Vol.2017-GI-37 No.1
2017/3/7

プレイヤー間の非対称条件を考慮したレーティングシステム

大竹 勇志[1,a)]　金井 崇[1,b)]

概要：レーティングシステムとは，対人競技においてプレイヤーの強さを定量的に評価する仕組みであり，ELO rating system [1] に端を発する．試合結果に応じてプレイヤーの強さを数値で表すことで，選手の番付や試合の組合せの決定ができる，ルールのバランス調整の目安となる，などのメリットがある．しかし今日までに提案されたレーティングシステムには，選手の体格や試合条件など，選手の技量と独立して存在し，かつ試合結果を左右するような要素の影響を除き，選手の技量を単独で評価することができないという欠点がある．
本研究では，Bradley-Terry モデルを拡張したモデルを複数提案し，プレイヤーの技量にあたる要素とそれ以外の要素による影響を分離して，それぞれを定量的に評価することを目指す．そしてそれらの拡張モデルを 2014 年の ATP プロテニス公式戦の試合結果に適用し，通常の評価との差異を確認する．

キーワード：Bradley–Terry モデル，レーティング，ランキング，最尤推定法，スポーツ，テニス

Yushi Otake[1,a)]　Kanai Takashi[1,b)]

1. はじめに

本研究の目的は，プレイヤーの技術以外の事項が勝敗に影響を与える可能性がある対人競技に対し，プレイヤーの技量と，それと独立に存在し勝敗を左右する要素の影響の大きさを，それぞれ定量的に評価できるようなレーティングシステムを提案することである．

選手の技量と独立して存在し，対人競技の試合結果に影響を与えるようなパラメータとしては，例えば選手の体格が挙げられる．ボクシングや柔道のような格闘技では体重が大きい選手の方が有利とされるため，体重別階級制が導入されている．これは，選手の技量と独立して存在し，試合に影響を与えるパラメータを除外することで，試合結果に選手の技量，実力を色濃く反映させるための試みであると見做すことができる．本研究で提案するようなレーティングシステムを用いれば，体格などの影響の大きさの評価や，それらを差し引いたときの選手の技量のみの評価が可能になることが期待される．

本研究では例として 2014 年の ATP プロテニス公式戦の試合結果を使用し，選手の身長の影響を示すパラメータと，その影響を差し引いた際の選手の技量を示すパラメータ双方を推定し，結果を検討する．

2. 関連研究

対人競技における過去の戦績を数理的に分析し，プレイヤーの実力を定量的に評価する「レーティング」に関する先行研究はいくつか存在する．実力評価の研究はチェ

1 東京大学大学院情報学環
The University of Tokyo Graduate School of Interdisciplinary Information Studies
a) yushi.otake2008115@gmail.com
b) kanai@graco.c.u-tokyo.ac.jp

スを題材とした ELO rating system [1] に端を発する．続いて，プレイヤーのパフォーマンスのばらつきを考慮した評価手法が現れ，例えばロジスティック分布を利用した Glicko ranking system [2] や，ベイズ推定を利用した TrueSkill(TM) [3] などが存在する．これらの評価手法では，各プレイヤーに「レート」と呼ばれるプレイヤーの実力を表す数値が与えられ，試合を行うごとに特定の更新式により両方の選手のレートが更新される．試合に勝つとレートが上昇し，負けるとレートが減少するが，強いプレイヤーに勝つほどレートの上昇幅が大きくなり，弱いプレイヤーに負けるほどレートの減少幅が大きくなる．

これらの評価手法の問題点として，たとえプレイヤーのパフォーマンスに関してある程度の不確実性を考慮していても，試合の結果はプレイヤーの実力に依るものであると仮定しているという点が挙げられる．よってプレイヤーの技術と関係がなく，しかも勝敗に影響を与える要素が存在する場合に，これらのレーティングシステムを適用すると正しい評価ができない．このような要素が存在する競技では，勝敗のみからプレイヤーの技量を抽出して定量的に評価することは困難である．もしその要素を予め除くことができればプレイヤーの技量を評価することができるが，要素の排除が難しい競技も存在し，この種の競技におけるプレイヤーの技術や，及び有利不利を生む要素による影響を評価する手法は未だ定まっていない．本研究では Bradley–Terry モデルをベースとした拡張モデルを用いることで，プレイヤーの技量にあたる要素と，それ以外の要素による影響を分離して，それぞれを定量的に評価することを目指す．

3. Bradley–Terry モデルの拡張

3.1 Bradley–Terry モデルの導出

Bradley–Terry モデル [4] は本来，官能検査における一対比較法のデータを解析するために提案されたものであるが，対人競技におけるプレイヤーの強さの推定にも用いることが可能である [5].

