

# ポーカーの統計を用いたプレイヤーの実力評価

大平 翼<sup>†</sup> 横山 昌平<sup>††</sup>

<sup>†</sup> 東京都立大学システムデザイン学部情報通信システムコース 〒191-006 東京都日野市旭が丘 6-6

<sup>††</sup> 東京都立大学院 システムデザイン研究科 〒191-0061 東京都日野市旭が丘 6-6

E-mail: <sup>†</sup>tohira-tsubasa@ed.tmu.ac.jp, <sup>††</sup>shohei@tmu.ac.jp

**あらまし** 日本では5カードを引いて役を作る、ドローポーカーが有名である。しかし近年、インターネット上での日本人ポーカープレイヤーの活躍により、ノーリミットホールデムと呼ばれるポーカーのゲームの1つが日本での人気を高めている。ノーリミットホールデムとは2枚の手札と5枚のコミュニティカードと呼ばれるすべてのプレイヤーが共有できるカードを組み合わせる役を作り勝負を行う。手役を公開する前に相手を下すことができるため、手役の強さ以外の部分で勝つことができ、技術的な要素が存在するゲームである。多くの試行回数をこなすことで、プレイヤー間の運の要素が均等化され技術差がでてくる。そのため、正確な実力を見積もるための試行回数というのは非常に多くなってくる。そこで本研究では、少ない試行回数から実力を推定するために、ポーカープレイヤーのプレーの傾向を表す統計を膨大なプレー履歴から集め、それらをニューラルネットワークに用いて学習させることで、本当に強いポーカープレイヤーであるかそうでないプレイヤーかを試行回数のデータから判別する分類器を構築した。完成した分類器は90%以上の正解率で、プレイヤーの実力を判別することができるようになった。

**キーワード** ソーシャルビッグデータ, 統計的学習, 機械学習

## 1 はじめに

ポーカーは世界中で楽しまれているトランプゲームの1つであり、カジノでも採用されているゲームであるポーカーは山札からプレイヤーに手札が配られ、その手役の強さで勝者を決定する。相手の手札を最後に見せ合うまではわからない不確定不完全情報ゲームである。不完全情報ゲームとは、プレイヤー同士に共有されていない情報があるゲームのことである。ポーカーでは、相手の手札とゲームの進行中では見ることができないので不完全情報ゲームと分類される。そのほかの不完全情報ゲームの例として麻雀、バックギャモン、人狼ゲームがあげられる。不完全情報ゲームとは対照的に、プレイヤー同士が全ての情報を共有しているゲームのことを完全情報ゲームという。完全情報ゲームの例として、チェス、将棋、オセロ、囲碁があげられる。ポーカーは、最後に手役を見せ合うまでに、ベットやコールなどのいくつかのアクションを行うことができ、手役が弱くてもブラフで相手を降ろし勝つといった、戦略的な部分を含んだゲームとなっている。運の要素と実力の要素が混合されており、カジノでもほかのゲームと比べてプレイヤーのスキルの割合が勝負の結果に大きく反映されるゲームだと考えられている。

2020年、カジノのポーカーで生計を立てている日本人たちの活躍がYouTubeでの活動を通して目立つようになり、日本でのポーカーの人気度が高まった。カジノでプレーできるポーカーには主に2種類の形式がある。1つ目はキャッシュゲーム形式のものである。キャッシュゲームでは、実際にチップと現金を交換してプレーするものである。2つ目は、トーナメント形式のものである。トーナメントでは、参加費を払うことで出場できる。内容は、参加したプレイヤーの中で優勝者を一人決めるというも

のである。賞金は、参加費の中からカジノ側の手数料を引いた分がプレイヤー割り当てられ、賞金の分配方法として、優勝者一人にすべての賞金が与えられるものや、成績上位10パーセントのプレイヤーに均等に分配するものなど様々なやり方がある。現在、日本国内にはカジノは存在していないが、現金を賭けないポーカーがプレーできるアミューズメントカジノ施設は多く存在している。日本では2016年に特定複合観光施設区域の整備の推進に関する法律が成立したことにより、カジノ施設が解禁された。[1] ポーカーはカジノゲームの1つとして導入される可能性がある。COVID-19の影響による自粛によって、自宅を過ごす時間が増え、オンラインポーカーのプレー人口が増加した。オンラインポーカーは、PCやスマートフォンに専用のソフトをインストールして、ソフト上でプレーする。実際の現金をかけてプレーできるサイトも存在しており、[2] オンラインポーカー上で数万ドル稼ぐプレイヤーも存在している。日本には現金での賭博を禁止する法律が存在するため、現金でのオンラインポーカーを行うべきではないが、世界では多くの国でプレーすることが可能である。オンラインポーカーをプレーするうえでの補助ツールとして、自分や対戦相手のプレー履歴の情報を収集し、成績やアクションの統計を出力してくれるソフトが存在している。このソフトは多くのオンラインプレイヤーに利用されており、オンラインポーカーで実力を上げるためには必須のツールだと考えられている。その理由は、ポーカーはどのタイミングで、どのような頻度で、どのようなアクションを行ったかということが戦略的に重要な部分であり、ソフトを使用することで、それらの情報を得ることができるからである。ここでの実力のあるポーカープレイヤーとは、ポーカーをプレーし続けた場合、所持しているチップの量が増え続けるプレイヤーのことを指す。実力のあるポーカープレイヤーであるかどうか

