

MATHEMATICAL ASPECTS OF VARIOUS METHODS FOR SAMPLING FROM CLASSICAL DISTRIBUTIONS

Ulrich Dieter
 Institute of Statistics
 Graz University of Technology
 A-8010 Graz, Lessingstrasse 27
 Austria

1. INTRODUCTION

Simulations were one of the first subjects J.v. Neumann had in mind for applying computers. He recommended some methods for generating $[0, 1]$ -uniformly distributed random numbers which are out of use today. However, his ideas for sampling from non-uniform distributions are still widely used. Since his 1951-paper is rather short, improvements and generalisations of his ideas have been found in the meantime. They will be discussed in this paper.

The most natural method for sampling from a distribution function $F(x)$ is *Direct Inversion*: $X = F^{-1}(U)$ has the distribution function $F(x)$ if F^{-1} is the inverse function to F and if U is uniformly distributed in $[0, 1]$. However, there are only a few distributions for which simple inversion functions are known: the *exponential* and the *Cauchy* distribution are the most prominent examples. The *Box-Muller* method for sampling from the *two-dimensional normal distribution* can also be considered as two-dimensional inversion. However, piecewise polynomial approximations of $F^{-1}(U)$ can always be constructed. Since they need large tables of constants the evolving algorithms are often slow if the parameters of the distribution function have to be changed frequently. For details see the papers of Ahrens and Kohrt (1981) and Chen and Asau (1974).

Calculation of the logarithm was slow in these early days. Therefore J.v. Neumann invented his famous *Comparison Method* for sampling from the exponential distribution. Forsythe (1972) and Ahrens and Dieter (1973) generalised it to the normal distribution; the resulting algorithms FL_n are still some of the fastest assembler methods for sampling from the normal distribution, but they need a table of size 4×2^n . For other distributions the situation is not very promising. Atkinson and Pearce (1976) tried to apply it to the *gamma-distribution*. Unfortunately, for every parameter of this distribution one needs a set of large tables, and in the case of changing parameters the resulting algorithm is rather slow. Generalisations of the comparison method did not yield

efficient procedures.

During the last twenty years it became clear that the *Acceptance-Rejection Method*, also introduced by J.v. Neumann, is the most adaptable method for sampling from complicated distributions. It works as follows:

Let $f(x)$ be a given probability density, and let $h(x)$ be a function such that $f(x) \leq h(x)$ within the range of $f(x)$. If the integral of $h(x)$ over this range is a finite number α , then $g(x) = h(x)/\alpha$ is a probability density function, and the following procedure is valid:

1. Take a random sample X from the distribution with probability density $g(x) = h(x)/\alpha$.
2. Generate a uniform random deviate U between zero and one. If $U \leq f(X)/h(X)$, accept X as a sample from the distribution $f(x)$. Otherwise reject X and go back to Step 1.

The ease of the method depends on the following properties of the hat-function $h(x)$:

- A. One has to select a hat function $h(x)$ from which it is easy to sample. Examples are normal, double-exponential and triangular densities.
- B. The parameters of the hat function have to be determined in such a way that the area α below $h(x)$ becomes minimal.

It will be shown that optimal hat functions can be calculated by analytical methods. Some of the published algorithms use hat-functions which are far away from the optimal ones.

In 1977 A.J. Kinderman and J.F. Monahan introduced a new method for sampling from a density $f(x)$, called *Ratio of Uniforms*. It consists of two parts: let m , s and t be real numbers and let k be equal to 1, but $k = 2$ might be another possible choice for some densities. First the *Table-Mountain-Function*

$$y = \begin{cases} \left(\frac{t}{m-x}\right)^{k+1} & x \in (-\infty, m-t] \\ 1 & x \in [m-t, m-t+s] \\ \left(\frac{s-t}{x-m}\right)^{k+1} & x \in [m-t+s, \infty) \end{cases}$$

is constructed by the transformation

$$X = m + \frac{sV - t}{U^{1/k}}, \quad Y = U^{1+1/k}$$

as a hat function for $f(x)/f$ where $f = \max f(x)$. Finally, the constants m , s and t have to be determined in such a way that the table-mountain-function is an optimal hat function; $m = 0$ and $k = 1$ lead to the original *Ratio of Uniforms* method. For many densities table-mountain-functions with shift $m \neq 0$ or asymmetrical shape $s \neq 2t$ may produce algorithms with even better performance.

