Analyzing Win Probabilities in the Dice Game 'Liar's Dice' Using Monte Carlo Simulations

Group Member: Jiajing Liang (liang82)
Yueyue Lin (yueyuel4)

Agenda

Part1 Introduction

Part2 Code design

Part3 Hypotheses

Part4 Validation

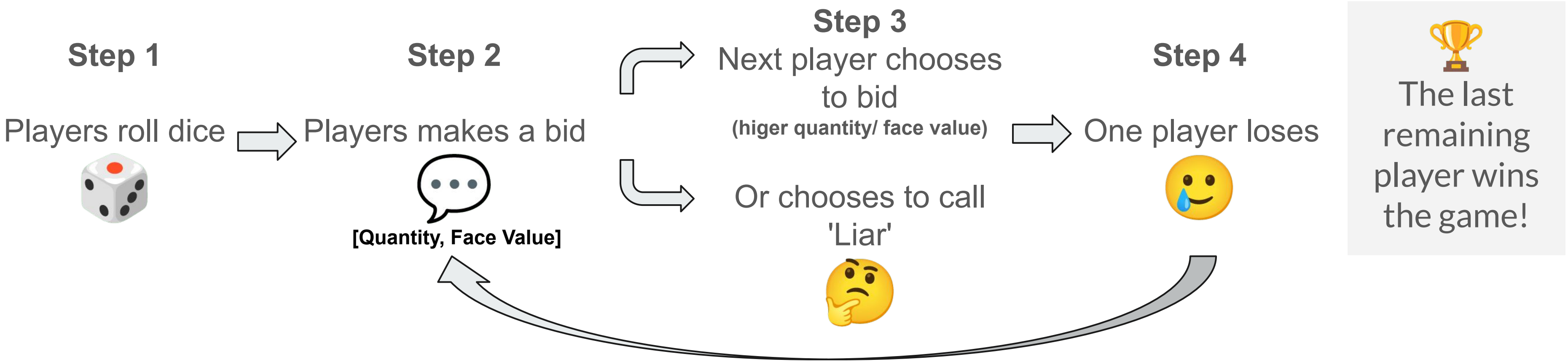
Part5 Experiment

Part6 Limitation & Future work

1 Introduction

Game Overview:

- Players bid on dice outcomes, considering **quantity** and **face value.**
- The bid can be challenged by the **next player** calling **"Liar."**
- **"1"** is a wildcard, contributing to any face value.
- If the **challenge** is correct, the bidder loses; otherwise, the challenger loses.

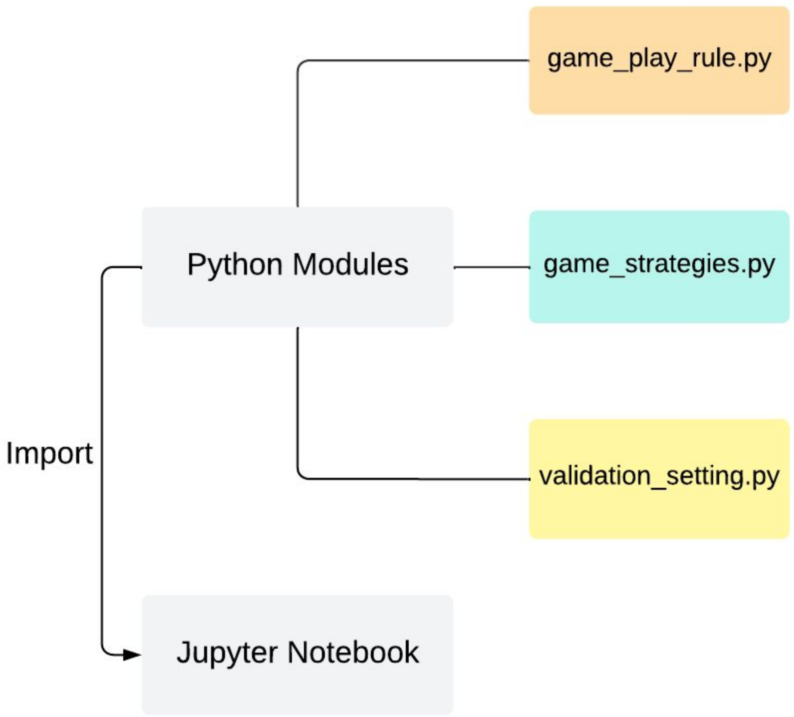


1 Introduction

Project Objectives:

1. Simulate the gameplay of "Liar's Dice" under varying conditions.
2. Analyze the impact of player strategies and rule changes on outcomes.
3. Validate the fairness and randomness of the game mechanics.

