# 引き分けを考慮したBradley-Terryモデルの適用法の研究

M2012MM035 坂巻翔太

指導教員：松田眞一

## 1　はじめに

将棋, スポーツ, 麻雀など対戦というものに勝敗は付き物であり, どちらが強いのかということに関心が寄せられている. スポーツには勝ち, 負けの他に引き分けというものが存在し, 試合を面白くする要因の一つである. 特にサッカーは他のスポーツと違い, 引き分けが多いスポーツである. Bradley-Terry モデルはスポーツのリーグ戦のような 2 つのチームが対戦して勝ち負けを争い, 直接対戦しないチームの「強さ」を推定できる.

今井・松田 [1] ではこの推定法を拡張した引き分けの試合を考慮した研究を行なっていたが, 引き分けを用いた場合の適用法の研究は行なわれていない.

本研究では今井・松田 [1] の研究を拡張し引き分けを考慮したときにこの推定法を適用することができるのかどうかの確認およびさらに拡張できるのかを検討する.

### 1.1　先行研究

平手 [5] ではサッカーの試合のデータを用いて, 上位 30 チームに 31〜50 位,51〜70 位,71〜100 位,101 以下の区切りに分けたチームの対戦結果を元にそれぞれのチームの強さを推定し, 世界ランキングと比較した.

## 2　Bradley-Terry モデル

個人もしくはチームが互いに対戦を行ない勝敗を競い, どちらか一方の個人もしくはチームが他方に勝つ確率をモデル化し,「強さ」を推定する時に Bradley-Terry モデルが用いられている.（以下では BT モデルと略す.）(竹内・藤野 [3])

### 2.1　引き分けを考慮した BT モデル

BT モデルとは複数のチームもしくは個人が対戦し, 一方が他方に勝つ確率を求めるために強さを推定することのできる統計的方法である. 特徴の一つとして, この推定方法を用いると直接対戦したことのないチーム間での強さも推定することができる. そこで勝敗の確率を次のように考えるとする. 勝敗を考える対象をチームと呼び, 全部で $m$ チームあるとする. 第 $i$ 番目のチームが第 $j$ 番目のチームに対して勝つ確率, 負ける確率, 引き分ける確率をそれぞれ $p_{ij}$,$q_{ij}$,$r_{ij}$ とすると

$$p_{ij} + q_{ij} + r_{ij} = 1, p_{ij} = q_{ji}, r_{ij} = r_{ji}$$

が成り立つ. 各チームに対応して $m$ 個の量 $\pi_1$,$\pi_2$, $\cdots$,$\pi_m$ が存在して, 全ての $i$,$j$ の組み合わせに対し, 確率 $\pi_{ij}$ というパラメータを導入し,

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j}(1 - r_{ij})$$

が成立しているとする. 一方, 引き分けについては次のように 2 つのパラメータ $\alpha$,$\beta$ に依存するモデルを考える.

$$r_{ij} = \alpha - \beta\left(\frac{\pi_i - \pi_j}{\pi_i + \pi_j}\right)^2$$

このモデルは 2 つのチームの強さに依存して引き分けの確率が決まると考えるもので, パラメータ $\beta$ の値が正であればチームの力の差が大きくなると引き分けにくく, 負であれば力が近いほど引き分けにくいことを表すことになる.(今井・松田 [1] 参照)

### 2.2　パラメータの推定

実際に第 $i$ 番目のチームと第 $j$ 番目のチームの対戦が $n_{ij}$ 回行なわれた場合に第 $i$ 番目のチームが勝つ回数の確率変数を $X_{ij}$, 負ける回数の確率変数を $Y_{ij}$, 引き分けとなる回数の確率変数を $Z_{ij}$ とおくとそれらは多項分布に従っていると考えられ, 次のように確率分布が定まる.

$$\Pr\{X_{ij} = x_{ij}, Y_{ij} = y_{ij}, Z_{ij} = z_{ij}; 1 \leqq i < j \leqq m\} = \prod\prod_{i<j}\frac{n_{ij}!}{x_{ij}!y_{ij}!z_{ij}!}p_{ij}^{x_{ij}}q_{ij}^{y_{ij}}r_{ij}^{z_{ij}}$$

