

1. イントロダクション

2014年にはブラジル、2010年には南アフリカで開催されたFIFAワールドカップ（以下「W杯」）。サッカーの大会としては最大級であり、その全世界における視聴率や経済波及効果は、スポーツの祭典として名高いオリンピックを凌ぐものがある。国のトップチーム同士のレベルの高い戦いは、世界各国の国民を熱狂させ、同時にプレイヤーも挑戦心をかき立てられる。直近の2014年ブラジル大会では、米大手IT企業のFacebookやTwitterが自社プラットフォームにて「WorldCup Trend」（Facebook）や「WorldCup Timeline」（Twitter）といったコンテンツをリリースしたのも記憶に新しい。また、Twitterにおいては、W杯の試合に関するツイートは全世界で約400万ツイートにのぼったという。（ITメディアニュース2014年6月16日より）

また、その経済効果も非常に大きいものであり、南アフリカ大会の場合、10のスタジアムの建設、増築改築が行われ、それらの費用だけでも約2000億円が投資された。直近のブラジル大会の場合は、約8兆円の経済効果が見込まれた。（朝日新聞デジタル2013年6月13日より）このように、国内のインフラが整備されるほか、観光客増加による国内外の消費拡大や治安改善など、W杯が開催国へもたらす影響は非常に大きい。そのため、官民間問わず大会招致に対する期待も大きく、世界中の多くの人々が注目をするのである。

前述した莫大な経済効果に加えて、最大級のスポーツの祭典ゆえに、当然、どの国が優勝するのかということへの関心も極めて大きく、世界の名だたるコンサルタントたちが、毎度のように様々な統計手法を用いてW杯優勝国予想などの報告書を発表する。その一つとして、外資系コンサルティングファームとして有名な株式会社プライスウォーターハウスクーパーズ（以下「PwC」）が、2014年ブラジル大会の開催前に発表した報告書がある。その中では、国の社会要因がサッカーの強さに与える影響を特定し、その結果からブラジルが優勝すると予測している。また、同様に世界有数の外資コンサルティングファームの株式会社ゴールドマンサックスも、出場国への経済効果予測に加え、国ごとの経済およびサッカーに関するデータを分析し、それらの結果からブラジル大会の結果シミュレーションを行っている。その結果は、PwC同様にブラジルが優勝すると予測している。

しかし、経済要因のみならず様々な要因が複雑に関連するサッカーというスポーツにおいて、その強さや試合結果を規定する要因を、統計モデルによって予測するのは容易なことではないといえよう。そこで、本研究においては、「各国のサッカーの強さを決定する社会要因は何か」というリサーチクエスションに回答することを本研究の目的とし、サッカーに関するデータのみならずFIFA加盟国の様々な社会要因に関するデータを用いて線形重回帰モデルを推定し、最小二乗法によりサッカーの強さを計る指標であるFIFAポイントに関する分析を行った。さらに、本研究においては、各国のW杯におけるパフォーマンスから二項プロビットモデルを想定し、最尤法により各国のパフォーマンス予測を行う。

本稿の構成は、2章で先行研究のレビューを行い、3章で本研究に用いられるデータについて説明し、記述統計を与える。4章では本研究のモデルおよび分析のフレームワーク（推定方法）を説明し、5章では推定方法とその解釈を、6章では本研究の結論を与える。

2. 先行研究

本章では、サッカーの強さに関して、計量的な分析を行っているものを中心に先行研究をレビューし、それらの反省点を踏まえ、本研究の独自性および分析のフレームワークの概略を与える。

前章で説明したPwCの発表した報告書（2014）では、直近6回のW杯で決勝リーグ（決勝グループリーグおよび決勝トーナメント）まで進んだ56カ国をサンプルとして、二つのモデルを想定し、分析を行っ

ている。

第一のモデルは、サンプルの国を行とした横断面データの重回帰分析であり、このモデルには、被説明変数にサッカーの強さを表す指標である「ワールドカップポイント」をとり、説明変数には「一人あたりGDP」、「人口」、「登録選手数」、「平均観客数」、「南米ダミー」、「欧州ダミー」、「開催国ダミー」が含まれる。「登録選手数」および「平均観客数」は、サッカーの国内人気を表す変数として設定されている。「南米ダミー」は、予測対象の2014年大会の開催国が南米大陸に位置するブラジルであることから、開催国と似たような気候区分に含まれる国々にはアドバンテージが存在するのではないかという仮説のもとで、南米大陸の国である場合には1を、そうでない場合には0を取るよう設定したダミー変数である。「欧州ダミー」についても同様で、FIFAランキングの高いいわゆる強豪国が多く存在する欧州各国の場合には1を、そうでない場合には0を取るダミー変数である。「開催国ダミー」は、W杯の開催国になったことのある国は、ホームのアドバンテージに加え、戦略ノウハウの定着など開催経験から得られるアドバンテージがあるという仮説のもと、過去3大会のうち開催国となった国である場合は1を、そうでない場合には0を取るよう設定したダミー変数である。この第一のモデルを最小二乗法により分析した結果、「一人あたりGDP」および「人口」は有意でなかったため、最終的なモデルから除外し、再び分析を行っている。その結果、「登録選手数」、「南米ダミー」、「欧州ダミー」、「開催国ダミー」が5%以下で統計的有意となり（符号はいずれも正）、「平均観客数」が10%以下で有意となった。

