

# プロ野球選手のチームへの貢献度に関する統計的解析

2007MI142 水野桂介

指導教員：木村美善

## 1 はじめに

私はプロ野球選手がチームにどれほど貢献しているかに興味を持ち、選手のデータからそれを知りたいと考えた。私が昔から好きだった阪神タイガースについて解析したく、12球団の2010年度データを取り解析を行うところから始めた。その結果などから、球団の主要打者に興味を持ち、解析を進めた。そこで、セイバーメトリクスという新しい野球の考え方の指標を用いることにした。セイバーメトリクスは昨今では有名になってきているが、私が行った統計的解析にどのように関わることができるのかを知りたいと考えた。分析結果をもとに主力選手の貢献順位をつけ、セイバーメトリクスにおける順位と比較していく。

## 2 データについて

主成分分析に用いたデータについては、[2]から、2010年度の12球団の投手、打者成績より、チームの成績に関わる打者12項目、投手12項目の全24項目を取り上げた。クラスター分析は主成分分析の結果を用いた。ロジスティック回帰分析に用いたデータについては、[3]から、阪神のチーム結果と打者7人の2010年度の成績より、 $y =$  勝敗、 $x_1 =$  打数、 $x_2 =$  得点、 $x_3 =$  安打、 $x_4 =$  打点、 $x_5 =$  四死球、 $x_6 =$  犠打、 $x_7 =$  盗塁、 $x_8 =$  本塁打の8項目の記録を用いて選手がどの項目において勝利に貢献しているかを調べた。

## 3 12球団の分析

### 3.1 主成分分析

主成分分析は、外的基準のない多次元データ解析法において最も基本的な手法である。本研究では、2010年度の12球団のデータ(上記)をもとに、解析を行った。この結果よりクラスター分析を行うことで、より細かい特徴を出す。

### 3.2 クラスター分析

クラスター分析とは、異なるものの混ざりあっている対象を、それらの何らかの意味で定義された類似度を手がかりにして似たものを集め、いくつかの均質的なものの集落に分類する方法である。本研究では、上記した12球団の主成分分析のデータから第一主成分得点から第四主成分得点を用いてウォード法で分類した。この結果より、阪神タイガースの特徴や同じ分類のチームとの違いを考察する。

### 3.3 分析結果

主成分分析の結果、累積寄与率が80パーセント以上になるところまで用いることにする。第一主成分寄与率(約40%)、第二主成分寄与率(約22%)、第三主成分寄与率(約16%)、第四主成分寄与率(約11%)となった。図1から阪神タイガースの番号は2で、第一主成分、第二主成とも

に負であり、図の左下に位置している。これより、阪神タイガースが2010年度では打撃を主として良い成績を収めたということがわかった。図2からクラスター分析は第1群:広島、横浜となった。これらのチームは他のチームより秀でたものがなく、全体的に成績の良くない。常に下位グループにいるであろうチームの集まり。第2群:中日、ヤクルト、日本ハム、楽天となった。これらのチームは投手面で他のチームより優れていて、年間の失点が少ない。しかし、打撃面ではあまり成績の良くないチームの集まり。第3群:ソフトバンクであった。このチームは打撃・投手面ともに成績が良く、12球団で最もバランスの良いチームであると言える。第4群:阪神、巨人、西武、ロッテ、オリックスとなった。これらのチームは打撃面で他のチームより優れていて、年間の得点が多い。しかし、投手面ではあまり成績の良くないチームの集まり。

### 3.4 考察

阪神タイガースは第一主成分、第二主成分ともに負であり、図1の左下の2にである。特に第二主成分は12球団で最も負の値である。これより、阪神タイガースが2010年度では打撃を主として良い成績を収めたということがわかった。阪神は第4群に分類され、打撃の良いチームであるとされた。第4群の中でも「打率」、「得点」、「安打」が最も良いことから、本塁打での一発攻勢ではなくヒットを積み重ねて点をとるチームであるとわかった。また、同じセ・リーグである巨人と比べてみると、阪神は出塁することに、巨人は本塁打や長打を打つことに長けている。同じ群、リーグであるにも関わらずチームのタイプが違うことがわかった。

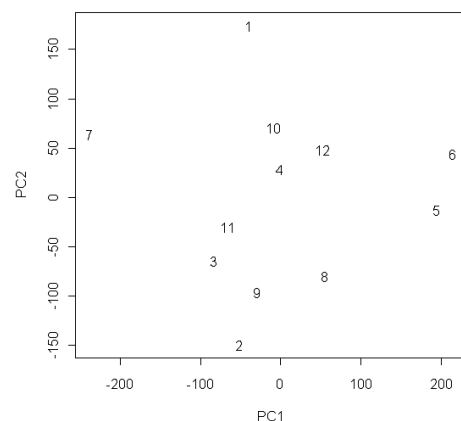


図1 第一主成分と第二主成分の散布図

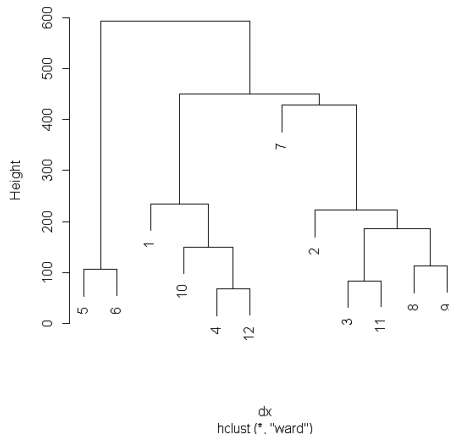


図 2 2010 年度 12 球団のデンドログラム (ワード法)