ただし

$$X_{ij} = x_{ij}, Y_{ij} = y_{ij}, Z_{ij} = z_{ij}; 1 \leqq i < j \leqq m$$

とする. これに対して, 前節のモデルを導入し, 対数尤度の最大化で方程式を立てると最終的に

$$\begin{cases} \pi_i = \dfrac{T_i}{\displaystyle\sum_{j\neq i}\frac{n_{ij} - z_{ij}}{\pi_i + \pi_j} - \sum_{j\neq i}\frac{4\beta\pi_j(\pi_i - \pi_j)}{\pi_i + \pi_j}\frac{(\alpha n_{ij} - z_{ij})P_{ij} - \beta n_{ij}Q_{ij}}{\{(1-\alpha)P_{ij} + \beta Q_{ij}\}\{\alpha P_{ij} - \beta Q_{ij}\}}} \\ \alpha = \dfrac{\displaystyle\sum\sum_{i<j} z_{ij} + \beta\sum\sum_{i<j}\frac{(n_{ij} - z_{ij})Q_{ij}}{(1-\alpha)P_{ij} + \beta Q_{ij}}}{\displaystyle\sum\sum_{i<j} n_{ij}} \\ \beta = \dfrac{\displaystyle\sum\sum_{i<j}(\alpha n_{ij} - z_{ij})}{\displaystyle\sum\sum_{i<j}\frac{z_{ij}Q_{ij}}{\alpha P_{ij} - \beta Q_{ij}}} \end{cases}$$

が得られる. ただし,

$$T_i = \sum_{j\neq i} x_{ij}$$

$$P_{ij} = (\pi_i + \pi_j)^2, Q_{ij} = (\pi_i - \pi_j)^2$$

とおく. この関係式を基に繰り返し計算で推定値を求める. ちなみに $\alpha$ は同じ強さの場合の引き分ける確率,$\beta$ は強さの差に対する引き分け係数である.(今井・松田 [1] 参照)

## 3　引き分けを考慮した BT モデルを用いた強さの推定

今井・松田 [1] で提案された引き分けを考慮した BT モデルの R 関数を用いて, 実際のサッカーデータにも実用できるのかを確かめ強さの推定を行なう.BT モデルの特徴を活かし, 直接対戦のないチーム間の強さを比較することができる.

### 3.1　試合のデータ

2010 年 1 月 1 日から 2012 年 6 月 30 日までに実施された FIFA 加盟国 207ヶ国の代表チームによる国際 A マッチの試合をデータとする. 集計した項目は, 対戦国, 試合の重要度 (WC 本選:1, 大陸選手権の本大会・FIFA コンフェデレーションズカップ:0.75, 大陸選手権の予選・FIFA ワールドカップ予選:0.63, 親善試合:0.25), 試合結果である. 上記で述べた項目を勝ち点制と勝ち数制の 2 つの方法でデータをまとめ,2 つの方法それぞれ勝ちと引き分けの時のデータを収集した.

対戦国については, 上位 30 チームを中心にデータを収集し, それ以下のチームについては,31〜40 位,41〜50 位,51〜70 位,71〜100 位,101 位〜という区分ごとで一まとまりとして考える. 試合の重要度については,FIFA 公式ランキング算出方法の割合を用いる. また, 今回使用するデータは FIFA ワールドカップ本大会での試合を基準とし,「勝ち」の場合を「1 勝」または「勝ち点 1」とし, 引き分けの場合は「0.5 勝」または「勝ち点 0.33」とする. 試合結果については, 対戦国それぞれに対して,「勝ち」,「負け」,「引き分け」の 3 通りとする. なお,90 分間で勝敗が着いた場合を「勝ち」,「負け」とし, 同点, 延長または PK で勝敗が着いた試合は「引き分け」とする. 以下本論文ではワールドカップのことを WC と略す.

### 3.2　FIFA ランキングデータ

FIFA ランキングデータとは国際 A マッチの成績をポイント化した FIFA 加盟国のランキングであり, 毎月 1 回発表される. ランキングの算出 (「Jcalco」FIFA ランキング一口メモ [6]) については勝ち点 (A), 試合の重要度 (B), 対戦国家間の強さ (C), 大陸連盟の強さ (D) の 4 つの値を元に計算している. 実際に FIFA ランキングが算出されるときの「勝ち点」は全ての国際 A マッチにおいて, 勝ち：3 点, 引き分け：1 点, 負け：0 点とする. ただし,PK 戦での勝敗については勝ち：2 点, 負け：1 点の勝ち点制である. 今回は勝ち点と試合の重要度を中心に考えるので対戦国家間の強さと大陸連盟間の強さは考慮しないものとする. また, 今回使用する FIFA ランキングデータについては,2012 年 7 月 4 日に更新されたランキングデータを用いている.

