http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 9 (1): 61-66, 2010
ISSN 1812-5638

Parameter Estimation of the Extended Generalized Gaussian Family Distributions using Maximum Likelihood Scheme

June-Yule Lee
Department of Marine Engineering, National Kaohsiung Marine University,
482, Chung-Chou 3rd Road, Chi-Jin District, Kaohsiung 805, Taiwan

Abstract: An extended generalized Gaussian distribution which can describe a family of symmetric and asymmetric distributions is considered. Parameter estimation of this function using maximum likelihood scheme is proposed. By measured the tail length and skewness of the observed data, the method integrates a pre-calculated table of initial values for parameters estimation. This allows a fast convergence of the presented model for real-time applications. The simulation results also show that the proposed scheme is an asymptotically unbiased estimator in terms of Cramer- Rao lower bound criterion.

Key words: Generalized Gaussian distribution, maximum likelihood, parameter estimation, Cramer-Rao lower bound

INTRODUCTION

In the field of science and engineering, the family of the Generalized Gaussian Distributions (GGD) are applied to model and detect the Gaussian and non-Gaussian signals (Miller and Thomas, 1972; Kassam, 1988). For example, in natural environments, the performance of communication systems can significantly degrade due to the atmospheric noise caused by the lightning activity in local and distant storm, or the ambient noise in the ocean. In seismic research (Walden, 1985), the non-Gaussian distribution has been used to model the amplitude distribution of the primary reflection coefficients. The aim is to estimate the reflection coefficient sequences as accurately as possible. In image processing, the fundamental work is fitting the image intensity histogram data to gain underlying knowledge of data source. Fan and Lin (2009) developed an expectation maximization algorithm to estimate the parameter of the GGD mixture model that best fit the observed image intensity histogram data. Other useful applications of the GGD also include, video coding (Sharifi and Leon-Garcia, 1995), handwritten character identification (Solihin and Leedham, 1999), speech recognition (Gazor and Zhang, 2003), etc. Therefore, the choice between different families of distribution is important and the estimation of the designed probability density function from observed data is the major task.

In this study, the parameter estimator for the extended generalized Gaussian distribution to include asymmetry is applied. The parameter estimation is determined by the maximum likelihood scheme with iteration method. The key factor of the iteration method depends on the choice of the initial and step values. Soderlind (2006) presented and discussed the approaches to step size selection in control theory and signal processing. Tlelo-Cuautle *et al*. (2007) proposed a procedure to control the step-size, where the step size is determined by the eigenvalues of its state matrixes. For saving the computing cost and improving the accuracy, we proposed a table look-up method.

THE MODEL

The generalized Gaussian distribution to include asymmetry proposed by Nandi and Mampel (1995) is:

$$p(x:\alpha_n,\beta_n,\alpha_p)=\begin{cases} A_p \exp\left(-\left(\dfrac{-x}{\beta_n}\right)^{\alpha_n}\right), & \text{for } x<0 \\ A_p \exp\left(-\left(\dfrac{\alpha_n\Gamma(1/\alpha_p)x}{\beta_n\alpha_p\Gamma(1/\alpha_n)}\right)^{\alpha_p}\right), & \text{for } x\ge 0 \end{cases} \quad (1)$$

and

$$A_p=\frac{\alpha_n}{2\beta_n\Gamma(1/\alpha_n)} \quad (2)$$

where, $\alpha_n>0$, $\alpha_p>0$, $\beta_n>0$ and Γ is gamma function. By setting:

$$\beta_n=\left(\frac{\Gamma(1/\alpha_n)}{\Gamma(3/\alpha_n)}\right)^{1/2} \quad (3)$$

then the variance of Eq. 1 is chosen as one. By changing the parameters in Eq. 1, one can obtain symmetric ($\alpha_n = \alpha_p$) and asymmetric ($\alpha_n \neq \alpha_p$) distributions. For symmetric cases, Fig. 1, (α_n , α_p) are (1,1) and (2,2) represent the Laplace distribution and Gaussian distribution, respectively. When α_n tends to 0 the distribution tends to impulse and when α_n tends to infinite the distribution tends to Uniform distribution. For asymmetric cases, (Fig. 2), the right skew distribution can be modeled by reducing the parameter α_p and as α_p decreases the tail length in the right side increases. Therefore, Eq. 1 provides a wide range of uni-modal distributions with symmetry and asymmetry.

