

陸上競技短距離走における日本人の記録の統計的分析

2016SS009 橋本将成

指導教員：松田 眞一

1 はじめに

日本人は本番に弱いと言われているが、オリンピックにどれだけコンディションを合わせられるかを日本人選手と外国人選手のこれまでの大会の記録をもとに統計的分析によって比較する。2020年に開催される東京オリンピックは1964年以来の自国開催ということもあり注目度が高く、陸上競技の花形といっても過言ではない100m走では、年々有望な選手が台頭してきている。100分の1のタイムを競い合う陸上競技の世界で、データを眺めるだけではわからない事柄を統計的分析によって明らかにしたい。

2 データについて

4年に一度開催されるオリンピックの間に、グランプリレースという大きな大会がある。ここでは、IAAF [2]に記載されている陸上競技100mの記録をもとに当時の開催大陸、開催年、風速に関するデータの分析を行う。先行研究として浅井 [1]で、男女100m走についての外国人選手の分析がある。そこでは、各レースの計測されたタイムの実測値と重回帰分析によって算出された予測値との誤差からそれぞれの選手の特徴が明らかになっている。しかし、日本人選手については分析を行っていないので、日本人選手と外国人選手の比較について分析をしていく。

本研究ではオリンピックに出た経験がある選手をはじめ、合計32名の選手を取り上げる。100m走では日本人選手7人と外国人選手9人、200m走では日本人選手6人、外国人選手10人の分析を行った。なお、100m走世界記録を保持しているウサイン・ボルト選手、100m走日本人記録を保持しているサニブラウン・アブデル・ハキーム選手は、十分なデータがなかったため分析から外すこととする。

3 100m, 200m 走について

紙面の都合上、分析結果が平均的な日本人選手、外国人選手を表1に示す。また、開催大陸と開催年は各選手、表2, 表3のようにダミー変数を対応させる。なお、表1と表2に関しては上側が100m走、下側が200m走を示す。

表1 選手名及びレース数

| 選手名 | レース数 | 選手名 | レース数 | 出身国 |
|-----|------|-----|------|------------|
| YR | 23 | SB | 21 | 中国 |
| TS | 20 | YB | 18 | ジャマイカ |
| IS | 16 | KG | 18 | トリニダード・トバゴ |
| YJ | 22 | LR | 15 | アイルランド |

表2 開催年ダミー変数の対応表

| 変数 | y0 | y1 | y2 | y3 | y4 | y5 | y6 |
|-------|------|------|------|------|------|------|------|
| TS | 2019 | 2018 | 2017 | 2016 | | | |
| YR | 2018 | 2017 | 2016 | 2014 | 2013 | | |
| SB | 2018 | 2017 | 2015 | 2014 | 2013 | 2012 | 2011 |
| YB | 2019 | 2018 | 2017 | 2016 | 2015 | 2012 | 2011 |
| IS | 2018 | 2017 | 2016 | 2015 | 2014 | 2013 | 2012 |
| YJ | 2019 | 2018 | 2017 | 2016 | 2015 | 2014 | |
| KG,LR | 2019 | 2018 | 2017 | 2016 | 2015 | 2014 | 2013 |

表3 開催大陸ダミー変数の対応表

| 変数 | t1 | t2 | t3 | t4 | t5 | t6 |
|-----|-------|-------|-----|---------|----|----|
| 大陸名 | ヨーロッパ | 南アフリカ | アジア | オーストラリア | 北米 | 南米 |

4 分析方法

分析方法には重回帰分析を用いる。重回帰分析とは、いくつかの説明変数を用いて目的変数を1次式で表現することである。(石村・石村 [3] 参照) その1次式がどの程度因果関係を説明できているか、最も当てはまりの良い重回帰式を見つけていく。今回は決勝のタイムに対して、どのような変数が影響を与えているのかを選手ごとに分析する。決勝のレースタイム ($r3$) を目的変数とし、決勝の風速 ($r3.Wind$)、準決勝のタイム ($r2$)、準決勝の風速 ($r2.Wind$)、開催年 ($y0, y1, \dots$)、開催大陸 ($t1, t2, \dots$) を説明変数とし重回帰分析、変数選択を行ったあと残差分析を行う。その際、外れ値と判断されたデータに関しては分析から外すこととし、再度重回帰分析・変数選択・残差分析を行う。また、風速は追い風が+, 向かい風を-とする。開催大陸で分ける必要がない時は、自国 (1)、他国 (0) とするダミー変数 (t) を用いる。開催年に関して分析に利用できるデータが1年間で1つしかないとき、その前後の年として扱うこととする。なお、各選手、開催大陸に関しては最もレース数が少ないダミー変数を分析から省き、開催年に関しては最も古い年に対応するダミー変数を省いて分析を行う。

5 重回帰分析結果

日本人選手と外国人選手の男子100m, 200mの分析を行った。重回帰分析を行い変数選択を行った結果、各選手の決定係数は概ね0.7程度になった。風速や開催大陸といった外的要因に影響を受けやすいことや、準決勝と決勝のタイムは連動していることが分かった。また、外国人選手は日本人選手に比べ、変数の係数にばらつきがあまりみられなかった。

表中に各変数の回帰係数を示し、()内は***が0.1%、**が1%、*が5%、.が10%で有意、無印は10%で有意でないことを示す。また、—はデータがないことを、()の数値は変数選択前の値を示す。

