

Jリーグにおけるレーティングとその統計分析

2020SS042 水野景介

指導教員：塩濱敬之

1 はじめに

本研究はJリーグについて、2018年から2023年の6年間の試合結果を基にMasseyのレーティング推定を行う[1]。推定したレーティングに基づいてJリーグのリーグ間、あるいは異なるシーズン間についてレーティング分布に違いがあるのかに関する統計分析を行う。

2 Masseyのレーティング

Masseyのレーティングは、スポーツなどの試合結果に基づくレーティングによく用いられる手法である。内容としては、プレイヤー*i*と*j*によって行われた試合を*k*、得点差を y_k 、両プレイヤーのレートそれぞれ r_i, r_j と置き、次式(1)が成り立つことを仮定している。また、 m プレイヤーによって n 試合行われた場合、次式(2)を得ることができる。

$$y_k = r_i - r_j \quad (1)$$

$$Xr = y + \varepsilon \quad (2)$$

X は $n \times m$ の行列であり、行列内の要素は $X_{ki} = 1, X_{kj} = -1$ 、その他の要素が0となっている。また、レーティングベクトル r は $\hat{r} = (X^T X)^{-1} X^T y$ により推定が可能である。また、 $X^T X = M, X^T y = p$ とそれぞれ置き換えると、 $\hat{r} = M^{-1} p$ となる。行列 M の対角成分は試合数、非対角成分は各チームとの対戦回数を表している。またベクトル p は得失点差を表している。しかし、 M は非正則行列であるため、 M の任意の1行を全て1にしたものを \tilde{M} 、それに対応する p の成分を0にしたものを \tilde{p} と置くことで、レーティングベクトルは次式により推定できる[2]。

$$\hat{r} = \tilde{M}^{-1} \tilde{p}.$$

3 Jリーグについて

3.1 概要

Jリーグ（邦文表記：日本プロサッカーリーグ、英文表記：JAPAN PROFESSIONAL FOOTBALL LEAGUE）は1993年発足の日本のプロサッカーリーグである。

現在Jリーグは3部制となっており、J1, J2, J3のリーグがある。部が増える背景として、リーグ内のレベルや競争力の向上（チームレベルの差が見られるようになったため、それを揃えて底上げを図った）を図ること、地方チームが増加傾向にあること、若手選手の育成機会・プレー機会の創出などが挙げられる。各リーグの開催期間は毎年2月頃から12月頃までとなっている。どの部についても、順位付けは勝ち点制が取られており、勝利チーム3ポイント、

敗北チーム0ポイント、引き分けの場合は両チーム1ポイントが与えられる。勝ち点が並んでいた場合は、得失点差・総得点数・直接対戦成績（勝ち点・得失点差・総得点数）・反則ポイント・抽選の順で評価し順位を決定する。対戦方法はホームとアウェイを入れ替えて総当たり戦を2回行う“ホーム&アウェイ方式”が取られている。リーグ間の昇格降格については、基本的に各リーグ間で2から3チームが入れ替わる。チーム数は近年3リーグとも20チーム前後となっており、チーム数の遷移は表1のようになっている。

表1 チーム数の遷移

シーズン	2017年	2018年	2019年	2020年
J1	18	18	18	18
J2	22	22	22	22
J3	17(3)	17(3)	18(3)	18(2)
シーズン	2021年	2022年	2023年	2024年
J1	20	18	18	20
J2	22	22	22	20
J3	15	18	20	20

3.2 課題

現在のJリーグには以下のような課題が存在する。リーグ間の収入格差が存在すること、すなわち、平均年俸は明確に公表されているものはないが、WEBサイトに記載されているものによるとJ1は3000から3500万円程度、J2は400から600万円程度、J3は300から400万円程度となっている。また、売上高格差が存在すること、すなわち、2022年度決算では、売上高リーグ1位の浦和レッズが81億2700万円であるのに対し、売上高リーグ最下位のY.S.C.C.横浜(J3)は1億5800万円と大きな格差が生じている。リーグ間格差が所属するチーム間の力量の差にどのように影響を与えるのか、そのような問題意識に答えるために、5年間に渡るJリーグのリーグ間格差の統計分析を行う。

3.3 データについて

使用するデータは、2018年から2023年の各年におけるJ1-J3の各リーグの全試合結果、合計6年間6210試合のデータを用いる[3]。データは、年(シーズン)、所属リーグ、ホームチーム名、アウェイチーム名、ホームチームの得点、アウェイチームの得点を用いた。

4 レーティング推定と分析

前項のデータを用いてレーティング推定を行い、等分散の検定を行った。各リーグ内のレーティングに対して独立

性を仮定した k 標本のレーティングの等分散の仮定に対する仮説検定を行う。リーグ間の等分散性、シーズン間の等分散性、リーグシーズン間の等分散性の帰無仮説を次のように設定した。

$$H_0^{(\text{League})} : \sigma_{J1}^2 = \sigma_{J2}^2 = \sigma_{J3}^2,$$

$$H_0^{(\text{Year})} : \sigma_{2018}^2 = \sigma_{2019}^2 = \dots = \sigma_{2023}^2,$$

$$H_0^{(\text{LY})} : \sigma_{J1,2018}^2 = \sigma_{J1,2019}^2 = \dots = \sigma_{J3,2023}^2.$$

