

競馬予想とは何か

競馬ファンの集団知と機械学習による予想の間に横たわるもの

Accuracy and Fine Structure of Predictions in a Racetrack Betting Market

守 真太郎
Mori Shintaro

北里大学 理学部 物理学科
Department of Physics, School of Science, Kitasato University
mori@sci.kitasato-u.ac.jp, <http://sharaku.sci.kitasato-u.ac.jp/mori>

久門 正人
Hisakado Masato

スタンダード & プアーズ
Standard & Poors
masato_hisakado@standardandpoors.com

キーワード: horse race, prediction, forecast, accuracy, scale invariance, logit model

概要

人と機械のどちらが賢いのか? 競馬市場参加者(以下、競馬ファンとする)は、競走馬の勝率について、オッズボードで情報を共有して最終オッズを形成する。オッズは得票率の逆数であり、「馬の勝率はその馬の得票率である」という合意が競馬ファンの間でなされるわけである。では、競馬ファン全体で形成したオッズの予想の精度はどの程度なのか? また、競馬ファンの予想にはどのような構造が隠されているのか? 本発表では、競馬ファンによる競馬予想と二つの確率モデルによる競馬予想との比較を行う。一つは過去の馬や騎手の実績をもとにスコアを算出し、勝率を計算するロジットモデル。一つはある競馬情報ベンダーが提供する過去の走破タイムと条件をもとに、現レースでの走破タイムを予想するニューラルネットのモデル(以下、Vモデル)。結論は、オッズの精度は本研究で用いたロジットモデルを10ポイント以上も上回り、機械より人のほうが賢い。得票率とロジットモデルによる馬のランキングの差異の解析から、得票率が1%を超える領域では、競馬ファンの予想とロジットモデルによる予想の微細構造は非常に類似し、両者の差異は精度のみである。一方、Vモデルの予想は、競馬ファンの予想やロジットモデルの予想とは微細構造が異なる。競馬ファンは、過去の実績を精査して馬の勝率を計算して予想している。また、オッズが100倍を越える万馬券の領域(得票率が1%以下)において、二つの確率モデルでは勝馬はランダムに分布し、馬の強さを正確に評価できていない。一方、競馬ファンは、そうした100回に1回も勝たないような馬であっても、ある程度正確に確率を評価し、馬の強さにもとづいたランキングを行っている。

1. 競馬予想の精度

オッズの精度を計るのに、ここでは Accuracy Ratio (以下、AR) を用いる [1]。ある事象が生起するかしないのかの予想を、生起確率や予想に用いた確率モデルでのスコアの大きなもの順に並べる。そして、確率(スコア、以下略)の大きいほうから、ある確率 p までの予測に対し、生起した事象の累積比率 $x_1(p)$ を y 軸に、生起しなかった事象の累積比率 $x_0(p)$ を x 軸にプロットした ROC(Receiver Operating Characteristic) カーブを描く。AR とは、ROC カーブの下側の面積 AUROC(Area Under ROC curve) を、ROC カーブが対角線に一致する場合はゼロ、予想が生起する事象と生起しない事象を確率軸上で完全に分離する場合、つまり ROC カーブが最初 y 軸に沿って $(0, 1)$ まで進み、次に x 軸にそって $(1, 1)$ まで進む場合に 1 になるよう規格化したものである。

$$AR = 2(AUROC - \frac{1}{2})$$

図1は、2000年以降のJRA主催35591レースに対して描いたROCカーブである。ARは67.6%である。この数値から、競馬ファンの予想精度の感じをつかむために、図2に週間天気予報のARの値をプロットした。 x 軸は、予報の発表時点から予報が対象とする日時までの時間。 y 軸にはARを示した。気象庁は、天気予報として、0時から6時、6時から12時などの6時間での降水確率を教える6時間予報と、1日24時間の降水確率を示す24時間予報の2種類を提供している。赤が24時間予報のAR、緑が6時間予報の結果である。朝5時発表の朝6時から12時の6時間予報が、予報発表時点と予報対象の日時の間が最短の1時間で、ARは約90%。午前11時発表の7日後の24時間予報だと、時間間隔は157時間となり、ARは40%強。予報対象の日時までの時間が長くなればなるほど、予報の精度が落ちていく。この結果から、競馬ファンの予想のAR=67.6%に対応する時間間隔は、概ね「明後日」であり、「明日の天気予報」に負けている。

