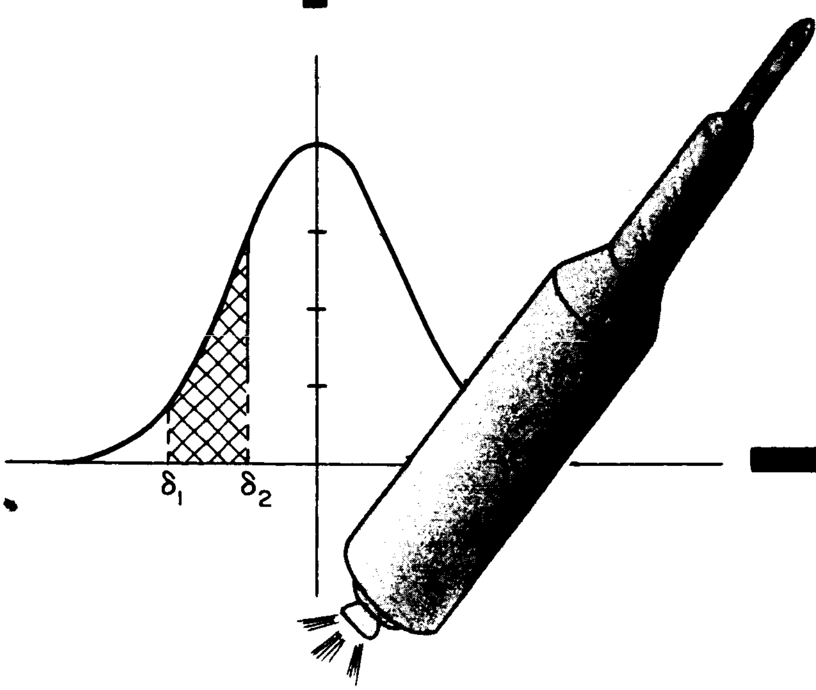
ON THE GENERATION OF PSEUDO-RANDOM NUMBERS FROM SEVERAL NON-UNIFORM DISTRIBUTIONS



TEXAS CENTER FOR RESEARCH
3100 PERRY LANE, AUSTIN, TEXAS

ON THE GENERATION OF PSEUDO-RANDOM NUMBERS FROM SEVERAL NON-UNIFORM DISTRIBUTIONS

by

Rosser J. Smith, III

I

Introduction

In many applications, it is either costly or difficult to secure a large random sample from a given statistical population. Consequently, it is often efficacious to use some kind of numerical simulation, or Monte Carlo Method. This usually requires a source of so-called "random numbers." Producing these random[1] numbers can be an inexpensive, simple process, thus facilitating the solution of one's problem or investigation.

It is the purpose of this paper to present several convenient methods for generating random numbers representing several of the fundamental statistical distributions. Most of the methods to be shown here are readily adapted to automated computation, and several will be given which are suitable for manual computation, where only a moderate sample size is required. It will be seen, however, that most of the methods require a supply of random numbers having a uniform distribution over the interval [0,1).

Several convenient methods for the generation and testing of uniformly distributed random numbers have been developed. These methods are described elsewhere [2,3,10] in detail[2], and so the balance of this paper will be devoted to methods for generating random numbers from other distributions.

---

1) These numbers are, strictly speaking, pseudo-random, because some sort of deterministic process generates them.

2) The generation, testing, and application of random digits are described in [9].

II

Methods for Generating Random Numbers

Since it is a relatively simple matter to generate uniformly distributed random numbers, it would be worthwhile to consider methods for transforming them into random numbers from other distributions. That is, one seeks a transformation T which utilizes a set $U = \{u_1, u_2, \ldots, u_n\}$ of independent, uniform variates to produce a set $X = T(U) = \{x_1, x_2, \ldots, x_m\}$ of independent variates from some other distribution. The transformation T is often found by equating the distribution functions of X and U.

Suppose Y is a random variable with distribution function H, and suppose one seeks a random variable $Z = z(Y)$ which will have distribution function K. That is, one requires that

$$K(z) = H(y) \tag{1}$$

for every admissable value of Y, so that

$$z(Y) = K^{-1}[H(y)] \tag{2}$$

is the desired transformation. Thus, if one can invert the distribution function K, it is possible to find a function of the random variable Y which will have the desired distribution.

As an example, suppose it is desired to produce a random variable X with an exponential density function,

$$k(x) = \theta e^{-\theta x}, \tag{3}$$

where $x \geq 0$ and $\theta > 0$. The distribution function of X is, therefore,

$$K(x) = \int_0^x k(t)dt = 1 - e^{-\theta x} \tag{4}.$$

Here, let Y denote a random variable with a uniform distribution over $[0,1)$, so that the density function of Y is

$$h(y) = 1,\ 0 \leq y < 1, \tag{5}$$

and the distribution function of Y is

$$H(y) = y \tag{6}.$$

To find the desired function z so that $X = z(Y)$ will have the exponential distribution, let

$$K(x) = H(y)$$

or
$$1 - e^{-\theta x} = y \tag{7}$$

so that
$$x = z(y) = \frac{-1}{\theta} \ln(1 - y) \tag{8}.$$

If one wishes to use (8) many times, the time saved by using u, rather than $1 - u$, could be considerable. For this reason, it is interesting to note that u and 1 - u are identically distributed. This is seen from a comparison of the Moment Generating Functions of u and 1 - u, respectively. That is, the Moment Generating Function of u is

$$M_u(t) = E[e^{ut}] = \int_0^1 e^{ut}\, du = (e^t - 1)/t,$$

where, for some $h > 0$, it is required that $-h < t < h$. Then, for $1 - u$,

$$M_{1-u}(t) = E[e^{(1-u)t}] = E[e^t\, e^{u(-t)}]$$

$$= e^t M_u(-t) = e^t \frac{e^{-t}-1}{-t} = \frac{e^t-1}{t},$$

so $M_u(t) = M_{1-u}(t)$, and therefore u and 1 - u are identically distributed. Thus, (8) can be rewritten as

$$x = -\frac{1}{\theta} \ln u. \qquad (9)$$

It is to be noted that inverting the distribution function is not always efficacious. As an example, suppose one seeks to use a uniform variate u to produce a standard normal variate X. As in (7), one equates the distribution functions,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt = u$$