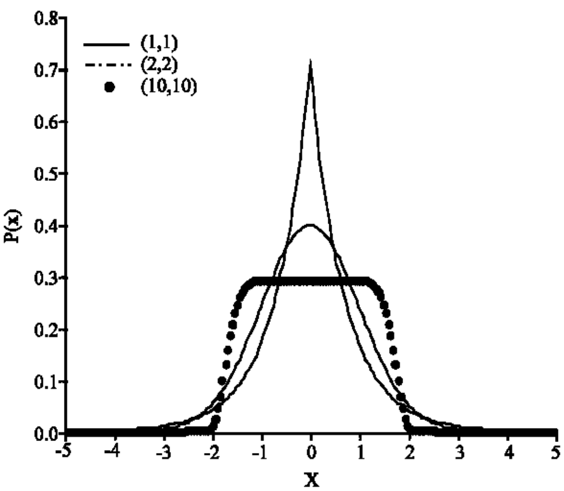

Fig. 1: Symmetric model for varying the parameters of ($\alpha_n = \alpha_p$)

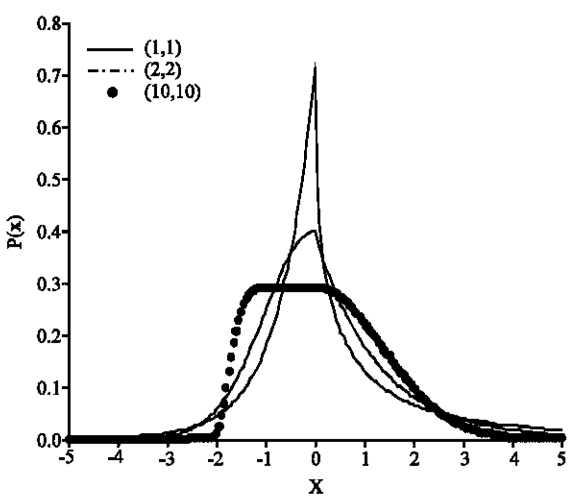

Fig. 2: Asymmetric model for varying the parameters of ($\alpha_n = \alpha_p$)

THE ESTIMATOR

Given a series samples x_i, where i = 1,2,3,...,N, the likelihood function of Eq. 1 is:

$$L = \left(A_p\right)^{N_l} \prod_{x_i<0} \exp\left(-\left(\frac{-x_i}{\beta_n}\right)^{\alpha_n}\right) \times \left(A_p\right)^{N_r} \prod_{x_i \geq 0} \exp\left(-\left(\frac{\alpha_n \Gamma(1/\alpha_p) x_i}{\beta_n \alpha_p \Gamma(1/\alpha_n)}\right)^{\alpha_p}\right) \tag{4}$$

where, N_l (N_r) is the number of x_i samples <0 (≥ 0). Taking the logarithm of Eq. 4, one gets:

$$\ln L = N_l \ln A_p - \sum_{x_i<0}\left(\frac{-x_i}{\beta_n}\right)^{\alpha_n} + N_r \ln A_p - \sum_{x_i \geq 0}\left(\frac{\alpha_n \Gamma(1/\alpha_p) x_i}{\beta_n \alpha_p \Gamma(1/\alpha_n)}\right)^{\alpha_p} \tag{5}$$

