MATH-414 – Stochastic simulation

Lecture 4: Monte Carlo Method

Prof. Fabio Nobile

EPFL

Outline

Setting

- Z: output of a stochastic model
- **Goal**: estimate $\mu = \mathbb{E}[Z]$
- other properties of the distribution of Z could be of interest as well (higher moments, quantiles, ...)

Assumptions:

- distribution of Z not known / not easily computable
- Z can be simulated (by simulating the stochastic process and evaluating its output)
- Typically $Z = \phi(U_1, U_2, \ldots, U_d)$ where $(U_1, \ldots, U_d)$ are all the uniform random variables used to simulate the stochastic process and ϕ represent the simulation algorithm.

Computing the expectation $\mu = \mathbb{E}[Z]$ can be seen as a high-dimensional integration problem

$$\mu = \mathbb{E}[Z] = \int_{[0,1]^d} \phi(u_1, \ldots, u_N) du_1 \ldots du_d$$

Monte Carlo method

The Monte Carlo method simply consists in

- Generating N i.i.d replicas $Z^{(1)}, \ldots, Z^{(N)}$ of Z (by simulation)
- estimating μ by a sample mean estimator

$$\hat{\mu}_N = \frac{1}{N} \sum_{i=1}^{N} Z^{(i)}$$

We assume hereafter that Z has finite second moments

$$\sigma^2 = \mathbb{V}\mathrm{ar}\left[Z\right] < \infty$$

Properties fo the Monte Carlo estimator

1. $\hat{\mu}_N$ is unbiased (i.e. $\mathbb{E}[\hat{\mu}_N] = \mu$)

$$\mathbb{E}[\hat{\mu}_N] = \frac{1}{N}\sum_{i=1}^{N} \underbrace{\mathbb{E}\left[Z^{(i)}\right]}_{=\mu,\ \forall i} = \mu$$

(expection is taken w.r.t. the joint distribution of the sample $Z^{(1)}, \ldots, Z^{(N)}$)

2. $\mathbb{V}\mathrm{ar}[\hat{\mu}_N] = \frac{\sigma^2}{N}$. Indeed:

$$\mathbb{V}\mathrm{ar}[\hat{\mu}_N] = \mathbb{E}\left[(\hat{\mu}_N - \mathbb{E}[\hat{\mu}_N])^2\right] = \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}(Z^{(i)} - \mu)\right)^2\right]$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N} \mathbb{E}\left[(Z^{(i)} - \mu)(Z^{(j)} - \mu)\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^{N} \underbrace{\mathbb{E}\left[(Z^{(i)} - \mu)^2\right]}_{=\sigma^2\ \forall i \text{ since } Z^{(i)} \text{ are iid}} + \frac{1}{N^2}\sum_{i\neq j} \underbrace{\mathbb{E}\left[(Z^{(i)} - \mu)(Z^{(j)} - \mu)\right]}_{=0 \text{ since } Z^{(i)}, Z^{(j)} \text{ are indept.}} = \frac{\sigma^2}{N}$$

Properties fo the Monte Carlo estimator

3. Almost sure convergence (from Strong Law of Large Numbers since $\mathbb{E}[Z] < \infty$)

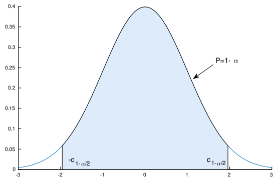
$$\hat{\mu}_N \xrightarrow{N\to\infty} \mu \quad \text{a.s.}$$

4. Asymptotic normality (from Central Limit Theorem since $\mathbb{V}\mathrm{ar}[Z] < \infty$)

$$\frac{\sqrt{N}(\hat{\mu}_N - \mu)}{\sigma} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1) \quad \text{as} \quad N \to \infty$$

Denoting c_α the α-quantile of the standard normal distribution

$$\mathbb{P}\left(\frac{\sqrt{N}|\hat{\mu}_N - \mu|}{\sigma} \le c_{1-\alpha/2}\right) \xrightarrow{N\to\infty} 1-\alpha$$



$$\Longrightarrow \qquad |\hat{\mu}_N - \mu| \le c_{1-\alpha/2}\frac{\sigma}{\sqrt{N}} \quad \text{asympt. with probability } 1-\alpha$$

