The Probability Transformation and Simulation

H. Krieger, Mathematics 156, Harvey Mudd College

Fall, 2008

The Basic Lemma: Let X be a random variable with cumulative distribution function F, (usually written $X \sim F$). Then

$$\forall x \in \mathbf{R}, \quad P(F(X) \leq F(x)) = F(x).$$

Proof: Since every distribution function is non-decreasing, we see that

$$\{F(X) \leq F(x)\} = \{X \leq x\} \cup \{X > x,\ F(X) = F(x)\}$$

and these events are mutually exclusive. But

$$P(X > x,\ F(X) = F(x)) = 0.$$

Consequently,

$$P(F(X) \leq F(x)) = P(X \leq x) = F(x).$$

Corollary (The Probability Transform): Let $X \sim F$ with F *continuous.* Then $F(X)$ has a uniform distribution on the interval $(0,1)$ (usually written $F(X) \sim U(0,1)$).

Proof: If F is continuous and $0 < p < 1$, $\exists x \in \mathbf{R}$ such that $F(x) = p$. But then,

$$P(F(X) \leq p) = P(F(X) \leq F(x)) = F(x) = p.$$

Simulation Lemma: Let $U \sim U(0,1)$ and let F be a distribution function. For $0 < p < 1$, define

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}.$$

Note that this inf is actually attained, *i.e.* a minimum, since a distribution function is continuous from the right. Then $X = F^{-1}(U) \sim F$.

Proof: If $x \in \mathbf{R}$ and $U \leq F(x)$, then $X = F^{-1}(U) \leq x$. On the other hand, from the definition of F^{-1}, we see that $U \leq F(F^{-1}(U))$. Consequently, if $F^{-1}(U) = X \leq x$, then

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

Therefore, the events $\{U \leq F(x)\}$ and $\{X \leq x\}$ are identical and, hence,

$$F(x) = P(U \leq F(x)) = P(X \leq x).$$

The Percentile Function: Let $X \sim F$. The function R defined on $(0,1)$ by $R(p) = F^{-1}(p)$ is called the *percentile function* of the random variable X. The Simulation Lemma tells us that an explicit knowledge of the percentile function for a random variable X reduces the problem of sampling from the distribution of X to that of sampling from a $U(0,1)$ distribution. In other words, if $U \sim U(0,1)$, then $R(U) \sim X$.

If F is continuous and strictly increasing on an open interval $(a,b) \subset \mathbf{R}$, where either a could be $-\infty$ or b could be $+\infty$ or both, with $\lim_{x\to a} F(x) = 0$ and $\lim_{x\to b} F(x) = 1$, then R becomes the ordinary inverse function of F on (a,b). If, in addition, F is absolutely continuous with derivative f, which is the probability density function describing this distribution, then the equation

$$F(R(p)) = p$$

and the chain rule show that

$$f(R(p)) = [R'(p)]^{-1}.$$

Therefore, the graph of the density function is given parametrically by

$$\{(R(p), f(R(p))) : p \in (0,1)\} = \{(R(p), [R'(p)]^{-1}) : p \in (0,1)\},$$

i.e. explicitly in terms of the percentile function and its derivative.

Another useful fact about percentile functions is that the moments of the distribution they describe can also be expressed explicitly. To see this, observe that if $X \sim R(U)$, where $U \sim U(0,1)$, then

$$E(|X|^k) = E(|R(U))|^k) = \int_0^1 |R(p)|^k dp.$$

Consequently, if this expectation is finite, we have for the kth moment the expression

$$E(X^k) = \int_0^1 [R(p)]^k dp.$$