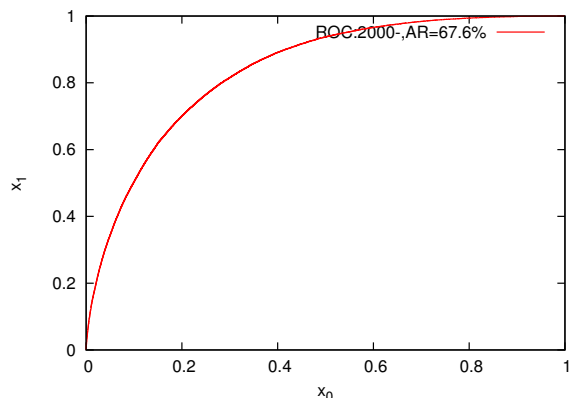


図1 2000年以降のJRA主催35591レース、出走頭数50万2342頭に対するROCカーブ。ARは67.6%

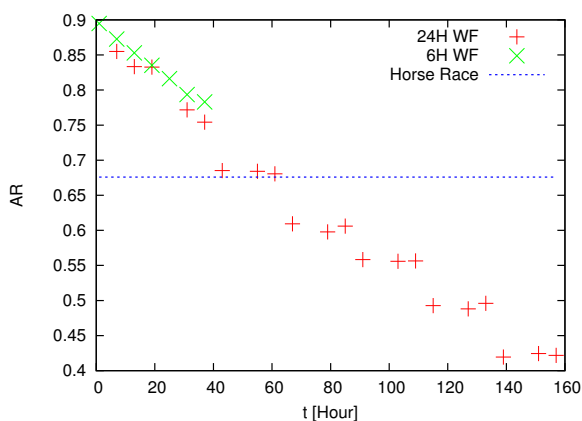


図2 天気予報発表時点から、その予報が対象とする期間の開始時点までの時間 t に対し、6時間予報、24時間予報のARをプロットしたもの。参考のため、競馬市場のARの67.6%を青の点線で示している。データは、2010年5月27日までの全国約70箇所約400日間の3万以上の気象庁発表の週間天気予報を用いた。

もっとも、ARの絶対値は予想対象に依存しない予想の精度の客観的な基準を与えるものではない。同じ対象に対しては、ARの大小で予想の優劣を比較できるが、異なる対象の場合は、比較できない。「競馬ファンの予想の精度は明後日雨が降るかどうかの天気予報と同じ」というのは正確な表現ではないが、競馬予想の難しさ、競馬ファンの予想精度のレベルについて、大体の感じはつかめる。

2. 競馬ファンの構成について

競馬ファンのオッズの予想精度はARで67.6%であることを述べた。では、この精度はどのようにして生まれてくるのだろうか。ここでは前回までのJWEINの発表(JWEIN08[2]、JWEIN09[3])を、一部訂正をしながら総括する。まず、競馬市場を経済物理的なアプローチで解析したとき、次の3つの特徴を持つことが分かった。

- 得票率のゆっくりとした収束 [4]
- ARのゆっくりとした収束 [4, 5]
- 万馬券の分布のスケール不変性 [6]

馬の得票率は投票の進行とともに変化する。そして、投票の終了した時点でそのレースに参加する競馬ファンの、出走馬の勝率に対するコンセンサスが成立したことになる。実際、得票率と勝率(ただし、ここでの勝率は、得票率がほぼ同じ馬を多数集め、その馬の中に含まれる勝馬の比率)の一致の度合はすばらしく「競馬市場はほとんど効率的」と言われる。

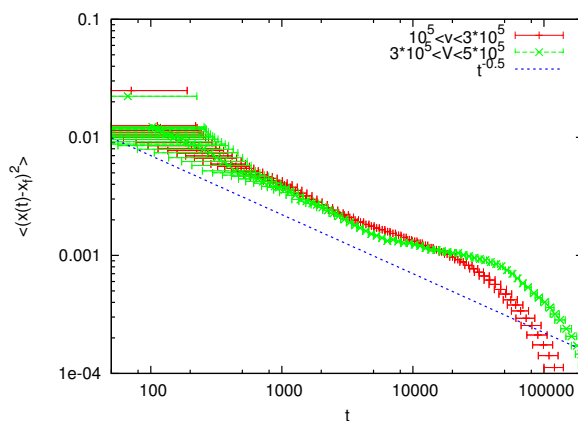


図3 投票数 t に対し得票率のゆらぎの自乗 $\langle (x(t) - x_f)^2 \rangle$ を全出走馬で平均したものをプロットした。総投票数が10万以上30万以下(赤)と30万以上50万以下(緑)を示している。

では、 t 票投票された時点での得票率 $x(t)$ は最終的な得票率 x_f にどのように収束するのか? 図3は、2008年のJRAの総票数が10万から30万、30万から50万のレースの得票率の最終得票率からの揺らぎの自乗を、レース出走馬で平均したものを t に対してプロットしたものである。投票の前半 ($t < 10^4$) において、両対数プロットは直線に乗り、べき乗則に従って収束していることが分かる。重要なのは、その指数が約0.5と小さいことである。もし、投票者が最終得票率 x_f と同じ確率で毎回投票した場合指数は1となり、急速に収束する。つまり、投票による得票率の収束は非常にゆっくりと起きることが分かる。JWEIN09[3]の発表では、得票率の分散(最終得票率からのゆらぎの自乗ではなく、出走馬全体での得票率の分散)がある時点まではゆっくりと減少し、それから増加に転じる点が理解できないとした。しかし、分散が投票の進行と共に増加するのは、投票と共に予想の精度が上昇(次に示す)することとほぼ同値である。何故なら予想精度の上昇とは勝ち馬の得票率の上昇と負け馬の得票率が減少を意味し、全体として得票率のパラツキが増加するからである。

一方、ARに関して同様なべき乗則を見ることが出来る。投票の最終結果を用いて計算したARの値を AR_f とし、投票数 t の時点でのARの値を $AR(t)$ とする。 $AR(t)$

