Statistics and Computing (1997) **7**, 173–181

Simulation of right and left truncated gamma distributions by mixtures

ANNE PHILIPPE

CNRS UPRES-A 60 85, Université de Rouen 76821 Mont Saint Aignan Cedex, France

Received August 1996 and accepted March 1997

We study the properties of truncated gamma distributions and we derive simulation algorithms which dominate the standard algorithms for these distributions. For the right truncated gamma distribution, an optimal accept–reject algorithm is based on the fact that its density can be expressed as an infinite mixture of beta distribution. For integer values of the parameters, the density of the left truncated distributions can be rewritten as a mixture which can be easily generated. We give an optimal accept–reject algorithm for the other values of the parameter. We compare the efficiency of our algorithm with the previous method and show the improvement in terms of minimum acceptance probability. The algorithm proposed here has an acceptance probability which is superior to $e/4$.

Keywords: minimum acceptance probability, mixture distribution, accept–reject algorithm

1. Introduction

The need for simulation of truncated gamma distributions appears in many settings when the parameter of the model is constrained by some form of ordering or censoring (see Gelfand *et al.*, 1992). Frequentist formulations have then to account for these constraints and, in a Bayesian framework, the restrictions transfer to the posterior distributions. The constraints can be so involved that both maximum likelihood and Bayesian estimations require simulations, either through EM type algorithms (see Dempster *et al.*, 1977; Geyer and Thomson, 1992) or Gibbs sampling (see Mengersen and Robert, 1996; Robert and Titterington, 1996). These different methods require large numbers of simulations and, therefore, it is necessary to provide efficient algorithms for truncated gamma distribution.

For truncated distributions, the accept–reject algorithms are the most commonly used simulation methods (see Devroye (1985) for the left truncated gamma distribution or Robert (1995) for the normal distribution). These algorithms are based on the following lemma (see Robert, 1996):

Lemma 1.1 Let f and g be two densities such that $f(x) \leq Mg(x)$ for every x in the support of f. The random variable x resulting from the algorithm:

1. `Generate` $z \sim g(z)$ `and` $u \sim \mathscr{U}_{[0,1]}$.
2. `If` $u \leq f(z)/Mg(z)$, `take` $x = z$; $[A_0]$
   `else repeat step 1.`

is distributed according to f.

The number of random variables z required to produce x is a geometric random variable $\mathscr{G}(M^{-1})$. As the efficiency of the algorithm depends on the choice of g, this instrumental distribution is usually chosen such that M is close to 1.

In Section 2, we study the properties of the right truncated gamma distribution. We show that it can be expressed as an infinite beta mixture. We then deduce an accept–reject algorithm based on finite mixtures of beta distributions. The acceptance probability depends on the number of components in the mixture through an explicit expression. Moreover, we deduce the number of components necessary to ensure that the acceptance probability is superior to a fixed threshold p.

In Section 3, we study the left truncated gamma distribution. In the particular case $a \in \mathbb{N}$, we establish some mixture properties which give new perspectives for building accept–reject algorithms for the other values of a. In the last section, a comparison with the classical algorithms

The different programs in C code are available on the electronic site: *http://www.dir.univ-rouen.fr/ura1378/*

proposed in Dagpunar (1978) and Devroye (1985) show the improvement brought by our new algorithm. Some of the proofs from Sections 2 and 3 are placed in the Appendix.

2. Right truncated gamma distributions

We denote by $\mathscr{TG}^-(a,b,t)$ the truncated gamma distribution with right truncation point t, i.e. the distribution with density

$$f^-(x|a,b,t) = \frac{b^a}{\gamma(a,bt)} \exp(-bx)x^{a-1}\, \mathbb{I}_{x\le t} \tag{1}$$

with $t>0$ and $\gamma(a,b) = \int_0^b \exp(-u)u^{a-1}\,\mathrm{d}x$. Without loss of generality, we can assume that $t=1$, since if $X \sim \mathscr{TG}^-(a,b,t)$ then $X/t \sim \mathscr{TG}^-(a,bt,1)$.

Proposition 2.1 The right truncated gamma distribution $\mathscr{TG}^-(a,b,1)$ can be expressed as an infinite mixture of beta distributions, i.e.

$$X \sim \sum_{k=1}^{+\infty} \frac{b^{a+k-1}\Gamma(a)\exp(-b)}{\Gamma(a+k)\gamma(a,b)}\, \mathscr{B}e(a,k) \tag{2}$$

where $\mathscr{B}e(\cdot,\cdot)$ denotes the beta distributions.

It is easy to see that when k goes to infinity, the weights of the mixture converge to zero. Therefore, Proposition 2.1 suggests an accept–reject algorithm $[A_{1n}]$ based on the following approximation of the density,

$$g_n(x) \propto \sum_{k=1}^{n} \frac{\omega_k}{\sum_{i=1}^n \omega_i} f_k(x)$$

where f_k is the density of the $\mathscr{B}e(a,k)$ distribution and $\omega_k = b^{k-1}/\Gamma(a+k)$ as in $f^-(x|a,b)$.

Proposition 2.2 For (a,b) fixed, the acceptance probability of the accept–reject algorithm $[A_{1n}]$ based on g_n is given by

$$P(n) = \frac{1 - \dfrac{\gamma(n,b)}{\Gamma(n)}}{1 - \dfrac{\gamma(a+n,b)\Gamma(a)}{\Gamma(a+n)\gamma(a,b)}} \tag{3}$$

Proposition 2.2 shows the relation between the acceptance probability and the number of components of the instrumental mixture. The probability $P(n)$ converges to 1 when n goes to infinity. Therefore, it is always possible to find n which ensures that $P(n) \ge p$ for fixed p. Given the complex expression of the acceptance probability (3), it is not possible to obtain an explicit expression of the number of components. However, in the following proposition, we make explicit the number of components n necessary to guarantee that $P(n) \ge p$ for an arbitrary $p \in [0,1]$.

Proposition 2.3 The number n of components necessary to ensure that $P(n) > p$ with $p \in [0.5,1]$ is approximated by

$$n(p) = \frac{1}{4}\left(t_p + \sqrt{t_p^2 + 4b}\right)^2 \tag{4}$$

where t_p is the $100\,p\%$ quantile of the normal distribution.

