Wolfgang Hörmann and Gerhard Derflinger

Fast Generation of Order Statistics

Paper

Fast Generation of Order Statistics

Wolfgang Hörmann, Gerhard Derflinger

Department of Applied Statistics and Data Processing
Wirtschaftsuniversität Wien

Preprint Series

Preprint 38
February 2001

http://statmath.wu-wien.ac.at/

Fast Generation of Order Statistics

Wolfgang Hörmann
IE-Department, Bogazici University Istanbul
and
Gerhard Derflinger
Department of Statistics, University of Economics and BA, Vienna

Generating a single order statistic without generating the full sample can be an important task for simulations. If the density and the CDF of the distribution are given it is no problem to compute the density of the order statistic. In the main theorem it is shown that the concavity properties of that density depend directly on the distribution itself. Especially for log-concave distributions all order statistics have log-concave distributions themselves. So recently suggested automatic transformed density rejection algorithms can be used to generate single order statistics. This idea leads to very fast generators. For example for the normal and gamma distribution the suggested new algorithms are between 10 and 60 times faster than the algorithms suggested in the literature.

Categories and Subject Descriptors: G.3 [**Mathematics of Computing**]: Probability and Statistics—*random number generation*

General Terms: Algorithms

Additional Key Words and Phrases: rejection method, transformed density rejection, order statistics, automatic algorithms, T-concave

1. INTRODUCTION

If $X_1, X_2, \ldots, X_n$ are iid random variables, then the order statistics for this sample are $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ where $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. Order statistics are an important notion of statistics and it is of practical importance in many applications to have a simple possibility to sample from order statistics. One simulation problem is the simulation of all order statistics, in other words the generation of an ordered sample. In Chapter V Devroye [1986] gives a detailed presentation of different methods to accomplish this task.

In this paper we restrict our interest to the case that we have to generate independent replications of a single order statistic. Of greatest practical importance are of course the maximum, the median and the minimum, but we will see that we can solve the general generation problem and need not distinguish between special cases.

The most popular method for generating a single order statistic seems to be the inversion method. There the order statistic of the uniform distribution (it is a beta distributed random variate) is generated first. Then this variate is transformed by the inverse of the cumulative distribution function (CDF) to get the order statistic of the desired distribution. It is well known that the inversion of the CDF is not an easy numerical task for many popular statistical distributions. And we get

The first author was supported by the Austrian Academy of Science, APART-scholarship.

additional accuracy problems if we want to generate the maximum of a large sample as the beta variates generated in the first step will all be very close to 1. Devroye [1986] (Chapter XIV.1) calls the inversion algorithm "virtually useless" unless the distribution function is explicitly invertible. But the "quick elimination" algorithm he suggests instead ([Devroye 1980] and [Devroye 1986]) suffers from two main draw-backs: It only works for the maximum or minimum and its execution time is not uniformly bounded but $O(\log(n))$. On the other hand the advantage of this algorithm is the fact that the CDF has to be inverted only once in the set-up. Of course there is always the possibility to generate the full sample to obtain a single order statistic but this is certainly very slow unless n is small.

The new idea of this paper is to generate order statistics by using one of the relatively recent automatic algorithms designed to generate from random variates with given density (eg. [Gilks and Wild 1992], [Hörmann 1995] or [Ahrens 1995]). It is well known that $U_{(i)}$, the i-th order statistic from a uniform sample of size n, has a beta distribution with parameters i and $n-i+1$ and thus the density

$$f_{U_{(i)}}(x) = k\,x^{i-1}(1-x)^{n-i},$$

where k is some normalization constant. For an arbitrary continuous distribution with density f and CDF F we can easily see (by the transformation theorem) that

$$f_{X_{(i)}}(x) = k\,f(x)F(x)^{i-1}(1-F(x))^{n-i}.$$

Of course the evaluation of such a density is time consuming for most distributions but at least it does not include the inversion of the CDF. To get a fast generator we can choose an automatic algorithm where the expected number of evaluations of the density is small.

