MATH 565: Monte Carlo Methods in Finance

Instructions:

a. *Show all your work to justify your answers.*

b. *Group work can be done in groups of no more than THREE students.*

c. *Individual work must be done on your own.*

d. *What you submit should represent your own work. If you use other sources, cite them.*

e. *You may use sample code from the class without citation.*

f. *Please include the results, conclusions, and published html files in your homework. Also include your group members' names in your homework and submit it to Blackboard using the following method* Math565 + HW1 + Group Number (e.g. Math565HW1Group1.pdf)

g. *MATLAB code must be submitted to Blackboard using the following method.* Math565 + HW1 + Group Number (e.g. Math565HW1Group1.m)

Assignment 1 Due 11:59 pm Tuesday, September 20, 2022 (Group Work)

1. Suppose that the Route 29 northbound bus is scheduled to arrive at the 32nd Street and State Street stop on the hour, at 15 minutes past the hour, at 30 minutes past the hour, and at 45 minutes past the hour. However, due to random fluctuations, it arrives anywhere between 1 minute early and 2 minutes late with uniform distribution. Assume that the arrivals of different buses are independent and identically distributed (IID). Compute by hand the following probabilities:

   a) The probability of four successive buses all not being late;
   b) The probability of needing to wait for more than 15 minutes for a bus if you arrive at the stop at noon; and
   c) The expected time required to wait for a bus if you arrive at the bus stop at a random time is 7.55. Create a Monte Carlo simulation to compute the exptected time. What answer do you compute using 10^6 runs?

   Solution:

   a) As the bus arrives any where between 1 minute early and 2 minutes late with uniform distribution, we know

   $$P(bus\ not\ being\ late) = \frac{1}{3}, \qquad P(bus\ being\ late) = \frac{2}{3}.$$

   And the arrivals of different buses are IID, thus

   $$P(Four\ successive\ buses\ all\ not\ being\ late) = P(bus\ not\ being\ late)^4 = \frac{1}{3^4} = \frac{1}{81}.$$

   b) The probability of needing to wait for more than 15 minutes for a bus if you arrive at the stop at noon is the probability of the bus around noon not being late and the next bus late.

   $$\begin{aligned} P(First\ not\ being\ late,\ second\ bus\ being\ late) &= P(bus\ not\ being\ late) \cdot P(bus\ being\ late) \\ &= \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}. \end{aligned}$$

c) Let Y be the arrival time of those passengers arriving between 12:02 PM and 12:17 PM. i.e., $Y \sim U[2,17]$. By symmetry, the expected waiting time for passengers arriving in this time interval is the same as for those arriving at any time. Let W be the amount of time required until the next bus comes.

```
%Y to denote the time you arrive at the stop between 12:02-12:17
%X to denote the time the bus arrive at the stop between
    12:14-12:17 and 12:29-12:32
n = 1e6;
y = 15*rand(n,1)+2;
x = 3*rand(n,2)-1; %X ~ U[-1,2]
x(:,1) = 15 + x(:,1);
x(:,2) = 30 + x(:,2);
%Calculate waiting time
w = (x(:,1)-y).*(x(:,1)>=y) + (x(:,2)-y).*(x(:,1)< y);
%Get the average
mu = mean(w);
disp(['The waiting time = ' num2str(mu,'%8.5f')])
```

2. Consider the estimator of $\mu = \mathbb{E}(Y)$ defined for even n as

$$\tilde{\mu}_n = \frac{1}{n}\left(\frac{3}{2}Y_1 + \frac{1}{2}Y_2 + \frac{3}{2}Y_3 + \frac{1}{2}Y_4 + \cdots + \frac{3}{2}Y_{n-1} + \frac{1}{2}Y_n\right), \qquad Y_i \overset{IID}{\sim} Y.$$

a) Show that $\tilde{\mu}_n$ is unbiased.

b) Derive $\mathrm{var}(\tilde{\mu}_n)$.

c) Which estimator of μ is better, $\tilde{\mu}_n$ or

$$\hat{\mu}_n = \frac{1}{n}(Y_1 + Y_2 + \cdots + Y_n), \qquad Y_i \overset{IID}{\sim} Y.$$

Why?

