

Importance Sampling

November 26, 2021

Importance sampling is an approximation method that aims at proposing a new distribution to sample from that has higher density on areas that contribute more to the expectation. That was a mouthful. Let's break that into smaller chunks.

First, let's quickly get this out of the way: importance sampling is an **approximation method**, not a sampling method. There is no real sampling procedure defined here.

Great! Now that that's out of the way, the next thing is to introduce two distributions: the *target distribution* $p(x)$ and the *proposal distribution* $q(x)$. The target distribution describes the true underlying distribution of the data, whereas the proposal distribution is the new distribution we will sample from. In importance sampling, we want to control how often we sample certain values. As stated, we usually want $q(x)$ to have high density on areas that contribute more to the expectation. What does this exactly mean?

Let's turn our attention to Figure 1. $p(x)$ is the true underlying distribution of the function $f(x)$ and $q(x)$ is our proposal distribution. Our goal is to characterize the function $f(x)$ (e.g. estimate the expectation). We can see that $p(x)$ will have trouble sampling from the right side of $f(x)$, which is where all the action is happening. Meanwhile, $q(x)$ will do a pretty good job at sampling from the important areas (where the action is happening); hence, **importance sampling**.

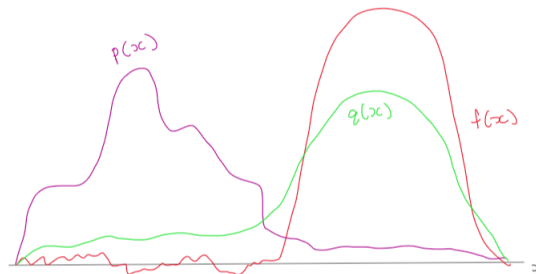


Figure 1: A good proposal distribution

Now it's important to note that, in the long run, it really doesn't matter which distribution we use, because Monte Carlo estimates are unbiased. However, it is **also** important to note that, if we sample from $q(x)$ instead of $p(x)$, our estimate will have a smaller variance, resulting in higher accuracy.

$$\sigma^2(I_n) = \frac{1}{n} \sigma^2\left(f(x) \frac{p(x)}{q(x)}\right),$$

where $\sigma^2(I_n)$ is the variance of our samples using $q(x)$ as our proposal distribution. As an aside, in some applications, importance sampling is used to solve the problem of sampling from an intractable density. In this case, typically $q(x)$ is chosen to be as close to the true (but intractable) distribution $p(x)$.

Now that we have the big idea, how do we do the derivation? It's actually really simple! We'll start with the expectation, then multiplying by $\frac{q(x)}{q(x)}$ and finally applying Monte Carlo approximations,

$$\begin{aligned} E(f(x)) &= \int f(x)p(x)dx \\ &= \int f(x)\frac{p(x)}{q(x)}q(x)dx \\ &\approx \frac{1}{n} \sum_{i=1}^n f(x_i) \frac{p(x_i)}{q(x_i)} \\ &= \frac{1}{n} \sum_{i=1}^n f(x_i) w(x_i) \end{aligned}$$

where

$$w(x) = \frac{p(x)}{q(x)}, \quad x_i \sim q(x)$$

1 Notes

- $q(x)$ is the proposal distribution
- essentially multiply by $\frac{q(x)}{q(x)}$

$$\begin{aligned} E(f(x)) &= \int f(x)p(x)dx \\ &= \int f(x)\frac{p(x)}{q(x)}q(x)dx \\ &\approx \frac{1}{n} \sum_{i=1}^n f(x_i)\frac{p(x_i)}{q(x_i)} \\ &= \frac{1}{n} \sum_{i=1}^n f(x_i)w(x_i) \qquad w(x) = \frac{p(x)}{q(x)} \qquad x_i \sim q(x) \end{aligned}$$

- can enhance accuracy by choosing a good $q(x)$

$$\sigma^2(I_n) = \frac{1}{n} \sigma^2\left(f(x)\frac{p(x)}{q(x)}\right)$$

- want to choose $q(x)$ large where $|f(x)|p(x)$ is large, because that's where all the action is happening