2 Code Design



`game_play_rule.py` is the core engine:

- It manages the gameplay mechanics.
- Other modules rely on its functions for game simulation.
- Functions:

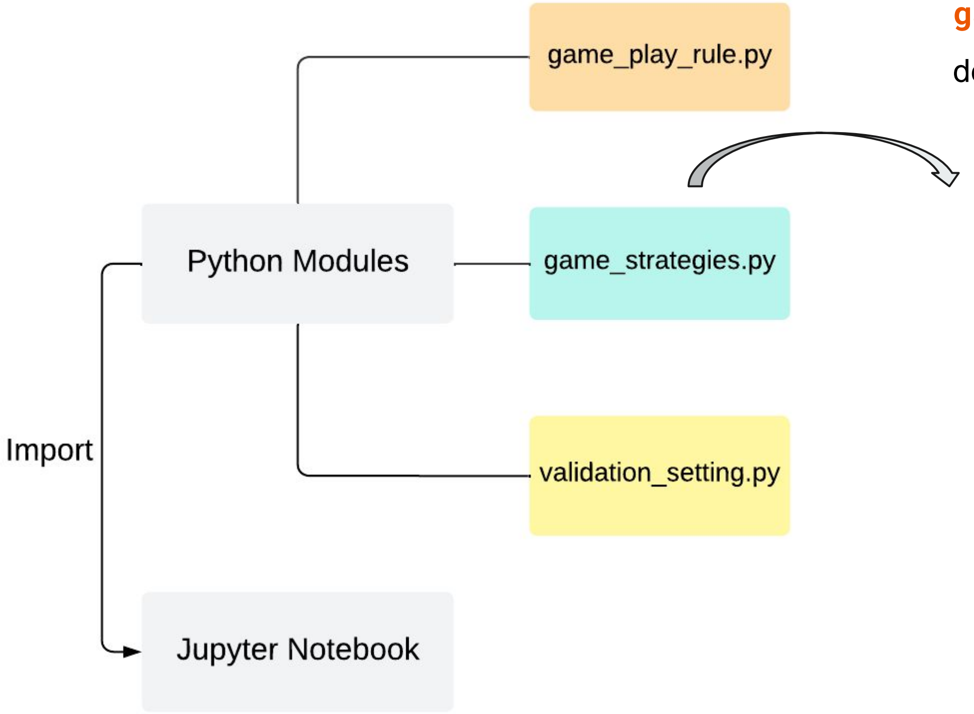
  `roll_dice`: Generates random dice outcomes for players.

  `valid_challenge`: Checks if a "Liar" call is valid.

  `update_all_dice`: Updates dice counts after each round.

  `simulate_game`: Runs the entire game loop.

2 Code Design



game_strategies.py defines how players make decisions:

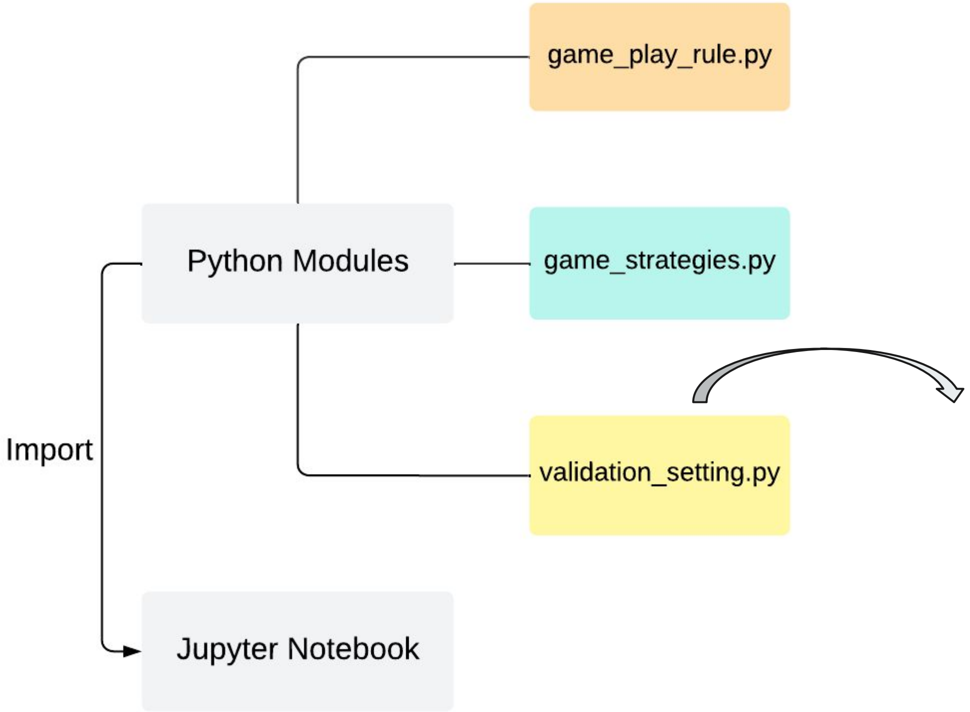
- It provides strategy classes for simulations.
- Provides decision-making logic for players during the game.
- Key Functions:

  `random_bid`: Generates random bids.

