Importance Sampling

Tuesday, January 16, 2018 9:16 AM

Importance sampling is used to find Expectations.

Given:

$$E_{p(x)}[f(x)] = \int f(x)\, p(x)\, dx$$

Assume that we could easily evaluate both $f(x)$ and $p(x)$ given x.

But somehow this integral is hard to compute. and $p(x)$ is hard to sample.

B/c else we could simply sample from $x \sim p(x)$ and compute:

$$E_{p(x)}[f(x)] \approx \frac{1}{L} \sum_{i=1}^{L} f(x_i)$$

Since $p(x)$ is difficult to sample from,

We could create a "proposal" distribution $q(x)$ that is easy to sample. With this new $q(x)$, we change the expectation formulation.

$$E_{p(x)}[f(x)] = \int f(x)\, p(x)\, dx$$

$$E_{q(x)}\left[f(x)\frac{p(x)}{q(x)}\right] = \int f(x)\frac{p(x)}{q(x)}\, q(x)\, dx \leftarrow$$

Now it is an expectation with respect to $q(x)$

$q(X)$

Now we can sample from $X_i \sim q(X)$ and compute the empirical expectation.

$$E_{p(x)}[f(X)] \approx \frac{1}{L}\sum_{i=1}^{L} f(X_i)\frac{p(X_i)}{q(X_i)}$$

<u>Example</u>

Let's look at a mixture of Gaussian with

$z \sim$ Bernoulli (p)

$X \sim N(0, \sigma^2)$

The goal is to find the posterior

$$p(z|X) = \frac{P(X|z)\,P(z)}{P(X)}$$

To calc the posterior $p(X) = \int P(X|z)P(z)\,dz$ must be computed. But since this integral can be difficult, we can use importance sampling to approximate it.

$$p(X) = E_{p(z)}[p(x|z)] = \int p(x|z)p(z)\,dz$$

$$E_{q(z)}\left[p(x|z)\frac{p(z)}{q(z)}\right] = \int p(x|z)\frac{p(z)}{q(z)}\,q(z)\,dz$$

Now we sample $z \sim q(z) \leftarrow$ proposal distribution.

$$E_{q(z)}\left[p(x|z)\frac{p(z)}{q(z)}\right] \approx \frac{1}{L}\sum_{i=1}^{L} p(x|z_i)\frac{p(z_i)}{q(z_i)}$$

$$E_{q(z)}\left[p(x|z)\frac{p(z)}{q(z)}\right] \approx \frac{1}{L}\sum_{i=1}^{L} p(x|z_i)\frac{p(z_i)}{q(z_i)}$$