

判別分析と時系列可視化による、「ぶよぶよ」における勝敗要因の特徴付け

中西蓮、小松歩夢、中原拓哉

大木公平、西川哲夫

武蔵野大学工学部数理工学科

〒135-8181 東京都江東区有明 3-3-3

TEL: 03-5530-7333, Email: te_nishi@musashino-u.ac.jp

1. はじめに

新型コロナウイルス感染が世界的に流行し、感症対策として在宅時間が増えた。その中でゲーム業界は、話題に上がることも多く、ゲーム機の売り上げが伸び、芸能人が動画配信サービスでゲームを行うようになるなど、需要が高まっている。ゲーム業界に注目が集まる中で、オンラインで楽しめるエンターテインメントとして E-sports はより発展していくものと考えられる。しかし、多様な大会やゲームがある中で、ゲーム内容に精通していないと楽しむことは難しい。解説を聞けばわかるものもあるが、素人目には戦況の理解が難しいものも多くあると考えられた[1]。

ここでは戦況における何が勝利に影響するかを知るために、相手に与えるおじゃまぶよ数や連鎖数、選手のタイプを調べることで、どの要素がどの程度勝利に影響しているのかを明らかにすることを目標とした。

2. 方法

2.1 提供データからの各種情報抽出

1) 使用データ

主催者より提供された以下の情報を用いた。

①対戦ごとのプレーの情報

大会毎での試合の情報を格納した json ファイル。ファイル内には、選手名、大会名、連鎖数、連鎖時間、消したぶよの個数、スコア、全消してきたかどうか、その試合での勝者などの情報が格納されている。

②各対戦の動画 URL の情報

試合動画(全 317 試合)URL の格納ファイル。

2) CSV ファイルの構築

上記①の json ファイルから各種情報を抽出し以下に示す 4 種類の CSV ファイルを作成した。情報の抽出と解析には Python とエクセルを用いて行った。

2)-1 コンペまとめファイルの構築

出場選手名、各選手の取得セット数・本数、勝者などの情報を抽出し、②の URL 情報を加え、1 行に試合関連情報をまとめた CSV ファイルを作成した。

2)-2 連鎖情報まとめファイルの構築

①の情報から連鎖に関する情報を取り出し、以下の 4 種類の CSV ファイルを作成した。

- 連鎖ごとに情報を格納(15,071 回)
- ゲームごとに情報を格納(2,476 ゲーム)
- セットごとに情報を格納 (721 セット)
- 試合ごとに情報を格納 (317 試合)

2)-3 トリガー情報まとめファイルの構築

①の情報から、連鎖をステップごとに分解したトリガーの情報を取り出し、トリガー毎に相手に与えるおじゃまの数、全消しの情報を加えたスコアなどの列を追加した CSV ファイルを作成した。

2)-4 選手ごとの情報のファイルの構築

前述した作成ファイルから試合ごと・ゲームごとに平均連鎖数や相手に与えるおじゃまの数の平均などの情報を、選手名をキーにして取り出し、1 行に選手の情報をまとめた CSV ファイルを作成した。

2.2 分析方法

1) 連鎖ごとの特徴分析

連鎖数が勝敗に影響していると考え、連鎖の特徴付けを行った。①勝敗毎の連鎖数頻度分布、及び②連鎖数毎の連鎖発生時刻の頻度分布解析を行った。

2) ゲームごとの特徴分析

大きな連鎖と相手におじゃまを与えることは、勝敗にとって重要であると考え、以下の分析を行った。

①最後のおじゃまぶよの個数差と連鎖ぶよの総個数差によるゲーム勝敗の判別分析

連鎖で消えたぶよ個数やおじゃまぶよ個数の勝敗への影響を調べるため、勝敗毎の最後のおじゃまぶよの個数差と連鎖ぶよの総個数差の頻度分布や散布図を作成しロジスティック回帰分析による判別分析を行った。

②おじゃまぶよ総個数差と連鎖ぶよの総個数差によるゲーム勝敗の判別分析

上記の分析をおじゃまぶよ総個数差に置き換えて同様な分析を行った。

3) ゲーム内の時系列分析

①連鎖開始順番及び連鎖開始時刻の分析

後打ち有利・先打ち不利の法則[2]が知られているが、これをデータで確認するために、連鎖開始の順番や、連鎖開始時刻の勝敗への影響を調べた。

②おじゃま数の時系列分析

ゲームの中で、連鎖状況がいかに行進し、ゲームの勝敗がいかにかに決まるかを調べるために、「おじゃま数」について、ゲーム中の時間的な変化の可視化を行った。各ゲームにおいて、選手毎の各時点での「おじゃま数」(図 1 A) と各時点までの積算「おじゃま数」(図 2 B) について、それぞれ時系列グラフを作成した。選手間の「おじゃま数」の差、及びその差と

