客隊三壘安打	-0.943	0.582	2.628	1	0.105	0.389
客隊雙殺打	-0.273	0.209	1.713	1	0.191	0.761
客隊犧牲短打	-0.090	0.384	0.055	1	0.815	0.914
客隊犧牲飛球	-0.073	0.348	0.045	1	0.833	0.929
客隊盜壘成功	-0.423	0.217	3.792	1	0.052	0.655
客隊盜壘刺殺	0.278	0.334	0.691	1	0.406	1.320
客隊失誤	0.619	0.272	5.164	1	0.023	1.856
客隊團隊打擊率	-15.564	3.893	15.988	1	0.000	<0.001
客隊壘打數	-0.189	0.048	15.484	1	0.000	0.828
客隊上壘率	-16.098	3.816	17.797	1	0.000	<0.001
客隊長打率	-8.585	2.130	16.244	1	0.000	<0.001
客隊攻擊指數	-4.987	1.173	18.088	1	0.000	0.007
主隊打數	0.030	0.056	0.294	1	0.588	1.031
主隊每打數平均打點	15.918	3.810	17.452	1	0.000	8184108.948
主隊二壘安打	0.197	0.152	1.678	1	0.195	1.218

表 4.2.3 比賽進行 40%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊三壘安打	-0.742	0.682	1.182	1	0.277	0.476
主隊雙殺打	-0.204	0.241	0.718	1	0.397	0.816
主隊犧牲短打	0.589	0.399	2.174	1	0.140	1.802
主隊犧牲飛球	0.933	0.500	3.476	1	0.062	2.541
主隊盜壘成功	0.268	0.200	1.788	1	0.181	1.307
主隊盜壘刺殺	-0.213	0.417	0.261	1	0.610	0.808
主隊失誤	-0.246	0.253	0.952	1	0.329	0.782
主隊團隊打擊率	11.688	3.219	13.185	1	0.000	119143.536
主隊壘打數	0.120	0.034	12.701	1	0.000	1.127
主隊上壘率	12.532	3.174	15.586	1	0.000	277099.874
主隊長打率	7.907	1.928	16.819	1	0.000	2715.857
主隊攻擊指數	2.942	0.732	16.131	1	0.000	18.949
客隊面對打擊數	0.079	0.044	3.177	1	0.075	1.082
客隊投球數	0.006	0.010	0.335	1	0.563	1.006
客隊好球數	-0.006	0.018	0.090	1	0.764	0.994
客隊被安打	0.214	0.069	9.461	1	0.002	1.238
客隊被全壘打	0.997	0.271	13.550	1	0.000	2.709
客隊四壞球	0.344	0.124	7.766	1	0.005	1.411
客隊四死球	-0.256	0.282	0.822	1	0.364	0.774
客隊奪三振	-0.099	0.077	1.649	1	0.199	0.905
客隊三振四壞比	-0.150	0.064	5.519	1	0.019	0.861
客隊暴投	0.111	0.255	0.191	1	0.662	1.118
客隊投手犯規	0.450	1.243	0.131	1	0.717	1.569

表 4.2.3 比賽進行 40%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊防禦率	0.408	0.098	17.522	1	0.000	1.504
客隊每局被上壘率	2.329	0.554	17.711	1	0.000	10.273
主隊面對打擊數	-0.175	0.051	11.765	1	0.001	0.839
主隊投球數	-0.022	0.009	5.466	1	0.019	0.978
主隊好球數	-0.019	0.016	1.430	1	0.232	0.982
主隊被安打	-0.315	0.082	14.858	1	0.000	0.730
主隊被全壘打	-0.684	0.230	8.821	1	0.003	0.505
主隊四壞球	-0.217	0.103	4.390	1	0.036	0.805
主隊四死球	-0.555	0.280	3.922	1	0.048	0.574
主隊奪三振	0.110	0.076	2.097	1	0.148	1.117
主隊三振四壞比	0.051	0.053	0.934	1	0.334	1.052
主隊暴投	-0.288	0.294	0.961	1	0.327	0.750
主隊投手犯規	x	x	x	x	x	x
主隊防禦率	-0.514	0.116	19.556	1	0.000	0.598
主隊每局被上壘率	-2.607	0.615	17.945	1	0.000	0.074

由表 4.2.4 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = 1.828 - 174.766 \times \text{客隊每打數平均打點} + 4.683 \times \text{客隊失誤} - 3.769 \times \text{主隊盜壘刺殺} + 3.933 \times \text{客隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.4 得到，客隊每打數平均打點的參數估計值為-174.766，勝算比估計值為 0.000，表示在其它條件固定下，客隊每打數平均打點每多一分打點，主隊獲勝的機率將低於原本的 0.001 倍(SPSS 報表跑出最多是小數點後三位，所

以當數值為 0.000 時，並非數值為 0，而是小於 0.001)；客隊失誤的參數估計值為 4.683，勝算估計值為 108.070，表示在其它條件固定下，客隊每多一次失誤，主隊獲勝的機率變成原本的 108.070 倍；主隊盜壘刺殺的參數估計值為-3.769，勝算比估計值為 0.023，表示在其它條件固定下，主隊每多一次盜壘刺殺，主隊獲勝的機率變成原本的 0.023 倍；客隊防禦率的參數估計值為 3.933，勝算估計值為 51.071，表示其它條件固定下，客隊每九局每多失一自責分，主隊獲勝的機率變成原本的 51.071 倍。Hosmer-Lemeshow 配適度的統計量 1.166，P 值為 0.997 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.4 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	1.828	1.973	0.859	1	0.354	6.221
客隊每打數平均打點	-174.766	71.255	6.016	1	0.014	0.000
客隊失誤	4.683	2.237	4.383	1	0.036	108.070
主隊盜壘刺殺	-3.769	1.801	4.380	1	0.036	0.023
客隊防禦率	3.933	1.656	5.639	1	0.018	51.071
配適度	卡方	1.166	df	8	P 值	0.997

### 三、比賽進行 70%(賽季前 168 場)

表 4.2.5 是主客場各項成績在比賽進行 70%時選入 56 個解釋變數的單變數邏輯斯迴歸，其中客隊打數、客隊每打數平均打點、客隊二壘安打、客隊犧牲短打、客隊盜壘成功、客隊失誤、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊每打數平均打點、主隊犧牲短打、主隊犧牲飛球、主隊失誤、主隊團隊打擊率、主隊壘打數、主隊上壘率、主隊長打率、主隊攻擊指數、客隊被安打、客隊被全壘打、客隊四壞球、客隊奪三振、客隊三振四壞比、客隊防禦率、客隊每局被上壘率、主隊面對打擊數、主隊投

球數、主隊好球數、主隊被安打、主隊被全壘打、主隊四壞球、主隊四死球、主隊奪三振、主隊三振四壞比、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.5 比賽進行 70%時各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.205	0.046	20.060	1	0.000	0.815
客隊每打數平均打點	-22.767	3.533	41.536	1	0.000	<0.001
客隊二壘安打	-0.391	0.115	11.541	1	0.001	0.676
客隊三壘安打	-0.635	0.44	2.079	1	0.149	0.53
客隊雙殺打	-0.133	0.159	0.704	1	0.401	0.875
客隊犧牲短打	-0.595	0.286	4.321	1	0.038	0.552
客隊犧牲飛球	-0.346	0.269	1.663	1	0.197	0.707
客隊盜壘成功	-0.310	0.158	3.866	1	0.049	0.734
客隊盜壘刺殺	0.045	0.243	0.035	1	0.852	1.046
客隊失誤	0.571	0.179	10.148	1	0.001	1.769
客隊團隊打擊率	-16.581	2.869	33.411	1	0.000	<0.001
客隊壘打數	-0.204	0.036	32.637	1	0.000	0.815
客隊上壘率	-18.230	3.009	36.714	1	0.000	<0.001
客隊長打率	-9.736	1.661	34.343	1	0.000	<0.001
客隊攻擊指數	-5.536	0.920	36.201	1	0.000	0.004
主隊打數	-0.043	0.041	1.094	1	0.296	0.958
主隊每打數平均打點	16.218	2.884	31.618	1	0.000	$1.105 \times 10^7$
主隊二壘安打	0.117	0.112	1.089	1	0.297	1.124
主隊三壘安打	0.047	0.384	0.015	1	0.903	1.048
主隊雙殺打	-0.145	0.173	0.706	1	0.401	0.865

表 4.2.5 比賽進行 70%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊犧牲短打	0.648	0.263	6.064	1	0.014	1.911
主隊犧牲飛球	1.113	0.345	10.397	1	0.001	3.044
主隊盜壘成功	0.187	0.148	1.591	1	0.207	1.205
主隊盜壘刺殺	0.029	0.308	0.009	1	0.925	1.029
主隊失誤	-0.429	0.194	4.899	1	0.027	0.651
主隊團隊打擊率	10.552	2.380	19.664	1	0.000	38243.941
主隊壘打數	0.112	0.024	21.185	1	0.000	1.119
主隊上壘率	12.596	2.462	26.179	1	0.000	295514.535
主隊長打率	7.910	1.533	26.616	1	0.000	2725.408
主隊攻擊指數	2.857	0.529	29.219	1	0.000	17.415
客隊面對打擊數	0.050	0.033	2.320	1	0.128	1.051
客隊投球數	0.000	0.007	0.004	1	0.949	1.000
客隊好球數	-0.020	0.014	2.080	1	0.149	0.980
客隊被安打	0.163	0.049	10.903	1	0.001	1.177
客隊被全壘打	0.813	0.196	17.237	1	0.000	2.255
客隊四壞球	0.295	0.084	12.180	1	0.000	1.343
客隊四死球	-0.019	0.225	0.007	1	0.932	0.981
客隊奪三振	-0.240	0.066	13.194	1	0.000	0.787
客隊三振四壞比	-0.208	0.064	10.485	1	0.001	0.812
客隊暴投	0.177	0.201	0.779	1	0.378	1.194
客隊投手犯規	0.606	1.235	0.241	1	0.624	1.833
客隊防禦率	0.342	0.068	25.401	1	0.000	1.408
客隊每局被上壘率	2.156	0.408	27.897	1	0.000	8.632

表 4.2.5 比賽進行 70%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊面對打擊數	-0.236	0.042	31.206	1	0.000	0.790
主隊投球數	-0.034	0.008	17.732	1	0.000	0.967
主隊好球數	-0.038	0.013	8.659	1	0.003	0.963
主隊被安打	-0.367	0.062	34.471	1	0.000	0.693
主隊被全壘打	-0.640	0.164	15.255	1	0.000	0.527
主隊四壞球	-0.299	0.081	13.677	1	0.000	0.742
主隊四死球	-0.428	0.202	4.511	1	0.034	0.652
主隊奪三振	0.120	0.059	4.128	1	0.042	1.128
主隊三振四壞比	0.109	0.054	4.134	1	0.042	1.116
主隊暴投	-0.380	0.204	3.459	1	0.063	0.684
主隊投手犯規	X	X	X	X	X	X
主隊防禦率	-0.545	0.087	39.332	1	0.000	0.580
主隊每局被上壘率	-3.293	0.526	39.118	1	0.000	0.037

由表 4.2.6 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=2.542-111.578\times\text{客隊每打數平均打點}+3.297\times\text{客隊失誤}-0.282\times\text{客隊三振四壞比}+2.233\times\text{客隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.6 得到，客隊每打數平均打點的參數估計值為-111.578，勝算比估計值為 0.000，表示其它條件固定下，客隊每打數平均打點每多一分打點，主隊獲勝的機率低於原本的 0.001 倍；客隊失誤的參數估計值為 3.297，勝算比估計值為 27.029，表示其它條件固定下，客隊失誤每增加一次，主隊獲勝的機率變成原本的 27.029 倍；客隊三振四壞比的參數估計值為-0.282，勝算比估計值為

0.754，表示其它條件固定下，客隊投手的三振四壞比每多一個單位，主隊獲勝的機率變成原本的 0.754 倍；客隊防禦率的參數估計值為 2.233，勝算比估計值為 9.327，表示其它條件固定下，客隊每九局每多失一自責分，主隊獲勝的機率變成原本的 9.327 倍。Hosmer-Lemeshow 配適度的統計量 0.530，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.6 比賽結果之參數估計與配適度

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
常數	2.542	1.515	2.816	1	0.093	12.700
客隊每打數平均打 點	-111.578	30.946	13.000	1	0.000	0.000
客隊失誤	3.297	1.107	8.875	1	0.003	27.029
客隊三振四壞比	-0.282	0.129	4.791	1	0.029	0.754
客隊防禦率	2.233	0.669	11.151	1	0.001	9.327
配適度	卡方	0.530	Df	8	P 值	1.000

#### 四、比賽進行 90%(賽季前 216 場)

表 4.2.7 是主客場各項成績在比賽進行 90%時選入 56 個解釋變數的單變數邏輯斯迴歸，其中客隊打數、客隊每打數平均打點、客隊二壘安打、客隊犧牲飛球、客隊失誤、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊每打數平均打點、主隊犧牲短打、主隊犧牲飛球、主隊團隊打擊率、主隊壘打數、主隊上壘率、主隊長打率、主隊攻擊指數、客隊被安打、客隊被全壘打、客隊四壞球、客隊奪三振、客隊三振四壞比、客隊防禦率、客隊每局被上壘率、主隊面對打擊數、主隊投球數、主隊好球數、主隊被安打、主隊被全壘打、主隊四壞球、主隊四死球、主隊奪三振、主隊三振四壞

比、主隊暴投、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.7 比賽進行 90%時各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.197	0.040	24.244	1	0.000	0.822
客隊每打數平均打點	-20.430	2.863	50.901	1	0.000	<0.001
客隊二壘安打	-0.424	0.103	16.961	1	0.000	0.654
客隊三壘安打	-0.651	0.374	3.04	1	0.081	0.521
客隊雙殺打	-0.153	0.146	1.090	1	0.296	0.858
客隊犧牲短打	-0.462	0.240	3.704	1	0.054	0.630
客隊犧牲飛球	-0.548	0.243	5.089	1	0.024	0.578
客隊盜壘成功	-0.279	0.144	3.767	1	0.052	0.756
客隊盜壘刺殺	-0.035	0.208	0.028	1	0.866	0.965
客隊失誤	0.456	0.149	9.364	1	0.002	1.577
客隊團隊打擊率	-17.908	2.661	45.302	1	0.000	<0.001
客隊壘打數	-0.203	0.032	41.205	1	0.000	0.816
客隊上壘率	-18.171	2.629	47.769	1	0.000	<0.001
客隊長打率	-9.518	1.447	43.244	1	0.000	<0.001
客隊攻擊指數	-5.345	0.789	45.883	1	0.000	0.005
主隊打數	-0.021	0.037	0.328	1	0.567	0.979
主隊每打數平均打點	17.803	2.728	42.578	1	0.000	$5.393 \times 10^7$
主隊二壘安打	0.134	0.103	1.685	1	0.194	1.144
主隊三壘安打	0.150	0.295	0.258	1	0.611	1.162
主隊雙殺打	-0.163	0.161	1.021	1	0.312	0.850
主隊犧牲短打	0.688	0.238	8.341	1	0.004	1.990

表 4.2.7 比賽進行 90%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊犧牲飛球	0.829	0.266	9.719	1	0.002	2.290
主隊盜壘成功	0.217	0.133	2.654	1	0.103	1.242
主隊盜壘刺殺	0.127	0.265	0.231	1	0.631	1.136
主隊失誤	-0.174	0.169	1.061	1	0.303	0.841
主隊團隊打擊率	12.069	2.157	31.320	1	0.000	174419.858
主隊壘打數	0.126	0.023	30.230	1	0.000	1.134
主隊上壘率	13.580	2.249	36.463	1	0.000	790475.263
主隊長打率	8.251	1.363	36.640	1	0.000	3829.869
主隊攻擊指數	3.111	0.501	38.538	1	0.000	22.442
客隊面對打擊數	0.053	0.030	3.150	1	0.076	1.054
客隊投球數	0.000	0.006	0.003	1	0.957	1.000
客隊好球數	-0.015	0.012	1.560	1	0.212	0.986
客隊被安打	0.201	0.045	19.779	1	0.000	1.223
客隊被全壘打	0.681	0.163	17.531	1	0.000	1.975
客隊四壞球	0.217	0.071	9.372	1	0.002	1.242
客隊四死球	0.023	0.196	0.013	1	0.908	1.023
客隊奪三振	-0.205	0.057	12.886	1	0.000	0.814
客隊三振四壞比	-0.133	0.043	9.406	1	0.002	0.875
客隊暴投	0.271	0.184	2.169	1	0.141	1.311
客隊投手犯規	-0.077	1.010	0.006	1	0.939	0.926
客隊防禦率	0.385	0.064	35.792	1	0.000	1.469
客隊每局被上壘率	2.278	0.373	37.347	1	0.000	9.760
主隊面對打擊數	-0.225	0.036	38.213	1	0.000	0.798

