RMSC5102 Final Review

LEUNG MAN FUNG, HEMAN

SPRING, 2020

Agenda

Review

- Monte Carlo Method
- Random Variable Generation
- Variance Reduction Technique
- Simulation in Action

Q&A

Monte Carlo Method

Standard Monte Carlo

HW2 5a: price a European call option

- Recall payoff function is $\max(S_T - K, 0)$
- Estimate $E[\max(S_T - K, 0)]$ by sample average $\frac{1}{n}\sum_{i=1}^{n} \max(S_T^{(i)} - K, 0)$

Algorithm

- 1) Generate $Z \sim N(0,1)$
- 2) Set $S_T = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z}$
- **3) Compute $\boldsymbol{\pi_i = \max(S_T - K, 0)}$**
- 4) Repeat 1 to 3 for $i = 1, \ldots, n$
- 5) Option price = $\frac{\mathrm{e}^{-\mathrm{rT}}}{\mathrm{n}} \sum_{i=1}^{n} \pi_i$

Things to Note

General algorithm (always refer to tutorial notes)

- 1) Generate random variable X_i
- 2) Calculate $h_i = h(X_i)$, where h is the target function
- 3) Repeat 1 and 2 for n times
- 4) $\hat{\theta} = \frac{1}{n}\sum_{j=1}^{n} h_j$ (remember to do discounting if necessary)

Be careful of…

- X_i is not necessarily Normal. Some students directly used the previous algorithm in midterm
- The target function $h(x)$ that you are interested in
  - We need to adjust for conditional probability in stratified sampling sometimes because h changes
- How to generate X_i. If you do not write R code, you need to use inverse transform
  - For any X_i that does not follow $N(0,1)$ or $U(a, b)$. This includes discrete uniform r.v. (to be discussed)

Random Variable Generation

Inverse Transform

If we know $X \sim F_X$ (i.e. the cdf), we can generate X out of $U \sim U(0,1)$

- Algorithm (discrete)
  - Generate $U \sim U(0,1)$
  - $X = x_j$ if $\sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^{j} p_i$
- Algorithm (continuous)
  - Generate $U \sim U(0,1)$
  - $X = F_X^{-1}(U)$ assuming the inverse exists

Do NOT use the continuous version for discrete uniform r.v.

- As argued in my Q&A, this is not appropriate in view of algorithm

Rejection Sampling

If we can simulate $Y \sim G_Y$ easily, we can use the proportional distribution as a basis to simulate X with pdf $f(x)$

Algorithm

- 1) Find $c = \max_y \frac{f(y)}{g(y)}$
- 2) Generate Y_i from a density g: $U_1 \sim U(0,1) \Rightarrow Y_i = G^{-1}(U_1)$
- 3) Generate $U_2 = U(0,1)$
- 4) If $\mathrm{U}_2 \leq \frac{1}{c} \cdot \frac{f(Y_i)}{g(Y_i)}$, set $X_i = Y_i$, otherwise return to 2

Number of iterations needed: $N \sim Geo\left(\frac{1}{c}\right) \Rightarrow E(N) = c$

Variance Reduction Technique

Antithetic Variables

If we are able to generate negatively correlated underlying random variables, the estimator can have lower variance as compared with independent samples

- This requires the target function $h(x)$ to be monotone
- Show $h'(x) \geq 0$ or $h'(x) \leq 0$ within the target range for monotonicity
- As $h(x)$ is monotone, $Cov\big(h(U), h(1-U)\big) \leq 0$ where $U \sim U(0,1)$
- As half of your variables are antithetic, you only need to generate $\frac{n}{2}$ numbers for n samples

Antithetic Variables

Algorithm:

- 1) Generate $U \sim U(0,1)$
- 2) Set $X_i = F^{-1}(U), Y_i = F^{-1}(1-U)$ (note: want X, Y same distribution but negative correlation)
- 3) Repeat 1 and 2 for n times
- 4) $\hat{\theta} = \frac{1}{2n}\sum_{i=1}^{n}[h(X_i) + h(Y_i)]$

Note:

- $F^{-1}(U)$ is monotone in general as cdf is monotone
- Hence $h[F^{-1}(U)]$ is monotone if $h(\cdot)$ is monotone

Stratified Sampling

If we have information about grouping in the population, then we may use conditional mean (mean of subgroup) as the sample from the population

Algorithm:

- Generate $V_{i,j} = \frac{1}{B}(U_{i,j} + i - 1)$ where $U_{i,j} \sim U(0,1)$ for $i = 1, \dots, B; j = 1, \dots, N_B$
- Set $X_{i,j} = F^{-1}(V_{i,j})$
- $\hat{\theta} = \frac{1}{B \times N_B} \sum_{j=1}^{N_B} \left[h\left(X_{1,j}\right) + \mathrm{h}\left(X_{2,j}\right) + \cdots + \mathrm{h}\left(X_{B,j}\right) \right]$ (average over subsamples and bins, remember to adjust for conditional probability)

Note:

- I have changed the representation to matrix elements $V_{i,j}$ for clearness
- i represents index of bins and j represent index of elements within a bin

Stratified Sampling

When to adjust for conditional probability?

