1 Inverse Transform Sampling

Give a uniform random variable $U \sim Unif(0,1)$ and a cumulative distribution function F:

$$F^{-1}(U) = X$$

X is a random variable and has the distribution function $F(x)$.

Proof. F is a monotone function, thus it is invertible. When defining the distribution function of X:

$$F_x(x) = P\{X \leq x\}$$
$$F_x(x) = P\{F^{-1}U \leq x\}$$
$$F_x(x) = P\{U \leq F(x)\}$$
$$F_x(x) = F(x)$$

□

When sampling from an exponential random variable using a uniform random variable when $\lambda = \frac{1}{2}$

$$F(x) = 1 - e^{-\lambda x}$$
$$u = 1 - e^{-x/2}$$
$$F^{-1}(U) = -2ln(1-u)$$

2 Rejection Sampling

Given two random variables with distributions $f(x)$ and $g(x)$ and a scaling factor c,

$$cg(x) \geq f(x)$$

When simulating a normal random variable with a probability distribution $f(x)$ by using an exponential random variable with a distribution $g(x)$,

$$cg(x) \geq f(x)$$
$$c \geq \frac{f(x)}{g(x)}$$
$$c \geq \frac{2e^{-x^2/2}}{\sqrt{2\pi}e^{-x}}$$
$$c = \sqrt{\frac{2e}{\pi}}, \qquad \frac{f(x)}{g(x)} = e^{-(x-1)^2/2}$$

Thus the absolute values of normal random samples can be determined by simulating an exponential distribution using uniform random variables by inverse transform sampling.

3 Box Muller Method

The Box Muller method takes uniformly distributed random numbers between (0,1) and turns them into numbers with a normal distribution. The uniformly-distributed random variables U_1 and U_2 will be used as inputs for the direction, θ, of the vector and length of the vector, r, within the graph.

$$\theta = 2\pi U_1$$

$$r = \sqrt{-2log(U_2)}$$

Proof. To show that $-2log(U_2)$ will give a squared vector length r^2 that is exponentially distributed: [5]

$$P[R \leq r^2] = P[-2log(U_2 \leq r^2)]$$

$$P[R \leq r^2] = P[log(U_2 \geq (-r^2/2)]$$

$$P[R \leq r^2] = 1 - P[U_2 < exp(-r^2/2)]$$

$$1 - \int_0^{exp(-r^2/2)} dt = 1 - exp(-r^2/2)$$

$$1 - e^{-r^2/2} \sim \int_{-\infty}^{\infty} \frac{1}{2} e^{-r^2/2} dr^2$$

□

This is the PDF of the Gaussian integral. And if we replace $r^2 = x$, then we would end up with:

The exponential distribution with a rate parameter of $\frac{1}{2}$ =

$$\int_0^{\infty} -\frac{1}{2} e^{(-x/2)} dx$$

$$-e^{-\infty/2} - (-e^0)$$

$$0 + 1$$

Where the area of the exponential distribution is also equal to 1, making it a PDF.

Because the Box Muller requires r and θ, we are also able to manipulate the Gaussian integral to switch to polar coordinates to give a normal distribution in context of the unit square.

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{\int_{-\infty}^{\infty} (e^{-x^2/2})(e^{-x^2/2}) dx}$$

$$\sqrt{\int_{-\infty}^{\infty} (e^{-x^2/2})(e^{-x^2/2}) dx}$$

$$\sqrt{\int_{-\infty}^{\infty}(e^{-x^2/2})(e^{-y^2/2})dx}$$

$$\sqrt{\int_{-\infty}^{\infty}e^{-(x^2+y^2)/2}dx}$$

We assume that the x and y variables are normally distributed. By definition, the length of the vector, is equal to $r^2 = x^2 + y^2$. The joint distribution $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$ also allows for the double integral where the first one will be within the cartesian plane and the second being within the unit circle:

$$\sqrt{\int_{0}^{2\pi}\int_{0}^{\infty}e^{-(r^2/2)}rdrd\Theta}$$

$$\sqrt{\int_{0}^{2\pi}\int_{0}^{\infty}e^{-(r^2/2)}rdrd\Theta}$$

$$\sqrt{2\pi\int_{0}^{\infty}e^{-(r^2/2)}dr}$$

$$\sqrt{2\pi\left[e^{-(r^2/2)}\right]_0^{\infty}}$$

where $u = -\frac{r^2}{2}$.