By using numerical methods, we can calculate the optimal number of components, i.e. the smallest integer $N_{\mathrm{op}}(p)$ such that $P(n) > p$ for $n > N_{\mathrm{op}}(p)$. However, in Table 1, we can note that the optimal value and (4) are quite similar. In particular, the number of components given by (4) is asymptotically optimal when a is large. However, it may be too pessimistic for small values of a.

Table 1. *Optimal number of components N_{op} and (4) for $p = 0.95$. Comparison of the acceptance probabilities of $[A_{1N}]$ (for both values of N) with the accept–reject algorithms based on the standard densities g. (NA: not available)*

a, b	$n(0.95)$	$N_{\mathrm{op}}(0.95)$	$P(n(0.95))$	$P(N_{\mathrm{op}}(0.95))$	$g \sim \mathscr{U}_{[0,1]}$	$g \sim \mathscr{G}a(a,b)$
0.1, 0.1	2	1	0.99	0.99	NA	0.83
1, 0.1	2	1	0.99	0.95	0.95	0.09
5, 0.1	2	2	0.99	0.99	0.20	0.00
10, 0.1	2	2	0.99	0.99	0.00	0.00
0.1, 1	4	2	0.99	0.97	NA	0.97
1, 1	4	4	0.98	0.98	0.63	0.63
5, 1	4	4	0.98	0.98	0.23	0.00
10, 1	4	4	0.97	0.97	0.10	0.00
0.1, 5	10	5	0.99	0.96	NA	0.99
1, 5	10	9	0.98	0.96	0.19	0.99
5, 5	10	10	0.96	0.96	0.63	0.55
10, 5	10	10	0.96	0.96	0.17	0.03
0.1,10	16	8	0.99	0.95	NA	0.99
1,10	16	15	0.97	0.96	0.10	0.99
5,10	16	16	0.95	0.95	0.50	0.97
10,10	16	16	0.95	0.95	0.43	0.54

In practice, we fix the acceptance probability, and the number of components $N = n(p)$ is given by Proposition 2.3. Since there exists a recurrence relation between the weights of the mixture, it is easy to build an algorithm to generate from the instrumental distribution. We take $\overline{\omega}_1 = \tilde{\omega}_1 = 1$ and for $k \in 1, \cdots, N-1$

$$\overline{\omega}_{k+1} = \overline{\omega}_k \frac{b}{a+k} \qquad \tilde{\omega}_{k+1} = \tilde{\omega}_k + \overline{\omega}_{k+1} \tag{5}$$

By taking $\tilde{\omega}_k = \tilde{\omega}_k / \tilde{\omega}_N$ and $\tilde{\omega}_0 = 0$, we get the following algorithm

1. `Generate` $u \sim \mathscr{U}_{[0,1]}$,
2. `Let` k `such that` $u \in [\tilde{\omega}_{k-1}, \tilde{\omega}_k]$, `and generate` $x \sim \mathscr{B}e(a, k)$ $[A_2]$
3. `Return` x

x is distributed according to the density g_N. Finally, the accept–reject algorithm $[A_{1N}]$ based on g_N provides an algorithm to simulate $\mathscr{T}\mathscr{G}^-(a, b, t)$.

1. `Generate` $u \sim \mathscr{U}_{[0,1]}$ `and` $x \sim g_N$
2. `Compute` $M^{-1} = \sum_{k=1}^{N} b^{k-1}/\Gamma(k)$
   `and` $\rho(x) = \left(e^{bx} \sum_{k=1}^{N} b^{k-1} (1-x)^{k-1} / \Gamma(k) \right)^{-1}$
3. `If` $uM \le \rho(x)$ `take` x; $[A_{1N}]$
   `else go back to Step 1.`

The interest of our algorithm is that the efficiency does not depend on the parameters of this distribution. Therefore it dominates and improves on the standard accept–reject algorithm. Indeed, if we consider the naive algorithm based on the gamma distribution, i.e. repeatedly simulating from the gamma distribution $\mathscr{G}a(a, b)$ until $x < 1$, the acceptance probability is equal to

$$P_g = \gamma(a, b)/\Gamma(a)$$

It is easy to see that the probability P_g converges to zero when a goes to infinity for fixed b's. Another standard algorithm is the accept–reject algorithm based on the uniform distribution. This algorithm is only defined when $a \ge 1$ and in this case, the acceptance probability is given by

$$P_u = \begin{cases} \dfrac{\gamma(a,b)}{b^a e^{-b}} & \text{if} \quad b \le a-1 \;\text{ and }\; a > 1 \\ \dfrac{\gamma(a,b)}{b e^{a-1} (a-1)^{a-1}} & \text{if} \quad b \ge a-1 \;\text{ and }\; a > 1 \\ \dfrac{1-e^{-b}}{b} & \text{if} \quad a = 1 \end{cases}$$

As in the previous algorithm, the performance of the algorithm depends on the parameter of the gamma distribution and the acceptance probability converges to zero when b goes to infinity for fixed a's. Therefore the efficiency of these algorithms is not ensured for every value of the parameter and in this sense the algorithm $[A_{1N}]$ dominates both algorithms. This fact is illustrated in Table 1.

3. Simulation of left truncated gamma distributions

We denote by $\mathscr{T}\mathscr{G}^+(a, b, t)$ the truncated gamma distribution with left truncation at t, i.e. the distribution with density

$$f^+(x|a, b, t) = \frac{b^a}{\Gamma(a, bt)} \exp(-bx) x^{a-1} \mathbb{I}_{x \ge t} \tag{6}$$

with $t > 0$ and $\Gamma(c, d) = \int_d^{+\infty} \exp(-u) u^{c-1}\, dx$. As for the right truncated distribution, we can assume without loss of generality that $t = 1$ and we denote by $f^+(\cdot|a, b)$ the corresponding density.

3.1. *a is a positive integer*

The following proposition gives a property of the left truncated gamma distribution when $a \in \mathbb{N}$.

