投手及び打者のクラスタリングに基づく対戦相性の解析

情報科学科　前田 涼　　指導教員：平尾 将剛，田坂 浩二

1　はじめに

プロ野球において，投手・打者間の対戦に“相性”はないだろうか．本研究では，投手・打者間の“対戦相性”に注目し，具体的に数値化することを目標とする．“対戦相性”の解析は個人対個人で行うのではなく，投手及び打者を特徴・傾向別（タイプ別）にグループ分けし，グループ対グループで行う．それにより，汎用性の高い解析結果を得ることを目指す．昨年度，平山ら [1] はクラスタリングを用いて，育成を視野に入れた投手の評価指標を作成した．この研究において，平山らは投手をクラスタリングすることにより分類している [1]．本研究では，まず投手及び打者をタイプ別にグループ分けする必要がある．そこで，投手及び打者をクラスタリングすることにより，タイプ別に分類・グループ分けする．この投手のグループと打者のグループの組み合わせにおいて，“対戦相性”の解析を行う．

本研究は，データスタジアム株式会社様，並びに情報・システム研究機構統計数理研究所様より貸与して頂いた日本プロ野球の 2015，2016 年度のデータを用いて行った．

2　解析方法

解析の手順を示す．解析には，投手及び打者の指標を標準化したものを用いる．

1. 【打者】指標を主成分分析する．（相関が強いものが存在するため．）累積寄与率をもとにクラスタリングに用いる主成分の数を決める．
2. 【投手】指標を用いてクラスタリングする．【打者】主成分を用いてクラスタリングする．（ユークリッド平方距離及びウォード法を用いて階層的クラスタリングを行い，クラスタに分割する．）
3. 【投手】&【打者】各クラスタにおいて，そのクラスタに属する選手の解析に用いた指標の平均値を計算する．それをもとに各クラスタの特徴・傾向を掴む．
4. 【投手×打者】投手のクラスタと打者のクラスタの組み合わせごとに対戦成績を分析し，「対戦傾向スコア」を作成する．
   (a) 投手のクラスタと打者のクラスタの組み合わせにおいて，４つの指標「打率」,「IsoD」,「IsoP」,「BB/K」の値を計算する．
   (b) ４つの指標に対して，各指標ごとに組み合わせ間で相対評価（５段階評価，値が大きいほど高評価）を行い，４つの指標をそれぞれ評価値に変換する．
   (c) ４つの指標の評価値の合計を「対戦傾向スコア」とする．
5. 【投手×打者】各クラスタの特徴・傾向と「対戦傾向スコア」の値より，考察を行う．

「対戦傾向スコア」の値が大きいとき打者有利であると考えられ，小さいとき投手有利であると考えられる．

3　解析結果

ここでは，2016 年度の結果について記述する．上記の手順に従って，主成分分析及びクラスタリングし「対戦傾向スコア」を計算した．組み合わせの数は 36（投手：６クラスタ，打者：６クラスタ）となった．得られた「対戦傾向スコア」の値を表１に示す．また，各クラスタの特徴・傾向を調べた．例えば，P5 はゴロ，P3 はフライ，P4 は奪三振，B1 はゴロや内野安打，B2 はフライや三振といった傾向にある．詳細や 2015 年度の結果については本論文を参照して欲しい．

表１　「対戦傾向スコア」 － 2016 年度－

		投手					
		P1	P2	P3	P4	P5	P6
打者	B1	10	9	11	11	10	12
	B2	11	17	11	11	16	12
	B3	11	16	15	9	15	13
	B4	17	14	12	12	15	16
	B5	8	12	12	11	8	14
	B6	10	11	8	10	13	9

4　考察

投手，打者ともに「対戦傾向スコア」の値に差があることから，有利な相手（比較的有利な相手），不利な相手（比較的不利な相手）が存在する，分かれていると考える．また，有利な箇所が多いクラスタ，不利な箇所が多いクラスタが存在することが読み取れる．

「対戦傾向スコア」の表と，各クラスタの特徴・傾向を合わせて見ることにより，投手・打者間の“対戦相性”を数値化でき，また良し悪しを判断することができると考えている．

また，2015 年度と 2016 年度の結果を比較すると，似た特徴・傾向をもつクラスタにおいて，「対戦傾向スコア」は似通っている部分と大きく異なる部分が存在することが読み取れる．

5　まとめと今後の課題

投手・打者間の“対戦相性”の良し悪しを「対戦傾向スコア」を作成することで数値化し，投手，打者のタイプ別に有利なタイプ，不利なタイプを調べた．「対戦傾向スコア」は各クラスタの特徴・傾向と合わせて見ることにより，采配の評価や，対戦データの少ない選手の“対戦相性”の推測などに役立てることができると考えている．

今後の課題としては，結果のより深い考察や，より良いクラスタリングの方法の調査，クラスタ分割における分割数の決め方の検討，「対戦傾向スコア」の計算方法の精密化などが挙げられる．結果のより深い考察の例としては，「対戦傾向スコア」において，２年間で大きく異なる箇所の原因の考察・調査が挙げられる．

参考文献

[1] 平山寛朗ら（以下７名），“育成を視野に入れた投手評価指標の提案”，統計数理研究所共同研究リポート 380，pp.23–26，2017