

2023 年度 卒業論文

自作競馬予想プログラム の実用性について

大東文化大学経営学部経営学科

学籍番号 20161393

向後 郁哉

目次

1. はじめに
2. 研究概要
 - 2-1 全体の流れ
 - 2-2 使用する自作プログラムについて
 - 2-3 回収率等の結果について
3. 2022 年のレースの結果
 - 3-1 最終的な着順と各データの相関
 - 3-2 各馬券の回収率
4. 2023 年のレース結果
5. まとめ
6. 参考文献

1. はじめに

まず始めに、ギャンブルや賭けといったものは基本やらないほうが良いとされている。ギャンブルをするにあたって胴元（運用するもの）が存在するが、掛け金のうち、胴元が取得する割合を控除率と呼ぶ。例えば、宝くじであれば約 55%とされており、1000 円分買えば平均して 450 円戻ってくるという計算になる。パチンコであれば 15%、競馬や競輪などの公営ギャンブルは 25%となっている。当然だが、ギャンブルは必ず胴元が儲かるようになっているのである。打ち手（ギャンブルをする人）である我々は基本的には勝てないのだ。しかし、中にはパチンコで生計を立てているパチプロやプロ馬券師という者も存在する。彼らは研究に研究を重ね、機械の仕組みや特性を理解し、独自の予想方法を立てるなど、回収率が 100%を超えるようにうまく立ち回っているのだ。素人がいきなり彼らのようになることは当然難しい。ビギナーズラックがあったとしても続かないのがオチである。

だが、もし誰でも回収率 100%以上を出せる方法があるとしたらどうだろうか。掛け金を出せば出すほど儲かる、そんな甘い話は現実的に可能なのだろうか。私はこんな夢のような話を実現可能にしたいと思い、この研究テーマを定めた。

そのうえで一つ決めなければならないことは、どのギャンブルを研究対象とするかである。まず宝くじだが、これは運の要素が大きく、買い続けても何らかのデータが得られるとは思えない。次にパチンコだが、私は打ったこともなければ、当然知識もない。ネットに転がっている情報を拾ったところで機械の特徴をつかめるとも思えないし、データが集まる前に私が破産してしまいそうだ。あと代表的なギャンブルとして残っているのは公営競技である。競馬、競輪、競艇、オートレースの 4 つがある。その中でも私は競馬を選んだ。私自身、競馬が好きだからというのも理由の一つであるが、回収率を大幅に増やせる可能性が高いのだ。競輪は最大 9 車、競艇は 6 艇、オートレースが 8 車なのに対し、競馬は最大 18 頭で争われる。その分組み合わせの数、配当にも大きな違いが出る。この中で 1 番当てやすいのは何かと言われれば競艇だろう。しかしながら、配当が安いということも頭に入れておかなければならない。賭け方次第では的中率 90%でも、回収率が 100%を切ってしまうことだってあり得る。それに比べると競馬は当てることは難しいが、高配当を狙える。ギャンブルは的中率よりも回収率が重視される。その高配当を回収できる予想方法があれば、理論上、必ず勝つことはできるはずだ。

競馬において重要視すべきデータは何なのか、どの買い方が一番勝つことができるのか。この 2 点に視点を置き、研究を進めていこうと思う。

2. 研究概要

2-1 全体の流れ

今回の研究では、2022年のGIレースのデータを用いて進めていく。検証方法は以下のとおりである。

1. 独自で設定した競馬予想プログラムにデータを当てはめ、指数を出力する
2. 実際の着順とデータの各項目との相関を出力し、どのデータの重要性が高いかを見比べる。
3. 指数の上位5頭の組み合わせで馬券を購入したと想定し、馬券ごとの回収率を出す。
4. 今年の秋から冬にかけて行われたGIレースでも同様に結果を出し、指数が高かったが馬券内に来なかった馬、逆に指数が低かったが馬券内に来た馬について考察する。

このような流れで進めていく。

2-2 使用する自作プログラムについて

今回の研究を始めるにあたり、予想を目に見える形で数値として表わす必要がある。競馬の予想方法は人によって様々で、データ分析をする人、調教やレース直前の様子などで馬の調子を見極める人、中には直感で予想する人もいる。それぞれの予想方法で馬券の買い方の軸となる1頭を決めるわけだが、そこに至るまでの考えをすべて言語化することは難しい。また、競馬は挙げだしたらキリがないほどデータの量が多い。「○○(場所)で○月に産まれた馬は来ない」なんてデータも存在する。そこで今回は、ある程度使用するデータの量を絞ることにした。私が思う、レースの結果に大きく関係してきそうな以下に示す15項目を厳選し、各データの重みづけをした。