- Try to write out the expectation you are trying to approximate (e.g. Ch6 p.43)
- Usually you need to when there is an indicator or you have restricted the support
- If you have time, you can try to simulate and compare with standard Monte Carlo
  - If the estimates differ a lot, probably you need to
  - Do ALL questions first

Control Variate

If we combine the estimate of our target unknown quantity with estimates of some known quantities, we can exploit the known information

Algorithm:

- Find μ_Y for Y with a known distribution (or estimate μ_Y via pilot simulation)
- Generate X_i, Y_i for $i = 1, \dots, n$
- Compute $\bar{X}, \bar{Y}, \hat{\sigma}_{XY}, \hat{\sigma}_Y^2$
- $\hat{\theta} = \bar{X} - \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_Y^2}(\bar{Y} - \mu_Y)$

Pilot simulation: we can run a simulation with small sample size (e.g. $m = 100$) and compute $\hat{\sigma}_{XY}, \hat{\sigma}_Y^2$ and $\mu_Y = \bar{Y}_m$ based on this pilot sample. Then we can use their values when we compute $\hat{\theta} = \bar{X}_n - \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_Y^2}(\bar{Y}_n - \mu_Y)$ for our target n samples

Control Variate

Properties of effective control: evaluable from simulation data, known mean and high correlation with the simulation variable. Possible candidates are underlying random variable (e.g. uniform when we use inverse transform) and martingale transform (will not be tested)

Note:

- The algorithm in last slide is one-off, i.e. it does not affect each sample
- So we should use control variate last if we were to combine the methods (e.g. HW4 Q3)

Importance Sampling

If certain values of the simulation variable have more impact on the parameter of interest (e.g. probability of a rare event)

- We can try to “emphasize” those values by sampling them more frequently and reduce variance
- This can be done by changing the probability measure using the likelihood ratio as weight

Algorithm:

- Find the likelihood ratio $\frac{f(x)}{g(x)}$ where $f(x)$ is the original target pdf
- Generate $X_i \sim G$ for $i = 1, \dots, n$
- $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \frac{h(X_i)f(X_i)}{g(X_i)}$

Maximum principle: choose g such that both $g(x)$ and $h(x)f(x)$ take maximum values at the same $x = x^*$ (not in our syllabus, just for your reference)

Importance Sampling

A possible candidate for $g(x)$ is the tilted density

Tilted density: $f_t(x) = \frac{e^{tx} f(x)}{M_X(t)}$ where $M_X(t)$ is the moment generating function of X

Choice of t in importance sampling

- Find the upper bound $\frac{h(x)f(x)}{f_t(x)} \leq \frac{h(x_t^*)f(x_t^*)}{f_t(x_t^*)}$ and minimize for t
  - The upper bound should only depends on t for minimization
- In other words, find $x = x_t^*$ such that $\frac{h(x)f(x)}{f_t(x)} \leq \frac{h(x_t^*)f(x_t^*)}{f_t(x_t^*)}$ for all x in the support
  - x_t^* has subscript t because it may depend on t
- Then minimize $\frac{h(x_t^*)f(x_t^*)}{f_t(x_t^*)}$ with respect to t

Simulation in Action

Down-and-in Call Option

Price of down-and-in option: $C_{di} = e^{-rT} E\left[(S_T - K)^+ \mathbb{I}\left(\min_{0 \le t \le T} S_t < V\right)\right]$

Algorithm:

- 1) Generate $Z \sim N(0,1)$
- 2) Set $S_{t_i} = S_{t_{i-1}} \exp\left[\left(r - \frac{1}{2}\sigma^2\right)(t_i - t_{i-1}) + \sigma\sqrt{t_i - t_{i-1}} Z\right]$
- 3) Repeat step 1 and 2 for $i = 1, \ldots, n$ where $t_0 = 0$ and $t_n = T$
- 4) If $\min_i S_{t_i} < V$, set $C_j = e^{-rT} \max(S_T - K, 0)$. Otherwise set $C_j = 0$
- 5) Repeat step 3 and 4 for $j = 1, \ldots, N$
- 6) The price is given by $\hat{\theta} = \frac{1}{N} \sum_{j=1}^{N} C_j$

Down-and-in Call Option

What if you already have the price of a vanilla call of the same parameters?