Proposition 3.1 If $X \sim \mathscr{T}\mathscr{G}^+(a, b, 1)$ and $a \in \mathbb{N}^*$ then $Y = X - 1$ follows a finite mixture of gamma $\mathscr{G}a(k, b)$ distribution, i.e.

$$Y \sim \sum_{k=1}^{a} \frac{b^{a-k} e^{-b}}{\Gamma(a, b)} \frac{(a-1)!}{(a-k)!} \mathscr{G}a(k, b)$$

Note that the sequence of weights ω_k satisfies the relation $w_{k+1} = ((a-k)/b)\omega_k$. Therefore, by considering the finite sequence (v_k, w_k) such that $v_1 = w_1 = 1$, and for $2 \le k \le a$,

$$v_k = v_{k-1}(a - k + 1)/b$$
$$w_k = w_{k-1} + v_k$$

we can generate the discrete random variable Z given by $P(Z = k) \propto v_k$ for $k \in 1, \cdots, a$. Then, since $\tilde{w}_k = v_k / w_a$, which is increasing, the algorithm is

1. `Generate` $u \sim \mathscr{U}_{[0,1]}$
2. `Let` k `such that` $u \in [\tilde{w}_{k-1}, \tilde{w}_k)$
   `and` $x \sim \mathscr{G}a(k, b)$ $[A_3]$
3. `Return` $x + 1$.

3.2. *Generalization*

In this section, we provide an accept–reject algorithm which generates the truncated gamma distribution $\mathscr{T}\mathscr{G}^+(a, b, 1)$ for a general a, using a distribution $\mathscr{T}\mathscr{G}^+(n, \tilde{b})$ with $n \in \mathbb{N}^*$ (since Proposition 3.1 gives an algorithm to generate from these densities).

First, the ratio of the densities $f^+(\cdot|a, b)/f^+(\cdot|n, \tilde{b})$ is bounded on $[1, +\infty]$ if and only if

$$b - \tilde{b} \ge 0 \quad \text{and} \quad a < n \tag{C_1}$$

or

$$b - \tilde{b} > 0 \quad \text{and} \quad a > n \ge 1 \tag{C_2}$$

Second, the bound of the ratio, when it exists, is obtained for

$$x_{\max} = \sup\{(a-n)/(b-\tilde{b}), 1\}$$

and we have

$$M = \sup_{x \geq 1} f^{+}(x|a,b)/f^{+}(x|n,\tilde{b})$$
$$= \begin{cases} \dfrac{\Gamma(n,\tilde{b})}{(\tilde{b}^{n})} \dfrac{b^{a}}{\Gamma(a,b)} e^{n-a} \left(\dfrac{a-n}{b-\tilde{b}}\right)^{a-n} & \text{if} \quad \dfrac{a-n}{b-\tilde{b}} \geq 1 \\ \dfrac{\Gamma(n,\tilde{b})}{\tilde{b}^{n}} \dfrac{b^{a}}{\Gamma(a,b)} e^{\tilde{b}-b} & \text{if} \quad \dfrac{a-n}{b-\tilde{b}} \leq 1 \end{cases}$$

3.2.1. *An accept–reject algorithm under (C_1).* Under (C_1), $a-n<0$ implies that $x_{\max}=1$ and the acceptance probability is

$$P^{C_1} = \frac{\Gamma(a,b)}{b^{a}} \frac{\tilde{b}^{n}}{\Gamma(n,\tilde{b})} e^{b-\tilde{b}}$$

Proposition 3.2 Under (C_1), the acceptance probability is maximal for the couple $(n,\tilde{b}) = (\lceil a \rceil, b)$ where $\lceil a \rceil$ is the smallest integer larger than a and this probability is equal to

$$P^{C_1} = \frac{\Gamma(a,b)}{\Gamma(\lceil a \rceil, b)} b^{\lceil a \rceil - a}$$

We denote by $[A_4]$ the corresponding algorithm.

The corresponding accept–reject algorithm is

1. Generate $x \sim \mathcal{TG}^{+}(\lceil a \rceil, b, 1)$ and $u \sim \mathcal{U}_{[0,1]}$.
2. Compute $\rho(x) = x^{a-\lceil a \rceil}$. $[A_4]$
3. If $u \leq \rho(x)$ take x;
   else go back to Step 1.

Note that, when b goes to infinity, the acceptance probability converges to 1. Indeed, for b large, we have the following approximation (see Abramowitz and Stegun, 1964, p. 263)

$$\Gamma(a,b) \sim e^{-b} b^{a-1} \left(1 + \frac{a-1}{b}\right)$$

which implies that

$$P^{C_1} \simeq \frac{1+(a-1)/b}{1+(\lceil a \rceil - 1)/b}$$
$$\simeq 1 - \frac{\lceil a \rceil - a}{b}$$

However, when b goes to zero, P^{C_1} goes to zero and this algorithm is not efficient for small values of the parameter b.

3.2.2. *An accept–reject algorithm under (C_2).* Let $S = [0, b-a+n]$, the acceptance probability can be expressed as

$$P^{C_2}(n,\tilde{b}) = \frac{\Gamma(a,b)}{b^{a}} \frac{\tilde{b}^{n}}{\Gamma(n,\tilde{b})} \left(e^{b-\tilde{b}} \mathbb{1}_{S}(\tilde{b}) + e^{a-n} \left(\frac{a-n}{b-\tilde{b}}\right)^{n-a} \mathbb{1}_{S^c}(\tilde{b})\right) \mathbb{1}_{\tilde{b} \leq b}$$

Following from the proof of Proposition 3.2 with fixed n, P^{C_2} is increasing on S. Therefore the optimal value of $\tilde{b}$ is in S^c if $b-a+n>0$, else $0 \leq \tilde{b} \leq b$ and the optimal couple $(n,\tilde{b})$ maximizes the expression

$$P = \frac{\Gamma(a,b)}{b^{a}} \frac{\tilde{b}^{n}}{\Gamma(n,\tilde{b})} e^{a-n} \left(\frac{a-n}{b-\tilde{b}}\right)^{n-a} \mathbb{1}_{S^c}(\tilde{b}) \mathbb{1}_{\tilde{b} \leq b}$$

For n fixed, it is easy to calculate the derivative $\partial P/\partial \tilde{b}$ and to show that it is equal to zero if and only if $\tilde{b}$ satisfies

$$\Gamma(n,\tilde{b}) = \frac{\tilde{b}^{n}(b-\tilde{b})e^{-\tilde{b}}}{a\tilde{b} - nb} \tag{7}$$

While it is not possible to find an explicit solution to this equation, note that some numerical methods or software such as *Maple V* or *Mathematica* can be used to find the numerical solution; however the cost in time is important and reduces the efficiency of the accept–reject algorithm. In the following lemmas, we give the optimal values of $(n,\tilde{b})$ for the extreme values of b.