In Section 2 we give a brief introduction into automatic algorithms. Section 3 contains the mathematics necessary to show that we can use the automatic algorithms to generate from order statistics. In Section 4 we compare the different methods to generate order statistics.

2. AUTOMATIC ALGORITHMS

As stated in the introduction there exist several recent automatic methods to generate from distributions with known density. All of them are based on the well known rejection method (also called acceptance-rejection method). There for a given density $f(x)$ a majorizing function (called **hat-function** $h(x)$) and a minorizing function (called **squeeze** $s(x)$) are constructed.

One simple idea is to use step functions as hat and squeeze. (see [Devroye 1986] Chapter VIII and [Ahrens 1995]). If we know the mode of a unimodal density and the interval where the density is bigger than 0, then we can decompose this interval into N intervals and use a constant as hat-function and a second constant as squeeze-function (lower bound) for the density. The resulting algorithm is simple; the area below the hat (and thus the expected number of iterations in the rejection algorithm) is $1+O(1/N)$ whereas the expected number of evaluations of the density is $O(1/N)$. This means that the sampling procedure will run fast if N is large. Of course before that we have to evaluate the density N times in a set-up step and we store all these values in a table. For all technical details, especially for the choice of the decomposition of the interval we refer to [Ahrens 1995].

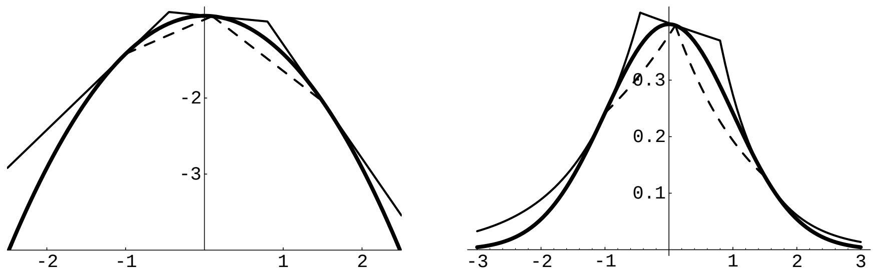

Fig. 1. Construction of a hat function for the normal density using transformed density rejection. The left hand side shows the transformed density with three tangents. The right hand side shows the density function with the resulting hat. Squeezes are drawn as dashed lines.

A second automatic method is based on an idea we call **transformed density rejection** (TDR) (see [Gilks and Wild 1992] and [Hörmann 1995]). It is also using the rejection principle. There the given density f is transformed by a strictly monotonically increasing transformation $T\colon (0,\infty) \to \mathbf{R}$ such that $T(f(x))$ is concave. We then say that f is *T-concave*; log-concave densities are an example with $T(x) = \log(x)$.

By the concavity of $T(f(x))$ it is easy to construct a majorizing function (hat) for the transformed density as the minimum of N tangents. Transforming this function back into the original scale we get a hat function $h(x)$ for the density f. For a fixed point of contact x_i we get

$$h_i(x) = T^{-1}\left(T(f(x_i)) + T(f(x_i))'(x - x_i)\right) \text{ and } h(x) = \min_{1\le i\le N} h_i(x).$$

By using secants between the points of tangency x_i and x_{i+1} of the transformed density we analogously can construct squeezes

$$s_i(x) = T^{-1}\left(T(f(x_i)) + \frac{T(f(x_{i+1})) - T(f(x_i))}{x_{i+1} - x_i}(x - x_i)\right) \text{ and } s(x) = \min_{1\le i\le N} s_i(x).$$

Figure 1 illustrates the situation for the standard normal distribution, $T(x) = \log(x)$ and $N = 3$ points of contact. We continue with a short formal description of the basic idea of transformed density rejection.