The *Acceptance-Complement* method has been introduced independently by Ahrens and Dieter (1982a,b), Kronmal and Peterson (1981), and Deák (1981); it needs the original acceptance-rejection procedure for tackling the area where the involved densities differ.

For classical discrete distributions efficient procedures based on acceptance-rejection, ratio-of-uniforms or acceptance-complement are available — see Ahrens and Dieter (1980), (1982b), (1989) and Stadlober (1989a,b). However, for finite distributions Walker's *Alias-Method* of 1977 yields often the fastest algorithm. It relies on the fact that every n -point distribution is an equiprobable mixture of $n - 1$ 2-point distributions.

In the paper several issues are discussed: (1) Details of the acceptance-rejection method and its generalisations. (2) The construction of optimal hat functions by analytical means. (3) The Ratio of Uniforms method as a special case of the acceptance-rejection procedure and optimal constants for it. (4) Conditions for optimal polynomial squeeze functions are derived. (5) Special distributions are considered as examples for the general procedure.

2. OPTIMAL HAT FUNCTIONS FOR THE ACCEPTANCE-REJECTION METHOD

In this section it is assumed that the hat function $h(x)$ touches $f(x)$ at two points L (left) and R (right) where $L < R$. Furthermore, we suppose that $h(x)$ depends on two parameters, called m and s . Thus we demand that

$$f(L) = \alpha g(L; m, s), \quad f(R) = \alpha g(R; m, s) \quad (1)$$

and $f(x) \leq \alpha g(x; m, s)$ for all other x . Since L and R are local maxima of $f(x)/g(x; m, s)$, we have the necessary conditions

$$\frac{f'(L)}{f(L)} = \frac{g'(L; m, s)}{g(L; m, s)} \quad (2)$$

and

$$\frac{f'(R)}{f(R)} = \frac{g'(R; m, s)}{g(R; m, s)} \quad (3)$$

If L and R are uniquely determined, they should satisfy the sufficient conditions

$$\frac{f''(L)}{f(L)} < \frac{g''(L; m, s)}{g(L; m, s)} \quad \text{and} \quad \frac{f''(R)}{f(R)} < \frac{g''(R; m, s)}{g(R; m, s)}.$$

Otherwise, the first derivative of $\ln(f(x)/g(x; m, s))$ has to be discussed in detail.

Equations (1), (2) and (3) are four equations for the determination of L , R , m , s and α . Assuming that L , R , and m can be expressed as functions of s , we have to minimize

$$\alpha(s) = \frac{f(L(s))}{g(L(s); m(s), s)} = \frac{f(R(s))}{g(R(s); m(s), s)}. \quad (4)$$

This leads to the necessary conditions

$$\begin{aligned} -\frac{d}{ds} \ln \alpha(s) &= -\frac{d}{dL} \ln \frac{f(L)}{g(L; m, s)} \frac{dL}{ds} \\ &\quad + \frac{d}{dm} \ln g(L; m, s) \frac{dm}{ds} + \frac{d}{ds} \ln g(L; m, s) = 0 \\ -\frac{d}{ds} \ln \alpha(s) &= -\frac{d}{dR} \ln \frac{f(R)}{g(R; m, s)} \frac{dR}{ds} \\ &\quad + \frac{d}{dm} \ln g(R; m, s) \frac{dm}{ds} + \frac{d}{ds} \ln g(R; m, s) = 0. \end{aligned}$$

In both equations the first expression after the equal sign is zero by (2) and (3). Solving both equations for dm/ds , and comparing, yields the fundamental relation

$$\frac{d \ln g(L; m, s)}{dm} \frac{d \ln g(R; m, s)}{ds} = \frac{d \ln g(R; m, s)}{dm} \frac{d \ln g(L; m, s)}{ds} \quad (5)$$

or, by observing

$$\frac{d}{dm} \ln g(L; m, s) = \frac{1}{g(L; m, s)} \frac{d}{dm} g(L; m, s)$$

the equivalent form

$$\frac{d g(L; m, s)}{dm} \frac{d g(R; m, s)}{ds} = \frac{d g(R; m, s)}{dm} \frac{d g(L; m, s)}{ds} \quad (6)$$

(2), (3), (4) and (5) or (6) contain five conditions for finding candidates L , R , m , s and α . Whether a solution will in fact lead to a local minimum of α has to be checked carefully in each special case.

