Estimating the Parameters of a Generalized Exponential Distribution

Syed Afzal Hossain

Department of Mathematics and Statistics

University of Nebraska at Kearney

A three-parameter generalized exponential distribution was suggested by Hossain and Ahsanullah [5]. Some important aspects of this distribution in the area of estimation remain unexplored in the earlier works. We discuss here the maximum likelihood (ML) method and the method of moments to estimate the parameters. The sufficient condition for the existence of a unique solution of the parameters obtained by the method of moments is derived. Finally, a popular real life numerical example is illustrated to investigate the application of both methods of estimation and the results obtained therein are compared with four similar well known three-parameter distributions.

1. Introduction

The estimation of parameters and drawing conclusions based on the estimated parameters is one of the important aspects of inferential statistics. The three-parameter gamma and Weibull distributions are commonly used in life time data analysis. These distributions have many desirable statistical properties. For more information on these distributions see Alexander [1] and Jackson [6].

Mudholkar [9] considered a three-parameter exponentiated Weibull distribution. This new family is suitable for modeling data that indicate non-monotone hazard rates and can be adopted for testing goodness of fit of Weibull as a submodel. It is a right skewed unimodal density function. The usefulness and flexibility of the family is illustrated by reanalyzing five classical data sets on bus-motor failures.

Gupta [3] proposed special cases of the exponentiated Weibull and exponentiated exponential models and compared their performances with the two-parameter gamma family and two-parameter Weibull family, mainly through data analysis and computer simulations. For more on exponentiated Weibull, beta Gumbel, and beta exponentiated distribution see Nadarajah [10], Nadarajah and Kotz [11], Nadarajah and Kotz [12], Nassar and Eissa [13], and Raqab and Ahsanullah [14].

A three-parameter generalized exponential distribution was suggested by Gupta and Kundu [4]. This distribution is a particular case of the exponentiated Weibull distribution originally proposed by Mudholkar, Srivastava, and Freimer [9]. Since its distribution function has closed form the

inference based on the censored data can be handled more easily than gamma family. One of its drawbacks being it is less flexible than the other two families for graduating tail thickness.

Recently, Kundu and Gupta [7] introduced the bivariate generalized exponential distribution so that the marginal have generalized exponential distributions. They reanalyzed one set of data and the results have shown that the bivariate generalized exponential distribution provides a better fit than the bivariate exponential distribution.

Hossain and Ahsanullah [5] introduced a new three-parameter generalized exponential distribution that represents a different type of generalization than Gupta and Kundu [4]. This distribution approaches a two parameter exponential distribution when the shape parameter approaches zero, whereas the distribution in Gupta and Kundu [4] approaches two-parameter exponential if the shape parameter approaches one.

In this article we have considered the estimation of parameters of the three-parameter generalized exponential distribution introduced by Hossain and Ahsanullah [5] by using the maximum likelihood estimation and the method of moments. A sufficient condition for the existence of unique solution for the parameters estimated by the method of moments is derived. A real time numerical example is analyzed and compared with four other similar distributions.

The rest of the article is organized as follows: Section 2 introduces the generalized exponential distribution and some its important properties. In section 3 we described the parameter estimation procedure using maximum likelihood method and method of moments. In section 4 we showed

the applications by analyzing a set of real time data and compared the results with similar three-parameter distributions.

2. Generalized Exponential Distribution

2.1 Distribution and density functions

A random variable X is said to have generalized exponential distribution (GE2) if it has the following distribution function

$$F(x) = \begin{cases} 0 & if & x < \delta \\ 1 - \left(1 - \alpha \dfrac{x-\delta}{\sigma}\right)^{\frac{1}{\alpha}} & if & \delta \leq x \leq \delta + \dfrac{\sigma}{\alpha} \\ 1 & if & x > \delta + \dfrac{\sigma}{\alpha} \end{cases}$$

where $x > \delta > 0, \ \alpha > 0, \ \text{and} \ \sigma > 0$.

The corresponding probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{\sigma}\left(1 - \alpha \dfrac{x-\delta}{\sigma}\right)^{\frac{1}{\alpha}-1} & if & \delta < x < \delta + \dfrac{\sigma}{\alpha} \\ 0 & & Otherwise \end{cases}$$

Hossain and Ahsanullah [5] has showed that $f(x)$ is approaches a two-parameter exponential distribution when $\alpha \to 0^+$. Therefore we will limit our discussions on α close to zero. However, later discussions will justify that $0 < \alpha < 1$.

