Lecture Notes II – Monte Carlo simulation

Marina Meila
mmp@stat.washington.edu

Department of Statistics
University of Washington

April 2025

MC: Calculating the expectations of a function by sampling

MC for computing an integral

MC for estimating a probability

MC for estimating a distribution

MC: Calculating the expectation of a function by sampling

Given a function $f(x)$ and a distribution F known (and its density $p(x) = F'(x)$).

- Let $\theta = \mathbb{E}[f(X)]$ be the parameter of interest

$$\theta \;=\; \mathbb{E}[f(X)] \;\equiv\; \mu_f \;\equiv\; \int_{-\infty}^{\infty} f(x)p(x)dx.$$

Idea Estimate θ by sample average $\hat{\theta}_N$

1. Sample $X_{1:N} \sim F$
2. $\hat{\theta}_N \;=\; \frac{1}{N}\sum_{i=1}^{N} f(X_i)$

Note Here we don't collect data, we sample from a **known** F

Example $f(x) = x$, $F = \exp(\lambda = 0.9)$ $\theta = \mathbb{E}[X] = \mu$, $\hat{\theta} = \hat{\mu} = \bar{X}$ sample mean

Mean and variance of $\hat{\theta}_N$

- $\mathbb{E}[\hat{\theta}_N] \;=\; \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N f(X_i)\right] \;=\; \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[f(X_i)\right] \;=\; \mu_f$ unbiased
- $\operatorname{Var}\hat{\theta}_N \;=\; \frac{1}{N}\operatorname{Var} f(X_1) \;=\; \frac{1}{N}\left(\int f^2(x)p(x)dx - \mu_f^2\right)$ $(*)$

MC for computing an integral

- **Wanted** $\mathcal{I} = \int_a^b f(x)dx$
- **Idea**

$$\mathcal{I} = (b-a)\int_a^b f(x)\frac{1}{b-a}dx \quad (1)$$

$$= Z\int_a^b f(x)\mathrm{uniform}(x)dx = Z\mu_f \quad (2)$$

- Now compute μ_f by sampling

  Algorithm

  1. Sample N samples $X_{1:N}$ $\sim$uniform$[a, b]$
  2. $\hat{\mathcal{I}} = \frac{1}{N}\sum_{i=1}^{N} f(X_i) \times (b-a)$

MC for estimating a probability

- **Wanted** $\theta =$ Probability of event $E \subset \mathbb{R}$ under a known distribution F
- **Idea**

  Algorithm
  1. Sample N samples $X_{1:N} \sim F$
  2. $\hat{\theta} = \frac{\sum_{i=1}^{N} I(X_i \in E)}{N}$

MC for estimating a distribution

- **Wanted** $\theta =$ CDF of unknown distribution F
- **Idea**

  Algorithm
  1. Get data $X_{1:n} \sim F$
  2. $\hat{F}(x) = \frac{\sum_{i=1}^{n} I(X_i \leq x)}{n}$