

予測点数から読み取る野球の「流れ」の分析

Analysis of "flow" of baseball read from predicted score

吉本 光希^{*1}, 内田 眞司^{*1}, 平 俊男^{*2}

Teruki YOSHIMOTO^{*1}, Shinji UCHIDA^{*1}, Toshio HIRA^{*2}

^{*1}奈良工業高等専門学校 情報工学科, ^{*2}奈良工業高等専門学校 機械工学科

Email: {yoshimoto, uchida}@info.nara-k.ac.jp, hira@mech.nara-k.ac.jp

あらまし: 本研究では野球の「流れ」を定量的に分析するために隠れマルコフモデルを用いて野球の試合展開をモデル化した。先行研究において走者とアウトカウントを状態とするマルコフ連鎖としてモデル化し「流れ」を分析した結果、選手の心理的要因を考慮することで「流れ」を説明できることが示唆されていた。そこで隠れマルコフモデルの潜在変数として選手の感情値を導入し、選手の心理的要因を考慮したモデルを構築した。本稿ではメジャーリーグの試合データを用いてモデルの予測精度を評価した結果を報告する。

キーワード: 野球, マルコフ連鎖, 心理的要因, メジャーリーグ

1. 研究背景・研究目的

野球などの団体スポーツには「流れ」というものが存在する。たったワンプレーや選手の気持ちで試合展開が大きく変わること「流れ」が変わるといふ。「流れ」を定義しようとこれまで様々な研究が行われてきた。先行研究^[1]では、野球のスコアブックから記録された試合経過のみに注目し、「流れ」を読み取ろうとランナーとアウトカウントを状態としたモデルを作成し、各状況の予測獲得点数を算出することで「流れ」を読み取ろうと試みた。

しかし、実際の試合の中ではスコアブックに現れる数値だけではなく、緊張や焦りといった選手の感情やチームの士気も「流れ」の構成要因の一つであると考えられる。スポーツにおいて感情などの心理的要因の影響は様々な論文で証明されている。バッターがストレスの掛かる状況下ではパフォーマンスが低下すること^[2]などの例があげられる。

本研究では記録された試合展開だけでなく、選手の心理的要因を考慮したモデルを構築し、予測得点を算出する。そして、従来モデルより予測精度を向上させ、より正確な「流れ」をつかむことを目的とする。

2. 先行研究

先行研究では、走者の有無 8 通り ($B=\{\emptyset, '1', '2', '3', '12', '13', '23', '123'\}$) とアウトカウント 3 通り ($O=\{0,1,2\}$) を変数として持つ 24 通りの状態 $U_i=(b,o)(i=1\sim 24)$ と 3 アウトの状態を示す吸収状態 U_{25} をあわせた 25 状態を持つマルコフモデルを構築した。マルコフモデルの模式図を図 2.1 に示す。

野球のグラウンドを模したダイヤモンドが 1 状態を表している。例として U_3 の状態はランナー 3 塁、アウトカウントなしを表す。その後、このマルコフモデルを用いてモンテカルロシミュレーションを行い予測獲得点数の算出を行った。その際、使用する

状態 i から状態 j への状態遷移で獲得する点数 t_{ij} は以下の数式 (1) を用いて算出する。

$$t_{ij} = (|b_i| + o_i + 1) - (|b_j| + o_j) \quad (1)$$

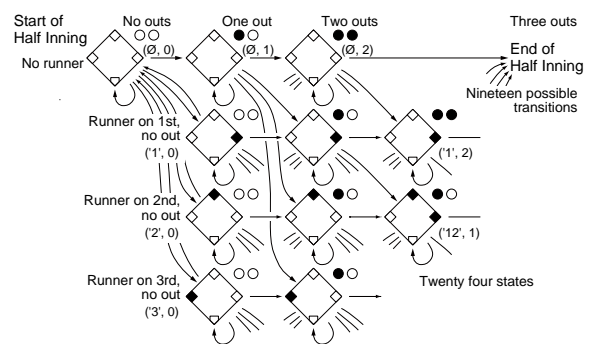


図 2.1 従来モデルの模式図

3. 研究方法

本研究では、メジャーリーグ、ボストン・レッドソックスの 2015 年から 2017 年までを使用する。データはメジャーリーグの試合結果を PBP(Play-By-Play)ファイルという形式で保管している Retrosheet^[3] からフリーウェアである Chadwick^[34] を用いてダウンロードしたデータを使用する。

そのデータを感情値という本研究で定義した閾値に従いデータを分類する。感情値は「先行研究で定義した「予測得点」に注目し、予測得点が実際の獲得点数を上回っている状況の場合 -1、0.5 ポイント以上下回っている場合 -1 それ以外の状況のとき 0 と定義した。