Algorithm: TDR
Required: density $f(x)$; transformation $T(x)$, construction points $c_1, \dots, c_n$.
Setup
Construct hat $h(x)$ and squeeze $s(x)$.
Compute intervals $I_1, \dots, I_n$.
Compute areas H_j below the hat for each I_j.
Generator
loop
  Generate I with probability vector proportional to $(H_1, \dots, H_n)$.
  Generate X with density proportional to $h|_I$ (by inversion).

Generate $U \sim U(0,1)$.
if $U\,h(X) \leq s(X)$ **then** /**evaluate squeeze**/
  return X.
if $U\,h(X) \leq s(X)$ **then** /**evaluate density**/
  return X.

It is obvious that the transformation T must have the property that the area below the hat is finite, and that generating a random variable with density proportional to the hat function by inversion must be easy (and fast). Thus we have to choose the transformations T carefully. Hörmann [1995] suggests the family T_c of transformations, where

$$T_0(x) = \log(x) \quad \text{and} \quad T_c(x) = \operatorname{sgn}(c)\ (x^c) \ \text{ for all } x > 0 \text{ (provided } c \neq 0\text{)}.$$

$\operatorname{sgn}(c)$ makes T_c increasing for all c. It is easy to see that for TDR the hat function is piecewise of the form $(a + b\ x)^{1/c}$ for $c \neq 0$ and $\exp(a + b\ x)$ for $c = 0$. This structure of the hat-functions also implies that for densities with unbounded domain we must have $c \in (-1, 0]$. For the choice of c it is important to note that for fixed f the area below the hat increases when c decreases. This can be understood when we compare the hat constructed for $c = 1$ (line segments that touch the density in the points of contact) and $c = 0$ (functions of the form $\exp(a + bx)$ that touch the density in the points of contact). On the other hand for $c = 1$ a T_1-concave density must be concave, for $c = 0$ we have T_0-concave which is simply log-concave. More general it is easy to prove that if f is T_c-concave, then f is $T_{c'}$-concave for every $c' \leq c$.

Because of computational reasons, the choice of $c = -1/2$ (if possible) is suggested. Then TDR can generate random variates of a larger family than the log-concave family, all $T_{-1/2}$-concave distributions. (All distributions of this family are unimodal with subquadratic tails.) TDR works best when the area below the hat and the area below the squeeze are as close as possible. Thus we have to find construction points to make this difference small. For the problem of finding appropriate construction points for the hat function Gilks and Wild [1992] have suggested the ingenious concept of adaptive rejection sampling. For TDR it works in the following way:

Start with (at least) two points on both sides of the mode and sample points x from the hat distribution. Add a new construction point at x whenever the density $f(x)$ has to be evaluated, i.e., when $s(X) < U\,h(X)$, until a certain stopping criterion is fulfilled. (U denotes a uniform random variate between 0 and 1.)

There exist also methods for finding construction points such that the expected number of evaluations of the density is minimized for given number of construction points, transformation and distribution [Derflinger et al. 2001]. A simple consideration gives that even for equally spaced construction points (which are far away from optimal) the area between hat-function and squeeze-function is $O(N^{-2})$ for $c > -1$ [Leydold and Hörmann 1998].

3. LOG-CONCAVE AND T-CONCAVE ORDER STATISTICS

If we want to be sure that we can use TDR algorithms for generating order statistics it is necessary to understand which order statistics have a T-concave distribution.

First we define the local concavity function of an arbitrary two times differentiable function $f(x)$ by

$$lc_f(x) := (f(x)/f'(x))' = 1 - \frac{f(x)f''(x)}{f'(x)^2}.$$

Clearly a density is T_c-concave if and only if $lc_f(x) \geq c$ for all x of the domain. For a fixed f and a fixed point x_0 the local concavity $lc_f(x_0)$ is a constant number. If we set $c = lc_f(x_0)$ we get $T_c(f(x_0))'' = 0$. So we can say that the local concavity of f in x_0 is the maximal real number c that allows that f is T_c-concave in x_0.