表 4.2.7 比賽進行 90%時各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊投球數	-0.036	0.007	24.184	1	0.000	0.965
主隊好球數	-0.041	0.012	12.624	1	0.000	0.959
主隊被安打	-0.376	0.056	44.990	1	0.000	0.687
主隊被全壘打	-0.580	0.144	16.329	1	0.000	0.560
主隊四壞球	-0.254	0.068	14.051	1	0.000	0.776
主隊四死球	-0.474	0.185	6.600	1	0.010	0.622
主隊奪三振	0.107	0.052	4.218	1	0.040	1.112
主隊三振四壞比	0.104	0.045	5.288	1	0.021	1.110
主隊暴投	-0.488	0.192	6.468	1	0.011	0.614
主隊投手犯規	-21.288	40192.970	0.000	1	1.000	<0.001
主隊防禦率	-0.521	0.074	48.960	1	0.000	0.594
主隊每局被上壘率	-3.096	0.441	49.200	1	0.000	0.045

由表 4.2.8 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.559 - 63.351 \times \text{客隊每打數平均打點} + 3.486 \times \text{客隊失誤} - 30.282 \times \text{客隊上壘率} + 47.446 \times \text{主隊每打數平均打點} + 3.515 \times \text{主隊犧牲飛球} - 4.16 \times \text{主隊盜壘刺殺} + 38.284 \times \text{主隊打擊率} + 2.059 \times \text{客隊防禦率} - 2.175 \times \text{主隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.8 得到，客隊每打數平均打點的參數估計值為-63.351，勝算比估計值為 0.000，表示在其它條件固定下，客隊每打數平均打點每多一分打點，主隊獲勝的機率低於原本的 0.001 倍；客隊失誤的參數估計值為 3.486，勝算比估計值為 32.669，表示在其它條件固定下，客隊失誤每增加一次，主隊獲勝的機率變成原本的 32.669 倍；客隊上壘率的參數估計值為-30.282，勝算比估計值為

0.000，表示在其它條件固定下，客隊上壘率每增加 1%，主隊獲勝的機率低於原本的 0.001 倍；主隊每打數平均打點的參數估計值為 47.446，勝算比估計值為  $4.031 \times 10^{20}$ ，表示在其它條件固定下，主隊每打數平均打點每多一分打點，主隊獲勝的機率變成原本的  $4.031 \times 10^{20}$  倍；主隊犧牲飛球的參數估計值為 3.515，勝算比估計值為 33.633，表示在其它條件固定下，主隊犧牲飛球每增加一個，主隊獲勝的機率變成原本的 33.633 倍；主隊盜壘刺殺的參數估計值為 -4.160，勝算比估計值為 0.016，表示在其它條件固定下，主隊盜壘刺殺每增加一個，主隊獲勝的機率變成原本的 0.016 倍；主隊打擊率的參數估計值為 38.284，勝算比估計值為  $4.231 \times 10^{16}$ ，表示在其它條件固定下，主隊打擊率每增加 1%，主隊獲勝的機率變成原本的  $4.231 \times 10^{16}$  倍；客隊防禦率的參數估計值為 2.059，勝算比估計值為 7.835，表示在其它條件固定下，客隊每九局每多失一自責分，主隊獲勝的機率變成原本的 7.835 倍；主隊防禦率的參數估計值為 -2.175，勝算比估計值為 0.114，表示在其它條件固定下，主隊每多失一自責分，主隊獲勝的機率變成原本的 0.114 倍。Hosmer-Lemeshow 配適度的統計量 0.107，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.8 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-0.559	4.409	0.016	1	0.899	0.572
客隊每打數平均打點	-63.351	26.898	5.547	1	0.019	0.000
客隊失誤	3.486	1.144	9.290	1	0.002	32.669
客隊上壘率	-30.282	12.594	5.782	1	0.016	0.000
主隊每打數平均打點	47.446	18.616	6.495	1	0.011	$4.031 \times 10^{20}$

表 4.2.8 比賽結果之參數估計與配適度(續)

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
主隊犧牲飛球	3.515	1.633	4.632	1	0.031	33.633
主隊盜壘刺殺	-4.160	1.637	6.461	1	0.011	0.016
主隊打擊率	38.284	19.087	4.023	1	0.045	$4.231 \times 10^{16}$
客隊防禦率	2.059	0.758	7.383	1	0.007	7.835
主隊防禦率	-2.175	0.893	5.928	1	0.015	0.114
配適度	卡方	0.107	df	8	P 值	1.000

### 五、比賽進行下半季 20%(賽季第 121-168 場)

由表 4.2.9 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -4.675 - 55.620 \times \text{客隊每打數平均打點} + 35.993 \times \text{主隊上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.9 得到，客隊每打數平均打點的參數估計值為 -55.620，勝算比估計值為 0.000，表示在其它條件固定下，客隊每打數平均打點每多一分打點，主隊獲勝的機率將低於原本的 0.001 倍；主隊上壘率的參數估計值為 35.993，勝算比估計值為  $4.282 \times 10^{15}$ ，表示在其它條件固定下，主隊上壘率每增加 1%，主隊獲勝的機率變成原本的  $4.282 \times 10^{15}$  倍。Hosmer-Lemeshow 配適度的統計量 2.987，P 值為 0.935 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.9 比賽結果之參數估計與配適度

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
常數	-4.675	3.664	1.628	1	0.202	0.009

表 4.2.9 比賽結果之參數估計與配適度(續)

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
客隊每打數平均打 點	-55.620	24.431	5.183	1	0.023	0.000
主隊上壘率	35.993	14.938	5.806	1	0.016	$4.282 \times 10^{15}$
配適度	卡方	2.987	df	8	P 值	0.935

## 六、比賽進行下半季 40%(賽季第 121-216 場)

由表 4.2.10 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = 13.990 - 68.169 \times \text{客隊上壘率} + 104.568 \times \text{主隊每打數平均打點} - 1.073 \times \text{客隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.2.10 得到，客隊上壘率的參數估計值為-68.169，勝算比估計值為 0.000，表示在其它條件固定下，客隊上壘率每增加 1%，主隊獲勝的機率低於原本的 0.001 倍；主隊每打數平均打點的參數估計值為 104.568，勝算比估計值為  $2.59 \times 10^{45}$ ，表示在其它條件固定下，主隊每打數平均打點每多一分打點，主隊獲勝的機率變成原本的  $2.59 \times 10^{45}$  倍；客隊防禦率的參數估計值為-1.073，勝算比估計值為 0.342，表示在其它條件固定下，客隊每九局每多失一自責分，主隊獲勝的機率變成原本的 0.342 倍，不過這與一般想法不同。Hosmer-Lemeshow 配適度的統計量 1.900，P 值為 0.984 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.2.10 比賽結果之參數估計與配適度

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
常數	13.990	4.508	9.632	1	0.002	1191068.045
客隊上壘率	-68.169	19.253	12.536	1	0.000	0.000
主隊每打數平均打 點	104.568	30.898	11.454	1	0.001	$2.59 \times 10^{45}$
客隊防禦率	-1.073	0.426	6.335	1	0.012	0.342
配適度	卡方	1.900	df	8	P 值	0.984

## 七、 前一場模型建立

表 4.2.11 是主客場各項成績在比賽進行 20%時選入前一場 56 個解釋變數的單變數邏輯斯迴歸，其中客隊每打數平均打點、客隊攻擊指數、客隊奪三振、主隊被全壘打、主隊防禦率的 P 值均小於 0.05，達到顯著水準。

表 4.2.11 比賽進行 20%時前一場各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	0.042	0.090	0.214	1	0.643	1.043
客隊每打數平均打點	-12.192	5.115	5.681	1	0.017	<0.001
客隊二壘安打	-0.156	0.213	0.537	1	0.464	0.856
客隊三壘安打	-0.216	0.719	0.09	1	0.764	0.806
客隊雙殺打	-0.392	0.364	1.157	1	0.282	0.676
客隊犧牲短打	-0.622	0.537	1.339	1	0.247	0.537
客隊犧牲飛球	-0.505	0.495	1.042	1	0.307	0.603
客隊盜壘成功	0.032	0.266	0.014	1	0.905	1.032

表 4.2.11 比賽進行 20%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊盜壘刺殺	0.048	0.451	0.011	1	0.916	1.049
客隊失誤	0.299	0.334	0.798	1	0.372	1.348
客隊團隊打擊率	-2.678	5.133	0.272	1	0.602	0.069
客隊壘打數	-0.100	0.056	3.216	1	0.073	0.905
客隊上壘率	-6.070	4.439	1.870	1	0.171	0.002
客隊長打率	-4.803	2.517	3.641	1	0.056	0.008
客隊攻擊指數	-2.379	1.180	4.066	1	0.044	0.093
主隊打數	0.031	0.089	0.117	1	0.732	1.031
主隊每打數平均打點	-0.799	3.729	0.046	1	0.830	0.450
主隊二壘安打	-0.008	0.252	0.001	1	0.974	0.992
主隊三壘安打	-21.343	20096.485	0.000	1	0.999	<0.001
主隊雙殺打	-0.734	0.393	3.486	1	0.062	0.480
主隊犧牲短打	0.217	0.575	0.142	1	0.706	1.242
主隊犧牲飛球	-0.799	0.778	1.053	1	0.305	0.450
主隊盜壘成功	-0.304	0.346	0.774	1	0.379	0.738
主隊盜壘刺殺	0.238	0.624	0.145	1	0.703	1.268
主隊失誤	-0.007	0.337	0.000	1	0.983	0.993
主隊團隊打擊率	0.808	3.828	0.045	1	0.833	2.244
主隊壘打數	0.002	0.039	0.002	1	0.961	1.002
主隊上壘率	2.085	3.648	0.327	1	0.568	8.047
主隊長打率	-0.940	2.129	0.195	1	0.659	0.390
主隊攻擊指數	0.261	0.900	0.084	1	0.772	1.298
客隊面對打擊數	0.038	0.061	0.387	1	0.534	1.038

表 4.2.11 比賽進行 20%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊投球數	-0.002	0.015	0.021	1	0.885	0.998
客隊好球數	0.001	0.027	0.001	1	0.978	1.001
客隊被安打	0.006	0.089	0.004	1	0.949	1.006
客隊被全壘打	-0.125	0.299	0.174	1	0.677	0.883
客隊四壞球	0.154	0.140	1.214	1	0.271	1.167
客隊四死球	-0.192	0.401	0.229	1	0.632	0.825
客隊奪三振	-0.577	0.196	8.641	1	0.003	0.562
客隊三振四壞比	-0.126	0.078	2.656	1	0.103	0.881
客隊暴投	-0.047	0.429	0.012	1	0.913	0.954
客隊投手犯規	-21.203	40192.970	0.000	1	1.000	<0.001
客隊防禦率	-0.007	0.087	0.006	1	0.940	0.993
客隊每局被上壘率	0.475	0.593	0.643	1	0.423	1.609
主隊面對打擊數	-0.078	0.071	1.205	1	0.272	0.925
主隊投球數	-0.009	0.015	0.344	1	0.558	0.991
主隊好球數	0.030	0.026	1.324	1	0.250	1.030
主隊被安打	0.008	0.112	0.004	1	0.947	1.008
主隊被全壘打	-0.631	0.309	4.167	1	0.041	0.532
主隊四壞球	-0.274	0.150	3.365	1	0.067	0.760
主隊四死球	-0.336	0.361	0.869	1	0.351	0.714
主隊奪三振	0.127	0.108	1.399	1	0.237	1.136
主隊三振四壞比	0.198	0.102	3.786	1	0.052	1.218
主隊暴投	0.038	0.357	0.011	1	0.915	1.039
主隊投手犯規	21.290	40192.970	0.000	1	1.000	1.762×10 <sup>9</sup>

表 4.2.11 比賽進行 20%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊防禦率	-0.320	0.131	5.918	1	0.015	0.726
主隊每局被上壘率	-0.956	0.733	1.701	1	0.192	0.384

由表 4.2.12 得到，前 48 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.535-0.731\times\text{客隊奪三振}+0.075\times\text{主隊好球數}-0.436\times\text{主隊防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.2.12 使用前一場建立模型-前 48 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.535	3.187	0.028	1	0.867	1.708
客隊奪三振	-0.731	0.247	8.743	1	0.003	0.481
主隊好球數	0.075	0.038	3.929	1	0.047	1.077
主隊防禦率	-0.436	0.172	6.468	1	0.011	0.646
配適度	卡方	10.246	df	7	P 值	0.175

表 4.2.13 是主客場各項成績在比賽進行 40%時選入前一場 56 個解釋變數的單變數邏輯斯迴歸，其中客隊每打數平均打點、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊打數、主隊二壘安打、主隊三壘安打、主隊壘打數、主隊上壘率、主隊攻擊指數、客隊面對打擊數、客隊投球數、客隊好球數、客隊四壞球、客隊三振四壞比、客隊每局被上壘率、主隊被全壘打、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.13 比賽進行 40%時前一場各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.014	0.054	0.069	1	0.793	0.986
客隊每打數平均打點	-7.192	2.825	6.481	1	0.011	0.001
客隊二壘安打	-0.177	0.157	1.269	1	0.260	0.838
客隊三壘安打	-0.798	0.537	2.206	1	0.138	0.45
客隊雙殺打	-0.284	0.217	1.710	1	0.191	0.753
客隊犧牲短打	-0.327	0.374	0.764	1	0.382	0.721
客隊犧牲飛球	-0.353	0.354	0.994	1	0.319	0.703
客隊盜壘成功	-0.365	0.209	3.062	1	0.080	0.694
客隊盜壘刺殺	0.088	0.324	0.073	1	0.787	1.092
客隊失誤	0.304	0.258	1.385	1	0.239	1.355
客隊團隊打擊率	-6.179	3.046	4.114	1	0.043	0.002
客隊壘打數	-0.071	0.033	4.572	1	0.033	0.931
客隊上壘率	-6.299	2.857	4.860	1	0.027	0.002
客隊長打率	-4.831	1.678	8.292	1	0.004	0.008
客隊攻擊指數	-1.752	0.691	6.429	1	0.011	0.173
主隊打數	0.141	0.059	5.644	1	0.018	1.152
主隊每打數平均打點	3.803	2.610	2.123	1	0.145	44.842
主隊二壘安打	0.498	0.171	8.442	1	0.004	1.645
主隊三壘安打	-2.274	1.091	4.345	1	0.037	0.103
主隊雙殺打	-0.441	0.238	3.440	1	0.064	0.643
主隊犧牲短打	0.036	0.367	0.009	1	0.922	1.036
主隊犧牲飛球	-0.188	0.429	0.191	1	0.662	0.829
主隊盜壘成功	-0.054	0.179	0.092	1	0.762	0.947

表 4.2.13 比賽進行 40%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊盜壘刺殺	0.159	0.424	0.141	1	0.708	1.172
主隊失誤	0.017	0.238	0.005	1	0.941	1.018
主隊團隊打擊率	4.356	2.517	2.994	1	0.084	77.919
主隊壘打數	0.066	0.028	5.657	1	0.017	1.068
主隊上壘率	6.242	2.527	6.099	1	0.014	513.666
主隊長打率	2.022	1.362	2.203	1	0.138	7.555
主隊攻擊指數	1.513	0.601	6.341	1	0.012	4.541
客隊面對打擊數	0.133	0.046	8.224	1	0.004	1.143
客隊投球數	0.025	0.010	5.711	1	0.017	1.025
客隊好球數	0.047	0.020	5.561	1	0.018	1.048
客隊被安打	0.109	0.057	3.611	1	0.057	1.115
客隊被全壘打	0.076	0.204	0.138	1	0.711	1.079
客隊四壞球	0.303	0.115	6.950	1	0.008	1.354
客隊四死球	0.121	0.294	0.169	1	0.681	1.129
客隊奪三振	-0.161	0.082	3.842	1	0.050	0.852
客隊三振四壞比	-0.151	0.063	5.653	1	0.017	0.860
客隊暴投	0.345	0.288	1.438	1	0.230	1.412
客隊投手犯規	-21.397	40192.970	0.000	1	1.000	<0.001
客隊防禦率	0.079	0.062	1.626	1	0.202	1.082
客隊每局被上壘率	1.010	0.401	6.340	1	0.012	2.745
主隊面對打擊數	-0.049	0.044	1.272	1	0.259	0.952
主隊投球數	-0.003	0.009	0.097	1	0.756	0.997
主隊好球數	0.011	0.016	0.486	1	0.486	1.011