全 m 人のプレイヤーが 1 対 1 の形式で互いに何回か試合を行い，その結果が既知であるとする．ただし，個々の試合では引き分けはなく，三すくみの関係が発生するような得意及び苦手関係はないものとする．このとき，Bradley–Terry モデル (以下 BT モデル) によって各プレイヤーの強さのパラメータを導くことを考える．

p_{ij} をプレイヤー i がプレイヤー j に勝つ確率とし，プレイヤー i が j に勝つことを $i \to j$ と表すことにする．プレイヤー i, j が直接試合を行わないとき，i, j の勝敗を別のプレイヤー k との試合結果によって決めるものとする．すなわち，

(1) $i \to k$ かつ $k \to j$ のとき，i は j に勝つと考える．
(2) $j \to k$ かつ $k \to i$ のとき，j は i に勝つと考える．
(3) それ以外の場合は引き分けとし，勝敗がつくまで試合を繰り返す．

このとき，プレイヤー i が j に勝つ確率と負ける確率の比は $p_{ik}p_{kj} : p_{jk}p_{ki}$ である．プレイヤー間に得意及び苦手の関係が存在しないので，結局，プレイヤー i が j と直接対戦して勝つ確率は

$$p_{ij} = \frac{p_{ik}p_{kj}}{p_{ik}p_{kj} + p_{jk}p_{ki}} \tag{1}$$

と表わせる．すなわち

$$\frac{p_{ij}}{p_{ji}} = \frac{p_{ik}p_{kj}}{p_{jk}p_{ki}} \tag{2}$$

となる．

全ての i, j, k について式 (2) が成り立つものとし，

$$\pi_i = \frac{p_{im}}{p_{mi}} \quad (i = 1, ..., m-1), \tag{3}$$

$$\pi_m = 1 \tag{4}$$

とすると，全ての $i, j\ (i \neq j)$ に対し

$$\frac{p_{ij}}{p_{ji}} = \frac{p_{im}p_{mj}}{p_{jm}p_{mi}} = \frac{\pi_i}{\pi_j}$$

すなわち

$$\begin{aligned} \frac{p_{ij}}{p_{ji}} &= \frac{\pi_i}{\pi_j} \quad (i \neq j) \\ \Longleftrightarrow p_{ij} &= \frac{\pi_i}{\pi_i + \pi_j} \quad (i \neq j) \end{aligned} \tag{5}$$

が成り立つ．BT モデルにおいては，π_i をプレイヤー i の強さを表すパラメータと考える．

3.2 拡張モデル 1 ： 一次元のパラメータを用いた Bradley–Terry モデルの拡張

身長や体重のような，一次元のパラメータが勝敗に影響を与えるかもしれない競技に対し，BT モデルを適用することを考える．前述の BT モデルをそのまま適用すると，プレイヤーの強さを表すパラメータ π_i にプレイヤー i の実力と，プレイヤー i の試合条件双方の影響が含まれてしまうため，π_i はプレイヤー i の実力や技術を示すパラメータとしては不適切となってしまう．

そこで，次のように BT モデルを拡張する．プレイヤー i に対応する一次元的なパラメータ d_i を考える．d_i は例えば，プレイヤー i の身長や体重など，競技の勝敗に影響を与えるかもしれない一次元的な要素の影響を表す．通常の BT モデル導出時と同様に，プレイヤー i, j が直接試合を行わないとき，別のプレイヤー k との勝敗によって i, j の勝敗を決するとする．拡張前の BT モデルと同じく，個々の試合では引き分けはなく，三すくみの関係が発生するような得意及び苦手関係はないものとすると，通常の BT モデルと同様にして式 (2)

$$\frac{p_{ij}}{p_{ji}} = \frac{p_{ik}p_{kj}}{p_{jk}p_{ki}}$$

が成り立つ．

全ての i, j, k について上記の関係が成り立つものとし，

$$\pi_i \cdot d_i = \frac{p_{im}}{p_{mi}} \quad (i = 1, ..., m-1), \tag{6}$$

$$\pi_m = 1 \tag{7}$$

とする．d_i はプレイヤー i の身長や体重など，プレイヤー毎に個別に対応し，試合に影響を与える可能性のある一次元的なパラメータを示す．

このとき，式 (6) より，全ての i, j $(i \neq j)$ に対し

$$\frac{p_{ij}}{p_{ji}} = \frac{p_{im}p_{mj}}{p_{jm}p_{mi}} = \frac{\frac{p_{im}}{p_{mi}}}{\frac{p_{jm}}{p_{mj}}} = \frac{\pi_i \cdot d_i}{\pi_j \cdot d_j}. \tag{8}$$