  `inform_bid`: Makes smarter bids based on the dice player have. Related with Preferred Dice Bid hypothesis.

  `make_action`: Decides to bid or call "Liar." Connected with our hypothesis of Threshold strategy.

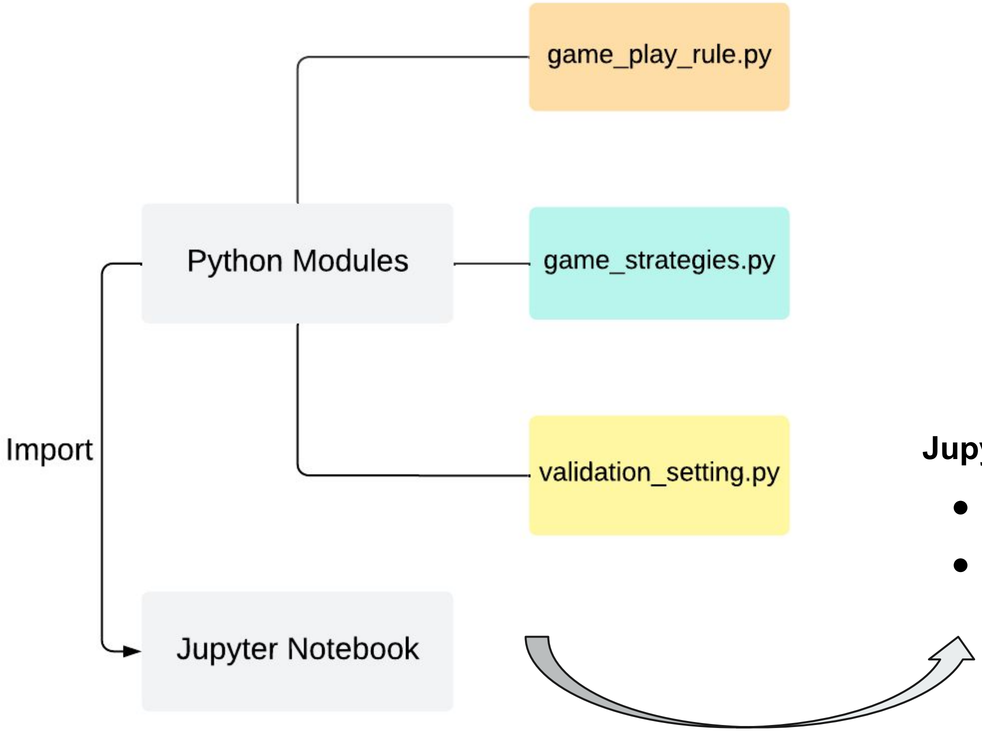
2 Code Design



`validation_setting.py` evaluates and visualizes game outcomes:

- Analyzes results to test fairness and validate hypotheses.
- Functions for win rates, challenge accuracy, and heatmaps.

2 Code Design



Jupyter Notebook:

- Combines modules to run experiments dynamically.
- Visualizes results.

3 Hypotheses

1 **First Caller Advantage**

Hypothesis: The first player to act has a higher win rate than other players because they set the initial bid and influence the strategy direction.

2 **Preferred Dice Bid**

Hypothesis: Players prioritizing bids based on the most frequent dice in their hand (preferred face value) achieve higher win rates.

3 **Threshold for Calling "Liar"**

Hypothesis:

Players using certain thresholds to decide when to challenge has a higher win rate than players who challenge randomly.

- ***Normal Threshold:*** 50% of the total number of dice in play.
- ***Optimal Threshold:*** based on a calculated probability ratio that accounts for the total dice and the player's own dice. (number of dice with the bid face value in own dice + 50% of the number of remaining dice)

4 **Optimal Strategy**

Hypothesis:

Players who combine thresholds with preferred dice bid achieve the highest win rates.

5 **Special Rule Impact**

Hypothesis: Modifying the rule (reducing one die instead of eliminating a player) affects game fairness and the win rate of players.