Lemma 3.3 For small values of b, the optimal couple $(n,\tilde{b})$ is $(\lfloor a \rfloor, b\lfloor a \rfloor/a)$ where $\lfloor a \rfloor$ is the largest integer smaller than a. And the acceptance probability is

$$P_{0}^{C_2} = \frac{\Gamma(a,b)}{a^{a}} \frac{\lfloor a \rfloor^{\lfloor a \rfloor}}{\Gamma(\lfloor a \rfloor, b\lfloor a \rfloor/a)} e^{a-\lfloor a \rfloor}$$

Lemma 3.4 For large values of b and $b > \lfloor a \rfloor - a$, the optimal couple $(n,\tilde{b})$ *is* $(\lfloor a \rfloor, b-a+\lfloor a \rfloor)$ where $\lfloor a \rfloor$ is the largest integer smaller than a. And the acceptance probability is

$$P_{\infty}^{C_2} = \frac{\Gamma(a,b)}{b^{a}} \frac{(b-a+\lfloor a \rfloor)^{\lfloor a \rfloor}}{\Gamma(\lfloor a \rfloor, b-a+\lfloor a \rfloor)} e^{a-\lfloor a \rfloor}$$

We thus obtained the asymptotically optimal couples $(n,\tilde{b})$. Note that the probabilities $P_{0}^{C_2}$ and $P_{\infty}^{C_2}$ are equal for $b=a$. Since it is not possible to obtain optimal couples for every value of (a,b), we propose to use the couple $(\lfloor a \rfloor, \lfloor a \rfloor b/a)$ if $b<a$ and $(\lfloor a \rfloor, b-a+\lfloor a \rfloor)$ otherwise. The corresponding accept–reject algorithm is

If $b \leq a$

1. Generate $x \sim \mathcal{TG}^{+}(\lfloor a \rfloor, b\lfloor a \rfloor/a, 1)$ and $u \sim \mathcal{U}_{[0,1]}$.
2. Compute $\rho_1(x) = x^{a-\lfloor a \rfloor} e^{-xb(1-\lfloor a \rfloor/a)}$ and
   $M = e^{-(a-\lfloor a \rfloor)}$.
3. If $\rho_1(x) > Mu$ take x;
   else go back to Step 1.

else $[A_5]$

1. Generate $x \sim \mathcal{TG}^{+}(\lfloor a \rfloor, b-a+\lfloor a \rfloor, 1)$ and $u \sim \mathcal{U}_{[0,1]}$.
2. Compute $\rho_2(x) = x^{a-\lfloor a \rfloor} e^{-x(a-\lfloor a \rfloor)}$ and $M = (a/b)^{a-\lfloor a \rfloor}$ $e^{-(a-\lfloor a \rfloor)}$.
3. If $\rho_2(x) > Mu$ take x, else go back to Step 1.

The acceptance probability of this algorithm is denoted by P_2.

Proposition 3.5 The algorithm $[A_5]$ satisfies the following properties:

(1) for every couple (a, b), P_2 is superior to $e/4$;
(2) for a fixed, P_2 converges to 1 when b goes to infinity;
(3) for b fixed, P_2 converges to 1 when a goes to infinity.

Note that in the case $a < 1$, $[A_5]$ is not available. For $a > 1$, as shown in Fig. 1, the algorithm $[A_5]$ does not necessarily dominate the algorithm $[A_4]$ for every b. However, its rate of convergence to 1 when b goes to infinity is superior. Moreover, for b close to 0, the acceptance probability of $[A_4]$ goes to zero while the acceptance probability of $[A_5]$ is bounded from below by $e/4$ (for every b). Therefore when $a > 1$, we consider that the algorithm $[A_5]$ is better than $[A_4]$.

Figure 1 compares the algorithm $[A_5]$ with the algorithm obtained under (C_2) by considering the optimal couple $(n, \tilde{b})$. Indeed, numerical methods give an approximation of the optimal couple which satisfies (7). It is obvious that for b close to zero and b large, the behaviour of both acceptance probabilities is similar. Indeed, the algorithm $[A_5]$ is obtained by taking the optimal couple for these values of the parameter. For the other values of the parameter, we can appreciate that the difference is not significant. Therefore, given that the gain in time brought by $[A_5]$ instead of the approximate solution is minimal, we can note the high efficiency of this algorithm.

4. Comparison with standard algorithms

Devroye (1985) proposes different accept–reject algorithms to generate the left truncated gamma distribution $\mathscr{TG}^+(a, b, 1)$. First, when $a < 1$, the algorithm provided by Devroye coincides with $[A_4]$. This algorithm is based on the truncated exponential distribution since $\lceil a \rceil = 1$. Second, for $a > 1$ and $b > a - 1$, the algorithm proposed by Devroye (1985) is

1. `Repeat`
   `Generate` $u \sim \mathscr{U}_{[0,1]}$ `and` $e, e' \sim \mathscr{E}xp(1)$.
   `Take` $x = b + \frac{e}{1-(a-1)/b}$ $[A_6]$
   `until` $x/b - 1 + \ln(b/x) < e'/\ (a-1)$.
2. `Return` x/b.

