Presentation Example: Importance Sampling

Jane Doe
John Galt

Academic Practice
DD/MM/2024

Motivation

- Many questions in economics involve expectations of random variables, e.g. Decision making under uncertainty, Forecasting, Statistical Inference, Asset Pricing...
- Calculating expectations of a continuous random variable $E[x]$ requires evaluation of integrals:

$$E[x] = \int x f(x) dx$$

- If the distribution $f(x)$ is complicated this integral is often not tractable or hard to evaluate.
- Therefore, it is often necessary to approximate the integral instead:

This presentation: Monte Carlo Simulations $\rightarrow$ Importance Sampling

Motivation

- Many questions in economics involve expectations of random variables, e.g. Decision making under uncertainty, Forecasting, Statistical Inference, Asset Pricing...
- Calculating expectations of a continuous random variable $E[x]$ requires evaluation of integrals:

$$E[x] = \int xf(x)dx$$

- If the distribution $f(x)$ is complicated this integral is often not tractable or hard to evaluate.
- Therefore, it is often necessary to approximate the integral instead:

This presentation: Monte Carlo Simulations $\rightarrow$ Importance Sampling

Motivation

- Many questions in economics involve expectations of random variables, e.g. Decision making under uncertainty, Forecasting, Statistical Inference, Asset Pricing...
- Calculating expectations of a continuous random variable $E[x]$ requires evaluation of integrals:

$$E[x] = \int x f(x) dx$$

- If the distribution $f(x)$ is complicated this integral is often not tractable or hard to evaluate.
- Therefore, it is often necessary to approximate the integral instead:

This presentation: Monte Carlo Simulations $\rightarrow$ Importance Sampling

Motivation

- Many questions in economics involve expectations of random variables, e.g. Decision making under uncertainty, Forecasting, Statistical Inference, Asset Pricing...
- Calculating expectations of a continuous random variable $E[x]$ requires evaluation of integrals:

$$E[x] = \int xf(x)dx$$

- If the distribution $f(x)$ is complicated this integral is often not tractable or hard to evaluate.
- Therefore, it is often necessary to approximate the integral instead:

This presentation: Monte Carlo Simulations $\rightarrow$ Importance Sampling

Outline of the Talk

1. Motivation ✓
2. Theory of Importance Sampling
3. Simulation Study
4. Outlook: Empirical Application for the Paper
5. Conclusion and Discussion

Theory of Importance Sampling

Main Idea

Approximate an integral over a complex target distribution $\pi(x)$ by weighted draws from a simpler proposal distribution $q(x)$ (see Doucet, Freitas, and Gordon (2001)).

Theory: Consider the Expected Value $E_\pi[x]$

$$
\begin{aligned}
E_\pi[x] &= \int x\pi(x)dx \\
&= \int x \underbrace{\frac{\pi(x)}{q(x)}}_{w(x)} q(x)dx \\
&= E_q[wx] \\
&\approx \frac{1}{T}\sum_{i=1}^{N} w_i x_i, \text{ with } x_i \sim q(x)
\end{aligned}
$$

Draws from $q(x)$ are reweighted using the **importance weights** $w(x) = \pi(x)/q(x)$:

- If x is more likely under $\pi(x)$ than under $q(x) \longrightarrow w(x) > 1$
- If x is less likely under $\pi(x)$ than under $q(x) \longrightarrow w(x) < 1$

More Theory: Why does it work?

Asymptotic Justification:

Strong Law of Large Numbers (SSLN):

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \rightarrow E[X] \quad \text{for} \quad N \rightarrow \infty \quad (\textit{almost surely}) \tag{1}$$

Assumptions for the SSLN in Equation(1):

1. $X_i \sim iid$ ✓
2. $E[X_i] = \mu < \infty \quad \forall i$ ✓
3. Finite variance ✓

If draws from $q(x)$ satisfy assumptions $1. - 3.$, the empirical average of weighted draws converges almost surely towards the expected value under $\pi(x)$:

$$\frac{1}{M}\sum_{i=1}^{M} w_i x_i \xrightarrow{a.s.} E_\pi[x]$$

Algorithm: Vanilla Importance Sampling

1. Draw M realizations x_i from the proposal $q(x)$
2. Compute and normalize the importance weights

$$w(x_i) = \frac{\pi(x_i)}{q(x_i)} \quad \text{and} \quad W_i = \frac{w(x_i)}{\sum_{i=1}^{M} w(x_i)}$$