If $\partial \ln L/\partial\alpha_n$, $\partial \ln L/\partial\beta_n$ and $\partial \ln L/\partial\alpha_p$ are equate to zero, then the maximum likelihood estimates α_n, β_n and α_p are obtained, i.e.,

$$\frac{\partial \ln L}{\partial \alpha_n} = N_l\left(\frac{1}{\alpha_n} + \frac{\Psi(1/\alpha_n)}{\alpha_n^2}\right) - \sum_{x_i<0}\left(-\frac{x_i}{\beta_n}\right)^{\alpha_n} \ln\left(-\frac{x_i}{\beta_n}\right) + N_r\left(\frac{1}{\alpha_n} + \frac{\Psi(1/\alpha_n)}{\alpha_n^2}\right)$$

$$-\sum_{x_i \geq 0} \frac{\beta_n \alpha_p^2 \left(I\alpha_n x_i\right)^{\alpha_p} \Gamma\left(1/\alpha_n\right)\left(I + \frac{\Psi\left(1/\alpha_n\right)}{\alpha_n}\right)}{\alpha_n \Gamma\left(1/\alpha_p\right)} = 0 \tag{6}$$

$$\frac{\partial \ln L}{\partial \beta_n} = -\frac{N_l}{\beta_n} + \sum_{x_i<0} \frac{1}{\beta_n} \alpha_n \left(-\frac{x_i}{\beta_n}\right)^{\alpha_n} - \frac{N_r}{\beta_n} + \sum_{x_i \geq 0} \frac{1}{\beta_n} \alpha_p \left(I\alpha_n x_i\right)^{\alpha_p} = 0 \tag{7}$$

$$\frac{\partial \ln L}{\partial \alpha_p} = \sum_{x_i \geq 0} \left(I\alpha_n x_i\right)^{\alpha_p} \left(\ln\left(I\alpha_n x_i\right) - \frac{I\beta_n \Gamma\left(1/\alpha_n\right)\left(\Psi\left(1/\alpha_p\right) + \alpha_p\right)}{\Gamma\left(1/\alpha_p\right)}\right) = 0 \tag{8}$$

Where,

$$I = \frac{\Gamma(1/\alpha_p)}{\beta_n \alpha_p \Gamma(1/\alpha_n)} \tag{9}$$

and $\Psi(\bullet)$ is digamma function defined as:

$$\Psi(\bullet) = \frac{\partial \ln \Gamma(\bullet)}{\partial(\bullet)} = \frac{1}{\Gamma(\bullet)} \frac{\partial \Gamma(\bullet)}{\partial(\bullet)} \tag{10}$$

To fit the observed data and approximate the solutions α_n, β_n and α_n, one need to solve (Eq. 6-8). In the following two iterated methods are developed.

Iterated method 1: Given a reasonable initial values α_{n0}, β_n and α_p, (Eq. 6-8) can be solved iteratively for α_n, β_n and α_p. First step, compute Eq. 6 with ($\alpha_{n0} \pm \Delta\alpha_n$) by fixed β_{n0} and

α_{p0}, then α_{n1} is selected by the value of $(\alpha_{n0}\pm\Delta\alpha_n)$ or $(\alpha_{n0}-\Delta\alpha_n)$ so that the likelihood function of Eq. 6 approaches zero. Second step, compute Eq. 7 with $(\beta_{n0}\pm\Delta\beta_n)$ by fixed α_{n1} and α_{p0}, then β_{n1} can be updated by the value of $(\beta_{n0}\pm\Delta\beta_n)$ or $(\beta_{n0}-\Delta\beta_n)$ so that Eq. 7 approaches zero. Third step, compute Eq. 8 with $(\alpha_{n0}\pm\Delta\alpha_n)$ by fixed α_{n1} and β_{n1}, then α_{p1} can be updated by the value of $(\alpha_{n0}+\Delta\alpha_n)$ or $(\alpha_{n0}-\Delta\alpha_n)$ so that Eq. 8 approaches zero. Thus, one iteration is completed for updating α_{n1}, β_{n1} and α_{n1}. This process continues until Eq. 6-8 simultaneously approach zero with some error tolerances. The accuracy of the estimation of this method depends on the choice of the step values $\Delta\alpha_n$, $\Delta\beta_n$ and $\Delta\alpha_p$. Therefore, small step value leads to accuracy of the estimation but requires more computing cost.

