

1 Inverse Transform Method

Assuming our computer can hand us, upon demand, iid copies of rvs that are uniformly distributed on $(0, 1)$, it is imperative that we be able to use these uniforms to generate rvs of any desired distribution (exponential, Bernoulli etc.). The first general method that we present is called the inverse transform method.

Let $F(x)$, $x \in \mathbb{R}$, denote any cumulative distribution function (cdf) (continuous or not). Recall that $F : \mathbb{R} \rightarrow [0, 1]$ is thus a non-negative and non-decreasing (monotone) function that is continuous from the right and has left hand limits, with values in $[0, 1]$; moreover $F(\infty) = 1$ and $F(-\infty) = 0$. Our objective is to generate (simulate) rvs X distributed as F ; that is, we want to simulate a rv X such that $P(X \leq x) = F(x)$, $x \in \mathbb{R}$.

Define the generalized inverse of F , $F^{-1} : [0, 1] \rightarrow \mathbb{R}$, via

$$(1) \quad F^{-1}(y) = \min\{x : F(x) \geq y\}, \quad y \in [0, 1].$$

If F is continuous, then F is invertible (since it is thus continuous and strictly increasing) in which case $F^{-1}(y) = \min\{x : F(x) = y\}$, the ordinary inverse function and thus $F(F^{-1}(y)) = y$ and $F^{-1}(F(x)) = x$. In general it holds that $F^{-1}(F(x)) \leq x$ and $F(F^{-1}(y)) \geq y$. $F^{-1}(y)$ is a non-decreasing (monotone) function in y .

This simple fact yields a simple method for simulating a rv X distributed as F :

Proposition 1.1 (The Inverse Transform Method) *Let $F(x)$, $x \in \mathbb{R}$, denote any cumulative distribution function (cdf) (continuous or not). Let $F^{-1}(y)$, $y \in [0, 1]$ denote the inverse function defined in (1). Define $X = F^{-1}(U)$, where U has the continuous uniform distribution over the interval $(0, 1)$. Then X is distributed as F , that is, $P(X \leq x) = F(x)$, $x \in \mathbb{R}$.*

Proof: We must show that $P(F^{-1}(U) \leq x) = F(x)$, $x \in \mathbb{R}$. First suppose that F is continuous. Then we will show that (equality of events) $\{F^{-1}(U) \leq x\} = \{U \leq F(x)\}$, so that by taking probabilities (and letting $a = F(x)$ in $P(U \leq a) = a$) yields the result: $P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$.

To this end: $F(F^{-1}(y)) = y$ and so (by monotonicity of F) if $F^{-1}(U) \leq x$, then $U = F(F^{-1}(U)) \leq F(x)$, or $U \leq F(x)$. Similarly $F^{-1}(F(x)) = x$ and so if $U \leq F(x)$, then $F^{-1}(U) \leq x$. We conclude equality of the two events as was to be shown. In the general (continuous or not) case, it is easily shown that

$$\{U < F(x)\} \subseteq \{F^{-1}(U) \leq x\} \subseteq \{U \leq F(x)\},$$

which yields the same result after taking probabilities (since $P(U = F(x)) = 0$ since U is a continuous rv.) ■

1.1 Examples

The inverse transform method can be used in practice as long as we are able to get an explicit formula for $F^{-1}(y)$ in closed form. We illustrate with some examples. We use the notation $U \sim \text{unif}(0, 1)$ to denote that U is a rv with the continuous uniform distribution over the interval $(0, 1)$.

1. *Exponential distribution:* $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, where $\lambda > 0$ is a constant. Solving the equation $y = 1 - e^{-\lambda x}$ for x in terms of $y \in (0, 1)$ yields $x = F^{-1}(y) = -(1/\lambda) \ln(1 - y)$. This yields $X = -(1/\lambda) \ln(1 - U)$. But (as is easily checked) $1 - U \sim \text{unif}(0, 1)$ since $U \sim \text{unif}(0, 1)$ and thus we can simplify the algorithm by replacing $1 - U$ by U :

Algorithm for generating an exponential rv at rate λ :

- i Generate $U \sim \text{unif}(0, 1)$.
 - ii Set $X = -\frac{1}{\lambda} \ln(U)$.
2. *Discrete rvs: discrete inverse-transform method:* Consider a non-negative discrete rv X with probability mass function (pmf) $p(k) = P(X = k)$, $k \geq 0$. In this case, the construction $X = F^{-1}(U)$ is explicitly given by: $X = 0$ if $U \leq p(0)$,

$$X = k, \text{ if } \sum_{i=0}^{k-1} p(i) < U \leq \sum_{i=0}^k p(i), \quad k \geq 1.$$

This is known as the *discrete inverse-transform method*. The algorithm is easily verified directly by recalling that $P(a < U \leq b) = b - a$, for $0 \leq a < b \leq 1$; here we use $a = \sum_{i=0}^{k-1} p(i) < b = \sum_{i=0}^k p(i)$, and so $b - a = p(k)$.