Clearly this does not provide a simple way to write $X = x(u)$. One must turn to numerical methods for such a case as this. (A numerical method has been described [8] to solve the above; it requires a fairly elaborate set of Chebyshev polynomials and requires more computation time than the methods to be given here.)

III

The Generation of Normal Variates

The normal distribution plays a fundamental role in the theory of statistics. A random variable X having this distribution, with mean μ and variance σ^2, has desnity function

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp[-(x - \mu)^2/(2\sigma^2)], \quad -\infty < x < \infty,$$

with $-\infty < \mu < \infty$ and $\sigma^2 > 0$. A very useful variate occurs for $\mu = 0$ and $\sigma^2 = 1$ and is said to be standardized. It is easily shown [7, p.124] that if X is a standardized normal variate, then $Y = \sigma X + \mu$ is a normal variate with mean μ and variance σ^2. Consequently, if one can generate a standardized normal variate, it is a simple matter to secure a normal variate with mean μ and variance σ^2, and so no loss of generality will occur in considering only standardized normal variates.

Since one is often interested in producing normal variates in the most expedient way, the central limit theorem has great appeal. It can be stated as follows [7, p.149]:

"Let $f(x)$ be a density with mean μ and finite variance σ^2.

Let $$\bar{X}_n = 1/n \sum_{i=1}^{n} x_i$$

be the mean of a random sample of size n from $f(x)$. If the random variable y_n is defined by

$$y_n = \frac{\bar{X}_n - \mu}{\sigma} \sqrt{n}, \tag{11}$$

then the density of y_n approaches the normal with mean zero and unit variance, as n increases without bound." Consequently, for sufficiently large sample size n, one can produce standardized normal variates (11) from any variates X with mean μ and finite variance σ^2. If u is a uniform variate with density function (5), then $\mu = 1/2$ and $\sigma^2 = 1/12$, and (11) becomes

$$y_n = \frac{\bar{u}_n - 1/2}{1/\sqrt{12}} \sqrt{n}$$

$$= \sqrt{3/n} \left(2\sum_{i=1}^{n} u_i - n\right) \qquad (12)$$

so that for large n, $\sum_{i=1}^{n} u_i$ has a normal distribution with mean $n/2$ and variance $n/12$. It has been noted [8] that $n = 12$ is a convenient choice[1], because then the desired unit variance would occur easily. Testing [8], however, has shown that this value is too small; while for $n \geq 50$ it has been observed that the form of $f(x)$ in the central limit theorem has little effect on the fidelity of the approximation. The figure below illustrates the distribution of 100 samples of size $n = 10$ from the uniform distribution (5).

1) In [8], it is noted that the IBM 704 requires about five milliseconds to produce a single normal variate from twelve uniform variates using this method.

Because of its time consumption and poor accuracy, the central limit theorem does not find wide application in the generation of normal variates. For those who require only moderate sample size, however, it can be very useful, or one can consult tables such as [9].

One can approach the problem of securing normal variates from uniform ones in a more elegant and accurate manner than that shown above.

Suppose that (u_1,u_2) denotes a pair of independent, uniform variates with density function (5). Consider the circle

$$x_1^2 + x_2^2 = \rho^2,$$

$$x_1 = \rho \mathrm{Cos}\theta$$

with

$$x_2 = \rho \mathrm{Sin}\theta,$$

so that $\theta = \text{Arctan } x_2/x_1$.

Let $\rho = -2 \ln u_1$ (13)

and $\theta = 2\pi\, u_2$,

so that $u_1 = e^{-\rho/2}$

and $u_2 = \theta/2\pi$.

The density function of ρ is

$$\phi(\rho) = h(u\ )\,|du_1/d\rho|$$

$$= 1\cdot|(-1/2)\ e^{-\rho/2}| = \frac{e^{-\rho/2}}{2}$$

$$= \frac{\rho^{(2/2)-1}\ e^{-\rho/2}}{2^{2/2}\ \Gamma(2/2)}$$

so that ρ has a Chi-Square distribution with two degrees of freedom.

Now, since $x_1^2 + x_2^2 = \rho = -2 \ln u_1$

and $\text{Arctan } x_2/x_1 = \theta = 2\pi\, u_2$,

it follows that $x_1 = (-2 \ln u_1)^{1/2} \text{Cos } 2\pi\, u_2$

and $x_2 = (-2 \ln u_1)^{1/2} \text{Sin } 2\pi\, u_2$, (14)

so that $u_1 = \exp[-1/2(x_1^2 + x_2^2)]$

and $u_2 = 1/(2\pi) \text{ Arctan } x_2/x_1$.

