Some Fast Methods of Generating Random Variables with Preassigned Distributions*
I. General Acceptance-Rejection Procedures

By

Herman Rubin
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #466

August 1976

*Research was supported in part by the Office of Naval Research Contract N00014-A-2260008, Project Number NR 042-216 and National Science Foundation Grant Nos. MCS 76-08316 and MPS74-A07836 A01 at Purdue University. Reproduction in whole or part is permitted for any purpose of the United States Government.

Some Fast Methods of Generating Random Variables with Preassigned Distributions*
I. General Acceptance-Rejection Procedures

By

Herman Rubin
Purdue University

1. Introduction. We give some methods which, despite their simplicity, have not found their way into the literature. They are based on two elementary observations concerning acceptance - rejection procedures. Of course, the procedure to be used in a specific situation depends on how many times the particular distribution, including the specific parameters, occurs. It also depends on the characteristics of the computer -- relative costs of computing and access, availability of instructions, number of registers, etc. The procedures given here should be at most 1-2 multiplications slower than Marsaglia-MacLaren-Bray procedures ([1], [2]) with comparable storage requirements. The procedures in section 4 are considerably more costly - In general they should take approximately one transcendental function time plus the cost of the uniform and exponential random variables used.

2. Preliminaries. The basic idea of an acceptance - rejection procedure is as follows. To obtain a random variable whose distribution has density of (with respect to the measure μ), one obtains a random variable Y whose density is bg where $f \leq g$, and then set X=Y with probability f(Y)/g(Y). This is usually done by comparing f(Y)/g(Y) with a uniform random variable or log (g(Y)/f(Y)) with an exponential random variable. There are also variations on this, some of which will be discussed here.

Three examples (not recommended procedures) will serve to illustrate this and one of the principles referred to in the introduction.

A. To obtain a Cauchy r.v. (density $\frac{1}{\pi}\ \frac{1}{1+x^2}$), let Y be uniform

*Research was supported in part by the Office of Naval Research Contract N00014-A-2260008, Project Number NR 042-216 and National Science Foundation Grant Nos. MCS 76-08316 and MPS74-A07836 A01 at Purdue University. Reproduction in whole or part is permitted for any purpose of the United States Government.

(-1,1) and let K be independent with $P(K = 0) = P(K = 1) = \frac{1}{2}$. Let U be uniform (U,1). Then if $U < (1 + Y^2)^{-1}$ set X=Y if K=0, and X=1/Y if K=1, else start over.

Alternative: Let $U^1 = (1+Y^2)U$. If $U^1 < 1$ form X as before and store U^1 as a uniform random variable, or else start over. Note that U^1 is independent of X given that $U^1 < 1$. Of course, this procedure runs a risk of accumlating roundoff, but this is not likely to be serious.

B. To obtain a normal r.v. ($f = e^{-\frac{1}{2}x^2}$), let Y be double exponential ($g = e^{-|x|}$), E exponential. Set X=Y if $E > \frac{1}{2}(|X| -1)^2$, else start over.

Alternative: Let $E^1 = E - \frac{1}{2}(|X| - 1)^2$. If $E^1 > 0$ form X as before and store E^1 as an exponential random variable. The same remarks apply, and roundoff is even less serious, especially if unnormalized arithmetic is used.

C. To enlarge an existing stack of independent exponential random variables, we can use the following procedure to generate exponential random variables, which is more efficient than the corresponding crude version of von Neumann [3]. Let N=0. Let U be uniform (0,1), E exponential. If $E^1 = E - U > 0$, E^1 and N + U are the exponential variables, else increase N by 1 and get new U and E and continue until acceptance. This procedure generates a net of one exponential variable per 3.7844 uniform random variables used. We recommend a more efficient version, similar to the general procedure for fixed distributions in Section 3.

The second observation is that it is in general not necessary to compute the tests precisely. Suppose that one wishes to obtain r.v.'s with density $a(1-x^2)^b$. Let Z be normal (0,1), $Y = Z/\sqrt{2b}$. If $|Z| \geq 1$, start over, else let E be exponential and compute $E^1 = E - bY^4/2(1-Y^2)$. If $E^1 > 0$, X=Y and store E^1 as an exponential. In all other cases the exponential is lost.

