

プロ野球全球データから見る統計分析と実用性

小林 裕哉[†] 大塚 真吾[‡]

[†] 神奈川工科大学 情報工学科 〒243-0292 神奈川県厚木市下荻野 433-3

E-mail: [†] s1521016@cce.kanagawa-it.ac.jp, [‡] otsuka@ic.kanagawa-it.ac.jp

あらまし 近年,スポーツにも IT 技術を活用しアスリートの支援や,一般の人の運動のサポートをする研究が増えている. その結果,「陸上競技」,「サッカー」,「水球」,「スポーツジム」などの様々な種目や運動施設で IT 技術が応用され,IT 技術の有用性を多くの場面で示している. そこで本研究では,スポーツの中でも「野球」に焦点を当て,試合の中で記録された詳細なデータから打者と投手の相性を分析する. その結果から,実際の試合での実用性を模索する.

キーワード 野球, データ分析, データマイニング, ピッチャー, バッター

1. はじめに

近年,IT 技術を活用したスポーツのデータ分析は革新的勢いで普及しており,様々な種目や運動施設等で導入されている. 来る 2020 年の東京オリンピックでも,施策のひとつである「オリンピック×IoT」では,スポーツの分野における計測機器の開発やデータ計測と可視化手法の高度化そして新サービスの提案など最新 ICT の効果的な活用が進められている[1]. オリンピックでは,これまでも先進的な通信技術や放送技術が使われてきた. 1960 年代には,電子技術の発達により,採点,協議結果計算の自動化が進み,1980 年代にはコンピュータ,1990 年代にはインターネット,2010 年以降はデジタル動画技術が取り入れられている[2]. オリンピックだけでなくパラリンピックでも IT 技術は導入されている. パラリンピックの場合,障がい当事者は障がいゆえの困難と,外国人である困難に直面する. そこでカギを握るのが,躍進目覚ましい IT(情報通信技術)であるといわれている[3]. 具体的なスポーツに IT 技術が使われている事例をいくつか示す. 「陸上競技」,特に跳躍種目はより高く,遠くへ跳ぶことを追求し,人間の限界を競う種目である.その中でパワーは持っているのに結果に結びつかない競技者にデジタルカメラやタブレット端末を使用し,跳躍フォームを撮,記録することで科学的解明を目指す研究がある[4]. 「サッカー」では,有酸素運動が疲労時のパスの正確性への影響を検証するために,実際の試合の動きを再現するシミュレーションプロトコルをもしいて,有酸素運動が疲労時の巧緻性において影響を与えることが示唆された実績がある[5]. また「水球」では,スポーツコードを活用した映像によるゲーム分析を行い,大会後

に数量データとして分析を行い,映像データベースを作成し,プレイの振り返りや対戦チームの特徴把握に生かす研究が行われている[6]. このようにスポーツと IT 技術は現在では深く関わっており,今後とも躍進していく関係性といえる.そこで本稿ではスポーツの中でも「野球」に焦点を当て,投手と打者の相性を試合データから分析していく.

2. 投手・打者の相性分析

2.1 使用するデータ

本研究を進めるにあたって使用したデータは,2015 年のプロ野球パ・リーグ全試合における「全球データ」である. 「全球データ」とは,投手がどのような投球をし,打者がどのように対応し,ヒットがあった場合には,どのような打球だったのか,また,どのような守備をしたのかを,一球ごとにスコアラーが記録したデータである. 試合は交流戦も含み,総試合数は 480 試合程度,レコード数(投球数)は 150,000 レコード程度のデータである. 球団,球場,選手個人についてはマスク処理が施されている. データは”投手 ID”,”球種”,”投球エリア”,”打者 ID”,”バッティング結果”,”飛球座標”などのフィールドが存在し,1 レコードは,1 球に対する結果を示している.

2.2 分析手法

投手,打者の相性を分析するにあたって,投手・打者それぞれに得意なコースや球種を絞り,「得意球」を設定する. 「得意球」は,二つの要素から絞り込む. ストライクゾーンを 9 つに区分することで,得意なコースを判断する要素と,”球種 ID”からどの球種が得意かを

判断する要素の二つである。球種については、「ストレート」、「カーブ」、「シュート」、「スライダー」、「フォーク」、「シンカー」、「チェンジアップ」、「特殊球」、「カット」の9種類で判断する。(図1参照) つまり9×9の81種の中から各投手・打者の得意球を判断する。投手については、上記の球の中でストライクになっているものをカウントし、カウントが多い上位3種を「得意球」とする。打者については、打った球の中でも、「単打」、「二塁打」、「エンタイトル2ベース」、「三塁打」、「本塁打」になっているものを投手同様にカウントし、上位3種を「得意球」とする。上位3種に重みをつけ比較することで、相性の判別を行う。

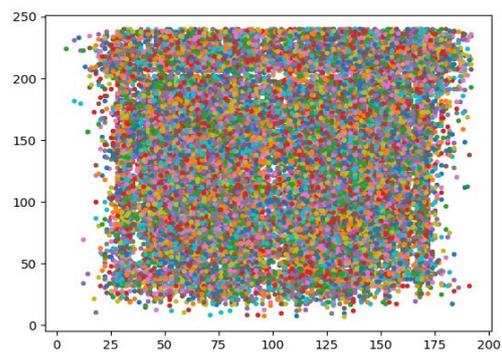


図3. ストライク球の座標プロット

球種ID	球種名	変化球F
0	無し	0
1	ストレート	0
2	カーブ	1
3	シュート	1
4	スライダー	1
5	フォーク	1
6	シンカー	1
7	チェンジアップ	1
8	特殊球	1
9	カット	1