The generalized exponential distribution introduced by Gupta and Kundu [4] is given by

$$F(x) = \begin{cases} 0 & \text{if } x < \delta \\ \left(1 - e^{-\alpha \frac{x-\delta}{\sigma}}\right)^{\alpha} & \text{if } x > \delta \end{cases}$$

The distribution is a two-parameter exponential distribution when $\alpha \to 1$.

To keep a clear distinction between GE2 and the Gupta and Kundu [4], the generalized exponential distribution in Gupta and Kundu [4] will be denoted as GE1.

The graphs of probability density function and distribution function for various values of parameters α, δ, and σ are shown in Figure 1 and Figure 2 respectively.
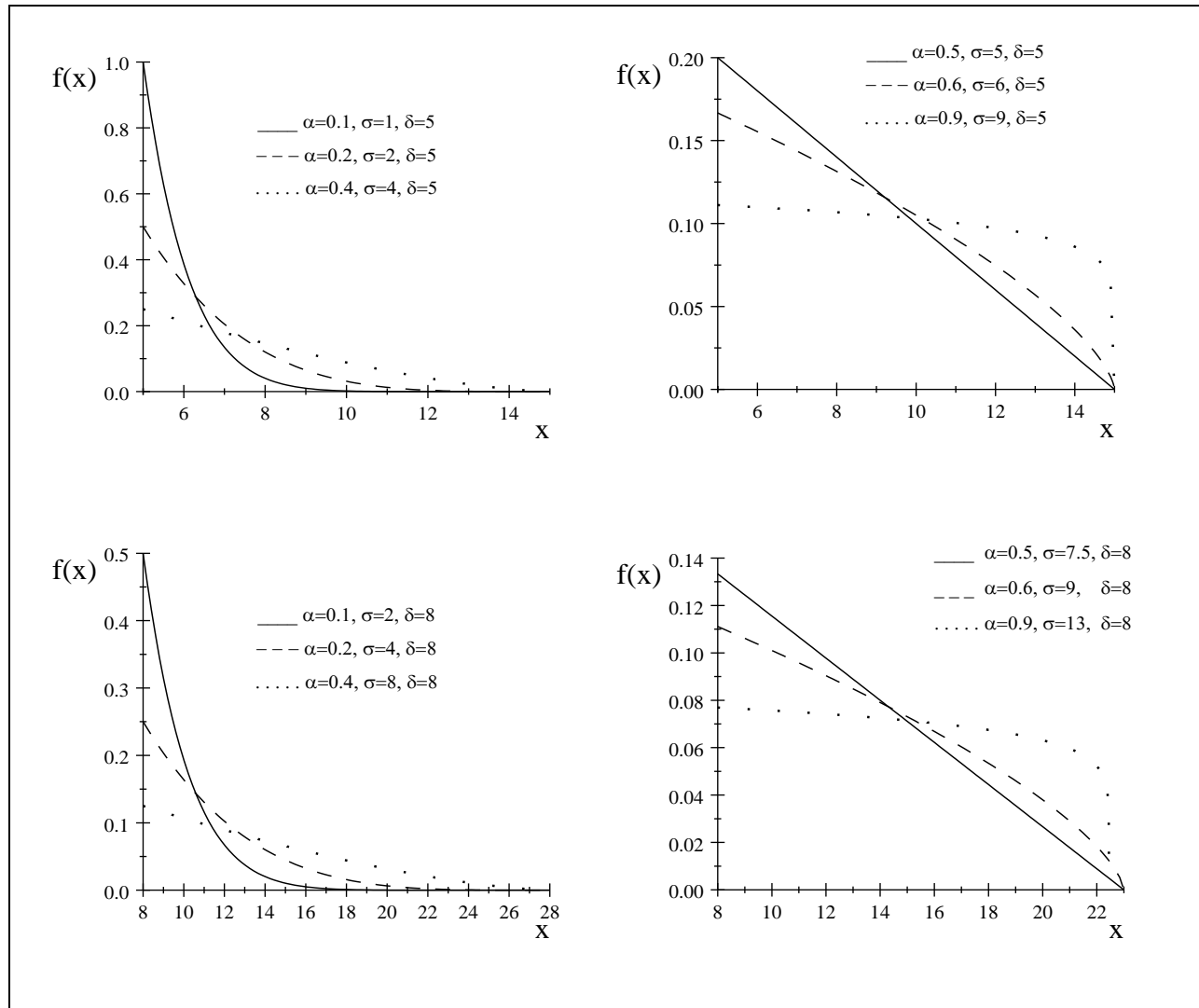
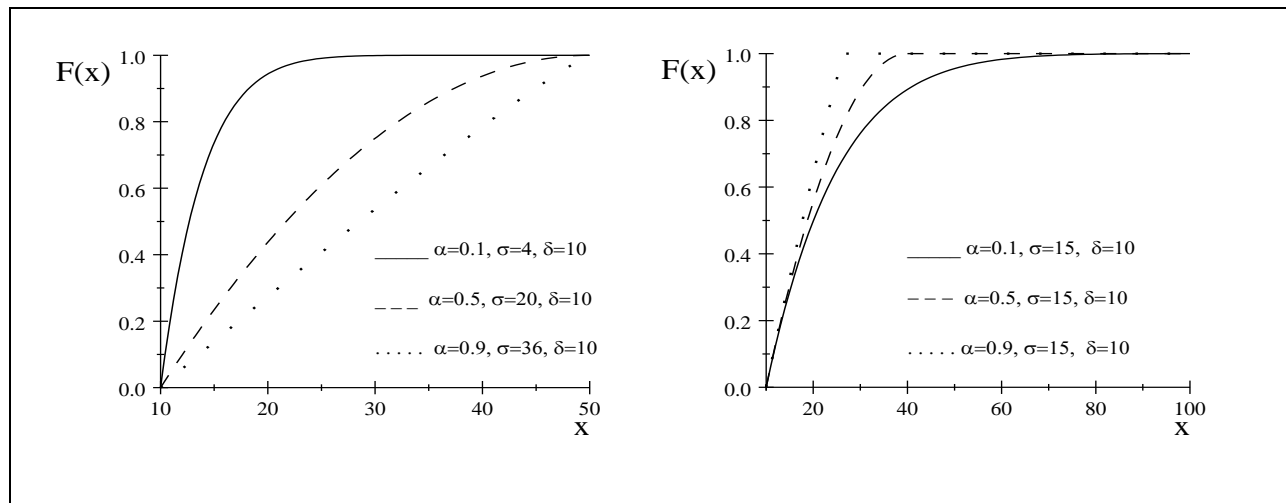Figure 1.   Plots of probability density function for various values of parameters