If not, compute $E - bY^4/2(1-\frac{2}{3}Y^2)$. If this is negative, start over. Otherwise (even this can be improved, but it is already rather rare), compute $E + bY^2 + b\log(1-Y^2)$, set X=Y if this is positive, and start over is this is negative.

3. Some procedures. The following procedures are written in a FORTRAN-like language. However, any implementation of these programs, especially the general ones, in anything other than assembler language, is nothing short of criminal.

Let f be a univariate density and let the domain of f be divided into N intervals. Let $g \geq f$ be a function such that for some α either $\int_{J_i} f(X)dx = \alpha$ or $\int_{J_i} g(X)dx = \alpha$, where the choice may depend in i. Furthermore, call J_i normal if g is constant on J_i. If J_i is normal, let $D(i) = \sup_{x=J_i} \log g(X)/f(X)$.

Let A(i) and B(i) be the endpoints of J_i if J_i is normal, and let C(i) = A(i) - B(i). Let TEXP always contain an exponential random variable, and let UNIF and EXPRV be uniform (0,1) and exponential random variables. Then the following "program" will produce random variables with density f.

```
      SUBROUTINE RANVAR (A,C,D,N,K)
C   K+1 is the number of abnormal cases, N the total number of cases, A, C,
    and D as above
 4    M + V = N * UNIF
C   M is the integer part.
      IF(L = M-K. LE.0)  GO TO "ABNORMAL"
C   ABNORMAL is the user--provided program to deal with the abnormal cases.
```

```
1     RANVAR = A(M) + V * C(M)
      IF (TEXP - D(M). LE.0)  GO TO 3
      TEXP = TEXP - D(M)
      RETURN

3     Z = TEXP
      TEXP = EXPRV
      IF (Z.GT.ALOG (g(RANVAR)/f(RANVAR)) RETURN
C  The test can be made by any convenient user-supplied method.
```

IF $(\alpha = \int_{J_M} f(X)dx)$ GO TO 2

```
      GO TO 4
2     V = UNIF
      GO TO 1
      END
```

Comments: (1) If N is a power of 2, N * UNIF can be done by such operations as shifts and string manipulation. On certain machines which do not have good communication between index and floating register, fixed point arithmetic should be used.

(2) If it is not desired to make a preliminary test or if some other test procedure is to be used, the two lines following 1 should be appropriately changed and the two lines starting with 3 should be replaced. Even for the normal $(f(x) = \exp(-\frac{1}{2}x^2))$ it may pay to make a preliminary test in the normal cases.

(3) For symmetric random variables, the sign bit can be extracted and put on at the end.

(4) It is usually more economical to have $\alpha = \int_{J_i} f(X)dx$ for all i, as this does not require the reloading of A(M), C(M), and D(M). However, setting $\alpha = \int_{J_i} g(X)dx$ for all i minimizes the expected number of trials, and in some cases it may be better to use $\alpha = \int_{J_i} g(x)dx$ for some i.

(5) Since the program is so short, it will pay to have special versions for common distributions such as the normal. Also, it may be slightly cheaper to produce many r.v.'s with a given distribution at the same time, as this might produce significant savings in the cost of the uniform and exponential r.v.'s needed.

(6) A procedure the author has found useful is to form an array whose first word contains the data N(or L,N $=2^L$) and K and which is followed by A, C, and D. For certain machines, it may be more convenient to have (A(i), C(i), D(i), T(i)), where T(i) contains data for the exact test in consecutive words.