・対象コースの経験

馬自身が今回走るコースの経験があるかどうかを示している。画像は大阪杯のレースデータなので、阪神競馬場芝2000mのコースを走ったことがあれば「1」、無ければ「0」となっている。(1 2ポイント、0 0ポイント)

・対象コースの複勝率

今回走るコースにおいて、3着以内にきた割合を示している。(複勝率÷10をポイントとして加算する)

・レース経験

対象レースを走ったことがあるかどうかを示している。レースによっては、リピーター（過去に同レースで走ったことがある）馬の好走率が高いこともある。（有は1 2ポイント、無は0 0ポイント）

・年齢

競走馬の成長曲線は様々だが、大体3, 4, 5歳がピークとなる。馬によってはそれ以上の年齢でも走ることがあるが、7歳以上となると基本的には好走は厳しい。（3歳4ポイント、4歳5ポイント、5歳5ポイント、6歳4ポイント、7歳2ポイント、8歳0ポイント、9歳-5ポイント）

・過去1年の勝利数、2着数、3着数

過去1年の間に、どれだけ好成績を残したかを示す。（勝利数×3、2着数×2、3着数×1ポイント）

・馬場状態における複勝率

馬が走る地を馬場というが、水分をどれほど含んでいるかによって柔らかさが変わり、4段階に分類される。芝であれば、踏みしめた際、馬場の表面がほとんど変化しない状態を「良」、踏みしめた際に水は出ないが、馬場の表面がやや凹む状態を「稍重」、表面に水は浮いていないが、踏みしめると水が浮く状態を「重」、表面や足跡に水が浮いている状態が「不良」となっている。（複勝率÷10をポイントとして加算する）

・産駒の複勝率

競馬はブラッドスポーツと呼ばれる側面を持ち、血統が重要視される。同じ父親を持つ馬がどれだけそのコースで走れているかを示している。（複勝率÷10をポイントとして加算する）

・ペース適正

マラソンでも同様のことがいえるが、レースや会場、出走する馬の脚質によってペースは大きく異なる。ハイペース、ミディアム、スローペースに分類される。展開を予想した際、その馬がペースに合うかどうかを示している。（1 5ポイント、0 0ポイント）

・騎手のコース複勝率

騎乗する人物が、そのコースをどれだけ得意としているかを示している。（複勝率÷10をポイントとして加算する）

・左・右回り複勝率

コースによってどちら向きに回るかが決まっており、馬によって得意不得意がある。
 (複勝率÷10をポイントとして加算する)

・同距離複勝率

馬がその距離を得意としているかどうかを示している。(複勝率÷10をポイントとして加算する)

・昨年同レース成績

昨年同レースを走った馬がいれば、その着順が記載されている。前年年大きく負けてしまった馬が巻き返すことは基本難しく、逆に好走した馬は再び好走する確率も高い。(6着以下は1ポイント、4、5着は3ポイント、3着5ポイント、2着7ポイント、1着9ポイント)

・調教評価

競走馬の訓練動画を視聴し、状態の良し悪しを見極める。調子が良さそうなら3、普通なら2、あまりよくなさそうな場合は1となっている。(3 10ポイント、25ポイント、10ポイント)

計算式

=COUNTIF(B3,"1")*2+COUNTIF(B3,"0")*0+C3/10+COUNTIF(D3,"1")*2+COUNTIF(D3,"0")*0+COUNTIF(E3,"3")*4+COUNTIF(E3,"4")*5+COUNTIF(E3,"5")*5+COUNTIF(E3,"6")*4+COUNTIF(E3,"7")*2+COUNTIF(E3,"8")*0+COUNTIF(E3,"9")*5+F3*3+G3*2+H3+I3/10+J3/10+COUNTIF(K3,"1")*5+COUNTIF(K3,"0")*0+L3/10+M3/10+N3/10+COUNTIF(O3,"0")*0+IF(O3>=6,"1")+COUNTIF(O3,"5")*3+COUNTIF(O3,"4")*3+COUNTIF(O3,"3")*5+COUNTIF(O3,"2")*7+COUNTIF(O3,"1")*9+COUNTIF(P3,"3")*10+COUNTIF(P3,"2")*5+COUNTIF(P3,"1")*4

また、2022 大阪杯でのデータを以下に示す。加点式になっており、その合計点数が予想となる。

2022大阪杯								良
馬番	対象コースの経験	対象コースの複勝率	レース経験	年齢	過去1年の勝利数	過去1年の2着数	過去1年の3着数	馬場状態における複勝率
1	1	50	0	6	2	1	0	33
2	1	100	0	4	1	1	0	50
3	0	0	0	6	0	2	0	24
4	1	100	0	4	6	0	0	88
5	1	50	0	5	2	1	2	53
6	0	0	0	4	3	1	0	100
7	0	0	0	5	1	0	0	63
8	1	100	0	5	0	1	1	75
9	1	100	0	4	4	1	1	71
10	0	0	0	6	0	1	0	82
11	1	50	0	4	1	1	0	60
12	1	80	0	6	2	1	2	55
13	0	0	0	6	1	0	0	53
14	1	100	1	5	1	1	1	67
15	1	0	0	7	1	1	0	38
16	1	0	1	9	2	0	0	30
17								
18								