Figure 2.   Plots of distribution function for various values of parameters

3. Estimation of the Parameters

3.1 Maximum Likelihood Method

In this section we briefly discuss the maximum likelihood estimators, see Bickel and Docksum [2]. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from $GE2(\alpha, \sigma, \delta)$; then the likelihood function, $L(\alpha, \sigma, \delta)$, is given by

$$L(\alpha, \delta, \sigma) = \frac{1}{\sigma^n} \prod_{i=1}^{n} \left(1 - \alpha \frac{x_i - \delta}{\sigma}\right)^{\frac{1}{\alpha} - 1}.$$

Clearly, the maximum likelihood estimate $\hat{\delta}$ of δ is

$$\hat{\delta} = Min\{x_1, x_2, \ldots, x_n\} \qquad \text{if} \qquad 0 < \alpha < 1. \tag{1}$$

We estimate the parameters α and σ by solving the following normal equations

$$\frac{d\ln(L(\alpha, \delta, \sigma))}{d\sigma} = -\frac{n}{\sigma} + \left(\frac{1}{\alpha} - 1\right) \sum_{i=1}^{n} \frac{\alpha \frac{x_i - \hat{\delta}}{\sigma^2}}{1 - \alpha \frac{x_i - \hat{\delta}}{\sigma}} = 0 \tag{2}$$

$$\frac{d\ln(L(\alpha, \delta, \sigma))}{d\alpha} = -\frac{1}{\alpha^2} \sum_{i=1}^{n} \ln\left(1 - \alpha \frac{x_i - \hat{\delta}}{\sigma}\right) - \left(\frac{1}{\alpha} - 1\right) \sum_{i=1}^{n} \frac{\frac{x_i - \hat{\delta}}{\sigma}}{1 - \alpha \frac{x_i - \hat{\delta}}{\sigma}} = 0 \tag{3}$$

It may be arduous to determine the necessary and sufficient conditions for the existence of unique solution to the parameters for nonlinear equations. However, two equations in two variables should be solvable with modern technology unless solution does not exist for the parameters. Note that the MLEs exist only if $0 < \alpha < 1$.

