

OPS の計算式における出塁率と長打率の最適な価値配分 ～一見手抜きだが便利な指標～

芝浦工業大学 数理科学研究会 BV17057 西脇 友哉

平成 30 年 11 月 4 日

1 研究背景

昨今のプロ野球におけるセイバーメトリクス^{*1}の浸透に伴い、OPS という指標が定着しつつある。(OPS = 出塁率 + 長打率) という計算式の発想が、複数の変量を総合して主成分を求める主成分分析^{*2}の原理に似ていると思ひ、研究の着手に至った。

2 概要

主成分分析により出塁率と長打率を総合して評価する。「OPS」「NOI」「GPA」^{*3}「第 1 主成分得点」「第 2 主成分得点」の 5 項目それぞれについてチームの 1 試合平均得点との相関係数を計算し、どの指標が優れているかについて考察する。

3 OPS とは

OPS は“On-base Plus Slugging”の略語で、出塁率と長打率を同時に評価出来る指標である。例年のリーグ平均は NPB, MLB 共に .700 程度で、数値が高い程優れた打者といえる。OPS が提唱された背景には打撃 3 部門^{*4}の指標としての欠陥があり、特に打率の短所を補う評価基準として認知度が高まっている。

他の指標には無い長所として、出塁率と長打率をそのまま足しただけという非常にシンプルな構造でありながら得点との連動性が高い点が挙げられる。

4 主結果

NPB の公式 HP から '11~'18 のペナントレースにおけるチームの出塁率及び長打率を引用し、得られた主成分から打撃指標の有用性を考察した。

4.1 主成分の計算

x_1 : 出塁率, x_2 : 長打率とし、ラグランジュの未定乗数法を用いて

$$Q(w_1, w_2) = \text{Var}(x_1) \cdot w_1^2 + \text{Var}(x_2) \cdot w_2^2 + 2w_1w_2\text{Cov}(x_1, x_2)$$
$$F(w_1, w_2) = Q(w_1, w_2) - \lambda(w_1^2 + w_2^2 - 1)$$

を $w_1^2 + w_2^2 = 1$ という条件の下で最大化する最適化問題を解き、主成分を求めた。データより $\text{Var}(x_1) = 0.000185$, $\text{Var}(x_2) = 0.000822$, $\text{Cov}(x_1, x_2) = 0.000291$ を Q, F の式に代入することで、第 1 主成分 y_1 , 及び第 2 主成分 y_2 について

$$y_1 = 0.36176x_1 + 0.93227x_2$$

$$y_2 = -0.93227x_1 + 0.36176x_2$$

を得た。出塁率と長打率を総合化した指標であるのは y_1 の方だが、 y_2 からは「出塁率と長打率のどちらに強みを持つチームであるか」を調べることが出来る。

4.2 考察 (主成分得点の有用性)

OPS と 1 試合平均得点の相関係数は $r_A = 0.968$, 第 1 主成分と 1 試合平均得点の相関係数は $r_B = 0.954$ であった。それぞれの母相関係数 ρ_A, ρ_B の差について両側検定を行ったところ有意水準 $\alpha = 0.05$ で帰無仮説 $H_0: \rho_A = \rho_B$ が受容されたため、 r_A, r_B に統計的な有意差は認められなかった。しかし、第 1 主成分に複雑な係数が付いた事を考慮した結果、OPS の方が指標としての有用性に優れていると判断した。

5 OPS を用いた勝率予測への展開

OPS の利便性の高さを生かすためにピタゴラス勝率^{*5}を用いて得点数と失点数から勝率を予測することを試みた。

まず判別分析の結果から、優勝出来るか否かの境界を年間勝率 .566 と定めた。更に OPS を説明変数、年間総得点数を目的変数として単回帰分析を行い、勝率の変動を OPS に帰着させた結果、以下の結論に至った。

NPB 平均程度の総失点数 (=551) のチームが優勝するには目安として 629 以上の総得点が必要で、これを達成するためにはチーム全体の OPS が .733 以上であることが望ましい。

今後の課題

データを収集する過程で、チーム単位では OPS が .700 を下回る例が大多数を占めることが分かった。これは規定打席未到達者^{*6}による影響が含まれていることが原因と考えられるが、今後は WAR^{*7}を用いてデータの対象を規定打席到達者に絞って考察を図りたい。

参考文献

- [1] NPB. jp 日本野球機構, <http://npb.jp/>, 最終アクセス: 2018.10.14
- [2] 小西貞則, 多変量解析入門—線形から非線形へ, 岩波書店, 2010
- [3] 下川敏雄, 実践のための基礎統計学, 講談社, 2016

*1 1970 年代に B.James より提唱された統計学的な野球の分析手法。

*2 主成分分析の手法については資料を参照されたい。

*3 $\text{NOI} = (\text{出塁率} + \text{長打率}/3) \cdot 1000$, $\text{GPA} = (1.8 \cdot \text{出塁率} + \text{長打率})/4$

*4 一般に打率, 打点, 本塁打を指す。

*5 $\text{ピタゴラス勝率} = (\text{総得点})^2 / \{(\text{総得点})^2 + (\text{総失点})^2\}$

*6 規定打席: 所属球団の試合数 $\times 3.1$ (小数点以下四捨五入)

*7 WAR: 「その守備位置の控え選手に比べてどれだけ勝利数の上積みへ貢献したか」を表すセイバーメトリクス指標。