RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

DECEMBER 1996

SAMPLING NONSTANDARD DISTRIBUTIONS VIA THE GIBBS SAMPLER

WORKING PAPER #9612-34

PAUL DAMIEN
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

AND

STEPHEN WALKER
IMPERIAL COLLEGE, 180 QUEEN'S GATE, LONDON

Sampling Nonstandard distributions via the Gibbs Sampler

Paul Damien[1]
and
Stephen Walker[2]

[1] Department of Statistics and Management Science, School of Business, University of Michigan, Ann Arbor, 48109-1234, USA.

[2] Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ.

Summary

Simple and fast algorithms to sample nonstandard distributions as appearing in Devroye (1986) are developed. They replace the need for rejection based approaches and are particularly appropriate for use in a Gibbs sampling context.

Key Words: Latent variables, Gibbs sampler, Uniform distribution.

1 Introduction

Let f be a continuous density function defined on the real line. The solution of generating a random variate X from f, starting with the assumption that it is possible to sample uniform random variables from the interval $(0,1)$ to provide an iid sequence of such variables, was developed in Damien and Walker (1996). To motivate their result stated as a theorem later, let $f(x) \propto \exp(-x\beta)I(a < x < b)$, where $\beta > 0$, $0 \leq a < b \leq \infty$ and I represents the Indicator function: write such a density as $e(\beta, a, b)$, let $f(x) \propto (1-x)^{\beta-1}$, where $\beta > 0$ and $0 \leq a < b \leq 1$: write such a density as $\beta(\beta, a, b)$. It is obvious that these densities can be sampled via the inverse transform method. Finally, let $U(a,b)$ denote the uniform distribution on (a,b). Now, the basic idea is to introduce a latent variable Y, construct the joint density of Y and X with marginal density for X given by f, and to use the Gibbs sampler (Smith and Roberts, 1993) to generate random variates from f. In particular, Damien and Walker (1996) show that all the conditional distributions in a such a Gibbs sampler will be on of the three distributions, $U(a,b)$, $\beta(\beta, a, b)$, $e(\beta, a, b)$. To this end they prove:

Theorem (Damien and Walker, 1996). If

$$f(x) \propto \prod_{l=1}^{L} g_l(x),$$

where the g_l are nonnegative invertible functions (not necessarily) densities — that is, if $g_l(x) > y$ then it is possible to obtain the set $A_l(y) = \{x : g_l(x) > y\}$ then it is possible to implement a Gibbs sampler for generating variates from f in which all but one conditional distributions will uniform distributions.

Damien and Walker (1996) exemplify the use of the Theorem by sampling all univariate continuous densities that appear in Johnson and Kotz. Also they illustrate the method for several Bayesian nonconjugate models and nonparametric models. Cumbus, Damien and Walker (1996) illustrate the use of the Theorem to simulate discrete and truncated multivariate densities, and Walker and Damien (1996) use the method to sample from a mixture of Dirichlet process model.

We consider four of the nonstandard distributions considered in chapter 9 of Devroye (1986) which can not be sampled directly. The Zipf distribution is used in linguistics and social sciences, while the Planck distribution is encountered in the physical sciences. It is known that provided one can efficiently simulate a Zipf random variate, it is straightforward to generate a Planck variate as well. Two other well known distributions are the generalised inverse Gaussian and Pearson IV distributions.

In this paper new algorithms to sample these distributions are developed. These algorithms are an alternative to the rejection algorithms which appear in Devroye (1986), and is based on the Theorem given earlier.

2 The Zipf distribution

The Riemann zeta function is given by

$$\zeta(a) = \sum_{i=1}^{\infty} \frac{1}{i^a},$$

where $a = s + jt$, s, t are real numbers, and j is the square root of minus one. Simple expressions for the zeta function are known in *special* cases. For example, if a is an *integer* then

$$\zeta(2a) = \frac{2^{2a-1}\pi^{2a}}{(2a)!} B_a,$$

where B_a is the ath Bernoulli number (Titchmarsh, 1951, p20).

