2010年度プロ野球球団別の統計的分析

2006MI003 安藤道太 指導教員:木村美善

はじめに

私は野球が好きで、野球中継や野球場へ試合を見に行く ことも多く一番身近なスポーツである. そこで、大学の専 攻として統計学を学んだことを活かし、プロ野球セントラ ルリーグの各球団の試合ごとのデータを用いて各球団に それぞれどのような特徴があるかを統計的に分析をする ことを卒業研究とした.

2 データについて

中日スポーツ([3]参照)から今年のプロ野球ペナント レースの交流戦以降のプロ野球球団の中日ドラゴンズ・ 阪神タイガース・読売ジャイアンツ・ヤクルトスワロー ズ・広島カープ・横浜ベイスターズの各試合約90試合の データを用いた. 変数は「勝敗 (勝ち=1, 負け=0)」「ホー ム, ビジター (ホーム=1, ビジター=0)」「打数」「得点」 「単打」「二塁打」「三塁打」「本塁打」「盗塁」「犠打飛」 「四死球」「三振」「残塁」「失策」「先発投球回」「継投投手 数」「失点」「被安打」「被本塁打」「与四死球」「奪三振」 である.

分析方法について

本研究では、ロジスティック回帰分析と重回帰分析を行っ た ([1] 参照).

4 密度推定曲線

図1~図6は各球団の勝敗の得失点差による密度推定曲 線である([2] 参照). 中日、ヤクルトは僅差での勝利が多く $1 \sim 2$ 点が頂点となる曲線を示した. 阪神は0 を境にほぼ 左右対称な曲線を表した. 最も特異なのは巨人であった. 僅差での試合が少なく ± 2 の2ヶ所を頂点にとった. 広島、 横浜はそれぞれ・2~・3点を頂点とする曲線となった.



図1 中日

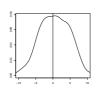


図 2 阪神

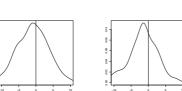


図 4 ヤクルト

図 5 広島



図3 巨人

図 6 横浜

曲線のプロットの結果から、僅差での勝利が多い中日・ ヤクルト、勝敗を問わず僅差での試合が少ない巨人、勝敗 が得点差に関係ない阪神、負けがグラフの頂点にある広 島・横浜の大きく4つに分けられる.

その中でも中日は突出して2点での勝利が多く,2点以 上の勝利は極端に少なくなることから他球団と比べ投手 力が安定しており、僅差の試合では数多く勝ってきたと考 えられる。これは、本塁打が出にくく点の入りにくいナゴ ヤドームが本拠地であった点も大きいだろう.

また,巨人に関しては2点差や3点差の試合より1点差 の試合が少なくなる傾向があり、不安定な戦い方をしてい たと言える. これは投手力が安定していなかった点と,強 力打線で勝ってきたチームであった点, さらにホームラン の出やすく比較的点の入りやすい東京ドームが本拠地で あった点などが考えられる.

ロジステック回帰分析

5.1 ロジステック回帰曲線

各球団の勝率と得点・失点との関係を調べるため、ロジ ステック回帰モデルに当てはめグラフに表した(図 7, 図 8).

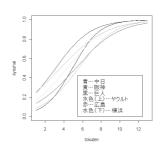


図 7 得点と勝率のロジステ ィック曲線

図8 失点と勝率のロジステ ィック曲線

プロットの結果、特に特徴的だったものは巨人である. 4 点~5点以上を取ると急激に勝率が上昇し、失点と勝率と の関係も他球団とは違い直線的な曲線を描いた. これは 先ほどの密度推定曲線で見られるように投手力は不安定 だが強力打線で打ち勝つ野球をしていると思われる.

また中日及びヤクルトは少ない得点で他球団よりも勝 率が高く、特に中日は他球団と比べ投手力が良く守り勝つ 野球をしていたと言える.

曲線について全体的に比べてみると勝率の低い下位球 団ほど少ない得点での勝率が低いことから、 先発投手・救 援投手を含めた投手力の差が勝率の高い上位球団と出て いるのだと感じた.

5.2 ロジステック回帰分析の結果

各球団の全試合の試合結果の勝敗を目的変数とし、その他の変数を従属変数とし変数選択法(ステップワイズ法)で残ったものでロジステック回帰分析を行った (表 1).

