

多変量解析法による日米野球選手の分析

－ ヤンキース松井選手の打撃分析 －

2000MM070 岡田 聡 2000MM103 山田 正氏

指導教員 木村 美善

1 はじめに

2003年、日本球界を代表する打者である松井秀喜選手がメジャーリーグ入りを果たし、ニューヨークヤンキースに移籍した。本研究では、松井選手が1年間を通してどのように成績を収めていったのかを多変量解析法を用いて分析する。また、メジャーに移籍の決まった松井稼頭央選手についてもメジャー1年目の成績予測を行う。使用したデータは [3], [4], [5] から入手した。

2 分析方法

分析には数量化 I, II, III 類, 重回帰分析法を使用し, [1], [2] を参照とした。

3 松井選手の6月と8月の打撃成績の比較

松井選手は6月は0.394という高打率を残したが, 8月には0.233とシーズンで最低の打率しか残せていない。6月と8月の間には1割6分ものひらきがあるのだが, 相手投手の攻め方に違いがあったのかどうかを分析した。

3.1 数量化 III 類を使用した分析

球種: 1「ストレート」2「変化球」

コース: 3「内角高め」4「内角中」5「内角低め」

6「真中高め」7「真中」8「真中低め」

9「外角高め」10「外角中」11「外角低め」

相手投手: 12「右腕」13「左腕」

をカテゴリーとして, 数量化 III 類を使って分析した。6月と8月の全打席データををひとまとめにし, 同時に数量化 III 類で分析し, プロット図を出力した。

また, 得られたプロット図の群の中で6月の打席, 8月の打席の数を数え, 6月と8月の間に投手の攻め方で違いがあったのかを見た。数が同じであった場合は6月と8月との間では攻め方には差がないものと考えた。

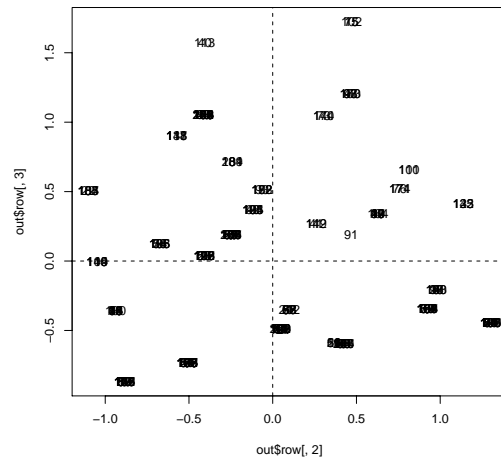
3.2 軸の意味付け

第一軸の固有値は, 0.4578であった。カテゴリーに付与する数量から正の大きな値を取ったカテゴリーは, 変化球, 外角低めであった。これらのカテゴリーは松井選手にとって苦手な組み合わせであると言える。負の方向に大きな値をとったのは, 真中, 内角高めであった。これらは松井選手の得意なコースのカテゴリーであると言える。

第二軸の固有値は, 0.4018であった。カテゴリーに付与する数値から正の大きな値を取ったカテゴリーは, 内角中, 左投手であった。

負の方向に大きな値をとったのは右投手であり, 右投手, 左投手で比べると対右投手の時の打率の方が良い。負の方向へ行くほど得意であると解釈できる。

3.3 サンプルスコアのプロット



3.4 考察

第一象限, 第二象限, 第三象限, 第四象限にいくつかの群ができた。第一象限, 第三象限のプロットが集まった群の打席数には6月と8月に差がないことから, 変化球, 真中の要素はさほど差が無いと言える。

第二象限のプロットが集まった群の打席数は6月, 8月でそれぞれ26:41であった。群が集まった部分の要素は真中高め, 真中低め, 左投手であった。

また, 第四象限で群が集まった部分の6月と8月の打席数は42:26であり, 第二象限での数のほぼ逆であることがわかった。ここでの要素は外角中, 外角低め, 右投手であった。

松井選手が外角低めに投げられたうち, 左投手が投げたのは全体の7%であった。8月は外角低めに投げられたうち左投手が投げたのは全体の33.3%であった。また, 外角中に左投手が投げた率は6月が全体の30%, 8月は全体の38%であった。いずれも8月の方に多く左投手が外角のコースを攻めて投げていることがわかる。松井選手の苦手な左投手が投げる回数が8月に比べて6月の方が少なかった。このことが6月と8月の打率の差になって表れてきたのではないかと考えられる。

実際のデータで言えば, 対右投手の外角の打率は0.14,

対左投手の外角の打率は 0.07 であり、苦手な外角の中でも対左投手の方が特に苦手であることがわかる。

4 打率予測

松井稼頭央選手は実際にどのくらいの成績を残すことができるかというのを昨シーズンの稼頭央選手の成績とメジャーリーグで好成績を収めた主な打者のデータを使って予測する。使用したデータは、松井秀喜選手を含める 2003 年度にメジャーリーグで活躍した選手を 40 人選んだ。

