PHYS 516- METHODS OF COMPUTATIONAL PHYSICS
ASSIGNMENT 2- MONTE CARLO BASICS

Anup V Kanale
February 1, 2017

1 Programming: Testing the Central Theorem of Monte Carlo Estimate

In this assignment, the dependence of Monte Carlo (MC) error is numerically tested on the sample size M.

$$Std\{\overline{f(x)}\} = \frac{Std\{f(x)\}}{\sqrt{M}} \tag{1}$$

Mean integral of π is used as the sample.

$$\frac{1}{M}\sum_{i=1}^{M}\frac{4}{1+r_n^2} = \overline{\frac{4}{1+r_n^2}} \approx \pi(r_n \in [0,1]) \tag{2}$$

The plot below shows the estimate of π as a function of the number of Monte Carlo trials.

1.1 Monte Carlo Estimate

The MC estimate of π is calculated, along with the error, using different number of trials and the results are shown in the plot below. The unbiased estimate of the standard deviation is given by

$$SD = \sqrt{\frac{\overline{f^2} - (\overline{f})^2}{M-1}} \tag{3}$$

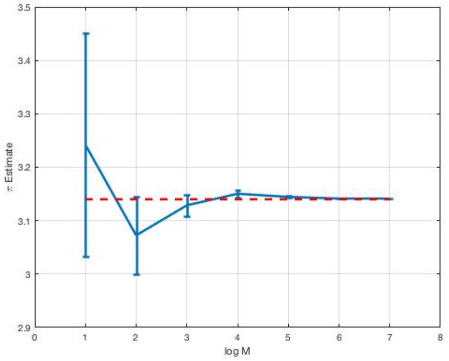
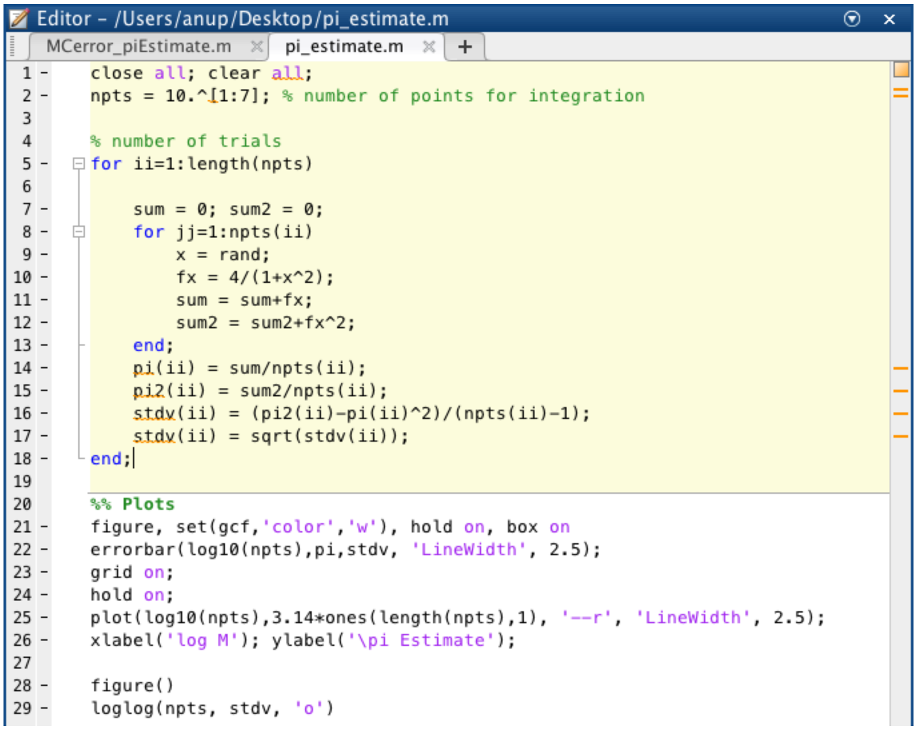


