Derivation of the Importance Sampling Algorithm

Here is the correct derivation of the importance sampling algorithm stated in Eqs. 10.59 to 10.60 and the correct statement of Algorithm 10.2.

Assume we wish to compute

$$E(g(\theta)) = \int g(\theta)f(\theta|x)d\theta.$$  \hfill (10.A1)

Given an importance density $h(\theta)$ that has the same support as $f(\theta|x)$, rewrite Eq. 10.A1 as

$$E(g(\theta)) = \int g(\theta)\frac{f(\theta|x)}{h(\theta)}h(\theta)d\theta$$

$$= f(x)^{-1}\int g(\theta)\frac{f(\theta)f(x|\theta)}{h(\theta)}h(\theta)d\theta$$

$$= f(x)^{-1}\int g(\theta)w(\theta|x)h(\theta)d\theta$$ \hfill (10.A2)

where $w(\theta|x) = h(\theta)^{-1}f(\theta)f(x|\theta)$. Similarly, we can write

$$1 = \int f(\theta|x)d\theta = f(x)^{-1}\int \frac{f(\theta)f(x|\theta)}{h(\theta)}h(\theta)d\theta$$

$$= f(x)^{-1}\int w(\theta|x)h(\theta)d\theta.$$ \hfill (10.A3)

We can use Eqs. 10.A2 and 10.A3 to write Eq. 10.A1 as

$$E(g(\theta)) = \frac{E(g(\theta))}{1} = \frac{f(x)^{-1}\int g(\theta)w(\theta|x)h(\theta)d\theta}{f(x)^{-1}\int w(\theta|x)h(\theta)d\theta}$$

$$= \frac{\int g(\theta)w(\theta|x)h(\theta)d\theta}{\int w(\theta|x)h(\theta)d\theta}. \hfill (10.A4)$$

We can approximate the numerator and denominator in Eq. 10.A4 respectively by $n$ draws from $h(\theta)$ as
\[ \int g(\theta)w(\theta|x)h(\theta)d\theta \approx \frac{1}{n} \sum_{i=1}^{n} g(\theta_i)w(\theta_i|x) \quad (10.A5) \]

\[ \int w(\theta|x)h(\theta)d\theta \approx \frac{1}{n} \sum_{i=1}^{n} w(\theta_i|x) . \]

Therefore, we have the importance sampling approximation of \( E(g(\theta)) \) as

\[ E(g(\theta)) \approx \frac{1}{n} \sum_{i=1}^{n} g(\theta_i)w(\theta_i|x) \quad (10.A6) \]

We can simulate Eq. 10.59 (10.A1) with the following algorithm

**Algorithm 10.2 (Importance Sampling)**

Sum = 0  
\( W = 0 \)  
For \( j = 1, \ldots, 10,000 \)

1. Draw \( \theta_j \) from \( h(\theta) \)
2. Compute \( w(\theta_j|x) = \frac{f(\theta_j)f(x|\theta_j)}{h(\theta_j)} \) and \( g(\theta_j) \)
3. Sum \( \leftarrow \) Sum + \( w(\theta_j|x)g(\theta_j) \)
4. \( W \leftarrow W + w(\theta_j|x) \)

Compute \( E[g(\theta)] \approx W^{-1}\text{Sum} \)