

複数年の成績を用いた日本プロ野球における年俸予測式の提案

Proposal for a Formula of Forecasting Annual Salaries in Japanese Professional Baseball considering Multi-Year Results of Players

2041098 橋本 黄輝
Kouki HASHIMOTO

指導教員 秋葉 知昭

In Japanese professional baseball, the weak point of players in contract negotiations has become a problem. Therefore, I will study to proposal of an annual salary forecasting formula. The difference from previous study is that the formula is based not on a single year's performance but on multiple years. This improved the forecasting and allowed us to confirm the evaluation of the continuity of performance in the actual contract negotiation. The forecasting formula using the simple average, which gave the best results, improved the mean absolute error by 9.05 million yen as the previous study.

1. 緒言

日本プロ野球(以下NPBとする)では、年俸交渉における球団に対する選手の立場の弱さが課題である。年俸予測式の提案により、選手が市場における自らの金銭的価値を容易に掴むことを可能にし、契約交渉における選手の立場向上を目指す。

2. 先行研究

NPBにおける年俸予測の研究は、渡辺[1]などがある。渡辺[1]は、説明変数に打撃指標を主に適用した重回帰分析[2]により、予測式を構築している。しかし、前年の成績のみを予測に適用したため、前年以外の成績が反映されていないという課題があった。本研究では、実際の契約交渉において成績の継続性が重視されているという仮定を立て、選手の複数年の成績を考慮した予測式の作成並びに予測向上を目指す。

3. 作者年俸予測手法の提案

3.1 変数設定

本研究にて使用する変数について説明する。

表1 年俸予測式作成に用いる変数

変数名	変数の設定
量的変数 : x_i ($i = 1, \dots, 14$)	打率, 試合数, 打席数, 打数, 安打, 本塁打, 打点, 盗塁, 四死球, 三振, 犠打, 併殺打, 出塁率, 長打率 ※ x_i (西暦): 当該年度シーズンの打撃成績を示す。
質的変数 : x_{15}	契約形態を示す $x_{15} \in \{0, 1\}$ 単年契約なら0, 複数年契約なら1を入力する2値変数
a_i	重回帰分析における偏回帰係数
b	重回帰分析における定数項

3.2 使用データ

2018年度から2021年度までの4シーズンにおいて、規定打席の3分の1を最低1回満たした選手176名に関して、打撃指標データ[3]と契約年数、年

俸といった契約情報[4]を集めた。

3.3 単純平均を用いた年俸予測式

量的変数だけに単純平均を適用する。最大4年分の成績を利用するが、例として2年間のデータを用いる予測式の作成について説明する。目的変数に2022年度の年俸を、説明変数には2020年, 2021年度2年間の成績を単純平均した成績データと2021年度の契約形態を反映した質的変数を用いて、重回帰分析を実施する。変数選択には変数増減法を用いた。得られた偏回帰係数, 定数項をもとに、年俸予測式は各偏回帰係数に対して2021年度, 2022年度の成績を単純平均した成績データと2022年度の契約を反映させた質的変数によって導出される。年俸予測式から算出される2023年度の予測年俸と2023年度の実年俸を比較し、年俸予測式の精度を評価する。

3.4 加重平均を用いた年俸予測式

加重平均を適用する範囲と変数選択方法は3.3節と同様である。式(1)を2年間のデータを用いた次期予測年俸(z)とする。重みの導出は後述する。なお、重回帰分析において選択されない説明変数の偏回帰係数は0と考える。

$$z = b + \sum_{i=1}^{14} a_i \times \frac{(x_{i(2022)} \times \alpha) + (x_{i(2021)} \times \beta)}{\alpha + \beta} + a_{15} \times x_{15} \quad (1)$$

例として打率の重み α の導出について説明する。各選手の打率と年俸の値を数直線上にプロットし、最小二乗法による回帰直線を求める。

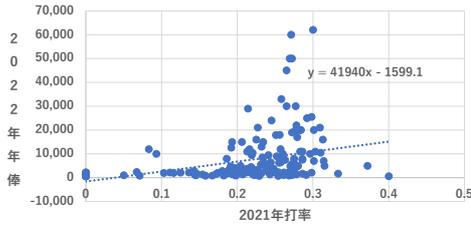


図1 2021年打率と2022年俸
 $y = 41940x - 1599.1$

(2)

$x = \text{打率} \times \alpha$ として、176名の年俸データを代入

$$\alpha = \frac{x}{\text{打率}} = \frac{y + 1599.1}{\text{打率} \times 41490}$$

(3)

