EEE 554 Project

"Probabilities of 5-Card Poker Hands: Monte Carlo Simulation vs. Analytical Calculations Using MATLAB"

Dhanraj Bhosale

1225506620

Abstract:

This project calculates the probabilities of 5-card poker hands drawn from a deck of 20 cards consisting of 4 suites of the first 5 ranks (1, 2, 3, 4, 5). The project considers four different categories of poker hands: four of a kind, one-pair, two-pair, and three of a kind. The analytical calculations are performed to estimate the probabilities of each category of poker hands. Monte Carlo simulation is used to estimate the probabilities, and the results are compared with the analytical calculations. The project is implemented using MATLAB programming language. The analytical calculations and Monte Carlo simulation both provide probabilities for each category of poker hands. The results show that the analytical calculations and Monte Carlo simulation provide similar probabilities.

Introduction:

In this report, we will develop a program in MATLAB to calculate the probabilities of different 5-card poker hands using Monte Carlo simulation. The poker game will be played with a deck of 20 cards, including four suits of the first 5 ranks (Ace, 2, 3, 4, and 5).

We will consider four types of poker hands: four of a kind, one-pair, two-pair, and three-of-a-kind. Additionally, we will calculate the actual probabilities of each hand using combinatorics, compare them with the probabilities obtained from the Monte Carlo simulation, and analyze the accuracy of Monte Carlo simulation based on the sample size.

Given:

Given 5-card poker hands.
Given desk of 20 = 4*5

So, The nCr function can be used to calculate hand frequencies; entering nCr with n=20 and r=5, $\binom{20}{5}$ = **15,504**

Analytical Calculations for Desk 20:

- **Four of a kind ->**
  A four-of-a-kind hand consists of four cards of the same rank, and one other card. We can choose the rank of the four cards in 5 ways. For each rank, we can choose four cards from the available four cards of that rank in 1 way. We can choose the rank of the fifth card in 4 ways, and we can choose any card from the available four cards of that rank in 4 ways. Therefore, the total number of possible four-of-a-kind hands is:

$$\binom{5}{1}\binom{4}{4}\binom{4}{1}\binom{4}{1} = 80$$

Probability = 80/15,504 = **0.005159958720330237**

- **One-Pair ->**
  A one-pair hand consists of two cards of the same rank, and three other cards, each of a different rank. We can choose the rank of the pair in 5 ways. We can choose two cards from the available four cards of that rank in C(4,2) ways. We can choose the ranks of the other three cards in 4 * 3 * 2 ways. For each rank, we can choose one card from the available four cards of that rank in 1 way. Therefore, the total number of possible one-pair hands is:

  $\binom{5}{1}\binom{4}{2}\binom{4}{3}\binom{4}{1}^3 = 7680$

  Probability = 7680/15,504 = **0.4953560371517028**

- **Two-Pair ->**
  A two-pair hand consists of two cards of one rank, two cards of another rank, and one other card of a different rank. We can choose the ranks of the two pairs in C(5,2) ways. For each pair, we can choose two cards from the available four cards of that rank in C(4,2) ways. We can choose the rank of the fifth card in 3 ways, and we can choose any card from the available four cards of that rank in 4 ways. Therefore, the total number of possible two-pair hands is:

  $\binom{5}{2}\binom{4}{2}^2\binom{3}{1}\binom{4}{1} = 4320$

  Probability = 4320/15,504 = **0.2786377708978328**

- **Three of a kind ->**
  A three-of-a-kind hand consists of three cards of the same rank, and two other cards, each of a different rank. We can choose the rank of the three cards in 5 ways. We can choose three cards from the available four cards of that rank in C(4,3) ways. We can choose the ranks of the other two cards in 4 * 3 ways. For each rank, we can choose one card from the available four cards of that rank in 1 way. Therefore, the total number of possible three-of-a-kind hands is:

$$\binom{5}{1}\binom{4}{3}\binom{4}{2}\binom{4}{1}^2 = 1920$$

<u>**Analytical Calculations for Desk 52:**</u>

- **Four of a kind ->** $\binom{13}{1}\binom{4}{4}\binom{12}{1}\binom{4}{1} = 624$

  Probability = 624/2,598,960 = **0.0002401**

- **One-Pair ->** $\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{1}^3 = 1{,}098{,}240$