積算連鎖ぶよ個数によって推定した選手1の勝利確率Pの時系列も表示した(結果の項で示す)。

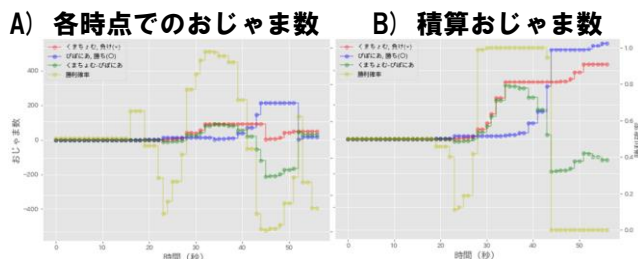


図1 1つのゲームにおける各選手のおじゃま数と選手間のおじゃま数の差の時系列

4) 選手毎の特徴分析

ゲーム毎の勝率などを用いて選手を特徴分析した

3. 分析結果と考察

3.1 連鎖ごとの特徴分析

① 勝敗毎の連鎖数頻度分析

連鎖数毎の頻度を勝利と敗北のゲーム毎に示した(図2)。勝利と敗北どちらでも、連鎖数の増加に伴い頻度が減少しているが、勝利ゲームでは、連鎖数が多いほど、敗北時の連鎖数より、その比率が大きくなることわかる。すなわち、連鎖数が多いことが、ゲームの勝敗に寄与しているものと考えられた。

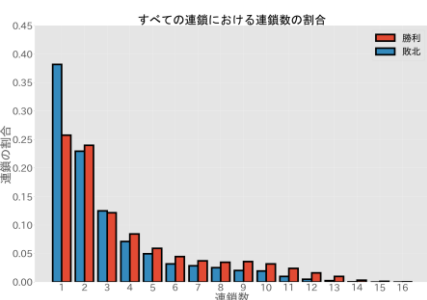


図2 勝敗毎の連鎖数相対頻度分布

② 連鎖数毎の、連鎖が発生した時刻の頻度分析

図3に、各連鎖の連鎖数と連鎖開始時刻の蜂群図(重なったプロットをばらつかせたプロット)を示す。縦軸が時間(秒)、横軸が連鎖数、赤い点が選手1の勝ち、青い点が選手2の勝ちを表す。大きな連鎖ほど、その連鎖数における最初の連鎖開始時刻が遅くなり、最後の連鎖開始時刻が早くなる傾向があった。これは、連鎖数が多い程、連鎖を生じさせるために必要な時間が長いことがその要因であるためと考えられた。

3.2 ゲームごとの特徴分析

各ゲームでの連鎖や時系列の特徴を分析した。

① 最後のおじゃまぶよの個数差と連鎖ぶよの総個数差によるゲーム勝敗の判別分析

ゲームの勝敗を決定する主なものは、その時点までに消した連鎖ぶよの個数と、その時点で相手に与えたおじゃまぶよの個数ではないかと考え、連鎖ぶよの総個数の差の分布(図4)、及びゲームの最後のおじゃまぶよ数の差の分布(図5)を、勝敗を分けて示した(赤が選手1の勝利、青が選手1の敗北を示す)。

