Random number generation for the generalized normal distribution using the modified adaptive rejection method

Hideo Hirose, Akihiro Todoroki
Department of systems Innovation & Informatics
Kyushu Institute of Technology
Iizuka, Fukuoka 820-8502 Japan

Abstract

When we want to grasp the characteristics of the time series signals emitted massively from electric power apparatuses or electroencephalogram, and want to decide some diagnoses about the apparatuses or human brains, we may use some statistical distribution functions. In such cases, the generalized normal distribution is frequently used in pattern analysts. In assessing the correctness of the estimates of the shape of the distribution function accurately, we often use a Monte Carlo simulation study; thus, a fast and efficient random number generation method for the distribution function is needed. However, the method for generating the random numbers of the distribution seems not easy and not yet to have been developed. In this paper, we propose a random number generation method for the distribution function using the the rejection method. A newly developed modified adaptive rejection method works well in the case of log-convex density functions.

Keywords: modified adaptive rejection method, exponential distribution, generalized normal distribution, envelop function, log-convex, log-concave

1   Introduction

When we want to grasp the characteristics of the time series signals emitted massively from electric power apparatuses or electroencephalogram, and want to decide some diagnoses about the apparatuses or human brains, we may use some statistical distribution functions. For example, we take a distribution pattern in certain time zone by making average in the zone, and we see the differences among those zones by discriminating the distribution patterns zone by zone. In such cases, the generalized gamma distribution or the generalized normal distribution is used [1,4]. In the cases of pattern recognition, the generalized normal distribution is frequently used in pattern analysts rather than the generalized gamma distribution which is preferred by statisticians. Here, we use the former case. The parameter estimation method for the generalized gamma distribution is seen in [5].

In estimating the shape (or the pattern) of the distribution function accurately, e.g., by using the maximum likelihood estimation method, quite a few data cases may be required for appropriateness of the estimation method as well as the real data cases. For a Monte Carlo simulation study, a fast and efficient random number generation method is needed. When only the density function is expressed in closed form (i.e., the cumulative distribution function is not expressed in closed form), the rejection method for the random number generation is frequently used [2,9]. Therefore, we use this method basically. The density function of the distribution shows two patterns: one is log-concavity and the other is log-convexity; in the former case a fundamental rejection method works, but in the latter case, because of difficulty in deciding the envelope function, some alternative method is required. The modified adaptive rejection method proposed here resolves this difficulty.

2 The generalized normal distribution

The density function of the generalized normal distribution is expressed as:

$$f(x|\mu,\sigma,\beta) = \frac{w(\beta)}{\sigma} \exp\left[-c(\beta)|\frac{x-\mu}{\sigma}|^{\frac{2}{1+\beta}}\right], \tag{1}$$

where,

$$c(\beta) = \left[\frac{\Gamma[3(1+\beta)/2]}{\Gamma[(1+\beta)/2]}\right]^{\frac{1}{1+\beta}}, \quad w(\beta) = \frac{\Gamma[3(1+\beta)/2]^{1/2}}{(1+\beta)\Gamma[(1+\beta)/2]^{3/2}}, \tag{2}$$

and

$$\sigma > 0, \quad -\infty < x < \infty, \quad \beta > -1, \Gamma[\cdot]: \text{ gamma function.}$$

In this paper, we deal with the random number generation in the case of $\mu = 0$ and $\sigma = 1$.