	M予想							
産駒の複勝率(過去3年)	ベース運正	騎手のコース複勝率	左・右回り複勝率	同距離複勝率	昨年同レース成績	騎教評価	合計点数	着順
25	1	44	54	30	0	2	47.6	6
39	1	17	57	50	0	2	53.3	13
31	1	0	38	50	0	2	32.3	15
35	1	65	100	89	0	3	87.7	5
29	1	17	47	55	0	2	52.1	10
37	1	0	100	100	0	2	59.7	9
17	1	0	40	75	0	2	37.5	16
39	1	0	86	82	0	2	58.2	1
42	1	6	86	100	0	2	72.5	3
25	1	30	78	80	0	1	40.5	4
29	1	39	83	33	0	2	51.4	8
40	1	36	50	48	0	2	56.9	12
32	1	45	73	50	0	2	42.3	11
39	1	67	67	75	1	3	80.5	2
17	1	9	36	10	0	2	30.0	7
39	1	26	53	40	11	2	34.8	14
							0.0	
							0.0	

2-3 回収率等の結果について

導き出された合計点数の上位5頭を対象に馬券を購入したと仮定して、馬券ごとの回収率や的中率などの結果を示していく。回収率とは、馬券の購入金額を100%とし、100%を上回ればプラス、下回ればマイナスとなる。以下は馬券の種類についての説明である。

- ・単勝…1着馬を当てる馬券
- ・複勝…3着以内に入線する馬を当てる馬券
- ・枠連…1, 2着の枠を当てる馬券
- ・馬連…1, 2着の馬を順不同で当てる馬券
- ・ワイド…3着以内の馬2頭の組み合わせを当てる馬券
- ・馬単…1, 2着の馬を当てる馬券
- ・3連複…1, 2, 3着の馬を順不同で当てる馬券
- ・3連単…1, 2, 3着の馬を当てる馬券

以上の8種類が馬券の種類となる。

3. 2022年のレース結果

3-1 最終的な着順と各データの相関

では、2022年に行われたレースを対象に、それぞれのデータの相関を見ていく。

	対象コースの経験	対象コースの複勝率(過去1年)	レース経験	年齢		
相関	0.019305267	-0.114519072	0.054533122	0.11281507		
	過去1年の勝利数	過去1年の2着数	過去1年の3着数	馬場状態における複勝率	産駒の複勝率(過去3年)	ペース適正
	-0.153931866	-0.085216886	-0.083367812	-0.207947743	-0.048479648	-0.03014502
	騎手のコース複勝率	左・右回り複勝率	同距離複勝率	昨年同レース成績	調教評価	合計点数
	-0.153734763	-0.146867193	-0.081471137	0.121057252	-0.34877	-0.29331

これは、2022年のすべてのGIレースの各データと着順の相関である。

表を見ていくと、ほとんどのデータに相関がないといえる。産駒の複勝率や騎手のコース複勝率は多少なりとも相関が出ると予想していたが、全くと言っていいほど関係がないことに驚いた。昨年の同レースの成績はこの中では比較的相関が強いほうではあるが、それでも非常に弱いものである。また、このデータは他に比べて対象とする馬が少ないため、その分他よりも大きい数字が出たのだと推測できる。指標となる合計点数に関しても、0.3をぎりぎり超えた程度の弱い相関であった。全頭の着順を完全的中することは不可能なので、指標の上位5頭ずつで相関を出してみれば結果が変わるかもしれない。

