Mashhad R. J. Math. Sci., Vol. 2(1)(2009) 69-80

Estimation of $P[Y < X]$ for generalized exponential distribution in presence of outlier*

P. Nasiri[†](⊠)
Department of Statistics, Tehran Payam-e Noor University, Tehran, Iran
and M. Jabbari Nooghabi
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

This paper deals with the estimation of $P(Y < X)$, where Y has generalized exponential distribution with parameters α and λ and X has mixture generalized exponential distribution (or marginal distribution of $X_1, X_2, ..., X_n$, in presence of one outlier with parameters β_1 and β_2) such that X and Y are independent. when the scale parameter (λ) is known the maximum likelihood estimator of $R = P(Y < X)$ is derived. Analysis of a simulated data set has also been presented for illustrative purposes.

Keywords and phrases: Maximum likelihood estimator, outlier, stress-strength model.

AMS Subject Classification 2000: Primary 62G05; Secondary 62E10.

1 Introduction

Recently the two-parameter generalized exponential (GE) distribution has been proposed by many authors. It has been studied extensively by Gupta and Kundu ([11]–[17]), Raqab [26], Raqab and Ahsanullah [27], Zheng [34] and Kundu *et*

*Received: 20 October 2008; Revised: 11 April 2009
†e-mail: pnasiri@hotmail.com

al. [21]. Note that the generalized exponential distribution is a submodel of the exponentiated weibull distribution introduced by Mudholkar and Srivastava [22] and later studied by Mudholkar *et al.* [24] and Mudholkar and Huston [23].

The two-parameter GE distribution has the following density function

$$f(x,\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}, \qquad x>0. \tag{1}$$

and the distribution function

$$F(x,\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha}, \qquad x>0 \tag{2}$$

Here $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively. For different values of the shape parameter, the density function can take different shape. For detail description of the distribution, one is referred to the original paper of Gupta and Kundu [11]. From now one GE distribution with the shape parameter α and scale parameter λ will be denoted by $GE(\alpha,\lambda)$.

Let the random variables $X_1, X_2, ..., X_{n-1}$ are independent, each having the probability density function f(x),

$$f_1(x,\beta_2) = \beta_2 e^{-x}(1-e^{-x})^{\beta_2-1}, \qquad x>0. \tag{3}$$

and the one random variable (As outlier) is also independent, has the probability density function $g(x)$.

$$f_2(x,\beta_1) = \beta_1 e^{-x}(1-e^{-x})^{\beta_1-1}, \qquad x>0 \tag{4}$$

The joint density of $X_1, X_2, ..., X_n$ is given as

$$\begin{aligned} f(x_1,x_2,...,x_n) &= \frac{1}{n}\prod_{i=1}^{n} f(x_i).\sum_{A_1}^{n}\frac{f_2(x_{A_1})}{f_1(x_{A_1})} \\ &= \frac{1}{n}{\beta_2}^{n} e^{-\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}(1-e^{-x_i})^{\beta_2-1}\sum_{A=1}^{n}\frac{\beta_1 e^{-x_{A_1}}(1-e^{-x_{A_1}})^{\beta_1-1}}{\beta_2 e^{-x_{A_1}}(1-e^{-x_{A_1}})^{\beta_2-1}} \\ &= \frac{1}{n}\beta_1{\beta_2}^{n-1} e^{-\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}(1-e^{-x_i})^{\beta_2-1}\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2} \\ ,x>0 \end{aligned} \tag{5}$$

(see [7] and [8])

From equation (5), the marginal distribution of X is,

$$h(x, \beta_1, \beta_2) = \frac{1}{n}\beta_1 e^{-x}(1-e^{-x})^{\beta_1 - 1} + \frac{n-1}{n}\beta_2 e^{-x}(1-e^{-x})^{\beta_2 - 1}, \qquad x > 0 \quad (6)$$