すなわち

$$\frac{p_{ij}}{p_{ji}} = \frac{\pi_i \cdot d_i}{\pi_j \cdot d_j} \quad (i \neq j)$$

$$\Longleftrightarrow p_{ij} = \frac{\pi_i \cdot d_i}{\pi_i \cdot d_i + \pi_j \cdot d_j} \quad (i \neq j) \tag{9}$$

が成り立つ．この拡張 BT モデルにおいては，π_i はプレイヤー i の技量を表すパラメータ，d_i は π_i に対応する補正項であり，$\pi_i \cdot d_i$ は最終的なプレイヤー i の強さを示す．

プレイヤー i に対して補正が不要である場合，補正項の影響をなくす必要がある．よって

$$d_i = 1 \quad (\pi_i \text{ を補正する必要がないとき}) \tag{10}$$

が成り立つ．

4. プレイヤーの強さ及び条件の影響力の推定

4.1 拡張モデル 1

m 人のプレイヤーが互いに何戦か試合を行ったとき，その結果から全ての π_i もしくは全ての d_i を推定することを考える．ただし π_i と d_i は必ず積の形 $\pi_i \cdot d_i$ の形で出現し，例えば片方を 2 倍にするともう片方を 1/2 倍することでプレイヤー i の強さについて同じ値を対応させることができるため，同時に推定することはできないことに注意する．

プレイヤー i がプレイヤー j と n_{ij} 回だけ試合を行ったとする．このとき $n_{ij} = n_{ji}$ である．プレイヤー i のプレイヤー j に対する勝数を確率変数

$$X_{ij} \quad (i \neq j; i, j = 1, ..., m) \tag{11}$$

で表す．個々の試合では引き分けはないものとしたので

$$X_{ij} + X_{ji} = n_{ij} = n_{ji} \quad (i \neq j) \tag{12}$$

となる．

ここで，X_{ij} が 2 項分布 $B(n_{ij}, p_{ij})$ に従うものとすれば

$$P\{X_{ij} = x_{ij}\} = \frac{n_{ij}!}{x_{ij}!x_{ji}!} p_{ij}^{x_{ij}} p_{ji}^{x_{ji}}, \quad (x_{ij} = 0, 1, ..., n_{ij}) \tag{13}$$

となる．したがって，それらの同時確率は

$$P\{X_{ij} = x_{ij}; i \neq j; i, j = 1, ..., m\} = \prod_{i<j} \left(\frac{n_{ij}!}{x_{ij}!x_{ji}!} p_{ij}^{x_{ij}} p_{ji}^{x_{ji}} \right). \tag{14}$$

ここで，式 (9) を代入すると，

$$\begin{aligned} &P\{X_{ij} = x_{ij}; i \neq j; i, j = 1, ..., m\} \\ &= \prod_{i<j} \left(\frac{n_{ij}!}{x_{ij}!x_{ji}!} \frac{(\pi_i \cdot d_i)^{x_{ij}} (\pi_j \cdot d_j)^{x_{ji}}}{(\pi_i \cdot d_i + \pi_j \cdot d_j)^{n_{ij}}} \right) \\ &= \prod_{i<j} \left(\frac{n_{ij}!}{x_{ij}!x_{ji}!} \frac{1}{(\pi_i \cdot d_i + \pi_j \cdot d_j)^{n_{ij}}} \right) \\ &\quad \cdot \prod_{i<j} \pi_i^{x_{ij}} \pi_j^{x_{ji}} \cdot \prod_{i<j} d_i^{x_{ij}} d_j^{x_{ji}} \\ &= \prod_{i<j} \left(\frac{n_{ij}!}{x_{ij}!x_{ji}!} \frac{1}{(\pi_i \cdot d_i + \pi_j \cdot d_j)^{n_{ij}}} \right) \\ &\quad \cdot \prod_{i=1}^{m} (\pi_i d_i)^{t_i} \end{aligned} \tag{15}$$

となる．ただし，

$$t_i = \sum_{j \neq i} x_{ij} \tag{16}$$

とおいた．

ここで，確率変数

$$T_i = \sum_{j \neq i} X_{ij} \tag{17}$$

を導入すると，この場合の尤度 L は

$$L = \text{const.} \times \prod_{i<j} (\pi_i \cdot d_i + \pi_j \cdot d_j)^{-n_{ij}} \cdot \prod_{i=1}^{m} (\pi_i d_i)^{T_i} \tag{18}$$

