EE511 - F17 (Silvester)

Projects 2 and 3 Monte Carlo Methods

Ruirui Huang

1. Estimate π by the area method and also find including confidence intervals on your estimate. Draw a graph of the successive values of the estimator as the number of samples increases. How many points do you need to use for your estimate to be within 1% of the true value of π (with probability 0.95)?

Source Code:

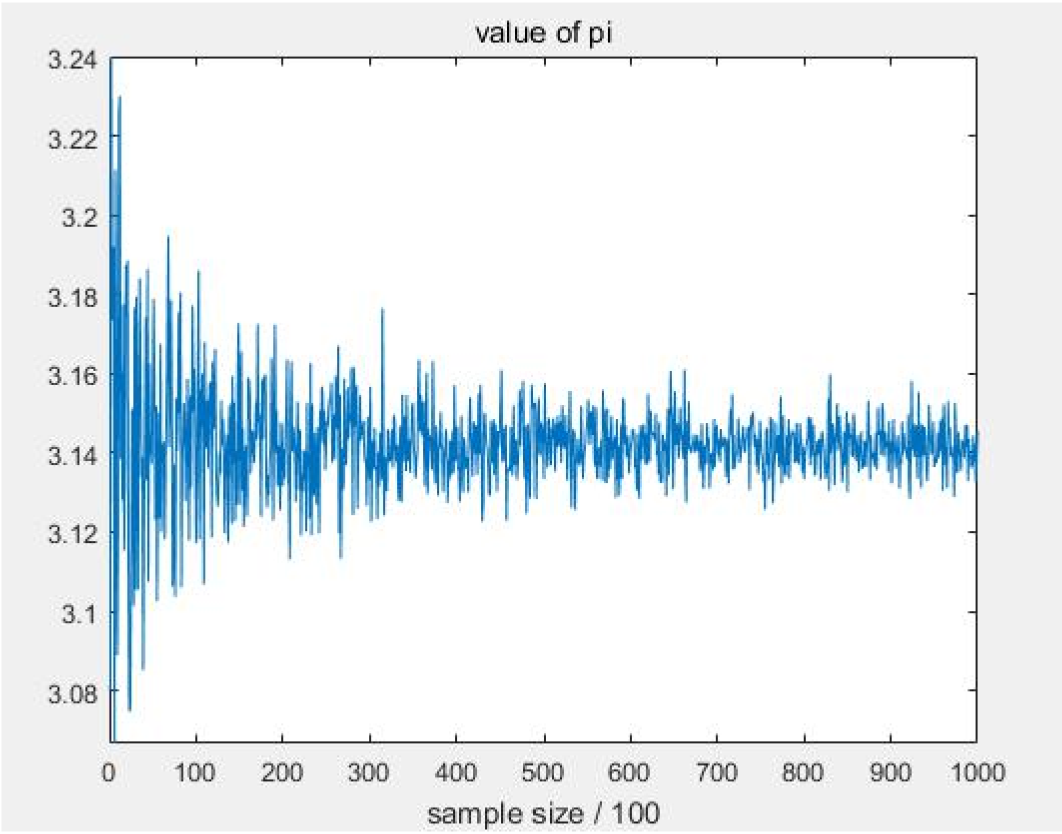
```
pi = zeros(1, 1000);                % 1000 different sample sizes
dots = 100;
index = 1; isExist = 0; n = 0;
while dots <= 100000
    p = zeros(1, dots);
    count = 0;
    for i = 1 : dots                % For each sample size, estimate pi
       x = rand(); y = rand();
       if(x^2 + y^2 <= 1)
        % To judge whether the point is inside or outside of the quarant
          p(i) = 1;
          count = count + 1;
       else
           p(i) = 0;
       end
    end
    pi(index) = 4 * mean(p);       % To calculate pi
    if(isExist == 0 && (1.96 * sqrt(var(p)) / sqrt(dots) <= 0.7854 * 0.01))
    % To be within 1% of the true value of pi(with probability 0.95)
        n = dots
        isExist = 1;
    end
    dots = dots + 100;         % Each time the sample size increase 100 dots
    index = index + 1;
end
pi(1000)
plot(pi); title("value of pi"); xlabel("sample size / 100"); axis([0 1000 min(pi)
max(pi)]);
```

