

競艇の順位を予想する

順位データの統計的モデリング

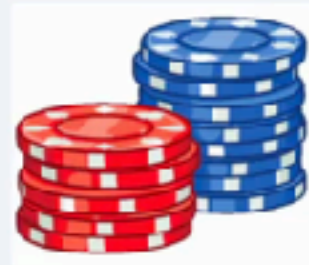
<https://github.com/isaomatsunami/Japan.R2023>

パッケージの背後にはfantasy sports

Home Packages About Blog Stats Snippets


Python package for working with Sports data.

recruitr-py



College Sports Recruiting - Python
Access College Sports Recruiting Data.


nwsipy



Women's Soccer - Python
nwsipy is a Python wrapper around the data generated by the nws1R package

R Packages

{sportsdataverse}



All Sports - R

The 'sportsdataverse' is a set of sports data packages that work in harmony because they share common data representations and API design. This package is designed to make it easy to install and load multiple 'sportsdataverse' packages in a single step.

[usheon] [efbfootD] [hcapD]

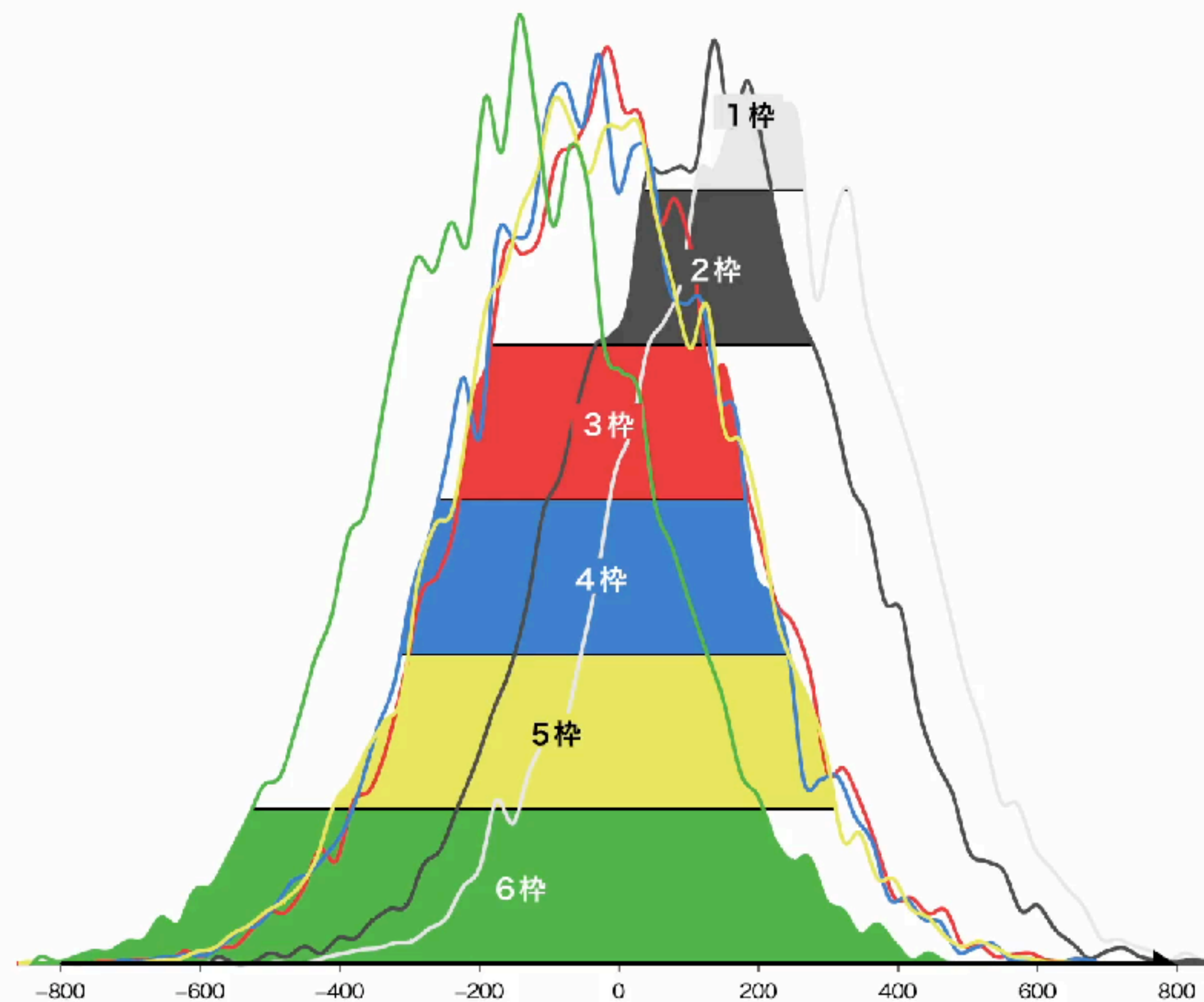
<https://www.sportsdataverse.org/packages>