表 1 ロジスティック回帰分析の結果

	Estimate	z value	P値	
(Intercept)	- 4.858	- 2.404	0.016	*
ホーム	1.854	4.287	1.81×10^{-5}	***
単打	0.279	3.485	0.001	***
二塁打	0.477	2.522	0.011	*
三塁打	1.646	3.147	0.001	**
本塁打	1.297	5.255	1.48×10^{-7}	***
盗塁	0.498	1.505	0.132	
犠打飛	0.917	4.250	2.14×10^{-5}	***
先発投球回	0.434	2.184	0.028	*
継投投手数	0.339	1.563	0.118	
被安打	- 0.408	- 4.712	2.46×10^{-6}	***
被本塁打	- 0.709	- 3.324	0.001	***
与四死球	- 0.289	- 2.657	0.007	**

分析の結果、勝利に働く要因として本塁打や単打が有効であるとは感じていたが、ホームでの試合の開催が後攻であるという有利さはあるがこれほどまで有利に効いていたのは意外であった。やはり、僅差の試合の終盤ではどうしても後攻の方が有利であり、ホームグランド特有の声援も関係しているのではないかと思われる。

個別の球団別に見ると、戦力の充実という意味も大きいが中日・阪神・巨人といった上位球団はホームでの試合に強く、それが勝率に表れていた。中日のような球場の広いホームグランドの場合、投手力を中心としたより少ない得点で勝つ野球を目指し、巨人のような球場の狭いホームグランドの場合、打力中心の打ち勝つ野球をするといったようにホーム球場の大きさに合わせて特化した野球を実践できたチームが上位に入る一つの条件のように感じた。

6 重回帰分析

6.1 重回帰分析法とは

得点を目的変数、その他の要素を説明変数として重回帰分析を行った。 重回帰分析では目的変数 y とそれに影響を与えると思われる説明変数 x_1, \dots, x_p について

 $Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_2 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_2 x_2 + b_3 x_3 + \dots + b_p x_p + b_1 x_1 + b_2 x_2 + b_2 x_2 + b_2 x_3 + b_2 x_1 + b_2 x_2 + b_3 x_2 + b_3 x_3 + b_2 x_2 + b_3 x_3 + b_3 x_2 + b_3 x_3 + b_3$

という回帰モデルを仮定する(ただし、 は誤差項).

6.2 重回帰分析による得点の要因の解析

各球団の全試合の得点を目的変数とし、その他の変数 を説明変数として変数選択 (ステップワイズ法)を行い、 残った変数で重回帰分析を行った (表 2).

表 2 重回帰分析の結果

	E 4: 4	. 1	D.店	
	Estimate	t value	P 値	
(Intercept)	- 2.072	- 5.928	9.93×10^{-9}	***
ホーム	0.508	2.816	0.005	**
単打	0.419	12.409	2.48×10^{-16}	***
二塁打	0.676	9.362	2.11×10^{-16}	***
三塁打	1.081	5.255	3.13×10^{-7}	***
本塁打	1.440	15.984	2.94×10^{-16}	***
盗塁	- 0.252	- 1.814	0.070	
犠打飛	0.351	4.144	4.64×10^{-5}	***
四死球	0.270	5.568	6.54×10^{-8}	***
失策	0.190	1.484	0.139	
継投投手数	- 0.156	- 2.534	0.011	*

この結果では、単打 1 本につき 0.41 点、本塁打 1 本につき 1.44 点ということになった.P 値はそれ程効いていないがホームでの試合につき 0.5 点が効いているということは後攻でのハンディは先攻と比べ 0.5 点分の価値があるという結果となった.

球団別では、巨人は本塁打での得点要因が高く、やはり空中戦を得意としているチームであった。反対に、中日は本塁打1本当たりの得点が少なくソロ本塁打が多いからだと思われる。これは、本塁打を狙える和田選手らの主軸に対してはピンチの場面では無理に勝負をしておらず、後続のバッターと勝負をしているからだと言える。 どの球団も、盗塁が得点に有効に効いていなかった。私の中では盗塁は有効な攻撃手段のひとつであると考えていたが、得点力とはあまり関係がなかった。1点を取りに行く際の攻撃の駆け引きの一つとして、盗塁は有効な手段であるが実際に得点に絡んでいるかと言えばそうではなかった。

7 おわりに

いくつかの分析法で解析してみた結果,長丁場のペナントレースでは投手力と打力のバランスがいかに大事なことか,ということを感じた.特に,先発投手や救援投手といった投手の総合力でどれだけ僅差の試合に勝てるかということが大事な要素に挙げられると感じた.やはり,野球は投手の占める割合が7~8割と言われているように,投手の占める比重がとても大きいと実感した.また,いかに本拠地の球場に適した野球を目指せるかというのも大きいと思った.

本研究を通じて分析・研究をすることの難しさを強く感じたが、研究を通じて知識を得たことや研究テーマを決めて取り組めたことは良かったと思う.

参考文献

- [1] 中村永友:多次元データ解析法. 共立出版, 2009
- [2] 舟尾暢男:The R Tips データ解析環境 R の基本技・グラフィックス活用集. オーム社, 2008
- [3] 中日スポーツ:(株) 中日新聞社