Josef Leydold

A Simple Universal Generator for Continuous and Discrete Univariate T-concave Distributions

Working Paper

Original Citation:

This version is available at: http://epub.wu.ac.at/1756/
Available in ePubWU: July 2006

ePubWU, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.
A Simple Universal Generator for Continuous and Discrete Univariate T-concave Distributions

Josef Leydold

Department of Applied Statistics and Data Processing
Wirtschaftsuniversität Wien

Preprint Series
Preprint 30
February 2000

http://statmath.wu-wien.ac.at/
A Simple Universal Generator for Continuous and Discrete Univariate T-concave Distributions

Josef Leydold
University of Economics and Business Administration, Department for Applied Statistics and Data Processing

We use inequalities to design short universal algorithms that can be used to generate random variates from large classes of univariate continuous or discrete distributions (including all log-concave distributions). The expected time is uniformly bounded over all these distributions. The algorithms can be implemented in a few lines of high level language code. In opposition to other black-box algorithms hardly any setup step is required and thus it is superior in the changing parameter case.

Categories and Subject Descriptors: G.3 [Probability and Statistics]: Random number generation
General Terms: Algorithms
Additional Key Words and Phrases: non-uniform random variates, universal method, ratio-of-uniforms method, transformed density rejection, discrete distributions, continuous distributions, log-concave distributions, T-concave distributions

1. INTRODUCTION

In the last decade several approaches have been introduced for so called universal (or black box) methods for generating non-uniform random variates. Recent papers propose methods where a hat function that approximates the respective probability density function or probability vector is constructed (see e.g. Ahrens [1993], Hörmann [1995], Ahrens [1995], Evans and Swartz [1998], Leydold [2000a], Leydold [2000b]; Hörmann and Derflinger [1996], Hörmann and Derflinger [1997]). These methods have (extremely) fast marginal generation time, but require a setup step, which is expensive compared to the average cost of generating one random variate. Although this setup step can be made short at the price of a much higher marginal generation time (e.g. Gilks and Wild 1992) the resulting algorithm are rather complex.

If only a few random variates are required methods like adaptive rejection sampling by Gilks and Wild [1992] or rejection from adjusted table-mountain-shaped hat functions (e.g. Hörmann 1995) have been suggested. However both require a rather expensive setup step and/or adaptation steps. Thus the approach by Devroye [1984] and Devroye [1987] is much more appropriate. It uses inequalities that holds for every log-concave distribution. It is based on the following theorems.

This work was supported by the Austrian Science Foundation (FWF), project no. P12805-MAT
Address: Augasse 26, A-1090 Vienna, Austria, email: Josef.Leydold@statistik.wu-wien.ac.at
Theorem 1 (Devroye [1986], §VII.2.5, Theorem 2.4). If \( f \) is a log-concave density with mode \( \mu = 0 \) and \( f(0) = 1 \), then writing \( q \) for \( F(0) \), where \( F \) denotes the c.d.f. of the distribution, we have

\[
f(x) \leq \begin{cases} 
\min(1, e^{1-x/(1-\phi)}), & (x \geq 0) \\
\min(1, e^{1+x/\phi}), & (x < 0)
\end{cases}
\]  

(1)

The area under the bounding curve in (1) is 2.

Remark 1. If \( F(\mu) \) is not known, a modified universal hat exists with area 4 (see Devroye [1986], §VII.2.3). In both cases these universal hats are not optimal. Devroye [1984] derives the properties of the optimal hat and provides a (rather expensive) generator for the corresponding density. The areas below the optimal bounding curves are \( \pi^2/6 \) and \( \pi^2/3 \), respectively, i.e., about 18% better.

Theorem 2 (Devroye 1987). For any discrete log-concave distribution with a mode at \( \mu \) and probabilities \( p_k \), we have

\[
p_{k+k} \leq p_k \min(1, e^{1-p_k/\phi}), \quad \text{for all } k.
\]

(2)

Remark 2. The expected number of iterations for a generator that utilizes equation (2) is \( 4 + p_k \). Devroye [1987] also gives some hints how this number can be decreased to \( 2 + p_k \) in special cases.

Algorithms that utilize these two theorems can be found in Devroye [1984] and Devroye [1987], respectively.

