

プロ野球の勝利に影響を与える要因の 分析

國本葉月¹、小坂圭史²

要旨

本稿ではプロ野球において、打撃力、走塁力、投手力、守備力の4要素がいかに関シーズンの勝率に影響を与えているのかを、重回帰分析によって調査した。説明変数には、セイバーメトリクスで用いられる指標を使用した。重回帰分析の結果として、打撃力が最も影響を与えている一方で、走塁力の影響については有意性がないという結果が判明した。また、得られた結果をさらに分析することで、野手の方が投手よりも影響力があるということなど、いくつかの興味深い考察を行うことができた。

1 慶應義塾大学経済学部

2 慶應義塾大学経済学部

1.はじめに

野球というスポーツで勝利を取めるには相手チームより多くの点を取ることが必要である。一方でサッカーやバスケットボールと違い、野球は攻撃と守備が明確に分かれているスポーツである。そのため場面によって役割が明確に分かれる。攻撃の際に必要なのは打撃、そして出塁した場合には走塁というものが要求される。守りでは相手選手への投球、そして打球がインフィールドに飛んだ時には守備が必要になる。

このように野球の大半は打撃、走塁、投球、守備の4要素で構成されている。これらの力が高ければ高いほどリーグ戦においてより多くの勝ち星を挙げることができるが一方で、この4要素がどれだけ勝利に寄与していて、どれがより重要なのかは打率や防御率などの野球でよく見られる数値からは判断できない。本稿ではこの4要素を統計的に分析し、その要素が勝利に強く関係しているかを分析していく。そして、得られた分析結果からさらに考察を行う。

2.方法

2.1 データ収集

分析に用いるデータは2017-2021年のNPBにおける5シーズン。全60チーム分のデータを使用する。各チームのデータはNPBの公式サイト(<https://npb.jp>)から収集した。

2.2 使用する変数

チームの打撃力を評価する指標としてはwOBA (weighted On Base Average)を使用する。wOBAは安打や四球に対して得点価値に応じた重みをかけて計算した出塁率である(蛭川, 2019)。wOBAは打率よりも得点との相関が強いため採用した。wOBAは以下の数式によって求めた。

$$wOBA = \frac{0.7 \times (\text{四球} + \text{死球}) + 0.9 \times \text{単打} + 1.3 \times \text{二塁打} + 1.6 \times \text{三塁打} + 2.0 \times \text{本塁打}}{\text{打数} + \text{四球} + \text{死球} + \text{犠飛}}$$

チームの走塁力を求める指標にはSpd(Speed score)を採用した。Spdは足の速さを求める指標で、計算式は株式会社DELTAが運営する1.02 Essence of baseball(<https://1point02.jp/op/index.aspx>)に掲載されているものを使用した。

$$Spd = (A + B + C + D) \div 4$$

$$A(\text{盗塁成功率}) = \{(\text{盗塁} + 3) \div (\text{盗塁} + \text{盗塁死} + 7) - 0.4\} \times 20$$

$$B(\text{盗塁企図}) = \text{SQRT}\{(\text{盗塁} + \text{盗塁死}) \div (\text{単打} + \text{四球} + \text{死球})\} \div 0.07$$

$$C(\text{三塁打割合}) = \text{三塁打} \div (\text{打数} - \text{本塁打} - \text{三振}) \div 0.02 \times 10$$

$$D(\text{得点割合}) = \{(\text{得点} - \text{本塁打}) \div (\text{安打} + \text{四球} + \text{死球} - \text{本塁打} - 0.1)\} \div 0.04$$

ここからは守りの要素となる数値について説明する。投手力を測る指標としてFIP(Fielding Independent Pitching)を選択した。求め方は企業や研究者によって多少差があるが本稿では1.02 Essence of baseballに掲載されている数式を使用した。

$$FIP = \{13 \times \text{被本塁打} + 3 \times (\text{与四球} - \text{故意四球} + \text{与死球}) - 2 \times \text{奪三振}\} \div \text{投球回} + \text{定数}$$

$$\begin{aligned} \text{定数} = & \text{リーグ全体の} [\text{失点率} - \{13 \times \text{被本塁打} + 3 \times (\text{与四球} - \text{故意四球} + \text{与死球}) \\ & - 2 \times \text{奪三振} \div \text{投球回}] \end{aligned}$$

