

# セイバーメトリクスを用いた野球チーム編成法

2014SS025 伊藤光希

指導教員：佐々木美裕

## 1 はじめに

近年、メジャーリーグ・ベースボールではセイバーメトリクスといわれる指標を用いた選手の能力分析を行っている。セイバーメトリクスとは、野球ライターであるビル・ジェームズによって1970年代に提唱されたもので、通称「野球統計学」と呼ばれている。書籍「マネー・ボール」では、セイバーメトリクスをどのようにチーム編成に用いているかについて具体例が詳述されている[1]。本研究では、日本のプロ野球選手にセイバーメトリクスを適用して「強い」チームの編成を行うことを試みる。ここで、「強い」チームとは勝率の高いチームのことを指すものとする。最終的に編成されたチームで、家庭用野球ゲーム「実況パワフルプロ野球」を用いた勝率のシミュレーションも行う。

## 2 セイバーメトリクス

セイバーメトリクスには、全部で約45種の指標がある。本研究では、その中から表1に示す4つの野手の指標と4つの投手の指標を用いる。この8つの指標を選んだ理由は、一般にこの8つの指標が主要な指標として用いられているからである。

表1: セイバーメトリクス一覧

野手		
指標	算出方法	測られる能力
OPS	出塁率 + 長打率	総合能力
IsoD	出塁率 - 打率	選球眼
IsoP	長打率 - 打率	純粋な長打力
P-S	(本塁打 × 盗塁 × 2) / (本塁打 + 盗塁)	打って走れる
投手		
指標	算出方法	測られる能力
WHIP	(被安打 + 与四球) / 投球回数	安定感
QS率	QS回数 / 登板試合数	試合を作る
K/BB	奪三振 / 与四球	制球力
LOB%	(被安打 + 与四死球 - 失点) / (被安打 + 与四死球 - 1.4 × 本塁打)	ピンチに強い

## 3 チームの指標の定義

セイバーメトリクスは、投手や野手の個人の能力を評価する指標であるが、これを用いて、チームの指標を次のように定義する。

チームに所属する野手の集合を  $I$  ( $|I| = n_1$ )、投手の集合を  $J$  ( $|J| = n_2$ ) とし、野手の指標の集合を  $F$ 、投手の指標の集合を  $H$  とする。また、野手  $i \in I$  の指標  $f \in F$  の値を

$b_{if}$ 、投手  $j \in J$  の指標  $h \in H$  の値を  $p_{jh}$  とする。このとき、このチームの指標  $B_f$  と  $P_h$  を

$$B_f = \frac{1}{n_1} \sum_{i \in I} b_{if} \quad (f \in F), \quad (1)$$

$$P_h = \frac{1}{n_2} \sum_{j \in J} p_{jh} \quad (h \in H) \quad (2)$$

と定義する。

## 4 重回帰分析による勝率の推定

式(1)(2)で求めたチームの指標を説明変数として勝率を推定する重回帰分析を行った結果について説明する。分析に用いたデータは、プロ野球2017年シーズンの各チーム全体の各記録(本塁打数、奪三振数など)から算出された指標値である[2]。

表1に示す8つの指標を説明変数とした場合、補正済  $R^2$  は0.89556と大きくなったものの、さらに、得られたパラメータの符号が直感的に不自然なものとなった。

そこで、野手と投手それぞれの指標のみで勝率の推定を行った。すなわち、野手の指標を用いた勝率  $v_f$  の推定式を

$$v_f = a + \sum_{f \in F} \alpha_f B_f \quad (3)$$

投手の指標を用いた勝率  $v_h$  の推定式を

$$v_h = b + \sum_{h \in H} \beta_h P_h \quad (4)$$

として、それぞれ重回帰分析を行った。その結果を表2に示す。

8つの指標を説明変数とする場合と比較すると補正済  $R^2$  はかなり小さくなったが、パラメータの符号はK/BB以外は直感的に自然である。K/BBの値は大きい方が投手の評価が高いことを意味するが、打者に打たせてアウトを取ることが勝率に貢献することも考えられるため、ここでパラメータの符号が負になることは著しく不自然な結果とは言えない。以上の考察から、表2の結果を用いて最適なチーム編成を考える。

表2: 重回帰分析結果

野手 ( $R^2 = 0.40927$ )			投手 ( $R^2 = 0.95230$ )		
指標 $f$	$\alpha_f$	$p$ 値	指標 $h$	$\beta_h$	$p$ 値
切片 (a)	-0.05791	0.89556	切片 (b)	-0.96612	0.40927
OPS	0.30294	0.86855	WHIP	-0.05444	0.87511
IsoD	2.70325	0.64675	QS率	0.00142	0.78793
IsoP	1.16413	0.73095	K/BB	-0.05407	0.56159
P-S	0.00149	0.48215	LOB%	0.02297	0.26201