もう少し細かく見ていこうと思う。

	対象コースの経験	対象コースの複勝率	レース経験	年齢	過去1年の勝利数	過去1年の2着数	過去1年の3着数
フェブラリーステークス	-0.130933459	-0.682439228	-0.234473368	0.57967033	-0.166451821	0.223035643	0.062647001
高松宮記念	0.005313064	-0.179674264	0.143873352	0.10678454	-0.200788967	-0.517593132	0.077918135
大阪杯	-0.365636212	-0.51004687	-0.040996003	0.29855685	-0.172838995	-0.141596999	-0.182869921
桜花賞	-0.331954493	-0.140575759	#DIV/0!	#DIV/0!	-0.054539568	-0.265605896	-0.089480906
皐月賞	0.295302053	-0.106161443	#DIV/0!	#DIV/0!	-0.612015905	0.240934712	0.323324246
天皇賞春	-0.15539822	-0.511099707	-0.15539822	0.50948841	-0.441171989	-0.063241024	-0.381811539
NHKマイルカップ	0.394472405	0.306659824	#DIV/0!	#DIV/0!	-0.419674387	0.129942418	-0.148035969
ヴィクトリアマイル	-0.096373885	-0.26245024	0.227120476	0.15146607	0.128236608	-0.095405279	0.571940742
オークス	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	-0.214020508	-0.095806126	0.086260729
日本ダービー	0.003725125	0.003725125	#DIV/0!	#DIV/0!	-0.104303492	0.178805987	-0.360068823
安田記念	-0.206056033	-0.410630216	0.15539822	0.62846951	-0.114885296	0.0279609	-0.224872398
宝塚記念	0.346657568	0.255247076	0.350611506	0.21691091	-0.132422618	-0.10943736	-0.043099712
スプリンターズステークス	-0.014625448	-0.031611665	0.108465229	-0.0930672	-0.086841684	-0.161874052	-0.266789188
秋華賞	-0.156315032	-0.196039212	#DIV/0!	#DIV/0!	-0.316631891	-0.414858432	-0.277883521
菊花賞	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	-0.085183284	0.063091518	-0.373415998
天皇賞秋	0.090189413	-0.036712291	0.405403757	0.53787685	-0.412348254	-0.484955788	0.159386529
エリザベス女王杯	-0.23226194	-0.29272651	0.05225627	0.03207192	0.132343523	0.063344165	-0.143088525
マイルチャンピオンシップ	0.011369511	-0.498284073	0.132951551	0.54893489	-0.304762019	-0.194964061	0.153083387
ジャパンカップ	-0.340733138	-0.594023947	-0.322750149	0.40998737	-0.230825895	0.307914995	-0.490037285
チャンピオンズカップ	0.043876345	-0.163259305	0	0.40829554	-0.178315738	-0.26647402	-0.031642559
阪神ジュベナイルフィリーズ	0.204439883	0.289121655	#DIV/0!	#DIV/0!	0.408879765	-0.412351079	0.231813038
朝日杯フューチュリティステークス	0	0	#DIV/0!	#DIV/0!	-0.195180015	0.288675135	0.056613852
有馬記念	0.669601384	0.229315284	0.248632624	0.56527676	-0.151822173	-0.081758743	-0.228924479
ホープフルステークス	0.009182431	0.009182431	#DIV/0!	#DIV/0!	-0.339950828	-0.159705674	-0.377486138

馬場状態における複勝率	産駒の複勝率(過去3年)	ペース適正	騎手のコース複勝率	左・右回り複勝率	同距離複勝率	昨年同レース成績	調教評価	合計点数
0.056967843	0.215293493	0.28022427	0.039032153	-0.091857868	-0.108393688	-0.16374758	-0.4354	-0.4898
-0.252035019	-0.214262965	-0.1285059	0.102341427	-0.116661844	-0.167954704	0.091695364	-0.2146	-0.4188
-0.498084432	-0.198619484	#DIV/0!	-0.302893286	-0.592772941	-0.358149839	0.275495862	-0.174	-0.5838
-0.243696782	0.406027511	#DIV/0!	0.225714643	-0.175991949	-0.045516003	#DIV/0!	-0.294	-0.2051
-0.47529525	-0.01479701	#DIV/0!	-0.309046475	-0.074075107	0.092011472	#DIV/0!	-0.676	-0.3927
-0.134489743	-0.245756349	#DIV/0!	0.017313749	-0.645043281	-0.511099707	0.191045581	-0.1137	-0.6649
-0.269783138	0.152755575	#DIV/0!	-0.215503131	0.178573868	-0.423362226	#DIV/0!	0.13629	-0.0961
-0.136049953	0.568325722	#DIV/0!	-0.104545384	-0.002756724	-0.063363226	0.219103662	-0.2268	-0.0501
-0.27716412	0.025707158	#DIV/0!	-0.414511988	-0.271138911	#DIV/0!	#DIV/0!	-0.1829	-0.4034
-0.136208695	-0.033446071	#DIV/0!	-0.442505798	-0.360048435	0.003725125	#DIV/0!	-0.7446	-0.4111
-0.69276162	0.084683082	#DIV/0!	-0.282229653	-0.723928198	-0.465864111	0.395260644	-0.3017	-0.5831
-0.389852877	-0.097337722	#DIV/0!	0.257121518	-0.158392818	0.353930304	0.210366904	-0.3962	0.02549
-0.358094947	-0.101981666	#DIV/0!	-0.365618247	-0.196904158	-0.092110654	0.140356496	0.04523	-0.2569
-0.618365135	-0.033568083	#DIV/0!	-0.268702059	-0.405208116	-0.00382575	#DIV/0!	-0.6468	-0.6329
-0.207632993	0.050350757	#DIV/0!	-0.048186942	0.073508079	#DIV/0!	#DIV/0!	-0.3918	-0.2484
-0.277625343	-0.006504446	#DIV/0!	-0.110228552	-0.308604039	0.079919619	0.345338003	-0.5662	-0.393
-0.133492479	0.204341588	#DIV/0!	-0.076484008	0.474290671	-0.303229794	0.105565464	0.22333	-0.1396
-0.273340462	-0.248674325	#DIV/0!	0.184663665	-0.506700403	-0.231209676	0.166298717	-0.5754	-0.4983
-0.121753945	-0.466583036	#DIV/0!	-0.348474486	-0.566715125	-0.234023189	-0.010172838	-0.5292	-0.6928
-0.49209722	0.312663377	#DIV/0!	-0.560510944	-0.36046318	-0.315882984	-0.022686711	-0.5694	-0.4931
0.015353155	-0.184380687	#DIV/0!	-0.002402786	0.536075853	-0.198224827	#DIV/0!	-0.251	0.16078
-0.424403515	-0.097453074	#DIV/0!	-0.350192417	0.167722385	-0.18683559	#DIV/0!	-0.5198	-0.2224
-0.529110137	-0.062310796	#DIV/0!	-0.213011119	-0.133119304	0.25984208	0.443683329	-0.6956	-0.108
0.007453995	-0.177899114	#DIV/0!	-0.291200421	-0.271995698	-0.259675552	#DIV/0!	-0.172	-0.2958