Experimental Results:

When the sample size is equal to 100000 dots, the estimated π is as below:

```
ans =

    3.1456
```



The dots needed to get my estimate be within 1% of the true value of π (with probability 0.95) is as below:

```
n =

       10300
```

Analysis:

Generate random variables x, y in (0, 1) as each possible dot's coordinate. Let P_i represent the outcome of each dot, with 1 if (x, y) is in the quadrant and 0 if not. Thus the mean value of P_i ($\hat{p}$) could be the estimate of the real probability $p = \frac{\pi}{4}$, and then $\pi = 4 * \hat{p}$.

From the graph above, we can see with the increase of sample size, the estimated value of π is becoming more and more accurate, and the range of the estimated value's deviation from the true value is becoming smaller and smaller.

It is common to assume that the distribution of $\hat{p}$ is Gaussian distribution with mean p and variance $s_{\hat{p}}^2$ - which is asymptotically valid for large n (Central Limit Theorem). So we can find a confidence interval where we expect $\hat{p}$ to fall $\Pr\left\{p - \beta s_{\hat{p}} \le \hat{p} \le p + \beta s_{\hat{p}}\right\} = 1-\alpha$, where $s_{\hat{p}}^2 = VAR(\frac{\sum_{i=1}^{n} P_i}{n-1}) = \frac{1}{n^2}\sum_{i=1}^{n} s_{P_i}^2 = \frac{1}{n}s_{P_i}^2$, namely $\Pr\left\{0.7854 - \frac{s_{P_i}\beta}{\sqrt{n}} \le \hat{p} \le 0.7854 + \frac{s_{P_i}\beta}{\sqrt{n}}\right\} = 1-\alpha$, where $s_{P_i}^2 = \frac{\sum_{i=1}^{n}(P_i - \hat{p})^2}{n-1}$ is the sample variance of P_i and β is selected so that the probability $1-\alpha$ meets a desired level of confidence.

For $1-\alpha = 0.95$, according to the standard table of Normal Distribution, we can get $\beta = 1.96$. And the estimate should within 1%, we can get $0.7854 + \frac{s_{P_i}\beta}{\sqrt{n}} = 1\% * 0.7854$. Then we can calculate the dots needed. Via MATLAB, we get when n is larger than 10300 (the minimal interval of my experiment is 100), my estimate can be within 1% of the true value of π (with probability 0.95).

2. Evaluate the following integral by a Monte Carlo approach: $\int_{(n-1)\pi}^{n\pi} \frac{\sin(x)}{x} dx$

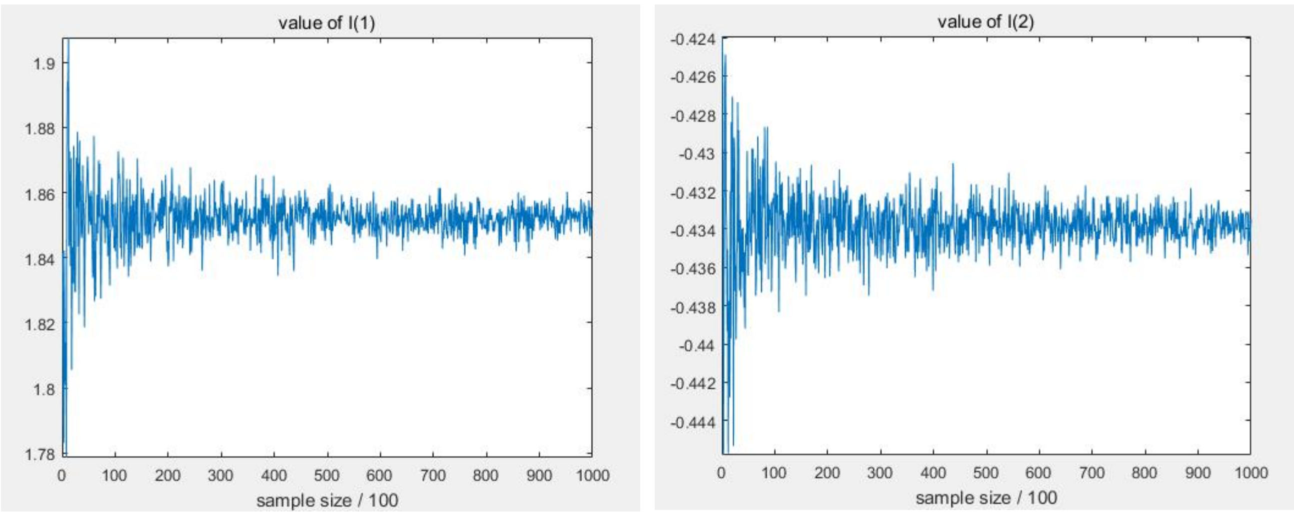
For n = 1,2,3,4,5 (You can reuse the random samples x (shifted by π for each consecutive interval) and values for sin(x) (which only change in sign from $n \to n+1$**) that you found for the case of n = 1 , for the latter integrals, due to the nature of the sin function.)**