4.1 数量化Ⅰ類による松井稼頭央選手の打率予測

40 人の選手データを基に来シーズンの松井稼頭央選手の打率を予測する。外的基準は各選手の打率とした。なお、日本のプロ野球とメジャーリーグでは試合数に違いがあり、稼頭央選手の出場試合は 140、今回使用したメジャーリーグの選手たちの平均出場試合数は 157 試合であったので試合数の条件を公平にするために稼頭央選手のデータにおける数値は、その値にそれぞれ $\frac{157}{140}$ をかけた数値を使用した。

4.2 数量化Ⅰ類分析結果

数量化Ⅰ類結果

アイテム		付与値	偏相関	範囲
安打数	1	- 0.0223	0.5878	0.0315
	2	0.0034		
	3	0.0055		
	4	0.0092		
二塁打数	1	- 0.0113	0.4037	0.0207
	2	- 0.0001		
	3	0.0094		
三塁打数	1	- 0.0026	0.3454	0.0146
	2	- 0.0026		
	3	0.0120		
本塁打数	1	0.0008	0.3600	0.0147
	2	- 0.0046		
	3	0.0100		
	4	- 0.0038		
打点数	1	0.0181	0.5119	0.0240
	2	0.0062		
	3	- 0.0059		
	4	0.0070		

4.3 松井稼頭央選手の打率予測

回帰式は、

$$y = 0.296$$

$$\begin{aligned} & - 0.00223x_{11} + 0.0034x_{12} + 0.0055x_{13} + 0.0092x_{14} \\ & - 0.0113x_{21} - 0.001x_{22} + 0.0094x_{23} \\ & - 0.0026x_{31} - 0.0026x_{32} + 0.012x_{33} \\ & + 0.0008x_{41} - 0.0046x_{42} + 0.010x_{43} - 0.0038x_{44} \\ & + 0.0181x_{51} + 0.0062x_{52} - 0.0059x_{53} + 0.007x_{54} \end{aligned}$$

この結果より、予測値(打率)yを求めると、0.329となる。

4.4 考察

稼頭央選手の打率を予想したところ 3 割 2 分 9 厘となった。稼頭央選手はホームランをたくさん打つというタイプの選手ではないが、足が速いので安打を多く打つことができる。メジャーリーグにおいては、高打率を残すには、ホームランを狙うよりも、安打を積み重ねていく方が良いのではないかとと思われる。そのことは、数量化Ⅰ類の分析結果における安打数の多い選手ほど打率を上げる影響があるという順序性から証明できた。ホームランは多い程打率が上がるというわけではない。ホームランバッターは大振りになり、打率が上がらないのではないかと考えられる。

5 打点の予測と要因を探る

この分析において、今期の松井選手の打点が 106 と出ていたが、予測される打点とはどんな要因(説明変数の種類)から得られたのか、また、予測される打点と実際の結果から出た打点とはどのような差があるのか、打点を目的変数とにおいて考察していく。松井選手の打点を他の選手の打撃成績から得るため、主成分分析に使用した選手から松井選手を除いた 22 選手のデータを用いて分析を行った。なお、説明変数の選択の方法として「後退消去法」を使用した。

\hat{a}_j : 回帰係数, $\sqrt{s^{ij} \frac{V_e}{n}}$: 標準誤差

$$F = \frac{\hat{a}_j^2}{s^{ij} \frac{V_e}{n}}$$

今回は上式の値を求めて自由度(1, n - p - 1)のF分布の限界値と比較し、 $|F| \geq F_{n-p-1}^1$ ならば仮説を棄却していく。Rを使った出力結果ですでにt-統計量が出ているため、 $t^2 = F$ を用いて検定を行っていく。

5.1 回帰係数の検定

帰無仮説 $H_0: a_j = 0$

対立仮説 $H_1: a_j \neq 0$ (j=1, 2, ..., 13)

とし、Fが最小値となるjに対して、 $H_0: a_j = 0$ が棄却されるまでFが最小値をとるjを一つずつ除きながら検定を続ける。 H_0 が棄却された時点で残ったjに対応する x_j が選択された説明変数である。

なお、変数増減の限界値は2と定めた。最終的に得られた結果は以下の通りで、出力結果は表1となる。

また、この表1から、標準誤差=3.405、決定係数=0.729、F検定の結果=4.375、有意水準=0.0093であることがわかる。

得られた回帰式は、

$$\begin{aligned} y = & 941.280 - 1.638x_1 - 0.992x_2 + 3.856x_3 - 0.361x_4 \\ & - 1.728x_5 + 0.080x_6 - 97.586x_7 - 2157.050x_8 \end{aligned}$$