From this, it follows that the joint density function of x_1 and x_2

$$f(x_1,x_2) = h(u_1,u_2)\ J(u_1,u_2/x_1,x_2)$$

$$= 1\cdot 1\cdot \begin{vmatrix} \partial u_1/\partial x_1 & \partial u_1/\partial x_2 \\ \partial u_2/\partial x_1 & \partial u_2/\partial x_2 \end{vmatrix}$$

$$= \frac{e^{-(x_1^2 + x_2^2)/2}}{-2\pi} = \frac{e^{-(x_1^2/2)}}{\sqrt{2\pi}}\ \frac{e^{-(x_2^2/2)}}{\sqrt{2\pi}}$$

so that x_1 and x_2 are independent, standardized normal variates.

The method (14) shown above was originally proposed by Box and Muller [1]; and if one's computing facilities can accurately and speedily evaluate the square root, logarithm, and trigonometric functions needed, the method is very satisfactory[1]. This method is especially valuable to one who needs only a few normal variates and proposes to use published tables and a calculator. Note also that this method is accurate in the tails of the normal distribution.

by now, it has been observed that the standardized normal population is not easily simulated. There are, however, distributions which are relatively easy to simulate, such as the familar uniform distribution (5). It would be very helpful if there were some way to decompose a complicated distribution, such as the standardized

1) In [8] it has been observed that the IBM 704 would require about 6.6 milliseconds to produce a normal variate with this method. Note that x_1 and x_2 can be combined so that $Z = 2^{-1/2}(x_1+x_2) = (-2 \ln u_1)^{1/2} \operatorname{Sin}(\pi/4)(8u_2+1)$ is a standardized normal variate.

normal one, into a set of simpler ones, for then one could simulate any complicated distribution by simulating its simpler components.[1]

Let U denote the set of all distribution functions, with X some random variable,

$$U = \{G_i(x) \mid i=1,2,\ldots,m,\ldots\} \ ,$$

and let R denote a set of real numbers

$$R = \{a_i \mid a_i = P[X \sim g_i(x)]\} \ ,$$

where $X \sim g_i(x)$ means "X has density function $g_i(x)$ and distribution function $G_i(x)$." Thus, $a_i \geq 0$ and $a_1+a_2+\ldots+a_m+\ldots = 1$. Thus, if $F \varepsilon U$ is some distribution function of X, then

$$\begin{aligned} F(t) &= P[X \leq t] \\ &= \sum_i P[X \leq t,\ X \sim g_i(x)] \\ &= \sum_i P[X \leq t \mid X \sim g_i(x)] \cdot P[X \sim g_i(x)] \\ &= \sum_i a_i\ G_i(t) \end{aligned}$$

and thus the density function of X can be written[2] as

$$f(x) = F'(x) = \sum_i a_i\ G_i'(x) = \sum_i a_i\ g_i(x) \qquad (15)$$

1) See [4,5] for additional information on this subject.

2) In the case of a discrete random variable, one would use a difference $g(x_i) = G(x_i) - G(x_{i-1})$ instead of the indicated differentation to produce the same result (15).

In practice, one tries to select the elements of U which will allow there to be more than a single non-zero a_i. This will be done below, where (15) will be used to approximate the standardized normal density function

$$f(x) = \frac{e^{-(x^2/2)}}{\sqrt{2\pi}}, \quad |x| \geq 0. \tag{16}$$

Figure 1 is a sketch of the graphs of $f(x)$ and $g_1(x)$, where

$$f_1(x) = \begin{cases} (3-x^2)/8, & |x| \geq 1 \\ (3-|x|)^2/16, & 1 \leq |x| < 3 \\ 0, & |x| \geq 3 \end{cases} \tag{17}$$

and $g_1(x)$ is the density function of

$$X = 2(u_1+u_2+u_3 - 1.5) \,, \tag{18}$$

where u_1, u_2 and u_3 are independent, uniform variates with density function (5). Note that $g_1(x)$ closely approximates $f(x)$ for $-3 < x < 3$. It is clear, however, that (18) is not a true normal variate; some sort of correction is needed.

Since (18) is generated easily, it is desirable that it be used as often as possible. That is, one desires a_1 in (15) to be as large as possible. One seeks the largest a_1 such that

$$\int_{-3}^{3}[f(x) - a_1 g_1(x)]dx \geq 0$$

is minimized. This reduces to minimizing

$$f(x) - a_1 g_1(x) \geq 0$$

for $-3 < x < 3$. That is, since (19) implies that

$$a_1 \leq f(x)/g_1(x)$$

one simply finds the minimal value of the ratio $f(x)/g_1(x)$ and gives this value to a_1.

Note that $$\frac{d}{dx}\left(\frac{f(x)}{g_1(x)}\right) = -\frac{f(x)[g_1'(x) + xg_1(x)]}{g_1^2(x)} = 0$$

for $|x| = 0,1,2$. The minimum value of the ratio occurs for $|x| = 2$,

so that $$a_1 = \frac{16e^{-2}}{\sqrt{2\pi}} \approx 0.86385\ 54642\ ,$$

and one can use (18) about 86% of the time to simulate a standardized normal variate. The correction will be necessary with probability

$$1-a_1 = \frac{\sqrt{2\pi} - 16e^{-2}}{\sqrt{2\pi}} \approx 0.13514\ 45358\ ,$$

or about 14% of the time.

