

NPBにおける打者評価指標が年俵に与える影響分析

2019SC019 岩瀬蓮

指導教員：河野浩之

1 打者評価指標と年俵の関係

NPBのプロ野球選手の給料は年俵制であり、シーズンの成績の評価から次期年俵が決められる。毎年シーズン終了後、そのシーズンの成績に基づいて年俵が提示される [1]。提示を受けた選手は、その提示額に納得することができなければ保留、交渉もしくは退団という選択に迫られる。本研究では年俵に納得いかずに退団や交渉する選手やその年俵に納得のできないファンに向け、実際選手の成績の何が年俵に影響を与えているのか分析する。

2 年俵と成績に関する影響分析

選手成績の年俵への影響分析の先行研究を表 1 に示す。渡部ら [3] はセ・リーグの 3 球団の成績を用いて年俵への影響分析を行い、 p 値が 0.05 以下となったのは FA 権、外国人、試合数の 3 つであった。課題としては全体的に有意確率が高く偶然を否定できないこと、今後の予定としてはデータを用いて年俵の予測が挙げられた。

Cheng ら [4] の研究では中国プロ野球リーグのレギュラープレイヤーとその給与の年間パフォーマンスの関係の分析を行い、優れたパフォーマンスは年俵決定に利用していない結果となった。課題として野手と投手の 2 カテゴリーのみでの分別で分析を行ったため、これらのプレイヤーをさらに分けて特定の役割に分けることが挙げられた。

表 1 年俵への要因分析の先行研究

	分析手法	結果
渡部ら [3]	重回帰分析	外国人, FA 権, 試合数が年俵に影響を与える
Cheng ら [4]	ステップワイズ重回帰分析	優れた成績は年俵へ影響を与えていない

3 条件分別を用いた精度改善手法と予測

本研究の提案として、年俵への影響分析の際に全選手、規定打席 1/3 以上の選手、新人王有資格選手、プロ通算年 10 年以上選手 (FA 権取得選手)、タイトル獲得経験選手、の条件に分けて重回帰分析する。理由として成績のばらつきが小さくなり有意的な変数が多くなると考察したためである。また多重共線性の問題を考え、相関の強い変数同士は一つにまとめて分析を行う。変数の数も確保しなければならないため、特に強い相関の相関係数 0.9 以上 [5] の変数を一つにまとめる。また本研究では先行研究では分析外であり、表す成績が特殊かつ近年注目度も高いセイバーメトリクス指標 [2] を追加する。分析後、渡部ら [3] の研究の今後の予定である年俵の予測を行う。

本研究で使用する抽出先データサイト、スクレイピングツール、分析ツールをまとめたものを表 2 に示す。

表 2 使用サイト、ツール

サイト名, ツール名	特徴
野球データ Freak	扱うデータも多く、シンプル他ではない詳しいデータが掲載
Octoparse	簡単、複雑なサイトにも対応 IP ベースでアクセス拒否されにくい
Exploratory	データを探索する分析ツールで 操作しやすく分析の種類、幅が広い

全 12 球団の野手の 2021 年シーズンの個人成績を用いる。2021 年度の各成績等を説明変数 (x)46 種類、2022 年度の各選手の推定年俵情報を目的変数 (y) として分析を行う。

4 年俵と打者評価の分析と結果

4.1 説明変数同士の相関関係

今回の実験で使用するデータは 46 種であるが、説明変数同士で強く相関し多重共線性の問題が出てきてしまうため相関の強い変数は組み合わせた一つの指標として分析を行う。相関係数が 0.9 以上となった変数同士とそれらをまとめる変数を表したものを表 3 に示す。

表 3 本研究に用いる変数

まとめる変数	相関係数 0.9 以上の変数
OPS	OPS, GPA, NOI, RC27, XR27, TA, 長打率
RC	XR, 得点, 塁打, 安打
IsoP	本塁打, SecA
打席数	打数

4.2 条件毎での重回帰分析結果と考察

係数はどの程度年俵に影響を与えているかを表している。変数 t 値はそれぞれの成績が年俵に与える影響の大きさを表し、絶対値が大きいほど影響が強いことを意味する。 p 値は説明変数の係数の有意確率を表す。有意確率はその結果が偶然起こる確率を示しており、5% (p 値 0.05) 以下であることが有意的な数値である。以下表 4 に代表して全選手での分析結果の一部を示す。決定係数は約 0.75 である。

表 4 全選手での重回帰分析結果 (一部)

変数	係数	t 値	p 値
OPS	2,600,182	2.089	0.041
FA 権	11,347	3.780	0.0003
併殺打	-1,539	-3.295	0.002