$$\sqrt{2\pi}$$

This can also be re-written as

$$\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}e^{-r^2/2}dx = 1$$

Because $r^2 = x^2+y^2$ and $e^{-(x^2+y^2)/2}$, $e^{-r^2/2}$ is the length of the vector. And since $\frac{1}{\sqrt{2\pi}}$ is the θ portion of the polar coordinates (the angle of the vector), we can use them to be our PDF for $P(X)$ and $P(Y)$.

$$P_x(X) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-z^2}{2}\right)$$

$$P_x(Y) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-z^2}{2}\right)$$

$U_1 = \frac{\theta}{\sqrt{2\pi}}$ turns into $\theta = \sqrt{2\pi U_1}$

$U_2 = e^{-r^2/2}$ turns into $r = \sqrt{-2ln(U_2)}$

Plug in 2 uniformly distributed random variables for U_1 and U_2. This will then give the r the length of the vector, and θ, the angle of the vector. You will then plug these values into:

$$z_0 = rcos(\theta)$$

$$z_1 = rsin(\theta)$$

The results are two normally distributed variables, z_0 and z_1, with z_0 being the transformation to the x coordinate and z_1 being the transformation to the y coordinate. Another way of writing the equation for the z_0 and z_1 variables is:

$$z_0 = R\cos(\theta) = \sqrt{-2ln(U_2)cos(2\pi U_1)}$$

$$z_1 = R\sin(\theta) = \sqrt{-2ln(U_2)sin(2\pi U_1)}$$

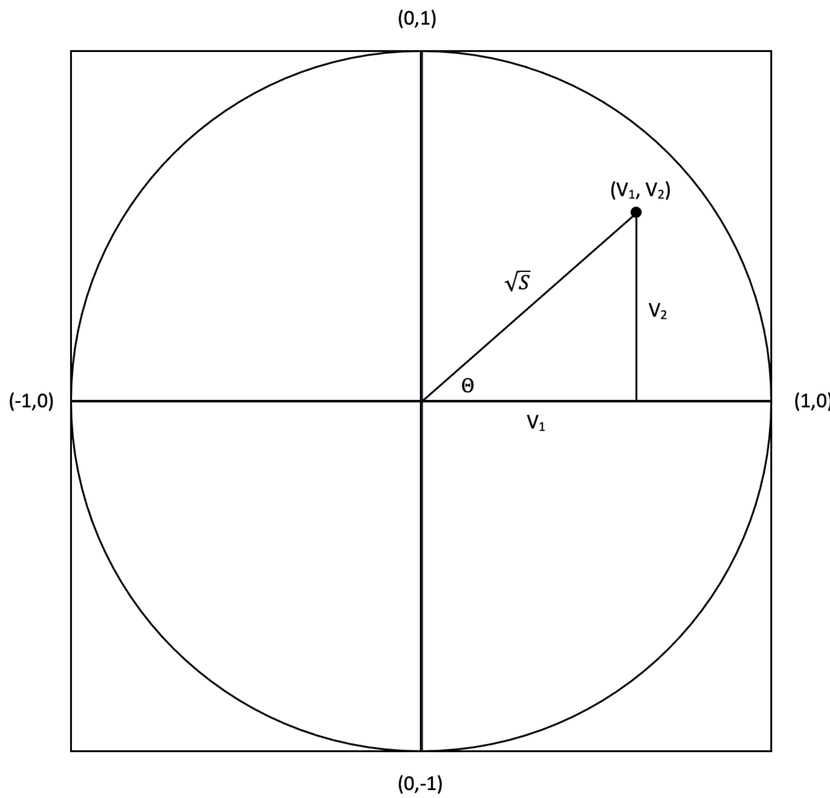
Figure 1: A demonstration of the Box Muller method. Since the X and Y axes are normally distributed, this proves that the Box Muller method was successful

`before.png` [3]

