修 士 学 位 論 文

題 名

統計的手法によるメジャーリーグ野球の打順最適化モデルの構築

指導教員　　福永　力　　教授

平成　30　年　1　月　5　日　　提出

首都大学東京大学院

理工学研究科　　数理情報科学　専攻

学修番号　　16878321

氏　名　　下木　健太

学位論文要旨 (修士 (理学))

論文著者名：下木 健太

論文題名:統計的手法によるメジャーリーグ野球の打順最適化モデルの構築

現在、様々な分野においてデータを有効活用し、知見を見出そうとする動きが益々盛んになってきている。これは、スポーツにおいても例外ではなく、日本にはスポーツのデータ解析・ソリューションの提供を強みとして、事業展開する企業も設立されている。野球に対してデータを活用し始めたのは、1970 年代にビル・ジェームズ氏が考案したセイバーメトリクス[*1]だといわれている。セイバーメトリクスとは、様々なデータを駆使して、野球の采配の妥当性などを議論するものであり、選手の能力を客観的に評価する指標も数多く考案され続けている。日本のプロ野球でも、1990 年代に東京ヤクルトスワローズの監督であった野村克也氏が提唱した「ID 野球[*2]」によってデータ活用が話題となり、データの重要性について広く認知されるようになった。以上より、野球に対してデータを活用し戦略分析をすることは、今後も重要なファクターとなっていくことに想像に難くないであろう。

野球において、打順はチームが勝利するために極めて重要な戦術として挙げられる。理由の 1 つとしては、打順が 1 つ前になるだけで年間 10 打席以上の打撃機会が増えるといわれているからである[*3]。しかし、打順について我々はある程度の共通認識を持っており、例えば日本のプロ野球では 1 番バッターは走塁能力に優れ出塁率が高い選手、2 番バッターは犠打などでランナーを進塁させ、3 番バッターは打率の高い選手、4 番バッターはホームランバッターを置くべきという考え方が広く一般的に信じられている。では、本当にそのような伝統的な打順構成が適切であるかについて検証するために、昨今様々な最適打順モデルが考案され、効果的な打順についての研究がなされている。しかし、それらのモデルが統計的に信頼できるか否かを検証したものはなかった。また、走塁能力を考慮した最適打順モデルについてはあまり研究がなされていない。

本研究において、著者が取り組んだことは大きく分けて以下の 2 つの項目である。

1. 現実の野球を模倣するモデルの作成
2. 最適打順の算出

項目 1 では、日本のプロ野球と比較して、多くのデータの入手が可能であるメジャーリーグ野

[*1] 英表記：sabermetrics
[*2] Important Data の略称のこと。identification ではない。
[*3] 柳川佳也、宮崎茂次、最適打順についての一考察、岡山大学

球を対象とし、過去の統計データを用いることで、選手の打撃能力・盗塁成功率・バッターの打撃結果別のランナーの進塁法則などのあらゆる事象を考慮した、野球の試合を再現するモデルの開発を行なった。そして、作成したモデルが統計的に信頼性があり、現実の野球を模倣しているといえるだろうという結論を得た。これは、日本のプロ野球についても同様の結果が得られることが示唆される。

項目 2 では、あるチームに着目し、項目 1 のモデルを用いてコンピュータシミュレーションによって考えられる全ての打順に対して、1 試合で獲得できる期待得点を算出した。また、それぞれの打順に対しての期待得点を算出する上で、計算量が膨大となってしまうことから CPU 上のマルチコアを用いた並列処理によって計算時間の節減を行ない、効率的にそれぞれの打順に対する期待得点を算出した。

以上より、コンピュータシミュレーションによって算出された期待得点の大きい打順にはどのような傾向が見られるかについて定量的な考察を行なった。特にランナーの走塁能力を考慮するか否かで最適打順にどのような影響を与えるかについても定量的な考察を行ない、その有効性をシミュレーションによって確認した。

目次

1　序論

1.1　研究内容

現在、様々な分野においてデータを有効活用しようという動きが益々盛んになってきている。これは、野球においても例外ではなく、他のスポーツと比較して起こりうるイベントが離散的であることから、数理的手法として取り扱いやすいため、古くより様々な人によって研究がなされてきた。野球に対してデータを活用し始めたのは、歴史が古く 1970 年代にビル・ジェームズ氏が考案したセイバーメトリクスだといわれている。セイバーメトリクスとは、様々なデータを駆使して、野球の采配の妥当性などを議論するものであり、選手の能力を客観的に評価する指標も多く考案されている。1990 年代に成績が低迷していたメジャーリーグ野球のオークランド・アスレチックを強豪チームへと作り上げていく様子を描いた、映画「マネー・ボール」は日本でも話題となった [5]。日本のプロ野球ではメジャーリーグ野球より遅れてデータ活用が話題となり、1990 年代に東京ヤクルトスワローズの監督であった野村克也氏が提唱した「ID 野球」によってデータ活用の重要性について広く認知されるようになった [3]。どちらの事例もデータを駆使して弱小チームを立て直すという点では類似している。以上より、野球に対してデータを活用し戦略分析をすることが、今後も重要なファクターとなることは想像に難くないであろう。

野球において打順はチームが勝利するために極めて重要な戦術として挙げられる。理由の 1 つとしては、打順が 1 つ前になるだけで年間 10 打席以上の打撃機会が増えることが挙げられる [1][2]。また、打順についてはある程度の共通認識を持っており、例えば日本のプロ野球では 1 番バッターは走塁能力に優れ出塁率が高い選手、2 番バッターは犠打などでランナーを進塁させ、3 番バッターは打率の高い選手、4 番バッターはホームランバッターを置くべきだと広く一般的に信じられている [9]。さらに、セイバーメトリクスの観点からは 2 番バッターに最も長打力のある選手を置くべきと考える意見もある [10]。昨今最適打順モデルの構築については様々な研究がなされており、例えば、D'Esopo and Lefkowitz モデル [4] や様々な最適打順決定モデルが提案されている [1][6][7][2]。しかし、いずれのモデルも現実との適合性については言及されていない。モデルがしっかり現実の野球を模倣しているか、また統計的に信頼性があるか否かを検証することは極めて重要であろう。さらに、提案されている最適打順モデルは走塁能力まで考慮されているものは少ない [1][6]。現実の野球において、チームが勝利するためには走塁・盗塁能力も言うまでもなく重要であることから、それらを考慮することができたらより有用なモデルであるといえよう。

本論文では、確率の観点から現実の野球を模倣した自作モデルの概要について論じ、そして自作モデルは統計的に信頼性があることを示す。次に、メジャーリーグのあるチームに対して自作モデルをベースに、コンピュータシミュレーションを用いて考えられる全ての打順に対して、それぞれ 1 万試合させることで 1 試合当たりの期待得点を求める。最後に、期待得点が大きい打順はどのような傾向が見られるかについて、また、走塁能力を考慮するか否かで最適打順にどのような影響を与えるかについて考察を述べる。

1.2　実験環境

1.2.1　扱うデータについて

今回は、詳細なデータが公開されていないなどの理由により、日本プロ野球のデータは研究対象外とする。日本にはスポーツに関して各種メディア向けスポーツデータの配信、チーム・選手向けのデータを活用した強化ソリューションなどを提供しているデータスタジアム株式会社[*1]があるが、しかしデータ販売は法人のみを対象としており、個人では手に入れることが不可能であった。また、日本統計学会スポーツ統計分科会では、スポーツデータ解析コンペティション[*2]の場で詳細なデータが提供されているが、提供されるデータは直近 2 年分のみと少なく、本研究においては不十分であった。一方で、メジャーリーグ野球のデータについてはインターネット上で非営利組織 Retrosheet[*3]によって公開されており、有志によってデータ入力がされており、詳細なデータのダウンロードが可能となっている。Retrosheet が提供しているデータの特徴として、データを読み解くことで試合がどのように推移したかがわかることが挙げられる。具体的には、1 打席ごとの対戦投手名・ランナーの状況・打球方向などがあり、変数の個数は 97 に及ぶ。得られるデータの一部を表 1 に示す。

表 1　Retrosheet データ 2001 年の一部

NUM	GAME_ID	TEAM_AWAY_ID	INN_CT	OUTS_CT	BALLS_CT	STRIKES_CT	PITCH_SEQ_TX
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
441	ANA200104150	SEA	8	0	0	0	X
442	ANA200104150	SEA	8	1	3	2	BCCBBX
443	ANA200104150	SEA	8	1	0	0	X
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

NUM	BAT_ID	PIT_ID	BASE1RUN_ID	BASE2RUN_ID	BASE3RUN_ID	EVENT_CD	RBI_CT
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
441	suzui001	hases001	mclem001			2	0
442	camem001	hases001		mclem001		20	0
443	marte001	hases001	camem001		mclem001	20	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

データの内容について説明する。このデータは 2001 年全てのチーム・試合の 1 打席毎の結果について記録したものであり、総数 NUM は約 18 万行に及ぶ。表 1 を詳しく見ていくと、例えば NUM= 441 は、2001 年 04 月 15 日に行なわれた、ホーム：アナハイム・エンゼルス、アウェイ：シアトル・マリナーズの試合で、イニング 8、アウトカウント 0、0 ボール 0 ストラ

[*1] http://www.datastadium.co.jp/

[*2] http://estat.sci.kagoshima-u.ac.jp/sports/compe07.htm

[*3] http://www.retrosheet.org/

イク、バッター:suzui001[*4]、ピッチャー:hases001、一塁ランナー:mclem001 の時の打撃結果は、EVENT_CD[*5]= 2、RBI_CT（打点）= 0 となる。つまり、このデータを詳しく見ていくことで、試合がどのように推移したかがわかるようになっている。

また、メジャーリーグ野球に関するデータを入手できるサイトとして、SeanLahman.com[*6]というものがある。これは、Sean Lahman 氏によって作成されたメジャーリーグ野球に関するデータベースであり、1871 年から現在までの様々なデータが記録されている。SeanLahman.com によって我々が入手できるデータセット名の一部を表 2 に示す。

表 2　Lahman の一部

データセット名	概要
Batting	年次ごとの各選手の打撃成績
Master	各選手の略歴
Pitching	年次ごとの各選手の投手成績
Salaries	各選手の年俸額
School	各選手の出身学校
Teams	年次ごとのチーム打撃成績・投手成績

以上から、メジャーリーグ野球の方がプロ野球に比べて様々なデータを豊富に用意することが可能であるため、メジャーリーグ野球を研究対象として選定した。

1.3　開発環境

1.3.1　ハードウェア

- OS : Windows 10 Pro
- 実装 RAM : 16.0GB
- CPU : Intel Core i7-6700 4 コア 8 スレッド

1.3.2　ソフトウェア

- 統合開発環境 : RStudio Version 1.0.136
- プログラミング言語 : R 言語

R 言語は、オープンソースのフリーソフトウェアであり、統計解析に関して強みを持つスクリプト

[*4] イチロー選手。オリックス・ブルーウェーブでプレーした後に、メジャーリーグへ移籍しシーズン歴代最多安打記録を更新した。

[*5] 打撃結果を表す EVENT_CD は 0 から 24 の整数値をとる。ちなみに EVENT_CD= 2 は凡打、EVENT_CD= 20 は単打を意味する。

[*6] http://www.seanlahman.com/baseball-archive/statistics/

言語である。S 言語に基づく商用システムである S-PLUS と本質的には類似しており、ニュージーランドのオークランド大学の Ross Ihaka と Robert Gentleman の両氏によって開発された。オープンソースであることから世界中の統計家やプログラマーによって、日々機能の拡張や様々なパッケージがリリースされている。また、Windows や Mac を始めとするほとんど全ての OS の下で稼働が可能である。先述の通り、R 言語は統計解析に特化していることから、2 次元プロットはもとより、線形（非線形）回帰分析・クラスター分析・ニューラルネットワーク等の複雑な統計モデルもパッケージを用いることで、比較的容易にプログラミングが可能になる。また、並列処理にも対応しておりマルチコアを有効活用することで計算時間の短縮も見込める。

