PHYS 516 Assignment 1
Monte Carlo Basics **Cathie Tsz Yan SO**

Part I – Theoretical Foundation

1. (*Monte Carlo sampling*)
 The Central Theorem states that the standard deviation of the finite sample mean $\bar{f}$ is

$$\sqrt{\frac{\overline{f^2} - \left(\bar{f}\right)^2}{M-1}}$$

where

$$\bar{f} = \langle f(x) \rangle = \frac{1}{M} \sum_{n-1}^{N} f(r_n)$$
$$\overline{f^2} = \langle f^2(x) \rangle = \frac{1}{M} \sum_{n-1}^{M} f^2(r_n)$$
$$M = no.of\ random\ numbers$$

Proof: To find the variance of the sample mean -

$$Var\{\overline{f(x)}\} = Var\left\{\frac{1}{M} \sum_{n-1}^{N} f(r_n)\right\} = \frac{1}{M^2} \sum_{m=1}^{M} \sum_{n=1}^{M} \langle [f(r_m) - \langle f(x) \rangle][f(r_n) - \langle f(x) \rangle] \rangle$$

Assumption – the random variables associated with f are uncorrelated. For any uncorrelated random variables A and B, we have $\langle AB \rangle = \langle A \rangle \langle B \rangle$. Thus,

$$\langle [f(r_m) - \langle f(x) \rangle][f(r_n) - \langle f(x) \rangle] \rangle = \begin{cases} [f(r_n) - \langle f(x) \rangle]^2 = Var\{f(x)\} & if\ m = n \\ \langle f(r_m) - \langle f(x) \rangle \rangle \langle f(r_n) - \langle f(x) \rangle \rangle = 0 & otherwise \end{cases}$$

which gives

$$Var\{\overline{f(x)}\} = \frac{Var\{f(x)\}}{M}$$

However, there is no way to to obtain a precise value for the population variance $Var\{f(x)\}$. We will define a new variable s_M. Then we will show $\langle s_M \rangle$ estimates $Var\{f(x)\}$.

$$s_M \equiv \frac{1}{M} \sum_{n=1}^{M} f^2(r_n) - \left[\frac{1}{M} \sum_{n=1}^{M} f(r_n)\right]^2$$

Hence,

$$\langle s_M \rangle = \frac{1}{M} \sum_{n=1}^{M} \langle f^2(x) \rangle - \frac{1}{M^2} \sum_{m=1}^{M} \sum_{n=1}^{M} \langle f(r_m) f(r_n) \rangle$$
$$= \langle f^2(x) \rangle - \frac{1}{M^2} \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n \neq m}}^{M} \langle f(r_m) f(r_n) \rangle - \frac{1}{M^2} \sum_{m=1}^{M} \langle f^2(x) \rangle$$

Again, imposing the uncorrelated assumption, we have $\langle f(r_m) f(r_n) \rangle = \langle f(x) \rangle \langle f(x) \rangle$. Therefore,

$$\langle s_M \rangle = \left(1 - \frac{1}{M}\right) \langle f^2(x) \rangle - \frac{1}{M^2} M \langle f(x) \rangle (M-1) \langle f(x) \rangle$$
$$= \frac{M-1}{M} (\langle f^2(x) \rangle - \langle f(x) \rangle^2) = \frac{M-1}{M} Var\{f(x)\}$$

Thus,

$$Std\{\overline{f(x)}\} = \sqrt{Var\{\overline{f(x)}\}} = \sqrt{\frac{Var\{f(x)\}}{M}} = \sqrt{\frac{\langle s_M \rangle}{M-1}} = \sqrt{\frac{\overline{f^2} - \left(\bar{f}\right)^2}{M-1}}$$

2. (*Non-uniform random number generation: transformation method*)

Since the number of samples in a fixed area under the Box-Muller transformation should remain the same, we have

$$N \times probability\ density \times area = N\, P(r_1, r_2)\, dr_1 dr_2 = N\, P(\zeta_1, \zeta_2)\, S(r_1, r_2, dr_1, dr_2)$$

where

$$\zeta_1 = \sqrt{-2\ln r_1}\cos(2\pi r_2)$$
$$\zeta_2 = \sqrt{-2\ln r_1}\sin(2\pi r_2)$$
$$S = \left|\frac{\partial(\zeta_1, \zeta_2)}{\partial(r_1, r_2)}\right| dr_1 dr_2$$