### 3.3　推定方法

勝ち数を国際 A マッチの試合の重要度に絡めて, 上位 30 チームと 31〜40 位,41〜50 位,51〜70 位,71〜100 位,101 位以下という区切りで以下の 4 通りのデータ方法で推定する.（服部・加藤 [4] 参照）ただし, 今回は勝ち数の結果のみ載せることとする.

1. 勝ち数制 (重要度あり)
2. 勝ち数制 (重要度なし)
3. 勝ち点制 (重要度あり)
4. 勝ち点制 (重要度なし)

なお, 今回はワールドカップの試合を基準として, 一まとめにした 4 区分のチームについては上位 30 チームとの対戦結果のみを考慮し, 一まとめにしたチーム間同士の対戦結果は考慮していない. また, 試合の重要度を絡めた場合の勝ち数制と勝ち点制の場合, それぞれの勝ちと引き分けの時の割合のポイントは以下の表 1 の通りであり, この値を用いて各チームの推定値を計算した.

表 1　試合の重要度

| | 勝ち数制 | | 勝ち点制 | |
|---|---|---|---|---|
| 勝敗 | 勝ち | 引き分け | 勝ち | 引き分け |
| a | 1 | 0.5 | 1 | 0.33 |
| b | 0.75 | 0.375 | 0.75 | 0.25 |
| c | 0.63 | 0.313 | 0.63 | 0.208 |
| d | 0.25 | 0.125 | 0.25 | 0.083 |

この表 1 の「引き分け」は,90 分間で同点, 延長もしくは PK 戦で勝敗が着いた場合も含めている. 表 1 にある a〜d は試合の重要度であり, 以下の通りである.(なお, 過去の研究平手 [5] の表とは異なっているが, 本論文の方の値が正しい)

- a:WC 本選
- b:大陸本大会・コンフェデ杯
- c:大陸予選・WC 予選
- d:親善試合

### 3.4　推定結果

今回は WC を含める場合＆含めない場合の勝ち数の重要度あり＆なしを載せる. WC を含めた場合の勝ち数（重要度あり）のときの順位相関係数,$\alpha,\beta$ は次のような結果を得られた. まず, 平手 [5] の時に一番良かった区分は 31〜40, 41〜50, 51〜70, 71〜100, 101〜 であり, その時の結果は 0.759,0.183,0.172 である. 31〜40, 41〜50, 51〜60, 61〜70, 71〜100, 101〜 の区分の時は 0.744,0.177,0.146 である.31〜50, 51〜70, 71〜100, 101〜 の区分の時は 0.749,0.182,0.171 である. 平手 [5] 同様に順位相関係数が一番良かったのは,31〜40, 41〜50, 51〜70, 71〜100, 101〜 の区分の時の 0.759 である. WC を含める場合の勝ち数（重要度なし）のときの順位相関係数,$\alpha,\beta$ は次のような結果を得られた.31〜40, 41〜50, 51〜60, 61〜70, 71〜100, 101〜 の区分の時は 0.669,0.144,0.111 である. 31〜40, 41〜50, 51〜70, 71〜100, 101〜 区分の結果は 0.649,0.144,0.112 である. 31〜50, 51〜70, 71〜100, 101〜 の区分の時は 0.672,0.143,0.11 である.

WC を含めない場合での一番良かった値のときの順位相関係数,$\alpha,\beta$ は次のような結果を得られた. 勝ち数 (重要度あり) のときの 31〜40, 41〜50, 51〜60, 61〜70, 71〜100, 101〜 の区分の時は 0.825,0.163,0.149 である. 勝ち数 (重要度なし) のときの 31〜50, 51〜70, 71〜100, 101〜 の区分の結果は 0.717,0.132,0.096 である.

