Stochastic Simulations

Autumn Semester 2024

Prof. Fabio Nobile *Assistant: Matteo Raviola*

Lab 04 – 3 October 2024

Stochastic process generation

Exercise 1

Let $X = [X_1, X_2, ..., X_n] \overset{\text{i.i.d}}{\sim} \mathcal{U}([-1,1]^n)$ be a random vector uniformly distributed over the n-dimensional square $\Gamma = [-1,1]^n$, and define the random variable $Z = \mathbb{1}_{\|X\|_{l^2}<1}$. Observe that

$$I = \mathbb{E}[Z] = \int_\Gamma \mathbb{1}_{\|x\|_{l^2}<1} p(x)\mathrm{d}x = \frac{1}{|\Gamma|}\left|B(0,1)\right|,$$

where $p(x)$ is the PDF of $\mathcal{U}([-1,1]^n)$, and $|B(0,1)|$ is the volume of the n-dimensional sphere with center 0 and radius 1.

1. Let $n = 2$. Use Monte Carlo to approximate the value of I:

$$\overline{I}_N := \frac{1}{N}\sum_{k=1}^{N} Z_k,$$

   For $N = 10, 100, 1000, 10000$, compute $\overline{I}_N$ as well as an approximate confidence interval and compare with the exact value I. In addition, plot the relative error $\frac{|\overline{I}_N - I|}{I}$ versus N in logarithmic scale and verify the convergence rate.

2. (On the choice of N). By *a priori* analysis (knowing that $Z \sim$ Bernoulli(p) with $p = \pi/4$), determine three lower bounds for $N(\alpha, \epsilon)$ with $\epsilon = 10^{-2}$ and $\alpha = 10^{-4}$ for ensuring that

$$\mathbb{P}\left(\left|\overline{I}_N - \pi/4\right| > \epsilon\right) < \alpha$$

   using Chebycheff's inequality (rigorous), the Berry-Esseen Theorem (rigorous) and the leap of faith

$$\frac{\overline{I}_N - \pi/4}{\sqrt{\mathrm{Var}(Z)/N}} \sim N(0,1).$$

   Discuss the advantages and disadvantages of using each bound.

3. An important property of the MC method is that, under very weak regularity assumptions, an $O(N^{-1/2})$ convergence rate holds independently of the dimensionality of the underlying problem. To illustrate this, consider approximating $\mathbb{E}[Z]$ as in the first point, for $n = 6$.

Exercise 2

Note: Refer to Section 4.4 of the lecture notes.
Consider the chemical reactions between three species S_1, S_2, S_3, which are determined by the following four reaction channels:

$$\begin{aligned} S_1 &\xrightarrow{c_1} \emptyset \ , \\ S_1 + S_1 &\xrightarrow{c_2} S_2 \ , \\ S_2 &\xrightarrow{c_3} S_1 + S_1 \ , \\ S_2 &\xrightarrow{c_4} S_3 \ . \end{aligned}$$

To simulate this system, consider the process $\boldsymbol{N}_t = (N_t^1, N_t^2, N_t^3) \in \mathbb{N}_0^3$, where N_t^i denotes the number of molecules of species S_i at time $t \geq 0$. In fact, this process is a time-continuous Markov chain with transition probabilities given by

$$\begin{aligned} \mathbb{P}\big(\boldsymbol{N}_{t+h} = \boldsymbol{N}_{t,1} = (N^1 - 1, N^2, N^3) \big| \boldsymbol{N}_t = = (N^1, N^2, N^3)\big) &= a_1(\boldsymbol{N}_t)h + o(h) \ , \\ \mathbb{P}\big(\boldsymbol{N}_{t+h} = \boldsymbol{N}_{t,2} = (N^1 - 2, N^2 + 1, N^3) \big| \boldsymbol{N}_t = (N^1, N^2, N^3)\big) &= a_2(\boldsymbol{N}_t)h + o(h) \ , \\ \mathbb{P}\big(\boldsymbol{N}_{t+h} = \boldsymbol{N}_{t,3} = (N^1 + 2, N^2 - 1, N^3) \big| \boldsymbol{N}_t = (N^1, N^2, N^3)\big) &= a_3(\boldsymbol{N}_t)h + o(h) \ , \\ \mathbb{P}\big(\boldsymbol{N}_{t+h} = \boldsymbol{N}_{t,4} = (N^1, N^2 - 1, N^3 + 1) \big| \boldsymbol{N}_t = (N^1, N^2, N^3)\big) &= a_4(\boldsymbol{N}_t)h + o(h) \ , \\ \mathbb{P}\big(\boldsymbol{N}_{t+h} = \boldsymbol{N}_{t,5} = (N^1, N^2, N^3) \big| \boldsymbol{N}_t = (N^1, N^2, N^3)\big) &= 1 - h\sum_{j=1}^{4} a_j(\boldsymbol{N}_t) + o(h) \ , \end{aligned}$$

