

2点差が本当に危険か

芝浦工業大学 数理科学研究会

辻村将吾

2018年5月20日

研究背景

私は、Jリーグの実況や解説陣あるいはサッカー日本代表の実況や解説陣の間で語られている「2点差が一番危険」について本当に2点差が危険なのかという疑問を抱いたため、このテーマにした。また、「危険」の定義は様々であるがここでは「2点差がついた時点で同点に追いつかれるまたは逆転される」という状況を危険という風に定義することにする。

1 研究の細かい説明

この研究では、「2点差がついた時点でのデータ」を10年分に調べて取っている。また、データは追いついた側を基準に見ている。その他注意としては

- 3点差以上からの同点や逆転はカウントしない(あくまで2点差がついた試合のみ)
- 点の取り合い(1-0⇒1-1⇒2-1⇒2-2など)の試合はカウントしない
- 今回扱うのは1部リーグのみ

また、今回の「2点差が本当に危険か」を検証するためにデータを比較することによって日本でだけ危険といわれているのかを調べていきたいと思っている

2 分析方法

以下では、研究の設定に基づいたデータを分析する方法を紹介する。

2.1 単回帰分析

1つの目的変数 y について1つの説明変数 x を使った式で表す分析方法を単回帰分析という。散布図のデータが関数関係

$$y = \beta_0 + \beta_1 x + \varepsilon$$

を表すとする。ただし、 β_0 は切片、 β_1 は傾き、 ε は誤差を表すとする。次に、推定値 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ を考える。実際の値 y_i と先ほど定義した \hat{y}_i との距離 $y_i - \hat{y}_i$ を残差といい、残差の二乗の和を最小にするような直線を求めることによって推定値を推測することができる。このような「2乗」したものが最小になるように定めることを最小2乗法と呼ぶ。実際にこの最小にす

る $\hat{\beta}_0$ と $\hat{\beta}_1$ を求めると、 $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$, $\hat{\beta}_0 = \bar{y} - \frac{S_{xy}}{S_{xx}}\bar{x}$ ただし、 $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$ $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ とする。

2.2 重回帰分析

先ほどの単回帰分析において、説明変数が複数ある場合に拡張することを考える。仮に、目的変数 y と2つの説明変数 x_1, x_2 があるとすると、つまり、 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ を考えたときに先ほどと同様に議論したときに、

$$\begin{pmatrix} \sum 1 & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{pmatrix}$$

を $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ について解いたものが最小にする $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ である。また、上の連立方程式のことを正規方程式と呼ぶ。

3 結果

Jリーグのデータにおいて $S_{xx} = 212.227$, $S_{xy} = 156.2699$ となりこの結果から重回帰分析を行うと

$$y = 0.142190x_1 + 0.00777x_2 + 1.234064x_4 + 0.083969x_5 + 3.522608$$

となる。ただし、 y :順位、 x_1 :2点差が付いた試合、 x_3 :同点に追いついた試合、 x_4 :逆転した試合、 x_5 :3点差以上付いた試合とする。

今後の課題

今後は、この結果を活かすことによってJリーグの優勝予想を立てていきたいと思っている。また、J1リーグのように一部リーグだけではなく日本代表の試合やJ2リーグの試合も含めて調査していきたい。実際に強さが均等なリーグのほど逆転する、または同点に追いつくことが難しいのかや海外のリーグのほうが同点に追いつくまたは逆転するのが難しいのかなども研究してみたいと思う。実際に研究をしてみた結果、寄与率などがかなり悪くなってしまったため線形回帰分析ではなく、非線形回帰分析によって分析すると結果が変わるのが興味深い。

参考文献

- [1] パーフェクト R, R サポーターズ, 技術評論社, 2017年。