### 3.5　考察

WC の試合を含めた試合で勝ち数, 勝ち点共に推定結果を求めたところ一番良かったのは勝ち数の重要度ありで

あった．全体的に影響力の高い WC の試合結果を抜くことによってだいぶ順位相関係数の値が良くなった．オランダやスウェーデンのように FIFA ランキングよりも良い推定結果が出たチームは WC などの重要な試合はもちろんほとんどの試合で勝っており, 強いチームと対戦しても引き分けなどの試合であまり負けていないからだと考えられる. 逆にイタリア, デンマーク, ロシアなどの FIFA ランキング順位よりも推定値の評価が低くなってしまったチームはランキング上位にも関わらず, 試合数の割に勝ち数が少ないためだと考えられる. 重要度なしでは全体的に推定値が高めになってしまっている. これは重要度なしで検定を行なうと FIFA が重要視している貴重な試合の戦績を無視しているとみなされ, 親善試合も WC も同じ扱いとなり一つ一つの試合に重みが無くなってしまったからだと考えられる. 以後は 9 区分のみ調べる.

## 4　適合度検定

適合度検定とはあるデータが予測され得る確率分布に従うかどうかを確かめるための統計的手法である．適合度検定の一般論について説明する.（白旗 [2] 参照）

$k$ 個の独立した項目 $A_1,\cdots,A_k$ があり,$p_i = \Pr\{A_i\}$ とする．帰無仮説として

$$H_0 : p_i = p_i(\theta), i = 1, \cdots, k$$

を考える．$\theta$ は未知パラメータとする．$H_0$ を仮定したときの $\theta$ の推定量を $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_k)$ とし,$n$ は十分大きいとする．$E(X_i) = np_i$ なので，統計量

$$\chi^2 = \sum_{i=1}^{k} \frac{(X_i - n\hat{p}_i)^2}{n\hat{p}_i}$$

を考えると,$\chi^2$ は帰無仮説から離れるほど大きな値を取る．$\chi^2$ は帰無仮説の下で近似的に自由度 $k - r - 1$ のカイ 2 乗分布 $\chi^2_{k-r-1}$ に従う. ここで $r$ はパラメータ $\theta$ の次元である.

### 4.1　検定に用いるデータについて

実際に行なわれた全試合を対象とし,WC を含めた試合＆含めなかった試合での勝つ確率, 負ける確率, 引き分ける確率を求めて, その値を区分けし, 区分ごとでの勝敗を集計して適合度検定を行なう. ちなみに各区分ごとでの各々確率はその区分内で集計した確率の平均の値である. WC の試合結果を含めなくてもその試合結果を予測することができるのか, また $\alpha, \beta$ がどのような動向を示すのかを検証する. 次にデータと検定法の詳細を示す．対戦する 2 チーム間で, 推定値が高い方の勝つ確率を $x$ とし, 勝率が低い方（他方が勝つ確率）の確率を $y$, 引き分ける確率を $z$ とする. 以下のようにして区分けする．観測度数は全試合結果を用いて, それぞれ「勝ち」を「1 勝」「勝ち点 1」,「負け」を「0 勝」,「勝ち点 0」,「引き分け」を「0.5 勝」「勝ち点 0.33」とし, 各区分ごとに試合結果を集計する.

$$\begin{cases} 0.95 < x \leq 1 \text{ かつ } (0 \leq y < 0.05) \text{ かつ } (0 \leq z < 0.05) \\ 0.85 < x \leq 0.95 \text{ かつ } (0.05 \leq y < 0.15) \text{ かつ } (0 \leq z < 0.15) \\ 0.75 < x \leq 0.85 \text{ かつ } (0.15 \leq y < 0.25) \text{ かつ } (0 \leq z < 0.25) \\ 0.65 < x \leq 0.75 \text{ かつ } (0.25 \leq y < 0.35) \text{ かつ } (0 \leq z < 0.35) \\ 0.55 < x \leq 0.65 \text{ かつ } (0.35 \leq y < 0.45) \text{ かつ } (0 \leq z < 0.45) \\ x \leq 0.55 \text{ かつ } (0.45 \leq y < 0.55) \text{ かつ } (0 \leq z < 0.55) \\ 0 \leq z \leq y < x \end{cases}$$

WC 以前のデータから WC の結果を予測する問題ではパラメータ推定に用いるデータと統計量に用いるデータとが独立であり，第 4 節の一般論において予測すべきパラメータ $\theta$ はない．よって，縦方向それぞれの自由度は 2 となり，それが 6 列あるので全部で 12 の自由度となる．したがって，その統計量と自由度 12 からカイ二乗値の有意水準 5 %の値は 21.03 である．一方，パラメータ推測と統計量の計算でデータが独立でない場合は自由度の計算が簡単には行えない．自由度は最大 12 であると分かるだけで自由度がどの程度減るのか明確ではない.