(7) This procedure can be modified for discrete distributions.

```
      SUBROUTINE EXPSTACK ( I, J, E)
C  I is the initial size, J is the terminal size, E is the stack array
      K = I
      Q = TEXP
 1    W =  0
 11   M + V = N * UNIF
      IF (M.EQ.0)  GO TO 10
 3    D = V * C(M)
      Q = Q - D
      IF (Q.LT.0)  GO TO 2
      K = K + 1
      E(K) = W + A(M) + D
      IF (K.NE.J)  GO TO 1
      TEXP = Q
      RETURN
 10   W = W + WO
      GO TO 11
```

```
2     IF (K.EQ.0)  GO TO "BACKUP"
C  "BACKUP" is the procedure to use when the stack gets empty.  This will
be rare.
      Q = E(K)
      K = K - 1
      GO TO 11 (if C(M) exp (-A(M)) is constant)
      V = UNIF
      GO TO 3
      END
```

Comments: (1) The stack procedure is especially well suited to the exponential. The facts that $\log e^{-x} = -x$ and $e^{-(x+y)} = e^{-x} e^{-y}$ make the rejection case (Q.LT.0) and the abnormal case (M.EQ.0) particularly simple. Because of this, the exponential is likely to be 1-2 multiplication or division times faster than anything else, including even normal. Also, exponential random variables are quite useful for input for other distributions (the abnormal case for the normal, e.g.).

4. A useful procedure - CONCAVE

Because of the large cost of computing the A, C, and D tables, these procedures are not useful for a few r.v.'s with a given distribution. For a large class of distributions, those with concave logarithm of density, the following type of procedure is useful. Assume $f(X) = ce^{\varphi(x)}$, $\varphi(0) = 0$, $\varphi(X) \leq 0$, and let $g(X) = ce^{\psi(x)}$, where

$$\psi(x) = \begin{cases} e^{\alpha(x-\beta)} & x < \beta, \\ 0 & \beta \leq x \leq \delta \\ e^{-\gamma(x-\delta)} & x > \delta. \end{cases}$$

Also let $\rho(x) \leq \varphi(x) \leq \sigma(x)$.

Let the array A start with $((\delta-\beta) + 1/\alpha + 1/\gamma, - 1/\alpha, \beta, 1/\gamma, \delta)$. The user may wish to have other items in the array for the computation of φ, ρ, σ, deciding whether the candidate is out of range, etc.

```
          FUNCTION CONCAVE (A, RHO, SIGMA, PHI)
1    X = A(1) * UNIF + A(2)
     IF(X.LT.0) GO TO 3
     Y = X - A(4)
     IF(Y.LT.0) GO TO 4
     U = TEXP
     CONCAVE = X(3) + Y
2    IF(TEXP = U - RHO(CONCAVE,A).GE.0) RETURN
     TEXP = EXPRV
     IF (U - SIGMA(CONCAVE,A).LE.0) GO TO 1
     IF (U - PHI(CONCAVE,A).LE.0) GO TO 1
     RETURN
3    SF = A(2)
     EP = A(3)
     GO TO 5
4    SF = A(4)
     EP = A(5)
C  THE FIRST FOUR STATEMENTS PLUS THE PRECEDING FIVE CODE VERY EFFICIENTLY
C  IN ASSEMBLER LANGUAGE
C
C  WE NOW GET THE EXPONENTIAL TAILS.  THE FOLLOWING CODE IS LIKELY TO BE
C  NEARLY OPTIMAL
```

```
5    U = TEXP
     Y = X/SF
     IF(TEXP = U - Y.LT.0) GO TO 6
     CONCAVE = EP + X
     GO TO 2
6    E = EXPRV + 1.
     U = EXPRV + E
     CONCAVE = EP + SF * E
     GO TO 2
     REENTRY OBDS
     GO TO 1
C    THIS DIFFERS FROM THE MAIN ENTRY IN THAT THE ARGUMENTS FROM THE
C    PREVIOUS CALL ARE TO BE USED.
     END
```

Notes:

(1) The call to SIGMA may have an out of bounds return.

(2) The code for exponential tails can usually be slightly improved by using Y as an input uniform to a version of the exponential procedure. The amount of improvement is usually small.

(3) In some cases, the half-normal can be used instead of the exponential. Unless the setup cost is much lower, this is unlikely to pay unless the analog of the modification suggested in note (2) is used.

