

## 要旨

本研究は、オンライン麻雀「雀魂」のプレイヤー間で語られる主観的な「地獄モード」という現象を、客観的なデータに基づき検証することを目的とする。**本稿はパイロットスタディとしての性格をもち、将来的な大規模ログ分析への足掛かりを示す。**分析対象は、特定プレイヤー1名が参加した10対局の対局ログである。「山の状態」を表現する複数の統計的指標を算出し、教師なし学習（K-Means クラスタリング、Isolation Forest）を用いて「局のパターン」を特定した。結果として、平均収支が著しく異なる「勝ちパターン」と「負けパターン」が対局ごとに定義され、その遭遇率がプレイヤー間で統計的に偏って分配される事例が複数観測された。このことから、「地獄モード」は単なる心理的バイアスではなく、客観的なデータとして観測可能な、統計的に不利な状況の偏りとして実在しうることが示唆された。

## 第1節 序論

### 1.1. 背景

近年、オンライン麻雀「雀魂」は世界的に多くのプレイヤーを魅了している。その一方で、プレイヤーコミュニティでは「地獄モード」と呼ばれる現象が頻繁に語られている。これは、単なる短期的な不運とは異なり、技量や戦術では抗うことが極めて困難な、理不尽とも言える不利な状況が持続するという主観的な体験を指す。

この現象は、国内最大級のQ&Aサイト「Yahoo!知恵袋」にも具体的な悩みとして投稿されている。例えば、「雀魂の地獄モードに入ってるかどうかを見分ける特徴みたいなのであれば教えてください」と題されたある質問<sup>1</sup>では、投稿者が『『そりゃ災難だと思って我慢しよう』の不運、逆連対がひたすら続く状態』を「地獄モード」と定義し、その客観的な指標を求めている。特筆すべきは、投稿者が自身のスキルに関する不毛な議論を避け、純粋に「状況の不利」を客観的に捉えたいという強い動機を持っている点である。

同様の声は、ソーシャルネットワーキングサービス「X」上でも多数観測される。本稿では、[調査期間]において「雀魂 地獄モード」等のキーワードを含む投稿の中から、この現象に対するプレイヤーの典型的な認識を示すものとして、以下の投稿例を挙げる。あるユーザーは、昇段と降段を繰り返す自身の対局履歴を振り返り、「これが雀魂。地獄モードのスイッチが入っているか、いないかだけ」と、勝敗が自身のスキルではなく、外部的な要因によって決定されているという認識を示している。また、「雀魂の地獄モードから抜け出せない」「地獄モード続行ですか…そうですか…もう無理」といった悲痛な投稿も散見され、この現象がプレイヤー体験（UX）を著しく損なう要因となっていることが伺える。

このように、多くのプレイヤーが「地獄モード」の存在を感覚的に認識し、それを客観視したいと願いながらも、そのための共通の言語や指標が存在しないのが現状である。

### 1.2. 研究目的

本研究の目的は、この主観的な「地獄モード」という現象を、プレイヤーの技量から独立した「山の状態」、すなわち牌の巡り合わせの統計的な偏りとして捉え、客観的なデータを用いてその存在を検証することである。我々は、牌の並びのランダム性や分布の複雑さといった一連の統計的指標を用いて「山の状態」を数値化し、機械学習の手法を用いて「不利な山の状態」のパターンを特定し、その出現頻度がプレ

プレイヤー間で偏って分配されるかを明らかにすることを目指す。本研究では、この「地獄モード」を、「教師なし学習によって特定された、平均獲得点が統計的に著しく低いゲーム状況のパターン」と操作的に定義する。本研究は個人でアクセス可能な限られた対局ログを用いたパイロットスタディであり、将来的な数万～数百万局規模の大規模データ解析への道筋を示すことを主眼としている。

### 1.3. 本稿の貢献

本稿は、これまで主観的にしか語られてこなかったオンライン麻雀における「運」や「流れ」といった要素に対し、教師なし学習（K-Means クラスタリング、異常検知）を用いた定量的な分析アプローチを提案する点に新規性がある。この手法により、個々の対局における「状況の有利・不利」を客観的に評価し、プレイヤー間の公平性についてデータに基づいた議論を行うための新たな枠組みを提供する。ひいては、このアプローチがオンラインゲームにおける運営とプレイヤー間の信頼性確保や、将来のゲームデザインにおける公平性の議論に貢献しうることを期待する。

