## Title
Quantification of Luck and Skill in Poker Via Simulations and Expected Value Computation



Author
Xu, Mingzhe

Publication Date
2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

# Quantification of Luck and Skill in Poker Via Simulations and Expected Value Computation

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Applied Statistics and Data Science

by

Mingzhe Xu

2024

ABSTRACT OF THE THESIS

Quantification of Luck and Skill
in Poker Via Simulations
and Expected Value Computation

by

Mingzhe Xu
Master of Applied Statistics and Data Science
University of California, Los Angeles, 2024
Professor Frederic R. Paik Schoenberg, Chair

Texas Hold'em is a globally popular poker game, enjoyed both recreationally among friends and professionally in casinos. Despite its widespread popularity, there remains a persistent debate about whether Texas Hold'em is primarily a game of luck or skill. This paper aims to address this debate through a rigorous statistical analysis that both defines and quantifies luck and skill in Texas Hold'em. By examining obfuscated real-world poker game hand histories, we compute the expected profit attributed to either luck or skill for the players at each betting round. We then use various visualizations to compare these aggregated results, which allows us to discover patterns and derive meaningful insights. Our findings contribute to a deeper understanding of the dynamics of Texas Hold'em, offering implications for players, strategists, and policymakers.

The thesis of Mingzhe Xu is approved.

Mahtash Maryam Esfandiari

Yingnian Wu

Frederic R. Paik Schoenberg, Committee Chair

University of California, Los Angeles

2024

TABLE OF CONTENTS

**1 Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **1**

**2 Texas Hold'em Rules** . . . . . . . . . . . . . . . . . . . . . . . . . . . **3**

**3 Data Overview** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **5**

3.1 Data Description . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

3.1.1 Raw Data Example . . . . . . . . . . . . . . . . . . . . . . . 6

3.2 Data Processing . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

**4 Methods** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **12**

4.1 Definition of Luck and Skill . . . . . . . . . . . . . . . . . . . . . 12

4.2 Expected Profit Calculation Example . . . . . . . . . . . . . . . . . 12

4.3 Implementation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

**5 Results** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **15**

5.1 Betting Round Pre-flop . . . . . . . . . . . . . . . . . . . . . . . . 15

5.2 Betting Round Flop . . . . . . . . . . . . . . . . . . . . . . . . . . 16

5.3 Betting Round Turn . . . . . . . . . . . . . . . . . . . . . . . . . . 17

5.4 Betting Round River . . . . . . . . . . . . . . . . . . . . . . . . . . 18

5.5 Overall Comparison . . . . . . . . . . . . . . . . . . . . . . . . . . 19

**6 Discussion** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **23**

6.1 Analysis of the Simulation Results . . . . . . . . . . . . . . . . . . 23

6.2 Limitations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24

**7 Appendix** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **26**

7.1 Data Filtering Code . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

7.2 Data Extraction Code . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27

7.3 Luck & Skill Calculation Code . . . . . . . . . . . . . . . . . . . . . . . . 31

7.4 Data Visualization Code . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

**References** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . **42**

LIST OF FIGURES


5.1 Percentage of Total Absolute Gains due to Luck and Skill in Pre-flop . . . . . . 15

5.2 Percentage of Total Absolute Gains due to Luck and Skill in Flop . . . . . . . . 16

5.3 Percentage of Total Absolute Gains due to Luck and Skill in Turn . . . . . . . . 17

5.4 Percentage of Total Absolute Gains due to Luck and Skill in River . . . . . . . . 18

5.5 Percentage of Total Absolute Gains due to Luck and Skill . . . . . . . . . . . . 19

5.6 Comparison of Total Absolute Gains Percentage of Luck and Skill in 4 Betting Rounds . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

5.7 Comparison of Total Absolute Gains of Luck and Skill in 4 Betting Rounds . . . 21

5.8 Cumulative Percentage of Total Absolute Gains Attributed to Luck and Skill . . 22

5.9 Cumulative Percentage of Total Absolute Gains Attributed to Luck and Skill w/ Confidence Bounds . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22

LIST OF TABLES

3.1 First Five Rows of Card Information . . . . . . . . . . . . . . . . . . . . . . . . 10

3.2 First Five Rows of Pot Information . . . . . . . . . . . . . . . . . . . . . . . . . 10

3.3 First Five Rows of Two Players' Bets . . . . . . . . . . . . . . . . . . . . . . . . 11

4.1 First Five Rows of Luck and Skill Calculation . . . . . . . . . . . . . . . . . . . . 14

5.1 Percentage of Total Absolute Gains in 4 Betting Rounds . . . . . . . . . . . . . 20

5.2 Total Absolute Gains in 4 Betting Rounds . . . . . . . . . . . . . . . . . . . . . . 21

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Jingyu Fu for locating the dataset and meticulously filtering the necessary data for analysis. Additionally, I extend my thanks to Professor Schoenberg for providing the function to compute the expected profit and for his assistance in running the simulations.

# CHAPTER 1

# Introduction

Poker games, such as Omaha, Five Card Draw, and Texas Hold'em, have long been immensely popular across the world. This popularity stems from the variety of gameplay options, the ability to accommodate multiple players simultaneously, and the inclusive nature of the games, where people can participate and enjoy regardless of their genders, ages, or geographical locations, as long as they know the rules. Among these, Texas Hold'em stands out as one of the most popular poker games. It is renowned for its straightforward yet comprehensive rules, making it suitable for both casual play among friends and competitive gambling in casinos. The game's combination of straightforward rules and deep strategic potential makes it accessible to beginners while remaining challenging for professional players.

In Texas Hold'em, winning is not solely determined by the luck of the dealt cards but also by psychological strategy and tactics, allowing players to gain an advantage even with less-than-ideal hands. Players must constantly read their opponents, make calculated bets, and decide when to 'bluff' or 'fold', adding layers of complexity to each round. This dynamic interplay between chance and strategy has given rise to an enduring and unresolved question: Is Texas Hold'em more a game of luck or skill? The debate continues to intrigue both casual players and professional gamblers, making Texas Hold'em a fascinating subject of study.

Many studies have explored the question of whether poker is dominated by skill or luck, yielding significant findings. One such study analyzed the performance consistency of poker superstars, concluding that skill plays a substantial role in the game [1]. Another study reinforced this view by demonstrating that poker players improve their skills through experience and strategic learning, indicating that poker is fundamentally a game of skill [2]. Additionally, Levitt and Miles provided empirical evidence from the World Series of Poker, highlighting that the consistent success of top players underscores the significant impact of skill on poker outcomes [5].

A later study statistically defined luck and skill using the concept of expected profit and analyzed 27 hands from the 2015 World Series of Poker, significantly influencing our understanding of the subject [6]. Building on this idea, another study conducted a detailed simulation and revealed that skill generally has a more substantial impact on outcomes than luck, with minimal correlation between the two [3].

In this paper, we will adhere to the definitions and quantifications of luck and skill as established by the aforementioned studies [6]. Using real-world hand history data, we will perform computations and conduct analyses to derive meaningful implications from the aggregated results.

Following the introduction, we will first explain the rules of Texas Hold'em, which are essential for understanding the subsequent computations. In Section 3, we will discuss the data source and our process for extracting useful information from the raw text data. Section 4 will cover how we define and calculate the amounts of luck and skill in each game along with concrete examples for better understanding. Section 5 will present the results using various visualizations. Finally, Section 6 will discuss our findings and address some limitations of our definitions and methods.

# CHAPTER 2

## Texas Hold'em Rules

Before we delve deeper into the discussion of luck and skill, let us first review the specific rules of Texas Hold'em for those who may be unfamiliar with the game. Understanding these rules is essential to comprehend the later statistical computations and fully appreciate the intricate balance of luck and skill that characterizes this beloved poker variant.

Texas Hold'em is typically played with two to ten players using a standard 52-card deck. The objective is to win chips by either having the best hand at the showdown or by forcing all other opponents to fold (quit) before the showdown. At the start of each hand, two players to the left of the dealer post the small blind and big blind, respectively. These are mandatory bets that initiate the game. The dealer position rotates clockwise after each hand to ensure fairness. Each player is then dealt two private cards by the dealer, known as hole cards, which they will combine with five community cards to form their best five-card hand.

The game consists of four betting rounds: pre-flop, flop, turn, and river. After receiving their hole cards, players decide whether to call the big blind, raise, or fold. Betting continues clockwise until all players have either matched the highest bet or folded. The dealer then reveals the first three community cards, known as the flop, followed by another round of betting starting with the player to the left of the dealer. Next, a fourth community card, the turn, is dealt, followed by a third round of betting. Finally, the fifth community card, the river, is dealt, leading to the final round of betting.