その後、分類したデータから、隠れマルコフモデルを構築し、試行回数 1000 回のモンテカルロシミュレーションを行う。その結果、算出された予測得点と実際の試合の結果を比較し、分析する。感情値を導入した提案モデルの概略図を図 3.1 に示す。左図は従来モデルの概略図から一部を取り出したものである。従来モデルでは左図の中の遷移確立 P_i, P_j は一定であり、シミュレーションの中で値が変わるこ

とはなかった。しかし、提案モデルではシミュレーションの中で遷移確立 P_i, P_j が変化する。遷移確立 P_i, P_j は、本研究で定義した感情値の値が変わると、右図に従って変化する。この遷移確立の変化を実際の試合での選手の動揺や心の余裕とみなし、モデルに選手の心理的要因を導入することを試みる。

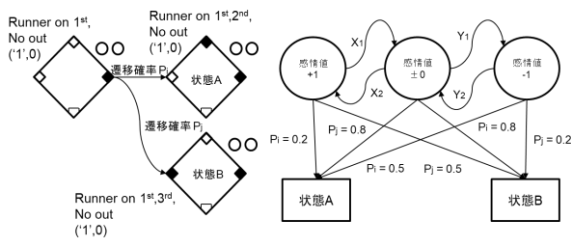


図 3.1 提案モデルの概略図

4. 研究結果

以下にそれぞれの感情値の場合での予測獲得点数を示す。

4.1 感情値の値が0の場合

以下の表 4.1.1 に感情値の値が0の場合の各状況の予測獲得点数を示す。列はランナーの状況，行はアウトカウントを示している。

表 4.1.1 感情値0の予測獲得点数

	0 outs	1 out	2 outs
無し	0.47	0.26	0.09
一塁	1.22	1.02	0.34
二塁	1.17	0.76	0.29
一二塁	2.07	1.27	0.73
三塁	0.82	0.47	0.2
一三塁	1.81	1.19	0.45
二三塁	1.43	0.87	0.45
満塁	2.36	1.61	0.56

4.2 感情値の値が+1の場合

以下の表 4.2.1 に感情値の値が+1の場合の各状況の予測獲得点数を示す。列はランナーの状況，行はアウトカウントを示している。

表 4.2.1 感情値+1の予測獲得点数

	0 outs	1 out	2 outs
無し	1.21	0.75	0.26
一塁	1.99	1.84	0.76
二塁	2.36	1.52	0.62
一二塁	3.03	2.07	1.32
三塁	1.81	1.24	0.51
一三塁	3.11	2.21	1.05
二三塁	2.84	1.77	1.05
満塁	3.57	2.54	1.39

4.3 感情値の値が-1の場合

以下の表 3.3.1 に感情値の値が-1の場合の各状況の予測獲得点数を示す。列はランナーの状況，行はアウトカウントを示している。

表 4.3.1 感情値が-1の予測獲得点数

	0 outs	1 out	2 outs
無し	0.47	0.26	0.09
一塁	1.23	1.01	0.36
二塁	1.18	0.77	0.29
一二塁	2.1	1.27	0.73
三塁	0.8	0.46	0.19
一三塁	1.82	1.18	0.44
二三塁	1.44	0.86	0.47
満塁	2.36	1.6	0.56

5. まとめ

上記の結果感情値0の状態の場合の予測獲得点数より、感情値+1の状態の場合の予測獲得点数が大幅に上昇していることが確認できた。これは、選手の心理的要因が予測獲得点数に影響していることの証明となる。一方、感情値が-1の場合の予測獲得点数は感情値0の場合の予測獲得点数と大きく変わることはなかった。このことから、選手の心理的要因は良い状況が発生した場合のほうが、悪い状況が発生したときと比べて試合結果に影響が大きいということがわかった。

今後の課題として、試合に大きく影響すると考えられ、かつ提案モデルで考慮していない打順を考慮したモデルを構築することがあげられる。

参考文献

- (1) 伴 瑞季, 平 俊男, 内田 眞司”スコアマイニングによる野球試合展開分析および「流れ」の検討”, <https://www.jstage.jst.go.jp/article/jsmemecj/2016/0/2016-J1220104/pdf> (2016)
- (2) 田中 ゆふ・田中 美史・関矢 寛史 (2012): “プレッシャーが野球打者の予測スキルに及ぼす影響”, https://kindai.repo.nii.ac.jp/index.php?action=pages_view_main&active_action=repository_action_common_download&item_id=12954&item_no=1&attribute_id=40&file_no=1&page_id=13&block_id=21(2012)
- (3) Chadwick: Software Tools for Scoring Baseball Games, <<http://chadwick.sourceforge.net/doc/index.html>> 2017年11月22日アクセス
- (4) Retrosheet, <http://www.retrosheet.org> 2017年5月20日アクセス