In this paper we introduce a new approach for universal bounding curves based on the ratio-of-uniforms method. The new algorithm is even simpler and can be applied to a larger class of distributions, including all log-concave distributions. As for Devroye’s algorithm the expected number of uniform random numbers does not depend on the particular distribution. In opposition to other black-box algorithms hardly any setup step is required. Thus it is superior in the changing parameter chase.

2. CONTINUOUS DISTRIBUTIONS

2.1 Ratio-of-uniforms

The ratio-of-uniforms method introduced by Kinderman and Monahan [1977] is a flexible method that can be adjusted to a large variety of distributions. It has become a popular transformation method to generate non-uniform random variates, since it results in exact, efficient, fast and easy to implement algorithms. It is based on the following (slightly modified) theorem.

Theorem 3 (Kinderman and Monahan 1977). Let \( f(x) \) be a positive integrable function with support \( [x_0, x_1] \) not necessarily finite. If \( (V, U) \) is uniformly distributed in

\[
A = \mathcal{A}(f) = \{(v, u) : 0 < u \leq \sqrt{f(v/u + \mu)}, \ x_0 < v/u + \mu < x_1\},
\]

then \( X = V/U + \mu \) has probability density function \( f(x)/\mathbb{f}f \).

For sampling random points uniformly distributed in \( A \), rejection from a convenient enveloping region is used. Kinderman and Monahan [1977] and others use
rejection from the minimal bounding rectangle, i.e., the smallest possible rectangle that contains $A$. It is given by (see Wakefield, Gelfand, and Smith [1991])

$$\mathcal{R}_{mbr} = \{(v, u): v^{-} \leq v \leq v^{+}, 0 \leq u \leq u^{+}\}$$

where

$$u^{+} = \sup_{x_{0} < x < x_{1}} \sqrt{f(x)}$$

$$v^{-} = \inf_{x_{0} < x \leq \mu} (x - \mu) \sqrt{f(x)}$$

$$v^{+} = \sup_{\mu \leq x < x_{1}} (x - \mu) \sqrt{f(x)}$$

2.2 A universal envelope

Assume $A$ is convex. Then it is easy to construct a universal bounding rectangle without computing these boundaries. Let $\mu$ be the mode of $f$, then $u^{+} = \sqrt{f(\mu)}$ and $A$ has the extremal points $(0, u^{+})$, $(v^{-}, u_{1})$ and $(v^{+}, u_{r})$, for respective $u_{1}$ and $u_{r}$. Define $A^{\pm} = \{(v, u) \in A: v > 0\}$ and analogously $A^{-}$. Then by the convexity of $A$, the triangle with vertices at $(0, 0)$, $(0, u^{+})$ and $(v^{+}, u_{r})$ is contained in $A^{+}$ and thus has a smaller area (see figure 1). Consequently

\[\frac{1}{2} u^{+} v^{+} \leq |A^{+}| \quad \text{and} \quad \frac{1}{2} u^{+} (-v^{-}) \leq |A^{-}|\]

where $|A^{+}|$ denotes the area of $A^{+}$. From the proof of theorem 3 (see e.g. Kinderman and Monahan [1977]) it follows immediately that

$$|A^{-}| = \frac{1}{2} \int_{x_{0}}^{\mu} f(x) \, dx = \frac{1}{2} F(\mu) \, ff$$

and

$$|A^{+}| = \frac{1}{2} \int_{\mu}^{x_{1}} f(x) \, dx = \frac{1}{2} (1 - F(\mu)) \, ff.$$
Hence if the cumulative distribution function $F(x)$ at mode $\mu$ is known, we find
\[ v^+ \leq (1 - F(\mu)) \left( \frac{ff}{\sqrt{f(\mu)}} \right) \quad \text{and} \quad v^- \geq -F(\mu) \left( \frac{ff}{\sqrt{f(\mu)}} \right). \]

Therefore we can summarize our result in the following theorem.