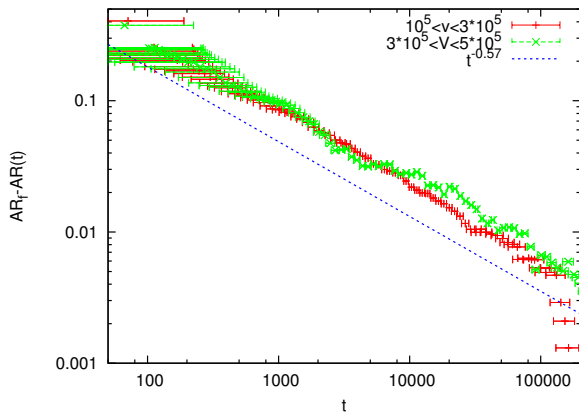


図 4 投票数 t に対し、 $AR_f - AR(t)$ の両対数プロット。総投票数が 10 万以上 30 万以下 (赤) と 30 万以上 50 万以下 (緑) を示している。

は t と共に増加し、投票回数が増えれば増えるほど、正しい情報がオッズに入ってくるのが分かる。図 4 は、 $AR_f - AR(t)$ を t に対して両対数プロットしたものである。プロットが直線に乗ることから、べき乗則に従ってゆっくりと投票の精度が上昇することが分かる。総投票数が 10 万から 30 万のデータでは最終的な AR の値は $AR_f = 68.3\%$ だが、投票の最初 10% (約 18000 票時点、出走 90 分前) で AR は 66.6% に達し、それ以降の 90% の投票では、1.7% しか精度は改善していない。また、得票率とは異なり、投票のほぼ最終段階までべき乗則に従っている。

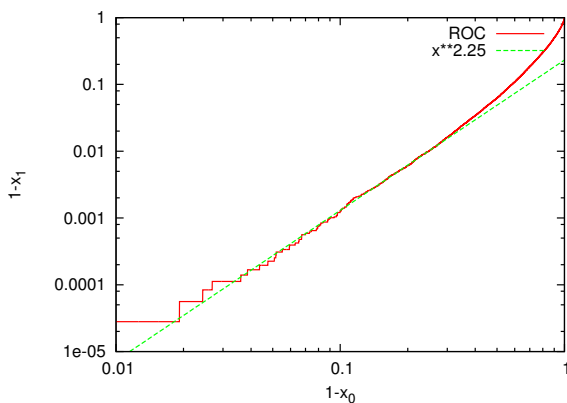


図 5 図 1 の ROC カーブを、右上の (1, 1) から両対数プロットで表示したもの。

第 3 の万馬券のスケール不変性は、得票率が 1% 未満の出走馬の中での勝ち馬の分布に関するものである。得票率が 1% 未満といっても、そうした馬は全体の約 26% を占める多数派であり、それらが勝ち馬に占める比率は約 1.3%。勝率に換算すると、1 レースに平均 14 頭出走することから 0.4% 未満である。この領域では、市場の効率性が成立せず、馬券は買われすぎであり、一攫千金を狙っ

た競馬ファンによるバイアス、Long-shot バイアスと呼ばれる [7]。万馬券候補の馬券は非常に多くても、実際の万馬券は非常にまれであることが分かる。もし、勝率が非常に小さな馬に対する競馬ファンの評価が不確かなものなら、勝ち馬は得票率軸でランダムに分布する。一方、勝率が小さいなりに、その中で馬の選別を行っているならば、得票率の小さい方から馬を並べたとき、勝ち馬の分布は一樣ではなくグラデーションを示すであろう。図 5 は、得票率が小さい方からの勝ち馬の累積比率、 $1 - x_1$ と負け馬の累積比率 $1 - x_0$ の両対数プロットである。図 1 の ROC カーブを右上の点 (1, 1) から、両対数プロットで拡大して見たものである。プロットが直線に乗っていることから、勝馬のグラデーションがべき乗則に従っていることが分かるが、重要なのは指数が 1 ではなく約 2 になっていることである。つまり、競馬ファンは、勝率が小さく、ランキングが難しいような馬であっても、強さを評価し、勝ち馬の分布に非自明なグラデーションパターンを描かせることに成功している。

では、こうしたべき乗則から何が分かるのか？ 我々は、馬の強さに関する情報を持ち込む「独立投票者」と、他人の投票結果をもとに、「人気のある馬=強い馬」という考えで投票する「ただ乗り投票者 (Herding Voter)」の 2 種類の投票者が存在する投票モデルを導入した [8]。勝率 w の馬に投票する独立投票者の比率は w に等しいとすると、得票率は多数回の投票の後、勝率に一致し、この意味で独立投票者は正しい情報を競馬市場に持ち込むことになる。この投票モデルに対し、得票率の最終得票率からのゆらぎの自乗が、べき乗則に従い、その指数が独立投票者の比率の 2 倍であること、また、数値的にはあるが、AR の収束もべき乗則に従うことを示した [4]。これらの結果から、競馬市場で、情報を持ち込む独立投票者の比率は約 4 人に 1 人であり、残りは他人の投票結果を参考に投票しているだけであることが分かる。ただし、比率の数字の解釈には注意が必要である。「独立投票者」対「ただ乗り投票者」の場合、前者の比率は 4 人に 1 人と決まるが、「裁定投票者」対「ただ乗り投票者」の場合、比率は決まらない。ここでいう裁定投票者とは、得票率と勝率の差に着目して投票確率が差に線形に依存して変化する投票者のことである。得票率と勝率が等しい場合は、勝率に等しい確率で投票する。差が正なら、その馬は買われすぎなので投票確率を勝率より減らす。差が負なら、勝率が得票率に勝るので、期待利得は、他の馬よりも相対的に高くなり、投票確率を勝率より増やす。この場合、系を記述するマスター方程式は、独立投票者のモデルと同じ構造となり、二つのモデルはパラメータの読み替えで対応がつく [4]。裁定投票者のモデルでは、投票確率を得票率により変化させるときの係数の値が余分なパラメータとして入り、結果として独立投票者の比率もそのパラメータに依存してしまう。独立投票者の比率がべき乗則の指数から分かっても、裁定投票者の比率