Figure 1: Monte Carlo Estimation of Pi

It is easily seen that the error decreases with increasing number of trials, and that beyond 10^5 trials, increasing M has negligible effect on the error.



```
close all; clear all;
npts = 10.^[1:7]; % number of points for integration

% number of trials
for ii=1:length(npts)

    sum = 0; sum2 = 0;
    for jj=1:npts(ii)
        x = rand;
        fx = 4/(1+x^2);
        sum = sum+fx;
        sum2 = sum2+fx^2;
    end;
    pi(ii) = sum/npts(ii);
    pi2(ii) = sum2/npts(ii);
    stdv(ii) = (pi2(ii)-pi(ii)^2)/(npts(ii)-1);
    stdv(ii) = sqrt(stdv(ii));
end;

%% Plots
figure, set(gcf,'color','w'), hold on, box on
errorbar(log10(npts),pi,stdv, 'LineWidth', 2.5);
grid on;
hold on;
plot(log10(npts),3.14*ones(length(npts),1), '--r', 'LineWidth', 2.5);
xlabel('log M'); ylabel('\pi Estimate');

figure()
loglog(npts, stdv, 'o')
```

Figure 2: Source Code

1.2 Monte Carlo Error

We next perform a numerical experiment to directly measure the standard deviation of the MC estimate. To do so, for each of the above M values, π is estimated N_{seed} times using N_{seed} different random-number seeds. The standard deviation σ_M of these N_{seed} estimates $\pi_1, \pi_2, \ldots, \pi_{N_{seed}}$

$$\sigma_M = \sqrt{\frac{1}{N_{seed}} \sum_{i=1}^{N_{seed}} \pi_i^2 - \left[\frac{1}{N_{seed}} \sum_{i=1}^{N_{seed}} \pi_i\right]^2} \tag{4}$$

The measured values of $\log_{10} \sigma_M$ are shown as a function of $\log_{10} M$ in the plot below. If the MC error decreases as $\sigma_M = C/\sqrt{M}$, then

$$log_{10}\sigma_M = log_{10}C - \frac{1}{2}log_{10}M$$

Using a least square fit, the equation of the line obtained is

$$log_{10}\sigma_M = -0.1653 - 0.505log_{10}M$$

We see that the unbiased estimates from the previous question neatly fall onto this fitted line.

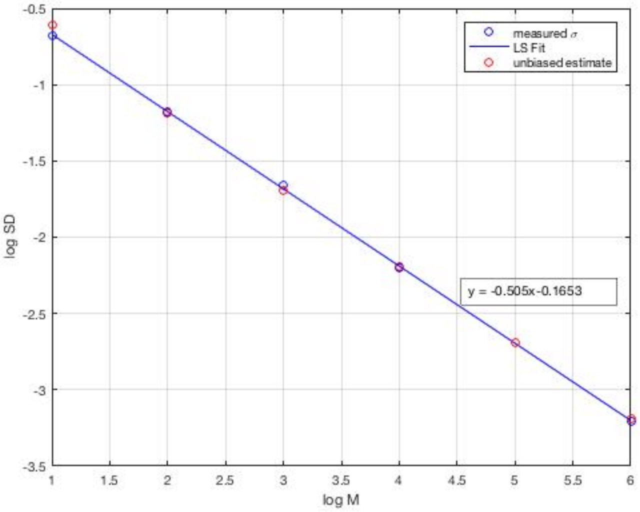
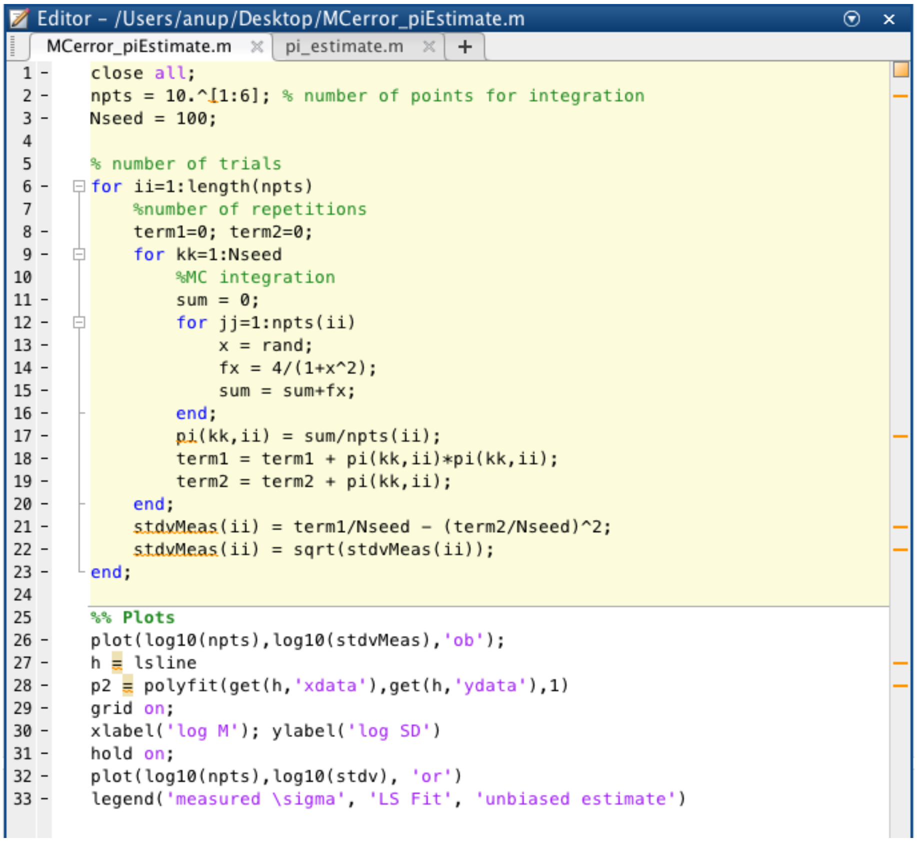