表 4.2.13 比賽進行 40%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊被安打	-0.098	0.064	2.329	1	0.127	0.907
主隊被全壘打	-0.486	0.213	5.214	1	0.022	0.615
主隊四壞球	-0.119	0.104	1.305	1	0.253	0.888
主隊四死球	-0.339	0.261	1.689	1	0.194	0.713
主隊奪三振	0.148	0.077	3.729	1	0.053	1.160
主隊三振四壞比	0.129	0.067	3.705	1	0.054	1.138
主隊暴投	0.292	0.270	1.168	1	0.280	1.339
主隊投手犯規	21.072	28420.722	0.000	1	0.999	1.418×10 <sup>9</sup>
主隊防禦率	-0.190	0.072	6.938	1	0.008	0.827
主隊每局被上壘率	-1.017	0.463	4.812	1	0.028	0.362

由表 4.2.14 得到，前 96 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=2.584-6.29\times\text{客隊長打率}+0.52\times\text{主隊二壘安打}-0.159\times\text{客隊三振四壞比}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.2.14 使用前一場建立模型-前 96 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	2.584	0.985	6.877	1	0.009	13.251
客隊長打率	-6.290	2.031	9.591	1	0.002	0.002
主隊二壘安打	0.520	0.195	7.091	1	0.008	1.682
客隊三振四壞比	-0.159	0.064	6.244	1	0.012	0.853
配適度	卡方	10.399	df	8	P 值	0.238

表 4.2.15 是主客場各項成績在比賽進行 70%時選入前一場 56 個解釋變數的單變數邏輯斯迴歸，其中客隊每打數平均打點、客隊團隊打擊率、客隊壘打數、客隊上壘率、客隊長打率、客隊攻擊指數、主隊二壘安打、客隊四壞球、客隊奪三振、客隊三振四壞比、主隊被安打、主隊被全壘打、主隊奪三振、主隊三振四壞比、主隊防禦率、主隊每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.2.15 比賽進行 70%時前一場各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.046	0.038	1.482	1	0.223	0.955
客隊每打數平均打點	-5.106	1.873	7.428	1	0.006	0.006
客隊二壘安打	-0.088	0.107	0.679	1	0.410	0.916
客隊三壘安打	-0.527	0.396	1.767	1	0.184	0.591
客隊雙殺打	-0.308	0.171	3.240	1	0.072	0.735
客隊犧牲短打	-0.356	0.280	1.613	1	0.204	0.700
客隊犧牲飛球	-0.209	0.258	0.656	1	0.418	0.812
客隊盜壘成功	-0.265	0.157	2.837	1	0.092	0.767
客隊盜壘刺殺	0.294	0.242	1.476	1	0.224	1.342
客隊失誤	0.320	0.181	3.107	1	0.078	1.377
客隊團隊打擊率	-4.866	2.020	5.802	1	0.016	0.008
客隊壘打數	-0.050	0.021	5.615	1	0.018	0.951
客隊上壘率	-4.437	1.930	5.282	1	0.022	0.012
客隊長打率	-3.204	1.082	8.763	1	0.003	0.041
客隊攻擊指數	-1.098	0.447	6.024	1	0.014	0.333
主隊打數	0.014	0.037	0.142	1	0.706	1.014
主隊每打數平均打點	0.506	1.887	0.072	1	0.789	1.658

表 4.2.15 比賽進行 70%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
主隊二壘安打	0.222	0.112	3.899	1	0.048	1.249
主隊三壘安打	-0.611	0.440	1.928	1	0.165	0.543
主隊雙殺打	-0.254	0.167	2.302	1	0.129	0.776
主隊犧牲短打	0.042	0.221	0.036	1	0.849	1.043
主隊犧牲飛球	0.179	0.292	0.374	1	0.541	1.196
主隊盜壘成功	-0.190	0.144	1.744	1	0.187	0.827
主隊盜壘刺殺	0.280	0.326	0.738	1	0.390	1.323
主隊失誤	-0.206	0.166	1.553	1	0.213	0.813
主隊團隊打擊率	0.359	1.883	0.036	1	0.849	1.432
主隊壘打數	0.014	0.019	0.521	1	0.470	1.014
主隊上壘率	2.077	1.832	1.286	1	0.257	7.982
主隊長打率	0.016	1.077	0.000	1	0.988	1.016
主隊攻擊指數	0.395	0.398	0.990	1	0.320	1.485
客隊面對打擊數	0.041	0.031	1.811	1	0.178	1.042
客隊投球數	0.012	0.007	2.786	1	0.095	1.012
客隊好球數	0.015	0.013	1.233	1	0.267	1.015
客隊被安打	0.008	0.041	0.038	1	0.846	1.008
客隊被全壘打	-0.113	0.152	0.558	1	0.455	0.893
客隊四壞球	0.161	0.074	4.701	1	0.030	1.174
客隊四死球	0.018	0.197	0.008	1	0.927	1.018
客隊奪三振	-0.141	0.061	5.269	1	0.022	0.868
客隊三振四壞比	-0.145	0.053	7.471	1	0.006	0.865
客隊暴投	0.287	0.209	1.878	1	0.171	1.332

表 4.2.15 比賽進行 70%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊投手犯規	-21.289	40192.970	0.000	1	1.000	<0.001
客隊防禦率	-0.003	0.045	0.004	1	0.952	0.997
客隊每局被上壘率	0.352	0.283	1.541	1	0.214	1.422
主隊面對打擊數	-0.056	0.030	3.499	1	0.061	0.945
主隊投球數	-0.011	0.007	2.829	1	0.093	0.989
主隊好球數	-0.012	0.012	0.959	1	0.327	0.988
主隊被安打	-0.108	0.043	6.367	1	0.012	0.898
主隊被全壘打	-0.296	0.140	4.478	1	0.034	0.743
主隊四壞球	-0.072	0.072	0.984	1	0.321	0.931
主隊四死球	-0.068	0.208	0.107	1	0.744	0.934
主隊奪三振	0.150	0.060	6.222	1	0.013	1.162
主隊三振四壞比	0.160	0.061	6.799	1	0.009	1.173
主隊暴投	0.041	0.193	0.045	1	0.833	1.042
主隊投手犯規	21.154	28420.722	0.000	1	0.999	$1.538 \times 10^9$
主隊防禦率	-0.120	0.045	7.024	1	0.008	0.887
主隊每局被上壘率	-0.821	0.303	7.326	1	0.007	0.440

由表 4.2.16 得到，前 168 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = 1.271 - 2.536 \times \text{客隊長打率} - 0.16 \times \text{客隊三振四壞比} + 0.161 \times \text{主隊三振四壞比}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.2.16 使用前一場建立模型-前 168 場

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
常數	1.271	0.626	4.127	1	0.042	3.564
客隊長打率	-2.536	1.164	4.749	1	0.029	0.079
客隊三振四壞比	-0.160	0.054	8.699	1	0.003	0.853
主隊三振四壞比	0.161	0.068	5.628	1	0.018	1.174
配適度	卡方	7.868	df	8	P 值	0.446

表 4.2.17 是主客場各項成績在比賽進行 90%時選入前一場 56 個解釋變數的單變數邏輯斯迴歸，其中客隊失誤、客隊四壞球、客隊三振四壞比、主隊奪三振的 P 值均小於 0.05，達到顯著水準。

表 4.2.17 比賽進行 90%時前一場各項成績的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊打數	-0.023	0.034	0.476	1	0.490	0.977
客隊每打數平均打點	-2.208	1.551	2.027	1	0.155	0.110
客隊二壘安打	-0.015	0.097	0.023	1	0.880	0.985
客隊三壘安打	-0.269	0.289	0.87	1	0.351	0.764
客隊雙殺打	-0.128	0.150	0.729	1	0.393	0.880
客隊犧牲短打	-0.295	0.252	1.365	1	0.243	0.745
客隊犧牲飛球	-0.038	0.224	0.029	1	0.866	0.963
客隊盜壘成功	-0.257	0.145	3.156	1	0.076	0.773
客隊盜壘刺殺	0.182	0.212	0.738	1	0.390	1.199
客隊失誤	0.385	0.160	5.756	1	0.016	1.469
客隊團隊打擊率	-1.486	1.648	0.813	1	0.367	0.226

表 4.2.17 比賽進行 90%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊壘打數	-0.017	0.017	0.999	1	0.318	0.983
客隊上壘率	-0.832	1.583	0.276	1	0.599	0.435
客隊長打率	-1.489	0.896	2.763	1	0.096	0.226
客隊攻擊指數	-0.311	0.369	0.710	1	0.399	0.733
主隊打數	0.015	0.034	0.204	1	0.651	1.015
主隊每打數平均打點	0.404	1.682	0.058	1	0.810	1.498
主隊二壘安打	0.177	0.097	3.310	1	0.069	1.194
主隊三壘安打	-0.471	0.388	1.471	1	0.225	0.625
主隊雙殺打	-0.225	0.158	2.045	1	0.153	0.798
主隊犧牲短打	-0.085	0.192	0.194	1	0.659	0.919
主隊犧牲飛球	0.307	0.243	1.587	1	0.208	1.359
主隊盜壘成功	-0.084	0.125	0.455	1	0.500	0.919
主隊盜壘刺殺	0.231	0.262	0.774	1	0.379	1.259
主隊失誤	-0.145	0.147	0.972	1	0.324	0.865
主隊團隊打擊率	-0.057	1.695	0.001	1	0.973	0.945
主隊壘打數	0.014	0.018	0.597	1	0.440	1.014
主隊上壘率	1.409	1.662	0.719	1	0.397	4.091
主隊長打率	-0.247	0.971	0.065	1	0.799	0.781
主隊攻擊指數	0.392	0.370	1.122	1	0.290	1.480
客隊面對打擊數	0.039	0.028	1.910	1	0.167	1.039
客隊投球數	0.009	0.006	1.910	1	0.167	1.009
客隊好球數	0.011	0.012	0.832	1	0.362	1.011
客隊被安打	0.004	0.037	0.014	1	0.907	1.004

表 4.2.17 比賽進行 90%時前一場各項成績的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
客隊被全壘打	-0.129	0.136	0.895	1	0.344	0.879
客隊四壞球	0.136	0.066	4.330	1	0.037	1.146
客隊四死球	0.013	0.172	0.006	1	0.940	1.013
客隊奪三振	-0.086	0.052	2.780	1	0.095	0.917
客隊三振四壞比	-0.086	0.039	5.021	1	0.025	0.917
客隊暴投	0.317	0.192	2.707	1	0.100	1.372
客隊投手犯規	0.664	1.233	0.291	1	0.590	1.943
客隊防禦率	-0.001	0.040	0.001	1	0.971	0.999
客隊每局被上壘率	0.288	0.258	1.248	1	0.264	1.334
主隊面對打擊數	-0.022	0.026	0.720	1	0.396	0.978
主隊投球數	-0.005	0.006	0.830	1	0.362	0.995
主隊好球數	-0.005	0.010	0.228	1	0.633	0.995
主隊被安打	-0.040	0.035	1.284	1	0.257	0.961
主隊被全壘打	-0.180	0.121	2.184	1	0.139	0.836
主隊四壞球	-0.003	0.062	0.002	1	0.961	0.997
主隊四死球	0.091	0.190	0.230	1	0.631	1.096
主隊奪三振	0.121	0.053	5.217	1	0.022	1.128
主隊三振四壞比	0.063	0.038	2.705	1	0.100	1.065
主隊暴投	0.117	0.175	0.449	1	0.503	1.124
主隊投手犯規	21.184	28420.722	0.000	1	0.999	$1.585 \times 10^9$
主隊防禦率	-0.048	0.037	1.641	1	0.200	0.953
主隊每局被上壘率	-0.247	0.244	1.025	1	0.311	0.781

由表 4.2.18 得到，前 216 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -1 + 0.386 \times \text{客隊失誤} - 0.084 \times \text{客隊三振四壞比} + 0.141 \times \text{主隊奪三振}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.2.18 使用前一場建立模型-前 216 場

變數	參數估計 值	S.E.	Wald	df	P 值	勝算比估計 值
常數	-1.000	0.467	4.579	1	0.032	0.368
客隊失誤	0.386	0.166	5.386	1	0.020	1.471
客隊三振四壞比	-0.084	0.039	4.741	1	0.029	0.919
主隊奪三振	0.141	0.055	6.461	1	0.011	1.151
配適度	卡方	5.144	df	8	P 值	0.742

## 八、驗證分析與比較

由表 4.2.19 得到，建模時間前 48 場時，預測 3 的預測能力最好，成功場次達到 8 場；建模時間前 96 場時，預測 2 及預測 5 並列的預測能力最好，成功場次達到 9 場；建模時間前 168 場時，預測 4 的預測能力最好，成功場次達到 6 場；建模時間前 216 場時，預測 1、預測 4、預測 5 預測能力並列最好，成功場次達到 7 場；建模時間下半季前 48 場時，預測 4 預測能力最好，成功場次達到 5 場；建模時間下半季前 96 場時，預測 4 預測能力表現最好，成功場次達到 6 場。

表 4.2.19 主客隊各項成績的分析整合表

建模場次 <sup>1</sup> (預測場次)	建模成功率 <sup>2</sup> (P 值 <sup>3</sup> )	預測方式	預測 <sup>4</sup> 成功 場數	預測方式	預測 成功 場數	預測方式 平均
前 48 場 (第 49-58 場)	91.50% (0.998)	預測 1	7	預測 4	7	6.83
		預測 2	7	預測 5	7	
		預測 3	8 <sup>5</sup>	預測 6	5 <sup>5</sup>	

表 4.2.19 主客隊各項成績的分析整合表(續)

建模場次 <sup>1</sup> (預測場次)	建模成功率 <sup>2</sup> (P 值 <sup>3</sup> )	預測方式	預測 <sup>4</sup> 成功 場數	預測方式	預測 成功 場數	預測方式 平均
前 96 場 (第 97-99 場、第 102- 108 場)	98.90% (0.997)	預測 1	<u>5</u>	預測 4	7	7.50
		預測 2	<u>9</u>	預測 5	<u>9</u>	
		預測 3	7	預測 6	8	
前 168 場 (第 169-178 場)	95.70% (1.000)	預測 1	5	預測 4	<u>6</u>	4.67
		預測 2	5	預測 5	5	
		預測 3	4	預測 6	<u>3</u>	
前 216 場 (第 217-226 場)	97.60% (1.000)	預測 1	<u>7</u>	預測 4	<u>7</u>	6.33
		預測 2	6	預測 5	<u>7</u>	
		預測 3	6	預測 6	<u>5</u>	
四時間點平均 <sup>6</sup>		預測 1	6	預測 4	6.75	6.33
		預測 2	6.75	預測 5	<u>7</u>	
		預測 3	6.25	預測 6	<u>5.25</u>	
下半季前 48 場	93.80% (0.935)	預測 1	<u>3</u>	預測 4	<u>5</u>	3.83
		預測 2	4	預測 5	4	
		預測 3	<u>3</u>	預測 6	4	
下半季前 96 場	93.80% (0.984)	預測 1	5	預測 4	<u>6</u>	4.67
		預測 2	5	預測 5	5	
		預測 3	5	預測 6	<u>2</u>	

註 1. 建立模型時使用的場次，預測部分的場次則為其後 10 場，代表建立模型場次為前 48 場，預測場次則為第 49 場至第 58 場，而後以此類推(其中第 100 及 101 場結果為平局，故將其從預測中刪去，所以前 96 場的預測場次為，第 97-99 場、第 102-108 場，以下用第 97-108 場簡稱)。