守備では DER(Defensive Efficiency Ratio)を使用。DER とはグラウンド上の打球の内どれだけアウトにできたかを調べる指標で、これによりそのシーズンの守備力を判断できる。守備に関しては他にも使われている指標があるが、NPB が公開しているデータから算出できるという面から DER を採用した。

$$\text{DER} = (\text{打席} - \text{安打} - \text{四球} - \text{死球} - \text{三振} - \text{失策}) \div (\text{打席} - \text{本塁打} - \text{四球} - \text{死球} - \text{三振})$$

以上の4つの指標を使用することでチームのデータを打撃力、走塁力、投手力、守備力に分けることができた。4指標のうち、wOBA、Spd、DER は大きいほど良い評価となる指標であり、FIP は小さいほど良い評価となる指標である。

2.3 分析方法

全5シーズン60チーム分のデータを集計後、wOBA、Spd、FIP、DERを算出。そして求めたwOBA、Spd、FIP、DERをそれぞれ標準化したものを説明変数とし、チームの勝率を従属変数とする重回帰モデルを最小二乗法によって推定。その偏回帰係数を比較して各要素の重要性についての検証を行う。

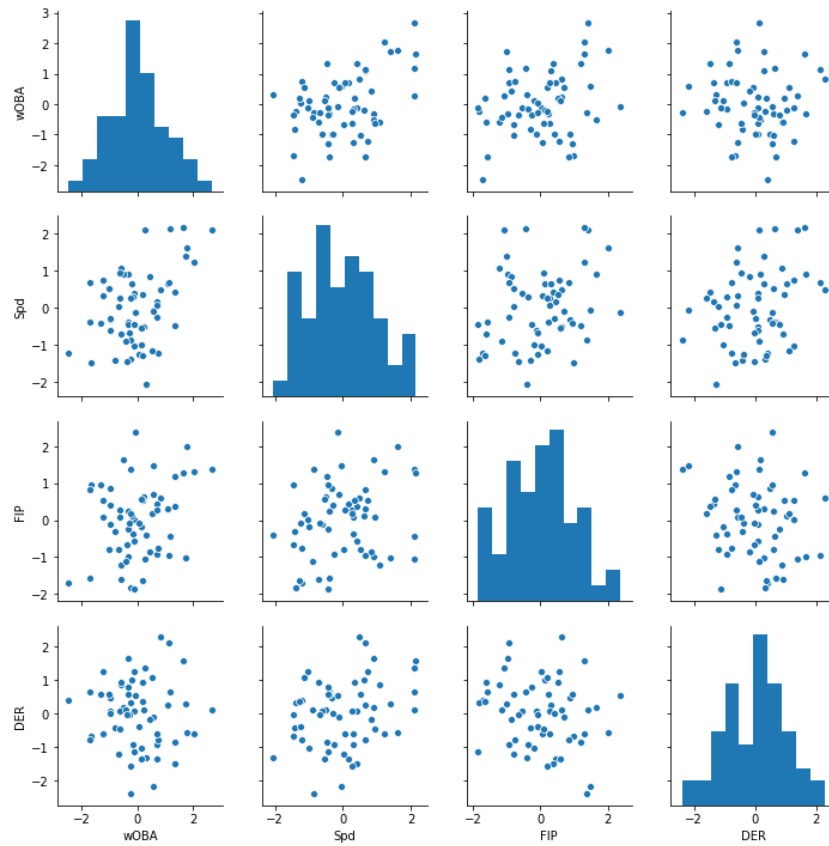
3.結果

はじめに、説明変数として用いた4指標について、それぞれの関連性について述べる。説明変数同士の相関係数は表1の通りであった。また、図1に標準化した説明変数の散布図を示す。結果として、2変数間ではそれほど高い相関関係は見られなかった。

