PROBABILITY AND RANDOM PROCESSES

PROJECT REPORT

Sai Charan Raghupatruni

—

1226175225

NAME : SAI CHARAN RAGHUPATRUNI
ASU ID : 1226175225

Table of Contents

1. PROBLEM STATEMENT

Make analytical analysis and write a program that calculates the probabilities of 5-card poker hands. Unlike regular poker, cards will be drawn from a deck of 20 = 4x5 cards including 4 suites of the first 5 ranks: 1(Ace), 2, 3, 4, 5. These hands whose probabilities you will calculate are

1. four of a kind
2. one-pair
3. two pair
4. 3-of-a-kind

2. INTRODUCTION

2.1. EVENT

In probability theory, an event is a subset of the sample space, which is the set of all possible outcomes of a random experiment.

An event is said to occur if the outcome of the experiment is one of the elements in the subset.

2.2. PROBABILITY

Probability is a way of quantifying uncertainty or the degree of belief in the occurrence of an event.

The probability of an event is expressed as a number between 0 and 1, with 0 indicating that the event is impossible and 1 indicating that the event is certain. The probability of an event can also be expressed as a percentage or a fraction.

$$\boxed{Probability = \frac{Number\ of\ favourable\ outcomes}{Total\ number\ of\ possible\ outcomes}}$$

2.3. COMBINATIONS

The number of ways of selecting "r" items from "n" items is given by

$$\boxed{{}^{n}_{r}C = \frac{n!}{(n-r)!\,r!}}$$

Where $m! = 1 * 2 * 3 * . . .* (m-1) * m$

2.4. HANDS IN POKER

Hands in poker consist of set of 5 cards in a hierarchical order. Here in the problem statement, we are given the following hands.

A. ***Probability of Four of a kind***
In a hand resulting a “four of a kind” hand consist of 4 cards of same number, different suites and one random card.

B. ***Probability of Three of a kind***
In a hand resulting a “three of a kind” hand consist of 3 cards of same number, different suites and two random non pair cards.

C. ***Probability of two pair***
In a hand resulting a “two pair” hand consist of 2 pairs of cards with the same number and a card not of same number as the pairs.

D. ***Probability of one pair***
In a hand resulting a “one pair” hand consist of 1 pair of cards with the same number and a card not of same number as the pair or non-pair cards

2.5. MONTE CARLO SIMULATION

Monte Carlo simulation is a computational technique used to estimate the probability of different outcomes in a system by running multiple random simulations.

In this technique, a model of a complex system is created using a set of mathematical equations, probability distributions, and other relevant factors. The inputs to the model are randomly generated, and the model is then run multiple times, producing a distribution of possible outcomes. By analysing this distribution, one can estimate the probability of different outcomes occurring.

2.6. CHEBYSHEV INEQUALITY

the Chebyshev inequality provides an upper bound on the probability that a random variable deviate from its mean by more than k standard deviations, regardless of the shape of the distribution. This is a very useful result, as it allows us to make statements about the likelihood of extreme values in a data set without making any assumptions about the distribution of the data.

Let X be a random variable with finite mean μ and finite variance σ^2. Then, for any positive number k:

$$\boxed{P(|X-\mu| \geq\ c) \leq \frac{var\ (X)}{k^2}}$$

3. ANALYTICAL CALCULATIONS

Given a deck of 20 = 4x5 cards including 4 suites of the first 5 ranks: 1(Ace), 2, 3, 4, 5
Therefore,

$$\boxed{\mathrm{N} = 20}$$

There are ${}^{\mathrm{N}}_{5}\mathrm{C}$ ways to pick 5 cards from a deck of 20 cards.

$$\boxed{{}^{\mathrm{N}}_{5}\mathrm{C} = 15504}$$

A. Probability of Four of a kind

The total number of ways this deck can contain “four of a kind” is.

$$ {}^{5}_{1}C * {}^{4}_{4}C * {}^{4}_{1}C * {}^{4}_{1}C = 80 $$

There are ${}^{5}_{1}C$ ways of selecting one of the 5 numbers and ${}^{4}_{4}C$ ways of selecting cards for four of a kind hand. Remaining card can be from remaining 4 numbers ${}^{4}_{1}C$ .In that, we can select ${}^{4}_{1}C$ ways.

$$ Probability\ of\ Four\ of\ a\ kind = \frac{\text{total number of ways this deck can contain “four of a kind”}}{\text{number of ways to pick 5 cards from a deck of 20 cards}} = \frac{80}{15504} = 0.0051599587 $$

B. Probability of Three of a kind

The total number of ways this deck can contain “three of a kind” is.

