SIMULATING THETA RANDOM VARIATES

Luc Devroye
School of Computer Science
McGill University

ABSTRACT. We develop an exact simple random variate generator for the theta distribution, which occurs as the limit distribution of the height of nearly all models of uniform random trees. Even though the density is only known as an infinite sum of functions, our algorithm does not require any summation.

KEYWORDS AND PHRASES. Random variate generation. Theta distribution. Simulation. Monte Carlo method. Expected time analysis. Probability inequalities.

1991 MATHEMATICS SUBJECT CLASSIFICATIONS: Primary 65C10. Secondary 65C05, 11K45, 68U20.

Authors' address: School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2A7. The authors' research was sponsored by NSERC Grant A3456 and FCAR Grant 90-ER-0291.

The properties of the theta distribution with distribution function

$$F(x) = \sum_{j=-\infty}^{\infty} \left(1 - 2j^2x^2\right) e^{-j^2x^2}$$
$$= \frac{4\pi^{5/2}}{x^3} \sum_{j=1}^{\infty} j^2 e^{-\pi^2 j^2/x^2}, \ x > 0$$

are described in Rényi and Szekeres (1967). The theta distribution occurs as the limit law for the height of a random rooted labeled free tree. It also is the limit law of the height of many other brands of random trees, such as random planted planar trees (or rooted ordered trees; see DeBruijn, Knuth and Rice, 1972). As shown in Flajolet and Odlyzko (1982), this is no coincidence. In fact, for all so-called simply generated families of trees, the theta distribution describes the limit law. The density of the distribution is plotted below. Its s-th moment, μ_s is given by

$$\mu_s = 2\,\Gamma(1 + s/2)(s - 1)\zeta(s) \ ,$$

where $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$ is the zeta function. In particular, the mean is $\sqrt{\pi}$ and the variance $\pi(\pi - 3)/2$.

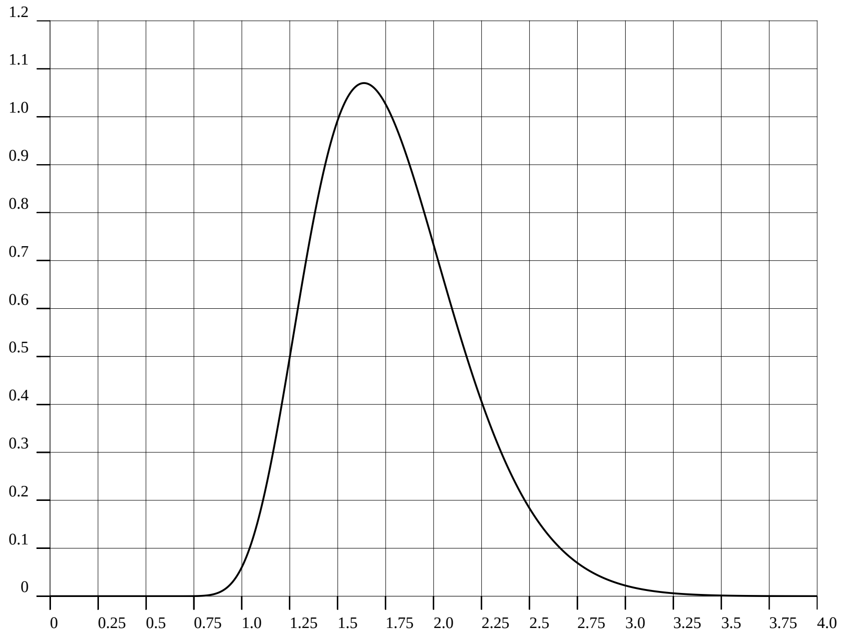

Figure 1. The density of the theta distribution.