EPFL

Confidence intervals

Define the asymptotic confidence interval of level $1-\alpha$

$$I_{\alpha,N} = \left[\hat{\mu}_N - c_{1-\alpha/2}\frac{\sigma}{\sqrt{N}},\ \hat{\mu}_N + c_{1-\alpha/2}\frac{\sigma}{\sqrt{N}}\right]$$

Then $\mathbb{P}\left(\mu \in I_{\alpha,N}\right) \xrightarrow{N\to\infty} 1-\alpha$

Problem: $I_{\alpha,N}$ is not directly computable (σ not known in general).
Solution: replace σ^2 with sample variance estimator

$$\hat{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z^{(i)} - \hat{\mu}_N\right)^2.$$

Since $\hat{\sigma}_N \to \sigma$ a.s. we have

$$\frac{\sqrt{N}(\hat{\mu}_N-\mu)}{\hat{\sigma}_N} = \underbrace{\frac{\sigma}{\hat{\sigma}_N}}_{\to 1 \text{ a.s.}}\ \underbrace{\frac{\sqrt{N}(\hat{\mu}_N-\mu)}{\sigma}}_{\xrightarrow{\text{d}}\mathcal{N}(0,1)} \xrightarrow{\text{d}} \mathcal{N}(0,1).$$

Computable (approximate) asymptotic confidence interval

$$\hat{I}_{\alpha,N} = \left[\hat{\mu}_N - c_{1-\alpha/2}\frac{\hat{\sigma}_N}{\sqrt{N}},\ \hat{\mu}_N + c_{1-\alpha/2}\frac{\hat{\sigma}_N}{\sqrt{N}}\right]$$

Adaptive Monte Carlo

Given an estimate $\hat{\sigma}^2$ of $\mathbb{V}\mathrm{ar}[Z]$ and the CLT, one can estimate the smallest sample size N needed to achieve a given tolerance *tol* with confidence level $1-\alpha$

$$N \geq \left(\frac{c_{1-\alpha/2}\hat{\sigma}}{tol} \right)^2$$

This suggest the following two stages algorithm

Algorithm: Two stages Monte Carlo.

Given: *tol*, α

1 Pilot run with $\bar{N}$ replicas $(Z^{(1)}, \dots, Z^{(\bar{N})})$; compute

$$\hat{\mu}_{\bar{N}} = \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} Z^{(i)}, \qquad \hat{\sigma}^2_{\bar{N}} = \frac{1}{\bar{N}-1} \sum_{i=1}^{\bar{N}} (Z^{(i)} - \hat{\mu}_{\bar{N}})^2$$

2 Set $N = \lceil \frac{c^2_{1-\alpha/2}\hat{\sigma}^2_{\bar{N}}}{tol^2} \rceil$

3 Generate a new sample $(Z^{(1)}, \dots, Z^{(N)})$

4 Output $\hat{\mu}_N$ and $\hat{I}_{\alpha,N}$.

One can further check that $|\hat{I}_{\alpha,N}| \leq 2tol$. If not, repeat the procedure.

Adaptive Monte Carlo

Alternatively one can add one new sample at the time until a stopping criterion is met

---

Algorithm: Sequential Monte Carlo.

Given: *tol*, α

Do a pilot run with $\bar{N}$ replicas $(Z^{(1)}, \ldots, Z^{(\bar{N})})$ and compute $\hat{\mu}_{\bar{N}}$, $\hat{\sigma}^2_{\bar{N}}$.

Set $N = \bar{N}$, $\hat{\mu}_N = \hat{\mu}_{\bar{N}}$, $\hat{\sigma}_N = \hat{\sigma}_{\bar{N}}$.

while $\dfrac{\hat{\sigma}_N c_{1-\alpha/2}}{\sqrt{N}} > tol$ **do**

  set $N = N + 1$

  generate $Z^{(N)}$ independent of $Z^{(i)}$, $i < N$

  recompute $\hat{\mu}_N$, $\hat{\sigma}^2_N$

end

Output $\hat{\mu}_N$ and $\hat{I}_{\alpha,N}$.

