Stochastic Simulation
Professor: Prof. Fabio Nobile
Assistant: Eliott Van Dieren

Week of 2025-10-27 to 2025-10-31

Lab 7 of Thursday 30th October 2025

Exercice 1.

Suppose we are given a control variate Y with known mean $\mathbb{E}(Y)$ and consider the usual modified random variable

$$\tilde{Z}_\alpha = Z + \alpha(Y - \mathbb{E}(Y)) ,$$

from which we aim at estimating $\mu = \mathbb{E}(Z)$. In fact, here we consider the following *one-shot algorithm* for estimating μ:

- Generate N i.i.d. replicas $(Z^{(i)}, Y^{(i)})$, $i = 1, \dots, N$.
- Estimate α_{opt} by $\hat{\alpha}_{\text{opt}} := -\hat{\sigma}^2_{Z,Y}/\hat{\sigma}^2_Y$, using the usual unbiased mean, variance, and covariance estimators based on the sample $(Z^{(i)}, Y^{(i)})_{i=1,\dots,N}$.
- Compute the control variate estimator of μ as
$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\Big(Z^{(i)} + \hat{\alpha}_{\text{opt}}(Y^{(i)} - \mathbb{E}(Y))\Big) .$$

1) Show that the estimator $\hat{\mu}$ is asymptotically normally distributed, in the sense that

$$\sqrt{N}\frac{\hat{\mu}-\mu}{\sigma_{\text{opt}}} \underset{N\to\infty}{\Rightarrow} \mathcal{N}(0,1) , \quad \text{where} \quad \sigma_{\text{opt}} = \sqrt{\mathrm{Var}(\tilde{Z}_{\alpha_{\text{opt}}})} .$$

Furthermore, explain why the asymptotic normality also holds when σ_{opt} is replaced by the usual empirical standard deviation based on a sample of realizations of $\tilde{Z}_{\alpha_{\text{opt}}}$.

Hint. Consider re-writing the estimator as the summation of the control variate estimator computed with the exact α_{opt} and a correction term involving $\hat{\alpha}_{\text{opt}} - \alpha_{\text{opt}}$ as follows:

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\Big(Z^{(i)} + \alpha_{\text{opt}}(Y^{(i)} - \mathbb{E}(Y))\Big) + (\hat{\alpha}_{\text{opt}} - \alpha_{\text{opt}})\Big(\frac{1}{N}\sum_{i=1}^{N} Y^{(i)} - \mathbb{E}(Y)\Big).$$

Then, recall Slutsky's theorem, which states that if ξ_n converges in distribution to ξ and η_n converges in probability to a constant c, then $f(\xi_n, \eta_n) \underset{n\to\infty}{\Rightarrow} f(\xi, c)$ for any continuous function $f\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Here, the symbol $\Rightarrow$ denotes convergence in distribution.

2) Implement the one-shot algorithm described above. Apply it to the examples considered in Lab 06, Exercise 1. That is, approximate the probability $p = \mathbb{P}(\boldsymbol{X} \in A)$ for the sets $A = \{\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2 \colon x_i \geqslant a,\ i = 1, 2\}$ with $a = 1, 3, 10$. Here, $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$ with $\Sigma = \left(\begin{smallmatrix} 4 & -1 \\ -1 & 4 \end{smallmatrix}\right)$.

a) First, explain why $Y = \mathbb{I}_{\{X_1+X_2 \geqslant 2a\}}$ for $\boldsymbol{X} = (X_1, X_2)$ could be a decent control variate for this problem.

b) Then perform simulations and investigate the variance reduction effect for the control variate Y. Moreover, use the result proved in point 1 to compute asymptotic 95% confidence intervals.

c) Can you think of other appropriate control variates?

Exercice 2.

Suppose we wish to construct a Brownian motion path $\{B_t \colon B_0 = 0,\ 0 \leqslant t \leqslant T\}$ that finishes at $T = 1$ in S distinct strata. To stratify standard Brownian motion on its endpoint, one first generates the S values at $T = 1$ and then samples the Brownian paths on the interval $[0, T]$ conditional upon these stratified terminal values.

1) Implement an algorithm that generates stratified standard Brownian motion using S equiprobable strata. Specifically, for each stratum Ω_j, $j = 1, \dots, S$, your algorithm should produce N_j stratified Brownian samples paths $B_{t_m}^{(i,j)}$, $i = 1, \dots, N_j$, evaluated in the discrete times $t_m = m/M$ with $m = 1, \dots, M \in \mathbb{N}$. Test your implementation for $S = 12$, $M = 1000$, and $N_j = 2$ by plotting the stratified samples paths. *Hint: Brownian bridge sampling.*

2) Consider the geometric Brownian motion process X_t that solves

$$dX = rX\,dt + \sigma X\,dW\,, \quad X(0) = X_0\,,$$

and which is given by $X_t = X_0 e^{Y_t}$, where $Y_t = (r - \sigma^2/2)t + \sigma W_t$ with W being a standard Wiener process. For $M \in \mathbb{N}$, let

$$\Psi(X_{t_0}, \dots, X_{t_M}) = \max_{0 \leqslant m \leqslant M} X_{t_m} - \min_{0 \leqslant m \leqslant M} X_{t_m}\,,$$

where $t_m = m/M$ as before. We want to estimate $\mu = \mathbb{E}[\Psi(X_{t_0}, \dots, X_{t_M})]$ for $X_0 = 6$, $r = 0.05$, $\sigma = 0.3$, and $M = 100$. Use your procedure developed in point 1 to estimate μ using stratified sampling with $S = 10$ strata. Moreover, compute the total number of samples N such that the asymptotic 99% confidence interval is smaller than $2\,tol$ for $tol = 10^{-2}, 10^{-3}$. Investigate both proportional and optimal sampling allocation in each strata.