3.2   Method of Moments

The *r*-th raw moment of the generalized exponential distribution, Hossain and Ahsanullah [5], is

given by $\sum_{j=0}^{r} \dfrac{\dfrac{r!}{(r-j)!}\sigma^{j}}{(\alpha+1)(2\alpha+1)\ldots(j\alpha+1)}\delta^{r-j}$, and will be used to determine the raw moments of the

distribution.   If a parametric family has *r* parameters, the moment equations are

$\dfrac{1}{n}\sum_{i=1}^{n} x_i^{j} = \mu_j'$, $j=1,2,\ldots,r$ where $\mu_j' = E\left(X^j \mid \theta\right)$ a function of the unknown parameter vector θ.

The method of moments estimator, see Bickel and Docksum [2], is the solution to these equations.

Theorem 3.1.   *The parameters obtained by the method of moments for a (α,δ,σ) generalized*

exponential distribution to have a unique solution, it is sufficient to show that the sample skewness

is bounded between 0 and 2, that is,

$$0 < \frac{\dfrac{1}{n}\sum (x-\bar{x})^3}{\left(\dfrac{1}{n}\sum (x-\bar{x})^2\right)^{\frac{3}{2}}} < 2$$

Proof.   Equating the sample raw moments with the corresponding population raw moments

(Hossain and Ahsanullah [5]) we obtain the following three equations in terms of the parameters

$\alpha, \delta,$ and σ .

$$\bar{x} = \delta + \frac{\sigma}{\alpha+1} \tag{4}$$

$$\frac{1}{n}\sum x^2 = \delta^2 + \frac{2\delta\sigma}{\alpha+1} + \frac{2\sigma^2}{(\alpha+1)(2\alpha+1)} \tag{5}$$

$$\frac{1}{n}\sum x^3 = \delta^3 + \frac{3\delta^2\sigma}{\alpha+1} + \frac{6\delta\sigma^2}{(\alpha+1)(2\alpha+1)} + \frac{6\sigma^3}{(\alpha+1)(2\alpha+1)(3\alpha+1)} \qquad (6)$$

Solving equations (4) and (5) simultaneously for δ and σ, we obtain

$$\sigma = (\alpha+1)(2\alpha+1)^{\frac{1}{2}}\left(\frac{1}{n}\sum x^2 - \bar{x}^2\right)^{\frac{1}{2}} = (\alpha+1)(2\alpha+1)^{\frac{1}{2}} s \qquad (7)$$

and $\qquad \delta = \bar{x} - (2\alpha+1)^{\frac{1}{2}} s \qquad (8)$

where $s = \left(\frac{1}{n}\sum x^2 - \bar{x}^2\right)^{\frac{1}{2}}$.

Substituting σ and δ from the equations (7) and (8) into the equation (6), we obtain

$$\frac{1}{n}\sum x^3 = \left(\bar{x} - (2\alpha+1)^{\frac{1}{2}} s\right)^3 + \frac{3\left(\bar{x} - (2\alpha+1)^{\frac{1}{2}} s\right)^2\left((\alpha+1)(2\alpha+1)^{\frac{1}{2}} s\right)}{\alpha+1}$$

$$+ \frac{6\left(\bar{x} - (2\alpha+1)^{\frac{1}{2}} s\right)\left((\alpha+1)(2\alpha+1)^{\frac{1}{2}} s\right)^2}{(\alpha+1)(2\alpha+1)} + \frac{6\left((\alpha+1)(2\alpha+1)^{\frac{1}{2}} s\right)^3}{(\alpha+1)(2\alpha+1)(3\alpha+1)}.$$

On further simplification, we obtain

$$(2\alpha+1)^{\frac{1}{2}}(1-\alpha) - (3\alpha+1)K = 0 \qquad (9)$$

where $K = \left(\dfrac{\dfrac{1}{n}\sum x^3 - \bar{x}^3 - 3s^2\bar{x}}{2s^3}\right)$.