**Theorem 4.** Suppose $A = A(f)$ is convex for a density function $f(x)$ with mode $\mu$. Let $F$ denote the c.d.f. of the distribution and let
\[
\mathcal{R} = \{ (v,u): v \leq v \leq u, 0 \leq u \leq u \}, \\
\mathcal{Q} = \{ (v,u): -v \leq v \leq u, 0 \leq u \leq u \}. \tag{6}
\]
where
\[
u_m = \sqrt{f(\mu)} \quad v_m = (\frac{ff}{\sqrt{f(\mu)}}) \quad v = -F(\mu) v_m \quad v = (1 - F(\mu)) v_m \tag{7}
\]
Then $\mathcal{A} \subset \mathcal{R} \subset \mathcal{Q}$ and
\[
|\mathcal{R}| = 2 |\mathcal{A}| \quad \text{and} \quad |\mathcal{Q}| = 4 |\mathcal{A}|. \tag{8}
\]

**Remark 3.** For the class of all distributions with convex sets $\mathcal{A}$, $\mathcal{R}$ is optimal, i.e., any other universal enveloping region must contain $\mathcal{R}$. Analogously for $\mathcal{Q}$ when $F(\mu)$ is not known.

Applying theorem 4 results in the following universal algorithm for distributions with convex set $\mathcal{A}$. It works with any multiple of the probability density function.

**Algorithm SRUC**

**Require:** p.d.f. $f(x)$, area $ff$, mode $\mu$; c.d.f. at mode $F(\mu)$ (optional)

1. \texttt{Setup} /*
2. \texttt{if} \ $F(\mu)$ \texttt{is provided} \texttt{then}
3. \texttt{else}
4. \texttt{repeat}
5. \texttt{Generate} \ $U$ \texttt{uniformly on} $(0, u_m)$.
6. \texttt{Generate} \ $V$ \texttt{uniformly on} $(v_m, v_r)$.
7. \texttt{until} \ $U \leq f(X)$.
8. \texttt{return} \ $X$.

**Remark 4.** By equation (8) the rejection constant of algorithm SRUC is 2 when $F(\mu)$ is known and 4 otherwise.

**Remark 5.** If only upper and lower bounds for $ff$, $F(\mu)$, $f(\mu)$ or $\mu$ are available, an accordingly modified version of algorithm SRUC still works.

2.3 A universal squeeze
When $F(\mu)$ is known we can also construct a universal squeeze.
Theorem 5. Suppose $A(f)$ is convex for a density function $f(x)$ with mode $\mu$. If $F(\mu)$ is given, where $F$ denotes the c.d.f. of the distribution, then there exists a set $S = S(f)$, such that $S \subset A$. We have $(V, U) \in S$ if and only if either

$$0 \leq \frac{V}{U} \leq \frac{v_r}{u_m} \quad \text{and} \quad U v_r + U u_m \leq v_r u_m, \quad (9)$$

or

$$0 \geq \frac{V}{U} \geq \frac{v_l}{u_m} \quad \text{and} \quad U v_l + U u_m \geq v_l u_m, \quad (10)$$

where $u_m, v_l$ and $v_r$ are as defined in theorem 4. Moreover

$$|S| = |A|/2. \quad (11)$$

Proof. Let $S$ denote the universal squeeze region and assume that $A^+ \neq \emptyset$. Let $\Delta$ be the triangle defined by the inequalities $v > 0\, v/u \leq v_r/u_m$ and $u v_r + u u_m \leq v_r u_m$. Hence its vertices are $(0, 0), (0, u_m)$ and $(v_r/2, u_m/2)$ (see figure 1). Define $S^+ = \{(v, u) \in S : v > 0\}$ and $R^+ = \{(v, u) \in R : v > 0\}$. Every straight line through a point $(V, U) \in R^+ \setminus A^+$ that does not intersect $A^+$, splits $R^+$ into two parts such that (i) $A^+$ and the edge $(0, 0)(0, u_m)$ are completely contained in the left hand part, and (ii) the area of left hand part is at least $|A^+|$ and hence cannot be smaller than $|R^+|/2$ (analogously to equation (8)). $S^+$ is then the intersection of the left hand parts of all such lines. Consequently $S^+$ must be contained in the triangles with respective vertices at $(0, 0), (u, u_m)$ and $(v_r, 0)$, $(0, 0), (u, u_m)$ and $(v_r, u_m)$. Since the intersection of these triangles is given by $\Delta$, we find $\Delta \subseteq S^+$. Now notice that $S^+$ is convex. Furthermore for every such straight line that intersects the boundary of $R^+$ in the points $(a, 0)$ and $(b, u_m)$ we must have $(a + b) \geq v_r/2$, since otherwise (ii) would be violated. Hence $(v_r/2, u_m/2) \in S^+$, thus $\Delta \subseteq S^+$ and equation (9) follows. Analogously we find $S^-$ and inequality (10) for the left hand part $R^-$. Obviously $|S| = |S^-| + |S^+| = |R^+|/4 = |A|/2$. \hfill \Box

Remark 6. For the class of all distributions with convex sets $A$, $S$ is optimal, i.e., any other universal squeeze region is contained in $S$.

