COMPUTER GENERATION OF INVERSE GAUSSIAN SAMPLES BY ENVELOPE REJECTION

Wen Jau Lee[1], S.H. Ong[2]

[1]Lot 8, Jalan Hi-Tech 2/3, Kulim Hi-Tech Park, 09000 Kulim, Kedah, Malaysia
[2]Institute of Mathematical Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia
*ongsh@um.edu.my (corresponding author)
Received on 2nd January 2011, accepted in revised form 23rd March 2011.

ABSTRACT   The paper proposes a method for computer sampling from the inverse Gaussian distribution with parameters σ, m by envelope rejection with the log-logistic distribution as the envelope distribution for $\sigma \geq m$. The proposed rejection algorithm is shown to have good efficiency.

(**Keywords**: inverse Gaussian, log-logistic distribution, computer sampling)

INTRODUCTION

The inverse Gaussian (IG) distribution with probability density function arises as the first-passage time distribution of a Brownian motion with positive drift. The cumulant generating function (logarithm of the characteristic function) of the IG distribution is the inverse of the cumulant generating function of the Gaussian (normal) distribution. The IG distribution is an important distribution in many applications. A good review of the IG distribution is given by Chhikara and Folks (1989)[1]. See also [2]. Michael et al (1976) [3] proposed a method to generate random variates from the IG distribution by transformations.

$$f(x) = \sqrt{\frac{\sigma}{2\pi x^3}} e^{-\frac{x-m}{2m^2 x}}, \qquad m, \sigma > 0, x > 0,$$

Cheng (1984) [4] described a method for the computer generation of an IG random sample with given sample mean and measure of dispersion. Other than these, there seems limited known literature on this subject.

In this paper we propose an acceptance-rejection (envelope rejection) method for the IG distribution by using the log-logistic distribution as the envelope distribution. Cheng (1977) [5] has considered the log-logistic envelope to develop a rejection algorithm for the gamma distribution. The gamma distribution is a well-known and important distribution with wide applications. Consequently, various methods of generating gamma samples have been proposed in the literature [6] and the interested reader may refer to [7] for further references. The organization of the paper is as follows. Section 2 describes the general rejection method for the computer generation of random variates and Cheng's [5] result for the gamma distribution. The main result of this paper is presented in Section 3 and Section 4 concludes.

RESULTS

The General Envelope Rejection Method

Envelope rejection method for computer sampling from a given distribution is based on a simple probability distribution inequality. Let $f(x)$ be a probability density function (pdf) from which random variates are to be generated. The basic envelope rejection algorithm ([7] p.41) assumes that a dominating pdf $g(x)$ has been chosen, and it is rather easy to generate from $g(x)$ ([7], section 3.2, p.43), such that

$$f(x) \leq Mg(x)$$

where $M \geq 1$ is a known constant. The mean number of iterations needed to generate a single variate x is M while the efficiency of the method is $1/M$. Thus M should be as close to one as possible.

If the ratio $T(x) = f(x)/Mg(x)$ is difficult or time consuming to evaluate, the 'squeeze' technique is often used, that is, easy to compute bounds are found such that

$$\ell_1(x) \le f(x)/Mg(x) \le \ell_2(x).$$

The general acceptance-rejection algorithm is then as follows with U denoting a uniform $(0, 1)$ random number; in symbol, $U \sim U\ 0,1$.

General Envelope Rejection Algorithm

(1) Generate $U \sim U\ 0,1$.

(2) Generate x from $g(x)$.

(3) If $U < \ell_1(x)$ go to (6).

(4) If $U > \ell_2(x)$ go to (1).

(5) If $U > T\ x\ = f(x)/Mg(x)$ go to (1).

(6) Accept x.

We next give Cheng's rejection method for the standard gamma distribution.

Envelope Rejection for the Gamma distribution
An important issue in the implementation of the rejection method is the determination of the easy-to-sample dominating (envelope) pdf $g(x)$ and the scaling constant M. The log-logistic distribution with pdf

$$g(\mu,\lambda;x) = \mu\lambda x^{\lambda-1} / \left(\mu + x^\lambda \right)^2 \mu,\lambda > 0,\ x > 0$$

has an invertible distribution function which facilitates generation of random variates by the inverse transform method. Cheng [5] derived the rejection method for the standard gamma distribution with the log-logistic distribution as envelope. This result is given below.

Let

$$f(\alpha;x) = x^{\alpha-1}e^{-x}/\Gamma(\alpha) \quad \text{and}$$

$$g(\mu,\lambda;x) = \mu\lambda x^{\lambda-1} / \left(\mu + x^\lambda \right)^2 \quad \text{be the}$$

standard gamma and the log-logistic pdf's respectively, where $\alpha > 1$. Then

Result ([5]):

$$f(x) \le M\ g(x) \tag{2.1}$$

where $\mu = \alpha^\lambda$, $\lambda = \sqrt{2\alpha-1}$ and $M = 4\alpha^\alpha e^{-\alpha} \big/ \Gamma\ \alpha\ \sqrt{2\alpha-1}$.