We shall consider two examples of possible hat functions $h(x)$ that touch given probability densities $f(x)$ at two locations L and R .

Triangular Hat Functions. The first example deals with densities that can be enclosed in an isosceles triangle $h(x)$, and whose corresponding density $g(x) = h(x)/\alpha$ depends on the parameters m and s as follows.

$$g(x; m, s) = \frac{1}{s} - \frac{1}{s^2} |x - m| \quad x \in [m - s, m + s].$$

Samples may be obtained as $X \leftarrow m + s(U_1 + U_2 - 1)$ where U_1 and U_2 are $[0, 1]$ -uniformly distributed.

First of all, we apply (2) and (3).

$$\frac{f'(L)}{f(L)} = \frac{1}{s - m + L}, \quad \frac{f'(R)}{f(R)} = -\frac{1}{s - R + m}, \quad (7)$$

where $L < m < R$. L and R are local maxima of $f(x)/g(x; m, s)$ if

$$f''(L) < 0 \quad \text{and} \quad f''(R) < 0$$

are satisfied. Now (4) reads

$$\alpha = \frac{f(L)s^2}{s - m + L} = \frac{f(R)s^2}{s - R + m} \quad (8)$$

and the minimization leads to the fundamental identity

$$s = R - L. \quad (9)$$

From the middle part of (8) m is calculated as

$$m = \frac{Rf(R) + Lf(L)}{f(R) + f(L)}, \quad (10)$$

which yields

$$\alpha = (R - L)(f(R) + f(L)). \quad (11)$$

Furthermore, since $f(x)$ touches an isosceles triangle at L and R , it is obvious that $f'(R) = -f'(L)$. This and (9) substituted into (7) yields

$$f'(R) = -f'(L) = -\frac{f(L)}{R - m} = -\frac{f(R)}{m - L}$$

and therefore

$$(L - R)f'(R) = (R - L)f'(L) = f(L) + f(R). \quad (12)$$

Now all parameters can be calculated: usually (12) determines L and R , and (9), (10) and (11) yield s , m , and α .

Double Exponential Hat Functions. Our second example of a hat function with two points of osculation, L and R , has infinite range; it is the double exponential (or Laplace) distribution with density

$$g(x; m, s) = \frac{1}{2s} \exp\left(-\frac{|m - x|}{s}\right).$$

Samples are obtained as $X \leftarrow m + TsE$ where E is a standard exponential deviate and T a random sign \pm .

This time (2) and (3) read

$$\frac{f'(L)}{f(L)} = \frac{1}{s}, \quad \frac{f'(R)}{f(R)} = -\frac{1}{s}, \quad (13)$$

where L and R , $L < m < R$, are local maxima of $f(x)/g(x; m, s)$ provided that

$$\frac{f''(L)}{f(L)} - \frac{1}{s^2} < 0 \quad \text{and} \quad \frac{f''(R)}{f(R)} - \frac{1}{s^2} < 0.$$

Now (4) may be written as

$$\ln \alpha(s) = \ln f(L(s)) + \ln 2s + \frac{m - L}{s} \quad (14)$$

and

$$\ln \alpha(s) = \ln f(R(s)) + \ln 2s + \frac{R - m}{s} \quad (15)$$

and minimization leads to the fundamental identity (6), which results in

$$2s = R - L. \quad (16)$$

Combining (13) and (16) determines L and R by solving the two equations

$$(R - L)f'(L) - 2f(L) = 0, \quad (17)$$

$$(R - L)f'(R) + 2f(R) = 0. \quad (18)$$

Subsequently, m is obtained from (14), (15) and (16) as

$$m = \frac{1}{2}(R + L) + \frac{R - L}{4} \ln \frac{f(R)}{f(L)}. \quad (19)$$

Adding the two representations of $\ln \alpha(s)$ in (14) and (15) yields

$$\alpha = (R - L)e^{\sqrt{f(L)f(R)}}. \quad (20)$$

The theory simplifies if $f(x)$ is symmetric about zero. This means $L = -R$, $f(L) = f(R)$, $f'(L) = -f'(R)$ and hence $m = 0$. The previous results (16) and (20) yield $s = R$ and $\alpha = 2eRf(R)$ and R is determined by (17) which now reads $Rf'(R) + f(R) = 0$. This determines all parameters. The sufficient condition becomes $R^2 f''(R) - f(R) < 0$. R is optimal if this is satisfied.