はパラメータの値に依存してしまう。

第3の勝ち馬の示すグラデーションのスケール不変性については、独立投票者が存在せず、ただ乗り投票者が馬の強さについて「わずかな情報」を最初に持っているとする投票モデルを解析した。その「わずかな情報」をもとに投票が行われることにより勝ち馬と負け馬がうまい具合に混ざり、スケール不変性をもつことを示した。得票率が1%未満の万馬券領域の馬の世界では、独立投票者の存在は1%未満であり、ほとんど無視できる。「わずかな情報」とは何なのか、それはよく分からないが、何らかの情報がインプットされないと、勝ち馬は得票率軸上ランダムに分布してしまう。競馬ファンは、競馬情報を熱心に研究し、その研鑽の成果が「わずかな情報」となり、勝ち馬の分布に非自明なグラデーションをもたらしているのだろう。

まとめると、情報を提供する独立投票者や裁定投票者の比率は、投票の初期段階では一定（独立投票者ならほぼ4人に1人）で、それが得票率やARの収束のべき乗則を生む。万馬券領域での勝馬と負け馬の混合のグラデーションが示すスケール不変性は、ただ乗り投票者が、独立投票者が提供するほんのわずかの情報をもとに投票を行い、馬を混ぜ合わせることで生まれる。

では、競馬ファンの予想能力はどの程度なのか？また、「情報」とはどのようなものなのか。以下のセクションでは、確率モデルで競馬予想を行った場合との比較を行い、競馬ファンの集団知の起源にせまる。

3. 確率モデルによる競馬予想

確率モデルでの競馬予想を取り上げ、その予想精度を検証する。一つは、以下に詳しく述べるロジットモデル。もう一つは、ニューラルネットを用いた走破タイム予想を行うモデルでVモデルと呼ぶ。このモデルはレースのクラスごとに別のモデルで構成されたものの総称である。

競馬とは、レースに出走した馬で勝ち馬を選択する過程である。その際、馬はさまざまな属性をもち、その属性に従ってレースが行われ、最も早く走った馬が勝ち馬となる。今、 H 頭の馬が出走するとし、そのうちの h 番目の馬の勝率 p_h を考える。馬は、 K 種類の観測可能な属性（クラス、スピード、成績など）を持ち、 $\vec{x}_h = (x_{h,1}, x_{h,2}, \dots, x_{h,K})$ で表すとする。その馬に騎乗する騎手もまた M 種類の観測可能な属性を持ち、 $\vec{y}_h = (y_{h,1}, y_{h,2}, \dots, y_{h,M})$ で表す。競馬の確率モデルとは、馬の勝率をこれらの属性の関数として表したものである。

$$p_h = p(\vec{x}_h, \vec{y}_h | \{\vec{x}_h, \vec{y}_h\}_{h=1, \dots, H})$$

Bolton and Chapman は、この関数形に次のものを用いた [9]。

$$p_h = \frac{1}{Z} e^{V_h} \quad (1)$$

$$Z = \sum_{h=1}^H e^{V_h} \quad (2)$$

$$V_h = \sum_{i=1}^K \theta_i^H x_i + \sum_{j=1}^M \theta_j^J y_j \quad (3)$$

p_h の式で分母に現れている Z は、出走する馬での確率の和が1になるための規格化因子である。また、指数関数の肩に乗っている関数 V_h は、馬の価値（効用）を表す。そして、レース数 R のデータに対し、各レース $r \in \{1, 2, \dots, R\}$ での勝ち馬の馬番号を h_r^* として、次の尤度関数を定義した。

$$\exp(L) = \prod_{r=1}^R p_{h_r^*}$$

関数 L は対数尤度関数であり、最尤法（ L を最大化）により馬の価値を表す関数 V_h の中の係数 θ_i^H, θ_j^J を決定する。

ここでは、尤度ではなくARを最大にするように係数を決定する。尤度最大とAR最大で得られる係数や確率の値は多少異なるが、今回の問題ではそれほど大きな差はない。我々が用いた馬の属性は全部で22種類、騎手の属性は4種類の計26種類である。その他、最終得票率のロジット、Vモデルの着順予想の数値 o を $(H-o)/(H-1)$ で1（1着）からゼロ（最後尾）に変換したものを、予想走破タイムをそのレースで走る馬の平均予想走破タイムから引いたもの、などの3種類の属性も後の議論のために用いた。以上の29種類の属性データを表にまとめている。