Now it comes down to solving a non-linear equation in only one unknown parameter α. Let

$f(\alpha) = (2\alpha+1)^{\frac{1}{2}}(1-\alpha) - (3\alpha+1)K$. Therefore,

$$f'(\alpha) = -3\alpha(2\alpha+1)^{-\frac{1}{2}} - 3K. \tag{10}$$

Equation (10) shows that $f(\alpha)$ is decreasing in α if $K > 0$. It can be easily shown that $f(0) = 1-K$

and $f(\infty) < 0$. Thus, $f(\alpha)$ has a unique solution in α if $1-K > 0$. Therefore, unique solution for

α exists and can be obtained by solving equation (9) if $0 < K < 1$, that is,

$$0 < \frac{\frac{1}{n}\sum x^3 - \bar{x}^3 - 3s^2\bar{x}}{2s^3} < 1.$$

$$\Rightarrow \quad 0 < \frac{\frac{1}{n}\sum(x-\bar{x})^3}{\left(\frac{1}{n}\sum(x-\bar{x})^2\right)^{\frac{3}{2}}} < 2$$

$$\Rightarrow \quad 0 < \text{Sample Skewness} < 2.$$

The unique solution for σ and δ can be obtained by solving equations (7) and (8) respectively.

This completes the proof.

□

4. Application

4.1 Data Analysis

In this section we analyze a set of data that arose in tests on the endurance of deep grove ball

bearings. They were discussed by Lieblein and Zelen [8], Gupta and Kundu [3], and Gupta and

Kundu [4]. The data are the number of revolutions in millions before failure for each of the 23

ball bearings in the life test and they are –

17.88   28.92   33.00   41.52   42.12   45.60   48.40   51.84   51.96

54.12   55.56   67.80   68.64   68.64   68.88   84.12   93.12   98.64

105.12   105.84   127.92   128.04   173.40

We have fitted four distributions, namely three-parameter Weibull, three-parameter gamma, and

three parameter GE1, and three-parameter GE2 to this data set.

4.1.1 ML estimates

From equation (1) we obtain the ML estimate for location parameter $\hat{\delta} = 17.88$. By solving the

equations (2) and (3) simultaneously we obtain ML estimates $\hat{\alpha} = 0.4341$ and $\hat{\sigma} = 77.33$.

4.1.2 Method of moments estimates

The sufficient condition for the existence of unique solution is given by $\dfrac{\dfrac{1}{n}\sum \left(x - \overline{x}\right)^3}{\left(\dfrac{1}{n}\sum \left(x - \overline{x}\right)^2\right)^{\frac{3}{2}}} = 0.94$.

Therefore, satisfies the condition of Theorem 1 for unique solution of all three parameters. Solving equations (4) and (5) for δ and σ in terms of α, and substituting these in equation (6), we obtain

$$\left(2\alpha+1\right)^{\frac{1}{2}}\left(1-\alpha\right)-\left(3\alpha+1\right)\left(\frac{\frac{1}{n}\sum x^3 - \bar{x}^3 - 3s^2\bar{x}}{2s^3}\right) = 0. \tag{11}$$

Solving equation (11) for α and substituting back into equations (7) and (8) we obtain the estimated parameters: $\hat{\alpha} = 0.2957$, $\hat{\sigma} = 59.9330$, and $\hat{\delta} = 25.9650$.

The following TABLE 1 appends the ML estimates of parameters, chi-squared statistics, and the Kolmogorov-Smirnov (K-S) statistics of four three-parameter distributions and the moments estimate of parameters for GE2. The TABLE 2 shows the observed and the expected frequencies. All the distributions fit well to this popular data set. The moment estimates of parameters of GE2 performed better than the other three-parameter distributions with ML estimated parameters except for GE1. The GE1 distribution fitted marginally better than GE2 for this data set.