2　単純なモデル

2.1　前提条件

野球において、ランナー・アウト別の全ての状況を考えると、24 通り存在する。これを表 3 に示す。

表 3　アウト・ランナー別の状況

	なし	1 塁	2 塁	3 塁	1,2 塁	1,3 塁	2,3 塁	満塁
0 アウト	1	2	3	4	5	6	7	8
1 アウト	9	10	11	12	13	14	15	16
2 アウト	17	18	19	20	21	22	23	24

3 アウト (攻撃が終了する) まで、常に選手が打席に入ったときには、アウト・ランナーの状況は表 3 の 1〜24 のいずれかであることがわかる。

まず、単純なモデルとしてバッターの打撃結果及びランナーの進塁規則を 6 通りとする。これらの規則を表 4 に示す。

表 4　進塁法則

バッターの打撃結果	ランナーの進塁法則
四死球	バッターは 1 塁へ、ランナーは若い番号の塁から埋めていく。
単打	バッターは 1 塁へ、1 塁ランナーは 2 塁へ、2 塁ランナーは 3 塁へ、その他のランナーは生還する。
2 塁打	バッターは 2 塁へ、1 塁ランナーは 3 塁へ、その他のランナーは生還する。
3 塁打	バッターは 3 塁へ、ランナーは全て生還する。
本塁打	バッター及び全てのランナーが生還する。
アウト	全てのランナーは進塁せず、アウトカウントを 1 つ増やす。(三振とみなす)

つまり、ランナーはバッターの塁打数に従って進塁する。これを単純なモデルとして定義する。表 4 の打撃結果について、データから各選手 i に対して確率ベクトル $\boldsymbol{P_i} = (p_W, p_S, p_D, p_T, p_H, p_O)$ を算出する。ここで、p_W は四死球、p_S は単打、p_D は 2 塁打、p_T は 3 塁打、p_H は本塁打、p_O はアウトとなる確率をそれぞれ表している。今回 2001 年のデータを対象とし、Lahman のデータベース Batting より必要なデータを入手する。得られるデータを表 5 に示す。

表 5　2001 年イチローの打撃データ

playerID	yearID	stint	teamID	lgID	G	AB	R	H	2B	3B	HR
suzukic01	2001	1	SEA	AL	157	692	127	242	34	8	8

playerID	RBI	SB	CS	BB	SO	IBB	HBP	SH	SF	GIDP	
suzukic01	69	56	14	30	53	10	8	4	4	3	

ここで、表 5 の変数名について補足する。stint はシーズン中において何番目のチームに所属したかの回数を表しており、例えば stint= 2 はシーズン中に移籍を 1 回したことになる。lgID は所属するチームのリーグ名を表しており、アメリカンリーグの AL 及びナショナルリーグの NL の 2 種類である。G は出場回数を表している。R は得点、RBI は打点を表しており、この 2 つは紛らわしい用語なのだが、R は自分がホームベースを踏んだ回数、RBI は打撃によってランナーがホームベースを踏んだ回数として考えればわかりやすい。AB は打数を表している。SB は盗塁成功回数を、CS は盗塁失敗回数を表している。SO は三振回数を表している。BB は四球 (フォアボール)、IBB は故意四球 (敬遠)、HBP は死球を表している。SH は犠打、SF は犠飛を表している。GIDP は併殺打 (ダブルプレー) の回数を表している。

また、打席数と打数の違いについて述べる。打席数と打数の関係を式で表現すると、

$$(\text{打席数}) = (\text{打数}) + (\text{四死球}) + (\text{犠打}) + (\text{犠飛}) + (\text{打撃妨害}) + (\text{走塁妨害})$$

となる。ここで、故意四球 (敬遠) は打席数に含まれないことに注意されたい。打率、本塁打率等を計算する時には、それぞれ分母に打数としており、これらは打撃成績を述べるときに注目される指標である。また、打撃ランキングの対象となるために必要とする規定打席 (= 所属球団の試合数 × 3.1) は打席数を指標として用いられる。

打撃妨害や走塁妨害は滅多に起こらないイベントであることより、今回は考慮しない。つまり、

$$(\text{打席数}) = (\text{打数}) + (\text{四死球}) + (\text{犠打}) + (\text{犠飛})$$

とする。(打席数) をＢＯＸで表し、以下のように確率ベクトル $\boldsymbol{P_i} = (p_W, p_S, p_D, p_T, p_H, p_O)$ を定義する。

- $p_W = \dfrac{BB + HBP}{BOX}$
- $p_S = \dfrac{H - (2B + 3B + HR)}{BOX}$
- $p_D = \dfrac{2B}{BOX}$
- $p_T = \dfrac{3B}{BOX}$
- $p_H = \dfrac{HR}{BOX}$
- $p_O = 1 - (p_W + p_S + p_D + p_T + p_H)$

$(\text{安打}) = (\text{単打}) + (2\text{塁打}) + (3\text{塁打}) + (\text{本塁打})$ より、$(\text{単打}) = (\text{安打}) - \{(2\text{塁打}) + (3\text{塁打}) + (\text{本塁打})\}$ となる。また、バッターが出塁する事象以外をアウトと定義している。

2.2 モデリング

モデリングする上で、いくつかの注意点について述べる。まず、スターティングメンバー (スタメン) をどのように選定するかという問題である。シーズン中 1 番バッターから 9 番バッターが全く同じ選手であるのは、まずありえないことであろう。なぜなら、もしスタメン選手がケガをしてしまったら、サブの選手が代わりに試合に出場したり、ある選手の調子の良し悪しによって打順が前後するのは十分考えられるからである。しかしそれらをモデルに組み込もうとすると煩雑になってしまうことが容易に想像できる。そこで今回は、シーズン中スタメンとして出場している上位 9 人から選定・打順を固定し、途中の選手交代はないものとする。つまり、シーズン中は同じ選手で打順は全く変わらないということである。また、ＤＨ制を採用しているアメリカンリーグのみを考慮し、対戦投手やアウェイの守備能力及び延長戦*7は考慮しないものとして、モデリングを行なうことにする。

次に、2.1 節に基づいてシミュレーションを行なう。今回扱うデータは 2001 年の Mariners と Bluejays とした。ちなみに、この年の Mariners は 116 勝 46 敗でアメリカンリーグ西地区 1 位に輝き、Bluejays は 5 チーム中 3 位であった。つまり、今回比較対象としているのは「最強なチーム」と「平均的なチーム」である。表 6 はそれぞれのチームの 2001 年における得点の度数分布表である。

表 6　度数分布表

得点	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Mariners	4	12	10	22	17	21	15	18	11	9	8	4	4	3	2	0	2
Bluejays	9	14	22	22	20	22	9	11	10	9	3	5	3	2	1	0	0

表 7 は単純モデルで必要となる確率ベクトルを計算し、確率行列としてまとめたものである。

次に、それぞれのチームの 2001 年の 162 試合*8において実際獲得した得点のヒストグラムと単純モデルによるシミュレート分布を比較してみる。ここで、シミュレーションの概要について述べる。バッターは表 7 に従って打撃結果が選択され、ランナーは表 4 に従って進塁するものとする。2.2 節に倣い、コンピュータシミュレーションによって 3 万試合させた。

*7 メジャーリーグでは日本プロ野球とは異なり、基本的に両チーム決着がつくまで試合が行われる。ちなみに日本プロ野球の延長戦は、2017 年現在において 12 イニングまでである。

*8 メジャーリーグのレギュラーシーズンは 162 試合と定められている。しかし、2001 年はアメリカ同時多発テロ事件の特例によって、162 試合より少ないチームもあった。日本プロ野球は毎年試合数が微妙に異なり、2001 年は 140 試合、2017 年は 143 試合であった。

表 7　2001 年の Mariners データから算出された確率行列 (表記の都合上、小数第 4 位未満で四捨五入している)

打順	playerID	p_W	p_S	p_D	p_T	p_H	p_O
1	suzukic01	0.0515	0.2602	0.0461	0.0108	0.0108	0.6206
2	mclemma01	0.1417	0.1786	0.0329	0.0185	0.0103	0.6181
3	martied01	0.1756	0.1377	0.0688	0.0017	0.0396	0.5766
4	olerujo01	0.1458	0.1753	0.0471	0.0015	0.0309	0.5994
5	boonebr01	0.0710	0.1870	0.0536	0.0043	0.0536	0.6304
6	camermi01	0.1248	0.1327	0.0474	0.0079	0.0395	0.6477
7	guillca01	0.1033	0.1683	0.0402	0.0076	0.0096	0.6711
8	bellda01	0.0608	0.1549	0.0549	0	0.0294	0.7
9	wilsoda01	0.0539	0.1691	0.0490	0.0025	0.0245	0.7010

図 1、図 2 では、それぞれ 2001 年の Mariners ・Bluejays の表 6 を基にした得点分布と単純なモデルによるシミュレート得点分布を示している。x 軸は得点、y 軸は確率密度を表している。

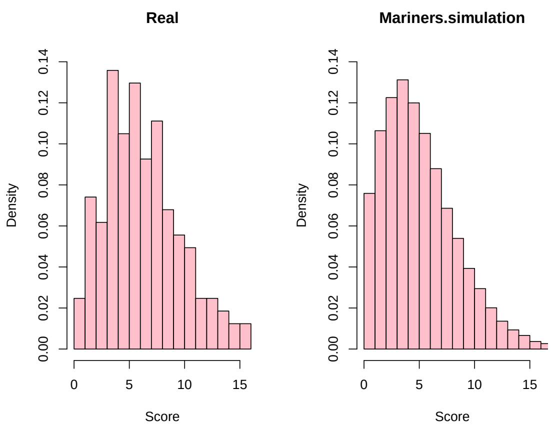

図 1　Mariners2001 の場合

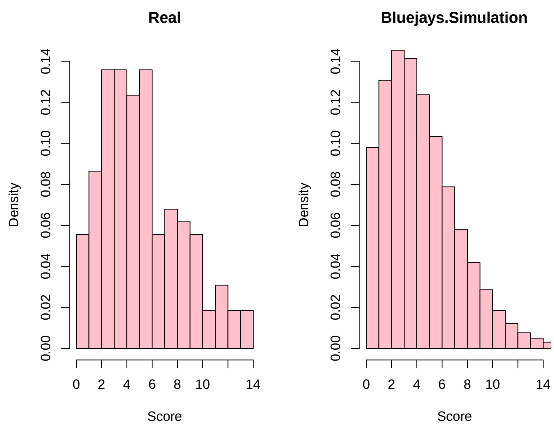

図 2　Bluejays2001 の場合

シミュレート分布の方が実際の得点分布と比較して、得点が過小傾向になっていることがわかる。今回の単純モデルが妥当であるかどうか、2.3 節で統計的検定を行なう。

2.3　モデルの妥当性

この節では、単純なモデルが統計的に妥当であるか検定を行なう。

2.3.1　適合度検定

仮定された理論上の確率分布に対して、標本から求められた度数が適合するか否かを検証することを適合度検定という。

今、k 種類のカテゴリー $A_1, A_2, \cdots, A_k$ に対して、実際に観測された個数を観測度数、仮定された理論上の確率分布に従う各カテゴリーに属する比率を理論確率、全観測数に理論確率を乗じたものを理論度数という。これらをまとめたものを表 8 に示す。