The acceptance probability of this algorithm is

$$P_3(a, b) = \left(1 - \frac{a-1}{b}\right)\Gamma(a, b)e^b b^{1-a}$$
$$\simeq 1 - \frac{(a-1)}{b^2} + O(b^{-3}) \qquad \text{for } b \text{ large}$$

We can compare the acceptance probability of this algorithm with $[A_5]$. The rate of convergence of both algorithms

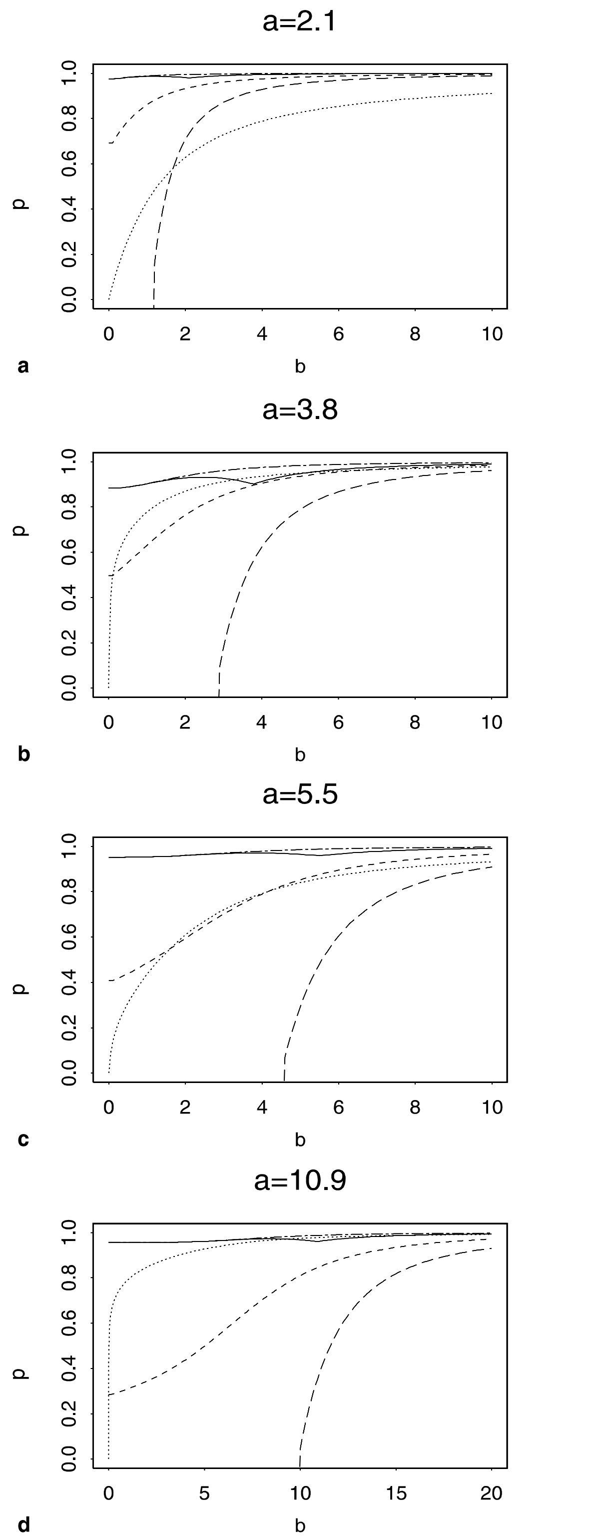

Fig. 1. *Representation of the acceptance probability of the algorithms $[A_4]$ (dots), $[A_5]$ (plain), the optimal algorithm under C_2 (dots and dashes), $[A_6]$ (long dashes) and $[A_7]$ (dashes) for different values of a. (a) a =2.1; (b) a =3.8; (c) a =5.5; (d) a =10.9. Note that $[A_6]$ is only defined for b > a−1.*

is in $O(b^{-2})$, but, for $a \geq 2$, $[A_5]$ dominates $[A_6]$. Indeed, for b large

$$P_2(a,b) - P_3(a,b) = b^{-2}(\lfloor a \rfloor - 1)$$

which is positive if $\lfloor a \rfloor > 1$, i.e. $a \geq 2$.

We can also compare $[A_5]$ with the accept–reject algorithm proposed by Dagpunar (1978) (denoted $[A_7]$). The generation of the distribution $\mathscr{TG}^+(a,b,1)$ (for $a > 1$ and $b > 0$) is based on the left truncated exponential distribution $y \sim e^{-cy}\mathbb{I}_{y \geq b}$; the optimal choice of c is

$$c_0 = \frac{b - a + \sqrt{(b-a)^2 + 4b}}{2b}$$

The accept–reject algorithm is

1. Generate $x \sim e^{-c_0 y}\mathbb{I}_{y\geq b}$ and $u \sim \mathscr{U}_{[0,1]}$.
2. Compute $\rho(x) = x^{a-1}e^{-x(1-c_0)}$
   and $M = (\frac{a-1}{1-c_0})^{a-1}e^{1-a}$. $[A_7]$
3. If $u \leq \rho(x)/M$ take x/b;
   else go back to Step 1

The acceptance probability is

$$\begin{aligned} P_4(a,b) &= \Gamma(a,b)c_0\left(\frac{1-c_0}{a-1}\right)^{a-1} e^{a-1+c_0 b} \\ &\simeq 1 - \frac{(a-1)(4a-5)}{2b^2} + O(b^{-3}) \quad \text{for } b \text{ large} \end{aligned}$$

Hence, for large b, it follows that

$$P_2(a,b) - P_4(a,b) \simeq b^{-2}(-a + \lfloor a \rfloor - (a-1)(4a-5))$$

is positive if $a \geq 2$. Therefore, $[A_7]$ is also dominated by $[A_5]$ asymptotically. We can appreciate the practical domination in Fig. 1.

When a is an integer, we generate the distribution from a mixture of gamma distributions and the acceptance probability is equal to 1. The cost of this algorithm is the computation of the weights of the mixture. However, it is necessary to carry out this step only once when we generate a sample. Therefore $[A_5]$ appears to be more efficient than $[A_6]$ and $[A_7]$ when we generate a large sample from $\mathscr{TG}^+(a,b,1)$.

The algorithm $[A_5]$ is based on the left truncated gamma distribution with an integer parameter whereas $[A_6]$ and $[A_7]$ use exponential distribution. It is clear that the cost of $[A_5]$ per value is higher than the simulation of the exponential distribution. However, when these distributions are used as instrumental distributions in the accept–reject algorithm, the difference is not significant.

Therefore this representation of the left truncated distribution by mixtures gives an accept–reject algorithm which is more efficient than standard algorithms. The existence of a minimum acceptance probability of $e/4 \simeq 0.67$ ensures the efficiency of $[A_5]$ for every value of the parameters, contrary to $[A_6]$ and $[A_7]$.