2. 建立模型時比賽結果與預測結果相同的百分比。
3. Hosmer-Lemeshow 適合度檢定的 P 值。
4. 使用 10 場進行預測之後的成功場次，7 代表 10 場比賽成功預測 7 場，5 代表 10 場比賽成功預測 5 場，以此類推。
5. 方框內表示該時間點內，預測成功場數最多，即成功率最高，而粗底線則代表預測成功數最少，成功率最低。
6. 表示上述 4 次建模(第 49-58 場、第 97-108 場、第 169-178 場、第 217-226 場)各預測的平均。

從表 4.2.11 到表 4.2.18 建立模型，分別得到前一場之邏輯斯迴歸模型，再由表 4.2.20 得到，建模成功率大約落在 60%-80%左右，跟前面相比有些差距，而在驗證的方面，在前 216 場驗證成功的場次最多，預測成功的場次到達 7 場。

表 4.2.20 主客隊各項成績—前一場的分析整合表

建模場次	建模成功率	P 值	預測成功場數
前 48 場	80.90%	0.175	5
前 96 場	76.60%	0.238	<u>4</u>
前 168 場	70.70%	0.446	<u>4</u>
前 216 場	61.80%	0.742	<u>7</u>

由表 4.2.19 和 4.2.20 得到，先以全季時間點來看，再把各個預測時間點跟預測方式分開來看，在預測第 97-108 場時，預測 2 及預測 5 的預測能力最佳，成功預測的場次來到 9 場；在預測第 169-178 場時，預測 6 的預測能力最差，成功預測的場次只有 3 次。以預測時間點區分，預測第 97-108 場時，平均預測成功最高，達到 7.5 次；預測第 169-178 場時，平均預測成功最低，只有 4.67 次。以預測方式來區分，預測 5 的預測能力最好，平均預測成功達到了 7 次；預測 6 的預測能力則是最差，平均預測成功只有 5.25 次；而全季 4 個時間點各個預測方式整體平均是 6.33 次。

將第 169-178 場、第 217-226 場及下半季第 49-58 場、下半季第 97-106 場進

行比較，不論以預測時間點還是預測方式來看，全季時間點的預測能力都優於下半季。

將前一場與預測能力最佳的預測 5 做比較，在第 49-58 場、第 97-108 場、第 169-178 場的部分，預測 5 優於前一場，在第 217-226 場則是相同。

在全季時間點的部分，平均預測成功次數來到 6.33 次，而前一場平均預測成功次數則是 5 次，前者優於後者。

### 第三節 以主客隊各項成績比建立模型與預測

#### 一、比賽進行 20%(賽季前 48 場)

表 4.3.1 是主客場各項成績比在比賽進行 20%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、被安打、被全壘打、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.1 比賽進行 20%時各項成績比的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-10.134	3.907	6.729	1	0.009	0.000
每打數平均打點	-7.395	2.449	9.116	1	0.003	0.001
二壘安打	-0.108	0.163	0.444	1	0.505	0.897
三壘安打	-0.454	0.386	1.387	1	0.239	0.635
雙殺打	-0.247	0.247	1.003	1	0.316	0.781
犧牲短打	-0.018	0.222	0.006	1	0.936	0.982
犧牲飛球	0.073	0.212	0.117	1	0.732	1.075
盜壘成功	-0.289	0.198	2.140	1	0.143	0.749
盜壘刺殺	0.220	0.234	0.885	1	0.347	1.246

表 4.3.1 比賽進行 20%時各項成績比的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
失誤	0.231	0.171	1.827	1	0.176	1.260
團隊打擊率	-4.602	1.465	9.867	1	0.002	0.010
壘打數	-3.594	1.202	8.937	1	0.003	0.027
上壘率	-5.354	1.593	11.295	1	0.001	0.005
長打率	-7.363	2.208	11.118	1	0.001	0.001
攻擊指數	-4.608	1.425	10.454	1	0.001	0.010
面對打擊數	10.832	3.434	9.947	1	0.002	50589.866
投球數	3.769	2.224	2.872	1	0.090	43.354
好球數	0.216	2.091	0.011	1	0.918	1.241
被安打	4.599	1.430	10.336	1	0.001	99.384
被全壘打	0.918	0.314	8.547	1	0.003	2.503
四壞球	0.339	0.182	3.488	1	0.062	1.404
四死球	-0.263	0.251	1.099	1	0.295	0.769
奪三振	-0.645	0.461	1.958	1	0.162	0.525
三振四壞比	-0.534	0.205	6.799	1	0.009	0.586
暴投	-0.171	0.172	0.980	1	0.322	0.843
投手犯規	10.579	14210.362	0.000	1	0.999	39309.502
防禦率	6.595	2.168	9.252	1	0.002	731.267
每局被上壘率	4.668	1.306	12.770	1	0.000	106.506

由表 4.3.2 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=5.549-4.608\times\text{攻擊指數}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.2 得到，攻擊指數的參數估計值為-4.608，勝算比估計值為 0.01，表示在其它條件固定下，客隊主隊相除後攻擊指數的商每增加一個單位，主隊獲勝的機率就變為原本的 0.01 倍。Hosmer-Lemeshow 配適度的統計量 8.167，P 值為 0.318 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.3.2 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	5.549	1.641	11.431	1	0.001	257.1
攻擊指數	-4.608	1.425	10.454	1	0.001	0.01
配適度	卡方	8.167	df	7	P 值	0.318

## 二、比賽進行 40%(賽季前 96 場)

表 4.3.3 是主客場各項成績比在比賽進行 40%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、盜壘成功、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、被全壘打、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.3 比賽進行 40%時各項成績比的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-7.318	2.199	11.075	1	0.001	0.001
每打數平均打點	-7.866	1.921	16.759	1	0.000	0.000
二壘安打	-0.255	0.120	4.511	1	0.034	0.775
三壘安打	-0.274	0.240	1.304	1	0.254	0.760
雙殺打	-0.095	0.125	0.575	1	0.448	0.910
犧牲短打	-0.169	0.188	0.810	1	0.368	0.844
犧牲飛球	-0.217	0.171	1.613	1	0.204	0.805
盜壘成功	-0.363	0.144	6.362	1	0.012	0.696

表 4.3.3 比賽進行 40%時各項成績比的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
盜壘刺殺	0.168	0.163	1.058	1	0.304	1.183
失誤	0.197	0.129	2.343	1	0.126	1.218
團隊打擊率	-5.602	1.209	21.466	1	0.000	0.004
壘打數	-2.159	0.527	16.795	1	0.000	0.115
上壘率	-6.292	1.304	23.272	1	0.000	0.002
長打率	-5.209	1.076	23.443	1	0.000	0.005
攻擊指數	-4.361	0.960	20.640	1	0.000	0.013
面對打擊數	11.279	2.526	19.932	1	0.000	79112.449
投球數	3.801	1.467	6.709	1	0.010	44.747
好球數	1.358	1.410	0.928	1	0.335	3.890
被安打	4.540	1.009	20.229	1	0.000	93.653
被全壘打	0.778	0.213	13.298	1	0.000	2.176
四壞球	0.251	0.139	3.262	1	0.071	1.285
四死球	0.065	0.142	0.212	1	0.645	1.067
奪三振	-0.666	0.349	3.631	1	0.057	0.514
三振四壞比	-0.401	0.147	7.506	1	0.006	0.669
暴投	0.012	0.126	0.009	1	0.924	1.012
投手犯規	0.225	0.621	0.131	1	0.717	1.252
防禦率	5.530	1.280	18.651	1	0.000	252.148
每局被上壘率	5.345	1.093	23.932	1	0.000	209.535

由表 4.3.4 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{p}{1-p}\right) = 3.326 - 6.843 \times \text{每打數平均打點} + 3.692 \times \text{每局被上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.4 得到，每打數平均打點的參數估計值為-6.843，勝算比估計值為 0.001，表示其它條件固定下，客隊主隊相除後每打數平均打點的商每增加一個單位，主隊獲勝的機率增加為原本的 0.001 倍；每局被上壘率的參數估計值為 3.692，勝算比估計值為 40.137，表示其它條件固定下，客隊主隊相除後每局被上壘率的商每增加一個單位，主隊獲勝的機率增加為原本的 40.137 倍。Hosmer-Lemeshow 配適度的統計量 9.500，P 值為 0.302 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.3.4 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	3.326	2.569	1.676	1	0.195	27.824
每打數平均打點	-6.843	2.128	10.337	1	0.001	0.001
每局被上壘率	3.692	1.577	5.478	1	0.019	40.137
配適度	卡方	9.500	df	8	P 值	0.302

### 三、比賽進行 70%(賽季前 168 場)

表 4.3.5 是主客場各項成績比在比賽進行 70%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、犧牲短打、犧牲飛球、盜壘成功、失誤、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、被全壘打、四壞球、奪三振、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.5 比賽進行 70%時各項成績比的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-5.217	1.466	12.662	1	0.000	0.005
每打數平均打點	-8.057	1.469	30.069	1	0.000	0.000

表 4.3.5 比賽進行 70%時各項成績比的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
二壘安打	-0.253	0.091	7.790	1	0.005	0.777
三壘安打	-0.245	0.190	1.667	1	0.197	0.782
雙殺打	-0.018	0.085	0.046	1	0.830	0.982
犧牲短打	-0.450	0.154	8.492	1	0.004	0.638
犧牲飛球	-0.425	0.150	8.021	1	0.005	0.654
盜壘成功	-0.254	0.096	6.931	1	0.008	0.776
盜壘刺殺	0.041	0.116	0.127	1	0.721	1.042
失誤	0.297	0.103	8.365	1	0.004	1.345
團隊打擊率	-5.002	0.795	39.567	1	0.000	0.007
壘打數	-2.759	0.483	32.623	1	0.000	0.063
上壘率	-6.878	1.049	42.960	1	0.000	0.001
長打率	-5.487	0.842	42.443	1	0.000	0.004
攻擊指數	-4.526	0.744	37.052	1	0.000	0.011
面對打擊數	12.703	2.093	36.856	1	0.000	328880.421
投球數	3.980	1.078	13.629	1	0.000	53.501
好球數	1.740	1.049	2.749	1	0.097	5.697
被安打	4.650	0.743	39.217	1	0.000	104.626
被全壘打	0.638	0.144	19.685	1	0.000	1.893
四壞球	0.457	0.150	9.278	1	0.002	1.580
四死球	0.172	0.116	2.184	1	0.139	1.187
奪三振	-0.813	0.263	9.565	1	0.002	0.443
三振四壞比	-0.563	0.136	17.063	1	0.000	0.570
暴投	0.078	0.100	0.604	1	0.437	1.081

表 4.3.5 比賽進行 70%時各項成績比的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
投手犯規	0.303	0.617	0.241	1	0.624	1.354
防禦率	3.262	0.530	37.900	1	0.000	26.092
每局被上壘率	5.867	0.903	42.244	1	0.000	353.082

由表 4.3.6 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=3.453-7.143\times\text{每打數平均打點}+3.803\times\text{每局被上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.6 得到，每打數平均打點的參數估計值為-7.143，勝算比估計值為 0.001，表示其它條件固定下，客隊主隊相除後每打數平均打點的商每增加一個單位，主隊獲勝的機率變成原本的 0.001 倍；每局被上壘率的參數估計值為 3.803，勝算比估計值為 44.826，表示其它條件固定下，客隊主隊相除後每局被上壘率的商每增加一個單位，主隊獲勝的機率變成原本的 44.826 倍。Hosmer-Lemeshow 配適度的統計量 3.514，P 值為 0.898 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.3.6 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	3.453	1.956	3.116	1	0.078	31.593
每打數平均打點	-7.143	1.633	19.136	1	0.000	0.001
每局被上壘率	3.803	1.307	8.465	1	0.004	44.826
配適度	卡方	3.514	df	8	P 值	0.898

#### 四、比賽進行 90%(賽季前 216 場)

表 4.3.7 是主客場各項成績比在比賽進行 90%時選入 28 個解釋變數的單變

數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、犧牲短打、犧牲飛球、盜壘成功、失誤、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、好球數、被安打、被全壘打、四壞球、四死球、奪三振、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.7 比賽進行 90%時各項成績比的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-5.046	1.236	16.673	1	0.000	0.006
每打數平均打點	-7.297	1.161	39.491	1	0.000	0.001
二壘安打	-0.276	0.081	11.643	1	0.001	0.759
三壘安打	-0.242	0.162	2.234	1	0.135	0.785
雙殺打	-0.002	0.078	0.001	1	0.978	0.998
犧牲短打	-0.415	0.132	9.815	1	0.002	0.660
犧牲飛球	-0.531	0.143	13.843	1	0.000	0.588
盜壘成功	-0.219	0.083	6.955	1	0.008	0.803
盜壘刺殺	-0.006	0.099	0.004	1	0.950	0.994
失誤	0.179	0.082	4.727	1	0.030	1.196
團隊打擊率	-5.273	0.729	52.282	1	0.000	0.005
壘打數	-2.904	0.444	42.805	1	0.000	0.055
上壘率	-7.064	0.980	51.942	1	0.000	0.001
長打率	-5.287	0.725	53.229	1	0.000	0.005
攻擊指數	-4.577	0.666	47.279	1	0.000	0.010
面對打擊數	11.085	1.671	44.011	1	0.000	65164.872
投球數	4.018	0.938	18.339	1	0.000	55.586
好球數	2.048	0.931	4.838	1	0.028	7.754
被安打	4.562	0.644	50.134	1	0.000	95.756

表 4.3.7 比賽進行 90%時各項成績比的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
被全壘打	0.537	0.116	21.536	1	0.000	1.712
四壞球	0.349	0.116	9.098	1	0.003	1.417
四死球	0.228	0.107	4.545	1	0.033	1.256
奪三振	-0.819	0.250	10.695	1	0.001	0.441
三振四壞比	-0.452	0.107	17.735	1	0.000	0.637
暴投	0.148	0.093	2.522	1	0.112	1.160
投手犯規	0.088	0.492	0.032	1	0.857	1.092
防禦率	3.464	0.504	47.254	1	0.000	31.930
每局被上壘率	5.580	0.770	52.487	1	0.000	265.110

由表 4.3.8 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=13.804-6.480\times\text{每打數平均打點}-16.660\times\text{面對打擊數}+8.726\times\text{每局被上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.8 得到，每打數平均打點的參數估計值為-6.480，勝算比估計值為 0.002，表示在其它條件固定下，客隊主隊相除後每打數平均打點的商每增加一個單位，主隊獲勝的機率變成原本的 0.002 倍；面對打擊數的參數估計值為-16.660，勝算比估計值為 0.000，表示在其它條件固定下，客隊主隊相除後面對打擊數的商每增加一個單位，主隊獲勝的機率將低於原本的 0.001 倍；每局被上壘率的參數估計值為 8.726，勝算比估計值為 6161.613，表示在其它條件固定下，客隊主隊相除後每局被上壘率的商每增加一個單位，主隊獲勝的機率變成原本的 6161.613 倍。Hosmer-Lemeshow 配適度的統計量 0.583，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.3.8 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	13.804	4.966	7.727	1	0.005	988635.791
每打數平均打點	-6.480	1.438	20.303	1	0.000	0.002
面對打擊數	-16.660	6.750	6.092	1	0.014	0.000
每局被上壘率	8.726	2.372	13.532	1	0.000	6161.613
配適度	卡方	0.583	df	8	P 值	1.000

### 五、比賽進行下半季 20%(賽季第 121-168 場)

由表 4.3.9 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=10.710-9.694\times\text{每打數平均打點}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.9 得到，每打數平均打點的參數估計值為-9.694，勝算比估計值為 0.000，表示在其它條件固定下，客隊主隊相除後每打數平均打點的商每增加一個單位，主隊獲勝的機率將低於原本的 0.001 倍。Hosmer-Lemeshow 配適度的統計量 1.754，P 值為 0.988 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.3.9 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	10.710	4.813	4.953	1	0.026	44816.425
每打數平均打點	-9.694	4.172	5.398	1	0.020	0.000
配適度	卡方	1.754	df	8	P 值	0.988