  Probability = 1,098,240/2,598,960 = **0.422569**

- **Two-Pair ->** $\binom{13}{2}\binom{4}{2}^2\binom{11}{1}\binom{4}{1} = 123{,}552$

Probability = 123,552/2,598,960 = **0.047539**

- **Three of a kind ->** $\binom{13}{1}\binom{4}{3}\binom{12}{2}\binom{4}{1}^2 = 54{,}912$

  Probability = 54,912/2,598,960 = **0.021128**

Monte Carlo iterations:

The accuracy of Monte Carlo simulations depends on the number of iterations utilized, which must be conducted in huge numbers. However, the number of iterations cannot be determined at random. As demonstrated in the preceding sections, the likelihood of the hands occurring is quite low, hence a high number of repetitions is anticipated. Additionally, we need a probability of precision greater than 0.00001 up to the fifth decimal point. The value of N must be carefully selected in order to obtain such precision.

Our main goal is to determine the value of N needed to calculate the likelihood of receiving four of a type with an error of less than 0.00001 and a 95% level of confidence. In a deck of 20 cards, the real likelihood of obtaining 4 of a kind is 0.00516. We employ the idea of confidence intervals and the central limit theorem to determine the value of N.

We have a chance of 0 or 1 for each iteration that indicates if a four-of-a-kind hand was created in that hand. The counter variable in the code receives a '1' for each good event that occurs.

As a result, the random variable X may be represented as:

$$X_i = \begin{cases} 1, if\ hand\ is\ made \\ 0, otherwise \end{cases}$$

As illustrated by X; above, the random variable used to choose the hands takes the form of Bernoulli trials, with our probability being either 1 or 0.

Since all of these random variables have the same iids and have a mean and standard deviation of o, we can express their sample mean as-

$$M_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

We use the following formula to obtain a probability of occurrence with an error of less than 0.00001 and with an accuracy of more than 95%:

$$P(|M_n - f| \geq \epsilon) \leq \alpha$$

Taking e=0.00001 and alpha= 0.05, we get formula

$$P(|M_n - f| \geq 0.00001) \leq 0.05$$

With the help of this formula, we are essentially attempting to calculate the likelihood of receiving a difference between Mn, or the probability we infer from our Monte Carlo simulations, and 'f,' or the actual analytical probability of the hand, which is less than 0.00001 with a probability of more than 95%.

The above equation will be used to determine the value of n using the Central Limit theorem. To achieve this, we change the form of this equation to that of the central limit theorem that we are aware of.

$$P(|M_n - f| \geq 0.00001) = \left|\frac{X_1 + X_2 + . . . + X_n}{n}\right| \geq 0.00001$$

We need to standardize above equation, to do so we multiply by $\sqrt{n}$ and divide by σ ,

So we get,

$$P(|M_n - f| \geq 0.00001) = \left|\frac{X_1 + X_2 + . . . + X_n}{\sqrt{n}\sigma}\right| \geq \frac{0.00001\sqrt{n}}{\sigma}$$

Due to the Central Limit theorem, which stipulates that if n is high enough, the standardized variable will converge into a Gaussian random variable, the aforementioned equation is now a standard random variable that is roughly Gaussian. As a result, it has a 0 mean and a 1 variance.

Thus, we may state that

$$P(|M_n - f| \geq 0.00001) \approx P(|Z| \geq \frac{0.00001\sqrt{n}}{\sigma})$$

AS X is a Bernoilli random variable, $\sigma = \sqrt{f * (1 - f)}$

For four of a kind, f = 0.00516. Hence, $\sigma = \sqrt{0.00516 * (1 - 0.00516)}$

$$\sigma = 0.07165$$

There for using this value in equation above

$$P(|M_n - f| \geq 0.00001) \approx P(|Z| \geq \frac{0.00001\sqrt{n}}{0.07165})$$

The right side can be denoted by variable $t = \frac{0.00001\sqrt{n}}{0.07165} = 1.3957 * 10^{-4}\sqrt{n}$

The above equation also written as,

$$P(|Z| \geq t) = 2(1 - \phi(y))$$

$$2\big(1 - \phi(t)\big) \leq 0.05$$

$$\phi(t) \geq 0.975$$

$t \geq 1.96$ Using Normal distribution tables

$$1.3957 * 10^{-4}\sqrt{n} \geq 1.96$$

$$n \geq 1.972 * 10^{8}$$

As we observe, N is very very large.