1. ENVELOPE REJECTION FOR THE IG DISTRIBUTION

Following Cheng [5], we use the log-logistic envelope for the IG distribution to obtain the following result.

Result 1
Let the IG and log-logistic pdf's respectively be given by

$$f(\sigma,m;x) = \sqrt{\frac{\sigma}{2\pi x^3}}\, e^{-\frac{(-m)}{2m^2 x}},\ \sigma,m > 0, x > 0,$$

and

$$g(\mu,\lambda;x) = \mu\lambda x^{\lambda-1} / \ \mu + x^\lambda\ ^2,\ \mu,\lambda > 0,\ x > 0.$$

Then $f(\sigma,m;x) \le Mg(\mu,\lambda;x)$

where

$$M = \frac{2e^{-\frac{(\sigma-1)^2}{2mT}} m \sqrt{\frac{2}{\pi}\sqrt{\frac{T^3\sigma^4}{m^3}}}}{T\lambda\sigma},\ \mu = \left(\frac{m}{\sigma T}\right)^\lambda,\ T = \frac{m+\sqrt{m^2+4\sigma^2}}{2\sigma^2} \tag{3.1}$$

and

$$\lambda = 2\sqrt{-\frac{1}{-3+\sqrt{-3-18mT+6T\sigma\sqrt{4+9m^2/\sigma^2}}}}.$$

Proof:
Let

$$T\left(x\right) = \frac{f\left(x\right)}{g\left(x\right)} = \sqrt{\frac{\sigma}{2\pi x^3}} e^{-\frac{x-m}{2m^2 x}} \bigg/ \frac{\mu\lambda x^{\lambda-1}}{\left(\mu+x^\lambda\right)^2} . \tag{3.2}$$

Differentiate (3.2) with respect to x, and equating to 0 gives

$$m^2\left(2x\lambda\left(\left(\frac{1}{x}\right)^\lambda - \mu\right) + \left(\left(\frac{1}{x}\right)^\lambda + \mu\right)\left(-\sigma\right)\right) + x^2\left(\left(\frac{1}{x}\right)^\lambda + \mu\right)\sigma = 0 \tag{3.3}$$

Let

$$x = \frac{m}{\sigma T_0} \tag{3.4}$$

where T_0 is a variable. Therefore,

$$\mu = x^\lambda = \left(\frac{m}{\sigma T_0}\right)^\lambda \tag{3.5}$$

Substitute (3.4) and (3.5) into (3.3) and solve for T_0. Consider the solution

$$T_0 = \frac{m + \sqrt{m^2 + 4\sigma^2}}{2\sigma^2} . \tag{3.6}$$

By equating the modes of $f\left(x\right)$ and $g\left(x\right)$, and after some algebra, we obtain

$$\left(\frac{\lambda-1}{1+\lambda}\right)^{\frac{1}{\lambda}} = \sigma T_0\left(\sqrt{1+\frac{9m^2}{4\sigma^2}} - \frac{3m}{2\sigma}\right) . \tag{3.7}$$

Since there is no closed form solution for λ, a numerical technique is employed to express λ in terms of m and σ. The following series expansion to the fourth order is obtained:

$$\left(\frac{\lambda-1}{1+\lambda}\right)^{\frac{1}{\lambda}} \approx 1 - 2\left(\frac{1}{\lambda}\right)^2 + \frac{4}{3}\left(\frac{1}{\lambda}\right)^4 . \tag{3.8}$$

By applying (3.7) in (3.8), and solving for λ, leads to

$$\lambda = 2\sqrt{-\frac{1}{-3 + \sqrt{-3 - 18mT_0 + 6T_0\sigma\sqrt{4 + 9m^2/\sigma^2}}}} . \tag{3.9}$$

To obtain the scaling constant M, substitute (3.4), (3.5), (3.6) and (3.9) into (3.2) to obtain

$$M = \frac{2e^{-\frac{\left(T\sigma-1\right)^2}{2mT_0}} m\sqrt{\frac{2}{\pi}}\sqrt{\frac{T_0^3\sigma^4}{m^3}}}{T_0\lambda\sigma} \tag{3.10}$$

The ratio $T\left(x\right) = f(x)/Mg(x)$ is given by

$$T\left(x\right) = e^{-\frac{mT_0\sigma-x}{2m^2T_0 x}} T_0 x^{4-\lambda}\left(x^\lambda + \mu\right)^2\left(\frac{\sigma}{x^3}\right)^{3/2} \bigg/ 4\mu\sqrt{\frac{T_0^3\sigma^4}{m}} . \tag{3.11}$$

Since (3.8) is a Laurent series approximation, the scaling constant M is adjusted by a small factor k to improve the envelope. We have

$$M' = \frac{2e^{-\frac{\left(T_0\sigma-1\right)^2}{2mT_0}} m\sqrt{\frac{2}{\pi}}\sqrt{\frac{T_0^3\sigma^4}{m^3}}}{T_0\lambda\sigma} * k$$

where $1.005 \le k \le 1.120$ is determined empirically. It is also shown empirically that $1.13 < M' < 1.16$. Hence the proposed rejection algorithm has good efficiency.