を判定するうえで重要となる点は、ポーカーの分散の影響である。ポーカーは1回の勝負では、その時にできた役の強さが勝敗に大きく影響を与える。1回のゲームであるプレイヤーがストレートフラッシュを完成させた場合、ワンペアのような弱い役しかできていないプレイヤーは、ブラフして相手降ろさない限り、ストレートフラッシュができたプレイヤーに勝負で勝つことはない。これは偶然できた役の強さに依存する運の要素が非常に強い勝負となる。しかし、通常ポーカーは何ゲームも行うので、プレイヤー同士ができる役の強さの頻度というものが均等化されていく。このようなプレイヤー同士の条件が平等な状況では、技術的な部分が勝負の結果に影響を与える。技術があるプレイヤーは、ワンペアで10枚分のチップを稼ぐことができるが、技術のないプレイヤーはストレートフラッシュを完成させていたとしても5枚分のチップしか稼げないということがある。役の強さは勝敗やチップの稼ぎに影響を与えるが、複数回ゲームを行うポーカーではチップの量を増やし続けるために技術的な部分が重要となってくる。よって、ポーカーは少ない試行回数では運の要素が強く働くため、正確な自身の強さを知るとは難しくなる。しかし、正確な実力を知るために行わなければならない試行回数はかなりの分量となってくる。普段オンラインポーカーをプレーする人がプレーするペースでは、1年以上かかってしまう可能性すらある。そこで本研究では、試行回数の少ないプレイヤーが実力のあるポーカープレイヤーであるかどうかを判別する。そのためにニューラルネットワークを用いて、運の影響が比較的に小さくなるように試行回数を多くこなしプレイヤーの統計と成績を利用して、分類器を構築する。

本研究では、ノーリミットホールデムを研究の対象とした。ノーリミットホールデムはポーカーの種類の1つであり、世界中で一番人気のあるゲームだと考えられている。カジノでプレーできるポーカーも基本的にはノーリミットホールデムである。以降、本文でのポーカーはノーリミットホールデムのことを指す。

## 2 関連研究

本研究は、機械学習を用いて、実力のあるポーカープレイヤーを判別するものである。よって本研究は、ボードゲームに機械学習を適用した研究やポーカーの研究と深く関連する。

### 2.1 完全情報ゲームと機械学習

ボードゲームと機械学習を組み合わせた研究はについて、[3]久保は、将棋 ai に使われる評価関数の特徴ベクトルを、熟練した指し手の棋譜を利用した機械学習で発見する手法を提案した。棋譜中に現れるすべての局面に対応する指し手と Minimax 探索の結果がよく一致する特徴ベクトルを求めるために、最適制御法を適用した。

[4]大森らは、棋士の棋風である攻めと受けそれぞれの棋譜を機械学習させることで、棋風が攻めか受けであるかの区別を行う研究をした。棋譜の攻めと受けの区別を行うために、指した手の駒と自分の王とのマンハッタン距離と指した手の駒と相手

の王とのマンハッタン距離と比較して、自分の王との距離が近い時は受けの手、相手の王との距離が近い時は攻めの手として、攻めの手が多い時は攻めの棋譜、受けの手が多い時は受けの棋譜として分けた。評価関数の機械学習には Bonanza を用いた。

### 2.2 不完全情報ゲームと機械学習

[5]尾崎らは牌譜を機械学習させ、麻雀プレイヤーの狙っている役の予測を行った。LSTM (Long Short-Term Memory) モデルに牌譜から抽出した盤面の情報を入力データとして学習させ、麻雀の役の1つである清一色の予測する手法を提案した。

