

サッカーの国対抗試合における勝敗の分析

2010SE119 美谷 亮

指導教員：白石 高章

1 はじめに

本研究では、2014 年に行われる第 20 回サッカー W 杯ブラジル大会での国対抗試合における勝敗の分析をすることを考える。日本が出場した 1998 年フランス大会、2002 年日本韓国大会、2006 年ドイツ大会、2010 年南アフリカ大会の直近過去 4 大会の W 杯の傾向と強豪国が多く参加した 2013 年コンフェデレーションズ杯のデータをもとに解析していく。

2 データについて

W 杯での日本が出場した、直近過去 4 大会のデータを Website[1] を介して集めた。また、2013 年コンフェデレーションズ杯での日本の全 3 試合の DVD を見て、データを集めた。次のように、経過時間について時間帯を区切り、表 2 を作成した。

前半 1: 0 ~ 15 分 2: 16 ~ 30 分 3: 31 ~ 45 分
後半 4: 46 ~ 60 分 5: 61 ~ 75 分 6: 76 ~ 90 分
なお、前半のロスタイムは 45 分に、後半のロスタイムは 90 分に含む。延長戦 7: 91 ~ 105 分 8: 106 ~ 120 分
また、DVD を見てデータを取る際に次の点について着目した。得点、失点、シュート数、枠内シュート数、フリーキック数、ファール数、カード枚数、オフサイド数、スローイン数、ゴールキック数、PK、交代数、パス数、パス成功率、ヘディング勝利数、ヘディングに勝利後に味方にパスが繋がった数、センタリング数、センタリング後に味方に繋がった数について着目して、データを集めた。なおパス数に関しては、1 本以上繋がったものをカウントし、1 本も繋がっていないものはカウントをしないものとする。

3 データ

3.1 W 杯での初戦の重要性

表 1 は、W 杯直近過去 4 大会における初戦の結果によるグループリーグ突破の割合を示したデータである。

表 1 W 杯での初戦の勝敗とグループリーグの結果について

試合数	勝ち	引分	負け	合計
初戦	46	36	46	128
GL 突破	39	21	4	64
割合	0.848	0.583	0.087	0.500

3.2 W 杯直近過去 4 大会における、全得点時間帯

表 2 は、日本の出場した W 杯直近過去 4 大会における全 256 試合においての、全得点である 624 得点の時間帯を

勝ち、引き分け、負けのチームに分類したデータである。

表 2 W 杯直近過去 4 大会における全得点時間帯

経過時間帯	1	2	3	4	5	6	7	8
勝ち	56	55	63	60	68	109	4	5
引き分け	18	17	17	28	16	26	0	0
負け	11	14	16	13	8	20	0	0
合計	85	86	96	101	92	155	4	5

3.3 W 杯直近過去 4 大会における、先制点での勝敗について

表 3 は、日本の出場した W 杯直近過去 4 大会における全 256 試合においての、先制点を得点したチームを勝ち、引き分け、負けに分類し、各割合を示したデータである。

表 3 W 杯直近過去 4 大会における、先制点での勝敗について

勝敗	勝ち	引分	負け	合計
試合数	159	47	25	231
割合	0.688	0.203	0.108	1.000

スコアレスドローが 25 試合。

4 データ解析結果

4.1 W 杯での初戦の重要性

初戦の結果で勝ちを収めることで、グループリーグを突破することに大きく近づくことができると考えられる。また、初戦で負けてしまった場合にグループリーグを突破できる確率は、とても小さくなってしまふことから、初戦が重要であると考えられる。

まず、W 杯直近過去 4 大会における、初戦での勝敗の結果によるグループリーグ突破に関しての解析を、表 1 をもとに行った。また、3 戦目の結果は 47 勝したチームがありその中でグループリーグ突破したのは 36 チームであった。帰無仮説は、初戦を勝ったチームと 3 戦目に勝ったチームでは、グループリーグを突破することに差がないとする。白石 [3] より、

$$T \equiv \frac{2\{\arcsin(\sqrt{\hat{p}_1}) - \arcsin(\sqrt{\hat{p}_2})\}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
 の式に、

$n_1 = 46$, $n_2 = 47$, $X = 39$, $Y = 36$, $\hat{p}_1 = 0.848$, $\hat{p}_2 = 0.766$, $\sqrt{\hat{p}_1} = 0.920$, $\sqrt{\hat{p}_2} = 0.875$

を代入し考える。計算をする中で、

$$\arcsin(\sqrt{\hat{p}_1}) = 1.168, \arcsin(\sqrt{\hat{p}_2}) = 1.065$$

で進めていくと、 $T = 0.995$ となり、棄却されないの、差がないという結果が出た。また、信頼区間 $\alpha = 0.05$ とし

，保守的な信頼区間

$$\frac{L}{KF_L^K(\alpha/2) + L} < p < \frac{K^*F_{L^*}^{K^*}(\alpha/2)}{K^*F_{L^*}^{K^*}(\alpha/2) + L^*}$$