Summary of Box Muller Method

Step 1. Generate U_1 and U_2 from $U \sim U(0, 1)$
Step 2. Let $z_0 = \sqrt{-2ln(U_2)cos(2\pi U_1)}$ and $z_1 = \sqrt{-2ln(U_2)sin(2\pi U_1)}$
Step 3. Return the standard independent normal random variables within the coordinate pair (z_0, z_1)

4 The Polar Method



where $S = V_1^2 + V_2^2$

Summary of Polar Method

Step 1. Generate U_1 and U_2 from $U \sim U(0,1)$
Step 2. Let $V_1 = 2U_1 - 1$ and $V_2 = 2U_2 - 1$
Step 3. If $S = V_1^2 + V_2^2 > 1$, then go back to Step 1.
Step 4. Return the standard independent normal random variables $X = \sqrt{\frac{-2\ln S}{S}}V_1$ and $Y = \sqrt{\frac{-2\ln S}{S}}V_2$

Why does the Polar Method Work?

Let $U_1 \sim U(0,1)$ and $U_2 \sim U(0,1)$. Let $V_1 = 2U_1 - 1$ and $V_2 = 2U_2 - 1$. Then $V_1 \sim U(-1,1)$ and $V_2 \sim U(-1,1)$. Hence, the point (V_1, V_2) is uniformly distributed in the "unit square" (square with an area of 4 centered at the origin).

Let $S = V_1^2 + V_2^2$. Suppose we only accept points if $0 \leq S \leq 1$. Since (V_1, V_2) is uniformly distributed in the "unit square" then by rejecting points outside of the unit circle(where $S > 1$) we have that these accepted points are uniformly distributed in the unit circle. In other words, this rejection gets us from the points being uniformly distributed in the square to the points being uniformly distributed in the unit circle.

Let's consider the CDF for $S = V_1^2 + V_2^2$. Our sample space for our accepted points then is the area of the unit circle, which is π. The area of a circle with radius $\sqrt{S}$ is πS. What is the probability for $S < t$? If $t < 0$, then $P(S < t) = 0$. If $t > 1$, then $P(S < t) = 1$. If $t \in [0, 1]$, then $P(S < t) = \frac{\pi t}{\pi} = t$ (Same as $\frac{\pi(\sqrt{t})^2}{\pi} = t$). Hence, the CDF for S is $F(t) = P(S \leq t) = \begin{cases} 0, \text{ if } t < 0 \\ t, \text{ if } 0 \leq t \leq 1 \\ 1, \text{ if } t > 1 \end{cases}$

Then the PDF of S is $f(t) = \frac{dF}{dt} = \begin{cases} 1, \text{ if } 0 \leq t \leq 1 \\ 0, \text{ otherwise} \end{cases}$ This PDF is the same as the uniform distribution PDF for $U(0, 1)$, so we have that $S \sim U(0, 1)$.

Derivation and Importance of Box-Muller Approach

We can change from one set of coordinates to another using the formula $f_Y(\mathbf{y}) = f_X(h(\mathbf{y})) \cdot |J|$ where J is the Jacobian matrix, $Y = g(X) = (g_1(x), g_2(y)) = (r, \theta)$, $h = g^{-1}$, and $X = g^{-1}(Y) = h(Y)$.

Let $X = R\cos\Theta$ and $Y = R\sin\Theta$ be standard normal independent random variables. Then $f_{R,\Theta}(r, \theta) = f_{X,Y}(x, y) \cdot |J_F(r, \theta)|$ where $F : \mathbb{R}^+ \times [0, 2\pi) \to \mathbb{R}^2$. Then

$$J_F(r, \theta) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix}$$

so we have that

$$|J_F(r, \theta)| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = (\cos\theta)(r\cos\theta) - (\sin\theta)(-r\sin\theta) =$$
$$r\cos^2\theta + r\sin^2\theta = r(\cos^2\theta + \sin^2\theta) = r(1) = r$$

Hence,

$$\begin{aligned}
f_{R,\Theta}(r,\theta) &= f_{X,Y}(x,y)\cdot r \\
&= f_X(x)f_Y(y)\cdot r(\text{because } X \text{ and } Y \text{ are independent}) \\
&= \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}\right)r \\
&= \frac{1}{2\pi}re^{-\frac{1}{2}(x^2+y^2)} \\
&= \frac{1}{2\pi}re^{-\frac{r^2}{2}} \\
&= \left(re^{-\frac{r^2}{2}}\right)\left(\frac{1}{2\pi}\right) \\
&= f_R(r)f_\Theta(\theta)
\end{aligned}$$

where R is Rayleigh distributed with $\sigma = 1$ and for $r \geq 0$ and Θ is uniformly distributed over $[0, 2\pi]$. So we have that R and Θ are independent.