The Zipf distribution has one parameter $a > 1$, and is defined by the probabilities

$$p_i = \frac{1}{\zeta(a) i^a}, \quad (i \geq 1).$$

Following the Theorem, the Gibbs sampler is based on the joint distribution given up to proportionality by

$$[i, u] \propto I(0 < i^a u \leq 1).$$

Then the full conditionals are given by

$$[u|i] = U(0, i^{-a}).$$

and

$$[j|u] = U(\{1,2,\dots,m_u\}),$$

where $m_u = \text{int}[u^{-1/a}]$, that is, $\Pr(j = j|u) = 1/m_u$ for each $j \in \{1,\dots,m_u\}$. We can construct the Markov chain $X_1, X_2, \dots$ such that $X_n \to_d x$ as $n \to \infty$ via

$$[X_{n+1}|X_n] = U(\{1,2,\dots,m_n\}),$$

where m_n is random and given by $\text{int}[X_n \tau_n^{-1/a}]$ where the τ_n are iid $U(0,1)$ and independent of the X_n. Here $\text{int}[x]$ denotes the largest integer less than or equal to x.

We ran the above algorithm for a number of values for a and in each case collected 20,000 random variates. Computing time for each run was approximately 10 seconds and the estimates for the reciprocal of the zeta function at $a = 2, 4$ and 6 are, respectively, 0.609, 0.923 and 0.981. A Monte Carlo approximation to the zeta function is given by

$$[\lim_{n \to \infty} n_1/n]^{-1},$$

where

$$n_1 = \sum_{k=1}^{n} I(X_k = 1),$$

and $X_1, X_2, \dots$ are the samples obtained from the Zipf distribution. The exact values, are $6/\pi^2$, $90/\pi^4$ and $945/\pi^6$ which are in good agreement with our estimates.

3 The Planck distribution

The Planck is a two parameter distribution with density given by

$$f(x) = \frac{b^{a+1}}{\Gamma(a+1)\zeta(a+1)} \frac{x^a}{e^{bx} - 1},$$

where $x > 0$, $a > 0$ is a shape parameter, $b > 0$ is a scale parameter. Our approach using the Theorem involves consideration of the joint density given up to proportionality by

$$f(x, u, v) \propto I(0 < x, u < x^a, v(\exp(bx) - 1) < 1).$$

where $a, b > 0$. Clearly the marginal distribution of x is as required. The full conditional distributions are all uniform and given by

$$f(u|x, v) = U(0, x^a),$$

$$f(v|u, x) = U(0, (\exp(bx) - 1)^{-1}),$$

and

$$f(x|v, u) = U(u^{1/a}, b^{-1} \log(1 - 1/v)).$$

We ran this algorithm with $a = 2$ and $b = 0.5$ and collected 10000 samples from the Planck distribution. Computing time was 13 seconds and a histogram representation of the sample is presented in Figure 1.

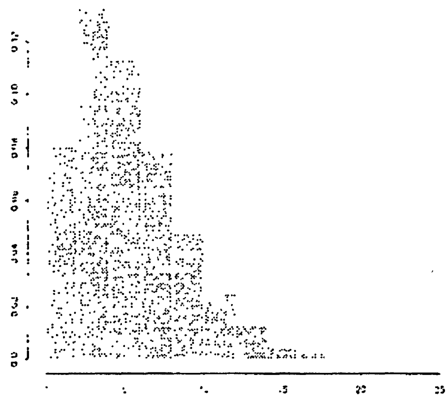

Figure 1: Histogram representation of samples obtained from the Planck distribution with parameters $a = 2$ and $b = 0.5$.