3. THE RATIO OF UNIFORMS METHOD

Let U and V be $[0, 1]$ -uniformly distributed random variables. Consider the transformation

$$X = m + \frac{sV - t}{U^{1/k}}, \quad Y = U^{1+1/k}.$$

Since the Jacobian

$$\frac{\partial(X, Y)}{\partial(U, V)} = \begin{vmatrix} -\frac{sV - t}{kU^{1+1/k}} & \frac{s}{U^{1/k}} \\ (1 + 1/k)U^{1/k} & 0 \end{vmatrix} = s(1 + 1/k)$$

is constant, the variables X, Y are again uniformly distributed in their domain D . It is bounded by the following curves:

The line $U = 0, V \in [0, 1]$ corresponds to $-\infty < X < \infty, Y = 0$.

The line $U = 1, V \in [0, 1]$ corresponds to $X \in [m - t, m - t + s], Y = 1$.

The line $U \in [0, 1], V = 0$ corresponds to $X = m - tU^{-1/k}, Y = U^{1+1/k}$, i.e. $X \in (-\infty, m - t], Y = \left(\frac{t}{m - X}\right)^{k+1}$.

The line $U \in [0, 1], V = 1$ corresponds to $X = m + (s - t)U^{-1/k}, Y = U^{1+1/k}$, i.e.

$$X \in [m - t + s, \infty), Y = \left(\frac{s - t}{X - m}\right)^{k+1}.$$

This means that the transformed area is bonded by the *table mountain function*

$$h(x) = \begin{cases} \left(\frac{t}{m-x}\right)^{k+1} & x \in (-\infty, m-t] \\ 1 & x \in [m-t, m-t+s] \\ \left(\frac{s-t}{x-m}\right)^{k+1} & x \in [m-t+s, \infty) \end{cases}$$

Its area D is equal to $s(1 + 1/k)$.

Assume that $f(x)$ is a density and $f = \max_x f(x)$.

To simplify the notation we use

$$\tilde{f}(x) = \frac{1}{f} f(x) = \frac{f(x)}{\max_x f(x)}.$$

For applying the acceptance-rejection method

$$\tilde{f}(x) \leq h(x)$$

must hold. This means

$$\begin{aligned} (m-x)^{k+1} \tilde{f}(x) &\leq t^{k+1}, & x \leq m-t \\ (x-m)^{k+1} \tilde{f}(x) &\leq (s-t)^{k+1}, & x \geq m-t+s \end{aligned}$$

Define

$$\begin{aligned} v_- &= \inf \{ (x-m)(\tilde{f}(x))^{1/(1+k)} \mid x \leq m-t \} \\ &= -\sup \{ (m-x)(\tilde{f}(x))^{1/(1+k)} \mid x \leq m-t \} \\ v_+ &= \sup \{ (x-m)(\tilde{f}(x))^{1/(1+k)} \mid x \geq m-t+s \} \end{aligned}$$

Our assumptions mean that v_- and v_+ are finite and $t = -v_-, s = v_+ - v_-$ are possible choices for s and t .

Sampling may be carried out in the following way.

Procedure RU

1. Generate $[0, 1]$ -uniform random numbers U and V and set $X \leftarrow m + (sV - t)U^{-1/k}, Y \leftarrow U^{1+1/k}$.
2. If $Y = U^{1+1/k} \leq \tilde{f}(X) = \tilde{f}(m + (sV - t)U^{-1/k})$ return X as a sample from $f(x)$. Else go to 1.

The special case $m = 0, s = v_+ - v_-, t = -v_-, k = 1$ as introduced by Kinderman and Monahan in 1977 reads as follows.

Procedure KM

1. Generate $[0, 1]$ -uniform random numbers U and V and set $U \leftarrow (\max_x f(x))^{1/2}U, V \leftarrow v_- + (v_+ - v_-)V$.
2. If $U^2 \leq f(V/U)$ return $X \leftarrow V/U$. Else go to 1.

Since the area below $\tilde{f}(x) = f(x)/\max_x f(x)$ is equal to $1/\max_x f(x)$, the expected number of trials is equal to

$$\alpha = 2(v_+ - v_-) \max_x f(x).$$

4. OPTIMAL CONSTANTS FOR THE RATIO OF UNIFORMS METHOD

So far optimal constants m, s, t have not been constructed. Since k is usually equal to 1, one has to minimize the area $s(1 + 1/k)$ of D , i.e. the quantity s .