Acknowledgements

The author is greatly indebted to C.P. Robert for a number of enlightening conversations and his insightful comments that helped to improve the exposition in the original manuscript. Comments from the Editor and the referees were instrumental in clarifying style and exposition.

Appendix

Proof of Proposition 2.1. The density can be rewritten as

$$f^-(x|a,b) = \frac{b^a \exp(-b)}{\gamma(a,b)} \exp(b(1-x))x^{a-1}\ \mathbb{I}_{x\leq 1}$$

and the term $\exp\{b(1-x)\}$ can be developed in a series expansion, i.e.

$$\exp\{b(1-x)\} = \sum_{k=0}^{+\infty} \frac{b^k}{k!}(1-x)^k$$

Therefore,

$$\begin{aligned} f^-(x|a,b) &= \frac{b^a \exp(-b)}{\gamma(a,b)} \sum_{k=0}^{+\infty} \frac{b^k}{k!}(1-x)^k x^{a-1}\ \mathbb{I}_{x\leq 1} \\ &= \frac{b^a \exp(-b)}{\gamma(a,b)} \sum_{k=1}^{+\infty} \frac{b^{k-1}\Gamma(a)}{\Gamma(a+k)} f_k(x) \end{aligned}$$

where f_k is the density of the beta distribution $\mathscr{Be}(a,k)$.

Proof of Proposition 2.2. We have for every $x \in \mathbb{R}$

$$\frac{f^-(x|a,b)}{g_n(x)} \propto \frac{\exp(-bx)}{\sum_{k=1}^{n} b^{k-1}(1-x)^{k-1}/\Gamma(k)}$$

and

$$\begin{aligned} \left(\frac{f^-(x|a,b)}{g_n(x)}\right)' &\propto \frac{\exp(-bx)}{\left(\sum_{k=1}^{n} b^{k-1}(1-x)^{k-1}/\Gamma(k)\right)^2} \\ &\quad \left(-b\sum_{k=1}^{n} b^{k-1}(1-x)^{k-1}/\Gamma(k) \right. \\ &\quad \left. + \sum_{k=2}^{n} b^{k-1}(1-x)^{k-2}/\Gamma(k-2)\right) \\ &\propto \frac{\exp(-bx)}{\left(\sum_{k=1}^{n} b^{k-1}(1-x)^{k-1}/\Gamma(k)\right)^2} \\ &\quad \left[-b^n(1-x)^{n-1}/\Gamma(n)\right] \leq 0 \end{aligned}$$

Therefore

$$M_n = \sup f^-(x|a,b)/g_n(x) = f^-(0|a,b)/g_n(0)$$
$$= \frac{\sum_{k=1}^{n} b^{a+k-1}/\Gamma(a+k)}{\sum_{k=1}^{n} b^{k-1}/\Gamma(k)} \frac{\Gamma(a)}{\gamma(a,b)}$$

According to Abramowitz and Stegun (1964, formulae 6.5.13, p.262 and 6.5.1, p.260), we have

$$\sum_{k=1}^{n} b^{k-1}/\Gamma(k) = e^b(1-\gamma(n,b)/\Gamma(n))$$

Moreover

$$\sum_{k=1}^{n} \frac{b^{a+k-1}}{\Gamma(a+k)} = \sum_{k=0}^{\infty} \frac{b^{a+k}}{\Gamma(a+k+1)} + \sum_{k=n}^{\infty} \frac{b^{a+k}}{\Gamma(a+k+1)}$$
$$= \sum_{k=0}^{\infty} \frac{b^{a+k}}{\Gamma(a+k+1)} + \sum_{k=0}^{\infty} \frac{b^{n+a+k}}{\Gamma(n+a+k+1)}$$
$$= \frac{\gamma(a,b)e^b}{\Gamma(a)} - \frac{\gamma(a+n,b)e^b}{\Gamma(a+n)}$$

see Abramowitz and Stegun (1964, formula 6.5.29, p.262) for the last equality. We get

$$P_n = \frac{1}{M_n} = \frac{1-\dfrac{\gamma(n+1,b)}{\Gamma(n+1)}}{1-\dfrac{\gamma(a+n+1,b)\Gamma(a)}{\Gamma(a+n+1)\gamma(a,b)}}$$

Proof of Proposition 2.3. Consider $p \in [0.5, 1]$. We want to invert the relation $P_n = p$. Note that

$$P_n \geq \frac{1-\dfrac{\gamma(n,b)}{\Gamma(n)}}{1-\dfrac{\gamma(a+n,b)}{\Gamma(a+n)}}$$
$$\simeq 1 - \frac{\gamma(n,b)}{\Gamma(n)} + \frac{\gamma(a+n,b)}{\Gamma(a+n)}$$
$$\geq 1 - \frac{\gamma(n,b)}{\Gamma(n)}$$
$$\simeq 1 - \Phi\left(\frac{b-n}{\sqrt{n}}\right)$$
$$= \Phi\left(\frac{n-b}{\sqrt{n}}\right)$$

where Φ is the normal probability function.

We obtain the root of the equation $\Phi(x) = p$ by solving

$$\frac{n-b}{\sqrt{n}} = t_p$$

where t_p is $100p\%$ quantile of the normal distribution. Therefore, it is easy to see that the positive root of this equation is

$$n(p) = \frac{1}{4}\left(t_p + \sqrt{t_p^2 + 4b}\right)^2$$

Proof of Proposition 3.1. The density g of Y can be expressed as

$$g(y|a,b) = f^+(y+1|a,b)$$
$$= \frac{b^a e^{-b}}{\Gamma(a,b)} e^{-by}(y+1)^{a-1} \mathbb{I}_{y\geq 0}$$
$$= \frac{b^a e^{-b}}{\Gamma(a,b)} e^{-by} \sum_{k=0}^{a-1} \binom{a-1}{k} y^k \mathbb{I}_{y\geq 0}$$
$$= \sum_{k=1}^{a} \omega_k f_k(x)$$

where

$$\omega_k = \frac{b^{a-k}e^{-b}}{\Gamma(a,b)} \frac{(a-1)!}{(a-k)!}$$

and f_k is the density of the gamma distribution $\mathcal{G}a(k,b)$.