If two or more players remain after the final betting round, a showdown occurs. Players reveal their hole cards, and the best five-card combination wins the pot. Hands are ranked according to standard poker hand rankings, from high card to royal flush. The hand rankings, from highest to lowest, are royal flush, straight flush, four of a kind, full house, flush, straight, three of a kind, two pair, one pair, and high card. If players have identical hands, the pot is split evenly.

In addition to these fundamental rules, strategies are also crucial to Texas Hold'em. Players must consider their position relative to the dealer, as being in a later position provides a strategic advantage by allowing them to see how their opponents act before making their own decisions. The strength of the hole cards, often referred to as "starting hands," also plays a significant role in determining a player's strategy. Players must evaluate the potential of their hole cards to form strong combinations with the community cards and may use bluffs to mislead their opponents into making poor decisions, such as folding stronger hands.

As we can see from the rules above, Texas Hold'em is not solely a game of chance but also one of skill and strategy. Players must consider their position, the strength of their hole cards, and their opponents' actions. The necessity of bluffing and reading opponents' behavior adds depth and complexity to the game. This interplay between luck and skill forms the core of Texas Hold'em, making it a compelling subject for deeper analysis and understanding.

# CHAPTER 3

# Data Overview

In this chapter, we discuss our data source, present an example of the data, and describe in detail how we filter out valid hands and extract the required information from the text.

## 3.1 Data Description

Online poker sites will automatically transcribe records of hands into .txt files as they are played. The data used in this paper is acquired from HandHQ.com, a business that observes online poker games and sells archives of these transcriptions for numerous poker rooms [4]. Many players use these data to supplement their own records, enhancing their understanding of opponents' behaviors and making more informed decisions at the table. By leveraging such detailed and extensive datasets, researchers can gain insights into the dynamics of poker beyond what is typically possible in live settings.

The entire dataset used in this research includes 17.25 million hands at NL50 (small-stake game), 6.79 million hands at NL200 (medium-stake game), and 2.87 million hands at NL1000 (high-stake game). These names refer to the standard buy-in for each game, which is one hundred big blinds (e.g., NL50 has a standard buy-in of $50, with big blinds of $0.50). This study encompasses databases with 212,224 players (NL50), 64,262 players (NL200), and 18,596 players (NL1000), all of whom have been anonymized to protect their identities. The hands analyzed span from December 2008 to April 2009, with all monetary values reported in United States dollars (USD).

### 3.1.1 Raw Data Example

The following text is an excerpt of the original text file.

```
Stage #3017254774: Holdem

No Limit $10, $2.50 ante - 2009-07-01 00:02:58 (ET)

Table: CLEO AVE (Real Money) Seat #1 is the dealer

Seat 1 - eXXdS46B0E4apgZgp7gHFw ($1,980 in chips)

Seat 2 - XrM1XlN29RxmLx3oZHhG0w ($1,492.50 in chips)

Seat 3 - X+u4T/E5ANkyZLKm1YjqwQ ($2,107.40 in chips)

Seat 4 - wyXD1O26Buq3VWHAij37Jg ($1,960 in chips)

Seat 5 - 3wT3m+GDGtVWU1KR2MWJ1Q ($1,146 in chips)

Seat 6 - /P+7Z0P/b7YiK60FW9dRAQ ($1,762.50 in chips)

eXXdS46B0E4apgZgp7gHFw - Ante $2.50

XrM1XlN29RxmLx3oZHhG0w - Ante $2.50

X+u4T/E5ANkyZLKm1YjqwQ - Ante $2.50

wyXD1O26Buq3VWHAij37Jg - Ante $2.50

3wT3m+GDGtVWU1KR2MWJ1Q - Ante $2.50

/P+7Z0P/b7YiK60FW9dRAQ - Ante $2.50

XrM1XlN29RxmLx3oZHhG0w - Posts small blind $5

X+u4T/E5ANkyZLKm1YjqwQ - Posts big blind $10

*** POCKET CARDS ***

wyXD1O26Buq3VWHAij37Jg - Folds

3wT3m+GDGtVWU1KR2MWJ1Q - Raises $40 to $40

/P+7Z0P/b7YiK60FW9dRAQ - Calls $40

eXXdS46B0E4apgZgp7gHFw - Calls $40

XrM1XlN29RxmLx3oZHhG0w - Folds

X+u4T/E5ANkyZLKm1YjqwQ - Folds
```

```
*** FLOP *** [Kc 7s 4h]
3wT3m+GDGtVWU1KR2MWJ1Q - Bets $75
/P+7Z0P/b7YiK60FW9dRAQ - Calls $75
eXXdS46B0E4apgZgp7gHFw - Folds
*** TURN *** [Kc 7s 4h] [6s]
3wT3m+GDGtVWU1KR2MWJ1Q - Bets $225
/P+7Z0P/b7YiK60FW9dRAQ - Calls $225
*** RIVER *** [Kc 7s 4h 6s] [9d]
3wT3m+GDGtVWU1KR2MWJ1Q - All-In $803.50
/P+7Z0P/b7YiK60FW9dRAQ - Calls $803.50
*** SHOW DOWN ***
3wT3m+GDGtVWU1KR2MWJ1Q - Shows [Ah Ac] (One pair, aces)
/P+7Z0P/b7YiK60FW9dRAQ - Shows [As Ad] (One pair, aces)
3wT3m+GDGtVWU1KR2MWJ1Q Collects $1,177 from main pot
/P+7Z0P/b7YiK60FW9dRAQ Collects $1,177 from main pot
*** SUMMARY ***
Total Pot($2,357) | Rake ($3)
Board [Kc 7s 4h 6s 9d]
Seat 1: eXXdS46B0E4apgZgp7gHFw (dealer) Folded on the FLOP
Seat 2: XrM1XlN29RxmLx3oZHhG0w (small blind) Folded on the POCKET CARDS
Seat 3: X+u4T/E5ANkyZLKm1YjqwQ (big blind) Folded on the POCKET CARDS
Seat 4: wyXD1O26Buq3VWHAij37Jg Folded on the POCKET CARDS
Seat 5: 3wT3m+GDGtVWU1KR2MWJ1Q won Total ($1,177) All-In HI:
($1,177) with One pair, aces [Ah Ac - P:Ah,P:Ac,B:Kc,B:9d,B:7s]
Seat 6: /P+7Z0P/b7YiK60FW9dRAQ won Total ($1,177) HI:($1,177)
with One pair, aces [As Ad - P:As,P:Ad,B:Kc,B:9d,B:7s]
```

The text above contains information about an entire hand, including four betting rounds, the community cards for each round, the amounts each player bets during each round, and the hole cards of the two players remaining in the game. With this information, we can do calculations on luck and skill. However, before performing these calculations, we need to parse the text file to filter out game descriptions similar to the one above and extract the specific information we need from each lengthy text.

## 3.2 Data Processing

The data processing can be generally divided into two parts: data filtering and data extraction. First, we need to search through the entire text file and filter out only the games that satisfy our criteria:

1. Each hand must have four complete betting rounds. If all other players fold except one before the final showdown, that hand will not be included.

2. Since our calculations will be based on the probability of winning between two players, we will only include hands that have exactly two players remaining in the final showdown.

3. We need to know the hole card information of the two players, which means they need to show their cards during the final showdown. If they do not show but mucked, that hand will also be excluded.

We began the filtering process by loading the data.table library for efficient data manipulation and file reading. The dataset was read using the fread function with the header argument set to FALSE and the separator specified as newline characters to properly read the poker hand data. To identify the start of each game, we searched for the keyword "Stage #", adding an ending index to capture the end of the last stage, ensuring complete coverage of all stages in the dataset. An empty character vector was initialized to store valid stages that include board information. We then extracted the content of each stage by its start and end indices within a loop, checking for the keyword "HI:" and excluding stages containing "[Does not show]". Stages meeting these criteria were considered valid and were appended to the vector. Finally, the filtered valid games were written to a new file named "filtered_poker.txt", creating a cleansed dataset for further analysis.

Next, to extract the required information from the text, we need to use different **REGEX** patterns to achieve that step by step. The general logic of each pattern is as follows:

1. First, we need the information of the five community cards. This information can be found in the 'SUMMARY' section after the word 'board', starting and ending with a square bracket. After extracting the whole sequence inside the square bracket, we can then split the string into separate cards using whitespace.

