

阪神タイガースの選手に関する統計的分析

2017SS020 石崎幸志

指導教員：阿部俊弘

1 はじめに

野球は犠打や盗塁など小技を効かせて1点を地道に稼ぐチームや小技をせず長打を狙い一気に得点するチーム、守備で守りぬくチームなど様々である。小さい頃から応援している阪神タイガースの攻め方や守り方などを分析していく。

2 データについて

データは、阪神タイガースの2013年から2019年までの1年間の選手データである[1, 2]。試合は、交流戦を含めたシーズンを通しての試合のみで、オープン戦やクライマックスシリーズ、日本シリーズの試合データは使用しない。今回活用するデータは、規定打席数1/2以上の野手の単打数、二塁打数、三塁打数、本塁打数、盗塁数、四球数、死球数、三振数、犠打数、併殺打数、失策数、守備率、打点数、得点数、規定投球数1/3の投手の被安打率(被本塁打を含まない)、被本塁打率、与四球率、与死球率、奪三振率、暴投率、防御率を使用する。

3 分析方法

統計ソフトRを用いて、データを標準化し重回帰分析、主成分分析、クラスター分析を行った。重回帰分析は、勝敗に関わる打点、得点、防御率を目的変数として、AIC(赤池情報量規準)による変数減少法で変数選択を行った。重回帰分析で残った変数に絞り、主成分分析は、相関係数行列を用いて分析を行い、クラスター分析は、ユークリッド距離によるワード法を用いて分析を行った[3, 4]。

4 重回帰分析

初めに打点を目的変数とし、変数選択で残った x_1 :単打数、 x_2 :二塁打数、 x_3 :三塁打数、 x_4 :本塁打数、 x_5 :盗塁数、 x_{10} :併殺打数を説明変数として、重回帰分析を行う[3, 4]。残差分析の結果、正規性もあり、外れ値も無くデータに誤りがないことが分かった。また、変数選択後にVIFを確認したところ、10未満で多重共線性は生じていない。

表1 打点を目的変数とした重回帰分析結果

打点	回帰係数	標準偏差	t値	p値
定数項	-0.001	0.037	-0.030	0.976
x_1	0.194	0.066	2.931	0.005
x_2	0.207	0.071	2.938	0.005
x_3	0.074	0.044	1.662	0.102
x_4	0.649	0.053	12.280	0.000
x_5	-0.123	0.045	-2.715	0.009
x_{10}	0.083	0.059	1.410	0.164

阪神タイガース野手データの打点における決定係数および自由度調整済決定係数は、 $R^2 = 0.925$, $R^{*2} = 0.917$ となる。よって、阪神タイガース野手の打点の約91%が単打、二塁打、三塁打、本塁打、盗塁、併殺打によって説明されると解釈することができる。

次に得点を目的変数とし、変数選択で残った x_1 :単打数、 x_3 :三塁打数、 x_4 :本塁打数、 x_5 :盗塁数、 x_6 :四球数、 x_8 :三振数、 x_9 :犠打数、 x_{12} :守備率を説明変数として、重回帰分析を行う[3, 4]。残差分析の結果、正規性もあり、外れ値も無くデータに誤りがないことが分かった。また、変数選択後にVIFを確認したところ、10未満で多重共線性は生じていない。

表2 得点を目的変数とした重回帰分析結果

得点	回帰係数	標準偏差	t値	p値
定数項	-0.001	0.050	-0.018	0.986
x_1	0.368	0.073	5.040	0.000
x_3	0.212	0.060	3.528	0.001
x_4	0.229	0.082	2.801	0.007
x_5	0.140	0.062	2.242	0.029
x_6	0.320	0.073	4.401	0.000
x_8	0.126	0.076	1.663	0.102
x_9	0.143	0.064	2.228	0.030
x_{12}	-0.086	0.054	-1.606	0.114

阪神タイガース野手データの得点における決定係数および自由度調整済決定係数は、 $R^2 = 0.870$, $R^{*2} = 0.850$ となる。よって、阪神タイガース野手の得点の約85%が単打、三塁打、本塁打、盗塁、四球、三振、犠打、守備率によって説明されると解釈することができる。

次に防御率を目的変数とし、変数選択で残った x_1 :被安打率(被本塁打を含まない)、 x_2 :被本塁打率、 x_3 :与四球率、 x_6 :暴投率を説明変数として、重回帰分析を行う[3, 4]。残差分析の結果、正規性もあり、外れ値も無くデータに誤りがないことが分かった。また、変数選択後にVIFを確認したところ、10未満で多重共線性は生じていない。

表3 防御率を目的変数とした重回帰分析結果

防御率	回帰係数	標準偏差	t値	p値
定数項	-0.002	0.052	-0.030	0.976
x_1	0.558	0.054	10.260	0.000
x_2	0.491	0.054	9.068	0.000
x_3	0.348	0.053	6.538	0.000
x_6	0.102	0.053	1.935	0.057