The main purpose of this paper is to focus on the inference of $R = P(Y < X)$, where $Y \sim GE(\alpha, \lambda)$ and X has mixture GE or marginal distribution of $X_1, X_2, ..., X_n$ with presence of one outlier. For simplify we consider $\lambda = 1$. The estimation of R is very common in the statistical literature. For example, if X is the strength of a component which is subject to a stress Y, when R is a measure of system performance and arises in the context of mechanical reliability of a system. We obtain the maximum likelihood estimator (MLE) of R. It may be mentioned here that related problems have been widely used in the statistical literature. The MLE of $P(Y < X)$, when X and Y have bivariate exponential distribution, has been considered by Awad *et al.* [2]. Church and Harris [4], Downtown [6], Govidarajulu [9], Woodward and Kelley [33] and Owen, Craswell and Hanson [25] considered the estimation of $P(Y < X)$, when X and Y are normally distributed. Similar problem for the multivariate normal distribution has been considered by Gupta and Gupta [10]. Kelley, Kelley and Schucany [18], Sathe and Shah [29], Tong [31], [32] considered the estimation of $P(Y < X)$ when X and Y are independent exponential random variables. Constantine and Karson [5] considered the estimation of $P(Y < X)$, when X and Y are independent gamma random variables. Sathe and Dixit [8] have been estimate of $P(Y < X)$ in the negative binomial distribution. Ahmad *et al.* [1] and Surles and Padgett [30] considered the estimation of $P(Y < X)$, where X and Y are Burr Type random variables. Baklizi and Dayyeh [3] have done shrinkage estimation of $P(Y < X)$ in exponential case.

The rest of the paper is organized as follows. In section 2, we derive the MLE of R. Analysis of a real life data set has been presented in section 3 and finally we draw conclusion in section 4.

2 Maximum likelihood estimator of R

Let $Y_1, Y_2, ..., Y_m$ be a random sample for Y with pdf

$$g(y,\alpha) = \alpha e^{-y}(1-e^{-y})^{\alpha-1}, \qquad y > 0 \tag{7}$$

and $X_1, X_2, ..., X_n$ be random sample for X with pdf

$$f(x,\beta_1,\beta_2) = \frac{1}{n}\beta_1 e^{-x}(1-e^{-x})^{\beta_1-1} + \frac{n-1}{n}\beta_2 e^{-x}(1-e^{-x})^{\beta_2-1}, \qquad x > 0 \tag{8}$$

Then

$$\begin{aligned}
R &= P(Y < X) \\
&= \int_0^\infty \int_0^x g(y,\alpha) f(x,\beta_1,\beta_2,) dydx \\
&= \int_0^\infty [\int_0^x \alpha e^{-y}(1-e^{-y})^{\alpha-1} dy] \times \\
&\quad [\frac{1}{n}\beta_1(1-e^{-x})^{\beta_1-1}\frac{n-1}{n}\beta_2 e^{-x}(1-e^{-x})^{\beta_2-1}]dx \\
&= \int_0^\infty \int_0^x [\frac{1}{n}\beta_1 e^{-x}(1-e^{-x})^{\beta_1-1}][\alpha e^{-y}(1-e^{-y})^{\alpha-1}]dydx \\
&+ \int_0^\infty \int_0^x [\frac{n-1}{n}\beta_2 e^{-x}(1-e^{-x})^{\beta_2-1}][\alpha e^{-y}(1-e^{-y})^{\alpha-1}]dydx \\
&= \int_0^\infty \frac{1}{n}\beta_1 e^{-x}(1-e^{-x})^{\alpha+\beta_1-1}dx \\
&+ \int_0^\infty \frac{n-1}{n}\beta_2 e^{-x}(1-e^{-x})^{\alpha+\beta_2-1}dx \\
&= \frac{1}{n}.\frac{\beta_1}{\alpha+\beta_1} + \frac{n-1}{n}.\frac{\beta_2}{\alpha+\beta_2}
\end{aligned} \tag{9}$$

Therefore, the MLE of R becomes

$$\hat{R} = \frac{1}{n}\frac{\hat{\beta}_1}{\hat{\alpha}+\hat{\beta}_1} + \frac{n-1}{n}\frac{\hat{\beta}_2}{\hat{\alpha}+\hat{\beta}_2}$$