本稿の構成は以下の通りである。まず第2節で、分析に用いたデータと『山の状態』を定義する統計的指標について述べる。第3節では、クラスタリングと異常検知を用いた分析手法を詳述する。第4節で分析結果を示し、プレイヤー間で観測された『山の状態』の偏りを明らかにする。最後に第5節で、結果を基に『地獄モード』の実在性について考察し、結論と今後の課題を述べる。

### 1.4. 先行研究

オンラインゲームにおける「運」と「公平性」に関する定量的研究は、主に二つの潮流に分類される。

第一に、**マッチメイキングの公平性**に関する研究である。これは、プレイヤーのスキルレーティング（MMR や ELO レート）に基づき、実力が拮抗したプレイヤー同士をマッチングさせるアルゴリズムの妥当性を検証するもので、多くの対戦型ゲームで議論されている。これはプレイヤー間のスキル差を調整する試みであり、本稿が対象とする、対局内での確率的な事象の偏り（運）とは異なるアプローチである。

第二に、**擬似乱数生成器（PRNG）の監査**に関する研究である。これは、ゲーム内で確率的な事象を生成するプログラムそのものを分析し、その出力が真にランダムであるかを統計的に検定するアプローチである。これらの研究は、ゲームの根幹をなすアルゴリズムの透明性を検証する上で重要である。しかし、多くのオンラインゲーム同様、「雀魂」の牌山生成アルゴリズムは公開されておらず、この手法を直接適用することは困難である。

これに対し、本稿は**アルゴリズムが非公開（ブラックボックス）である状況下で、その出力結果である「局の状況」そのものをデータとして捉え、そこに潜む統計的な偏りから公平性を問う**という、異なる視点のアプローチを提案する。

## 第2節 方法

### 2.1. 分析対象データ

#### 2.1.1. サンプルング方法と対象

本研究の分析対象は、オンライン麻雀「雀魂」の段位戦「玉の間」において、特定の1名のプレイヤー（プレイヤーID: Toru1968、段位は[豪1]から[豪3]の範囲）が経験した、特定の期間における対局ログである。

サンプルとして、当該プレイヤーが「雀傑3」から「雀豪1」へ昇格した直後から、時系列に沿って連続した10対局を選定した。この期間を選定した理由は、当該プレイヤーが「雀豪1」昇格後、25対局を経

て「雀傑3」へ降格したという事実に基づいている。この一連の期間は、プレイヤーが著しい不振、すなわちコミュニティで語られる「地獄モード」の渦中にあったと強く推察されるため、本研究の目的に合致した、極めて特徴的なケーススタディの対象であると言える。

### 2.1.2. 比較対象

これらの対局において、分析対象プレイヤーToru1968 と対戦した他の3名は、対局ごとに異なる不特定多数のプレイヤーである。したがって、本研究で行うプレイヤー間の比較は、Toru1968 と、その各対局に参加した「不特定の対戦相手」との間で、遭遇した『山の状態』に統計的な差が存在するかを検証するものである。

## 2.2. 「山の状態」を定義する統計的指標

本研究では、「山の状態」という抽象的な概念を客観的に数値化するため、各局・各プレイヤーが経験した状況を、以下の統計的指標（特徴量）に変換した。各指標は、麻雀における「運」や「流れ」の特定の側面を捉えるという理論的根拠に基づいて選定されている。

