

Importance Sampling and Independent Metropolis-Hastings with Unbounded weights

22/10/2026
Orsay.

Many are Selected - One is Chosen!

→ we can sample from q , we want to approximate π , known up to normalized cts.

write: $\pi(dx) = w(x) q(dx)$ $0 < w < \infty$ $q(w) = 1$.

Throughout $f: \mathbb{R}^d \rightarrow \mathbb{R}$ $\|f\|_\infty < \infty$.

↳ Both algo use w up to cts.

Pl:

→ Given $z_1, \dots, z_N \stackrel{iid}{\sim} q$, a selection rule returns:

$$Y_n = z_I \quad I \in \{1, \dots, N\}$$

The rule may inspect the samples $(z_i)_{1 \leq i \leq n}$, evaluate the weights $(w(z_i))_{1 \leq i \leq n}$, flip a few coins, but return one of the samples. We measure

$$\varepsilon_n = \| \mathbb{E}(Y_n) - \pi \|_{TV}, \quad \| \mu - \nu \|_{TV} = \sup_{0 \leq f \leq 1} | \mu(f) - \nu(f) |.$$

From rejection to approximate Selection:

if $w \leq n$ rejection sampling accepts $z \sim q$ w.p. $w(z)/n$, exact draw!
known

When w unbounded selection rule:

$$P(Y_n \in A) \leq N q(A)$$

$$\varepsilon_n \geq \frac{\pi(w > 2N)}{2}$$

Rule 1: SIR & SNIS:

$$P(I=i | z_{1:N}) = \frac{w(z_i)}{\sum_{i=1}^N w(z_i)} \quad \text{SIR}$$

$$\hat{F}_N = \frac{\sum_{i=1}^N w(z_i) f(z_i)}{\sum_{i=1}^N w(z_i)}$$

$$\mathbb{E}[f(z_I) | z_{1:N}] = \hat{F}_N, \quad \varepsilon_N = \sup_{0 \leq f \leq 1} | \mathbb{E}[\hat{F}_N] - \pi(f) |$$

Rule 2: INH: One step from x , initialize from q .

Draw $z \sim q$, $U \sim \text{Uniform}(0,1)$ accept z w.p. $\alpha(x, z) = \min \left\{ 1, \frac{w(z)}{w(x)} \right\}$

final state $q^{\otimes n-1}$

The Two rules use same $(z_i)_i, (w(z_i))_i$ differently.

Variance ordered: if $q(w^2) < \infty$:

SNIS vs INH ergodic average

$$\sigma_{\text{SNIS}}^2 = q(w^2(f - \pi(f))) \leq \sigma_{\text{INH}}^2$$

[Delyon & Leliandier 2018]

but bias still matters!

Consider $\Theta = \mathbb{E}_1 \left[\mathbb{E}_{2,1} [f(\theta_1, \theta_2) | \theta_1] \right]$, $m(\theta_1) = \mathbb{E}_{2,1} [f(\theta_1, \theta_2) | \theta_1 = \theta_1]$

for each outer draw $\theta_{1,j}$ approximate using $\hat{m}_N(\theta_{1,j})$, using N proposals (SNIS or INH) averaged over outer replicates $\cap$:

$$\hat{\theta}_{\cap, N} = \frac{1}{\cap} \sum_{j=1}^{\cap} \hat{m}_N(\theta_{1,j}) \quad C = \cap N$$

$$b_n(\theta_1) = \mathbb{E}[\hat{m}_N(\theta_1)] - m(\theta_1) \Rightarrow \mathbb{E}[\hat{\theta}_{\cap, N}] - \Theta = \mathbb{E}_1[b_n(\theta_1)]$$

$$\text{if } |\mathbb{E}_1[b_n(\theta_1)]| \lesssim N^{-\gamma} \quad \gamma > 0.$$

$$\forall (\hat{m}_N(\theta_1)) \leq V$$

$$\Rightarrow \text{NSE} \lesssim \frac{N}{C} + N^{-2\gamma}$$

Balancing:

$$\text{NSE} = O\left(C^{-\frac{2\gamma}{2\gamma+1}}\right).$$

what about bias ordering?

$q(w^p) < \infty$ for some $p > 1$.

$$\text{Thm} \quad \mathbb{E} \left[|\hat{F}_N - \pi(f)|^S \right] \leq C \begin{cases} n^{-S/2} & p \geq 2, \quad 1 \leq S \leq p. \\ n^{-S(p-1)/2} & 1 < p < 2, \quad 1 \leq S \leq p. \end{cases}$$

In particular: if $q(w^2) < \infty$ $|\mathbb{E}[\hat{F}_N] - \pi(f)| \leq \frac{12q(w^2)}{N}$.

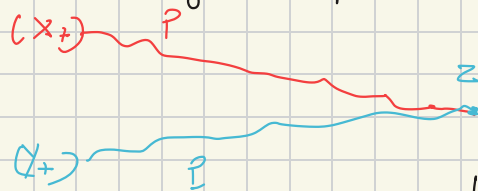
if $q(w^{2+\delta}) < \infty$, $q(w^{-\delta}) < \infty$ for some $\delta > 0$.

$$\lim_{N \rightarrow \infty} N \left(\mathbb{E}[\hat{F}_N] - \pi(f) \right) = -q(w^2(f - \pi(f))).$$

The bias of SNIS / SIR in TV is $O(N^{-\delta})$.