for h sufficiently small, where $\boldsymbol{N}_{t,k}, k \in \{1, ..., 5\}$ indexes the possible transitions. Here, the so-called propensity functions are

$$a_1(\boldsymbol{N}) = c_1 N^1 \ , \quad a_2(\boldsymbol{N}) = c_2 \frac{N^1(N^1 - 1)}{2} \ , \quad a_3(\boldsymbol{N}) = c_3 N^2 \ , \qquad a_4(\boldsymbol{N}) = c_4 N^2 \ ,$$

with $\boldsymbol{N} = (N^1, N^2, N^3)$.

1. Try to construct the transition matrix corresponding to the above transition probabilities and note the challenges. Is it possible to simulate the chemical reaction without the explicit Q matrix? **Hint:** Think back to how you simulated the process in Exercise 2.1.

2. Utilise the following algorithm to simulate the chemical reaction system. Plot a time series for each species' number of molecules for $t \in [0, T]$, $T = 0.2$, for the reaction rates

$$c_1 = 1 \ , \quad c_2 = 5 \ , \quad c_3 = 15 \ , \quad c_4 = \frac{3}{4} \ ,$$

using $\boldsymbol{N}_0 = (400, 800, 0)$ as initial number of molecules. Repeat the simulation for the same reaction rates $c_1, \ldots, c_4$ also for $T = 5$.

Exercise 3

Let $\{N_t \in \mathbb{N}_0 \colon t \geq 0,\ N_0 = 0\}$ be a Poisson process with rate λ.

Algorithm 1: Reaction simulation

1: Set $\boldsymbol{N}_0 = (N_0^1, N_0^2, N_0^3)$, $J_0 = 0$
2: **for** $n = 1, 2, ...$ **do**
3: Compute $\lambda = \sum_{j=1}^{4} a_j(\boldsymbol{N}_{J_{n-1}})$
4: Generate $S_n \sim \text{Exp}(\lambda)$ and set $J_n = J_{n-1} + S_n$
5: Generate $I \in \{1, 2, 3, 4\}$ with probability mass function $\mathbb{P}(I = j) = \frac{a_j(\boldsymbol{N}_{J_{n-1}})}{\sum_{l=1}^{4} a_l(\boldsymbol{N}_{J_{n-1}})}$, which is the probability that the j^{th} reaction happens.
6: Set $\boldsymbol{N}_t = \boldsymbol{N}_{J_{n-1}} \forall t \in [J_{n-1}, J_n)$ and $\boldsymbol{N}_{J_n} = \boldsymbol{N}_{t,I}$
7: **end for**

1. Show that, conditional on the event $\{N_T = n\}$, the jump times $J_1, \ldots, J_n$ have joint density function

$$f_{J_1,\ldots,J_n}(j_1, \ldots, j_n) = n! T^{-n} \mathbb{I}(0 \leq j_1 \leq \cdots \leq j_n \leq T) \ .$$

In other words, show that conditional on $\{N_T = n\}$, the jump times $J_1, \ldots, J_n$ have the same distribution as an ordered sample of size n from the uniform distribution on $[0, T]$.