第二のモデルは、過去6回大会の試合成績より得たパネルデータから順序プロビットモデルを想定し、最尤法により分析したものである。このモデルには、被説明変数に各国の大会におけるパフォーマンスから階層を組んだ「各国のパフォーマンス」をとり、説明変数に「一人あたりGDP」、「人口」、「登録選手数」、「平均観客数」、「南米ダミー」、「欧州ダミー」、「ホスト国ダミー」、「ホスト大陸ダミー」、「前大会パフォーマンスダミー」、「前々大会パフォーマンスダミー」を取っている。「各国のパフォーマンス」については、パフォーマンスの水準を表す観測不能な潜在変数を想定し、観測可能な「各国のパフォーマンス」と潜在変数との間に「優勝」ならば5を、「決勝進出」ならば4を、「準決勝進出」ならば3を、「準々決勝進出」ならば2を、「ベスト16進出」ならば1を、「ベスト16未満」ならば0となるような潜在変数の値に応じた仮定を設定している。分析結果は、「一人あたりGDP」、「人口」、「平均観客数」は非有意であったため、これらは最終的なモデルから除外している。再分析の結果は、「登録選手数」、「南米ダミー」、「欧州ダミー」、「ホスト国ダミー」が有意水準5%で統計的有意であった（符号はいずれも正）。また、「ホスト大陸ダミー」は10%で有意となった。

しかし、PwCの分析においては、第一のモデルにおける重回帰分析において、サンプル数が56と極めて少ないことから、バイアスが生じている可能性が否めない。また、順序プロビットモデルを想定した第二のモデルにおいて、観測可能なパフォーマンスによって組んだ階層は、パフォーマンスが高まるほどその分布数が少なくなっており、分析結果の信頼性に欠ける点が指摘される。すなわち、ある1大会において、「優勝」というパフォーマンスを残す国は1カ国のみであり、6大会分のパネルデータを得られたとしても「優勝」というパフォーマンスを残す国は6カ国しか存在せず、結果がバイアスする可能性がある。さらに、両モデルにおいて分析対象を決勝リーグに進出した国々に限定しているため、それらの分析結果は決勝リーグに進出した国々の中での傾向を示しているにすぎないという致命的な問題が指摘できる。本研究が目的としている一般的なサッカーの強さを決定づける要因の特定に対しては、有益な分析結果ではないといえよう。

しかし、PwCの手法は、分析のフレームワークとしては非常に有用であるため、本研究においては、PwCの分析フレームワークを参照する。すなわち、本研究においては2つのモデルを与え、一方は線形重回帰モデルの最小二乗法による分析、他方は順序プロビットモデルの最尤法による分析を行う。

また、同様に世界的に有名な外資コンサルタント株式会社ゴールドマンサックスも、W杯優勝国予想に関する報告書（2014）を発表している。この報告書では、1960年の第1回大会以降の親善試合を除く国際試合（大会予選、1次リーグ予選、決勝トーナメントすべての試合）約14000試合の試合結果を対象に

、モンテカルロ法により2014年大会の試合結果をシミュレーションしている。他にも、参加する32カ国の過去の成績、サッカーの状況、経済の状況などを国別に説明している。試合の結果予測については過去の国際試合の勝敗および得点数、失点数を回帰分析したもので、決勝はブラジル対アルゼンチンの対戦になるとしており、その結果、3対1でブラジルが優勝すると予測している。なお、回帰モデルには、二者間の実力の測定値を算出するイロレーティングと呼ばれる算出法から求めた国ごとの能力が被説明変数に、説明変数には「直近10試合の得点数」、「直近5試合の失点数」、「対象組み合わせの過去大会における結果」、「対象組み合わせの試合がホームか否か」、「対象組み合わせの試合がホーム大陸か否か」などが含まれている。

しかし、ゴールドマンサックスの分析においては、サンプルが過去の国際試合の結果を対象としているため、チームの属性をあらわす身長や体重などは分析に含まれていない。また、監督の能力や個々の選手の実力など、個人的な属性についても考慮されていないため、試合の結果予測としては非常に有力である一方、「サッカーの強さ」をあらわす分析としてふさわしいとは言い難い。