Algorithm SRUC can be easily extended to make use of theorem 5.

2.4 T-concave distributions

Stadlober [1989b] and Dieter [1989] have clarified the relationship of the ratio-of-uniforms method to the ordinary acceptance/rejection method. It can be viewed as rejection from a table-mountain shaped density (see figure 2). Leydold [2000a] has shown a deeper connection to the so called transformed density rejection method (see Hörmann [1995] for a description of this method). Moreover a full characterization of all distributions with convex region $A$ is derived.

Theorem 6 (Leydold [2000a]). $A(f)$ is convex if and only if $f(x)$ is T-concave with transformation $T(x) = -1/\sqrt{x}$, i.e., if and only if $-1/\sqrt{f(x)}$ is a concave function.

Notice that this class of T-concave distributions includes all log-concave distributions [Hörmann 1995].

We further can use this connection to derive universal upper and lower bounds for T-concave distributions.
\textbf{Theorem 7.} For any \( T \)-concave density \( f \) with \( T(x) = -1/\sqrt{x} \) and mode \( \mu \) let \( u_m, v_m, v_l \) and \( v_r \) be defined as in theorem 4 and let \( x_m = v_m/u_m, x_l = v_l/u_m \) and \( x_r = v_r/u_m \). Define

\[
\hat{h}(x) = \begin{cases} 
  f(\mu) & \text{for } -x_m \leq x - \mu \leq x_m \\
  \frac{\nu^2}{(x-\mu)^2} & \text{otherwise}
\end{cases}
\]

\[
h(x) = \begin{cases} 
  \frac{\nu^2}{(x-\mu)^2} & \text{for } x - \mu < x_l \\
  f(\mu) & \text{for } x_l \leq x - \mu \leq x_r \\
  \frac{\nu^2}{(x-\mu)^2} & \text{for } x - \mu > x_r
\end{cases}
\]

\[
s(x) = \begin{cases} 
  \left( \frac{v_m u_m}{v_l + u_m (x-\mu)} \right)^2 & \text{for } x_l \leq x - \mu < 0 \\
  \left( \frac{v_m u_m}{v_l + u_m (x-\mu)} \right)^2 & \text{for } 0 \leq x - \mu < x_r \\
  0 & \text{otherwise}
\end{cases}
\]

\[
\hat{s}(x) = \begin{cases} 
  f(\mu) / A & \text{for } x_l \leq x - \mu \leq x_r \\
  0 & \text{otherwise}
\end{cases}
\]

Then

\[
\hat{h}(x) \geq h(x) \geq f(x) \geq s(x) \geq \hat{s}(x) \quad \text{for all } x.
\]

\textbf{Proof.} \((v, u) \rightarrow (v/u + \mu, -1/\mu)\) maps \( \mathcal{A}(f) \) one-to-one onto the region \( \mathcal{T}(f) = \{(x, y); y \leq T(f(x)) = -1/\sqrt{T(x)}; x_l < x < x_l\} \) \cite{leydold2008}. Moreover a straight line \( ax + bu = c \) in \( \mathcal{A}(f) \) is mapped onto the line \( a(x - \mu) + cy = -b \) in \( \mathcal{T}(f) \), and consequently to a curve

\[
y = \left( \frac{c}{b + a(x - \mu)} \right)^2
\]

in the original scale, since \( T^{-1}(x) = -1/x^2 \). (If \( c = 0 \), the line \( ax + bu = 0 \) through the origin is mapped into a line parallel to the \( y \)-axis.) We then get upper bounds by the respective boundaries of \( \mathcal{R} \) and \( \mathcal{Q} \) (with \( u > 0 \)) in theorem 4. These are
given by $v = u, u = u_m, v = v_r,$ and $v = -v_m, u = u_m, v = v_m,$ respectively. Thus
the upper bounds $h(x)$ and $\hat{h}(x)$ follow from (17). Analogously we get $s(x)$ by
theorem 5 and eq. (9), (10) and (17). The last inequality follows from the fact,
that $s(x + \mu) = s(x_r + \mu) = (um / 2)^2 \leq s(x)$ for all $x \in [x_l + \mu, x_r + \mu]$.