Note that $\int_{-\infty}^{\infty}[f(x)-a_1g_1(x)]dx = 1-a_1$

so that $$\frac{f(x)-a_1g_1(x)}{1-a_1}$$

can be called a "residual" density function. The graph of $f(x)-a_1g_1(x)$ is sketched in Figure 2. Note that the "residual" density function can be well approximated by $g_2(x)$, the (triangular) density function of $X = 1.5(u_1+u_2-1)$,

where $$g_2(x) = \begin{cases} (6-4\ x\ )/9, & 0 \leq |x| < 1.5 \\ 0 & , |x| \geq 1.5 \end{cases}$$

Note that (19) is easily generated. As with (18), one would like to use it as often as possible. It is therefore desirable to find a_2 so that

$$f(x) - a_1 g_1(x) - a_2 g_2(x) \geq 0$$

is minimized for $-1.5 < x < 1.5$. That is, one must find the minimal value of

$$\frac{f(x) - a_1 g_1(x)}{g_2(x)}$$

and assign it to a_2. To find x such that $-1.5 < x < 1.5$ and

$$\frac{d}{dx}\left(\frac{f(x) - a_1 g_1(x)}{g_2(x)}\right) = 0 ,$$

it is expedient to use a numerical method to find that $|x|$ = 0.87386 312884, 2.0. The first value provides the minimum, and so a_2 = 0.11081 79673; and (19) can be used about 11% of the time as a standardized normal variate.

Figure 3 is a sketch of $f(x) - a_1g_1(x) - a_2g_2(x)$. Note that

$$\int_{-\infty}^{\infty}[f(x) - a_1g_1(x) - a_2g_2(x)]dx = 1-a_1-a_2$$

$$\approx 0.02532\ 65685$$

so that one now has the "residual" density function

$$g_3(x) = \frac{f(x) - a_1g_1(x) - a_2g_2(x)}{1-a_1-a_2} .$$

As g_3 is multimodel, it would be tedious to search for one or more simple variates whose densities would satisfactorily approximate g_3 over the interval $-3 < x < 3$. Consequently, a rejection technique will be employed to simulate g_3 over this interval.

Over the finite interval $-3 < x < 3$, note that the maximum value of $g_3(x)$ is approximately 0.3181471173, and therefore if $x = 6u_1-3$ and $y = 0.3181471173\ u_2$, (x,y) is distributed uniformly over the rectangle (with area A = 1.9088827038) enclosing the relevant portion g_3. Moreover, of the points (x,y) comprising the rectangle, x can be taken as a variate with density function g_3 only if $y \leq g_3(x)$; because only those points (x,y) lying on and under the curve g_3 contribute to valid probability statements about X. Consequently one generates the points (x,y) until $y \leq g_3(x)$, and then with probability a_3, one lets $X = x$.[1]

The probability a_4 associated with generating x 3 is relatively small, $a_4 = 1-a_1-a_2-a_3 = 0.02532\ 65681 - a_3$. (20)
Consequently, it is reasonable to use (14) for the tails of the distribution. Since the variates secured by (14) are independent, standardized, normal variates

$$a_4 = P[|X| \geq 3] = 2 \int_3^\infty \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt$$

$$\simeq 0.00269\ 97960\ 63\ ,$$

so that from (20) one obtains

$$a_3 = 0.02262\ 677245\ .$$

Thus, for $|X| \geq 3$, one generates x_1 and x_2 according to (14) until at least one of them has absolute value greater than or equal to three, and then the appropriate one is taken as X. The probability that a pair

1) a_3 will be calculated later. Note that the probability that a pair (x,y) will provide an acceptable X is $(1.908882704)^{-1} = p_3 =$ 0.5238666566, or about two uniform variates will be needed for each acceptable X when this rejection technique is used.

and then the appropriate one is taken as X. The probability that a pair (u_1,u_2) will provide at least one normal variate $|X| \geq 3$ is

$$a_4(2-a_4) \approx 0.00539\ 2421227,$$

so that one can expect to produce about 185 pairs (u_1,u_2) to secure at least one standardized normal variate X when (14) is used.

According to (15), then, f(x) can be written as

$$f(x) = a_1g_1(x) + a_2g_2(x) + a_3g_3(x) + a_4g_4(x),$$

and the desired variate X is generated according to the following rule:

1. With probability $a_1 = 0.08638554642$, let

$$X = 2(u_1+u_2+u_3-1.5)$$

2. With probability $a_2 = 0.1108179673$, let

$$X = 1.5(u_1+u_2-1)$$

3. With probability $a_3 = 0.02262677245$, generate pairs (x,y) until $y \leq g_3(x)$, and then let $X = x$, where $x = 6u_1-3$ and $y = 0.3181471173u_2$ with

$$g_3(x) = \begin{cases} 15.75192787e^{-x^2/2}-4.263583239(3-x^2)-1.944694161(1.5-|x|) & \text{for } |x| < 1 \\ 15.75192787e^{-x^2/2}-2.1317916185(3-|x|)^2-1.944694161(1.5-|x|) & \text{for } 1 \leq |x| < 1.5 \\ 15.75192787e^{-x^2/2}-2.1317916185(3-|x|)^2, & \text{for } 1.5 \leq |x| < 3 \\ 0, & \text{for } |x| \geq 3 \end{cases}$$

4. With probability $a_4 = 0.002699796063$, generate pairs (x_1,x_2) until either $|x_1|$ or $|x_2|$ is greater than or equal to three, and then let that one be X, where

$$x_1 = (-2 \ln u_1)^{1/2} \text{Cos } 2\pi\, u_2$$

and $$x_2 = (-2 \ln u_1)^{1/2} \text{Sin } 2\pi\, u_2.$$

To obtain some idea of the efficacy of this technique it is helpful to calculate the expected number E[N] of uniform variates needed to produce a single normal variate. To do so, let n_i denote the number of uniform variates one expects to use in the i^{th} step of the process, i = 1,...,4. Then

$$E[N] = \sum_{i=1}^{4} a_i n_i$$

$$= 3a_1 + 2a_2 + Aa_3 + a_4/a_4(2-a_4)$$

$$\approx 4.079381839,$$

so that one can expect to generate about four uniform variates for each normal variate produced.[1]

The above technique has been shown in detail, applied here to the standardized normal distribution. In theory, it can be applied to any distribution. It can be seen, however, that one would find it handy to compile a catalogue of density functions of relatively easily generated variates, such as those used in steps 1 and 2 of the above technique.