2. Second, we need to extract the hole cards of the remaining two players. We search for patterns starting with 'Shows' followed by a whitespace and a square bracket with the hole cards inside. We search for all matching strings, then use indexing and whitespace to split each card.

3. Next, we need to get the total bets in the pot before each betting round. To do this, we first split the text of each hand into different segments representing each betting round using key words such as 'TURN', 'RIVER'. Then we extract all the numbers with a dollar sign before it combined with certain action words like bets, raises, and so on. After getting all the numbers, we sum them up and use a cumulative sum to get the total money in the pot for each round.

4. Lastly, we need the amount of bets the two players put into the pot each round. We first locate which two players stay till the end using their seat numbers. Then we locate every line containing their unique index and use the same pattern used previously to extract how much money each player bets in each betting round.

Therefore, for each hand, we should get nine cards (4 hole cards + 5 community cards), five values for the total bets in the pot, and five values for each of the players' individual bets (10 values in total). In addition, we transform each card into an integer ranging from 1 to 52 where 1 = 2c, 2 = 3c, ..., 52 = As. In this paper, we process two raw text files and eventually obtain 277 hands that satisfy the criteria and provide all the information required for later computations. The following tables (Table 3.1 to Table 3.3) present the first few rows of the information extracted from the original text files.

| c1 | c2 | c3 | c4 | c5 | c6 | c7 | c8 | c9 |
|---|---|---|---|---|---|---|---|---|
| 39 | 13 | 52 | 26 | 12 | 45 | 29 | 44 | 21 |
| 23 | 12 | 8 | 51 | 2 | 37 | 25 | 38 | 40 |
| 3 | 10 | 9 | 30 | 4 | 6 | 45 | 21 | 36 |
| 33 | 19 | 49 | 35 | 42 | 9 | 36 | 11 | 26 |
| 3 | 13 | 15 | 26 | 12 | 38 | 39 | 4 | 29 |

Table 3.1: First Five Rows of Card Information

| pot0 | pot1 | pot2 | pot3 | pot4 |
|---|---|---|---|---|
| 30.00 | 150.00 | 300.00 | 750.00 | 2357.00 |
| 15.00 | 35.00 | 95.00 | 215.00 | 1195.00 |
| 15.00 | 20.00 | 60.00 | 150.00 | 450.00 |
| 15.00 | 100.00 | 240.00 | 240.00 | 240.00 |
| 15.00 | 60.00 | 150.00 | 150.00 | 450.00 |

Table 3.2: First Five Rows of Pot Information

| Abet0 | Bbet0 | Abet1 | Bbet1 | Abet2 | Bbet2 | Abet3 | Bbet3 | Abet4 | Bbet4 |
|---|---|---|---|---|---|---|---|---|---|
| 2.50 | 2.50 | 40.00 | 40.00 | 75.00 | 75.00 | 225.00 | 225.00 | 803.50 | 803.50 |
| 10.00 | 0.00 | 0.00 | 10.00 | 30.00 | 30.00 | 60.00 | 60.00 | 490.00 | 490.00 |
| 5.00 | 10.00 | 5.00 | 0.00 | 20.00 | 20.00 | 45.00 | 45.00 | 150.00 | 150.00 |
| 5.00 | 10.00 | 45.00 | 40.00 | 70.00 | 70.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 10.00 | 5.00 | 20.00 | 25.00 | 45.00 | 45.00 | 0.00 | 0.00 | 150.00 | 150.00 |

Table 3.3: First Five Rows of Two Players' Bets

For simplicity of naming, c1 and c2 represent the hole cards of Player A, and c3 and c4 represent the hole cards of Player B. Cards c5 through c9 are the five board cards. In addition, pot0 represents the total bets in the pot before the cards are dealt, while Abet0 and Bbet0 stand for the bets of Player A and Player B, respectively, before the cards are dealt. The same naming convention applies to other variables, with 1-4 each representing their corresponding information in betting rounds 1 through 4 (pre-flop, flop, turn, and river).

(Please check the Appendix for the complete R code)

# CHAPTER 4

# Methods

## 4.1 Definition of Luck and Skill

Before performing any calculations, we first need to define luck and skill in Texas Hold'em. In this paper, we use the definitions described in Professor Schoenberg's paper "Quantification of Luck and Skill in Texas Hold'em" [6]. To properly define luck and skill, we first need to introduce the concepts of *equity* and *expected profit* in poker.

In poker, *equity* represents a player's expected share of the pot, assuming no additional betting and based solely on the cards already revealed in the hand. This can be simply calculated by multiplying the total bets in the pot at that moment by the probability of a player winning the round. A player's *expected profit* for a specific segment of a poker hand is calculated as the increase in their equity minus the expenses incurred during that segment.

Following these definitions, we can define skill as the expected profit gained during the betting rounds and luck as the expected profit gained from the dealing of the cards [6]. In the following section, we will demonstrate in detail how to calculate the expected profit attributed to luck and skill for one specific betting round.

## 4.2 Expected Profit Calculation Example

Since there are four betting rounds, we need to perform four calculations per player to determine their expected profit for each round. For the pre-flop round, where no community cards are dealt, calculating the probability of winning by hand is nearly impossible because it involves considering $\binom{48}{5} = 1,712,304$ different combinations and determining how many result in a win for one player. This complexity also somewhat applies to the flop round, where nearly a thousand different combinations need to be considered. Consequently, we will rely on computational resources (computers) to handle these calculations.

For simplicity and clarity, consider the turn round where the four community cards are Spade 5, Club 10, Spade 9, and Spade 2. Player A has Spade A and Spade 7, while Player B holds Heart 5 and Heart 2. In this case, Player A already has a flush without revealing the fifth card, whereas Player B has two pairs. Therefore, Player A is more likely to win this hand intuitively. The only scenario where Player B could win this hand is if the fifth community card in the river round is either Club 5, Diamond 5, Club 2, or Diamond 2, which would give Player B a full house to beat Player A's flush. Thus, the probability of Player A winning is $1 - \frac{4}{44} = \frac{10}{11}$.

Suppose there are a total of 5000 chips in the pot at the moment, Player A's equity will be $\frac{10}{11} \times 5000 \approx 4545$ chips. In this example, we are unable to calculate the expected profit due to luck because we do not know the probability of Player A winning before the fourth community card is dealt. Suppose we denote the previous probability of winning as $p_0$ and the current one as $p_1$, we can simply calculate it by $5000 \times (p_1 - p_0)$.

Now, suppose during the turn round, both players raise 10000 chips, then the pot increases from 5000 chips to 25000 chips, and Player A's expected share of the pot also increases from 4545 to approximately 22727 chips, for an increase of 18,182 chips. The cost for Player A on the turn round is 10000 chips, so the expected profit due to Player A's skill for this round is 18182 chips $-$ 10000 chips $=$ 8182 chips.

## 4.3 Implementation

In the R function we developed (see Appendix), we apply the same logic as described above. For each round in each hand, we first calculate the probability of winning for one player, multiply it by the total bets in the pot, and then subtract the cost to get the expected profit for that round. We sum these values over the four rounds to get the total expected profit due to luck and skill. This process is iterated through every hand we collect to obtain aggregated and more reliable results.

The function itself takes nine cards, five total bets in the pot, and five bets from each player for each round as input. It returns the expected profit due to luck and skill for each round, as well as the absolute sum of the expected profit for luck and skill, respectively.

The table below (Table 4.1) shows the first few rows from the output of our luck and skill calculation function.

| luck0 | skill0 | luck1 | skill1 | luck2 | skill2 | luck3 | skill3 | lucktotal | skilltotal |
|---:|---:|---:|---:|---:|---:|---:|---:|---:|---:|
| -6.66 | 3.34 | 13.21 | 0.00 | 0.12 | 0.00 | 0.00 | 45.00 | 6.66 | 48.34 |
| 19.69 | 6.99 | -49.24 | 0.00 | -4.48 | -634.74 | 934.29 | 0.00 | 900.25 | 627.75 |
| -0.46 | -3.49 | -2.49 | 0.00 | -1.06 | 0.00 | -2.50 | 0.00 | 6.51 | 3.49 |
| -4.59 | -6.57 | 32.17 | 28.89 | 45.10 | 0.00 | 0.00 | 0.00 | 72.68 | 22.32 |
| 1.06 | 3.69 | 2.62 | 0.00 | -0.09 | 14.55 | 8.18 | 0.00 | 11.77 | 18.23 |

Table 4.1: First Five Rows of Luck and Skill Calculation

# CHAPTER 5

# Results

In this section, we present a series of visualizations for each betting round, as well as aggregated results for comparison and further analysis.