When $|\beta| < 1$, the shape of the density function is bell-shaped, thus the rejection method works very well if we select the exponential distribution function as the envelope function. When $\beta > 1$, the shape of the density function is reverse-J-shaped when $x \geq 0$. In this case, it is easily proved that the exponential density function can no longer cover the density function of the generalized normal distribution; the shape of the logarithmic transformed density function of the generalized normal distribution when $\beta > 1$ shows the log-convexity. Typical density functions are shown in Figure 1.
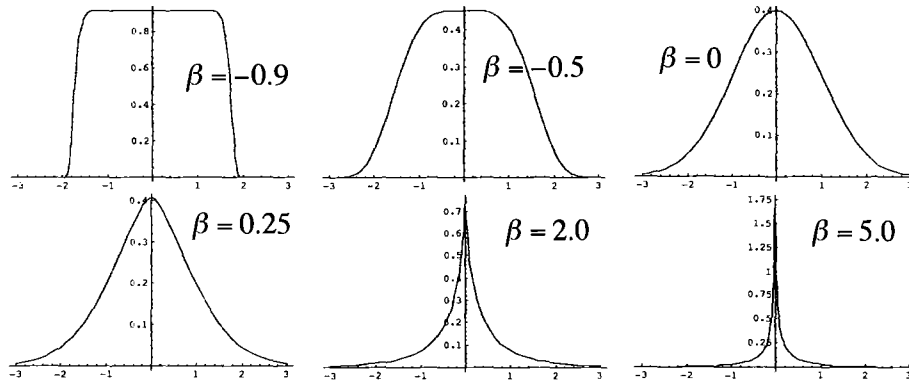

Figure 1: Typical density functions of the generalized normal distribution

3 Random number generation using the rejection method

It is easily seen by the following formula of Bayes theorem how the rejection method works in generating the random numbers.

$$\begin{aligned}
P(x|u \leq f(x)/\{cg(x)\}) &= \frac{P(x|u \leq f(x)/\{cg(x)\}|x)g(x)}{\int P(x|u \leq f(x)/\{cg(x)\}|x)g(x)dx} \\
&= \frac{f(x)/\{cg(x)\}g(x)}{\int [f(x)/\{cg(x)\}g(x)]dx} = \frac{f(x)}{\int f(x)dx} = f(x)
\end{aligned} \tag{3}$$

where, $f(x)$ and $g(x)$ are the garget density function and the covering envelop function respectively; u denotes the uniform random nnumber.

3.1 In the case of $|\beta| < 1$

If some envelope function can easily be determined for the target density function, an ordinary rejection method is expected to work. The exponential distribution, $k\lambda \exp(-\lambda x)$, is a primary candidate for the envelope function as seen in Figure 2. By using a quite neat optimization method developed, we could find the optimal values of the parameters in the exponential distribution function as shown in table 1 [6].
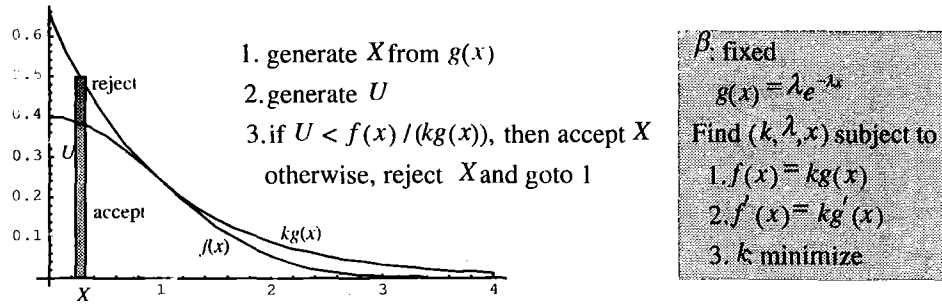
Figure 2: The envelope function for the rejection method when $|\beta| < 1$

Table 1: Optimal parameter values for the envelope function of the exponential distribution.

β	k	λ	β	k	λ
-0.90	2.28660	0.65656	0.10	1.26997	1.03684
-0.80	2.05364	0.70596	0.20	1.22887	1.07448
-0.70	1.88653	0.74766	0.30	1.19151	1.11301
-0.60	1.75679	0.78582	0.40	1.15735	1.15253
-0.50	1.65152	0.82218	0.50	1.12596	1.19310
-0.40	1.56359	0.85768	0.60	1.09699	1.23481
-0.30	1.48855	0.89290	0.70	1.07013	1.27771
-0.20	1.42347	0.92820	0.80	1.04515	1.32187
-0.10	1.36628	0.96384	0.90	1.02183	1.36735
0.00	1.31549	1.00000			

3.2 In the case of $\beta > 1$

When $\beta > 1$, a totally new method is needed; we propose here to use the adaptive rejection method to generate the random number which is originally proposed by Gilks [3] for log-concave distributions. Since the density function is no longer log-concave when $\beta > 1$, we modify the adaptive rejection method to be applicable in the case of log-convex [7,8].