※レース経験、年齢、ペース適正、昨年同レース成績に関しては、年齢限定戦、メディアムペース（早くも遅くもない通常のペース）が要因ですべての馬が同数値になり、#DIV/0が出ている箇所あり。

この表はレースごとに各データと着順の相関をまとめたものである。年齢で相関が出ていない部分は、全馬同じ年齢の世代戦であるというのが理由である。

左から順にみていくと、対象コースの複勝率についてはフェブラリーステークス、大阪杯、天皇賞春、ジャパンカップで強めの相関が出ていた。フェブラリーステークスで優勝したカフェファラオは前年度の覇者であり、このコースで3勝無敗であった。天皇賞春で2着のディーブポンドは前年も2着で、大阪杯で優勝したポタジェはコースの複勝率が100%、ジャパンカップで優勝したヴェラアズールも同様であった。しかしながら、それ以外の馬もコース適正があったかという、そうとも限らない。例えば、フェブラリーステークスで2着に入線したテイエムサウズダンはこのコースでの勝利経験はなく、3着に入線したソダシに関しては、初体験のコースだった。いずれの馬も重賞で多くの勝利を重ねており、実力は確かなものがある。GIにおいてはやはり実力者が馬券に絡んでくるのがセオリーといえるだろう。

続いて年齢だが、こちらは多くのレースで高い相関が出ていた。年齢が高いとやはり勝つのは厳しいといえる。しかし、年齢の部分で唯一負の相関に寄っていたスプリンターズステークスでは、7歳馬のジャンダルムが優勝した。同年の3月にも同コースで勝利を取っており、衰えていなかったといえるだろう。次に過去1年間の1, 2, 3着数だが、意外にも高い相関はなく、あまり結果には直結していないといえる。

馬場状態における複勝率では、3レースで高い相関が出ていた。いずれのレースも良馬場開催で、勝ち馬、入線馬共に良馬場で高い複勝率を誇っていた。しかし、一般的に適性が必要とされている重馬場での開催であった高松宮記念とエリザベス女王杯では、高い相関は見られなかった。2つのレースの共通点といえば、どちらも3連単20万円以上の大荒れになったということだろうか。

次に産駒の複勝率だが、こちらもあまり高い相関は見られなかった。逆にヴィクトリアマイルでは、むしろ実績の低い種牡馬のほうが好走しているということになる。GIレベルの高いレースになると、同じ父親を持つ馬が多数出走する。また、現役時代強かった競走馬であれば、年間200頭もの子供たちが誕生する。いくら成績を出す馬が多くても、まったく走らないといった子も当然いるため、一概にこのデータだけで実力を図ることはできないというのも相関が低い理由の一つなのだろう。

次に騎手のコース複勝率についてだが、こちらも高い相関は見られなかった。唯一高い相関が出たチャンピオンズカップでは、このコースを得意としている福永ジョッキーと横山典弘ジョッキーが2、3着に入線している。GIレベルの高いレースになると、一流のジョッキーたちが集うことになる。いくらそのコースが得意な若手騎手がいたとしても、強い馬に乗ったトップクラスのジョッキーには勝つことは難しい。このデータに関しては、GI以外のレースで見ると面白いかもしれない。