### 六、比賽進行下半季 40%(賽季第 121-216 場)

由表 4.3.10 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=9.655-6.921\times\text{每打數平均打點}-13.375\times\text{投球數}+10.523\times\text{每局被上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.3.10 得到，每打數平均打點的參數估計值為-6.921，勝算比估計值為 0.001，表示在其它條件固定下，客隊主隊相除後每打數平均打點的商每增加一個單位，主隊獲勝的機率變成原本的 0.001 倍；投球數的參數估計值為-13.375，勝算比估計值為 0.000，表示在其它條件固定下，客隊主隊相除後投球數的商每增加一個單位，主隊獲勝的機率變成將低於的 0.001 倍；每局被上壘率的參數估計值為 10.523，勝算比估計值為 37168.455，表示在其它條件固定下，客隊主隊相除後每局被上壘率的商每增加一個單位，主隊獲勝的機率變成原本的 37168.455 倍。Hosmer-Lemeshow 配適度的統計量 2.560，P 值為 0.988 大於 0.959，因此表示不違反模型配適資料的假設。

表 4.3.10 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	9.655	5.258	3.371	1	0.066	15591.808
每打數平均打點	-6.921	2.511	7.597	1	0.006	0.001
投球數	-13.375	6.213	4.635	1	0.031	0.000
每局被上壘率	10.523	3.924	7.193	1	0.007	37168.455
配適度	卡方	2.560	df	8	P 值	0.959

## 七、前一場模型建立

表 4.3.11 是主客場各項成績比在比賽進行 20%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中攻擊指數的 P 值均小於 0.05，達到顯著水準。

表 4.3.11 比賽進行 20%時各項成績比前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.099	2.677	0.001	1	0.970	0.906

表 4.3.11 比賽進行 20%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
每打數平均打點	-0.400	0.311	1.653	1	0.199	0.670
二壘安打	-0.115	0.177	0.421	1	0.516	0.891
三壘安打	0.118	0.328	0.128	1	0.720	1.125
雙殺打	0.049	0.226	0.048	1	0.827	1.051
犧牲短打	-0.163	0.226	0.520	1	0.471	0.850
犧牲飛球	-0.144	0.216	0.443	1	0.506	0.866
盜壘成功	-0.013	0.207	0.004	1	0.949	0.987
盜壘刺殺	0.098	0.235	0.173	1	0.677	1.103
失誤	0.103	0.152	0.458	1	0.499	1.108
團隊打擊率	-0.613	0.604	1.027	1	0.311	0.542
壘打數	0.048	0.068	0.502	1	0.479	1.049
上壘率	-1.083	0.752	2.073	1	0.150	0.339
長打率	-0.318	0.486	0.429	1	0.513	0.727
攻擊指數	-0.913	0.458	3.973	1	0.046	0.401
面對打擊數	3.596	2.340	2.361	1	0.124	36.439
投球數	0.956	1.733	0.305	1	0.581	2.602
好球數	-1.168	1.804	0.419	1	0.517	0.311
被安打	0.080	0.740	0.012	1	0.914	1.083
被全壘打	0.280	0.186	2.269	1	0.132	1.323
四壞球	0.565	0.306	3.415	1	0.065	1.759
四死球	0.094	0.207	0.205	1	0.651	1.098
奪三振	-0.995	0.508	3.834	1	0.050	0.370
三振四壞比	-0.234	0.122	3.704	1	0.054	0.791

表 4.3.11 比賽進行 20%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
暴投	0.038	0.201	0.035	1	0.851	1.038
投手犯規	-19.860	20085.137	0.000	1	0.999	0.000
防禦率	0.663	0.417	2.534	1	0.111	1.941
每局被上壘率	1.050	0.646	2.643	1	0.104	2.857

由表 4.3.12 得到，前 48 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.763 + 0.565 \times \text{四壞球}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.3.12 使用前一場建立模型-前 48 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-0.763	0.444	2.957	1	0.086	0.466
四壞球	0.565	0.306	3.415	1	0.065	1.759
配適度	卡方	8.589	df	6	P 值	0.198

表 4.3.13 是主客場各項成績比在比賽進行 40%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中打數、二壘安打、盜壘成功、團隊打擊率、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、被全壘打、四壞球、奪三振、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.13 比賽進行 40%時各項成績比前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-5.483	2.019	7.373	1	0.007	0.004
每打數平均打點	-0.316	0.169	3.496	1	0.062	0.729
二壘安打	-0.429	0.167	6.616	1	0.010	0.651
三壘安打	-0.150	0.228	0.435	1	0.509	0.861

表 4.3.13 比賽進行 40%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
雙殺打	-0.021	0.143	0.022	1	0.881	0.979
犧牲短打	-0.204	0.191	1.134	1	0.287	0.816
犧牲飛球	-0.138	0.167	0.680	1	0.410	0.871
盜壘成功	-0.321	0.146	4.823	1	0.028	0.726
盜壘刺殺	0.122	0.164	0.554	1	0.457	1.129
失誤	0.087	0.121	0.513	1	0.474	1.091
團隊打擊率	-1.420	0.521	7.422	1	0.006	0.242
壘打數	-0.029	0.038	0.595	1	0.440	0.971
上壘率	-2.050	0.625	10.760	1	0.001	0.129
長打率	-0.876	0.361	5.881	1	0.015	0.416
攻擊指數	-0.618	0.240	6.659	1	0.010	0.539
面對打擊數	7.320	1.980	13.671	1	0.000	1510.688
投球數	3.229	1.355	5.682	1	0.017	25.264
好球數	1.975	1.325	2.222	1	0.136	7.208
被安打	1.087	0.464	5.492	1	0.019	2.965
被全壘打	0.296	0.136	4.724	1	0.030	1.345
四壞球	0.583	0.234	6.234	1	0.013	1.791
四死球	0.200	0.149	1.809	1	0.179	1.221
奪三振	-0.944	0.377	6.249	1	0.012	0.389
三振四壞比	-0.185	0.089	4.331	1	0.037	0.831
暴投	0.065	0.144	0.201	1	0.654	1.067
投手犯規	-20.138	18026.537	0.000	1	0.999	0.000
防禦率	0.504	0.214	5.553	1	0.018	1.656

表 4.3.13 比賽進行 40%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
每局被上壘率	1.459	0.476	9.405	1	0.002	4.301

由表 4.3.14 得到，前 96 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -7.166 - 0.827 \times \text{二壘安打} + 0.239 \times \text{壘打數} + 12.722 \times \text{面對打擊數} - 4.343 \times \text{好球數}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.3.14 使用前一場建立模型-前 96 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-7.166	2.469	8.421	1	0.004	0.001
二壘安打	-0.827	0.270	9.408	1	0.002	0.437
壘打數	0.239	0.075	10.124	1	0.001	1.270
面對打擊數	12.722	3.351	14.409	1	0.000	334919.324
好球數	-4.343	2.090	4.318	1	0.038	0.013
配適度	卡方	5.938	df	8	P 值	0.654

表 4.3.15 是主客場各項成績比在比賽進行 70%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中上壘率、面對打擊數、投球數、被安打、四壞球、奪三振、三振四壞比、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.3.15 比賽進行 70%時各項成績比前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-1.678	1.217	1.902	1	0.168	0.187
每打數平均打點	-0.192	0.107	3.207	1	0.073	0.826
二壘安打	-0.138	0.081	2.881	1	0.090	0.871
三壘安打	-0.103	0.175	0.343	1	0.558	0.902

表 4.3.15 比賽進行 70%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
雙殺打	-0.109	0.099	1.227	1	0.268	0.897
犧牲短打	-0.220	0.139	2.516	1	0.113	0.802
犧牲飛球	-0.103	0.129	0.627	1	0.428	0.903
盜壘成功	-0.163	0.092	3.143	1	0.076	0.849
盜壘刺殺	0.116	0.117	0.982	1	0.322	1.123
失誤	0.137	0.095	2.097	1	0.148	1.147
團隊打擊率	-0.392	0.266	2.168	1	0.141	0.676
壘打數	-0.031	0.037	0.668	1	0.414	0.970
上壘率	-0.993	0.392	6.417	1	0.011	0.370
長打率	-0.213	0.174	1.498	1	0.221	0.808
攻擊指數	-0.208	0.116	3.222	1	0.073	0.812
面對打擊數	3.286	1.166	7.938	1	0.005	26.737
投球數	2.494	0.935	7.121	1	0.008	12.109
好球數	1.672	0.988	2.866	1	0.090	5.323
被安打	0.688	0.314	4.799	1	0.028	1.989
被全壘打	0.088	0.088	0.988	1	0.320	1.092
四壞球	0.357	0.136	6.914	1	0.009	1.429
四死球	0.050	0.111	0.205	1	0.650	1.052
奪三振	-0.725	0.247	8.587	1	0.003	0.484
三振四壞比	-0.188	0.070	7.184	1	0.007	0.828
暴投	0.124	0.114	1.192	1	0.275	1.132
投手犯規	-20.218	17918.712	0.000	1	0.999	0.000
防禦率	0.112	0.073	2.389	1	0.122	1.119

表 4.3.15 比賽進行 70%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
每局被上壘率	0.815	0.294	7.660	1	0.006	2.259

由表 4.3.16 得到，前 168 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -1.863 + 2.963 \times \text{投球數} - 0.841 \times \text{奪三振}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.3.16 使用前一場建立模型-前 168 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-1.863	0.976	3.638	1	0.056	0.155
投球數	2.963	0.977	9.200	1	0.002	19.349
奪三振	-0.841	0.261	10.397	1	0.001	0.431
配適度	卡方	4.285	df	8	P 值	0.831

表 4.3.17 是主客場各項成績比在比賽進行 90%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中失誤、四壞球、奪三振、三振四壞比的 P 值均小於 0.05，達到顯著水準。

表 4.3.17 比賽進行 90%時各項成績比前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.871	1.034	0.709	1	0.400	0.419
每打數平均打點	-0.076	0.089	0.745	1	0.388	0.926
二壘安打	-0.106	0.071	2.242	1	0.134	0.900
三壘安打	-0.047	0.130	0.128	1	0.720	0.954
雙殺打	-0.045	0.086	0.277	1	0.599	0.956
犧牲短打	-0.162	0.126	1.652	1	0.199	0.850

表 4.3.17 比賽進行 90%時各項成績比前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
犧牲飛球	-0.101	0.118	0.739	1	0.390	0.904
盜壘成功	-0.160	0.084	3.637	1	0.057	0.852
盜壘刺殺	0.079	0.102	0.601	1	0.438	1.082
失誤	0.172	0.085	4.078	1	0.043	1.188
團隊打擊率	-0.123	0.223	0.305	1	0.581	0.884
壘打數	-0.014	0.033	0.181	1	0.671	0.986
上壘率	-0.340	0.317	1.153	1	0.283	0.712
長打率	-0.083	0.144	0.332	1	0.565	0.920
攻擊指數	-0.080	0.082	0.973	1	0.324	0.923
面對打擊數	1.922	0.966	3.954	1	0.047	6.832
投球數	1.464	0.780	3.525	1	0.060	4.322
好球數	1.027	0.854	1.446	1	0.229	2.794
被安打	0.302	0.221	1.864	1	0.172	1.353
被全壘打	0.013	0.075	0.029	1	0.865	1.013
四壞球	0.233	0.103	5.080	1	0.024	1.263
四死球	0.004	0.097	0.001	1	0.971	1.004
奪三振	-0.615	0.228	7.269	1	0.007	0.541
三振四壞比	-0.107	0.052	4.288	1	0.038	0.898
暴投	0.142	0.103	1.889	1	0.169	1.152
投手犯規	-0.011	0.536	0.000	1	0.984	0.989
防禦率	0.004	0.050	0.005	1	0.944	1.004
每局被上壘率	0.331	0.221	2.246	1	0.134	1.392

由表 4.3.18 得到，前 216 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.431+0.215\times\text{四壞球}-0.588\times\text{奪三振}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.3.18 使用前一場建立模型-前 216 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.431	0.335	1.656	1	0.198	1.538
四壞球	0.215	0.103	4.351	1	0.037	1.240
奪三振	-0.588	0.229	6.610	1	0.010	0.555
配適度	卡方	6.787	df	8	P 值	0.560

## 八、驗證分析與比較

由表 4.3.19 得到，建模時間前 48 場時，預測 2 的預測能力最好，成功場次達到 9 場；建模時間前 96 場時，預測 2 的預測能力最好，成功場次達到 9 場；建模時間前 168 場時，預測 4 的預測能力最好，成功場次達到 7 場；建模時間前 216 場時，預測 1、預測 4、預測 5 的預測能力並列最好，成功場次達到 7 場；建模時間下半季前 48 場時，預測 4 預測能力最好，成功場次達到 6 場；建模時間下半季前 96 場時，預測 1 和預測 4 的預測能力並列最好，成功場次達到 7 場。

表 4.3.19 主客隊各項成績比的分析整合表

建模場次 (預測場次)	建模成功率 (P 值)	預測方式	預測 成功 場數	預測方式	預測 成功 次數	預測方式 平均
前 48 場 (第 49-58 場)	89.4% (0.318)	預測 1	7	預測 4	8	7
		預測 2	9	預測 5	3	
		預測 3	8	預測 6	7	

表 4.3.19 主客隊各項成績比的分析整合表(續)

建模場次 (預測場次)	建模成功率 (P 值)	預測方式	預測 成功 場數	預測方式	預測 成功 次數	預測方式 平均
前 96 場 (第 97-108 場，不含 100、101)	92.6% (0.302)	預測 1	<u>5</u>	預測 4	6	7.33
		預測 2	<u>9</u>	預測 5	8	
		預測 3	8	預測 6	8	
前 168 場 (第 169-178 場)	93.90% (0.898)	預測 1	5	預測 4	<u>7</u>	4.83
		預測 2	5	預測 5	5	
		預測 3	4	預測 6	<u>3</u>	
前 216 場 (第 217-226 場)	94.30% (1.000)	預測 1	<u>7</u>	預測 4	<u>7</u>	6.17
		預測 2	6	預測 5	<u>7</u>	
		預測 3	<u>5</u>	預測 6	<u>5</u>	
四個時間點平均		預測 1	6	預測 4	7	6.33
		預測 2	<u>7.25</u>	預測 5	<u>5.75</u>	
		預測 3	6.25	預測 6	<u>5.75</u>	
下半季前 48 場	93.80% (0.988)	預測 1	5	預測 4	<u>6</u>	4.83
		預測 2	5	預測 5	5	
		預測 3	5	預測 6	<u>3</u>	
下半季前 96 場	93.80% (0.959)	預測 1	<u>7</u>	預測 4	<u>7</u>	6
		預測 2	<u>5</u>	預測 5	6	
		預測 3	<u>5</u>	預測 6	6	

從表 4.3.11 到表 4.3.18 建立模型，分別得到前一場之邏輯斯迴歸模型，再由表 4.3.20 得到，建模成功率大約落在 60%-70%左右，跟前面相比有些差距，而

在驗證的方面，在前 96 場與前 216 場驗證成功的場次最多，預測成功的場次到達 7 場。

表 4.3.20 主客隊各項成績比-前一場的分析整合表

建模場次	建模成功率	P 值	預測成功次數
前 48 場	66%	0.198	6
前 96 場	72.30%	0.654	7
前 168 場	65.90%	0.831	5
前 216 場	60.80%	0.560	7

由表 4.3.19 和表 4.3.20 得到，先以全季時間點來看，再把各個預測時間點跟預測方式分開來看，在預測第 49-58 場和第 97-108 場時，這兩個時間點的預測 2 預測能力最佳，成功預測的場次來到 9 場；在預測第 49-58 場的預測 5 以及第 169-178 場的預測 6 的預測能力最差，成功預測的場次只有 3 次。以預測時間點區分，預測第 97-108 場時，平均預測成功最高，達到 7.33 次；預測第 169-178 場時，平均預測成功最低，只有 4.83 次。以預測方式來區分，預測 2 的預測能力最好，平均預測成功達到了 7.25 次；預測 5 和預測 6 的預測能力則是最差，平均預測成功有 5.75 次。

將第 169-178 場、第 217-226 場及下半季第 49-58 場、下半季第 97-106 場進行比較，第 169-178 場與下半季第 49-58 場的平均預測成功率是相同的，第 217-226 場的預測成功率則是略優於下半季第 97-106 場。

將前一場與預測能力最佳的預測 2 做比較，在第 49-58 場、第 97-108 場的部分，預測 2 的平均略優於前一場，在第 169-178 場則是相同，而第 217-226 場，則是前一場優於預測 2。