表 8　適合度検定の概要

カテゴリー	A_1	A_2	$\cdots$	A_k	合計
観測度数	O_1	O_2	$\cdots$	O_k	n
理論確率	p_1	p_2	$\cdots$	p_k	1
理論度数	np_1	np_2	$\cdots$	np_k	n

理論度数について、$E_i = np_i$ と書き換えると、検定統計量 χ^2 は、全ての i について $E_i \geqq 5$ のとき

$$\chi^2 = \sum_{i=1}^{k} \frac{(E_i - O_i)^2}{E_i} \tag{1}$$

は自由度 $k-1$ の χ^2 分布に近似的に従うことが知られている[*9]。もし、全ての i について観測度数 O_i と理論度数 E_i が近い値であれば、χ^2 は小さな値になることがわかる。つまり、観測分布と理論分布が近ければ χ^2 は小さくなるということである。

2.3.2　統計的検定

単純なモデルが実際の得点分布に近いかどうか適合度検定を行なう。

1. **仮説を立てる**
   帰無仮説 H_0 は「実際の得点分布とシミュレート得点分布に差がない。」とすると、対立仮説 H_a は「実際の得点分布とシミュレート得点分布に差がある。」となる。今後は帰無仮説が正しいと仮定した上で、議論が進んでいく。
2. **有意水準を設定する**
   今回は $\alpha = 0.1$ とする。
3. **観測度数と理論度数から、検定統計量 χ^2 を求める**
   表 6 の度数分布表が観測度数、単純モデルによるシミュレート結果が理論度数となる。表 8 に倣い、Mariners のデータに対して、それらをまとめたものを表 9 に示す。

[*9] 統計学入門 p.245-p.247、東京大学出版会

表 9　Mariners における単純なモデル

得点	0	1	2	3	4	5	6	7	8
観測度数	4	12	10	22	17	21	15	18	11
理論確率	0.0759	0.1064	0.1225	0.1312	0.1200	0.1051	0.0879	0.0686	0.0539
理論度数	12.2904	17.2368	19.8504	21.2544	19.4346	17.0262	14.2398	11.1132	8.7372
得点	9	10	11	12	13	14	15	16〜	合計
観測度数	9	8	4	4	3	2	0	2	162
理論確率	0.0393	0.0294	0.0201	0.0136	0.0093	0.0066	0.0037	0.0065	1
理論度数	6.3666	4.7682	3.2562	2.1978	1.5120	1.0746	0.5940	1.0476	162

続いて、検定統計量 χ^2 を求めたいところであるが、少し工夫を要する。なぜなら、全ての得点 i について $E_i \geqq 5$ を満たしていないからである。そこで、今得点は 1 点刻みであるので、$E_i \geqq 5$ を満たすように得点の幅を調整する。その結果を表 10 に示す。

表 10　Mariners における得点の幅を調整した単純なモデル

得点	0	1	2	3	4	5	6	7	8
観測度数	4	12	10	22	17	21	15	18	11
理論確率	0.0759	0.1064	0.1225	0.1312	0.1200	0.1051	0.0879	0.0686	0.0539
理論度数	12.2904	17.2368	19.8504	21.2544	19.4346	17.0262	14.2398	11.1132	8.7372

得点	9	10〜11	12〜	合計
観測度数	9	12	11	162
理論確率	0.0393	0.0495	0.0397	1
理論度数	6.3666	8.0244	6.4260	162

これで、適合度検定をすることが可能となった。式 (1) と表 10 より、検定統計量 $\chi^2_{Mariners}$ を求めると、

$\chi^2_{Mariners} = 24.5$ となる。Bluejays についても同様にして、$\chi^2_{Bluejays} = 19.2$ を得る。

4. **検定統計量から結論を導く**

   以上をまとめたものを表 11 に示す。

   モデルの期待値とは、単純なモデルで 3 万試合行ない、その平均値を表している。実際の期

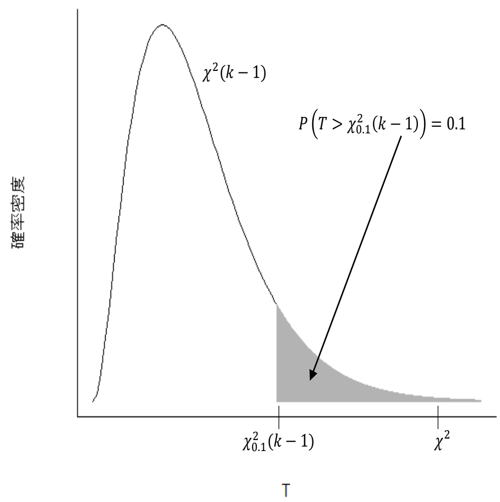

図 3　自由度 $k-1$ の χ^2 分布の確率密度関数と上側 10 パーセント点 $(\alpha = 0.1)$ の $\chi^2_{0.1}(k-1)$

表 11　適合度検定の結果 (単純なモデル)

チーム名	χ^2 値	自由度	p 値	モデルの期待値	実際の期待値
Mariners2001	24.5	11	0.011	4.60	5.72
Bluejays2001	19.2	10	0.004	3.91	4.73

待値とは、表 8 より $i \geq 0$ を満たす整数 i 点に対して

$$\sum_i i \times \frac{O_i}{162}$$

で求められる。p 値は $P(T > \chi^2\text{値})$ を指している。ここで、T は自由度 $k-1$ のカイ 2 乗分布に従う確率変数である。つまり、χ^2 が大きいほど p 値は小さくなることがわかり、χ^2 値は x 軸に対して原点から離れることになる。

χ^2 値は Mariners のとき 24.5、Bluejays のとき 19.2 となりこれは両チームともに $\chi^2_{0.1}(k-1)$ より大きくなっている。ここで、p 値は、自由度 $k-1$ のカイ 2 乗分布の密度関数を $\chi^2(t;k-1)$ で表せば、

$$(p\text{ 値}) = \int_{\chi^2}^{\infty} \chi^2(t;k-1)dt$$

とも表現することができ、p 値は Mariners のとき 0.011、Bluejays のとき 0.004 となりこれは両チームともに、有意水準 α より小さくなっている。つまり表 11 より、どちらのチームも $\alpha \geqq (p\text{ 値})$ であることから、棄却域に χ^2 値が含まれていることがわかり、帰無仮説が棄却されることになる。従って、対立仮説「実際の得点分布とシミュレート得点分布に差がある。」が採択されることになる。

2.4　考察

これまで扱ってきた単純なモデルと、実際の得点分布は乖離していることが統計的検定でわかった。これはランナーの進塁法則が、実際の野球と比較して簡略化しすぎているからだと考えられる。例えば、ランナー 2 塁のときにバッターの打撃結果が単打ならばランナーが生還することもあろうが、単純モデルではそれらが考慮されていない。もし、バッターの塁打数以上にランナーが進むのもありうることを考慮できれば、現実の野球を模倣することができ、単純なモデルと比較して得点分布の当てはまりが良くなるであろう。次章ではランナーの余剰進塁を始め、アウトも三振と凡打に分けて考慮するなど、もう少し現実の野球に近づけるためのモデリング法について論ずる。

3　改良モデル

単純なモデルは、統計的に信頼できるモデルではないことを前章で述べた。この章では、現実の野球に近づけるためにモデルをより精緻化することを考える。これを改良モデルと呼ぶことにする。

3.1　モデルの精緻化

3.1.1　余剰進塁

単純モデルでは、ランナーはバッターの塁打数に従って進塁していた。しかし、現実の野球ではランナーはバッターの塁打数より 1 つ多く進塁することもある。これを余剰進塁という。ここで、余剰進塁は以下の打撃結果の場合にあり得ることに注意されたい。

- 単打
- 二塁打
- 凡打 (アウト)

凡打についての詳細は、3.1.2 節で述べる。例えば、ランナー 1 塁のときに打撃結果が単打であったら、単純なモデルではランナー 1,2 塁となるが、余剰進塁を考慮することでランナー 1,2 塁と 1,3 塁の場合があり得ることになる。次に余剰進塁の部分をどのようにモデリングするかについて述べる。ここでは、1996~2005 年の 10 年分の Retrosheet データを用いて余剰進塁が起こる確率を求める。

方針としては以下の流れである。

表 12　余剰進塁の例

NUM	BAT_ID	PIT_ID	BASE1RUN_ID	BASE2RUN_ID	BASE3RUN_ID	EVENT_CD	RBI_CT
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
30339	suzui001	nomoh001*10				20	0
30340	mclem001	nomoh001	suzui001			20	0
30341	marte001	nomoh001	mclem001		suzui001	3	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

ここで、NUM= 30340 においてランナー 1 塁のときに EVENT_CD= 20 となっている。EVENT_CD は打撃結果を表しており、EVENT_CD= 20 は単打を示している。NUM= 30341 を参照すると、ランナーは 1,3 塁となっており、NUM= 30340 の 1 塁ランナーが NUM= 30341 では 3 塁まで到達していることがわかる。なぜなら、これらの RUN_ID が一致しているからである。これはつまり余剰進塁したということになる。以上のようにして、ランナー 1 塁かつ打撃結果が単

打の場合を全数として、そのうちランナー 1,3 塁となっている個数を求めれば、ランナー 1 塁かつ単打での余剰進塁を算出することができる。同様にしてその他のランナー・打撃結果についても余剰進塁が起こる確率を求め、それらをまとめたものを表 13 と表 14 に示す。

表 13　単打の余剰進塁

	打撃結果								
	ランナー	なし	1 塁	2 塁	3 塁	1,2 塁	1,3 塁	2,3 塁	満塁
打席時	1 塁					$*$	.2926		
	2 塁		$\underline{.5622}$				$*$		
	3 塁		$\underline{1.0}$						
	1,2 塁					$\underline{.3186}$	$\underline{.2712}$		$*$
	1,3 塁					$\underline{*}$	$\underline{.2709}$		
	2,3 塁		$\underline{\underline{.5418}}$				$\underline{*}$		
	満塁					$\underline{\underline{.3023}}$	$\underline{\underline{.2498}}$		$\underline{*}$

表 14　2 塁打の余剰進塁

	打撃結果								
	ランナー	なし	1 塁	2 塁	3 塁	1,2 塁	1,3 塁	2,3 塁	満塁
打席時	1 塁			$\underline{.4055}$				$*$	
	2 塁			$\underline{\underline{1.0}}$					
	3 塁			$\underline{\underline{1.0}}$					
	1,2 塁			$\underline{\underline{\underline{.4644}}}$				$\underline{*}$	
	1,3 塁			$\underline{\underline{\underline{.4222}}}$				$\underline{*}$	
	2,3 塁			$\underline{\underline{\underline{1.0}}}$					
	満塁			.4076[*11]				$\underline{\underline{*}}$	

表 13 と表 14 は、数値が入ってる部分は、それぞれデータから算出された余剰進塁する確率を表しており、$* = 1 -$ (余剰進塁の確率) は、バッターの塁打数に従って進塁する確率としており、例えば表 13 のランナー 1,2 塁の $*$ は、$* = 1 - (.3186 + .2712)$ にて求められる。また、下線部分は 1 点、二重下線部分は 2 点の打点を表している。

[*10] 野茂英雄投手、日本人プレイヤーとしてメジャーリーグで活躍し、トルネード投法で知られる。
[*11] 3 打点である。(表 14 を参照)

3.1.2　凡打

単純なモデルではアウトは三振のみとしていたが、改良モデルでのアウトは三振と凡打の 2 種類を考える。つまり､各選手 i に対して打撃結果の確率ベクトル $\boldsymbol{P_i} = (p_W, p_S, p_D, p_T, p_H, p_{so}, p_{poor})$ を再定義する。$p_W \sim p_H$ については単純なモデルと同様であり、p_{so} と p_{poor} については以下のように定義する。ここで、p_{so} は三振、p_{poor} は凡打となる確率である。