となるので，対数尤度 l は

$$\begin{aligned} l &= \log L \\ &= \sum_{i=1}^{m} T_i (\log \pi_i + \log d_i) \\ &\quad - \sum_{i<j} n_{ij} \log(\pi_i \cdot d_i + \pi_j \cdot d_j) + \text{const.} \end{aligned} \tag{19}$$

となる．

4.2 $\boldsymbol{\pi}, \boldsymbol{d}$ の最尤推定

前述の通り，π と d は必ず積の形 $\pi \cdot d$ の形で出現するため，両者を同時に推定することはできない．よって片方ずつ推定する．

4.2.1 $\boldsymbol{d}$ が既知であるとき

拡張モデル 1 の場合，対数尤度は

$$l = \sum_{i=1}^{m} T_i \log \pi_i$$

$$-\sum_{i<j} n_{ij} \log(\pi_i \cdot d_i + \pi_j \cdot d_j) + \text{const.} \qquad (20)$$

となる．この l を最大化するような π_i を求めて推定値とする．

4.2.2　π が既知であるとき

拡張モデル 1 の場合，対数尤度は

$$l = \sum_{i=1}^{m} T_i \log d_i$$
$$-\sum_{i<j} n_{ij} \log(\pi_i \cdot d_i + \pi_j \cdot d_j) + \text{const.} \qquad (21)$$

となる．この l を最大化するような π_i を求めて推定値とする．

5.　テストデータによる実験

予め一定値に定めたパラメータのもとで BT モデルに従い人工的な試合結果データを作成し，そのデータのみを使ってパラメータをどの程度正しく推定できるかどうか確認する．

プレイヤーを 100 人とし，$\pi_1 = 0.01, \pi_2 = 0.02, ..., \pi_{100} = 1.00$ とする．d の値はテスト毎に定める．式 (9) に示される確率に従ってプレイヤーが勝つ確率を定めた状態で，互いに何試合かさせてその勝敗をデータとして記録する．その勝敗データのみを用いて各プレイヤーの π_i を前述の方法により最尤推定し，最初に定めた $\pi_1 = 0.01, \pi_2 = 0.02, ..., \pi_{100} = 1.00$ と一致しているかどうか比較する．尤度関数の最大化は統計ソフトの R 上で行い，焼きなまし法を用いる．また，複数回推定したそれぞれの結果の平均を推定値として使用し，推定時の各 π_i の初期値はテスト毎に定める．

5.1　拡張モデル 1 に則したテストデータの作成と結果

2 種類のデータを作成する．選手間に有利不利が現れるよう，次の表のように d を定める．

表 1　テストデータの d_i の値

	プレイヤー 1,...,50	プレイヤー 51,...,100
テストデータ 1	0.5	1.5
テストデータ 2	1.5	0.5

互いに 1000 試合させたときの勝数分布は以下のようになった．d による補正の影響が見られる．

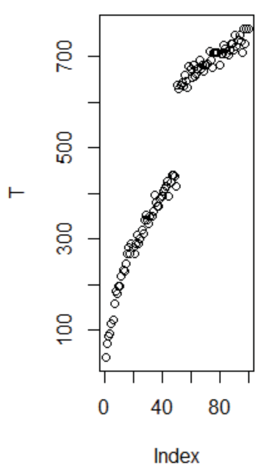

図 1　テストデータ 1 勝数分布

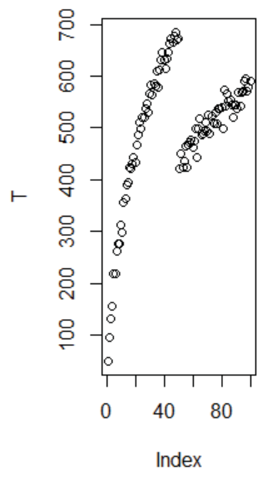

図 2　テストデータ 2 勝数分布

初期値として各 π_i に 0 から 1 までの乱数値を与え，100 回の推定結果を平均して推定値を導き出した．このときの結果は次のようになった．テストデータ 1, 2 の推定結果の平均二乗誤差の平方根 (RMSE) はそれぞれ 0.188, 0.186 となり，全体的に少し大きめに推定されているが，初期パラメータを概ね正しく推定できている．

