1 A Note on how to randomly sample from the $f(x)$ distribution

erm, Oct 26, 2009

Many software packages have commands to draw random samples from the standard normal distribution and the uniform distribution on the 0 to 1 (hereafter the unit uniform).

However, commands and software do not generally exist for taking random samples from other distributions.

1.1 Assume the problem is take a random sample from the $f_X(x)$ distribution where one has the ability (e.g. Mathematica) to take a random sample from the uniform distribution on the 0 - 1 interval.

Let $F_X(X)$ denote the cdf of $f_X(X)$

One can show that one can obtain a random draw from $f_X(X)$ by obtaining a random draw from the unit uniform distribution, $f_U(U)$, then plugging that draw into $F_X^{-1}(U)$ where $F_X^{-1}(.)$ is the inverse of $F_X(X)$. That is, $F_X^{-1}(U)$ is $h = F_X(X)$ solved for X and then evaluated at U.

1.1.1 Example 1:

.

Take a random sample from the (negative) exponential distribution (MGB 3rd edition, p. 112). If X has a negative exponential distribution

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & if \quad 0 \leq x \leq \infty \\ 0 & if \quad x < 0 \end{cases}$$

where $\lambda > 0$. The corresponding cdf is

$$h = F_X(x) = 1 - e^{-\lambda x}$$

Solve $h = 1 - e^{-\lambda x}$ for x, Solution is: $\left\{x = -\frac{\ln(-h+1)}{\lambda}\right\}$. So

$$F_X^{-1}(u) = -\frac{\ln(-u+1)}{\lambda}$$

In words, if u is a random draw from the unit uniform, then $-\frac{\ln(-u+1)}{\lambda}$ is a random draw from the exponential distribution.[1]

One draws a random sample of N observations from the unit uniform, then converts each using the formula $-\frac{\ln(-u+1)}{\lambda}$

1.1.2 Example 2:

I often use this "rule" to take to draw a random sample from a population with an Extreme Value distribution. The cdf for the Extreme Value distribtuion is

$$F_X(x) = \Pr(X \leq x) = \exp^{-e^{-x}}$$

That is $h = \exp^{-e^{-x}} \Rightarrow \ln h = -e^{-x} \Rightarrow -\ln h = e^{-x} \Rightarrow \ln(-\ln h) = -x \Rightarrow$

$$x = -\ln(-\ln h)$$

So, if if u is a random draw from the unit uniform, then $x = -\ln(-\ln u)$ is a random draw from the Extreme Value distribution.

1.2 Some Background

This method for taking random samples is an application of Theorem 12 (sec 5.2) of MGB page 202. It is called the Probability Integral Transform.

Theorem 12 says "If X is a random variable with a continuous cummulative distribution function $F_X(x)$, then $U = F_X(X)$ is uniformly distribution over the interval $(0,1)$. Conversely, if U is uniformly distributed on the interval $(0,1)$, then $X = F_X^{-1}(U)$ has the cummulative distribution function $F_X(.)$.

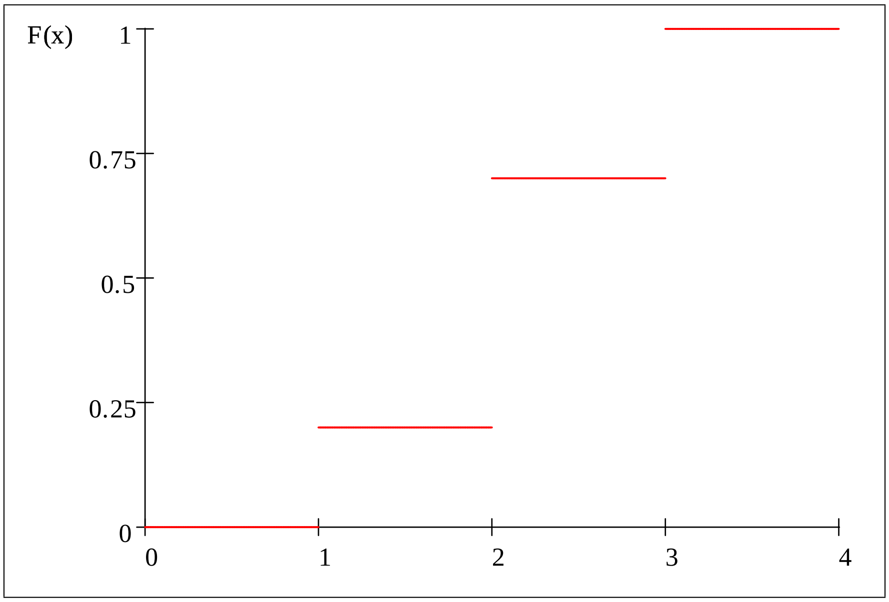
1.3 Does this technique work if $f_X(x)$ is a discrete distribution?