4 The generalised inverse Gaussian distribution

Our algorithm is particularly appropriate for the generalised inverse Gaussian (GIG) distribution. According to Devroye there are two difficulties associated with a rejection based algorithm to sample the GIG distribution: firstly, it is required to calculate the modified Bessel function of the third kind: and, secondly, the expected number of iterations in the rejection algorithm is large.

The GIG distribution has three parameters with density function given up to a constant of proportionality by

$$f(x) \propto x^{\lambda-1}\exp(-0.5(a/x + bx)),$$

where $x > 0$, $\lambda \in (-\infty, +\infty)$, and $a, b > 0$. Note that the gamma and inverse Gaussian distributions are a special case of the GIG. We define the joint density $f(x, u, v, w)$ by

$$f(x, u, v, w) \propto I\left(u < x^{\lambda-1}, v < \exp(-0.5a/x), w < \exp(-0.5bx)\right).$$

The full conditionals are given by

$$f(u|v, w, x) = U(0, x^{\lambda-1}),$$

$$f(v|w, x, u) = U(0, \exp(-0.5a/x)),$$

$$f(w|x, u, v) = U(0, \exp(-0.5bx)),$$

and

$$f(x|u, v, w) = \begin{cases} U\left(\max\left\{-a/(2\log v), u^{1/(\lambda-1)}\right\}, -2/b\log w\right) & \text{if } \lambda > 1 \\ U(-a/(2\log v), -2/b\log w) & \text{if } \lambda = 1 \\ U\left(-a/(2\log v), \min\left\{1/u^{1/(1-\lambda)}, -2/b\log w\right\}\right) & \text{if } \lambda < 1. \end{cases}$$

5 The Pearson IV distribution

The Pearson family of distributions has 12 members and all but one can be sampled directly using standard distributions. The exception is the Pearson IV distribution given up to a constant of proportionality by

$$f(x) \propto \left(1 + (x/a)^2\right)^{-b} \exp(-c\arctan(x/a))$$

where x is defined on $(-\infty, +\infty)$, $a > 0$, $b > 1/2$ and c is a real. We define the joint density

$$f(x, u, v) \propto I\left(u < (1 + (x/a)^2)^{-b}, v < \exp(-c\arctan(x/a))\right).$$

The full conditonals are given by

$$f(u|v,x) = U\left(0,(1+(x/a)^2)^{-k}\right),$$

$$f(v|x,u) = U(0,\exp(-c\arctan(x/a))),$$

and

$$f(x|u,v) = U\left(\max\left\{-a\sqrt{1/u^{1/k}-1},\, a\tan(-1/c\log v)\right\},\, a\sqrt{1/u^{1/k}-1}\right),$$

where we have assumed without loss of generality that $c > 0$.

6 Conclusions

In this paper, we have provided a new and simple way of generating random variates from the Zipf, Planck, GIG, and Pearson IV distributions. The method described requires no more than a Gibbs sampler involving uniform full conditional distributions, obtained after the introduction of latent variables.

Acknowledgements

This work was supported in part by an EPSRC ROPA grant. Support for travel was provided in part by the United Kingdom Engineering and Research Council and The University of Michigan. This paper was completed when the first author was visiting Imperial College, London.

References

Cumbus, C., Damien, P. and Walker, S.G. (1996) Sampling truncated Poisson and multivariate normal densities via the Gibbs sampler. *Submitted for publication.*

Damien, P. and Walker, S.G. (1996) Sampling probability densities via uniform random variables and a Gibbs sampler. *Submitted for publication.*

Devroye, L. (1986) *Non-Uniform Random Variate Generation.* Springer-Verlag, New York.

Smith, A.F.M., and Roberts, G. (1993) Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods (with discussion). *Journal of the Royal Statistical Society, Series B*, 55, 3-24.

Titchmarsh, E.C. (1951) *The Theory of the Riemann Zeta Function*. Oxford University Press, Oxford, 1951.

Walker, S.G. and Damien, P. (1996) Sampling a Dirichlet process mixture model. em Submitted for publication.