Iterated method 2: In this section the gradient method for iteration is developed. In the j-th iteration the Taylor series of Eq. 11 is applied to update the parameters α_{nj}, β_{nj} and α_{pj}.

$$\theta^{(j+1)} = \theta^{(j)} - V(\theta^{(j)})^{-1}\text{grad } \ln L(\theta^{(j)}) \tag{11}$$

where, $\theta = (\alpha_n, \beta_n, \alpha_p)^T$, grad In L = $(\partial \text{In } L/\partial\alpha_n, \partial \text{In } L/\partial\beta_n, \partial \text{In } L/\partial\alpha_p)^T$ and:

$$V(\theta) = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial\alpha_n^2} & \frac{\partial^2 \ln L}{\partial\alpha_n\partial\beta_n} & \frac{\partial^2 \ln L}{\partial\alpha_n\partial\alpha_p} \\ \frac{\partial^2 \ln L}{\partial\beta_n\partial\alpha_n} & \frac{\partial^2 \ln L}{\partial\beta_n^2} & \frac{\partial^2 \ln L}{\partial\beta_n\partial\alpha_p} \\ \frac{\partial^2 \ln L}{\partial\alpha_p\partial\alpha_n} & \frac{\partial^2 \ln L}{\partial\alpha_p\partial\beta_n} & \frac{\partial^2 \ln L}{\partial\alpha_p^2} \end{pmatrix} \tag{12}$$

is a Fisher information matrix. The inverse of the Fisher information matrix plays a key role for the iteration of Eq. 11. Given N samples of the observed data, the Fisher matrix can be obtained from Eq. 12. For each iteration, the expected value of the information matrix is calculated with updating parameters α_{nj}, β_{nj} and α_{pj}. The estimation of the parameter is obtained as the value $V(\theta^{(j)})^{-1}$ converge to zero with some error tolerances.

Cramer-Rao lower bound: It is important to understand the unbiasedness of an estimator. The most common way to decide whether or not an estimator's distribution is unbiased is by looking at its expected value. The Cramer-Rao Lower Bound (CRLB) for the standard deviation of every unbiased estimator is defined as (Hogg and Tanis, 1989):

$$\text{std}(\hat{\alpha}_n) \geq \left\{NE\left[\left(\partial \ln L/\partial\alpha_n\right)^2\right]\right\}^{-1/2} \tag{13}$$

$$\text{std}(\hat{\beta}_n) \geq \left\{NE\left[\left(\partial \ln L/\partial\beta_n\right)^2\right]\right\}^{-1/2} \tag{14}$$

$$\text{std}(\hat{\alpha}_p) \geq \left\{NE\left[\left(\partial \ln L/\partial\alpha_p\right)^2\right]\right\}^{-1/2} \tag{15}$$

where, N is the samples of the every realizations and the E[•] is the expected value. Note that Eq. 13-15 are equivalence to the inverse of the square root of the diagonal elements in Eq. 12.

RESULTS AND DISCUSSION

In the following simulation, the digamma function of Eq. 10 is computed using the algorithm in Char *et al.* (1991) and for further details of the algorithms one can refer to Abramowitz and Stegun (1972). The step value for method 1 is 0.005 and the error tolerance for method 2 is 0.05. All simulation results are based on 50 Monte Carlo runs.

To obtain a good initial value, the tail length and skewness of the observed data are measured and indicated by Q_0 (tailness) and Q_2 (skewness) (Lee, 2003):

$$Q_0 = \frac{\bar{U}(0.05) - \bar{L}(0.05)}{\bar{U}(0.5) - \bar{L}(0.5)} \tag{16}$$