### 2.3 ポーカー AI

ポーカーの AI に関する研究については、[6], [7]Brown らはノーリミットホールデムの AI である Libratus と Pluribus を開発した。Libratus はヘッズアップノーリミットホールデムの AI で Pluribus は多人数のノーリミットホールデムの AI である。これらの AI はゲームの抽象化を行うことでゲームの状態数を減らし、MCCFR(Monte Carlo counterfactual regret minimization) と呼ばれるアルゴリズムを使用して Blueprint と呼ばれる戦略を作成。この戦略は初期ラウンドのみ適応され、後半のラウンドではリアルタイムでサブゲームを解くことで戦略を作成した。また、Blueprint にはないベットサイズを人間がとってきた場合、そのサイズを分析し Blueprint に分析した戦略を追加し人間のプレイヤーに適応した。Libratus は4人のヘッズアップノーリミットホールデムのスペシャリストと合計12万ハンド対戦し、人間のプレイヤーに対して勝利した。Pluribus は実力を測る実験として、5体の AI と1人の人間での対戦形式と、1体の AI と5人の人間での対戦形式で実力を測定した。Pluribus は両方の形式において、良い成績を残した。

## 3 準備

本章では、提案手法の前提知識について述べる。まず、本研究の対象ゲームとなっているノーリミットホールデムのルールについて説明する。次に、ポーカーにおける統計、本研究で使った統計を収集するためのソフト、ポーカープレイヤーの実力評価のための指標、試行回数についてまとめる。

### 3.1 ノーリミットホールデムのルール

#### 3.1.1 アクション

本小節ではゲーム中のアクションを表す名称を述べる。

- ベット:チップを賭けること
- レイズ:ベットに上乗せしたチップを賭けること
- コール:ベットやレイズにのること
- チェック:チップを賭けず、様子見すること
- フォールド:ベットやレイズに対して降りること

#### 3.1.2 ゲームの基本

ノーリミットホールデムでは、まずプレイヤーそれぞれに2枚の手札が配られる。それから、コミュニティカードが合計5枚、場にオープンされる。コミュニティカードとは、ゲームに参加しているすべてのプレイヤーが共有して使用できるカードの

ことである。プレイヤーは手札の2枚と、コミュニティカードの計7枚から、一番強い役となる5枚を選び勝負を行う。ゲーム中にはベッティングラウンドと呼ばれるアクションを行うための時間がある。ベッティングラウンドが1回のゲームで4回存在する。それぞれのラウンドで、ベットやチェックなどのアクションを行う。1回目のラウンドをプリフロップ、2回目のラウンドをフロップ、3回目のラウンドをターン、4回目のラウンドをリバーとそれぞれ名称がある。また、フロップ、ターン、リバーの総称をポストフロップという。コミュニティカードは2回目のラウンドで3枚オープンされ、3回目のラウンドで4枚目、4回目のラウンドで5枚目と、順番にカードがオープンされる。5枚目のコミュニティカードがオープンされ、プレイヤーが役の勝負を行う段階のことをショーダウンと呼ぶ。ショーダウンでは残っているすべてのプレイヤーが手札を見せ合い、一番役の強いプレイヤーがその場にあるチップを総取りする。また、各ベッティングラウンドにおいて、残りのプレイヤーが1人となり、それ以外のプレイヤーがフォールドした状態となった場合、ショーダウンに向かうことなく、残ったプレイヤーが場のチップを総取りできる。フォールドしたプレイヤーは手札を公開する義務はない。すべてのラウンドにおいて、ベット額の上限はない。またレイズの回数についても制限はない。

### 3.1.3 ポジション

プレイヤーの座席位置のことをポジションと呼ぶ。本小節では、各ポジションの名称、位置、ルールについて述べる。

まず、ボタンと呼ばれるポジションのプレイヤーが決められる。ボタンの左隣のポジションをスモールブラインド、スモールブラインドの左隣を、ビッグブラインドとそれぞれ呼ぶ。ポジションは1回のゲームが終わるごとに、右隣のプレイヤーに移動する。スモールブラインドとビッグブラインドは、プリフロップで決められた額をそれぞれ強制的にベットをしなければならない。強制ベット額は、スモールブラインドはビッグブラインドがベットする額の半分をベットする。強制ベット額は変動することではなく、1回のゲームが終わるごとに、新たなスモールブラインドとビッグブラインドが強制ベットを行う。ビッグブラインドが強制的にベットする額が、ゲームでの最小ベット額となる。

### 3.1.4 ゲームの進行

まず、プリフロップでの流れについて説明する。アクションはビッグブラインドの左隣のプレイヤーから始まる。アクションは時計回りで行っていく。プリフロップでレイズが発生した場合、それに対してさらなるレイズが発生せず、ほかのプレイヤーのうちの誰かがコールした場合、レイズしたプレイヤーとコールしたプレイヤーがポストフロップに進む。レイズに対して他すべてのプレイヤーがフォールドした場合、レイズしたプレイヤーが場にあるチップすべてを獲得してゲームは終了する。レイズに対してさらなるレイズが入った場合は、上記のレイズが1回入った場合と同じ流れが繰り返される。プリフロップで、プレイヤーが全てのチップを賭けた場合、チップを賭けあったプレイヤー同士で手札をその段階で公開し、コミュニティカードが5枚すべてオープンされ役の勝負を行う。レイズが発生せず、誰