Proof of Proposition 3.2. For $n \in \mathbb{N}$ fixed, we consider the function

$$\phi(\tilde{b}) = \frac{\tilde{b}^n e^{-\tilde{b}}}{\int_{\tilde{b}}^{\infty} e^{-x}x^{n-1}\,\mathrm{d}x}$$
$$= \left(\int_1^{\infty} e^{-\tilde{b}(x-1)}x^{n-1}\,\mathrm{d}x\right)^{-1}$$

Since $\forall x \in [1,\infty]$ the function $\tilde{b} \to e^{-\tilde{b}(x-1)}x^{n-1}$ is non-increasing, ϕ is non-decreasing and the condition $\tilde{b} \leq b$ implies that the optimal choice of $\tilde{b}$ is b. We can rewrite the acceptance probability as

$$P^{C_1} = \frac{\Gamma(a,b)}{\Gamma(n,b)} b^{n-a}$$
$$= \frac{\int_b^{\infty} e^{-x}(x/b)^{a-1}\,\mathrm{d}x}{\int_b^{\infty} e^{-x}(x/\tilde{b})^{n-1}\,\mathrm{d}x}$$

For $x \geq b$ fixed, the sequence $\left((x/b)^{n-1}\right)_n$ is non-decreasing, and hence the following sequence $\left(\int_b^{\infty} e^{-x}(x/b)^{n-1}\,\mathrm{d}x\right)_{n\in\mathbb{N}}$ is a non-decreasing sequence. Therefore the acceptance probability is decreasing in n and the condition $n > a$ implies that the optimal choice of n is $\lceil a \rceil$, the smallest integer larger than a.

Proof of Lemma 3.3. When b is small, the condition $0 < \tilde{b} < b$ implies that $\tilde{b}$ is small too. Therefore, the acceptance probability can be approximated by

$$p = \frac{\Gamma(a,b)}{b^a} \frac{\tilde{b}^n}{\Gamma(n)} e^{a-n} \left(\frac{a-n}{b-\tilde{b}}\right)^{n-a}$$

When we differentiate p with respect to $\tilde{b}$, it is easy to show that p is maximal for $\tilde{b} = nb/a$. Therefore the maximum acceptance probability is

$$P^{C_2} = \frac{\Gamma(a,b)}{\Gamma(n,nb/a)}\frac{n^n}{a^a}e^{a-n}$$

In order to prove that the optimal value in n is the largest integer smaller than a, we consider the sequence

$$u_n = \frac{\Gamma(a,b)}{\Gamma(n,nb/a)}\frac{n^n}{a^a}e^{a-n}$$

and prove that it is increasing for $n < a$.

For every n, we can write

$$u_n = \frac{\Gamma(a,b)e^a a/b}{\int_1^\infty \exp(-n(bx/a-1))(bx/a)^{n-1}\,\mathrm{d}x}$$

To show that u_n is decreasing, it suffices to prove that

$$h(y) = \int_1^\infty \exp(-y(bx/a-1))(bx/a)^{y-1}\,\mathrm{d}x$$

is decreasing. Note that

$$\begin{aligned}\frac{\partial}{\partial y}h(y) &= \int_1^\infty \frac{\partial}{\partial y}\left[\exp(-y(bx/a-1))(bx/a)^{y-1}\right]\mathrm{d}x\\ &= \int_1^\infty \exp(-y(bx/a-1))(bx/a)^{y-1}\\ &\quad (1+\ln(xb/a)-xb/a)\,\mathrm{d}x.\end{aligned}$$

Since the function $1+\ln(z)-z$ is non-positive, the function h is decreasing and the sequence $(u_n)_n$ increasing.

Proof of Lemma 3.4. For large values of b, (C_2) implies that $\tilde{b}$ is large too. Therefore, the acceptance probability can be approximated by (see Abramowitz and Stegun, 1964, 6.5.32, p.263)

$$\begin{aligned}p &\simeq \frac{\Gamma(a,b)}{e^{-\tilde{b}}\tilde{b}^{n-1}}\frac{\tilde{b}^n}{b^a}e^{a-n}\left(\frac{a-n}{b-\tilde{b}}\right)^{a-n}\\ &= \frac{\Gamma(a,b)}{b^a}\tilde{b}e^{\tilde{b}}e^{a-n}\left(\frac{a-n}{b-\tilde{b}}\right)^{a-n}\end{aligned}$$

When we differentiate p with respect to $\tilde{b}$, it is easy to show that the derivative is equal to zero if

$$-\tilde{b}^2 + (b-1-a+n)\tilde{b} + b = 0$$

and the root of interest is

$$\begin{aligned}b_0 &= 1/2\left(b-1+a-n+\sqrt{(b-1-a+n)^2+4b}\right)\\ &\simeq b-a+n\end{aligned}$$

We can consider $b-a+n$ as the asymptotic optimal value of $\tilde{b}$ if $b \geq n-a$ and the acceptance probability is

$$P^{C_2} = \frac{\Gamma(a,b)}{b^a}\frac{(b-a+n)^n}{\Gamma(n,b-a+n)}e^{a-n}$$

In order to obtain the optimal choice of n, we consider the sequence

$$\begin{aligned}v_n &= \frac{\Gamma(a,b)}{b^a}\frac{(b-a+n)^n}{\Gamma(n,b-a+n)}e^{a-n}\\ &= \frac{\int_1^\infty e^{a-bx}x^{a-1}\,\mathrm{d}x}{\int_1^\infty e^{n(1-x)+(a-b)x}x^{n-1}\,\mathrm{d}x}\end{aligned}$$

Consider the function $k(y) = \int_1^\infty e^{y(1-x)+(a-b)x}x^{y-1}\,\mathrm{d}x$. Then

$$\begin{aligned}\frac{\partial}{\partial y}k(y) &= \int_1^\infty \frac{\partial}{\partial y}e^{y(1-x)+(a-b)x}x^{y-1}\,\mathrm{d}x\\ &= \int_1^\infty e^{y(1-x)+(a-b)x}x^{y-1}(\ln(x)+1-x)\,\mathrm{d}x \geq 0\end{aligned}$$

which implies that k is decreasing and the sequence $(v_n)_{n\in\mathbb{N}}$ is increasing. Therefore the condition $n < a$ implies that the optimal choice of n is the largest integer smaller than a.

