

# プロ野球球団別戦力の統計的分析

## —様々な分析法を用いて—

2006MI081 小牧祐介

指導教員：木村美善

### 1 はじめに

3年次より統計学を専門的に学ぶようになり、ゼミでは野球について様々な分析をしていたので、卒業論文では野球データと統計学を用いて、研究に取り組みたいと考えた。また個人的に好きな阪神というチームは、どのような特徴があるのか、得意、または苦手チームなどはあるのか、など様々な疑問について統計学を用いて分析することができないだろうか、と考えたことが本研究をテーマとして選んだきっかけである。本研究では、阪神を中心にセリーグ球団の統計的分析を行った。

### 2 データ

「日本野球機構オフィシャルサイト」[3]から、Bradley-Terry モデル、移動平均グラフでは、2005年度～2008年度成績の「勝敗」を用いた。また、重回帰分析では2008年度の「打数」、「一二三塁打」、「打点」、「三振」、「被四死球」、「本塁打」、「打者」、「被安打」、「奪三振」、「四死球」、「自責点」、「勝敗」、「屋内外」、「ホーム」、「先取点」の全15項目を用いた。

### 3 Bradley-Terry モデル

最初にチームの「強さ」というものについて考えると、一般に、スポーツにおいては「強い」方が必ず勝つとは限らない。実際「強い」はずのチームが負ける「番狂わせ」がなく、一方が常に勝つものと決まっていれば、ゲームの興味はなくなってしまう。したがって、あるチームが必ず勝つということではなく、そのチームの勝つ「確率」が大きいことを意味すると思えなければならない。そこで勝敗の確率を問題とする。いま、全部で  $m$  チームが互いに何回か対戦するものとし、そのうちのチーム  $i$  が  $j$  に勝つ確率を

$$p_{ij} \quad (i, j; i, j = 1, 2, \dots, m)$$

と表す。ここでは簡単のために引き分けはないものと仮定する。そうすると、

$$p_{ij} + p_{ji} = 1 \quad (i, j)$$

が成り立つ。各チームに対応して  $m$  個の量  $p_1, p_2, \dots, p_m$  が存在して、すべての  $i, j$  の組み合わせに対し、確率  $p_{ij}$  が

$$p_{ij} = \frac{e^{p_i}}{e^{p_i} + e^{p_j}} \quad (i, j)$$

と表されるものと想定する。このとき、 $p_i$  はチーム  $i$  の「強さ」を表すものと解釈できる。この考え方は、Bradley and Terry によるもので、Bradley-Terry モデルと呼ばれる。

### 4 カモ・苦手関係

Bradley-Terry モデルの「カモ・苦手」関係の量的表現を用いて各年度ごとの「カモ・苦手」について探る ([2], 参照)。

#### 4.1 2008年度カモ・苦手関係表

表1 2008年度セリーグ

	巨人	阪神	中日	広島	ヤク	横浜
巨人	X	0.34	-0.88	-0.77	0.91	0.48
阪神	-0.34	X	1.08	0.06	-0.12	-0.71
中日	0.88	-1.08	X	0.34	-0.65	0.46
広島	0.77	-0.06	-0.34	X	0.05	-0.41
ヤク	-0.91	0.12	0.65	-0.05	X	0.12
横浜	-0.48	0.71	-0.46	0.41	-0.12	X

#### 4.2 カモ・苦手関係の考察

この数値が大きな正の値をとるときは「カモ」にし、大きな負の値をとるときは「苦手」にしているといえる。ここでもおもしろいのは、上位チームが下位チームにたとえ勝ち越し、引き分けていても「苦手」にしている。つまり正の値ではなく負の値を示していることである。表1の阪神 - 横浜 (13.5勝 - 10.5勝) がその例である。これは、正確にいえば下位チームが上位チームを「カモ」にしたというのではなく、実力以上に善戦したことを表している。阪神が「カモ」にした相手として2006年横浜、2007年巨人、2008年中日あげられる。特に2006年の横浜と2008年の中日は1.141、1.076と非常に「カモ」にしていた。この3チームに共通していた点は、失点が少ないということである。また、阪神が「苦手」にした相手として2006年中日、2007年広島あげられる。特に2007年の広島には-1.031と非常に「苦手」としていた。この2チームに共通していた点は、失点が多いということである。2005年度～2008年度の4年間のデータを用いて分析した結果、阪神は特定の1チームに「カモ・苦手」があるわけではないことが分かった。阪神が優勝できない主な理由のひとつとして実力どおりに勝てないことが考えられる。下位チームに善戦されている年度が4年中、3年もあったので、下位チームに実力どおり勝つことが優勝につながると考えられる。また、失点が少ないチームを「カモ」にし、失点が多いチームを「苦手」としていることで、阪神は打撃成績より、投手成績のほうが勝敗に強く影響していると考えられる。

### 5 移動平均法

移動平均グラフを用いてチームごとの調子の波の激しさを調べた。 $y$ 軸の中心に書かれている数字は、そのシーズンの勝率である。[1]を参考にし、8ゲームを移動平均