レースで確率表現をしたい

カジノ	ブラックジャック ルーレット スロット	確率計算可能	リスクの受け手
	ポーカー 麻雀	確率計算可能	場所・ディーラーの 提供
レース	競馬 競艇 ドッグレース	オッズ	リスクなし

10R:各枠の推定パラメータ

弱い↔強い



データ:2022/11/24から11/18の女子限定3360レース

1 枠

2 枠

3 枠

4 枠

5 枠

6 枠

<https://race.chunichi.co.jp/boat/daily/08/today>

3連単:10R

閉じる

舟券	確率	基準 (75%相当)	オッズ 10:52時点
1-2-3	6.0%	12.5	↓18.8
1-2-4	5.5%	13.7	↓17.1
1-2-5	5.5%	13.7	↓10.6
2-1-3	4.4%	17.0	↓31.8
2-1-5	3.9%	19.0	↑18.3
2-1-4	3.9%	19.2	↓24.6
1-3-2	3.6%	20.5	↑37.6
1-5-2	3.5%	21.2	↓16.9
1-4-2	3.0%	24.6	↓28.6
1-2-6	3.0%	25.3	↑142.3
2-1-6	2.2%	34.2	↑138.6
2-3-1	2.1%	35.9	↑57.5
1-3-4	2.0%	37.7	↑43.7
2-5-1	1.9%	38.9	↑28.3
2-4-1	1.9%	39.7	↓35.8
1-3-5	1.9%	40.3	↓40.5
1-5-4	1.8%	41.2	↓43.4
1-4-3	1.8%	41.7	↓61.1
1-4-5	1.7%	43.6	↓42.0

2連単:10R

閉じる

舟券	確率	基準 (75%相当)	オッズ 10:52時点
1-2	19.9%	3.8	↑4.6
2-1	14.4%	5.2	↓6.5
1-3	8.3%	9.0	↑16.7
1-5	7.7%	9.7	↑11.5
1-4	7.5%	10.0	↑14.7
2-3	4.0%	18.7	↓18.4
2-5	3.9%	19.3	↓13.1
2-4	3.8%	19.9	↑7.8
3-1	3.5%	21.7	26.3
4-1	3.1%	24.2	↑46.0
1-6	2.9%	25.5	↓33.4
5-1	2.9%	25.9	↓20.4
3-2	2.8%	26.7	↓21.6
5-2	2.3%	32.2	↓16.7
4-2	2.0%	36.8	↓21.6
2-6	1.5%	51.7	↓73.6
4-3	1.0%	72.8	↓122.7
5-3	1.0%	73.5	↓122.7
3-4	1.0%	77.3	↑23.0

単勝:10R

舟券	確率	基準 (75%相当)	オッズ 10:52時点
1	46.4%	1.6	↑3.2
2	27.6%	2.7	↓1.9
3	8.5%	8.8	↑4.8
5	7.5%	9.9	↑9.7
4	7.4%	10.2	↑9.7
6	2.6%	29.2	↑9.7

<https://race.chunichi.co.jp/boat/daily/08/today>

順位データからパラメータを推定する

<https://github.com/isaomatsunami/Japan.R2023>



当てることと 勝つことの間 競艇の統計的モデリング

競艇ほど当てることと勝つことの違いを教えてくれるギャンブルはない。1 枠に賭ければ52%も当たるが、資金は着実に減っていく。勝つための戦略は存在しないのだろうか？