## 5 野球チーム編成モデル

本節では、野手および投手の候補を所与として、勝率を最大化する野球チームを編成するモデルを定式化する。はじめに、3節で定義した記号に加え、以下の記号を定義する。

$K$ : 野手の役割 (ポジション) の集合

$M$ : 投手の役割 (先発・中継ぎ・抑え) の集合

$r_l$ : 役割  $l \in K \cup M$  の適性がある選手の必要人数

$$s_{il} = \begin{cases} 1: & \text{選手 } i \in I \cup J \text{ に役割 } l \in K \cup M \text{ の適性がある} \\ 0: & \text{上記以外} \end{cases}$$

次に、以下の変数を定義する。

$$x_{il} = \begin{cases} 1: & \text{選手 } i \in I \cup J \text{ を役割 } l \in K \cup M \text{ に割り当てる} \\ 0: & \text{上記以外} \end{cases}$$

ここで、 $I$  を野手の候補集合、 $J$  を投手の候補集合、 $I \cap J = \emptyset$  とすると、候補から野手と投手を選択する問題は、独立に定式化できる。はじめに、式 (1)(3) より、野手を選択する問題は次のように定式化できる。

$$\max. \sum_{i \in I} \sum_{f \in F} \sum_{l \in K} \alpha_f b_{if} x_{il} \quad (5)$$

$$\text{s.t.} \sum_{i \in I} s_{il} x_{il} = r_l \quad (l \in K) \quad (6)$$

$$\sum_{l \in K} s_{il} x_{il} \leq 1 \quad (i \in I) \quad (7)$$

$$x_{il} \in \{0, 1\} \quad (i \in I, l \in K) \quad (8)$$

目的 (5) は、予測勝率の最大化を表している。制約条件 (6) は、ポジション  $l \in K$  の適性のある野手を  $r_l$  人選択することを示している。制約条件 (7) は、各選手に割り当てられるポジションは高々 1 つであることを示している。制約条件 (8) は変数の 0-1 制約である。

同様に、式 (2)(4) より、投手を選択する問題は次のように定式化できる。

$$\max. \sum_{i \in J} \sum_{h \in H} \sum_{l \in M} \beta_h p_{ih} x_{il} \quad (9)$$

$$\text{s.t.} \sum_{i \in J} s_{il} x_{il} = r_l \quad (l \in M) \quad (10)$$

$$\sum_{l \in M} s_{il} x_{il} \leq 1 \quad (i \in J) \quad (11)$$

$$x_{il} \in \{0, 1\} \quad (i \in J, l \in M) \quad (12)$$

目的 (9) は、予測勝率の最大化を表している。制約条件 (10) は、役割  $l \in M$  の適性のある投手を  $r_l$  人選択することを示している。制約条件 (11) は、各投手に割り当てられる役割は高々 1 つであることを示している。制約条件 (12) は変数の 0-1 制約である。

## 6 計算実験

2017 年シーズンに年間を通して活躍した選手を候補選手とするために、規定打席 (443 打席) に達した野手、規定

投球回 (143 回) に達した投手を候補選手とする。候補となった野手は 46 人、投手は 20 人である。この候補選手の集合から 16 人の野手と 12 人の投手を選択する。選択する選手の人数に関しては日本代表チームの構成メンバーに準じた構成とした。Gurobi を用いて最適化を行った結果、 $v_f = 0.55281$ ,  $v_h = 0.60670$  となるチームを編成することができた。

## 7 勝率シミュレーション結果

家庭用野球ゲーム「実況パワフルプロ野球」シリーズでは、実在するプロ野球選手の昨シーズンの成績 (今回の研究では 2017 年度成績) に基づき各選手の能力値が自動的に設定される。さらに、各チームの選手およびポジションや打順などを任意に設定し、対戦をシミュレーションできるゲームモードが存在する。これを用いて、プロ野球 1 シーズンを 10 年間シミュレーションし、最適編成したチームの勝率を求めた。対戦する 5 チームについては、2 つのパターンを設定した。5 チームとも読売ジャイアンツ (2017 年シーズンの勝率が .500 に近かったチーム) と仮定した場合であり、2 つめは、5 チームとも福岡ソフトバンクホークス (2017 シーズンで最も勝率が高かったチーム) と仮定した場合である。最適編成チームの順位や勝率などの結果を表 3 に示す。

表 3: 最適編成チームの結果

対戦 5 チーム	最適編成チームの結果			
	順位	勝	負	分 勝率
読売ジャイアンツ	1	989	419	22 .70241
福岡ソフトバンク	1	838	575	17 .59306

2017 年シーズンにセ・リーグで優勝した広島東洋カープの勝率が .633、パ・リーグで優勝した福岡ソフトバンクホークスの勝率が .657 であることから、勝率の高いチームが編成できたと考えられる [2]。

## 8 おわりに

提案するモデルで編成したチームは、予測勝率も高く、シミュレーションの結果でも実際のプロ野球界の強いチームよりも強かった。

本研究の最適化モデルは役割別に選手を選択するのみであったが、野手のスターティングメンバーや打順、投手の先発ローテーションなどについても考慮できるような最適化モデルを作成できると良い。

## 参考文献

- [1] 中山宥: 「マネーボール ～奇跡のチームをつくった男～」, ランダムハウス講談社, 東京, 2004.
- [2] 2017 年度版 データで楽しむプロ野球 <http://baseballdata.jp>