

Random Numbers and Monte Carlo Integration

Vignesh M Pai (20211132)

1 Transforming One Dimensional Probability Distributions

We start with a random number generator X that uniformly gives a random number in the range $[0, 1]$. The problem is to transform this number by a bijective function $f : [0, 1] \rightarrow [a, b]$ such that the resulting random function $Y = f(X)$ has probability distribution function ρ . In other words, we need to satisfy the equation

$$p(Y \in [y, y + dy]) = \rho(y)dy \quad (1)$$

We know that

$$p(Y \in [f(x), f(x) + f'(x)dx]) = p(X \in [x, x + dx]) = dx \quad (2)$$

substituting $y = f(x)$ and $dy = f'(x)dx$, we get

$$p(Y \in [y, y + dy]) = \frac{dy}{f'(x)} = \frac{dy}{f'(f^{-1}(y))} \quad (3)$$

Let the inverse of f be g , then

$$f'(f^{-1}(y)) = f'(g(y)) = \frac{1}{g'(y)} \quad (4)$$

Substituting back to (1), we get that

$$\begin{aligned} g'(y) &= \rho(y) \\ f^{-1}(y) &= \int_a^y \rho(y)dy \end{aligned}$$

Hence, to find f , ρ needs to be integrable and the integral needs to be invertible. Let's calculate f for some common values of ρ .

1.

$$\begin{aligned} \rho(x) &= \lambda e^{-\lambda x}, \quad x \geq 0 \\ g(y) &= 1 - e^{-\lambda y} \\ f(x) &= -\frac{\ln(1-x)}{\lambda} \end{aligned}$$

2 Transforming the Gaussian Distribution

Suppose we wish to calculate f for ρ given by

$$\rho(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Clearly, this is not integrable, hence calculating f is not so simple. However, we know that $\rho'(x, y) = \rho(x)\rho(y)$ is integrable once transformed to polar coordinates.

$$\begin{aligned}\rho'(x, y)dxdy &= \rho'(r \cos(\theta), r \sin(\theta))rdrd\theta \\ &= \frac{1}{2\pi} \exp\left(-\frac{r^2}{2}\right) rdrd\theta \\ &= -d\left(\exp\left(-\frac{r^2}{2}\right)\right) d\left(\frac{\theta}{2\pi}\right)\end{aligned}$$

If we choose the change of variables

$$\begin{aligned}U &= \exp\left(-\frac{r^2}{2}\right) \\ V &= \frac{\theta}{2\pi}\end{aligned}$$

then note that U, V lie in the range $[0, 1]$. If U, V set to be random variables, and we write x, y as

$$\begin{aligned}x &= r \cos(\theta) = \sqrt{-2 \ln U} \cos(2\pi V) \\ y &= r \sin(\theta) = \sqrt{-2 \ln U} \sin(2\pi V)\end{aligned}$$

then x, y will each be random variables with the distribution ρ . The proof of this fact essentially lies in the change of variables we performed to derive U, V .

It turns out trigonometric functions are expensive to compute. Therefore, in practice, we can use the Marsaglia polar method. Let u, v be random variables in $[-1, 1]$ (they represent $\sqrt{U} \cos \theta$ and $\sqrt{U} \sin \theta$), let $s = u^2 + v^2$ and continue until $s \leq 1$. We now write x, y as

$$\begin{aligned}x &= \sqrt{-2 \ln s} \frac{u}{\sqrt{s}} = \sqrt{\frac{-2 \ln s}{s}} u \\ y &= \sqrt{-2 \ln s} \frac{v}{\sqrt{s}} = \sqrt{\frac{-2 \ln s}{s}} v\end{aligned}$$

It can be verified that x, y have distribution ρ .

Note that if a random variable X has distribution ρ , then $\sigma X + \mu$ has Gaussian distribution with mean μ and variance σ^2 .

3 Monte Carlo Integration

Suppose we wish to calculate the following integral

$$I = \int_0^1 J(x)dx$$

Let X be the uniform random variable in $[0, 1]$ as before, then we can approximate I as

$$I \approx \langle J \rangle = \frac{1}{N} \sum_{i=1}^N J(x_i)$$

where x_1, \dots, x_N are N samples of X . By CLT, it is obvious that the distribution of I will be a Gaussian with standard deviation

$$\sigma = \frac{\sigma_J}{\sqrt{N}} = \sqrt{\frac{\langle J^2 \rangle - \langle J \rangle^2}{N}}$$

the averages of J, J^2 could be computed using $J(x_i)$.

For integrals with different limits, we would need to transform our random variables. Suppose our limits are a, b and ρ is a probability distribution function on $[a, b]$. Then we can write

$$I = \int_a^b J(y)dy = \int_a^b \frac{J(y)}{\rho(y)} \rho(y)dy$$

Define $\tilde{J}(y) = J(y)/\rho(y)$ and suppose Y is a random variable with distribution ρ , then we can approximate I as

$$I \approx \frac{1}{N} \sum_{i=1}^N \tilde{J}(y_i)$$

where y_1, \dots, y_N are N samples of Y . Usually, Y is obtained by a transformation f of X .

$$I \approx \frac{1}{N} \sum_{i=1}^N \tilde{J}(f(x_i))$$

where x_1, \dots, x_N are N samples of X . When the distribution ρ is chosen to be of a similar shape to J , it is known as importance sampling.