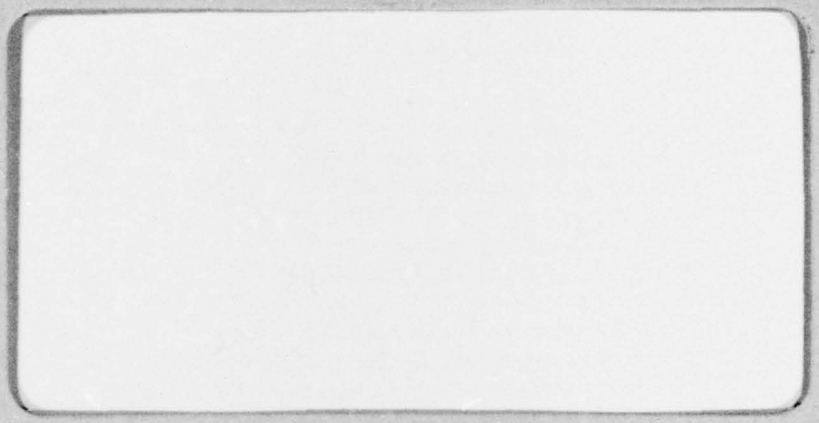
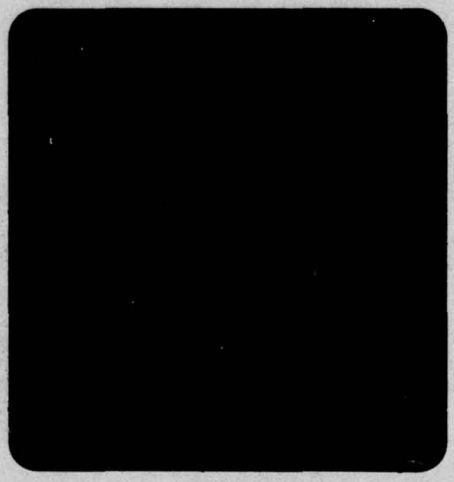
AD A052570

SOUTHERN METHODIST UNIVERSITY

(12)



AD No. ____
DDC FILE COPY



DDC
RECEIVED
APR 12 1978
B

DEPARTMENT OF OPERATIONS RESEARCH AND ENGINEERING MANAGEMENT
SCHOOL OF ENGINEERING AND APPLIED SCIENCE
DALLAS, TEXAS 75275

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DDC
RECEIVED
APR 12 1978
B

(9) Technical Report, OREM-78002 (14)

(12)

(6) REJECTION USING PIECE-WISE LINEAR MAJORIZING FUNCTIONS IN RANDOM VARIATE GENERATION

(10) Bruce W. Schmeiser
and
Mohamed A. Shalaby

Department of Operations Research and Engineering Management
Southern Methodist University
Dallas, Texas 75275

(11) Jan[redacted]78 (12) 14 p.

ABSTRACT

In the context of random variate generation on digital computers, the use of piece-wise linear majorizing functions in conjunction with the general rejection algorithm is proposed. Based on previous results obtained in the generation of beta variates, the expected advantages and disadvantages of applying the concept to other distributions are discussed, as is the use of minorizing functions for fast acceptance of values. Areas of potential application are also discussed.

This work is supported in part by the Office of Naval Research under Contract N00014-77-C-0425. (15)

This paper will be included in the Proceedings of Computer Science and Statistics: Eleventh Annual Symposium on the Interface, North Carolina State University, March 6-7, 1978.

410 597 mt

I. INTRODUCTION

Much literature in the last twenty years has been devoted to process generation on digital computers. Process generation is the creation of a sequence of observations having the properties of some desired distribution or process. Almost always process generation is a transformation of one or more uniform (0,1) values to the desired distribution. Common methods for performing the transformation include using the inverse distribution function transformation, rectangular approximation, special properties, composition, and rejection. (See for example, [2,5,6].) In the past most interest has centered on the first three methods. Recently composition and rejection have received much attention. In a wide variety of cases, rejection is fast, easy to code, and requires little memory.