かがビッグブラインドの強制ベットにコールした場合、ポストフロップに進む。ビッグブラインドまですべてのプレイヤーがフォールドした場合、ビッグブラインドは自分が強制ベットしていたチップとスモールブラインドが強制ベットしていたチップを獲得する。

ポストフロップでは、基本的にプリフロップと同じ流れでゲームを進めるのだが、アクションはスモールブラインドから始めなければならない。

#### 3.1.5 役について

この小節では、役の種類や役の強さについて説明する。まず、同じ役同士での勝負となった場合、ランクの高いカードを使っている役のほうが勝つ。カードのランクは、Aが一番強く、二番目にKと続き、Q, J, T, 9, 8, 7, 6, 5, 4, 3, 2とランクが弱くなっていく。以下にノーリミットホールデムの役を強い順で挙げる。

- ストレートフラッシュ

5枚のカードのランクが連番であり、スートが全て同じ役のこと。

- フォーオブアカインド

5枚中4枚のカードが同じランクである役のこと

- フルハウス

5枚中3枚のカードのランクが同じであり、残り2枚のカードのランクも同じである役のこと。

- フラッシュ

5枚のカードのスートが全て同じである役のこと

- ストレート

5枚のカードのランクが連番である役のこと

- スリーオブアカインド

5枚中3枚のカードのランクが同じで、残り2枚のランクは異なる役のこと

- ツーペア

5枚中2枚のカードのランクが同じで、残りの3枚のうち2枚のカードのランクが同じ役のこと。

- ワンペア

5枚中2枚のカードのランクが同じであり、残りの3枚のランクは全て異なる役のこと

- ハイカード

上記の役に該当しない5枚のこと

## 3.2 統計について

本研究ではポーカーにおける、特定の状況下における特定のプレーの頻度を表す統計を使用する。統計の例として、プリフロップでの参加率、ポストフロップでのベット率、コール率、フォールド率などが挙げられる。これらの統計はプレイヤーのチップの増減に影響を与えるアクションに関するものであり、プレイヤーのポーカーに関する技術的な部分を反映していると考えられる。これらの統計はプレイヤーごとで大きく異なるので、ポーカーの実力があるプレイヤーかそうでないプレイヤーかを分けるときに利用できる要素だと考える。これらの統計を機械学習を使う上での特徴量として利用することで、長期的に勝て

るプレイヤーかそうでないプレイヤーかを判別する分類器を構築する。

### 3.3    トラッキングソフト

オンラインポーカーサイトではプレーごとにその時の状況やアクション内容が詳細に記録されたハンド履歴がテキストファイルとして作られる。このハンド履歴は、自由にダウンロードすることができる。プレイヤーはトラッキングソフトと呼ばれるハンド履歴を読み込み、あらゆるポーカーに関する情報を出力してくれるソフトを使用して、自分や相手の分析を行う。トラッキングソフトは、プレーごとにおけるチップの増減、プレーに関する統計を自動で出力してくれる。またそれらを日付ごとや、役の強さなどでフィルタリングして統計を出すこともできる。本研究では、トラッキングソフトの1つである [8]Holdem manager3 を使用して、プレイヤーごとの統計をまとめ、学習に使用するデータセットを作成した。

### 3.4    実力判別の指標

ポーカーにおいて、実力を判断する指標として、最低ベット額を基準に 100 プレーあたりにチップをどれだけ増やしているかというものがある。この指標の単位は  $[bb/100]$  と表される。bb とは最低ベット額を 1 とした時の単位である。最低ベットが 1 ドルと 2 ドルのテーブル両方で最低ベット額を 1bb と表記することができる。本研究では  $[bb/100]$  の値のことを win rate と呼び、これをプレイヤーの実力とする。

### 3.5    試行回数について

本研究では、5 万回のプレーしたデータから、プレイヤーが長期で負けないプレイヤーかどうかを判定する。5 万回のプレー回数を選択した理由としては、多くのプレイヤーが区切りをつけるのによく使用するプレー量だからである。ただし、5 万回のプレー回数は実力を見積もるには少ないと考える。正確な実力を見積もるためにはどれだけ試行回数を正確に示すことは不可能だが、本研究ではできるだけポーカーの分散の要素を減らすために、30 万回以上のプレー回数のあるプレイヤーの統計と win rate を学習用データに使った。