式(3)より、選手個々で得られる打率に対する重み α を式(1)にそのまま適用する方法と、得られた選手個々の重みの平均を算出し、その値を式(1)における打率に対する重み α に適用する方法の2種類により年俸予測式を作成する。

3.5 k-means法による年俸予測式の再構築

k-means法を176名の2021年成績に対して適用する。クラスター数はエルボー法を用いて $k=3$ とした。分類ごとに年俸予測式を再度構築する。図2は各クラスタリングの平均値グラフである。

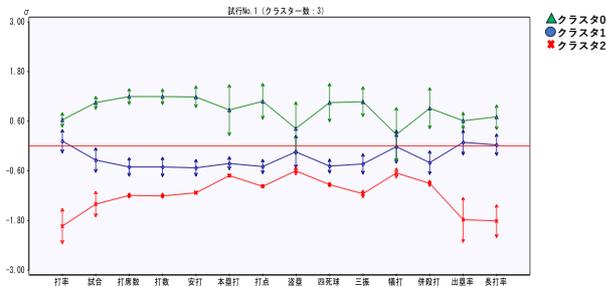


図2 各クラスタリングの平均値グラフ

4. 結果・考察

4.1 当該選手一括での年俸予測式の評価

MAE, MAPEにより、渡辺[1]と提案手法を評価する。単純平均を用いた年俸予測式において、渡辺[1]と比べ大きな改善が見られた。特に成績データを4年間用いた式ではMAEで約800万円、MAPEで約8.0%改善した。更に、適用する成績データの年数を長くするほど、両評価値共に向上したという特徴的な結果が読み取れた。平均した重みを用いた加重平均による年俸予測式についても同様の傾向が見られた。実際の契約更改においても前年だけでなく、

過去数年度の成績を含めた成績の継続性が評価されて年俸交渉がされているという仮定を、間接的にはあるが、証明した結果と言える。故に、提案手法における年俸予測式は、契約交渉における実情を反映した年俸予測式が確立出来たと言える。

表2 年俸予測式の評価

	渡辺[1]	単純平均	加重平均(個々)	加重平均(平均)
1年	MAE=3455.7 MAPE=53.31%			
2年		MAE=3438.1 MAPE=59.10%	MAE=3668.6 MAPE=68.37%	MAE=3403.8 MAPE=57.60%
3年		MAE=2802.6 MAPE=47.22%	MAE=3491.5 MAPE=75.98%	MAE=2861.9 MAPE=48.34%
4年		MAE=2656.6 MAPE=45.72%	MAE=3565.0 MAPE=76.81%	MAE=2860.6 MAPE=51.93%

4.2 クラスタリングを考慮した年俸予測式の評価

最も予測が上手くいったのは単純平均の成績データを4年分用いたのものであった。再構築前の渡辺[1]と比べるとMAEは905万円、MAPEは15.86%改善した。

表3 年俸予測式の評価

	渡辺[1]	単純平均	加重平均(個々)	加重平均(平均)
1年	MAE=3022.4 MAPE=42.59%			
2年		MAE=3001.1 MAPE=43.64%	MAE=3538.3 MAPE=62.92%	MAE=3122.5 MAPE=51.80%
3年		MAE=2630.4 MAPE=39.66%	MAE=3427.2 MAPE=75.28%	MAE=2680.9 MAPE=42.12%
4年		MAE=2550.7 MAPE=37.45%	MAE=3364.4 MAPE=61.47%	MAE=2943.5 MAPE=51.83%

5. 結言

本研究では、年俸予測の作成に際して、契約交渉における成績の継続性についての評価に注目し、適用する成績データの期間を長くすることで予測改善を図った。結果、単純平均を用いた予測式では、渡辺[1]と比べ、大きな改善が見られた。その上、成績の継続性が年俸に与える影響を間接的にはあるが、説明できた。他方、加重平均を用いた予測式については芳しい結果を得られなかった。今回は最小二乗法による近似直線を活用し、重みを導出したが、他にも次元数をあげた曲線や指数近似などの利用による重み導出の最適化が出来れば、予測改善が進むと考える。

文献

- [1] 渡辺静香：日本のプロ野球の年俸と成績の関係分析(2015)
- [2] 永田靖, 棟近雅彦：多変量解析入門(2001)
- [3] 野球データ Freak(2024年1月最終閲覧)
- [4] Baseball Schole(2024年1月最終閲覧)