また、対立仮説は、母分散は等しいとはいえない(少なくとも1つの母分散は異なる)とする。等分散の検定にはいくつかの手法が存在し、今回はバートレット検定とルビーン(Levene) 検定の2つの手法を用いて行った[4]。帰無仮説の下で、次の検定統計量検定の結果を表2から4に記す。

表2 バートレット検定

従属関数	独立関数	p 値	統計量
レート	リーグ	0.6973	0.7211
レート	年	0.1575	6.6182
レート	リーグ, 年	0.2308	17.494

表3 ルビーン検定

従属関数	独立関数	p 値	統計量
レート	リーグ	0.9465	0.0550
レート	年	0.1071	1.9208
レート	リーグ, 年	0.2253	1.2705

この結果から、すべての検定下ですべての帰無仮説に対して等分散の帰無仮説は棄却されない。したがって、Jリーグ6年間の各リーグについてレーティング(力量)分布のばらつきに大きな違いが存在しないことが分かった。また、因子をリーグ、各シーズンとした2元配置分散分析を用いたところ平均レーティングにも統計的に有意な差がないことが分かった。

次に、2年毎にデータを集計して、レーティング推定を行った。2年毎にまとめてレーティング推定を行うことで、リーグ間の昇格降格を含み、Jリーグ全体の特徴や傾向を調べることができる。図1は2年毎のレーティングを各リーグ毎に推移を示す箱ひげ図、図2は各リーグにおける2年毎のレーティングの平均値の推移を表した折れ線グラフである。これを見ると、図1からは上からJ1, J2, J3の順になっており、シーズンによる昇降は少なく、レートの値が取る範囲は基本的に大きくは重なっていないことがわかる。また図2からは、J1は緩やかに上昇、J2とJ3は昇降を繰り返しており、この6年間ではリーグ間格差は緩やかに拡大しているように見える。この結果の理由を考察すると、まずJ3は2020年から2021年にかけて下降しているが、この年はチーム数が15チームとオープン参加のチームが無くなった為、例年に比べて2,3チーム少なくなっている。また2021年から2022年にかけては上昇しているが、これはチーム数が18に戻り、2022年から新規参入の

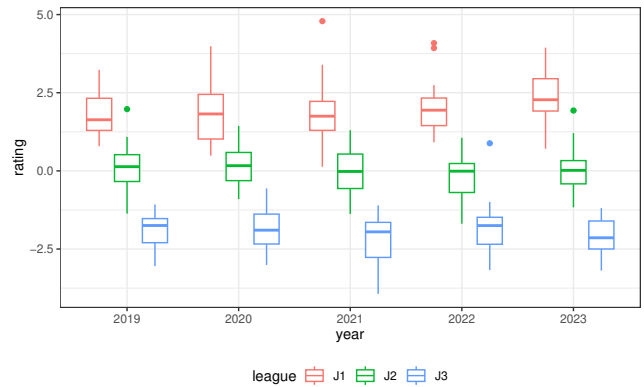


図1 2年毎のレーティングの推移を示す箱ひげ図

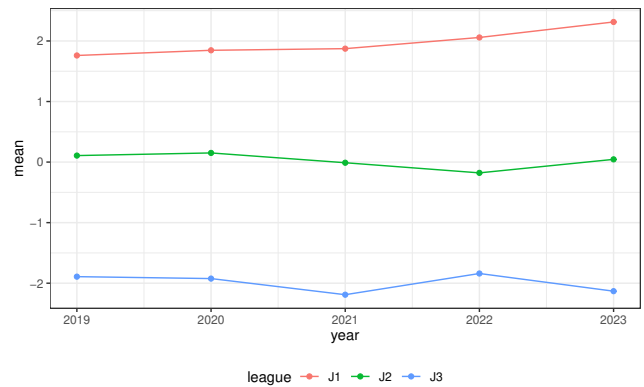


図2 各リーグにおける2年毎のレーティングの平均値推移の折れ線グラフ

いわきFCが優勝しており、相当な突き上げがあったのではないかと考える。また、各リーグ間で概ね2点ほどの差が存在し、Jリーグでは1試合平均得点が2点台であるため、リーグ間でのこの点差は非常に大きいことがわかる。

5 まとめ

本研究結果から、J1-J2, J2-J3間には点差にして概ね2点程度の格差が存在すること。一方、各リーグ内やシーズン間には有意な差は無いことが分かった。

参考文献

- [1] Massey, K. (1997), Statistical models applied to the rating of sports teams, Bluefield College.
- [2] 黒木裕鷹・塩濱敬之:『Masseyのレーティング指標の統計的性質とそのネットワーク分析への応用』, 南山大学理工学部システム数理学科, 2022
- [3] Jリーグ公式サイト <https://www.jleague.jp/>
- [4] G.K. カンジ(原著), 池谷裕二(翻訳), 久我奈穂子(翻訳), 田栗正章(監修)(2009)「逆」引き 統計学 実践統計テスト 100, 講談社.