- `p_runs`（連検定の  $p$  値）麻雀における「流れ」という概念は、特定の種類の牌が連続して訪れる経験と関連付けられることが多い。そこで本研究では、ツモ牌の系列における連続性（連）を直接的に検定する「連検定」を採用した。この検定の  $p$  値が統計的に有意な水準にある場合、そのツモ系列は完全なランダムとは考えにくく、プレイヤーが「良い流れ」または「悪い流れ」と認識する現象の、客観的な兆候であると仮定する\*\*（計算式の詳細は付録 A.1 を参照）\*\*。
- `entropy_bits`（エントロピー）情報理論におけるエントロピーとは、ある確率分布が持つ不確かさや情報量の尺度である。本研究では、プレイヤーが経験したツモ牌の種類の分布に対してエントロピーを計算した。我々の仮説はこうである。もし手役の完成に必要な牌に偏ってツモが来ている「有利な状況」ならば、ツモ牌の分布は偏り、エントロピーは低くなるはずである。逆に、手がまとまりにくい「不利な状況」では、分布は均一に近づき、エントロピーは高くなる。このように、エントロピーを用いることで、「分布の偏り」という普遍的な尺度で、「手の進みややすさ」という麻雀特有の状況を評価する\*\*（計算式の詳細は付録 A.3 を参照）\*\*。
- `cramers_v`（クラマーの  $V$  連関係数）「地獄モード」では、自身の捨て牌が他家のロン牌になる、あるいは他家の鳴きを誘発するといった経験が語られる。この指標は、プレイヤーの捨て牌と他家の行動といった、2つの事象間の関連の強さを測る。この値が高い場合、プレイヤーの行動とゲームの展開の間に、偶然とは考えにくい何らかの強い関連が存在する可能性を示唆する\*\*（計算式の詳細は付録 A.2 を参照）\*\*。
- `markov_p`（マルコフ連鎖テストの  $p$  値）牌の出現順序が、完全に独立したランダムではなく、何らかの構造を持つ可能性を検証するために採用した。この指標は、ツモの系列が単純な確率モデル（マルコフ性）から逸脱している度合いを評価する。 $p$  値が低い場合、その系列が単純な確率的連鎖ではない、より複雑な構造を持つことを示唆する\*\*（計算式の詳細は付録 4 を参照）\*\*。
- その他の記述的指標 `riichi_turn_by_tsumo`（ツモ補正リーチ巡目）、`tsumo_counts`（ツモ回数）、`naki_counts`（鳴き回数）といった指標は、「山の状態」そのものではなく、その結果として現れるゲーム展開の文脈を捉えるために使用する。これらは、後のクラスタリング分析において、各パターンがどのような性質を持つかを解釈する際の重要な補助情報となる。

なお、本研究で採用した個々の統計指標と、その局の最終的な獲得ポイント（`points_gained`）との間

に強い線形相関関係がないことは、予備的な分析で確認済みである（最も強い相関でも `riichi_turn` の +0.242 に留まる）。この事実は、単一の指標だけでは「山の状態」の有利・不利を単純に測ることは困難であることを示唆している。それゆえに本研究では、単一の指標に頼るのではなく、複数の指標を組み合わせた多次元的な特徴空間で局の状態を評価する、**K-Means クラスタリング**のような**多変量解析**のアプローチが不可欠であると判断した。

### 2.3. データの前処理

分析に先立ち、データの前処理を行った。鳴き (`naki`) や立直 (`riichi`) が行われなかった局において、該当する指標 (`naki_counts`, `riichi_turn` 等) は欠損値 (NaN) となる。今回は「行動の有無」を明確に表現するため、最も解釈が明快な 0 による補完を選択した。この処理により、全ての局を同じ条件で分析することが可能となる。

## 第3節 分析手法

### 3.1. 特徴量の標準化 (Feature Standardization)

第2節で定義した各統計的指標は、それぞれ単位や数値の範囲（スケール）が大きく異なる。このようなスケールが異なるデータを直接分析にかけると、数値の大きい指標が結果に不当に大きな影響を与えてしまう。

そこで、本研究では分析に先立ち、全ての特徴量を標準化した。具体的には、各特徴量の平均が 0、標準偏差が 1 になるように変換 (Z-score normalization) を行った。これにより、全ての指標を公平な基準で評価し、後のクラスタリングや異常検知の精度を高めることが可能となる。

### 3.2. 「山の状態」のパターン分類 : K-Means クラスタリング

「山の状態」に、繰り返し出現するいくつかの典型的な「パターン」が存在するかを検証するため、教師なし学習の代表的な手法である K-Means クラスタリングを採用した。本手法は、多次元のデータを、事前に定めた  $k$  個のクラスター（グループ）に分類する。

本研究では、 $k$  の値を 3 に設定した。これは、麻雀におけるプレイヤーの基本的な戦略状況が、大きく分けて「攻め（有利）」「守り（不利）」「中間・様子見」という 3 類型に分類されることに着想を得ている。