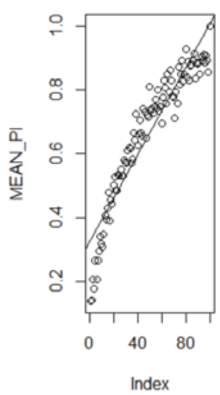

図 3　テストデータ 1 π_i 推定結果

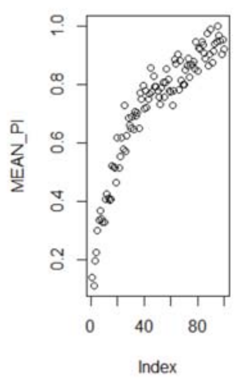

図 4　テストデータ 2 π_i 推定結果

続いて π_i を既知として，初期パラメータと等しい $\pi_1 = 0.01, \pi_2 = 0.02, ..., \pi_{100} = 1.00$ を与え，d_i を推定する．テストデータ 1，2 について，初期値として $d_1 =, ..., = d_{100} = 1$ を与えて 100 回の推定の平均値を取ったところ，以下のような結果となった．RMSE はそれぞれ 0.291, 0.296 となった．

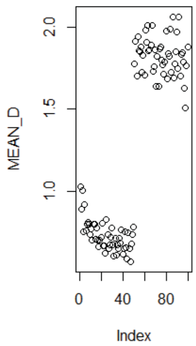

図 5　データ 1 d_i 推定結果 4

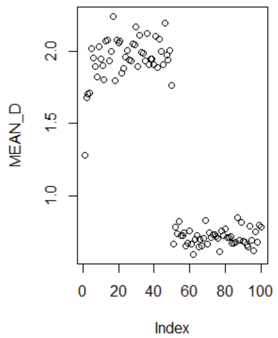

図 6　データ 2 d_i 推定結果

最後に，コートなどの条件により，選手が発揮する強さが試合毎に変化する場合を考える．条件 1, 2 により，次のように d_i が変化するとする．

表 2　条件による d_i の値

	条件 1	条件 2
$d_1, ..., d_{50}$	1.5	0.5
$d_{51}, ..., d_{100}$	0.5	1.5

条件 1, 2 をランダムに選んで 1000 回試合をさせたときの勝数の分布は，次のようになった．

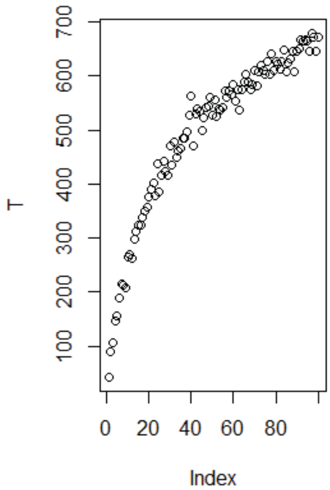

図 7　テストデータ 3 勝数分布

このままでは推定が行えないので，条件 1, 2 で行った試合をそれぞれ集めたテストデータ 4-1, 4-2 を作成する．$\pi_1 = 0.01, \pi_2 = 0.02, ..., \pi_{100} = 1.00$ を与え，d_i を推定した結果は以下のようになった．RMSE はそれぞれ 0.121, 0.124 となった．

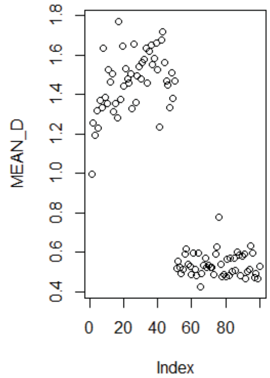

図 8　データ 4-1 d_i 推定結果

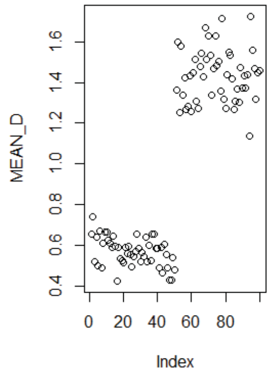

図 9　データ 4-2 d_i 推定結果

6. テニスデータによる実験

ATP Professional Tennis Tournament Data [6] から借用した，2014 年に行われたプロテニス国際大会の試合結果データを用いて，実験を行う．試合数は 2601 試合であり，内 5 セットマッチ (3 本先取) の試合が 507 試合あり，残りの 2094 試合は 3 セットマッチ (2 セット先取) の試合である．

6.1 点数計算

例えば 3 セットマッチの試合を行う場合，最終的な敗者が 1 セット勝っている場合と 1 セットも勝っていない場合の 2 つの場合が存在し，これらを混同することは正確な評価を妨げてしまうため，データが存在するならば敗者が勝利した本数も考慮することが望ましい．

