

# RMSC5102 Simulation Methods for Risk Management Science and Finance

Tutorial Notes

Spring, 2020

LEUNG Man Fung, Heman

## Contents

I) Probability and statistics .....	3
Discrete random variables.....	3
Binomial distribution .....	3
Poisson distribution.....	4
Continuous random variables.....	4
Uniform distribution.....	4
Normal distribution .....	5
Some remarks .....	5
II) Financial derivative .....	6
Forward.....	6
Option.....	6
Binomial tree.....	6
Black–Scholes–Merton model .....	6
III) Stochastic calculus .....	8
Brownian motion .....	8
Stochastic integral .....	8
IV) Simulation methods.....	10
Theoretical support .....	10
Standard Monte Carlo .....	10
Inverse transform.....	10
Rejection sampling.....	11
V) Variance reduction.....	12
Antithetic variables .....	12
Stratified sampling .....	12

## I) Probability and statistics

### Discrete random variables

Random variables: numeric quantities that take different values with specified probabilities

Discrete random variable: a R.V. that takes value from a discrete set of numbers

Probability mass function: a pmf assigns a probability to each possible value  $x$  of the discrete random variable  $X$ , denoted by  $f(x) = P(X = x)$

$$\sum_{i=1}^n f(x_i) = 1 \text{ (total probability rule)}$$

Cumulative distribution function: a cdf gives the probability that  $X$  is less than or equal to the value  $x$ , denoted by  $F(x) = P(X \leq x)$

Expected value:  $\mu = E(X) = \sum_{i=1}^n x_i P(X = x_i)$  (the idea is “probability weighted average”)

Variance:  $\sigma^2 = Var(X) = \sum_{i=1}^n (x_i - \mu)^2 P(X = x_i)$  (the idea is “probability weighted distance from mean”)

$$\text{Alternatively } Var(X) = E(X^2) - [E(X)]^2$$

Translation/rescale:  $E(aX + b) = aE(X) + b$ ,  $Var(aX + b) = a^2 Var(X)$

Linearity of expectation:  $E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i)$

### Binomial distribution

Factorial:  $n! = n \times (n - 1) \times \dots \times 1$ , note that  $0! = 1$

Permutation (order is important):  $P_k^n = \frac{n!}{(n-k)!}$

Combination (order is not important):  $C_k^n = \frac{n!}{k!(n-k)!}$ , also denoted as  $\binom{n}{k}$

Binomial distribution: probability distribution on the number of successes  $X$  in  $n$  independent experiments, each experiment has a probability of success  $p$ , then  $X \sim B(n, p)$

Pmf:  $P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$  for  $x = 0, 1, 2, \dots, n$

Mean:  $E(X) = np$

Variance:  $Var(X) = np(1 - p)$

### Poisson distribution

Poisson distribution: probability distribution on the number of occurrence  $X$  (usually of a rare event) over a period of time or space with rate  $\lambda$ , then  $X \sim Po(\lambda)$ . Useful in modelling jump

$$\text{Pmf: } P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \text{ for } x = 0, 1, 2, \dots$$

$$\text{Mean: } E(X) = \lambda$$

$$\text{Variance: } Var(X) = \lambda$$

### Continuous random variables

Continuous random variable: a R.V. that takes value over an interval of numbers

Probability density function: a pdf specifies the probability of the random variable falling within a particular range of values, denoted by  $f(x)$

$$P(a \leq X \leq b) = \int_a^b f(x)dx, \text{ which is the area under the curve from } a \text{ to } b$$

$$P(X = a) = \int_a^a f(x)dx = 0 \text{ for all } a$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \text{ (total probability rule)}$$

Cumulative distribution function: a cdf gives the probability that  $X$  is less than or equal to the value  $x$ , denoted by  $F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$

$$P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a) \text{ (by the fundamental theorem of calculus)}$$

$$\text{Expected value: } \mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx \text{ (the idea is "probability weighted average")}$$

$$\text{Variance: } \sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 \text{ (the idea is "probability weighted distance from mean")}$$