- $p_{so} = \dfrac{SO}{BOX}$
- $p_{poor} = 1 - (p_W + p_S + p_D + p_T + p_H + p_{so})$

つまり、バッターが出塁する事象と三振以外を凡打になるとしている。アウトについて三振と凡打に分ける利点として、例えばランナー 2 塁のときに凡打 (内野ゴロや外野フライ等) であったら、ランナーは 3 塁へ進塁することがあるであろう。凡打を導入することで、それらを考慮することが可能になる。またランナー 1 塁のときに凡打であったら、ランナーとバッターの両者がアウトになる併殺打 (ダブルプレー) もあり得るであろう。そしてランナー 3 塁のときに凡打 (外野フライ) であったら、タッチアップの可能性もある。以上を考慮した凡打での進塁法則を表 15 に示す。また、1996~2005 年の 10 年分の Retrosheet データを用いて凡打の場合での余剰進塁及び併殺打が起こる確率を求めたものを表 16 に示す。

表 15　凡打での進塁法則

ランナー	ランナーの進塁法則
1 塁	ダブルプレーか、そのままか、ランナーは 2 塁へ
2 塁	そのままか、ランナーは 3 塁へ
3 塁	タッチアップか、そのまま
1,2 塁	ランナー 1 塁とバッターのダブルプレーか、そのままか、ランナーは 2,3 塁へ
1,3 塁	ランナー 3 塁生還かつランナー 1 塁とバッターのダブルプレーか、タッチアップか、そのまま
2,3 塁	タッチアップか、そのまま
満塁	ランナー 3 塁生還かつランナー 1 塁とバッターのダブルプレーか、タッチアップか、ホームゲッツーか 、そのまま

表 16 は表 15 の項目ごとに順番を対応させている。また、数値$^{\#}$ の部分は併殺打、下線部分は 1 打点を表している。数値が入ってる部分は、3.1.1 節と同様に算出し、$* = 1 -$ (余剰進塁または併殺打の確率) は、ランナーの状況はそのままでアウトカウントが 1 つ増える確率としている。例えば、ランナー 1 塁のときに打撃結果が凡打であった場合、併殺打の確率は 0.2298、バッターがア

表 16　凡打の余剰進塁及び併殺打

	打撃結果								
	ランナー	なし	1 塁	2 塁	3 塁	1,2 塁	1,3 塁	2,3 塁	満塁
打席時	1 塁	$.2298^{\#}$	$*$	.1933					
	2 塁			$*$	.4659				
	3 塁	.5218			$*$				
	1,2 塁				$.2326^{\#}$	$*$		.1846	
	1,3 塁	$.2563^{\#}$	.2979				$*$		
	2,3 塁			.2951				$*$	
	満塁				$.1662^{\#}$	.2817		$.0596^{\#}$	$*$

ウトになりランナーが 2 塁へ進塁する確率は 0.1933 であるから、ランナーの状況が変化しない確率 $*$ は、$* = 1 - (0.2298 + 0.1933)$ となる。

3.2　モデリング

3.1 節では改良モデルを作成するための準備を行なった。前章と同様に 2001 年の Mariners 及び Bluejays に対して、確率ベクトル $P_i = (p_W, p_S, p_D, p_T, p_H, p_O, p_{so}, p_{poor})$ を算出する。表 17 は Mariners のデータから算出された確率行列である。

表 17　2001 年の Mariners データから算出された確率行列 (凡打も考慮)

打順	playerID	p_W	p_S	p_D	p_T	p_H	p_{so}	p_{poor}
1	suzukic01	0.0515	0.2602	0.0461	0.0108	0.0108	0.0718	0.5488
2	mclemma01	0.1417	0.1786	0.0329	0.0185	0.0103	0.1725	0.4456
3	martied01	0.1756	0.1377	0.0688	0.0017	0.0396	0.1549	0.4217
4	olerujo01	0.1458	0.1753	0.0471	0.0015	0.0309	0.1031	0.4963
5	boonebr01	0.0710	0.1870	0.0536	0.0043	0.0536	0.1594	0.4710
6	camermi01	0.1248	0.1327	0.0474	0.0079	0.0395	0.2449	0.4028
7	guillca01	0.1033	0.1683	0.0402	0.0076	0.0096	0.1702	0.5010
8	bellda01	0.0608	0.1549	0.0549	0.0000	0.0294	0.1157	0.5843
9	wilsoda01	0.0539	0.1691	0.0490	0.0025	0.0245	0.1691	0.5319

単純なモデルと同様にして、それぞれのチームの 2001 年の 162 試合において実際獲得した得点のヒストグラムと、改良モデルによる 3 万試合でのシミュレート分布を比較してみる。ここで改良モデルと単純なモデルの違いは、余剰進塁と併殺打を考慮している点であった。図 4、図 5 で

は、それぞれ 2001 年の Mariners・Bluejays の表 6 を基にした得点分布と改良モデルによるシミュレート得点分布を示している。

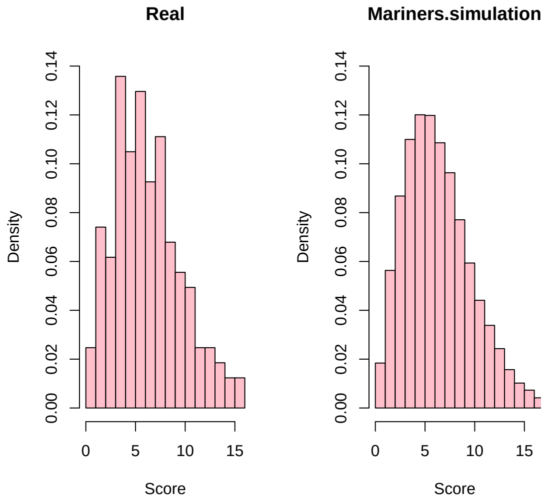

図 4　Mariners2001 の場合

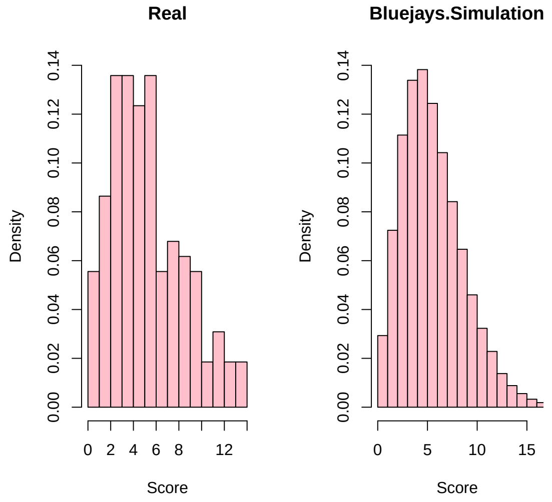

図 5　Bluejays2001 の場合

単純なモデルと比較して、改良モデルによるシミュレート分布は実際の得点分布を表しているように見え、良いモデルであることが示唆されている。3.3 節では改良モデルの妥当性について統計的検定を行ない、改良モデルの有意性について見る。

3.3　モデルの妥当性

この節では、改良モデルが統計的に妥当であるか検定を行なう。

表 18　Mariners における得点の幅を調整した改良モデル

得点	0∼1	2	3	4	5	6	7	8	9	10	11	12∼	合計
観測度数	16	10	22	17	21	15	18	11	9	8	4	11	162
理論確率	0.0747	0.0868	0.1100	0.1201	0.1198	0.1086	0.0963	0.0771	0.0594	0.0441	0.0338	0.0693	1
理論度数	12.1068	14.0616	17.8146	19.4508	19.4076	17.5932	15.6060	12.4848	9.6174	7.1442	5.4810	11.2320	162

表 19　適合度検定の結果 (改良モデル)

チーム名	χ^2 値	自由度	p 値	モデルの期待値	実際の期待値
Mariners2001	5.32	11	0.91	5.89	5.72
Bluejays2001	9.58	10	0.48	5.08	4.73

統計的検定の流れは、単純なモデルのときと同様である。表 18 より χ^2 値を求め、改良モデルでの p 値 と有意水準 α の関係を比較してみると、どちらのチームも (p 値) $\geqq \alpha$ であることから、棄却域に χ^2 値が含まれないことがわかり、帰無仮説が棄却されないことになる。従って、帰無仮説「実際の得点分布とシミュレート得点分布に差がない。」が採択されることになる。ここで、帰無仮説が正しいと言っているが、統計学的観点から厳密に言い換えると、帰無仮説を棄却するに足る証拠はないと意味していることに注意されたい。

3.4　考察

3.3 節では、有意水準 α において、「改良モデルの得点分布と実際の得点分布に差はない。」という帰無仮説が棄却されないということを見た。つまり、改良モデルは現実の野球のモデリングができているとしても無理はないということである。また、改良モデルは現実の野球の得点分布をモデリングできていることから、打順を任意に入れ替え、それぞれの期待得点を比較しても一般性は失われないであろう。後の章ではこれらを仮定した上で、期待得点をあげるためにどのような打順にするのが最適であるかについて論ずる。

4　最適打順

打順は様々なパターンが考えられる。ここでは期待得点が最も大きくなる打順を最適打順と呼ぶことにする。

4.1　最適打順を求める上での問題点

改良モデルを用いて、期待得点が最も大きくなる打順を求めるときにおいて問題点がある。それは、計算量が膨大になってしまうことである。野球において打者は 9 人であることから、考えられる打順の総数は 9!通りある。それぞれの打順に対して、1 万試合のシミュレーションを行なうのは現実的ではない。また、シミュレーション回数を減らしてしまうと、期待得点の精度において問題となってしまう。そこで、シミュレーションの際にマルチコアを有効活用することで計算時間の短縮を図り、これらの問題点を解決する。

4.1.1　foreach パッケージ

R 言語で並列処理を行なうツールとして、foreach パッケージが知られている。これは Revolution Analytics 社[*12]によって開発されており、for ループの処理を並列化できるものである。ここでは for ループの部分が試合数に相当する。試合数 (シミュレーション回数) の部分を複数のコアに分担させることで、計算時間の短縮を図る。

4.1.2　並列処理の流れ

並列処理の流れのイメージ図を図 6 に示す。図 6 は core 数=4 の場合を表している。CPU 内の各コアに対してシミュレーションの分担をさせることで計算時間の節減が可能になる。

4.1.3　foreach のスケーラビリティ

バッターについて、最初の 4 人について 4!通り調べ以降 5 人を固定したとき、それぞれ 1 万試合行ない期待得点を求め終える時間が並列処理によってどのように推移するか比較を行なった。結果を図 7 に示す。