Stable update rules for $\hat{\mu}_N$ and $\hat{\sigma}^2_N$:

$$\hat{\mu}_{N+1} = \frac{N}{N+1}\hat{\mu}_N + \frac{1}{N+1} Z^{(N+1)}, \qquad \hat{\sigma}^2_{N+1} = \frac{N-1}{N}\hat{\sigma}^2_N + \frac{1}{N+1}\left(Z^{(N+1)} - \hat{\mu}_N\right)^2$$

It can be shown that $\lim_{tol \to 0} \mathbb{P}\left(|\hat{\mu}_{N(tol)} - \mu| \leq tol\right) = 1 - \alpha$

Non asymptotic error bound – Chebyshev

CLT gives only an asymptotic result for $N \to \infty$. For small sample sizes, other more robust bounds can be used.

Bound based on Chebyshev inequality $\mathbb{P}\left(|Y - \mathbb{E}[Y]| > a\right) \leq \frac{\mathbb{V}\mathrm{ar}[Y]}{a^2}$

Applied to $Y = \hat{\mu}_N$ and with $\mathbb{V}\mathrm{ar}[Y]/a^2 = \alpha$ gives

$$\mathbb{P}\left(|\hat{\mu}_N - \mu| > \frac{\sigma}{\sqrt{N\alpha}}\right) \leq \alpha$$

Computable (approximate) confidence interval of level $1-\alpha$

$$\hat{I}_{\alpha,N}^{Cheb} = [\hat{\mu}_N - \frac{1}{\sqrt{\alpha}}\frac{\hat{\sigma}_N}{\sqrt{N}},\ \hat{\mu}_N + \frac{1}{\sqrt{\alpha}}\frac{\hat{\sigma}_N}{\sqrt{N}}].$$

Compare with CLT result $\quad \hat{I}_{\alpha,N} = \left[\hat{\mu}_N - c_{1-\alpha/2}\frac{\hat{\sigma}_N}{\sqrt{N}},\ \hat{\mu}_N + c_{1-\alpha/2}\frac{\hat{\sigma}_N}{\sqrt{N}}\right]$

Notice that $c_{1-\alpha/2} \ll \frac{1}{\sqrt{\alpha}}$ for small α.

Non asymptotic error bound – Berry-Essén

The Berry-Essén bound quantifies the deviation of the cdf of $\frac{\sqrt{N}(\hat{\mu}_N-\mu)}{\sigma}$ from a standard normal cdf Φ – Requires bounded third moments

$$\sup_x \left| \mathbb{P}\left(\frac{\sqrt{N}(\hat{\mu}_N-\mu)}{\sigma} \leq x \right) - \Phi(x) \right| \leq k \frac{\mathbb{E}\left[|Z-\mu|^3\right]}{\sqrt{N}\sigma^3}, \quad (k \approx 0.4748)$$

hence

$$\mathbb{P}\left(\frac{\sqrt{N}|\hat{\mu}_N-\mu|}{\sigma} \leq x \right) \geq \underbrace{2\Phi(x) - 2k\frac{\mathbb{E}\left[|Z-\mu|^3\right]}{\sqrt{N}\sigma^3} - 1}_{\geq 1-\alpha}$$

Given estimates $\hat{\sigma}_N \approx \mathrm{std}[Z]$ and $\hat{\gamma}_N^3 \approx \mathbb{E}\left[|Z-\mu|^3\right]$, and

$$\hat{x}_\alpha : \qquad \Phi(\hat{x}_\alpha) - \frac{k\hat{\gamma}_N^3}{\sqrt{N}\hat{\sigma}_N^3} = 1 - \frac{\alpha}{2} \qquad \text{(corrected quantile)}$$

Computable confidence interval: $\hat{I}_{\alpha,N}^{BE} = [\hat{\mu}_N - \hat{x}_\alpha \frac{\hat{\sigma}_N}{\sqrt{N}},\ \hat{\mu}_N + \hat{x}_\alpha \frac{\hat{\sigma}_N}{\sqrt{N}}]$