## 4    提 案 手 法

### 4.1    データセット

データセットの作成において、オンラインポーカーサイト「pokerstars」での 6 人制のノーリミットホールデム、最低ベット額が 0.1 ドルのキャッシュゲームのテーブルのプレー履歴を利用した。このプレー履歴は [8]HHDEALER で購入した。データセット作成に利用したプレー履歴の数は 8 千万である。学習には、30 万回以上のプレー回数がある 50 人のプレイヤーの統計と win rate を利用した。ニューラルネットワークに使用するデータは 5 万回プレーにおける統計情報である。そのため、30 万回以上のプレー回数のあるプレイヤーの統計を期間で区切り、5 万回プレーごとに統計を出した。表 1 に「pokerstars」で取得できる、プレー履歴の例を示す。

```
PokerStars Zoom Hand #188554302719: Hold'em No Limit ($0.02/$0.05) - 2018/07/10 11:23:05 JST
[2018/07/09 22:23:05 ET]
Table 'Donati' 6-max Seat #1 is the button
Seat 1: arestantik13 ($5.07 in chips)
Seat 2: technokok ($7.52 in chips)
Seat 3: TOQOOR ($5.13 in chips)
Seat 4: aaaa1365 ($10.38 in chips)
Seat 5: A.M.Knuckle ($5.05 in chips)
Seat 6: Pharaohly ($8.89 in chips)
technokok: posts small blind $0.02
TOQOOR: posts big blind $0.05
*** HOLE CARDS ***
Dealt to TOQOOR [Ac 3s]
aaaa1365: folds
A.M.Knuckle: folds
Pharaohly: folds
arestantik13: folds
technokok: calls $0.03
TOQOOR: checks
*** FLOP *** [Qh 8h 8d]
technokok: bets $0.10
TOQOOR: calls $0.10
*** TURN *** [Qh 8h 8d] [Ah]
technokok: checks
TOQOOR: bets $0.10
technokok: calls $0.10
*** RIVER *** [Qh 8h 8d Ah] [4s]
technokok: checks
TOQOOR: bets $0.20
technokok: folds
Uncalled bet ($0.20) returned to TOQOOR
TOQOOR collected $0.48 from pot
TOQOOR: doesn't show hand
*** SUMMARY ***
Total pot $0.50 | Rake $0.02
Board [Qh 8h 8d Ah 4s]
Seat 1: arestantik13 (button) folded before Flop (didn't bet)
Seat 2: technokok (small blind) folded on the River
Seat 3: TOQOOR (big blind) collected ($0.48)
Seat 4: aaaa1365 folded before Flop (didn't bet)
Seat 5: A.M.Knuckle folded before Flop (didn't bet)
Seat 6: Pharaohly folded before Flop (didn't bet)
```

図 1 プレー履歴

このようなプレー履歴を 8 千万プレー分をトラッキングソフトに読み込ませた。トラッキングソフトが出力するプレイヤーの統計情報は表??のような形となる。

### 4.2    特徴量の選択

学習に使用する統計を選択するうえで、統計の正確性を考慮して、5 万プレーでサンプル数が 100 以上集まる統計だけを特徴量として学習させる。サンプル数が少ないと正しい統計が取れていない可能性があり、うまく学習を行うことができない。学習に使用する特徴量の候補として、プリフロップでの統計を 12 個、ポストフロップでの統計を 32 個選択した。プリフロップの統計とその内容を図 2 に示す。

統計	内容
VPIP	参加率
Limp	リンプでの参加率
PFR	レイズでの参加率
Cold Call	レイズに対してコールで参加する確率
3bet In Position	インポジションで3ベットする確率
3bet Out of Position	アウトオブポジションで3ベットする確率
Fold to 3Bet IP	インポジションで3ベットに対してフォールドする確率
Fold to 3Bet OOP	アウトオブポジションで3ベットに対してフォールドする確率
4Bet IP	インポジションで4ベットする確率
4Bet OOP	アウトオブポジションで4ベットする確率
Fold to 4Bet	4ベットに対してフォールドする確率
5Bet (Raise 4Bet)	5ベットする確率