4core を用いた並列処理の方が、逐次処理と比較して約 3 分の 1 の時間 (58 秒) で期待得点を算出することができた。ちなみに、4core で 5!通りの期待得点を求め終える時間は 292 秒であった。ここで、計算時間はパターン数に比例するという仮定をおけば、x!通り調べるのに

$$292 \times \frac{x!}{5!}$$

という式から、おおよその時間を算出することができる。それらをまとめたものを表 20 に示す。

[*12] http://www.r-analytics.jp/

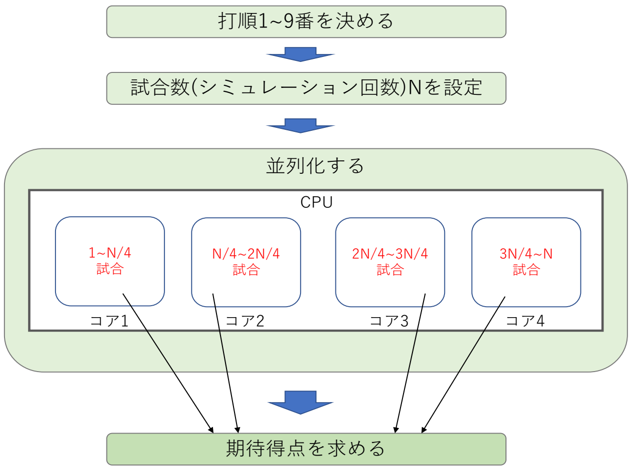

図 6　並列処理の流れ

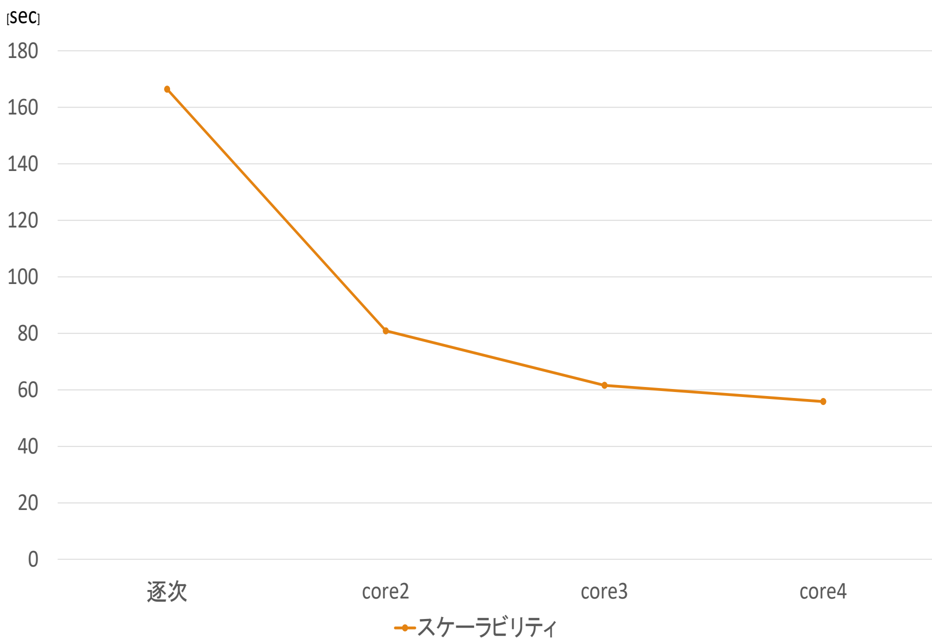

図 7　4!通りでの core 数によるスケーラビリティ

計算時間の都合上、今回は 9 番バッターのみを固定して残りの 8!通り (上位 8 人) を調べ、それぞれの打順の期待得点を求めることにする。

4.2　走塁能力を加味しない最適打順

まず走塁能力を加味しない最適打順とは、単に改良モデルに従うものとして求めた最適打順のことである。この節では、改良モデルにて特に当てはまりが良かった Mariners のデータについて、

表 20　4core でのパターン数別おおよその計算時間

パターン数	おおよその計算時間
6!	29min
7!	204min
8!	27.2h
9!(全パターン)	10days

表 21 の 9 番バッターを固定し、1∼8 番の打順を全てのパターン 8!通りについて、改良モデルに従ったシミュレーションを 1 万試合行ない、それぞれの期待得点を算出し、期待得点が大きい打順にはどのような傾向があるか考察を行なう。

表 21　2001 年の Mariners データから算出された確率行列 (再掲)

打順	playerID	p_W	p_S	p_D	p_T	p_H	p_{so}	p_{poor}
1	suzukic01	0.0515	0.2602	0.0461	0.0108	0.0108	0.0718	0.5488
2	mclemma01	0.1417	0.1786	0.0329	0.0185	0.0103	0.1725	0.4456
3	martied01	0.1756	0.1377	0.0688	0.0017	0.0396	0.1549	0.4217
4	olerujo01	0.1458	0.1753	0.0471	0.0015	0.0309	0.1031	0.4963
5	boonebr01	0.0710	0.1870	0.0536	0.0043	0.0536	0.1594	0.4710
6	camermi01	0.1248	0.1327	0.0474	0.0079	0.0395	0.2449	0.4028
7	guillca01	0.1033	0.1683	0.0402	0.0076	0.0096	0.1702	0.5010
8	bellda01	0.0608	0.1549	0.0549	0.0000	0.0294	0.1157	0.5843
9	wilsoda01	0.0539	0.1691	0.0490	0.0025	0.0245	0.1691	0.5319

4.2.1　セイバーメトリクス指標

最適打順を考察する上で必要になるセイバーメトリクス指標について述べる。まず、我々がよく耳にする打率について述べておこう。打率 (AVG[*13]) は以下の計算式で定義される。

$$(\text{打率}) = \frac{(\text{安打})}{(\text{打数})}$$

ちなみに、シーズン中において打率が最も高い選手は首位打者[*14]と呼ばれ、これに加えて本塁打王、打点王の 3 つのタイトルを独占した選手は三冠王と称され栄誉ある記録といわれている。しかし、打点については個人の得点能力というよりは、チームの打撃結果に影響されるものである。つ

[*13] average
[*14] 規定打席 (= 所属球団の試合数 × 3.1) を満たしている必要がある。

まり、打率・本塁打・打点といわれる指標は直感的でわかりやすいが、必ずしも個人の能力を表しているわけではないことに注意されたい。セイバーメトリクスとは、これらの指標以上に個人の能力を示す指標を用いて評価するものである。

次に、セイバーメトリクス指標の出塁率 (OBP*15) は以下の計算式で定義される。

$$(\text{出塁率}) = \frac{(\text{安打}) + (\text{四死球})}{(\text{打数}) + (\text{四死球}) + (\text{犠飛})}$$

この指標の利点として、打率と比較してより選手の出塁する確率を表すことができている。なぜなら、打率は安打のみに焦点を当てており四死球は考慮されていないからである。例えば、ランナーなしのときに打撃結果が単打または四死球であったら、結果的には同じランナーの状況になるので、それら両方を考慮できる出塁率は有用であることがわかる。

さらに、長打率 (SLG*16) は以下のように定義される。

$$(\text{長打率}) = \frac{(\text{単打}) \times 1 + (2\ \text{塁打}) \times 2 + (3\ \text{塁打}) \times 3 + (\text{本塁打}) \times 4}{(\text{打数})}$$

長打力のある選手は、長打率が大きくなることがわかる。また、出塁率と長打率の和は OPS*17と呼ばれ、

$$(OPS) = (\text{出塁率}) + (\text{長打率})$$

OPS という指標は、1984 年ごろにセイバーメトリクスの立役者ビル・ジェームズ氏らによって考案され、チームの得点と OPS の相関係数は 0.95 となることから、選手の得点能力をとても良く表せている指標として知られている [5, p.28-29]。表 22 は、各選手ごとに打率及び先述のセイバーメトリクス指標をまとめたものである。

打率が高いからといって、出塁率も高くなるとは限らないという事実が見て取れる。

4.2.2 考察

シミュレーションによってそれぞれの打順について期待得点を求めた。最高点は 6.0046 点、最低点は 5.7353 点となった。図 8、図 9、図 10、図 11 は期待得点の上位 30 位及び打順を示したものであり、表 22 を参照してそれぞれの指標の上位・中位・下位 3 人を赤・緑・青で色分けした。

以上から、得られた最適打順の傾向を考察していく。まず、出塁率 (OBP) において打席番号 1∼9 番について赤色が占める割合を求めると、1 番バッターを先頭に降順になった。これは出塁率の高い選手は上位打線に配置し、打撃機会を増やすことで得点力が上がるものだと考えられる。また、OPS の大きい選手は 2∼5 番に集中しており、2∼5 番で OPS 上位 3 位が連なっていることが多い。そして、下位打線には打撃成果が乏しい選手が中心となった。

*15 on base percentage
*16 slugging percentage
*17 on base plus slugging percentage

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.0046	martied01	mclemma01	olerujo01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
2	5.9897	camermi01	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
3	5.9895	mclemma01	olerujo01	martied01	boonebr01	bellda01	camermi01	suzukic01	guillca01	wilsoda01
4	5.9864	mclemma01	olerujo01	martied01	camermi01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
5	5.9829	camermi01	martied01	mclemma01	suzukic01	olerujo01	boonebr01	bellda01	guillca01	wilsoda01
6	5.9824	camermi01	mclemma01	suzukic01	olerujo01	martied01	boonebr01	guillca01	bellda01	wilsoda01
7	5.9812	martied01	suzukic01	mclemma01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	5.9792	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	camermi01	guillca01	wilsoda01
9	5.979	mclemma01	suzukic01	martied01	boonebr01	olerujo01	camermi01	bellda01	guillca01	wilsoda01
10	5.9783	suzukic01	camermi01	olerujo01	martied01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
11	5.9782	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
12	5.9776	mclemma01	martied01	boonebr01	camermi01	olerujo01	suzukic01	bellda01	guillca01	wilsoda01
13	5.9775	olerujo01	martied01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	camermi01	wilsoda01
14	5.9744	olerujo01	mclemma01	martied01	suzukic01	boonebr01	guillca01	camermi01	bellda01	wilsoda01
15	5.9737	mclemma01	martied01	suzukic01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
16	5.9732	mclemma01	olerujo01	martied01	boonebr01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
17	5.973	martied01	camermi01	mclemma01	suzukic01	boonebr01	olerujo01	guillca01	bellda01	wilsoda01
18	5.9728	olerujo01	suzukic01	mclemma01	martied01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
19	5.9727	olerujo01	martied01	boonebr01	camermi01	mclemma01	suzukic01	bellda01	guillca01	wilsoda01
20	5.9726	martied01	suzukic01	boonebr01	olerujo01	camermi01	mclemma01	guillca01	bellda01	wilsoda01
21	5.972	mclemma01	martied01	olerujo01	boonebr01	suzukic01	camermi01	bellda01	guillca01	wilsoda01
22	5.972	camermi01	olerujo01	mclemma01	martied01	suzukic01	boonebr01	bellda01	guillca01	wilsoda01
23	5.9718	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
24	5.9701	mclemma01	olerujo01	camermi01	martied01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
25	5.9698	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
26	5.9698	olerujo01	martied01	mclemma01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
27	5.9697	martied01	olerujo01	suzukic01	mclemma01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
28	5.9695	mclemma01	martied01	olerujo01	suzukic01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
29	5.9694	mclemma01	olerujo01	camermi01	martied01	boonebr01	bellda01	guillca01	suzukic01	wilsoda01
30	5.9693	mclemma01	martied01	olerujo01	boonebr01	camermi01	guillca01	suzukic01	bellda01	wilsoda01