## 4 選手についての解析

### 4.1 ロジスティック回帰分析

確率のように範囲  $[0, 1]$  の値をとる目的変数を説明変数群  $(x_1, x_2, \dots, x_p)$  で説明するとき、通常の重回帰分析では説明変数群の合成変数の取りうる値が  $[0, 1]$  の外にでてしまい適切でないなので、説明変数群の合成変数をロジスティック関数にすることで、その値域の範囲を  $[0, 1]$  に収めることができる。本研究では、阪神の主要打者 7 人 (規定打席以上もしくは 144 試合出場) の全打席のデータを取り、分析を行った。主要選手がどの項目において貢献し、それがチームの結果に影響しているかを説明する。

### 4.2 分析結果からの考察

2010 年度の 7 人の主力選手を分析した結果、各々の選手が各々の分野でチームに貢献していることがわかった。中でもマートン選手はほとんどの分野で勝利に貢献している、2010 年度の阪神タイガースの中核をなす選手であることがわかった。逆に金本選手は出場機会が多かったものの、チームに貢献する結果を残すことが明らかにできていないことがわかった。ここであえて順位をつけるとするならば、1 位マートン選手、2 位鳥谷選手、3 位ブラゼル選手、4 位新井選手、5 位城島選手、6 位平野選手、7 位金本選手となった。これは各々の最適モデルでみた結果、チームの勝率が良かった順番でつけた。

### 4.3 セイバーメトリクス

野球には打率や打点、防御率など多くの価値基準・指標基準が存在するが、統計的視点から選手の評価方法や戦略を考え、分析する手法全体をセイバーメトリクスという。本研究では、 $OPS$ 、 $ISO(ISO_P)$ 、 $SPD$ 、 $BABIP$ 、 $wRC$ 、 $wRAA$ 、 $wOBA$  の数値を用いて、選手に順位をつけていく。

### 4.4 計算結果からの考察

表 1 から順位をポイントにおいて考えると、1 位はマートン選手、2 位はブラゼル選手、3 位新井選手、4 位鳥谷選

手、5 位城島選手、6 位平野選手、7 位金本選手となった。これは打者として優秀かどうかを表している。したがって、阪神タイガースの主要選手の中で打者として最も優秀であるのは、マートン選手であった。また逆に打者として優秀でないのは、金本選手であった。

表 1 セイバーメトリクスの計算結果

| $OP$  | $IS$  | $SP$ | $BA$  | $wRC$ | $wRA$ | $wO$  |
|-------|-------|------|-------|-------|-------|-------|
| 0.894 | 0.15  | 4.2  | 0.378 | 119   | 43.8  | 0.4   |
| 0.848 | 0.174 | 5.7  | 0.336 | 104.5 | 31.2  | 0.38  |
| 0.859 | 0.174 | 3.5  | 0.348 | 104   | 31.9  | 0.382 |
| 0.859 | 0.204 | 2.9  | 0.298 | 97.8  | 30    | 0.382 |
| 0.902 | 0.277 | 1.6  | 0.335 | 100.3 | 32.6  | 0.387 |
| 0.82  | 0.071 | 4.2  | 0.393 | 97.8  | 10.4  | 0.345 |
| 0.731 | 0.17  | 1.4  | 0.28  | 100.3 | 1.5   | 0.329 |

## 5 分析結果からの選手の考察

### 5.1 順位の違いについて

分析とセイバーメトリクスの順位が同じなのが 1 位のマートン選手、5 位の城島選手、6 位の平野選手、7 位の金本選手であった。この選手達は打者としての能力通りにチームに貢献していたと考えられる。鳥谷選手に関しては、打者としての能力以上にチームに貢献している。理由はチームの中で足を使って貢献している人が鳥谷選手以外ではないからだと考えられる。分析の結果を見ても明らかで、チームにとって欠かすことのできない存在だとわかった。ブラゼル選手は打点や本塁打といった、点を取る能力に関してはチームの中で絶対必要な 4 番打者と言える。

## 6 おわりに

本研究で阪神タイガースについて分析した結果、選手がそれぞれの持ち味を生かしてチームに貢献していることがわかった。また、セイバーメトリクスを用いたことで、能力以上にチームに貢献する選手がいるということがわかった。その選手、本研究で言えば鳥谷選手が、ある意味最もチームに貢献したと言ってもいいのではないかと感じた。研究結果から言えば、マートン選手がチームに最も貢献しているが、私は鳥谷選手も同等にチームに貢献した選手だと思った。チームに必要な選手とは、能力の高い選手はもちろんであるが能力以上に貢献できる選手、つまりチームの勝利を一番に考えることのできる選手だと感じた。

## 参考文献

- [1] 丹後俊郎・山岡和枝・高木晴良:統計ライブラリー「ロジスティック回帰分析」— SAS を利用した統計解析の実例—新日本印刷・渡辺製本出版、東京、1996.
- [2] プロ野球データ Freak: <http://baseball-data.com/>
- [3] Sponichi Annex 2010 年プロ野球 阪神データ: <http://www.sponichi.co.jp/baseball/npb/2010/team/tigers/data.html>