Comp. & Maths. with Appls. Vol. 7, No. 6, pp. 547–552, 1981
Printed in Great Britain.

0097–4943/81/060547–06$02.00/0

ON THE COMPUTER GENERATION OF RANDOM VARIABLES WITH A GIVEN CHARACTERISTIC FUNCTION

LUC DEVROYE
School of Computer Science, McGill University, 805 Sherbrooke Street West, Montreal, Canada H3A 2K6

(*Received March* 1980)

Communicated by Patrick L. Odell

Abstract—We consider the problem of the computer generation of a random variable X with a given characteristic function when the corresponding density and distribution function are not explicitly known or have complicated explicit formulas. Under mild conditions on the characteristic function, we propose and analyze a rejection/squeeze algorithm which requires the evaluation of one integral at a crucial stage.

1. INTRODUCTION

Consider the problem of the computer generation of a random variable X with a given continuous distribution function F. It is well-known that when U is a uniform (0, 1) random variable, then $F^{-1}(U)$ has distribution function F (principle of *inversion*). Often F is hard to invert but the density f of X is given in analytical form. One may then combine one or more of the following techniques to generate X on a computer: the *rejection* method, the *composition* method, the *Forsythe-von Neumann* method[1, 2], the *squeeze* method[3], the *ratio-of-uniforms* method[4] or the *partial integration* method[5, 6].

In some applications, statisticians are given the characteristic function ϕ of X and the computation of either F or f from ϕ is hard. Often one is not willing to construct a gigantic table of values for F and/or f, and then use an interpolation type algorithm for the generation of random numbers (e.g. Ref. [7]). In this note, we will give a couple of direct methods for the computer generation of X when ϕ is given, and we will put mild conditions on the class of characteristic functions considered here.

2. MAIN RESULTS

Let the random variable X have density f and characteristic function

$$\phi(t) = E(e^{itX}) = \int e^{itx} f(x)\, dx.$$

To generate X on the computer, we will derive an integrable function g that dominates f: $g \geq f$, and use the rejection principle. For this derivation, we will need some conditions on ϕ because the tail behavior of f is related to the smoothness of $\phi(t)$ near $t = 0$. We have:

Inequality 1. If the characteristic function ϕ of a random variable X is twice differentiable, and ϕ, ϕ' and ϕ'' are absolutely integrable and absolutely continuous, then X has a density f satisfying:

$$f(x) \leq \min\left(c, \frac{k}{x^2}\right) \tag{1}$$

where

$$c = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\phi(t)|\, dt, \tag{2}$$

$$k = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\phi''(t)|\, dt. \tag{3}$$

Proof. By the relation

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\phi(t)\,dt \tag{4}$$

which is valid whenever ϕ is absolutely integrable [8], and by partial integration (which is allowed since ϕ and its first two derivatives are absolutely continuous and absolutely integrable), we have

$$f(x)=\frac{1}{2\pi xi}\int_{-\infty}^{+\infty} e^{-itx}\phi'(t)\,dt=-\frac{1}{2\pi x^2}\int_{-\infty}^{+\infty} e^{-itx}\phi''(t)\,dt, \tag{5}$$

from which (3) follows trivially. Also, (2) is an immediate consequence of (4).

Thus, for a large class of characteristic functions, f is bounded from above by

$$g(x)=\min\left(c,\frac{k}{x^2}\right). \tag{6}$$

The area under g is easily seen to be $A=4\sqrt{(kc)}$. The smaller A is, the sharper the inequality $f(x)\le g(x)$ is.

LEMMA 1.

Let $k, c>0$ be arbitrary positive numbers. When V_1 and V_2 are i.i.d. uniform $(-1, 1)$ random variables, then

$$\sqrt{\frac{k}{c}}\frac{V_1}{V_2}$$

has density $A^{-1}g(x)$ where g is defined in (6), and $A=4\sqrt{(kc)}=\int_{-\infty}^{+\infty} g(x)\,dx$.

Proof. When $x<1$, we have $P(|V_1/V_2|<x)=x/2$, and when $x>1$, we have $P(|V_1/V_2|<x)=1-1/2x$. Thus, the density of $|V_1/V_2|$ evaluated at x is min $(1/2, 1/2x^2)$. The generalization towards the density of $\sqrt{(k/c)}.\ V_1/V_2$ is trivial.

In principle, we are now able to generate X by the rejection method provided that we are able to compute the integral (4) with any desired accuracy. The basic algorithm is outlined below.

