The automatic generation of one- and multi-dimensional distributions with transformed density rejection

Josef Leydold and Wolfgang Hörmann
University of Economics and Business Administration
Department for Applied Statistics and Data Processing
Augasse 2-6, A-1090 Vienna, Austria
e-mail: Josef.Leydold@wu-wien.ac.at

Keywords: Random number generation, rejection method

ABSTRACT

A rejection algorithm, called "transformed density rejection", is presented. It uses a new method for constructing simple hat functions for a unimodal density f. It is based on the idea of transforming f with a suitable transformation T such that $T(f(x))$ is concave. The hat function is then constructed by taking the pointwise minimum of tangents which are transformed back to the original scale. The resulting algorithm works very well for a large class of distributions and is fast. The method is also extended to the two- and multidimensional case.

INTRODUCTION

Different automatic (also called universal or black-box) methods have been suggested to sample from univariate log-concave distributions. The main advantage of these methods is that one algorithm coded and tested only once can do the same job as a library of algorithms designed for standard distributions. In this paper we want to report that this tools are not only available for one dimensional distributions, but can be extended to the multidimensional case.

ONE-DIMENSIONAL DISTRIBUTIONS

To design a universal algorithm utilizing the rejection method it is necessary to find an automatic way to construct a hat function for a given density. Transformed density rejection introduced under a different name in [2] and generalized in [3] is based on the idea that the density f is transformed by a monotone T in such a way that $g(x) = T(f(x))$ is concave. Then it is simple to construct a hat $l(x)$ for $g(x)$ as the pointwise minimum of N tangents on $g(x)$ touching in N points. Figure 1 illustrates the idea for the normal distribution with $N = 3$ and $T(x) = \log(x)$. The hat and the density in the transformed scale are shown on the left hand side, the original scale on the right hand side. As g is concave it is clear that we have $g(x) \leq l(x)$ for all x. Transforming $l(x)$ back into the original scale we get $h(x) = T^{-1}(l(x))$ with $f(x) \leq h(x)$ as majorizing function or hat for f. To generate the random variates is then done in a standard way: Decomposition is used to decide between the different regions of the hat and then ordinary rejection is applied, using inversion to sample from the dominating distribution.

To design automatic random variate generators based on this idea we have some freedom in choosing the transformation T, the number of design points N and the location of these design points.

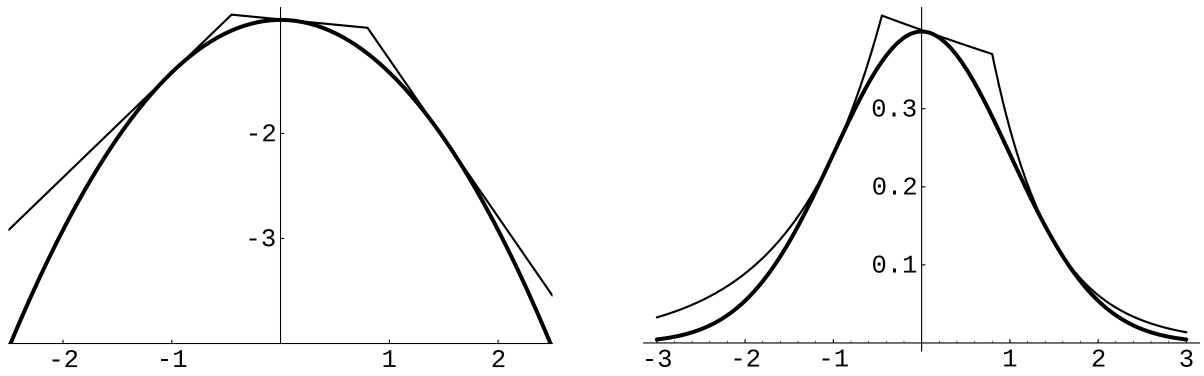

Figure 1: *hat function for univariate normal density*

It is necessary that the transformation T is differentiable and that $T'(x) > 0$. In addition it is of practical importance that $T^{-1}(x)$ is simple and that $F(x) = \int T^{-1}(x)\ dx$ is simple and easy to invert. For distributions with unbounded domain it is also necessary that $\lim_{x \to -\infty}$ is bounded. $T(x) = \log(x)$ (the case considered in [2]) and $T(x) = -(x^c)$ with $-1 < c < 0$ (suggested in [3]) fulfill these conditions. The hats constructed with these transformations are piecewise exponential or piecewise of the form $(const{+}x)^{1/c}$. The algorithm can only be used if $T(f(x))$ is concave. For $T(x) = \log(x)$ this is the case for arbitrary log-concave densities, for $T(x) = -(x^{1/2}) = -1/\sqrt{x}$ we get a generalisation of log-concave called T-concave in [3].

An automatic algorithm of practical importance can be designed when transformed density rejection is used with three points of contact, the mode and one point to the left and to the right of it. The constructed hat has table-mountain shape and it is possible to give easy conditions to minimize the area below the hat.

Of course a much better fit of the hat can be obtained if the number of design points is increased. It is possible to choose these points of contact close to optimal with some afford or to use a stochastic method. This means that after starting with two points an additional point of contact is added if a random variate has to be rejected during generation till the maximum number of design points N is reached. It is not difficult to prove that the area below the hat constructed in this way is $1 + O(1/N^2)$. Experience shows that the area below the hat is not strongly affected by the starting points and is not so far from the smallest obtainable area.