在全季時間點的部分，平均預測成功次數來到 6.33 次，而前一場平均預測成功次數則是 6.25 次，前者略優於後者。

## 第四節 以主客隊各項成績差建立模型與預測

### 一、 比賽進行 20%(賽季前 48 場)

表 4.4.1 是主客場各項成績差在比賽進行 20%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、被安打、被全壘打、四壞球、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.1 比賽進行 20%時各項成績差的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.279	0.116	5.765	1	0.016	0.757
每打數平均打點	-41.529	13.596	9.330	1	0.002	0.000
二壘安打	-0.126	0.198	0.402	1	0.526	0.882
三壘安打	-0.746	0.648	1.325	1	0.250	0.474
雙殺打	-0.224	0.287	0.607	1	0.436	0.799
犧牲短打	-0.218	0.352	0.383	1	0.536	0.804
犧牲飛球	0.021	0.376	0.003	1	0.955	1.021
盜壘成功	-0.501	0.267	3.523	1	0.061	0.606
盜壘刺殺	0.336	0.394	0.729	1	0.393	1.400
失誤	0.458	0.248	3.414	1	0.065	1.581
團隊打擊率	-18.234	5.548	10.801	1	0.001	0.000
壘打數	-0.310	0.095	10.595	1	0.001	0.734
上壘率	-16.984	4.746	12.803	1	0.000	0.000
長打率	-17.180	5.320	10.427	1	0.001	0.000
攻擊指數	-7.070	2.068	11.691	1	0.001	0.001
面對打擊數	0.258	0.084	9.355	1	0.002	1.294

表 4.4.1 比賽進行 20%時各項成績差的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
投球數	0.024	0.014	2.853	1	0.091	1.024
好球數	-0.004	0.022	0.036	1	0.849	0.996
被安打	0.403	0.127	10.104	1	0.001	1.497
被全壘打	1.086	0.341	10.132	1	0.001	2.963
四壞球	0.359	0.132	7.399	1	0.007	1.431
四死球	-0.185	0.309	0.356	1	0.551	0.831
奪三振	-0.093	0.083	1.233	1	0.267	0.912
三振四壞比	-0.154	0.066	5.495	1	0.019	0.857
暴投	-0.112	0.256	0.192	1	0.661	0.894
投手犯規	21.158	28420.722	0.000	1	0.999	$1.545 \times 10^9$
防禦率	1.448	0.466	9.656	1	0.002	4.256
每局被上壘率	3.340	0.925	13.032	1	0.000	28.206

由表 4.4.2 得到，前 48 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.58+1.448 \times \text{防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.2 得到，防禦率的參數估計值為 1.448，勝算比估計值為 4.256，表示在其它條件固定下，客隊主隊相減後防禦率的差每增加一個單位，主隊獲勝的機率就變為原本的 4.256 倍。Hosmer-Lemeshow 配適度的統計量 5.774，P 值為 0.566 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.2 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.58	0.562	1.066	1	0.302	1.787

表 4.4.2 比賽結果之參數估計與配適度(續)

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
防禦率	1.448	0.466	9.656	1	0.002	4.256
配適度	卡方	5.774	df	7	P 值	0.566

## 二、 比賽進行 40%(賽季前 96 場)

表 4.4.3 是主客場各項成績差在比賽進行 40%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、盜壘成功、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、被全壘打、四壞球、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.3 比賽進行 40%時各項成績差的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.217	0.067	10.512	1	0.001	0.805
每打數平均打點	-46.799	10.786	18.828	1	0.000	0.000
二壘安打	-0.261	0.119	4.796	1	0.029	0.770
三壘安打	-0.358	0.404	0.785	1	0.375	0.699
雙殺打	-0.100	0.169	0.348	1	0.555	0.905
犧牲短打	-0.437	0.294	2.219	1	0.136	0.646
犧牲飛球	-0.444	0.290	2.350	1	0.125	0.642
盜壘成功	-0.414	0.173	5.736	1	0.017	0.661
盜壘刺殺	0.255	0.263	0.939	1	0.333	1.290
失誤	0.378	0.179	4.442	1	0.035	1.459
團隊打擊率	-21.125	4.395	23.103	1	0.000	0.000
壘打數	-0.272	0.059	21.405	1	0.000	0.762

表 4.4.3 比賽進行 40%時各項成績差的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
上壘率	-20.002	3.921	26.020	1	0.000	0.000
長打率	-12.897	2.719	22.502	1	0.000	0.000
攻擊指數	-7.116	1.534	21.509	1	0.000	0.001
面對打擊數	0.279	0.063	19.470	1	0.000	1.322
投球數	0.025	0.009	7.016	1	0.008	1.025
好球數	0.012	0.015	0.583	1	0.445	1.012
被安打	0.443	0.096	21.459	1	0.000	1.558
被全壘打	0.972	0.228	18.170	1	0.000	2.644
四壞球	0.330	0.096	11.755	1	0.001	1.392
四死球	0.173	0.202	0.733	1	0.392	1.189
奪三振	-0.103	0.056	3.460	1	0.063	0.902
三振四壞比	-0.113	0.045	6.138	1	0.013	0.894
暴投	0.203	0.183	1.233	1	0.267	1.225
投手犯規	0.450	1.243	0.131	1	0.717	1.569
防禦率	1.470	0.335	19.260	1	0.000	4.351
每局被上壘率	3.926	0.796	24.302	1	0.000	50.705

由表 4.4.4 得到，前 96 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.625-2.034\times\text{雙殺打}+2.345\times\text{失誤}+3.154\times\text{防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.4 得到，雙殺打的參數估計值為-2.034，勝算比估計值為 0.131，表示其它條件固定下，客隊主隊相減後雙殺打的差每增加一個單位，主隊獲勝的機率變成原本的 0.131 倍；失誤的參數估計值為 2.345，勝算比估計值為 10.437，

表示其它條件固定下，客隊主隊相減後失誤的差每增加一個單位，主隊獲勝的機率變成原本的 10.437 倍；防禦率的參數估計值為 3.154，勝算估計值為 23.430，表示其它條件固定下，客隊主隊相減後防禦率的差每增加一個單位，主隊獲勝的機率變成原本的 23.430 倍。Hosmer-Lemeshow 配適度的統計量 0.687，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.4 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.625	0.781	0.642	1	0.423	1.869
雙殺打	-2.034	0.840	5.866	1	0.015	0.131
失誤	2.345	0.927	6.399	1	0.011	10.437
防禦率	3.154	1.118	7.966	1	0.005	23.430
配適度	卡方	0.687	df	8	P 值	1.000

### 三、比賽進行 70%(賽季前 168 場)

表 4.4.5 是主客場各項成績差在比賽進行 70%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、犧牲短打、犧牲飛球、盜壘成功、失誤、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、被全壘打、四壞球、奪三振、三振四壞比、暴投、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.5 比賽進行 70%時各項成績差的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.164	0.045	13.306	1	0.000	0.849
每打數平均打點	-50.270	9.093	30.564	1	0.000	0.000
二壘安打	-0.303	0.091	10.986	1	0.001	0.739
三壘安打	-0.351	0.288	1.486	1	0.223	0.704

表 4.4.5 比賽進行 70%時各項成績差的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
雙殺打	-0.014	0.119	0.013	1	0.908	0.986
犧牲短打	-0.731	0.217	11.324	1	0.001	0.482
犧牲飛球	-0.784	0.237	10.911	1	0.001	0.456
盜壘成功	-0.260	0.112	5.336	1	0.021	0.771
盜壘刺殺	0.016	0.188	0.008	1	0.930	1.017
失誤	0.496	0.136	13.263	1	0.000	1.643
團隊打擊率	-20.736	3.137	43.700	1	0.000	0.000
壘打數	-0.276	0.045	38.167	1	0.000	0.758
上壘率	-21.711	3.197	46.128	1	0.000	0.000
攻擊指數	-6.958	1.138	37.413	1	0.000	0.001
面對打擊數	0.306	0.051	36.005	1	0.000	1.358
投球數	0.026	0.007	14.109	1	0.000	1.026
好球數	0.017	0.011	2.344	1	0.126	1.017
被安打	0.428	0.068	39.956	1	0.000	1.534
被全壘打	0.775	0.150	26.563	1	0.000	2.171
四壞球	0.371	0.073	25.518	1	0.000	1.449
四死球	0.271	0.158	2.947	1	0.086	1.312
奪三振	-0.173	0.046	14.371	1	0.000	0.841
三振四壞比	-0.184	0.048	14.523	1	0.000	0.832
暴投	0.295	0.143	4.238	1	0.040	1.344
投手犯規	0.606	1.235	0.241	1	0.624	1.833
防禦率	1.044	0.173	36.634	1	0.000	2.842
每局被上壘率	4.033	0.615	42.988	1	0.000	56.422

由表 4.4.6 得到，前 168 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.171 + 3.595 \times \text{盜壘刺殺} + 3.384 \times \text{失誤} - 0.219 \times \text{好球數} - 0.516 \times \text{三振四壞比} \\ + 3.173 \times \text{防禦率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.6 得到，盜壘刺殺的參數估計值為 3.595，勝算比估計值為 36.432，表示其它條件固定下，客隊主隊相減後盜壘刺殺的差每增加一次，主隊獲勝的機率變成原本的 36.432 倍；失誤的參數估計值為 3.384，勝算比估計值為 29.499，表示其它條件固定下，客隊主隊相減後失誤的差每增加一次，主隊獲勝的機率變成原本的 29.499 倍；好球數的參數估計值為 -0.219，勝算比估計值為 0.803，表示其它條件固定下，客隊主隊相減後好球數的差每增加一次，主隊獲勝的機率變成原本的 0.803 倍；三振四壞比的參數估計值為 -0.516，勝算比估計值為 0.597，表示其它條件固定下，客隊主隊相減後三振四壞比的差每增加一個單位，主隊獲勝的機率變成原本的 0.597 倍；防禦率的參數估計值為 3.173，勝算比估計值為 23.870，表示其它條件固定下，客隊主隊相減後防禦率的差每增加一個單位，主隊獲勝的機率為原本的 23.870。Hosmer-Lemeshow 配適度的統計量 0.221，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.6 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-0.171	0.616	0.077	1	0.781	0.843
盜壘刺殺	3.595	1.327	7.341	1	0.007	36.432
失誤	3.384	1.120	9.135	1	0.003	29.499
好球數	-0.219	0.085	6.616	1	0.010	0.803
三振四壞比	-0.516	0.180	8.180	1	0.004	0.597
防禦率	3.173	0.866	13.417	1	0.000	23.870
配適度	卡方	0.221	df	8	P 值	1.000

#### 四、 比賽進行 90%(賽季前 216 場)

表 4.4.7 是主客場各項成績差在比賽進行 90%時選入 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、犧牲短打、犧牲飛球、盜壘成功、失誤、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、好球數、被安打、被全壘打、四壞球、四死球、奪三振、三振四壞比、暴投、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.7 比賽進行 90%時各項成績差的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.158	0.038	17.729	1	0.000	0.854
每打數平均打點	-52.243	8.546	37.371	1	0.000	0.000
二壘安打	-0.341	0.084	16.640	1	0.000	0.711
三壘安打	-0.374	0.235	2.531	1	0.112	0.688
雙殺打	-0.017	0.109	0.024	1	0.877	0.983
犧牲短打	-0.662	0.186	12.674	1	0.000	0.516
犧牲飛球	-0.842	0.211	15.883	1	0.000	0.431
盜壘成功	-0.251	0.100	6.363	1	0.012	0.778
盜壘刺殺	-0.067	0.159	0.176	1	0.675	0.935
失誤	0.316	0.110	8.313	1	0.004	1.372
團隊打擊率	-21.796	2.929	55.364	1	0.000	0.000
壘打數	-0.277	0.040	48.420	1	0.000	0.758
上壘率	-21.984	2.957	55.268	1	0.000	0.000
長打率	-12.987	1.767	54.002	1	0.000	0.000
攻擊指數	-6.943	1.007	47.565	1	0.000	0.001
面對打擊數	0.267	0.041	42.948	1	0.000	1.306
投球數	0.026	0.006	18.861	1	0.000	1.027

表 4.4.7 比賽進行 90%時各項成績差的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
好球數	0.020	0.010	4.451	1	0.035	1.020
被安打	0.436	0.061	51.201	1	0.000	1.547
被全壘打	0.671	0.124	29.251	1	0.000	1.957
四壞球	0.289	0.058	24.662	1	0.000	1.335
四死球	0.315	0.143	4.830	1	0.028	1.370
奪三振	-0.156	0.040	14.865	1	0.000	0.856
三振四壞比	-0.133	0.035	14.637	1	0.000	0.876
暴投	0.396	0.135	8.556	1	0.003	1.485
投手犯規	0.363	0.917	0.157	1	0.692	1.438
防禦率	1.055	0.156	45.633	1	0.000	2.873
每局被上壘率	3.875	0.534	52.747	1	0.000	48.170

由表 4.4.8 得到，前 216 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -2.434 + 2.475 \times \text{盜壘刺殺} + 2.472 \times \text{失誤} - 1.174 \times \text{面對打擊數} + 1.84 \times \text{防禦率} \\ + 11.319 \times \text{每局被上壘率}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.8 得到，盜壘刺殺的參數估計值為 2.475，勝算比估計值為 11.883，表示在其它條件固定下，客隊主隊相減後盜壘刺殺的差每增加一個單位，主隊獲勝的機率變成原本的 11.883 倍；失誤的參數估計值為 2.472，勝算比估計值為 11.845，客隊主隊相減後失誤的差每增加一個單位，主隊獲勝的機率變成原本的 11.845 倍；面對打擊數的參數估計值為 -1.174，勝算比估計值為 0.309，表示在其它條件固定下，客隊主隊相減後面對打擊數的差每增加一個單位，主隊獲勝的機率變成原本的 0.309 倍；防禦率的參數估計值為 1.840，勝算比估計值為 6.296，

表示在其它條件固定下，客隊主隊相減後防禦率的差每增加一個單位，主隊獲勝的機率變成原本的 6.296 倍；每局被上壘率的參數估計值為 11.319，勝算比估計值為 82400.071，表示在其它條件固定下，客隊主隊相減後每局被上壘率的差每增加一個單位，主隊獲勝的機率變成原本的 82400.071 倍。Hosmer-Lemeshow 配適度的統計量 8.985，P 值為 0.344 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.8 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-2.434	0.907	7.208	1	0.007	0.088
盜壘刺殺	2.475	0.883	7.849	1	0.005	11.883
失誤	2.472	0.669	13.664	1	0.000	11.845
面對打擊數	-1.174	0.366	10.295	1	0.001	0.309
防禦率	1.840	0.515	12.766	1	0.000	6.296
每局被上壘率	11.319	3.304	11.738	1	0.001	82400.071
配適度	卡方	8.985	df	8	P 值	0.344

## 五、比賽進行下半季 20%(賽季第 121-168 場)

由表 4.4.9 得到，下半季前 48 場的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.906-73.056\times\text{每打數平均打點}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.9 得到，每打數平均打點的參數估計值為-73.056，勝算比估計值為 0.000，表示在其它條件固定下，客隊主隊相減後每打數平均打點的差每增加一個單位，主隊獲勝的機率將低於原本的 0.001 倍。Hosmer-Lemeshow 配適度的統計量 0.052，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.9 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.906	1.225	0.547	1	0.459	2.475
每打數平均打點	-73.056	43.336	2.842	1	0.092	0.000
配適度	卡方	0.052	df	7	P 值	1.000

## 六、 比賽進行下半季 40%(賽季第 121-216 場)

由表 4.4.10 得到，此分析的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.142 - 69.447 \times \text{每打數平均打點} - 0.497 \times \text{奪三振}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

由表 4.4.10 得到，每打數平均打點的參數估計值為-69.447，勝算比估計值為 0.000，表示在其它條件固定下，客隊主隊相減後每打數平均打點的差每增加一個單位，主隊獲勝的機率將低於原本的 0.001 倍；奪三振的參數估計值為-0.497，勝算比估計值為 0.608，表示在其它條件固定下，客隊主隊相減後奪三振的差每增加一個單位，主隊獲勝的機率變成原本的 0.608 倍。Hosmer-Lemeshow 配適度的統計量 0.372，P 值為 1.000 大於 0.05，因此表示不違反模型配適資料的假設。

表 4.4.10 比賽結果之參數估計與配適度

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	-0.142	0.538	0.069	1	0.792	0.868
每打數平均打點	-69.447	22.160	9.821	1	0.002	0.000
奪三振	-0.497	0.228	4.744	1	0.029	0.608
配適度	卡方	0.372	df	8	P 值	1.000