続いて左・右回りの複勝率についてだが、6つのレースで高い相関が出た。特に安田記念では非常に高い相関が出ており、左回りに強い馬達が上位を独占する形となった。次に同距離の複勝率だが、意外にも高い相関は見られなかった。唯一相関が高かったのは天皇賞春であるが、そもそも芝3200mで行われるGIはこのレースしかなく、データの母数が少ないことと、昨年好成績の馬が好走したことで相関が高くなっている。GIレベルのレースではその距離のスペシャリストたちが集うため、それまで同距離で負けていない馬でも、実力差などの様々な要因で3着入線を逃すこともある。こちらのデータも先ほどの騎手のコース複勝率と同様、GI以外のレースに当てはめれば高い相関が得られると思われる。

次に昨年と同レース成績だが、こちらはデータの母数が少ないために相関が高くなりやすくなっている。しかしながら、前年好走している馬の好走率はやはり高いものになっているため、馬券を購入する際は抑えるべきデータであるといえる。最後に合計点数との相関だが、他のデータに比べて著しく低いものはなく、悲観するほど悪い予想ではないといえるだろう。

この結果から言えることは、1つのデータを見るだけでは競馬の予想など到底でき

ないということだろう。ほかのデータとの兼ね合いや単純な馬の実力など、ここにはないものも予想に含めなければ、競馬で勝つことは難しい。

3-2 各馬券の回収率

それでは自作したプログラムで出力した指標の上位5頭で馬券を購入したと想定し、的中率、回収率を見ていこうと思う。

	単勝	複勝	枠連	馬連	ワイド	馬単	3連複	3連単
フェブラリーステークス	510	540	700	2640	1230	4690	0	0
高松宮記念	2780	660	0	0	0	0	0	0
大阪杯	5870	2800	4790	10980	12260	39630	50990	537590
桜花賞	1450	360	2550	0	0	0	0	0
皐月賞	0	140	0	0	0	0	0	0
天皇賞春	0	380	450	0	500	0	0	0
NHKマイルカップ	0	0	1190	0	0	0	0	0
ヴィクトリアマイル	570	420	1770	2010	730	4010	0	0
オークス	650	960	1420	8150	2890	12750	0	0
日本ダービー	420	720	420	730	2850	1440	4570	15770
安田記念	820	470	720	1740	830	3740	0	0
宝塚記念	420	180	0	0	0	0	0	0
スプリンターズステークス	0	510	0	0	0	0	0	0
秋華賞	570	280	1040	0	320	0	0	0
菊花賞	410	160	0	0	0	0	0	0
天皇賞秋	260	600	1640	3330	1210	4930	0	0
エリザベス女王杯	0	0	0	0	0	0	0	0
マイルチャンピオンシップ	0	0	0	0	0	0	0	0
ジャパンカップ	450	540	470	940	1470	1920	2360	9850
チャンピオンズカップ	0	760	0	0	1920	0	0	0
阪神ジュベナイルフィリーズ	0	0	0	0	0	0	0	0
朝日杯フューチャリティステークス	310	270	560	550	260	1070	0	0
育馬記念	0	0	0	0	0	0	0	0
ホープフルステークス	0	480	0	0	0	0	0	0
合計	15490	11230	17720	31070	26470	74180	57920	563210
増減	3490	-770	820	7070	2470	26180	9920	419210
的中率	58.3%	79.2%	54.2%	37.5%	50.0%	37.5%	12.5%	12.5%
回収率	129%	94%	105%	129%	110%	155%	121%	391%

こちらが今回の結果である。なんと複勝以外のすべての馬券で回収率100%以上を達成することができた。回収率100%越えを達成できた理由として、大阪杯の的中が大きい。このレースは3連単が50万円を超える波乱のレースとなったが、1着の8番人気ポタジェを5番手評価、2着の3番人気レイパパレを2番手評価、3着の7番人気9番アリーヴォを3番手評価していた。これらの馬を高く評価できた理由としては、3頭ともに2000m、右回りの複勝率がよかったことが挙げられる。今回私が作成したプログラムではそういった複勝率の部分に重きを置いていることもあり、それが