We continue with two lemmas necessary to prove the theorems below.

LEMMA 1. *If $f(x)$ is a T_c-concave density with $-1 < c \leq 0$, then the corresponding cumulative distribution function $F(x)$ is $T_{c/(c+1)}$-concave.*

PROOF. It is not difficult to see that for arbitrary f, T_c-concavity can be characterised by:

$$\left(\frac{f(x)}{f'(x)}\right)' \geq c. \tag{1}$$

As T_c-concavity implies unimodality we can assume that f has a single mode which will be denoted by m. At first we prove the lemma for $x \leq m$. Integration of (1) between t and x, $t < x$, gives after multiplication by $f'(t)$, which is positive for $t < m$.

$$f'(t)\frac{f(x)}{f'(x)} - f(t) \geq cf'(t)(x-t).$$

We integrate this over t from the lower bound of the support up to x, using integration by parts for the right-hand side:

$$f(x)\frac{f(x)}{f'(x)} - F(x) \geq cF(x), \tag{2}$$

where $f(u) = 0$ has been assumed in the case of a finite lower bound u. If this is not the case then we consider a sequence of differentiable functions $\phi_k(x)$ with $\phi_k(u) = 0$ which converges towards $f(x)$. Display (2) can be easily transformed into

$$lc_F(x) \geq \frac{c}{c+1}$$

This completes the proof for $x \leq m$.

The case of $x > m$ is easy: From $F''(x) = f'(x) < 0$ it follows that $F(x)$ is T_1-concave, i.e. $F(x)$ is T_c-concave for all $c \leq 1$. $\square$

LEMMA 2. *For all two times differentiable functions f, f_1, f_2 we have:*

$$lc_{f_1 f_2}(x)) = \frac{f_2(x)^2 f_1'(x)^2\, lc_{f_1}(x) + f_1(x)^2 f_2'(x)^2\, lc_{f_2}(x)}{(f_1'(x)f_2(x) + f_1(x)f_2'(x))^2}, \tag{3}$$

$$lc_{f^n}(x) = \frac{lc_f(x)}{n}. \tag{4}$$

PROOF. Both results can be checked by straightforward algebra. □

Now we start with our first main result for log-concave distributions:

THEOREM 3. *For a continuous, log-concave distribution all order statistics have a log-concave distribution.*

PROOF. The theorem can be shown using a result of Prekopa [1973] that states that all marginal distributions of a log-concave distribution are again log-concave. Together with the formula of the multidimensional distribution of order statistics this implies our theorem. Nevertheless we give the following elementary proof.
Lemma 1 implies that $F(x)$ is log-concave. Since $(1 - F(-x))$ is the CDF of $-X$, $1 - F(x)$ is log-concave as well. The logarithm of the density of the i-th order statistic has the form:

$$\log\left(f_{X_{(i)}}(x)\right) = \log k + \log f(x) + (i-1)\log F(x) + (n-i)\log(1-F(x))$$

This is a linear combination (with non-negative coefficients) of concave functions and therefore concave itself. □

The situation becomes much more difficult if we consider the case $c < 0$. Nevertheless we can prove the T-concavity property of the minimum and the maximum:

THEOREM 4. *For a continuous, T_c-concave distribution with $-0.5 \le c \le 0$ the distribution of the maximum (or minimum) of n iid random variates is T_c-concave again.*
For $c < -0.5$ and $c > 0$ the statement does not hold in general.