Coupling to the rescue:

INH rejection probability: $r(x) = 1 - \int \min\{1, \frac{w(z)}{w(x)}\} q(dz)$.



Run Two chains using same proposals and uniforms

$$\text{Let } z = \inf\{t \geq 0, X_t = Y_t\}.$$

Then: For distinct starting points $x, y, \forall t \geq 1$

$$\|P^t(x, \cdot) - P^t(y, \cdot)\|_{TV} = \left(\max\{r(x), r(y)\}\right)^t = P_{x,y}^t(z > t)$$

First instance of maximal at every time is Pitman 7.6 for MCMC.

Rates:

if $q(w^p) < \infty, p > 1$: a CRN coupling with $x_0 \sim q, y_0 \sim \pi$

$t \rightarrow \infty$

$$P(z > t) \leq \frac{C}{(1+t)^{p-1}} \text{ and } \|q^{p^t} - \pi\|_{TV} \leq \frac{C}{(1+t)^{p-2}}$$

C etc depends on $p, q(w^p)$.

with N proposal, INH returns $X_{N-1} \sim q^{p^{N-1}}$

$$\sup_{0 \leq f \leq 1} |\mathbb{E}_q[f(X_{N-1})] - \pi(f)| \leq C N^{-(p-1)}$$

Bias comparison at same proposal budget.

SNIS SIR	$ \mathbb{E}[\hat{F}_n] - \pi(f) \sim N^{-1}$
-------------	--

INH	$ \mathbb{E}_q[f(X_{N-1})] - \pi(f) = O(N^{-(p-1)})$
-----	---

If $p > 2$ INH has lower bias!

The exponent $(p-1)$ is sharp for INH up to polylogarithmic terms.

Removing The bias altogether:

Telescoping To coupled chains

$$b_t = \mathbb{E}[f(x_t)] \rightarrow \pi(f).$$

$$\pi(f) = b_0 + \sum_{t=1}^{\infty} b_t - b_{t-1}$$

[Glynn & Rhee 2014, Jacob, Oleary, Atchade 2020, Middleton 19].

Construct γ with same initialization and kernels X

$$\mathbb{E}[f(x_t) - f(\gamma_{t-1})] = b_t - b_{t-1}$$

$$Z = \{ \inf_{t \geq 0} : X_t = \gamma_{t-1} \}$$

Coupling with π on γ :

$$f(x_t) - f(\gamma_{t-1}) = 0 \quad t \geq 2$$

$$H = f(x_0) + \sum_{t=1}^{Z-1} f(x_t) - f(\gamma_{t-1})$$

$$\mathbb{E}[Z] < \infty$$

$$\mathbb{E}[H] = \pi(f).$$

Apply This to SIVIS:

PIPH: IPH on

$$\bar{q} = q^{\otimes N} \quad \hat{Z}_N(x) = \frac{1}{N} \sum_{i=1}^N w(x_i) \quad \bar{\pi} = \hat{Z}_N \bar{q}.$$

$$\frac{d\bar{\pi}}{d\bar{q}} = \hat{Z}_N \quad \text{acceptance proba} \quad 1 \wedge \frac{\hat{Z}_N(z)}{\hat{Z}_N(x)}.$$

$$\text{and } \mathbb{E}_{\bar{\pi}}[\hat{F}_N] = \mathbb{E}_{\bar{q}}[\hat{Z}_N \hat{F}_N] = \mathbb{E}_{\bar{q}}\left[\frac{1}{N} \sum_{i=1}^N w(x_i) f(x_i)\right] = \pi(f).$$

↳ SIVIS unbiased under $\bar{\pi}$.

Applying The previous construction gives

$$\hat{F}_N = \hat{F}_N(x_0) + \sum_{t=1}^{\hat{Z}_N-1} \hat{F}_N(x_t) - \hat{F}_N(\gamma_{t-1}).$$

$x_0 = (x_0^1, \dots, x_0^N) \sim q^{\otimes N}$ ← SIVIS

Bias correction.

Then:

For $1 \leq p \leq 1$, if $q(w^p) < \infty$, $p > 1$, Then, $X_0 \sim \bar{q}$, $X_0 \sim \bar{q}$ undegenerated

$$\text{Coupling: } \mathbb{E}[H_N] = \pi(f) \quad ; \quad \mathbb{E}[H_N^s] < \infty \quad 1 \leq s < p.$$

$p > 2 \Rightarrow f$ finite variance.

$$\text{for } p \geq 2: \mathbb{P}(Z_N > t) \leq \frac{C}{\sqrt{N} + t} \quad \mathbb{E}[Z_N] \leq 1 + \frac{C}{\sqrt{N}}$$

Cost:

$$\mathbb{E}[\text{weight evaluations}] = N \mathbb{E}[Z_N + 1] = 2N + o(\sqrt{N}).$$

$p \geq 2$.

□