

男子プロテニスの試合に関する統計的分析

2011SE091 伊藤秀悟

指導教員：木村美善

1 はじめに

テニスはサーブを打つ側が有利なスポーツと言われている。主な理由としては、サーブを自分のタイミングで打つことが出来るという理由が挙げられることが多い。そこで本研究では、男子プロテニスの試合ごとのデータを用いて、分析を行い、勝利に対してどのような要素が有効に働いているか統計的に分析を行う。

2 データ

スポーツナビ ([3] 参照) から 2014 年のグランドスラムと呼ばれる全豪オープン、全仏オープン、ウィンブルドン選手権、全米オープンの 4 つの男子プロテニスの大会において途中棄権となった試合を除いた試合のデータを用いる。変数は「勝敗 (勝ち=1, 負け=0)」、「サービスエース数」、「ファーストサーブ成功率」、「ダブルフォルト数」、「サーブポイント成功率」、「レシーブポイント成功率」、「ウィナー数」、「ブレイク率」、「最高サーブ速度」、「ファーストサーブ平均速度」、「セカンドサーブ平均速度」の合計 11 個である。

3 分析方法

分析方法として、ロジスティック回帰、主成分分析及びクラスター分析を行った ([1], [2] 参照)。

4 ロジスティック回帰分析

説明変数群 x_1, x_2, \dots, x_n を用いて、確率のような範囲 $[0, 1]$ の値をとる目的変数を説明するとき、通常重回帰分析では、説明変数群の合成変数の取りうる値が $[0, 1]$ の外に出てしまい、 $(-\infty, \infty)$ となり適切ではない。なので下式のように説明変数群の合成変数をロジスティック関数にすることで、その値域を範囲 $(0, 1)$ に収めることができる。

$$y = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}$$

目的変数 y を勝敗 (勝ち=1, 負け=0) とし、11 個の説明変数を用いて、 y を予測した。まず 4 つの大会のデータ全てを用いて、どのような特徴があるか分析を行った。その後、それぞれの大会ごとに分析を行い、大会ごとの傾向や特徴の分析を行う。

4.1 4大会のロジスティック回帰分析

4 つの大会の全てのデータと全ての説明変数を用いて、ロジスティック回帰分析を行った。分析を行った結果、多重共線性は見られなかったが、より良い分析を行う為に、AIC に基づく変数減少法により残った変数を用いて、再度ロジスティック回帰分析を行い、結果を表 1 に示す。

表 1 を見ると、勝利に働く要因としてサーブポイント率は有効であるとは感じていたが、レシーブポイント率やエース数の方が有効に効いていることが分かる。また、最高サーブの速さが有効に働いており、サーブが速いほどサービスエースが取りやすくなり、勝利に有効に働いたのではないかと思われる。またセカンドサーブの速度は遅くなることによって、勝利に対して有効に働くので、ファーストサーブでは、スピードが重要であるが、セカンドサーブは、スピードではなく、コースやボールのスピンのによって攻めて行くことが、重要であると考察出来る。しかし、セカンドサーブが遅くなることによって、甘いコースに入った場合、相手のリターンによってポイントを取られてしまうリスクも増えるため、他の変数と比較して勝敗にあまり大きな影響を与えていないことが考察出来る。

表 1 ロジスティック回帰分析結果

	推定値	z 値	p 値
(Intercept)	-20.925	-11.333	2.00×10^{-16}
サーブポイント率	0.095	2.829	4.67×10^{-3}
レシーブポイント率	0.301	14.294	2.00×10^{-16}
ウィナー数	0.101	10.053	2.00×10^{-16}
ブレイク率	-0.022	-3.981	6.86×10^{-5}
最高サーブ速度	0.033	7.601	2.94×10^{-14}
平均 2nd サーブ速度	-0.015	-1.677	9.35×10^{-2}

4.2 各大会のロジスティック回帰分析

同様の分析方法で、大会ごとの分析を行ったところ、4 つの大会全てにおいて「サーブポイント成功率」「レシーブポイント成功率」の 2 つの要素が、勝利に大きな影響を与えており、サーブ及び、レシーブの獲得が勝利に対してどちらとも有効であった。

各大会の特徴として、コートの特徴の違いが挙げられる。全豪オープンと全米オープンの 2 つの大会のコートは、ハードコートが採用されている。この 2 つの大会では、サーブポイント率が他の 2 つの大会と比べて特に大きな影響を与えており、サーブがバウンド後によく跳ねて、球速が落ちないというハードコートならではの特徴がよく表れていたと読み取れる。また同じハードコートでも、全米オープンのコートは全豪オープンのコートよりもコートが固くなっている為、全米オープンの方が全豪オープンと比較してサーブポイントが与える影響が大きくなっている。全豪オープンのみ、変数選択によりダブルフォルトが変数として残っている。セカンドサーブを確実に入れるのではなく、ダブルフォルトの可能性も承知でセカンドサーブでも攻めていく姿勢が、ダブルフォルトが多くなってしまっても、勝ちに有効に働いていると考察することが出来る。

