Invers CDF of Continuous Distribution

We know that the CDF of a continuous distribution is not always in the form of an elementary function. In such cases, we can use numerical methods to find the inverse of the CDF.

The Newton-Raphson method is one approach to finding the inverse of the CDF. It is an iterative method used to find the root of a function. The formula for the Newton-Raphson method is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{1}$$

where x_{n+1} is the next approximation, x_n is the current approximation, $f(x_n)$ is the function value at x_n, and $f'(x_n)$ is the derivative of the function at x_n.

Let's consider any continuous distribution X with the CDF $F(x)$,

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)\, dt \tag{2}$$

Next, we set the CDF to be R and aim to find x such that $F(x) = R$. We can rewrite the equation as

$$F(x) - R = 0 \tag{3}$$

Thus, we define a new function $g(x) = F(x) - R$. Applying the Newton-Raphson method (1), we get the iterative formula as

$$x_{n+1} = x_n - \frac{F(x_n) - R}{f(x_n)} \tag{4}$$

Direct Transformation

We know that the PDF of the standard normal distribution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{5}$$

For any random variable X with PDF $f(x)$ and a function $u(X)$, the expected value of $u(X)$ is given by

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) \cdot f(x)\, dx \tag{6}$$

Next, we are interested in finding the MGF of the sum of independent random variables X_i such that $Y = \sum_{i=1}^{n} X_i$. The MGF of Y is given by

$$M_Y(t) = E[e^{tY}] = E\left[e^{t\sum_{i=1}^{n} X_i}\right] = E\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} E[e^{tX_i}] = \prod_{i=1}^{n} M_{X_i}(t)$$

Thus, we conclude that the MGF of the sum of independent random variables is the product of the MGFs of each random variable:

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) \tag{7}$$

Chi-Square and Normal Distribution

Consider independent random variables $Z_1, Z_2, \ldots, Z_n$ with $Z_i \sim N(0,1)$ for $i = 1, 2, \ldots, n$. The MGF of Z_i^2 is given by

$$M_{Z_i^2}(t) = E[e^{tZ_i^2}] = \int_{-\infty}^{\infty} e^{tz^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(t-\frac{1}{2})z^2}\, dz$$

Let $u = \sqrt{t - \frac{1}{2}} z$, then $du = \sqrt{t - \frac{1}{2}}\, dz$. Thus, we get

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{u^2} \frac{1}{\sqrt{t - \frac{1}{2}}}\, du = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{t - \frac{1}{2}}} = \frac{1}{\sqrt{1 - 2t}}$$

So, the MGF of Z_i^2 is

$$M_{Z_i^2}(t) = \frac{1}{\sqrt{1 - 2t}} \tag{8}$$

This result is similar to the MGF of a chi-square distribution with 1 degree of freedom. Therefore, we conclude that $Z_i^2 \sim \chi^2(1)$.

Now, consider the sum of independent random variables $Z_1^2, Z_2^2, \ldots, Z_n^2$. Since Z_i are independent, Z_i^2 are also independent. Then we can use (7) to find the MGF of the sum of independent random variables $Y = \sum_{i=1}^{n} Z_i^2$.

$$M_Y(t) = \prod_{i=1}^{n} M_{Z_i^2}(t) = \prod_{i=1}^{n} \frac{1}{\sqrt{1 - 2t}} = \left(\frac{1}{\sqrt{1 - 2t}} \right)^n \tag{9}$$

Equation (9) is the MGF of a chi-square distribution with n degrees of freedom. Thus, we conclude that

$$\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n) \tag{10}$$

Relation with Uniform Distribution

The PDF of a chi-square distribution with n degrees of freedom is given by

$$f(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)} \tag{11}$$

Let $B^2 = Z_1^2 + Z_2^2 \sim \chi^2(2)$ from (10). Then, from (11), the CDF of B^2 is given by

$$F_{B^2}(x) = P(B^2 \leq x) = \int_{-\infty}^{x} \frac{e^{-t/2}}{2}\, dt = 1 - e^{-x/2} \tag{12}$$

Now, on the other hand, let $R \sim U(0,1)$. Then the CDF of R is

$$F_R(x) = P(R \leq x) = x \tag{13}$$

If we want to understand the relation between B^2 and R, we can start from (13) to match (12). Manipulating equation (13):

$$P(R > x) = 1 - P(R \leq x) = 1 - x$$

Substitute x with $e^{-t/2}$, then we get

$$\begin{aligned} P(R > e^{-t/2}) &= 1 - e^{-t/2} \\ P(\ln R > -t/2) &= 1 - e^{-t/2} \\ P(-2\ln R < t) &= 1 - e^{-t/2} \end{aligned} \tag{14}$$

We see that (14) matches (12). Thus, we conclude that

$$B = \sqrt{-2\ln R} \tag{15}$$

Polar Coordinates

Since $B^2 = Z_1^2 + Z_2^2$ resembles the equation of a circle, we can interpret B as the radius of a circle centered at $(0,0)$ with radius B. So we can construct relation between B and Z_1, Z_2 by the polar coordinate of circle

$$Z_1 = B\cos(\theta) \tag{16}$$

$$Z_2 = B\sin(\theta) \tag{17}$$

If $R_1, R_2 \sim U(0,1)$ then from (15) we can get $B = \sqrt{-2\ln R_1}$. In other hand, for $0 \leq \theta \leq 2\pi$ then rewrite the equation as $\theta = 2\pi R_2$ and its obvious because $0 \leq R_2 \leq 1$. So the random variables from (16) and (17) can be formulated as

$$Z_1 = \sqrt{-2\ln R_1}\cos(2\pi R_2) \tag{18}$$

$$Z_2 = \sqrt{-2\ln R_1}\sin(2\pi R_2) \tag{19}$$