図 8　AVG で色分け

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.0046	martied01	mclemma01	olerujo01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
2	5.9897	camermi01	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
3	5.9895	mclemma01	olerujo01	martied01	boonebr01	bellda01	camermi01	suzukic01	guillca01	wilsoda01
4	5.9864	mclemma01	olerujo01	martied01	camermi01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
5	5.9829	camermi01	martied01	mclemma01	suzukic01	olerujo01	boonebr01	bellda01	guillca01	wilsoda01
6	5.9824	camermi01	mclemma01	suzukic01	olerujo01	martied01	boonebr01	guillca01	bellda01	wilsoda01
7	5.9812	martied01	suzukic01	mclemma01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	5.9792	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	camermi01	guillca01	wilsoda01
9	5.979	mclemma01	suzukic01	martied01	boonebr01	olerujo01	camermi01	bellda01	guillca01	wilsoda01
10	5.9783	suzukic01	camermi01	olerujo01	martied01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
11	5.9782	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
12	5.9776	mclemma01	martied01	boonebr01	camermi01	olerujo01	suzukic01	bellda01	guillca01	wilsoda01
13	5.9775	olerujo01	martied01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	camermi01	wilsoda01
14	5.9744	olerujo01	mclemma01	martied01	suzukic01	boonebr01	guillca01	camermi01	bellda01	wilsoda01
15	5.9737	mclemma01	martied01	suzukic01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
16	5.9732	mclemma01	olerujo01	martied01	boonebr01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
17	5.973	martied01	camermi01	mclemma01	suzukic01	boonebr01	olerujo01	guillca01	bellda01	wilsoda01
18	5.9728	olerujo01	suzukic01	mclemma01	martied01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
19	5.9727	olerujo01	martied01	boonebr01	camermi01	mclemma01	suzukic01	bellda01	guillca01	wilsoda01
20	5.9726	martied01	suzukic01	boonebr01	olerujo01	camermi01	mclemma01	guillca01	bellda01	wilsoda01
21	5.972	mclemma01	martied01	olerujo01	boonebr01	suzukic01	camermi01	bellda01	guillca01	wilsoda01
22	5.972	camermi01	olerujo01	mclemma01	martied01	suzukic01	boonebr01	bellda01	guillca01	wilsoda01
23	5.9718	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
24	5.9701	mclemma01	olerujo01	camermi01	martied01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
25	5.9698	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
26	5.9698	olerujo01	martied01	mclemma01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
27	5.9697	martied01	olerujo01	suzukic01	mclemma01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
28	5.9695	mclemma01	martied01	olerujo01	suzukic01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
29	5.9694	mclemma01	olerujo01	camermi01	martied01	boonebr01	bellda01	guillca01	suzukic01	wilsoda01
30	5.9693	mclemma01	martied01	olerujo01	boonebr01	camermi01	guillca01	suzukic01	bellda01	wilsoda01

図 9　OBP で色分け

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.0046	martied01	mclemma01	olerujo01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
2	5.9897	camermi01	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
3	5.9895	mclemma01	olerujo01	martied01	boonebr01	bellda01	camermi01	suzukic01	guillca01	wilsoda01
4	5.9864	mclemma01	olerujo01	martied01	camermi01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
5	5.9829	camermi01	martied01	mclemma01	suzukic01	olerujo01	boonebr01	bellda01	guillca01	wilsoda01
6	5.9824	camermi01	mclemma01	suzukic01	olerujo01	martied01	boonebr01	guillca01	bellda01	wilsoda01
7	5.9812	martied01	suzukic01	mclemma01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	5.9792	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	camermi01	guillca01	wilsoda01
9	5.979	mclemma01	suzukic01	martied01	boonebr01	olerujo01	camermi01	bellda01	guillca01	wilsoda01
10	5.9783	suzukic01	camermi01	olerujo01	martied01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
11	5.9782	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
12	5.9776	mclemma01	martied01	boonebr01	camermi01	olerujo01	suzukic01	bellda01	guillca01	wilsoda01
13	5.9775	olerujo01	martied01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	camermi01	wilsoda01
14	5.9744	olerujo01	mclemma01	martied01	suzukic01	boonebr01	guillca01	camermi01	bellda01	wilsoda01
15	5.9737	mclemma01	martied01	suzukic01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
16	5.9732	mclemma01	olerujo01	martied01	boonebr01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
17	5.973	martied01	camermi01	mclemma01	suzukic01	boonebr01	olerujo01	guillca01	bellda01	wilsoda01
18	5.9728	olerujo01	suzukic01	mclemma01	martied01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
19	5.9727	olerujo01	martied01	boonebr01	camermi01	mclemma01	suzukic01	bellda01	guillca01	wilsoda01
20	5.9726	martied01	suzukic01	boonebr01	olerujo01	camermi01	mclemma01	guillca01	bellda01	wilsoda01
21	5.972	mclemma01	martied01	olerujo01	boonebr01	suzukic01	camermi01	bellda01	guillca01	wilsoda01
22	5.972	camermi01	olerujo01	mclemma01	martied01	suzukic01	boonebr01	bellda01	guillca01	wilsoda01
23	5.9718	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
24	5.9701	mclemma01	olerujo01	camermi01	martied01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
25	5.9698	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
26	5.9698	olerujo01	martied01	mclemma01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
27	5.9697	martied01	olerujo01	suzukic01	mclemma01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
28	5.9695	mclemma01	martied01	olerujo01	suzukic01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
29	5.9694	mclemma01	olerujo01	camermi01	martied01	boonebr01	bellda01	guillca01	suzukic01	wilsoda01
30	5.9693	mclemma01	martied01	olerujo01	boonebr01	camermi01	guillca01	suzukic01	bellda01	wilsoda01

図 10　SLG で色分け

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.0046	martied01	mclemma01	olerujo01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
2	5.9897	camermi01	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
3	5.9895	mclemma01	olerujo01	martied01	boonebr01	bellda01	camermi01	suzukic01	guillca01	wilsoda01
4	5.9864	mclemma01	olerujo01	martied01	camermi01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
5	5.9829	camermi01	martied01	mclemma01	suzukic01	olerujo01	boonebr01	bellda01	guillca01	wilsoda01
6	5.9824	camermi01	mclemma01	suzukic01	olerujo01	martied01	boonebr01	guillca01	bellda01	wilsoda01
7	5.9812	martied01	suzukic01	mclemma01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	5.9792	mclemma01	olerujo01	martied01	boonebr01	suzukic01	bellda01	camermi01	guillca01	wilsoda01
9	5.979	mclemma01	suzukic01	martied01	boonebr01	olerujo01	camermi01	bellda01	guillca01	wilsoda01
10	5.9783	suzukic01	camermi01	olerujo01	martied01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
11	5.9782	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	bellda01	guillca01	wilsoda01
12	5.9776	mclemma01	martied01	boonebr01	camermi01	olerujo01	suzukic01	bellda01	guillca01	wilsoda01
13	5.9775	olerujo01	martied01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	camermi01	wilsoda01
14	5.9744	olerujo01	mclemma01	martied01	suzukic01	boonebr01	guillca01	camermi01	bellda01	wilsoda01
15	5.9737	mclemma01	martied01	suzukic01	olerujo01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
16	5.9732	mclemma01	olerujo01	martied01	boonebr01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
17	5.973	martied01	camermi01	mclemma01	suzukic01	boonebr01	olerujo01	guillca01	bellda01	wilsoda01
18	5.9728	olerujo01	suzukic01	mclemma01	martied01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
19	5.9727	olerujo01	martied01	boonebr01	camermi01	mclemma01	suzukic01	bellda01	guillca01	wilsoda01
20	5.9726	martied01	suzukic01	boonebr01	olerujo01	camermi01	mclemma01	guillca01	bellda01	wilsoda01
21	5.972	mclemma01	martied01	olerujo01	boonebr01	suzukic01	camermi01	bellda01	guillca01	wilsoda01
22	5.972	camermi01	olerujo01	mclemma01	martied01	suzukic01	boonebr01	bellda01	guillca01	wilsoda01
23	5.9718	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	mclemma01	guillca01	wilsoda01
24	5.9701	mclemma01	olerujo01	camermi01	martied01	suzukic01	boonebr01	guillca01	bellda01	wilsoda01
25	5.9698	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
26	5.9698	olerujo01	martied01	mclemma01	suzukic01	boonebr01	camermi01	bellda01	guillca01	wilsoda01
27	5.9697	martied01	olerujo01	suzukic01	mclemma01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
28	5.9695	mclemma01	martied01	olerujo01	suzukic01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
29	5.9694	mclemma01	olerujo01	camermi01	martied01	boonebr01	bellda01	guillca01	suzukic01	wilsoda01
30	5.9693	mclemma01	martied01	olerujo01	boonebr01	camermi01	guillca01	suzukic01	bellda01	wilsoda01

図 11　OPS で色分け

表 22　2001 年の Mariners データから算出された打率及びセイバーメトリクス指標

打順	playerID	AVG	OBP	SLG	OPS
1	suzukic01	0.3497	0.3815	0.4566	0.8381
2	mclemma01	0.2861	0.3843	0.4059	0.7902
3	martied01	0.3064	0.4234	0.5426	0.9660
4	olerujo01	0.3024	0.4012	0.4720	0.8732
5	boonebr01	0.3307	0.3723	0.5778	0.9501
6	camermi01	0.2667	0.3528	0.4796	0.8325
7	guillca01	0.2588	0.3333	0.3553	0.6886
8	bellda01	0.2596	0.3030	0.4149	0.7179
9	wilsoda01	0.2653	0.3050	0.4032	0.7082

4.3　走塁能力を加味した最適打順

野球において、打撃成績は重要になるのは勿論のこと、走塁能力もチームが勝利する上において重要なことである。理由の 1 つとしては、今回は研究対象としていないが、走塁能力に優れていれば守備範囲が広がり相手のチームに得点を与えにくくすることに直結するからである。この節ではモデリングの手法として、盗塁成功率に着目することで走塁能力について定義し、走塁能力を加味しない場合と比較して、どのような違いが見られるか 4.2 節と同様にして考察を行なっていく。

4.3.1　得点期待値

走塁能力を加味する上で必要となる得点期待値の概念について述べる。得点期待値とは、あるアウトカウント・ランナー状況 (例えば、1 アウトのランナー 1,2 塁など) がイニング中において初めて発生してから、イニング終了までに獲得された得点を求めシチュエーションごとの期待値として定義したものである。得点期待値の具体的な求め方を表したものを表 23 に示す。

表 23　得点期待値の求め方の例

打者名	###	A	B	C	D	E	F	G	H	I
打撃結果	(イニングの始まり)	単打	三振	単打	二塁打	四球	二塁打	三振	単打	三振
アウトカウント	(0)	0	1	1	1	1	1	2	2	3
ランナー状況	(なし)	1 塁	1 塁	1,2 塁	2 塁	1,2 塁	2 塁	2 塁	1 塁	なし
現得点	(1)	1	1	1	3	3	5	5	6	6

表 23 はある 1 イニングにおける打撃結果・アウト・ランナー状況の例を表しており、このチー

ムはイニング始まり時において既に 1 点獲得しており、このイニングは打者 A から始まり打者 I で終了していることがわかる。得点期待値の定義に従って計算すると、それぞれの得点期待値は

- アウトカウント 0・ランナーなし : $6-1=5$ 点
- アウトカウント 0・ランナー 1 塁 : $6-1=5$ 点
- アウトカウント 1・ランナー 1 塁 : $6-1=5$ 点
- アウトカウント 1・ランナー 1,2 塁 : $6-1=5$ 点
- アウトカウント 1・ランナー 2 塁 : $6-3=3$ 点
- アウトカウント 2・ランナー 1 塁 : $6-6=0$ 点
- アウトカウント 2・ランナー 2 塁 : $6-5=1$ 点

となることがわかる。上記の例について、例えば 0 アウトランナーなしの状況が初めて起こった時の得点は 1 点であり、イニング終了時は 6 点であることから、この状況の得点期待値は $6-1=5$ となる。また、1 アウトランナー 2 塁の状況は、打者名 D 及び F の 2 通りあるが、当該状況が初めて発生してから何点獲得できるかということより打者名 D を参照すると、現得点は 3 点であり、そしてイニング終了時の得点は 6 点であるから得点期待値は $6-3=3$ となる。今回は具体例として 1 イニングのみを対象としたが、過去のメジャーリーグのデータを用いることで、全てのアウトカウント・ランナー状況別 (= 24 通り) の得点期待値を求めることができる。表 24 は 1996~2005 年のデータから、全てのアウトカウント・ランナー状況別の得点期待値を算出したものである。