- i.e. $C_v = e^{-rT}E[(S_T - K)^+]$
- Possible to evaluate $E\left[\mathbb{I}\left(\min_{0\le t\le T} S_t < V\right)\right]$ and combine with C_v directly
  - This probably will not be tested. Just to stimulate your thinking
- Your target function h becomes $\mathbb{I}\left(\min_{0\le t\le T} S_t < V\right)$ in that case
- This is kind of like adjusting for conditional probability

Path-dependent Option

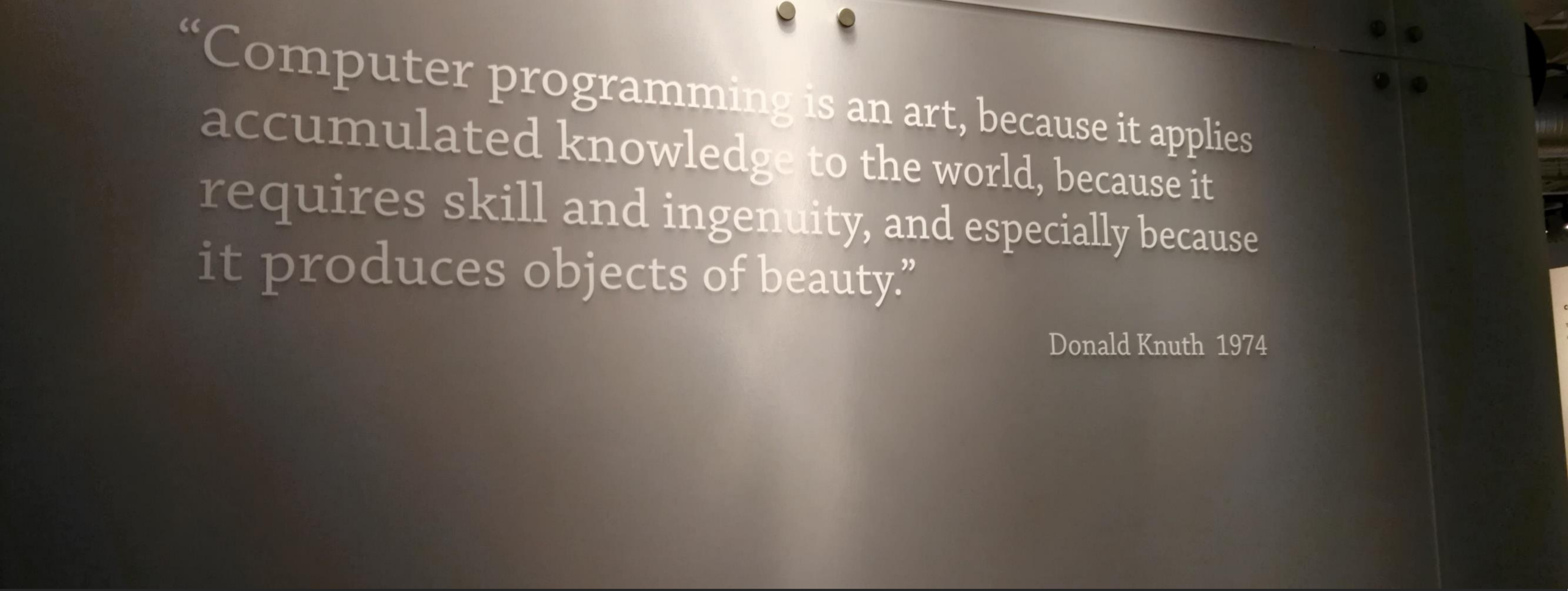
Problem with discretization

- The discretized process does not have the correct transition density
  - First order Euler scheme has normal increments
  - Second order Milstein scheme has non-central chi square increments
  - Optimal tradeoff between n and N exists for the two schemes. See Duffie and Glynn (1995)
    - Content of the paper will not be tested
- The discretized process may incorrectly evaluate the payoff
  - E.g. Asian option
  - Possible solution: Brownian bridge. See Beaglehole, Dybvig and Zhou (1997)

Path-dependent Option

American option

- Problem with branching paths
- Possible solution: linear regression. See Longstaff and Schwartz (2001)
- Just for your interest
  - Consider taking courses from Prof. Wong :P He taught us this in undergrad



Q&A

Thank you for taking RMSC5102! Write me (or department) an email if you like my tutorials :P

Let's keep in touch :)