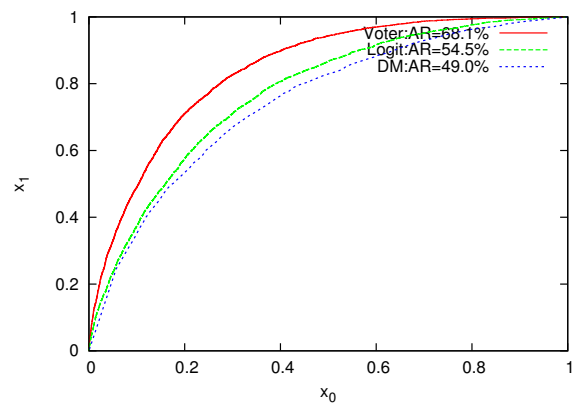


図6 20000レース以降の15592レースに対するROCカーブ。競馬ファンの予想 (Voter:赤)、ロジットモデル (Logit:緑)、Vモデルの予想着順のスコアを用いたモデル (DM:青)のROCカーブとARの値を示した。

係数を決定するために、2000年以降の35591レースのうち、10000レースから19999レースの1万レースをAR最大化のための学習データに用いた。また、ロジットモデルの評価には、20000レース以降の15592レースの結果を用いた。26種類の属性のうち、係数が大きく、AR

表 1 ロジットモデルに用いた属性一覧。ロジットモデルに用いた 26 種類の属性に限定して AR を最大化した結果。得票率はロジットに変換後、標準化（平均ゼロ、分散 1 に）。データは、予想するレースから過去 5000 レースのデータをもとに、馬や騎手の実績データを計算。また、直近データは、馬については前 3 走または過去 4 から 6 走の 3 レースでの結果、騎手については過去 100 レース（騎乗回数ではない）での結果をもとに計算した。ポスト位置、V モデルの着順予想は 0 から 1 にスコア化（コース 1 が 1、最外側のコースがゼロ、1 着が 1 で最終着がゼロ）。予想走破タイムは各レースでの平均走破タイムを引き算してスコア化。距離変化、馬体重変化、負担重量変化など、過去に出走せずデータがない馬については全てゼロとする。データは、最適化に用いるレース全体のデータを用いて、標準化。ただし、平均獲得賞金、平均走破タイム、ベスト走破タイムについては、各レース毎で標準化ののち、全レースデータで再度標準化している。ポスト位置*馬場状況は、馬場状況の数値（1 = 良馬場から 4 = 悪）を 2.5 から引いたもの掛けたもの。内コースは馬場がよい場合には有利になることを取り入れたもの。3 列目に、AR を最大化したときの係数の値、4 列目は、検証に用いたデータで各係数をゼロ（他の係数はそのまま）としたときの、AR の変化を示している。

変数	係数	ΔAR
1 着率	-0.01	0%
2 着率	0.07	0.2%
3 着率	0.01	0%
4 着率	0.03	0%
5 着率	0.01	0%
6 着率	0.09	0.3%
7 着率	0.08	0.1%
前 3 走の 1、2、3 着率	0.29	3.1%
前 3 走の 4、5 着率	0.15	0.9%
前 4、5、6 走での 1、2、3 着率	0.05	0%
平均獲得賞金	0.26	2.2%
ポスト位置スコア	0.01	0%
馬体重変化/平均馬体重	0.04	0%
負担重量変化/平均負担重量	0.01	0%
距離変化/平均距離	-0.18	0.8%
平均走破タイム (/Km)	-0.14	0.3%
ベストタイム (/Km)	-0.04	0%
ベスト距離からの変化/平均距離	-0.02	0%
出走回数	-0.01	0%
ペア回数	0.01	0%
ポスト位置スコア*馬場状況	0.01	0%
騎手とのペア率	0.01	0%
騎手の勝率	0.17	0.7%
騎手の 2、3 着率	0.10	0.2%
騎手の 4、5 着率	0.04	0%
騎手の最近の 1、2、3 着率	0.02	0%
予想順位（スコア化）	0	NA
予想走破タイム（スコア化）	0	NA
得票率のロジット	0	NA

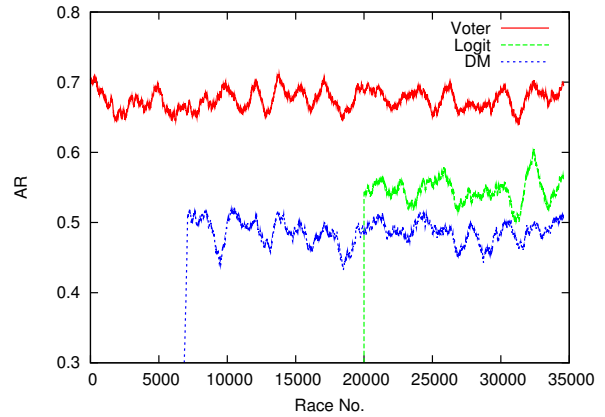


図 7 前述の 3 つのモデルでの AR の変化。ロジットモデルでは、検証用のサンプルである 20000 レース以降について示してある。

の上昇に貢献しているもの (ΔAR が 0.3%以上) は、前 3 走の成績、平均獲得賞金、距離変化、平均走破タイム、騎手の勝率、6 着率の 7 種類である。6 着率の寄与は不思議だが、これは前 3 走の結果が 5 着内の実績までを考慮しているためであると考えられる。また、AR の値は、54.5%となった。一方、競馬ファンの予想、つまり得票率を用いて勝ち馬を予想した場合の AR は 68.1%、V モデルの予想着順を用いて予想した場合、AR は 49.0%となった。(図 6 を参照のこと。) 図 7 には、100 レース刻みで 1000 レース分のデータを用いて計算した AR の変化を示した。