and

$$Q_2 = \frac{\bar{U}(0.05) - m(0.25)}{m(0.25) - \bar{L}(0.05)} \tag{17}$$

where, $\bar{U}(\alpha_R)$ and $\bar{L}(\alpha_L)$ are the means of right α_R and left α_L intervals and $m(\alpha)$ is α-trimmed mean. The Q_0 for Gaussian distribution is 2.58 and for Laplace distribution is 3.30, and Q_2 is 1 for all symmetric cases. Distributions with $Q_2>1$ have relatively longer right tail (right skewness), while those with $0<Q_2<1$ have relatively longer left tail (left skewness). By choosing the parameter values (with $\alpha_n = \alpha_p$ for symmetry and $\alpha_n \neq \alpha_p$ for asymmetry), the Q_0-Q_2 plot (Fig. 3) for various tail length and skewness can be pre-calculated which corresponds to the parameters under the Eq. 1. For example, the values for Gaussian model the (Q_0, Q_2) is (2.58, 1) and the corresponding $(\alpha_n, \beta_n, \alpha_p)$ is (2, 1.14, 2). The results shown in Fig. 3 are various Q_0-Q_2 plots which cover the tailness up to 5.5 and right skewness up to 10. Using the table look-up of Q_0-Q_2 plots and the corresponding $(\alpha_n, \beta_n, \alpha_p)$, a reasonable initial value of estimated parameter can be interpolated.

The samples applied in Table 1 are Gaussian distributions, the results shown both iterated methods are reach the exact values of the parameters with low bias. As the sample size increases the standard deviation of the estimation decreases. It appears that the estimation is asymptotically unbiased and satisfies the Cramer-Rao lower bound, where the standard deviations approach but greater than the CRLB (Fig. 4). The results are shown in Table 2 and Fig. 5 are for the samples of Laplace distributions. They are also in agreement with the theoretical values of the parameters.

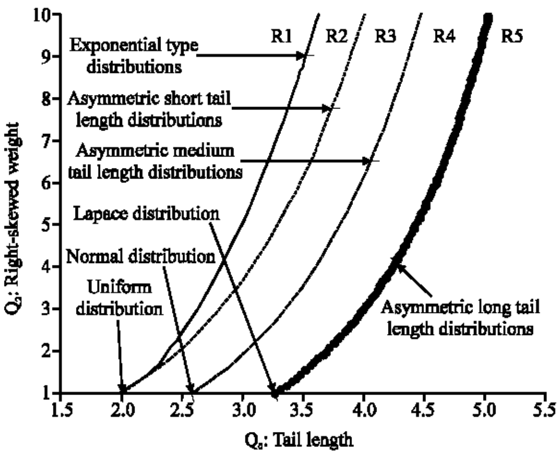

Fig. 3: The Q_0-Q_2 plots

The parameter estimation of extended GGD model include asymmetric distribution is achieved in this study. This differs from the previous study of GGD mode, where more data can be fitted in the proposed model. Furthermore, the pre-calculated table of the Q_0-Q_2 plots save the computing cost and improve the accuracy of the estimation. The key factor of the iteration method depends on the choice of the initial and step values, in this study a table look-up method is provided for fast and accuracy estimation.

Table 1: Gaussian distribution with 50 realizations

Samples	Parameters	α_n	β_n	α_p
	Exact values	2.000	1.414	2.000
128	Method 1	2.121±0.432	1.407±0.092	1.995±0.343
	Method 2	2.154±0.564	1.406±0.122	2.026±0.423
256	Method 1	2.013±0.196	1.409±0.068	2.034±0.326
	Method 2	2.100±0.234	1.404±0.095	2.041±0.383
512	Method 1	2.017±0.215	1.414±0.564	2.034±0.199
	Method 2	2.018±0.261	1.411±0.793	2.034±0.243
1024	Method 1	1.995±0.156	1.405±0.049	1.983±0.150
	Method 2	1.991±0.172	1.402±0.054	1.981±0.170
2048	Method 1	2.026±0.105	1.420±0.027	2.013±0.093
	Method 2	2.028±0.125	1.419±0.037	2.015±0.110
4096	Method 1	2.006±0.071	1.414±0.019	1.997±0.064
	Method 2	2.005±0.086	1.143±0.027	1.999±0.077