### Uniform distribution

Uniform distribution: if  $X$  follows uniform distribution on the interval  $[a, b]$ , then it has the same probability density at any point in the interval and we denote it by  $X \sim U(a, b)$ . Basic R.V. in probability integral transform

$$\text{Pdf: } f(x) = \frac{1}{b-a} \text{ for } a \leq x \leq b, \text{ otherwise } 0$$

$$\text{Cdf: } F(x) = \int_a^x \frac{1}{b-a} dt = \left[ \frac{t}{b-a} \right]_a^x = \frac{x-a}{b-a} \text{ for } a \leq x \leq b$$

$$\text{Mean: } E(X) = \frac{a+b}{2}$$

$$\text{Variance: } Var(X) = \frac{(b-a)^2}{12}$$

### Normal distribution

Normal distribution: if  $X$  follows normal distribution with mean  $\mu$  and variance  $\sigma^2$ , then  $X \sim N(\mu, \sigma^2)$ . Often used to represent continuous random variable with unknown distributions

$$\text{Pdf: } f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \text{ for } -\infty < x < \infty$$

Standard normal distribution:  $Z \sim N(0,1)$

Cdf of standard normal: denoted as  $\Phi(z) = P(Z \leq z)$

$$P(a \leq Z \leq b) = P(Z \leq b) - P(Z \leq a) = \Phi(b) - \Phi(a)$$

$\Phi(-z) = 1 - \Phi(z)$  by symmetric property

Percentile of standard normal:  $\Phi(1.645) = 0.95$ ,  $\Phi(1.96) = 0.975$

Standardization: if  $X \sim N(\mu, \sigma^2)$ , then  $\frac{X-\mu}{\sigma} \sim N(0,1)$

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

### Some remarks

Variance of sum:  $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

Tower rule of expectation:  $E(X) = E[E(X|Y)]$

Law of total variance (EVE):  $Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$

Sum of poisson: if  $X \sim Po(\lambda_1)$ ,  $Y \sim Po(\lambda_2)$  independently, then  $X + Y \sim Po(\lambda_1 + \lambda_2)$

Sum of normal: if  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$  independently, then  $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Square of standard normal: if  $X \sim N(\mu, \sigma^2)$ , the  $Z^2 = \left[\frac{X-\mu}{\sigma}\right]^2 \sim \chi_1^2$

Sum of chi square: if  $X \sim \chi_n^2$ ,  $Y \sim \chi_m^2$ , then  $X + Y \sim \chi_{n+m}^2$

## II) Financial derivative

### Forward

Payoff:  $S_t - K$  (long),  $K - S_t$  (short)

Pricing:  $f = S - Ke^{-r(T-t)}$ ,  $F = Se^{r(T-t)}$

With known cash income:  $f = S - I - Ke^{-r(T-t)}$ ,  $F = (S - I)e^{r(T-t)}$ ,  $I = PV(\text{income})$

With known dividend yield:  $f = Se^{-q(T-t)} - Ke^{-r(T-t)}$ ,  $F = Se^{(r-q)(T-t)}$

Minimum variance hedge ratio:  $h^* = \rho \times \frac{\sigma_S}{\sigma_F} \Rightarrow N_F^* = h^* \times N_A$  (since  $h = \frac{N_F}{N_S}$ )

### Option

Upper bounds:  $C_E \leq C_A \leq S$ ,  $P_E \leq Ke^{-r(T-t)}$ ,  $P_A \leq K$

Lower bounds:  $\max(S - Ke^{-r(T-t)}, 0) \leq C_E \leq C_A$ ,  $\max(Ke^{-r(T-t)} - S, 0) \leq P_E \leq P_A$

Put-call parity:  $C_E - P_E = S - I - Ke^{-r(T-t)}$  (idea is call - put = forward)

Put call inequality:  $S - K \leq C_A - P_A \leq S - Ke^{-r(T-t)}$

European-American relationship:  $P_A > P_E$ ,  $C_A = C_E$  (for non-dividend-paying)

### Binomial tree

Risk neutral probability:  $q = \frac{e^{r\delta t} - d}{u - d}$ ,  $u = e^{\sigma\sqrt{\delta t}}$ ,  $d = u^{-1} = e^{-\sigma\sqrt{\delta t}}$