図1. 球種 ID 内訳

1球結果ID	結果名
0	なし
1	ファウル
2	単打
3	死球
4	二塁打
5	エンタイトル2ベース
6	三塁打
7	本塁打
8	失策出塁
9	凡打
10	犠打
11	犠飛

図2. 1球結果一部抜粋

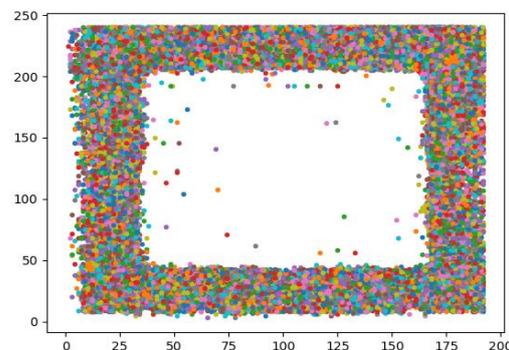


図4. ボール球の座標プロット

2.3 分析結果

分析を行うにあたって、ストライクゾーンの範囲が選手の体格によって変化する問題がある。そこで全投球のうち、ストライクカウントとボールカウントに分けてプロットを行った。(図3,4参照) その結果、ストライクエリアとボールエリアが明確に分かれていないことが分かった。得意球を判別する上で明確な基準が必要のため、使用するデータに合わせるように設定した。投手の得意球はストライクカウントとなった球のみを使用するので、ボール球を考慮せず、ストライク側のエリアでストライクゾーンを設定した。また設定したエリアを9つに分割し、番号をつけ新たなフィールドとしてレコードに付与した。(図5参照)

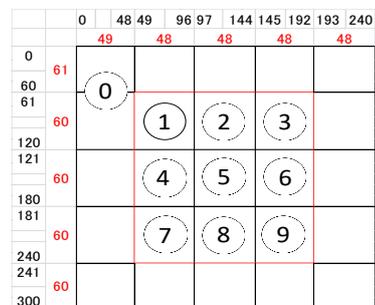


図5. ストライクエリアの区分

ストライクゾーンを上記のように設定し、手順通り分析を行った。得意球の重みづけに関しては、順位に差があるときにのみ点数をつけ有利不利が分かるように設定した。有利不利は比率で表し、最大で3:1の有利性が示せるものとした。また、順位が同率の場合や、得意球に該当がない場合には、点数をつけず、1:1となるようにした。実際の相性比較を行った結果を以下に示す。投手ID“10000”の選手と打者ID“10002”と“10006”の選手を比較し、相性分析を行ってみる。(表

1,2 参照) “10002”との相性は 1:1 で五分であり,“10006”との相性は 2:1 で投手側が抑える確率の方が高いという結果が得られた.(図 6,7 参照)

参 考 文 献

- [1] 山本 雄平, 田中 成典, 姜 文淵, 中村 健二, 田中 ちひろ, 清尾 直輝, “アメリカンフットボールの可視化システムの開発および選手のプレー分析に関する研究” 情報処理学会論文誌 vol.59 No.5 p.1334-1350 Nov.2018
- [2] 田崎 雅彦, “オリンピックのための情報処理” 情報処理 Vol.55 No.11 Nov.2014
- [3] 江藤俊太郎 “IT(情報通信技術)が実現する2020年東京パラリンピックレガシーとは” 2014
- [4] 小賦 肇, 片峯 隆, 長野 史尚, 岡部 優真, “陸上競技跳躍種目における指導法について: ICT 機器を用いた跳躍種目の指導法” 名桜大学総合研究 no.27 p.73-79
- [5] 清水 万輝, 高野 綾斗, 井上 功一, 榎木 泰介, “サッカー選手における有酸素運動が疲労時のパスの正確性に与える影響” 大阪教育大学紀要. 人文社会科学・自然化学 No.66 p.13-20 2018
- [6] 南 隆尚, “水球日本代表男子チームにおけるデータの活用術” 日本体育学会第 68 回大会 2017

表 1. 投手の得意球

投手 ID	得意球 1	得意球 2	得意球 3
10000	コース 5/ ストレート	コース 6/ ストレート	コース 7/ フォーク

表 2. 打者の得意球

打者 ID	得意球 1	得意球 2	得意球 3
10002	コース 5/ ストレート	コース 2/ ストレート	コース 1/ ストレート
10006	コース 4/ ストレート	コース 5/ ストレート	コース 1/ ストレート

```
投手ID:
10000
打者ID:
10002
投手 1:1 打者
```

```
投手ID:
10000
打者ID:
10006
投手 2:1 打者
```

図 6. “10002” との比較結果 図 7. “10006” との比較結果

2.4 考察

今回の分析では,選手個人の相性を比率で算出することができたが,実際の相性や結果と一致するとは言えない.実際には他に多くの要因が重なり,相性の有無が定まるため,相性を判断するにはさらに多くの要因を絡め,データ数も増やす必要がある.

3. おわりに

本研究はプロ野球における投手・打者の相性を見る一つの方法の提案であった. 今後,今回出した結果が実際のデータと比較してどのくらいの正当性があるのかを照らし合わせる必要がある. また,投球時の腕や左右の打席など,その他の要因も含めて,相性を考え検