5.1 決定係数, 風速, 準決勝のタイム, 準決勝の風速

表4 重回帰分析の結果(100m, 200m 走)

| 選手名 | 決定係数 | $r3_Wind$ | $r2$ | $r2_Wind$ |
|-----|-------|-------------|------------|------------|
| TS | 0.834 | -0.051(**) | 0.146 | |
| YR | 0.748 | -0.062(***) | 0.797(***) | 0.043(*) |
| SB | 0.823 | -0.048(*) | 0.650(***) | |
| YB | 0.685 | | 1.079(***) | |
| IS | 0.891 | -0.241(***) | | |
| YJ | 0.726 | -0.044(*) | 0.344(*) | |
| KG | 0.783 | | 0.389(***) | |
| LR | 0.923 | -0.113(*) | 0.754(***) | 0.208(***) |

100m 走において日本人選手は決勝の風速, 外国人選手は準決勝のタイムでより強く有意を示した。決勝の風速の係数は日本人選手と外国人選手でそれぞれ近い値が得られた。YB 選手は, 決定係数が一番低くその中でも, 準決勝のタイムが良ければ決勝のタイムがいいので最も相関が高くなったと考えられる。一方, 200m 走ではどちらかしか残らない選手が存在した。日本人選手は風の影響を大きく受け, 外国人選手は体力を温存して走るためこのような結果になったと考えられる。また, 準決勝の風速が有意を示した選手は準決勝を全力で走っていると考えられる。

5.2 開催年

表5 重回帰分析の結果(開催年)

| 選手名 | $y0$ | $y1$ | $y2$ | $y3$ | $y4$ | $y5$ | $y6$ |
|-----|------------|-------------|------------|-----------|-----------|------|------|
| TS | -0.115(*) | -0.151(**) | -0.206(**) | | — | — | — |
| YR | | -0.063 | | | | | — |
| SB | -0.191(**) | | | | -0.070(.) | — | |
| YB | | 0.107(.) | | | — | | |
| IS | | | | — | 0.154(**) | | — |
| YJ | -0.370(**) | -0.529(***) | -0.195 | | — | | — |
| KG | -0.489(*) | -0.348(**) | | | | — | |
| LR | — | | | 0.525(**) | 0.295(*) | — | |

100m, 200m 走に共通して $y0$ に近づくほど負の相関が大きくなっていることから, 年々タイムを伸ばしていることが分かる。特に外国人選手のほうが負の相関が大きい。外国人選手はオリンピック等の大会で自己ベストやシーズンベストを更新していることから, 調整能力が高く本番に強いといえる。日本人選手はこのような姿を目指していくことが求められる。また, ほとんどの選手が直近のダミー変数で負の相関を示していることから 2020 年のオリンピックに合わせてきていることが分かる。

5.3 開催大陸

表6 重回帰分析の結果(開催大陸)

| 選手名 | t | $t1$ | $t2$ | $t3$ | $t4$ | $t5$ | $t6$ |
|-----|-------------|-----------|------|-----------|------|-----------|-----------|
| TS | -0.078 | | | | | | |
| YR | (0.003) | | | | | | |
| SB | | -0.217(.) | — | -0.248(*) | | — | — |
| YB | | -0.073(.) | — | | | | |
| IS | -0.299(***) | | | | | | |
| YJ | | -0.089 | | | | | |
| KG | | | — | — | | -0.382(.) | -0.444(*) |
| LR | -0.182 | | — | — | — | — | — |

自国開催のレースの方がいい結果が出やすい選手もいれば, 特に変わりのない選手も同数であった。外国人選手は日本人選手より, $t1$ で相関を示す選手が多い。陸上はヨーロッパで栄えているため, トップ選手でも積極的にレースに挑戦していることが分かった。

6 まとめ

日本人選手と外国人選手に共通して, 風速と同程度準決勝のタイムが有意を示していた。100m 走と 200m 走は同じ単距離走でも走り方や様々な要因が影響するので, 違いがみられた。実力を考えると決勝進出は可能であるので, 本番に強くなるためにどんなレースでもマルチに結果を残していくことが求められる。本研究では, 納得できる結果が多く得られたものの, 技術的な差であると判断してしまう部分があった。区間ごとに注目したりストライドやピッチ数といったひとつひとつのレースに注目すれば, 理想の走りや理論上の最高タイムが導きだせたり違ったものが見えてくると考えられる。

7 おわりに

ここ数年で, 9 秒台の記録を出せる選手が現れてきた。日本人選手のレベルが全体的に上がっていることが見て取れる。ベストな走りができれば決勝進出はそう遠くないように思える。毎回, 風速が違えばタイムも違うため, 単純にタイムを眺めているだけではわからなかった多くのことを知ることができた。この研究で陸上競技がさらに好きになったので, 今後も 100m 走日本人初の決勝進出と 200m での記録更新を期待して, 見届けていきたい。

参考文献

- [1] 浅井洋希『陸上競技における記録の統計的分析』, 数理情報学部情報システム数理学科卒業論文, 2010.
- [2] IAAF, GrandPrix-Results-: <https://www.iaaf.org/results/2019> (2019/04 閲覧)
- [3] 石村貞夫・石村光資郎:『入門はじめての多変量解析』, 東京図書, 2007.