$$ {}^{5}_{1}C * {}^{4}_{3}C * {}^{4}_{2}C * {}^{4}_{1}C * {}^{4}_{1}C = 1920 $$

There are ${}^{5}_{1}C$ ways of selecting one of the 5 numbers and ${}^{4}_{3}C$ ways of selecting cards for three of a kind hand. Remaining cards can be from remaining 4 numbers ${}^{4}_{2}C$ .In that, we can select ${}^{4}_{1}C$ ways each remaining card.

$$ Probability\ of\ Three\ of\ a\ kind = \frac{\text{total number of ways this deck can contain “Three of a kind”}}{\text{number of ways to pick 5 cards from a deck of 20 cards}} = \frac{1920}{15504} = 0.1238390093 $$

C. Probability of two pair

The total number of ways this deck can contain “two pair” are.

$$ {}^{5}_{2}C * {}^{4}_{2}C * {}^{4}_{2}C * {}^{3}_{1}C * {}^{4}_{1}C = 4320 $$

There are ${}^{5}_{2}C$ ways of selecting two of the 5 numbers and ${}^{4}_{2}C$*${}^{4}_{2}C$ ways of selecting cards for two pair hand. Remaining cards can be from remaining 3 numbers ${}^{3}_{1}C$ .In that, we can select ${}^{4}_{1}C$ ways for the remaining card.

$$ Probability\ of\ two\ pair = \frac{\text{total number of ways this deck can contain “two pair”}}{\text{number of ways to pick 5 cards from a deck of 20 cards}} = \frac{4320}{15504} = 0.2786377709 $$

D. Probability of one pair

In a hand resulting a “one pair” hand consist of 1 pair of cards with the same number and a card not of same number as the pair or non-pair cards

The total number of ways this deck can contain “one pair” are.

$$ {}^{5}_{1}C * {}^{4}_{2}C * {}^{4}_{3}C * {}^{4}_{1}C * {}^{4}_{1}C * {}^{4}_{1}C = 7680 $$

There are ${}^{5}_{1}C$ ways of selecting one of the 5 numbers and ${}^{4}_{2}C$ ways of selecting cards for two pair hand. Remaining cards can be from remaining 4 numbers ${}^{4}_{3}C$ .In that, we can select. ${}^{4}_{1}C * {}^{4}_{1}C * {}^{4}_{1}C$ ways for the remaining cards.

$$ Probability\ of\ one\ pair = \frac{\text{total number of ways this deck can contain “one pair”}}{\text{number of ways to pick 5 cards from a deck of 20 cards}} = \frac{7680}{15504} = 0.4953560372 $$

4. ALGORITHM FOR MONTE CARLO SIMULATION

i. Initialize an array “cards” consisting of numbers from 1 to 19. Using the remainder function get the remainders of each value in “cards” when divided by 5. Now put these remainder values in “rank” array.

ii. This “rank” array is equivalent to a sorted 20-card deck. Now, initialize various counter variables to count the occurrence of hands.

iii. Say, N iterations must be performed, a loop is initiated. In which for every iteration, a shuffled deck is generated, and the first five cards are selected. This is basically like picking 5 random cards from a 20-card deck.

iv. The selected hand is checked for an outcome. The hand is tested for number of unique elements in the array and the maximum frequency of occurrence of each element in that specific iteration. The hand is determined based on the test results as follows.

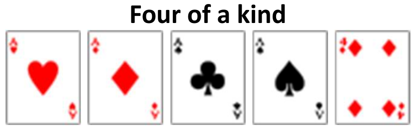
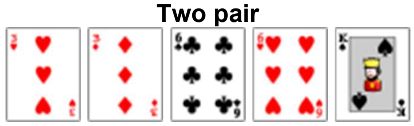
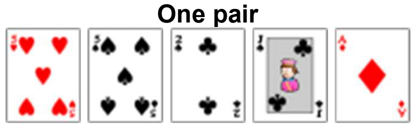
Hand	Number of unique elements (m)	maximum frequency of occurrence (max_freq)
Four of a kind 	2	4
Three of a kind 	3	3
Two pair 	3	2
One pair 	4	2

Table 4.1. Unique number of cards and maximum frequency of occurrence for given hands

v. The values of the counters are updated according to the response from the tests.

vi. Finally, the counters are divided by total iterations to get the probabilities of each hand.

5. SIMULATION RESULTS

The algorithm is implemented using MATLAB program. 10^{10} iterations are performed. The computed outputs are displayed on the command window screen as follows. The program took **8 hours 4 minutes and 22 seconds** to complete simulation.