### 4.2　検定結果のまとめ

4 パターンの中で適合度検定の結果が一番良かったのは勝ち点の重要度ありであり, この時の WC 以前データを用いて WC のデータに対する適合度検定を行なった. この時の P 値は 0.852 である．以下に勝ち点の重要度あり（WC 以前データによる WC データでの適合度検定結果）の適合度検定の結果を表 2 に示す.

表 2　検定結果　勝ち点の重要度あり（WC 以前データによる WC データでの適合度検定結果）

| | 9 割以上 | 9 割 | 8 割 | 7 割 | 6 割 | 5 割以下 |
|---|---|---|---|---|---|---|
| $x$ 勝ち数 | 62.89 | 85.63 | 55.24 | 45.8 | 36.65 | 27.91 |
| $y$ 勝ち数 | 0 | 8.9 | 10.39 | 17.14 | 13.39 | 19.88 |
| $z$ 引き分け | 1.123 | 1.659 | 6.558 | 5.27 | 5.688 | 5.073 |
| $w$ 試合数 | 64.013 | 96.189 | 72.188 | 68.21 | 55.728 | 52.863 |
| $x$ 確率 | 0.969 | 0.915 | 0.798 | 0.693 | 0.601 | 0.493 |
| $y$ 確率 | 0.015 | 0.053 | 0.14 | 0.22 | 0.298 | 0.396 |
| $z$ 確率 | 0.016 | 0.031 | 0.062 | 0.086 | 0.101 | 0.112 |

また, 勝ち数, 勝ち点の重要度あり, なしの $\alpha, \beta$, 適合度検定の結果は次の表 3 に示す.

表 3　WC 前予測の WC 含めないデータ

| WC 含めない | α | β | 適合度 |
|---|---|---|---|
| 勝ち数あり | 0.163 | 0.149 | 7.614 |
| 勝ち数なし | 0.132 | 0.096 | 24.164 |
| 勝ち点あり | 0.113 | 0.104 | 7.085 |
| 勝ち点なし | 0.092 | 0.061 | 16.672 |

### 4.3　$\alpha,\beta$ について

$\alpha,\beta$ について次のような結果が得られた.$\alpha,\beta$ の相対誤差は図 1 の通りである. 図の縦線は「適合度検定の値」, 横線は「$\alpha - \beta$」の値である. この図より,$\alpha,\beta$ の差が小さいほど適合度検定の値が良くなり, 差が大きくなるほどあてはまり具合が悪くなるという関係性がわかる．また,WC

を含めた結果の方が比較的安定しやすいようだが,WC を除いた試合でも当てはまり具合は良いようだ.

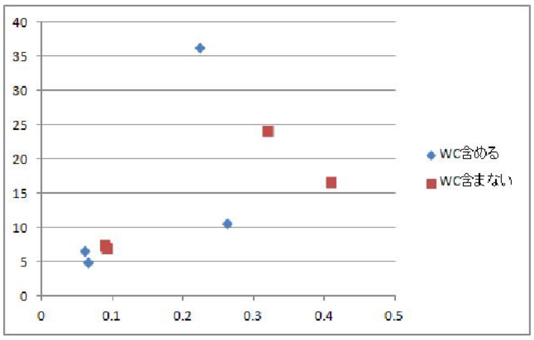


図 1 検定結果 $\alpha,\beta$ の相対誤差

### 4.4 考察

BT モデルの適用法を行なった平手 [5] の研究結果とは異なり, 引き分けを考慮した BT モデルの場合では WC を含める場合, 含めない場合共に勝ち点の重要度ありのときの適合度検定の値が良くなった. また,WC を含める場合, 含めない場合の勝ち数, 勝ち点の重要度なしはともに順位相関係数も適合度検定のあてはまり具合も重要度を考慮した場合と比べ, 低い値となった. また, 前節 4.3 で述べたように $\alpha,\beta$ の相対誤差は低いほど適合度検定のあてはまり具合がよくなるようである.