Consider the following example:

Assume $f(x) = .2$ if $x = 1$, $f(x) = .5$ if $x = 2$, $f(x) = .3$ if $x = 3$, and zero otherwise. Picture the df and cdf of this discrete distributino

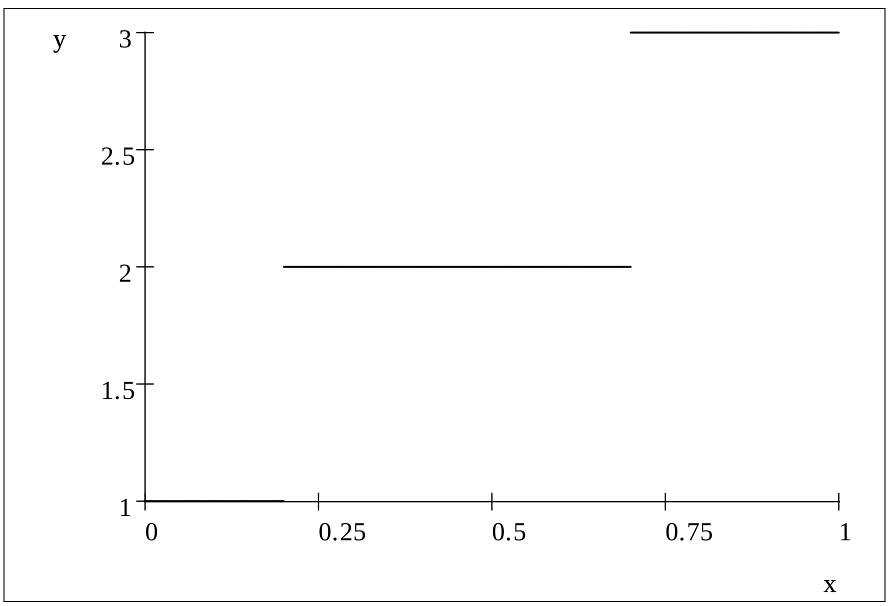
$$\begin{cases} 0 & \text{if} \quad x < 1 \\ .2 & \text{if} \quad 1 \leq x < 2 \\ .7 & \text{if} \quad 2 \leq x < 3 \\ 1 & \text{if} \quad x \geq 3 \end{cases}$$

[1] Note that since u is between zero and one, $\ln(-u+1) < 0$, so $-\frac{\ln(-u+1)}{\lambda} > 0$.



So, one way to take a random sample from this density function is to invert the cdf to get

$$\begin{cases} x = 1 & \text{if} \quad 0 < y \leq .2 \\ x = 2 & \text{if} \quad .2 < y \leq .7 \\ x = 3 & \text{if} \quad y > .7 \end{cases}$$



So, when we take a random sample from the unit uniform density, we transform them into 1's, 2's and 3's using the above rule. What did I just demonstrate?

That the *probability integral transform* works, at least sometimes, even if the density function is discrete.

1.4 Does this technique work if $f_X(x)$ is a mixture of a continous and discrete distributions?