---

1) One could compute the expected computing time for each normal variate by replacing n_i with t_i, the expected computing time for the i^{th} process.

IV

Other Distributions

1. The Chi-Square Distribution

If X is a random variable with the Chi-Square Distribution with n degrees of freedom, it has density function

$$f(x) = \frac{x^{(n/2)-1} e^{-x/2}}{\Gamma(n/2)\ 2^{n/2}} \qquad (21)$$

where $x > 0$ and $n = 1,2,\ldots$ [7, p.226]. This is a special form of the Gamma Distribution [7, p.126].

For this distribution, there are several special cases of n to observe before moving to an asymptotic distribution.

Recall that in obtaining (14) it was found that $X = -2 \ln u$ had a Chi-Square Distribution with $n = 2$. Consequently, if

$$X = -2 \ln(u_1 u_2 \ldots u_m) \qquad (22)$$

then X has a Chi-Square Distribution with $n = 2m$ [7, p.244]. Thus, (22) can be used to generate chi-square variates with even degrees of freedom.

The square of a standardized normal variate has a Chi-Square Distribution with $n = 1$ [7, p.243]. Thus, if Y is a standardized normal variate which is independent of $u_1, u_2, \ldots, u_m$, then using (22), $X + Y^2$ has a Chi-Square Distribution with $2m + 1 = n$ and can be used to generate chi-square variates with odd degrees of freedom.

If one has a supply of standardized normal variates Y, there is a convenient method for generating chi-square variates X having fairly large degrees of freedom n. It is the so-called Wilson-Hilferty transformation [6,11]:

$$Y = [(X/n)^{1/3} + 2/(9n) - 1] [9n/2]^{1/2}, \tag{23}$$

or $$X = n[Y\sqrt{2/(9n)} + 2/(9n) + 1]^3. \tag{24}$$

It has been found [6] that (23) converges to a normal variate for smaller n than the familiar transformation[1]

$$Y = \sqrt{2X} - \sqrt{2n-1}$$

or (25)

$$X = 1/2[Y = \sqrt{2n-1}]^2$$

which is often seen given for $n > 25$. In the case of (23), a numerical investigation [6] has shown that for $n \geq k$ the maximum absolute error $|E|$ between the standardized normal distribution function and that of (23) is:

k	$\|E\|$
1	0.03443
2	.01218
3	.00692
5	.00353
10	.00148
12	.00119
13	.00109
14	.00103
15	.00092

From consulting the table, one can decide at which value of n he will cease using (22) and use (23).

---

1) See: Fisher, R. A. Statistical Methods for Research Workers (Tenth Edition), Oliver & Boyd, London (1948), p.81.

2. The Beta Distribution

If X is a random variable having the Beta Distribution, with parameters p and q, then X has density function

$$f(x) = \frac{\Gamma(p + q + 2)}{\Gamma(p + 1)\Gamma(q + 1)} x^p (1 - x)^q$$

where $0 \leq x < 1$, $p > -1$, and $q > -1$ [7, p.129]. To generate X compute

$$s = [u_1 \Gamma(p + 2)/\Gamma(p + q + 3)]^{1/(p + 1)} \tag{26}$$

and $$t = [u_2 \Gamma(q + 2)]^{1/(q + 1)}$$

until $$s + t \leq 1,$$

and then $$X = s/(s + t) \tag{27}$$

has the desired Beta Distribution.

To verify (27), let $z = s + t$ so that

$$s = xz$$

and $$t = z(1 - x)$$

from which one can obtain

$$u_1 = \frac{\Gamma(p + q + 3)}{\Gamma(p + 2)} x^{p + 1} z^{p + 1}$$

and $$u_2 = \frac{z^{q + 1}(1 - x)^{q + 1}}{\Gamma(q + 2)} .$$

To find the desired joint distribution of X and Z, it is necessary to compute $J(u_1,u_2/x,z)$, where

$$J(u_1,u_2/x,z) = \begin{vmatrix} \partial u_1/\partial x & \partial u_1/\partial z \\ \partial u_2/\partial x & \partial u_2/\partial z \end{vmatrix}$$

$$= \frac{\Gamma(p + q + 3)}{\Gamma(p + 1)\Gamma(q + 1)} z^{p + q + 1} x^p(1 - x)^q$$

so that the joint density of X and Z is

$$h(x,z) = \phi(u_1,u_2)\cdot|J(u_1,u_2/x,z)|$$

$$= J(u_1,u_2/x,z).$$

Restricting $0 \leq Z \leq 1$ according to (26), the density function of X is

$$f(x) = \int_0^1 h(x,z)dz$$

$$= \frac{\Gamma(p + q + 3)}{\Gamma(p + 1)\Gamma(q + 1)} x^p(1 - x)^q \int_0^1 z^{p + q + 1}dz$$

$$= \frac{\Gamma(p + q + 2)}{\Gamma(p + 1)\Gamma(q + 1)} x^p(1 - x)^q, \; 0 \leq x < 1,$$

which implies X has the proper Beta distribution.