## 七、 前一場模型建立

表 4.4.11 是主客場各項成績差在比賽進行 20%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中四壞球、奪三振、三振四壞比的 P 值均小於 0.05，達到顯著水準。

表 4.4.11 比賽進行 20%時各項成績差前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	0.022	0.081	0.075	1	0.784	1.023
每打數平均打點	-4.718	3.092	2.329	1	0.127	0.009
二壘安打	-0.090	0.197	0.211	1	0.646	0.914
三壘安打	0.494	0.588	0.708	1	0.400	1.640
雙殺打	0.197	0.287	0.473	1	0.492	1.218
犧牲短打	-0.488	0.364	1.799	1	0.180	0.614
犧牲飛球	-0.120	0.368	0.107	1	0.744	0.887
盜壘成功	0.124	0.246	0.252	1	0.616	1.132
盜壘刺殺	-0.208	0.388	0.287	1	0.592	0.812
失誤	0.129	0.210	0.379	1	0.538	1.138
團隊打擊率	-1.352	3.363	0.162	1	0.688	0.259
壘打數	-0.047	0.037	1.629	1	0.202	0.954
上壘率	-4.144	2.914	2.022	1	0.155	0.016
長打率	-1.795	1.736	1.070	1	0.301	0.166
攻擊指數	-1.320	0.769	2.946	1	0.086	0.267
面對打擊數	0.080	0.057	1.977	1	0.160	1.083
投球數	0.004	0.012	0.136	1	0.712	1.004
好球數	-0.019	0.020	0.849	1	0.357	0.981
被安打	0.001	0.077	0.000	1	0.992	1.001

表 4.4.11 比賽進行 20%時各項成績差前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
被全壘打	0.318	0.229	1.938	1	0.164	1.375
四壞球	0.231	0.113	4.198	1	0.040	1.260
四死球	0.126	0.288	0.192	1	0.661	1.134
奪三振	-0.252	0.099	6.459	1	0.011	0.777
三振四壞比	-0.169	0.070	5.907	1	0.015	0.844
暴投	-0.038	0.262	0.021	1	0.885	0.963
投手犯規	-21.203	28414.293	0.000	1	0.999	0.000
防禦率	0.134	0.081	2.726	1	0.099	1.143
每局被上壘率	0.667	0.464	2.068	1	0.150	1.948

由表 4.4.12 得到，前 48 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.068-0.169\times\text{三振四壞比}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.4.12 使用前一場建立模型-前 48 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.068	0.328	0.043	1	0.836	1.070
三振四壞比	-0.169	0.070	5.907	1	0.015	0.844
配適度	卡方	2.972	df	7	P 值	0.888

表 4.4.13 是主客場各項成績差在比賽進行 40%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中打數、每打數平均打點、二壘安打、團隊打擊率、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、四壞球、奪三振、三振四壞比、防禦率、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.13 比賽進行 40%時各項成績差前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.150	0.058	6.599	1	0.010	0.861
每打數平均打點	-5.855	2.129	7.562	1	0.006	0.003
二壘安打	-0.405	0.135	8.993	1	0.003	0.667
三壘安打	0.031	0.390	0.006	1	0.937	1.031
雙殺打	0.053	0.169	0.100	1	0.751	1.055
犧牲短打	-0.284	0.287	0.979	1	0.322	0.753
犧牲飛球	-0.140	0.274	0.263	1	0.608	0.869
盜壘成功	-0.164	0.146	1.261	1	0.261	0.849
盜壘刺殺	-0.063	0.256	0.061	1	0.805	0.939
失誤	0.120	0.163	0.546	1	0.460	1.128
團隊打擊率	-5.846	2.194	7.098	1	0.008	0.003
壘打數	-0.083	0.026	10.305	1	0.001	0.920
上壘率	-7.141	2.125	11.290	1	0.001	0.001
長打率	-3.293	1.130	8.484	1	0.004	0.037
攻擊指數	-1.836	0.525	12.212	1	0.000	0.159
面對打擊數	0.180	0.049	13.295	1	0.000	1.197
投球數	0.022	0.009	5.827	1	0.016	1.022
好球數	0.021	0.015	1.963	1	0.161	1.021
被安打	0.139	0.052	7.232	1	0.007	1.149
被全壘打	0.280	0.146	3.647	1	0.056	1.323
四壞球	0.240	0.086	7.805	1	0.005	1.272
四死球	0.256	0.201	1.615	1	0.204	1.291
奪三振	-0.161	0.060	7.213	1	0.007	0.851

表 4.4.13 比賽進行 40%時各項成績差前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
三振四壞比	-0.173	0.057	9.231	1	0.002	0.841
暴投	0.008	0.181	0.002	1	0.963	1.008
投手犯規	-21.157	23148.410	0.000	1	0.999	0.000
防禦率	0.148	0.054	7.657	1	0.006	1.160
每局被上壘率	1.147	0.341	11.304	1	0.001	3.150

由表 4.4.14 得到，前 96 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.325+0.17\times\text{面對打擊數}-0.142\times\text{奪三振}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.4.14 使用前一場建立模型-前 96 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.325	0.241	1.819	1	0.177	1.384
面對打擊數	0.170	0.050	11.521	1	0.001	1.185
奪三振	-0.142	0.065	4.737	1	0.030	0.867
配適度	卡方	8.028	df	8	P 值	0.431

表 4.4.15 是主客場各項成績差在比賽進行 70%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中每打數平均打點、二壘安打、失誤、壘打數、上壘率、長打率、攻擊指數、面對打擊數、投球數、被安打、四壞球、奪三振、三振四壞比、每局被上壘率的 P 值均小於 0.05，達到顯著水準。

表 4.4.15 比賽進行 70%時各項成績差前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.051	0.036	1.997	1	0.158	0.951
每打數平均打點	-2.771	1.306	4.502	1	0.034	0.063

表 4.4.15 比賽進行 70%時各項成績差前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
二壘安打	-0.169	0.084	4.096	1	0.043	0.844
三壘安打	-0.008	0.278	0.001	1	0.976	0.992
雙殺打	-0.026	0.119	0.047	1	0.828	0.974
犧牲短打	-0.210	0.184	1.312	1	0.252	0.810
犧牲飛球	-0.209	0.201	1.083	1	0.298	0.811
盜壘成功	-0.027	0.102	0.068	1	0.794	0.974
盜壘刺殺	0.054	0.190	0.080	1	0.778	1.055
失誤	0.246	0.121	4.171	1	0.041	1.280
團隊打擊率	-2.762	1.466	3.549	1	0.060	0.063
壘打數	-0.033	0.015	4.894	1	0.027	0.968
上壘率	-3.373	1.372	6.044	1	0.014	0.034
長打率	-1.592	0.748	4.531	1	0.033	0.204
攻擊指數	-0.712	0.299	5.679	1	0.017	0.491
面對打擊數	0.077	0.028	7.588	1	0.006	1.080
投球數	0.017	0.006	7.511	1	0.006	1.017
好球數	0.019	0.011	2.963	1	0.085	1.019
被安打	0.071	0.033	4.596	1	0.032	1.073
被全壘打	0.108	0.096	1.264	1	0.261	1.114
四壞球	0.130	0.055	5.522	1	0.019	1.139
四死球	0.047	0.152	0.095	1	0.757	1.048
奪三振	-0.143	0.044	10.595	1	0.001	0.867
三振四壞比	-0.178	0.047	14.323	1	0.000	0.837
暴投	0.112	0.140	0.641	1	0.424	1.118

表 4.4.15 比賽進行 70%時各項成績差前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
投手犯規	-21.183	23195.921	0.000	1	0.999	0.000
防禦率	0.061	0.032	3.711	1	0.054	1.063
每局被上壘率	0.575	0.210	7.511	1	0.006	1.777

由表 4.4.16 得到，前 168 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.125-0.178\times\text{三振四壞比}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.4.16 使用前一場建立模型-前 168 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.125	0.167	0.555	1	0.456	1.133
三振四壞比	-0.178	0.047	14.323	1	0.000	0.837
配適度	卡方	9.838	df	8	P 值	0.277

表 4.4.17 是主客場各項成績差在比賽進行 90%時選入前一場 28 個解釋變數的單變數邏輯斯迴歸，其中失誤、奪三振、三振四壞比的 P 值均小於 0.05，達到顯著水準。

表 4.4.17 比賽進行 90%時各項成績差前一場的單變數邏輯斯迴歸

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
打數	-0.028	0.030	0.897	1	0.344	0.972
每打數平均打點	-1.297	1.115	1.354	1	0.245	0.273
二壘安打	-0.104	0.072	2.049	1	0.152	0.901
三壘安打	-0.006	0.219	0.001	1	0.977	0.994
雙殺打	0.042	0.109	0.146	1	0.702	1.043
犧牲短打	-0.084	0.158	0.281	1	0.596	0.920

表 4.4.17 比賽進行 90%時各項成績差前一場的單變數邏輯斯迴歸(續)

解釋變數	參數估計值	S. E.	Wals	df	P 值	勝算比估計值
犧牲飛球	-0.181	0.174	1.088	1	0.297	0.834
盜壘成功	-0.071	0.092	0.591	1	0.442	0.931
盜壘刺殺	-0.008	0.159	0.002	1	0.962	0.992
失誤	0.243	0.106	5.277	1	0.022	1.275
團隊打擊率	-0.697	1.190	0.343	1	0.558	0.498
壘打數	-0.016	0.013	1.606	1	0.205	0.984
上壘率	-1.077	1.125	0.916	1	0.338	0.341
長打率	-0.637	0.638	0.995	1	0.318	0.529
攻擊指數	-0.341	0.255	1.783	1	0.182	0.711
面對打擊數	0.044	0.023	3.490	1	0.062	1.044
投球數	0.010	0.005	3.570	1	0.059	1.010
好球數	0.010	0.009	1.290	1	0.256	1.010
被安打	0.025	0.027	0.891	1	0.345	1.025
被全壘打	0.043	0.087	0.244	1	0.621	1.044
四壞球	0.077	0.048	2.593	1	0.107	1.080
四死球	-0.039	0.137	0.082	1	0.774	0.961
奪三振	-0.104	0.038	7.544	1	0.006	0.902
三振四壞比	-0.081	0.029	7.865	1	0.005	0.922
暴投	0.083	0.127	0.426	1	0.514	1.087
投手犯規	-0.414	0.913	0.205	1	0.651	0.661
防禦率	0.024	0.027	0.813	1	0.367	1.024
每局被上壘率	0.250	0.173	2.107	1	0.147	1.285

由表 4.4.18 得到，前 216 場建立的邏輯斯迴歸估計模型如下：

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)=0.056+0.231\times\text{失誤}-0.078\times\text{三振四壞比}$$

其中  $\hat{p}$  為主隊比賽獲勝的估計機率。

表 4.4.18 使用前一場建立模型-前 216 場

變數	參數估計值	S.E.	Wald	df	P 值	勝算比估計值
常數	0.056	0.142	0.158	1	0.691	1.058
失誤	0.231	0.108	4.603	1	0.032	1.260
三振四壞比	-0.078	0.029	7.340	1	0.007	0.925
配適度	卡方	10.387	df	8	P 值	0.239

## 八、 驗證分析與比較

由表 4.4.19 得到，建模時間前 48 場時，預測 2、預測 4、預測 5 的預測能力並列最好，成功場次達到 7 場；建模時間前 96 場時，預測 2、預測 3、預測 5 的預測能力並列最好，成功場次達到 8 場；建模時間前 168 場時，預測 1、預測 4、預測 5、預測 6 的預測能力並列最好，成功場次達到 5 場；建模時間前 216 場時，預測 1、預測 2、預測 3、預測 6 預測能力並列最好，成功場次達到 8 場；建模時間下半季前 48 場時，預測 4 預測能力最好，成功場次達到 6 場；建模時間下半季前 96 場時，預測 4 的預測能力最好，成功場次達到 8 場。

表 4.4.19 主客隊各項成績差的分析整合表

建模場次 (預測場次)	建模成功率 (P 值)	預測方式	預測 成功 次數	預測方式	預測 成功 次數	預測方式 平均
前 48 場	68.10% (0.566)	預測 1	<u>6</u>	預測 4	<u>7</u>	6.5
		預測 2	<u>7</u>	預測 5	<u>7</u>	
		預測 3	<u>6</u>	預測 6	<u>6</u>	

表 4.4.19 主客隊各項成績差的分析整合表(續)

建模場次 (預測場次)	建模成功率 (P 值)	預測方式	預測 成功 次數	預測方式	預測 成功 次數	預測方式 平均
前 96 場	69.10% (1.000)	預測 1	<u>5</u>	預測 4	6	7
		預測 2	<u>8</u>	預測 5	<u>8</u>	
		預測 3	<u>8</u>	預測 6	7	
前 168 場	62.20% (1.000)	預測 1	<u>5</u>	預測 4	<u>5</u>	4.5
		預測 2	4	預測 5	<u>5</u>	
		預測 3	<u>3</u>	預測 6	<u>5</u>	
前 216 場	56.60% (0.344)	預測 1	<u>8</u>	預測 4	<u>6</u>	7.33
		預測 2	<u>8</u>	預測 5	<u>6</u>	
		預測 3	<u>8</u>	預測 6	<u>8</u>	
四個時間點平均		預測 1	<u>6</u>	預測 4	<u>6</u>	6.33
		預測 2	<u>6.75</u>	預測 5	6.5	
		預測 3	6.25	預測 6	6.5	
下半季前 48 場	70.80% (1.000)	預測 1	5	預測 4	<u>6</u>	4.67
		預測 2	5	預測 5	5	
		預測 3	4	預測 6	<u>3</u>	
下半季前 96 場	60.40% (1.000)	預測 1	7	預測 4	<u>8</u>	5.83
		預測 2	5	預測 5	6	
		預測 3	5	預測 6	<u>4</u>	

從表 4.4.11 到表 4.4.18 建立模型，分別得到前一場之邏輯斯迴歸模型，再由表 4.4.20 得到，建模成功率大約落在 50%-70%左右，跟前面相比有些差距，而

在驗證的方面，在前 216 場驗證成功的場次最多，預測成功的場次達到 8 場。

表 4.4.20 主客隊各項成績差-前一場的分析整合表

建模場次	建模成功率	P 值	預測成功次數
前 48 場	68.10%	0.888	7
前 96 場	69.10%	0.431	7
前 168 場	62.20%	0.277	4
前 216 場	56.60%	0.239	8

由表 4.4.19 和表 4.4.20 得到，先以全季時間點來看，再把各個預測時間點跟預測方式分開來看，在預測第 97-108 場且使用預測 2、預測 3、預測 5 的方式時，以及預測 217-226 場且使用預測 1、預測 2、預測 3、預測 6 的方式時，預測能力最佳，成功預測的場次來到 8 場；在預測第 169-178 場時，預測 3 的預測能力最差，成功預測的場次只有 3 次。以預測時間點區分，預測第 217-226 場時，平均預測成功最高，達到 7.33 次；預測第 169-178 場時，平均預測成功最低，只有 4.5 次。以預測方式來區分，預測 2 的預測能力最好，平均預測成功達到了 6.75 次；預測 1 和預測 4 的預測能力則是最差，平均預測成功有 6 次。

將第 169-178 場、第 217-226 場及下半季第 49-58 場、下半季第 97-106 場進行比較，下半季第 49-58 場的預測成功率是略優於第 169-178 場，而第 217-226 場的預測成功率則是略優於下半季第 97-106 場。

將前一場與預測能力最佳的預測 2 做比較，在第 97-108 場的部分，預測 2 優於前一場，而第 49-58 場、第 169-178 場、第 217-226 場則是相同預測成功場數。