Assume that $\tilde{f}(x)$ touches $h(x)$ at a point $L < m - t$ and at $R > m - t + s$. This means

$$\begin{aligned} t &= \max \{ (m-x)(\tilde{f}(x))^{1/(1+k)} \mid x \leq m-t \} \\ s-t &= \max \{ (x-m)(\tilde{f}(x))^{1/(1+k)} \mid x \geq m-t+s \} \end{aligned}$$

or

$$t = (m-L)(\tilde{f}(L))^{1/(1+k)} \quad (21)$$

$$s-t = (R-m)(\tilde{f}(R))^{1/(1+k)}. \quad (22)$$

Logarithmic differentiation yields

$$\frac{\tilde{f}'(L)}{\tilde{f}(L)} = \frac{k+1}{m-L}, \quad \frac{\tilde{f}'(R)}{\tilde{f}(R)} = -\frac{k+1}{R-m}. \quad (23)$$

This shows that L and R are functions of m . Finally, adding (21) and (22) leads to

$$s = (m-L)(\tilde{f}(L))^{1/(1+k)} + (R-m)(\tilde{f}(R))^{1/(1+k)}. \quad (24)$$

The derivative of s with respect to m becomes

$$\begin{aligned} \frac{ds}{dm} &= (\tilde{f}(L))^{1/(1+k)} - (\tilde{f}(R))^{1/(1+k)} \\ &+ \frac{\partial}{\partial L} ((m-L)(\tilde{f}(L))^{1/(1+k)}) \frac{dL}{dm} \\ &+ \frac{\partial}{\partial R} ((R-m)(\tilde{f}(R))^{1/(1+k)}) \frac{dR}{dm} \\ &= (\tilde{f}(L))^{1/(1+k)} - (\tilde{f}(R))^{1/(1+k)}, \end{aligned}$$

since the two other expressions are zero by the optimality of L and R . Hence

$$\tilde{f}(L) = \tilde{f}(R), \text{ i.e. } f(L) = f(R) \quad (25)$$

is a necessary condition for an optimum. (23) and (25) will determine L, R and m , and (21) and (24) yield t and s . Usually, the calculation simplifies by noticing that

$$R-L = (k+1) \left(\frac{\tilde{f}(L)}{\tilde{f}'(L)} - \frac{\tilde{f}(R)}{\tilde{f}'(R)} \right) \quad (26)$$

and

$$m = \frac{L\tilde{f}(L) - R\tilde{f}(R)}{\tilde{f}'(L) - \tilde{f}'(R)}; \quad (27)$$

both are consequences of (23).

The gamma distribution is treated as an example. Its density is

$$f(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x}, \quad x \geq 0, \quad a \geq 1.$$

Its logarithmic derivative becomes

$$\frac{\tilde{f}'(x)}{\tilde{f}(x)} = \frac{a-1-x}{x}.$$

Hence (26) yields

$$\begin{aligned} R-L &= (k+1) \left(\frac{L}{a-1-L} - \frac{R}{a-1-R} \right) \\ &= \frac{(k+1)(a-1)(L-R)}{(a-1-L)(a-1-R)} \end{aligned}$$

or

$$(a-1-L)(a-1-R) + (k+1)(a-1) = 0.$$

Setting

$$L = a-1-\lambda, \quad R = a-1+\rho,$$

leads to

$$\lambda\rho = (k+1)(a-1).$$

Now m is calculated from (27) as

$$\begin{aligned} m &= \frac{RL}{a-1} = \frac{(a-1-\lambda)(a-1+\rho)}{a-1} \\ &= a-1+\rho-\lambda - \frac{\lambda\rho}{a-1} \\ &= a-1+\rho-\lambda - (k+1). \end{aligned}$$

Finally, $\rho-\lambda$ is calculated from (25) which means

$$\begin{aligned} 0 &= R-L + (a-1) \ln L/R \\ &= \rho + \lambda + (a-1) \left(\ln \frac{1-\lambda/(a-1)}{1+\rho/(a-1)} \right) \\ &= \frac{1}{2(a-1)} \left((\rho^2 - \lambda^2) - \frac{2(\rho^3 + \lambda^3)}{3(a-1)} + \frac{\rho^4 - \lambda^4}{2(a-1)^2} - \frac{2(\rho^5 + \lambda^5)}{5(a-1)^3} + \dots \right). \end{aligned}$$