We have shown that $X = R\cos\Theta$ and $Y = R\sin\Theta$ are standard independent normal random variable when $R \sim \text{Rayleigh}(1)$ and $\Theta \sim U(0, 2\pi)$. Now we want to use the Inverse Transform Method to simulate draws from R and Θ in order to simulate a draws from X and Y.

Since $R \sim \text{Rayleigh}(1)$, then the CDF of R is $F_R(r) = 1-e^{-\frac{r^2}{2}}$ for $r \geq 0$. Let $U \sim U(0,1)$. Then using the Inverse Transform Method, we can let $U = F_R(r)$. So we have that $U = 1 - e^{-\frac{r^2}{2}}$ which implies that $r = \sqrt{-2\ln(1-U)}$. Let $U_1 = (1-U)$. Since $U \sim U(0,1)$, then $U_1 \sim U(0,1)$, so $r = \sqrt{-2\ln U_1}$ is a simulated draw from the Rayleigh(1) distribution.

Now we apply the Inverse Transform Method again to simulate values for Θ. Since $\Theta \sim U(0, 2\pi)$, then the CDF for Θ is $F_\Theta(\theta) = \begin{cases} 0, \text{ for } \theta < 0 \\ \frac{\theta}{2\pi}, \text{ for } 0 \leq \theta \leq 2\pi \\ 1, \text{ for } \theta > 2\pi \end{cases}$.

Let $U_2 \sim U(0,1)$. Again, using the Inverse Transform Method $U_2 = F_\Theta(\theta)$. So we have that $U_2 = \frac{\theta}{2\pi}$ which implies that $\theta = 2\pi U_2$. So $\theta = 2\pi U_2$ is a simulated draw from $U(0, 2\pi)$.

Putting all of these ideas together and plugging in these simulated values for R and Θ, we have that

$$X = R\cos\Theta = \sqrt{-2\ln U_1}\cos(2\pi U_2)$$

and

$$Y = R\sin\Theta = \sqrt{-2\ln U_1}\sin(2\pi U_2)$$

are standard independent normal random variables.

From the Box-Muller approach, we know that $X = R\cos\Theta = \sqrt{-2\ln U_1}\cos(2\pi U_2)$ and $Y = R\sin\Theta = \sqrt{-2\ln U_1}\sin(2\pi U_2)$ are standard independent normal random variables. From basic right triangle trigonometry, we have that
$\cos\Theta = \dfrac{V_1}{\sqrt{V_1^2+V_2^2}} = \dfrac{V_1}{\sqrt{S}}$
and $\sin\Theta = \dfrac{V_2}{\sqrt{V_1^2+V_2^2}} = \dfrac{V_2}{\sqrt{S}}$.

Using the fact that $S \sim U(0,1)$ and using these trig substitutions, we have that

$$X = R\cos\Theta = \sqrt{-2\ln S}\left(\frac{V_1}{\sqrt{S}}\right) = \sqrt{\frac{-2\ln S}{S}}V_1$$

and

$$Y = R\sin\Theta = \sqrt{-2\ln S}\left(\frac{V_2}{\sqrt{S}}\right) = \sqrt{\frac{-2\ln S}{S}}V_2$$

are standard independent normal random variables.