TABLE 1

Estimates of parameters for similar three-parameter distributions

	$\hat{\alpha}$	$\hat{\sigma}$	$\hat{\delta}$	χ^2	K – S
Gamma (MLE)	2.7316	0.0441	10.2583	0.950	0.107
Weibull (MLE)	1.5979	0.0156	14.8479	1.321	0.118
GE1 (MLE)	4.1658	0.0314	4.7476	0.675	0.103
GE2(MLE)	0.4341	77.3300	17.8800	4.289	0.112
GE2(Moments)	0.2957	59.9330	25.9650	1.557	0.185

TABLE 2

Observed and expected frequencies

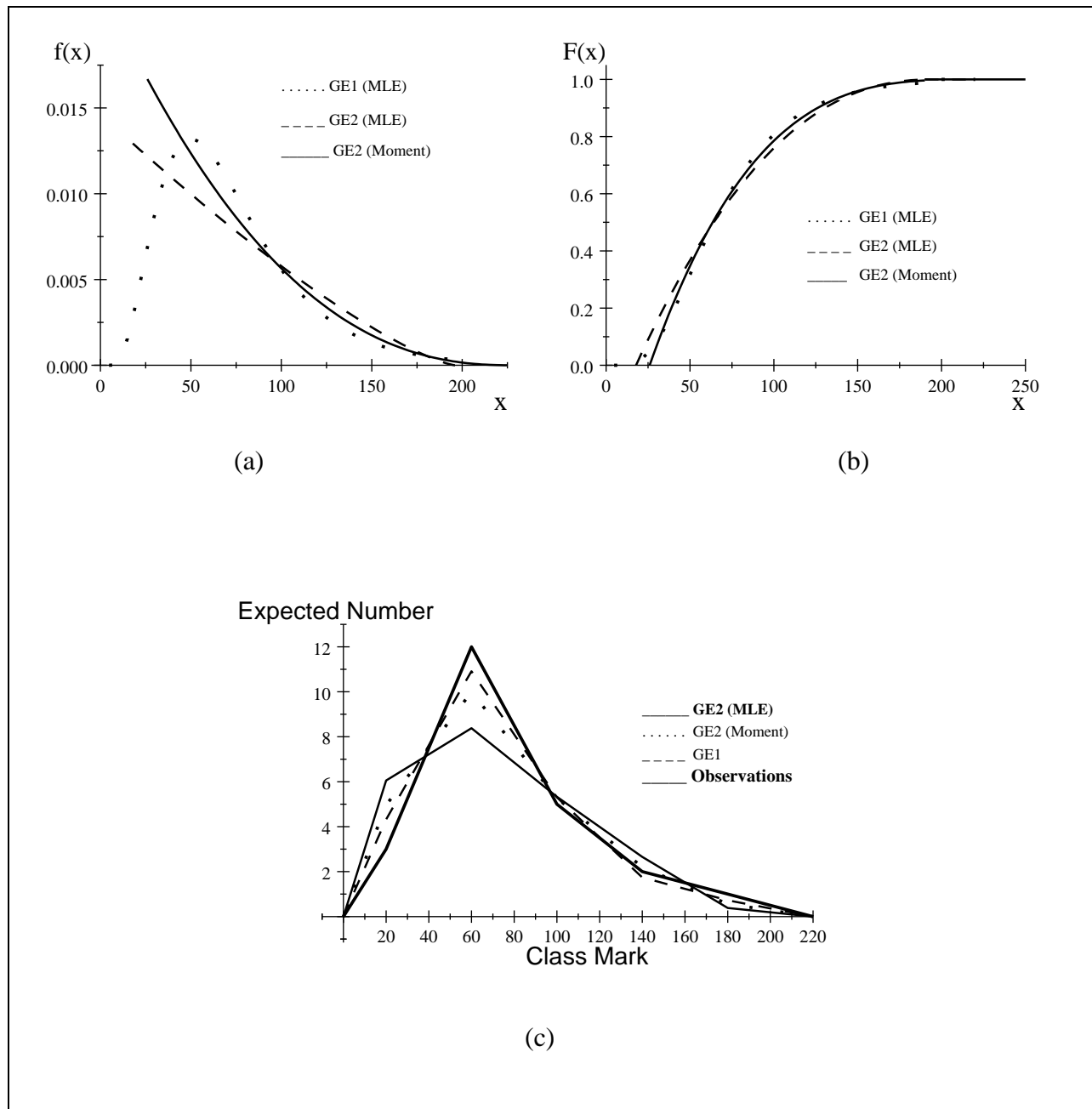| Intervals	Observed	Gamma	Weibull	GE1	GE2(Moments)	GE2(ML)
0 – 40	3	4.493	4.738	4.322	4.956	6. 053
40 – 80	12	10.658	10.016	10.913	9.984	8. 382
80 – 120	5	5.423	5.627	5.303	5.268	5. 331
120 – 160	2	1.805	1.992	1.739	2.201	2. 655
160 – 200	1	0.621	0.627	0.723	0.560	0.381