(4) This procedure can be modified for discrete distributions. However, the setup cost can be quite high and a procedure with

$$g(i) = \begin{cases} 1 & L \le i \le U \\ 2^{-k} & U + (k-1)V < i \le U + kV, \\ & L - (k-1)V > i \ge L - kV, \end{cases}$$

may be enough cheaper to set up to be preferred. For binomial with large variance, a procedure of this general type is certainly good.

5. If an exponential tail is not present, the coefficient (α or γ) is set infinite. The procedure uses only $1/\alpha$ and $1/\gamma$ and works correctly if either or both is 0.

6. The expected number of trials for the "best" procedure is $c(\delta - \beta + \frac{1}{\alpha} + \frac{1}{\gamma}) \leq \frac{1}{1-e^{-1}} = 1.582$; for approximately normal distributions it is near $(4/\pi)^{\frac{1}{2}} = 1.12838$.

The expected number of trials is minimized by solving the equations.

(a) $\varphi(X) = -1$,

(b) $\varphi(Y) = -1$,

(c) $X < 0 < Y$,

(d) $\varphi'(X) = \alpha$,

(e) $\varphi'(Y) = -\gamma$,

(f) $\beta = X + 1/\gamma$,

(g) $\delta = Y - 1/\gamma$.

However, if "=" is replaced by "$\leq$" in (a) and (b), the resulting procedure is correct. If Y-X is increased by 1%, the additional cost will be much less than one multiplication time.

5. Examples. For the normal and exponential, the more complicated methods in the previous section are quite good. We do not believe that they are optimal; for the exponential in particular some intricate bit handling methods, some of which do not even have roundoff error, may well be better.

For the gamma with shape parameter > 1, the logarithm $C + a \log t - t$ of the density is concave and the general concave procedure can be used. If we set $t = a(1 + u)$, an upper approximation for $\delta + \frac{1}{\gamma}$ is

$$\hat{y} = \frac{1.0053814}{a} + (\sqrt{.5a} + \tfrac{1}{6})^{-1}$$

and a lower approximation for $\beta - 1/\alpha$ is

$$\hat{X} = \max(-1, -(\sqrt{.5a} + \tfrac{1}{6})^{-1} + \frac{1/3}{a + .1133594}).$$

If $X = -1$, the lower exponential part is absent.

Also, $\log(1 + t) - t \leq -t^2/2(1 + \frac{2}{3}t)$,

and

$$\log(1 + t) - t \geq \begin{cases} -t^2/(2 + t) & t \geq 0, \\ -t^2/2(1 + t) & t \leq 0. \end{cases}$$

Alternatively, if we set $t = (a + \frac{2}{3})(1 + v)^3$, the density becomes

$$\log(v) = k + 3\beta \log(1 + v) + \beta - \beta(1 + v)^3,$$

$\beta = a + \frac{2}{3}$. This is valid for $a > -\frac{2}{3}$. Now

$$3 \log(1 + v) + (1 + v)^3 = -\tfrac{9}{2}v^2 - \tfrac{3}{4}v^4 + \ldots \leq -\tfrac{9}{2}v^2$$

Also if we set $\psi(v) = 3 \log(1+v) + 1 - (1 + v)^3 + \frac{9}{2}v^2$,

we have

$$\psi(v) \leq -\frac{.75\ v^4}{1 + .8v},$$

$$\psi(v) \geq \begin{cases} \dfrac{-.75\ v^4}{1 + .75v} > -.75v^4, & v \geq 0 \\ \dfrac{-.75v^4}{1 + v}, & v \leq 0. \end{cases}$$

This provides a procedure starting with normal random variables which has a low rejection probability, particularly if a is large. However, we question whether the three extra multiplications involved and the higher cost of the normal make it particularly good.

Notice that for $v \geq 0$ we give two lower bounds for ψ. The rejection probability is small enough that for $a \geq -.5$ it is likely to pay to use $-.75\, v^4$ first. For $a < -.5$ the procedure is not very good although it works to $a = -\frac{2}{3}$.