4 Validation

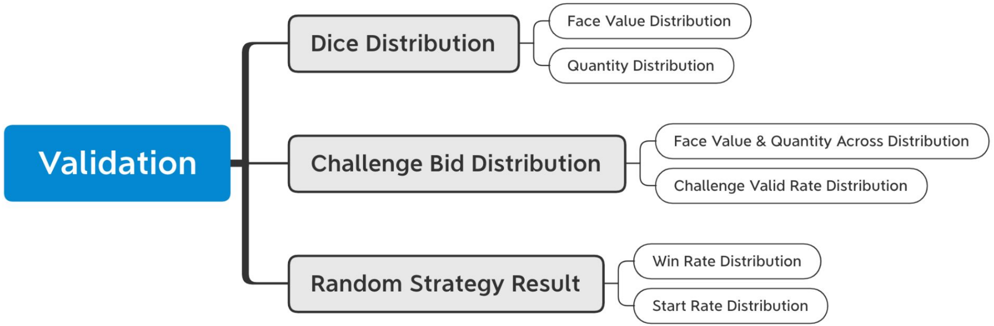
Hidden Assumptions

1. Dice are randomly rolled.
2. Without any specific strategy, players choose to challenge the bid randomly.
3. Without any specific strategy, every player has the same chance to make the first bid.
4. Without any specific strategy, every player has an equal win rate.

4 Validation

The validation part includes 3 main validations for our hidden assumptions.

- Dice Distribution is to check if the dice are rolled randomly
- Challenge Bid Distribution is to check if players challenge the bid randomly
- Random Stratgy Result: The win rate distribution and the start rate distribution should be uniform distribution without any specific strategy



* The game is validated on the game with 5 players and each player has 5 dice .

AS EXPECTED?

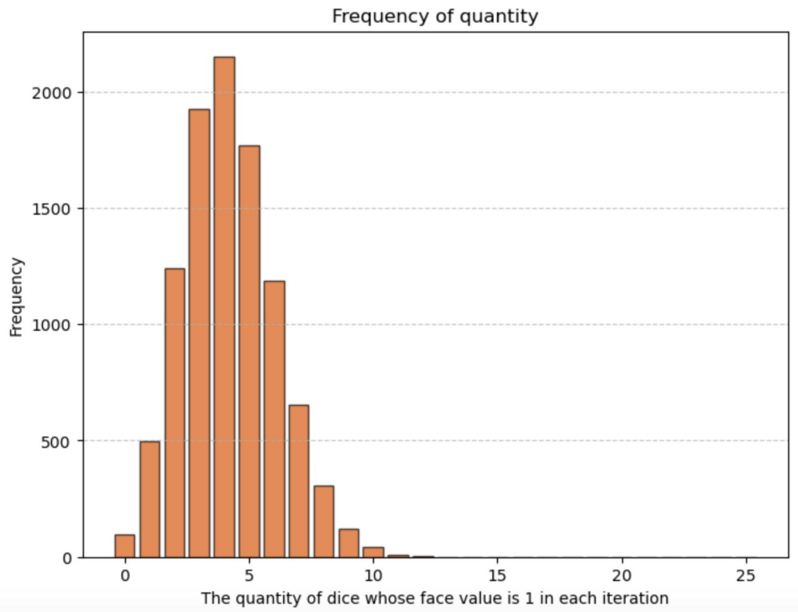
NOT AS EXPECTED?

4 Validation - Dice Distribution

✅ Rolled randomly, the Face Value Distrubution should be a Uniform Distribution.

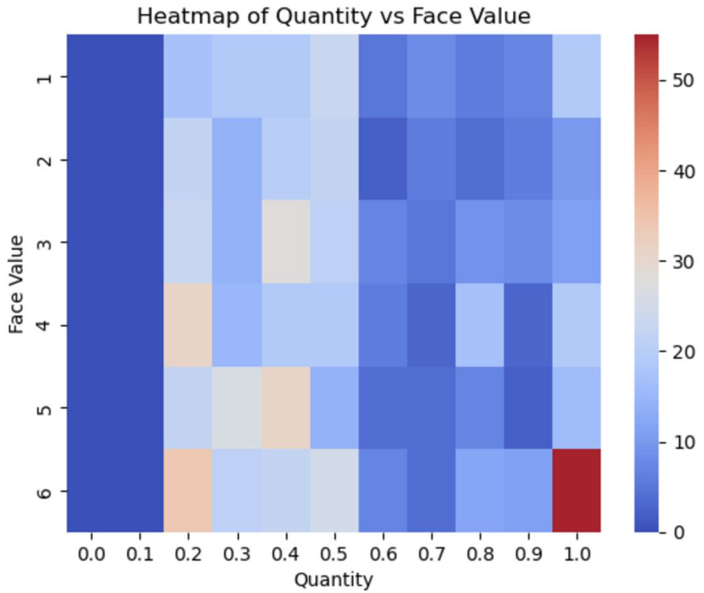
Face_value	Occurrence	Occurrence rate
1	4021	16.08%
2	4258	17.03%
3	4133	16.53%
4	4153	16.61%
5	4261	17.04%