図 8 には、万馬券領域での勝ち馬の分布を調べるため、 $(1 - x_0, 1 - x_1)$ の両対数プロットを示している。 $1 - x_0$ が 0.28 以下の万馬券領域に着目すると、競馬ファンの予想では、図 5 と同様にスロープが約 2 であり、勝ち馬の分布は単にランダムではなく、馬の強さもある程度反映した分布、グラデーションを示す。V モデルの予想着順を用いた予想では、万馬券領域に対応する部分での勝ち馬の分布が見えない。その理由は、予想着順では、馬のランク付けが荒すぎるため、最下位のランクにすでに 10%近い馬がいるからである。そこで、予想走破タイムのみをロジットモデルのスコアに用いて AR を最大にするパラメータを決定し (V モデルの AR とほぼ同じ値が得られる)、ROC カーブの両対数プロットを描いている。両対数プロットは傾きがほぼ 1 の直線に乗る。つまり、V モデルでは、万馬券領域での馬の識別能力はなく、諦めてしまっていることが分かる。ロジットモデルも同様に、傾きはほぼ 1 であり、万馬券領域で馬の勝率を微細に評価することに成功していない。

以上の結果をまとめると、ロジットモデルによる競馬予想、ニューラルネットによる走破タイムの予想に基づく競馬予想はどちらも競馬ファンの予想精度に AR で 10%以上もの差で負けている。ロジットモデルに V モデルの予想着順のスコアを追加し AR を最大化すると 57%まで精

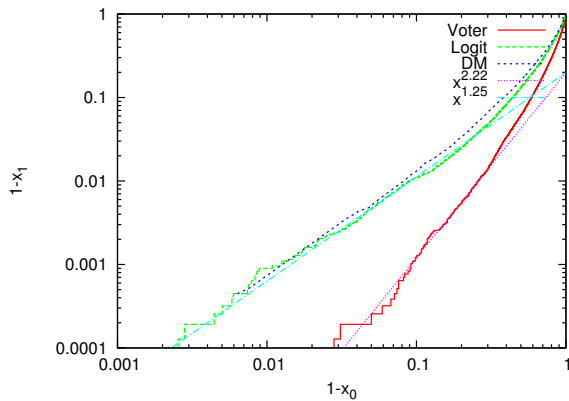


図 8 前述の 3 つのモデルに対する ROC カーブ (図 6) に対し、 $(1-x_0, 1-x_1)$ を両対数プロットしたもの。競馬ファンの予想では、スロープは約 2 だが、V モデル、ロジットモデルでは約 1 となる。V モデルでは、万馬券領域の勝ち馬の分布が分からないため、予想走破タイムをロジットモデルのスコアに用いた場合の ROC カーブを代わりにプロットしている。

度を上げることは可能だが、それでも競馬ファンの予想精度とは 10% も劣る。つまり、現時点で機械学習は、競馬ファンの予想に、天気予報で 1 日分負けている。競馬ファンの予想精度が天気予報で明後日雨が降るかどうか分かる程度なら、機械学習による予想は明々後日の天気予報のレベルということになる。もっとも、この事実だけで機械が人間に劣るという証明にはならない。さらによいスコアを発見して馬の効用関数にインプットすることで AR を高めることは可能である。しかし、競馬ファンの予想に関して驚くべき点は、万馬券領域での勝ち馬の分布に関するものである。勝率が 1% に満たない非常にまれな現象でも、競馬ファンはそれを評価し、一様ではなくグラデーションを勝ち馬の分布を導くことに成功している。一方、ロジットモデルや V モデルでは、AR を最大化するように最適化しているためでもあるが、テール部分では勝ち馬は一様に分布し、馬の勝率の評価を諦めてしまっていることが分かる。競馬ファンは万馬券を獲得するために、ありとあらゆる情報を探し、万馬券領域での馬の勝率の評価にある程度成功している。もちろん、この動機は「万馬券を獲得すること」につきものだろうが、その情報収集能力、予想精度はスゴイという他はない。

4. モデルの微細構造の比較

ここまでは競馬ファンの予想 (オッズ)、ロジットモデル、V モデルのそれぞれのモデルを単独で分析してきた。この章ではそれぞれのモデルが他のモデルと比較してどのような関係にあるのかを探っていく。これまで検証してきた AR は基本的にはバルクの序列の比較であり、また万馬券領域の勝ち馬の分布は、テール部分での序列の比較であった。この章ではさらに細部の序列化につい

