Inverse CDF method and probability integral transform

Jeffrey W. Miller

June 19, 2016

This instructional note covers (a) the method of generating a random variable by mapping a uniform random variable through the generalized inverse of the cumulative distribution function (CDF), and (b) the fact that mapping a continuous random variable through its CDF yields a uniform random variable.

1 The result

Let F be a CDF on $\mathbb{R}$, and let $G(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$ for $u \in (0, 1)$. The function G is referred to as the generalized inverse of F. The following result is basic and well known (Dudley, 2002, Prop 9.1.2 and Exercise 9.1.5).

Theorem.

(a) If $U \sim \mathrm{Uniform}(0, 1)$ *then* $G(U) \sim F$.

(b) If $X \sim F$ *and* F *is a continuous function, then* $F(X) \sim \mathrm{Uniform}(0, 1)$.

Before proving the theorem, we make some clarifying remarks and establish some lemmas. A CDF is a function $F : \mathbb{R} \to [0, 1]$ such that $F(x) = \mathbb{P}(X \leq x)$ for some real-valued random variable X. Equivalently, $F : \mathbb{R} \to [0, 1]$ is a CDF if and only if:

1. F is nondecreasing (i.e., $x \leq y$ implies $F(x) \leq F(y)$),
2. F is right-continuous (i.e., $F(x) = \lim_{y \searrow x} F(y)$),
3. $F(x) \to 1$ as $x \to \infty$, and
4. $F(x) \to 0$ as $x \to -\infty$.

The notation $\lim_{y \searrow x}$ denotes the limit as y converges to x from above. Note that $G(u) \in \mathbb{R}$ for any $u \in (0, 1)$, by items 3 and 4 above.

2 Proof

Lemma 2.1. *For any $u \in (0,1)$, we have $u \leq F(G(u))$.*

Proof. Let $u \in (0,1)$. Since $G(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$ and $G(u) \in \mathbb{R}$, there exists a sequence $x_1, x_2, \ldots \geq G(u)$ such that $F(x_n) \geq u$ for all n and $x_n \to G(u)$ as $n \to \infty$. Therefore, $u \leq \lim_{n\to\infty} F(x_n) = F(G(u))$ since F is right-continuous. □

Lemma 2.2. *Let $u \in (0,1)$ and $x \in \mathbb{R}$. Then $u \leq F(x)$ if and only if $G(u) \leq x$.*

Proof. If $u \leq F(x)$ then $G(u) \leq x$ by the definition of G. Meanwhile, if $G(u) \leq x$ then $u \leq F(G(u)) \leq F(x)$ by Lemma 2.1 and since F is nondecreasing. □

Lemma 2.3. *If F is continuous, then $F(G(u)) = u$ for any $u \in (0,1)$.*

Proof. Let $u \in (0,1)$. By Lemma 2.1, $F(G(u)) \geq u$. So, we just need to show that $F(G(u)) \leq u$. Let $x \in \mathbb{R}$ such that $F(x) = u$, which we can do by the Intermediate Value Theorem (Rudin, 1976, 4.23) since F is continuous, $F(x) \to 0$ as $x \to -\infty$, and $F(x) \to 1$ as $x \to \infty$. Then $G(u) \leq x$ by Lemma 2.2. Hence, $F(G(u)) \leq F(x) = u$. □

Proof of theorem. (a) Suppose $U \sim \text{Uniform}(0,1)$. Then for any $x \in \mathbb{R}$,

$$\mathbb{P}(G(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x),$$

where the first step is by Lemma 2.2 and the second step is since $\mathbb{P}(U \leq u) = u$ for any $u \in [0,1]$. Therefore, $G(U) \sim F$.

(b) Suppose $X \sim F$ and F is continuous. Let $U \sim \text{Uniform}(0,1)$. Then $G(U)$ has the same distribution as X by part (a). Thus, for any $u \in (0,1)$,

$$\mathbb{P}(F(X) \leq u) = \mathbb{P}(F(G(U)) \leq u) = \mathbb{P}(U \leq u) = u,$$

where the second step is by Lemma 2.3. Therefore, $F(X) \sim \text{Uniform}(0,1)$. □

References

R. M. Dudley. *Real Analysis and Probability.* Cambridge University Press, 2002.

W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill New York, 1976.