The fact that the density is only known as an infinite sum renders random variate generation rather difficult, and this has prompted us to write this note. There are two different sums that describe the theta density,

$$f(x) = \sum_{j=1}^{\infty} f_j(x) = \sum_{j=1}^{\infty} g_j(x) \ ,$$

where

$$f_j(x) \stackrel{\text{def}}{=} 2\left(4j^4x^3 - 6j^2x\right)e^{-j^2x^2} \ , \ x > 0 \ .$$

and

$$g_j(x) \stackrel{\text{def}}{=} 4\pi^{5/2}\left(\frac{2\pi^2j^4}{x^6} - \frac{3j^2}{x^4}\right)e^{-\pi^2j^2/x^2}, \ x > 0 \ .$$

Note that each summand f_j, g_j has one sign change. Also, each summand is of the form $a_j - b_j$ where $a_j, b_j \geq 0$. If the positive portions (the a_j's) together define a mixture modulo a constant, we could invoke von Neumann's rejection method. However, in both cases, this strategy fails as $\sum_j \int a_j = \infty$.

A possible solution is given by the series method first developed in Devroye (1981), and further explained in Devroye (1986). In the series method, the rejection principle is applied, and the acceptance condition—which is of the form $T < f(X)$ for random variables T and X—is verified through convergent upper and lower bounds of f. The convergent bounds are only computed until T drops below the lower bound or exceeds the upper bound, an event that has probability one. In the present case, this strategy would be applicable without any problems. However, the method developed in this paper is conceptually simpler, easier to analyze, and shorter in implementation. We will work out the details for the theta distribution, but the general principle should be useful in many other contexts. First, we note the following.

LEMMA 1. *Let* $\gamma = 16e^{-3\pi} = 0.0012911228\ldots$ *Then*

$$f_j(x) \leq \gamma^{j-1} f_0(x) \ , \ x \geq \sqrt{\pi} \ ,$$

where

$$f_0(x) = 8x^3e^{-x^2} \ .$$

Similarly,

$$g_j(x) \leq \gamma^{j-1} g_0(x) \ , \ 0 \leq x \leq \sqrt{\pi} \ ,$$

where

$$g_0(x) = \frac{8\pi^{9/2}}{x^6}e^{-\pi^2/x^2} \ .$$

PROOF. To see this, it suffices that the ratio of positive portions of f_{j+1} over f_j is not more than γ for all j. For general $j \geq 1$, this ratio is

$$\begin{aligned}\frac{4(j+1)^4 x^3 e^{-(j+1)^2 x^2}}{4j^4 x^3 e^{-j^2 x^2}} &= \left(\frac{j+1}{j}\right)^4 e^{-(2j+1)x^2} \\ &\leq 16 e^{-3x^2} \\ &\leq 16 e^{-3\pi} \ , \ x \geq \sqrt{\pi} \ .\end{aligned}$$

The ratio between consecutive positive portions of the g_j's is

$$\left(\frac{j+1}{j}\right)^4 e^{-(2j+1)\pi^2/x^2} \leq 16 e^{-3\pi^2/x^2} \leq 16 e^{-3\pi} \ , \ 0 \leq x \leq \sqrt{\pi} \ . \square$$

By geometric summation, we conclude the following:

$$\begin{aligned}f(x) &\leq \begin{cases} \frac{1}{1-\gamma} f_0(x) & (x \geq \sqrt{\pi}) \ ; \\ \frac{1}{1-\gamma} g_0(x) & (x \leq \sqrt{\pi}) \ . \end{cases} \\ &= \frac{1}{1-\gamma} \min(f_0(x), g_0(x)) \ .\end{aligned}$$

This is directly useful to develop a rejection method.