Hence

$$\begin{aligned} \rho - \lambda &= \frac{2}{3(a-1)} \frac{\rho^3 + \lambda^3}{\rho + \lambda} - \frac{1}{2(a-1)^2} \frac{\rho^4 - \lambda^4}{\rho + \lambda} + \\ &\quad \frac{2}{5(a-1)^3} \frac{\rho^5 + \lambda^5}{\rho + \lambda} - \dots \\ &= \frac{2}{3(a-1)} (\rho\lambda + (\rho - \lambda)^2) - \\ &\quad \frac{1}{2(a-1)^2} (\rho - \lambda)(2\rho\lambda + (\rho - \lambda)^2) + \\ &\quad \frac{2}{5(a-1)^3} ((\rho\lambda)^2 + 3\rho\lambda(\rho - \lambda)^2 + (\rho - \lambda)^4) \\ &\quad - \dots \end{aligned}$$

A second order approximation yields

$$\rho - \lambda = \frac{2}{3}(k+1) + \frac{4(k+1)^2}{5 \times 27(a-1)}. \quad (28)$$

Hence

$$m = a-1 - \frac{1}{3}(k+1) + \frac{4(k+1)^2}{5 \times 27(a-1)}. \quad (29)$$

Kinderman and Monahan (1980) develop a generator for the gamma distribution with parameters $a \geq 1$, starting with a shifted gamma density with mode 0 in order to obtain a good fit of the table-mountain function. This means $m = a-1$ in our theory, whereas the optimal value is close to $m = a-5/3$. A similar trick is used by Monahan (1987) for generating samples from the χ -distribution with parameters $a \geq 1$.

Recently, the ratio of uniforms was applied to discrete distributions. Ahrens and Dieter (1989) used it for the Poisson distribution and Stadlober (1989a,b) generated binomial and hypergeometric random variates by this procedure.

Other transformations of uniform variables are considered by Barbu (1982, 1984). $X = V/\sqrt{U}$ is applied to Student-t and normal distributions, whereas $X = \frac{U-V}{U}$ is applied to gamma and beta distributions. However, the efficiencies α of the resulting gamma and beta generators are not bounded. Vaduva (1985) generalized these transformation methods to multivariate distributions and applied them to multivariate normal, Student-t and Dirichlet distributions. Stefanescu and Vaduva (1987) investigated multidimensional transformations of the form $(X_1, \dots, X_n) = (V_1/\sqrt[n]{U}, \dots, V_n/\sqrt[n]{U})$. They showed that $k=2$ leads to a better efficiency α than $k=1$ in the cases of the univariate q -exponential distribution and the uni- and multivariate normal and Student-t distributions. However, this improvement does not speed up sampling since the square root is too slow. These examples were the reason to consider exponents $k \neq 1$. Probably, no exponent $k \neq 1$ is of any practical interest.

5. SQUEEZE FUNCTIONS

Step 2 of the acceptance-rejection method in Section 2 can be improved if some simple bound $b(x)$ on $f(x)/h(x)$ is known. This was extensively used by Ahrens and Dieter (1974). According to G. Marsaglia (1977) such expressions $b(x)$ are called 'squeeze functions' or simply *squeezes*. Let

$$b(x) \leq q(x) = f(x)/h(x) \quad \text{for all } x \in \mathbb{R}.$$

The rejection test in Step 2 then changes to

2'. Generate a uniform random variable U between zero and one. If $U \leq b(X)$ accept X as a sample from the target distribution $f(x)$. If $U > q(X) = f(X)/h(X)$ reject X and go back to 1. Otherwise accept X .

The squeeze function $b(x)$ should be easy to calculate and close to $q(x)$. Polynomials in x fulfill this condition and we shall try to determine their coefficients.