The log-logistic distribution function has a closed form expression

$$G(\mu, \lambda; x) = \frac{x^\lambda}{\alpha^\lambda + x^\lambda}.$$

To generate a log-logistic variate, set

$$G(\mu, \lambda; x) = U = \frac{x^\lambda}{\mu + x^\lambda}$$

and solve to get $x = \left[\mu U / \left(1 - U \right) \right]^{\frac{1}{\lambda}}$ where μ, λ are defined in (3.5) and (3.9).

Figure 1 shows the plots of $f(\sigma, m; x)$ and $M'g(\mu, \lambda; x)$. It is seen that the log-logistic distribution is a good envelope.
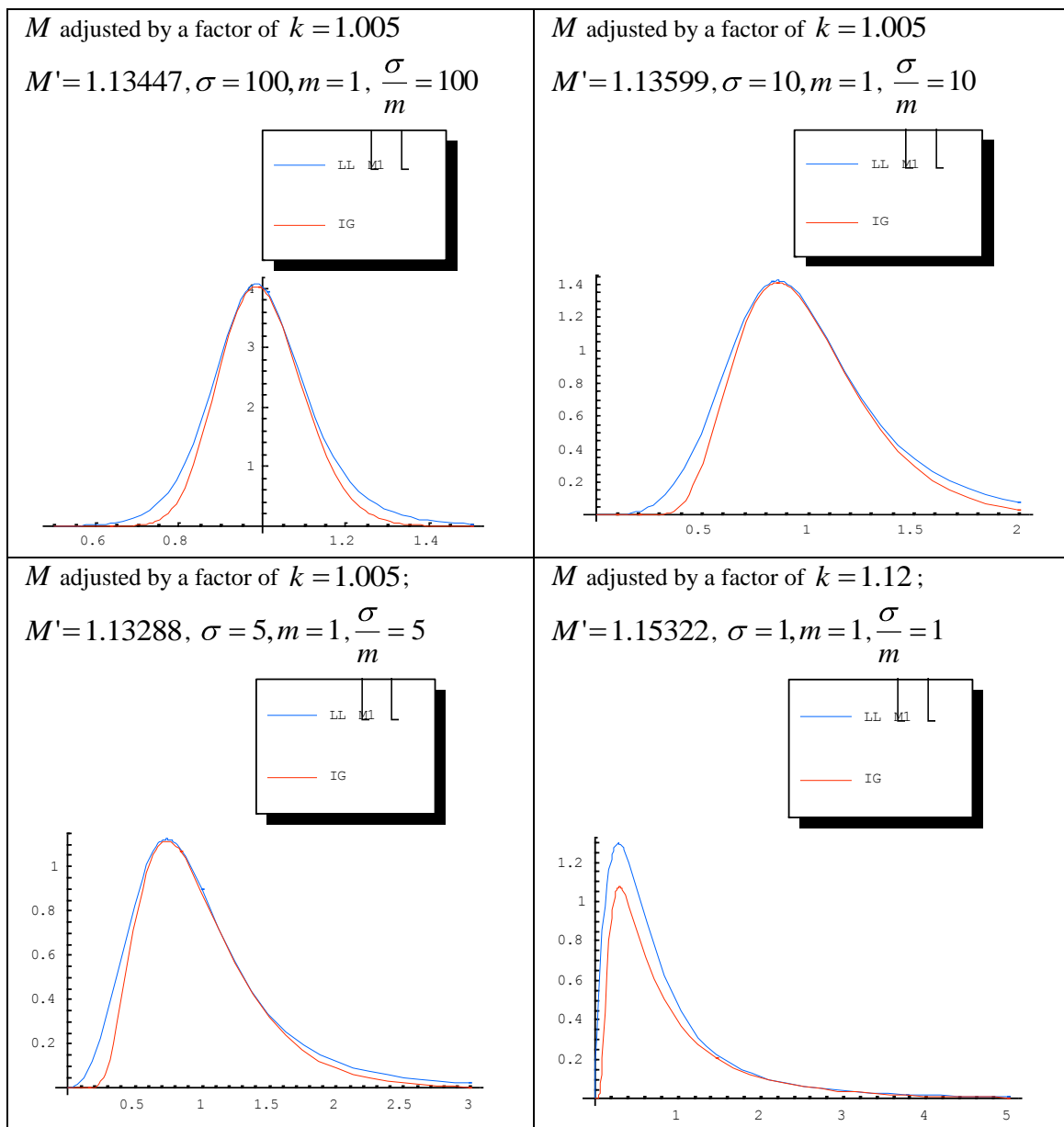