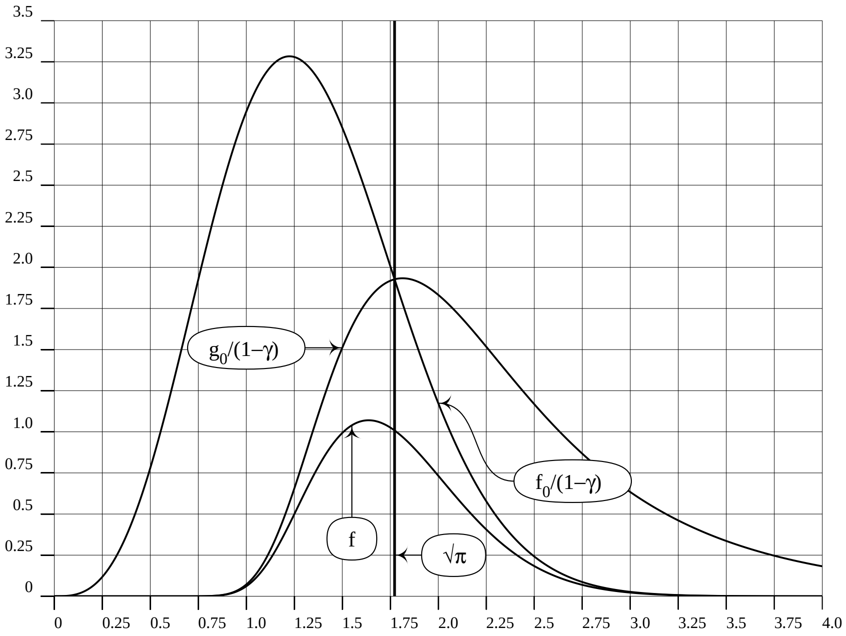

Figure 2. The theta density and the bounding curves f_0 and g_0 are shown. We bound the theta density by $f_0/(1-\gamma)$ on $[\sqrt{\pi}, \infty)$ and by $g_0/(1-\gamma)$ on $[0, \sqrt{\pi}]$. The area under $f_0 + g_0$ is 7. The area under $\min(f_0, g_0)$ is about 1.55. The mean of the theta density is $\sqrt{\pi}$. By design, we have $f_0(\sqrt{\pi}) = g_0(\sqrt{\pi})$.

As $\int_0^\infty f_0 = 4$ and $\int_0^\infty g_0 = 3$, and as $f_0/4$ is the density of the square root of a gamma (2) random variable and $g_0/3$ is the density of $\pi/\sqrt{Y}$, where Y is gamma (5/2), we see that the following algorithm is valid.

```
repeat
  repeat generate V uniform [0,1]
    if V ≤ 4/7
      then X ← √G_2, A ← [X ≥ √π]
      else X ← π/√G_{5/2}, A ← [X ≤ √π]
      (note: G_2 and G_{5/2} are gamma with parameters 2 and 5/2)
  until A
  generate U uniform [0,1]
  if X ≥ √π then Accept ← [U f_0(X)/(1 − γ) ≤ f(X)]
            else Accept ← [U g_0(X)/(1 − γ) ≤ f(X)]
until Accept
return X
```

We recognize the standard rejection method. The gamma variates G_2 and $G_{5/2}$ may be obtained from standard sources such as Best (1978) or Ahrens and Dieter (1982). However, one might as well use $E_1 + E_2$ for the gamma (2) random variate and $E_1 + E_2 + N^2/2$ for the gamma (5/2) variate, where N is standard normal, and E_1 and E_2 are exponential random variables. The main trouble in the algorithm is with the verification of the acceptance condition, which involves the evaluation of the infinite sum $f(X)$. But here is the contribution of this paper: *we may replace the acceptance condition by an equivalent acceptance condition that does not require any summation of a series.* In the range $[\sqrt{\pi}, \infty)$, $f_j(x) \geq 0$ for all j, while on $[0, \sqrt{\pi}]$, $g_j(x) \geq 0$ for all j. As $f_j \leq \gamma^{j-1} f_0$ and $g_j \leq \gamma^{j-1} g_0$, we obtain an equivalent acceptance test by selecting a random $J = j$ with probability $(1-\gamma)\gamma^{j-1}$ for $j \geq 1$, and having picked a piece, to accept X with probability $f_J(X)/\gamma^{J-1} f_0(X)$ or $g_J(X)/\gamma^{J-1} g_0(X)$, depending upon whether $X \geq \sqrt{\pi}$ or not. Observe that if $x \geq \sqrt{\pi}$,