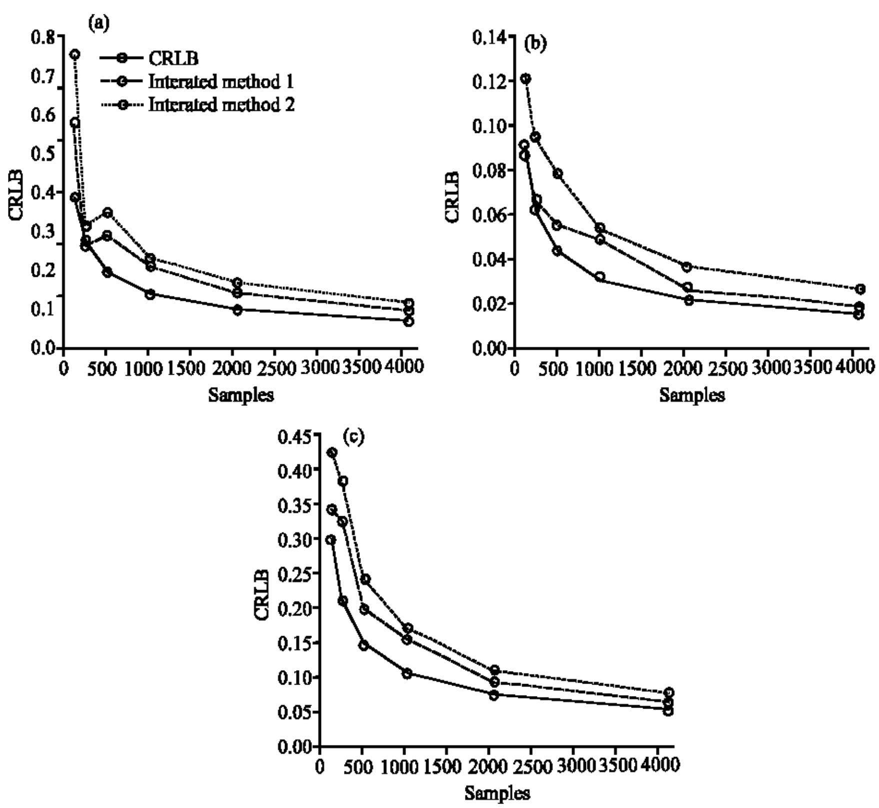

Fig. 4: Standard deviation of the estimated parameters of Gaussian samples

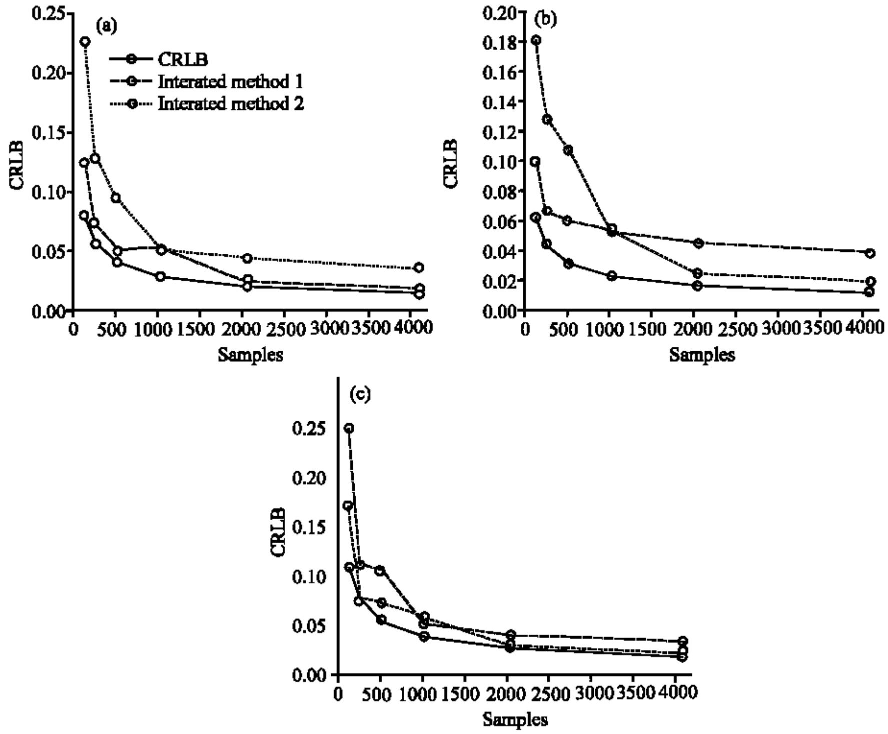

Fig. 5: Standard deviation of the estimated parameters of Laplace samples

Table 2: Laplace distribution with 50 realizations

Samples	Parameters	α_n	β_n	α_p
	Exact values	1.000	0.707	1.000
128	Method 1	1.055±0.125	0.766±0.100	1.087±0.170
	Method 2	1.129±0.224	0.827±0.182	1.154±0.249
256	Method 1	1.008±0.072	0.714±0.066	1.015±0.079
	Method 2	1.018±0.128	0.719±0.128	1.020±0.112
512	Method 1	1.007±0.051	0.708±0.060	0.999±0.074
	Method 2	1.011±0.094	0.710±0.108	1.002±0.107
1024	Method 1	1.008±0.052	0.718±0.054	1.018±0.059
	Method 2	1.006±0.051	0.717±0.052	1.015±0.052
2048	Method 1	1.000±0.024	0.708±0.024	1.002±0.030
	Method 2	0.999±0.044	0.707±0.045	1.002±0.040
4096	Method 1	1.001±0.018	0.708±0.019	1.000±0.023
	Method 2	1.001±0.035	0.707±0.038	0.999±0.035

CONCLUSION

The maximum likelihood estimator is applied to fit three-parameter probability density function. The results showed that the proposed estimator is asymptotically unbiased for generalized Gaussian family distributions. They are useful for modeling and detecting of the observed data in real-time applications.

ACKNOWLEDGMENT

The author would like to acknowledge the valued comments made by Prof. A.K. Nandi.

REFERENCES

Abramowitz, M. and I.A. Stegun, 1972. Handbook of Mathematical Functions. Dover publications, New York, ISBN: 0486612724.

Char, B.W., K.O. Geddes, G.H. Gannet, B.L. Leong, M.B. Monagan and S.M. Watt, 1991. Maple V Manual. Springer-Verlag, New York.