In this paper the use of piece-wise linear rejection functions and methods for fast acceptance of observations are discussed for the univariate continuous case. The beta distribution is used as an example, drawing on the results of Schmeiser and Shalaby [8]. Discussion centers on concepts necessary for generalizing the results to other distributions. The general rejection algorithm is discussed in Section II and specialized to the piece-wise linear case in Section III. Discussed are limitations in Section IV, fast acceptance in Section V, and potential applications in Section VI.

II. THE GENERAL REJECTION ALGORITHM

In this section a general form of random variate generation using rejection is given. This general form is specialized to the piece-wise linear special case. Implications and applicability of the piece-wise linear approach are discussed.

The common rectangular rejection algorithm can be generalized as follows. Let $p(x)$ be the density function from which random variates are to be generated. Let $t(x)$ be a

majorizing function of $p(x)$; i.e., $t(x) \geq p(x)$ for all x. Corresponding to $t(x)$ is the density function $r(x) = t(x) / k$, where $k = \int_{-\infty}^{\infty} t(x)\, dx$. Figure 1 illustrates the relationship of $p(x)$, $t(x)$ and $r(x)$. The general rejection algorithm for generating variates from $p(x)$ is

1. Generate a value x from $r(\cdot)$.
2. Generate a value u from the rectangular distribution over the interval $[0,1]$.
3. If $u \leq p(x) / t(x)$, accept x by setting $y = x$. Otherwise go to step 1.

Proposition: The algorithm provides values of y from the distribution having density function $p(\cdot)$.

Proof: Let A denote the event that step 3 results in acceptance. In any given iteration

$$P(A \mid x) = p(x) / t(x) = p(x) / [k\, r(x)]$$

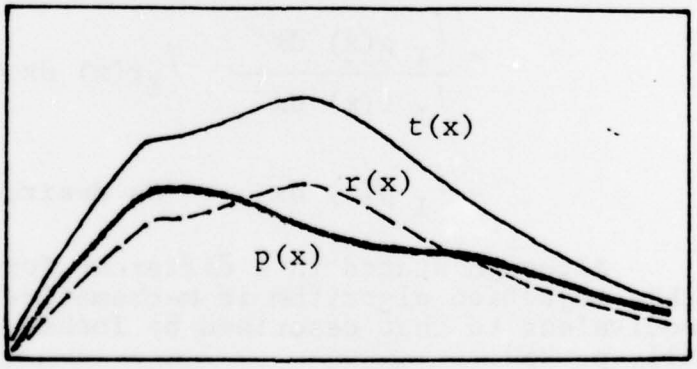


Figure 1. The functions used in the general rejection algorithm for generating random variates from $p(x)$.

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

and therefore

$$P(A \mid X_\varepsilon I) = \frac{P(A \text{ and } X_\varepsilon I)}{P(X_\varepsilon I)}$$

$$= \frac{\int_I P(A \mid x)\, r(x)\, dx}{\int_I r(x)\, dx}$$

$$= \frac{\int_I [p(x)/t(x)]\, r(x)\, dx}{\int_I r(x)\, dx}$$

$$= \frac{\int_I p(x)\, dx}{\int_I t(x)\, dx}$$

and

$$P(A) = \int_{-\infty}^{\infty} P(A \mid x)\, r(x)\, dx = 1/k$$

Let Y be the random variable resulting from the algorithm. It is necessary to show that $P(Y_\varepsilon I) = \int_I p(x)\, dx$ for any interval I. The proof follows directly from

$$P(Y_\varepsilon I) = P(X_\varepsilon I \mid A)$$

$$= P(A \mid X_\varepsilon I)\, P(X_\varepsilon I) \,/\, P(A)$$

$$= \frac{\int_I p(x)\, dx}{\int_I t(x)\, dx} \quad \int_I r(x)\, dx \,/\, [1/k]$$

$$= \int_I p(x)\, dx \qquad \text{as desired.} \blacksquare$$

Although stated in a different form, this rejection algorithm is mathematically equivalent to that described by Tocher [11, p. 25].

For a given majorizing function $t(x) = k\, r(x)$, k is selected to be as small as possible while still maintaining $t(x) \geq p(x)$ for all x. This results in maximizing the probability P(A) that in any given iteration the value x generated in step 1 will be accepted in step 3.