テニスにおいては，6 ゲームを先に取ったプレイヤーが 1 セットを獲得し，3 セットマッチの場合は 2 セット，5 セットマッチの場合は 3 セットを先に獲得したプレイヤーが勝者となる．以上のような理由から，試合 m におけるプレイヤー i の勝数 T_{im} は次のように計算する．[*1]

$$T_{im} = g + 6s + 6f(m) \cdot g(i, m). \quad (22)$$

ただし g はプレイヤー i がその試合で勝利したゲーム数，s はプレイヤー i がその試合で勝利したセット数，関数 $f(m)$ は各試合に勝利するのに必要なセット数を示す．たとえば 3 セットマッチの場合 $f(m) = 2$, 5 セットマッチの場合 $f(m) = 3$ となる．関数 $g(m, i)$ はプレイヤー i が勝利したかどうかを示し，勝利した場合 $g(i, m) = 1$，敗北した場合 $g(i, m) = 0$ となる．

式 (12) より，試合数は試合をした 2 人のプレイヤーの勝

[*1] これはデュースを考慮していないが，デュースを考慮すると 1 セットや 1 ゲームの価値が無限大に発散してしまうため，今回は考慮しないこととする．

利数の和となるため，プレイヤー i, j 間の試合 m のデータ処理上の試合数 n_m は以下のように計算する．

$$n_m = T_{im} + T_{jm}. \tag{23}$$

6.2 ATP ファイナルに出場した選手に関する π, d 推定

ATP ファイナルに出場している 9 名 [*2] の間の試合のみを抜き出し，モデルを適用する．全員を対象にせず，一部の試合のみを抜き出して分析を行うのは，結果の解釈がしやすいからである．

まず，$d_i = 1(\forall i)$ とする．すなわち通常の BT モデルを適用する．結果は以下の表のようになった．

表 3 $d = 1$ ATP ファイナル分析結果

プレイヤー	π	$Point_i$
Berdych T.	0.406150779	0.371382637
Cilic M.	0.668363304	0.414201183
Djokovic N.	0.985602553	0.656207367
Federer R.	1	0.620224719
Ferrer D.	0.334865512	0.323924731
Murray A.	0.483295134	0.361182519
Nishikori K.	0.750665581	0.545300593
Raonic M.	0.486701954	0.376126126
Wawrinka S.	0.743049506	0.589242054

ただし，n_i をプレイヤー i の総試合数としたとき，

$$Point_i = \frac{T_i}{n_i}, \tag{24}$$

$$T_i = \sum_m T_{im}. \tag{25}$$

である．T_i はプレイヤー i の総勝数を示すが，T_i は試合数を重ねると大きくなる傾向があるため，1 ゲームあたりの勝数である $Point_i$ を比較対象として用いる．

表を見ると，Federer R. 選手がトップの π_i を獲得しているが，彼の $Point_i$ は 9 人中 2 位である．Federer R. 選手が他の選手に比べ強い選手との試合を多く経験したため，このような逆転が発生したと考えられる．

次に，各選手の身長をパラメータとして組み込むことを考える．各選手の身長 h_i と $Point_i$ の関係は以下の表のようになった．$h_i, Point_i$ の相関係数は -0.26 となり，弱い負の相関が認められた．

何らかの理由により，身長が大きい選手ほど有利だと判明したとする．このとき 9 人の選手の平均身長 (約 187.7cm) を $\bar{h}$ とおき，以下のように各選手の d_i を定める．

[*2] 本来の出場資格者はランキング上位 8 名なのだが，この年は出場者のうち Raonic 選手が 3 試合目を怪我のため棄権し，代わりに Ferrer 選手が出場したため，9 名となっている．

表 4 ATP ファイナル h_i と $Point_i$ の関係

プレイヤー	h_i	$Point_i$
Berdych T.	196	0.371382637
Cilic M.	198	0.414201183
Djokovic N.	188	0.656207367
Federer R.	185	0.620224719
Ferrer D.	175	0.323924731
Murray A.	190	0.361182519
Nishikori K.	178	0.545300593
Raonic M.	196	0.376126126
Wawrinka S.	183	0.589242054

$$d_i = \frac{h_i}{\bar{h}}. \tag{26}$$

この d_i のもとで π_i を推定すると，以下のような結果になった．

表 5 ATP ファイナル h_i 補正を行った π_i 推定結果

プレイヤー	π_i(補正後)	π_i(補正前)	h_i
Berdych T.	0.392665833	0.406150779	196
Cilic M.	0.647218616	0.668363304	198
Djokovic N.	0.994166618	0.985602553	188
Federer R.	1	1	185
Ferrer D.	0.357453095	0.334865512	175
Murray A.	0.482459687	0.483295134	190
Nishikori K.	0.778953632	0.750665581	178
Raonic M.	0.478190403	0.486701954	196
Wawrinka S.	0.749167658	0.743049506	183

表 (3) の結果に比べ，π_i の値に若干の変化が見られる．例えば身長が最も低い Ferrer D. 選手の π_i は，補正により少し上昇している．また，Murray A. 選手と，彼より身長の高い Raonic M. 選手の π_i は，補正前では Raonic M. 選手の方が大きいが，補正後は逆転している．