The algorithm for the log-concave distribution is:
1) Select k nodes, x_1, x_2, $\cdots$, x_k, in the support domain T_k of the target density function.
2) Determine $u_k(x)$, $J_k(x)$, $S_k(x)$ such that:
$$u_k(x) = h(x_j) + (x - x_j)h'(x_j)$$
$$S_k(x) = \exp[u_k(x)]/\int_D \exp[u_k(x')]dx'$$
$$l_k(x) = \{(x_{j+1} - x)h(x_j) + (x - x_j)h(x_{j+1})\}/(x_{j+1} - x_j). \quad (l_k(x) = -\infty \quad (x < x_1, x > x_k))$$
3) Generate a random number X of $S_k(x)$, and a uniform random number W,

   a) if $W \leq \exp(l_k(X) - u_k(X))$, then accept X, and go to 2), otherwise,

   b) if $W \leq \exp(h(X) - u_k(X))$, then accept X, and include X in T_k as x_{k+1} (T_k becomes T_{k+1}), then go to 2), otherwise.

   c) reject X, and go to 2).

We modify the algorithm to suit to log-convex distribution. The algorithm for the log-convex distribution is:
2') Determine $u_k(x)$, $J_k(x)$, $S_k(x)$ such that:
$$S'_k(x) = \exp[l_k(x)]/\int_D \exp[l_k(x')]dx' \quad (l_k(x) = h(x) \quad (x > x_k))$$
3') Generate a random number X of $S'_k(x)$, and a uniform random number W,

   a) if $W \leq \exp(u_k(X) - l_k(X))$, then accept X, and go to 2), otherwise,

   b) if $W \leq \exp(h(X) - l_k(X))$, then accept X, and include X in T_k as x_{k+1} (T_k becomes T_{k+1}), then go to 2), otherwise,

   c) reject X, and go to 2). Other parts are the same as in log-concave case.

The envelope functions for log-concave and log-convex distribution functions are illustrated in Fugure3. We can see that a density function of the exponential distribution cannot cover the logarithmic transformed density function of the generalized normal distribution function.
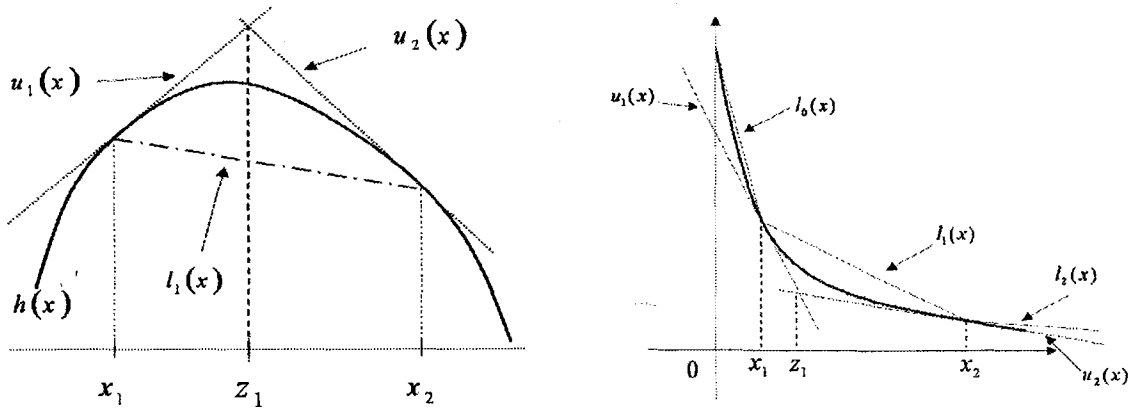

Figure 3: The envelope functions for log-concave and log-convex density functions