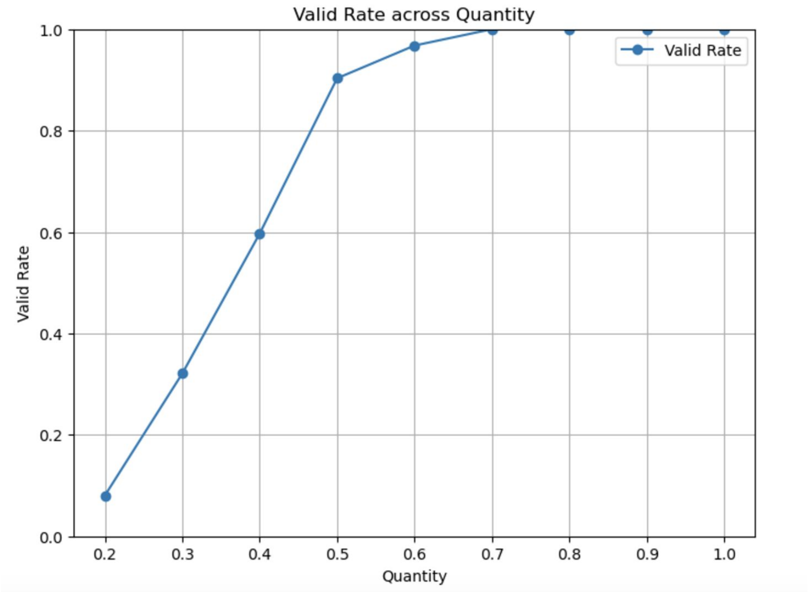
✅ Rolled randomly, The Frequence of quantity Distribution should be a binomial Distribution.



4 Validation - Challenge Bid Distribution

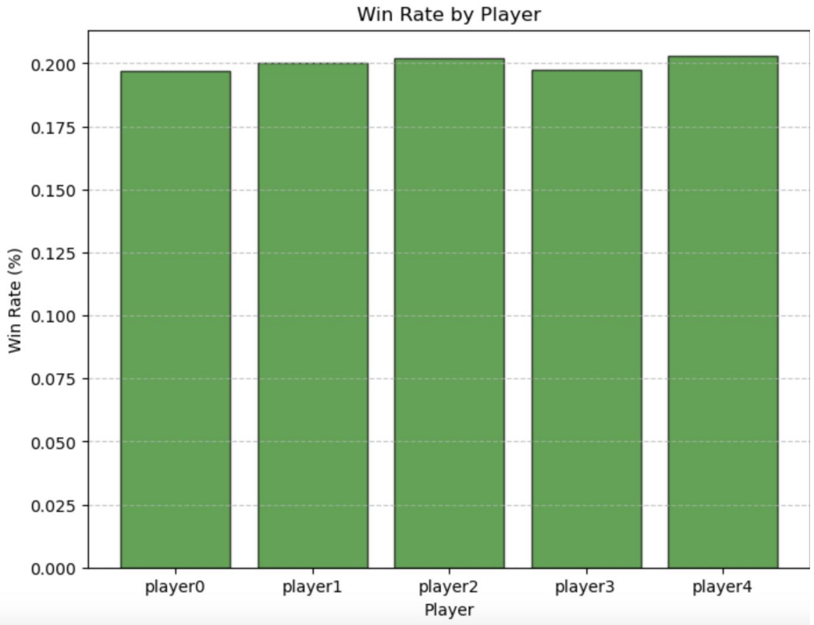
- ✅Players is likely to challenge when the quantity of bid is small, but almost randomly.
- It is reasonable that players never challenge when the quantity is smaller than 10% of total dice because the first bid should be bigger than that under game rule
- The valid rate increases as the quantity of bid raises, as expected.



