Ann. Inst. Statist. Math.
28 (1976), Part A, 399-401

MOMENTS OF THE TIME TO GENERATE RANDOM VARIABLES BY REJECTION

J. ARTHUR GREENWOOD

(Received Sept. 27, 1973; revised Mar. 5, 1974)

The rejection technique, as presented by Butcher [1], generates a random variable η with the density function

$$f(y)=ad(y)g(y)\ , \tag{1}$$

where $a>0$, $0\leqq g(y)\leqq 1$, and the algorithm T yields a random variable with density function $d(y)$, by the following steps: Operate the algorithm T, yielding the number Y, and compute $g(Y)$. Operate a uniform random number generator, yielding the number X, $0\leqq X\leqq 1$. If $X\leqq g(Y)$, accept Y; otherwise reject Y and start again. If the X and Y obtained in successive attempts are independent, then the probability that the nth attempt is successful,

$$\Pr[N=n]=p(1-p)^{n-1}\ , \qquad n>0\ , \tag{2}$$

where

$$p=\int d(y)g(y)dy=1/a\ .$$

The geometric distribution (2) has factorial moment generating function

$$\begin{aligned}\sum_{n=1}^{\infty} p(1-p)^{n-1}(1+u)^n &= p(1+u)\sum_{m=0}^{\infty}[(1-p)(1+u)]^m \\ &=(1+u)/(1+u-au) \\ &=1+a\sum_{j=1}^{\infty}(a-1)^{j-1}u^j\ ,\end{aligned}$$

so that

$$\mathrm{E}\binom{N}{j}=a(a-1)^{j-1}\ ;$$

in particular

$$\mathrm{E}(N)=a\ , \qquad \mathrm{Var}(N)=2\mathrm{E}\binom{N}{2}+\mathrm{E}(N)-[\mathrm{E}(N)]^2=a(a-1)\ .$$

If the algorithm T requires time t_1, computation of g requires t_2, and generating X requires t_3, then the mean and variance of time to generate η are $a(t_1+t_2+t_3)$ and $a(a-1)(t_1+t_2+t_3)^2$.

Sibuya [2] proposes to accelerate the process by generating X first, and using the same X in the inequality $X \leqq g(Y)$ until some Y is accepted. The moments of time to generate under this proposal seem not to have been published. Conditional on X, the distribution of N is again geometric:

$$\text{(3)} \qquad \Pr[N=n \mid X=x]=\pi_x(1-\pi_x)^{n-1}, \qquad n>0,$$

where

$$\pi_x=\int H[d(y)-x]g(y)dy=1/\alpha_x ,$$

and

$$H(z)=1, \quad z \geqq 0; \qquad H(z)=0, \quad z<0 .$$

Integration over the uniform distribution of X yields the unconditional distribution of N,

$$\Pr[N=n]=\int_0^1 \pi_x(1-\pi_x)^{n-1}dx ,$$

$$\text{(4)} \qquad \mathrm{E}\binom{N}{j}=\int \alpha_x(\alpha_x-1)^{j-1}dx ,$$

$$\mathrm{E}(N)=\int_0^1 \alpha_x dx , \quad \mathrm{Var}(N)=\int_0^1 \alpha_x(2\alpha_x-1)dx-[\mathrm{E}(N)]^2 .$$

The behaviour of (4) depends on the asymptotic form of α_x for x near 1: if, as with many useful rejection schemes,

$$d(y_0)=1 , \qquad d(y_0+\Delta y)=1-c(\Delta y)^2+O(\Delta y)^3 ,$$

then, for x close to 1, π_x is approximately proportional to $2c^{-1/2}(1-x)^{1/2}$, and $\mathrm{Var}(N)$ fails to exist. The two examples below were chosen for easy explicit evaluation of π_x and do not necessarily represent desirable rejection techniques.

First example. To generate a random variable η with the density function

$$\text{(5)} \qquad f(y)=6y(1-y) , \qquad 0<y<1 .$$

Take

$$g(y)=1 , \quad d(y)=4y(1-y) , \quad a=3/2 .$$

Then

$$(6) \qquad p=2/3\,, \quad \mathrm{E}(N)=3/2\,, \quad \mathrm{Var}(N)=3/4\,.$$

For the accelerated technique, conditional on X,

$$\pi_x=(1-x)^{1/2}\,, \quad \alpha_x=(1-x)^{-1/2}\,,$$

and so

$$\Pr[N=n]=4/[n(n+1)(n+2)]\,, \quad \mathrm{E}(N)=2\,, \quad \mathrm{Var}(N)=\infty\,.$$

The expected time to generate is $1.5(t_1+t_2+t_3)$ under the standard technique, and $2(t_1+t_2)+t_3$ under the accelerated technique. Here $t_1=t_3$, so that the expected times are respectively $3t_1+1.5t_2$ and $3t_1+2t_2$; the accelerated technique is not advantageous.

Second example. To generate a random variable η with the density function (5). Take

$$g(y)=2(1-2|y-1/2|)\,, \quad d(y)=(1+2|y-1/2|)/2\,, \quad a=3/2\,;$$

so that the distribution of N under the standard technique is given by (6). For the accelerated technique, conditional on X,

$$\pi_x=1\,, \quad x\leqq 1/2\,, \qquad \pi_x=2-2x\,, \quad x\geqq 1/2\,,$$

$$\alpha_x=1\,, \quad x\leqq 1/2\,, \qquad \alpha_x=(1-x)^{-1}/2\,, \quad x\geqq 1/2\,;$$

$$\Pr[N=1]=3/4\,, \quad \Pr[N=n]=1/[2n(n+1)]\,, \qquad n>1,\ \mathrm{E}(N)=\infty\,.$$

In practical computing, the effect of infinite variance (a fortiori of infinite mean) is that runs of satisfactory generation are interrupted by runs of very long time to generate; thus, it can happen that an established and apparently error-free computer routine will *intermittently* violate the time limit assigned to a job (whenever a value of X close to 1 happens to be generated).

Institute of Marine and Atmospheric Sciences, City University of New York

References

[1] Butcher, J. C. (1961). Random sampling from the normal distribution, *Computer J.*, **3**, 251-253.

[2] Sibuya, M. (1962). Further consideration on normal random variable generator, *Ann. Inst. Statist. Math.*, **14**, 159-165.