COMP 480/580 — Probabilistic Algorithms and Data Structure Oct. 12, 2023

Lecture 15 Basic Sampling

Lecturer: Anshumali Shrivastava

Scribe By: Claire Huo(ch107), Weijie Huang(wh31), and Yudi Yang(jy103)

1 Intro: Estimate an Area

Problem: Estimate the ratio of a (random) area in a square. We uniformly sample from the region, and whether the sample belongs to the region.

Example: to estimate:

$$\int_0^1 e^{-x^2}\,\mathrm{d}x$$

There are multiple ways to do this:

1. Pick x and y randomly in $[0, 1]$ and check if it lies in the area.
2. Pick x and compute $y = f(x)$.
3. Riemann Integral.

Those solutions are very different. The sampling solution is independent of the messiness of the area. The Riemann Integral solution is not. E.g., a higher dimension problem.

2 How to Generate Distributions in Programs

We first need a source of randomness by assuming a uniform number generation.

To generate Normal Distribution, there are a few ways to do this:

1. We can add up a lot of uniform numbers.
2. We can convert our uniform number generator since we know how Normal Distribution looks like. $y = f(x)$.

Today's lecture focus on the methods that generate new distributions from a known distribution generator.

3 Inversion Method

Purpose: Generate random numbers from a given probability distribution.

Algorithm 1 Inversion Method

for $i = 1$ **to** N **do**
  Draw a uniform random number:
    Sample $y_i \sim U[0, 1]$
  Transform to the desired distribution:
    Compute $x_i = F^{-1}(y_i)$
end for

This method requires us to find the inverse of the Cumulative Distribution Function (CDF) of the distribution F^{-1}. This does not require what F is but only need F^{-1}, but finding the inverse may be as hard as sampling. This means that this method does not always work.
Proof:

$$\begin{aligned} &Pr(u \leq x) \\ =&Pr(F^{-1}(y) \leq x) \\ =&Pr(y \leq F(x)) \\ =&F(x) \end{aligned}$$

4 Rejection Sampling

Purpose: Generate samples from a distribution for which direct sampling is difficult.

Algorithm 2 Rejection Sampling

for $i = 1$ **to** N **do**
  Loop:
  Draw a sample:
    Sample $x \sim g(x)$
  Draw a uniform random number:
    Sample $u \sim U[0, 1]$
  Acceptance Check:
  if $u \leq \frac{f(x)}{M \cdot g(x)}$ **then**
      Accept x and **Exit** the loop.
  else
      Continue the loop (reject x and try again).
  end if
end for

Assuming we have access to proposal distribution g, how can we sample from target distribution f?

Constraint: $\exists M$ such that $f \leq Mg$. This means that if for some values x, $g(x) = 0$ while $f(x) > 0$, we cannot use g as the proposal distribution.

Recursively, sample $x \sim g$ and $u \sim U[0, 1]$. If $u \leq \frac{f(x)}{Mg(x)}$, then return x, else continue the loop.

Intuitively, if at certain x, $g(x)$ is large while $f(x)$ is small, then $u \leq$ is a very small number, which will likely be rejected.

Changing the distribution by changing the acceptance / rejection condition.

The hardness is to pick optimal M and g. Ideally, we have to choose the smallest possible M and g close enough to f. Otherwise, a large portion of sampling will be rejected.

5 Importance Sampling

Purpose: Estimate properties of a particular distribution, while only having samples generated from a different distribution than the distribution of interest.

Algorithm 3 Importance Sampling

for $i = 1$ **to** N **do**

  Draw a sample from the proposal distribution:

    Sample $x_i \sim g(x)$

  Calculate weighted function value:

    Compute $F(x_i) = \frac{w(x_i)f(x_i)}{g(x_i)}$

end for

Estimation:

  $\hat{I} = \frac{1}{N}\sum_{i=1}^{N} F_i$

Importance Sampling is not a sampling method. It is an estimation method.
We want to sample from $f(x)$, but we do not know what $f(x)$ is.
If we want to estimate

$$\int_{x \sim f(x)} w(x)f(x)\,\mathrm{d}x$$

$w(x)$ is the weight of any $g(x)$. This is the expected value of any g. $E_{x\sim f(x)}(w(x))$
$x_n \to \frac{\sum_{i=1}^{n} w(x_i)}{n}$. Sample $x_1, x_2, \ldots, x_n \sim g(x)$, $\frac{1}{n}\sum_{i=1}^{n}[\frac{w(x_i)xf(x_i)}{g(x_i)}] = \int \frac{w(x)f(x)}{g(x)} g(x)\,\mathrm{d}x$. $\frac{f(x_i)}{g(x_i)}$ is the importance weight.