て見てみる。例えばモデル A とモデル B を比較する場合どれだけ付加的な序列化の情報をプラス出来ているかを考える。もしモデル B がモデル A と別の視点で付加的な情報を加えられるのであれば、たとえモデル A の AR がモデル B の AR より高くても、モデル B はモデル A とは異なった価値のあるモデルであると考えられる。一方、付加的な情報が与えられないのであればそれはモデル自体が包含関係にある。

延べ 3000 頭の競走馬のレース・データを用いて A モデルと B モデルで比較を行う。まず A モデルで対象馬をすべてを序列化した後、それを 300 頭ずつに 10 等分する (行)。その 10 等分された各行の中の 300 頭を B モデルによってさらに序列化し、30 頭ずつに 10 等分する (列)。すると 100 のます目が出来上がり、各ます目の中に 30 頭が区分されることになる。これは A モデルで荒く序列化し、細部を B モデルで序列化することに相当する。ここで、ます目ごとに勝ち馬頭数をカウントし、それぞれの列について勝ち馬頭数の和をとったものをグラフにすると、グラフの傾きが急であるほど A モデルの序列化に対し B モデルが情報を付加していると考えられる。つまり異なる視点で馬を評価できるということである。

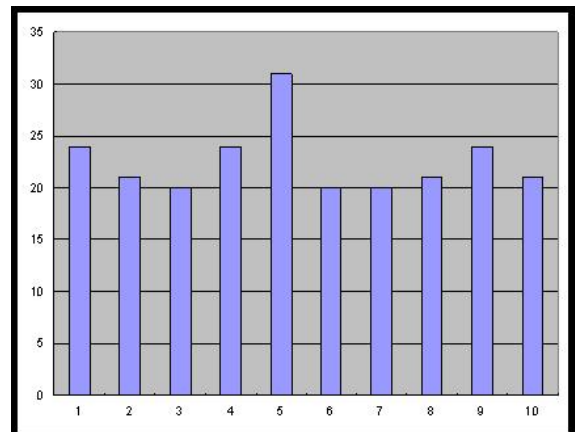


図 9 オッズで序列化したのち細部をロジットモデルで序列化したものの。

図 9、図 10 の 2 つのグラフは上記の分析手法を“オッズにロジットモデルと V モデルが付加的な情報を加えているか?”について適用した結果である。図 9 ではほとんど傾いていない。これはロジットモデルは細部の序列がオッズほどできていないことを意味する。このことからロジットモデルとオッズは包含関係にあることが分かる。一方、図 10 は V モデルがオッズで荒く序列化したものに付加的な情報を加えていることを示す。図 11 は参考のため細部もオッズで序列化したものである。これらの傾きはほぼ同様であることから、V モデルの細部の序列化はオッズ並であることが分かるがオッズは V モデルよりも AR が高いことから V モデルはバルクの精度が弱く、細部の精度はオッズ並であることが分かる。総合す

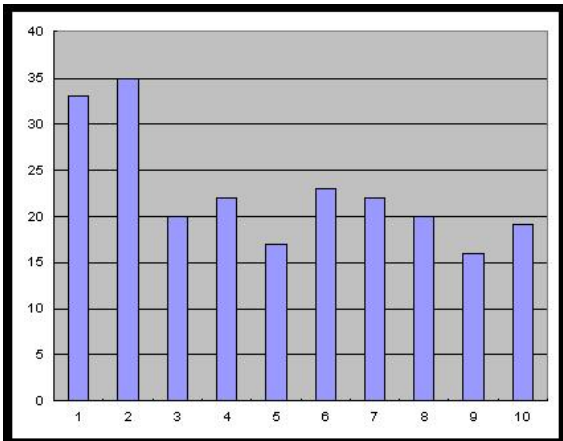


図 10 オッズで序列化したのち細部をVモデルで序列化したもの

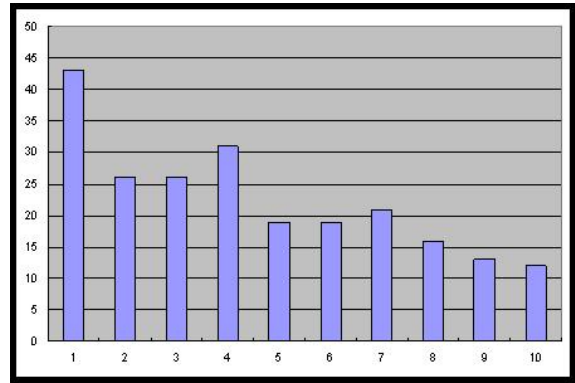


図 13 ロジットモデルで序列化したのち細部をVモデルで序列化したもの

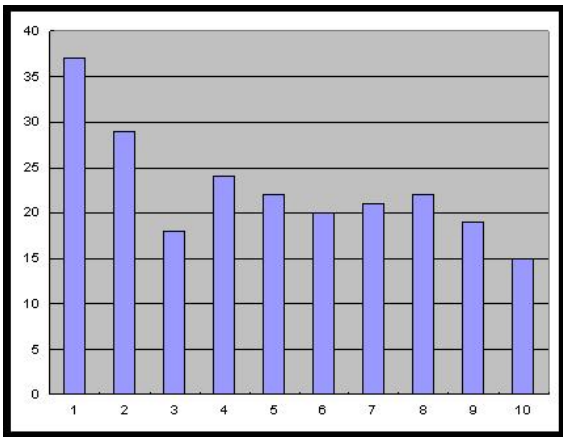


図 11 細部もオッズで序列化したもの

るとロジットモデルはバルク部分はオッズに比べ、やや劣るもののそれなりの精度はあるが細部の序列が弱い。一方Vモデルはバルク部分は弱い但細部の序列化精度はある、とすることが分かる。