Calculating in similar way,

Number of iterations required for 95%

Hand	95%
Four of a kind	$1.972 * 10^{8}$
Three of a kind	$4.168 * 10^{9}$
Two Pairs	$7.72 * 10^{9}$
One Pair	$9.6 * 10^{9}$

Questions:

1. **Do your analytical probability calculations match the estimates your are getting from Monte Carlo simulation? Make sure that you make N large enough to get 5 digits correct after the decimal place. For each category, report the value of N you used to ensure that your estimate of the probability of each category is correct to 5 decimal places. Report also the estimated probability with Monte Carlo simulations and the actual probability computed analytically. What is the relationship between N and the accuracy of your estimates through Monte Carlo simulation?**

   The analytical probability calculations should match the estimates obtained through Monte Carlo simulations, although there may be some small differences due to the random nature of the simulation.

   In my simulation 4 out of 5 analytical probability calculations match the estimates I am getting from Monte Carlo simulation.

   Please find simulated output for each category with N value in section above.

- **Four of a kind ->**

  Value matched with N = 7000000

```
Monte Carlo simulation results for N = 7000000 hands
Four of a kind: 0.0051631
One pair: 0.49547
Two pair: 0.27836
Three of a kind: 0.12399
Analytical calculation results:
Four of a kind: 0.00516
One pair: 0.49536
Two pair: 0.27864
Three of a kind: 0.12384
```

- **One-Pair ->**

  Value matched With N = 70000000

```
Monte Carlo simulation results for N = 70000000 hands
Four of a kind: 0.0051689
One pair: 0.49536
Two pair: 0.27854
Three of a kind: 0.12385
Analytical calculation results:
Four of a kind: 0.00516
One pair: 0.49536
Two pair: 0.27864
Three of a kind: 0.12384
```

- **Two-Pair ->**

  Probability = 4320/15,504 = **0.2786377708978328**

```
Monte Carlo simulation results for N = 200000000 hands
Four of a kind: 0.0051666
One pair: 0.49538
Two pair: 0.27859
Three of a kind: 0.12384

Analytical calculation results:
Four of a kind: 0.00516
One pair: 0.49536
Two pair: 0.27864
Three of a kind: 0.12384
```

- **Three of a kind ->**

  Value matched With N = 200000000

  Probability = 1920/15,504 = **0.1238390092879257**

```
Monte Carlo simulation results for N = 200000000 hands
Four of a kind: 0.0051666
One pair: 0.49538
Two pair: 0.27859
Three of a kind: 0.12384

Analytical calculation results:
Four of a kind: 0.00516
One pair: 0.49536
Two pair: 0.27864
Three of a kind: 0.12384
```

The accuracy of the Monte Carlo simulation improves as the number of hands generated (N) increases.

To ensure that the estimates are correct to 5 decimal places, a large enough value of N must be used. In general, the required value of N depends on the expected probability of the event. For rare events, a larger value of N is needed to get accurate estimates.

2. How are the probabilities of the different categories different for the case where we have 20 cards, compared to the case where we have 52 cards?

According to the probabilities calculated for 52 in above section-

The probabilities of the different categories are different for the case where we have 20 cards compared to the case where we have 52 cards. In the case of 52 cards, there are more cards and more ranks, so the number of possible poker hands is larger. The probabilities of each category are also different because the number of cards and ranks affects the number of ways in which we can form the different poker hands. For example, in the case of 52 cards, the probability of getting a four of a kind is much smaller compared to the case of 20 cards, since there are more cards to choose from and the likelihood of getting 4 cards of the same rank is smaller. On the other hand, the probability of getting a two pair is larger in the case of 20 cards compared to the case of 52 cards, since there are fewer ranks to choose from and it is easier to get two pairs of the same rank.

MATLAB Program:

For MATLAB programming, I have tried to make it more and more efficient. Using it I was able to match 4 out of 5 values by N = 200000000 in just 45 minutes.