有意的な変数は9つであった。規定打席1/3以上選手は5つ、新人王有資格者は4つ、プロ通算年10年以上選手は5つ、タイトル獲得経験者は3つであり有意確率の低い変数を多くすることができた。各条件で多く出た影響あり変数は「FA権」、影響なし変数は「二塁打」となった。「FA権」は8シーズン試合に出続けることで獲得できることから長く活躍することが高年俵へ繋がると考察できる。セイバーメトリクスでは「OPS」、「RC」が影響あり、「IsoD」、「IsoP」が影響なしの結果となった。また有意的な成績と得点との相関係数から得点との関わりがある成績が年俵に影響を与えていることが分かった

4.3 先行研究との結果の比較

渡部ら [3] の研究では有意的とされた指標は外国人、FA権、試合数である。本研究での三つの変数との比較を以下の表5に示す。三つの変数で係数は同じような傾向となった。特に外国人はどちらも p 値は大きく 0.05 を下回る結果となった。FA権は係数の値、 p 値ともに渡部らの研究より影響を与えているデータとなった。逆に試合数は渡部らの研究に比べて有意確率は高くなった。新人王有資格者での分析では試合数は年俵への影響ありと分析した。

表5 渡部ら [3] の研究 (上段) と本研究 (下段) の比較

外国人	p 値	FA権	p 値	試合数	p 値
12,872	0.002	4212	0.06	-124	0.045
9,529	0.015	11,347	0.0003	-63	0.093

5 有意的な変数による予測

5.1 有意的な変数による年俵の予測式

4章の分析で有意的な数値として出た変数(成績)を用いて、中日ドラゴンズの選手の年俵を予測する。2021年シーズン全打席全選手での重回帰分析を行い、その結果から作成した重回帰式を以下に示す。 x_1 から x_{15} までFA権、IsoD、IsoP、OPS、RC、タイトル、三塁打、二塁打、併殺打、外国人、打席数、打点、打率、敬遠、試合数の成績である。

$$\text{年俵予測} = (1.14 \times 10^4)x_1 + (-2.23 \times 10^6)x_2 + (-2.21 \times 10^6)x_3 + (2.25 \times 10^6)x_4 + \dots + (-5.23 \times 10)x_{15} + 172.9$$

5.2 中日ドラゴンズの選手の年俵の予測結果と考察

予測式を用いて各選手の年俵の予測とその評価をしていく。予測する選手は中日ドラゴンズの打者でかつ一軍出場登録経験のある選手を予測の対象として分析する。また2021、2022シーズンに入団した選手と2021シーズンに退団、引退した選手は分析対象外とする。2023年度の予測結果を表6に示す。制度の確認としてそれぞれ年度の誤差を算出し誤差平均値を求める。2022年の予測の結果の平均

誤差の値は約2,900万円、2023年の予測の結果の平均誤差の値は約2,600万円であった。また2022年の予測の決定係数は約0.75となり、2023年の予測の決定係数は約0.66となった。

表6 中日ドラゴンズの野手(一部)の年俵予測(/万)

選手名	予測年俵(2023)	実年俵(2023)
大島洋平	41,654	25,000
岡林勇希	3,888	4,000
高橋周平	4,786	6,750
大野将太	11,260	2,200

決定係数からある程度予測することができたと考えられるが、実際年俵は数百万や数千万単位であるため、誤差が約2500万円もあると年俵の参考にはならない。データと実際選手の置かれている状況が特殊である場合は予測が難しくなっている。例を挙げると、大野将太選手はFA権を行使して中日ドラゴンズに入団したが実際の成績はかなり低い、そのため実際年俵は低いのに対し、予測年俵はFA権所持の為年俵を高く予測している。

6 結び

重回帰分析の結果、各条件で全体的に年俵に影響を与えていると分析した変数は「FA権」である。得点との相関の数値から得点との関わりがある成績が年俵に影響を与えていることが分かった。年俵の予測では、決定係数0.7程の予測はできた。今後の予定としては、「FA権」の6シーズンへ検討されているため、価値の変化はあるのか。予測では有意的な変数を増加させることで予測精度向上の期待ができるが全ての変数を有意的にするのは難しいためより多くの有意的な変数での重回帰が必要である。

参考文献

- [1] 日本プロフェッショナル野球協約2022, 日本野球機構, pp.1-38,
- [2] 鳥越規央, “統計学が見つけた野球の心理-最先端のセイバーメトリクスが明らかにしたもの-,” 講談社, pp.1-230, 2022/3/20.
- [3] 渡辺静香, 朝日弓未, “日本プロ野球選手の年俵に影響を与える要因分析,” 日本計算機統計学会シンポジウム論文集, pp.1-2, 2016.
- [4] Chih-Cheng Chen, Chih-Jen Chuang, Tian-Shaing Kuo, Wen-Ing Chen, “Relationship between player performance and salary in a professional baseball league,” European Journal of Education Studies. Vol.4, No.3, pp.1-19, 2018.
- [5] 川瀬雅也, 松田史生 “相関と相関係数,” 公益社団法人, 日本生物工学会, 生物工学会誌, 第95巻8号, pp.1-4, 2017.