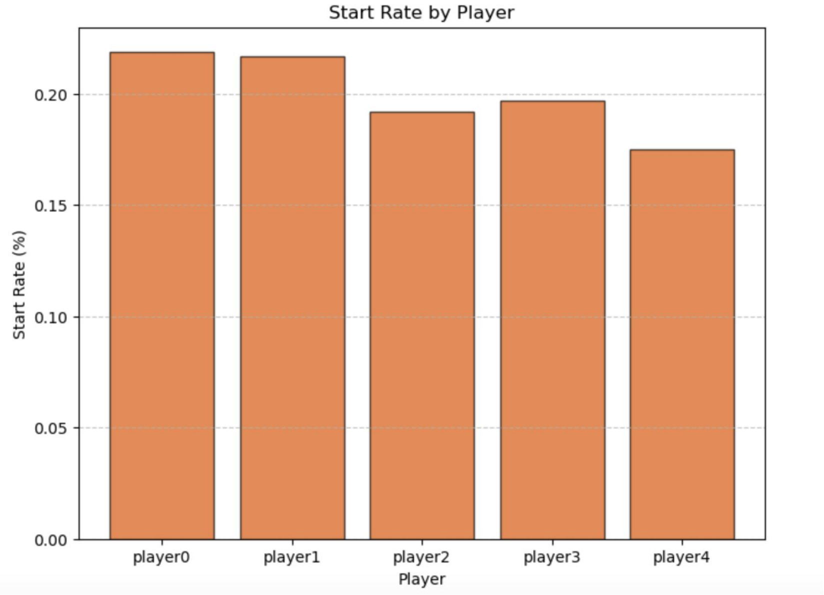


4 Validation - Random Strategy Result

✅The Win Rate Distribution is a Uniform Distribution as expected



✅The Start Rate Distribution is a Uniform Distribution as expected

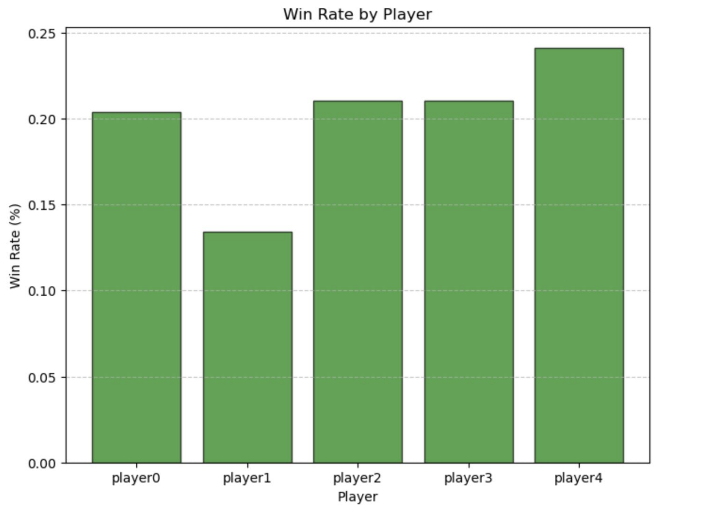


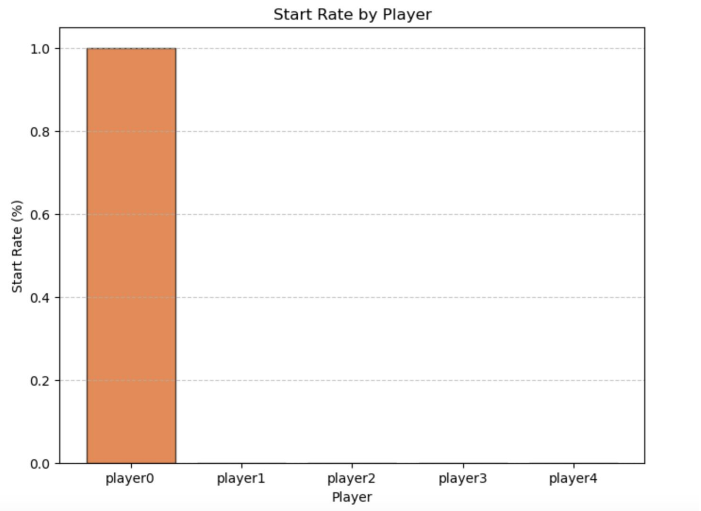
5 Experiment - Hypothesis 1 Result

Expected result : The player choose to bid first(player0) would have a higher win rate than other players

Actual result : ❌

Contrary to the hypothesis, the next player in sequence (Player1) shows a significantly lower win rate compared to others. Player1 is at a disadvantage maybe because they must directly respond to the first bid, often with limited options. This position forces riskier decisions, leading to lower win rates.





5 Experiment - Hypothesis 2 Result

Expected result : The player prioritizing bids based on the most frequent dice in their hand (preferred face value) achieve higher win rate.

Actual result : ✅ **As Expected.** Players using the "Preferred Dice Bid" strategy showed consistently higher win rates compared to those using random bidding.

