

プロ野球投手の年俵と成績に関する統計的分析

2008MI253 塚原一翔

指導教員：木村美善

1 はじめに

野球選手の年俵は個人の成績によって決まっている。活躍した年度は上昇し、そうでなければ減俵になる。本研究では、球界を代表する6人ずつの先発投手と抑え投手、8人の若手投手の年俵について、その年度の活躍や勤続年数がどの程度、年俵に対して影響しているかについて分析した。また、個別の投手を比較して、どの球団の選手が待遇がいいのか、同じ球団でも球団側の評価方法の違いによって選手の年俵に違いは出るのかなども分析した。

2 データについて

選手は先発投手として、杉内俊哉、石川雅規、上原浩治、川上憲伸、黒田博樹、和田毅の6人、近年先発として結果を出している若手投手として、ダルビッシュ有、内海哲也、金子千尋、涌井秀章、成瀬善久、田中将大、久保康友、岸孝之の8人、抑え投手として、岩瀬仁紀、馬原孝浩、藤川球児、小林雅英、久保田智之、豊田清の6人を対象とした。先発投手、若手投手のデータは、「年俵」を目的変数とし、「登板試合数」、「勝数」、「負数」、「奪三振数」、「防御率」、「勤続年数」の6項目を説明変数として用いて分析し、抑え投手は、「奪三振数」の項目を「セーブ数」に変更して分析した。年度ごとの成績や年俵については、こちら、プロ野球人事部-[3]を参照した。

3 分析方法について

重回帰分析とロバスト重回帰分析を行った。([1],[2],[4]参照)

4 重回帰分析

4.1 先発投手

先発投手のデータをまとめたうえでダミー変数を加えて重回帰分析してみた。

表1 先発投手の分析結果

	回帰係数	t 値	p 値
Intercept	-5524.49	-1.745	0.08874
登板試合数	151.69	0.755	0.45490
勝利数	315.14	1.351	0.18440
奪三振数	17.51	0.570	0.57214
勤続年数	3285.39	14.927	2×10^{-16}
杉内俊哉	-4925.71	-2.588	0.01340
石川雅規	-8798.12	-2.932	0.00555
上原浩治	2289.54	1.274	0.20996
川上憲伸	-5905.79	-3.155	0.00304
黒田博樹	-14086.63	-6.739	4.33×10^{-8}

基準は和田投手である。項目ごとの回帰係数を比較すると、勤続年数の値が大きくなっている。これは毎年実績

を出して長く活躍することが難しいので、毎年結果を出すような選手は球団が大きく評価しているのだと考えられる。各選手の回帰係数を比較してみると、上原投手が正の値であるが、それ以外の投手は負の値になっている。特に黒田投手の負の値が非常に大きくなっている。これは黒田投手が所属していた広島という球団の財政問題が大きく関係している。広島は他の球団に比べて財政状況があまり良くなく、選手の年俵も抑えられがちである。過去に広島で年俵が2億円を超えた選手は前田智徳、金本知憲の2選手のみであり、黒田投手の年俵は破格である。なので他の投手と比較すると成績に見合った金額をもらっていないという結果になる。また、同じソフトバンクの杉内、和田両投手を比較すると、回帰係数にも約5000万円の差が付いており、杉内選手はあまり金額をもらっていないという結果になっている。これはソフトバンクは個人の成績を年度ごとに査定しており、過去の実績よりその年度の成績を評価する傾向があるからである。よって、怪我であまり活躍できなかった年度がある杉内投手より、毎年コンスタントに活躍する和田投手のほうが年俵が多くなっているのだと考えられる。

4.2 若手投手

若手投手のデータをまとめたうえでダミー変数を加えて重回帰分析してみた。

表2 若手投手の分析結果

	回帰係数	t 値	p 値
Intercept	1882.82	0.548	0.587727
登板試合数	-246.79	-1.553	0.131714
勝利数	133.60	0.494	0.625232
負け数	-554.59	-2.260	0.031803
奪三振数	42.41	1.617	0.116998
勤続年数	3467.75	8.403	3.86×10^{-9}
ダルビッシュ有	5931.83	2.653	0.013004
内海哲也	-3979.74	-1.616	0.117282
金子千尋	-4888.57	-1.903	0.067373
涌井秀章	1279.37	0.579	0.566906
成瀬善久	-10148.04	-4.077	0.000342
田中将大	3599.25	1.507	0.142912
久保康友	-1346.81	-0.584	0.563808

基準は岸投手となっている。変数では勤続年数の影響力が非常に大きくなっている。若手投手でありながら勤続年数が影響しているのは、新人時代からすでに結果を出して、毎年どんどん年俵が上がっていった選手が多いからではないかと考えられる。選手ごとを比較しても毎年良い成績を残し、年俵がはね上がっているダルビッシュ、