Fan, S.K. and Y. Lin, 2009. A fast estimation method for the generalized gaussian mixture distribution on complex images. Comput. Vision Image Understand., 113: 839-853.

Gazor S. and W. Zhang, 2003. Speech probability distribution. IEEE Signal Proc. Lett., 10: 204-207.

Hogg, R.V. and E.A. Tanis, 1989. Probability and Statistical Inference. 3rd Edn., Maxwell Macmillan, London, ISBN: 0029465001.

Kassam, S.A., 1988. Signal Detection in Non-Gaussian Noise. Springer-Verlag, New York

Lee, J.Y., 2003. Low bias mean estimator for symmetric and asymmetric unimodal distributions. Comput. Method Programs Biomed., 72: 99-107.

Miller, J.H. and J.B. Thomas, 1972. Detectors for discrete-time signals in non-gaussian noise. IEEE Tran. Inform. Theory, 18: 241-250.

Nandi, A.K. and D. Mampel, 1995. An extension of the generalized gaussian distribution to include asymmetry. J. Franklin Inst., 332: 67-75.

Sharifi, K. and A. Leon-Garcia, 1995. Estimation of shape parameter for generalized Gaussian distributions in subb and decompositions of video. IEEE Trans. Circuits Syst. Video Technol., 5: 52-56.

Soderlind, G., 2006. Time-step selection algorithms: Adaptivity, control and signal processing. Applied Numerical Mathematics, 56: 488-502.

Solihin, Y. and C.G. Leedham, 1999. Integral ratio: A new class of global thresholding techniques for handwriting images. IEEE Trans. Pattern Anal. Mach. Intell., 21: 761-768.

Tlelo-Cuautle, E., J.M. Munoz-Pacheco and J. Martinez-Carballido, 2007. Frequency scaling simulation of Chua's circuit by automatic determination and control of step-size. Applied Mathematics Comput., 194: 486-491.

Walden, A.T., 1985. Non-gaussian reflectivity, entropy and deconvolution. Geophysics, 50: 2862-2888.