Figure 3: Unbiased Estimates and Measured values of Standard Deviation



```
close all;
npts = 10.^[1:6]; % number of points for integration
Nseed = 100;

% number of trials
for ii=1:length(npts)
    %number of repetitions
    term1=0; term2=0;
    for kk=1:Nseed
        %MC integration
        sum = 0;
        for jj=1:npts(ii)
            x = rand;
            fx = 4/(1+x^2);
            sum = sum+fx;
        end;
        pi(kk,ii) = sum/npts(ii);
        term1 = term1 + pi(kk,ii)*pi(kk,ii);
        term2 = term2 + pi(kk,ii);
    end;
    stdvMeas(ii) = term1/Nseed - (term2/Nseed)^2;
    stdvMeas(ii) = sqrt(stdvMeas(ii));
end;

%% Plots
plot(log10(npts),log10(stdvMeas),'ob');
h = lsline
p2 = polyfit(get(h,'xdata'),get(h,'ydata'),1)
grid on;
xlabel('log M'); ylabel('log SD')
hold on;
plot(log10(npts),log10(stdv), 'or')
legend('measured \sigma', 'LS Fit', 'unbiased estimate')
```

Figure 4: Source Code

2 Non-uniform random number generation: Box-Muller Transformation

A Box-Muller algorithm generates a normally distributed random number. This algorithm is similar in principle to the Greenland Phenomenon.

First a set of independent uniform random numbers r_1 and r_2 are generated in the range $(0, 1)$. The following transformations are then applied to get two new independent random numbers.

$$\zeta_1 = \sqrt{-2\ln r_1}\cos(2\pi r_2) \tag{5}$$

$$\zeta_2 = \sqrt{-2\ln r_1}\sin(2\pi r_2) \tag{6}$$

Now let us try to express r_1 and r_2 in terms of ζ_1 and ζ_2. Summing and squaring Eq(5) and (6),

$$\therefore \zeta_1^2 + \zeta_2^2 = -2\ln r_1 \tag{7}$$

$$\implies r_1 = \exp\left\{-\frac{\zeta_1^2 + \zeta_2^2}{2}\right\} \tag{8}$$

Similarly, dividing Eq.(6) by Eq.(5),

$$r_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{\zeta_2}{\zeta_1}\right) \tag{9}$$