3. The F Distribution:

If X is a random variable having the F distribution, with m and n degrees of freedom, then X has the density function

$$f(x) = \frac{\Gamma((m + n)/2)}{\Gamma(m/2)\Gamma(n/2)} (m/n)^{m/2} \frac{x^{(m - 2)/2}}{(1 + mx/n)^{(m + n)/2}}$$

where $x \geq 0$, $m > 0$, and $n > 0$. (In practice, m and n are positive integers.)

There is a well-known transformation [7, p.244]:

$$y = \frac{mx/n}{1 + mx/n} \tag{28}$$

where y is a Beta variate with parameters $p = (m/2) - 1$ and $q = (n/2) - 1$, and x has the F distribution, with m and n degrees of freedom.

Since a method for generating Beta variates has been given, it is convenient to rewrite (28) as

$$x = ny/m(1 - y). \tag{29}$$

Referring to (27), (29) can be written as

$$x = (n/m) \frac{s/(s + t)}{1 - x/(s + t)}$$

$$= ns/(mt)$$

$$= n/m \left[\frac{u_1\Gamma(p + 2)}{\Gamma(p + q + 3)}\right]^{1/(p + 1)} [u_2\Gamma(q + 2)]^{-1/(q + 1)}$$

or $$x = n/m\left[\frac{((m + 2)/2)}{\Gamma((m + n + 2)/2)}\right]^{2/m} [\Gamma((n + 2)/2)]^{-2/n} \left(\frac{u_1^{2/m}}{u_2^{2/n}}\right) \tag{30}$$

Consequently, subject to $x + t \leq 1$ in (26), or equivalently

$$\left[\frac{u_1\Gamma((m + 2)/2)}{\Gamma((m + n + 2)/2)}\right]^{2/m} + [u_2\Gamma((n + 2)/2]^{2/n} \leq 1,$$

x has the F distribution with m and n degrees of freedom.

4. Student's t Distribution:

If X is a random variable having Student's t distribution, with n degrees of freedom, then X has density function

$$f(x) = \frac{\Gamma((n + 1)/2)}{\sqrt{2\pi}\ \Gamma(n/2)} (1 + x^2/n)^{-(n + 1)/2}$$

where $-\infty < x < \infty$, and $n > 0$. (In practice, n is usually a positive integer.) [7, p.233].

It is easy to show that if X has Student's t distribution, with n degrees of freedom, then X^2 has the F distribution with one and n degrees of freedom [7, p.233]. Therefore, if Y has this F distribution, then the desired t variate X can be generated by

$$X = \pm\sqrt{Y} \qquad (31).$$

The generation of the F distribution was discussed in the previous section, and so, referring to (30), Y can be generated by

$$Y = \frac{n\pi(u_1 u_2^{-1/n})^2}{[(n + 1)\Gamma((n + 1)/2)]^2\ [(n/2)\Gamma(n/2)]^{2/n}} \qquad (32)$$

subject to

$$\frac{\pi u_1^2}{(n + 1)^2[\Gamma((n + 1)/2)]^2} + [\frac{n u_2 \Gamma(n/2)}{2}]^{2/n} \leq 1 \qquad (33).$$

Consequently, the desired variate x is given by (31), or equivalently,

$$x = \pm \frac{\sqrt{n\pi}\ u_1 u_2^{-1/n}}{(n + 1)\Gamma((n + 1)/2)[(n/2)\Gamma(n/2)]^{1/n}} \qquad (34),$$

subject to (33).

To assign the proper sign to (34), it is necessary that it be positive with probability 0.5. Referring to (5), it is seen that

$$P[u_3 \geq 0.5] = 0.5$$

so that $$P[u_3 - 0.5 \geq 0] = 0.5$$

or $$P[\frac{u_3 - 0.5}{|u_3 - 0.5|} = +1] = 0.5$$

$$= P[\frac{u_3 - 0.5}{|u_3 - 0.5|} = -1]$$

so that multiplying (34) by $(u_3 - 0.5)/|u_3 - 0.5|$ will affix the proper sign.

Recall that the three distributions in this section are derived from the Beta distribution, and each of the generation methods essentially requires that $s + t \leq 1$ in (26). Consequently one would desire the probability that a point (u_1, u_2) would provide a valid variate. For the case of the Beta variate (27) this probability is, from (26),

$$P[s + t \leq 1] = P[u_1 \leq w]$$

$$= \int_0^1 \int_0^w 1 \, du_1 \, du_2 = \int_0^1 w \, du_2,$$

where $$w = \frac{\Gamma(p + q + 3)}{\Gamma(p + 2)} (1 - [u_2 \Gamma(q + 2)]^{1/(q + 1)}):$$

Consequently, after some rearranging, this becomes

$$P[x + t \leq 1] = \frac{\Gamma(p + q + 3)}{\Gamma(q + 1)\Gamma(p + 2)} \int_0^{[\Gamma(q + 2)]^{1/(q + 1)}} Z^q (1-Z)^{p+t} dZ \quad (35)$$

and it is to be noted that, for various values of p and q, (35) has been tabulated, because the integral is essentially the density function for the Beta distribution, with parameters q and $p + 1$.[1]

To generate the F distribution, with m and n degrees of freedom, referring to (28) and (35), $p = (m/2)-1$ and $q = (n/2)-1$ so that (35) becomes

$$P[s + t \leq 1] = \frac{\Gamma((m + n + 2)/2)}{\Gamma(n/2)\Gamma((m + 2)/2)} \int_0 [\Gamma((n+2)/2)]^{2n} Z^{(n-2)/2}(1-Z)^{m/2} dZ$$

$$= \frac{(m + n)\Gamma((m + n)/2)}{m\Gamma(m/2)\Gamma(n/2)} \int_0 [(n/2)\Gamma(n/2)]^{2/n} Z^{(n-2)/2}(1-Z)^{m/2} dZ, \quad (36)$$

which is associated with the Beta distribution, with parameters $(n/2)-1$ and $m/2$.