何らかの理由によりプレイヤーの π_i を推定できたと仮定する．今回は便宜上，表 (3) の π_i を使って d_i を推定する．π_i の値が正しいとすると，d_i の差は選手間の隠れた有利不利を表すはずである．結果は以下のようになった．

表 6 ATP ファイナル d_i 推定結果

プレイヤー	d_i	h_i
Berdych T.	0.796531091	196
Cilic M.	1.023119294	198
Djokovic N.	1.205155714	188
Federer R.	1.238714567	185
Ferrer D.	0.805986763	175
Murray A.	0.913603719	190
Nishikori K.	1.044427105	178
Raonic M.	0.922513751	196
Wawrinka S.	1.049947995	183

d_i, h_i の相関係数は -0.11 であり，相関は認められなかった．π_i の値を正しいとするならば，身長が π_i 推定時に大きな影響を及ぼしているとは考えにくいということが分かる．

6.3 全選手に関する π, d 推定

全ての選手に関して π, d を推定する．都合により，身長が判明しなかった選手が含まれる 14 試合は除外した．$T_i, n_i, Point_i$ を計算し，$h_i, Point_i$ の相関係数を導き出したところ，0.12 という値が得られた．相関はほとんど認められない．

ATP ファイナルの場合と同様に，まず $d_i = 1(\forall i)$ と仮定して π_i を推定し，推定結果を用いて d_i を推定する．選手の数が多いため正確な分析は困難であるが，$Point_i, \pi_i$ の相関係数は 0.10 であった．勝数と相関がほとんどないため，独自の評価方法に従っていると言える．また，$Point_i, d_i$ の相関係数は 0.11 であった．

ATP ファイナルに出場した 9 人のプレイヤーの結果を抜き出して次に載せた．"(MC)"と表示している列には ATP ファイナルの分析で載せた値を再掲している．全選手を分析対象としたためか，π_i に関しては ATP ファイナルで上位の選手ほど下がり，下位の選手ほど上がっている．ATP ファイナル出場者は全体の中でもハイレベルの選手であることを考えると，下位の選手が上がるのは妥当と言える．また，d_i に関しては全体的な上昇が見られた．これも ATP ファイナル出場者が比較的上位の選手であることを考慮すると，納得できる結果である．

表 7 全選手推定結果から ATP ファイナル選手抜き出し

プレイヤー	h_i	$Point_i$	π_i(全選手)	π_i(MC)
Berdych T.	196	0.633957553	0.880616038	0.406150779
Cilic M.	198	0.620127389	0.766509314	0.668363304
Djokovic N.	188	0.731825114	0.934043489	0.985602553
Federer R.	185	0.70692395	0.817853751	1
Ferrer D.	175	0.622790253	0.763402491	0.334865512
Murray A.	190	0.642268541	0.854982367	0.483295134
Nishikori K.	178	0.646428571	0.761805345	0.750665581
Raonic M.	196	0.616499736	0.872258337	0.486701954
Wawrinka S.	183	0.620597115	0.81192525	0.743049506

プレイヤー	d_i(全選手)	d_i(MC)
Berdych T.	1.8076729	0.796531091
Cilic M.	2.160893416	1.023119294
Djokovic N.	2.507295299	1.205155714
Federer R.	2.330586916	1.238714567
Ferrer D.	1.861269969	0.805986763
Murray A.	2.145548087	0.913603719
Nishikori K.	2.205395209	1.044427105
Raonic M.	1.615728231	0.922513751
Wawrinka S.	1.661621132	1.049947995

6.4 テニスコートを考慮した π, d 推定

テニスコート毎に試合を分け，それぞれのデータに拡張モデル 1 を適用することで，両選手に共通の試合条件が選手に与える影響を定量的に評価することを試みる．ATP ファイナルに出場した選手同士の試合は，ハード，クレイ，グラスの 3 種類のコートで行われているため，この試合データに拡張モデル 2 を適用して分析を行う．π_i の値は何らかの方法により既知であるとし，今回は便宜上表 (3) の結果を選手の π_i の推定値として用いる．d_i の初期値を全て 1 とし，100 回推定した結果の平均は以下のようになった．

表 8 ATP ファイナル コート影響推定

プレイヤー	d_{hard}	$Point_{hard}$
Berdych T.	0.768132181	0.445134576
Cilic M.	1.030431628	0.672527473
Djokovic N.	1.36158603	0.942771084
Federer R.	1.209112932	0.892638037
Ferrer D.	0.842305885	0.411134904
Murray A.	0.986751381	0.49471831
Nishikori K.	1.020100974	0.712034384
Raonic M.	0.852148688	0.515
Wawrinka S.	0.929430301	0.682326622