The shapes of the simulated distribution function is diagnosed using the Kolmogorov-Smirnov test, and we can determine that the random numbers follow the specified distributions with a certain significance level. Histograms when $\beta = 2$ and $\beta = 0$, in typical cases, agree well with the theoretical density shapes, which are shown in Figure 4.
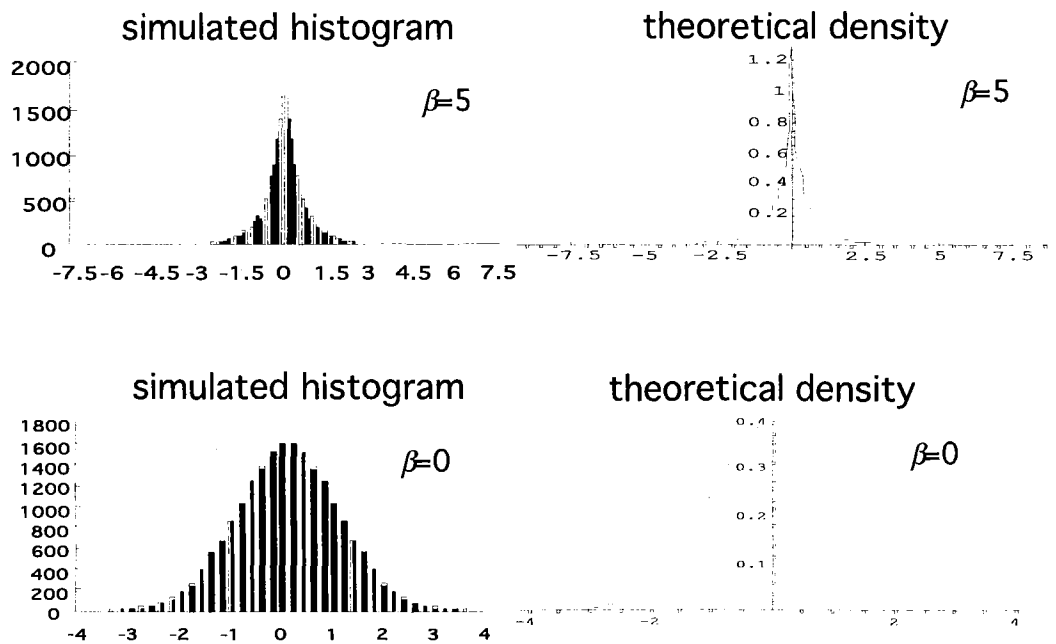

Figure 4: The density functions and its simulated histograms

4   Computing time

The simple rejection method when $|\beta| \leq 1$ uses the exponential density function as the envelop function. Although the computing time for the exponential density function is short, the simple rejection method

requires the longer computing time for the evaluation of the generalized normal density function than that for the exponential density function in each random number generation. The computing time for one random number generation is dominantly determined by the computing time for the generalized normal density function divided by the rejection rate. The rejection rate is defined here by the mean number of rejection times for one random number generation.

In using the adaptive rejection method, the computing time is dominantly determined by the computation for the covering and squeezing functions because the rejection rate becomes smaller and smaller as the number of random numbers is increasing. Figure 5 shows the update region, which will be, as the number of the random number generation becomes larger, smaller.
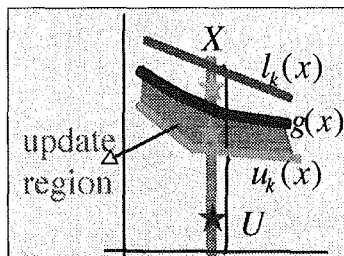

Figure 5: The update region in the adaptive rejection method

Table 2 shows the computing time for one random number generation in the simple rejection method and in the adaptive rejection method. We used a usual Windows Machine of 2.8 GHz clock. Only when the number of random numbers is smaller than 10, the rejection method is comparative to the adaptive rejection method; in almost all the cases, the adaptive rejection method shows the superiority over the simple rejection method for the computing time. For example, in the case of the normal distribution ($\beta = 0$) with 10,000 random numbers generation, the computing time in the adaptive rejection method is about 0.522 microsecond with only 0.10% rejection rate, while the computing time in the simple rejection method is about 1.35 microsecond with 31% rejection rate. In the table, it can be seen that the rejection rate remains a constant value even if the number of random number generation becomes larger in the simple rejection method, while the rejection rate becomes smaller as the number of random number generation becomes larger. It is recommended to generate the random numbers all at once in the adaptive rejection method rather than to generate each time with small amount of random numbers. When the number of generated random numbers is large enough, the computing time in the modified adaptive rejection method is expected to be less than half of the computing time in the simple rejection method.