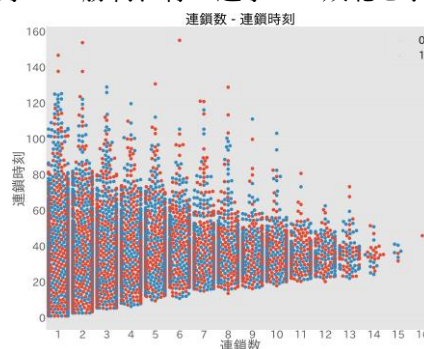


図3 各連鎖の連鎖数と連鎖開始時刻の蜂群図

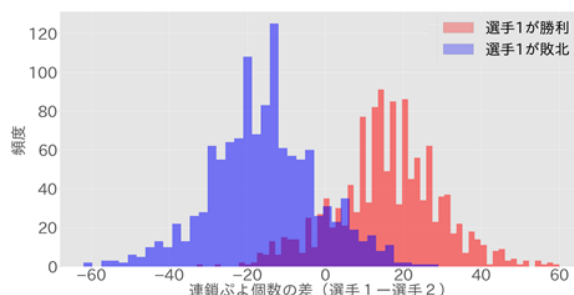


図4 ゲーム毎の連鎖ぶよの総個数の差の頻度分布

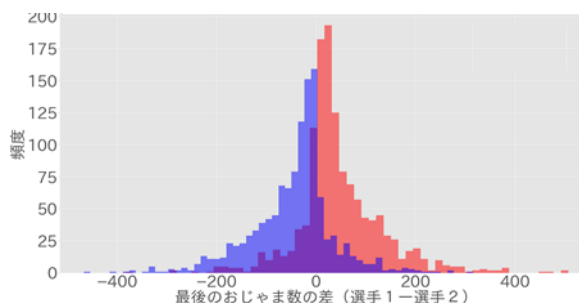


図5 ゲームの最後のおじゃまぶよ数の差の分布

これらのグラフから、連鎖ぶよの総個数の差とゲームの最後のおじゃまぶよ数の差が、ゲームの勝利に大きく貢献していることがわかった。次に、これら2つの量を組み合わせることで、勝敗の判別能力の向上を検討した。図6に連鎖ぶよの総個数の差とゲームの最後のおじゃまぶよ数の差を散布図に表した。

2つの変数を組み合わせたロジスティック回帰分析によると、選手1が勝利する確率Pは、

$$P = \frac{1}{1 + e^{c + a\Delta oj + b\Delta py}}$$
 で与えられる。ここで、 $P=0.5$ の境界線は、 $y = \frac{-bx-c}{a}$ 、 $a = -0.0109$ 、 $b = -0.152$ 、 $c = -0.0354$ 、 Δoj : おじゃまぶよ数の差、 Δpy : 連鎖ぶよの総個数の差 である。判別精度として、正解率=0.90、適合率=0.90、再現率=0.89 が得られた。

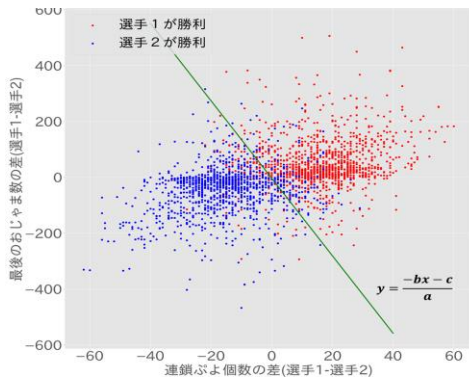


図6 連鎖ぶよの総個数の差とゲームの最後のおじゃまぶよ数の差の散布図

②おじゃまぶよ総個数差と連鎖ぶよの総個数差によるゲーム勝敗の判別分析

最後のおじゃまぶよ数だけでなく、おじゃまぶよの総数の影響を見るために、勝敗毎におじゃまぶよの総数の差の分布を調べたところ、図7に示すように、頻度分布が勝敗毎にきれいに分離した。そこで、最後のおじゃまぶよ数を、おじゃまぶよの総数に置き換えて①と同様な分析を行った(図8)。

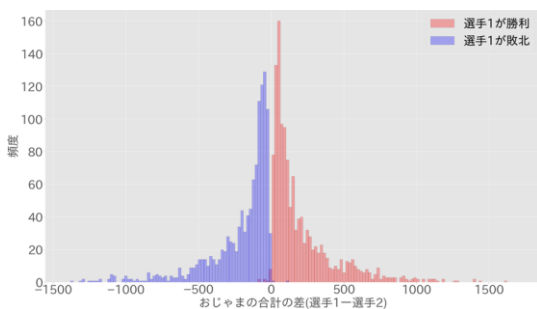


図7 おじゃまぶよ総数の差の分布

ロジスティック回帰分析を同様に行った結果、選手1が勝利する確率Pとして、 $P = \frac{1}{1 + e^{c + a\Delta oj + b\Delta py}}$
 $a = -0.0805, b = -0.159, c = -0.0440$ が得られ、判別精度として、正解率=0.997、適合率=0.993、再現率=1.0という非常に高い値が得られた。