Remark 7. The proof of theorem 3 uses the fact that $(v, u) \mapsto (v/u + \mu, u^2)$ maps
$A(f)$ one-to-one onto the region $\{(x, y); y \leq f(x), x_0 < x < x_1\}$, and that it has
constant Jacobian 2. An immediate consequence is that
\[
\frac{1}{4} \int \hat{h}(x) \, dx = \frac{1}{2} \int h(x) \, dx = \int f(x) \, dx = 2 \int s(x) \, dx = 4 \int \hat{s}(x) \, dx.
\] (18)
Moreover
\[
\int_{x_l + \mu}^{x_r + \mu} h(x) \, dx = \int_{x_l + \mu}^{\infty} h(x) \, dx.
\] (19)
Completely analogously results hold for the left hand tail of $h(x)$ and for $\hat{h}(x)$.

We can now use eq. (16) to compile a universal generator for $T$-concave distribu-
tions based on the acceptance/rejection technique. Algorithm STDG generates a
random variate with probability proportional to the hat function by inversion.
Squeezes are omitted.

Algorithm STDG

Require: p.d.f. $f(x)$, area $ff$, mode $\mu$; c.d.f. at mode $F(\mu)$ (optional)

/* Setup */
1. $u_m \leftarrow \sqrt{f(\mu)}$, $v_m \leftarrow \sqrt{ff}/u_m$. 
2. if $F(\mu)$ is provided then
3. $A \leftarrow 2 \sqrt{f}$. /* Area below hat */
4. $a_l \leftarrow F(\mu)/A$, $a_r \leftarrow A/2 + a_l$.
5. $u \leftarrow -F(\mu) v_m$, $v \leftarrow v_m + u$.
6. else
7. $A \leftarrow 4 \sqrt{f}$
8. $a_l \leftarrow A/4$, $a_r \leftarrow 3A/4$.
9. $u \leftarrow -v_m$, $v \leftarrow v_m$.
/* Generate */
10. repeat
11. Generate $U$ uniformly on $(0, A)$.
12. if $U < a_l$ then
13. $X \leftarrow -u_l^2 /U + \mu$. /* Compute $X$ by inversion */
14. $Y \leftarrow U^2 /u_l^2$. /* Compute $h(X)$, */
15. else if $U \leq a_r$ then
16. $X \leftarrow v_l /u_m + (U - a_l) /u_m^2 + \mu.
17. $Y \leftarrow f(\mu)$.
18. else /* $U > a_r$ */
19. $X \leftarrow v_r^2 / (u_m v_r - (U - a_r)) + \mu.
20. $Y \leftarrow (A-U)^2 /v_r^2$.
21. Generate $V$ uniformly on $(0, 1)$.
22. until $VY \leq f(X)$.
23. \textbf{return } X.

\textit{Remark 8.} Obviously algorithm \textsc{stdr} is more complex (and slower) than algorithm \textsc{sruc}. However it has two advantages:

1. The rejection constant can be decreased when the domain of density is given by \((x_0, x_1) \subseteq \mathbb{R}\). Just replace \((0, A)\) in step 11 by \((A_1, A_r)\), where \(A_1 = \int_{-\infty}^{x_0} h(x) \, dx\) and \(A_r = \int_{x_1}^{\infty} h(x) \, dx\).

2. \textsc{stdr} does not suffer from the same (possible) defects when using linear congruential generators (with bad lattice structure) as have been reported for the ratio-of-uniform methods [Hörmann 1994a; Hörmann 1994b]. (However, this does not guarantee the absence of other deficiencies.)

2.5 The mirror principle

Devroye [1984] suggests the usage of a hat function for \(f(x) + f(-x)\) when \(F(\mu)\) is not known to reduce the expected number of uniform random numbers. To apply this idea to our situation we need the following result.