In a given area, the number of samples is preserved, i.e.,

$$NP(r_1, r_2)dr_1dr_2 = NP'(\zeta_1, \zeta_2)S(r_1, r_2, \zeta_1, \zeta_2)$$

The new transformed area can be expressed as

$$S = \left\|\frac{\partial(\zeta_1, \zeta_2)}{\partial(r_1, r_2)}\right\| dr_1dr_2$$

The probability of finding a point in the (r_1, r_2) space is 1, since the distribution is uniform. Using this fact, and the above two equations we can write

$$P'(\zeta_1, \zeta_2) = \left\|\frac{\partial(\zeta_1, \zeta_2)}{\partial(r_1, r_2)}\right\|^{-1} \underbrace{P(r_1, r_2)}_{=1} \tag{10}$$

$$= \underbrace{\left\|\frac{\partial(r_1, r_2)}{\partial(\zeta_1, \zeta_2)}\right\|}_{J} \tag{11}$$

The determinant in Eq.(11) is known as the Jacobian and is is evaluated as follows:

$$J = \left\|\begin{matrix} \frac{\partial r_1}{\partial \zeta_1} & \frac{\partial r_1}{\partial \zeta_2} \\ \frac{\partial r_2}{\partial \zeta_1} & \frac{\partial r_2}{\partial \zeta_2} \end{matrix}\right\| = \left|\frac{\partial r_1}{\partial \zeta_1}\frac{\partial r_2}{\partial \zeta_2} - \frac{\partial r_2}{\partial \zeta_1}\frac{\partial r_1}{\partial \zeta_2}\right| \tag{12}$$

Let us now calculate the terms in the Jacobian using Eq.(8) and Eq.(9),

$$
\begin{align}
\frac{\partial r_1}{\partial \zeta_1} &= -\zeta_1 \exp\left\{-\frac{\zeta_1^2+\zeta_2^2}{2}\right\} = -\zeta_1 r_1 \tag{13}\\
\frac{\partial r_1}{\partial \zeta_2} &= -\zeta_2 \exp\left\{-\frac{\zeta_1^2+\zeta_2^2}{2}\right\} = -\zeta_2 r_1 \tag{14}\\
\frac{\partial r_2}{\partial \zeta_1} &= -\frac{\zeta_2}{2\pi(\zeta_1^2+\zeta_2^2)} \tag{15}\\
\frac{\partial r_2}{\partial \zeta_2} &= \frac{\zeta_1}{2\pi(\zeta_1^2+\zeta_2^2)} \tag{16}
\end{align}
$$

Substituting these in Eq. (11)

$$
\begin{align}
P'(\zeta_1,\zeta_2) &= \left|-\frac{\zeta_1^2 r_1}{2\pi(\zeta_1^2+\zeta_2^2)} - \frac{\zeta_2^2 r_1}{2\pi(\zeta_1^2+\zeta_2^2)}\right| \tag{17}\\
&= \frac{r_1}{2\pi} \tag{18}\\
&= \frac{1}{2\pi}\exp\{\zeta_1^2+\zeta_2^2\} \tag{19}\\
&= \sqrt{\frac{1}{2\pi}}\exp\{\zeta_1^2\}\sqrt{\frac{1}{2\pi}}\exp\{\zeta_2^2\} \tag{20}\\
\therefore P'(\zeta_1,\zeta_2) &= P'(\zeta_1)P'(\zeta_2) \tag{21}
\end{align}
$$

which means that the two variables are uncorrelated. Therefore, a general Gaussian random number distribution can be expressed as