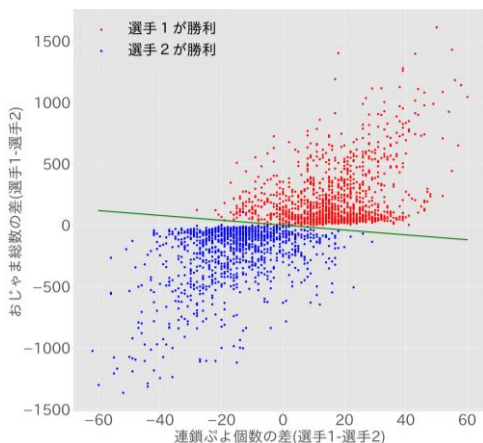


図8 連鎖ぶよの総個数の差とおじゃまぶよ総数の差の散布図

3.3 ゲーム内の時系列分析

①連鎖開始順番及び連鎖開始時刻の分析

後打ち有利・先打ち不利の法則[2]の影響を見るために、先攻・後攻の勝敗への影響を調べた(図9)。

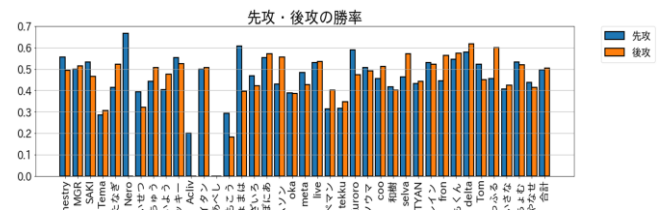


図9 先攻・後攻の勝敗への影響

図9より、選手毎では、先攻・後攻によって勝率が数%変動するが、全体としては後攻の勝率は50.36%であり、母比率50%の比率の検定によるP値は約0.6であり、統計的に有意ではなかった。

次に、おじゃまぶよの開始時刻の分布を勝敗毎に調べた(図10)。このように、22秒を境として、早い時間帯と遅い時間帯での分布に系統的な違いがあり、遅いほど勝率が高いことが分かった。22秒未満と以上のゲーム勝率を求めると、それぞれ47.0%と57.5%となり、母比率50%の比率の検定によると、P値は約 10^{-4} 以下となり、勝率の違いはランダムなばらつきの結果とはいえないことが分かった。これには、後打ち有利の考え方が反映している可能性があると考えられる。

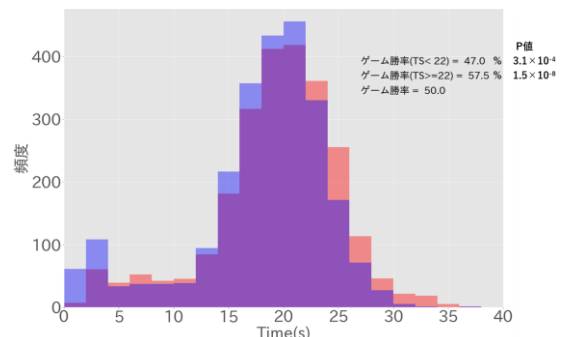


図10 おじゃまぶよの開始時刻の分布

選手毎のおじゃまぶよの開始時刻の分布を図11に示す。選手毎に分布の特徴が異なることが分かる。特に選手「くまちょむ」は、10秒以下に高いピークが存在し、ユニークな分布を示している。

②おじゃま数の時系列分析

ゲームの中で連鎖状況がいかに行進し、ゲームの勝敗がいかにかに決まるかを調べるために、「おじゃま数」について、ゲーム中の時間的な変化の可視化を行った。図12に時系列データの例を示す。おじゃま数を赤と青の線で、おじゃま数の差を緑の線で、ロジスティック回帰によって推定した選手1の勝利確率Pを黄色の線で、時間軸に沿って表示しており、これらにより、各時点での「おじゃま数」と、推定勝利確率をモニタ

一できる。

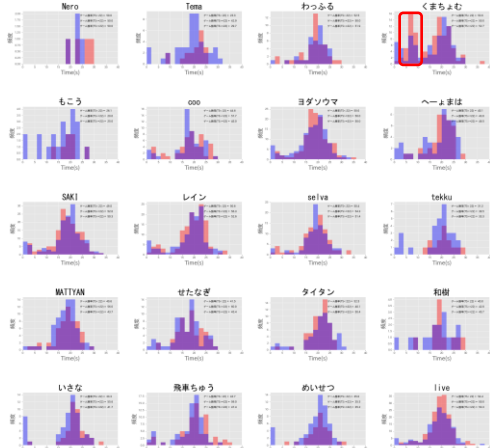


図 11 選手毎のおじゃまぶよの開始時刻の分布

図 12 の A)は、一方のおじゃまぶよが先行し、他方が約 10 秒遅れて追いかけるが、間に合わなかった例、B)は、両方が交互におじゃまぶよを繰り返す例、逆転を繰り返した例である。C)と D)は、A)と類似しているが、C)は最後に逆転した例で、D)はさらに最後に逆転した例である。全ゲームを見ると、C)、D)程度のシンプルなパターンが多く、A)、B)のように極端にシンプルまたは複雑なパターンは多くなかった。おじゃまぶよを交互に繰り返す場合が多く、その間隔は 5~10 秒程度が多かった。全く同じパターンの出現が良くみられ、パターンの類型化が可能であると考えられた。