図 2 プリフロップの統計

ポストフロップの統計とその内容を図 3 に示す。

統計	内容
WTSD%	フロップを見たうえでショーダウンまで行く確率
W\$WSF	フロップを見たときの勝率
W\$SD%	ショーダウンでの勝率
A\$g	(ベット数+レイズ数) ÷ (コール数)
River Call Efficiency	(リバーでコールして勝った額) / (リバーでコールした額)
Flop CBet IP%	フロップインポジションでのCベット確率
Flop CBet OOP%	フロップアウトポジションでのCベット確率
Flop Fold vs Cbet OOP	フロップアウトオブポジションでCベットにたいしてフォールドする確率
Flop Raise Cbet OOP	フロップアウトオブポジションでCベットにたいしてレイズする確率
Flop Fold vs Cbet IP	フロップインポジションでCベットにたいしてフォールドする確率
Flop Raise Cbet IP	フロップインポジションでCベットにたいしてレイズする確率
Bet Turn OOP vs Missed Flop CBet OOP	アウトオブポジションでフロップCベットをされなかったときに、ターンでベットする確率
Turn CBet IP%	ターンインポジションでのCベット確率
Turn CBet OOP%	ターンアウトオブポジションでのCベット確率
Bet Flop vs Missed CBet	フロップでCベットされなかったときに、ベットする確率
Fold Turn IP Skip CBet	インポジションでフロップCベットせずに、ターンのベットに対してフォールドする確率
Delayed Turn CBet% OOP	アウトオブポジションでフロップCベットせず、ターンでベットする確率
Check Fold Flop% As PFR	プリフロップをレイズで参加し、アウトオブポジションのフロップでチェックフォールドする確率
Check Raise Flop% As PFR	プリフロップをレイズで参加し、アウトオブポジションのフロップでチェックレイズする確率
Delayed Turn CBet% IP	インポジションでフロップCベットせず、ターンでベットする確率
Turn Fold vs Cbet OOP	アウトオブポジションでターンのCベットに対してフォールドする確率
Bet River OOP vs Missed Turn CBet OOP	アウトオブポジションでターンCベットされなかったときに、リバーでベットする確率
Turn Fold vs Cbet IP	インポジションでターンのCベットに対してフォールドする確率
Skip Turn CB and Fold To River Bet IP	インポジションでターンのCベットをせずに、リバーのベットに対してフォールドする確率
Check Fold Turn OOP Skip CBet	アウトオブポジションでターンCベットせずに、チェックフォールドする確率
Check Raise Turn OOP Skip CBet	アウトオブポジションでターンCベットせずに、チェックレイズする確率
Bet Turn IP vs Missed CBet	インポジションでターンCベットされなかったときに、ベットする確率
Flop CBet-Fold IP TOT	インポジションでフロップCベットをし、レイズに対してフォールドする確率
Flop CBet-Fold OOP TOT	アウトオブポジションでフロップCベットをし、レイズに対してフォールドする確率
River CBet OOP%	リバーアウトオブポジションでのCベット確率
River CBet IP%	リバーインポジションでのCベット確率

図 3 ポストフロップの統計

Cベットとは、continuation betの略称のことである。Cベットは、プリフロップで最後にレイズを行ったプレイヤーがフロップで引き続きベットを行うプレーのことを指す。またフロップでCベットを行ったプレイヤーが、ターンやリバーでも引き続きベットを行うこともCベットと呼ぶ。Cベットはポーカーの戦術としてはとても一般的なものなので、Cベットを行う頻度は重要度が高い。

### 4.3 ラベリング

学習のために収集したプレイヤーのデータにおいて、win rateの平均が3.72[bb/100]となった。本研究では、30万回以上のプレー回数でのwin rateが3.72[bb/100]よりも大きな値である場合、実力のあるプレイヤーであるとみなす。区切った5万回プレーごとのそれぞれのデータに対して、実力のあるポーカープレイヤーの統計かどうかのどちらかにラベリングした。ラベリングは30万回以上のプレーで勝っているプレイヤーであれば、分割した5万回のプレーでwin rateが3.72[bb/100]以下

だとしても実力のあるプレイヤーの統計としてラベリングを行う。こうすることで5万回のプレー回数という少ない試行回数による分散の影響に惑わされず、5万回プレーの情報に対して正しいラベル付けを行うことができる。

### 4.4 ニューラルネットワーク

機械学習のモデルの1つである、ニューラルネットワークを構築して、実力のあるプレイヤーかどうかを判別する分類器を構築する。本研究で使用するニューラルネットワークの構造は表2に示す。

表 1 ニューラルネットワークの設定

入力層の入力数	選択した統計数
隠れ層の数	2 層
隠れ層の入力数	50 個
出力層の入力数	50 個
入力層, 隠れ層の活性化関数	relu
出力層の活性化関数	sigmoid
最適化アルゴリズム	RMSprop
目的関数	二値交差エントロピー

ニューラルネットワークを利用した、プレイヤーの実力判定の分類器の構造を図4に示す

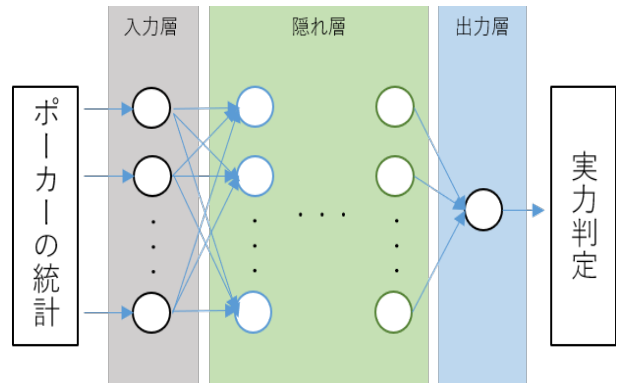


図 4 ニューラルネットワークの構造

### 4.5 Permutation Importance

本研究ではどの統計が結果に寄与しているのかを調べるために、Permutation Importanceという手法を行う。この手法は、テストデータの1つの特徴量の値をランダムに入れ替え、モデルに適用して正解率を出し、どのように特徴量が正解率に影響を与えているかを確認するものである。テストデータの特徴量の並び替えを図5に示す

