Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 181 (2021) 51–56
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

Random variate generation for the truncated negative gamma distribution ☆

Luc Devroye *

School of Computer Science, McGill University, Canada

Received 4 January 2019; received in revised form 21 August 2020; accepted 1 September 2020
Available online 12 September 2020

Abstract

We provide a uniformly efficient and simple random variate generator for the truncated negative gamma distribution restricted to any interval.

Keywords: Random variate generation; Simulation; Monte Carlo method; Expected time analysis

1. Introduction

In this note, we derive a uniformly fast random variate generator for the family of densities that are proportional to

$$f(x) = x^{-\lambda} \exp(-x), \ x \in [s, t),$$

where $0 < s < t \le \infty$ and $\lambda \ge 1$ are the parameters. For $\lambda < 1$, we obtain the standard gamma $(1 - \lambda)$ distribution, for which many good algorithms are available (see [2], [6] and [4]). For this reason we will call this the negative gamma family. Since it must be a density, we cannot have $s = 0$.

This distribution appears in the astrophysics literature where it is known as the power law with cut-off, or the power law with exponential cut-off. Our nomenclature stresses the tight connection with the gamma distribution.

Deriving uniformly fast algorithms for multi-parameter families of distributions becomes harder as the number of parameters grows. With three parameters, λ, s and t, one must be very careful. Our method is based on the fact that after an exponential transformation, the distribution is log-concave on its support.

The algorithms here are designed for situations in which one or more of the parameters change on each call. If they are static, then there are various other methods that should be considered, including table methods and adaptive rejection sampling [5].

☆ Research sponsored by NSERC Grant A3456.
* Correspondence to: School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2K6.
E-mail address: lucdevroye@gmail.com.

https://doi.org/10.1016/j.matcom.2020.09.005
0378-4754/

2. A transformed negative gamma distribution

The following “trick” helps in a large number of examples. If X is negative gamma, then $Y = \log(X/s)$ has a log-concave density proportional to

$$\exp\left(-(\lambda-1)(y+\log s) - se^{y}\right),\ y \in [0, \log(t/s)).$$

It is understood that when $t = \infty$, then the support of this density is $[0, \infty)$. This density has a unique mode at the origin, is monotonically decreasing on the positive halfline, and is log-concave. It is convenient to normalize so that the value of the function we will be dealing with is 1 at the origin. So, Y has density proportional to $\exp(h(y))$, where

$$h(y) = -(\lambda-1)y - s\left(e^{y}-1\right),\ y \in [0, \log(t/s)).$$

We recall that all log-concave densities for which the density is available in black box format, and for which the location of the mode (or a mode) is known, one can generate random variates by the rejection method thanks to a universal inequality given in [1]. That method breaks down when one only knows the density up to a normalization constant. In this case, our normalization constant involves the incomplete gamma integral, for which only approximations are known. One can get around the unknown normalization constant quite easily—the path to that was sketched in section 7.2.6 of Devroye [2], and in particular in Theorem 2.6 (page 299) and the algorithm on page 301, but indirectly also in the work of Hörmann et al. [6], and Leydold and Hörmann [7,8]), some automated methods worked out in the more recent papers of Devroye [3,4]). However, universal algorithms are never as efficient as specific designs.

Let $z > 0$ be fixed. We will only use the following two inequalities in the rejection method:

$$h(y) \le \begin{cases} 0 & \text{for all } y \ge 0, \\ h(z) + (y-z)h'(z) = h(z) - (y-z)\left((\lambda-1) + se^{z}\right) & \text{if } y \ge z. \end{cases}$$

For general log-concave densities, Leydold and Hörmann [7,8] and indirectly Devroye [3,4] establish that the exponential tail should start when $h(y) \approx -1$. Doing so leads quite easily to the correct recipe for a uniformly fast algorithm. A useful choice of the threshold z is obtained as the minimum of three values. Define

$$\begin{aligned} z_0 &= \log(t/s), \\ z_1 &= \log\left(1 + \frac{1}{2s}\right) \\ z_2 &= \frac{1}{2(\lambda-1)} \text{ (which is } \infty \text{ if } \lambda = 1). \end{aligned}$$