上手くはまったといえる。

4. 2023 年のレース結果

先ほどと同様に 2023 年の秋競馬を予想していく。

	対象コースの経験	対象コースの複勝率	レース経験	年齢	過去1年の勝利数	過去1年の2着数	過去1年の3着数
エリザベス女王杯	#DIV/0!	#DIV/0!	-0.270030862	0.4082912	-0.018073601	-0.474252007	0.252538136
マイルチャンピオンシップ	0.388290137	0.470678724	-0.194911726	-0.1519109	0.217757514	0.120127144	0.048650426
ジャパンカップ	-0.457749747	-0.681754942	0.032301791	0.5918318	-0.65769967	-0.153260433	0.034532067
チャンピオンズカップ	-0.098198051	0.187967167	0.032732684	-0.1128837	-0.197859424	0.016126236	-0.231455025
阪神ジュベナイルフィリーズ	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	-0.196213429	-0.203619063	-0.436932572
朝日杯フューチュリティステークス	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	-0.36279073	0.304074893	0.21127377

馬場状態における複勝率	産駒の複勝率(昨年)	ペース適正	騎手のコース複勝率	左・右回り複勝率	同距離複勝率	昨年同レース成績	調教評価	合計点数
-0.443360539	-0.044765914	0.077151675	-0.376458277	-0.16510061	0.0269061	-0.06789951	0.073193	-0.3224
0.279258371	0.31675911	#DIV/0!	-0.03600509	-0.151347907	0.086047691	-0.00686665	-0.328691	-0.4653
-0.672748395	-0.286873831	-0.690641334	-0.42606544	-0.627739429	-0.404013114	0.258996336	-0.727142	-0.9158
-0.287881985	0.392183678	#DIV/0!	0.442675953	-0.335964498	0.308163744	0.220924848	-0.199667	0.08058
-0.479101163	-0.509395601	#DIV/0!	-0.698313569	0.158737988	-0.585300461	#DIV/0!	-0.317879	-0.7106
-0.305769225	0.274030702	#DIV/0!	-0.04875263	-0.084122897	-0.073791827	#DIV/0!	-0.438849	-0.3277

※レース経験、年齢、ペース適正、昨年同レース成績に関しては、年齢限定戦、ミディアムペース（早くも遅くもない通常のペース）が要因ですべての馬が同数値になり、#DIV/0 が出ている箇所あり。

こちらが 2023 年の秋競馬における、各データと着順の相関である。特に目を引くのがジャパンカップで、指標と着順の相関が 0.9 と非常に高いものになった。また、阪神ジュベナイルフィリーズでも 0.7 という強い相関を記録し、母数は少ないものの予想精度は高いといえる。

続いて、回収率も見ていく。

	単勝	複勝	枠連	馬連	ワイド	馬単	3連複	3連単
エリザベス女王杯	240	310	570	0	380	0	0	0
マイルチャンピオンシップ	0	840	0	0	1780	0	0	0
ジャパンカップ	130	430	180	180	880	260	600	1130
チャンピオンズカップ	0	0	0	0	0	0	0	0
阪神ジュベナイルフィリーズ	0	450	0	0	770	0	0	0
朝日杯フューチュリティステークス	270	450	0	0	590	0	0	0
合計	640	2480	750	180	4400	260	600	1130
増減	-2360	-520	-3650	-5820	-1600	-11740	-11400	-34870
的中率	50.0%	83.3%	33.3%	16.7%	83.3%	16.7%	16.7%	16.7%
回収率	21%	83%	17%	3%	73%	2%	5%	3%

こちらが 2023 年のエリザベス女王杯からのデータになる。回収率を見てみると最高でも複勝の 83% とマイナスになってしまっている。チャンピオンズカップに至っては、指名馬が 1 頭も馬券内に来ていないという結果になってしまった。しかし、良い

点もいくつかある。まずは的中率である。データの母数は少ないものの、チャンピオンズカップ以外では指標上位5頭のうち2頭が3着以内に入線している。また、ジャパンカップでは指標1番手評価のイクイノックスが1着、2番手評価のリバティアイランドが2着、3番手評価のスターズオンアースが3着と、指標の上位3頭で完璧に的中させることができた。5番手評価のドウデュースも4着に好走し、プログラムの予想通りに決まったといえるだろう。では、なぜこれほどまでに回収率が悪いのだろうか。理由は今年のレース傾向にある。昨年は1番人気がなかなか勝てないどころか馬券外になるレースも多く、堅く決着するレースは少なかった。今年はそんな昨年の傾向から変わり、1番人気が高い確率で入線している。