## 5.1 Betting Round Pre-flop

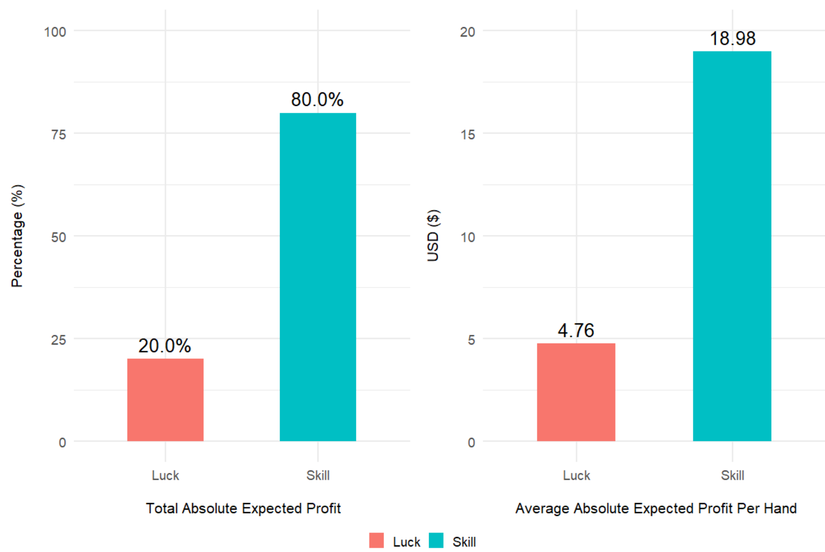


Figure 5.1: Percentage of Total Absolute Gains due to Luck and Skill in Pre-flop

In the pre-flop round, there is a significant gap between luck and skill, with approximately 80% of the equity gained in this round attributed to the player's skill. Additionally, the average absolute expected profit due to skill per hand is around \$19, compared to \$4.76 due to luck.

## 5.2 Betting Round Flop

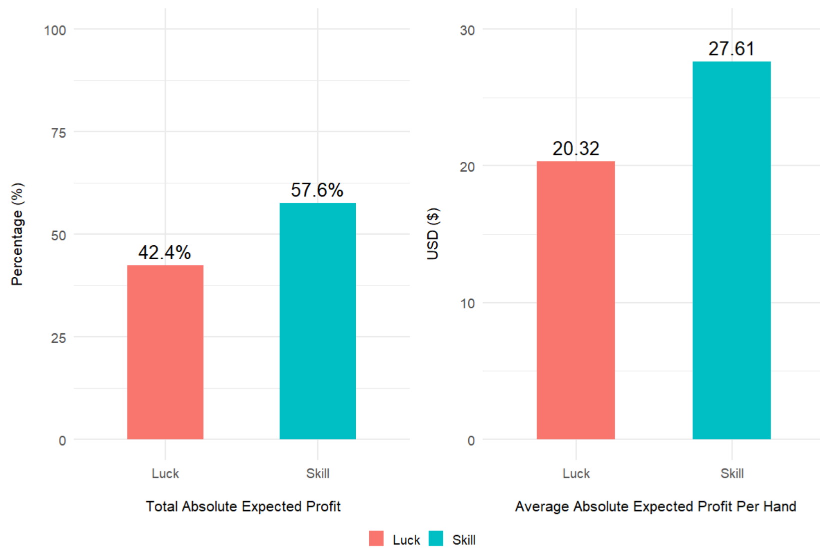


Figure 5.2: Percentage of Total Absolute Gains due to Luck and Skill in Flop

During the flop round, the difference between luck and skill narrows as three additional community cards are dealt. Combining all 277 hands collected, 42.4% of the total expected profit is attributed to luck, while 57.6% is attributed to skill. The average absolute expected profit per hand is \$20.32 for luck and \$27.61 for skill, demonstrating an overall increasing trend compared to the previous round.

## 5.3 Betting Round Turn

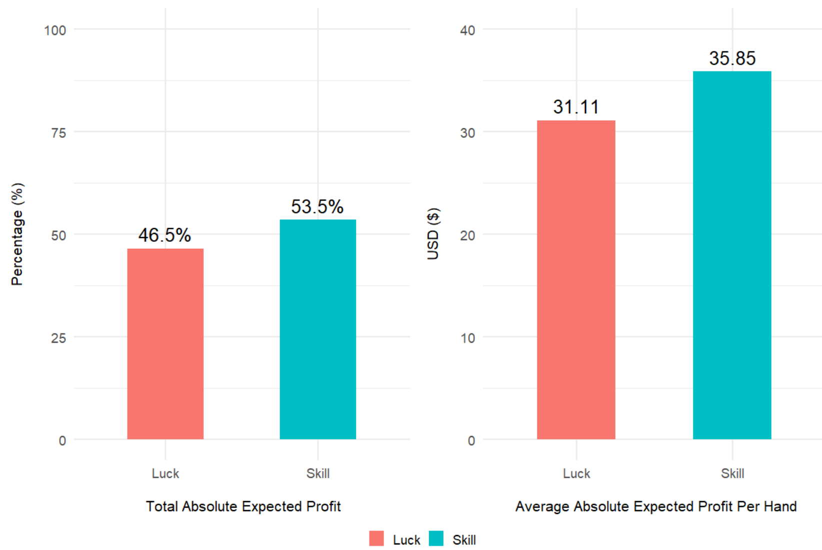


Figure 5.3: Percentage of Total Absolute Gains due to Luck and Skill in Turn

In the turn round, the gap between luck and skill narrows further, with 46.5% of the total expected profit due to luck and 53.5% due to skill. The average absolute expected profit per hand increases for both luck and skill, with $31.11 for luck and $35.85 for skill.

## 5.4 Betting Round River

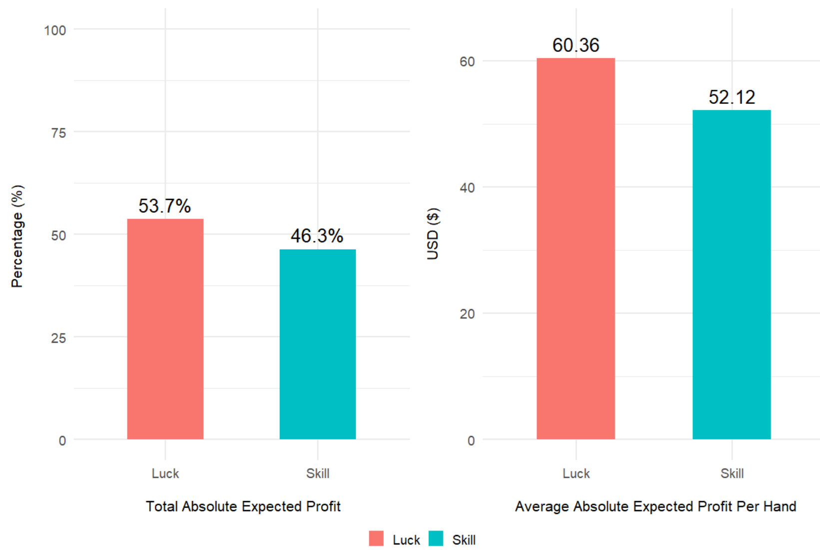


Figure 5.4: Percentage of Total Absolute Gains due to Luck and Skill in River

In the river round, the previous trend reverses as luck becomes more dominant: 53.7% of the total expected profit in the last betting round is due to luck, while 46.3% is due to skill. The average absolute expected profit per hand is also the highest for both categories among the four betting rounds, with $60.36 for luck and $52.12 for skill.

## 5.5 Overall Comparison


Figure 5.5: Percentage of Total Absolute Gains due to Luck and Skill

Overall, combining all four rounds, 43.8% of the total absolute expected profit is attributed to luck, while 56.2% is attributed to skill. In our analysis, skill appears to play a more important role in winning Texas Hold'em games.