$$\mathbf{P}\left\{X \text{ is accepted} | X = x\right\} = \sum_{j=1}^{\infty} (1-\gamma)\gamma^{j-1} \frac{f_j(x)}{\gamma^{j-1} f_0(x)} = \frac{(1-\gamma) f(x)}{f_0(x)} ,$$

just as in the algorithm given above. A symmetric observation is valid for the g_0 side of the line. This way of avoiding infinite sums in the absence of true mixtures may find many other applications for difficult densities. We summarize the modified algorithm:

```
repeat
  repeat generate V uniform [0,1]
    if V ≤ 4/7
      then X ← √G₂, A ← [X ≥ √π]
      else X ← π/√G_{5/2}, A ← [X ≤ √π]
  until A
  J ← 0
  repeat generate W uniform [0,1], J ← J + 1 until W ≤ 1 − γ
        (J is one plus a geometric (1 − γ))
  generate U uniform [0,1]
  if X ≥ √π then Accept ← [Uγ^{J−1} f₀(X) ≤ f_J(X)]
             (equivalently, Accept ← [Uγ^{J−1} ≤ (J⁴ − (3/2)J²/X²) e^{−(J²−1)X²}])
             else Accept ← [Uγ^{J−1} g₀(X) ≤ g_J(X)]
             (equivalently, Accept ← [Uγ^{J−1} ≤ (J⁴ − 3J²X²/(2π²)) e^{−π²(J²−1)/X²}])
until Accept
return X
```

The expected complexity of this algorithm is appropriately measured by the expected number of random variables V needed per returned X. It is easily seen that this is precisely $7/(1-\gamma)$. If we take the expected number of U or J random variables needed per returned X, the situation is much better. This is nothing but $1/(1-\gamma)$ times

$$\begin{aligned}\int_0^{\sqrt{\pi}} g_0(x)\,dx + \int_{\sqrt{\pi}}^{\infty} f_0(x)\,dx &= 3\mathbf{P}\{G_{5/2} \geq \pi\} + 4\mathbf{P}\{G_2 \geq \pi\} \\ &= 0.838892269\ldots + 0.715897785\ldots \\ &\stackrel{\text{def}}{=} p + q \\ &= 1.55479005\ldots\ .\end{aligned}$$

Additional savings may be achieved by eliminating the inner loop of the algorithm. In that case, with probability $p/(p+q)$, we must generate X with density proportional to $g_0 I_{[0,\sqrt{\pi}]}$ and with the complimentary probability, X must have density proportional to $f_0 I_{[\sqrt{\pi},\infty)]}$. This requires efficient algorithms for the right tail of a gamma distribution. One is referred for this to pages 420–425 of Devroye (1986). Finally, in the definition of `Accept`, $J = 1$ with probability $1 - \gamma \approx 0.999$, and thus, both calculations of exponents may be avoided most of the time.

References.

J. H. Ahrens and U. Dieter, "Generating gamma variates by a modified rejection technique," *Communications of the ACM*, vol. 25, pp. 47–54, 1982.

D. J. Best, "Letter to the editor," *Applied Statistics*, vol. 27, p. 181, 1978.

N. G. de Bruijn, D. E. Knuth, and S. O. Rice, "The average height of planted plane trees," in: *Graph Theory and Computing*, ed. R.-C. Read, pp. 15–22, Academic Press, New York, 1972.

L. Devroye, "The series method in random variate generation and its application to the Kolmogorov-Smirnov distribution," *American Journal of Mathematical and Management Sciences*, vol. 1, pp. 359–379, 1981.

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

P. Flajolet and A. Odlyzko, "The average height of binary trees and other simple trees," *Journal of Computer and System Sciences*, vol. 25, pp. 171–213, 1982.

A. Meir and J. W. Moon, "On the altitude of nodes in random trees," *Canadian Journal of Mathematics*, vol. 30, pp. 997–1015, 1978.

A. Rényi and G. Szekeres, "On the height of trees," *Journal of the Australian Mathematical Society*, vol. 7, pp. 497–507, 1967.