Solution:

a)

$$\begin{aligned}
\mathbb{E}(\tilde{\mu}_n) &= \mathbb{E}\left[\frac{1}{n}\left(\frac{3}{2}Y_1 + \frac{1}{2}Y_2 + \frac{3}{2}Y_3 + \frac{1}{2}Y_4 + \cdots + \frac{3}{2}Y_{n-1} + \frac{1}{2}Y_n\right)\right] \\
&= \frac{1}{n}\left(\frac{3}{2}\mathbb{E}[Y_1] + \frac{1}{2}\mathbb{E}[Y_2] + \frac{3}{2}\mathbb{E}[Y_3] + \frac{1}{2}\mathbb{E}[Y_4] + \cdots + \frac{3}{2}\mathbb{E}[Y_{n-1}] + \frac{1}{2}\mathbb{E}[Y_n]\right) \qquad \text{as } Y_i \overset{IID}{\sim} Y \\
&= \frac{1}{n}\left(\frac{3}{2}\cdot\mu\cdot\frac{n}{2} + \frac{1}{2}\cdot\mu\cdot\frac{n}{2}\right) \\
&= \mu
\end{aligned}$$

b) Let $\mathrm{var}(Y) = \sigma^2$.

$$\begin{aligned}
\mathrm{var}(\tilde{\mu}_n) &= \mathrm{var}\left[\frac{1}{n}\left(\frac{3}{2}Y_1 + \frac{1}{2}Y_2 + \frac{3}{2}Y_3 + \frac{1}{2}Y_4 + \cdots + \frac{3}{2}Y_{n-1} + \frac{1}{2}Y_n\right)\right] \\
&= \frac{1}{n^2}\left(\frac{9}{4}\mathrm{var}(Y_1) + \frac{1}{4}\mathrm{var}(Y_2) + \frac{9}{4}\mathrm{var}(Y_3) + \frac{1}{4}\mathrm{var}(Y_2) + \cdots + \frac{9}{4}\mathrm{var}(Y_{n-1}) + \frac{9}{4}\mathrm{var}(Y_n)\right) \\
&= \frac{1}{n^2}\left(\frac{9}{4}\cdot\sigma^2\cdot\frac{n}{2} + \frac{1}{4}\cdot\sigma^2\cdot\frac{n}{2}\right) \\
&= \frac{5}{4n}\sigma^2
\end{aligned}$$

c) $\hat{\mu}_n$ is better, as $\text{var}(\hat{\mu}) < \text{var}(\tilde{\mu})$.

3. Let Y be a Bernoulli random variable with unknown mean μ, i.e., $P(Y = 0) = 1 - \mu$ and $P(Y = 1) = \mu$. Let $\hat{\mu}_n = (Y_1 + \cdots + Y_n)/n$ be the sample mean of IID Y_i that have the same distribution as Y. How large must n be to ensure that the root mean square error (RMSE) in approximating μ by $\hat{\mu}_n$ is no larger than 0.01? Your final answer should not depend on μ.

   Solution: As Y is a Bernoulli random variable with unknown mean μ, we have

$$\text{E}(Y) = \mu, \qquad \text{var}(Y) = \mu(1-\mu).$$

   Thus

$$\text{E}(\hat{\mu}_n) = \text{E}\left[(Y_1 + \cdots + Y_n)/n\right] = \frac{1}{n}\text{E}\left(Y_1 + \cdots + Y_n\right) = \frac{1}{n} n\,\text{E}(Y) = \mu,$$

   is an unbiased estimator.

$$\text{var}(\hat{\mu}_n) = \text{var}\left[(Y_1 + \cdots + Y_n)/n\right] = \frac{1}{n^2} n \cdot \text{var}(Y) = \frac{\mu(1-\mu)}{n}.$$

   Therefore,

$$RMSE = \sqrt{\text{var}(\hat{\mu}_n)} \le \sqrt{\frac{\mu(1-\mu)}{n}} \le 0.01 \Longrightarrow n \ge \frac{\mu(1-\mu)}{0.01^2} \ge 2500,$$

   as $\mu(1-\mu) \ge \frac{1}{4}$ when $0 < \mu < 1$.

4. Let Y be a random variable with unknow mean μ and variance σ^2. Let

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{\mu}_n)^2, \qquad Y_1, Y_2, \ldots \overset{IID}{\sim} Y.$$

   (a) If $\hat{\sigma}^2_{1000} = 4.3$, how many samples should be used to estimate μ with an absolute error tolerance of 0.02 with a confidence level of 99% based on a Central Limit Theorem (CLT) approximation?