表1：説明変数同士の相関係数

	wOBA	Spd	FIP	DER
wOBA	1	0.429761	0.263762	-0.04138
Spd	0.429761	1	0.240518	0.274469
FIP	0.263762	0.240518	1	-0.19223
DER	-0.04138	0.274469	-0.19223	1

図1：標準化した説明変数の散布図



次に、多変数間の線形関係について調べるために、VIF(Variance Inflation Factor)統計量の計算をおこなった。表2は、その結果を示したものである。全てのVIF統計量が基準値である10を大幅に下回っているため、多重共線性が存在する可能性は低いと言える。

表2：各説明変数のVIF統計量

説明変数	VIF 統計量
wOBA	1.294647
Spd	1.441857
FIP	1.172363
DER	1.194756

標準化したwOBA、Spd、FIP、DERを説明変数とし、勝率を従属変数とした重回帰分析の結果は表3の通りであった。wOBA(x_1)、Spd(x_2)、FIP(x_3)、DER(x_4)とすると、得られた重回帰式は、

$$y = 49.9833 + 4.7414x_1 + 0.4749x_2 - 1.7898x_3 + 2.1333x_4$$

となった。修正済決定係数は0.587となり、まずまずの当てはまりであった。また、p値は、wOBA(x_1)、FIP(x_3)、DER(x_4)に関しては、5%有意で説明力がある変数であるが、Spd(x_2)は説明力がない変数であると判断された。説明力があるとされた3つの指標のうち、勝率への寄与が最も大きかったのはwOBA(打撃)であり、次いでDER(守備)、FIP(投手)であった。

表3：3変数での重回帰分析結果

	係数	標準誤差	t	P-値	下限 97.5%	上限 97.5%
切片	49.9833	0.572	87.412	0.000	48.837	51.129
wOBA	4.7414	0.651	7.288	0.000	3.438	6.045
Spd	0.4749	0.687	0.692	0.492	-0.901	1.851
FIP	-1.7898	0.619	-2.891	0.005	-3.031	-0.549
DER	2.1333	0.625	3.413	0.001	0.881	3.386

続いて、説明力がないと判断された Spd を除く、3 指標のみでの重回帰分析を行なった。
4 変数の場合と同様に、標準化したデータを用いた。ただし、区別するため、wOBA(x'_1)、
FIP(x'_3)、DER(x'_4)とする。得られた重回帰式は、

$$y = 49.9833 + 4.9269x'_1 - 1.6944x'_3 + 2.2897x'_4$$

となった。修正済決定係数は 0.591 であった。その他の結果は表 4 の通りである。

表 4 : Spd を除く 3 変数での重回帰分析結果

	係数	標準誤差	t	P-値	下限 97.5%	上限 97.5%
切片	49.9833	0.569	87.822	0.000	48.843	51.123
wOBA	4.9269	0.590	8.350	0.000	3.745	6.109
FIP	-1.6944	0.601	-2.820	0.007	-2.898	-0.491
DER	2.2897	0.580	3.948	0.000	1.128	3.452

以下考察では、3 変数の分析結果の数値を用いることとする。

4.考察

まず、回帰式による勝率から実際の勝率を引いた勝率差分について、チームの順位毎に異なる傾向が見られたことについて述べる。表5は過去5年間(2017~2021)の順位別の平均勝率差分を、表6は同時期の1位と6位のチームについて、それぞれの勝率差分を示したものである。1位のチームの勝率差分はマイナスの値をとる傾向があり、6位のチームはプラスの値をとる傾向があるということがわかる。これは1位のチームは実力以上に勝率が高い傾向があり、6位のチームはその反対の傾向があることを示している。

表5：順位別の平均勝率差分(2017~2021)

順位	平均勝率差分
1位	-3.43
2位	-2.25
3位	-1.04
4位	2.18
5位	1.32
6位	3.23

表6：1位のチームと6位のチームの勝率差分(2017~2021)

