Introduction to Monte Carlo

Kari Heine (University of Bath)

SLS, Bath, 12th November 2018

Problem statement

We are interested in integrals of the form

$$\mathbb{E}[\varphi(X)] = \int \varphi(x)\pi(\mathrm{d}x), \tag{1}$$

where π is a probability measure defined on $(\mathbb{X}, \mathcal{X})$, $\varphi : \mathcal{X} \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R})/\mathcal{X}$-measurable.

Typically these integrals are intractable

Monte Carlo

Often it is simpler (than to evaluate $\mathbb{E}[\varphi(X)]$) to draw a sample

$$\{X^1, \ldots, X^N\} \overset{\text{iid}}{\sim} \pi$$

in which case we can straightforwardly calculate

$$\pi^N(\varphi) = \frac{1}{N} \sum_{i=1}^{N} \varphi(X^i) \approx \mathbb{E}[\varphi(X)] \tag{2}$$

There are many theoretically sound results on the approximation '$\approx$' in (2).

Validity of Monte Carlo

$$\mathbb{E}\left[\pi^N(\varphi)\right] = \mathbb{E}[\varphi(X)] \qquad \text{(Unbiased)}$$

$$\mathbb{V}\left[\pi^N(\varphi)\right] = \frac{1}{N}\mathbb{V}[\varphi(X)] \qquad \text{(Error)}$$

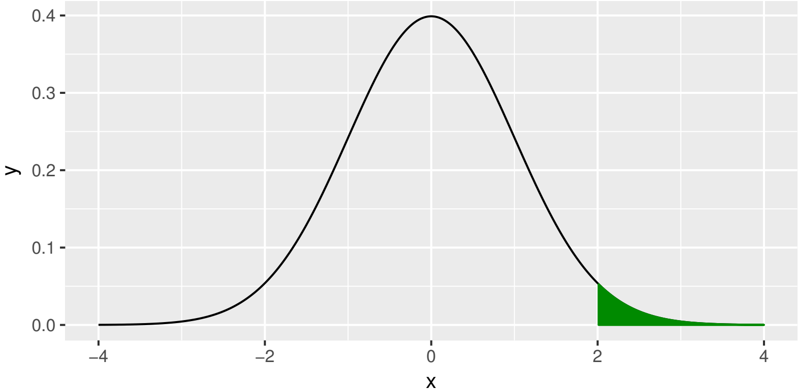
$$\pi^N(\varphi) \xrightarrow[N\to\infty]{\text{a.s.}} \mathbb{E}[\varphi(X)] \qquad \text{(Convergence)}$$

$$\sqrt{N}\left(\pi^N(\varphi) - \mathbb{E}[\varphi(X)]\right) \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \mathbb{V}[\varphi(X)])$$

(Central limit theorem)

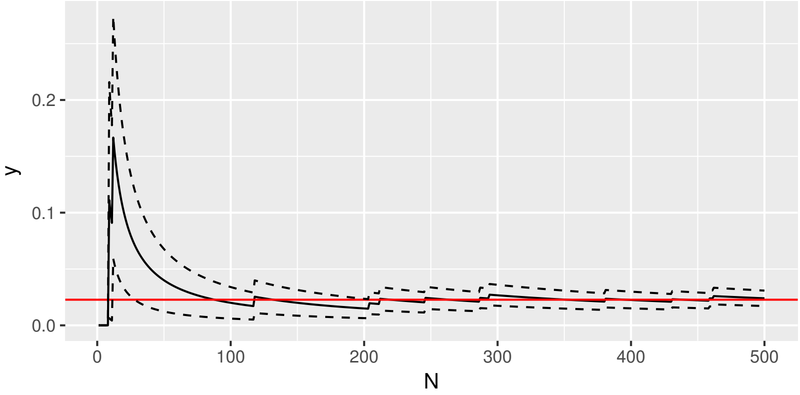
Problems with Monte Carlo

$$\mathbb{P}[X > 2] = \mathbb{E}[\mathbb{I}(X > 2)] = \int \mathbb{I}(x > 2)\pi(\mathrm{d}x), \quad \text{where} \quad X \sim \mathcal{N}(0, 1)$$



Problems with Monte Carlo

$$\mathbb{P}[X > 2] = \mathbb{E}[\mathbb{I}(X > 2)] = \int \mathbb{I}(x > 2)\pi(\mathrm{d}x), \quad \text{where} \quad X \sim \mathcal{N}(0, 1)$$