M adjusted by a factor of $k = 1.005$

$M' = 1.13447, \sigma = 100, m = 1, \dfrac{\sigma}{m} = 100$

M adjusted by a factor of $k = 1.005$

$M' = 1.13599, \sigma = 10, m = 1, \dfrac{\sigma}{m} = 10$

M adjusted by a factor of $k = 1.005$;

$M' = 1.13288, \sigma = 5, m = 1, \dfrac{\sigma}{m} = 5$

M adjusted by a factor of $k = 1.12$;

$M' = 1.15322, \sigma = 1, m = 1, \dfrac{\sigma}{m} = 1$

Figure 1 : Plots showing $f(\sigma, m; x)$ and $M'g(\mu, \lambda; x)$. The envelope rejection algorithm for the IG distribution is given as follows with U denoting a uniform (0, 1) random number.

Rejection Algorithm IG

Compute constants and $T(x)$.

(1) Generate $U, U_1 \sim U(0,1)$.

(2) Generate x from log-logistic distribution $G(\mu, \lambda)$ where

$$x = \mu U_1 / \left(1 - U_1\right)^{\frac{1}{\lambda}}.$$

(3) If $U > T(x)$ go to (1).

(4) Accept x.

CONCLUSION

The paper studies the generation of IG variates for case $\sigma \geq m$ through a fast and, efficient rejection method using the log-logistic as the dominating distribution. The method requires a small adjustment to the scaling constant M for tighter envelope. The constant M' is close to 1 and shows that the proposed algorithm is very efficient. The scope of the present paper does not cover for the long tail situations when $\sigma < m$. This will be considered elsewhere.

The result of this paper can be applied to generate random samples from the Poisson-inverse Gaussian distribution which is a mixed Poisson distribution [8]. In the formulation of a mixed Poisson distribution, the Poisson mean is allowed to fluctuate as a random variable. If this random variable has the inverse Gaussian distribution, then this results in the Poisson-inverse Gaussian distribution. Thus to generate variates from the Poisson-inverse Gaussian distribution, we begin by generating values from the inverse Gaussian distribution, and using these values as the Poisson mean, we generate variates from the Poisson distribution. Kemp and Kemp (1990, 1991)[9,10] have given efficient algorithms to generate Poisson random variates. The idea of generating variates by the mixture method has been considered for bivariate discrete distributions [11], and for the negative binomial distribution [12].

REFERENCES

1. Chhikara, R. and Folks, T.L. (1989). The inverse Gaussian distribution: theory, methodology, and applications. New York, Marcel Dekker.

2. Seshadri, V. (1994).The Inverse Gaussian Distribution: A Case Study in Exponential Families. Oxford University Press.

3. Michael, J. R., Schucany, W. R. and Haas, R. W (1976). Generating Random Variates Using Transformations with Multiple Roots. Amer. Statist. 30:88-90.

4. Cheng, R. C. H. (1984). Generation of Inverse Gaussian Variates with Given Sample Mean and Dispersion. Appl. Statist. 33:309-316

5. Cheng, R.C.H. (1977). The Generation of Gamma Variables with Non-integral Shape Parameter. Appl. Statist. 26(1): 71-75.

6. Kundu, D. and Gupta, R.D. (2007). A convenient way of generating gamma random variables using generalized exponential distribution. Computat. Statist. Data Anal. 51:2796 – 2802.

7. Devroye, L. (1986). Non-uniform Random Variate Generation. New York, Springer.

8. Ong, S.H. (1998) A note on the mixed Poisson formulation of the Poisson-inverse Gaussian distribution. Commun. Stat. Simul. Comput. 27:67-78.

9. Kemp, C.D. and Kemp, A.W. (1990). A composition-search algorithm for low-parameter Poisson generation. J. Statist. Comput. Simul. 35:239-244.

10. Kemp, C.D. and Kemp, A.W. (1991). Poisson Random Variate Generation. Appl. Statist. 40(1):143-158.

11. Ong, S.H. (1992). The computer generation of bivariate binomial variable with given marginals and correlation. Commun. Stat. Simul. Comput. 21: 285-299.

12. Ong, S.H. and Lee, W.J. (2008) Computer generation of negative binomial variates by envelope rejection. Computat. Statist. Data Anal. 52: 4175-4183