並び替え			
プレイヤー名	特徴量1	特徴量2	特徴量3
A	1	10	100
B	2	20	200
C	3	30	300
D	4	40	400

プレイヤー名	特徴量1	特徴量2	特徴量3
A	1	30	100
B	2	40	200
C	3	20	300
D	4	10	400

図 5 テストデータの特徴量の並び替え

## 5 検証と結果

プリフロップに関係する統計 12 個、ポストフロップに関係する統計 32 個、計 44 個を特徴量として、ニューラルネットワークの学習を行った。検証結果は 4 分割交差検証によるものである。学習したモデルにテストデータを適用した場合の正解率を表 2 に示す。モデルは 90 %以上の正解率を出すことができた。

表 2 全統計モデルの正解率

acc	0.936363622546196
-----	-------------------

Permutation Importance を行った結果を表 3 に正解率が小さい順で示す。

Permutation Importance を実行した結果を統計の種類別に分け、正解率の平均値を出した。以下のように分類し、それぞれの正解率の平均値を表 4 に示す。

- プリフロップに関係する統計
- ポストフロップに関係する統計
- フロップに関係する統計
- ターンに関係する統計
- リバーに関係する統計
- インポジションに関係する統計
- アウトオブポジションに関係する統計
- ベットやレイズに関係する統計
- コールやフォールドに関係する統計

Permutation Importance の実行で正解率に大きく影響を与えた特徴量のみを使って、学習を行った。選択する特徴量は Permutation Importance による正解率が低かった上位 10,15,20 個を選択した。その結果を表 5 に示す

Permutation Importance による正解率が低かった上位 20 個を使用した分類器は、すべての統計を特徴量として学習させた分類器と近い正解率となった。

## 6 考 察

各ストリート、フロップ、ターン、リバーの 3 つに分類した場合における統計の比較だとリバー、ターン、フロップの順で正解率に与える影響が大きくなった。これは、ストリートが後半になるにつれて、すでに行われたアクションは多くなる。つまりプレイヤー同士が、対峙している相手に対して得られる情報が増えている状態となる。このような状況では、技量のあるプレイヤーが情報を生かして対戦相手に対する優位性を大きくすることができる。またストリートが後半であるゆえに、コミュニティカードが多く公開されている。これは得られる情報がさらに増えたということになる。つまり技術の要素がより影響を与える後半のストリートに関する統計は、win rate に対して重要度が高くなっていると考えられる。

ポジションの違いによる統計の比較ではインポジションよりもアウトオブポジションの統計のほうが正解率に与える影響が大きくなっている。インポジション、アウトオブポジションと

表 3 Permutation Importance による統計別の正解率

W\$SD %	0.872727
River CBet OOP %	0.895455
VPIP	0.906818
River CBet IP %	0.913636
Check Fold Turn OOP Skip CBet	0.915909
Agg	0.918182
Flop CBet OOP %	0.920455
Fold Turn IP Skip CBet	0.925000
WTSD %	0.925000
Fold to 3Bet IP	0.929545
Fold to 3Bet OOP	0.929545
River Call Efficiency	0.929545
Delayed Turn CBet % OOP	0.929545
Bet River OOP vs Missed Turn CBet OOP	0.931818
5Bet (Raise 4Bet)	0.931818
PFR	0.931818
4Bet OOP	0.934091
Bet Flop vs Missed CBet	0.934091
Turn CBet IP %	0.934091
W\$WSF	0.934091
Turn CBet OOP %	0.934091
Check Fold Flop % As PFR	0.934091
Check Raise Flop % As PFR	0.934091
3bet In Position	0.936364
Flop CBet IP %	0.936364
Turn Fold vs Cbet OOP	0.936364
Skip Turn CB and Fold To River Bet IP	0.936364
4Bet IP	0.936364
Delayed Turn CBet % IP	0.936364
Flop Raise Cbet OOP	0.938636
Check Raise Turn OOP Skip CBet	0.938636
3bet Out of Position	0.938636
Turn Fold vs Cbet IP	0.938636
Bet Turn OOP vs Missed Flop CBet OOP	0.938636
Flop Raise Cbet IP	0.938636
Cold Call	0.938636
Flop Fold vs Cbet OOP	0.940909
Flop CBet-Fold OOP TOT	0.940909
Flop Fold vs Cbet IP	0.943182
Bet Turn IP vs Missed CBet	0.943182
Fold to 4Bet	0.943182
Limp	0.945455
Flop CBet-Fold IP TOT	0.945455