3. *Bernoulli (p) and Binomial (n, p) distributions:*

Suppose we want to generate a Bernoulli (p) rv X , in which case $P(X = 0) = 1 - p$ and $P(X = 1) = p$ for some $p \in (0, 1)$. Then the discrete inverse-transform method yields:

Algorithm for generating a Bernoulli (p) rv X :

- i Generate $U \sim \text{unif}(0, 1)$.
- ii Set $X = 0$ if $U \leq 1 - p$; $X = 1$ if $U > 1 - p$.

Suppose we want X to have a binomial (n, p) distribution, that is,

$$p(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad 0 \leq k \leq n.$$

One could, in principle, use the discrete inverse-transform method with these $p(k)$, but we also can note that X can be represented (in distribution) as the sum of n iid Bernoulli (p) rvs, Y_1, \dots, Y_n ;

$$X = \sum_{i=1}^n Y_i,$$

the number of successes out of n iid Bernoulli (p) trials.

Alternative algorithm for generating a binomial (n, p) rv X :

- i Generate n iid rvs $U_1, U_2, \dots, U_n \sim \text{unif}(0, 1)$.
- ii For each $1 \leq i \leq n$, set $Y_i = 0$ if $U_i \leq 1 - p$; $Y_i = 1$ if $U_i > 1 - p$. (This yields n iid Bernoulli (p) rvs.)
- iii Set $X = \sum_{i=1}^n Y_i$.

The advantage of this algorithm is its simplicity, we do not need to do the various computations involving the $p(k)$. On the other hand, this algorithm requires n uniforms for each copy of X versus only one uniform when using the discrete inverse-transform method. Thus we might not want to use this algorithm when n is quite large.

In fact, when n is very large, and p is small, it follows (e.g., can be proved; there is a theorem lurking here), that the distribution of X is very approximately the Poisson distribution with mean np . This motivates our next example.

4. *Poisson distribution with mean α :*

In this case

$$p(k) = P(X = k) = e^{-\alpha} \frac{\alpha^k}{k!}, \quad k \geq 0.$$

We could thus use the discrete inverse-transform method, but of course it involves computing (in advance) pieces like $\frac{\alpha^k}{k!}$. Here we present an alternative algorithm that makes use of properties of a Poisson process at rate α . The trick is to recall that if $\{N(t) : t \geq 0\}$ is the counting process of a Poisson process at rate α , then $N(1)$ has a Poisson distribution with mean α . Thus if we can simulate $N(1)$, then we can set $X = N(1)$ and we are done. Let $Y = N(1) + 1$, and let $t_n = X_1 + \dots + X_n$ denote the n^{th} point of the Poisson process; the X_i are iid with an exponential distribution at rate α . Note that $Y = \min\{n \geq 1 : t_n > 1\} = \min\{n \geq 1 : X_1 + \dots + X_n > 1\}$, a stopping time. Using the inverse transform method to generate the iid exponential interarrival times X_i , we can represent $X_i = -(1/\alpha) \ln(U_i)$. We then can re-write (recalling that $\ln(xy) = \ln(x) + \ln(y)$)

$$\begin{aligned} Y &= \min\{n \geq 1 : \ln(U_1) + \dots + \ln(U_n) < -\alpha\} \\ &= \min\{n \geq 1 : \ln(U_1 \cdots U_n) < -\alpha\} \\ &= \min\{n \geq 1 : U_1 \cdots U_n < e^{-\alpha}\} \end{aligned}$$

We thus can simulate Y by simply consecutively generating independent uniforms U_i and taking the product until the product first falls below $e^{-\alpha}$. The number of uniforms required yields Y . Then we get $X = N(1) = Y - 1$ as our desired Poisson. Here then is the resulting algorithm:

Alternative algorithm for generating a Poisson rv X with mean α :

- i Set $X = 0, P = 1$
- ii Generate $U \sim \text{unif}(0, 1)$, set $P = UP$
- iii If $P < e^{-\alpha}$, then stop. Otherwise if $P \geq e^{-\alpha}$, then set $X = X + 1$ and go back to *ii*.