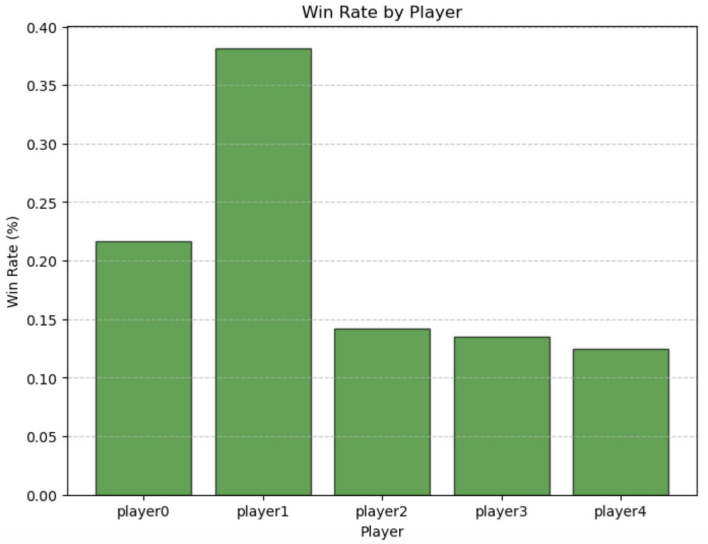
	Player	Wins	Win Rate
0	player0	323	0.323
1	player1	157	0.157
2	player2	194	0.194
3	player3	151	0.151
4	player4	175	0.175

5 Experiment - Hypothesis 3 Result

Expected result : Players using certain thresholds to decide when to challenge have higer win rate than players who challenge randomly.

Actual result : ✅ As Expected.
Players using the "Threshold" strategy showed higher win rates compared to those challenge randomly. The player using Optimal_threshold strategy has a highest win rate.

Player	Wins	Win Rate	
player0	2167	0.2167	Normal_threshold
player1	3816	0.3816	Optimal_threshold
player2	1419	0.1419	Random
player3	1351	0.1351	
player4	1247	0.1247	



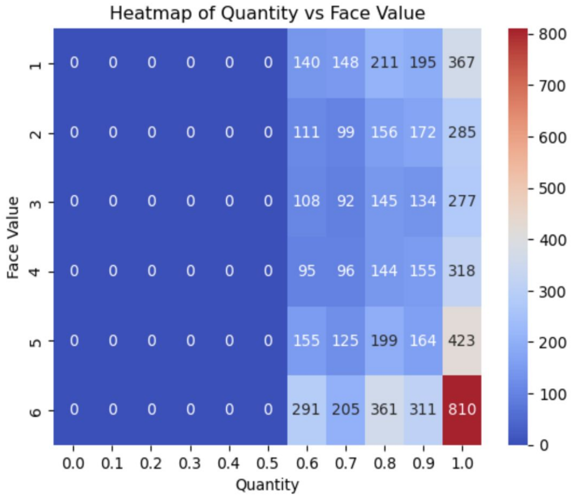
5 Experiment - Hypothesis 3 validation

Expected challenge bid distribution:

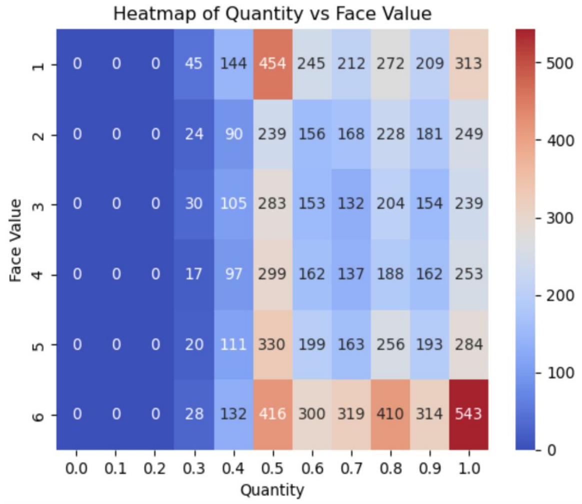
✅The player using normal_threshold strategy would not challenge if the quantity is smaller than 50% of total dice;

✅The player using optimal_threshold stratgey would show less challenges than random player if the quantity is small