この  $k$  値の妥当性を検証するため、エルボー法およびシルエットスコアを用いた評価を実施した。 $k=2$  から  $k=5$  までの範囲で分析を行った結果、統計的な指標と解釈可能性の両面において、 $k=3$  が最もバランスの取れた結果を示すことが確認されたため、この値を最適値として採用した。

### 3.3. 特異な「山の状態」の検出 : Isolation Forest

「地獄モード」は、単なる不利なパターンだけでなく、他の大多数の局とは全く異なる「極端な不運」として現れる可能性がある。このような外れ値 (アノマリー) を検出するため、本研究では Isolation Forest アルゴリズムを採用した。

分析にあたり、異常と見なすデータの割合 (contamination) は  $0.1$  (10%) に設定した。これは、1 回の半荘 (はんちゃん) が一般的に約 8 局から 12 局で構成されることを考慮し、その中でゲーム全体の流れや最終的な勝敗を決定づけるような『特異な局』が、平均して 1 局程度発生するという、経験則に基づいた仮説である。

さらに、この設定値の妥当性を検証するため、 $0.05$  から  $0.15$  の範囲で感度分析を実施した。その結果、`contamination=0.1` の設定が、統計的な外れ値であり、かつ平均獲得ポイントが他の大多数の局と著しく

乖離する『特異な局』（我々の分析目的に合致する局）を最も明確に捉えることが確認されたため、この値を採用した。

3.4. プレイヤー間の偏りの分析

上記の分析により、全ての局は「どのパターン（クラスター0, 1, 2）に属するか」というラベルと、「異常か正常か」というフラグを持つことになる。最終段階として、これらのラベルとフラグをプレイヤー毎に集計し、各プレイヤーがどのパターンにどれくらいの割合で遭遇したかを算出する。

3.5. 再現性と透明性の担保

本研究の分析は、完全な再現性を担保するため、分析に用いた全データおよび本節で詳述した手法を実装した Python 解析コードの全てを、公開リポジトリ（GitHub）上で公開している。読者は、以下の URL から全ての資源にアクセスし、本稿の結果を独自に検証することが可能である。

URL: <https://github.com/Torul968/mahjong-performance-analysis>

3.6. 手法の有効性検証（シミュレーションによる堅牢性の確認）

本研究で提案する分析手法が、実際にデータ内の統計的な偏りを正しく検出できるか検証するため、擬似データを用いたシミュレーションを実施した。

シミュレーションでは、まず 800 局分（4 人×200 局）の「完全に公平なデータセット」を生成した。このデータセットでは、全プレイヤーの各統計指標は、同一の正規分布から生成される。次に、同様に 800 局分の「意図的に偏らせたデータセット」を生成した。ここでは、特定の 1 プレイヤー（A さん）の特定の指標（entropy\_bits, p\_empirical）の分布の平均値をずらすことで、統計的に不利な状況を再現した。両データセットに本稿の分析手法を適用した結果を、それぞれ表 X および表 Y に示す。

プレイヤー名 担当局数 異常検知された局数 異常検知率 クラスタ 0 割合 クラスタ 1 割合 クラスタ 2 割合

A さん	200	16	8.0%	35.0%	38.0%	27.0%
B さん	200	23	11.5%	37.5%	36.0%	26.5%
C さん	200	19	9.5%	31.0%	35.5%	33.5%
D さん	200	22	11.0%	30.5%	31.5%	38.0%

表 X 公平なデータセット

プレイヤー名 担当局数 異常検知された局数 異常検知率 クラスタ 0 割合 クラスタ 1 割合 クラスタ 2 割合

A さん	200	32	16.0%	86.0%	7.0%	7.0%
B さん	200	18	9.0%	16.0%	42.5%	41.5%
C さん	200	16	8.0%	16.0%	42.0%	42.0%
D さん	200	14	7.0%	15.5%	42.5%	42.0%

表 Y 偏らせたデータセット

表 X が示す通り、「公平なデータセット」においては、各プレイヤーのクラスター遭遇率や異常検知率に大きな差は見られなかった。これは、本手法が偏りのないデータから、誤って偏りを検知しないことを示している。

