

「今打つ確率」を考える

1. 研究動機

近年プロ野球においては、セイバーメトリクスと呼ばれるデータを用いた選手のパフォーマンスの分析が盛んに行われており、様々な指標が現在知られている。ところが、過去のデータを分析しただけのものがほとんどで、そこから先はどうなるのかを推測するという指標は見受けられない。そこで、次の打席でヒットを打つ確率（以降「打つ確率」と表現）はどのようにして求められるのかについて研究してみた。

2. 研究方法

①対象とする選手：柳田悠岐選手（ソフトバンク）

ここ数年で最も活躍しているバッターの1人である。特にパ・リーグMVPを受賞した昨シーズンはほとんどの打撃成績においてリーグの上位に名を連ねており、まさに2020年のプロ野球界最強打者といえる存在であった。また、プロ入りから結果を残し続けており、継続的で安定的なデータを取ることが可能であったため、柳田選手を選んだ。

年度	試合	打席	打数	安打	単打	二塁打	三塁打	本塁打	塁打	得点	打点	打率	出塁率	長打率	得点圏
2015	138	605	502	182	116	31	1	34	317	110	99	0.363	0.469	0.631	0.413
2016	120	536	428	131	78	31	4	18	224	82	73	0.306	0.446	0.523	0.314
2017	130	551	448	139	77	30	1	31	264	95	99	0.310	0.426	0.589	0.379
2018	130	550	475	167	97	29	5	36	314	95	102	0.352	0.431	0.661	0.389
2019	38	157	128	37	23	6	1	7	66	17	23	0.289	0.420	0.516	0.333

表1. 2015～2019年の柳田選手の主な成績

②収集データの概要

2020年の全打席について打席結果を調べ、各打席の直近20打席の打率を求めた。また後述する「打つ確率」の計算式に基づき、各打席の相手投手の左右・被打率、得点圏の走者の有無、得点差状況（ビハインド/同点/リード）、さらに柳田選手の通算の得点圏走者あり・なし、対右投手・左投手、ビハインド・同点・リード時の打率とプロ入りからの通算打率を計算した。さらに球場別のヒットの出やすさに関する独自の指標^(※)を算出した。

(※) 2020年シーズンにおける各球場の単打、二塁打、三塁打、本塁打のパークファクターに、総安打数に占める単打/二塁打/三塁打/本塁打数の割合をそれぞれ掛け合わせ合計したものを各球場のヒットの出やすさの指標として用いた。