   (b) Again assume that $\hat{\sigma}^2_{1000} = 4.3$ as in part (a), and suppose that $\hat{\mu}_n = 12.456$ for n chosen according to part a). If you compute another $\hat{\mu}_n$ with the same n, but for a sample independent of the previous one, would you expect it to be 12.456 also? Explain why. Would you be surprised if this new $\hat{\mu}_n$ were 13.456? Explain why.

   Solution:

   (a) Choosing a standard deviation inflation factor when we have CLT confidence intevals of half-width $2.58 \times 1.2 \times \sqrt{4.3}/\sqrt{n}$. Setting this to be 0.02, we get

$$n = \left\lceil \frac{(2.58)^2 \times (1.2)^2 \times 4.3}{(0.02)^2} \right\rceil = 103042.$$

   (b) We expect that independent $\hat{\mu}_n$ with the same n should be somewhat less than ± 0.02 away from the original. Since $\hat{\mu}_n$ is random, we would not expect two different $\hat{\mu}_n$ to agree to three digits, and we would be surprised if the new $\hat{\mu}_n$ differed by more than 0.04 from the old one. Thus our answers are "no" and "yes", respectively.

5. Use Central Limit Theorem confidence intervals to compute the answer to 1c) by Monte Carlo simulation to an absolute tolerance of 1 second with 99% confidence.
   (Hint: You can use `meanMC_CLT`) *Solution:* From the problem, we have $\alpha = 0.01, \varepsilon = 1/60$. Suppose, we generate 1000 samples at first, then we get the sample standard deviation $\hat{\sigma}$. Thus by Central Limit Theorem

$$n = n_\mu = \left\lceil \left(\frac{2.58\mathcal{C}\hat{\sigma}}{\varepsilon}\right)^2 \right\rceil,$$

   where $\mathcal{C} = 1.2$ is an inflation factor.

```
%Problem 5 is solved directly by meanMC_CLT
[sol, out] = meanMC_CLT(@(n) w(n), 1/60);
disp('The exact expection of waiting time = 7.55.')
disp(['The waiting time based on meanMC_CLT = ' num2str(sol,'%8.5f')
    ])
disp(['The sample size = ' num2str(out.nSample,'%8d')])

function waitingtime = w(n)
%Y to denote the time you arrive at the stop between 12:02-12:17
%X to denote the time the bus arrive at the stop between 12:02-12:32
y = 15*rand(n,1)+2;
x = 3*rand(n,2)-1; %X~U[-1,2]
x(:,1) = 15 + x(:,1);
x(:,2) = 30 + x(:,2);
%Calculate waiting time
waitingtime = (x(:,1)-y).*(x(:,1)>=y) + (x(:,2)-y).*(x(:,1)< y);
end
```

The output is

```
The exact expection of waiting time = 7.55.
The waiting time based on meanMC_CLT = 7.54693
The sample size = 643249
```

Assignment 1 Computer Problem

Table of Contents

1 c)

```
%Y to denote the time you arrive at the stop between 12:02-12:17
%X to denote the time the bus arrive at the stop between 12:02-12:32
n = 1e6;
y = 15*rand(n,1)+2;
x = 3*rand(n,2)-1; %X~U[-1,2]
x(:,1) = 15 + x(:,1);
x(:,2) = 30 + x(:,2);
%Calculate waiting time
w = (x(:,1)-y).*(x(:,1)>=y) + (x(:,2)-y).*(x(:,1)< y);
%Get the average
mu = mean(w);
disp(['The waiting time  = ' num2str(mu,'%8.5f')])
```

```
The waiting time  = 7.55694
```

Problem 5 is solved directly by meanMC_CLT

```
[sol, out] = meanMC_CLT(@(n) wt(n), 1/60);
disp('The exact expection of waiting time = 7.55.')
disp(['The waiting time based on meanMC_CLT = ' num2str(sol,'%8.5f')])
disp(['The sample size = ' num2str(out.nSample,'%8d')])


function waitingtime = wt(n)
%Y to denote the time you arrive at the stop between 12:02-12:17
%X to denote the time the bus arrive at the stop between 12:02-12:32
y = 15*rand(n,1)+2;
x = 3*rand(n,2)-1; %X~U[-1,2]
x(:,1) = 15 + x(:,1);
x(:,2) = 30 + x(:,2);
%Calculate waiting time
waitingtime = (x(:,1)-y).*(x(:,1)>=y) + (x(:,2)-y).*(x(:,1)< y);
end
```

```
The exact expection of waiting time = 7.55.
The waiting time based on meanMC_CLT = 7.55303
The sample size = 666000
```

Published with MATLAB® R2022a