

# レアルマドリードの UEFA チャンピオンズリーグにおける 勝因に関する統計的分析

2016SS025 川西健太

指導教員：白石高章

## 1 はじめに

私は、幼少期からサッカーをしていて、サッカー観戦にも熱狂している。その中で、毎年 UEFA チャンピオンズリーグを観戦していて、私はレアルマドリードをよく応援している。そこで、過去4年で3度の優勝をしているレアルマドリードについて、その強さの要因はなにか、どのようなことが試合の勝因に大きく関与しているのか知りたいと思った。サッカーは90分間の中で目まぐるしく展開が変わり勝敗が決まる。試合の経過時間や時間帯ごとの得失点など、様々なことが関わっていると思われる。本研究ではある勝敗に何が影響しているのかを探るために、過去4年間のレアルマドリードのデータ全47試合で統計的に分析を行った。

## 2 データについて

本研究では、レアルマドリードの過去4年間のチャンピオンズリーグの試合のデータ(文献[1])を使用した。変数としてこれらを使用する。

- x1 : 勝敗 (0,1:引き分けは負けとする)
- x2 : 0~15 分の得点
- x3 : 16~30 分の得点
- x4 : 31~45 分の得点 (ロスタイムによる得点含む)
- x5 : 46~60 分の得点
- x6 : 61~75 分の得点
- x7 : 76~90 分の得点 (ロスタイムによる得点含む)
- x8 : 延長戦での得点
- x9 : 先制点 (0,1)
- x10 : home and away(0,1)
- x11 : シュート数
- x12 : 枠内シュート数
- x13 : ファウル数
- x14 : オフサイド数
- x15 : 前半リード (0,1)
- x16 : ゲーム支配率
- x17 : コーナーキック数
- x18 : パス数
- x19 : 逆転 (0,1)
- x20 : 0~15 分の失点
- x21 : 16~30 分の失点
- x22 : 31~45 分の失点 (ロスタイムによる失点含む)
- x23 : 46~60 分の失点
- x24 : 61~75 分の失点
- x25 : 76~90 分の失点 (ロスタイムによる失点含む)

## 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(文献[2],[3],[4])を用いた。

## 4 クラスタ分析

レアルマドリードの過去4年の UEFA チャンピオンズリーグの試合のデータを用いて、クラスター分析を標準化をしてウォード法で行った。

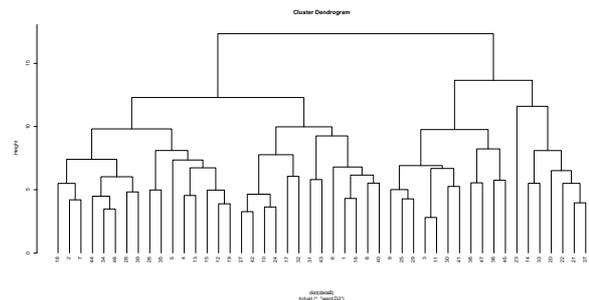


図1 デンドログラム(ウォード法)

### 4.1 分析結果

図1で、距離の10のところではクラスターを大きく5つに分けてみると、左から1つ目の群は、試合立ち上がりでの失点がなく、先制点、または前半リードしている試合が多く、そういった試合では勝率は8割を超える。2つ目の群は、パス数が比較的多く、セットプレーも多いことからシュート数も多く、勝ちにつながっている。3つ目の群は、先制点も取れず前半リードできていない試合では引き分けか負けになっており、パス数も少ないことから自分たちのリズムでサッカーができていない。4つ目の群は、唯一の延長戦であり、シュート数も多く5得点を挙げており、終始攻めのサッカーをしている。5つ目の群は、逆転している試合群であり、先制点は取れていないが点の取り合いとなっており、レアルマドリードは接戦をものにする勝負強さがある。

### 4.2 考察

以上のことから、レアルマドリードの勝因として、先制点、さらに終盤での失点を避け、最後まで集中力が大事になっており、終始攻め続けるのがチームの特徴である。試合におけるゲーム支配率が高いほど、シュート数も増え、落ち着いて自分たちのサッカーができ、得点に絡むプレーが多くでることから持ち味の得点力が活かせるので、勝ちにつながっていると考えられる。

## 5 ロジスティック回帰分析

$x_1$  を目的変数とし、その他の  $x_2$  から  $x_{25}$  の変数を AIC の値が最も改善される変数選択を行い、でてきた 6 つのデータを用いて分析を行った。

表 1 ロジスティック回帰分析の結果

変数	回帰係数	標準偏差	t 値	p 値
$x_7$	0.210	0.099	2.121	0.040
$x_9$	0.759	0.096	7.944	0.000
$x_{18}$	0.196	0.086	2.276	0.028
$x_{19}$	0.658	0.104	6.316	0.000
$x_{22}$	-0.082	0.092	-2.707	0.010
$x_{23}$	-0.094	0.115	-1.944	0.059