Proof of Proposition 3.5. (1) When b goes to zero, the acceptance probability $P_0^{C_2}$ converges to

$$p(a) = \frac{\Gamma(a)}{\Gamma(\lfloor a\rfloor)}e^{a-\lfloor a\rfloor}\frac{\lfloor a\rfloor^{\lfloor a\rfloor}}{a^a}$$

To prove that the acceptance probability is superior to $e/4$, we show that for all a, $p(a) \geq e/4$ and all couples (a,b), $P^{C_2}(a,b) \geq p(a)$. First, p is a non-continuous function and it is easy to see that

(a) for $n \geq 1$ and $n \in \mathbb{N}$, $\lim_{\substack{a\to n\\ a\geq n}} p(a) = p(n) = 1$;

(b) for $n \geq 2$ and $n \in \mathbb{N}$, $\lim_{\substack{a\to n\\ a\leq n}} p(a) = e\left(\frac{n-1}{n}\right)^n = \eta_n$.

It is easy to show that the sequence η_n is increasing, therefore, we have

$$\inf_{n\geq 2}\eta_n = \eta_2 = e/4$$

Now, it suffices to show that p is decreasing on all sets $[n, n+1[$, where $n \in \mathbb{N}$. Consider the function

$$g(\epsilon) = p(n+\epsilon) = \frac{\Gamma(n+\epsilon)}{\Gamma(n)}\frac{e^\epsilon n^n}{(n+\epsilon)^{(n+\epsilon)}}$$

Its derivative is

$$g'(\epsilon) = \frac{\Gamma(n+\epsilon)}{\Gamma(n)}\frac{e^\epsilon n^n}{(n+\epsilon)^{(n+\epsilon)}}(\Psi(n+\epsilon)-\ln(n+\epsilon))$$

and we can write (see Abramowitz and Stegun, 1964, 6.3.21, p.259)

$$\begin{aligned}\Psi(n+\epsilon)-\ln(n+\epsilon) &= -\frac{1}{2(n+\epsilon)}\\ &\quad -2\int_0^\infty \frac{t}{(t^2+(n+\epsilon)^2)(e^{2\pi t}-1)}\mathrm{d}t\end{aligned}$$

With this equality, it is clear that g' is negative and g is decreasing. So for all a, $p(a) \geq p(\lceil a\rceil) \geq e/4$.

Now we show that $P^{C_2}(a,b) \geq p(a)$. Note that, for $\epsilon \in [0,1]$ fixed, the function $t(a,\epsilon) = \Gamma(a,\epsilon a)/\Gamma(a)$ is increasing in a. This property implies that for $b \leq a$, we have

$$
\begin{aligned}
P_0^{C_2}(a,b)/p(a) &= \frac{\Gamma(a,b)\Gamma(\lfloor a \rfloor)}{\Gamma(a)\Gamma(\lfloor a \rfloor, b\lfloor a \rfloor/a)} \\
&= \frac{t(a,b/a)}{t(\lfloor a \rfloor, b/a)} \geq 1
\end{aligned}
$$

Hence, for $b < a$, P_2 which is equal to $P_0^{C_2}$ is superior to $p(a)$. Moreover the function $P_\infty^{C_2}$ is an increasing function in b which implies that for $b > a$,

$$P_2(a,b) = P_\infty^{C_2}(a,b) \geq P_\infty^{C_2}(a,a) = P_0^{C_2}(a,a) \geq p(a)$$

Therefore for every (a,b), we have $P_2(a,b) \geq p(a) \geq e/4$.
(2) For b large we have (see Abramowitz and Stegun, 1964, 6.5.32, p.263)

$$\Gamma(a,b) \simeq e^{-b} b^{a-1} \left(1 + \frac{a-1}{b} + \frac{(a-1)(a-2)}{b^2}\right)$$

therefore,

$$
\begin{aligned}
P_2 &= P_\infty^{C_2} \\
&\simeq \frac{b-a+\lfloor a \rfloor}{b} \\
&\quad \frac{1+(a-1)/b+(a-2)(a-1)/b^2}{1+(\lfloor a \rfloor-1)/(n-a+\lfloor a \rfloor)+(\lfloor a \rfloor-2)(\lfloor a \rfloor-1)/(b-a+\lfloor a \rfloor)^2} \\
&\simeq 1-\frac{(a-\lfloor a \rfloor)}{b^2}
\end{aligned}
$$

Thus, the acceptance probability converges to 1 when b goes to infinity.
(3) For every fixed b, $P_2(a,b) > p(a)$. Since p converges to 1 when a goes to infinity, the acceptance probability also converges to 1.

References

Abramowitz, M. and Stegun, I. (1964) *Handbook of Mathematical functions*. Dover, New York.

Dagpunar, J.S. (1978) Sampling of variates from a truncated gamma distribution. *Journal of Statistical Computation and Simulation*, **8**, 59–64.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society* (Ser. B), **39**, 1–38.

Devroye, L. (1985) *Non-Uniform Random Variate Generation*. Springer-Verlag, New York.

Gelfand, A.E., Smith, A.F.M. and Lee, T.M. (1992) Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *Journal of the American Statistical Association*, **87**, 523–32.

Geyer, C.J. and Thompson, E.A. (1992) Constrained Monte Carlo maximum likelihood for dependent data (with discussion). *Journal of the Royal Statistical Society* (Ser. B), **54**, 657–99.

Mengersen, K. and Robert, C.P. (1996) Testing for mixtures: a Bayesian entropy approach. In *Bayesian Statistics 5*, J.O. Berger, J.M. Bernardo, A.P. Dawid, D.V. Lindley and A.F.M. Smith (eds). Oxford University Press, London.

Robert, C.P. (1995) Simulation of truncated normal variables. *Statistics and Computing*, **5,** 121–5.

Robert, C.P. (1996) *Méthodes de Monte Carlo par chaînes de Markov*. Economica, Paris.

Robert, C.P. and Titterington, D.M. (1996) Reparametrisation strategies for hidden Markov models and Bayesian approaches to maximum likelihood estimators. Doc. travail CREST #9624, Insee, Paris.