1位のチーム

年度	チーム	勝率差分
2021	S	-1.5
2021	B	-4.1
2020	G	-2.1
2020	H	-7.3
2019	G	-6.4
2019	L	-2.4
2018	C	-7.2
2018	L	-5.2
2017	C	2.5
2017	H	-0.6

6位のチーム

年度	チーム	勝率差分
2021	DB	7.7
2021	L	-0.2
2020	S	3.8
2020	B	6.8
2019	S	3
2019	B	-1.4
2018	T	13.2
2018	E	7.4
2017	S	-5.2
2017	M	-2.8

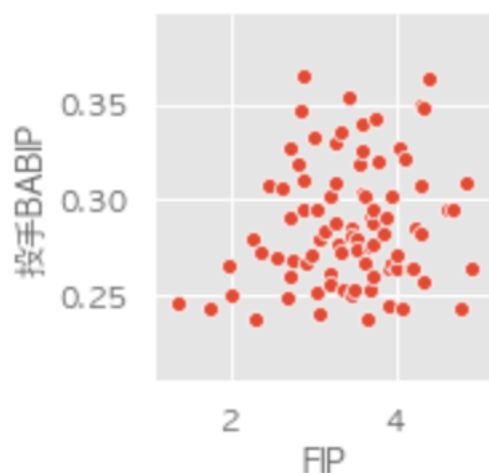
今回用いた指標の計算式の中には、得点圏打率のような、場面ごとの数値は使われていないため、そのチームのチャンスの場面での強さなどは考慮できていない。これを踏まえた上で、1位と6位で明らかに回帰直線からの振れの方向が違うことを考えると、勝ち負けに伴うチームのムードが、重要な場面での強さに関連しているのではないかと推測した。

次に、守備指標 DER と投手力指標 FIP の関連性についての分析を行った。FIP は守備の関与しない四死球、奪三振、被本塁打により計算される指標である。この指標が良いチームの投手陣はコントロールが良く、打たれにくいと言える。一般的に、そのようなチームの場合、インフィールドの打球は打ち損じたものが増え、アウトにできる可能性が高くなると考えられるだろう。つまり、グラウンド上の打球の内どれだけアウトにできたかを調べる指標である DER は高くなるはずだということである。しかしながら、DER と FIP の相関係数は-0.19223 であり、関連性があるとは言えない結果であった。この結果を受け、チーム単位での関連性だけでなく、投手一人一人についても同様の結果が得られるかをさらに調べた。ここでは、新しく投手 BABIP という指標を用いた。

$$\text{BABIP} = (\text{被安打} - \text{被本塁打}) \div (\text{打者} - \text{与四死球} - \text{奪三振} - \text{被本塁打})$$

この指標はグラウンド上に飛んだ打球のうち安打になった割合を示すものである。2021年のシーズンの投手データのうち、60 イニング以上投げた 86 投手に限定して、各投手の FIP と BABIP の相関関係を調べた。各投手の FIP と BABIP は共に、SPAIA に掲載されているものを用いた。結果として、相関係数は 0.1617 となり、各投手単位でも同様に関連性のなさが示された。図 2 に散布図を示す。

図 2 : FIP と投手 BABIP の散布図(2021 年シーズンの 86 投手)



この結果から、打球がグラウンド上に飛んだ場合、それがヒットになるかどうかは投手力とは別物であるということが言える。

最後に、守備力が投手力に依存しないという分析結果を前提とし、勝率への影響力について投手と野手の比較を行った。打撃力は野手のみが関わっており、投手力は投手のみが関わっている。一方で守備力は投手、野手どちらも関わるものである。そのため、重回帰式の守備 $FIP(x'_4)$ の係数を、投手分と野手分に分割しようと考えた。ここで、守備に関わった回数として、守備機会を用いた。守備機会は、