Then set

$$z = \min\left(z_0, z_1, z_2\right),\ w = \operatorname{argmin}\left(z_0, z_1, z_2\right).$$

Rejection will be based on

$$e^{h(y)} \le e^{g(y)} \stackrel{\text{def}}{=} \begin{cases} 1 & 0 \le y \le z, \\ \exp\left(h(z) - a(y-z)\right) & \text{if } z \le y. \end{cases}$$

where we define

$$a \stackrel{\text{def}}{=} (\lambda-1) + se^{z}.$$

We first establish that this is uniformly fast. We did not attempt to optimize the rejection constant. Slightly moving z can improve the bound of Theorem 1:

Theorem 1. *Let N be the number of loops in the rejection method based on the above inequality. Then*

$$\sup_{\substack{0<s\le t\le\infty \\ \lambda\ge 1}} \mathbb{E}\{N\} \le e + 2.$$

Proof. Observe that the integral under the dominating curve e^g is

$$z + \frac{e^{h(z)}}{a}\mathbb{1}_{[w>0]}.$$

The integral of e^h over $[0, z]$ is at least $z\exp(h(z))$. Thus,

$$\mathbb{E}\{N\} \le \frac{z + \frac{e^{h(z)}}{a}\mathbb{1}_{[w>0]}}{z\exp(h(z))} = \exp(-h(z)) + \frac{1}{az}\mathbb{1}_{[w>0]} \stackrel{\text{def}}{=} I + II.$$

If $z = z_1$, then $az \ge sz_1e^{z_1} = (s + 1/2)\log\left(1 + \frac{1}{2s}\right) \ge \frac{1}{2}$. If $z = z_2$, then $az \ge (\lambda - 1)z_2 = 1/2$. So, $II \le 2$. Now, we always have

$$-h(z) = (\lambda - 1)z + s\left(e^z - 1\right) \le (\lambda - 1)z_2 + s\left(e^{z_1} - 1\right) = 1.$$

Therefore, $\mathbb{E}\{N\} \le e + 2$. $\square$

3. The rejection algorithm for our example

To apply the rejection method, we need the integrals of e^g over $[0, z]$ and $[z, \infty)$, respectively. The former is z. The latter is given by

$$b \stackrel{\text{def}}{=} \frac{1}{a}e^{h(z)} = \frac{1}{a}\exp\left(-(\lambda - 1)z - s\left(e^z - 1\right)\right).$$

Furthermore, we note that a random variate with density proportional to e^g on $[z, \infty)$ is simply generated as $z+E/a$, where E is a standard exponential random variable. Finally, in the algorithm below, if W is a candidate point generated from e^h, and V is a uniform [0, 1] random variable, then we replace the rejection step $Ve^{h(W)} \le e^{g(W)}$ by the condition $E^* > h(W) - g(W)$, where we used the fact that V is distributed as e^{-E^*}, where E^* is exponentially distributed (see the various lines that involve E^*). This leads to the following rejection algorithm.

[Algorithm for the negative gamma density on $[s,t]$ of parameter $\lambda \ge 1$.]

[Set-up]

$z_0 \leftarrow \log(t/s)$
$z_1 \leftarrow \log(1 + 1/(2s))$
$z_2 \leftarrow \frac{1}{2(\lambda-1)}$ (∞ if $\lambda = 1$)
$z \leftarrow \min(z_0, z_1, z_2)$
$w \leftarrow \arg\min(z_0, z_1, z_2)$
$a \leftarrow (\lambda - 1) + se^z$
$b \leftarrow \frac{1}{a}\exp(-(\lambda - 1)z - s(e^z - 1))$