Assume that $q(x)$ attains its maximum 1 at the point $x = 0$. ($q(0) = 1$, $q'(0) = 0$). Let $q_k = q^{(k)}(0)$ be the first non-zero derivative of $q(x)$. Then

$$q(x) = 1 + \frac{x^k}{k!}q_k + \frac{x^{k+1}}{(k+1)!}q_{k+1} + \frac{x^{k+2}}{(k+2)!}q_{k+2} + \dots$$

near $x = 0$. We try to find out under what circumstances

$$b(x) = 1 - b_k x^k$$

could serve as a squeeze function. For this, b_k has to satisfy

$$-b_k \leq \frac{1}{k!} \left(q_k + \frac{x}{k+1} q_{k+1} + \frac{x^2}{(k+1)(k+2)} q_{k+2} + \dots \right)$$

or

$$-b_k \leq \frac{1}{k!} \left(q_k - \frac{k+2}{k+1} \frac{q_{k+1}^2}{4q_{k+2}} + \frac{q_{k+2}}{(k+1)(k+2)} \left(x + \frac{q_{k+1}}{2q_{k+2}}(k+2) \right)^2 \right) + \dots$$

Hence

$$q_{k+2} > 0$$

is a necessary condition for a squeeze. If this is not fulfilled, we can still try

$$b(x) = 1 - b_k x^k - b_{k+1} x^{k+1} \quad \text{where} \quad b_k = -\frac{1}{k!} q_k,$$

and the former condition changes to $q_{k+3} > 0$. If this is again not satisfied, one has to carry on. Eventually there may be some m -th power for which

$$q_0 = 1, \quad q_\nu = 0 \quad \text{for} \quad 1 \leq \nu \leq k-1,$$

and

$$q_k \neq 0, \quad q_{k+m+2} > 0 \quad \text{for some} \quad m \geq 0.$$

Now let

$$b(x) = 1 - b_k x^k - \dots - b_{k+m} x^{k+m}$$

where

$$b_{k+\nu} = -\frac{1}{(k+\nu)!} q_{k+\nu} \quad \text{for} \quad 0 \leq \nu \leq m-1$$

and

$$1 - b_k x^k - \dots - b_{k+m} x^{k+m} \leq q(x) \quad \text{for all} \quad x \in \mathbb{R}.$$

This means that

$$(1 - q(x) - b_k x^k - \dots - b_{k+m-1} x^{k+m-1}) x^{-k-m} \leq b_{k+m}.$$

is required. Therefore b_{k+m} has to be equal to

$$\max\{x^{-k-m}(1 - q(x) - b_k x^k - \dots - b_{k+m-1} x^{k+m-1})\}.$$

The maximum b_{k+m} is reached at a point x_b where

$$(k+m)(1 - q(x_b)) + x_b q'(x_b) - \sum_{\nu=0}^{m-1} (m-\nu) b_{k+\nu} x_b^{k+\nu} = 0.$$

If the second derivative is negative at x_b , b_{k+m} is indeed optimal. Therefore one has to check whether the condition

$$(k+m)(k+m+1)x_b^{-k-m-2}(1 - q(x_b)) + 2(k+m)x_b^{-k-m-1}q'(x_b) - q''(x_b)x_b^{-k-m} - \sum_{\nu=0}^{m-1} (\nu-m)(\nu-m-1)b_{k+\nu}x_b^{\nu-m-2} < 0$$

is satisfied. For an example of the general theory see Dieter (1989).

REFERENCES

- Ahrens, J.H. and Dieter, U. (1973), Extension of Forsythe's method for random sampling from the normal distribution. *Mathematics of Computation* 27, 927-937.
- Ahrens, J.H. and Dieter, U. (1974), Computer methods for sampling from gamma, beta, Poisson and binomial distributions. *Computing* 12, 223-246.
- Ahrens, J.H. and Dieter, U. (1980), Sampling from binomial and Poisson distributions: A method with bounded computation times. *Computing* 25, 193-208.
- Ahrens, J.H. and Dieter, U. (1982a), Generating gamma variates by a modified rejection technique. *Communications of the ACM* 25, 47-54.
- Ahrens, J.H. and Dieter, U. (1982b), Computer generation of Poisson deviates from modified normal distributions. *ACM Transactions on Mathematical Software* 8, 163-179.
- Ahrens, J.H. and Dieter, U. (1989), A convenient sampling method with bounded computation times for Poisson distributions. *American Journal of Mathematical and Management Sciences* (to appear).
- Ahrens, J.H. and Kohrt, K.D. (1981), Computer methods for efficient sampling from largely arbitrary statistical distributions. *Computing* 26, 19-31.
- Atkinson, A.C. and Pearce, M.C. (1976), The computer generation of beta, gamma and normal random variables. *Journal of the Royal Statistical Society A*, 139, 431-461.