Compared with automatic methods suggested in the literature the algorithms based on transformed density rejection are quite fast and comparatively simple.

MULTIDIMENSIONAL DISTRIBUTIONS

Even more important is the fact that the idea can be generalized to higher dimensions as well as there are no universal algorithms for multivariate distributions available in the literature. In the important book [1] it is even stressed that no general inequalities for multivariate log-concave densities are available, a fact which makes the design of black-box algorithms, similar to those developed in [1] for the univariate case, impossible.

The idea of transformed density rejection remains exactly the same as outlined above. Again the algorithm is only applicable to T-concave distributions, i.e. distributions, where $T(f(x))$ is concave. But now instead of tangents we have to use tangent plains. Figure 2 shows the situation for the binormal distribution with the four points of contact $(\pm 1, \pm 1)$ (the transformed scale is shown above and the original scale below). Of course the realization of the algorithm becomes much more complicated

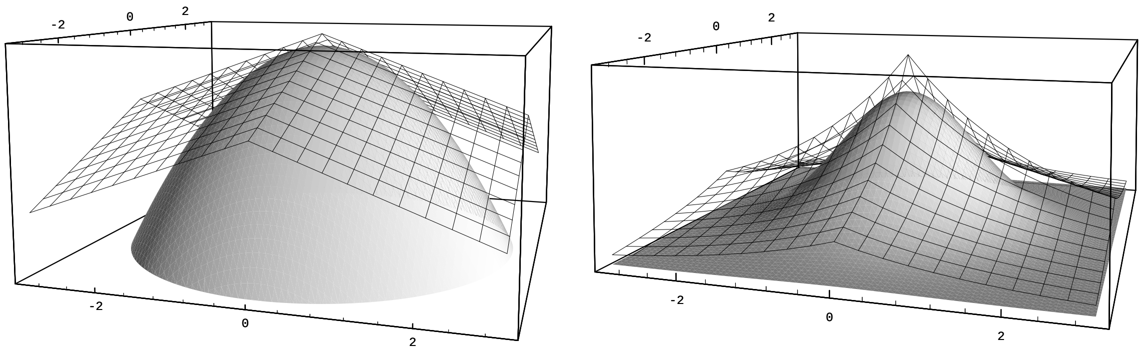

Figure 2: *hat function for bivariate normal density*

since the piecewise definition of the hat is no longer on intervals but on polytopes of dimension n and the hat function is exponential (or like $1/(const + x)^{1/c}$) in direction of the gradient and constant on hyperspaces orthogonal to the gradient. We have to compute the marginal density along this gradient of the hat and generate a univariate random number to get such an orthogonal hyperspaces. Unluckily the marginal looks like $A(t) \cdot T^{-1}(\alpha + \beta\, t)$ where $A(t)$ is piecewise a polynomial which gives the volume of the intersection of the orthogonal hyperspace with the polytope. For dimension n the degree of this polynomial is $n - 1$. Thus we have to use a black box algorithm for log-concave distributions. The last step is to get a random point uniformly distributed on this cut polytope.

Because of the complexity of the method, it seems to be impossible to find a simple choice with reasonable acceptance probability for small N. Therefore the stochastic choice of taking the rejected points as additional design points after starting with some arbitrary points seems to be the only possibility. But it is not as easy as in one dimension to get starting points which produce a hat with bounded volume below it. The volume below the hat is $1 + O(N^{-2/n})$ which shows that for higher dimensions we have to expect a bad rejection constant.

This idea is used to design a universal generator for log-concave distributions in [4]. The resulting algorithm is tested with bivariate normal and beta distributions and works well.

The use of the same method in practice for higher dimensions is more difficult. The algorithm still works (see [5]) but its complexity depends on the number of vertices of the polytopes, which increases rapidly with dimension n. This results in a very slow algorithm. Moreover it is sensitive for round-off errors, so that the algorithm works for dimensions at most 5. A way to avoid these troubles is to construct at first the polytopes as simple as possible, i.e. simple cones, and then construct the hat function on each of these domains. Thus the hat function is not continuous any more (which increases the rejection constant), but the complexity of the algorithm becomes small. Hence the R^n is split into such simple cones with the vertices at the mode of the distribution. The points of contact of the tangent hyperplanes are located on the center lines of these cones. In [6] an algorithm for log-concave distributions is presented. Computational experiments with normal distributions show, that the time for the generation of random tuples with respect to the head function is linear in n. Round-off errors does not cause a considerable loss of significance. Of course rejection constant and the necessary number of cones are big when the dimension is high. But the algorithm still works for the normal distribution in dimension 10.

REFERENCES

[1] L. Devroye. *Non-uniform Random Variate Generation*. Springer, New York, 1986.

[2] W. R. Gilks and P. Wild. *Adaptive rejection sampling for Gibbs sampling.* Applied Statistics, 41:337–348, 1992.

[3] W. Hörmann. *A rejection technique for sampling from T-concave distributions.* ACM Transactions on Mathematical Software, 21:182–193, 1995.

[4] W. Hörmann. *A universal generator for bivariate log-concave distributions.* Preprint, 1996.

[5] J. Leydold and W. Hörmann. *A sweep plane algorithm for generating random tuples.* Preprint, 1997.

[6] J. Leydold. *A rejection technique for sampling from log-concave multivariate distributions.* Preprint, 1997.