Remark: The function Ψ is related to the value of a look-back option whose payoff is equivalent to buying at the minimum and selling at the maximum price on the time interval $[0, T = 1]$. As given here, Ψ omits the (constant) discount factor e^{-rT} that compensates for waiting until time T to collect the payoff.

3) Repeat the previous point, but now consider only $N_j = 2$ samples per stratum Ω_j, $j = 1, \dots, S$, and investigate the estimator's variance decay as a function of S.

Exercice 3.

Consider the problem of estimating $\mu = \mathbb{E}(Z)$ for $Z = \psi(X)$ and $X \sim U(0, 1)$.

1) Show that the *randomized midpoint quadrature* estimator

$$\hat{\mu}_S := \frac{1}{S}\sum_{j=1}^{S}\psi\left(\frac{j-1+U_j}{S}\right) ,$$

with $U_1, \dots, U_S$ i.i.d. $U(0,1)$, corresponds to a stratified sampling estimator of μ. *Hint: consider uniform strata.*

Optional: explain why $\hat{\mu}_S$ is called *randomized midpoint quadrature* estimator.

2) Suppose that $\psi \in C^1([0,1])$ Show that the estimator $\hat{\mu}_S$, which is a Monte Carlo type estimator, converges with *super-canonical rate* (i.e. faster than $S^{-1/2}$). Specifically, show that

$$\sqrt{\mathbb{E}[(\mu-\hat{\mu}_S)^2]} \leqslant cS^{-3/2} ,$$

for an appropriate positive constant $c < \infty$ independent of S. Determine also the constant c.

Exercice 4.

1) Consider the random variable $Z = 4\,\mathbb{I}_{\{U_1^2+U_2^2\leqslant 1\}}$ with $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \mathcal{U}(0,1)$, so that $\mathbb{E}(Z)=\pi$. Consider the control variates $\tilde{Z}_{\alpha,i} = Z - \alpha(Y_i - \mathbb{E}(Y_i))$ where the controls Y_i are given by:

$$Y_1 := \mathbb{I}_{\{U_1+U_2\leqslant 1\}} , \quad Y_2 := \mathbb{I}_{\{U_1+U_2\geqslant\sqrt{2}\}} , \quad \text{and} \quad Y_3 := (U_1+U_2-1)\mathbb{I}_{\{1<U_1+U_2\leqslant\sqrt{2}\}} .$$

Estimate their respective expected variance reduction $\mathbb{V}\text{ar}(\tilde{Z}_{\alpha,i})/\mathbb{V}\text{ar}(Z)$ using $N = 10^6$ simulations.

2) Consider again the random variable $Z = 4\,\mathbb{I}_{\{U_1^2+U_2^2\leqslant 1\}}$ as in point 1. We now wish to use multiple control variates and compare their variance reduction to the single control variate case. Consider the control variate $\tilde{Z}_{\boldsymbol{\alpha}} = Z - \boldsymbol{\alpha}\cdot(\boldsymbol{Y}-\mathbb{E}(\boldsymbol{Y}))$ where $\boldsymbol{\alpha}\in\mathbb{R}^d$ and $\boldsymbol{Y}$ is a d-dimensional control vector. Perform simulations and report the expected variance reduction $\mathbb{V}\text{ar}(\tilde{Z}_{\boldsymbol{\alpha}})/\mathbb{V}\text{ar}(Z)$ for each of the following control vectors:

$$\boldsymbol{Y}^1 := (Y_1, Y_2)^T , \quad \boldsymbol{Y}^2 := (Y_1, Y_3)^T , \quad \boldsymbol{Y}^3 := (Y_2, Y_3)^T , \quad \text{and} \quad \boldsymbol{Y}^4 := (Y_1, Y_2, Y_3)^T .$$

Here, the random variables Y_i, $i = 1, 2, 3$ are as described in point 1.

3) Implement a one-shot control variate algorithm for the control vector with the best variance reduction.

(Optional) Exercice 5.

Let Z be a random variable of which we would like to estimate the mean value and let Y be a suitable control variate. It the mean of Y is known, we can build a Control Variate Monte Carlo estimator as

$$\hat{\mu}_{CV} = \frac{1}{N}\sum_i (Z^{(i)} - \alpha(Y^{(i)} - E[Y])), \text{ with } (Z^{(i)}, Y^{(i)}) \sim \text{i.i.d } (Z, Y) \tag{5.1}$$

Consider now that case in which $\mathbb{E}[Y]$ is not known and we need to estimate it via sampling.

1) A first idea is to estimate $\mathbb{E}[Y]$ by the sample average estimator $\widehat{\mu}_Y = \frac{1}{N}\sum_j Y^{(j)}$ using the same sample as in Eq. (5.1). Show that the resulting estimator is unbiased but its variance is not smaller than (actually equal to) the one of a crude Monte Carlo estimator on Z.

2) A second idea is to estimate $\mathbb{E}[Y]$ with an independent Monte Carlo estimator using a sample size N_Y. Let us denote by C_Z the cost of generating $Z^{(i)}$ and by C_Y the cost of generating $Y^{(i)}$, which we assume smaller than C_Z, and rename the sample size N used in Eq. (5.1) as N_Z. For a given total budget $C = N_Z(C_Z + C_Y) + N_Y C_Y$ for this control variate estimator, determine the optimal choice of N_Z and N_Y and the minimal variance achievable by the above strategy.

3) Compare then the variance obtained with that of a crude Monte Carlo estimator that uses a sample size N the exhausts the same total budget $C = NC_Z$.