プレイヤー	d_{clay}	$Point_{clay}$
Berdych T.	0	0
Cilic M.	0.198776799	0.175
Djokovic N.	0.468326094	0.640866873
Federer R.	1.301668358	0.542857143
Ferrer D.	0.866441948	0.289940828
Murray A.	0	0
Nishikori K.	0.876346903	0.753246753
Raonic M.	0.412096161	0.246753247
Wawrinka S.	2.569035229	0.803191489

プレイヤー	d_{grass}	$Point_{grass}$
Berdych T.	0.656313239	0.222222222
Cilic M.	1.39779045	0.521212121
Djokovic N.	2.01923662	0.648241206
Federer R.	1.207928064	0.625827815
Ferrer D.	0	0
Murray A.	0	0
Nishikori K.	0.224614864	0.244094488
Raonic M.	0.755196154	0.482993197
Wawrinka S.	0.73892061	0.30952381

7. 本研究のまとめ

本研究では，プレイヤーの実力と独立したパラメータが存在する場合に，勝数や勝率による評価とは異なる，より適切な評価を行うレーティングシステムとして，Bradley–Terry

モデルをベースとした拡張モデルを提案し，人工的なテストデータでの実験により正常な推定ができることを確認した．さらに，2014 年の ATP プロテニス公式戦の試合結果データに拡張モデルを実際に適用し，BT モデルによるレーティングを行うことで対戦相手まで考慮した評価ができること，新たに追加した補正項の影響が確かに現れること，推定済のパラメータを使用して新たなパラメータの推定ができることを確認した．また，コート毎に試合データを分割することで，コートが選手に与える影響を示すパラメータを推定できることも確認した．

8. 今後の課題

8.1 別の手法による π, d の推定

本研究で提案した手法では，π, d を同時に推定することができず，片方のパラメータを推定するには，もう片方のパラメータが既知である必要がある．今回は便宜上，$d_i = 1 (\forall i)$ として推定した π_i を用いて d_i を推定したが，別の手法により π_i, d_i を推定する方が望ましい．π_i の推定に関しては，勝率や勝数などの一般的な値から推定値を導いたり，同じような条件下で試合をしたプレイヤーの集団のうち，最も強いプレイヤーの π_i を 1 と仮定するなどの方法が考えられる．d_i の推定に関しては，同じような条件下で試合をしたプレイヤーの集団を抜き出し，別の条件下にあるプレイヤーの集団との勝率を比較することで推定するという方法が挙げられる．集団の構成人数が多ければ多いほど，集団の平均的な技量，π は同じとみなせるので，勝率の比を d の比とみなせる．

8.2 d_i 推定の精度向上

図 (5), (6) より，d_i の推定に関しては π_i の推定よりも大きな誤差が生まれている．d_i の推定に関しては焼きなまし法による推定があまり適していない恐れがあり，追加の検証を要する．

参考文献

[1] Arpad E. Elo, “The Rating of Chessplayers Past and Present, ” Arco Publishing, New York, 1978.

[2] Mark E. Glickman. (2013) “Example of the Glicko-2 system.” <http://www.glicko.net/glicko/glicko2.pdf>, (accessed 2017-01-09).

[3] Ralf Herbrich, Tom Minka, and Thore Graepel, “TrueSkill(TM): A Bayesian Skill Rating System”, in Advances in Neural Information Processing Systems 20, MIT Press, January 2007.

[4] Bradley, R. A. and Terry, M. E. : “Rank analysis of incomplete block designs. 1. The method of paired comparisons, ” Biometrika 39, 324-335 (1952).

[5] 竹内 啓, 藤野 和健：スポーツの数理科学—もっと楽しむための数字の読み方–, 共立出版. (1988).

[6] ATP & WTA Professional Tennis Tournament Data. (2011), <https://datahub.io/ja/dataset/atp-wta-professional-tennis-tournament-data>, (accessed 2017-01-09).

[7] 大竹 勇志 :Bradley–Terry モデルを用いた対戦型ゲームにおけるハンデ設定, (2015).

[8] Rmi Coulom. (2007) “Computing Elo Ratings of Move Patterns in the Game of Go.” <https://www.remi-coulom.fr/Amsterdam2007/MMGoPatterns.pdf>, (accessed 2017-01-13).

[9] David R. Hunter. (2004) “MM algorithms for generalized Bradley-Terry models.” <http://sites.stat.psu.edu/ dhunter/papers/bt.pdf>, (accessed 2017-01-13).