For the χ distribution, density $C\, x^{a+1} e^{-\frac{1}{2}x^2}$, the same idea can be applied with $x = (2a+1)(1+v)$. Since $\log \varphi(u) = b(-u + \log(1+u) - \frac{1}{2}u^2)$, the approximating functions for the gamma are easily modified. We also have simple formula

$$\hat{y} = \left(1 + \frac{.1770955}{\sqrt{b} + .4247652}\right)/b,$$

$$\hat{x} = \begin{cases} -1, & b \leq .2375, \\ 3b - 1.07125, & .2375 \leq b \leq 1, \\ \left(1.001497 - \dfrac{.175833}{\sqrt{b} - .23736}\right)/\,b, & b \geq 1, \end{cases}$$

where $b = 2a + 1$.

Another example is the beta distribution. If we take the density in the form

$$f(t) = c\left(1 + \frac{t}{B}\right)^B \left(1 - \frac{t}{A}\right)^A, \qquad -B < t < A$$

we find that

$$\log f(t) \leq \log c - \frac{(A+B)t^2}{2AB + \frac{4}{3}(A-B)t - t^2}$$

and can use this to form the procedure. The approximation to the optimal procedure of this type is fairly good unless one of A or B is small (<2/3) and the other is not.

Let us consider the binomial with large variance. Let s denote the number of successes, f the number of failures, $s^* = s + \frac{1}{2}$, $f^* = f + \frac{1}{2}$, $S = (n+1)p$, $F = (n+1)(1-p)$. We can write

$$\log s! = s^* \log s^* - s^* + \frac{1}{2} \log 2\pi - \eta(s^*)$$

where

$$\eta(s^*) \sim \frac{1}{24\ s^*} - \frac{7}{2880\ s^*} + \frac{31}{40320 s^{*5}} + \ldots,$$

Then

$$\log P = c - s^*\log \frac{s^*}{S} - f^*\log \frac{f^*}{F} + \eta(s^*) + \eta(f^*)$$

If $x = s^* - S$, $s^*\log \frac{s^*}{S} + f^*\log \frac{f^*}{F} \geq \frac{(n+1)x^2}{2SF + \frac{2}{3}(S-F)x - x^2/3}$

The limit as $F \to \infty$ is valid and hence we also have a procedure for the Poisson with large variance.

The setup cost is quite high, so if only a few binomials are wanted, we can use further acceptance-rejection procedures as

$$\sum_{n=0}^{\infty} e^{-\alpha-\beta n} \leq \frac{1-\frac{\alpha}{2} + \frac{1}{12}\alpha^2}{1 + \frac{\alpha}{2} + \frac{1}{12}\alpha^2} \left(\frac{1}{\beta} + \frac{1}{2} + \frac{\beta}{12}\right),$$

and

$$\log \frac{1 - \frac{\alpha}{2} + \frac{\alpha^2}{12}}{1 + \frac{\alpha}{2} + \frac{\alpha^2}{12}} = 1 + \sum_{n=1}^{\infty} (-1)^n \left(\frac{\alpha^{6n-1}}{(6n-1)12^{3n-1}} + \frac{\alpha^{6n+1}}{(6n+1)12^{3n}}\right)$$

$$\sum e^{-\beta n} = \frac{1}{\beta} + \frac{1}{2} + \frac{\beta}{12} \left(1 - \frac{\beta^2}{60} + \frac{\beta^4}{2520} + \ldots\right),$$

provided $\alpha \leq \sqrt{12}$ and $\beta < 2\pi$. For $\alpha = \beta = 1$, the rejection probability for this part of the approximation is $< .00233$.

Caveat. In many situations, better procedures can be found or these procedures can be modified for greater efficiency. Finding good procedures is an art, not a science.

REFERENCES

[1] Marsaglia, G., MacLaren, M. D., Bray, T. A. "A fast procedure for generalizing normal random variables", Comm. Assn. Comp. Mach. 7 (1964) pp. 4-10.

[2] Marsaglia, G., MacLaren, M. D., Bray, T. A. "A fast procedure for generalizing exponential random variables. Comm. Assn. Comp. Mach. 7 (1964) pp. 298-300

[3] von Neumann, John. "Various techniques used in connection with random digits", Monte Carlo Method, pp. 36-38, Natl. Bur. of Stand. Appl. Math. Ser. No. 12, Washington, 1951.