ゲーム数とした。

### 5.1 阪神の移動平均グラフ

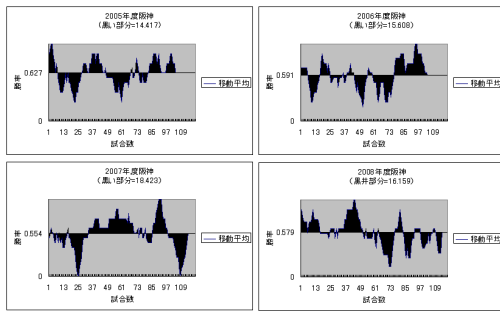


図 1 阪神移動平均

### 5.2 移動平均法の考察

このグラフは、シーズンにおける各チームの好調時と不調時を極めて分かりやすく表示している。次の合計値を計算することにより波の激しさの度合いを測定することができる。

黒い部分 = 合計値 [差異 (変動する勝率 - シーズンの勝率)]

この合計値は基本的にグラフの黒い部分のサイズを測ることであり、グラフの中に黒い部分が多く見られるとしたら、そのチームは極めて波が激しいシーズンを送ったことになる。図 1 のように各チームの黒い部分のサイズを比較すると、阪神は他チームと比べて波の激しいチームといえる。熱狂的なファンがいる阪神のようなチームは、ホーム・ビジターが勝率に影響するため波が激しくなるという仮説を立て、ホーム・ビジターで分けて勝率を比べたところ、どの年のどのチームもホームのほうが5%以上勝率が高かった。つまりこの仮説は成り立たず、全てのチームに対して、ホーム・ビジターが勝率に影響があるということである。次に、阪神は打撃成績より投手成績の方が勝敗に影響しているというカモ・苦手関係の結果に着目した。投手成績は個人の出来、不出来で試合内容に強く影響する。よって阪神は投手の出来、不出来が原因で調子の波が激しくなると考えられる。

## 6 重回帰分析

2008 年度データを用い、勝敗を従属変数とし、打数、一二三塁打、打点、三振、被四死球、本塁打、打者、被安打、奪三振、四死球、自責点、屋内外、ホーム、先取点、を説明変数とした。そのうち、屋内外 (外=1, 内=0)、ホーム・ビジター (ホーム=1, ビジター=0)、先取点 (取る=1, 取られた=0) をダミー変数としている。移動平均法から得られた結果をもとに、2008 年度のシーズン勝率である 0.579 に 0.1 を足したもの (=0.679) 以上の期間を好調、0.1 引いたもの (=0.479) 以下の期間を不調、この間の期間を普通として、それぞれの期間を重回帰分析し比較した。

## 6.1 好調時

表 2 好調時

	回帰係数	標準誤差	p 値	
(Intercept)	-0.26008	0.47000	0.582588	
打点	0.09465	0.01311	3.37e-09	***
三振	-0.02019	0.01333	0.136328	
打者	0.05012	0.01815	0.008125	**
被安打	-0.06581	0.02445	0.009762	**
四死球	-0.14994	0.03836	0.000291	***
自責点	-0.04524	0.02421	0.067794	.

## 6.2 重回帰分析の考察

表 2 より、変数選択の結果好調時には三振が選ばれているが他の 2 つの状態では選ばれていない。好調時には積極的に打ちにいっていることがわかる。しかし、三振の係数は負であり敗北に効いていることに違いはなかった。普通時には、一二三塁打、本塁打が他の状態と違って効いていた。しかし他の期間と比べて当てはまるデータが多いため、このような結果が得られたと考えられる。不調時には内外が他の状態と違って効いていた。好調時、普通時、不調時で効いている変数が違うものがあるということが分かった。予想では好調時には一二三塁打や本塁打、奪三振など勝利に関係ありそうな変数が効いてくると考えていたが、自分の予想通りには結果が出ず、はっきりしなかった。

## 7 おわりに

本研究では、統計的方法を用いて野球データを分析したが、やり遂げたという達成感よりも分析をすることの難しさの方を強く感じた。データの取り方、多面的な結果の解釈の必要性など、統計の奥深さを感じた。データの取り方や分析方法を変えてもうまく結果が出なかったり、納得のいく結果が得られないことも多かった。しかし、失敗を繰り返すことによって新たな分析方法や、知識が身についたことは非常によかった。本研究では、阪神を中心に様々な分析を行った。阪神が「カモ・苦手」としているチームがわかったこと、打撃成績より投手成績の方が勝敗に影響していること、阪神は他チームと比べると調子の波が激しいチームであるといったことなどは新しい発見であった。また、ビジターよりホームの方が勝率があがることも改めて分かった。しかし好調時、普通時、不調時ごとの比較分析では効いてくる変数に違いがあることはわかったが、自分の予想を裏付けるための十分な分析と結果が得られなかったことは心残りである。

## 参考文献

- [1] J. アルバート/J. ベネット、後藤寿彦監修、加藤貴昭訳：『メジャーリーグの数理科学上、下』、シュプリンガー・フェアラーク東京株式会社、東京、2005。
- [2] 竹内啓、佐藤和建：『スポーツの数理科学』、共立出版株式会社、1988。
- [3] 日本野球機構オフィシャルサイト  
<http://www.npb.or.jp/>