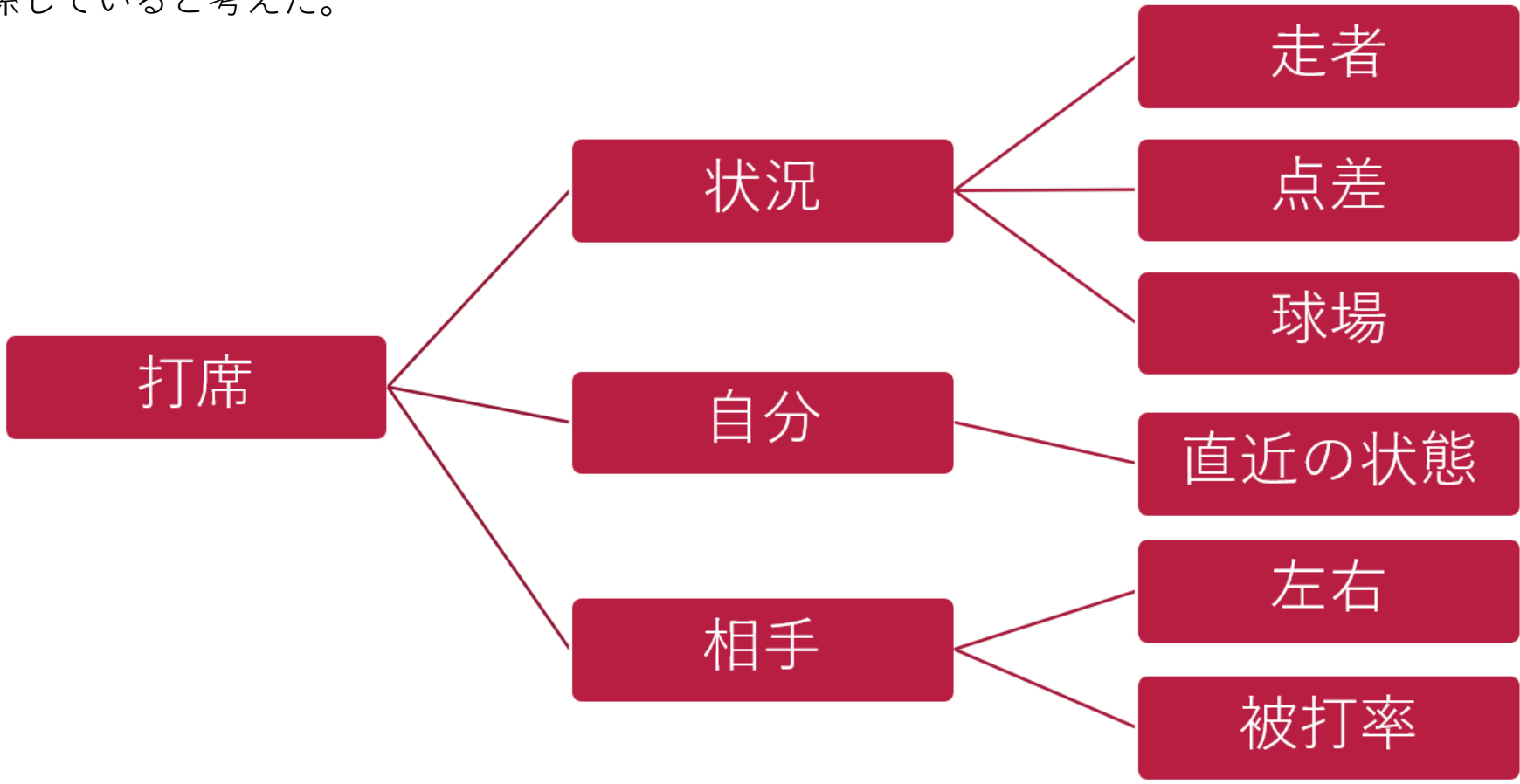
		2019	2018	2017	2016	2015	2014	2013	2012	2011	計	打率
得点圏	打数	33	113	116	121	126	145	78	49	2	783	0.342
	安打	11	44	44	38	52	48	22	9	0	268	
その他	打数	95	362	332	307	376	379	220	146	3	2220	0.311
	安打	26	123	95	93	130	118	66	39	0	690	
右	打数	104	351	342	317	378	349	213	133	3	2190	0.315
	安打	27	123	103	98	135	109	62	32	0	689	
左	打数	24	124	106	111	124	175	85	62	2	813	0.331
	安打	10	44	36	33	47	57	26	16	0	269	
ビハインド	打数	39	146	111	106	129	192	127	70	3	923	0.298
	安打	10	37	44	29	45	56	39	15	0	275	
同点	打数	43	154	136	161	178	141	60	59	1	933	0.340
	安打	14	66	38	51	66	46	18	18	0	317	
リード	打数	46	175	201	161	195	191	111	66	1	1147	0.319
	安打	13	64	57	51	71	64	31	15	0	366	
計	打数	128	475	448	428	502	524	298	195	5	3003	0.319
	安打	37	167	139	131	182	166	88	48	0	958	

PayPay	ZOZO	メット	楽天生命	札幌ドーム	京セラ
0.9806	0.9368	1.0229	0.9826	1.0056	1.0381

表2,3 集計した各データ（黄色と緑の数値を主に使用）

③「打つ確率」の予測式

バッターが打席に入る時、ヒットを打つかどうかにつながると考えられる主要要素を樹形図に書きだして、走者、点差、球場といった打席状況、自分の状態、相手投手の左右と被打率の6つの要素が関係していると考えた。



これをもとに、次の式によって「打つ確率」を求め、これがその打席時点でのシーズン打率を超えていればヒットを打つのではないかと予測した。

直近20打席の打率

×

相手投手の被打率

2020シーズンのパ・リーグの選手の平均打率(=0.250)

×

対右/左投手の通算打率

通算打率

×

得点圏走者がある時/ない時の通算打率

通算打率

×

ビハインド/同点/リード時の通算打率

通算打率

×

球場のヒットの出やすさの指標

その時点でのシーズン打率

⇒ヒットを打つ

(緑文字 = 「打つ確率」)

(注)