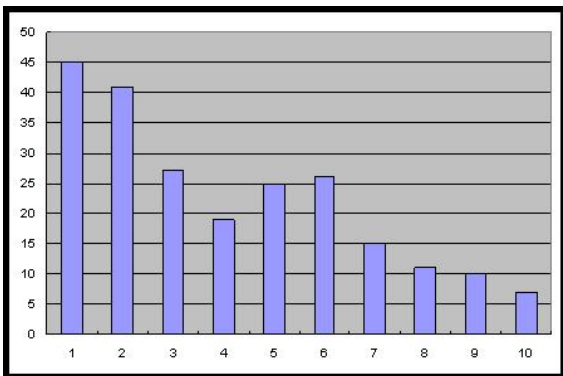


図 12 Vモデルで序列化したのち細部をロジットモデルで序列化したもの

以下の図 12 と図 13 の 2 つのグラフは上記の分析手法を“ロジットモデルとVモデルは互いに付加的な情報を

加えているか？”について適用した結果である。図 12 が右下がりになっていることから、ロジットモデルはVモデルに付加的な情報を加えていることが分かる。同様に図 13 ではVモデルはロジットモデルに付加的な情報を加えていることが分かる。これらからロジットモデルとVモデルは互いに相補的な関係あることが分かる。つまり視点が異なっているのである。このようなモデルは併用することによりパフォーマンスを上げることが可能である。実際、ロジットモデルのスコアにVモデルの予想着順スコアを追加したモデルでは、ARは約2%上昇した。

5. ま と め

人と機械のどちらが賢いのか？今回の発表では、競馬予想を対象にして、競馬ファンの予想、機械学習による予想の精度の比較を行った。我々の用いたモデルとの比較では、人は機械より賢い。ARは10%以上高く、また万馬券領域での勝ち馬のグラデーションは、勝率が1%を切るような希な現象であっても、競馬ファンがある程度正確に馬の強さを評価し、識別していることを示している。もちろん、その背後には「万馬券を獲る」という強い意志があるのだろうが、その情報収集能力、勝率の計算能力は驚くべきものである。もちろん、今回用いたモデルで、有効なスコアを導入できればARを上げることは可能だろう。とくに、機械学習では、過去の実績を基に勝率を計算するため、新馬戦などのレースは不得意である。そうしたレースは除外する。またレースの種別毎にモデルを分ける、スコアの線形和ではなく、指数関数項を追加する努力をすればARを数%は上げられるだろうが、それでも10%の差は大きい。また、テール部分についても、どうモデルを最適化すれば勝ち馬のグラデーションをつけられるのか謎である。AR最大化、尤度最大化などでは、順当に強い馬を強いと評価し、バルク部分で精度を上げようとするため、どうしてもテールでは精度が失われてしまう。万馬券狙いに絞った予想モデルとは、どのような手法で最適化するのか。オッズに比

例した重みをつけて、それを AR や尤度に導入する等、さまざまな工夫が考えられる。こうした努力の果てに、いつか機械が人をを超えることができるのか。先は長そうである。

謝 辞

本研究は科研費 21654054 (挑戦的萌芽研究) の助成を受けました。

◇ 参 考 文 献 ◇

- [1] B.Enleman, E.Hayden and D.Tasche, *Testing rating accuracy*, WWW.RISK.NET (2003).
- [2] 守、久門、「多数決とスケール不変性」,JWEIN08 講演論文集,p106.
- [3] 守、久門、「多数決と相転移」,JWEIN09 講演論文集.
- [4] S.Mori and M.Hisakado, Component Ratios of Independent and Herding Betters in a Racetrack Betting Market, preprint arXiv:1006.4884.
- [5] S.Mori and M. Hisakado: *Emergence of Scale Invariance and Efficiency in Racetrack Betting Market*, Proc. of the 9th Asia-Pacific Complex Systems Conference Complex 09, (Mitsugu Matsushita, Yuji Aruka, Akira Namatame, and Hiroshi Sato eds)pp 258-267.
- [6] S.Mori and M.Hisakado, Exact Scale Invariance in the Mixing of Binary Candidates in Voting Model, *J.Phys.Soc.Jpn*,**79**(2010)034001.
- [7] D. B. Hausch, V. SY. Lo and W. T. Ziemba, *Efficiency of Racetrack Betting Market, 2008 Edition* World Scientific
- [8] M. Hisakado and S.Mori, Phase Transition and Information Cascade in a Voting Model, *J.Phys.A,Math.Theor.* 43 (2010) 315207.
- [9] R.N.Bolton and R.G.Chapman, *Searching for Positive Returns at the Track : A Multinomial Logit Model for Handicapping Horse Races*, *Management Science* 32,1040-1059.