在全季時間點的部分，平均預測成功次數來到 6.33 次，而前一場平均預測成功次數則是 6.5 次，後者優於前者。

## 第五節 選入的解釋變數統計

表 4.5.1 是主客隊各項成績—打擊成績與投手成績，這裡共計選入 36 個解釋變數(不分主客隊)。

表 4.5.1 主客隊各項成績選入的解釋變數統計

	主客隊各項成績	主客隊各項成績(前一場)
前 48 場	客隊暴投 客隊每局被上壘率 主隊防禦率	客隊奪三振 主隊好球數 主隊防禦率
前 96 場	客隊每打數平均打點 客隊失誤 主隊盜壘刺殺 客隊防禦率	客隊長打率 主隊二壘安打 客隊三振四壞比
前 168 場	客隊每打數平均打點 客隊失誤 客隊三振四壞比 客隊防禦率	客隊長打率 客隊三振四壞比 主隊三振四壞比
前 216 場	客隊每打數平均打點 客隊失誤 客隊上壘率 主隊每打數平均打點 主隊犧牲飛球 主隊盜壘刺殺 主隊打擊率 客隊防禦率 主隊防禦率	客隊失誤 客隊三振四壞比 主隊奪三振
下半季前 48 場	客隊每打數平均打點 主隊上壘率	
下半季前 96 場	客隊上壘率 主隊每打數平均打點 客隊防禦率	

表 4.5.2 是主客隊各項成績選入變數的統計次數，這裡統計是不計算主客隊，其中防禦率、每打數平均打點及三振四壞比為前 3 多的選入變數。

表 4.5.2 主客隊各項成績選入的解釋變數統計次數

選入變數	次數	選入次數	變數
防禦率	7	每打數平均打點	6
三振四壞比	5	失誤	4
長打率	2	上壘率	3
奪三振	2	盜壘刺殺	2
暴投	1	每局被上壘率	1
好球數	1	犧牲飛球	1
打擊率	1		

表 4.5.3 是主客隊各項成績比—打擊成績比與投手成績比，這裡共計選入 21 個解釋變數，詳細統計如下表。

表 4.5.3 主客隊各項成績比選入的解釋變數統計

	主客隊各項成績比	主客隊各項成績比(前一場)
前 48 場	攻擊指數	四壞球
前 96 場	每打數平均打點 每局被上壘率	二壘安打 壘打數 面對打擊數 好球數
前 168 場	每打數平均打點 每局被上壘率	投球數 奪三振
前 216 場	每打數平均打點 面對打擊數 每局被上壘率	四壞球 奪三振
下半季前 48 場	每打數平均打點	
下半季前 96 場	每打數平均打點 投球數 每局被上壘率	

表 4.5.4 是主客隊各項成績比選入變數的統計次數，其中每打數平均打點及每局被上壘率為前 2 多的選入變數。

表 4.5.4 主客隊各項成績比選入的解釋變數統計次數

選入變數	次數	選入次數	變數
每打數平均打點	5	每局被上壘率	4
四壞球	2	面對打擊數	2
投球數	2	奪三振	2
攻擊指數	1	二壘安打	1
壘打數	1	好球數	1

表 4.5.5 是主客隊各項成績差—打擊成績差與投手成績差，這裡共計選入 23 個解釋變數，詳細統計如下表。

表 4.5.5 主客隊各項成績差選入的解釋變數統計

	主客隊各項成績差	主客隊各項成績差(前一場)
前 48 場	防禦率	三振四壞比
前 96 場	雙殺打 失誤 防禦率	面對打擊數 奪三振
前 168 場	盜壘刺殺 失誤 好球數 三振四壞比 防禦率	三振四壞比
216 場	盜壘刺殺 失誤 面對打擊數 防禦率 每局被上壘率	失誤 三振四壞比
下半季前 48 場	每打數平均打點	
下半季前 96 場	每打數平均打點 奪三振	

表 4.5.6 是主客隊各項成績差選入變數的統計次數，其中三振四壞比、防禦率及失誤為前 3 多的選入變數。

表 4.5.6 主客隊各項成績差選入的解釋變數統計次數

選入變數	次數	選入次數	變數
三振四壞比	4	防禦率	4
失誤	4	面對打擊數	2
奪三振	2	盜壘刺殺	2
每打數平均打點	2	雙殺打	1
好球數	1	每局被上壘率	1

表 4.5.7 是三種解釋變數處理方式選入變數的統計總次數，由其統計出的變數可以得到，在用邏輯斯迴歸建立模型的這個步驟中，三種解釋變數處理方式(包含前一場的方式)選入共 22 種變數，其中每打數平均打點、防禦率跟三振四壞比(不分主隊客隊)選入的次數最多，可以認為是影響勝負的重要變數，也可以得知場次越多，選入的變數也越多。

表 4.5.7 選入解釋變數的統計總次數

變數	次數	變數	次數	變數	次數
每打數平均打點	13	防禦率	11	三振四壞比	9
失誤	8	每局被上壘率	6	奪三振	6
盜壘刺殺	4	面對打擊數	4	上壘率	3
好球數	3	長打率	2	二壘安打	2
四壞球	2	投球數	2	暴投	1
犧牲飛球	1	攻擊指數	1	雙殺打	1
壘打數	1	打擊率	1		

# 第五章 結論與建議

## 第一節 結論

### 一、模型建立與預測間的差異

主客隊各項成績(打擊成績與投手成績)在第四章第二節的驗證可以得知，預測 5(兩隊上次對戰數據與前一場數據(分主客場)的平均)的方式在驗證中表現最佳，10 場預測平均準確率來到了 7 場，其中在預測第 97-108 場則是達到 9 場的準確率；預測 2(兩隊上次對戰數據)和預測 4(前一場數據(分主客場))的準確率則是在預測 5 之後，10 場預測平均準確率是 6.75 次。

主客隊各項成績比(打擊成績比與投手成績比)在第四章第三節的驗證可以得知，預測 2 的方式在驗證中表現最佳，10 場預測平均準確率來到了 7.25 場，其中在預測第 49-58 場以及第 97-108 場則是達到了 9 場的準確率；預測 4 的準確率則是排在預測 2 之後，10 場預測平均準確率是 7 次。

主客隊各項成績差(打擊成績差與投手成績差)在第四章第四節的驗證可以得知，預測 2 的方式在驗證中表現最佳，10 場預測平均準確率來到了 6.75 場，其中在預測第 97-108 場以及第 217-226 場則是達到了 8 場的準確率；預測 5 和預測 6(近五場比賽的平均)的準確率則是排在預測 2 之後，10 場預測平均準確率是 6.5 次。

### 二、預測場次的差別

在第四章第二節到第四節，分別都對賽季 20%(全季前 48 場)、賽季 40%(全季前 96 場)、賽季 70%(全季前 168 場)及賽季 90%(全季前 216 場)建立模型，且對其後 10 場(第 49-58 場、第 97-108 場(第 100、101 場結果為和局，故不做預測)、第 169-178 場、第 217-226 場)進行預測驗證，從驗證結果可以得知，第一種解釋變數處理及第二種解釋變數處理預測的第 97-108 場，10 場預測平均準確

率都是最高的，而主客隊各項成績差則是第 217-226 場的 10 場預測平均準確率最高。

在賽季 70%及 90%的部分與下半季 20%(下半季前 48 場，預測下半季第 49-58 場)、下半季 40%(下半季前 96 場，預測下半季第 97-106 場)做比較，只有主客隊各項成績差的下半季 40%預測準確率優於賽季 90%、主客隊各項成績比的賽季 70%跟下半季 20%預測準確率相同，其餘皆是賽季優於下半季，下一節會對此做討論。

在對前一場做解釋變數處理時，主客隊各項成績、主客隊各項成績比、主客隊各項成績差的 10 場平均預測成功場次分別是 5 場、6.25 場、6.5 場，全賽季最高平均預測場次分別是各項成績預測 5 的 7 場、各項成績比預測 2 的 7.25 場、各項成績差預測 2 的 6.75 場，很明顯全賽季都優於前一場，所以可以判定原本的解釋變數處理方式比較好。

## 第二節 討論與建議

### 一、選入變數之討論

在第四章第五節的部分對選入之變數進行統計，共選入了 22 種變數，其中選入次數超過兩次的共有 14 種，超過五次的共有 6 種，而十次以上的則是只有 2 種。跟打擊方面相關的變數共有 10 種(每打數平均打點、盜壘刺殺、上壘率、長打率、二壘安打、犧牲飛球、攻擊指數、雙殺打、壘打數、打擊率)，而跟投球與守備方面相關的則有 10 種(防禦率、三振四壞比、失誤、每局被上壘率、奪三振、面對打擊數、好球數、四壞球、投球數、暴投)，其中奪三振、四壞球、四死球在第四章第一節有提到，因為只取自單方面的資料，所以在此先算在投球方面的統計。

選入變數最高的三個為每打數平均打點、防禦率、三振四壞比，分別在此討論這三個變數。每打數平均打點在此解釋為，每個打數可以獲取分數越高，比賽

越容易獲勝，在第四章有提到本研究在選入變數時將其從每打點平均打數更改，也就是說，這個變數的數值越高，代表當比賽球隊打數越高時，得到的分數越多。從上一章也可以得知，當為主客隊各項成績時，這若客隊的這個數值越高，主隊也越不容易獲勝，反之亦然；當為主客隊各項成績比時，客隊主隊相除之後的商越大，主隊越不容易獲勝；當為主客隊各項成績差時，客隊主隊相減之後的差越大，主隊越不容易獲勝。

防禦率為自責分 $\times 9$ /投球局數，當此數值越高時，若局數固定時，投手失掉的自責數越多，也越不容易獲得比賽勝利。從上一章也可以得知，當為主客隊各項成績時，這若客隊的這個數值越高，主隊也越容易獲勝，反之亦然；當為主客隊各項成績比時，客隊主隊相除之後的商越大，主隊越容易獲勝；當為主客隊各項成績差時，客隊主隊相減之後的差越大，主隊越容易獲勝。

三振四壞比是投手奪三振與四壞球的比例，靠著三振讓人出局的數目越多，以及降低四壞球讓人上壘的機率，這個值越高，越容易獲得比賽勝利。從上一章也可以得知，當為主客隊各項成績時，這若客隊的這個數值越高，主隊也越不容易獲勝，反之亦然；當為主客隊各項成績比時，客隊主隊相除之後的商越大，主隊越不容易獲勝；當為主客隊各項成績差時，客隊主隊相減之後的差越大，主隊越不容易獲勝。

## 二、解釋變數處理與場次的建議

解釋變數處理採用的是打擊成績與投手成績(分主客隊)、打擊成績比與投手成績比及打擊成績差與投手成績差，其中打擊成績比與投手成績比(主客各項成績比)的部分，採取的是將客隊與主隊的變數進行相除，此時就會產生當主隊變數為零時，整個變數就會不存在，若產生以上情況，本研究採用的是將資料為零的客隊主隊變數各加上 0.5，但調整過後的變數有可能因此變得極大，本研究認為不太恰當，可惜未能找到更好的處理方法，希望未來可以找到更佳的处理方式。

建立模型的場次部分一開始採用只有 4 個時間點，分別為賽季 20%、賽季 40%、賽季 70%以及賽季 90%，之後在對其後 10 場進行預測，但在賽季 70%及 90%的部分，預測成功率很明顯低於前兩個時間點，尤其是賽季 70%，很明顯遠低於其它三個時間點，在此懷疑是否因為是下半季的原因(中職球季分為上下半季，若為半季冠軍，可以在季後進行冠軍賽)，所以球隊進行調整，將主力選手以及未上場之選手調換，以維持冠軍賽的戰力，因此本研究將下半季的資料獨立出來進行建模與預測，再與原本全季的資料進行比對。結果為全季預測優於下半季預測，因此判定推論為錯誤的，原因可能在於場次的選擇，或是模型建立的方式，若將預測的場次從 10 場改成 20 場，或是將模型建立的時間點從 70%改成 60%或 80%，有可能提升預測準確率。

### 三、其餘建議

本研究採取邏輯斯迴歸選用中職四隊在 2017 年賽季共 236 場比賽資料，這樣會先產生一些問題，先以球隊來看，若是四支球隊，這樣以比賽組合來看就會有 6 種組合，再加上主客場的話，會高達 12 種組合，因此其餘研究大部分採用對兩支球隊跨球季的研究，否則在建立模型或進行預測時，是否將組合納入，這是需要考慮的。而中華職業棒球大聯盟的選擇，因為中職比賽有和局的結果判定，這對場次為有些許的影響，本研究採用的是邏輯斯迴歸，對結果只有兩個變數，所以和局的部分可能要進行刪去，但若採用其它的統計分析方法，或是其它的職業棒球聯盟，也許就可以改善這個問題。

本研究最後得知的三種解釋變數處理方式各項成績、各項成績比、各項成績差的最高平均預測準確率分別是預測 5(兩隊上次對戰數據與前一場數據(分主客場)平均)的 7 場、預測 2(兩隊上次對戰數據)的 7.25 場、預測 2 的 6.75 場，雖然各項成績比的最高平均預測準確率最高，但因為需要進行資料調整，所以暫不推薦，而各項成績優於各項成績差，並且各項成績將所有資料都投入模型中

比較能反映整個數據的差異，所以最為推薦。

另外為何使用 6 種預測的原因，在於進行預測時，無法得知該場比賽的解釋變數，所以只能透過已完成的賽事進行假設，在研究中使用的 6 種驗證方式中，預測 2(兩隊上次對戰數據)平均預測成功場次最佳，所以在此認為最適當。

# 參考文獻

## 一、中文部分

- 中華民國棒球協會 (2006)。台灣棒球百年史。台北市：中華民國棒球協會。
- 方敬硯 (2017)。利用資料探勘技術建立運動彩券投注之預測模式-以職業棒球聯盟為例(碩士論文)。取自臺灣碩博士論文系統。(系統編號 105YUNT0031028)
- 王濟川、郭志剛 (2003)。Logistic 迴歸模型—方法及應用。台北市：五南。
- 吳仁 (1996)。棒球操典。台北市：聯經。
- 李俊德 (2012)。預測職棒比賽結果—以中華職棒二十一年為例(碩士論文)。取自臺灣碩博士論文系統。(系統編號 100NPTT5507002)
- 林文斌、鄧元湘 (2000)。長打率在現代棒球中扮演的角色，德明學報，16，315-324。
- 林家鳳 (2016)。中華職棒球員攻擊戰術模型之建置(碩士論文)。取自臺灣碩博士論文系統。(系統編號 104NTPU0337020)
- 施致平、黃蕙娟、倪瑛蓮 (2010)。中華職棒比賽勝負預測模式之建構。體育學報，43(2)，115-130。
- 張美惠(譯) (2008)。瘋足球，迷棒球-職業運動經濟學(原作者：S. Szymanski 及 A. Zimbalist)。臺北市：時報文化，15-61。
- 許璫方 (2010)。Probit 迴歸在職棒比賽結果預測的應用(碩士論文)。取自臺灣碩博士論文系統。(系統編號 098NHCT5507009)
- 陳麒文、李天行、梁玉秋 (2014)。建構美國職棒大聯盟的勝隊預測模式：以人工類神經網路方式。大專體育學刊，16(2)，167-181。
- 馮瑞祥 (2013)。運用資料探勘技術於職棒比賽勝負預測之研究-以美國職棒大聯盟為例(碩士論文)。取自臺灣碩博士論文系統。(系統編號 100PCCU1396051)

葛學文 (2014)。美國職棒大聯盟之球隊績效和球員能力資料採礦(碩士論文)。

取自臺灣碩博士論文系統。(系統編號 102THU00337001)

鄧文揚 (2013)。中華職棒球隊勝負預測模型之建立(碩士論文)。取自臺灣碩博

士論文系統。(系統編號 101HCU00396003)

瘦菊子 (2003)。燦爛的球季—台灣職棒風雲。台北市：聯合文學。

## 二、外文部分

Bock, J. R. (2015). Pitch sequence complexity and long-term pitcher performance.

Sports, 3(1), 40-55.

Steve, S. (1995). *Practical Biostatistical Methods*. California: Duxbury Press.

## 三、網站部分

中華職棒大聯盟。中華職棒大聯盟全球資訊網 The Official Site of CPBL 【關於

中職】。取自網址 <http://www.cpbl.com.tw/footer/about/>

中華職棒大聯盟。中華職棒大聯盟全球資訊網 The Official Site of CPBL 【球

隊】【球員成績】。取自網址

[http://www.cpbl.com.tw/web/team\\_playergrade.php?&team=E02&gameno=01](http://www.cpbl.com.tw/web/team_playergrade.php?&team=E02&gameno=01)

台灣運彩。台灣運彩【關於運動彩券】。取自網址

<https://www.sportslottery.com.tw/zh/web/guest/about-sports>

竹板凳(2016)。台灣棒球維基館。取自網址

<http://twbsball.dils.tku.edu.tw/wiki/index.php/%E6%A3%92%E7%90%83>