- 「打つ確率」は、‘直近の状態’である直近20打席の打率をベースにし、そこに各打席の‘状況’（‘相手’も含む）のそれぞれの要素から求められる、この打席でのヒットの打ちやすさをかけたものである。
- 2020シーズンのパ・リーグの選手の平均打率(=0.250)

→基本的にはバッターの打率の平均 = ピッチャーの被打率の平均と考えられることから、この数値（平均打率）が2020シーズンにパ・リーグの投手がリーグの平均的な打者にどれくらいヒットを打たれたかを表していると考えてよく、さらには相手投手の被打率をこの数値で割ることで、相手投手の打たれやすさ、つまり柳田選手にとっては相手投手の打ちやすさを表す。

3. 結果

仮定した内容がどのくらいの精度で正しいのかを調べるために、2020シーズンの柳田選手の全515打席について「打つ確率」を求めてヒットを打つかどうかの予測を行い、それと実際の打席結果を照らし合わせて的中率を調べたところ、ヒットを打ったか打たなかったかが一致したのは268打席であり、**的中率は約52%**であった。

	球場	左右	被打率	ランナ	点差	PF (改左右)		被打率	状況	得点圏	打ちや	20打席	打率	予測式	予測	実際	的中
1	ソ	右	0.270	0	0	0.981	0.315	1.080	0.340	0.311	1.084	0.000	0.000	0.000	0	0	1
2	ソ	右	0.270	0	0	0.981	0.315	1.080	0.340	0.311	1.084	0.000	0.000	0.000	0	0	1
3	ソ	右	0.270	0	0	0.981	0.315	1.080	0.340	0.311	1.084	0.000	0.000	0.000	0	0	1
4	ソ	右	0.160	1	0	0.981	0.315	0.640	0.340	0.342	0.707	0.000	0.000	0.000	0	0	1
5	ソ	右	0.246	1	0	0.981	0.315	0.984	0.340	0.342	1.087	0.000	0.000	0.000	0	0	1
6	ソ	右	0.246	0	-	0.981	0.315	0.984	0.298	0.311	0.866	0.000	0.000	0.000	0	0	1
7	ソ	右	0.246	0	-	0.981	0.315	0.984	0.298	0.311	0.866	0.000	0.000	0.000	0	1	0
8	ソ	右	0.209	1	0	0.981	0.315	0.836	0.340	0.342	0.924	0.250	0.250	0.231	0	0	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
508	ロ	右	0.270	0	0	0.937	0.315	1.080	0.340	0.311	1.035	0.125	0.340	0.129	0	1	0
509	ロ	右	0.270	1	0	0.937	0.315	1.080	0.340	0.342	1.140	0.188	0.341	0.214	0	1	0
510	ロ	右	0.270	0	-	0.937	0.315	1.080	0.298	0.311	0.908	0.250	0.343	0.227	0	0	1
511	ロ	右	0.228	0	-	0.937	0.315	0.912	0.298	0.311	0.767	0.235	0.342	0.180	0	0	1
512	ソ	右	0.228	0	0	0.981	0.315	0.912	0.340	0.311	0.915	0.235	0.341	0.215	0	0	1
513	ソ	左	0.349	1	0	0.981	0.331	1.396	0.340	0.342	1.622	0.235	0.341	0.382	1	1	1
514	ソ	右	0.303	0	4	0.981	0.315	1.212	0.319	0.311	1.142	0.294	0.343	0.336	0	0	1
515	ソ	右	0.273	1	4	0.981	0.315	1.092	0.319	0.342	1.133	0.235	0.342	0.267	0	0	1

図4. 検証結果の一部

4. 考察

① ‘直近の状態’について

検証結果について考察する前に、検証するためにデータを取る過程で気づいたことがあったので考察しておく。

右のグラフは、「打つ確率」を求めるにあたって用いた直近20打席の打率の推移に加え、直近10,30,40打席の打率の推移とシーズン通算の打率の推移をまとめたものである。これを見ると、直近何打席の打率であるかにかかわらず、直近の打率が高い部分と低い部分があり、山と谷がおおよそ一致していることがわかる。この打率の高低は柳田選手のシーズン内における好不調の波を表していて、このグラフから1シーズン中に約10回、つまりおよそ50打席ごとに波が来ていることがわかる。これに関しては理由がわからなかったため、今後の課題として理由を探っていきたい。また、ある選手のシーズン中の調子はどうであったかを知りたいとき、単に打率の高低や月間打率で判断するよりも、一定期間ごとの打率を複数重ね合わせて、その推移から判断したほうがより正確に状態を判断することができるといえるだろう。