上述のように、サッカーの強さに関する計量分析は様々になされているが、その中で、本研究が与える貢献は、主に3点あげられる。第一に、PwCのフレームワークを参照し、サンプルの数を拡大することで、サッカーの強さに影響を与える要因により一般性を持たせる点である。本研究におけるサンプルは、FIFA加盟国208カ国中で有効なデータを得ることができた153カ国であり、PwCの分析と比較してサンプル数の少なさによるバイアスを小さくすることができるといえよう。本研究のサンプルにはW杯出場経験のない国々も多く含まれており、PwCの分析において指摘されるサンプルセクションバイアスを回避すると予想できる。第二に、パフォーマンスの予測に関して、階層に属する分布数を考慮し、階層を組み直して分析することで、より信頼性のある予測が可能となる点である。すなわち、観測可能なパフォーマンスから「決勝トーナメントのグループリーグ進出」および「予選敗退」という二項の階層を組み、二項プロビットモデルを分析することで、より正確に大会におけるパフォーマンス予測を行う。第三に、説明変数の設定を厳密に行うことで、サッカーの強さに影響を与える要因をより多角的に測定する点である。先行研究においては、各国の経済要因については言及されているものの、国の平均身長や平均体重、教育水準など、スポーツにおける重要な要素は欠落していた。本研究においては、「国の社会的要因に関する説明変数」としてこれらの変数を加え、より正確に、サッカーに影響を与える要素を考慮し、説明変数の欠落によるバイアスを回避することが可能であろう。

3. データと記述統計

本研究の分析で扱われるデータは主にサッカーに関する変数と国の社会的要因に関する変数とに大別される。サッカーに関する変数については、W杯を主催する国際サッカー連盟（FIFA）の公式ホームページ掲載のデータおよび、過去に開催されたすべてのW杯に関連するデータを記録したホームページであるFIFARANKING.net掲載のデータを用いる。国の社会的要因に関する変数については、世界銀行（WB）の公式データベースや、国際通貨基金（IMF）の公式データベース（World Economic Outlook Database）など、主に国連専門機関の提供するデータを用いた。

3. 1. 被説明変数

本研究において、国のサッカーの強さを表す変数である「FIFAポイント（ pt_i ）」については、FIFARANKING.net掲載のデータのうち、2012年4月発表のFIFAポイントを利用した。なお、FIFAポイントは実際にナショナルチームのサッカーの強さのランキングであるFIFAランキングの算出に用いられており、数値が大きいほど、国際試合における成績が高く、サッカーが強いといえることができる。

3. 2. 説明変数

3. 2. 1. サッカーに関する変数

身長に関する説明変数 (hgt_i) は、世界の平均身長が国別に掲載されているwebページaverageheight.coより各国の男性の平均身長を利用し、単位は日本において標準的なcm単位に換算した。体重に関する説明変数は、FIFA加盟国の平均体重を網羅するデータが存在しなかったため、肥満の指標として利用されるBMIの国別平均BMIを代理変数とし、「平均BMI (bmi_i)」を設定した。体重の代理変数としての平均BMIについては、WHOによる調査をソースとしているBioMed Central社の報告書(2012)によって得られたデータを利用した。「開催国ダミー ($exbids_i$)」については、2014年までで全20回開催されている過去大会のうち、データを得られた2012年よりも以前に、一度でも開催国になったことがある国の場合は1を、そうでない場合には0を取るダミー変数を設定することで、開催国のアドバンテージの有無を図る。

3. 2. 2. 国の社会的要因に関する変数

人口、GDP、気候に関する説明変数は主に世界銀行のデータベースより2012年のデータを利用し、それぞれ「人口 (pop_i)」、「一人当たりGDP ($gdppc_i$)」、「年平均降水量 ($prec_i$)」を設定した。なお、GDPについては、多重共線性によるバイアスを解消するため一人当たりに換算されたデータを採用した。気候に関する説明変数は、サンプルとなる国々の平均気温を網羅するデータが存在しなかったため、世界銀行のデータおよび気象庁のホームページを参照し、年平均降水量を代理変数として利用した。単位については、人口は1万人、GDPは米ドル、年平均降水量はmmとして再計算した。

教育水準に関する説明変数は、国連開発計画 (UNDP) の報告書 (2011) より「各国の不平等調整済み人間開発指数 ($ihdi_i$)」(以下、「IHDI」)を代理変数として利用した。なお、IHDIは出生時平均余命、平均就学年数、就学予測年数、一人当たりジニ係数の関数である人間開発指数 (HDI) に対して、各国の所得不平等を考慮し再計算したものであり、長寿で健康な生活、知識へのアクセス、人間らしい生活水準という人間開発の三要素の達成度を図る指数である。一般に、IHDIの値が高いほど、前述の人間開発の三要素が達成されており先進的な生活が実現されていると言える。本研究で採用した2011年発表の報告書中で、IHDIが最も高かったのはノルウェーの0.943であった。