Algorithm

(1) Generate V_1 and V_2 i.i.d. uniform $(-1,+1)$, and U uniform $(0, 1)$ independent of V_1 and V_2. Set $X\leftarrow\sqrt{(k/c)}\ V_1/V_2$. If $|V_1|<|V_2|$, go to 3.

(2) If $kU<f(X)X^2$, exit with X. Otherwise, go to 1. Here f is evaluated with the aid of formula (4).

(3) If $cU<f(X)$, exit with X. Otherwise, go to 1. Here again, f is evaluated with the aid of formula (4).

Remark 1. The average number of times step 1 is executed is

$$A=4\sqrt{(kc)}=\frac{2}{\pi}\left[\int|\phi|\int|\phi''|\right]^{1/2}.$$

This is also the average number of evaluations of the integral (4). One should keep in mind however that no inversion is necessary as, say, in the solution of

$$\int_{-\infty}^{X} f(x)\,dx=U$$

when U is given and f is given.

Remark 2. When an explicit solution of (4) is possible, we are back in the case "f is known", and the algorithm given above reduces to the ratio-of-uniforms method.

Remark 3. When X is a symmetric random variable, it is known that ϕ is even and that its imaginary part vanishes. Thus, (4) can be simplified to

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos(tx)\phi(t)\,dt. \tag{7}$$

Remark 4. For optimal efficiency, $\int|\phi|\int|\phi''|$ should be small. When X has characteristic function ϕ, then $X + m$ has characteristic function $e^{itm}\phi(t)$, and we may ask ourselves the question what the optimal value is for m. Clearly, $\int|\phi|$ is not affected by m. But $\phi''(t) = -\int e^{itx}x^2 f(x)\,dx$ ([8], p. 199) at least whenever $E(X^2) < \infty$, so that $\phi''(0) = E(X^2)$ and $|\phi''(t)| \le E(X^2)$, all t. When X is replaced by $X + m$, then $|\phi''(t)| \le E((X+m)^2)$, and this is minimal when $m = -E(X)$. Therefore, generally speaking, the efficiency of the algorithm will be enhanced by centering at the mean.

Remark 5. Inequality (1) is not applicable for the important class of characteristic functions that are real (hence, even) and whose gradient and second derivative exist except possibly at $t = 0$. For example, the Cauchy characteristic function $\exp(-|t|)$ falls in this category. It is not hard to establish however that when ϕ is real, ϕ, ϕ' and ϕ'' are absolutely continuous and absolutely integrable on $(0, \infty)$, inequality (1) remains valid with the same c but a different k:

$$k = \frac{1}{\pi}\left[|\phi'(0)| + \int_0^{\infty} |\phi''(t)\,dt\right]. \tag{8}$$

Indeed, by partial integration on $(0, \infty)$, we have

$$f(x) = \frac{1}{\pi}\int_0^{\infty} -\frac{\sin(tx)}{x}\phi'(t)\,dt = -\frac{\phi'(0)}{\pi x^2} - \frac{1}{\pi}\int_0^{\infty}\frac{\cos(tx)}{x^2}\phi''(t)\,dt \le \frac{k}{x^2}. \tag{9}$$

The derivative $\phi'(0)$ is the derivative from the right at 0. If in addition ϕ is convex on $(0, \infty)$, and if ϕ'' decreases monotonically on $(0, \infty)$, then

$$f(x) \le \frac{|\phi'(0)|}{\pi x^2}. \tag{10}$$

We finally notice that if ϕ vanishes off $(-\alpha, +\alpha)$ for some finite α, and if ϕ is real, and ϕ, ϕ' and ϕ'' exist and are absolutely continuous and absolutely integrable on $(0, \alpha)$ (right derivatives are considered at $t = 0$ and left derivatives at $t = \alpha$), then inequality (1) is valid with

$$k = \frac{1}{\pi}\left[|\phi'(0)| + |\phi'(\alpha)| + \int_0^{\infty} |\phi''(t)|\,dt\right]. \tag{11}$$

An example of a characteristic function in this class is $1 - |t|$, $|t| < 1$.

Remark 6. For the evaluation of (4) in practice, we refer to the literature on the evaluation of Fourier integrals and numerical integration in general (see for instance Ref. [9]).

3. IMPROVEMENTS AND EXAMPLES

We need a fast and good numerical integration algorithm for the evaluation of the integral (4). Fortunately, in many cases, a quick analysis of ϕ will allow us to avoid evaluating (4) with large probability. For example, when ϕ can be written as

$$\sum_{i=1}^{n} a_i\phi_i$$

where $a_1, \ldots, a_n$ is a probability vector and $\phi_1, \ldots, \phi_n$ are characteristic functions, then *composition* can be used to generate X on a computer. First generate a $\{1, \ldots, n\}$-valued random variable I with probability vector $(a_1, \ldots, a_n)$, and then generate X with characteristic function ϕ_l. Savings will result when ϕ_i is easy from a simulation viewpoint (e.g., ϕ_i has a known density) whenever a_i is large.

