

二塁走者がフライ性の単打によって本塁に生還する可能性を規定する要因の分析

身体運動科学研究領域

5021A039-0 福本 健人

研究指導教員：矢内 利政 教授

【緒言】プロ野球界では弾道トラッキングシステムが多数の球場に常設されたことにより、投球されたボール及び打球に関するあらゆる運動学的データが収集可能となっている。しかし、これらを集約したビッグデータを用いて、得点を増加させるために重要な走塁のパフォーマンス向上に着目した研究は少ない。走塁のパフォーマンスが重要な場面として、走者が二塁に出塁しており、外野手が第一捕球者となった単打が発生した場面が挙げられる。この場面は書籍において、最も走塁判断が難しい場面だと述べられている(澤宮, 2018)。そこで本研究では、走者が二塁におり、外野手が第一捕球者であるフライ単打が発生した場面において、二塁走者の生還成否を規定する走者・守備者及び打球特徴の要因を、2段階の分析により導出することとした。研究の目的は、以下の2つとなる。1つ目は、二塁走者が本塁生還する可能性を決定する力学的要因を系統的に示す deterministic model (Hay, 1988)を構築することとした。2つ目は、deterministic model において定義された決定要因を用いた統計モデルを作り、予測精度を高める影響が大きい要因を導出することとした。

第一部 deterministic model の構築

【方法】本研究では、二塁走者の生還率を出力とした deterministic model を構築した。deterministic model は、力学的要因を系統的に並べることで構築される。そして末端となった要因の内、本研究ではデータを取得できない要因を排除した。最終的に残った末端要因を、二塁走者の生還する可能性の決定要因として取り扱うこととした。

【結果】構築された deterministic model の末端要因から、データを取得できない要因を排除した結果、「打球初速度」、「鉛直打球角度」、「水平打球角度」、「打球飛距離」、「送球精度」、「インニング」、「アウトカウント」、「塁状況」、「点差」、「二塁走者の走力」が決定要因となった。

第二部 統計モデルによる予測

【方法】本研究で用いるデータは、2016年から2021年のMLB公式戦のデータで、MLB公式ウェブサイト「Baseball

Savant」と、MLB公認で FanGraphs 社が公開しているウェブサイト「FanGraphs」から取得した。二塁走者が本塁に生還したか否かを目的変数として、deterministic model の末端要因の内、取得可能であった12個の説明変数を用いた予測モデルを一般化加法モデル(GAM)で構築した。GAMは数理モデリングにおいて用いられる手法の1つで、説明変数ごとに目的変数との関係を推定し、それぞれの関係性を足し合わせ、全説明変数と目的変数の関係とする手法である。非線形分布の変数を説明変数として使用できる点、説明変数と目的変数の関係の解釈性が高い点から、本研究ではGAMを採用した。取得したデータの内、8割を学習データ、2割をテストデータとし、予測正解率の算出にはテストデータを用いた。モデリングに用いた変数とその取得方法、回帰方法は、表に記したものになる。外野手の送球精度は、外野手の送球による得点抑止力を表す指標で、FanGraphs に提示されている「arm」と呼ばれる指標を用いた。そして、統計モデルの予測精度の高さを評価する指標である赤池情報量規準(AIC)を基準とした変数選択によって、ベストフィットモデルを求めた。また、決定要因となった変数の予測精度を比較するため、以下の式で説明変数のAIC寄与の大きさ(AIC_dif)を算出した。

$$AIC_dif_i = AIC_{subi} - AIC_{best} - 2$$

ここでAIC_dif_iは説明変数iのAIC_difの値、AIC_{subi}はベストフィットモデルから説明変数iを除いた際のAICの値、AIC_{best}はベストモデルのAICの値を示す。

表：モデリングに用いた変数表

変数名	単位	取得方法	回帰方法
二塁走者の生還有無(scored_second)	点	算出	—
アウトカウント(outs_when_up)	アウト	Baseball Savant	ロジスティック回帰
インニング(inning)	回	Baseball Savant	ロジスティック回帰
塁状況(runner_situation)	—	算出	ロジスティック回帰
点差(run_difference)	点	算出	平滑化スプライン
二塁走者の走力(sprint_speed)	m/s	Baseball Savant	平滑化スプライン
打球飛距離(hit_distance_sc)	m	Baseball Savant	平滑化スプライン
打球初速度(launch_speed)	km/h	Baseball Savant	平滑化スプライン
鉛直打球角度(launch_angle)	°	Baseball Savant	平滑化スプライン
水平打球角度(spray_angle)	°	算出	平滑化スプライン
外野手の送球による得点抑止力(arm)	点	FanGraphs	平滑化スプライン
打球の落下地点と捕球者の初期位置間の距離(fielder_hit_distance)	m	算出	平滑化スプライン
打球の落下地点と捕球者の初期位置の成す角度(fielder_hit_angle)	°	算出	平滑化スプライン