Some points that I took care of-

- Developed logic so I just needed 1-D array for Desk
- Designed conditions in such way that, decision tree should require minimal comparison
- Also arranged the IF ELSE condition in such way that the max probability conditions should be most outside to decrease number of comparisons
- Arranged IF-ELSE in such way that second else condition checks 2 underlying conditions

```
clear all; close all; clc;

% Define deck of cards
deck = [1,1,1,1,2,2,2,2,3,3,3,3,4,4,4,4,5,5,5,5];

% Define number of hands to generate
N = 400000;

% Initialize counts for each category of poker hand
dbhosal1_four_of_a_kind_count = 0;
dbhosal1_one_pair_count = 0;
dbhosal1_two_pair_count = 0;
dbhosal1_three_of_a_kind_count = 0;
tic
% Simulate N hands of poker
for n=1:N
    % Draw 5 cards randomly from the deck
    ranks_only = deck(randperm(20,5));

    dbhosal1_uniq_ranks_len = length(unique(ranks_only));
    [modeValue, modeCount] = mode(ranks_only);

    % One pair
    if dbhosal1_uniq_ranks_len == 4
        dbhosal1_one_pair_count = dbhosal1_one_pair_count + 1;
    % Two pair
    elseif dbhosal1_uniq_ranks_len == 3
        if modeCount == 2
            % Two pairs of ranks
            dbhosal1_two_pair_count = dbhosal1_two_pair_count + 1;
        elseif modeCount == 3
            % Three of a kind
            dbhosal1_three_of_a_kind_count = dbhosal1_three_of_a_kind_count + 1;
        end
    % Four of a kind
    elseif modeCount == 4
        % At least one pair
        dbhosal1_four_of_a_kind_count = dbhosal1_four_of_a_kind_count + 1;
    end
```

```
end
toc
% Compute probabilities using Monte Carlo simulation
four_of_a_kind_prob = dbhosal1_four_of_a_kind_count / N;
one_pair_prob = dbhosal1_one_pair_count / N;
two_pair_prob = dbhosal1_two_pair_count / N;
three_of_a_kind_prob = dbhosal1_three_of_a_kind_count / N;

% Display results
disp(['Monte Carlo simulation results for N = ' num2str(N) ' hands']);
disp(['Four of a kind: ' num2str(four_of_a_kind_prob)]);
disp(['One pair: ' num2str(one_pair_prob)]);
disp(['Two pair: ' num2str(two_pair_prob)]);
disp(['Three of a kind: ' num2str(three_of_a_kind_prob)]);

% Analytical calculations
actual_N = nchoosek(20,5);
% Four of a kind: Choose rank (5 options) and any of 4 suits (C,D,H,S)
four_of_a_kind_prob_actual = 5 * 4 * 4/ actual_N;
% One pair: Choose rank (5 options), choose 2 suits (C,D,H,S) for pair, choose 3
ranks for other cards (4 options), choose suit for each (4 options each)
one_pair_prob_actual = 5 * nchoosek(4,2) * nchoosek(4,3) * 4^3 / actual_N;
% Two pair: Choose 2 ranks (5 options), choose 1 suit (C,D,H,S) for each pair
% (4 options), choose rank for fifth card (3 options), choose suit for fifth card (4
options)
two_pair_prob_actual = nchoosek(5,2) * nchoosek(4,2)^2 * 3 * 4 / actual_N;
% Three of a kind: Choose rank (5 options), choose 3 suits (C,D,H,S) for that rank,
choose 2 ranks for other cards (4 options), choose suit for each (4 options each)
three_of_a_kind_prob_actual = 5 * nchoosek(4,3) * nchoosek(4,2) * 4^2 / actual_N;

% Display analytical results
disp(' ');
disp('Analytical calculation results:');
disp(['Four of a kind: ' num2str(four_of_a_kind_prob_actual)]);
disp(['One pair: ' num2str(one_pair_prob_actual)]);
disp(['Two pair: ' num2str(two_pair_prob_actual)]);
disp(['Three of a kind: ' num2str(three_of_a_kind_prob_actual)]);
```

Conclusion:

In this project, we have used Monte Carlo simulation and analytical calculations to estimate the probabilities of four categories of poker hands drawn from a deck of 20 cards. The results show that both methods provide similar probabilities. The Monte Carlo simulation is a powerful tool for estimating probabilities in complex systems, especially when analytical calculations are difficult or impossible to perform.