For a given density function p(x) the choice of majorizing function $t(x) = k\, r(x)$

plays a central role in determining whether or not the resulting algorithm is efficient. The majorizing function must both have nearly the same shape as p(x) (thereby resulting in a small value of k) and a density function r(x) which is amenable to variate generation (via any technique, but probably not rejection).

The reason for early disfavor of rejection was the selection of a uniform distribution for r(x). (In fact, many textbooks discuss only this special case.) The rectangular assumption restricts consideration to distributions having a finite range or to approximations obtained by truncation. While such approximations may be made as accurate as desired in theory, great inefficiency results from using a rectangular distribution to model tails having small probabilities.

In the last several years the use of non-rectangular rejection regions has appeared more frequently in the literature. The gamma distribution especially has been the topic of several papers [1,9,10,12]. All of these papers have used density functions r(x) corresponding to well-known distributions. The basic results of these papers is the identification of suitable functions r(x) and the determination of k such that valid and efficient algorithms result.

III. PIECE-WISE LINEAR MAJORIZING FUNCTIONS

While the use of common theoretical distributions for r(x) has been fruitful, another approach which is very general and often easy to apply is to use piece-wise linear majorizing functions. Piece-wise linearization calls for partitioning the range of the random variable into segments such that t(x) is linear over each segment. The usual rectangular rejection region is a special case corresponding to only one segment. Another special case, briefly discussed by Lewis [5, p. 82], is the use of "regular parts," which is a discontinuous piece-wise linear majorizing function having only rectangular segments. More generally, however, the linear segments may lie at the angle providing the best fit to p(x). As an example, Schmeiser and Shalaby [8] used a

piece-wise linear majorizing function in considering rejection methods for the beta distribution. Figure 2 illustrates the algorithm for a particular beta density function.

Step 1 of the algorithm requires generation of variates from the density $r(x) = t(x)/k$. Now the piece-wise linear $r(x)$ is composed of a mixture of rectangular, triangular, and trapezoidal densities. Note in Figure 2 that the trapezoidal densities are each composed of a rectangular lower density and a triangular upper density. Thus, $r(x) = \sum_{i=1}^{n} \alpha_i r_i(x)$ where $\sum_{i=1}^{n} \alpha_i = 1$ and $0 \leq \alpha_i \leq 1$ for all i, $i = 1,2,\ldots, n$ and each $r_i(x)$ is either a rectangular or a triangular density.

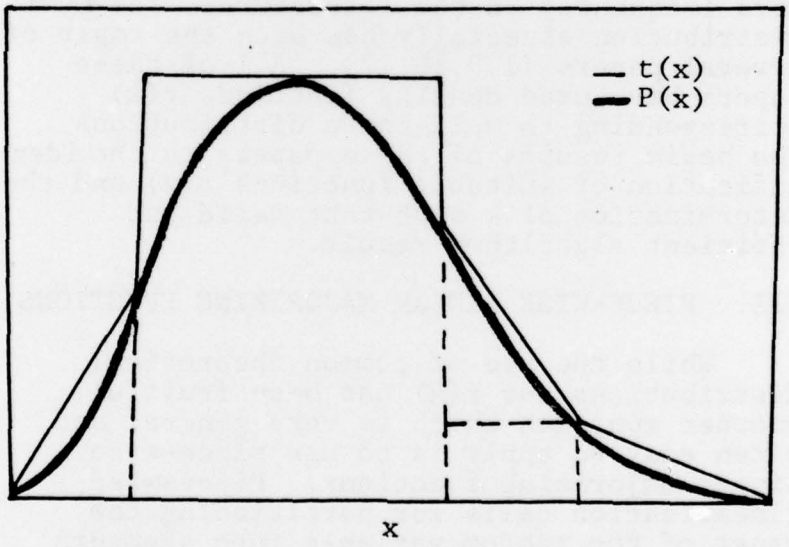


Figure 2. Rejection algorithm using a piece-wise linear majorizing function $t(x)$.