## 5 まとめ

平手 [5] 同様に引き分けを考慮した BT モデルでは, 勝ち数, 勝ち点の重要度ありとなしでは「勝ち数・重要度あり」のパターンが最も相関がよかった. 適合度検定のあてはまり具合まで考慮すると一番良かったのは「勝ち点の重要度あり」のときだった. どのチームも直接対戦のないチームがあるに加えて, 試合の重要度や対戦国間の強さのポイントを考慮してないが, それなりに FIFA ランキングに近い結果になったと思う. 今回のモデルにおいても試合の重要度がランキングに大きな影響を及ぼすことがわかった. WC の試合結果を含めた試合&除いた試合でのデータを集め解析して見たところ, 全体的に影響力の高い WC の試合結果を抜くことによってだいぶ順位相関係数の値が良くなった.WC の試合結果は格上チームが格下チームに負けるなどの番狂わせが起こった時に大きな影響を与えることがあり, そのせいで FIFA ランキングとの関係性が薄れてしまうことがある.

平手 [5] のときに推定値が良かったスペインは今回のモデルでは少し推定順位が下がってしまった. これはほとんどの試合ではそこそこ勝っているが, 特に WC などの重要な試合で上位チームとの対戦時に引き分けが多い事が原因であると考えられる.

適合度検定では,WC の試合結果を含めた試合&除いた試合ともに重要度がある方があてはまり具合が良かった.WC 以前のデータからも WC の試合予想はできたと思う. ただし, 全試合を通して長い目で見れば今回のモデルでも有効性は見られるが,WC 本大会の個々の試合においては強いチームが弱いチームに負けてしまうという波乱が起こる試合もあり, 個別の試合結果まで当てられるわけではない.WC 本大会の影響力が強く実際の FIFA ランキングではそれほど順位が高くないチームでも推定値が良くなってしまうようだ. また, 今回のデータでは直近 1 年ごとの割合ポイントを使用しなかったため, 実際の FIFA ランキングとは異なった結果になってしまったと考えられる. また,WC 本大会のあった 2010 年は今回使用したデータの中でも試合数が多いうえに, 試合の重要度も高くなってしまったため実際の FIFA ランキングとは結果が異なってしまったと考えられる.

今回の研究において $\alpha,\beta$ の符号がすべて正であった. 同じぐらいの強さのチーム同士の方がチーム間の強さの差が離れているチーム同士の試合よりも引き分けやすいという結果を得ることができた. また,$\alpha,\beta$ の符号の値が近ければ近いほど適合度検定のあてはまり具合も良く符号がの値の差が離れれば離れるほど当てはまり具合も悪くなるという関係性がわかった. また,WC を含めた結果の方が比較的安定しやすいようだが,WC を除いた試合でも当てはまり具合は良いようだ. 全体として, 引き分けを考慮した BT モデルでも WC 予測が可能であることが分かった. 引き分けのある試合でも今回の推定法は有効であると言える.

## 6 おわりに

親善試合から WC までの試合結果を収集する際に勝ち数&勝ち点の 2 パターンに対して重要度あり&なしの計 4 パターンの結果をまとめるのに時間がかかり大変だった. しかし, そのおかげで引き分けを用いた BT モデルの適用法が実際の FIFA ランキング加盟国のサッカー試合を用いて有効であることが分かった. また $\alpha,\beta$ についての動向を調べるのがとても難しかった. 今回は $\alpha,\beta$ の符号が負の時の動向を調べることが出来なかったので調べてみたい. 引き分け係数の値のパターンについて十分確かめることができたとは言えないが, 全体的にそれなりに良い結果になったと思う.

## 参考文献

[1] 今井寛・松田眞一:「引き分けを考慮した BT モデルの性能評価」, 南山大学紀要アカデミア数理情報編, 第 3 巻,pp.35-45,2003.

[2] 白旗慎吾:「統計解析入門」, 共立出版,1992.

[3] 竹内啓・藤野和建:応用統計数学シリーズ「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.

[4] 服部匡志・加藤明:「プロテニスプレーヤーの強さの統計的研究」, 南山大学数理情報学部数理情報学科卒業論文要旨集,1997.

[5] 平手良典:「サッカーにおける Bradley-Terry モデルの適用法の研究」, 南山大学大学院数理情報研究科数理情報専攻修士論文,2013.

[6] 「Jcalco」FIFA ランキング一口メモ http://www.fifaworldranking.com/wrhow.html,2001.