The Student's t distribution, with n degrees of freedom, is obtained from the F distribution, with one and n degrees of freedom. Consequently, $m = 1$ in (36), and

$$P[s + t \leq 1] = \frac{(n+1)\Gamma((n+1)/2}{\sqrt{\pi}\ \Gamma(n/2)} \int_0 [(n/2)\Gamma(n/2)]^{2/n} Z^{(n-2)/2}(1-Z)^{1/2} dZ \quad (37)$$

for the Student's t distribution. Here (37) is associated with the Beta distribution, with parameters $(n/2)-1$ and $1/2$, so that a published table could be consulted for this case as well as the preceding two, (35) and (36).

1) This distribution has been extensively tabulated as "The Incomplete Beta" distribution, so (35) can be evaluated with the aid of a table. See: Biometrika Tables For Statisticians, Vol. I (E. S. Pearson and H. O. Hartley, editors), Cambridge University Press, London, (1954), pp. 142-156.

In certain cases, at least one of the parameters in the Incomplete Beta integral is an integer, and this can simplify an integral of the form

$$P = \frac{\Gamma(p + q + 2)}{\Gamma(p + 1)\Gamma(q + 1)} \int_0^x t^p \ (1 - t)^q \ dt. \qquad (38)$$

When q is a positive integer, it is helpful to note that

$$(1 - t)^q = \sum_{K=0}^{q} \binom{q}{K} (-t)^K$$

so that (38) can be written as the finite series

$$P = \frac{\Gamma(p + q + 2)}{\Gamma(p + 1)q!} x^{p + 1} \sum_{K=0}^{q} \binom{q}{K} \frac{(-x)^K}{p + 1 + K}$$

If it occurs that p is a positive integer, repeated integration by parts will yield the following finite series:

$$P = 1 - \sum_{K=0}^{p} \binom{p + q + 1}{K} x^K (1 - x)^{p + q + 1 - K}$$

If, however, both p and q are positive integers, the repeated integration by parts can be used to find that

$$P = \sum_{K=p+1}^{p+q+1} \binom{p + q + 1}{K} x^K (1 - x)^{p + q + 1 + K}$$

For nonintegral values of the parameters, one will have to devise an appropriate technique for (38), or perhaps a published table will be helpful.

5. The Gamma Distribution:

If the random variable X has a Gamma Distribution with parameters λ and k, then X has density function

$$f(x) = \frac{\lambda}{\Gamma(k)} (\lambda x)^{k-1} e^{-\lambda x} \qquad (39)$$

where $x \geq 0$, $k > 0$, and $\lambda > 0$. Note that for $k = n/2$ and $\lambda = 1/2$, this becomes the Chi-Square Distribution (21) with n degrees of freedom. The distribution (39) can be characterized by its Moment Generating Function [7, p.129)], where $t < \lambda$,

$$M_X(t) = E[e^{Xt}] = (1 - \frac{t}{\lambda})^{-k} . \qquad (40)$$

Recall that if Y is a random variable having the Exponential Distribution with mean $E[Y] = 1/\lambda$, then Y has density function (3), or

$$h(y) = \lambda e^{-\lambda y}$$

where $y \geq 0$ and $\lambda > 0$. This distribution is characterized by the following Moment Generating Function [7, p.119]:

$$M_Y(t) = (1 - \frac{t}{\lambda})^{-1} . \qquad (41)$$

If $Y_1, Y_2, \ldots, Y_k$ are k independent random variables, each having density function $h(y)$ above, then the Moment Generating Function for

$X = \sum_{i=1}^{k} Y_i$ is given by [7, p.121]:

$$M_X(t) = \prod_{i=1}^{k} M_{Y_i}(t)$$

$$= \prod_{i=1}^{k} (1 - \frac{t}{\lambda})^{-1}$$

$$= (1 - \frac{t}{\lambda})^{-k} ,$$

which is exactly (40). Consequently, this X has density function (39).

Recall that if u_i has the Uniform Distribution over [0,1) , then

$$Y_i = -\frac{1}{\lambda} \ln u_i \qquad (9)$$

has the Exponential Distribution with density function $h(y)$ above. Thus, where $u_1, u_2, \ldots, u_K$ are k independent random variables, each having the Uniform Distribution over [0,1], it follows that

$$X = \sum_{i=1}^{k} Y_i = -\frac{1}{\lambda} \ln \prod_{i=1}^{k} u_i \qquad (42)$$

is a gamma variate with parameters λ and k.

6. The Poisson Distribution:

If the random variable K has the Poisson Distribution, then K has density function

$$f(k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad (43)$$

where $\lambda > 0$ and k is a non-negative integer. Here, then, $f(k) = P[K = k]$.