植民地に関するダミー変数 ($udummy_i$, $frdummy_i$, $spdummy_i$) については、FIFAランキング上位にいる国3カ国 (イギリス、フランス、スペイン) の植民地になったことのある国の場合は1を、そうでない場合は0を取るダミー変数を、3カ国それぞれについて設定することで、移民や植民などによる影響を図る。

3. 3. 記述統計

以上の変数について、記述統計を表1に示す。各国のサッカーの強さを表すFIFAポイントについては、サンプルの平均値が471.0で、最小値はブルネイの4.0、最大値はスペインの1442.0であった。当時の日本のFIFAポイントは753.0であるため、日本はサンプル中では比較的上位に位置するといえる。また、「開催国ダミー ($exbids_i$)」によりサンプルを開催国と非開催国に分類すると、開催国の平均的なFIFAポイントは999.94、非開催国の平均的なFIFAポイントは411.02であった。全サンプルの平均値が471.0ということとを考慮すると、非開催国よりも開催国の方が、サッカーの強さを表すFIFAポイントは高くなる傾向があることがわかる。

さらに、「植民地ダミー ($udummy_i$, $frdummy_i$, $spdummy_i$)」によりサンプルをイギリス植民地だった国、フランス植民地だった国、スペイン植民地だった国、上記3カ国の植民地とならなかった国の4つに分類する。その結果、イギリス植民地だった国々の平均FIFAポイントは348.9、フランス植民地だった国々の平均FIFAポイントは411.8、スペイン植民地だった国々の平均FIFAポイントは626.4、上記いずれの国の植民地にならなかった国々の平均FIFAポイントは509.5であった。全サンプルの平均値が471.0であることから、スペイン植民地であった国々はFIFAポイントが高くなる傾向にあることがわかった。本研究に用いた2012年4月時点のFIFAランキングにおいて、スペイン

が首位であることから、この事実は非常に興味深いといえよう。

各国の平均身長については、サンプルの平均は170.8cmであり、各国の平均BMIについては、サンプルの平均は23.87であった。一般にBMIは、18.5以上25.0未満の場合、「普通体重」として認められるため、サンプルの平均値は「普通体重」を捉えているとみなすことができる。人口については、平均値が4,520.8万人、最小値がアイルランドの32.07万人、最大値が中国の135,069.5万人であった。およそ13億というこの中国の数値は全世界人口の約6分の1に匹敵する。

3. 4. 相関関係

さらに、FIFAポイント、すなわちサッカーの強さと、サッカーおよび国の社会要因に関する説明変数との相関関係を把握するために、各変数による相関行列を表2に示す。まず、「平均身長」、「平均BMI」、「一人あたりGDP」については、いずれにおいても、FIFAポイントとの間に弱い正の相関がみられた。一方、「年平均降水量」については、FIFAポイントと間に非常に弱い負の相関がみられた。ここで、被説明変数であるFIFAポイントではなく、「IHDl」と各変数との相関について着目すると、「FIFAポイント」、「平均身長」、「平均BMI」、「一人あたりGDP」との間に正の相関がみられる（「FIFAポイント」とは0.486、「平均身長」とは0.559、「平均BMI」とは0.688、「一人あたりGDP」とは0.705）。

<表1 変数の記述統計>

変数名	標本数	平均値	標準偏差	最小値	最大値
FIFAポイント	153	471.0	315.407	4.0	1,442.0
平均身長	153	170.8	5.398	156.0	183.8
平均BMI	153	23.87	1.887	19.85	28.39
人口	153	4,520.8	15,109.695	32.07	135,069.5
一人あたりGDP	153	13,617	19,100.413	251	103,859
年平均降水量	153	1,099	754.478	51	3,240
IHDl指数	153	0.6571	0.184	0.2840	0.9430

<表2 相関関係>

変数名	FIFA ポイント	平均身長	平均BMI	人口	一人あたり GDP	年平均 降水量	IHDl指 数
FIFAポイント	1.000						
平均身長	0.545	1.000					
平均BMI	0.364	0.360	1.000				
人口	0.000	-0.133	-0.139	1.000			
一人あたりGDP	0.384	0.576	0.387	-0.050	1.000		
年平均降水量	-0.092	-0.356	-0.065	-0.013	-0.022	1.000	
IHDl指数	0.486	0.559	0.688	-0.012	0.705	-0.034	1.000

4. モデルと分析のフレームワーク

本研究においては、各国のサッカーの強さに影響を与える要因の特定および、次期W杯におけるパフォーマンスの予測を目的としており、それぞれ、線形重回帰モデルの最小二乗法による分析および、各国のパフォーマンスに従い序列を組んだ順序プロビットモデルの最尤法による分析が想定される。そのため、本研究においては二つのモデルを想定する。