において，初戦に勝つことでグループリーグを突破する信頼区間を求めた．結果として， $0.711 < p < 0.937$ となり，95% 信頼区間において，71% 以上の確率で突破できることが考えられる．この2つのデータの結果によって，W 杯初戦を勝ったチームでは，グループリーグでの初戦が重要ということが分かる．

4.2 W 杯直近過去4大会における，全得点時間帯

過去4大会の全得点数は624得点である．その中でも，得点が圧倒的に多かった時間帯は76-90分で，得点数は155得点である．また，そのうちの109得点は勝ったチームが得点している．このことから，残り15分という時間帯は，真の実力が試される時間帯である．それまでの時間はディフェンスの集中力が保たれていても，残り15分で集中力が切れる，または体力が無くなることがある．よって，勝つチームはこの残り15分で得点することで，勝ちを収めることが言えると考えられる．

以下より，藤井 [2] を参照し R を使用し解析を行った．表2のW 杯直近過去4大会における全得点時間帯のデータの解析をしていく．表2での得点時間帯と勝ち，引き分け，負けにおいてカイ二乗検定を行った．結果は，カイ二乗検定量 9.3094，自由度 10，p 値 0.503 である．よって，p 値が有意水準 0.05 よりも大きいので帰無仮説は棄却できない．この検定の結果より，得点時間帯におけるの勝敗については関係がないということになる．さらに，表2での時間帯別での合計得点において，得点が入る時間帯が一樣であるかについての適合度検定を行った．結果は，カイ二乗検定量 34.0439，自由度 5，p 値 2.334e-06 である．2.334e-06 は非常に小さい数値となり，p 値が有意水準 0.05 よりも小さいので，帰無仮説は棄却される．この検定結果より，時間帯別での得点は一樣でないという結果となった．また，試合終了直前の76-90分の時間帯が最も得点が多い．

4.3 W 杯直近過去4大会における，先制点での勝敗について

先制点を得点したチームは，159勝25敗している．このことから，先制点を得点したチームは試合の流れを自分たちのペースで作ることができ，良い流れで試合をコントロールすることができる．よって，勝利に大きく近づくことができると考えられる．

次に，表3を解析していく．先制点を入れたときの勝つ確率と，先制点を入れたときの勝つか引き分ける確率についての信頼区間を求める．勝つ確率は，正確な保守的な信頼区間では $0.6043 < p < 0.7672$ となり漸近的な信頼区間は $0.6095 < p < 0.7665$ である．この2つの信頼区間は極めて近い．また，勝つか引き分ける確率は正確な保守的

な信頼区間では $0.8286 < p < 0.938$ と漸近的な信頼区間は $0.8385 < p < 0.9435$ である．よって，先制点を入れたチームが勝つ確率は60%以上である．また，勝つか引き分ける確率は80%以上となる．この結果より，先制点を取ることによって引き分け以上になる確率が高くなるので，先制点を与えることなく，先制点を取るようすることで勝利に大きく近づくことができる．

次は，本稿の表4.21と表4.22に2013年コンフェデレーションズ杯でのグループリーグの3試合におけるデータが記載されている．ここでそのデータの中で，パスの本数について解析していく．それぞれのデータにおいて，時間帯別のパスの本数について一樣であるという帰無仮説についての適合度検定を行った．

結果として日本の場合では，カイ二乗検定量 47.2215，自由度 5，p 値 5.12e-09 である．5.12e-09 は，非常に小さな数値となるので，p 値が有意水準 0.05 よりも小さいので帰無仮説は棄却される．また，日本と対戦した3チームでの場合では，カイ二乗検定量 57.9811，自由度 5，p 値 3.174e-11 である．3.174e-11 も，非常に小さな数値となるので，p 値が有意水準 0.05 よりも小さいので帰無仮説は棄却される．この検定結果より，時間帯別のパスの本数は一樣ではないという結果となった．これは，日本や他のチームにおいても，1試合を通してパスを安定して回すことで，試合を支配することはできないという結果である．なので，日本が今大会において，強豪国と試合をしていくなかで，日本のペースの時間帯があることが考えられる．

5 おわりに

本研究を通して，サッカーの国対抗試合における勝敗の分析を進めてきた．その中で考えていた以上に解析結果を求めることが困難であった．ここで，解析結果から考察をまとめた．まず，W 杯初戦に勝利することで大きくグループリーグ突破に繋がることが分かった．そして，そのW 杯初戦に勝利するためには，先制点を取ることで勝利に大きく近づくことが重要であることも分かった．DVD からのデータでは，日本は前半の終わりに近づいた31-45分にパス数が大きく減っていることから，相手の時間帯になることが分かる．そして，試合の終わりに近づくとつれて，日本のパスの本数が増えていることから，試合の終盤では試合を支配することができていることがわかる．この時間帯まで失点することなく試合を進めることができれば，日本が勝利するチャンスが増えてくるのではないかと考えられる．

参考文献

- [1] ワールドカップのデータベース【WorldCup's world】
<http://members.jcom.home.ne.jp/wcup/>
- [2] 藤井 良宜：『R で学ぶデータサイエンス カテゴリーカル データ解析』 共立出版
- [3] 白石 高章：『統計科学の基礎-データと確率の結びつきがよくわかる数理』 日本評論社