対照的に、表 Y が示す通り、「偏らせたデータセット」においては、A さんのクラスター遭遇率が他のプレイヤーと著しく異なっている。これは、本手法がデータに内在する統計的な偏りを正しく検出し、特定

のプレイヤーに紐づける能力を持つことを証明している。

以上の検証により、本研究で用いる分析手法は、実際の対局データのサンプルサイズの問題とは独立して、ゲーム状況の偏りを検出する上で堅牢かつ有効であると結論付ける。なお、この検証に用いた擬似データ生成スクリプトも、第 3.5 節で言及した GitHub リポジトリで公開し、本検証の完全な再現性を保証する。

4.1. 「山の状態」パターンの特定

分析対象とした 10 対局、合計 100 局以上の各プレイヤーの状況データに対し、第 3 節で述べた手法を適用した。その結果、K-Means クラスタリングは、対局ごとに特徴的な 3 つの「山の状態」のパターン（クラスター）を一貫して抽出した。

各クラスターは、局の平均獲得ポイント（points\_gained）に基づき、以下の 3 つの類型に大別された。

- 「勝ちパターン」：平均獲得ポイントが大きくプラスになるパターン。多くの場合、効果的な鳴きや、高打点での和了（あがり）と関連していた。
- 「負けパターン」：平均獲得ポイントが大きくマイナスになるパターン。「地獄モード」の構成要素であり、リーチが実らない、あるいは失点を重ねる状況がこれに該当する。
- 「中間パターン」：平均獲得ポイントが 0 点に近い、膠着状態や平凡な展開を示すパターン。

なお、どのクラスターが「勝ち」「負け」に該当するかは、対局ごとにその対局内のデータに基づいて動的に決定された。

第 4 節 結果

4.2. 各対局における「山の状態」の偏り

本分析における最も重要な発見は、特定された「山の状態」の各パターンが、プレイヤー間で均等に分配されず、対局ごとに極端な偏りをもって出現したことである。以下に、10 対局の中から特徴的な 2 つのケースを示す。

ケーススタディ 1：ある対局の分析

この対局では、クラスタ 1 が平均+1488 点の「勝ちパターン」、クラスタ 2 が平均-1300 点の「負けパターン」として特定された。この対局における各パターンの遭遇率（表 1）を見ると、プレイヤー間で著しい格差が確認できる。

表 1: ケーススタディ 1 におけるパターン遭遇率

プレイヤー名	「勝ちパターン」遭遇率	「負けパターン」遭遇率
:---	:---	:---
D さん	58.3%	8.3%
A さん	8.3%	58.3%
B さん	41.7%	25.0%
C さん	25.0%	50.0%

表 1 が示す通り、この対局では D さんが「勝ちパターン」を 58.3%という高確率で経験したのに対し、A さんは「負けパターン」を 58.3%という極めて高い確率で経験していた。

ケーススタディ 2：別の対局の分析

この対局では、クラスタ 0 が平均+8614 点の「圧勝パターン」、クラスタ 1 が平均-2933 点の「大敗パターン」として特定された。この対局においても、プレイヤー間で深刻な偏りが見られた（表 2）。

表 2: ケーススタディ 2 におけるパターン遭遇率

プレイヤー名	「圧勝パターン」遭遇率	「大敗パターン」遭遇率
:---	:---	:---
C さん	37.5%	0.0%
A さん	12.5%	62.5%
B さん	12.5%	12.5%
D さん	25.0%	37.5%

この対局では、A さんが「大敗パターン」に集中的に遭遇する一方で、C さんは一度も「大敗パターン」を経験することなく、「圧勝パターン」を最も多く経験していた。

4.3. 「異常検知」された局の性質

10 対局を通じて「異常」と検出された局の性質は、対局ごとに異なっていた。ある対局では極端な高得点局が、また別の対局では極端な低得点局が「異常」と判断される傾向が見られた。これは、「異常」が単に「幸運」や「不運」のどちらかを指すのではなく、「大多数の平凡な局から統計的に逸脱した、特異な展開」を客観的に捉えていることを示唆している。