まず、各国のサッカーの強さは、各国の平均身長や平均体重、開催国となるアドバンテージなど様々な要因と関係していると考えることができる。本研究においては、「モデル1」として、各国のサッカーの強さすなわち「FIFAポイント (pt_i)」とそれら様々な要因との関係性を以下のように定式化する。これを以下、式 (1) とする。

$$pt_i = \beta_0 + \beta_1 hgt_i + \beta_2 bmi_i + \beta_3 pop_i + \beta_4 gdppc_i + \beta_5 gdpsq_i + \beta_6 prec_i + \delta w_i + \varepsilon_i$$

前章でも解説したように、「IHDI ($ihdi_i$)」は、他の説明変数との相関関係がみられる。これは、「IHDI ($ihdi_i$)」が出生時平均余命、平均就学年数、就学予測年数、一人当たりジニ係数など多くの変数の値に依存する変数であり、本研究で用いる国の社会要因を表す変数と相関しやすい傾向にあるためだと推測される。そのため、本研究においては、多重共線性によるバイアスを回避するため、本研究のモデルにおいては「IHDI ($ihdi_i$)」を削除した。

また、「一人あたりGDP ($gdppc_i$)」については、FIFAポイントとの関係性を考慮した場合、その関係性は直線的な関係にならないと判断し、「一人あたりGDP ($gdppc_i$)」を二乗した値である「一人あたりGDPの二乗項 ($gdpsq_i$)」を加えた。すなわち、一定の値まではFIFAポイントと一人あたりGDPとは一次関数的な正の相関関係を持つが、ある一点を超えた時には負の相関関係を持ち、曲線的な関係になると考えられるということである。これは、「一人あたりGDP ($gdppc_i$)」が最も大きいルクセンブルクのFIFAポイントが283.0であり、決して高い順位に位置していないという事実からも認められよう。なお、 w_i は各国の属性を表すベクトルであり、具体的には「開催国ダミー ($exbids_i$)」および「植民地ダミー ($ukdummi_i, frdummy_i, spdummy_i$)」が含まれる。そのほか、 β_0 は定数項、 ε_i は誤差項である。

また、各国のパフォーマンスもサッカーの強さと同様に、様々な要因と関係しているといえよう。本研究においては、「モデル2」として、各国のW杯におけるパフォーマンスとそれら様々な要因との関係性を以下のように定式化する。これを以下、式 (2) とする。

$$perform_i^* = \beta_0 + \beta_1 hgt_i + \beta_2 bmi_i + \beta_3 pop_i + \beta_4 gdppc_i + \beta_5 gdpsq_i + \beta_6 prec_i + \delta w_i + \varepsilon_i$$

ここで、 $perform_i^*$ は、大会における各国のパフォーマンスの水準を表す、実際には観測できない潜在的な変数である。一方、各国のパフォーマンスに関して実際に観測できる変数「各国のパフォーマンス」($perform_i$)については、本研究において利用した2012年4月のFIFAランキング以前で直近のW杯(2010年南アフリカ大会)でのパフォーマンスを、序列を組み階層化した数値を利用することができる。すなわち、「決勝トーナメントのグループリーグへ進出」の場合は1、「予選敗退」の場合は0の値をとるようにを設定した。また、本研究においては、これら二つの変数の間に以下のような仮定を設定する。この仮定を、以下、式 (3) とする。

$$perform_i = \begin{cases} 1 & \text{iff } perform_i^* < \mu \\ 0 & \text{iff } \mu \leq perform_i^* \end{cases}$$

ここで μ は閾値パラメータであり、具体的には、2014年大会のパフォーマンスの結果に関する条件である。すなわち、国のパフォーマンスの水準が μ 未満の場合には「予選敗退」、 μ 以上の場合には「決勝トーナメントのグループリーグ進出」というパフォーマンスを残すという仮定である。これら両階層に分布するデータの記述統計は以下の表3および表4の通りである。

<表3 変数の記述統計（「決勝トーナメントのグループリーグ進出」した国について）>

変数名	標本数	平均値	標準偏差	最小値	最大値
FIFAポイント	31	889.0	268.172	256	1442

平均身長	31	174.2	4.897	163.8	183.8
平均BMI	31	24.88	1.358	21.93	27.82
人口	31	5227.32	6833.204	205.72	31387.37
一人あたりGDP	31	23818.84	18703.813	1220	67436
GDPの二乗項	31	917074423.03	1121053027.580	1488400	4547614096
年平均降水量	31	1027.94	434.891	89	1976

<表4 変数の記述統計（「予選敗退」した国について）>

変数名	標本数	平均値	標準偏差	最小値	最大値
FIFAポイント	122	365.6	225.817	4	1114
平均身長	122	169.9	5.168	156	182.4
平均BMI	122	23.62	1.917	19.85	28.39
人口	122	4341.25	16561.723	32.07	135069.5
一人あたりGDP	122	11025.09	18316.170	251	103859
GDPの二乗項	122	457034552.57	1510604354.344	63001	10786691881
年平均降水量	122	1116.93	814.996	51	3240