Now, to compute the MLE of R, we first consider the joint distribution of $X_1, X_2, ..., X_n$with presence of one outlier in (5), so

$$\begin{aligned}
L(\alpha,\beta_1,\beta_2) &= g(y_1,y_2,...,y_m).f(x_1,x_2,...,x_n) \\
&= \alpha^m e^{-\sum_{i=1}^m y_i}\pi_{i=1}^m(1-e^{-y_i})^{\alpha-1} \\
&\times \frac{1}{n}\beta_1\beta_2^{n-1}e^{-\sum x_i}\Pi_{i=1}^n(1-e^{-x_i})^{\beta_2-1}\sum_{A_1=1}^n(1-e^{-x_{A_1}})^{\beta_1-\beta_2}.
\end{aligned}$$

The Log-likelihood function of the observed sample is

$$\begin{aligned}\ln L(\alpha,\beta_1,\beta_2) &= m\ln(\alpha) - \sum_{i=1}^{m} y_i + (\alpha-1)\sum_{i=1}^{m}\ln(1-e^{-y_i}) \\ &+ \ln[\frac{\beta_1\beta_2^{n-1}}{n} e^{-\sum_{i=1}^{n} x_i}\Pi_{i=1}^{n}(1-e^{-x_i})^{\beta_2-1}\sum_{A_1=1}^{n}(1-e^{-x_i})^{\beta_1-\beta_2}]\end{aligned} \tag{10}$$

The MLE's of α,β_1 and β_2 say $\hat{\alpha},\hat{\beta}_1$ and $\hat{\beta}_2$, respectively, which is obtained as the solutions of

$$\frac{\partial \ln L}{\partial\alpha} = \frac{m}{\alpha} + \sum_{i=1}^{m}\ln(1-e^{-y_i}) = 0$$

or

$$\frac{m}{\alpha} = -\sum_{i=1}^{m}\ln(1-e^{-y_i})$$

Hence

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{m}\ln(1-e^{-y_i})} \tag{11}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial\beta_1} &= \frac{1}{\beta_1} + \frac{\frac{\partial}{\partial\beta_1}\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}}{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}} \\ &= \frac{1}{\beta_1} + \frac{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2} ln(1-e^{-x_{A_1}})^{\beta_1-\beta_2}}{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}} = 0\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial\beta_2} &= \frac{n-1}{\beta_2} + \sum_{i=1}^{n}\ln(1-e^{-x_i}) + \frac{\frac{\partial}{\partial\beta_2}\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}}{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}} \\ &= \frac{n-1}{\beta_2} + \sum_{i=1}^{n}\ln(1-e^{-x_i}) - \frac{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}\ln(1-e^{-x_{A_1}})^{\beta_1-\beta_2}}{\sum_{A_1=1}^{n}(1-e^{-x_{A_1}})^{\beta_1-\beta_2}} \\ &=\end{aligned} \tag{13}$$

Form (13), (14), $\hat{\beta}_1$ and $\hat{\beta}_2$ can be obtained as the solution of non-linear equations.

For $\beta_1 = \beta_2 = \beta$ in homogenous case, $\hat{\alpha}$ and $\hat{\beta}$ can be obtained as

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{m}\ln(1-e^{-y_i})} \qquad \hat{\beta} = \frac{-n}{\sum_{i=1}^{m}\ln(1-e^{-y_i})}. \tag{14}$$

These equations proposed by Kundu and Gupta (2006).

3 Numerical experiments and discussions

In order to have some idea about Bias and Mean Square Error (MSE) of MLE, we perform sampling experiments using Maple.
To have inference about R, we consider the following small sample sizes; (m,n)=

$$(15,15),(20,20),(25,25),(15,20),(20,15),(15,25),(25,15),(20,25),(25,20).$$

Here, we take $\alpha = 1.50$ and $\beta_1 = 2.5$ and $\beta_2 = 2.75$, respectively. As we know, the generated sample size n from $f(x,\beta_1,\beta_2)$, (n-1) sample generated from the equation (3) and one sample generated from the equation (4). All the results are based on 1000 replications. Here we present a complete analysis of a simulated data. The data has been generated using $m = n = 20$, $\alpha = 1.5, \beta_1 = 2.5$ and $\beta_2 = 2.75$.