$$(\text{守備機会}) = (\text{刺殺数}) + (\text{補殺数}) + (\text{失策数})$$

によって計算される。2021年の12チームの守備機会は1イニングあたり4.203回。規定投球回に到達した投手のデータを参考にすると、投手の1イニングあたりの守備機会は0.229回となる。このことから、インフィールドの打球のうち投手が関わるのは5.46%であると分かった。ここから回帰式を野手、投手で分けると以下の式のように表せる。

$$y = 49.983 + (4.927x'_1 + 2.165x'_4) + (0.125x'_4 - 1.694x'_3)$$

左の括弧内が野手の影響力、右の括弧内が投手の影響力を表している。係数を見ると野手の影響力の方が強いと分かる。しかし試合中、野手は同時に8もしくはDHを含む9人が同時に出場しているのに対し、投手は1人しか出場できないという差がある。ここで左の括弧内を8で割り、野手1人あたりの影響力を調べると以下のようになる。

$$y_1 = 0.616x'_1 + 0.27x'_4$$

野手1人あたりの影響力と同じように、先発投手1人あたりの影響力を調べる。2021年の規定投球回に達した投手のイニング数から求めると、1人の先発投手は全体の12.3%を占めていると分かった。このことから先発投手1人あたりの影響力は以下の式になる。

$$y_2 = 0.015x'_4 - 0.208x'_3$$

それぞれのデータを標準化したことを考慮して係数を比較すると、野手1人の方が先発投手1人よりもシーズンにおける勝率に影響を与えていることが分かる。この結果を、セイバーメトリクスで使われるWAR(Wins Above Replacement)という指標の傾向と照らし合わせた。WARとは、控えレベルの選手に比べてどれだけ勝ち星を増やしたかということを示し、打撃、守備など、全ての要素を含んだ総合力として用いられる指標である。また、ポジションに関係なく比較できるという特徴を持つ。表7はDELTA社が公開している、過去2年間のWAR上位選手を順に並べたものである。WARは野手の方が高くなる傾向があり、その傾向は表7からも見て取ることができる。これは、先発投手1人の影響力より野手1人の方が影響力が大きいという今回の分析結果と通じていることが分かる。

表7：過去2年間のWAR上位選手(2020、2021)

2021年 WAR					
	野手	野手 WAR		投手	投手 WAR
1	鈴木誠也	8.7		山本由伸	8.6
2	森友哉	7		上沢直之	5.2
3	柳田悠岐	7		N・マルティネス	5.1
4	村上宗隆	6.8		伊藤大海	4.7
5	山田哲人	6.1		柳裕也	4.6

2020年 WAR					
	野手	野手 WAR		投手	投手 WAR
1	柳田悠岐	8.4		山本由伸	5
2	坂本勇人	5.7		大野雄大	5
3	近本光司	5.5		菅野智之	4.8
4	浅村栄斗	5.4		森下暢仁	4.2
5	鈴木誠也	5.3		西勇輝	3.7

5.終わりに

本論では、プロ野球の勝率に関する各要素の影響力について論じてきた。チーム全体で見ると打撃力が最も影響を与えている一方で、走塁力は影響を与えていないと言いう結果であった。また、投手に比べて野手の方が勝率への影響が大きいという結果も得られた。一方で、今回の分析では考慮することができなかった場面ごとの状況や、勝負強さのようなものも勝率に影響を与えている可能性があるため、より詳細なデータを追加して分析をする必要がある。

参考文献

蛭川皓平 (2019) 『セイバーメトリクス入門 脱常識で野球を科学する』 水曜社

WEB上の資料

NPB.jp 日本野球機構、<https://npb.jp>

1.02 - Essence of Baseball | DELTA Inc. <https://1point02.jp/op/index.aspx>

末木新(2017)「高校野球における試合の勝敗に影響を与える要因：投手力・打撃力・守備力の比較」 https://www.jstage.jst.go.jp/article/jjpehss/advpub/0/advpub_16083/_article/-char/ja/

スパイア【SPAIA】 | スポーツ×AI×データ解析 総合メディア <https://spaia.jp>