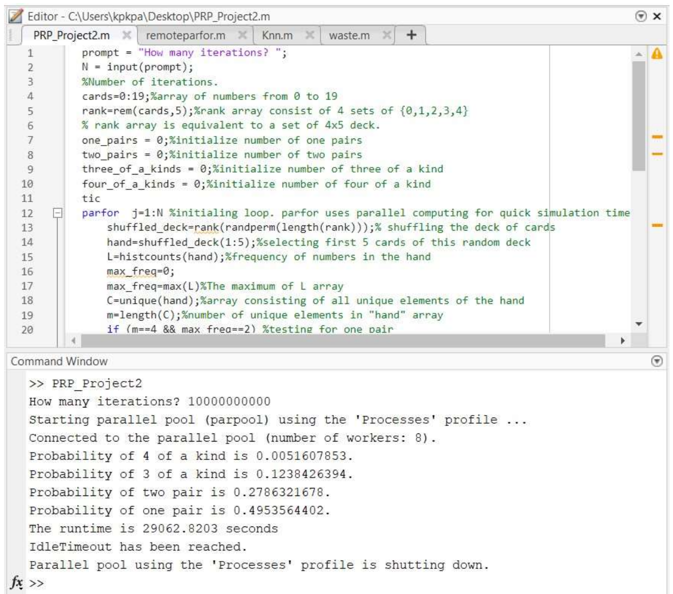


Figure 5.1. Output of Simulation

6. COMPARISION OF ANALYTICAL RESULTS AND COMPUTED RESULTS

Q. Do analytical probability calculations match the estimates from the Monte Carlo simulation?

The accuracy of the simulation increases monotonically with the increase in the number of iterations.

Here, N iterations is adjusted in such a way that the computed probabilities are accurate up to 5 decimal places ($\in \leq 0.000001$).

Hand	Analytical Probability of hand	Computed Probability of hand	Difference of the Probabilities
Four of a kind	0.0051599587	0.0051607853	0.0000008266
Three of a kind	0.1238390093	0.1238426394	0.0000036301
Two pair	0.2786377709	0.2786321678	0.0000056031
One pair	0.4953560372	0.4953564402	0.000000403

Table 6.1. Comparison of analytical results and computed results

Q. 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?

For the 20-card deck, there are 4 suits (hearts, diamonds, clubs, and spades) with cards numbered from Ace to 5. Whereas, in 52-card deck, there are 4 suits with cards numbered from Ace to King.

Since in a 20-card deck, only fewer cards are available, there are only 15,504 possible shuffles. In a 52-deck card there are ${}^{52}_{5}C$ (= 2598960) possible combinations. The denominator while calculating probabilities consist of number of possible hands in the deck. Therefore, the probabilities of the given hands are more in the 20-card deck.

The table below lists the probabilities of various hands in 20-card deck as well as 52 card deck.

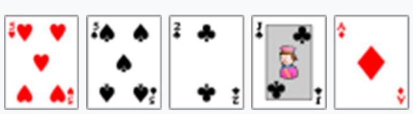
Hand	Expression for Probability of given hand	Probabilities for 52-Card Deck (N=52)	Probabilities for 20-Card Deck (N=20)
Four of a kind 	$\frac{{}^{\frac{N}{4}}_{1}C * {}^{4}_{4}C * {}^{\frac{N}{4}-1}_{1}C * {}^{4}_{1}C}{{}^{N}_{5}C}$	0.00240	0.00515
Three of a kind 	$\frac{{}^{\frac{N}{4}}_{1}C * {}^{4}_{3}C * {}^{\frac{N}{4}-1}_{2}C * {}^{4}_{1}C * {}^{4}_{1}C}{{}^{N}_{5}C}$	0.02113	0.12383
Two pair 	$\frac{{}^{\frac{N}{4}}_{2}C * {}^{4}_{2}C * {}^{4}_{2}C * {}^{\frac{N}{4}-2}_{1}C * {}^{4}_{1}C}{{}^{N}_{5}C}$	0.04754	0.27860
One pair 	$\frac{{}^{\frac{N}{4}}_{1}C * {}^{4}_{2}C * {}^{\frac{N}{4}-1}_{3}C * {}^{4}_{1}C * {}^{4}_{1}C * {}^{4}_{1}C}{{}^{N}_{5}C}$	0.42257	0.49535

Table 6.2. Probability comparisons for 52-card and 20-card decks

Q. What is the relationship between N and the accuracy of your estimates through Monte Carlo simulation?

The relationship between the number of iterations and the accuracy of estimates through Monte Carlo simulation is a positive correlation. As the number of iterations increases, the accuracy of estimates tends to improve.