となる。

表 1 説明変数 8 個の出力結果

項目	回帰係数	標準誤差	t-統計量	p-値
試合数	- 1.638	0.681	- 2.406	0.0317
打席	- 0.992	0.297	- 3.342	0.0053
安打	3.856	0.965	3.998	0.0015
二塁打	- 0.361	0.144	- 2.505	0.0264
三塁打	- 1.728	0.505	- 3.423	0.0045
塁打数	0.080	0.037	2.138	0.0521
出塁率	- 97.586	64.497	- 1.513	0.1542
打率	- 2157.050	546.370	- 3.948	0.0017
定数項	941.280	178.641	5.269	0.0002

t-統計量が一番低い出塁率の値を検定すると、 $t^2 = 2.289 > F_{13}^1(\alpha) = 2$ となり仮説は棄却される。よって、この説明変数は落さなくてよいことになり、以下の回帰式が求まる。決定係数は 0.729 となり、この項目の全体のおよそ 73 % が説明できることになる。

5.2 重回帰による打点予測の考察

回帰式に使用された項目は表 1 の結果の 8 種類であった。つまり、これらの項目が今シーズンの打点において影響度が強いものであったと考えられる。本塁打数が少なかった松井選手にとって、回帰式に直接関係のなかった本塁打は打点を求める上では関係がなかったということがいえる。他の選手と比べて松井選手は 16 本という少なさが目立ったものの、打点に影響をほとんど与えなかったという結果が得られた。

得られた回帰式に松井選手の成績を当てはめると、予測打点は 98.24 という実際にあった打点 106 と大きく異なった。他の 22 選手の予測打点も実際の打点と比べて、±1 から 4 となっていたことから、予測を大きく上回っていた松井選手の打点は数字以上の活躍を残せたといえる。予測より上回っていた要因として、説明変数の「試合数」「打席」が全試合出場を果たした松井選手にとって、特に打点に対して大きく働いていたと考えられる。

6 勝ち試合の貢献度

ここでは、松井選手がチームの勝ちに貢献したのかどうかを数量化 II 類を用いて分析していく。しかし、ヤンキースというチーム自体が大差で勝った試合が比較的に多いため、単純に打点が勝ちに結び付いただけでは本来の貢献度は証明されにくい。そこで、チームの勝った試合で松井選手の打点があった打席だけに条件をしばり、松井選手の打点の貢献度を分析した。状況にリードやビハインドといった要素は含まずに試合の緊迫状況といった形として扱った。つまり、打席に立ったときの打点の価値を測るための分析を行った。

各アイテムは、外的基準 1「5 点差以上勝ち」2「1~4 点差勝ち」として、
打撃結果: 1「凡打」2「ヒット」3「2, 3 塁打」
4「ホームラン」

ランナー: 1「なし」2「1 塁」3「2 塁」4「3 塁」
アウトカウント: 1「ノーアウト」2「ワンアウト」
3「ツーアウト」

状況: 1「同点, 1 点差」2「2~4 点差」
3「5 点差以上」

打順: 1「4, 5 番」2「それ以外」
として分析を行った。

表 2 勝ちの種類による分析結果

アイテム		付与値	偏相関	範囲
結果	1	0.3685	0.1073	0.6255
	2	0.0427		
	3	- 0.2570		
	4	- 0.090		
ランナー	1	0.1584	0.4019	3.1602
	2	2.827		
	3	- 0.3330		
	4	- 0.3030		
アウト カウント	1	- 0.9851	0.2919	1.3322
	2	0.3026		
	3	0.3470		
状況	1	- 0.1306	0.0855	0.3493
	2	0.2187		
	3	0.0894		
打順	1	- 0.3585	0.2355	0.8963
	2	0.5378		
外的基準	1	0.5740		
	2	- 0.4078		

6.1 勝ち試合の考察

外的基準の「1~4 点差勝ち」の方向に示したアイテムは、「2, 3 塁打」「ホームラン」「2 塁」「3 塁」「ノーアウト」「同点, 1 点差」「4, 5 番」の 7 つであった。中でも、偏相関係数と範囲が最も高かったのはランナーで、次いでアウトカウントとなった。特に接戦でランナーが得点圏にいるときは非常によく打っている。また、ノーアウトからは先頭打者として長打で出塁、もしくはホームランを打っている状況が高いことが考えられ、打順が 4, 5 番と出ていることはクリーンナップとしての接戦での役割を果たせた結果となった。このような結果から勝負強さは先に予測されていたが、試合の緊迫している状況での打点を得ていることが証明された。

7 6 月好調の要因について

数量化 III 類の追加として 6 月の好調の要因をデータを用いて分析を行ってみた。調べていくと、5 月の後半から 7 月の中旬にかけての間、松井選手はセンターを守っていたことが分かった。そこで守備位置の交代がバッティングに影響があったのかを調べるために分析を行った。