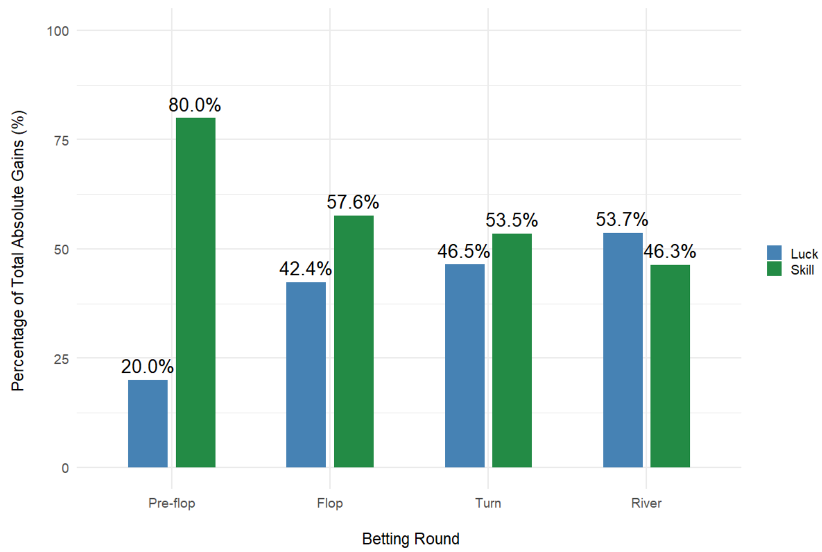


Figure 5.6: Comparison of Total Absolute Gains Percentage of Luck and Skill in 4 Betting Rounds

| Betting Round | Luck | Skill |
|---|---|---|
| Pre Flop | 20% | 80% |
| Flop | 42.4% | 57.6% |
| Turn | 46.5% | 53.5% |
| River | 53.7% | 46.3% |

Table 5.1: Percentage of Total Absolute Gains in 4 Betting Rounds

From the graph and table above, we can observe that the percentage of total absolute gains due to luck gradually increases through the four betting rounds, while the percentage of skill decreases.

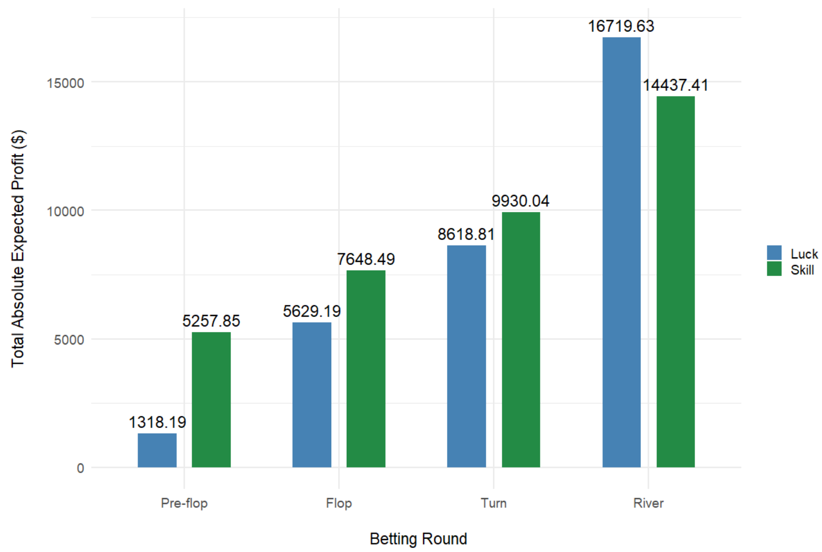


Figure 5.7: Comparison of Total Absolute Gains of Luck and Skill in 4 Betting Rounds

| Betting Round | Luck | Skill |
|---|---|---|
| Pre Flop | 1318.2 | 5257.9 |
| Flop | 5629.2 | 7648.5 |
| Turn | 8618.8 | 9930.0 |
| River | 16719.6 | 14437.4 |

Table 5.2: Total Absolute Gains in 4 Betting Rounds

The bar graph and table above show the total absolute expected profit in each betting round. We observe an overall rapidly increasing pattern through the four rounds, indicating that players tend to bet more as the game progresses. Notably, from pre-flop to flop, the total absolute expected profit for luck quadruples, and from turn to river, it nearly doubles.

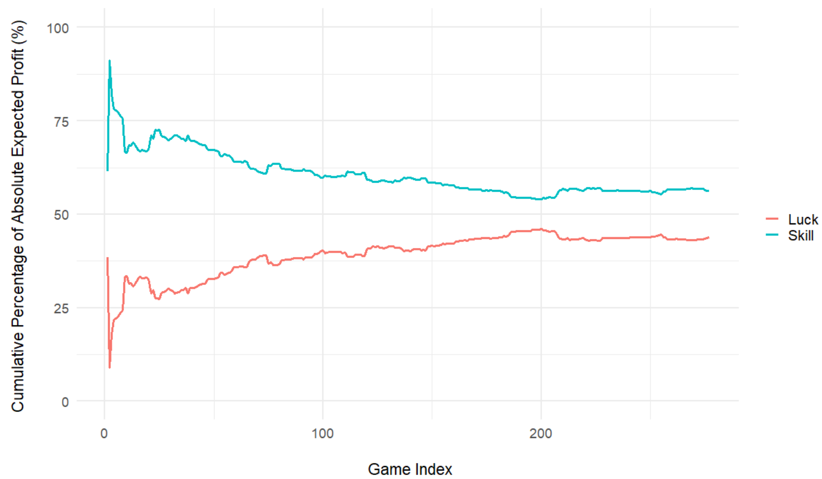


Figure 5.8: Cumulative Percentage of Total Absolute Gains Attributed to Luck and Skill

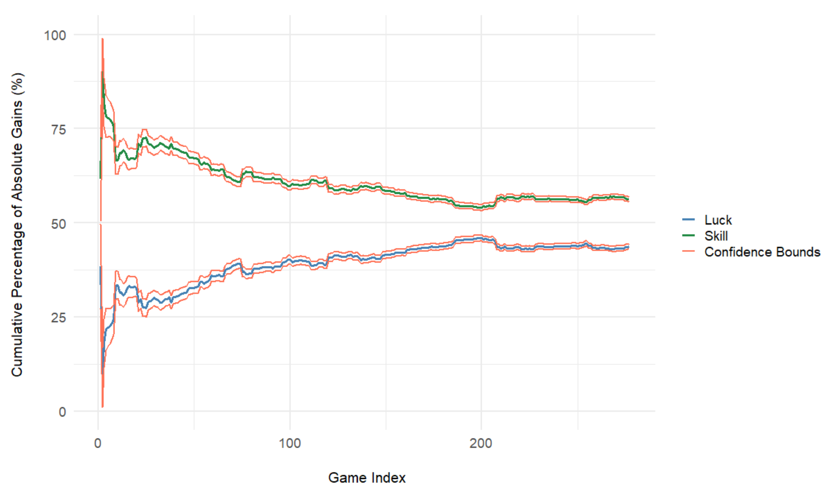


Figure 5.9: Cumulative Percentage of Total Absolute Gains Attributed to Luck and Skill w/ Confidence Bounds

The two graphs above show the cumulative percentage of the total expected profit attributed to luck and skill across all 277 hands with and without the 95% confidence bounds. Despite the initial zigzag patterns, both lines converge as the number of games increases.

# CHAPTER 6

# Discussion

## 6.1 Analysis of the Simulation Results

Our results in the previous chapter indicate a dynamic shift in the balance between luck and skill across the four betting rounds in Texas Hold'em. Initially, in the pre-flop round, skill is dominant, contributing approximately 80% of the total absolute expected profit. This is because the pot contains minimal money before the betting round, limiting the impact of luck on equity changes. In the flop round, with three community cards revealed, the influence of luck increases to 42.4%. The variability introduced by community cards reduces the impact of individual strategic decisions, requiring players to adapt their strategies based on new information.

During the turn round, with a fourth community card, the contributions of luck and skill become nearly balanced, at 46.5% and 53.5%, respectively. Players must reassess their strategies, as the evolving board can alter hand strengths. In the final river round, luck surpasses skill, accounting for 53.7% of the total expected profit, as the complete set of community cards largely determines outcomes. However, the strategic groundwork laid in earlier rounds will still influence the final results, underscoring the importance of both luck and skill throughout the game. This interplay between luck and skill highlights the complex nature of Texas Hold'em, where strategic decisions and adaptability play critical roles alongside the inherent uncertainty of the game.

Additionally, both the average expected profit per hand and total expected profit each round show an increasing trend as the game progresses. In the pre-flop round, the average absolute expected profit due to skill is around \$19, compared to \$4.76 due to luck. By the river round, these values rise to \$60.36 for luck and \$52.12 for skill. The total expected profit increases from \$1318 to \$16720 for luck, which is over 10 times larger, and from \$5258 to \$14437 for skill. This growth in expected profit reflects the compounding effect of strategic decisions and the escalating stakes as players commit more chips to the pot in each subsequent round. As the game progresses, the pots become larger, and the decisions more critical, leading to greater absolute gains from both luck and skill.