This is because Monte Carlo simulation relies on generating a large number of random samples to approximate a numerical result. The more samples are generated, the more precise the estimate becomes. This is because the "law of large numbers" states that as the sample size increases, the sample mean approaches the true population mean.

the Chebyshev inequality can be used to explain the relationship between the number of iterations and the accuracy of estimates in Monte Carlo simulation.

we want to estimate the expected value of a random variable X using Monte Carlo simulation, and we have generated n independent samples $X_1, X_2, X_3, \ldots, X_n$ from the distribution of X. Let M_u be the true expected value of X, and let sigma be the standard deviation of X. Then, the sample mean, and standard deviation of samples are as follows.

$$M_n = \frac{X_1 + X_2 + X_3, \ldots, + X_n}{n}$$

$$\sigma_{samples} = \frac{\sigma_X}{\sqrt{n}}$$

Since the process here is a Bernoulli process,

$$\sigma_X{}^2 = M_u * (1 - M_u)$$
$$therfore\ the\ maximum\ of\ \sigma_X = 0.25$$

From Chebyshev inequality, we have.

$$P(|M_n - M_u| \geq k) \leq \frac{\sigma_{samples}{}^2}{k^2}$$

Goal is to find the number of iterations n such that $|M_n - M_u| < 0.00001$ with 75% confidence.
Comparing the goal with the above equation.
We have,

$$k = 0.00001$$

$$\frac{\sigma_{samples}{}^2}{k^2} \leq 0.25 \implies \frac{\sigma_X{}^2}{n * k^2} \leq 0.25$$

$$the\ maximum\ of\ \sigma_X{}^2\ is\ 0.25. \quad Therfore$$

$$\boxed{\frac{0.25}{n * (0.00001)^2} \leq 0.25 \implies n = 10000000000}$$

We need 10^{10} iterations to get an error less than 0.00001 with 75% confidence.

7. CONCLUSION

The probabilities of the given hands (four of a kind, three of a kind, two pair, one pair) are analytically calculated and the Monte Carlo simulation is performed for 10^{10} iterations to obtain the hands probabilities. The required accuracy $\in \leq 0.000001$ is achieved.

REFERENCES

1. Steven M Kay, (2012) Intuitive Probability and Random Processes using MATLAB.
2. Wikipedia, Poker Probability. Available at https://en.wikipedia.org/wiki/Poker_probability
3. MathWorks, Parallel Computing Toolbox. Available at https://www.mathworks.com/products/parallel-computing.html

APPENDIX

CODE:

```
prompt = "How many iterations? ";
N = input(prompt)
%Number of iterations.
cards=0:19;%array of numbers from 0 to 19
rank=rem(cards,5);%rank array consist of 4 sets of {0,1,2,3,4}
% rank array is equivalent to a set of 4x5 deck.
one_pairs = 0;%initialize number of one pairs
two_pairs = 0;%initialize number of two pairs
three_of_a_kinds = 0;%initialize number of three of a kind
four_of_a_kinds = 0;%initialize number of four of a kind
tic
parfor  j=1:1:N %initialing loop. parfor uses parallel computing for quick
simulation time.
    shuffled_deck=rank(randperm(length(rank)));% shuffling the deck of cards
    hand=shuffled_deck(1:5);%selecting first 5 cards of this random deck
    L=histcounts(hand);%frequency of numbers in the hand
    max_freq=0;
    max_freq=max(L)%The maximum of L array
    C=unique(hand);%array consisting of all unique elements of the hand
    m=length(C);%number of unique elements in "hand" array
    if (m==4 && max_freq==2) %testing for one pair
one_pairs = one_pairs+1;
continue
    end
    if(m==3 && max_freq==2)%testing for two pair
      two_pairs = two_pairs+1;
   continue
    end
     if(max_freq==3 && m==3)%testing for Three of a kind
      three_of_a_kinds = three_of_a_kinds+1;
   continue
   end
    if (max_freq==4)%testing for Four of a kind
       four_of_a_kinds =four_of_a_kinds+1;
       continue
    end
end
X = ['the runtime is',num2str(toc),' seconds']
Prob_4_kind= four_of_a_kinds/N;
fprintf('Probability of 4 of a kind is %f.\n',Prob_4_kind);
Prob_3_kind= three_of_a_kinds/N;
fprintf('Probability of 3 of a kind is %f.\n',Prob_3_kind);
Prob_2pair= two_pairs/N;
fprintf('Probability of two pair is %f.\n',Prob_2pair);
Prob_1pair= one_pairs/N;
fprintf('Probability of one pair is %f.\n',Prob_1pair);
disp(X)
```