各アイテムは、外的基準を1「レフト」2「センター」
として、

打撃結果: 1「凡打」2「ヒット」3「2,3塁打」
4「ホームラン」

打点: 1「あり」2「なし」

ランナー: 1「なし」2「1塁」3「2塁」4「3塁」

状況: 1「同点、1点差」2「2~4点差」
3「5点差以上」

打順: 1「4,5番」2「それ以外」

勝敗: 1「5点差以上勝ち」2「1~4点差勝ち」3「負け」
として分析を行った。

7.1 守備位置の影響について

分析の結果、外的基準「センター」の方向性を示した
アイテムは「ヒット」「2,3塁打」「ホームラン」、打点
「あり」、ランナー「なし」「2塁」「同点、1点差」「2
~4点差」、打順「それ以外」「1~4点差勝ち」であっ
た。偏相関係数と範囲は「打順」が最も外的基準に影響
しているが、これは4番を打っていたウィリアムズ選手
の欠場で、この時期は打順の組み替えが多かったことが
考えられる。総合的に評価しても、センターを守っていた
場合の打撃成績は凡打、打点なしといったアイテムは
示さなかったことから、レフトの成績に比べて全て良い
結果が得られた。

センターに良い結果が出た要因として考えられること
は、松井選手は日本にいたときレギュラー時の守備位置
は長年センターを守っていた。打撃能力を買われてメ
ジャー移籍したものの、開幕からレフトを守ってきた。
一般に慣れている守備位置についての方が、余計な不安要
素がなくなり打撃にも影響があるといわれる。松井選手
は開幕時から今まで守ったことのなかったレフトの守備
に不安があり、打撃にも影響があったのではないかと考
えられる。

8 2選手のランナー別の打撃比較

重回帰分析と数量化II類の考察として、今シーズ
の松井選手に近い成績を残したテハーダ選手との比較を
行ってみた。

なお、2選手の得点圏打率は、松井選手が0.335、テハー
ダ選手が0.288であった。

表3 松井選手の成績

ランナー	打数	安打	本塁打	打率	打点
1,2塁	69	20	1	0.290	21
2塁	30	10	1	0.333	8
1,3塁	26	10	0	0.385	16
2,3塁	12	2	0	0.167	9
3塁	13	6	1	0.462	9
満塁	23	10	2	0.435	27
通算	623	179	16	0.287	106

表4 テハーダ選手の成績

ランナー	打数	安打	本塁打	打率	打点
1,2塁	42	13	3	0.310	17
2塁	46	12	1	0.262	10
1,3塁	17	5	1	0.294	16
2,3塁	12	5	1	0.417	11
3塁	19	4	1	0.211	7
満塁	17	5	2	0.294	17
通算	636	177	27	0.278	106

8.1 2選手の比較

松井選手は特に満塁での打率0.435、打点27という好
成績が非常に目立っている。この成績は全メジャー選手
の中でも5位の成績であり、打点27は2位という素晴
らしい成績となっている。一方テハーダ選手は、ホーム
ランこそ松井選手より10本以上多いものの、得点圏の
時にあまり効果がでていないため、松井選手と比べチャ
ンスにあまり強くないことが表れている。当初シーズン
全体の成績や、他の選手たちと比較しても打点106とい
う成績が取れたことは考えにくかったのだが、これだけ
得点圏に打つことができれば納得がいくものになり、や
はり「勝負強さ」が強く表れる結果となった。

9 おわりに

メジャー1年目の松井秀樹選手の成績を分析していっ
たが、苦手なコース、勝負強さ、守備位置の影響などと
いった直接数字だけでは分からなかったことが表れ、優
勝に貢献できた要因や来年の課題などが明らかになった。
また、これから移籍していく日本人選手にとっての課題
となる日本野球とメジャーの違いが分かりやすく表れた
のではないかと感じた。

謝辞

本研究にあたって、多大な助言を頂き、また熱心に御
指導頂きました木村美善教授、安藤雅和先生、その他御
協力頂きましたすべての方に深く感謝致します。

参考文献

- [1] 岡本憲明・寺本直人：イチローの打撃分析，南山大
学経営学部情報管理学科卒業論文要旨集，2001
- [2] 田中豊・脇本和昌：多変量統計解析法，現代数学社，
1998
- [3] スポーツナビ松井秀樹特集，
[http://sportsnavi.yahoo.co.jp/baseball/mlb/players/
matui/03season/index.html](http://sportsnavi.yahoo.co.jp/baseball/mlb/players/matui/03season/index.html)
- [4] メジャーリーグベースボールMLB.com，
<http://mlb.mlb.com/NASApp/mlb/index.jsp>
- [5] nikkansports.com，
<http://www.nikkansports.com/news/mlb/matui.html>