We recall further that when $X_1, \dots, X_n$ are independent random variables with characteristic functions $\phi_1, \dots, \phi_n$, then $\sum_{i=1}^{n} X_i$ has characteristic function

$$\prod_{i=1}^{n} \phi_i,$$

and that $aX_1 + b$ has characteristic function $e^{itb}\phi_1(at)$.

The problem with the composition method here is the establishment that each ϕ_i is a valid characteristic function (e.g. suppose ϕ is given, and that ϕ_1 is an "easy" characteristic function, what is the largest value for a_1 such that $(\phi - a_1\phi_1)/(1-a_1)$ is a valid characteristic function?). Bochner's theorem ([8], p. 207) which states that ϕ is a characteristic function iff $\phi(0) = 1$, ϕ is continuous and ϕ is nonnegative definite is of little use *to us*.

In such situations, the squeeze principle may be applied. As an example, consider a real, nonnegative characteristic function ϕ that is unimodal at $t = 0$ and has two strictly monotone tails (e.g. $\phi(t) = (1+t^2)^{-1}$). By the inequalities

$$\cos tx \le 1 - \frac{t^2x^2}{2} + \frac{t^4x^4}{24} \le 1 - bt^2x^2, |tx| < \frac{\pi}{2}$$

and

$$\cos tx \ge 1 - \frac{t^2x^2}{2}, \text{all } x,$$

where

$$b = \frac{1}{2}\left(1 - \frac{\pi^2}{48}\right),$$

we have

$$f(x) \le h_2(x) = \frac{1}{2\pi}\int_{|tx|<(\pi/2)} (1 - bt^2x^2)|\phi(t)|\,\mathrm{d}t \tag{12}$$

and

$$f(x) \ge h_1(x) = \frac{1}{2\pi}\int_{|tx|<(3\pi/2)} \left(1 - \frac{t^2x^2}{2}\right)|\phi(t)|\,\mathrm{d}t. \tag{13}$$

We have squeezed f between two functions, h_1 and h_2. In many examples, h_1 and h_2 are explicitly known because ϕ and $t^2\phi$ have often known indefinite integrals whereas $\cos(tx)\phi$ has not. Notice also that $h_2 \le c$ where c is defined in (2). The basic algorithm with squeezing is:

(1) Generate X with density $A^{-1}g$, and generate U uniform $(0,1)$ independent of X. Set $T \leftarrow Ug(X)$.

(2) If $T \le h_1(X)$, exit with X.

(3) If $T > h_2(X)$, go to 1.

(4) If $T \le f(X)$ (where f is computed using (4)), exit with X. Otherwise, go to 1.

Since h_1 and h_2 are explicitly known, we will avoid computing the integral in (4) some of the time.

Example 1

Some limit distributions. The weighted Cramer-von Mises statistic (see Ref. [10], for a survey on this topic) has a limit distribution that depends upon the weight function that is used in the definition of the statistic. Some of the limit distribution functions are known in an infinite series format (see Refs. [11] or [12]). At least in two instances, the limit characteristic functions are very simple while the corresponding distribution functions are not:

$$\phi_1(t) = \left[\frac{\sqrt{(2t)}}{\sinh\sqrt{(2t)}}\right]^{1/2} = \left[1 + 2\frac{|t|}{3!} + 2\frac{|t|^2}{5!} + \cdots\right]^{-1/2},$$

and

$$\phi_2(t) = \left[\frac{-2\pi it}{\cos\left(\frac{\pi}{2}\sqrt{(1+8it)}\right)}\right]^{1/2}.$$

The characteristic function ϕ_1 varies near $t = 0$ as $1 - (|t|/6)$, and therefore, inequality 1 is not straightforwardly applicable. However, remark 5 applies here, and (1) remains valid with c as in (2) and k as in (8).

