Speeding up Monte Carlo methods

When is this method useful?

When is this method useful?

- Reliability analysis

  → estimating small probabilities

When is this method useful?

- Reliability analysis

  → estimating small probabilities

- Estimating properties

  → if you can define them as an expectation

Failure probabilities as expectations

Failure probabilities as expectations

$$X \sim P(X)$$

Failure probabilities as expectations

$$X \sim P(X)$$

$$p_{fail} = \mathcal{P}(g(X) < 0)$$

Failure probabilities as expectations

$$X \sim P(X)$$

$$p_{fail} = \mathcal{P}(g(X) < 0)$$

$$= \mathbb{E}_{X \sim P}\left[\mathbb{1}_{g(X)<0}\right]$$

The finite-sample approximation is *unbiased*...

The finite-sample approximation is *unbiased*...

...and the law of large numbers gives us *convergence*

The finite-sample approximation is *unbiased*...

...and the law of large numbers gives us *convergence* , i.e.

The finite-sample approximation is *unbiased*...

...and the law of large numbers gives us *convergence* , i.e.

$$\mathbb{E}_{X\sim P}[f(X)] = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

when $x_i \sim P$

Hundreds, thousands, millions of samples are common

Hundreds, thousands, millions of samples are common

When $f(x)$ is expensive to evaluate, this can get out of hand quickly

$\rightarrow$ for example, when it means solving a PDE

Let's rewrite this:

$$\mathbb{E}_{X \sim P}\left[f(X)\right] = \int f(x)p(x)dx$$

Let's rewrite this:

$$\mathbb{E}_{X \sim P}[f(X)] = \int f(x)p(x)dx$$

$$= \int f(x)\frac{p(x)}{q(x)}q(x)dx$$

Let's rewrite this:

$$\mathbb{E}_{X\sim P}[f(X)] = \int f(x)p(x)dx$$

$$= \int f(x)\frac{p(x)}{q(x)}q(x)dx$$

$$= \mathbb{E}_{X\sim Q}\left[f(X)\frac{p(X)}{q(X)}\right]$$

Let's rewrite this:

$$\mathbb{E}_{X\sim P}[f(X)] = \int f(x)p(x)dx$$

$$= \int f(x)\frac{p(x)}{q(x)}q(x)dx$$

$$= \mathbb{E}_{X\sim Q}\left[f(X)\frac{p(X)}{q(X)}\right]$$

$\rightarrow \frac{p(x)}{q(x)}$ are called *importance weights*

Key concept: rewrite an expectation w.r.t. a probability distribution as another expectation w.r.t. a new probability distribution

Key concept: rewrite an expectation w.r.t. a probability distribution as another expectation w.r.t. a new probability distribution

Why though?

A different perspective: achieve faster convergence by decreasing variance

A different perspective: achieve faster convergence by decreasing variance

→ Importance sampling is a *variance reduction* technique

When you're approximating $\int f(x)p(x)dx$, focus on areas where $|f(x)p(x)|$ is large