Summary of Polar Method

Step 1. Generate U_1 and U_2 from $U \sim U(0,1)$
Step 2. Let $V_1 = 2U_1 - 1$ and $V_2 = 2U_2 - 1$
Step 3. If $S = V_1^2 + V_2^2 > 1$, then go back to Step 1.
Step 4. Return the standard independent normal random draws $X = \sqrt{\frac{-2\ln S}{S}}V_1$ and $Y = \sqrt{\frac{-2\ln S}{S}}V_2$

5 Relationship with Chi-Square Distribution

Given two iid normal variables X and Y obtained via Marsaglia polar method, we show that the square of their sum is sampled from a chi-square distribution with 2 degrees of freedom. First we demonstrate that the square of a standard normal variable may be be expressed as $X^2 \sim \chi^2(k=1)$. Given standard normal variable X we define Z as a random variable such that

$$Z = X^2$$

The cumulative distribution function of Z is then given by

$$F_Z(z) = P(Z \leq z)$$

$$F_Z(z) = P(X^2 \leq z) = P(-z^{1/2} \leq X \leq z^{1/2})$$

This can otherwise be expressed as

$$F_Z(z) = \int_{-z^{1/2}}^{z^{1/2}} f_x(x)dx$$

$f_x(x)$ is the PDF of a standard normal variable, and so given by

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

We will now show that the PDF of Z, $f_z(z)$ is equal to that of a Chi-square distribution.

$$f_z(z) = \frac{d}{dz} F_z(z) = \frac{d}{dz} \int_{-z^{1/2}}^{z^{1/2}} f_x(x) dx$$

The Leibniz Integral Rule allows us to evaluate the integral as

$$f_x(\sqrt{z})(\sqrt{x}\frac{d}{dx}) - f_x(-\sqrt{z})(-\sqrt{x}\frac{d}{dx})$$

Which can be simplified to

$$\frac{1}{\sqrt{2\pi}\sqrt{z}} \exp\left(\frac{-z}{2}\right)$$

Since $\Gamma(1/2) = \sqrt{\pi}$.

The PDF of the chi-squared distribution is defined as

$$f(x;k) = \begin{Bmatrix} \frac{x^{\frac{k}{2}-1} e^{\frac{-x}{2}}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}, & \text{if } x > 0; \\ 0, & otherwise \end{Bmatrix}$$

With degrees of freedom parameter k set to 1, the PDF of the chi-squared distribution simplifies to

$$f(x;1) = \frac{x^{\frac{-1}{2}} e^{-x/2}}{\sqrt{2\pi}} = f_z(z)$$

Thus, the PDF of a squared normal distribution is equal to the PDF of a chi-square distribution with 1 degree of freedom.

It can be shown that the sum of chi-squared random variables is also a chi-squared random variable. This implies that the sum of two squared normal variables, as obtained via inverse transform sampling, is also a chi-square distribution (with $k = 2$). A chi-squared distribution with k=2 degrees of freedom is defined by a PDF of

$$f(x;2) = \frac{e^{\frac{-x}{2}}}{2}$$

An exponential distribution parameterized with $\lambda = \frac{1}{2}$ is given by

$$f(x;\frac{1}{2}) = \frac{1}{2} e^{\frac{-x}{2}}$$

Consequently the sum of the two iid normal variables generated by Polar method, $X^2 + Y^2$, may be represented by $Q \sim \chi^2(k = 2) \equiv Exp(\lambda = \frac{1}{2})$. By repeatedly generating random normal variables using the method described above (or by extending the process to some region larger than $\mathbb{R}^2$), we can map to arbitrary parameterizations of the chi-square distribution.

References

[1] Ross, S. M. (2019). *A first course in probability* (10th ed., pp. 450–465). Pearson.

[2] Weisstein, Eric W. "Gaussian Integral." From *MathWorld*–A Wolfram Web Resource.https://mathworld.wolfram.com/GaussianIntegral.html

[3] Barros, L. M. (2020, January 28). Box-Muller Transform. Box. Retrieved March 4, 2023, from https://stackedboxes.org/2020/01/28/box-muller-transform/

[4] Kim, J. (2015, January 22). *The principles for how to generate random samples from a Gaussian distribution - Rhea.* www.projectrhea.org; Project Rhea.https://www.projectrhea.org/rhea/index.php/The_principles_for_how_to_generate_random_samples_from_a_Gaussian_distribution

[5] Vandrevala, C. (2015, May 10). The Box-Muller Algorithm. The box-muller algorithm. Retrieved February 25, 2023, from https://blog.cupcakephysics.com/computational20physics/2015/05/10/the-box-muller-algorithm.html