5. 分析結果と解釈

5. 1. 「モデル1」の分析結果

まず、式（1）に関して、被説明変数であるFIFAポイントと、各説明変数との単回帰分析の結果を表5に示す。「人口」および「年平均降水量」以外の説明変数はすべて1%以上の水準で統計的有意であった。次に、式（1）に関して、各変数を重回帰分析した結果を表6に示す。重回帰分析により統計的有意となった変数は0.1%以下の水準で「平均身長」および「開催国ダミー」（いずれも符号は正）、5%以下の水準で「スペイン植民地ダミー」（符号は正）の3つであった。これより、（1）平均身長が高いほど、FIFAポイントは高くなり、サッカーの強さにも正の影響を与え、（2）開催国となるアドバンテージはサッカーの強さに正の影響を与えるほか、（3）歴史的にスペインの植民地であった国々はサッカーが強い傾向があるということがいえる。

平均身長については、回帰係数が28.11であることから、国の平均身長が1cm高くなると、FIFAポイントは28.11ポイント上昇する傾向にあるということが言える。一般に、空中に浮いたルーズボールの競り合いなどでは、身長が高い方がボール奪取の可能性が高いというのは、想像に容易いため、これは「背が高い方がサッカーには有利」ということがデータにより裏付けられたといえよう。

開催国のアドバンテージに関しては、ホームとアウェーとの試合を比較した時、観客数や選手のコンディションを考慮すると明らかにホームの試合の方が選手のパフォーマンスが高くなることは想像に容易い。これより、W杯という非常に規模の大きい大会において、多くの自国民の声援を味方につけることが可能な開催国すなわちホームの状態プレーをできるということは、サッカーの強さに影響を与え、その結果、アドバンテージが存在するということが容易に想像できるであろう。この結果は、先行研究のPwCの分析結果と整合的であった。

スペインの植民地ダミーについては、非常に興味深い考察が与えられよう。具体的には、FIFAランキング首位に位置することの多いスペインの植民地だった国はFIFAポイントが高い傾向にあるということは、歴史的にスペイン人にはサッカーに対する有利性が存在する可能性を示唆している。スペインの国内一部リーグ「リーガエスパニョーラ」のクラブチームの形態をみると、スペインのクラブチームの下部組織は、他国よりも非常に充実している。これより、国内チームがサッカーに投資している度合いが高いということ

、および、有力な若手を非常に早い段階から育成しようという意図が読み取れる。このようなスペインのサッカーに対する姿勢が植民地においても根付いたのではないかと考察できる。

また、「開催国ダミー」を利用し、開催国および非開催国それぞれについて分析した結果についても、似たような結果を得ることができた。本研究で用いるサンプルから、開催国のみを抽出し、最小二乗法による回帰を行った結果、5%以下の有意水準で「身長」が統計的に有意となった（符号は正）。すなわち、W杯開催国だけに着目しても、身長が高いほどFIFAランキングが高く、サッカーは強いといえる。同様に、非開催国についての結果も、0.1%以下の水準で「身長」が、0.5%以下の水準で「スペインの植民地ダミー」が有意であった（いずれも符号は正）。これは、非開催国においては、身長が高いことおよびスペインの植民地であったことがFIFAランキングに正の影響を与えているという解釈を与えることができる。これらの解釈は上述の全サンプルにおける分析結果に対する解釈と類似しているといえよう。

5. 2. 「モデル2」の分析結果

まず、式（3）の仮定のもとで、式（2）を最尤法により推定した結果を表7に示す。統計的に有意な変数は「平均身長」および「開催国ダミー」、「イギリス植民地ダミー」（いずれも符号は正）の3つであり、最小二乗法による「モデル1」の分析結果と多少の変化が見られた。これより、平均身長が高く、開催国であり、イギリスの植民地であった国々はパフォーマンスがよくなる確率が高いといえる。

中でもイギリス植民地のダミーについては、非常に興味深い結果となった。「サッカー（フットボール）の母国」として名高いイングランドは、イギリス連邦を構成する一国であり、イングランド国内の1部リーグは世界的に見ても非常にレベルが高く、現在においてもその伝統的なサッカーの形は、少なくともイギリス連邦内においては根付いているといえることができるであろう。事実として、世界一のクラブチームを決めるFIFAクラブワールドカップ大会において、ヨーロッパを代表して選出されるクラブチームはイングランドの国が多い（2000年より開催し、イングランドのチームがヨーロッパ代表となったのは2005年、2008年、2011年）。さらに、イギリス連邦を構成するスコットランドやウェールズなどの国々からも、世界各国の有名チームへと優秀なサッカー選手を輩出していることから、母国としての伝統的なサッカーはイングランドのみならずイギリス連邦内に根付いているといえることができるだろう。このような特性を持つイギリスの植民地であることにより、植民地の国々においてもその伝統的なサッカーの様式が根付いた可能性があると考えることができる。

