Statistical Computing

This Department will carry articles of high quality on all aspects of computation in statistics.
Papers describing new algorithms, programs, or statistical packages will not contain listings of the program, although the completely documented program must be available from the author. Review of the paper will always include a running test of the program by the referee.
The description of a program or package in this Department should not be construed as an endorsement of it by the American Statistical Association or its Committees, nor is any warranty implied about the validity of the program.
The Editorial Committee will be pleased to confer with authors about the appropriateness of topics or drafts of possible articles.

Generating Random Variates Using Transformations with Multiple Roots

JOHN R. MICHAEL, WILLIAM R. SCHUCANY AND ROY W. HAAS*

ABSTRACT

The general approach to generating random variates through transformations with multiple roots is discussed. Multinomial probabilities are determined for the selection of the different roots. An application of the general result yields a new and simple technique for the generation of variates from the inverse Gaussian distribution.

1. *Introduction*

Occasionally it is possible to generate variates, x, from a distribution of interest by a simple application of the inverse probability integral transformation. If the cumulative distribution function, F, has a closed form expression for its inverse, F^{-1}, then it is often efficient to use $x = F^{-1}(u)$, where u is a variate from an acceptable uniform (0, 1) generator. When this is not the case, it is sometimes possible to produce a transformation to the variable of interest from another variable for which a random number generator already exists. For example, Box and Muller [1] have shown how normal variates can be produced from uniform variates using a direct transformation.

In some other instances a known relationship may be of the form

$$v = g(x), \tag{1}$$

and a value of x is sought for each value of v that is generated. When a single-valued inverse does not exist, more than one value of x satisfies (1).

For a specific observation, v_0, suppose that there are k distinct roots of (1) denoted $x_1, x_2, \ldots, x_k$. (Note that k may depend upon v_0.) The problem is how to determine the multinomial probabilities for choosing each of the k roots.

If X and V are discrete random variables then probabilities can be associated with each of the k roots. The conditional probability with which the ith root should be chosen, $p_i(v_0)$, is easily seen to be

$$p_i(v_0) = P[X = x_i | V = v_0] = \frac{P[X = x_i,\ V = v_0]}{P[V = v_0]} = \frac{P[X = x_i]}{\sum_{j=1}^{k} P[X = x_j]}. \tag{2}$$

For the continuous case, a similar expression will be developed for an interval about v_0. Then the limit will be taken as the interval shrinks to the point v_0. The result is not generally a simple ratio of the likelihood of the ith root to the sum of the likelihoods of the k roots.

2. *Main Result*

Suppose X and V are absolutely continuous random variables. Let $f(x)$ and $F(x)$ denote the density function and the cumulative distribution function of X, respectively. Let g be such that the first derivative of g, g', exists, is continuous, and is nonzero, except on a closed set of values for X with probability zero. Consider the interval $(v_0 - h, v_0 + h)$, where $h > 0$. According to the inverse function theorem, for h sufficiently small, the inverse image of $(v_0 - h, v_0 + h)$ is comprised of k disjoint intervals about the k distinct roots. Let the interval containing the ith root, x_i, be denoted (y_{i1}, y_{i2}). If $p_i^h(v_0)$ is the probability with which an observation should be chosen from the ith interval given that V is in the interval $(v_0 - h, v_0 + h)$, then, similar to (2),

$$p_i^h(v_0) = \frac{P[y_{i1} < X < y_{i2}]}{\sum_{j=1}^{k} P[y_{j1} < X < y_{j2}]} = \frac{F(y_{i2}) - F(y_{i1})}{\sum_{j=1}^{k} [F(y_{j2}) - F(y_{j1})]}.$$

Since selection is to be made among the k points $x_1, x_2, \ldots, x_k$ (having observed the point v_0), and these points are the limits $\lim_{h\to 0} [(y_{j1}, y_{j2})] = x_j$ and $\lim_{h\to 0} [(v_0 - h, v_0 + h)] = v_0$, then $p_i(v_0) = \lim_{h\to 0} [p_i^h(v_0)]$ will yield the conditional probability with which the ith root should be selected. Hence

$$p_i(v_0) = \lim_{h\to 0} p_i^h(v_0)$$

$$= \left\{1 + \sum_{j=1, j\neq i}^{k} \lim_{h\to 0} \left[\frac{F(y_{j2}) - F(y_{j1})}{F(y_{i2}) - F(y_{i1})}\right]\right\}^{-1}$$

---

* Dept. of Statistics, Southern Methodist Univ., Dallas, TX 75275. This research was supported by ONR Contract N00014-75-C-0439.

$$= \left\{ 1 + \sum_{j=1, j \neq i}^{k} \lim_{h \to 0} \left[\frac{(y_{j2} - y_{j1})/h}{(y_{i2} - y_{i1})/h} \cdot \frac{[F(y_{j2}) - F(y_{j1})]/(y_{j2} - y_{j1})}{[F(y_{i2}) - F(y_{i1})]/(y_{i2} - y_{i1})} \right] \right\}^{-1}$$

$$= \left\{ 1 + \sum_{j=1, j \neq i}^{k} \left| \frac{g'(x_i)}{g'(x_j)} \right| \cdot \frac{f(x_j)}{f(x_i)} \right\}^{-1}. \tag{3}$$