\textbf{Lemma 8.} Let \(g_1(x)\) and \(g_2(x)\) be two non-negative functions with respective bounding rectangles \(R_1\) and \(R_2\) for \(A(g_1)\) and \(A(g_2)\) with common left lower vertex \((0, 0)\) and the respective right upper vertices \((v_1, u_1)\) and \((v_2, u_2)\). Then

\[ A(g_1 + g_2) \subseteq \left\{ (v, u) : 0 \leq v \leq \sqrt{v_1^2 + v_2^2}, 0 < u \leq \sqrt{u_1^2 + u_2^2} \right\}. \]

\textbf{Proof.} Let \((v, u) \in A(g_1 + g_2)\). Obviously \(v \geq 0\). By equation (5) \(u^2 \leq \sup(g_1(x) + g_2(x)) \leq \sup g_1(x) + \sup g_2(x) \leq u_1^2 + u_2^2\) and \(v^2 \leq \sup(x - \mu)^2 (g_1(x) + g_2(x)) \leq \sup(x - \mu)^2 g_1(x) + \sup(x - \mu)^2 g_2(x) \leq v_1^2 + v_2^2\) as proposed. \(\square\)

\textbf{Theorem 9.} For any \(T\)-concave density \(f\) with \(T(x) = -1/\sqrt{x}\) and mode \(\mu\) let

\[ \hat{R} = \{(v, u) : -\nu_m \leq v \leq \nu_m, 0 < u \leq \sqrt{2}\nu_m\} \]

where \(\nu_m\) and \(\nu_m\) are as defined in theorem 4. Then

\[ A(f(x) + f(-x)) \subseteq \hat{R} \quad \text{and} \quad |\hat{R}| = 4\sqrt{2} |A(f)|. \]

\textbf{Proof.} Let \(R(p) = \{(v, u) : -p\nu_m \leq v \leq (1-p)\nu_m, 0 < u \leq u_m\}\). Then by theorem 4 we find \(A(f(x)) \subseteq R(F(\mu))\) and \(A(f(-x)) \subseteq R(1 - F(\mu))\). Hence by lemma 8 \(A(f(x) + f(-x)) \subseteq \bigcup_{p \in [0, 1]} \{(v, u) : -p(v) \leq v \leq p(v), 0 < u \leq \sqrt{2}\nu_m\} = \hat{R}\), where \(v(p) = \sqrt{p^2 \nu_m^2 + (1-p)^2 \nu_m^2}\). Equation (21) follows immediately from equation (8). \(\square\)

\textit{Remark 9.} \(\hat{R}\) is not optimal. However the optimal envelope \(R_{\text{opt}}\) is not rectangular and contains the rectangles \(\{(v, u) : -\nu_m/2 \leq v \leq \nu_m/2, 0 < u \leq \sqrt{2}\nu_m\}\) and \(\{(v, u) : -\nu_m \leq v \leq \nu_m, 0 < u \leq \nu_m\}\). Thus we find the estimate \(|R_{\text{opt}}| \geq (1 - \frac{1}{\sqrt{2}}(\sqrt{2} - 1)) \cdot |\hat{R}| \geq 0.79 \cdot |\hat{R}|\).

Using theorem 9 we can compile the following algorithm. It reduces the expected number of uniform random variates at the expense of more evaluations of \(f(x)\).
Algorithm \texttt{SRUCM}

\textbf{Require:} p.d.f. \( f(x) \), area \( Jf \), mode \( \mu \).

\texttt{// Setup \ )
1: \( u_m \leftarrow \sqrt{T(\mu)} \), \( v_m \leftarrow (Jf) / u_m \).

\texttt{// Generator \ )
2: \textbf{loop}
3: Generate \( U \) uniformly on \((0, \sqrt{2} u_m)\).
4: Generate \( V \) uniformly on \((-v_m, v_m)\).
5: \( X \leftarrow V / U \).
6: \textbf{if} \( U^2 \leq f(X + \mu) \) \textbf{then}
7: \textbf{return} \( X + \mu \).
8: \textbf{if} \( U^2 \leq f(X + \mu) + f(-X + \mu) \) \textbf{then}
9: \textbf{return} \(-X + \mu \).

Remark 10. By equation (21) the rejection constant of algorithm \texttt{SRUCM} is \( 2 \sqrt{2} \) in opposition to 4 in algorithm \texttt{SRUC} when \( F(\mu) \) is not known.