The generation of variates from a piece-wise linear $r(x)$ (using the composition method) requires the generation of a variate

from $r_i(x)$ with probability α_i. The rectangular densities may be easily generated using $x = a + (b - a)\,u$ where u is a uniform (0,1) variate and a and b are the bounds of the rectangular density function. The triangular densities require $x = a + (b - a) \max(u_1, u_2)$ when $r_i(x)$ has a positive slope and $x = a + (b - a) \min(u_1, u_2)$ when $r_i(x)$ has a negative slope, where u_1 and u_2 are independently generated uniform (0, 1) variates. Since generation from a piece-wise linear r(x) requires no exponential level operations, step 1 can be executed quite rapidly.

In addition the probability of acceptance in step 3 can be made close to one, since a piece-wise linear majorizing function can be made to fit any density function p(x) arbitrarily well by increasing the number of segments. Here a trade-off developes between few segments resulting in simple coding (with associated minimal memory requirements) and many segments resulting in longer code, more memory requirements, but higher probability of acceptance. The use of even a few segments provides a considerably better fit than the simple rectangular region. For example, in reference [8] three and five segments are used in two of the beta generation algorithms. While the single segment provides an algorithm which is not competitive for many beta parameter values, five segments are the nucleus of the fastest algorithm available for many parameter values.

IV. LIMITATIONS

The applicability of the rejection technique is dependent only upon the selection of the majorizing function. For a particular density p(x), the minimal value of k such that $k\,r(x) \geq p(x)$ for all x is central to the applicability of the rejection technique. This inequality implies two conditions:
1) $k\,r(x)$ must be greater than zero whenever $p(x)$ is greater than zero and
2) $\lim_{x \to a} k\,r(x) = \infty$ whenever $\lim_{x \to a} p(x) = \infty$.
Since k must be greater than one and finite these two conditions become 1) r(x) must be greater than zero whenever p(x)

is greater than zero and 2) $\lim_{x \to a} r(x) = \infty$ whenever $\lim_{x \to a} p(x) = \infty$. Since the piece-wise linear majorizing function cannot be non-zero at infinity nor be infinite, the piece-wise approach is applicable only as an approximation to distributions having densities with one or more infinite values and to distributions with infinite length ranges.

These are two theoretically important restrictions. For example, the beta distribution with parameters less than one, the gamma distribution with shape parameter less than one, and some members of a general family of distributions of Schmeiser and Deutsch [7] have points at which p(x) is infinite. In addition, many distributions have infinite length ranges, most commonly $(-\infty, \infty)$ and $[0, \infty)$.

However these restrictions are not important in practice. First consider the problem of $\lim_{x \to a} p(x) = \infty$. Few computers have accuracy beyond 10^{-8}, nor do many applications require more accuracy. If ε is the minimum discernable accuracy, then an approximation using the finite values $r(a + \varepsilon)$ or $r(a - \varepsilon)$ rather than the infinite r(a) will probably be quite acceptable. The second problem of infinite length range may be overcome by including so much of the range that the excluded portion will never be missed. For example, consider the normal distribution. While having a range of $(-\infty, \infty)$, the probability of observing a point more than ten standard deviations from the mean is only 1.524 x 10E-23. Of course, if this is unacceptable then one hundred standard deviations can be included with little additional cost.

Although the theoretical limitations are not important, the use of a piece-wise linear majorizing function does require that the density function p(x) yields r(x) with a reasonable amount of effort. For distributions having only a single shape this is not important, since an appropriate r(x) may be determined once and for all. For example, see Kinderman and Ramage's [4] normal variate

generator. However, families of distributions such as the gamma and beta include multiple shapes. A generator designed to generate values from any member of the family must be able to quickly determine an appropriate majorizing function. For example the beta generators of reference [8] use the equations for the location of the mode and points of inflexion to locate the junctures of the piece-wise linear segments. The points of inflexion are critical since the convexity or concavity of the density function at various points is necessary to prove that indeed $k\ r(x) \geq p(x)$ for all x.

V. FAST ACCEPTANCE

In addition to choosing r(x) such that the probability of acceptance is high, another function b(x) may be chosen to reduce the time necessary to determine whether or not $u \leq p(x) / t(x)$ in step 3. If b(x) is substantially faster to evaluate than p(x) and if $b(x) \leq p(x)$ for all x, then the algorithm may be made faster by replacing step 3 with

3(a). If $u\ t(x) \leq b(x)$, accept x by setting $y = x$.