3. *Applications*

a. *Symmetric Distributions*

Consider first the case in which the random variable of interest, X, with mean μ and variance σ^2, has a symmetric distribution and satisfies the requirements at the beginning of Section 2. Suppose that random observations of V are easily generated, where $V = g(X) = |(X - \mu)/\sigma|$. It is now desired to produce random observations of X. For each variate, v_0, we must choose between the two roots $x_1 = (\mu + \sigma v_0)$ and $x_2 = (\mu - \sigma v_0)$. According to (3), since $f(x_1) = f(x_2)$, and $|g'(x_1)| = |g'(x_2)|$, then $p_1(v_0) = p_2(v_0) = \frac{1}{2}$. The obvious solution is to choose the roots with equal probabilities. We have now justified generating Laplace (double exponential) variates from simple exponential variates or even normal variates from $\chi_{(1)}$ observations. Without the symmetry of f and g above, the correct probabilities may not be so obvious, but (3) can still simplify nicely as the next example demonstrates.

b. *Inverse Gaussian Distribution*

The standard form for the density of the inverse Gaussian distribution [3] is given by

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[\frac{-\lambda(x - \mu)^2}{2\mu^2 x}\right], \quad x > 0, \mu > 0, \lambda > 0.$$

The cumulative distribution function as given by Chhikara and Folks [2] is expressed in terms of cumulatives of the standard normal and is not easily inverted.

Following Shuster [4] we may write

$$V = g(X) = \frac{\lambda(X - \mu)^2}{\mu^2 X} \sim \chi^2_{(1)}. \tag{4}$$

Observations from $\chi^2_{(1)}$ are easily generated as the squares of standard normals. For each chi-square variate, v_0, we must solve (4) for x to obtain a corresponding observation from the inverse Gaussian distribution. For any $v_0 > 0$ there are exactly two roots of the associated quadratic equation which can always be expressed as

$$x_1 = \mu + \frac{\mu^2 v_0}{2\lambda} - \frac{\mu}{2\lambda}\sqrt{4\mu\lambda v_0 + \mu^2 v_0^2}$$

and

$$x_2 = \mu^2/x_1 \tag{5}$$

since the relationship which exists between the roots of any quadratic equation implies here that $x_1 x_2 = \mu^2$.

The difficulty in generating observations with the desired distribution now lies in choosing between the two roots. From the previous section it has already been argued that x_1 should be chosen with probability

$$p_1(v_0) = 1 - p_2(v_0) = \left\{ 1 + \left| \frac{g'(x_1)}{g'(x_2)} \right| \cdot \frac{f(x_2)}{f(x_1)} \right\}^{-1}.$$

Using (5) it can be shown that $f(x_2)/f(x_1) = (x_1/\mu)^3$ and $g'(x_1)/g'(x_2) = -(\mu/x_1)^2$. Hence the smaller root, x_1, should be chosen with probability

$$p_1(v_0) = \frac{\mu}{\mu + x_1}. \tag{6}$$

So for each random observation from a chi-square distribution with one degree-of-freedom, v_0, the smaller root is calculated. An auxillary Bernoulli trial is then performed with $p_1(v_0) = \mu/(\mu + x_1)$. If the trial results in a "success", x_1 is chosen; otherwise, the larger root, $x_2 = \mu^2/x_1$, is chosen.

A typical FORTRAN subroutine for generating the observation might contain code similar to the following:

```
C     V HAS A CHISQUARE(1)
C     DISTRIBUTION
      W = MU*V
C     C = MU/(2.*LAMBDA) NEED NOT
C     BE COMPUTED FOR EACH LOOP
      X1 = MU + C*(W -
          SQRT(W*(4.*LAMBDA + W)))
      P1 = MU/(MU + X1)
C     Y HAS A UNIFORM(0, 1)
C     DISTRIBUTION
      X = X1
      IF (Y .GE. P1) X = MU*MU/X1
C     THE DESIRED VARIATE IS
C     RETURNED IN X
```

The subroutine which we actually implemented returns independent pairs of inverse Gaussian variates since the Box-Muller routine returns pairs of standard normal variates, which are squared to obtain pairs of independent chi-square variates.

In testing the above procedure several large samples were generated with various combinations of the parameters μ and λ. The Kolmogorov-Smirnov goodness-of-fit test was applied to each sample. There was no evidence to indicate that the generated observations did not come from the inverse Gaussian distribution.

When we first encountered the problem of generating random observations from the inverse Gaussian distribution, several other methods were proposed for determining the probability with which the smaller root should be chosen. Each method had some intuitive appeal, but contained a subtle error which could have been avoided by adhering to the procedure developed in Section 2. The correct probability, given by (6), is a simple expression, but quite difficult to produce intuitively.

An alternative general method which has been used successfully for other distributions is that of polynomial approximation to the inverse cumulative. The difficulty in this instance is that μ is not a location parameter and hence a different polynomial would be required for each value of μ. On the other hand, the method of Section 2 is more efficient as well as being exact.

Acknowledgment

We are grateful to the two referees for their valuable comments on an earlier version of this paper.

REFERENCES

1. Box, G. E. P. and Muller, M. E. (1958). A note on the generation of random normal deviates. *Ann. Math. Statist.* **29**, 610–11.
2. Chhikara, R. S. and Folks, J. L. (1974). Estimation of the inverse Gaussian distribution function. *J. Amer. Statist. Assoc.* **69**, 250–4.
3. Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions-1*. Boston: Houghton-Mifflin Co., Ch. 15.
4. Shuster, J. (1968). On the inverse Gaussian distribution function. *J. Amer. Statist. Assoc.* **63**, 1514–16.