	1番人気の勝率	//の複勝率
2022	16.67%	33.33%
2023	41.67%	72.73%
※2023は朝日杯FSまでの22レース		

2022年はG I 24レース中1番人気は僅か4勝、3着以内で見ても8回と苦戦しているのに対し、今年は朝日杯の時点ですでに1番人気は10勝し、昨年の複勝率よりも高い状態となっている。内訳をみると10勝のうち3勝が世界最強と名高いイクイノックスと、今年3冠牝馬に輝いたリバティアイランド、2勝がダート界の新星レモンポップであり、この3頭だけで8勝している。また、1番人気に2回支持されたスターズオンアースもそれぞれ2、3着と結果を残している。つまり、1番人気になるべくしてなった馬が、着実に好成績を収めているのだ。当然この馬たちは人気するためオッズが低くなり、相当な穴馬が来ないと配当が高くつくことはない。今年は如何に買い目を絞り、金額を厚く張るかがカギになる年だったのだ。今年のほかのレースも予想すれば結果は変わるかもしれないが、全体的に配当が低いため、回収率は100%を切ることが予想される。

5. まとめ

今回私は自作競馬予想プログラムの実用性についてというテーマで研究を進めてきた。自分でもまさか競馬についての研究をすることになろうとは思ってもしなかった。しかし、今思えばこの研究をしてよかったと思う。もしこの論文に目を通し、競馬の魅力が少しでも伝わったなら本望である。

2022年の競馬データをもとに、今回の研究を進めてきたが、やはり競馬の予想は難しいと感じた。単勝や複勝、ワイドなどの馬券は当たりやすいが配当は低くなる。逆に3連系の馬券は、的中率は低い配当は大きい。今回の研究では、それなりに高い的中率で回収率が少しでも100%を超えることが目標であった。今回、この研究は間違いなく結果が出ているといえる。しかしながら、競馬などのギャンブルをする人は少し勝った程度では満足しない性質で、大勝ちしたくなるものなのだ。今回作成したプログラムを5年間ほど運用し、それでも回収率が100%を超えるのなら、おそらく自分で買うよりもこちらを信用して賭け続けたほうが確実に所持金は増える。小勝ちを続けたほうが効率は間違いなく良いのだ。だからこそ、このプログラムで大阪杯の3連単50万を的中できたのは正直驚いている。通常こういった予想プログラムは、的中率が高く、小勝ちを重ねるものが多い。実際私もそういったものになるだろうと予測していた。予想とは違う形になったが、ある意味作成した人の個性が出たプログラムになったといえるだろう。

ひとつひとつのデータを相関で見ても特に強い数値は現れないが、データが組み合わさることで一つの予想として成り立つのだと思う。前述したように競馬には様々なデータが存在する。どのデータを重要視するかまたは軽視するのか、そういうところも含めて競馬の予想というのは実に奥深く面白いものである。今回のプログラムをさらに発展させ、データを組み合わせた際の相関や、新たなデータ、そのレースの特徴を裏付けるデータなど、まだまだ精度は上げられそうである。目標を1つ設定するのなら、高い確率で3連単を当てられる夢のようなプログラムを作りたい。

私が子供の頃はたまにテレビ中継の競馬を観戦する程度で、そこまで興味がなかった。しかし、2012年の有馬記念で見たある1頭が私を競馬の魅力へと引きずり込んだ。大学生になってから競馬を本格的に見はじめ、馬券を購入できるようになるまではひとりでの予想を楽しんでいた。2021年、たまたまテレビで観戦していたG I 優駿牝馬。出走表を特にみることもなく、ただレースを眺め、1頭の馬が先頭でゴールした。私が有馬記念で見たあの馬の産駒だったのだ。そこからもう私はたまたまなく競馬が好きになった。競馬は紛れもなくギャンブルであり、疎遠されてしまう趣味なのか

もしれない。しかし、筋書きのない一つのドラマであるという側面もあることを覚えておいてもらいたい。競馬場にはお金を増やすためでなく、推し馬を応援するために来る人も中にはいるのだ。馬という命を扱っている以上競馬をただのギャンブルとしてではなく、スポーツとして、命を軽んじることがないように、これからも好きであり続けたい。

6. 参考文献

控除率について

<https://honkawa2.sakura.ne.jp/5680.html>

データ

ネット競馬 <https://www.netkeiba.com/>

KeibaList <https://keibalist.com/>

指導教員からの講評

自分自身の興味から、適切なテーマを選択し、精力的に分析を行うことができました。競馬予想は、すでに多くのプログラムがあり、なかなかそれを超えることは難しいかもしれませんが、といて、諦めていては何も進みませんので、まずはこうしたテーマに前向きに取り組んだ姿勢を評価したいと思います。今回、まとめられた卒業論文を見て、「引き際を考えることも重要ではないか」という気が強くしました。例えば、参加者の懐事情や、置かれた状況、過去の戦績などを考慮し、確実なところであれば勝負に出ても良いが、そうでない場合には勝負を踏みとどまらせるようなプログラムです。一般に負けが混むとハイリスク・ハイリターンを求めがちですが、それは懐事情にもよるはずなので、そのへんを冷静にサジェストできるようなものはあり得るかもしれません。