3. DISCRETE DISTRIBUTIONS

Stadlober [1989a] has shown that the ratio-of-uniforms method is well suited for generating from discrete distributions. Indeed, considerations from §2 can also be used to design a universal algorithm for discrete distributions. However some modifications are necessary.

A discrete distribution with probability vector \( p_i \), with support \( I \subseteq \mathbb{Z} \), is called \( T\)-concave if

\[ p_i \geq \frac{1}{2} (T(p_{i-1}) + T(p_{i+1})) \quad \text{for all } i \in I \]  \hspace{1cm} (22)

For log-concave distributions we have \( T(x) = \log(x) \) and \( p_i^2 \geq p_{i-1} p_{i+1} \). Obviously \( p_i \) is unimodal. Denote its mode by \( \mu \). For the following assume that \( p_i \) is \( T\)-concave with transformation \( T(x) = -1 / \sqrt{x} \). Let

\[ f_p(x) = \begin{cases} p_{\lfloor x \rfloor} \text{ for } [x] \in I \\ 0 \text{ otherwise} \end{cases} \]  \hspace{1cm} (23)

where \([x]\) denotes the largest integer not greater than \( x \). Since \( f_p \) is a step function, \( \overline{A(f_p)} \) cannot be convex. Consider the convex hull \( C \) of \( A \). Because of inequality (22), \( C \) contains the points \((i + 1 - \mu \sqrt{p_{i-1}}) \sqrt{p_{i+1}} \) for all \( i \geq \mu \), and \((i - \mu) \sqrt{p_{i-1}} \sqrt{p_{i+1}} \) for all \( i \leq \mu \), with \( i \in I \subseteq \mathbb{Z} \) (use transformation \((v, u) \mapsto (v / u + \mu, u^2) \), see remark after the proof of theorem 7). These are the “spikes” of \( A \). Let \((v^+, u_r)\) be the right extremal point of \( A \) (see (5)).

\textbf{Lemma 10.} Let \( \Delta \) be the triangle with vertices at \((0,0)\), \((0, u^+)\) and \((v^+, u_r)\). Then \( |\Delta| \leq |A^+| \), where \( A^+ = \{(v, u) \in A : v > 0 \} \).

\textbf{Proof.} Notice that the edges \((0,0)\)\((0, u^+)\) and \((0,0)\)\((v^+, u_r)\) are always contained in the closure of \( A \). Moreover the third edge \((0, u^+)\)\((v^+, u_r)\) is also contained in \( A \) whenever \( u_r = 0 \) by inequality (22). Then \( \Delta \subseteq A \) and the proposition follows. Now assume \( u_r > 0 \). Edge \((0, u^+)\)\((v^+, u_r)\) is contained in the quadrangle \( Q \) with vertices in \((0, u^+), (u^+, u^+), (v^+, u_r) \) and \((v^+ - u_r, u_r)\). Figure 3 shows the
“worst case” where equality holds in (22) for all $i \in (\mu, \mu + u^+)$. Notice that edge $(0, u^+)(v^+, u_r)$ splits $Q$ into two parts with $|\Delta \cap Q| \leq |Q|/2$, because $u_r \leq u^+$. Moreover $Q$ in figure 3 can be partitioned into quadrangles each with two sides parallel to the $v$-axis. In each of these quadrangles the region that is contained in $\mathcal{A}$ is larger than its complement. Hence $|\mathcal{A} \cap Q| \geq Q/2$ and the proposition follows. □

![Diagram](image)

**Fig. 3.** Quadrangle $Q$ and edge $(0, u^+)(v^+, u_r)$

It obvious that an analogous result holds for $\mathcal{A}^-$ and we arrive at the following proposition. Notice that $\sup_{i<\mu} p_i = p_{\mu-1}$.