so, by inverting the equations above,

$$r_1 = \exp\left(-\frac{\zeta_1^2 + \zeta_2^2}{2}\right)$$

and

$$P(\zeta_1, \zeta_2) = \left|\frac{\partial(\zeta_1, \zeta_2)}{\partial(r_1, r_2)}\right|^{-1}$$
$$= \left|\frac{\partial \zeta_1}{\partial r_1}\frac{\partial \zeta_2}{\partial r_2} - \frac{\partial \zeta_1}{\partial r_2}\frac{\partial \zeta_2}{\partial r_1}\right|^{-1}$$
$$= \left|-\frac{r_1^{-1}\cos(2\pi r_2)}{\sqrt{-2\ln r_1}} 2\pi\sqrt{-2\ln r_1}\cos(2\pi r_2) - 2\pi\sqrt{-2\ln r_1}\sin(2\pi r_2)\frac{r_1^{-1}\sin(2\pi r_2)}{\sqrt{-2\ln r_1}}\right|^{-1}$$
$$= \left|-\frac{2\pi}{r_1}\right|^{-1} = \frac{r_1}{2\pi} = \frac{1}{2\pi}\exp\left(-\frac{\zeta_1^2 + \zeta_2^2}{2}\right)$$
$$= \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{\zeta_1^2}{2}\right)\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{\zeta_2^2}{2}\right) = P(\zeta_1)P(\zeta_2)$$

i.e. separable.

$$P(\zeta_{1\ or\ 2}) = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{\zeta_{1\ or\ 2}^2}{2}\right)$$

Therefore, the Box-Muller algorithm generates (two) normally distributed random number(s) $\zeta_{1\ or\ 2}$ with unit variance.

 Cathie Tsz Yan SO

Part II – Numerical Test

1. (*Monte Carlo estimate*)

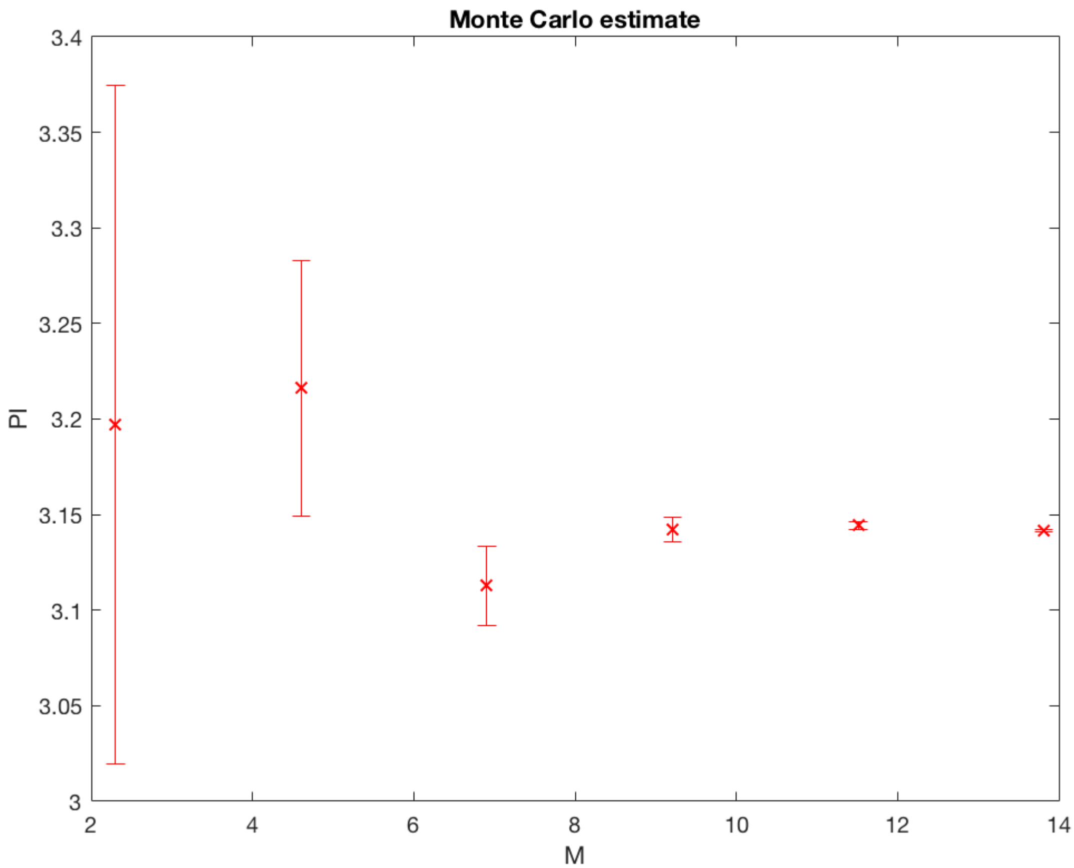
Source code:

```
/* Monte Carlo integration of PI by sample mean */
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <math.h>

int main() {
  double x, pi, sum = 0.0, pi2, sum2 = 0.0, stdv, fx;
  int try, ntry;
  printf("Input the number of MC trials\n");
  scanf("%d",&ntry);
  srand((unsigned)time((long *)0));
  for (try=0; try<ntry; try++) {
    x = rand()/(double)RAND_MAX;
    fx = 4.0/(1.0 + x*x);
    sum += fx;
    sum2 += fx*fx;
  }
  pi = sum/ntry;
  pi2 = sum2/ntry;
  stdv = sqrt((pi2-pi*pi)/(ntry-1));
  printf("MC estimate for PI = %f += %e\n", pi, stdv);
  return 0;
}
```

Results:

M = [10 10^2 10^3 10^4 10^5 10^6];
PI = [3.197029 3.216067 3.112710 3.142317 3.144274 3.141827];
STDV = [1.776670e-01 6.674424e-02 2.055114e-02 6.455923e-03 2.033267e-03 6.433500e-04];

Plot on next page.