いうのはプレイヤーの相対的な位置関係のことを示す言葉で、各ストリートの開始時に、ポジションの位置的に、先に行動を決定しなければいけないプレイヤーのポジションをアウトオブポジション、アウトオブポジションのプレイヤー行動を受けて、行動の開始が一番最後になるプレイヤーのポジションをインポジションと呼ぶ。インポジションのプレイヤーはアウトオブポジションのプレイヤーの行動を見た後にアクションを行うことができるので、相手に対する情報をより多く得た状態で判断を下すことができる。そのためインポジションのプレイヤーはア

表 4 統計の種類別正解率

リバーに関係する統計	0.9176135
ポストフロップに関係する統計	0.930131968
アウトオブポジションに関係する統計	0.93068175
ベットやレイズに関係する統計	0.932142857
ターンに関係する統計	0.9339015
インポジションに関係する統計	0.9351516
コールやフォールドに関係する統計	0.935551929
プリフロップに関係する統計	0.935950364
フロップに関係する統計	0.936983545

表 5 重要度が高い統計による正解率

10	0.8977272659540176
15	0.915909081697464
20	0.9340909123420715

ウトオブポジションのプレイヤーよりも有利な状態で簡単にアクションを決定することができる。一方アウトオブポジションのプレイヤーはインポジションのプレイヤーの行動を確認することができないため、アクションの判断が難しくなる。そのためミスを起こしやすい。技量のあるプレイヤーはアウトオブポジションでのプレーが上手であると考えられるため、意思決定が簡単なインポジションに比べ、技術差が出やすいシチュエーションであるといえる。そのためアウトオブポジションに関係する統計は、win rate に対して重要度が高くなっていると考えられる。

ベットとレイズに関係する統計と、コールとフォールドに関する統計の比較については、ベットやレイズに関係する統計のほうが正解率に与えている影響が大きい。ベットやレイズといったアクションは、強い手役を持っているときに稼ぐプレイや、弱い手役を持っているときに相手をフォールドさせて稼ぐことにつながる。つまり、これら技術はチップを稼ぐうえでとても重要な技術である。強い手役を持っている場合は相手がそこそこ強い手役を持っていた場合に最大限稼げるようにベットやレイズの金額を決定したり、自分の手役が弱く見えるようにして相手にコールやレイズをさせる戦略などが挙げられる。ブラフを行う場合はいかに自分の手役が強く見えるようにするか、もしくは相手が弱い手役しかない状況を特定してベットやレイズをするなどの戦略が考えられる。これらの戦略は技術力のあるプレイヤーがうまく実行できると考えられる。そのため win rate に対して重要度が高くなっていると考えられる。

## 7 おわりに

本研究では、ポーカーの統計がプレイヤーの技量に関する情報を含んでいるという考えのもと、ポーカーの統計を利用して少ない試行回数での実力評価を行うための分類器を作成した。完成した分類器は判別の精度が 90 % を超えるため十分に利用可能なものであると思われる。ただし、本研究では十分な試行回数と少ない試行回数を筆者の考えのもとに決定しており、本当に試行回数が多いか少ないかというのは定かではない。そのた

め、さらに少ない試行回数での判定を行える可能性があると考えられる。これは、今後検証すべき課題であると考ええる。

## 文 献

- [1] カジノ法案 政府、ポーカー解禁へ 集客目指す, 毎日新聞, 2018-5-14, <https://mainichi.jp/articles/20180514/ddm/003/010/063000c>
- [2] HighstakesDB (最終閲覧日:2020 年 12 月 24 日), <https://www.highstakesdb.com/>
- [3] 保木 邦仁. 局面評価の学習を目指した探索結果の最適制御. 第 11 回ゲームプログラミング ワークショップ, pp. 78-83, (2006)
- [4] 大森 翔太朗, 金子 知適. 機械学習を用いた将棋における棋風の学習の研究. 情報処理学会研究報告, vol.2015-GI-34, No. 6, pp. 1-7, (2015)
- [5] 尾崎 大夢, ライエル グリムエルゲン. 機械学習を用いた麻雀の役予測手法の提案, 第 25 回ゲームプログラミング ワークショップ, pp. 77-80 (2020)
- [6] N. Brown, T. Sandholm, Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. Science 359, pp. 418–424, (2018)
- [7] N. Brown, T. Sandholm, Superhuman AI for multiplayer poker. Science 365, pp. 885–890, (2019)
- [8] Holdem Manager 3, (最終閲覧日:2020 年 12 月 24 日) <https://www.holdemmanager.com/hm3/>
- [9] HH DEALER, (最終閲覧日:2020 年 12 月 24 日) <https://www.hhdealer.com/>