[Generation]

if $w = 0$ then repeat
    $W \leftarrow zU$, where U is uniform $[0,1]$
    generate an exponential random variable E^*
    Accept $\leftarrow [E^* > (\lambda - 1)W + s\left(e^W - 1\right)]$
    until Accept
  else repeat
    if $V < z/(z + b)$ (where V is uniform on $[0,1]$)
      then $W \leftarrow zU$, where U is uniform $[0,1]$
        generate an exponential random variable E^*
        Accept $\leftarrow [E^* > (\lambda - 1)W + s\left(e^W - 1\right)]$
      else $W \leftarrow z + E/a$, where E is exponential
        generate an exponential random variable E^*
        Accept $\leftarrow [W \le \log(t/s)]$
          and $[E^* > se^z\left(e^{W-z} - 1 - (W - z)\right)]$
    until Accept
return $X \leftarrow se^W$

4. A special case: $\lambda = 1$

For the one parameter density proportional to

$$f(x) = \frac{1}{x} e^{-x}, x \geq s > 0,$$

the algorithm becomes much simpler:

```
[Algorithm for the negative gamma density on [s,∞) of parameter λ = 1.]
[Set-up]
z ← log(1 + 1/s)
b ← 1/((s+1)e)
[Generation]
repeat
if V < z/(z + b) (where V is uniform on [0,1])
  then W ← zU, where U is uniform [0,1]
       generate an exponential random variable E*
       Accept ← [E* > s(e^W − 1)]
  else W ← z + E/(s + 1), where E is exponential
       generate an exponential random variable E*
       Accept ← [E* > (s + 1)(e^(W−z) − 1 − (W − z))]
until Accept
return X ← se^W
```

In the algorithm above, we tacitly replaced z_1 by the choice of z given in the first line of the algorithm. One can verify that the bounding method used in Theorem 1 gives the better estimate $\mathbb{E}\{N\} \leq e + 1$.

5. The gamma density with parameter in (0, 1]

There are many methods for generating gamma random variables with arbitrary parameters. We are interested though in the case of gamma random variables with parameter $b = 1 - \lambda \in (0, 1]$, i.e., having density

$$\frac{x^{-\lambda} e^{-x}}{\Gamma(1-\lambda)},$$

but restricted to the interval $[s, t] \subseteq [0, \infty)$. This is a three parameter family of distributions. The purpose is, once again, to derive a uniformly fast rejection method.

First we note that the transformed random variable

$$Y = X^b$$

has density proportional to

$$\exp\left(-y^{1/b}\right), y \geq 0,$$

when X is gamma (b). It is much simpler to deal with Y, as its density is monotonically decreasing and log-concave on the positive halfline. We note that this same transformation has also been suggested by Tanizaki [10]. We will restrict Y to $[s^b, t^b]$. Note that the density of Y is proportional to $\exp(h(y))$ where

$$h(y) = s - y^{1/b}.$$

This normalized form has $h(s^b) = 0$, which facilitates the further development. Set

$$z_0 = t^b,$$
$$z_1 = (1 + s)^b,$$
$$z = \min(z_0, z_1).$$

On $[s^b, z]$, we will apply rejection with as bounding curve 1. On $[z, t^b]$, if this interval is not empty, we use the bound

$$h(y) \leq h(z) - a(y - z),$$

with $a = -h'(z) = (1/b)z^{(1/b)-1}$. The algorithm is as follows:

```
[Algorithm for the negative gamma density on [s,t] of parameter λ ∈ [0,1).]
[Set-up]
  b ← 1 − λ
  z_0 ← t^b
  z_1 ← (1 + s)^b
  z ← min (z_0, z_1)
  w ← arg min (z_0, z_1)
  a ← (1/b) z^{(1/b)−1}
  γ ← (1/a) exp (s − z^{1/b})
[Generation]
if w = 0 then repeat
            W ← s^b + (t^b − s^b)U, where U is uniform [0,1]
            generate an exponential random variable E*
            Accept ← [E* > W^{1/b} − s]
            until Accept
      else repeat
            if V < (z − s^b)/((z − s^b) + γ) (where V is uniform on [0,1])
              then W ← s^b + (z − s^b)U, where U is uniform [0,1]
                   generate an exponential random variable E*
                   Accept ← [E* > W^{1/b} − s]
              else W ← z + E/a, where E is exponential
                   generate an exponential random variable E*
                   Accept ← [W ≤ t^b]
                       and [E* > W^{1/b} − z^{1/b} − a(W − z)]
            until Accept
return X ← W^{1/b}
```

