

# グランパスのサッカーの勝因に関する統計的分析

2010SE019 藤野翔太

指導教員：木村美善

## 1 はじめに

私は、Jリーグの名古屋グランパスをよく応援している。サッカーの試合を観戦するにあたり、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。それを知ることで今まで漠然と見ていた試合が、より見応えのあるものになると考えた。また、あまり試合を見ない人でもどのような点が勝因に関わっているかを知ることが出来れば、楽しくサッカーの試合を見ることが出来るのではと考えた。では、どのようなことが勝因に関わってくるのだろうか。サッカーは90分間で勝敗が決まる。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。本研究では勝敗に何が影響しているかを探るために2012年と2013年のJリーグの名古屋グランパスのデータ全68試合で判別分析、主成分分析、クラスター分析を行った。

## 2 データについて

Jリーグの公式HPや名古屋グランパスのHPの情報から作成した2012年のJリーグの名古屋グランパスの対戦データを用いる([2], [3]参照)。また判別式の有効性を調べるために、2011年のJリーグの5試合を使用した。変数として $x_1$ (勝敗)(ただし、今回の研究では勝因を知りたいので引き分けは負けとして分析した)、 $x_2$ (0~15分の得点)、 $x_3$ (16~30分の得点)、 $x_4$ (31~45分の得点)、 $x_5$ (46~60分の得点)、 $x_6$ (61~75分の得点)、 $x_7$ (76~90分の得点)、 $x_8$ (先制点)、 $x_9$ (前半リード)、 $x_{10}$ (home and away)、 $x_{11}$ (シュート数)、 $x_{12}$ (コーナーキック数)、 $x_{13}$ (フリーキック数)、 $x_{14}$ (0~15分の失点)、 $x_{15}$ (16~30分の失点)、 $x_{16}$ (31~45分の失点)、 $x_{17}$ (46~60分の失点)、 $x_{18}$ (61~75分の失点)、 $x_{19}$ (76~90分の失点)、また先制点を取った時間帯の変数 $x_{20}$ (0~15分での先制点)、 $x_{21}$ (16~30分での先制点)、 $x_{22}$ (31~45分での先制点)、 $x_{23}$ (46~60分での先制点)、 $x_{24}$ (61~75分での先制点)、 $x_{25}$ (76~90分での先制点)を使用している。

## 3 判別分析

得られる判別式の有効性を調べるためにランダムに5試合取り除き、全63試合のデータを使い分析した(8, 17, 22, 35, 60試合目を取り除く)。 $x_1$ (勝敗)を目的変数 $y$ とし、 $x_2, \dots, x_{25}$ を説明変数として分析を行った。変数選択の過程で変数間に相関が見つかったため、不要な変数を除いた( $x_9, x_{16}, x_{20}, x_{22}$ の前半リード、31~45分の失点、0~15分での先制点、31~45分での先制点を取り除いた)。外れ値は無く、残差プロット図やテコ比とクックの距離も考慮したが、特に大きく外れているものが無かつ

表1 判別分析結果

変数	係数	平均(負)	平均(勝)
$x_2$	1.018	0.111	0.373
$x_3$	0.000	0.166	0.185
$x_4$	1.104	0.027	0.333
$x_5$	0.469	0.250	0.296
$x_6$	1.971	0.111	0.407
$x_7$	0.968	0.277	0.518
$x_8$	1.303	0.333	0.888
$x_{10}$	0.465	0.333	0.629
$x_{11}$	-0.042	10.30	10.59
$x_{12}$	-0.015	4.444	3.703
$x_{13}$	-0.026	13.66	12.29
$x_{14}$	-1.690	0.166	0.037
$x_{15}$	-1.094	0.055	0.037
$x_{17}$	-0.788	0.361	0.037
$x_{18}$	-0.894	0.444	0.074
$x_{19}$	-0.617	0.444	0.333
$x_{21}$	0.731	0.055	0.111
$x_{23}$	-0.731	0.083	0.074
$x_{24}$	-1.059	0.055	0.037
$x_{25}$	-0.381	0.027	0.074

た。したがって、このまま判別分析を行うと次の表1の結果を得る。表1は、判別式の係数と勝った試合と負けた試合の変数毎の平均である。分析結果より、平均の中間から判別式に必要な判別点:0.719が求められた。判別式 $z = 1.018x_2 + \dots - 0.381x_{25} - 0.719$ の値が正であると勝ち、負であると負けと判別される。表1の係数より、先制点と61~75分の得点、0~15分の失点が大きく影響していることが分かる。その他にも0~15分の得点、31~45分の得点、16~30分の失点も影響力がある。判別式の有効性を調べるため、あらかじめ取り除いていた5試合と2011年のJリーグの5試合、全10試合のデータを判別式 $z = 1.018x_2 + \dots - 0.381x_{25} - 0.719$ に代入し、判別を行った。その結果、誤判別された試合はなかった。したがって、この判別式は有効であると思われる。