Hints: *Use the joint distribution of the holding times $S_1, \ldots, S_{n+1}$ to first derive the joint distribution of the jump times, where $S_{i+1} = J_{i+1} - J_i$. Then compute the conditional distribution of the jump times given that $N_T = n$, using the fact that $\{N_T = n\} = \{J_n \leq T < J_{n+1}\}$ a.s.*

2. Use the property above to propose an algorithm to generate the process N_t, $t \in (t_1, t_2)$, conditional upon $N_{t_1} = n_1$ and $N_{t_2} = n_2 > n_1$. Such a process is called *Poisson bridge*.

Exercise 4

Let $\{N_t, t \geq 0, N_0 = 0\}$ be a non-homogeneous Poisson process with rate $\lambda : [0, \infty) \mapsto \mathbb{R}_+$. In addition, define $\Lambda(t) = \int_0^t \lambda(s) \mathrm{d}s$, and let $\{\tilde{N}_t, t \geq 0, \tilde{N}_0 = 0\}$ be a homogeneous Poisson process with rate one.

1. Show that the non-homogeneous Poisson process can be obtained as $N_t = \tilde{N}_t \circ \Lambda(t)$, i.e., $N_t = \tilde{N}_{\Lambda(t)}$.

2. Simulate a non-homogeneous Poisson process with rate function $\lambda(t) = \sin^2(t)$ on the interval $[0, 50]$.

Exercise 5 (Optional)

1. Generate a random walk $\{X_n \in \mathbb{Z},\ n \in \mathbb{N}_0,\ X_0 = 0\}$ with transition probabilities

$$\mathbb{P}(X_{n+1} = j | X_n = j-1) = \mathbb{P}(X_{n+1} = j | X_n = j+1) = a \ , \quad \mathbb{P}(X_{n+1} = j | X_n = j) = 1-2a \ ,$$

for some $0 < a \leq 1/2$.

2. Consider the rescaled process $Y_{t_i} := \sqrt{\Delta t/(2a)}X_i$ for $i = 0, \ldots, n$ with $t_i = i\Delta t$. Compare this process with the process W_{t_i}, $i = 0, \ldots, n$, where W_t denotes a Wiener process with $W_0 = 0$. That is, show that both processes "look similar" in the limit as $\Delta t \to 0$ by plotting multiple realizations of both processes for $n = \lceil 1/\Delta t \rceil$.

3. (**Optional:**) More theoretical analysis of the observed phenomenon:

   (a) Consider the spatial mesh $x_m = m\Delta x = m\sqrt{\Delta t/(2a)}$ for $m \in \mathbb{Z}$ and the following notation for the rescaled process' probability mass function at time t_i:

   $$\bar{u}(t_i, x_m) := \mathbb{P}(Y_{t_i} = x_m | Y_0 = 0), \quad m \in \mathbb{Z}, i = 0, 1, \ldots$$

   Use the discrete Chapman–Kolmogorov formula

   $$\mathbb{P}(Y_{t_{i+1}} = x_m | Y_0 = 0) = \sum_k \mathbb{P}(Y_{t_{i+1}} = x_m | Y_{t_i} = x_k)\mathbb{P}(Y_{t_i} = x_k | Y_0 = 0) \qquad (1)$$

   to derive a difference equation for $\bar{u}(t_{i+1}, x_m)$ in terms of $\bar{u}(t_i, \cdot)$.

   (b) Show that the difference equation obtained in 3a corresponds to a finite difference approximation of the one dimensional heat equation

   $$u_t(t, x) = \frac{u_{xx}(t, x)}{2}, \quad x \in \mathbb{R}, t > 0,$$

   on a uniform grid $x_i = i\Delta x$ and $t_j = j\Delta t$ with $\Delta t = 2ax^2$, using a second order centered finite difference stencil in space and a first order forward Euler scheme in time.

   (c) For the standard Wiener process with $\mathbb{P}(W_0 = 0) = 1$, we denote the probability density function at time $t > 0$ by

   $$u(t, x) := \frac{e^{-x^2/(2t)}}{\sqrt{2\pi t}}, \qquad x \in \mathbb{R}.$$

   For all $t > 0$ and $x \in \mathbb{R}$, show that the density satisfies the same heat equation introduced in point 3b.