Example 2

Convex characteristic functions. Polya has shown that every even real function ϕ, convex on $(0, \infty)$ with $\phi(0) = 1$ is a characteristic function (see Ref. [8], p. 217, Ex. 13). The proof is based on the approximation from below of ϕ by sums of triangular characteristic functions. The standard triangular characteristic function is

$$\phi(t) = 1 - |t|, \quad |t| < 1.$$

It is known to induce the density

$$f(x) = \frac{1 - \cos(x)}{\pi x^2}.$$

The random variate generation problem is trivial here. Nevertheless, we will use the triangular density as an example to test the tightness of some of the inequalities given above. It is clear that remark 5 applies here so that inequality (1) is valid with c as in (2), and k as in (11). Simple integration then yields

$$c = \frac{1}{2\pi}, \quad k = \frac{2}{\pi}.$$

The average work needed per variate is proportional to $A = 4\sqrt{(kc)} = 4/\pi$. Thus, the bound $f \le g$ is very tight here. As a matter of fact, since A is so small, we may as well use the basic algorithm outlined in this note with of course the integral in (4) replaced by the explicit expression for f. This gives:

(1) Generate V_1, V_2 i.i.d. uniform $(-1, 1)$ random variates, and set $X \leftarrow 2\,V_1/V_2$. Generate U uniform $(0, 1)$ independent of V_1, V_2. If $|X| < 2$, go to 3.

(2) If $2U < 1 - \cos X$, exit with X. Otherwise, go to 1.

(3) If $UX^2 < 2\,(1 - \cos X)$, exit with X. Otherwise, go to 1.

Since we have $1 - \cos(x) \ge (x^2/2) - (x^4/24)$, it is clear that we may quickly accept in step 2 when $2U < X^2/2 - (X^4/24)$, and that we may quickly accept in step 3 when $U < 1 - (X^2/12)$. Thus we can effectively avoid the cosine evaluation most of the time. Since all real convex characteristic functions are compositions (not necessarily finite!) of triangular characteristic functions, one might be able to use the given algorithm for $\phi(t) = 1 - |t|$ as a building block in a more sophisticated composition-type algorithm for convex characteristic functions in general.

Example 3

Famous densities. The bounds derived here are surprisingly tight for many common unimodal densities. For example, the standard *normal* density has characteristic function

$\exp(-t^2/2)$. One can easily check that $c = (1/\sqrt{(2\pi)})$, and that $k = (1/\sqrt{(2\pi)})E(|X^2 - 1|)$ where X is standard normally distributed. So that $A = 4\sqrt{(kc)}$ is once again close to 1.

When X is *gamma* distributed with parameters n and 1, then $X - n$ has a density with the mode at 0. Its characteristic function is

$$\phi(t) = \frac{e^{-itn}}{(1 - it)^n}.$$

It is a straightforward exercise to show that the values for c and k vary with n in such a way that $A = 4\sqrt{(kc)}$ tends to the "normal" A as $n \to \infty$.

The *Cauchy* density $f(x) = \pi^{-1}(1 + x^2)^{-1}$ has characteristic function $e^{-|t|}$. By inequality 1 and remark 5 on convex characteristic functions (see (10)), we see that (1) is valid with $c = 1/\pi$ and $k = 1/\pi$. The important constant A equals $4/\pi$, as in the case of the triangular characteristic function.

REFERENCES

1. G. E. Forsythe, Von Neumann's comparison method for random sampling from the normal and other distributions. *Math. Comput.* **26**, 817–826 (1972).
2. J. von Neumann, Various techniques used in connection with random digits. *Monte Carlo Method, National Bureau of Standards Series* **12**, 36–38, (1951).
3. G. Marsaglia. The squeeze method for generating gamma variates. *Comput. Math. Applic.* **3**, 321–325 (1977).
4. A. J. Kinderman and J. F. Monahan, Computer generation of random variables using the ratio of uniform deviates. *ACM Trans. Math. Software*, **3**, 257–260, 1977.
5. G. A. Mikhailov, On modelling random variables for one class of distribution laws. *Theory Prob. Applic.* **10**, 681–682 (1965).
6. I. Lux, A special method to sample some probability density functions. *Computing* **20**, 183–188 (1978).
7. E. L. Butler, Algorithm 370. General random number generator. *Comm. ACM* **13**, 49–52 (1970).
8. M. Loeve, *Probability Theory*. Van Nostrand, Princeton, New Jersey (1963).
9. P. J. Davis and P. Rabinowitz, *Numerical Integration*. Blaisdell, London (1967).
10. D. A. Darling, The Kolmogorov–Smirnov, Cramer-von Mises tests. *Ann. Math. Stat.* **26**, 1–20 (1955).
11. N. V. Smirnov, On the Cramer-von Mises criterion. *Uspehi Matem. Nauk* **4**, 196–197 (1949).
12. T. W. Anderson and D. A. Darling, Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *Ann. Math. Stat.* **23**, 193–212 (1952).