PROOF. It is enough to consider the maximum as the minimum has obviously the same T_c-concavity properties as the maximum.
The density of the maximum of n variates is $n\,F^{n-1}(x)f(x)$. Using (3) and (4) we get

$$lc_{F^{n-1}f}(x) = \frac{f(x)^4(n-1)lc_F(x) + F(x)^2 f'(x)^2 lc_f(x)}{((n-1)f(x)^2 + F(x)f'(x))^2}$$

As $lc_F(x) = 1 - F(x)f'(x)/f(x)^2$ we can replace $F(x)$ by $(1 - lc_F(x))f(x)^2/f'(x)$ and plug this into the above equation. Cancelling $f(x)^4$ finally results in:

$$lc_{F^{n-1}f}(x) = \frac{(n-1)lc_F(x) + (1 - lc_F(x))^2 lc_f(x)}{(n - lc_F(x))^2}$$

The assumptions for f imply that $lc_f(x) \ge c$ for a fixed c with $-0.5 \le c \le 0$. So we get the simple bound:

$$lc_{F^{n-1}f}(x) \ge \frac{(n-1)lc_F(x) + (1 - lc_F(x))^2 c}{(n - lc_F(x))^2} = b(lc_F(x))$$

We now interpret this bound as a function in $lc_F(x)$ and write $b(lc_F(x))$. Using the result of Lemma 1 it is clear that it is enough to consider values of $lc_F(x) \ge c/(c+1)$. We show in Lemma 5 below that $b(lc_F(x)) \ge c$ for $lc_F(x) \ge c/(c+1)$, $-0.5 \le c \le 0$ and $n \ge 2$. As the case $n = 1$ is trivial this completes the proof.
It is easy to find examples that show that the Theorem is only true for $-0.5 \le c \le 0$. Take eg. $c = -0.501$ and $f(x) = (x+1)^{1/-0.501}$ for $x \ge 0$; or for $c = 1$ $f(x) = 2x$. □

LEMMA 5. *For $x \geq c/(c+1)$, $-0.5 \leq c \leq 0$ and $n \geq 2$ we have:*

$$b(x) := \frac{(n-1)x + (1-x)^2 c}{(n-x)^2} \geq c$$

PROOF. Obviously $b(x)$ has a positive pole for $x = n$. Looking at the first derivative of b

$$b'(x) = \frac{(n-1)(n-2c+x(1+2c))}{(n-x)^3}$$

we can see that the numerator is always positive for $-0.5 \leq c \leq 0$, $x \geq c/(c+1)$ and $n \geq 2$ as:

$$n - 2c + (1+2c)x \geq 2 - 2c + (1+2c)c/(c+1) \geq 2$$

Thus $b'(x) \geq 0$ between $c/(c+1)$ and the pole $x = n$ and $b'(x) \leq 0$ right of the pole of the function. Thus $b(x)$ cannot have a local extremum (with the exception of the pole of course). Together with the values of $b(x)$ for $x = c/(c+1)$ and $x \to \infty$

$$b(c/(c+1)) = c/(n + c(n-1)) \geq c \qquad \text{and} \lim_{x\to\infty} b(x) = c$$

the proof is completed. □

We have invested a lot of time to find the proof of the above theorem for all order statistics. Unfortunately the structure of the density of the order statistic $k\, f(x) F(x)^{n_1} (1-F(x))^{n_2}$ makes a proof similar to the one above impossible. Nevertheless due to the considerations done during our fruitless attempts to prove the result and due to our extensive numerical experimentation we are convinced that the following assertion is true:

For any T_c-concave distribution with $-0.5 \leq c \leq 0$ all order-statistics have a T_c-concave distribution.

4. COMPARISON OF METHODS

We start with shortly describing possible algorithms for generating the r-th order statistic of a sample of size n from a continuous "original distribution". We also mention what is necessary for using these algorithms: There are the following three known methods:

Naive method: Generate the full sample and order it to find the required order statistic. Thus a generator for the original distribution and sorting or a faster algorithm for finding the r-th order statistic is needed. Finding maximum or minimum reduces to n comparisons.

Inversion: Generate the corresponding uniform order statistic and transform it with the inverse CDF. A generator for the beta distribution and an algorithm to invert the CDF of the original distribution is needed. We used TDR to generate from the beta distribution and numerical inversion using Newton's method and a table of size 1000.