Importance sampling

Obviously

$$\mathbb{E}_\pi[\varphi(X)] = \int \varphi(x)\pi(x)\mathrm{d}x = \int \varphi(x)\frac{\pi(x)}{\gamma(x)}\gamma(x)\mathrm{d}x = \mathbb{E}_\gamma\left[\varphi(X)\frac{\pi(X)}{\gamma(X)}\right]$$

so we can construct another approximation

$$\pi_{\mathrm{IS}}^N(\varphi) = \frac{1}{N}\sum_{i=1}^N \varphi(X^i)\frac{\pi(X^i)}{\gamma(X^i)} \approx \mathbb{E}_\gamma\left[\varphi(X)\frac{\pi(X)}{\gamma(X)}\right] = \mathbb{E}_\pi[\varphi(X)],$$

where

$$\{X^1, \ldots, X^N\} \overset{\mathrm{iid}}{\sim} \gamma$$

Validity of importance sampling

We have immediately

$$\mathbb{E}\left[\pi_{\mathrm{IS}}^N(\varphi)\right] = \mathbb{E}[\varphi(X)] \qquad \text{(Unbiased)}$$

$$\pi_{\mathrm{IS}}^N(\varphi) \xrightarrow[N\to\infty]{\text{a.s.}} \mathbb{E}[\varphi(X)] \qquad \text{(Convergence)}$$

but considerations on the variance of the $\pi_{\mathrm{IS}}^N(\varphi)$ are somewhat more subtle. The variance may indeed be infinite, depending on the function π/γ.

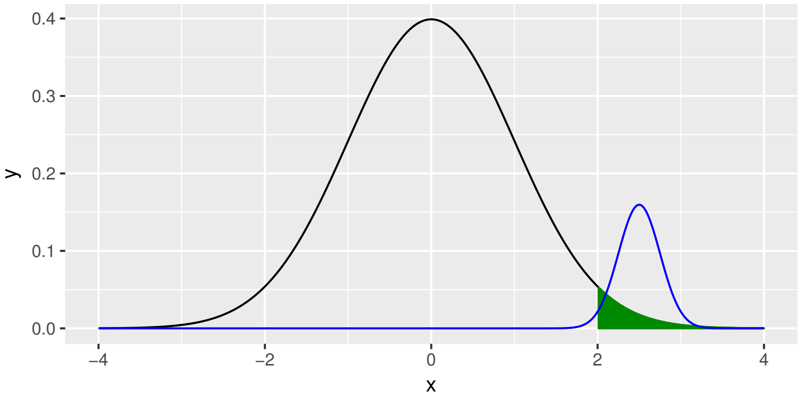
Theorem 1

The choice of γ which minimises the variance of $\pi_{\mathrm{IS}}^N(\varphi)$ is

$$\gamma(x) = \frac{|\varphi(x)|\pi(x)}{\int |\varphi(x)|\pi(x)\mathrm{d}x}$$

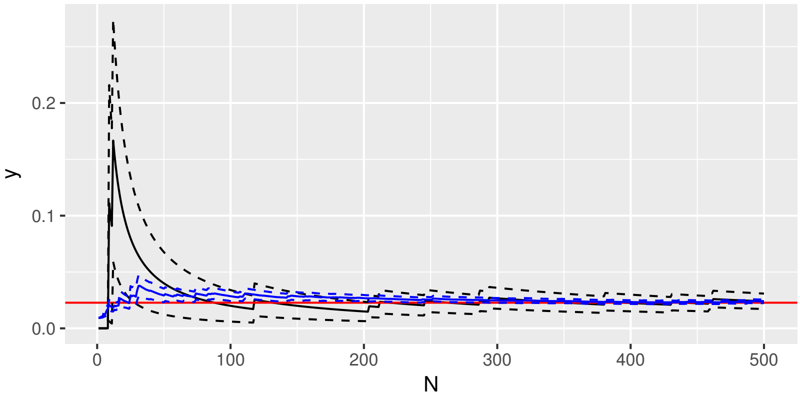
Importance sampling for the problem above

$$\mathbb{P}[X > 2] = \mathbb{E}[\mathbb{I}(X > 2)] = \int \mathbb{I}(x > 2)\pi(\mathrm{d}x), \quad \text{where} \quad X \sim \mathcal{N}(0, 1)$$