It is to be recalled that for some positive integer k

$$Y_k = \sum_{i=1}^{k} x_i ,$$

where $X_1,\ldots,X_n$ are independent random variables identically distributed according to (3) with parameter λ, has the Gamma Distribution (39). It is interesting to note the distribution of K, where K is defined according to

$$Y_K = \sum_{i=1}^{K} X_i \leq 1 < \sum_{i=1}^{K+1} X_i . \qquad (44)$$

If H is the distribution function of K, then

$$H(k) = P[K \leq k] , \qquad (45)$$

and the density function of K is found by

$$h(k) = H(k) - H(k-1) , \qquad (46)$$

where (45) is found from (39) according to

$$P[K \geq k] = \int_0^1 \frac{\lambda(\lambda y)^{k-1}}{(k-1)!} e^{-\lambda y} dy$$

$$= 1 - H(k-1) .$$

Thus $$H(k) = \int_1^\infty \frac{\lambda(\lambda y)^k}{k!} e^{-\lambda y} dy ,$$ and

$$h(k) = \int_1^\infty (\frac{\lambda^{k+1} y^K e^{-\lambda y}}{k!} - \frac{\lambda^K y^{K-1} e^{-\lambda y}}{(k-1)!} dy$$

$$= \frac{\lambda^K}{k!} \int_1^\infty (\lambda e^{-\lambda y}\, y^K - ky^{K-1}\, e^{-\lambda y})\, dy$$

$$= -\frac{\lambda^K}{k!} \int_1^\infty \frac{d}{dy}(e^{-\lambda y})\, dy$$

$$h(k) = -\frac{\lambda^K}{k!} y^K\, e^{-\lambda y} \Big|_{y=1}^{\infty} = -\frac{\lambda^K}{k!} (0 - 1^K\, e^{-\lambda \cdot 1})$$

$$= e^{-\lambda} \frac{\lambda^k}{k!}$$

so that K has the Poisson Distribution. Thus, to generate pseudo-random numbers from (43), one can compute Y_k in (44), where it is to be recalled that

$$X_i = -\frac{1}{\lambda} \ln u_i$$

has the distribution (3), where u_i is an easily generated uniform variate. Thus, (44) is

$$Y_k = -\frac{1}{\lambda} \ln \prod_{i=1}^{k} u_i \leq 1 < Y_{k+1}$$

and then k is distributed according to (43). This can be somewhat simplified to yield k of the form

$$\prod_{i=1}^{k+1} u_i < e^{-\lambda} \leq \prod_{i=1}^{k} u_i \ . \qquad (47)$$

7. The Binomial Distribution:

If the random variable X has the Binomial Distribution, then X has density function [7, p.64]

$$f(x) = \binom{n}{x} p^x (1 - p)^{n-x} ,$$

where $x = 0,1,\ldots,n$, and $0 \leq p \leq 1$. The distribution function of X is

$$F(x) = \sum_{t=0}^{x} f(t) = \sum_{t=0}^{x} \binom{n}{t} p^t (1 - p)^{n-t}$$

$$= 1 - \sum_{t=x+1}^{n} \binom{n}{t} p^t (1 - p)^{n-t} ,$$

where it is to be noted that $F(x) = P[X \leq x]$, and $f(x) = P[X = x]$.

Since X can take on only a finite number of values, it is feasible to evaluate the distribution function F for each admissable value, and for each value of X, $0 \leq F(x) \leq 1$. This suggests a simple technique for the generation of pseudo random numbers from the Binomial Distribution.[1]

Let u denote a random variable having the Uniform Distribution over $[0,1)$. Then for every u, there exists some value of X such that

$$F(x - 1) \leq u < F(x)$$

so that letting $X = x$ will provide the desired pseudo random number.

1) It will be noted that the technique can be applied to any random variable which can take on only a finite number of values.

Bibliography

1. Box, G.E.P. and Muller, M.E., "A Note on the Generation of Random Normal Deviates." Ann. Math. Stat. 29 (1958), pp.610-611.

2. Gannon, L.J. and Schmittroth, L.A., "Computer Generation and Testing of Random Numbers." A.E.C. Research and Development Report, IDO-16921, Mathematics and Computers, TID-4500 (22 Ed.) August, 1963.

3. International Business Machines Corporation, "Random Number Generation and Testing." Manual C20-8011, New York, 1959.

4. Marsaglia, G., "Expressing a Random Variable in Terms of Uniform Random Variables." Ann. Math. Stat. 32 (1961), pp.894-898.

5. Marsaglia, G., and Bray. T.A., "A Convenient Method for Generating Normal Variables." SIAM Review. 6 (1964), pp.260-264.

6. Mathur, R.K., "A Note on the Wilson-Hilferty Transformation of χ^2." Bull. Calcutta Stat. Assoc. 10 (1961), pp.103-105.

7. Mood, A.M. and Graybill, F.A., Introduction to the Theory of Statistics (2nd Ed.). McGraw-Hill Book Co., New York, 1963.

8. Muller, M.E., "A Comparison of Methods for Generating Normal Deviates on Digital Computers." Journ. Assoc. Comp. Mach. 6 (1959), pp.376-383.

9. Rand Corporation, A Million Random Digits with 100,000 Normal Deviates. The Free Press Publishers, Glencoe, Illinois, (1955).

10. Smith, R.J., "On the Generation and Testing of Uniformly Distributed Pseudo-Random Numbers." Report TCM-5, February 1, 1966. Texas Center for Research, Austin, Texas.

11. Wilson, E. B. and Hilferty, M.M., "The Distribution of Chi-Square." Proc. Nat. Acad. Sci., U.S.A., 17 (1931) pp. 684-688.