全仏オープンは、クレイコートという土のコートが採用されており、他のコートと比べて球速がバウンド後に遅くなり、サービスエースを狙うことが困難になるため、ラリーによる乱打戦になる場合が多い。全仏オープンでも、サービスポイント率は大きな影響を与えていたが、レシーブポイント率が他の大会に比べて大きな影響を与えている点や変数選択法によりエース数が残り、ラリーでのポイントの獲得が他の大会に比べて勝敗に大きな影響を与えている点からクレイコートの特徴が読み取れる。

ウィンブルドン選手権は、グラスコートという芝のコートを採用している。グラスコートは、他のどの種類のコートよりもボールのバウンド後に弾道が低く、球速が速くなるという特徴があり、ファーストサーブによるサービスエースが狙いやすくなる。分析結果からサービスエース数やファーストサーブ成功率などファーストサーブに関係するデータが他の大会と比較して大きな影響を与えている点から、グラスコートのサービスエースが狙いやすいという特徴を読み取ることが出来る。

5 主成分分析

4つの大会の全ての試合に関する変数の内の「勝敗」を除いたデータを用いて、大会ごとに主成分分析を行った。4つの大会全てで、第1主成分から第5主成分までの累積寄与率が約80%となったが、主成分の特徴に差が現れた。

大会ごとの主成分分析の結果を比較してみると、全仏オープンと全豪オープンの2つの大会は、第1主成分でファーストサーブに関する特徴が現れており、試合においてサービスゲームをキープすることの重要性、特にファーストサーブでポイントを獲得することの重要性が分かる。またウィンブルドン選手権と全米オープンの2つの大会においては、第1主成分では、全ての変数が正の方向を表す結果となった。この二つの大会は、他の二つの大会に比べてサービス側が有利な大会であるということが、ロジスティック回帰分析から分かっている。サーブが有利な大会であるからこそレシーブを獲得することで大きく勝利に近づくことになるので、この2つの大会は、ファーストサーブに関する要素だけでなく、全ての変数が正の方向を示す形となったと考察出来る。全仏オープン、ウィンブルドン選手権、全米オープンの3つの大会の第2主成分は、サーブで攻めているがミスをしてしまったことを表す軸となっている。全豪オープンのみ第2主成分がサービスゲームをキープ出来ているかどうかを表す軸となっている。第3主成分は、全豪オープン、全米オープンでは、ラリーによる特徴が、全仏オープン、ウィンブルドン選手権では、サーブによるポイント獲得に関する特徴が現れている。第4主成分は、全米オープンのみ、レシーブゲームに関する特徴がみられ、残りの3つの大会では、サービスゲームで攻めていく姿勢が見て取れる。第5主成分については、ウィンブルドン選手権でレシーブゲームに関する特徴がみられたが、他の3つの大会では、サーブに関する特徴が現れる結

果となった。

どの大会においても、第1主成分から第5主成分までの間でサーブに関する特徴が多く現れており、大会により差はあるもののサーブが勝利するために重要であると考えられる。

6 クラスタ分析

主成分得点を用いて、どのような特徴を持った試合の群があるか大会ごとにウォード法により分析を行った。

6.1 クラスタ分析結果

大会毎のクラスタ分析の結果、どの大会も「ストレート勝ちに近い勝利を収めた選手の群」「ストレートに近い負けをした選手の群」「接戦の末に勝利した選手の群」「接戦の末に敗れた選手の群」の4つの群に分類出来た。大会ごとの比較を行うと、ウィンブルドン選手権は、他の3つの大会に比べ、「ストレート勝ちに近い勝利を収めた選手の群」の数が少なくなっている。この大会は、サーブ側が有利となる芝のコートで行われるので、相手のサーブをブレイクすることが困難となり、3セット連取してのストレート勝ちが難しい為であると考えられる。逆にサーブ側が、他の大会ほど有利ではない全仏オープンでは、「ストレート勝ちに近い勝利を収めた選手の群」「ストレートに近い負けをした選手の群」の2つの群の選手が多く、相手のサーブをブレイクしやすい環境にあると考えられる。

6.2 選手の分析

クラスタ分析によるデンドログラムを見ると、何回戦の試合であるかという点に関する特徴は見られなかった。

全仏オープンでのラファエル・ナダルの様に勝ちすんだ試合結果の獲得セット数が安定していた選手は、図の下の方で結合をしており、安定した成績を残しながら、勝ち進んでいることがデンドログラムから見て取れた。

7 おわりに

本研究を終えて、勝利に影響を与える要素としては、どの大会でもサーブに関する要素が勝利に対して大きな影響を与えていた。自分のサーブは獲得しなければならないもので、相手に取られてしまうと勝利から大きく遠ざかり、自分のサービスゲームを獲得した上で、レシーブを獲得すると勝利に大きく近づくことが出来るという結果から、自分のサーブがいかに重要であるか知ることが出来た。またコートの特徴が、大会毎の分析により現れていて、試合をするコートに合ったプレーをすることも重要である。

参考文献

- [1] 藤井良宣:カテゴリカルデータ解析, 共立出版, 東京, 2010.
- [2] 中村永友:多次元データ解析法, 共立出版, 東京, 2009.
- [3] スポーツナビ: <http://sports.yahoo.co.jp/> 2014年12月.