

FIFA ワールドカップにおける勝因に関する統計的解析

2018SS039 丸山朋弘

指導教員：白石高章

1 はじめに

小学1年から高校3年まで12年間サッカーを続けてきて、サッカーに興味がある。その中でも、サッカー大会の最高峰に位置付けられ、4年に1度開催するFIFAワールドカップでの勝因が何なのか知りたいと考えた。本研究では勝敗に影響する要因を探るため、過去3大会分のデータ全48試合を使って統計的に分析を行った。

2 データ

本研究では、文献[1]と[2]を用いてワールドカップの試合のデータを集めた。集めたデータを用いて、ワールドカップで優勝するために必要な要因を考え、本文では18個の変数を用いて研究した。

x_1 :結果(勝ち:2, 引き分け:1, 負け:0とする)(ペナルティキック戦による決着は引き分けとする)

x_2 :ゴール数

x_3 :先制点(先制する:1, 先制しない:0)

x_{3+i} :15(i-1)~15i分の得点 ($i=1, \dots, 8$)

ただし x_6, x_9, x_{10}, x_{11} はロスタイムの得点を含む

x_{12} :枠内シュート数

x_{13} :枠外シュート数

x_{14} :コーナーキック数

x_{15} :ボール支配率(味方チームがボールに触れてから、敵に奪われるまでの時間の割合)

x_{16} :オフサイド数

x_{17} :イエローカード数

x_{18} :レッドカード数

3 分析方法

分析方法として、クラスター分析、ロジスティック回帰分析、因子分析、主成分分析(文献[3],[4])を用いた。ワールドカップでは、決勝トーナメント16試合を3大会分、合計48試合が行われ、96チーム分のデータを用いた。同じ国名のチームの試合がいくつかあるが、選手の疲労によるチームの状況や対戦相手によって戦い方を変えるため、試合ごとで分けて分析する。

4 クラスター分析

ワールドカップの過去3大会分の試合のデータを用いて、クラスター分析をウォード法で行った。

4.1 分析結果

図1において、距離20のところではクラスターを3つの群に分けると、左から1つ目の群は、106~120分での得点数が3つの群のどの時間帯よりも多く、試合終了間際の得点で勝率を上げている。しかし、シュートやコーナーな

どのチャンス数が少ないことやカードを多くもらっていることから勝率を伸ばしきれていない。2つ目の群は、ゴール数、先制点の要素が低いため勝率が低い。カード数が少ないことは勝率を高くする要因にも思えるが、相手に強気で当たるプレーができていないともとれる。3つ目の群は、ゴール数、先制点、支配率の重要視されやすい要素すべてが高いことに加えて、シュートやコーナーなどのチャンス数も高いため勝率が高い。オフサイド数も高いが敵陣地でプレーをする時間が多いからだと考えられる。

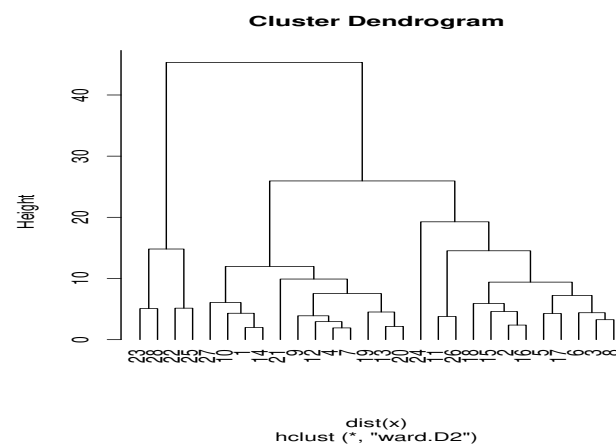


図1 デンドログラム(ウォード法)

4.2 考察

以上のことから、勝因としてカード数よりもシュートやコーナーなどのチャンス数の方が勝敗に関わり、更に先制点とゴール数が勝率を高めることに大きく関わっていることが分かる。また、オフサイドやカードをもらうこと自体はマイナス要素であるが、攻撃するときについてくるものでもあるので一概に悪い要素とは言えない。

5 ロジスティック回帰分析

x_1 を目的変数とし、 x_2 から x_{18} の変数をAICの値が最も改善されるように変数選択を行うことで、でた12個のデータを用いて分析を行った。

表1 変数減少後

変数	回帰係数	標準偏差	t値	p値
Intercept	0.128	0.115	1.111	0.270
x_2	0.838	0.310	2.706	0.008
x_3	0.494	0.102	4.832	6.11×10^{-6}
x_4	-0.754	0.361	-2.088	0.040
x_5	-0.824	0.328	-2.515	0.014
x_6	-0.795	0.326	-2.439	0.017
x_7	-0.658	0.364	-1.806	0.075
x_8	-0.644	0.301	-2.137	0.036
x_9	-0.752	0.310	-2.426	0.017
x_{10}	-0.627	0.351	-1.788	0.077
x_{11}	-0.768	0.357	-2.152	0.034
x_{13}	-0.021	0.011	-1.909	0.060
x_{14}	0.031	0.017	1.818	0.073

5.1 分析結果

12 個の変数の中で、他の変数と比較してゴール数は大きな係数となった。よって勝率を高める要因の変数であることが分かる。枠外シュート数は低い相関関係を示しており、チャンスをもものに出来ないことが勝率を下げる要因であることが分かる。逆にコーナーキック数は得点に繋がりがやすく、先制点の次に高い相関関係を示している。