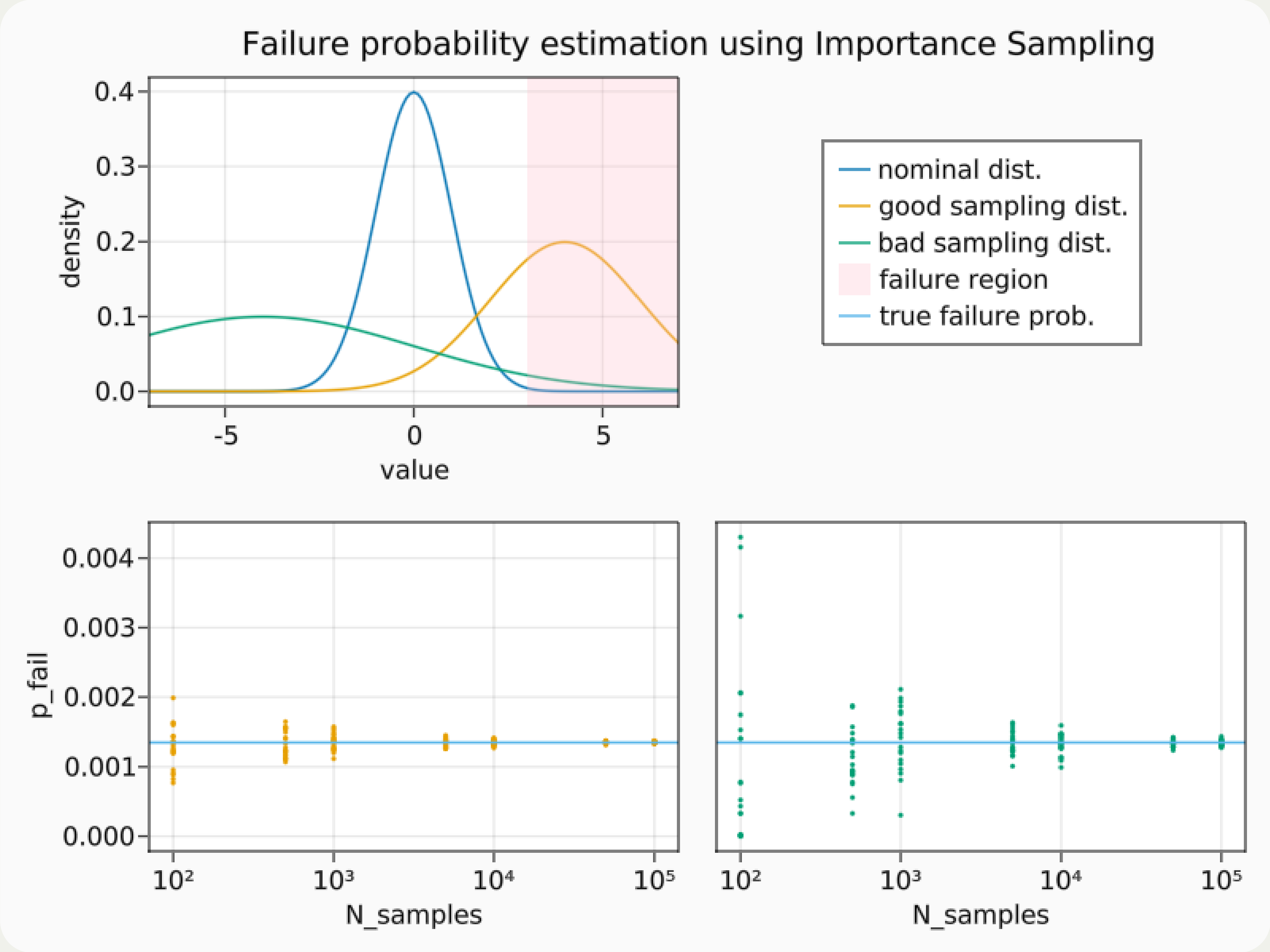
When you're approximating $\int f(x)p(x)dx$, focus on areas where $|f(x)p(x)|$ is large

Sampling from P may not be what you want to do then...

When you're approximating $\int f(x)p(x)dx$, focus on areas where $|f(x)p(x)|$ is large

Sampling from P may not be what you want to do then...

...so you should sample from a different distribution!



Failure probability estimation using Importance Sampling
density
value
nominal dist.
good sampling dist.
bad sampling dist.
failure region
true failure prob.
p_fail
N_samples
N_samples

How do you choose a sampling distribution?

How do you choose a sampling distribution?

Multifidelity Importance Sampling by Peherstorfer et al. (2016)

How do you choose a sampling distribution?

Multifidelity Importance Sampling by Peherstorfer et al. (2016)

→ Use a Reduced Order Model

A quick overview of Multifidelity Importance Sampling (MFIS):

A quick overview of Multifidelity Importance Sampling (MFIS):

- Generate a surrogate model

A quick overview of Multifidelity Importance Sampling (MFIS):

- Generate a surrogate model
- Find examples of failures (using the surrogate)

A quick overview of Multifidelity Importance Sampling (MFIS):

- Generate a surrogate model
- Find examples of failures (using the surrogate)
- Fit a density Q to those failures

A quick overview of Multifidelity Importance Sampling (MFIS):

- Generate a surrogate model
- Find examples of failures (using the surrogate)
- Fit a density Q to those failures
- Evaluate the high-fidelity model ($X \sim Q$)

Failure mode: MFIS and Importance Sampling can *slow down* convergence

In conclusion, IS and MFIS can

- reduce computational requirements
- reduce usage of resources like electricity
- save you time

Open source implementation will be available soon!

github.com/RickDW/MFIS

→ Includes *fully automatic* surrogate generation from PDEs and more

→ Fits right into the scientific ML ecosystem of Julia!

julia