第5節 考察

5.1. 結果の要約と研究目的との関連

本研究の結果、K-Means クラスタリング等の教師なし学習を用いることで、麻雀の各局を平均収支が大きく異なる複数の「状態パターン」に分類できることが示された。さらに、これらの有利・不利なパターンが、対局内でプレイヤー間に著しく偏って分配される事例が、分析した 10 対局の中から複数確認された。この結果は、序論で立てた研究目的、すなわち「主観的な『地獄モード』という現象が、客観的なデータとして観測可能か」という問いに対し、肯定的な答えを与えるものである。

5.2. 「地獄モード」の実在性に関する考察

本研究が示した「負けパターンへの遭遇率の偏り」は、プレイヤーが「何をしても勝てない」と感じる「地獄モード」の正体の一端を捉えている可能性が高い。

重要なのは、この分析がプレイヤーのスキル（例：何を切るか）から独立した、純粋な「状況（配られる牌の性質）」を対象としている点である。これは、「地獄モード」が単なる技量不足や心理的バイアスだけでなく、統計的に証明可能な「理不尽な不運」として実在しうることを強く示唆している。

さらに、本稿で観測された「偏り」の発生源について、二つの可能性が考えられる。一つは、牌山生成アルゴリズムが特定のプレイヤーに対して統計的に不利な状況を意図的または結果的に作り出している可能性（サーバー起因の偏り）。もう一つは、同卓した他家の一人が極端な幸運（天国モード）を経験し、その副次的影響として他のプレイヤーが相対的に不利な状況に置かれる可能性（相対的な偏り）である。しかし、本研究の立場は、この二つを区別することに本質的な意味はない、というものである。プレイヤーの体験（Player Experience）という観点から見れば、その原因がサーバーにあるか、あるいは他家の幸運にあるかに関わらず、「個人の技量では覆すことが極めて困難な、統計的に不利な状況」に置かれているという事実が変わりはないからだ。あるプレイヤーが和了（アガリ）に向かう有効牌は、山のどこに

あろうと、あるいは他家の手の中にあろうと、そのプレイヤーにとっては等しく「ツモれない牌」である。したがって、本稿における「地獄モード」の証明とは、その発生源を特定することではなく、原因が何であれ「統計的に勝利期待値の低い局が、特定のプレイヤーに偏って分配されるという事象は、実際に発生しうる」という事実を、客観的なデータで明らかにすることである。

この発見は、プレイヤー間のフラストレーションを軽減し、「運」の要素についてより建設的な議論を促す一助となるだろう。

### 5.3. 本研究の限界

本研究の意義は大きい一方で、いくつかの限界も存在する。

本研究の位置づけ 本稿はあくまで限られたケーススタディであるため、結果は大規模データに必ずしも直接適用されない。しかし、本手法の適用可能性と実装フローを示す点で重要な第一歩となる。

第一に、本研究は特定の1名のプレイヤーが参加した10対局という限られたデータに基づいている。この結果を「雀魂」全体、あるいは麻雀というゲーム全体に一般化するためには、より大規模で多様なプレイヤー層を含むデータでの検証が不可欠である。

第二に、「山の状態」を定義するために用いた統計指標は、運の偏りの一部を捉えるものに過ぎない可能性がある。例えば、本稿の指標では捉えきれない、他の種類の「不運」も存在しうる。

### 5.4. 今後の課題と展望

今後の課題として、まず**数千、数万単位の対局データを用いた大規模分析**が挙げられる。これにより、今回見られたパターンの一般性や、段位による偏りの違いなどを検証できるだろう。

また、本稿では意図的に切り離れた**プレイヤーのスキル（段位や具体的な打牌選択）と、「山の状態」との相互作用**を分析することも興味深いテーマである。例えば、「不利な山の状態」に遭遇した際に、熟練者はどのようにして失点を最小化するのか、といった分析が考えられる。

将来的には、これらの知見を基に、局の開始時点の指標から、その局がどのパターンに分類されるかを予測するモデルを構築し、プレイヤーに「今は耐えるべき時です」といった客観的な情報を提供するツールの開発にも繋がるかもしれない。

### 5.5. 実証研究としての位置づけと運営側への提言

本稿は少数対局を対象としたケーススタディであり、限界を認識したうえで、より広範なログへの適用可能性を示す**基盤的検証研究**と位置づけている。

この分析が示唆する「統計的な偏り」の一般性を検証するための大規模な分析は、数万、数百万単位の対局ログへのアクセスを必要とし、これは一個人の研究の範疇を大きく超える。ゲーム全体の公平性を大規模データで証明する能力と責任は、本質的にゲームの運営企業側にあると我々は考える。