5.2 考察

勝率を上げるためにゴール数を多くする。そのためにコーナーキックのような得点しやすい状況を作る必要がある。枠外シュートでは、チャンスを逃した上に相手ボールにしてしまうことから、得点からはより遠ざかってしまう。以上のことから、より多くの得点できる機会を作ったうえで得点をする、つまり決定力が勝率を上げるために必要な要素であると考えられる。

6 因子分析

表 2 因子負荷量の結果

変数	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
x ₁	0.186	0.484	0.303	0.221	0.092	0.103
x ₂	0.021	0.558	0.565	0.333	0.425	0.267
x ₃	0.077	0.562	0.194	0.142	0.107	0.018
x ₄	-0.273	0.929	-0.145	-0.174	-0.082	-0.018
x ₅	-0.036	0.232	0.107	0.083	0.958	0.063
x ₆	-0.029	-0.024	0.563	0.130	-0.036	-0.073
x ₇	-0.033	0.016	-0.003	0.994	0.054	-0.044
x ₈	-0.162	0.206	0.654	-0.048	0.115	-0.047
x ₉	0.017	0.094	0.002	-0.033	0.067	0.990
x ₁₀	0.562	0.180	0.167	-0.075	-0.092	-0.119
x ₁₁	0.323	-0.047	0.269	-0.130	-0.086	-0.148
x ₁₂	0.770	0.131	0.034	-0.054	0.123	0.079
x ₁₃	0.302	0.004	-0.043	0.120	-0.197	-0.051
x ₁₄	0.831	-0.052	-0.107	0.088	-0.020	0.139
x ₁₅	0.444	0.009	-0.094	0.010	0.029	0.005
x ₁₆	0.251	-0.191	-0.143	-0.116	0.019	0.035
x ₁₇	-0.091	-0.297	-0.003	0.058	-0.088	-0.035
x ₁₈	-0.007	-0.037	-0.226	0.081	-0.053	-0.117

因子負荷量の plot 図と分析結果を用いて考察を行った。factor1 は、91~105, 105~120, 枠内シュート, 枠外シュート, コーナー, 支配率, オフサイドが高い。91~105, 枠内シュート, コーナーは 0.5 を超えている。factor1 は、支配率を高めシュートを多く打ちチャンスをたくさん経験しておくことが延長戦での得点に繋がることが分かる。つまり相手のチームに慣れた試合であることが分かる。factor2 は、結果、先制、0~15 が高い。先制、ゴール、0~15 は 0.5 を超えている。factor2 は、先制しやすい時間帯が前半の序盤であることを示しており、先制がゴールを増やす要因となり、たくさんのゴールが結果を勝利に繋げる事を示している。つまり早い時間に先制点を取ることで勢いをつけ有利に進めた試合であることが分かる。factor3 は、ゴール数、31~45 の得点、61~75 の得点が 0.5 よりも大きな値が出ている。factor3 は、前半の終盤や後半の中盤に得点が多く入りやすいことを示している。factor4 は、46~60 の得点がとても高い値を出しており、後半の序盤での得点が得点数を伸ばす要因となり勝ちに繋がりがやすいことを示している。factor5 は、16~30 の得点がとても高い値を出しており、前半の中盤での得点が得点数を伸ばす要因となることを示しているが、勝敗には関わりにくい要因であることを示している。factor6 は、76~90 の得点がとても

高い値を出しており、91~105, 106~110 は負の値を出している。76~90 の得点で勝ち越している場合は勝利につながるが、引き分けて延長を行った場合は得点に繋がらないことを示している。

7 主成分分析

第 8 主成分までで累積寄与率 70 % を超えるので、寄与率の高かった 2 つの主成分について考察した。

7.1 分析結果

第 1 主成分 (寄与率 18.6 %) オフサイド数, イエローカード数, レッドカード数がマイナスの係数を取っているため、「試合中の攻守が切り替わる軸」とする。

第 2 主成分 (寄与率 14.8 %) 枠内シュート数, コーナーキック数が大きくマイナスの係数を取っているため、「試合中の得点に繋がる軸」とする。

第 3 主成分 (寄与率 8.5 %) 第 4 主成分 (寄与率 7.3 %)

第 5 主成分 (寄与率 7.1 %) 第 6 主成分 (寄与率 6.7 %)

第 7 主成分 (寄与率 5.5 %) 第 8 主成分 (寄与率 5.1 %)

7.2 考察

主成分分析からは、主成分の偏り方が均一ではなく傾向を見つけることが困難であった。

8 終わりに

本研究を通して FIFA ワールドカップの分析を進めてきた。枠外シュートは得点に繋がらないため敗北する要因になりやすいが、延長までいくと得点に繋がる要因になりうる、という関係性に驚いた。カードは敗北に一番関わってくる要素であるが、得点するために攻撃を重視した結果もらうことが多く、その攻撃が得点に繋がってれば勝利に近づけると考えられる。0~15 分の間の立ち上がりに先制点を取ることが、勝率を上げる一番の要因である。要因として、相手チームより先に得点をする、勢いをつけたまま試合を進められることや拮抗した試合になれば以降得点が動かないことが考えられる。次に、勝利に繋がりがやすいのが前半の終盤、後半の中盤での得点だが前半立ち上がりよりも得点を取りやすい。両チームが本領発揮した状態での得点であるため、試合結果として表れやすいと考えられる。これらのことを、FIFA ワールドカップ 2022 を観戦するときや今後プレーする際にも活かしていきたい。

参考文献

- [1] SoccerD.B.
<https://soccer-db.net/>
- [2] ワールドカップのデータベース
<https://worldcdb.com/>
- [3] 青木繁伸:『R による統計的解析』。オーム社出版、東京、2009。
- [4] 中村栄友:『R で学ぶデータサイエンス 2 多次元データ解析法』。共立出版、東京、2009。