**Theorem 11.** Let $p_i, i \in \mathbb{Z}_+$ be a $T$-concave probability vector of a discrete distribution, with $T(x) = -1/\sqrt{x}$ and mode $\mu$. Let $F$ denote the c.d.f. of the distribution and let $\mathcal{R}_d = \mathcal{R}_d^- \cup \mathcal{R}_d^+$ with

$$\mathcal{R}_d^- = \{(v, u): F(\mu) - 1 \sum p_i / \sqrt{p_i - 1} \leq v \leq 0, 0 \leq u \leq \sqrt{p_i - 1} \}$$

$$\mathcal{R}_d^+ = \{(v, u): 0 \leq v \leq 1 - F(\mu), \sum p_i / \sqrt{p_i}, 0 \leq u \leq \sqrt{p_i} \}$$

and $\mathcal{Q}_d = \mathcal{Q}_d^- \cup \mathcal{Q}_d^+ with

$$\mathcal{Q}_d^- = \{(v, u): - \sum p_i / \sqrt{p_i - 1} \leq v \leq 0, 0 \leq u \leq \sqrt{p_i - 1} \}$$

$$\mathcal{Q}_d^+ = \{(v, u): 0 \leq v \leq \sum p_i / \sqrt{p_i}, 0 \leq u \leq \sqrt{p_i} \}$$

Then $\mathcal{A} \subset \mathcal{R}_d \subset \mathcal{Q}_d$ and

$$|\mathcal{R}_d^-| = 2 |\mathcal{A}|$$

and

$$|\mathcal{Q}_d| = 4 |\mathcal{A}|$$

**Remark 11.** We set $\mathcal{R}_d^- = \emptyset$ whenever $p_{\mu-1} = 0$.

**Algorithm SRCUD**

**Require:** $p_i, i \in \mathbb{Z}_+$, $\sum p_i$, mode $\mu$, c.d.f. at mode $F(\mu)$ (optional)

```plaintext
1: u_l \leftarrow \sqrt{p_{\mu-1} - 1}, \quad u_r \leftarrow \sqrt{p_\mu}.
2: if $F(\mu)$ is provided then
```


3: \( u_l \leftarrow -F(\mu) \sum p_i/u_l, \quad v_r \leftarrow (1 - F(\mu - 1)) \sum p_i/u_r. \)
4: **else**
5: \( u_l \leftarrow -\sum p_i/u_l, \quad v_r \leftarrow \sum p_i/u_r. \)
6: / * Generator */
7: **repeat**
8: Generate \( V \) uniformly on \((v_l, v_r)\).
9: **if** \( V < 0 \) **then**
10: Generate \( U \) uniformly on \((0, u_l)\).
11: **else**
12: \( I \leftarrow [V/U] + \mu. \)
13: **until** \( I^2 \leq p_I. \)
14: **return** \( I. \)

**Remark 12.** The rejection constant of algorithm **SRUUD** is 2 when \( F(\mu) \) is known and 4 otherwise.

**Remark 13.** Obviously the mirror principle can also be applied to the discrete case.

4. COMPUTATIONAL EXPERIENCES

We have coded versions of **srouc** (with and without using the universal squeeze), **sroucm** (i.e., using the mirror principle) and of **stdr** (with and without using squeeze \( s(x) \)) where we have restricted the domain of the hat to the domain of the given \( p.d.f. \) It is obvious that the expected number of uniform random numbers for new algorithms is quite high and the marginal generation time is higher than that for specialized algorithms or fast universal algorithms that require a more expensive setup step (e.g., algorithm **arou** in Leydold [2000a], or **utdr** in [Hörmann 1995], that uses a semi-empirical rule to construct a hat function). However, when only a couple of random numbers are requested, the new algorithms are superior both in generation time and the size and complexity of their codes.

To get an idea of the performance we ran several tests. The expected number of uniform random numbers for both **srouc** and **stdr** is 8 when the c.d.f. at the mode is not known and 4 otherwise. When using the mirror principle in the first case, it could be reduced to 5.66. However its usage is only recommended when the generation of uniform random numbers is (very) expensive compared to the evaluation of the p.d.f. As expected **srouc** is a little bit faster than **stdr**. However when the domain of the head function is restricted to the domain of the p.d.f. **stdr** requires less iterations and less uniform random numbers (e.g. 3.26 for the beta(5,7) distribution) and is thus faster.

When comparing the total times (including setup) for generating 10 random variates we found that both algorithm have about the same performance as **utdr** when c.d.f. at the mode is known (and thus are faster, when less random variates are required). If the c.d.f. at the mode is not known, they are at least faster than **arou** or the algorithm by Gilks and Wild [1992].