Inverse Random Sampling for Conditional Weibull and Pareto Distribution

Tuan Nguyen Doan

January 16, 2021

There are times where it is obviously hard to randomly sampling from a particular distribution, even though we may have a precise analytical formulation of the distribution. In this errata, I look at **Inverse Random Sampling** as an easy way to draw random sample for cases where the analytical form of the distribution i.e. precise pdf, cdf are relatively not too complicated.

The trick here is to use an idea called **probability integral transform (PIT)**, which states that if X is a continuous random variable with a cdf $F(X)$, then the random variable $Y = F_X(X)$ has a uniform distribution on $[0, 1]$.

Inversely, if $Y \sim U[0, 1]$ and if X has a cumulative distribution F_X, then the random variable $F_X^{-1}(Y)$ has the same distribution as X.

1 Simple case with Exponential

Let's start with a vanilla exponential distribution, whose pdf is $f(t) = \lambda e^{(\lambda t)}$ and cdf is $F(t) = 1 - e(\lambda t)$.

We want to sample from this distribution. We know that by **PIT**, whatever the value of t, we get that $F(t) = 1 - e(\lambda t) = u$ follows a $U[0, 1]$ distribution. With a bit of manipulation, we have:

$$u = 1 - e(\lambda t)$$

$$\frac{1}{-\lambda}\log(1 - u) = t$$

Now we can simulate values from the exponential distribution by sampling random values from $U[0, 1]$ and plugging them into u from the equation above.

2 Conditional Weibull simulation

In my article, I talked about randomly sampling time remaining in human life after an t_1 years, where human life time follows a Weibull distribution with shape parameter κ and scale parameter λ.

Weibull distribution has the property that:

$$P(X > t) = e^{-(\frac{t}{\lambda})^k}$$

The reader can verify that the distribution of **total** human life $X = t_1 + t$ time conditional to $X > t_1$ is:

$$P(X > t_1 + t | X > t_1) = e^{-\frac{1}{\lambda^\kappa}((t_1+t)^k - t_1^k)}$$

$$F(X = t_1 + t | X > t_1) = 1 - e^{-\frac{1}{\lambda^\kappa}((t_1+t)^k - t_1^k)}$$

Similar to above, we let $F(X) = U$ where U is uniform on (0,1) and take the inverse $X = F^{-1}(U)$:

$$X = t + t_1 = (t_1^\kappa - \lambda^\kappa \log(1 - U))^{\frac{1}{\kappa}}$$

so the remaining time life t is simply $(t_1^\kappa - \lambda^\kappa \log(1 - U))^{\frac{1}{\kappa}} - t_1$

3 Conditional Pareto simulation

If X follows the Pareto distribution with shape parameter α and minimum possible value t_m, then the event that it is greater than or equal to a particular number $t_1 > t_m$, is a Pareto distribution with the same Pareto index α but with minimum t_1 instead of t_m.

The conditional cdf would thus be:

$$F(X = t_1 + t | X > t_1) = 1 - (\frac{t_1}{x})^\alpha$$

Again, all we do is equate to uniform distribution and solve for t:

$$u = 1 - (\frac{t_1}{x})^\alpha$$

$$x = t + t_1 = t_1(1 - u)^{-\frac{1}{\alpha}}$$

$$t = t_1(1 - u)^{-\frac{1}{\alpha}} - t_1$$