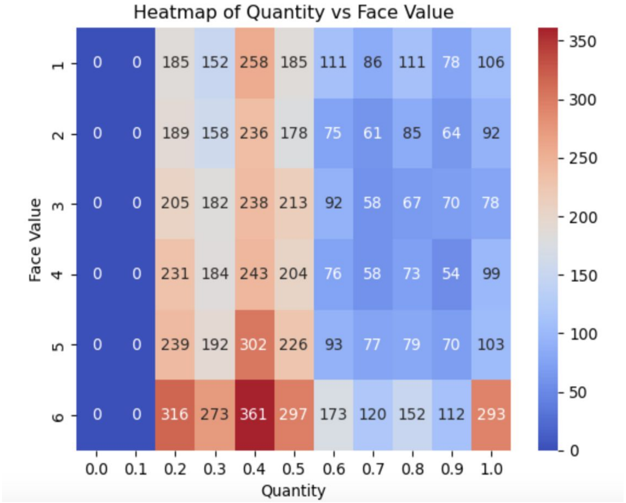
Player with normal_threshold Strategy



Player with optimal_threshold Strategy



Player with Random Strategy

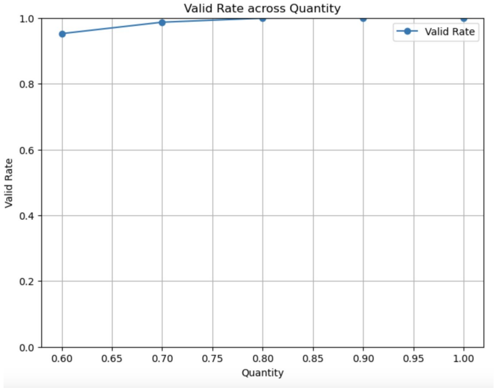


5 Experiment - Hypothesis 3 validation

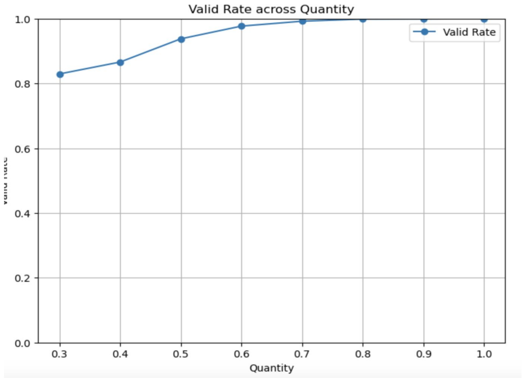
Expected challenge valid rate :
✅Players using specific strategies are expected to have a higher challenge valid rate compared to random strategies.

player	valid_calls	invalid_calls	total_calls	valid_call_rate
player0	6438	54	6492	99.17%
player1	9577	285	9862	97.11%
player2	5423	2890	8313	65.24%
player3	5116	2431	7547	67.79%
player4	5218	2568	7786	67.02%

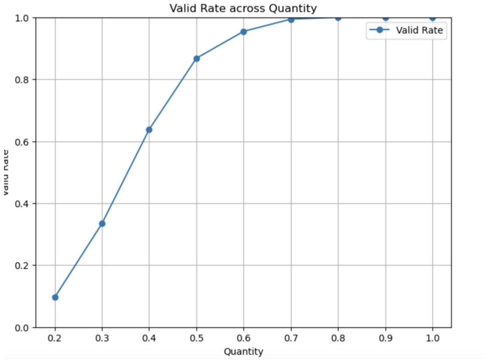
Player0 with normal_threshold Strategy



Player1 with optimal_threshold Strategy



Player2 with Random Strategy



5 Experiment - Hypothesis 4

Expected result : Players who combine thresholds with preferred dice bid achieve the highest win rates.

Actual result : ✅ As Expected. Players using the preferred dice bid strategy combining with optimal threshold showed the highest win rates compared to those used other strategies.

Player	Wins	Win Rate
player0	819	0.0819
player1	1648	0.1648
player2	870	0.0870
player3	2737	0.2737
player4	3926	0.3926

Player Strategy Assignment

- player0 : Normal threshold strategy
- player1 : Optimal threshold strategy
- player2 : Prefer Dice Bid strategy
- player3: Normal threshold + Prefer Dice Bid strategy
- player4: Optimal threshold+ Prefer Dice Bid strategy

5 Experiment - Hypothesis 5

Expected result : Modifying the rule (reducing one die instead of eliminating a player if a challenge happens) affects game fairness and the win rate of players.

Actual result : ✅ As Expected. The win rate of different strategies changes, and the “good” strategies performs better , the “bad” strategies performs worse under special rule than under normal rule.

Player	Wins	Win Rate
player0	195	0.0195
player1	604	0.0604
player2	677	0.0677
player3	4620	0.4620
player4	3904	0.3904

Player Strategy Assignment

- player0 : Normal threshold strategy
- player1 : Optimal threshold strategy
- player2 : Prefer Dice Bid strategy
- player3: Normal threshold + Prefer Dice Bid strategy
- player4: Optimal threshold+ Prefer Dice Bid strategy

6 Limitation & Future work

Limitation:

1. **Simplified Strategies:** The current model assumes players strictly follow predefined strategies, which limits the exploration of adaptive or mixed strategies seen in real-world gameplay.
2. **Unrealistic Assumptions:** Errors such as misreading dice, miscalculations, or hesitation are not considered, which may differ from actual player behavior.
3. **Limited Rule Variability:** Only a small range of rule modifications was tested, and their impact on gameplay dynamics and fairness may not be fully explored.

Future work:

1. **Increase Experiment Repetition for Consistency :** To ensure the results are statistically robust and not influenced by randomness. Run 1,000 games per simulation, then repeat this process to observe variability.
2. **Track Strategy Evolution:** Analyze how the strategies of surviving players change after a player losed. Identify patterns or shifts in their bidding and challenging behaviors.

Thanks