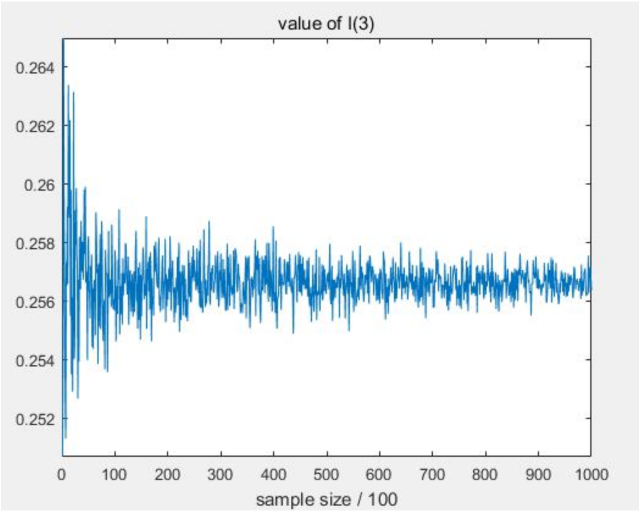
Source Code:

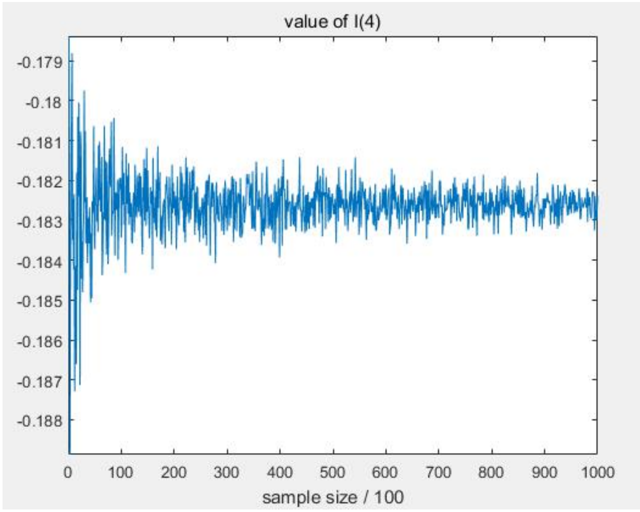
```
dots = 100;
count = 1;
E1 = zeros(1, 1000); E2 = zeros(1, 1000); E3 = zeros(1, 1000);
% 1000 different sample sizes
E4 = zeros(1, 1000); E5 = zeros(1, 1000);
while dots <= 100000
    x = rand(1, dots);
    y = sin(pi * x);
```