3. Resample the draws $\{x_i\}_{i=1}^{M}$ using a multinomial distribution with corresponding probabilities $\{W_i\}_{i=1}^{M}$:

$$x_i \sim \mathcal{MN}\left(\{x_i\}_{i=1}^{M}, \{W_i\}_{i=1}^{M}\right)$$

4. Approximate expectations under the target distribution $\pi(x)$ with

$$E_\pi[x] \approx \frac{1}{M}\sum_{i=1}^{M} x_i$$

Importance Sampling: Example

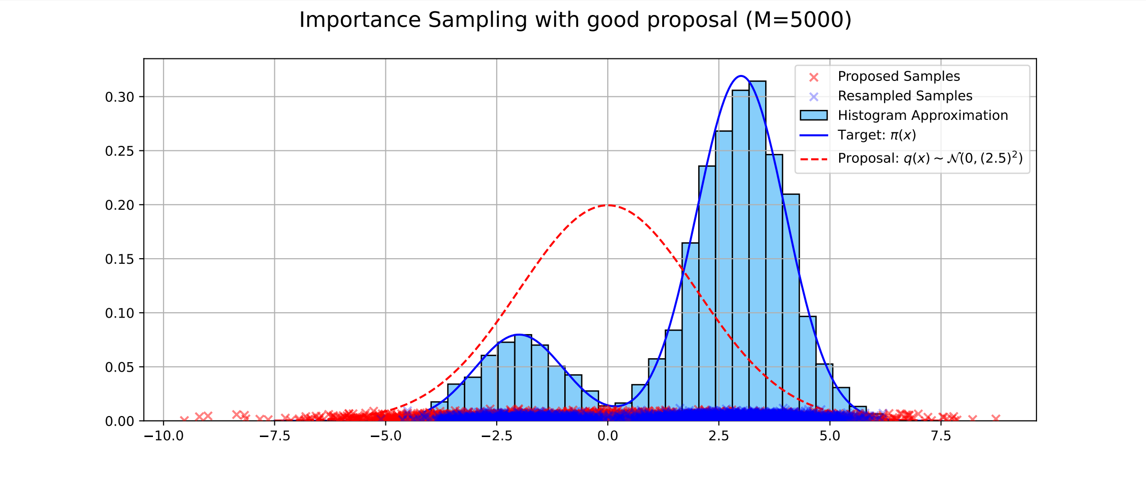

Figure 1: Approximation of a mixed Gaussian with $E[x] = 2$ using 5000 draws from a univariate normal distribution with good coverage. The histogram approximation accurately captures the two modes. The estimated mean $\hat{\mu} = 2.05$ is almost identical to the true value.

Importance Sampling: Sample Size

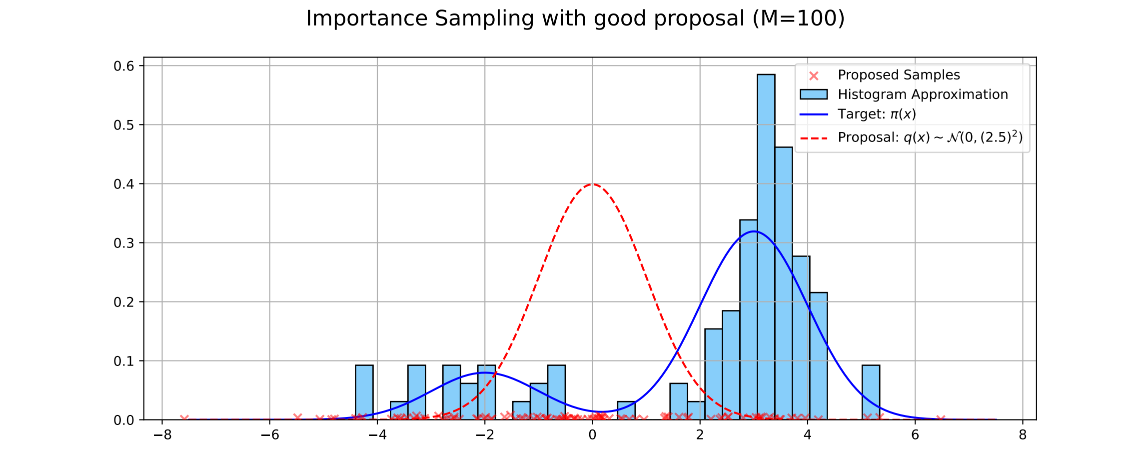

Figure 2: Approximation of a mixed Gaussian with $E[x] = 2$ using 100 draws from a univariate normal distribution. Due to low M the histogram approximation is still sparse. However, with a value of $\hat{\mu} = 2.1$, the estimated mean is already close.