3(b). If $u\ t(x) \leq p(x)$, accept x by setting $y = x$. Otherwise go to step 1.

The theory underlying the algorithm has not changed, since step 3(a) accepts x only if 3(b) would accept x anyway. However, the use of step 3(a) often makes the evaluation of p(x) unnecessary. Since density functions often include time consuming operations such as exponential and gamma functions, the savings due to this minorizing function can be substantial. In particular, a piece-wise linear minorizing function is often easy to determine and is always fast to evaluate.

Sometime a minorizing function is not necessary. If a trapezoidal region is formed by t(x) which is composed of triangular and rectangular regions, then the rectangular region is often entirely under p(x). In this case no check is necessary in step 3, the value of x being accepted automatically.

Note also that the minorizing function b(x) may be used with any majorizing function, not just the piece-wise linear functions discussed in Sections II and III.

VI. POTENTIAL APPLICATIONS

There are a number of distributions to which the above concepts can be applied. The gamma distribution generators currently available involve several logarithmic operations. The use of piece-wise linear majorizing and minorizing functions would almost certainly be faster. Pearson distributions other than the gamma may also be amenable to the piece-wise linear approach. Two families which have well-known, but slow, generators, the Weibull and lognormal, could also be generated using this approach. The F and t distributions, classically generated using their relationships to the normal or the beta distributions, could be more quickly generated using the above techniques. In addition, the J-shaped beta family could benefit from such techniques. (Jöhnk's algorithm [3] for U-shaped and the algorithms discussed in Schmeiser and Shalaby [8] for bell-shaped beta distributions probably preclude much faster times using the above techniques.)

It is also possible that the piece-wise linear techniques can be applied to discrete distributions such as the Poisson and binomial. Current generators for these two distributions require times proportional to the mean of the Poisson and to the number of trials for the binomial. The commonly used normal approximations could be avoided by the use of rejection techniques.

VII. SUMMARY AND CONCLUSIONS

The commonly used rectangular rejection region has been generalized and the most general form of rejection has been specialized to the use of piece-wise linear majorizing functions. The implications of the piece-wise linear approach have been discussed and potential applications have been mentioned.

generator. However, families of distributions such as the gamma and beta include multiple shapes. A generator designed to generate values from any member of the family must be able to quickly determine an appropriate majorizing function. For example the beta generators of reference [8] use the equations for the location of the mode and points of inflexion to locate the junctures of the piece-wise linear segments. The points of inflexion are critical since the convexity or concavity of the density function at various points is necessary to prove that indeed $k\ r(x) \geq p(x)$ for all x.

V. FAST ACCEPTANCE

In addition to choosing r(x) such that the probability of acceptance is high, another function b(x) may be chosen to reduce the time necessary to determine whether or not $u \leq p(x) / t(x)$ in step 3. If b(x) is substantially faster to evaluate than p(x) and if $b(x) \leq p(x)$ for all x, then the algorithm may be made faster by replacing step 3 with

3(a). If $u\ t(x) \leq b(x)$, accept x by setting $y = x$.

3(b). If $u\ t(x) \leq p(x)$, accept x by setting $y = x$. Otherwise go to step 1.

The theory underlying the algorithm has not changed, since step 3(a) accepts x only if 3(b) would accept x anyway. However, the use of step 3(a) often makes the evaluation of p(x) unnecessary. Since density functions often include time consuming operations such as exponential and gamma functions, the savings due to this minorizing function can be substantial. In particular, a piece-wise linear minorizing function is often easy to determine and is always fast to evaluate.

Sometime a minorizing function is not necessary. If a trapezoidal region is formed by t(x) which is composed of triangular and rectangular regions, then the rectangular region is often entirely under p(x). In this case no check is necessary in step 3, the value of x being accepted automatically.

Note also that the minorizing function b(x) may be used with any majorizing function, not just the piece-wise linear functions discussed in Sections II and III.

