

< 修 士 論 文 >

競馬において人気と結果に
乖離のあるレースの予測
(要 旨)

滋 賀 大 学 大 学 院
デ ー タ サ イ エ ン ス 研 究 科
デ ー タ サ イ エ ン ス 専 攻

修了年度：2021年度

学籍番号：6020123

氏 名：山田 宏輝

指導教員：梅津 高朗

提出年月日：2022年1月12日

背景・目的

近年、再度盛り上がりつつある競馬であるが、これの結果を予想することは困難である。下馬評道理決まるレース、そうでないレース様々あるが、人それぞれ好みのレースを探して的中を目指している。しかし、人気な馬が勝つと予想してそのまま人気の馬が勝つて的中したとしても、払戻は穴馬が勝ったときと比べて小さくなる。そして、払戻金が投票した金額を下回ってしまうと回収率は上がらない。そこで、なるべく人気な馬だけが来るような堅いレースではなく穴馬も来るような荒れるレースを見つけようと考えた。

本論文は競馬の予想自体に目を向けるのではなく、競馬を予想するうえで重要な回収率を上げることが目標に、レースで上位に入った馬の人気がどのような条件で変化するのかわかるとすることを目的とする。

データ

日本中央競馬会が提供している JRA-VAN データラボにて公開されている「TARGET frontier JV」から取り込んだ公式のデータを使用する。また、JRA の東京競馬場の過去 15 年間で行われた 7477 レースを分析対象とする。障害レースはレース回数が少ないため今回は除外した。

分析の方法

荒れるレースを分析するために過去のレースで 1, 2 着に入った馬の平均人気がどのような要素で変化するのかわかるとするために重回帰分析を用いる。よって、本論文では 1, 2 着馬の平均人気を目的変数として、その他レース開始前の時点で正式に公開されている情報を説明変数とする。

分析結果

今回の分析を行うにあたって日本中央競馬会が提供している JRA-VAN データラボにて公開されている「TARGET frontier JV」から公式のデータを取り込んだ。対象としたデータは東京競馬場で過去 15 年間に行われたレースであるが、障害コースのレースも含まれていた。当然中央競馬のレースの一つではあるが、一般的に行われているのは芝とダートのコースで行われるレースであるので障害コースのレースを除外することを決めた。障害コースのレースを除いた状態でレースは全部で 7477 レースで、これが今回のデータの総数である。また、レース開始前の時点で正式に公開されている情報として「距離」、「馬場」、「出走頭数」、「馬場状況」を取り入れた。このとき、「距離」に関してはメートル単位で表記されることが一般的であるのでそのままメートル単位でデータを取った。「馬場」に関しては前述のとおり障害コースを除いた芝コースとダートコースの 2 つのコースであるのでダミー変数を用いてデータとした。「出走頭数」はレースに登録されていた頭数をデータとした。「馬場状況」に関しては良馬場開催が基本となるため、稍重馬場、重馬場、不良馬場

は馬場状況の悪い状態として 1 つにまとめ、良馬場開催かそれ以外の状況の開催かで分けてダミー変数を用いてデータとした。これらを説明変数としてRを用いて、重回帰分析を行った。説明変数の中で、「距離」のデータは明らかにほかのデータに対して大きいためダミー変数を除いた説明変数は標準化を行ったうえで重回帰分析をした。分析の結果から、「距離」は荒れるか否かに影響があるとは言えなかった。一方で、その他の変数である「馬場」、「出走頭数」、「馬場状況」は影響があると判断した。「馬場状況」に関しては良馬場開催であると 1,2 着の平均人気が下がるような結果を示した。「出走頭数」に関しては頭数が多いほど 1,2 着の平均人気上がるような結果を示した。「馬場」に関しては、芝コースでの開催の場合であると 1,2 着の平均人気上がるような結果を示した。

結論・考察

今回の分析では荒れるレースの要因を 1,2 着馬の平均人気を用いることによって明らかにした。分析の結果から「馬場」はダートコースよりも芝コースで、「出走頭数」はより多い場合で、「馬場状況」は良馬場以外の状況での開催でレースは人気と結果に乖離が生まれるといえる。しかし、今回の重回帰分析では決定係数の値が 0.04125、調整済み決定係数の値が 0.04073 と低いため回帰モデルとしてはまだ十分に説明できているとはいえない。データ数は十分にあったため、回帰モデルがうまく表せなかった原因として考えられるのは今回使用した変数が少なかったことが考えられる。より精度を上げるには複数の変数を追加する必要があると考えられる。しかし、レース開始前の時点で公開されている情報から荒れるレースに影響がある要素を明らかにしたことは進歩である。