5 Conclusion

The generalized normal distribution shows a variety of shapes of the density function from log-concave to log convex. Because the cumulative distribution function cannot be expressed by a closed form, some iterative method such as the rejection method is needed to generate the random numbers. When the absolute value of the shape parameter is smaller than or equal to one, the conventional simple rejection method works with the exponential density function as the covering function. When the shape parameter is larger than one, however, the density function of it cannot be covered by the typical exponential density function. In such log-convexity density function, a modified adaptive rejection method is found to be useful. The larger the number of random numbers, the more efficient in generating each random number. When the number of generated random numbers is large enough, the computing time in the modified adaptive rejection method is expected to be less than half of the computing time in the simple rejection method. The developed method is applicable in the Monte Carlo simulation use.

Table 2: Computing time in the adaptive and simple rejection methods.

β	n	adaptive rejection rate	adaptive rejection time	simple rejection rate	simple rejection time
	10	54.0	1.42	79.0	1.95
	100	7.50	0.612	64.0	1.60
−0.50	1000	1.50	0.531	64.0	1.61
	10000	0.43	0.523	65.0	1.60
	100000	0.10	0.521	64.0	1.62
	10	10.0	1.51	39.0	1.49
	100	4.50	0.610	33.0	1.39
0.00	1000	1.01	0.527	30.0	1.36
	10000	0.40	0.528	31.0	1.37
	100000	0.10	0.522	31.0	1.35
	10	9.00	1.50	9.0	1.47
	100	7.20	0.612	13.0	1.23
0.50	1000	2.20	0.529	12.0	1.21
	10000	0.43	0.524	14.0	1.22
	100000	0.10	0.525	12.0	1.22
	10	68.0	1.63		
	100	11.0	0.659		
2	1000	3.00	0.612		
	10000	0.46	0.590		
	100000	0.09	0.589		
	10	93.0	1.55		
	100	29.4	0.731		
5	1000	4.57	0.654		
	10000	0.73	0.641		
	100000	0.11	0.650		
	10	1628	1.02		
	100	288	0.888		
15	1000	26.8	0.740		
	10000	3.38	0.733		
	100000	0.34	0.723		

rate: [%] number of rejections for one random number, time: [microsecond].

References

[1] Chang, S.G., Spatially adaptive wavelet thresholding with context modeling for image denoising, *IEEE Transactions on Image Processing*, **9**, pp. 1523-1531 (2000).

[2] Fushimi, M., *Random number*, University of Tokyo Press (1989).

[3] Gilks, W.R., Adaptive rejection sampling for Gibbs sampling, *Applied Statistics*, **41**, 377-348 (1992).

[4] Mallat, S.G., A theory for multiresolution signal decomposition - The wavelet presentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, pp. 674-693 (1989).

[5] Hirose, H.: Maximum likelihood parameter estimation by model augmentation with applications to the extended four-parameter generalized gamma distribution, *Mathematics and Computers in Simulation*, **54**, pp.81-97 (2000)

[6] Hirose, H. and Todoroki, A., Random number generation method for the generalized normal distribution, *Proceedings of Convention of IEEJ in Kyushu*, No.10-2A-10 (2003.9).

[7] Todoroki, A. and Hirose, H., Random number generation for the generalized normal distribution using the adaptive rejection method, *Proceedings of 18th Annual Meeting of the Japanese Society of Computational Statistics*, pp.97-100 (2004.5)

[8] Todoroki, A. and Hirose, H., Random number generation for the generalized normal distribution using the rejection method, *Proceedings of the Third International Conference on Information (Info2004)*, Tokyo, pp.43-46 (2004.11-12)

[9] Ross, S., *Simulation*, 3rd edi., Academic Press (2002).