Importance Sampling: Proposal Choice

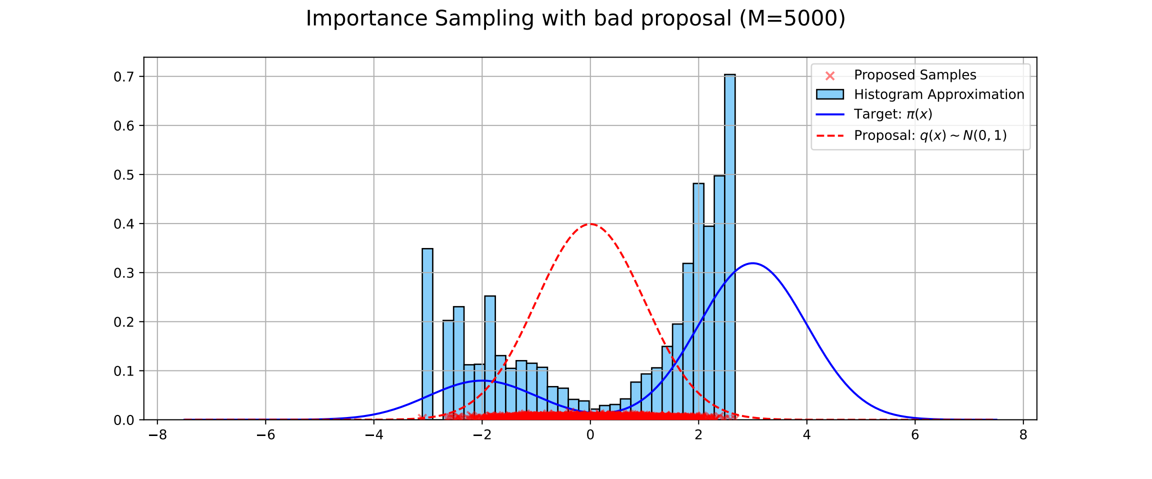

Figure 3: Example of a bad proposal that does not properly cover the modes of the target distribution. The histogram approximation can't capture the features of $\pi(x)$ well. Yet, the estimated mean $\hat{\mu} = 2.19$ is also not too far off.

How does the Sample Size and Proposal Choice affect the Accuracy of Importance Sampling?

Evaluating Estimation Accuracy of Importance Sampling

Simulation Set-up:

- Target Distribution: $\pi(x) \sim 0.8 \times \mathcal{N}(3,1) + 0.2 \times \mathcal{N}(-2,1) \longrightarrow \mu = 2, \sigma^2 = 4$
- Proposal Distributions: $q_1(x) \sim \mathcal{N}(0,(2.5)^2)$ and $q_2(x) \sim \mathcal{N}(0,1)$

Results for $N = 1000$ Simulations and increasing sample size M:

1. Good Proposal (q_1):

M	**50**	**100**	**500**	**1000**	**5000**
Mean($\hat{\mu}$)	1.956	1.935	1.989	2.007	2.001
Var($\hat{\mu}$)	0.26	0.14	0.025	0.012	0.003

2. Bad Proposal (q_2):

M	**50**	**100**	**500**	**1000**	**5000**
Mean($\hat{\mu}$)	0.611	0.888	1.355	1.47	1.757
Var($\hat{\mu}$)	1.445	1.319	0.761	0.596	0.305

Estimates from $q_1(x)$ are far more precise with smaller variance (Factor of 100 for $M = 5000$).

Typical Applications in Economics

1. Herbst and Schorfheide (2014): Approximate distribution of parameters of a dynamic macro model using Importance sampling
2. Glasserman and Li (2005): Analyze probabilities and risks of credit portfolios
3. Doucet, Freitas, and Gordon (2001): Analyze volatility in stock markets.

In the application for the seminar paper we plan to use importance sampling to evaluate downside risks for the stock indices DAX and S&P 500.

Conclusion

Importance Sampling

1. ... provides a convinient methods to approximate Expectations when closed form solutions are hard/impossible to obtain
2. ... conceptually simple and easy to implement on modern computers
3. ... widely applicable to many econometric problems.

Drawbacks:

1. Choice of correct proposal often not easy.
2. Method becomes slow in high dimensions.

Outlook:

1. Application of the method to evaluate credit risks.