Stochastic Simulations

Autumn Semester 2024

Prof. Fabio Nobile *Assistants: Matteo Raviola*

Lab 06 – 17 October 2024

Variance Reduction Techniques (cont.)

Exercise 1

Consider the discrete time random walk $\{X_n \in \mathbb{Z} \colon X_0 = 0, n \in \mathbb{N}\}$ with transition probabilities:

$$\begin{aligned} p_{i,i+1} &\equiv \mathbb{P}(X_{n+1} = i+1 | X_n = i) = a, \\ p_{i,i-1} &\equiv \mathbb{P}(X_{n+1} = i-1 | X_n = i) = 1-a, \quad n \geq 0,\ i \in \mathbb{Z},\ a \in (0,1), \end{aligned}$$

and define the stopping time $\tau_N := \inf\{n \colon X_n = K\}$ for a given constant $K \in \mathbb{N}$. We aim at estimating $\mathbb{P}(\tau_K < T)$, for some given $T \in \mathbb{N}$.

1. Set $K = 4$, $a = 1/3$, $T = 10$. Compute a Monte Carlo estimate of $\mathbb{P}(\tau_K < T)$.
2. For the same values as in the previous point, estimate $\mathbb{P}(\tau_K < T)$ using the antithetic variate variance reduction technique and compare your results to those in point 1.

Exercise 2

Suppose that we want to compute $p = \mathbb{P}(\boldsymbol{X} \in A)$, where $\boldsymbol{X}$ is a d-dimensional Gaussian random vector with mean $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix Σ. If the (Borel) set $A \subseteq \mathbb{R}^d$ contains the mean $\boldsymbol{\mu}$, then the event $\boldsymbol{X} \in A$ is typically not rare, and the use of importance sampling is generally unnecessary. If, on the other hand, $\boldsymbol{\mu} \notin A$ and p is small, then one may wish to consider the use of importance sampling. Let $\mathbb{P}^*$ denote the optimal (yet impractical) sampling measure with density g^*, that is

$$d\mathbb{P}^* = g^*\, d\boldsymbol{x} = \frac{1}{\mathbb{P}(\boldsymbol{X} \in A)} \mathbb{I}_A \phi_\Sigma \, d\boldsymbol{x} \ ,$$

where ϕ_Σ is the density of the $\mathcal{N}(\mathbf{0}, \Sigma)$ distribution. Given the rapid decay of $\phi_\Sigma(\boldsymbol{x})$ as $\|\boldsymbol{x}\|_2 \to \infty$, most of the mass of $\mathbb{P}^*$ is typically located at the maximizer $\boldsymbol{x}^*$ of ϕ_Σ over A (which we assume to exist uniquely). This suggests using an importance sampling distribution $\tilde{\mathbb{P}}$ with density g that concentrates most of its mass near $\boldsymbol{x}^*$, that makes g easily computable, and from which realizations can efficiently be generated.

1. One such importance sampling distribution is the Gaussian distribution centered in $\boldsymbol{x}^*$ with covariance matrix Σ. Describe the importance sampling algorithm to estimate p.

2. Implement your algorithm for $d = 2$; take $A = \{\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2 \colon x_i \geq a,\ i = 1, 2\}$, and
$$\Sigma = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix} .$$
Then carry out the following points for $a = 1, 3, 10$:
   (a) First, try to provide simulation estimates of $p = \mathbb{P}(\boldsymbol{X} \in A)$ and the associated 95% confidence interval using the (naive) crude Monte Carlo method.
   (b) Next, find the point $\boldsymbol{x}^*$ that maximizes the $\mathcal{N}(\mathbf{0}, \Sigma)$ density over A and repeat point (a), with the crude Monte Carlo method replaced by importance sampling, where the importance distribution is $\mathcal{N}(\boldsymbol{x}^*, \Sigma)$.
   (c) In point (b), experiment with importance distributions of the form $\mathcal{N}(\boldsymbol{x}^*, \delta\Sigma)$ for different $\delta > 0$.

Exercise 3

Suppose that we wish to compute $\mu = \mathbb{E}\big[\Psi_\tau(X_0, \ldots, X_\tau)\mathbb{I}_{\{\tau<\infty\}}\big]$, where τ is a stopping time adapted to the discrete time, discrete state Markov chain $\{X_n \in \mathbb{Z}^d \colon n \in \mathbb{N}_0\}$ with initial probability distribution p_0 (i.e. $X_0 \sim p_0$) and with Markov transition probabilities $p_{i,j} \in [0, 1]$ such that
$$p_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i) \ , \forall\, i, j \in \mathbb{Z}^d \ .$$
Instead of relying on the evolution of this Markov chain, it is natural to try to use importance measures that preserve the Markov property (so as to guarantee that the paths can be simulated efficiently under the importance measure). That is, one replaces the transition probabilities $p_{i,j}$ by other Markov transition probabilities $q_{i,j} \in [0, 1]$, which dominate $p_{i,j}$ (i.e. $q_{i,j} = 0 \Rightarrow p_{i,j} = 0$); analogously for the initial distribution. Moreover, we require that the stopping time τ is almost surely finite for the Markov process with transition probabilities $q_{i,j}$, which appears natural from a practical point of view.

1. Prove that $\mu = \mathbb{E}\big[\Psi_\tau(X_0, \ldots, X_\tau)\mathbb{I}_{\tau<\infty}\big]$ can be written as:
$$\mu = \mathbb{E}_q\big[\Psi_\tau(X_0, \ldots, X_\tau) w(X_0, \ldots, X_\tau)\big] \ , \quad \text{with} \quad w(X_0, \ldots, X_m) = \frac{p_0(X_0)}{q_0(X_0)} \prod_{j=1}^{m} \frac{p_{X_{j-1},X_j}}{q_{X_{j-1},X_j}} \ .$$
2. Implement the importance sampling algorithm for discrete time Markov processes to the random walk $\{X_n \in \mathbb{Z} \colon X_0 = 0, n \in \mathbb{N}\}$ with transition probabilities:
$$p_{i,i+1} \equiv \mathbb{P}(X_{n+1} = i{+}1 | X_n = i) = \frac{1}{2} = \mathbb{P}(X_{n+1} = i{-}1 | X_n = i) \equiv p_{i,i-1} \ , \quad n \geq 0,\, i \in \mathbb{Z} \ .$$
Consider the stopping time $\tau_N := \inf\{n \colon X_n = N\}$ for a given constant $N = 4$ and apply the algorithm to estimate $\mathbb{P}\big(\tau_N < T\big)$ with $T = 10$. For the the importance sampling, consider the random walk with transition probabilities
$$q_{i,i+1} \equiv \mathbb{P}(X_{n+1} = i + 1 | X_n = i) = \alpha > 1 - \alpha = \mathbb{P}(X_{n+1} = i - 1 | X_n = i) \equiv q_{i,i-1} \ .$$
Experiment with different values for α.