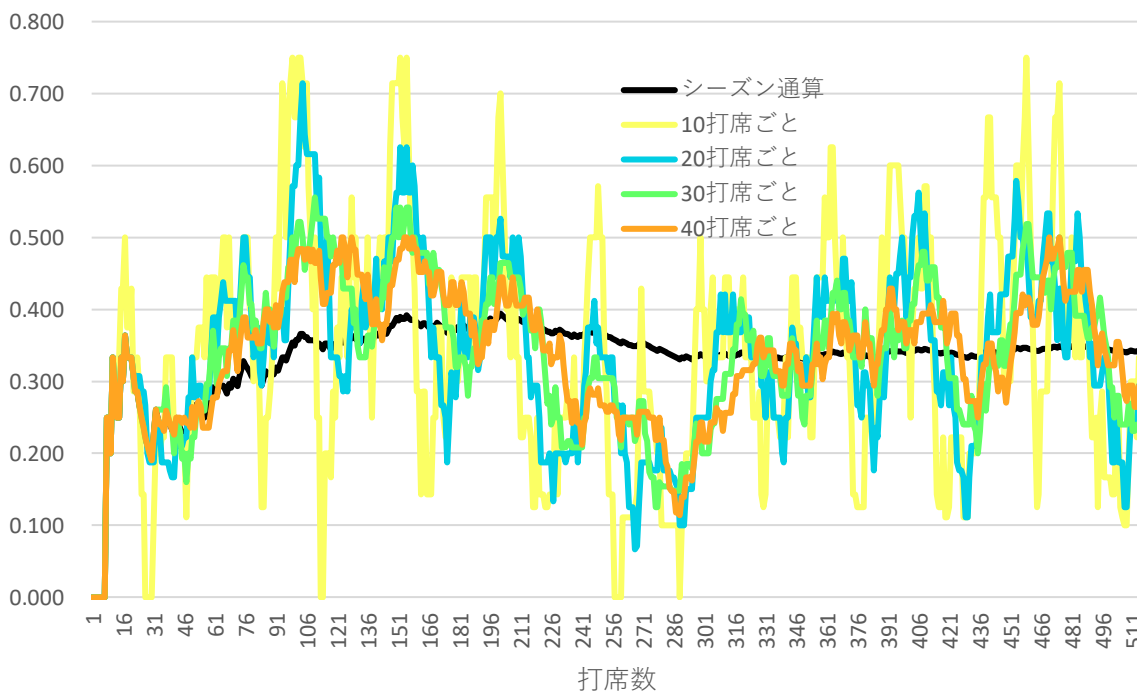


図5. 一定打席ごとの打率の推移

② 「打つ確率」の検証

最後に「打つ確率」の検証に入る。結果のところでも書いたとおり、的中率はおおよそ52%だった。この数値がどのくらい精度の高いものであったかを判断するにあたって、右のグラフは、ヒットを打った時と打たなかったときの「打つ確率」の分布を調べその数をまとめたグラフである。これを見ると、ヒットを打った時もヒットを打たなかったときも、同じように「打つ確率」が分布していることがわかる。これから、「打つ確率」が高くてもヒットを打たないときは打たないし、また同様に「打つ確率」が低くてもヒットを打つときは打つ、つまり残念ながら「打つ確率」と実際の結果との関連性は高くなかったということが分かった。その理由について考えると、まず一番可能性があるのは、直近の打率ではその打席における状態を表し切れていない、ということである。ただ①の考察の通り、直近の打率の推移からおおよそその調子の変化がわかるという点については間違っていないと考えるため、この数値が全体の「打つ確率」に影響を与えすぎたか、その打席におけるヒットの打ちやすさが「打つ確率」に与える影響が小さすぎたかのどちらかだと考えられる。そのあたりの割合の調節がうまくいかなかったことはまず反省点としてあげられる。次に、「打つ確率」の予測式に含まれていない要素が与える影響である。樹形図で考えた要素以外にも、相手の調子であったり、その時のチーム全体の状態（例えば、連勝中であれば『今日も勝とう』という気持ちが働いていつもより集中して打席に臨めてヒットが打てるかもしれない）などがあるが、すべての要素を考えた立式をすることは今回はできなかった。さらに柳田選手の場合は選球眼が非常によく全打席の16%にあたる85打席が四死球による出塁であるため、これが大きな影響を与えた可能性が高い。今回は四球の場合をヒットを打たなかったほうに含めており、選球眼の良さや相手投手の四球の出しやすさを計算に加えたり、四球としてヒットの有無とは独立した場合分けをしったりなどすればもう少し精度が上がったのではないだろうか。最後に、「打つ確率」の予測にあたって、一部に2020年シーズンのデータを使用していることである。本来は2020年の打席の予測であるので、2019年までのデータをもとにするべきであるが、球場、パ・リーグ選手の平均打率、相手投手の被打率は2020年のものを使用している。シーズン中であれば、その都度データが更新されるのでそれを利用して予測できるが、今回はシーズン終了後にデータを得たので2020年を通じての値になってしまった。また2019年のデータを使うことも考えたが、今シーズンから加入したり一軍で活躍し始めた選手もいる中で前のシーズンの成績を使うと正確性はより低くなってしまうと考えたため、今回はこのような結果になった。