The shifting balance between luck and skill in Texas Hold'em highlights the game's complex dynamics. Early rounds emphasize strategic decision-making, giving skilled players an advantage in evaluating hole cards and predicting opponents' behaviors. As the game progresses, players need to adjust strategies based on community cards and changing probabilities. In later stages, the rising influence of luck underscores the game's inherent uncertainty, ensuring that even less experienced players have a chance to win.

Overall, our analysis demonstrates that Texas Hold'em is a complex game where both luck and skill play significant roles. Skilled players can leverage their strategic insights to maximize their expected profits, but the element of luck ensures that even less experienced players have a chance to win. The balance between luck and skill not only shapes the strategies employed by players but also contributes to the enduring popularity and dynamic nature of Texas Hold'em poker.

## 6.2 Limitations

Our research also has its limitations and room for improvement in the future. Firstly, we included only 277 hands in this analysis. Adding more data could lead to different results, potentially showing a more even distribution of luck and skill. The main factor preventing us from including more data is the computational expense of our luck and skill calculation function. As mentioned before, in the pre-flop round, we need to calculate $\binom{48}{5}$ possible outcomes to get the probability of winning, which makes even running one iteration take at least ten minutes on a local machine. In the future, we might leverage tools like cloud computing to speed up our calculations or simplify our methodologies by approximating the probability in the pre-flop round.

Secondly, the calculations performed in this paper are not extremely rigorous in that our function only calculates the probability of winning between two players. It does not consider cases where more than two players are involved in a specific round. The results would be more accurate if we could develop a more advanced algorithm that accounts for the number of players left in each betting round.

Lastly, our definitions of luck and skill have their drawbacks. One issue is that luck and skill are often correlated in practice, which may lead to confounded results when analyzing their distribution. Therefore, we might need a more sophisticated approach to both define and quantify luck and skill statistically.

# CHAPTER 7

# Appendix

## 7.1 Data Filtering Code

```
library(data.table)
poker <- fread('your path to the raw text file', header = FALSE, sep = "\n")
lines <- as.character(poker$V1)
# Identify indices for game starts
stage_starts <- grep("Stage #", lines)
# Add an ending index for the last stage to capture its end
stage_starts <- c(stage_starts, length(lines) + 1)
# Prepare to capture valid stages with board information
valid_data <- character(0)
for (i in 1:(length(stage_starts) - 1)) {
  # Extract current stage content
  stage_lines <- lines[stage_starts[i]:(stage_starts[i+1]-1)]
  # Check for 'HI:' within current stage and filter out [Does not show\\]""
  if (any(grep("HI:", stage_lines)) &&
      !any(grep("\\[Does not show\\]", stage_lines))) {
    valid_data <- c(valid_data, stage_lines)
  }
}
writeLines(valid_data, con = "filtered_poker.txt")
```

### 7.2 Data Extraction Code

```
text_content <- read_file("filtered_poker.txt")
# Splitting the content into games using the pattern
# that looks ahead for "Stage #"
# This regex ensures we only split
# where "Stage #" is at the start of a new line
games <- strsplit(text_content, "(?<=\n)(?=Stage #)", perl=TRUE)[[1]]
# Filter out any entries that are just new lines or don't contain game data
games <- games[sapply(games, nchar) > 1]
#function for transforming each card into an integer
transform_card <- function(card) {
  # Define the mappings for suits and card numbers
  suits <- c(c = 1, d = 2, h = 3, s = 4)
  numbers <- c('2' = 2, '3' = 3, '4' = 4, '5' = 5, '6' = 6, '7' = 7,
               '8' = 8, '9' = 9, '10' = 10, 'J' = 11,
               'Q' = 12, 'K' = 13, 'A' = 14)
  suit_char <- substr(card, nchar(card), nchar(card))
  num_char <- substr(card, 1, nchar(card) - 1)
  num_val <- unname(numbers[num_char])
  suit_val <- unname(suits[suit_char])
  # Return the value
  13*(suit_val-1) + num_val - 1
}
betting_regex <- "\\b(Ante|Posts\\s+(small|big)\\s+blind|Bets|
Calls|Raises|All-In)\\s+\\$([0-9]+\\.?[0-9]*)"
# Function to calculate cumulative pot
calculate_cumulative_pot <- function(text_segment) {
```

```
  matches <- str_extract_all(text_segment, betting_regex)
  matches <- unlist(matches)
  amounts <- as.numeric(gsub("[^0-9.]", "", matches))
  sum(amounts, na.rm = TRUE)
}
find_max_bet <- function(text_segment) {
  # Regex pattern to capture bets and specifically handle "Raises to"
  # This pattern looks for monetary amounts that may include decimals.
  bets <- str_extract_all(text_segment, "\\b(Posts\\s+(small|big)\\s+
  blind|Bets|Calls|Ante)\\s+\\$(\\d+\\.?\\d*)|Raises\\s+\\$\\d+\\.?
  \\d*\\s+to\\s+\\$(\\d+\\.?\\d*)")
  bets <- unlist(bets)
  # Extract only the amounts, specifically
  # capturing the final amounts from Raises
  amounts <- as.numeric(gsub(".*\\$(\\d+\\.?\\d*).*", "\\1", bets))
  if (length(amounts) == 0) {  # If no bets were made (all checks)
    return(0)
  } else {
    return(max(amounts, na.rm = TRUE))  # Find the maximum bet
  }
}
results_df <- tibble(
  c1 = numeric(), c2 = numeric(), c3 = numeric(), c4 = numeric(),
  c5 = numeric(), c6 = numeric(), c7 = numeric(), c8 = numeric(),
  c9 = numeric(),
  pot0 = numeric(), pot1 = numeric(),
  pot2 = numeric(), pot3 = numeric(), pot4 = numeric(),
```

```
  Abet0 = numeric(), Bbet0 = numeric(),
  Abet1 = numeric(), Bbet1 = numeric(),
  Abet2 = numeric(), Bbet2 = numeric(),
  Abet3 = numeric(), Bbet3 = numeric(),
  Abet4 = numeric(), Bbet4 = numeric()
)
for (i in seq_along(games)) {
  game_text <- games[i]
  player_hands <- str_extract_all(game_text,
  "(?<=\\[)[\\dJQKA]{1,2}[cdhs]{1}
  (?: [\\dJQKA]{1,2}[cdhs]{1})(?=\\])")[[1]]
  player1 <- player_hands[1]
  player1 <- strsplit(player1, " ")[[1]] %>%
  sapply(transform_card) %>% unname()
  player2 <- player_hands[2]
  player2 <- strsplit(player2, " ")[[1]] %>%
  sapply(transform_card) %>% unname()
  c1 <- player1[1]; c2 <- player1[2]; c3 <- player2[1]; c4 <- player2[2]
  board_cards <- str_extract(game_text, "(?<=Board \\[)[\\dJQKA]{1,2}[cdhs]
  (?: [\\dJQKA]{1,2}[cdhs]){4}")
  board_cards <- strsplit(board_cards, " ")[[1]] %>%
  sapply(transform_card) %>% unname()
  c5 <- board_cards[1]; c6 <- board_cards[2]; c7 <- board_cards[3]
  c8 <- board_cards[4]; c9 <- board_cards[5]
  segments <- strsplit(game_text,
  "\\*\\*\\* (POCKET CARDS|FLOP|TURN|RIVER|SHOW DOWN)
   \\*\\*\\*")[[1]]
```

```
pot0 <- calculate_cumulative_pot(segments[1])
pot1 <- calculate_cumulative_pot(paste(segments[1:2], collapse = ""))
pot2 <- calculate_cumulative_pot(paste(segments[1:3], collapse = ""))
pot3 <- calculate_cumulative_pot(paste(segments[1:4], collapse = ""))
pot4 <- calculate_cumulative_pot(paste(segments[1:5], collapse = ""))
summary_text <- str_split(game_text,
"\\*\\*\\* SUMMARY \\*\\*\\*")[[1]][2]
indices <- str_extract_all(summary_text,
"Seat \\d+:\\s*([^\\s]+)\\s.*?\\[.+\\]")
extracted_indices <- lapply(indices,
function(x) str_match(x, "Seat \\d+:\\s*([^\\s]+)")[,2])[[1]]
player1_index <- extracted_indices[1] %>%
str_replace_all("[+]", "\\\\\\0")
player2_index <- extracted_indices[2] %>%
str_replace_all("[+]", "\\\\\\0")
before_cards_text <- str_split(segments[1], "\n")[[1]]
after_cards_text <- str_split(segments[2], "\n")[[1]]
player1_before_cards <- before_cards_text[str_detect(before_cards_text,
player1_index)] %>% paste(collapse = "\n")
player2_before_cards <- before_cards_text[str_detect(before_cards_text,
player2_index)] %>% paste(collapse = "\n")
Abet0 <- calculate_cumulative_pot(player1_before_cards)
Bbet0 <- calculate_cumulative_pot(player2_before_cards)
player1_after_cards <- after_cards_text[str_detect(after_cards_text,
player1_index)] %>% paste(collapse = "\n")
player2_after_cards <- after_cards_text[str_detect(after_cards_text,
player2_index)] %>% paste(collapse = "\n")
```

```
  Abet1 <- calculate_cumulative_pot(player1_after_cards)
  Bbet1 <- calculate_cumulative_pot(player2_after_cards)
  Abet2 <- find_max_bet(segments[3]); Bbet2 <- find_max_bet(segments[3])
  Abet3 <- find_max_bet(segments[4]); Bbet3 <- find_max_bet(segments[4])
  Abet4 <- find_max_bet(segments[5]); Bbet4 <- find_max_bet(segments[5])
  results_df <- results_df %>%
    add_row(
      c1 = c1, c2 = c2, c3 = c3, c4 = c4,
      c5 = c5, c6 = c6, c7 = c7, c8 = c8,
      c9 = c9,
      pot0 = pot0, pot1 = pot1,
      pot2 = pot2, pot3 = pot3, pot4 = pot4,
      Abet0 = Abet0, Bbet0 = Bbet0, Abet1 = Abet1, Bbet1 = Bbet1,
      Abet2 = Abet2, Bbet2 = Bbet2, Abet3 = Abet3, Bbet3 = Bbet3,
      Abet4 = Abet4, Bbet4 = Bbet4
    )
}
results_df <- na.omit(results_df)
saveRDS(results_df, file = "poker_info.rds")
```