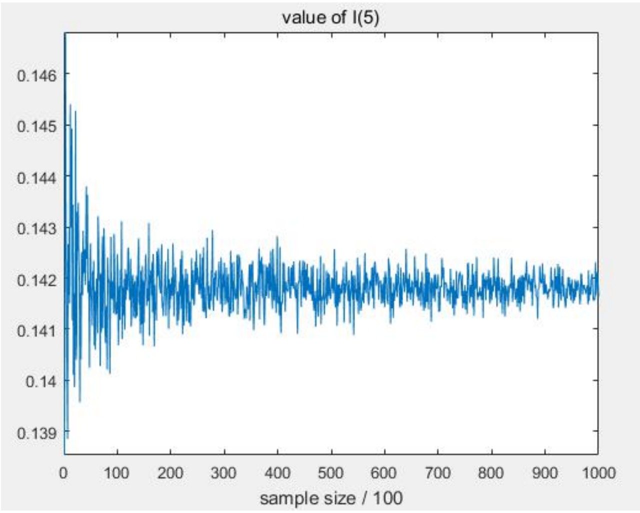
```
    p1 = zeros(1, dots); p2 = zeros(1, dots); p3 = zeros(1, dots);
    p4 = zeros(1, dots); p5 = zeros(1, dots);
    for i = 1 : dots
    % For each sample size, calculate corresponding formula
        p1(i) = y(i) / x(i); p2(i) = - y(i) / (x(i)+1); p3(i) = y(i) / (x(i)+2);
        p4(i) = - y(i) / (x(i)+3); p5(i) = y(i) / (x(i)+4);
    end
    E1(count) = mean(p1); E2(count) = mean(p2); E3(count) = mean(p3);
    E4(count) = mean(p4); E5(count) = mean(p5);
    count = count + 1;
    dots = dots + 100;
end
 figure(1); plot(E1); title("value of I(1)"); xlabel("sample size / 100"); axis([0
1000 min(E1) max(E1)]);
 hold on; figure(2); plot(E2); title("value of I(2)"); xlabel("sample size / 100");
axis([0 1000 min(E2) max(E2)]);
 hold on; figure(3); plot(E3); title("value of I(3)"); xlabel("sample size / 100");
axis([0 1000 min(E3) max(E3)]);
 hold on; figure(4); plot(E4); title("value of I(4)"); xlabel("sample size / 100");
axis([0 1000 min(E4) max(E4)]);
 hold on; figure(5); plot(E5); title("value of I(5)"); xlabel("sample size / 100");
axis([0 1000 min(E5) max(E5)]);
```

Experimental Results:









Analysis:

Supposing we wish to evaluate $I = \int_0^1 g(x)dx$. If the random variable U is Uniform(0, 1), then $E(g(u)) = \int_0^1 g(u) * 1du = I$. Thus to evaluate $\int_{(n-1)\pi}^{n\pi} \frac{\sin(x)}{x} dx$, we can change it to $\int_0^1 \frac{\sin \pi(y+n-1)}{y+n-1} dy$ and y is Uniform(0, 1). According to the nature of sin function $\sin(\alpha+\pi) = -\sin(\alpha)$, $\sin \pi(y+n-1) = (-1)^{n-1} \sin(\pi y)$, thus we could simplify the evaluation.

Based on this approach, evaluate the following integral $D(n) = \int_0^{n\pi} \frac{\sin(x)}{x} dx$.

For n = 10,100,1000. What do you observe? Can you speculate on the value of

$\int_0^{\infty} \frac{\sin(x)}{x} dx$ **(this is called the Dirichlet Integral)**

Since $D(n) = \int_0^{n\pi} \frac{\sin(x)}{x} dx = \int_0^{\pi} \frac{\sin(x)}{x} dx + \int_{\pi}^{2\pi} \frac{\sin(x)}{x} dx + \int_{2\pi}^{3\pi} \frac{\sin(x)}{x} dx + \ldots + \int_{(n-1)\pi}^{n\pi} \frac{\sin(x)}{x} dx$, we can add the result of all periods' results. Because of the limitation of computer, only 1000 samples are used to evaluate each period's result.

Source Code and Experimental Results:

```
x = rand(1, 1000);
y = sin(pi * x);
sum = 0;
for n = 1 : 101001000
    p = zeros(1, 1000);
    for i = 1 : 1000
        p(i) = (-1)^(n-1) * (y(i) / (x(i)+(n-1)));
    end
    sum = sum + mean(p);    % sum of all periods of Integration
end
mean(p)
sum
```

```
ans =                sum =

  -6.3623e-08          1.5719
```

When n is large enough, $\int_{(n-1)\pi}^{n\pi} \frac{\sin(x)}{x} dx$ is close to 0. Thus the sum of all periods' integration is close to a constant value.