推定確率はおおむね現実に合致

モデルが推定した確率と実際の勝率

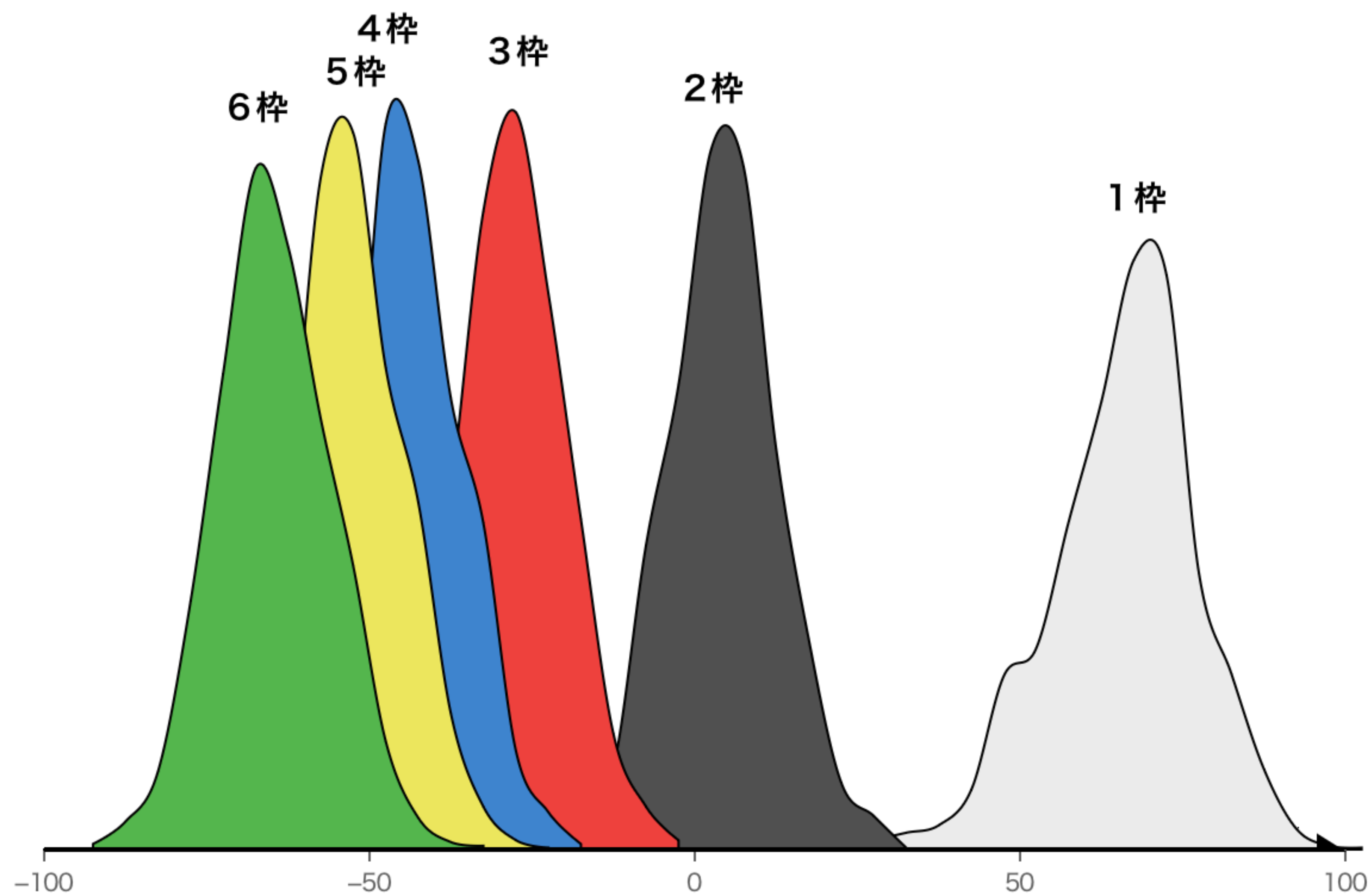
推定確率	総舟券数	当選券	実際の勝率
~1%	210804	950	0.5%
1~2%	64829	875	1.3%
2~3%	33981	847	2.5%
3~4%	18626	617	3.3%
4~5%	11965	524	4.4%
5~10%	30191	2025	6.7%
10~15%	13311	1636	12.3%

問題点:A級とB級が重なる

- A級上位選手とB級上位選手は同じ位置に配置されてしまう。競艇ではA級選手しか出場できない大会があり、トップ選手はB級選手と交流することが極めて稀。このため、2つのグループが重ねて配置されてしまう。
- 野球や将棋でも恐らく同様
- ただし、同じ理由（対戦することが稀）で舟券の予想には影響がない。

1 枠は格段に有利

「枠の力」の推測値の分布



対象:23/3/31までの20000レース

問題点:パラメータに干渉がある

- ・ 枠順の影響力を考えると、どの枠が1着になっても2着以降は若い枠になる確率が高いはずなのに、**現実のデータでは3、4枠が1着になると4枠、5枠が2着になる確率が比較的高い**。これは、3、4枠が1着になる場合、内側の1、2枠がチキンレースになってターンに失敗していることが多く、後続の4、5枠が上位に食い込むことができるためだ。
- ・ 3、4枠が1位になる時だけ、4、5枠のパラメータがブーストされる「パラメータの干渉」をどうモデル化したらいいか？