また、「モデル2」では、誤差項 ε が正規分布すると仮定し、その分散を1として標準化すると、閾値 μ と、式（2）の定数項パラメーター β_0 との差を推定することが可能である。これより、 $\beta_0 = 0$ として標準化を行うと、観測不能なパフォーマンスの水準に対する閾値の値を求めることができ、その値は、23.948となっている。

ここで、日本の2014年のデータを用いて、2018年W杯の日本のパフォーマンス予測を行う。前章で述べたように「モデル2」においては、多重共線性の問題が解消されており、分析結果より統計的に有意な変数を取り出すと、「モデル2」における回帰式は

$$perform_i^* = 0.124 hgt_i + 2.191 exbids_i + 0.734 ukdummy_i$$

となる。2014年版の世界の男性の平均身長ランキングにおいて、日本は170.7cmの29位であった。（livedorニュースBLOGOS

2014年12月11日）また、2002年日韓大会において開催経験があり「開催国ダミー（ $exbids_i$ ）」は1を取る。同様に、イギリス植民地となった経験はないことから「イギリスの植民地ダミー（ $ukdummy_i$ ）」は0を取る。これらを回帰式に代入すると、パフォーマンスの水準 $perform_i^*$ は、およそ23.358となる。本研究が与えた結果と照らし合わせると、この水準の実現値23.358は閾値23.948よりも小さい値である。これより、最

新のデータを用いて2018年大会の日本のパフォーマンスを予測すると、 $perform_i^*$ の値は0となり、これは「予選敗退」を意味している。日本は過去大会において、1994年アメリカ大会以前は一度も予選突破をした経験がないことから、整合性を持ち得る。1998年フランス大会以降はグループリーグまでは進出できているものの、近年のFIFAランキングの推移を参照すると2011年から継続して下降の傾向が続いていることから、この結果は整合的であると言わざるを得ないであろう。

<表5 推定結果（単回帰分析）>

変数名	回帰係数	標準偏差	t値
平均身長(<i>hgt</i>)	31.828	3.987	7.983
<i>hgt</i> の定数項	-4963.964	681.272	-7.286
平均BMI(<i>bmi</i>)	60.81	12.67	4.800
<i>bmi</i> の定数項	-980.09	303.38	-3.231
人口(<i>pop</i>)	4.233e-06	1.699e-03	0.002
<i>pop</i> の定数項	471.6	26.79	17.604
一人あたりGDP(<i>gdppc</i>)	6.340e-03	1.241e-03	5.109
<i>gdppc</i> の定数項	385.3	29.11	13.239
年平均降水量(<i>prec</i>)	-0.03851	0.03388	-1.137
<i>prec</i> の定数項	-513.98657	45.15493	11.383
IHDI (<i>ihdi</i>)	833.02	122.00	6.828
<i>ihdi</i> の定数項	-75.70	83.24	-0.909

<表6 推定結果（重回帰分析）>

変数名	回帰係数	標準偏差	t値
(定数項)	-462.3	820.6	-5.634
平均身長(<i>hgt</i>)	28.11	4.736	5.935
平均BMI(<i>bmi</i>)	8.608	12.53	0.687
人口(<i>pop</i>)	9.72E-04	1.20E-03	0.810
一人あたりGDP(<i>gdppc</i>)	2.44E-03	3.10E-03	0.788
GDPの二乗項(<i>gdpsq</i>)	-4.05E-08	3.25E-08	-1.180
年平均降水量(<i>prec</i>)	0.022	0.027	0.811
開催国ダミー(<i>exbids</i>)	410.0	65.61	6.248
イギリス植民地ダミー(<i>ukdummy</i>)	-67.04	42.09	-1.593
フランス植民地ダミー(<i>frdummy</i>)	29.38	57.32	0.513
スペイン植民地ダミー(<i>spdummy</i>)	157.2	66.61	2.361

<表7 推定結果（重回帰分析）>

変数名	回帰係数	標準偏差	z値
平均身長(<i>hgt</i>)	0.123	0.431	2.87
平均BMI(<i>bmi</i>)	0.455	0.111	0.41
人口(<i>pop</i>)	4.46E-06	1.06E-05	0.42
一人あたりGDP(<i>gdppc</i>)	1.90E-05	2.05E-05	0.93
GDPの二乗項(<i>gdpsq</i>)	-3.92E-10	2.88E-10	-1.36
年平均降水量(<i>prec</i>)	2.37E-05	2.43E-04	0.10
開催国ダミー(<i>exbids</i>)	2.191	0.526	4.17
イギリス植民地ダミー(<i>ukdummy</i>)	0.734	0.373	1.97
フランス植民地ダミー(<i>frdummy</i>)	0.277	0.467	0.59
スペイン植民地ダミー(<i>spdummy</i>)	0.792	0.581	1.36
閾値パラメータ(<i>/cutI</i>)	23.948	7.287	

6. 結論

本研究は、サッカーの強さに影響を与える様々な要因について、特にサッカーに関する要因および国の社会的要因という側面から分析した。本研究から得られた結果は以下のようにまとめることができる。