VI. POTENTIAL APPLICATIONS

There are a number of distributions to which the above concepts can be applied. The gamma distribution generators currently available involve several logarithmic operations. The use of piece-wise linear majorizing and minorizing functions would almost certainly be faster. Pearson distributions other than the gamma may also be amenable to the piece-wise linear approach. Two families which have well-known, but slow, generators, the Weibull and lognormal, could also be generated using this approach. The F and t distributions, classically generated using their relationships to the normal or the beta distributions, could be more quickly generated using the above techniques. In addition, the J-shaped beta family could benefit from such techniques. (Jöhnk's algorithm [3] for U-shaped and the algorithms discussed in Schmeiser and Shalaby [8] for bell-shaped beta distributions probably preclude much faster times using the above techniques.)

It is also possible that the piece-wise linear techniques can be applied to discrete distributions such as the Poisson and binomial. Current generators for these two distributions require times proportional to the mean of the Poisson and to the number of trials for the binomial. The commonly used normal approximations could be avoided by the use of rejection techniques.

VII. SUMMARY AND CONCLUSIONS

The commonly used rectangular rejection region has been generalized and the most general form of rejection has been specialized to the use of piece-wise linear majorizing functions. The implications of the piece-wise linear approach have been discussed and potential applications have been mentioned.

The rejection algorithm using the piece-wise linear majorizing function has both advantages and disadvantages compared to other rejection methods. The disadvantage is that the determination of the piece-wise linear majorizing function can require a non-trivial set-up cost. The expected advantages are

1. applicability to a wide variety of distributions,
2. fast and easy generation of x in step 1,
3. high probability of acceptance in step 3, and
4. fast acceptance in step 3(a).

REFERENCES

1. Fishman, G.S. "Sampling from the gamma distribution on a computer," Communications of the ACM, 19, 7 (July 1976), 407-409.
2. Jannson, B. Random Number Generators, Almquist and Wiksell, Stockholm, 1966.
3. Jöhnk, M.D. "Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen," Metrika, 8 (1964), 5-15.
4. Kinderman, A.J. and Ramage, J.G. "Computer generation of normal random variables," Journal of the American Statistical Association, 71 (December 1971), 893-896.
5. Lewis, T.G. Distribution Sampling for Computer Simulation, D.C. Heath and Co., Lexington, Mass., 1975.
6. McGrath, E.J. and Irving, D.C. Techniques for Efficient Monte Carlo Simulation. Volume II. Random Number Generation for Selected Probability Distributions, NTIS, 5285 Port Royal Road, Springfield, VA 22151, 1973.
7. Schmeiser, B.W. and Deutsch, S.J. "A versatile four-parameter family of probability distributions, suitable for simulation," AIIE Transactions, 9, 2 (June 1977), 176-182.
8. Schmeiser, B.W. and Shalaby, M.A. "Rejection methods for beta variate generation," Technical Report IEOR77014, Southern Methodist University, Dallas, TX 75275.
9. Tadikamalla, P.R. "Computer generation of gamma random variables," Communications of the ACM, forthcoming.
10. Tadikamalla, P.R. "Computer generation of gamma random variables -- II," Technical Report, Department of Business Administration, Eastern Kentucky University, Richmond, KY 40475.
11. Tocher, K.D. The Art of Simulation, English Universities Press, London, 1963.
12. Wallace, N.D. "Computer generation of gamma random variates with non-integral shape parameters," Communications of the ACM, 17 (1974), 691-695.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER OREM 78002	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Rejection Using Piece-Wise Linear Majorizing Functions in Random Variate Generation		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bruce W. Schmeiser Mohamed A. Shalaby		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0425
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Dallas, TX 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Analysis Program Office of Naval Research Arlington, VA 22217		12. REPORT DATE January 1978 13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Distribution of this document is unlimited.

DISTRIBUTION STATEMENT A
Approved for public release; Distribution Unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Simulation
Process Generation
Rejection

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

In the context of random variate generation on digital computers, the use of piece-wise linear majorizing functions in conjunction with the general rejection algorithm is proposed. Based on previous results obtained in the generation of beta variates, the expected advantages and disadvantages of applying the concept to other distributions are discussed, as is the use of minorizing functions for fast acceptance of values. Areas of potential application are also discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*