### 3.1 考察

判別分析より、先制点、前半45分までの得失点が大きく試合を分けていると言える。名古屋グランパスが試合に勝つ確率を上げる要因は、0~30分までの失点を抑えた上で、先制点を取ることであると言える。判別された試合の中で引き分けだが、勝ちと判断された試合がある。これは名古屋グランパスが試合終了間際に先制点を決めているに

も関わらずその後失点し、引き分けているからであると考えられる。判別分析の結果は、私が予想していたものとは違っていた。なぜなら、後半終了間際の76分から90分での失点あまり影響がないからだ。これは名古屋グランパスが試合に勝つときは、先行逃げ切りというチームスタイルのような点でこのような分析結果になったのであろうと予測出来る。名古屋グランパスの試合を観戦する時は、これらの点を意識して見る事で今までと違った見方が出来るのではないだろうか。

## 4 主成分分析

判別分析で使用したデータを使い主成分分析を行った。第6主成分までを分析してみたところ、累積寄与率は約54%であった。主成分得点より第1主成分は勝敗、前半での得点、失点、先制点が多いことより、前半での経過が勝敗に大きく影響している試合を表す軸であると考えられる。第2主成分は得点とシュート数、コーナーキックが多いことより、チャンスの場面での得点した試合を表す軸であると考えられる。プロット図より、左下の群は、先制点を取っているが、46~60分での失点が多いと考えられることより、勝ち負けが混在している。右下の群は、先制点が取れず失点が多いと考えられことより、負けた試合が群れている。左上の群は、早い時間帯で先制点を取ることができ、試合を優位に運んでいると考えられることより、勝った試合が群れている。右上の群は、先制点を取られたが、早い時間帯で追いつき、また離されるといった苦しい展開であると考えられることにより、負け試合が群れている。

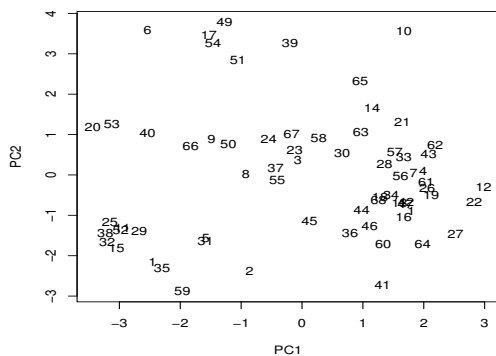


図1 第1主成分得点と第2主成分得点

### 4.1 クラスタ分析

主成分分析から第6主成分までの主成分得点を用いてクラスタ分析を最近隣法、最遠隣法、重心法、メディアン法、加重平均法、ウォード法、可変法で行った。その中で綺麗に群分けできたウォード法で分析をした。

図2で距離の85のところではクラスタを大きく4つに分けて見ると、左から1つ目の群は、失点が無く、多く得点している試合が見られる。2つ目の群は、早い時間帯で先制点を取っている試合が多く見られる。3つ目は、先

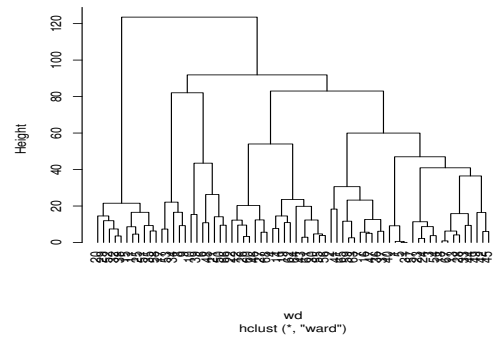


図2 デンドログラム (ウォード法)

制点を取られ、負けもしくは引き分けている試合が多く見られる。4つ目は、失点も少なく得点も少ない試合が多く見られる。

### 4.2 考察

主成分分析とクラスタ分析により、試合の分類をすることが出来た。それにより、先制点を取ることが試合の分岐点になりうる1番大きな要因であると考えられる。しかし、今回の主成分は、1つ1つの寄与率が低く情報量が少ないということも考慮しなくてはならない。

## 5 おわりに

本研究で2012年と2013年の名古屋グランパスの試合の分岐点になったのは、前半での得失点や先制点であり、これほど大きく影響しているとは思わなかった。また、61~75分での得点の影響力が大きいのは、もし、2点差で負けていても75分までに1点差に追いつくことが出来たら試合終了までモチベーションを保つことが出来るからだと考える。今回、この研究をしていることで少なくとも普段のサッカーを見る視点が変わってきた。現在行われている名古屋グランパスの試合も先制点や前半での得点などを意識して見てもらいたい。しかし、2013年のJリーグで名古屋グランパスの監督が交代してしまった。サッカーの戦術は、監督の采配が大きく影響する。戦術が変われば、先行逃げ切りのチームスタイルも変わってしまうと予測できる。本研究で得られた結果が2014年以降のグランパスの試合でも有効であるか調べてみたい。

## 参考文献

- [1] 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，2009。
- [2] 名古屋グランパス公式サイト  
<http://nagoya-grampus.jp/>
- [3] Jリーグ公式サイト  
<http://www.j-league.or.jp/>
- [4] 西尾裕：J1リーグ優勝に関する統計的分析-南山大学数理工学部情報システム数理学科卒業論文，2010。