

機械学習を用いた競泳における理想的なレースプランの分析

2018SS040 正木達也

指導教員：三浦英俊

1 はじめに

水泳はレース中のレースプランが重要であり駆け引きが醍醐味である。文書 [1, 2] のような泳速度やストローク変数についての研究はされてきた。しかし、レースにおけるペース配分の研究はほとんど見られない。水泳の大会では、自分と近いタイムの選手と同じレースで争う場合が多い。そのため、自分と同等の力を持つ周りの選手たちに勝つためには、レースプランの熟慮が必要だといえる。

2 研究の背景と目的

水泳の大会のうち、日本選手権や日本学生選手権のようなレベルの高い大会となると、参加するための条件として標準記録が設定される。それを突破できる選手たちが出場する大会では、予選を勝ち残った上位 8 名が A 決勝に進出し、9 位から 16 位の選手が B 決勝に進出する。その後、それぞれで行ったレースによって最終的な順位が確定するというのが基本である。本研究では A 決勝に残ることができる可能性が高いレースプランを理想的なものと考え、水泳には基本的に 3 種類のレースプランが存在する。文書 [3] より前半にリードを取り、逃げ切る先行逃げ切り型や後半にラストスパートをする追い込み型、全体的に一定のペースを維持するイーブンペース型がある。

本研究の目的は、男子 200 m 自由形における最適なレースプランの発見である。Seiko リザルト速報 [4] の男子 200 m 自由形データを利用して研究を行う。200 m 自由形の距離を 4 分割し、50 m ごとのタイムを計算し、全体を占める割合を求める。R を用い決定木でデータ分析を行い、A 決勝に残るためにはどのようなレースプランを実行すればよいかを考える。

3 データの説明

本研究では国内選手のうち参加標準記録を突破した選手（以降は「トップ選手」と表現する）が出場する日本選手権水泳競技大会やジャパンオープン。国内の大学生選手のうち参加標準記録を突破した選手（以降は「大学生選手」と表現する）が参加できる日本学生選手権水泳競技大会を中心に数個の大会をまとめてデータ収集をした。トップ選手の A 決勝、大学生選手の A 決勝、トップ選手の予選、大学生選手の予選で分析を行う。データ数はトップ選手の A 決勝は 32 個、大学生選手の A 決勝は 32 個、トップ選手の予選は 174 個、大学生選手の予選 180 個である。

予選の決定木分析を行う上で上位の選手と下位の選手では大きなタイムの差が生まれてしまう。タイムに差がありすぎることによって結果に偏り生じてしまうので、予選通過タイムより 5 秒以上遅い選手を取り除いて分析を行う。

4 記号の導入とラベルの説明

200 m 自由形のタイムを 4 分割して以下のように合計を 100% とする割合を記号で表す。

t_1 : 0m から 50 m のタイムの割合

t_2 : 50m から 100 m のタイムの割合

t_3 : 100m から 150 m のタイムの割合

t_4 : 150m から 200 m のタイムの割合

決勝と予選でそれぞれ 3 種類のラベルを付ける。

a: 決勝で優勝した選手

b: 決勝で 2,3 位の選手

c: 決勝で 4 位から 8 位の選手

d: 予選で 1 位から 8 位の選手

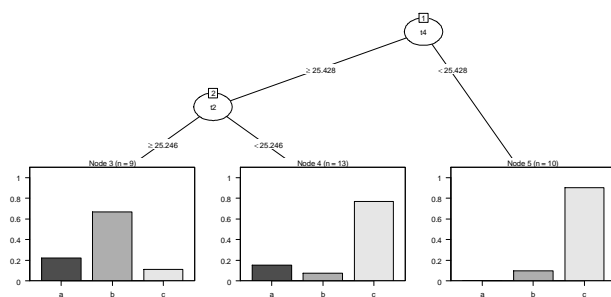
e: 予選で 9 位から 16 位の選手

f: 予選敗退の選手

5 決定木

文献 [5, 6] より決定木とはラベル付きデータが観測されているとき、木構造の推論規則を用いて入力に対して出力の予測をするものである。決定木の学習方法は、葉と呼ばれる箱に存在するデータが定数を満たすかどうかで 2 つの箱に分けラベルをそれぞれ付けます。その際に発生する損失であるエントロピーを計算します。その計算結果からエントロピーが最小になるよう新たな葉を作成する。さらにエントロピーの値を適切に設定することで決定木が複雑になりすぎぬよう調節する。この工程を再帰的に繰り返されることで木構造が構成される。各ノードに対する簡易的な条件で予測が繰り返し行われるので結果の意味を理解しやすいという点から決定木を採用した。