阪神タイガース投手データの防御率における決定係数およ

び自由度調整済決定係数は、 $R^2 = 0.790, R^{*2} = 0.774$ となる。よって、阪神タイガース投手の失点の約 77% が、被安打率 (被本塁打を含まない)、被本塁打率、与四球率、与死球率、奪三振率、暴投率によって説明されると解釈することができる。

5 主成分分析

5.1 打点主成分分析

第 3 主成分で累積寄与率が 84.26% となり、第 3 主成分までの分析を行う。

第 1 主成分 (寄与率 46.26%)... 併殺打が多いが単打も長打も打てる選手を表す軸

第 2 主成分 (寄与率 27.47%)... 盗塁も三塁打も打てる足の速い選手を表す軸

第 3 主成分 (寄与率 10.53%)... 盗塁は出来ないが、三塁打を打てる足の速い選手を表す軸

5.2 得点主成分分析

第 4 主成分で累積寄与率が 80.35% となり、第 4 主成分までの分析を行う。

第 1 主成分 (寄与率 34.97%)... 三振が多いが出塁でき本塁打も打てる選手を表す軸

第 2 主成分 (寄与率 22.29%)... 足が速く小技が効く選手を表す軸

第 3 主成分 (寄与率 13.60%)... 守備が得意な選手を表す軸

第 4 主成分 (寄与率 9.49%)... 三塁打の打てる足の速い選手を表す軸

5.3 防御率主成分分析

第 3 主成分で累積寄与率が 82.46% となり、第 3 主成分までの分析を行う。

第 1 主成分 (寄与率 34.13%)... 安打を打たれてしまう選手を表す軸

第 2 主成分 (寄与率 29.03%)... 失投が多く被本塁打や暴投が多い選手を表す軸

第 3 主成分 (寄与率 19.30%)... 制球力が無く与四球や暴投が多い選手を表す軸

6 クラスタ分析

6.1 打点クラスタ分析

第 1 群：北條 (2018) ~ ナバー口 (2018)... 全体的に成績が低い数値の群

第 2 群：陽川 (2018) ~ 大和 (2013)... 全体的に成績が平均的の群

第 3 群：マートン (2013) ~ 福留 (2015)... 二塁打、本塁打、併殺打が多く、盗塁が少ない群

第 4 群：近本 (2019) ~ 鳥谷 (2015)... 単打、三塁打、盗塁が多い群

6.2 得点クラスタ分析

第 1 群：ゴメス (2014) ~ マートン (2014)... 本塁打、三振が多く、三塁打、犠打少ない群

第 2 群：近本 (2019) ~ 鳥谷 (2015)... 単打、三塁打、四球が多い群

第 3 群：植田 (2018) ~ 大和 (2013)... 盗塁、犠打が多く、本塁打、三振が少ない群

第 4 群：北條 (2018) ~ 福留 (2014)... 単打、盗塁が少ない群

6.3 防御率クラスタ分析

第 1 群：ガルシア (2019) ~ 能見 (2013)... 被安打率が高く、暴投率が少ない群

第 2 群：岩田 (2019) ~ 藤川 (2016)... 被本塁打率が高い群

第 3 群：渡辺亮 (2012) ~ 加藤康介 (2012)... 被本塁打率が低い群

第 4 群：能見 (2018) ~ ドリス (2017)... 暴投率が高く、被安打率、与四球率が低い群

第 5 群：藤浪 (2017) ~ 藤川 (2018)... 与四球率、暴投率が高く、被安打率が低い群

7 考察

二塁打や本塁打が多い長距離バッターが単打や四球が多い出塁率の高い選手をホームに返せている。また、盗塁出来る足の速い選手の打点が少ない事、守備率の高い選手の得点が少ない事が課題である。打点に大きく関わっている選手は、併殺打になってしまっても思いっきり振っていく選手である。得点に大きく関わっている選手は、三振になってしまっても単打、四球や本塁打といった出塁出来る選手である。防御率に一番関わっている選手は、予想通り被安打率、被本塁打率が多い選手であった。二番目、三番目に関わっている選手は、被本塁打率、暴投率が高い選手と与四球率、暴投率が高い選手であり、被本塁打、与四球や暴投といった失投してしまう選手が大きく防御率に関わっている。

8 終わりに

本研究を通して、今までとは違った視点で阪神タイガースの選手を見ることにより、試合観戦では分からない事が知れた。これからの試合観戦では様々な視点から見て、理解を深め楽しんでいきたい。

参考文献

- [1] 畑農鋭矢・水落正明 (2017) データ分析をマスターする 12 のレッスン
- [2] プロ野球データ Fresk <https://baseball-data.com/>
- [3] 豊田秀樹 (2012) 回帰分析入門ー R で学ぶ最新データ解析ー
- [4] 永田靖・棟近雅彦 (2013) 多変量解析法入門