- Barbu, G. (1982), On computer generation of a random variables by transformations of uniform variables. *Bulletin Mathématique de la Société des Sciences Mathématiques de la République Socialiste de Roumanie* 26, 129–139.
- Barbu, G. (1984), A new fast method for computer generation of gamma and beta random variables by transformations of uniform variables. *Economic Computation and Economic Cybernetics Studies and Research (Bucharest)* 19, 69–81
- Chen, H.C. and Asau, Y. (1974), On generating random variates from an empirical distribution. *AIEE Transactions* 6, 163–166.
- Deák, I. (1981), An economical method for random number generation and a normal generator, *Computing* 27, 113–121.
- Devroye, L. (1986), *Non-uniform random variate generation*. Springer, New York.
- Dieter, U. (1989), Optimal acceptance–rejection methods for sampling from various distributions. *American Journal of Mathematical and Management Sciences* (to appear).
- Forsythe, G.E. (1972), Von Neumann's comparison method for random sampling from the normal and other distributions. *Mathematics of Computation* 26, 817–826.
- Kinderman, A.J. and Monahan, J.F. (1977), Computer generation of random variables using the ratio of uniform deviates. *ACM Transactions on Mathematical Software* 3, 257–260.
- Kinderman, A.J. and Monahan, J.F. (1980), New methods for generating Student's t and gamma variables. *Computing* 25, 369–377.
- Kronmal, R.A. and Peterson, A.V. (1981), A variant of the acceptance–rejection method for computer generation of random variables. *Journal of the American Statistical Association* 76, 446–451. (Corrigendum (1982) *Journal of the American Statistical Association* 77, 954.)
- Kronmal, R.A. and Peterson, A.V. (1984), An acceptance–complement analogue of the mixture–plus–acceptance–rejection method for generating random variables. *ACM Transactions on Mathematical Software* 10, 271–281.
- Knuth, D.E. (1981), *The art of computer programming, Vol. II: seminumerical algorithms*, Addison–Wesley, Reading, Mass.
- Marsaglia, G. (1977), The squeeze method for generating gamma variates. *Journal of Computational and Applied Mathematics* 3, 321–325.
- Monahan, J.F. (1987), An algorithm for generating chi random variables. *ACM Transactions on Mathematical Software* 13, 168–172.
- Neumann, J. v. (1951), Various techniques used in connection with random digits. Monte Carlo methods. *National Bureau of Standards AMS*, 12, 36–38
- Stadlober, E. (1989a), Binomial random variate generation: A method based on ratio of uniforms. *American Journal of Mathematical Management Sciences* (to appear).
- Stadlober, E. (1989b), Sampling from Poisson, binomial and hypergeometric distributions: Ratio of uniforms as a simple and fast alternative. *Mathematisch–Statistische Sektion 303*. Forschungsgesellschaft Joanneum, Graz, Austria.
- Stefanescu, St. and I. Vaduva (1987), On computer generation of random vectors by transformations of uniformly distributed vectors. *Computing* 39, 141–153.
- Vaduva, I. (1985), Computer generation of random vectors based on the transformation of uniformly distributed vectors. *Proceedings of the 7th Conference on Probability Theory, Brasov 1982*, VNU Science Press, Utrecht, 589–598.
- Walker, A.J. (1977), An efficient method for generating discrete random variables with general distributions. *ACM Transactions on Mathematical Software* 3, 253–256.

AUTHOR'S BIOGRAPHY

Ulrich DIETER has been a full professor at the Institute of Statistics, Graz University of Technology, Austria, since 1973. He received his Dr. rer. nat. (Ph.D.) in 1958 and his habilitation in 1965 in Mathematics, both from Kiel University. He spent the year 1966 at Stanford, where he was again in 1974 and 1981/82. From 1967 till spring 1973 he worked at the Institute of Statistics at the Technical University Karlsruhe as an associate professor. His current research interests are random variate generation, both uniform and non-uniform, sequential analysis, theory of gambling and number theory. He is member of AMS, ASA, DMV, GAMM, IMS, ÖMG.

Ulrich Dieter
 Institute of Statistics
 Graz University of Technology
 A-8010 Graz, Lessingstraße 27, Austria
 (0043)316-7061 6475
 statistik@rech.tu-graz.ada.at