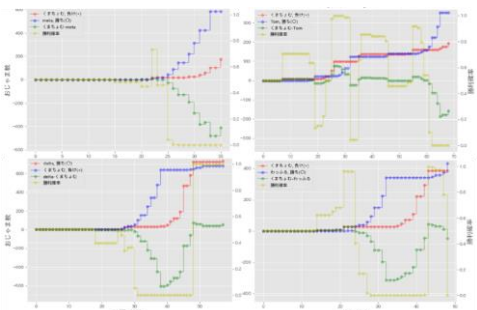


図 12 おじゃま数の時系列の典型的パターン

3.4 選手毎の特徴分析

①平均連鎖数とゲーム勝率の関係

選手毎の平均連鎖数とゲーム勝率の散布図を示す (図 13)。連鎖数が多いほどゲーム勝率が高く、相関係数は 0.61 であった。

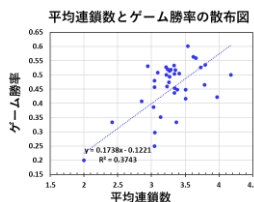


図 13 平均連鎖数とゲーム勝率の散布図

②ゲーム勝率と試合勝率の関係

選手毎のゲーム勝率は、多くの場合試合勝率と異なる (図 14)。ゲーム勝率が 50%以上では、試合勝率の方が大きいグループ、50%以下では試合勝率が小さいグループの 2 つに分かれ

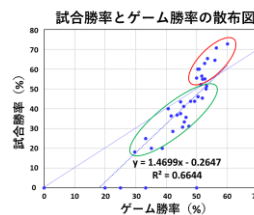


図 14 ゲーム勝率と試合勝率の散布図

た。ゲームの組み立て方等が関係している可能性が考えられ、今後の検討課題である。

③おじゃまぶよ数の差と連鎖ぶよの個数差による勝率の推定

確率 P で各選手の推定勝率を求め、実際の勝率に対し

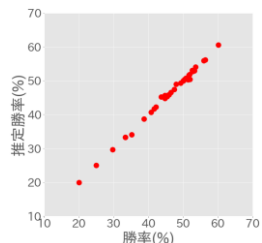


図 15 おじゃまぶよ数の差などによる勝率の推定

④連鎖開始時刻の違いによるゲーム勝率の違い

連鎖開始時刻が 22 秒未満と以上でのゲーム勝率を、元のゲーム勝率に対してプロットした(図 16)。ゲーム勝率が 50%以上では、22 秒未満と以上での差が小さいが、50%以下では、22 秒以上の方がゲーム勝率が大きかった。

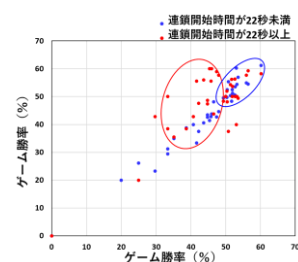


図 16 連鎖開始時刻の違いによるゲーム勝率の違い

4.結論と今後の課題

ぶよぶよの 317 試合のデータから勝利に貢献する要因を分析した結果、以下のことがわかった。

- 1) 連鎖数が多いほど、連鎖開始時刻が遅くなり、また勝率が高くなった。
- 2) 各ゲームで、相手に与えるおじゃまぶよ総数が多い程、また連鎖ぶよ数が多い程、勝率が高かった。そこで、相手に与えるおじゃまぶよ総数差と連鎖ぶよ総数差を説明変数としてロジスティック回帰分析を行うことにより、各選手の勝利確率を予測できた。
- 3) 連鎖開始時刻が 22 秒未満と 22 秒以上の場合のゲーム勝率は、後者が統計的に有意に高いことがわかった。また、おじゃまぶよ数の典型的な時系列パターンが存在することがわかった。今後の課題としては、戦略のパターンとその効果の分析や盤面情報の分析などが考えられる。

5.謝辞

本研究で用いたデータは、『情報・システム研究機構 統計数理研究所 医療健康データ科学研究センター』及び『株式会社 SEGA』より提供していただきました。深く感謝いたします。

参考文献

- [1] 日本での「eスポーツ」の認知度と理解度は? <https://esports-world.jp/column/91>
- [2] ぶよぶよ：後打ち理論は正しいか? <https://toorun12.theblog.me/posts/834317/>