Pricing:  $f = e^{-r\delta t}[qf_u + (1 - q)f_d]$

Backward induction: start from payoff as terminal prices (American: take max between payoff and f)

### Black-Scholes-Merton model

Black-Scholes equation:  $\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} = rV$

Black-Scholes formula:  $C(S_t, t) = \Phi(d_1)S_t - \Phi(d_2)Ke^{-r(T-t)}$ ,

$$P(S_t, t) = Ke^{-r(T-t)} - S_t + C(S_t, t) = \Phi(-d_2)Ke^{-r(T-t)} - \Phi(-d_1)S_t$$

$$\text{where } d_1 = \frac{1}{\sigma\sqrt{T-t}} \left[ \ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t) \right], d_2 = d_1 - \sigma\sqrt{T-t}$$

Implied volatility: the value of volatility when back-solving an option pricing model (such as BS) with current market price

### III) Stochastic calculus

#### Brownian motion

Wiener process:  $W_t$  is called a Wiener process if the following holds

Stationary increment:  $W_t - W_s \sim N(0, t - s)$

Independent increment:  $W_{t_4} - W_{t_3} \perp W_{t_2} - W_{t_1}$

Starts at zero:  $P(W_{t_0} = 0) = 1$

Properties:  $Cov(W_s, W_t) = \min(s, t)$ ,  $[dW_t]^2 = dt$  (quadratic variation), nowhere differentiable

Itô's process:  $X_t$  is an Itô's process if it is solution to the following stochastic differential equation

$$\begin{cases} dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = a \end{cases}$$

Where  $\mu(t, X_t)$  is known as the drift function and  $\sigma(t, X_t)$  is known as the volatility function. **You may think  $dX_t \approx X_{t+\delta t} - X_t$  and  $dt \approx \delta t$  (useful in simulation)**

#### Stochastic integral

Definition:  $\int_0^T f(s, W_s) dW_s = \lim_{\delta(\pi) \rightarrow 0} \sum_{j=0}^{N(\pi)-1} f(t_j, W_{t_j})(W_{t_{j+1}} - W_{t_j})$

**Itô's lemma:**  $df(t, X_t) = \left[ \frac{\partial f}{\partial t} + \mu(t, X_t) \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2(t, X_t) \frac{\partial^2 f}{\partial x^2} \right] dt + \sigma(t, X_t) \frac{\partial f}{\partial x} dW_t$

**Geometric Brownian motion:**  $dS_t = rS_t dt + \sigma S_t dW_t \Rightarrow S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t}$

**Consequently,**  $S_{t+\delta t} = S_t e^{(r - \frac{1}{2}\sigma^2)\delta t + \sigma\sqrt{\delta t}Z}$  where  $Z \sim N(0,1)$

Finding stochastic integral: "guess" the function such that it will contain the integrand in its SDE. Use Itô's lemma to find SDE of the guess and then integrate both sides

Solving SDE: "guess" a solution and use Itô's lemma to verify that the solution satisfies the SDE (the following table is borrowed from Prof. Yau Chun Yip's notes on Stochastic Calculus)



Name	SDE	Solution ( $X_t$ )
Ornstein-Uhlenbeck(OU) process	$dX_t = -\alpha X_t dt + \sigma dW_t$	$ce^{-\alpha t} + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s$
Mean reverting OU	$dX_t = (m - \alpha X_t) dt + \sigma dW_t$	$\frac{m}{\alpha} + \left(c - \frac{m}{\alpha}\right) e^{-\alpha t} + \sigma \int_0^t e^{\alpha(s-t)} dW_s$
Geometric Brownian motion	$dX_t = aX_t dt + bX_t dW_t$	$ce^{(a-b^2/2)t + bW_t}$
Brownian bridge	$dX_t = \frac{b-X_t}{1-t} dt + dW_t$	$a(1-t) + bt + (1-t) \int_0^t \frac{dW_s}{1-s}$
	$dX_t = \left(\sqrt{1+X_t^2} + \frac{1}{2}X_t\right) dt + \sqrt{1+X_t^2} dW_t$	$\sinh(c+t+W_t)$
	$dX_t = X_t^3 dt + X_t^2 dW_t$	$\frac{1}{c-W_t}$
	$dX_t = -\frac{1}{2}X_t dt + \sqrt{1-X_t^2} dW_t$	$\sin(c+W_t)$
	$dX_t = -\frac{1}{1+t}X_t dt + \frac{1}{1+t} dW_t$	$(c+W_t)/(1+t)$
	$dX_t = rdt + \alpha X_t dW_t$	$ce^{\alpha W_t - \frac{1}{2}\alpha^2 t} + r \int_0^t e^{\alpha(W_t-W_s) - \frac{1}{2}\alpha^2(t-s)} ds$

