Importance Sampling
Latest Revision: January 21, 2014

Frank Schorfheide[1]

1 Importance Sampling

- A particle filter is a sequential importance sampler, which also involves some resampling steps.
- To understand the particle filter, it is important to understand importance sampling.
- Let $\pi(\theta) = f(\theta)/Z$. Monte Carlo approximation of
$$\mathbb{E}_\pi[h(\theta)] = \int h(\theta)\pi(\theta)d\theta = \frac{1}{Z}\int h(\theta)w(\theta)g(\theta)d\theta,$$
where
$$w(\theta) = \frac{f(\theta)}{g(\theta)}.$$
- Define
$$\bar{h}_N = \frac{\frac{1}{N}\sum_{i=1}^N h(\theta^i)w(\theta^i)}{\frac{1}{N}\sum_{i=1}^N w(\theta^i)},$$
where the "particles" θ^i's are drawn from the distribution with density $g(\cdot)$.
- Notation: $\mathbb{E}_\pi$ and $\mathbb{E}_g$ for expectations under $g(\cdot)$ and $\pi(\cdot)$.
- Note that
$$\mathbb{E}_g[hf/g] = \int hZ\pi d\theta = Z\mathbb{E}_\pi[h].$$
- Likewise,
$$\mathbb{E}_g[f/g] = \int Z\pi d\theta = Z\mathbb{E}_\pi[1] = Z.$$
- Provided that
$$\mathbb{E}_g[|hf/g|] < \infty \quad \text{and} \quad \mathbb{E}_g[|f/g|] < \infty$$

[1]Department of Economics, 3718 Locust Walk, Room 525 McNeil Building, Telephone (215) 898 8486. Email: *schorf@ssc.upenn.edu*, URL: *http://www.econ.upenn.edu/~schorf.*

- we can deduce from the Strong Law of Large Numbers and the Continuous Mapping Theorem that
$$\bar{h}_N \stackrel{a.s.}{\longrightarrow} \mathbb{E}_\pi[h].$$

- Recall
$$\bar{h}_N = \frac{\frac{1}{N}\sum_{i=1}^N h(\theta^i)w(\theta^i)}{\frac{1}{N}\sum_{i=1}^N w(\theta^i)},$$

- Re-normalize the weights: define $v(\theta) = w(\theta)/Z$.

- We will derive
$$\begin{aligned} \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^N h(\theta^i)v(\theta^i) - \mathbb{E}_\pi[h]\right) &\Longrightarrow N(0,?) \\ \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^N v(\theta^i) - 1\right) &\Longrightarrow N(0,?) \end{aligned}$$

- Then use δ method:
$$\begin{aligned} \sqrt{N}(\bar{h}_N - \mathbb{E}_\pi[h]) &= \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^N h(\theta^i)v(\theta^i) - \mathbb{E}_\pi[h]\right) \\ &\quad -\mathbb{E}_\pi[h]\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^N v(\theta^i) - 1\right) + o_p(1). \\ &\Longrightarrow N(0,?) \end{aligned}$$

First,
$$\begin{aligned} \mathbb{V}_g[hv] &= \mathbb{E}_g\left[(hv - \mathbb{E}_\pi[h])^2\right] \\ &= \mathbb{E}_g\left[(h(\pi/g) - \mathbb{E}_\pi[h])^2\right] \\ &= \mathbb{E}_g[h^2(\pi/g)^2 - 2h\mathbb{E}_\pi[h](\pi/g) + \mathbb{E}_\pi^2[h]] \\ &= \mathbb{E}_\pi[(\pi/g)h^2] - \mathbb{E}_\pi^2[h] \end{aligned}$$

Second,
$$\begin{aligned} \mathbb{V}_g[v] &= \mathbb{E}_g\left[(v-1)^2\right] \\ &= \mathbb{E}_g\left[(\pi/g - 1)^2\right] \\ &= \mathbb{E}_g[(\pi/g)^2 - 2(\pi/g) + 1] \\ &= \mathbb{E}_\pi[(\pi/g)] - 1 \end{aligned}$$

- This leads to

$$\begin{aligned}
&\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} h(\theta^i)v(\theta^i) - \mathbb{E}_\pi[h]\right) \\
&\quad \Longrightarrow \quad N\bigg(0, (\mathbb{E}_\pi[(\pi/g)h^2] - \mathbb{E}^2_\pi[h])\bigg) \\
&\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} v(\theta^i) - 1\right) \\
&\quad \Longrightarrow \quad N\bigg(0, (\mathbb{E}_\pi[(\pi/g)] - 1)\bigg)
\end{aligned}$$

- Important regularity conditions for Lindeberg-Levy CLT:

$$\sup_\theta \; \pi/g < \infty \quad \text{and} \quad \mathbb{E}_g[h^2] < \infty.$$

- We also need the covariance...

Third,

$$\begin{aligned}
COV_g(hv, v) &= \mathbb{E}_g\bigg[(h(\pi/g) - \mathbb{E}_\pi[h])(\pi/g - 1)\bigg] \\
&= \mathbb{E}_g[h(\pi/g)^2 - \mathbb{E}_\pi[h](\pi/g) - h(\pi/g) + \mathbb{E}_\pi[h]] \\
&= (\mathbb{E}_\pi[(\pi/g)h] - \mathbb{E}_\pi[h])
\end{aligned}$$

- At the end of the day...

$$\sqrt{N}(\bar{h}_N - \mathbb{E}_\pi[h]) \Longrightarrow N(0, \Omega(h))$$

- where

$$\begin{aligned}
\Omega(h) &= (\mathbb{E}_\pi[(\pi/g)h^2] - \mathbb{E}^2_\pi[h]) + \mathbb{E}^2_\pi[h](\mathbb{E}_\pi[(\pi/g)] - 1) \\
&\quad -2\mathbb{E}_\pi[h](\mathbb{E}_\pi[(\pi/g)h] - \mathbb{E}_\pi[h]) \\
&= \mathbb{E}_\pi[(\pi/g)h^2] + \mathbb{E}^2_\pi[h]\mathbb{E}_\pi[(\pi/g)] - 2\mathbb{E}_\pi[h]\mathbb{E}_\pi[(\pi/g)h] \\
&= \mathbb{E}_\pi[(\pi/g)(h^2 + \mathbb{E}^2_\pi[h] - 2h\mathbb{E}_\pi[h])] \\
&= \mathbb{E}_g[(\pi/g)^2(h - \mathbb{E}_\pi[h])^2] \\
&= \mathbb{V}_g[(\pi/g)(h - \mathbb{E}_\pi[h])]
\end{aligned}$$

- Using the crude approximation

$$\mathbb{E}_\pi[(\pi/g)(h - \mathbb{E}_\pi[h])^2] \approx \mathbb{V}_\pi[h](\mathbb{V}_g[\pi/g] + 1)$$

  we get a measure of effective sample size

$$ESS = N\frac{\mathbb{V}_\pi[h]}{\Omega(h)} = \frac{N}{1 + \mathbb{V}_g[\pi/g]}.$$