

サッカーワールドカップにおける統計的分析

2015SS007 深田拓臣

指導教員：白石高章

1 はじめに

私は8歳から11年間サッカーをしていたため、4年に1度開催されるサッカーワールドカップに興味を持っていた。サッカーはとても人気なスポーツで、世界中の国で様々な大会が開催されている。その中でもワールドカップは世界中で注目される大会であり、日本では全試合がテレビで放送されている。私はサッカーの試合の勝敗を左右する要因を知ること、新しい角度からワールドカップを見られるのではないかと考えた。本研究では過去のワールドカップの試合のデータを用いて分析を行い、どのような要因が勝敗に影響するか統計的に分析を行う。

2 ワールドカップ分析にあたって

サッカーは敵のゴール内にボールを入れることで得点が得られるスポーツであるため、勝利にはボールを保持する時間が非常に大切になってくると考えられる。[1]によると2010年のワールドカップに出場した32チームのボール支配率上位16チームの内10チームが決勝トーナメントに進出している。しかし、同サイトによると4年後の2014ワールドカップでは上位16チームの内、7チームしか決勝トーナメント進出に進出できてない。これより2010と2014では、支配率の重要性が異なることが想定され、戦術にはブームが存在しサッカーは戦術変化への対応が求められるスポーツであると考えられる。そこで、2018ワールドカップではどのような要素が重要視されるのか、支配率以外にもいくつかの変数を用いて分析する。

3 データ

本研究では、ワールドカップの結果データを文献[2]と[3]を用いて集めた。このデータを用いて、ワールドカップ本大会でグループリーグを突破するための要因、優勝するための要因を考えた。本研究では主に2018FIFAワールドカップロシア大会に出場した32チームの試合結果を利用している。

本文では次の変数を活用した。 x_1 :1試合におけるシュート数、 x_2 :1試合におけるゴール数、 x_3 :試合の結果(勝利:2,引き分け:1,敗戦:0とする)(決勝トーナメントでのペナルティキック(以下PK)戦での決着は引き分けとする)、 x_4 :1試合における敵陣ペナルティエリア内侵入回数、 x_5 :1試合におけるボール支配率、 x_6 :1試合で出したイエローカード数、 x_7 :1試合で出したレッドカード数、 x_8 :1試合で獲得したコーナーキックの数、 x_9 :1試合で獲得したフリーキックの数、 x_{10} :1試合で獲得したPKの数、 x_{11} :1試合におけるデュエルの勝率、 x_{12} :1試合におけるインターセプトの数、 x_{13} :1試合における空中戦勝利率、 x_{14} :1試合にオフサイド

になった数。

4 分析方法

分析方法として数量化2類、ロジスティック回帰、因子分析を行った。分析は[4],[5]を参考にした。ワールドカップでは、グループリーグの48試合、決勝トーナメントの16試合の合計64試合が行われる。そのため128チームのデータを利用している。また、同じチームの試合がいくつかあるが、対戦相手やグループリーグでのチーム状況に依って戦い方を変えるため、試合ごとに分けて分析する。

5 数量化2類

目的変数を x_3 とした数量化2類をRで実行した。

表1 変数の判別係数

変数	判別係数	負け平均	勝ち平均
シュート	0.017	11.569	13.317
ゴール	0.951	0.490	1.870
ペナ内侵入	-0.019	19.294	21.571
支配率	-0.014	48.510	50.987
イエロー	-0.060	1.941	1.584
レッド	-1.169	0.059	0.013
CK	0.025	4.431	4.883
FK	-0.047	14.314	14.792
PK	-0.421	0.157	0.273
デュエル勝率	0.138	47.686	51.532
インターセプト	-0.019	10.824	10.247
空中戦勝利率	-0.023	48.000	51.326
オフサイド	-0.129	1.392	1.317

表2 判別結果総括表

	負け	勝ち
負け	40	11
勝ち	10	67

「ゴール」の判別係数が0.951になり、他の変数と比較すると圧倒的に大きくなった。1試合の中でもゴールの回数は多くないため、いかに多くの得点を取るかどうかで試合結果が決まると言える。「レッドカード」の判別係数は-1.169と大きな負の値になった。レッドカードが出る試合は多くないが、レッドカードを受けてしまったチームは大きな不利を受けることになる。「デュエル勝率」には正の相関がある。判別係数は0.138となった。デュエル勝率はそのものの値が大きくなるため、この判別係数はとても大きな値だと考えられる。