Vector valued output

- Output of stochastic model: $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^\top$
- Goal: estimate $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{Z}] = (\mathbb{E}[Z_1], \ldots, \mathbb{E}[Z_m])^\top$

Monte Carlo estimator:

- Generate N iid replicas $\boldsymbol{Z}^{(1)}, \ldots, Z^{(N)}$ of $\boldsymbol{Z}$
- compute $\hat{\boldsymbol{\mu}}_N = \frac{1}{N}\sum_{i=1}^N \boldsymbol{Z}^{(i)}$

Assuming bounded second moments, with covariance matrix
$C = \mathbb{E}\left[(\boldsymbol{Z}-\boldsymbol{\mu})(\boldsymbol{Z}-\boldsymbol{\mu})^T\right]$

CLT: $\sqrt{N}(\hat{\boldsymbol{\mu}}_N - \boldsymbol{\mu}) \xrightarrow{\mathrm{d}} \mathcal{N}(0, C)$ and $N(\hat{\boldsymbol{\mu}}_N - \boldsymbol{\mu})^\top C^{-1}(\hat{\boldsymbol{\mu}}_N - \boldsymbol{\mu}) \xrightarrow{\mathrm{d}} \chi^2_m$

C can be replaced by sample covariance matrix
$\hat{C}_N = \frac{1}{N-1}\sum_{i=1}^N (\boldsymbol{Z}^{(i)} - \hat{\boldsymbol{\mu}}_N)(\boldsymbol{Z}^{(i)} - \hat{\boldsymbol{\mu}}_N)^\top$

Computable asymptotic confidence region of level $1-\alpha$

$$\hat{I}_{\alpha,N} = \{\boldsymbol{y} \in \mathbb{R}^m : (\hat{\boldsymbol{\mu}}_N - \boldsymbol{y})^\top \hat{C}_N^{-1}(\hat{\boldsymbol{\mu}}_N - \boldsymbol{y}) \le \frac{\chi^2_{m;1-\alpha}}{N}\}$$

where $\chi^2_{m;1-\alpha}$ is the $1-\alpha$ quantile of the χ^2_m distribution.

Delta method

- Output of stochastic model: $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^\top$
- Goal: estimate $\zeta = f(\mathbb{E}[Z_1], \ldots, \mathbb{E}[Z_m])$
  with $f : \mathbb{R}^m \to \mathbb{R}$ a smooth function

Monte Carlo estimator:

- Generate N iid replicas $\boldsymbol{Z}^{(1)}, \ldots, Z^{(N)}$ of $\boldsymbol{Z}$
- compute $\hat{\mu}_{i,N} = \frac{1}{N}\sum_{k=1}^{N} Z_i^{(k)}$
- estimate $\hat{\zeta}_N = f(\hat{\mu}_{1,N}, \ldots, \hat{\mu}_{m,N})$

Notice that in general $\hat{\zeta}$ is biased.

Error estimation can be based on first order Taylor expansion (delta method)

$$\hat{\zeta}_N - \zeta = f(\hat{\boldsymbol{\mu}}_N) - f(\boldsymbol{\mu}) = \nabla f(\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}_N - \boldsymbol{\mu}) + o(\|\hat{\boldsymbol{\mu}}_N - \boldsymbol{\mu}\|).$$

Then

$$\sqrt{N}(\hat{\zeta}_N - \zeta) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \nabla f(\boldsymbol{\mu}) C \nabla f(\boldsymbol{\mu})^\top).$$

Computable asymptotic confidence interval of level $1-\alpha$

$$\hat{I}_{\alpha,N} = [\hat{\zeta}_N - \Delta_N,\ \hat{\zeta}_N + \Delta_N], \qquad \Delta_N = \frac{c_{1-\alpha/2}}{\sqrt{N}}\sqrt{\nabla f(\hat{\boldsymbol{\mu}}_N)\hat{C}_N \nabla f(\hat{\boldsymbol{\mu}}_N)^\top}$$

EPFL