### 7.3 Luck & Skill Calculation Code

```
source("holdemcode.r")
poker_info <- readRDS("poker_info.rds")
poker_info <- as.matrix(poker_info)
luckskill <- function(c1,c2,c3,c4,c5,c6,c7,c8,c9,Abet0,Bbet0,pot0,
                      Abet1,Bbet1,pot1,Abet2,Bbet2,pot2,Abet3,Bbet3,
```

```
                    pot3,Abet4,Bbet4,pot4){
## c1-c9 are the cards. c1-c2 are the cards of player A,
## c3-c4 are the cards of player B, and c5-c9 are the board.
## Abet0, Bbet0 are the amounts put in by players A and B
## before the cards are dealt and pot0 is the total in the pot,
## For i = 1,2,3,4, Abeti, Bbeti, poti are amounts put in and
## total in the pot after betting round i.
## 1. See what A's probability of winning is pre-flop.
n = choose(48,5)
result = rep(0.5,n); a = c(c1,c2,c3,c4)
b = c(1:52)[-a]; i = 0
for(i1 in 1:44){for(i2 in ((i1+1):45)){for(i3 in ((i2+1):46)){
  for(i4 in ((i3+1):47)){for(i5 in ((i4+1):48)){
  flop1 = c(b[i1],b[i2],b[i3],b[i4],b[i5])
  ## Player A's cards and the board together.
  flop2 = switch2(c(flop1,a[1:2]))
  ## Player B's cards and the board together.
  flop3 = switch2(c(flop1,a[3:4]))
  b1 = handeval(flop2$num,flop2$st)
  b2 = handeval(flop3$num,flop3$st)
  i = i+1; if(b1 > b2) result[i] = 1; if(b1 < b2) result[i] = 0}}}}}
p0 = mean(result)
## In the end we have basically the % chance of A winning,
## with ties split 50-50.
## 2. Calculate A's luck gain before the flop.
luck0 = pot0 * p0 - Abet0
## 3. Calculate A's skill gain before the flop
```

```
## = expected profit after first betting round minus expected profit
## right before the first betting round.
skill0 = pot1 * p0 - Abet1 - pot0 * p0
## 4. See what A's probability of winning is after the flop is dealt.
n = choose(45,2)
result = rep(0.5,n); a = c(c1,c2,c3,c4,c5,c6,c7)
b = c(1:52)[-a]; i = 0
for(i1 in 1:44){for(i2 in ((i1+1):45)){
  flop1 = c(b[i1],b[i2])
  ## Player A's cards and the board together.
  flop2 = switch2(c(flop1,a[1:2],c5,c6,c7))
  ## Player B's cards and the board together.
  flop3 = switch2(c(flop1,a[3:4],c5,c6,c7))
  b1 = handeval(flop2$num,flop2$st)
  b2 = handeval(flop3$num,flop3$st)
  i = i+1; if(b1 > b2) result[i] = 1; if(b1 < b2) result[i] = 0}}
p1 = mean(result)
## 5. Calculate A's luck gain when the flop is dealt.
luck1 = pot1 * p1 - pot1 * p0
## 6. Calculate A's skill gain after the flop.
skill1 = pot2 * p1 - Abet2 - pot1 * p1
## 7. See what A's probability of winning is after the turn is dealt.
n = 44
result = rep(0.5,n); a = c(c1,c2,c3,c4,c5,c6,c7,c8)
b = c(1:52)[-a]; i = 0
for(i1 in 1:44){
  flop1 = b[i1]
```

```
  ## Player A's cards and the board together.
  flop2 = switch2(c(flop1,a[1:2],c5,c6,c7,c8))
  ## Player B's cards and the board together.
  flop3 = switch2(c(flop1,a[3:4],c5,c6,c7,c8))
  b1 = handeval(flop2$num, flop2$st)
  b2 = handeval(flop3$num, flop3$st)
  i = i+1;  if(b1 > b2) result[i] = 1; if(b1 < b2) result[i] = 0}
p2 = mean(result)
## 8. Calculate A's luck gain when the turn is dealt.
luck2 = pot2 * p2 - pot2 * p1
## 9. Calculate A's skill gain after the turn.
skill2 = pot3 * p2 - Abet3 - pot2 * p2
## 10. See who won, after the river was dealt.
result = rep(0.5,n); a = c(c1,c2,c3,c4,c5,c6,c7,c8,c9)
## Player A's cards and the board together.
flop2 = switch2(c(a[1:2], a[5:9]))
## Player B's cards and the board together.
flop3 = switch2(c(a[3:9]))
b1 = handeval(flop2$num, flop2$st)
b2 = handeval(flop3$num, flop3$st)
p3 = (b1 > b2) + .5*(b1 == b2) ## 1 if A won, and 0.5 if it was a tie.
## 11. Calculate A's luck gain when the river is dealt.
luck3 = pot3 * p3 - pot3 * p2
## 12. Calculate A's skill gain after the river.
skill3 = pot4 * p3 - Abet4 - pot3 * p3
## 13. Add it up.
lucktotal = abs(luck0 + luck1 + luck2 + luck3)
```

```
  skilltotal = abs(skill0 + skill1 + skill2 + skill3)
  return(c(luck0,skill0,luck1,skill1,luck2,skill2,
           luck3,skill3,lucktotal,skilltotal))
}
for(i in 1:nrow(poker_info)){
  g_i <- poker_info[i,]
  luckskill_list[[i]] <- luckskill(g_i[1],g_i[2],g_i[3],
                                   g_i[4],g_i[5],g_i[6],
                                   g_i[7],g_i[8],g_i[9],
                                   g_i['Abet0'],g_i['Bbet0'],
                                   g_i['pot0'],g_i['Abet1'],g_i['Bbet1'],
                                   g_i['pot1'],g_i['Abet2'],g_i['Bbet2'],
                                   g_i['pot2'],g_i['Abet3'],g_i['Bbet3'],
                                   g_i['pot3'],g_i['Abet4'],g_i['Bbet4'],
                                   g_i['pot4'])
}
saveRDS(luckskill_list, "output.rds")
```