Theorem 2. *Let N be the number of loops in the rejection method based on the above algorithm. Then*

$$\sup_{\substack{0\le s\le t\le\infty \\ \lambda\in[0,1)}} \mathbb{E}\{N\} \le \frac{e^2}{e-1}.$$

Proof. Observe that the integral under the dominating curve e^h is

$$z - s^b + \frac{e^{h(z)}}{a}\mathbb{1}_{[w>0]}.$$

The integral of e^h over $[s^b, z]$ is at least $(z - s^b)\exp(h(z))$. Thus,

$$\mathbb{E}\{N\} \le \frac{z - s^b + \frac{e^{h(z)}}{a}\mathbb{1}_{[w>0]}}{(z - s^b)\exp(h(z))} = \exp(-h(z)) + \frac{1}{a(z - s^b)}\mathbb{1}_{[w>0]} \stackrel{\text{def}}{=} I + II.$$

If $z = z_1$, then

$$\begin{aligned}
a(z - s^b) &= \frac{z_1^{1/b-1}(z_1 - s^b)}{b} \\
&= \frac{(s+1)((1+s)^b - s^b)}{b(s+1)^b} \\
&= \frac{s+1}{b}\left(1 - \frac{1}{(1+1/s)^b}\right) \\
&\ge \frac{s+1}{b}\left(1 - \exp(-b/(s+1))\right).
\end{aligned}$$

Since $s+1 \geq b$, the expression is at least $(e-1)/e$. So, $II \leq e/(e-1)$. Now, we always have

$$-h(z) = z^{1/b} - s \leq z_1^{1/b} - s = 1.$$

So, $I \leq e$. □

Acknowledgements

The distribution was pointed out to me by astrophyscist B.T. Ravishankar of the Indian Institute of Space Science and Technology, and UR Rao Satellite Centre, who needed a fast generator for it in order to generalize the work of Magdziarz and Zdziarski [9]. Ravishankar implemented and successfully tested all algorithms shown in this note, but declined having his name as coauthor of the paper. The author would also like to thank both referees.

References

[1] L. Devroye, A simple algorithm for generating random variates with a log-concave density, Computing 33 (1984) 247–257.
[2] L. Devroye, Non-Uniform Random Variate Generation, Springer-Verlag, New York, 1986.
[3] L. Devroye, A note on generating random variables with log-concave densities, Statist. Probab. Lett. 82 (2012) 1035–1039.
[4] L. Devroye, Random variate generation for the generalized inverse Gaussian distribution, Stat. Comput. 24 (2014) 239–246.
[5] W.R. Gilks, P. Wild, Adaptive rejection sampling for Gibbs sampling, Appl. Stat. 42 (1992) 337–348.
[6] W. Hörmann, J. Leydold, G. Derflinger, Automatic Nonuniform Random Variate Generation, Springer-Verlag, Berlin, 2004.
[7] J. Leydold, W. Hörmann, Black box algorithms for generating non-uniform continuous random variates, in: W. Jansen, J.G. Bethlehem (Eds.), Compstat 2000, 2000, pp. 53–54.
[8] J. Leydold, W. Hörmann, Universal algorithms as an alternative for generating non-uniform continuous random variates, in: G.I. Schuler, P.D. Spanos (Eds.), Monte Carlo Simulation, 2001, pp. 177–183.
[9] P. Magdziarz, A.A. Zdziarski, Angle-dependent Compton reflection of X-rays and gamma-rays, Mon. Not. R. Astron. Soc. 273 (1995) 837–848.
[10] H. Tanizaki, A simple gamma random number generator for arbitrary shape parameters, Econ. Bull. 3 (2008) 1–10.