以上の考察を踏まえると、「打つ確率」は現段階では正確であるとは言い難い。ただ、改善点はかなり多く見つかったので、これらの一つずつ改善していけばかなり精度が上がるのが期待される。

「その打席で打つ確率」が求められればより野球を見ることが楽しくなることは間違いないので、ぜひ新たな指標として利用できる程度にまで研究したいと思った。

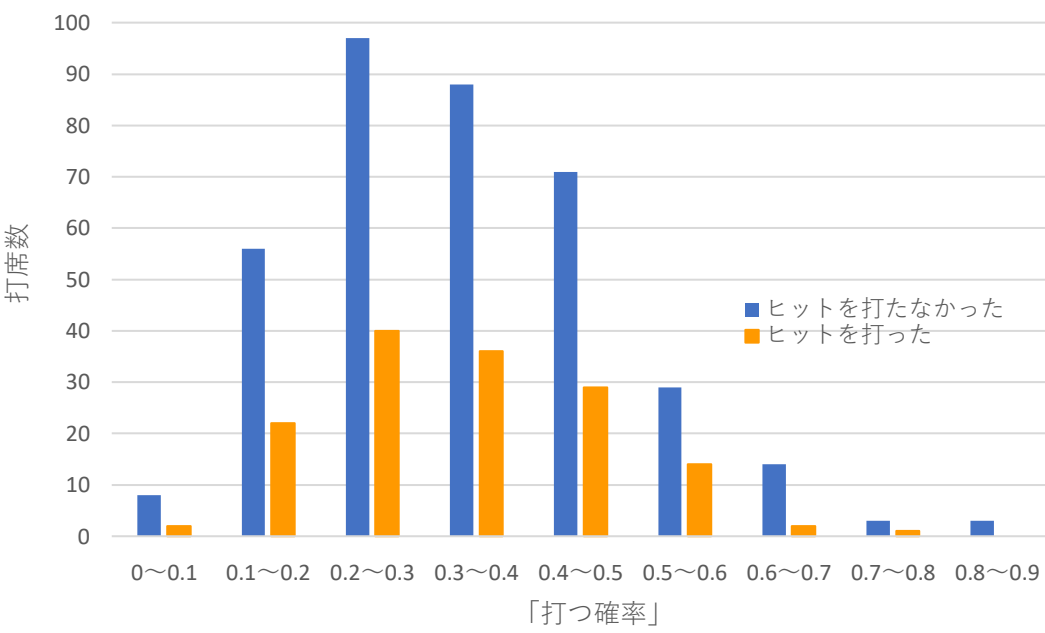


図6. 「打つ確率」と実際の結果の関係

5. 参考文献

〈参考サイト〉
データで楽しむプロ野球 <https://baseballdata.jp/>
プロ野球 - スポーツナビ <https://baseball.yahoo.co.jp/npb/>

『情報・システム研究機構 統計数理研究所 医療健康データ科学研究センター』
『データスタジアム 株式会社』
『Yahoo!スポーツデータ』
以上のご支援いただいた関係者の皆様に感謝いたします。