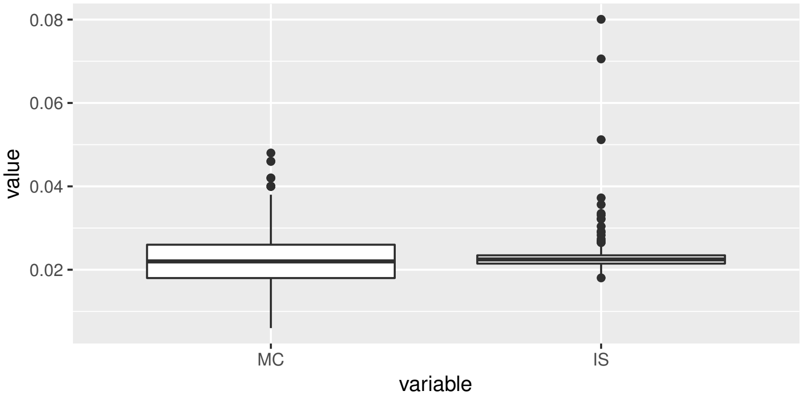
Importance sampling for the problem above

$$\mathbb{P}[X > 2] = \mathbb{E}[\mathbb{I}(X > 2)] = \int \mathbb{I}(x > 2)\pi(\mathrm{d}x), \quad \text{where} \quad X \sim \mathcal{N}(0, 1)$$

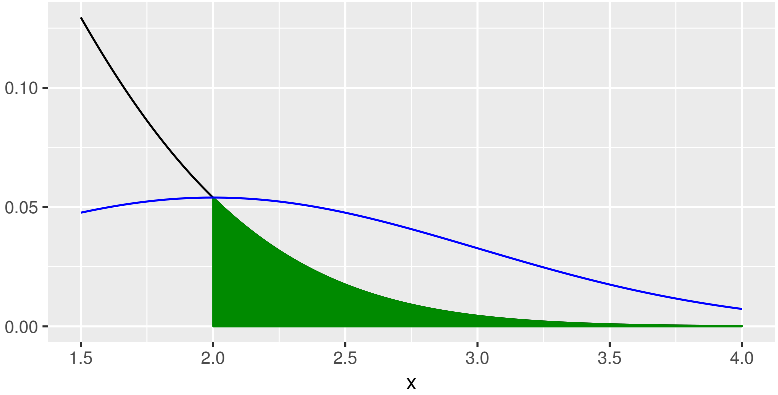


Importance sampling for the problem above

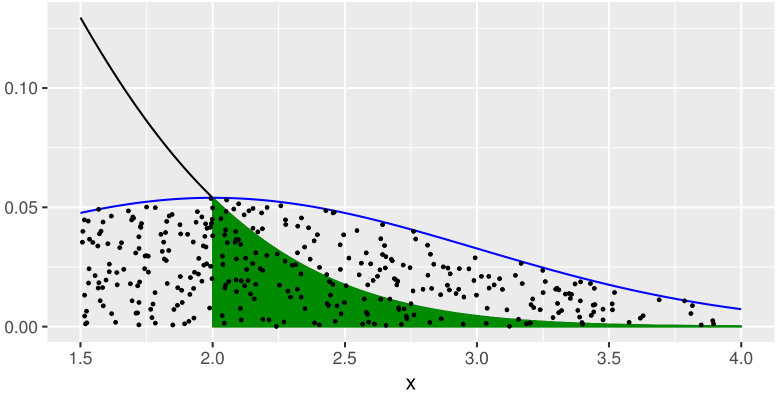
$$\mathbb{P}[X > 2] = \mathbb{E}[\mathbb{I}(X > 2)] = \int \mathbb{I}(x > 2)\pi(\mathrm{d}x), \quad \text{where} \quad X \sim \mathcal{N}(0, 1)$$