From the result above, we can speculate that when n is close to infinity, the value of the Dirichlet Integral is π/2.

3. Find (by a Monte Carlo approach) the probabilities of the different possible poker hands (see lecture notes).

Source Code

```
deck = zeros(1, 52);
for i = 1: 52
    deck(i) = i;
end
straightFlush = 0; fourofaKind = 0; fullHouse = 0; flush = 0; straight = 0;
threeofaKind = 0; twoPairs = 0; onePair = 0; highCard = 0;
color = zeros(1, 5); number = zeros(1, 5);
for times = 1 : 5000000
    for card = 1 : 51       %random shuffle of the deck
       index = floor((52-card) * rand());
       temp = deck(card); deck(card) = deck(index+card); deck(index+card) = temp;
    end
    for k = 1 : 5                    % get the first 5 of the deck as the possible hand
        color(k) = ceil(deck(k) / 13);          % get the color of the cards
        number(k) = mod(deck(k)-1, 13) + 1;      % get the number of the cards
    end
    handsColor = sort(color); handsNumber = sort(number);
    if((handsColor(5) == handsColor(1)) && (handsNumber(5) - handsNumber(1) == 4))
        straightFlush = straightFlush + 1;
    elseif((handsNumber(4) == handsNumber(1)) || (handsNumber(5) == handsNumber(2)))
        fourofaKind = fourofaKind + 1;
    elseif((handsNumber(1) == handsNumber(3) && handsNumber(4) == handsNumber(5)) ||
             (handsNumber(1) == handsNumber(2) && handsNumber(3) == handsNumber(5)))
        fullHouse = fullHouse + 1;
    elseif(handsColor(5) == handsColor(1))
        flush = flush + 1;
    elseif(handsNumber(5) - handsNumber(1) == 4)
        straight = straight + 1;
    elseif((handsNumber(1) == handsNumber(3)) || (handsNumber(2) == handsNumber(4))
            || (handsNumber(3) == handsNumber(5)))
        threeofaKind = threeofaKind + 1;
    elseif((handsNumber(1) == handsNumber(2) && handsNumber(3) == handsNumber(4)) ||
            (handsNumber(1) == handsNumber(2) && handsNumber(4) == handsNumber(5)) ||
            (handsNumber(2) == handsNumber(3)  && handsNumber(4) == handsNumber(5)))
        twoPairs = twoPairs + 1;
    elseif((handsNumber(1) == handsNumber(2)) || (handsNumber(2) == handsNumber(3))
        || (handsNumber(3) == handsNumber(4)) || (handsNumber(4) == handsNumber(5)))
        onePair = onePair + 1;
    else
        highCard = highCard + 1; end
```

```
end
p1 = straightFlush / 5000000
p2 = fourofaKind / 5000000
p3 = fullHouse / 5000000
p4 = flush / 5000000
p5 = straight / 5000000
p6 = threeofaKind / 5000000
p7 = twoPairs / 5000000
p8 = onePair / 5000000
p9 = highCard / 5000000
p = p1 + p2 + p3 + p4 + p5 + p6 + p7 + p8 + p9      % the total possibility is 1
```

Experimental Results:

```
p1 =

   1.5400e-05


p2 =

   2.3220e-04


p3 =

    0.0015


p4 =

    0.0020


p5 =

    0.0277


p6 =

    0.0190


p7 =

    0.0428


p8 =

    0.4057


p9 =

    0.5011


p =

     1
```

Analysis:

According to lecture note, by changing the certain card of the sequence and the randomly picked card and repeating for 51 times, with the certain card plus one index each time, we can get the cards shuffled. We can use the first 5 cards as possible hands since they are chosen randomly. For each kind of hand, there are different restriction on the colors and numbers of the cards. The more strict restriction, the lower possibility of the certain hand showing up.