田中, 涌井の3投手は回帰係数も正の値になっており, これらが勤続年数の影響力が強い原因ではないかと思われる。久保, 内海, 金子, 成瀬の4投手は編回帰係数が負の値になっている。これらの投手は年俵がダウンした年度があったり, 成績に対する年俵が低い年度があったからであると考えられる。特に成瀬投手の値は非常に大きな負の値になっており, これは個別の分析のときの結果でも出ているように, 16勝した年度の年俵が低すぎるからだと考えられる。また, その年度以降も結果を残しているがそれほど年俵が大幅増加する年度がないからである。Cookの距離は全ての値が非常に小さく, 突出した影響力を持つ年度のデータがなく安定している。

5 ロバスト回帰分析

抑え投手6選手のデータをまとめてダミー変数を用いて重回帰分析してみると, 1年度だけ外れ値と思われるデータが出てきた。なので, ロバスト回帰分析によって詳しく分析してみる。

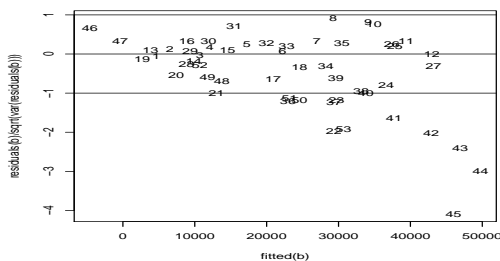


図1 抑え投手のロバスト回帰分析図

ロバスト回帰分析するとやはり1つだけ外れ値になるデータ(45)が出てきた。次に重回帰分析で, この値だけ外して分析すると同じような分析結果が得られた。これより, (45)のデータは外れ値であると考えられる。このデータは豊田投手が移籍によって年俵が大幅ダウンした年度のデータである。選手が移籍すると, 年俵が大幅に変化したりする場合もあり, 分析結果に大きな影響を与える場合があるが, ロバスト回帰分析を使うと, それらの年度だけには大きく影響されることなく分析することができる。

(41)から(45)までの数値の残差が非常に大きな負の値になっている。この選手は豊田投手であり, 彼は成績に対してあまり年俵をもらっていないことになる。(8)から(12)はゆるやかに下がっているが残差は正の値になっているので成績に対して多くの年俵をもらっていることになる。この選手は岩瀬投手であり, 球団に優遇されていると考えられる。

6 ロバスト回帰分析後の重回帰分析

ロバスト回帰分析によって外れ値が検出できたので, これを基に外れ値を除いてダミー変数を用いて重回帰分析を再び行くと表3のような結果になった。

表3 抑え投手の分析結果

	回帰係数	t 値	p 値
Intercept	-9668.93	-4.136	0.000161
勝利数	417.23	1.810	0.077257
セーブ数	334.19	7.320	4.44×10^{-9}
勤続年数	2865.42	11.535	9.52×10^{-15}
岩瀬仁紀	8385.46	4.033	0.000221
馬原孝浩	-708.46	-0.282	0.779171
藤川球児	-947.49	-0.427	0.671616
小林雅英	2909.62	1.248	0.218705
豊田清	-9598.90	-3.801	0.000450

基準は久保田投手である。項目ごとにみると, 勤続年数の値が非常に大きくなっている。これは抑え投手という勝敗に直結するポジションで, 1度の失敗で信頼がすごく落ちてしまうなかで, 毎年安定した成績を出し続けるのが難しいからであり, 長い年数結果を出し続けた選手は飛躍的に年俵が上昇するからである。重回帰分析の結果より, ロバスト回帰分析で豊田投手の残差が非常に大きな負の値になっているのは, 勤続年数の回帰係数の値が大きく, 年俵に大きく影響しているからであると考えられる。豊田投手は他の選手より5年くらい勤続年数が長い, 移籍などにより年俵が上昇していないので残差の値が大きな負の値になったと考えられる。また, 岩瀬投手の回帰係数の値が他の選手に比べて非常に大きく, 岩瀬投手が毎年良い結果を残しているからであると考えられる。

7 おわりに

実績のある先発投手はある程度の成績でも, 今までの実績が考慮されるのでそこまで目立った活躍がなくてもある程度の年俵をもらうことができるが, 若手投手は実績があまりないので, 結果を残せない年度があると減俵になったりしていた。しかし, 若手投手は年俵も低いので良い成績を残すと一気に跳ね上がったりすることもある。抑え投手は先発投手とは違い, 連投もあり, 精神的にも肉体的にも先発投手とは比較にならないくらい負担があると考えられる。なので年俵も上昇率も先発投手に比べて大きいように感じた。本研究では2000年代以前のデータがあまり見つからなかったため, 過去の選手については分析できなかった。そのようなデータがあれば現代の選手との比較もできたと思う。

参考文献

- [1] 神谷美紀:ロバスト回帰の研究, 南山大学数理情報学部数理科学科卒業論文, 2005
- [2] 中澤港:Rによる統計解析の基礎, REASON Education Japan, 2003.
- [3] こちら, プロ野球人事部:
<http://home.a07.itscom.net/kazoo/pro/pro.htm>.
- [4] Rで学ぶロバスト推定:
<http://www.slideshare.net/sfchaos/r-7773031>.