- (1) サッカーの強さと、国の社会要因に関する変数「平均身長」、「平均BMI」、「一人当たりGDP」との間にはそれぞれ弱い正の相関関係が認められる。
- (2) サッカーの強さを調べる「モデル1」において統計的有意であった変数は「平均身長」、「開催国ダミー」、「スペインの植民地ダミー」で、いずれも符号は正であった。
- (3) パフォーマンスを予測する「モデル2」において統計的有意であった変数は「平均身長」、「開催国ダミー」、「イギリス植民地」で（いずれも符号は正）、「モデル1」とは分析結果に差が生じた。
- (4) 開催国において有意となった変数は「平均身長」のみで（符号は正）、非開催国において有意となった変数は「平均身長」および「スペインの植民地ダミー」であり（いずれも符号は正）、これは全サンプルにおけるポイントの強さを調べる「モデル1」の傾向と類似している。
- (5) 最新のデータを用いると、次期の2018年大会において日本のパフォーマンスは「予選敗退」と予測される。

本研究の結果は、一般的なサッカーの強さは、国の平均身長、主催国か否かということ、歴史的にスペインの植民地であることによりアドバンテージを受けるということを意味している。さらに、W杯の主催経験がある国とそうでない国との間で、サッカーの強さのみならず、強さに影響を与える要因が異なるということも同時に明らかにした。すなわち、ナショナルチームのサッカーの強さを高めるためには、大会主催の経験を持つことによるアドバンテージが存在するため、大会誘致への投資を積極的に行うことが必要であり、可能であるならば、国の平均身長を高めていくような政策、すなわち健康的な生活を国民が送れるような厚生を提供することが必要であるといえよう。同時に、植民地支配に関しては様々な歴史的解釈が関連するため本稿においては割愛するが、FIFAランキング首位に位置し「無敵艦隊」として名高いスペインのサッカーに倣うことが求められているといえるだろう。

また、ナショナルチームの次期大会におけるパフォーマンスを高めるためには、サッカーの強さと同様に大会主催経験を持つこと、国の平均身長を高めるようにすることに加え、「サッカーの母国」として有名なイギリス流のサッカーを参考とすることが望ましいといえるだろう。

最後に、本研究における課題を示す。本研究の分析に用いたサンプルの数は153であり、これはFIFA加盟国すべてを網羅するものではない。故に、先行研究よりも、より一般性を持つ分析結果を提示することができたものの、完全な一般性を持たず、分析結果にバイアスが生じている可能性も指摘されうる。したがって、FIFA加盟国208カ国すべてを網羅できるようなデータを入手し分析する必要性があるだろう。また、本研究においては、平均体重の代理変数として平均BMIを、（最終的にはモデルから削除したものの）教育水準の代理変数としてIHDIを、気候要因の代理変数として年平均降水量を利用するなど、多くの代理変数を用いている。そのため、本研究が目的としていたサッカーの強さに影響を与える要因について、厳密な意味で分析が行われていないという指摘もあり得るだろう。そのため、より詳細なデータを用い、説明要因を的確に捉えることができるモデルを構築する必要があるだろう。

7. 参考文献

朝日新聞デジタル2013年6月13日

<http://www.asahi.com/international/update/0613/TKY201306130005.html>

アクセス日時：2015年1月7日17:39

ITメディアニュース2014年6月16日

<http://www.itmedia.co.jp/news/spv/1406/16/news073.html>

アクセス日時：2015年1月7日17:21

PricewaterhouseCoopers

“The PwC World Cup Index: what can the dismal science tell us about the beautiful game?”(June 2014)

Goldman Sachs

“The World Cup and Economics 2014”(May 2014)

FIFA official homepage

<http://www.fifa.com/>

アクセス日時：2014年12月18日16:05

FIFARANKING.com

<http://fifaranking.net/>

アクセス日時：2014年12月11日17:30

The World Bank DataBank

<http://databank.worldbank.org/data/home.aspx>

アクセス日時：2014年12月11日11:38

IMF Data

<http://www.imf.org/external/data.htm>

アクセス日時：2014年12月10日15:02

averageheight.co

<http://www.averageheight.co/average-male-height-by-country>

アクセス日時：2014年12月4日17:32

BioMed Central

“The weight of nations: an estimation of adult human biomass”(2012)

WHO Global Infobase Team

“The SuRF Report 2 Surveillance of chronic disease risk factors: country-level data and comparable estimates”(2005)

Japan Meteorological Agency

<http://www.jma.go.jp/jma/menu/menureport.html>

アクセス日時：2014年12月11日17:36

国連開発計画(UNDP)

「人間開発報告書2011（統計表）」（2011）

livedoorニュースBLOGOS 2014年12月11日

<http://blogos.com/article/101034/>

アクセス日時：2015年 1 月8日17:46