6 トップ選手と大学生選手決勝の決定木



使用データ数は 32 個である。これを見ると上位 3 名には t_4 が 25.4% 以上かつ t_2 が 25.2% 以上である傾向が強いと言える。

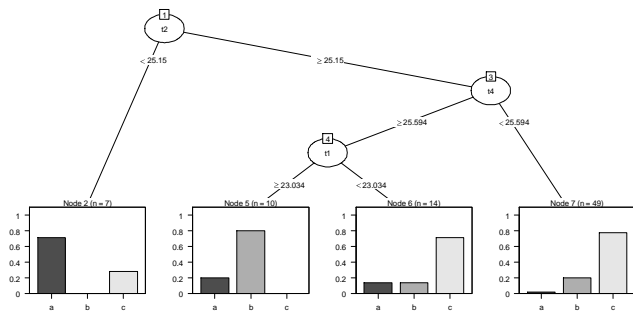


図2 大学生選手の決勝の決定木

図2は大学生選手の決勝での決定木を表したものである。トップ選手での分析と同様に3種類のラベルを付けて決定木分析を行った。使用データ数は32個である。これを見ると t_4 が 25.4% 以上かつ t_2 が 25.2% 以上であると上位3位である傾向が強いと言える。

7 トップ選手と大学生選手予選の決定木

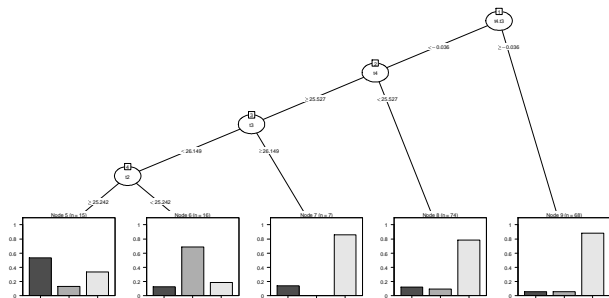


図3 トップ選手の予選についての決定木

図3はトップ選手の予選を突破するための決定木を表したものである。以下のように3種類のラベルを付けて決定木分析を行った。 d はA決勝に進出した選手、 e はB決勝に進出した選手、 f は予選通過できなかった選手。使用データ数は174個である。これを見ると、A決勝に残る理想的なレースプランは t_1 が 23.4% 未満、 t_2 が 25.3% 以上、 t_3 が 26.0% 未満、 t_4 が 25.6% 未満のレースプランであると言える。

大学生選手予選の決定木は複雑なため割愛する。トップ選手での分析と同様に3種類のラベルを付けて決定木分析を行った。使用データ数は180個である。大学生の場合理想的なレースプランは t_2 が 25.2% 以上、 t_3 が 26.1% 未満、 t_4 が 25.5% 以上のレースプランであると言える。

8 予選通過のレースプラン

トップ選手と大学生選手の t_1, t_2, t_3, t_4 の予選を通過しやすい値を比較する。

表1 トップ選手と大学生選手の予選を突破する傾向の強いレースプランの比較 (数値は%)

理想	t1	t2	t3	t4
トップ	~23.4	25.3~	~26.0	~25.7
大学生		25.2~	~26.1	25.5~

表1はトップ選手と大学生選手のそれぞれの $t_1 \sim t_4$ の予選を突破する基準値を表したものである。これを見るとトップ選手と大学生選手の t_2, t_3 の値がほとんど同値であることが読み取れる。そこから中盤のレースプランは似ていることがわかる。トップ選手の後半に注目すると、 t_3 から t_4 にかけてタイムを上げていることから追い込み型であると言える。それに対して大学生選手は t_3 から t_4 にかけてタイムをキープしているという点からイーブンペース型であると言える。

9 終わりに

本研究では決定木を用いて数個の大会をまとめて、男子200m自由形の理想的なレースプランを考えた。それにより予選とA決勝で結果を残すには、違うタイプのレースプランが必要であることがわかった。今後の課題としては予選で実力が飛びぬけている選手が力を温存して泳いでしまう場合がある。力を温存してしまうことで結果に偏りが生じてしまう可能性がある。工夫として、事前に提出されるタイム(エントリータイム)と実際に泳いだタイムを比べ、差が大きすぎる場合には取り除いて分析を行う必要がある。

参考文献

- [1] 生田泰志, 松田有司, 山田陽介, 来田宣幸, 小田伸午: クロール泳における泳速度, ストローク頻度およびストローク長の変化と筋活動の関係 2010
- [2] 松井健, 寺田晶裕, 立貞栄司, 本部洋介, 生田泰志, 若吉浩二, 野村照夫: 競泳レースにおける5m毎の泳速度とストローク変数の変化-日本選手権200m自由形種目における泳力別比較- 1998
- [3] 佐藤大典: 『競泳男子400m自由形におけるレースパターンとパフォーマンスの関係』
- [4] 競泳リザルト速報サービス (<http://swim.seiko.co.jp/>)
- [5] 金森敬文: 『Rによる機械学習入門, 12(1), 191-193, 2017
- [6] Sebastian Raschka, Vahid Mirjalili (株式会社クイープ 訳): 『Python機械学習プログラミング達人データサイエンティストによる理論と実践』. 株式会社廣済堂, 東京, 2018.