【結果と考察】ベストフィットモデルには11つの変数が選択され、その予測正解率は76.3%となった。4回に1回の割合で誤った予測が起こる予測モデルとなった。図1か

ら、決定要因となった変数は全てモデルに含まれていた。そのため、決定要因となった全ての変数は統計的に予測精度を高めているといえる。説明変数の AIC 寄与の大きさが最も大きかったのは、アウトカウント(AIC_dif = 2254)であった。決定要因となった変数の内説明変数の AIC 寄与の大きさが大きかったのは、順に水平打球角度(AIC_dif = 968)、二塁走者の走力(AIC_dif = 259)、外野手の送球による得点抑止力(AIC_dif = 157)、鉛直打球角度(AIC_dif = 139)であった。



図 1 : deterministic model におけるベストフィットモデルに含まれた各変数の AIC 寄与の大きさ

2 アウト時に予測値が大きい正の値を取ることが示された(図 2a)。この結果は 2 アウト時の生還率が高いことを意味する。2 アウトの時には打球がヒットになるかアウトになるかの打球判断を二塁走者が行う必要がなく、打球のインパクトの瞬間に本塁へのスタートを切れることが、生還率を大きく高めているためだと考えられる。二塁走者の走力が 8.3m/s 以上の場合に予測値は正の値を取ることが示された(図 2b)。この結果は、8.3m/s 以上の走力を持つ選手が二塁走者であれば、走力が高い選手ほど生還率が高まることを示すものである。水平打球角度は $\pm 15^\circ$ と $\pm 45^\circ$ で予測値が正の値となり、 $\pm 30^\circ$ で負の値となった(図 2c)。この結果は、外野手の正面に飛んだ打球はその打球の勢いに関わらず生還率を低下させ、外野手の間やライン際に飛んだ打球は、生還率を高めることを示すものである。鉛直打球角度については 20° 以上の値で予測値は負の値となる(図 2d)ことから、鉛直打球角度は大きいほど生還率を低下させると示された。

【実践における活用】 本研究における成果は、二塁走者の走塁を指示する三塁コーチによって活用されることが特に望まれ、勝負の分かれ目となる 1 点を、奪い取る可能性を高めるための方策を提示できると考えられる。アウトカウント、二塁走者の走力や外野手の送球による得点抑止力は、プレー開始までに予め把握しておくことが、三塁コーチが精度の高い走塁指示をするために必要となる。また、水平打球角度や鉛直打球角度の値の変化を迅速かつ正確に察知することが、三塁コーチが精度の高い走塁指示をするために重要だと示唆された。これらの値は実践において、三塁コーチがコーチボックスの位置から瞬時に把握するこ

とは難しい。また、打球がヒットになるかアウトになるかの打球判断を二塁走者が行う必要がない 2 アウトの場面では、生還率が大きく高まることが示された。これは、2 アウトであるために走者が打球のインパクトの直後に本塁へのスタートを躊躇なく切れるためと考えられる。アウトカウントが 0 や 1 の場合においても二塁走者が早いタイミングでスタートを切り、本塁生還率を高めるためには、選手及び三塁コーチは普段の練習から、水平打球角度、鉛直打球角度や打球がヒットになるかアウトになるかどうかを、迅速かつ正確に判断する練習を行う必要があると考えられる。また、8.3m/s 以上の走力を持つ選手が二塁走者であれば、生還率を高めることが示されたことから、三塁コーチはチームの選手に対し、8.3m/s 以上の走力を身に付けるよう指導すべきと考えられる。

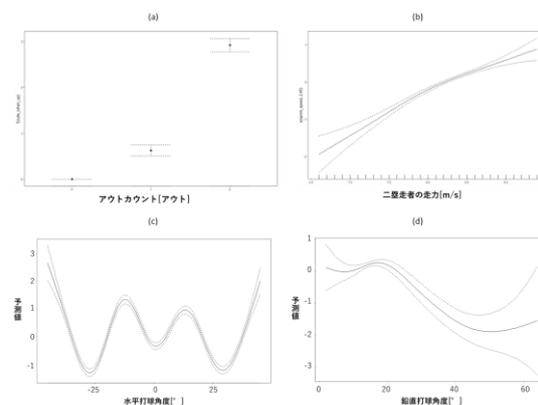


図 2 : 各説明変数の値と予測値の関係

予測値が正の値だと生還率を高める影響をその変数が与えており、負の値だと生還率を低下させる影響をその変数が与えている

【結論】 deterministic model を構築し、二塁走者の生還率の決定要因として 10 の変数を定義した。実測値を用いたベストフィットモデルにはその内の 9 つの変数が選択され、その予測正解率は 76.3%となった。生還率に大きな影響を及ぼす要因として、アウトカウント、二塁走者の走力、水平打球角度、外野手の送球による得点抑止力、鉛直打球角度があげられた。この結果は、二塁走者の生還率を高めるためには水平打球角度と鉛直打球角度を元に打球がヒットになるか否かを、迅速かつ正確に判断する能力が二塁走者および三塁コーチには必要であることを示唆するものである。