$$
P'(\zeta) = \sqrt{\frac{1}{2\pi}}\exp\{\zeta^2\}
$$

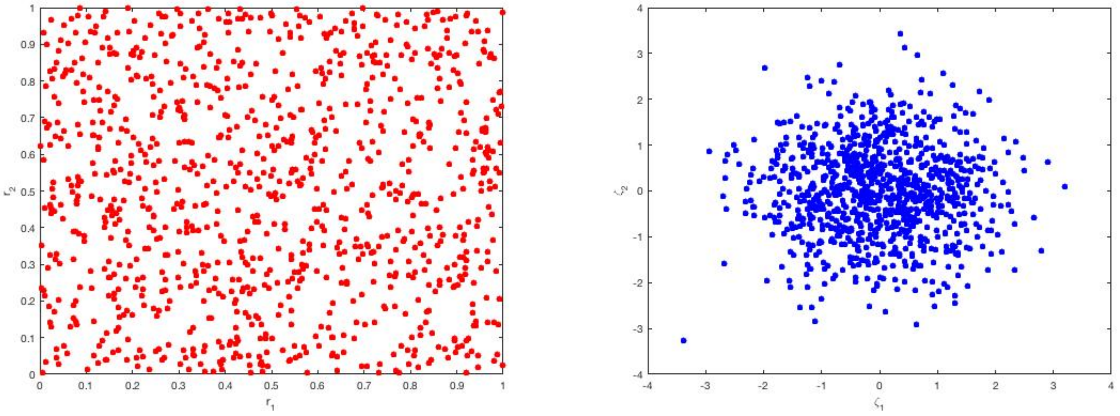
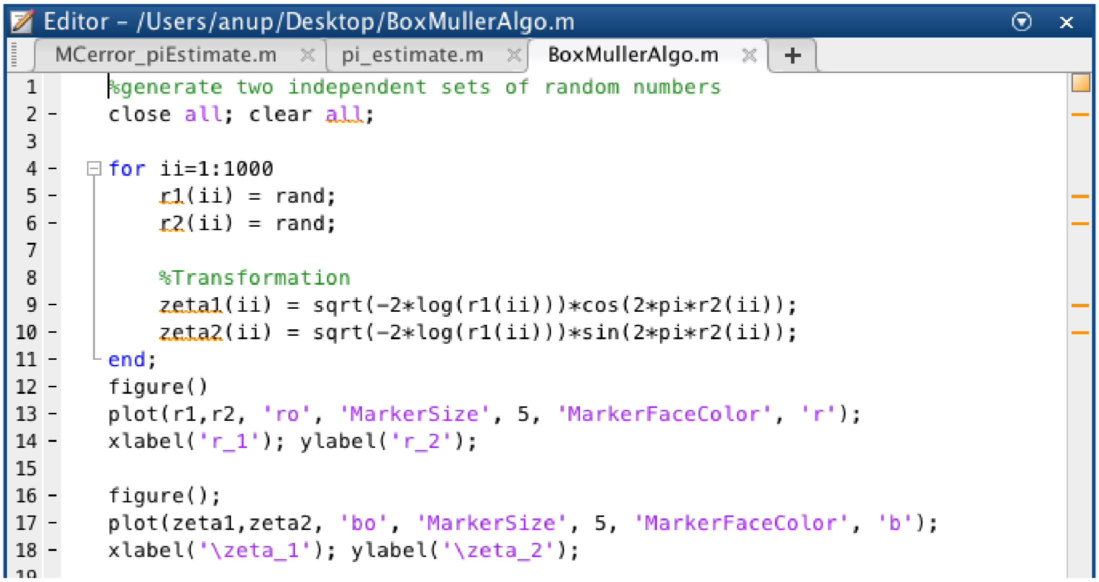


Figure 5: $r_1 - r_2$ space and the transformed $\zeta_1 - \zeta_2$ space



```
%generate two independent sets of random numbers
close all; clear all;

for ii=1:1000
    r1(ii) = rand;
    r2(ii) = rand;

    %Transformation
    zeta1(ii) = sqrt(-2*log(r1(ii)))*cos(2*pi*r2(ii));
    zeta2(ii) = sqrt(-2*log(r1(ii)))*sin(2*pi*r2(ii));
end;
figure()
plot(r1,r2, 'ro', 'MarkerSize', 5, 'MarkerFaceColor', 'r');
xlabel('r_1'); ylabel('r_2');

figure();
plot(zeta1,zeta2, 'bo', 'MarkerSize', 5, 'MarkerFaceColor', 'b');
xlabel('\zeta_1'); ylabel('\zeta_2');
```

Figure 6: Source Code