### 5.1 分析結果

表の結果から、「先制点」が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。また「76~90 分の得点」と「逆転」も回帰式に大きな影響を与えていることから、リアルマドリードは攻撃的サッカースタイルだといえる。試合中盤による失点についての変数が低い相関関係を示しているのも、リアルマドリードは失点してからの切り替えが早く、試合の終盤に得点が生まれる強く粘り強い戦い方をしていると考えられる。チーム全体が最後まで試合をあきらめないという精神的強さも読み取れる。これらのことがリアルマドリードの勝因にも繋がっており、勝率の高さを物語っている。また、パス数の回帰係数が低い値を取っているが、これはほかの変数に対してデータの数が多いためである。実際、P 値も低い値を取っている。リアルマドリードが試合に勝つ要因は、前半の終盤と後半の立ち上がりの失点を 0 に抑え、先制点を取ることである。つまり、試合の立ち上がりでの得点、また、試合の終盤における集中力がリアルマドリードの特徴ではないかと考える。今後リアルマドリードの試合を観戦するときには、これらの点を意識することで今までと違った見方ができると感じた。

## 6 因子分析

分析結果と因子負荷量の plot 図を用いて考察を行った。factor1 では、勝敗と先制点が高い値を取っている、また、0~15 分の失点が負の高い値を取っていることから、試合において 1 点目を取る取られるが勝敗を大きく分けることを示している。factor2 では、ゲーム支配率とパス数が高い値を取っており、ファウル数が負の値を取っている。ボールを保持する時間が長いほどファウルする機会が少ないことが要因としてあげられる。factor3 では、シュート数、枠内シュート数が高い値を取っていることから、得点シーンを示していることが分かる。factor4 では、逆転だけが低い値を取っており、先制点が負の値を取っているため、逆転をした試合において先制点は取れず取られている傾向にあることを示している。factor5 では、試合の立ち上がり

表 2 因子負荷量の結果

変数	factor1	factor2	factor3	factor4	factor5	factor6	factor7	factor8	factor9	factor10
$x_1$	0.816		0.104	0.189	0.103			0.231		
$x_2$		0.134			0.663					0.125
$x_3$	0.291			-0.165		-0.116	0.667		0.103	
$x_4$			0.207				0.569			-0.171
$x_5$	0.117		0.102				0.100	0.975		
$x_6$	0.300		0.183	-0.293		-0.214				
$x_7$	0.303			0.299		0.507	-0.113	-0.174	-0.133	
$x_8$			0.219	0.186	-0.180	0.234	-0.168		0.620	0.153
$x_9$	0.646		0.187	-0.553	0.318		0.198	0.100		
$x_{10}$			0.222		0.208					0.914
$x_{11}$		0.356	0.687						0.182	0.186
$x_{12}$	0.120		0.916		0.171	0.170	0.215	0.152		0.125
$x_{13}$	0.206	-0.280			0.119	-0.391		-0.179		-0.193
$x_{14}$	0.191		-0.337	0.343	-0.401	0.280	0.233	0.126	-0.191	0.169
$x_{15}$	0.365			-0.243	0.654	-0.110	0.574		0.149	
$x_{16}$		0.908		-0.128					0.121	
$x_{17}$		0.274	0.127	-0.117		0.125	0.236	-0.323	-0.102	0.274
$x_{18}$	0.180	0.930	0.122		0.227	0.123				
$x_{19}$	0.128	-0.245		0.892	-0.226	0.149		0.133	0.110	
$x_{20}$	-0.516		-0.114					0.103		0.273
$x_{21}$	-0.234			0.247	-0.367	-0.231		-0.108	-0.160	-0.140
$x_{22}$	-0.243				0.183	0.223		-0.226		
$x_{23}$					0.100	0.870	-0.107	-0.107	0.163	
$x_{24}$	-0.408		0.103				-0.203		-0.145	
$x_{25}$					0.130				0.761	-0.127

での得点が勝ちにつながることを示している。factor6 では、後半立ち上がりの失点と終盤での得点が高い値を取っている。ファウル数が負の値を取っているのは、失点してから得点を取りにいかないといけないうえ、前がかりになる際にファウルをしてしまうシーンを示している。factor7 は立ち上がり以降の前半の結果を示している。factor8 では、後半立ち上がりの得点のみが突出して高い値となっている。factor9 では、試合において終盤での得点の動きを示している。factor10 では、ホーム&アウェイが突出して高い値を取っており、リアルマドリードのホームでの強さを示している。

## 7 おわりに

本研究を通してリアルマドリードの UEFA チャンピオンズリーグの分析を進めてきた。リアルマドリードの勝敗に対して分岐点となったのは先制点と終盤での試合の進め方であり、ここまで大きく影響しているとは思わなかった。また、シュートやゴールのシーンだけでなく、得失点の時間帯、ボールのポゼッションやホームの重要性など、普段は見られにくい点も重要視する必要があることが分かった。これらのことを今後プレーする際にも活かしたい。また、昨シーズンから絶対的エースであるクリスティアーノ・ロナウドが退団したことで持ち味の得点力も低下していると予測できるので、本研究で得られた結果が昨年以降のリアルマドリードでも有効であるか調べたい。

## 参考文献

- [1] SoccerD.B レアルマドリード  
<https://soccer-db.net/>
- [2] [R によるデータサイエンス] クラスター分析  
<https://financial-it-engineer.hatenablog.com/>
- [3] 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
- [4] 青木繁伸: 『R による統計的解析』, オーム社, 2009 年