6 ロジスティック回帰分析

AIC の値が最も改善されるように変数を削除をする変数減少法で変数選択を行った。目的変数は x_3 とした。

表 3 変数減少後

変数	回帰係数	標準偏差	p 値
Intercept	-0.89	0.322	0.007
ゴール	0.22	0.03	2.42×10^{-11}
レッド	-0.28	0.195	0.16
FK	-0.01	0.008	0.17
デュエル勝率	0.03	0.009	0.0003
空中戦勝利率	0.01	0.004	0.10

「ゴール」、「レッド」、「FK」、「デュエル勝率」、「空中戦勝利率」の 5 つが残った。その中でも「ゴール」は他の変数と比較しても大きな係数となった。よって勝利するのに最も重要な変数であることがわかる。「デュエル勝率」には正の相関がある。ゴールほどではないが比較的係数が大きい変数であり、引き分け以上の結果を得るために重要な変数であることがわかる。「空中戦勝利率」は正の相関が出ている。空中戦はデュエルと似ていて選手間での争い事であるのでこのような結果が得られたと考えられる。

7 デュエルの重要性

説明変数のなかで試合中のプレーとなる変数は「シュート」、「デュエル」、「インターセプト」、「空中戦勝利率」、「オフサイド」の 5 変数である。そして、ロジスティック回帰分析で取捨選択して残った変数とプレーの変数と共通する変数は「デュエル勝率」と「空中戦勝利率」の 2 つであった。「空中戦勝利率」は、重回帰分析とロジスティック回帰の 2 つの分析で回帰係数の符号が異なるため、勝敗を分ける重要な変数とは言えないと考えられる。よって良い試合結果を得るために必要となるプレーは 1 つ 1 つのデュエルの場面で勝つことだと言える。近年世界中で「デュエル」という言葉が使われ、日本代表メンバーでもデュエルに強い選手が招集されるようになった。デュエルの重要性が世界中で証明されてきている。

8 因子分析

factor1 では、シュート、ペナルティエリア内侵入、支配率、コーナーキックが 0.5 よりも大きな値が出ている。factor1 は、ボール支配からシュートまでの回数、コーナーキックの回数と得点のチャンスとなるシーンを示している。factor2 は、ゴールと勝ち負けが大きな値を示しているため、試合の結果を示している。factor3 は、デュエル勝率と空中戦勝利率が高い値を表している。factor3 は、フィジカルの強さを示していることがわかる。factor4 は、フリーキックとデュエル勝利率が高い値を示している。デュエルでは 1 対 1 での争い場面でありファールになること

も多いシーンである。自分がファールを受けて自分ボールになったときのデュエルは勝利扱いとなるため、factor4 はファールを受けたりデュエルで勝ったりと、自分のボールで始められる場面の多さを示している。また、小さな値ではあるがイエローカードが -0.29 と負の値を示しているのは、ファールの際にイエローカードをもらうことがあるためである。factor5 では、PK のみが突出して高い値となっている。factor6 では、支配率が大きな値を示しており、インターセプトが負の値を示している。ボールを保持する時間が長いほどインターセプトをする機会が少ない事が要因として挙げられる。

表 4 因子分析結果

変数	factor1	factor2	factor3	factor4	factor5	factor6
シュート	0.905	0.104		0.117		
ゴール		0.645			0.184	
勝ち負け		0.970		0.200		
ペナ内侵入	0.870		0.117	0.119		
支配率	0.543			-0.309		0.516
イエロー	-0.179	-0.126		-0.290	-0.116	
レッド		-0.124				0.170
CK	0.723		0.118		0.208	
FK				0.445		0.226
PK		0.142			0.981	
デュエル勝率	0.237	0.185	0.543	0.777		
インターセプト					-0.115	-0.550
空中戦勝利率	0.177		0.970			
オフサイド						0.283

9 終わりに

本研究を通してサッカーワールドカップの分析を進めてきた。その中で解析結果を求めることは想像以上に困難なものであった。しかし本研究を通してサッカーの試合の味方を変えることが出来た。シュートやゴールのシーンだけでなく、普段は見られにくいデュエルや空中戦といったフィジカル面でのプレーも重要視する必要があることが分かった。今回のこの研究より重要視される変数に着目して、今後のプレーする際に活かしたい。

参考文献

- [1] フットボール ラボ ,
<http://www.football-lab.jp/column/entry/537> .
- [2] ワールドカップのデータベース ,
<https://worldcdb.com/>
2018 年 8 月閲覧.
- [3] 朝日新聞 DIGITAL 2018 サッカーワールドカップ,
<https://www.asahi.com/worldcup/schedule/>
2018 年 8 月閲覧.
- [4] 青木繁伸:『R による統計的解析』. オーム社出版,
東京, 2009.
- [5] 中村永友:『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.