Quick elimination (**QE**:) ([Devroye 1980] and [Devroye 1986]) It works only for maxima or minima. An algorithm to sample from the given distribution, restricted to a half-open interval, is necessary. (We used transformed density rejection algorithms to accomplish this task).

Using Theorem 3 we know that all order statistics of a log-concave distribution have a log-concave distribution themselves. Theorem 4 states that the maximum and minimum of $T_{-1/2}$-concave distributions is $T_{-1/2}$-concave again. So we can use transformed density rejection to generate order statistics from log-concave distributions, and maxima and minima of T-concave distributions. As most of the important standard distributions (like the gamma, beta and normal distribution) are log-concave this algorithm (called **OSTDR** here) should be useful for many applications. Required is the possibility to evaluate the density and the CDF of the original distribution. For our implementation of TDR (see [Leydold et al. 2001]) the only necessary additional knowledge is a starting point within the domain of the order statistic. If n is not too large and we know (as it is typically the case) the moments and the mode of the original distribution it is possible to use simple heuristics to find this point: use eg. the mode of the original distribution plus two standard deviations as starting point if you want to generate the maximum. We had no problems in our experiments using such heuristics for n up to 1000. If n is larger the standard deviation of the order statistic becomes very small and the starting points could lie in regions where the density of the order statistic is too close to zero. If this happens we can compute the expectation of the corresponding uniform order statistic, which is $(r-1)/(n-1)$ and use a point close to $F^{-1}((r-1)/(n-1))$ as starting point. This means that we have to do one approximate numerical inversion in the set-up.

A second possibility is to use step functions as hat and squeeze for the density of the required order statistic. We call this algorithm (also not suggested in the literature) **OSSTEP**. The density and the CDF of the original distribution and the mode of the order statistic are required. The algorithm is restricted to densities with bounded domain. If we want to use it for densities with unbounded domain we have to know "save" cut off points for the order statistic such that the probability for the order statistic to lie outside of these points is computational negligible. To find such points we can use the simple fact that for any order statistic the probability to be bigger (or smaller) than a certain value is bounded by n times the probability of the original distribution.

To compare the characteristics of these five algorithms we first look at the time complexity. Clearly the execution time of the Naive method grows at least linearly with n and for QE we have $O(\log n)$ whereas the other three algorithms have uniformly bounded execution times with respect to n. Typically for most standard distributions by far the most time consuming operation is the evaluation of the CDF F of the original distribution. For the inversion algorithm, we are using a numerical root finding method (like Newton's algorithm or regula falsi). This means that we have to evaluate F several times to generate a single random variate from the order statistic distribution. For Algorithms OSTDR and OSSTEP the number of necessary evaluations of the density of the order statistic (which includes an evaluation of F) strongly depends on the number of construction points that are used. We use the fraction ρ of the area below the hat divided through the area below the squeeze to describe the characteristic of the used algorithm. The expected number of evaluations of the density (and therefore of F) is equal to $\rho - 1$. As we can specify the required ρ in the set-up this means that with Algorithms OSTDR and

		OSTDR	OSSTEP	NINV	QE	NAIVE
		Normal distribution				
n	r	ρ=0.995	ρ=0.99			
20	10	1.0(7)	1.1(58)	10.6(12)		
20	20	1.0(7)	1.1(58)	19.2(12)	7.4(1)	20(1)
100	50	1.0(7)	1.1(58)	10.7(12)		
100	100	1.0(7)	1.1(58)	22.8(12)	9.9(1)	100(1)
1000	500	1.0(7)	1.1(58)	10.8(12)		
1000	1000	1.0(7)	1.1(58)	32.7(12)	13.5(1)	1000(1)
		Gamma(10) distribution				
		ρ=0.9975	ρ=0.995			
20	10	1.0(26)	1.2(260)	39.5(73)		
20	20	1.0(26)	1.2(260)	51.7(73)	7.5(2)	20(1)
100	50	1.0(26)	1.2(260)	40.7(73)		
100	100	1.0(26)	1.2(260)	62.6(73)	10.2(2)	100(1)
1000	500	1.0(26)	1.2(260)	41.3(74)		
1000	1000	1.0(26)	1.2(260)	66.0(74)	13.6(2)	1000(1)