Integrating factor: add  $e^{rt}$  to both sides of a SDE (target: cancel some terms)

Martingale property:  $E \left[ \int_0^T f(t, W_t) dW_t \middle| \mathcal{F}_s \right] = \int_0^s f(t, W_t) dW_t$

In particular,  $E \left( \int_0^T f(t, W_t) dW_t \right) = 0$

Itô isometry:  $E \left[ \left( \int_0^T f(t, W_t) dW_t \right)^2 \right] = \int_0^T E[f(t, W_t)^2] dt$

Similarly,  $E \left[ \left( \int_0^T f(t, W_t) dW_t \right) \left( \int_0^T g(t, W_t) dW_t \right) \right] = \int_0^T E[f(t, W_t)g(t, W_t)] dt$

Product rule:  $d(X_t Y_t) = X_t dY_t + Y_t dX_t + d[\sigma(t, X_t)W_t, \bar{\sigma}(t, Y_t)\bar{W}_t]$

## IV) Simulation methods

### Theoretical support

Sample mean:  $\bar{X}_n = \sum_{i=1}^n X_i$

Sample variance:  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

Law of large numbers (WLLN): Let  $X_1, \dots, X_n$  be i.i.d. random variables with mean  $\theta$  and variance  $\sigma^2$ , then for any given  $\epsilon > 0$ ,  $P(|\bar{X}_n - \theta| > \epsilon) \rightarrow 0$  as  $n \rightarrow \infty$

Central limit theorem (CLT, Lindeberg–Lévy): Let  $X_1, \dots, X_n$  be i.i.d. random variables with mean  $\theta$  and finite variance  $\sigma^2$ , then  $\bar{X}_n \xrightarrow{d} N\left(\theta, \frac{\sigma^2}{n}\right)$  as  $n \rightarrow \infty$

### Standard Monte Carlo

Idea: take average of independent replications/scenarios of the reality/future

Algorithm:

- 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)

### Inverse transform

Idea: if we know  $X \sim F_X$  (i.e. the cdf), we can generate  $X$  out of uniform random numbers

Algorithm (discrete):

- 1) Generate  $U \sim \text{Uniform}(0,1)$
- 2)  $X = x_j$  if  $\sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^j p_i$

Algorithm (continuous):

- 1) Generate  $U \sim \text{Uniform}(0,1)$
- 2)  $X = F_X^{-1}(U)$  assuming the inverse exists

### [Rejection sampling](#)

Idea: 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  $g$ :  $U_1 \sim \text{Uniform}(0,1) \Rightarrow Y_i = G^{-1}(U_1)$
- 3) Generate  $U_2 \sim \text{Uniform}(0,1)$
- 4) If  $U_2 \leq \frac{1}{c} \cdot \frac{f(Y_i)}{g(Y_i)}$ , set  $X_i = Y_i$ , otherwise return to 2

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

## V) Variance reduction

### Antithetic variables

Idea: 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

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)  $\hat{\theta} = \frac{1}{2n} \sum_{i=1}^n [h(X_i) + h(Y_i)]$

Useful corollary: if  $h(x)$  is monotone, then  $\text{Cov}(h(U), h(1 - U)) \leq 0$  where  $U \sim U(0,1)$

### Stratified sampling

Idea: 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:

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