## 7.4 Data Visualization Code

```
library(tidyverse)
library(ggpubr)
poker_list <- readRDS("outputrds")
# convert into a matrix
poker <- do.call(rbind, poker_list) %>% as.data.frame()
colnames(poker) <- c("luck0","skill0","luck1","skill1","luck2",
                     "skill2","luck3","skill3","lucktotal","skilltotal")
```

```
sums <- colSums(abs(poker))
betting_round1 <- sums[1:2]; betting_round2 <- sums[3:4]
betting_round3 <- sums[5:6]; betting_round4 <- sums[7:8]
total <- sums[9:10]
# Calculate total percentages
total_sum <- sum(total)
percentages <- total / total_sum * 100
percentage_df <- data.frame(
  Column = c("Luck", "Skill"),
  Percentage = percentages
)
ggplot(percentage_df, aes(x = Column, y = Percentage, fill = Column)) +
  geom_bar(stat = "identity", width = 0.5) +
  xlab("\n Total Absolute Expected Profit") +
  geom_text(aes(label = sprintf("%.1f%%", Percentage)),
            vjust = -0.5, color = "black") +
  scale_y_continuous(limits = c(0, 100)) + labs(fill = "") +
  ylab("Percentage (%)\n") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5),
        axis.title.x = element_text(size = 13),
        axis.title.y = element_text(size = 13))
# same syntax for betting round 1-4, so only demonstrate round 1
total_bet1 <- sum(betting_round1)
percentages <-  betting_round1 / total_bet1 * 100
percentage_df <- data.frame(
  Column = c("Luck", "Skill"),
```

```
  Percentage = percentages
)
gain_df <- data.frame(
  Column = c("Luck", "Skill"),
  total = betting_round1 / 277
)
round1_per <- ggplot(percentage_df, aes(x = Column,
y = Percentage, fill = Column)) +
  geom_bar(stat = "identity", width = 0.5) +
  xlab("\n Total Absolute Expected Profit") +
  geom_text(aes(label = sprintf("%.1f%%", Percentage)),
            vjust = -0.5, color = "black") +
  scale_y_continuous(limits = c(0, 100)) + labs(fill = "") +
  ylab("Percentage (%)\n") +
  theme_minimal() +
  theme(axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12))
round1_sum <- ggplot(gain_df, aes(x = Column, y = total, fill = Column)) +
  geom_bar(stat = "identity", width = 0.5) +
  xlab("\n Average Absolute Expected Profit Per Hand") +
  geom_text(aes(label = round(total,2)), vjust = -0.5, color = "black") +
  labs(fill = "") + scale_y_continuous(limits = c(0, 20)) +
  ylab("USD ($)\n") +
  theme_minimal() +
  theme(axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12))
ggarrange(round1_per, round1_sum, ncol = 2,
```

```
         common.legend = TRUE, legend = "bottom")
# combine 4 betting rounds into one graph
tab <- rbind(betting_round1, betting_round2, betting_round3, betting_round4)
luck_4_percent <- round(tab / rowSums(tab) * 100, 3)
data <- data.frame(
  Round = factor(c("Pre-flop", "Flop", "Turn", "River"),
                 levels = c("Pre-flop", "Flop", "Turn", "River")),
  Luck = luck_4_percent[,1],
  Skill = luck_4_percent[,2]
)
data_long <- pivot_longer(data, cols = c(Luck, Skill),
                          names_to = "variable",
                          values_to = "value")
ggplot(data_long, aes(x = Round, y = value, fill = variable)) +
  geom_bar(width = 0.5, stat = "identity",
           position = position_dodge(width = 0.6)) +
  labs(x = "\n Betting Round",
       y = "Percentage of Total Absolute Gains (%) \n") +
  theme_minimal() + scale_y_continuous(limits = c(0, 100)) +
  theme(axis.title.x = element_text(size = 13),
        axis.title.y = element_text(size = 13)) +
  geom_text(aes(label = sprintf("%.1f%%", value)),
            position = position_dodge(width = 0.6),
            vjust = -0.5, color = "black",size = 4,
            hjust = 0.5) + labs(fill = "") +
  scale_fill_manual(values = c("Luck" = "steelblue", "Skill" = "#238b45"))
data <- data.frame(
```

```
  Round = factor(c("Pre-flop", "Flop", "Turn", "River"),
                levels = c("Pre-flop", "Flop", "Turn", "River")),
  Luck = tab[,1],
  Skill = tab[,2]
)
data_long <- pivot_longer(data, cols = c(Luck, Skill),
                          names_to = "variable", values_to = "value")
a <- ggplot(data_long, aes(x = Round, y = value, fill = variable)) +
  geom_bar(width = 0.5, stat = "identity",
           position = position_dodge(width = 0.7)) +
  labs(x = "\n Betting Round",
       y = "Total Absolute Expected Profit \n") +
  theme_minimal(base_size=18) + scale_y_continuous(limits = c(0, 17000)) +
  theme(axis.title.x = element_text(size = 17),
        axis.title.y = element_text(size = 17)) +
  geom_text(aes(label = round(value,2)),
            position = position_dodge(width = 0.7),
            vjust = -0.5, color = "black",size = 6,
            hjust = 0.5) + labs(fill = "") +
  scale_fill_manual(values = c("Luck" = "steelblue", "Skill" = "#238b45"))
# cumulative percentage
poker <- poker %>%
  mutate(
    luck_cumsum = cumsum(lucktotal),
    skill_cumsum = cumsum(skilltotal),
    total = luck_cumsum + skill_cumsum,
    luck_per = luck_cumsum / total * 100,
```

```
    skill_per = skill_cumsum / total  * 100,
    Index = 1:nrow(poker))
cumsum_df <- poker %>%
  select(Index, luck_per, skill_per) %>%
  pivot_longer(
    cols = c("luck_per", "skill_per"),
    names_to = "Source",
    values_to = "Cumulative Percentage") %>%
    mutate(Source = recode(Source,
           luck_per = "Luck",
           skill_per = "Skill"))
ggplot(cumsum_df, aes(x = Index, y = `Cumulative Percentage`,
                     color = Source)) + geom_line(linewidth = 1) +
  scale_y_continuous(limits = c(0, 100)) +
  labs(x = "\n Game Index",
       y = "Cumulative Percentage of Absolute Expected Profit (%) \n",
       color = "") + theme_minimal() +
  theme(legend.text = element_text(size = 12))
# w/ confidence bounds
s <- sd(poker$luck_per); n <- seq_along(poker$luck_per)
poker$luck_lower <- poker$luck_per - 1.96 * s / sqrt(n)
poker$luck_upper <- poker$luck_per + 1.96 * s / sqrt(n)
s <- sd(poker$skill_per); n <- seq_along(poker$skill_per)
poker$skill_lower <- poker$skill_per - 1.96 * s / sqrt(n)
poker$skill_upper <- poker$skill_per + 1.96 * s / sqrt(n)
ggplot(poker, aes(x = Index)) +
  geom_line(aes(y = luck_per, color = "Luck"), linewidth = 1.2) +
```

```
geom_line(aes(y = luck_lower, color = "Confidence Bounds"),
linewidth = 0.8) +
geom_line(aes(y = luck_upper, color = "Confidence Bounds"),
linewidth = 0.8) +
geom_line(aes(y = skill_per, color = "Skill"), linewidth = 1.2) +
geom_line(aes(y = skill_lower, color = "Confidence Bounds"),
linewidth = 0.8) +
geom_line(aes(y = skill_upper, color = "Confidence Bounds"),
linewidth = 0.8) +
scale_color_manual(values = c("Luck" = "steelblue",
                              "Skill" = "#238b45",
                              "Confidence Bounds" = "coral1"), name = "",
                   breaks = c("Luck", "Skill", "Confidence Bounds"),
                   labels = c("Luck", "Skill", "Confidence Bounds")) +
labs(x = "\n Game Index",
     y = "Cumulative Percentage of Absolute Gains (%) \n") +
theme_minimal() +
scale_y_continuous(limits = c(0, 100)) +
theme(axis.title.x = element_text(size = 13),
      axis.title.y = element_text(size = 13))
```

## REFERENCES

[1] R. Croson, P. Fishman, and D. Pope. Poker superstars: Skill or luck? *Chance*, 21:28, 2008.

[2] M.A. Dedonno and D.K. Detterman. Poker is a skill. *Gaming Law Review*, 12(1):31–36, 2008.

[3] Zhaoyu Guo, Irwin Khuu, Kevin Zhu, Jeffrey S. Rosenthal, and Frederic Paik Schoenberg. Distinguishing luck from skill through statistical simulation: A case study. *Communications in Statistics - Simulation and Computation*, 51(5):2537–2559, 2022.

[4] HandHQ.com. Obfuscated datamined hand histories, 2011.

[5] S. D. Levitt and T. J. Miles. The role of skill versus luck in poker: evidence from the world series of poker. *Journal of Sports Economics*, 15:31–44, 2014.

[6] Frederic Paik Schoenberg. Quantification of luck and skill in texas hold'em. *Significance*, 15(6):30–33, 2018.