Table 1. Marginal execution times in μ-seconds and set-up times in milli-seconds (in brackets)

OSSTEP we can – at the expense of a longer set-up – reduce the expected number of evaluations of F to values close to 0. Here is an important difference between OSTDR and OSSTEP. Writing N for the total number of design points, we have $\rho - 1 = O(1/N^2)$ for OSTDR and $\rho - 1 = O(1/N)$ for OSSTEP. This means that for the same value of ρ we expect a much longer set-up and much larger tables for OSSTEP than for OSTDR.

Using the facilities of our UNURAN-library [Leydold et al. 2001] and the five methods described above, we generated order statistics of the normal and the gamma distribution. We experimented to find a value of ρ such that the generation of 10^6 variates including set-up is as fast as possible. We know that the timing results are strongly influenced by hardware, compiler, uniform generator distribution and so on. Nevertheless we report some of our timing results in Table 1. We can clearly see that for the two new algorithms the marginal execution time is not influenced by n and r or by the distribution. Only the set-up is slower for the gamma distribution, as the evaluation of the CDF is much slower than for the normal distribution. We can see that the two new algorithms are between ten and 60 times faster than numerical inversion. Compared with the quick elimination algorithm QE the factor depends on the sample size and is for our examples around ten. Note that QE only works for maxima and minima and also depends on a fast method to generate from the truncated original distribution.

If we compare the marginal execution times of the two new algorithms we can see that OSTDR is only about ten percent faster than OSSTEP; but the necessary number of design points is much larger for OSSTEP which results in a very slow set-up. This problem is even increasing if we consider heavy tailed distributions.

5. CONCLUSION

We have presented the necessary mathematics to show that we can use our recently developed universal algorithms to generate a single order statistic. Our computer experiments show that – depending on the numerical difficulties associated with the CDF of the desired distribution – the new method is between 10 and 60 times faster than the methods proposed in the literature.

REFERENCES

AHRENS, J. H. 1995. A one-table method for sampling from continuous and discrete distributions. *Computing 54*, 2, 127–146.

DERFLINGER, G., HÖRMANN, W., AND TIRLER, G. 2001. On the optimal selection of design points for transformed density rejection. Technical report, Institut für Statistik, WU Wien, Augasse 2–6, A-1090 Wien, EU.

DEVROYE, L. 1980. Generating the maximum of independent identically distributed random variables. *Comput. Math. Appl. 6*, 305–315.

DEVROYE, L. 1986. *Non-Uniform Random Variate Generation*. Springer-Verlag, New-York.

GILKS, W. R. AND WILD, P. 1992. Adaptive rejection sampling for Gibbs sampling. *Applied Statistics 41*, 2, 337–348.

HÖRMANN, W. 1995. A rejection technique for sampling from T-concave distributions. *ACM Trans. Math. Software 21*, 2, 182–193.

LEYDOLD, J. AND HÖRMANN, W. 1998. A sweep-plane algorithm for generating random tuples in simple polytopes. *Mathematics of Computation 67*, 224, 1617–1635.

LEYDOLD, J., HÖRMANN, W., JANKA, E., AND TIRLER, G. 2001. *UNU.RAN User Manual*. Augasse 2–6, A-1090 Wien, EU: Institut für Statistik, WU Wien. http://statistik.wu-wien.ac.at/unuran.

PREKOPA, A. 1973. On logarithmic concave measures and functions. *Acta Scientiarium Mathematicarum Hungarica 34*, 335–343.