本研究で得られた分析結果を基に、分析対象プレイヤーが運営事務局に問い合わせたところ、「全てのプレイヤーが同様の条件で対局を行っている点から、公平な条件であると認識しております」との回答が得られた。この回答は、個別のデータに基づく統計的な偏りの可能性を考慮せず、形式的な公平性をもって、実質的な公平性が担保されていると見なす、運営側の現在のスタンスを示すものである。

オンラインゲーム、特にeスポーツとして競技性を高めるタイトルにおいては、その公正さがプレイヤーの信頼の基盤となる。運営企業には、受動的な姿勢に留まらず、能動的に大規模データの監査を行い、その結果を公表するなどして、コミュニティとの信頼関係を構築していくことが、今後の健全な発展のために強く望まれる。



## 5.6. 先行研究との比較と本研究の独創性

序論で述べた通り、ゲームの公平性に関する先行研究の多くは、マッチメイキングの最適化や、既知あるいはリバースエンジニアリングされた乱数生成アルゴリズムの監査といった\*\*「プロセスの検証」\*\*に焦点を当ててきた。

これに対し、本研究の独創性は、アルゴリズムが不明なブラックボックスを対象に、その\*\*「出力結果（＝統計的に偏った状況）」\*\*から、プレイヤーが体験する実質的な公平性を問うという新しい分析フレームワークを提示した点にある。ポーカーにおける運とスキルの分離に関する研究とも異なり、本稿は特定のプレイヤーに不利な状況が「パターン」として偏在しうることを実証した。

このアプローチは、「雀魂」に限らず、ガチャの確率やアイテムのドロップ率など、アルゴリズムが非公開である多くのオンラインゲームの「運」の要素について、プレイヤー側から客観的な議論を行うための、汎用的な手法となりうる可能性がある。

### 引用

<sup>1</sup> 「雀魂の地獄モードに入ってるかどうかを見分ける特徴みたいなのであれば教えてください。」. Yahoo! 知恵袋. 投稿日: 2025 年 7 月 10 日. 最終閲覧日: 2025 年 7 月 14 日.

### 章外 付録

#### 付録 A : 主要な分析指標の計算式

##### A.1. p\_runs (連検定)

データ系列のランダム性を検定するために用いられる連検定 (Runs Test) の p 値。サンプルサイズが大きい場合、検定統計量  $z$  は以下の式で近似されます。

$$z = \frac{R - \mu_R}{\sigma_R}$$

ここで、

- $R$ : 観測された連の数
- $\mu_R$ : 期待される連の数
- $\sigma_R$ : 連の数の標準偏差

---

##### A.2. cramers\_v (クラマーの V 連関係数)

2つのカテゴリカル変数間の関連の強さを測る指標 (0 から 1)。カイ二乗統計量を基に計算されます。

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(r-1, c-1)}}$$

ここで、

- $\chi^2$ : カイ二乗統計量
- $n$ : 全サンプルサイズ
- $r$ : 分割表の行数
- $c$ : 分割表の列数

---

##### A.3. entropy\_bits (エントロピー)

情報理論における、確率変数が持つ不確かさ（情報量）の尺度。底を 2 とする対数を用いることで、単位は「ビット」となります。

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2(p(x_i))$$

ここで、

- $X$ : 確率変数（例：牌の種類）
  - $p(x_i)$ : 事象  $x_i$  が発生する確率
- 

#### A. 4. markov\_p（マルコフ連鎖テスト）

事象の系列が、直前の状態にのみ依存するという単純なマルコフ性を持つかどうかの検定（適合度検定）から得られる  $p$  値。多くの場合、カイ二乗統計量が用いられます。

$$\chi^2 = \sum_{i,j} n_{ij} p_{ij} (n_{ij} - n_{ij} p_{ij})^2$$

ここで、

- $n_{ij}$ : 状態  $i$  から状態  $j$  へ遷移した観測度数
- $n_i$ : 状態  $i$  の総観測度数
- $p_{ij}$ : 状態  $i$  から状態  $j$  へ遷移する期待確率