Figure 3.   Plots of (a) estimated probability density function, (b) estimated distribution function, and (c) expected values under different models and observed values.

5. Conclusion

This paper offers a new family of three-parameter generalized exponential distribution. The distribution can be used as an alternative to analyzing skewed data. Since the distribution function is in closed form, the inference based on the censored data can be handled by this model more easily than the gamma family distributions.

The model is bounded on both sides by positive values and has an increasing hazard rate function. This makes the model suitable for demographers and actuaries to determine the force of mortality for older age groups.

The Generalized exponential distribution (GE2) is a right skewed unimodal density function. Its hazard function is monotonically increasing for $0 < \alpha < 1$. The one real life example used in this article shows that this distribution can quite effectively be used to analyze lifetime data in place of generalized gamma, generalized Weibull and generalized exponential (GE1) distributions. This three-parameter GE2 distribution is always as flexible as the three-parameter GE1, Weibull or gamma distributions.

References

[1]   Alexander, G. N., 1962. "The use of the gamma distribution in estimating the regulated output from the storage" *Trans. Civil Engineering, Inst. of Engineers*, Australia 4, 29 –   34.

[2]   Bickel, P. J. and Docksum, K. A., 2001. *Mathematical Statistics: Basic Ideas and Selected Topics*, 2nd ed., 1, Prentice Hall, New Jersey.

[3]   Gupta, R. D. and Kundu, D., 2001. "Exponentiated exponential family: an alternative to gamma and Weibull distributions." *Biometrical Journal*, 43, 117-130.

[4]   Gupta, R. D. and Kundu, D., 1999. "Generalized exponential distributions." *Australia and New Zealand Journal of Statistics*. 41(2), 173-188.

[5]   Hossain, S. A. and Ahsanullah, M., 2010. "On Generalized Exponential Distributions." *Advances and Applications in Statistical Sciences*. Vol. 4, Issue 1, 1- 22.

[6]   Jackson, O. A. Y., 1969. "Fitting a gamma or log-normal distribution to fiber-diameter measurements of wool tops." *Applied Statist*ics, 18, 70 – 75.

[7]   Kundu, D. and Gupta, R. D., April 2009. "Bivariate generalized exponential distribution." *Journal of Multivariate Analysis*, Volume 100, Issue 4, 581-593.

[8]   Lieblein, J. and Zelen, M., 1956. Statistical investigation of the fatigue life of deep groove ball bearings. J. Res. Nat. Bur. Standards 57, 273 – 316.

[9]   Mudholkar, G. S., Srivastava, D. K. and Freimer, M., 1995. "The exponentiated Weibull family: a reanalysis of the bus motor failure data." *Technometrics* 37, 436-445.

[10]   Nadarajah, S., 2005. "Exponentiated beta distribution." *Computers and Mathematics with Applications*, 49, 1029-1035.

[11]   Nadarajah, S. and Kotz, S., 2006. "The beta exponentiated distribution." *Reliability Engineering and System Safety*, 91, 689-697.

[12]   Nadarajah, S. and Kotz, S., 2004. "The beta Gumbel distribution." *Mathematical Problems in Engineering*. 323-332.

[13]   Nassar, M. M. and Eissa, F. H., 2003. "On the exponentiated Weibull distribution." *Communications in Statistics: Theory and Methods*, 32(7), 1371-1336.

[14]   Raqab, M. Z. and Ahsanullah, M., 2001. "Estimation of the location and scale parameters of generalized exponential distribution based on order statistics." *Journal of Statistical Computation and Simulation*. Vol. 69, 109-123.