表 24　得点期待値

		ランナー							
		なし	1 塁	2 塁	1,2 塁	3 塁	1,3 塁	2,3 塁	満塁
ア	0 アウト	.5333	.9226	1.1661	1.5233	1.4364	1.8438	2.0303	2.3296
ウ	1 アウト	.2849	.5504	.6991	.9392	.9775	1.1901	1.4051	1.5629
ト	2 アウト	.1100	.2374	.3307	.4391	.3769	.5121	.5782	.7443

例えば、表 24 より、ノーアウト満塁の状況になったらイニング終了までの得点期待値は 2.3296 点だということがわかる。また、1 イニングの期待得点は 0.5333 点ということもわかる[*18]。

4.3.2　前提条件

ここで、盗塁のルールについて説明する。盗塁はランナーが 1 塁のときのみ可能性があるとして、三盗や本盗、ダブルスチール等は考慮しない。そして、そのランナーが盗塁をするか否かの判断の基準は表 24 を用いることにする。例えば、ノーアウトランナー 1 塁のときに盗塁成功率 q_0 がどのくらいであれば期待得点が上昇するかを考える。表 24 より、以下の式で評価をすることがで

[*18] イニングの始まりは、0 アウト・ランナーなしであるから。

きる。

$$0.9226 < 1.1661q_0 + 0.2849(1 - q_0) \tag{2}$$
$$q_0 > 0.7237 \tag{3}$$

上の例では、盗塁がもし成功すればノーアウトランナー 2 塁、失敗すれば 1 アウトランナーなしとなる。つまり、盗塁をするべきか否かを盗塁成功率の観点から判断していることになる。1 アウト・2 アウトについても同様にして求めると、$\boldsymbol{Q} = (q_0, q_1, q_2) = (0.7237, 0.7476, 0.7179)$ が閾値となり、これより大きい盗塁成功率を持つランナーであれば、盗塁をさせるべきという判断になる。

ここで、各選手の盗塁成功率 p_{steal} を

- $p_{steal} = \frac{SB}{SB+CS}$

と定義する。ここで、SB は盗塁成功数、CS は盗塁失敗数を表している。また、盗塁試行回数が極端に少ない選手は $p_{steal} = 0$ としている。

表 25　盗塁成功率も含んだ 2001 年の Mariners データから算出された確率行列

打順	playerID	p_W	p_S	p_D	p_T	p_H	p_{so}	p_{poor}	p_{steal}
1	suzukic01	0.0515	0.2602	0.0461	0.0108	0.0108	0.0718	0.5488	0.8
2	mclemma01	0.1417	0.1786	0.0329	0.0185	0.0103	0.1725	0.4456	0.8478
3	martied01	0.1756	0.1377	0.0688	0.0017	0.0396	0.1549	0.4217	0
4	olerujo01	0.1458	0.1753	0.0471	0.0015	0.0309	0.1031	0.4963	0
5	boonebr01	0.0710	0.1870	0.0536	0.0043	0.0536	0.1594	0.4710	0
6	camermi01	0.1248	0.1327	0.0474	0.0079	0.0395	0.2449	0.4028	0.8718
7	guillca01	0.1033	0.1683	0.0402	0.0076	0.0096	0.1702	0.5010	0
8	bellda01	0.0608	0.1549	0.0549	0.0000	0.0294	0.1157	0.5843	0
9	wilsoda01	0.0539	0.1691	0.0490	0.0025	0.0245	0.1691	0.5319	0

走塁能力を加味した最適打順モデルとは、今までの改良モデルに加えて盗塁成功率に着目することで、走塁能力も考慮しようというものである。

4.3.3　考察

4.2 節と同様にして、それぞれの打順に対して期待得点を求める。走塁能力を加味したモデルでは最高点は 6.2967 点、最低点は 5.991 点となった。図 12、図 13、図 14、図 15 は期待得点の上位 30 位及び打順を示したものであり、それぞれの指標の上位・中位・下位 3 人を赤・緑・青で色分けした。なお、表 25 の盗塁成功率が $\boldsymbol{Q}$ の全ての要素 (閾値) に対して上回る選手を斜体・太字にして表示している。

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.2967	mclemma01	guillca01	martied01	boonebr01	olerujo01	camermi01	suzukic01	bellda01	wilsoda01
2	6.2711	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
3	6.2702	mclemma01	suzukic01	guillca01	martied01	boonebr01	olerujo01	camermi01	bellda01	wilsoda01
4	6.2698	mclemma01	suzukic01	boonebr01	martied01	olerujo01	camermi01	guillca01	bellda01	wilsoda01
5	6.2677	camermi01	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
6	6.2603	mclemma01	olerujo01	guillca01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
7	6.2588	mclemma01	suzukic01	olerujo01	martied01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	6.2554	mclemma01	suzukic01	olerujo01	martied01	guillca01	boonebr01	camermi01	bellda01	wilsoda01
9	6.2551	martied01	camermi01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
10	6.2542	mclemma01	suzukic01	olerujo01	guillca01	martied01	boonebr01	camermi01	bellda01	wilsoda01
11	6.2529	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
12	6.2527	guillca01	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
13	6.2526	camermi01	boonebr01	guillca01	martied01	olerujo01	mclemma01	suzukic01	bellda01	wilsoda01
14	6.2523	mclemma01	olerujo01	martied01	boonebr01	guillca01	camermi01	suzukic01	bellda01	wilsoda01
15	6.2493	mclemma01	olerujo01	boonebr01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
16	6.2485	mclemma01	suzukic01	boonebr01	martied01	guillca01	olerujo01	camermi01	bellda01	wilsoda01
17	6.248	camermi01	martied01	olerujo01	mclemma01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
18	6.2477	martied01	olerujo01	boonebr01	camermi01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
19	6.2477	camermi01	suzukic01	guillca01	boonebr01	olerujo01	mclemma01	martied01	bellda01	wilsoda01
20	6.2462	martied01	olerujo01	mclemma01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
21	6.2461	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
22	6.2444	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
23	6.2441	mclemma01	suzukic01	boonebr01	martied01	olerujo01	guillca01	camermi01	bellda01	wilsoda01
24	6.244	mclemma01	olerujo01	guillca01	boonebr01	camermi01	martied01	suzukic01	bellda01	wilsoda01
25	6.2439	mclemma01	suzukic01	boonebr01	martied01	camermi01	olerujo01	guillca01	bellda01	wilsoda01
26	6.2432	camermi01	olerujo01	guillca01	martied01	mclemma01	suzukic01	boonebr01	bellda01	wilsoda01
27	6.2426	mclemma01	suzukic01	boonebr01	olerujo01	martied01	camermi01	guillca01	bellda01	wilsoda01
28	6.2416	mclemma01	guillca01	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	wilsoda01
29	6.2404	mclemma01	suzukic01	guillca01	martied01	camermi01	boonebr01	olerujo01	bellda01	wilsoda01
30	6.2402	camermi01	suzukic01	martied01	olerujo01	boonebr01	mclemma01	guillca01	bellda01	wilsoda01

図 12　AVG で色分け (走塁能力を加味)

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.2967	mclemma01	guillca01	martied01	boonebr01	olerujo01	camermi01	suzukic01	bellda01	wilsoda01
2	6.2711	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
3	6.2702	mclemma01	suzukic01	guillca01	martied01	boonebr01	olerujo01	camermi01	bellda01	wilsoda01
4	6.2698	mclemma01	suzukic01	boonebr01	martied01	olerujo01	camermi01	guillca01	bellda01	wilsoda01
5	6.2677	camermi01	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
6	6.2603	mclemma01	olerujo01	guillca01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
7	6.2588	mclemma01	suzukic01	olerujo01	martied01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	6.2554	mclemma01	suzukic01	olerujo01	martied01	guillca01	boonebr01	camermi01	bellda01	wilsoda01
9	6.2551	martied01	camermi01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
10	6.2542	mclemma01	suzukic01	olerujo01	guillca01	martied01	boonebr01	camermi01	bellda01	wilsoda01
11	6.2529	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
12	6.2527	guillca01	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
13	6.2526	camermi01	boonebr01	guillca01	martied01	olerujo01	mclemma01	suzukic01	bellda01	wilsoda01
14	6.2523	mclemma01	olerujo01	martied01	boonebr01	guillca01	camermi01	suzukic01	bellda01	wilsoda01
15	6.2493	mclemma01	olerujo01	boonebr01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
16	6.2485	mclemma01	suzukic01	boonebr01	martied01	guillca01	olerujo01	camermi01	bellda01	wilsoda01
17	6.248	camermi01	martied01	olerujo01	mclemma01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
18	6.2477	martied01	olerujo01	boonebr01	camermi01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
19	6.2477	camermi01	suzukic01	guillca01	boonebr01	olerujo01	mclemma01	martied01	bellda01	wilsoda01
20	6.2462	martied01	olerujo01	mclemma01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
21	6.2461	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
22	6.2444	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
23	6.2441	mclemma01	suzukic01	boonebr01	martied01	olerujo01	guillca01	camermi01	bellda01	wilsoda01
24	6.244	mclemma01	olerujo01	guillca01	boonebr01	camermi01	martied01	suzukic01	bellda01	wilsoda01
25	6.2439	mclemma01	suzukic01	boonebr01	martied01	camermi01	olerujo01	guillca01	bellda01	wilsoda01
26	6.2432	camermi01	olerujo01	guillca01	martied01	mclemma01	suzukic01	boonebr01	bellda01	wilsoda01
27	6.2426	mclemma01	suzukic01	boonebr01	olerujo01	martied01	camermi01	guillca01	bellda01	wilsoda01
28	6.2416	mclemma01	guillca01	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	wilsoda01
29	6.2404	mclemma01	suzukic01	guillca01	martied01	camermi01	boonebr01	olerujo01	bellda01	wilsoda01
30	6.2402	camermi01	suzukic01	martied01	olerujo01	boonebr01	mclemma01	guillca01	bellda01	wilsoda01

図 13　OBP で色分け (走塁能力を加味)