Monte Carlo estimate
PI
M
3.4
3.35
3.3
3.25
3.2
3.15
3.1
3.05
3
2
4
6
8
10
12
14

2. (*Monte Carlo error*)

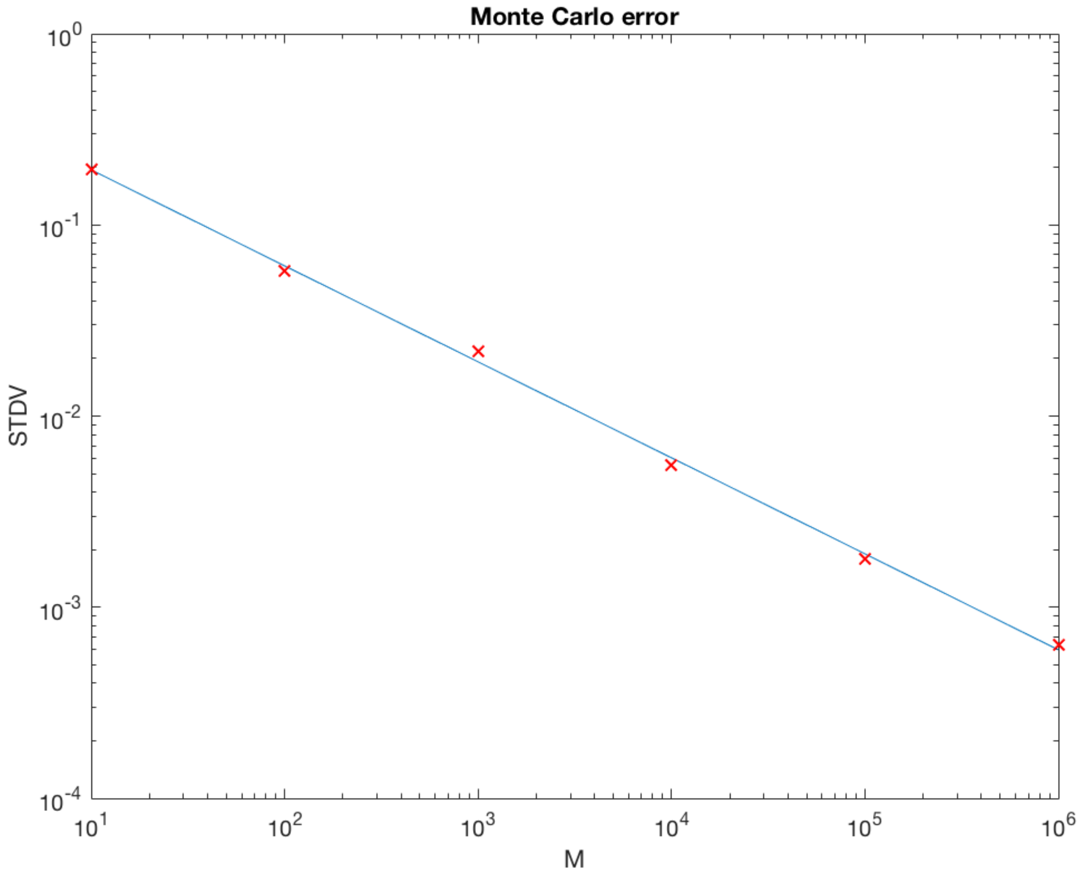
Source code:

```
/* Monte Carlo integration of PI by sample mean */
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <math.h>
#define NSEED 100

int main() {
  double x, pi, sum = 0.0, pi_av = 0.0, pi2_av = 0.0, stdv;
  int try, ntry, outer;
  printf("Input the number of MC trials\n");
  scanf("%d",&ntry);
  srand((unsigned)time((long *)0));

  for (outer=1; outer <= NSEED; outer++) {
    sum = 0.0;
    for (try=0; try<ntry; try++) {
    x = rand()/(double)RAND_MAX;
    sum += 4.0/(1.0 + x*x);
    }
    pi = sum/ntry;
    pi_av += pi;
    pi2_av += pi*pi;
  }
  pi_av /= NSEED;
  pi2_av /= NSEED;
  stdv = sqrt(pi2_av-pi_av*pi_av);
  printf("Ntry = %d: Stdv estimate for PI = %e\n", ntry, stdv);
  return 0;
}
```

Results:
M = [10 10^2 10^3 10^4 10^5 10^6];
sM = [1.964829e-01 5.781135e-02 2.187134e-02 5.564133e-03 1.788560e-03 6.341625e-04];

Estimated value of power: -0.5022

Plot on next page.


Monte Carlo error
STDV
M
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^1
10^2
10^3
10^4
10^5
10^6