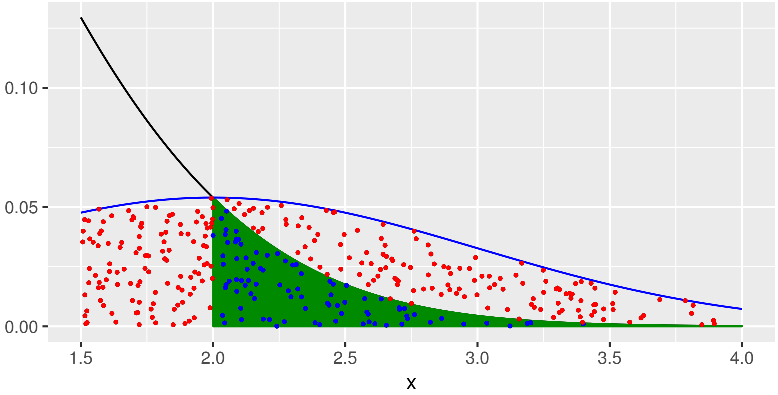
Accept reject method



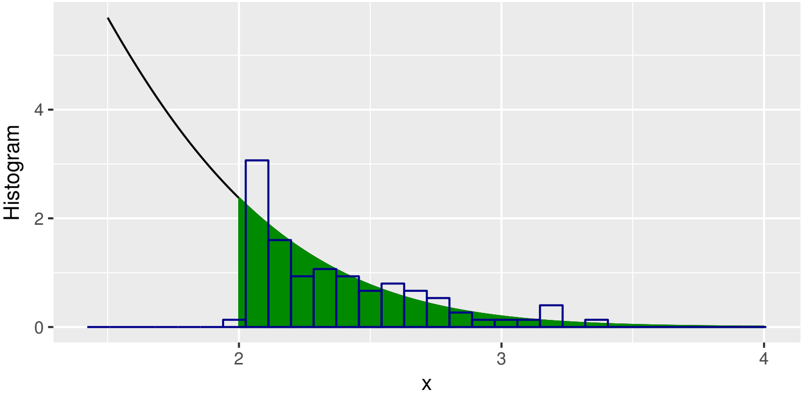
Accept reject method



Accept reject method



Accept reject method



Multilevel Monte Carlo (MLMC)

Consider a situation where we are interested in an expectation

$$\mathbb{E}[\varphi(X)] = \int \varphi(x)\pi(\mathrm{d}x),$$

where $\varphi : \mathbb{X} \to \mathbb{R}$ is uniformly Lipschitz, but X cannot be simulated exactly from π. Instead we have approximations $\{\pi_1, \ldots, \pi_L\}$ of different levels of accuracy of π that are easy to sample from. We also assume that for $\ell_1 < \ell_2$, π_{ℓ_1} is less expensive to simulate from than π_{ℓ_2} but also a worse approximation of π.

Multilevel Monte Carlo

Define

$$\widehat{P}_\ell = \varphi(X_\ell), \quad \text{where} \quad X_\ell \sim \pi_\ell \text{ and } \ell \in \{1, \ldots, L\}.$$

Clearly

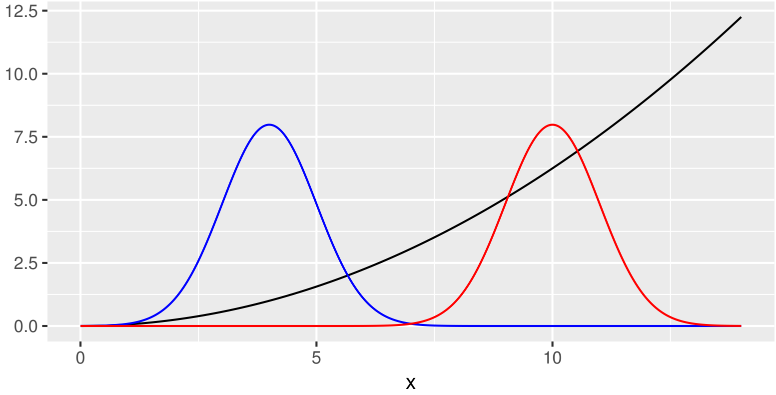
$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$$

For each of the expectations above we can construct the approximations

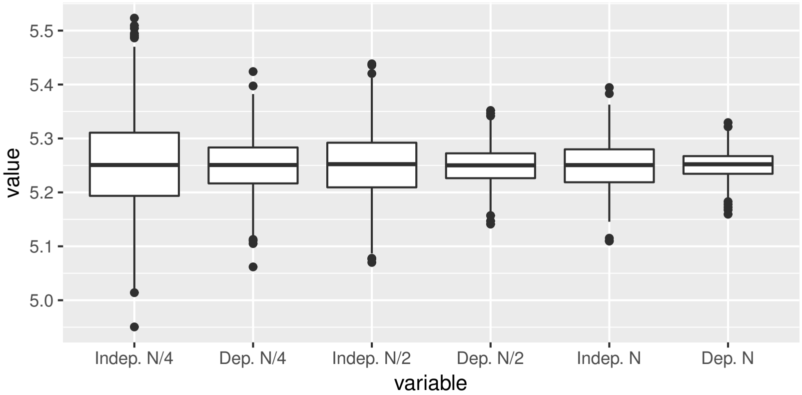
$$\widehat{Y}_0 = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \widehat{P}_0^i \quad \text{and} \quad \widehat{Y}_\ell = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} (\widehat{P}_\ell^i - \widehat{P}_{\ell-1}^i)$$

Variance reduction by control variates

$$\mathbb{E}[(X_2/4)^2 - (X_1/4)^2], \quad \text{where} \quad X_1 \sim \mathcal{N}(4,1), \quad X_2 \sim \mathcal{N}(10,1)$$



Variance reduction by control variates



$N = 1000$