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.2967	***mclemma01***	guillca01	martied01	boonebr01	olerujo01	***camermi01***	***suzukic01***	bellda01	wilsoda01
2	6.2711	***mclemma01***	boonebr01	olerujo01	martied01	***camermi01***	***suzukic01***	guillca01	bellda01	wilsoda01
3	6.2702	***mclemma01***	***suzukic01***	guillca01	martied01	boonebr01	olerujo01	***camermi01***	bellda01	wilsoda01
4	6.2698	***mclemma01***	***suzukic01***	boonebr01	martied01	olerujo01	***camermi01***	guillca01	bellda01	wilsoda01
5	6.2677	***camermi01***	martied01	olerujo01	boonebr01	***mclemma01***	***suzukic01***	guillca01	bellda01	wilsoda01
6	6.2603	***mclemma01***	olerujo01	guillca01	martied01	boonebr01	***camermi01***	***suzukic01***	bellda01	wilsoda01
7	6.2588	***mclemma01***	***suzukic01***	olerujo01	martied01	boonebr01	***camermi01***	guillca01	bellda01	wilsoda01
8	6.2554	***mclemma01***	***suzukic01***	olerujo01	martied01	guillca01	boonebr01	***camermi01***	bellda01	wilsoda01
9	6.2551	martied01	***camermi01***	olerujo01	boonebr01	***mclemma01***	***suzukic01***	guillca01	bellda01	wilsoda01
10	6.2542	***mclemma01***	***suzukic01***	olerujo01	guillca01	martied01	boonebr01	***camermi01***	bellda01	wilsoda01
11	6.2529	***mclemma01***	olerujo01	martied01	boonebr01	***camermi01***	***suzukic01***	guillca01	bellda01	wilsoda01
12	6.2527	guillca01	***mclemma01***	olerujo01	martied01	boonebr01	***camermi01***	***suzukic01***	bellda01	wilsoda01
13	6.2526	***camermi01***	boonebr01	guillca01	martied01	olerujo01	***mclemma01***	***suzukic01***	bellda01	wilsoda01
14	6.2523	***mclemma01***	olerujo01	martied01	boonebr01	guillca01	***camermi01***	***suzukic01***	bellda01	wilsoda01
15	6.2493	***mclemma01***	olerujo01	boonebr01	martied01	***camermi01***	***suzukic01***	guillca01	bellda01	wilsoda01
16	6.2485	***mclemma01***	***suzukic01***	boonebr01	martied01	guillca01	olerujo01	***camermi01***	bellda01	wilsoda01
17	6.248	***camermi01***	martied01	olerujo01	***mclemma01***	boonebr01	***suzukic01***	guillca01	bellda01	wilsoda01
18	6.2477	martied01	olerujo01	boonebr01	***camermi01***	***mclemma01***	***suzukic01***	guillca01	bellda01	wilsoda01
19	6.2477	***camermi01***	***suzukic01***	guillca01	boonebr01	olerujo01	***mclemma01***	martied01	bellda01	wilsoda01
20	6.2462	martied01	olerujo01	***mclemma01***	boonebr01	***camermi01***	***suzukic01***	guillca01	bellda01	wilsoda01
21	6.2461	***mclemma01***	***camermi01***	olerujo01	martied01	boonebr01	***suzukic01***	guillca01	bellda01	wilsoda01
22	6.2444	martied01	olerujo01	boonebr01	***mclemma01***	***suzukic01***	guillca01	***camermi01***	bellda01	wilsoda01
23	6.2441	***mclemma01***	***suzukic01***	boonebr01	martied01	olerujo01	guillca01	***camermi01***	bellda01	wilsoda01
24	6.244	***mclemma01***	olerujo01	guillca01	boonebr01	***camermi01***	martied01	***suzukic01***	bellda01	wilsoda01
25	6.2439	***mclemma01***	***suzukic01***	boonebr01	martied01	***camermi01***	olerujo01	guillca01	bellda01	wilsoda01
26	6.2432	***camermi01***	olerujo01	guillca01	martied01	***mclemma01***	***suzukic01***	boonebr01	bellda01	wilsoda01
27	6.2426	***mclemma01***	***suzukic01***	boonebr01	olerujo01	martied01	***camermi01***	guillca01	bellda01	wilsoda01
28	6.2416	***mclemma01***	guillca01	olerujo01	martied01	***camermi01***	***suzukic01***	boonebr01	bellda01	wilsoda01
29	6.2404	***mclemma01***	***suzukic01***	guillca01	martied01	***camermi01***	boonebr01	olerujo01	bellda01	wilsoda01
30	6.2402	***camermi01***	***suzukic01***	martied01	olerujo01	boonebr01	***mclemma01***	guillca01	bellda01	wilsoda01

図 14　SLG で色分け (走塁能力を加味)

順位	期待得点	1番バッター	2番バッター	3番バッター	4番バッター	5番バッター	6番バッター	7番バッター	8番バッター	9番バッター
1	6.2967	mclemma01	guillca01	martied01	boonebr01	olerujo01	camermi01	suzukic01	bellda01	wilsoda01
2	6.2711	mclemma01	boonebr01	olerujo01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
3	6.2702	mclemma01	suzukic01	guillca01	martied01	boonebr01	olerujo01	camermi01	bellda01	wilsoda01
4	6.2698	mclemma01	suzukic01	boonebr01	martied01	olerujo01	camermi01	guillca01	bellda01	wilsoda01
5	6.2677	camermi01	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
6	6.2603	mclemma01	olerujo01	guillca01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
7	6.2588	mclemma01	suzukic01	olerujo01	martied01	boonebr01	camermi01	guillca01	bellda01	wilsoda01
8	6.2554	mclemma01	suzukic01	olerujo01	martied01	guillca01	boonebr01	camermi01	bellda01	wilsoda01
9	6.2551	martied01	camermi01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
10	6.2542	mclemma01	suzukic01	olerujo01	guillca01	martied01	boonebr01	camermi01	bellda01	wilsoda01
11	6.2529	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
12	6.2527	guillca01	mclemma01	olerujo01	martied01	boonebr01	camermi01	suzukic01	bellda01	wilsoda01
13	6.2526	camermi01	boonebr01	guillca01	martied01	olerujo01	mclemma01	suzukic01	bellda01	wilsoda01
14	6.2523	mclemma01	olerujo01	martied01	boonebr01	guillca01	camermi01	suzukic01	bellda01	wilsoda01
15	6.2493	mclemma01	olerujo01	boonebr01	martied01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
16	6.2485	mclemma01	suzukic01	boonebr01	martied01	guillca01	olerujo01	camermi01	bellda01	wilsoda01
17	6.248	camermi01	martied01	olerujo01	mclemma01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
18	6.2477	martied01	olerujo01	boonebr01	camermi01	mclemma01	suzukic01	guillca01	bellda01	wilsoda01
19	6.2477	camermi01	suzukic01	guillca01	boonebr01	olerujo01	mclemma01	martied01	bellda01	wilsoda01
20	6.2462	martied01	olerujo01	mclemma01	boonebr01	camermi01	suzukic01	guillca01	bellda01	wilsoda01
21	6.2461	mclemma01	camermi01	olerujo01	martied01	boonebr01	suzukic01	guillca01	bellda01	wilsoda01
22	6.2444	martied01	olerujo01	boonebr01	mclemma01	suzukic01	guillca01	camermi01	bellda01	wilsoda01
23	6.2441	mclemma01	suzukic01	boonebr01	martied01	olerujo01	guillca01	camermi01	bellda01	wilsoda01
24	6.244	mclemma01	olerujo01	guillca01	boonebr01	camermi01	martied01	suzukic01	bellda01	wilsoda01
25	6.2439	mclemma01	suzukic01	boonebr01	martied01	camermi01	olerujo01	guillca01	bellda01	wilsoda01
26	6.2432	camermi01	olerujo01	guillca01	martied01	mclemma01	suzukic01	boonebr01	bellda01	wilsoda01
27	6.2426	mclemma01	suzukic01	boonebr01	olerujo01	martied01	camermi01	guillca01	bellda01	wilsoda01
28	6.2416	mclemma01	guillca01	olerujo01	martied01	camermi01	suzukic01	boonebr01	bellda01	wilsoda01
29	6.2404	mclemma01	suzukic01	guillca01	martied01	camermi01	boonebr01	olerujo01	bellda01	wilsoda01
30	6.2402	camermi01	suzukic01	martied01	olerujo01	boonebr01	mclemma01	guillca01	bellda01	wilsoda01

図 15　OPS で色分け (走塁能力を加味)

以上から、走塁能力も加味した最適打順の傾向を考察していく。まず、1 番バッターについて出塁率と走塁能力が高い選手となる傾向がみられ、長打率はそこまで重要でないことがいえる。3 番バッターは OPS が大きく、長打率はそれほど大きくない選手となる傾向がある。つまりヒットを打つのが得意な選手である。4 番バッターは全ての指標について優れており、SLG も優れている割合が極めて高かった。つまり、ホームランを打つのが得意な選手である。また、2∼5 番で OPS 上位 3 位が連なっていることが多かった。下位打線は 4.2.2 節と同様に打撃結果が乏しい選手が中心となった。

4.4　議論

4.3 節では、走塁能力を加味した最適打順について述べた。この節では、Mariners2001 にモデルを当てはめることを行なう。表 26 はそれぞれのモデルでの得点期待値をまとめたものである。

表 26　それぞれのモデルから算出された Mariners2001 の得点期待値

単純モデル	改良モデル	走塁能力を加味したモデル	実際の期待値
4.60	5.89	6.14	5.72

表 26 から得られる考察について述べる。まず改良モデルより走塁能力を加味したモデルの方が得点期待値が大きくなった。なぜなら、走塁能力を加味したモデルは 1996 年から 2005 年のデータから算出された得点期待値を参照し、これが大きくなる場合のみ盗塁を試行するという仮定をしているからである。

今までのモデルは、走塁能力を含め基本的に攻撃に着目したものである。しかし、ピッチャーの利き手・キャッチャーの守備能力・犠打等を考慮することでこれらのモデルの得点期待値は下がることが容易にわかり、より実際の期待値に近づくことが示唆される。

4.5　結言

これまで、期待得点の観点から最適打順について述べてきた。以上をまとめると、出塁率が高い選手はなるべく若い番号の打順に置くことをベースにして、1 番バッターは出塁率及び走塁能力が優れている選手で、長打率はそれほど重要ではない。2 番バッターは走塁能力が優れている選手、3 番バッターは OPS 及び安打率に優れ、4 番バッターは最強打者、5 番バッターは 3 番バッターには劣るが OPS がある程度良い選手、6 番バッター以降は出塁率順とすることで期待得点を大きくすることができるという結論が得られた。

改良モデルでは 1 年間を通じて得られたデータから、確率値でそれぞれの打撃結果を割り当てていた。しかし、現実の野球では相手のピッチャーの能力によって打撃結果が左右されることは言うまでもない。また、ピッチャーが左投げであったら、1 塁ランナーの盗塁成功率が仮に高くても、盗塁の試行回数は激減するであろう。以上より、相手の投手及び守備能力を考慮したモデルを開発

することが今後の課題である。

5　謝辞

本研究にて統計モデリングや野球等に関しての知見をさらに深化させることができました。これらの研究の機会を与えて下さった指導教員の福永力教授には、日々の生活を含めて大変お世話になりました。ここに深く感謝の意を表します。また、統計モデリングに関して、お忙しい中様々な助言をして頂いた社会科学研究科経営学専攻の小方浩明准教授に深く感謝の意を表します。さらに、研究室の仲間や友人等、本研究に関して助言をして頂いた方々に深く感謝の意を表します。

参考文献

[1] 穴太克則・高野健大（2015）得点圏打率・盗塁・併殺を考慮した最適打順決定モデルについて：FA 打者トレード戦略の検討、数理解析研究所講究録 1939：p.133-142

[2] 柳川佳也・宮崎茂次 (2000) 最適打順についての一考察、日本生産管理学会 7 巻 2 号 p.193-196

[3] 野村克也 (2008) 野村再生工場-叱り方、褒め方、教え方、角川書店

[4] D’Esopo, D.A. and Lefkowitz, B., “The distributionof runs in the game of baseball”, In Optimal Strategies in Sports, 1977

[5] 鳥越規央 (2014) 勝てる野球の統計学-セイバーメトリクス、岩波書店

[6] 鳥越規央・薄井一樹・時光順平（2011）セイバーメトリクスによる最適打順決定モデルとそのシミュレーション、数理解析研究所講究録 1758：p.1-14

[7] 春原達郎（2012）D’Esopo & Lefkowitz モデルに基づく得点分布を考慮した野球の打順分析、中央大学大学院研究年報理工学研究科篇第 42 号

[8] 大澤清・合田憲人、打順最適化問題の高速化手法、情報処理学会論文誌第 48 巻第 3 号 p.1443-1454

[9] http://d.hatena.ne.jp/hiroakit/20091027/1256609214

[10] データで見る「最強打者」の打順／鳥越規央の野球視角、http://p.kanzensokuhou.jp/m/torigoe/20160308/01.sphtml

[11] 福島真太朗 (2014)R によるハイパフォーマンスコンピューティング、ソシム株式会社

[12] J.Albert・J.Bennett(2004) メジャーリーグの数理科学上巻・下巻、丸善出版

[13] 東京大学教養学部統計学教室 (1991) 統計学入門、東京大学出版会