The data has been truncated after two decimal places and it has been presented below. The Y values are,

0.74	1.41	0.86	0.20	0.72	3.11	0.73	0.44
1.31	0.86	0.27	2.27	0.88	1.32	4.41	1.17
0.86	2.19	0.53	0.08				

and the corresponding X values are,

1.12	5.30	0.65	1.46	1.27	0.74	1.51	0.81
1.79	2.11	1.33	1.50	1.57	1.26	0.49	2.93
0.85	0.85	1.73	1.83				

Now, we obtain the MLE of $\hat{\alpha} = 1.671$, $\hat{\beta}_1 = 0.315$ and $\hat{\beta}_2 = 2.18$. Therefore, $\hat{R} = 0.5457$.

4 Conclusions

In this paper, we have addressed the problem of estimating $P(Y < X)$ for the Generalized Exponential distribution with presence of one outlier, when the scale parameter is know. The results are given in table 1, 2 and 3. It is observed that the maximum likelihood estimator R work quit well. We report the average estimates and the MSEs based on 1000 replications. The results are reported in the following Tables. In this case, as expected when m=n and m, n increase then the average biases and the MSEs decrease . For fixed m as n increases the MSEs decrease and also for fixed n as m increases the MESs decreases.

Table 1

$\alpha = 1.5, \beta_1 = 2.5, \beta_2 = 2.75$

(n,m)	R	$\hat{R}$	Bias $\hat{R}$	MSE $\hat{R}$
(15,15)	0.6455882354	0.3358738948	-0.3097143406	0.1837793490
(20,20)	0.6459558824	0.3335933151	-0.3123625673	0.1882900337
(25,25)	0.6461764706	0.3448594453	-0.3013170253	0.1811570295
(15,20)	0.6455882354	0.3489987431	-0.2965894923	0.1767666919
(20,15)	0.6459558824	0.3433466835	-0.3026091989	0.1804719689
(15,25)	0.6455882354	0.3547599579	-0.2908282775	0.1719679412
(25,15)	0.6461764706	0.3489031863	-0.2972732843	0.1791734199
(20,25)	0.6459558824	0.3373103044	-0.3086455780	0.1854977684
(25,20)	0.6461764706	0.3412694950	-0.3049069756	0.1847561247

Table 2

$\alpha = 1.5, \beta_1 = 2.5, \beta_2 = 2.75$

(n,m)	$\hat{\alpha}$	Bias $\hat{\alpha}$	MSE $\hat{\alpha}$
(15,15)	1.642093877	0.142093877	0.2526986879
(20,20)	1.580778465	0.080778465	0.1554931754
(25,25)	1.566022760	0.066022760	0.1135263151
(15,20)	1.575140968	0.075140968	0.1467095745
(20,15)	1.618209236	0.118209236	0.2182551913
(15,25)	1.563346657	0.063346657	0.1069492224
(25,15)	1.606350563	0.106350563	0.2132163252
(20,25)	1.562139231	0.62139231	0.1176980818
(25,20)	1.581781725	0.081781725	0.1455059822

Table 3

$\alpha = 1.5, \beta_1 = 2.5, \beta_2 = 2.75$

(n,m)	$\hat{\beta}_1$	Bias $\hat{\beta}_1$	MSE $\hat{\beta}_1$
(15,15)	2.257869912	-0.242130088	383.2417437
(20,20)	1.239779251	-1.260220749	14.25371765
(25,25)	3.353135569	0.853135569	1509.645815
(15,20)	6.881500401	4.381500401	19357.50094
(20,15)	1.228829792	-1.271170208	20.41377913
(15,25)	1.634796843	-0.865203157	33.74240776
(25,15)	2.291221924	-0.208778076	628.5131946
(20,25)	14.49625222	11.99625222	86704.72425
(25,20)	4.531374567	2.031374567	5162.227553

Table 4

$\alpha = 1.5, \beta_1 = 2.5, \beta_2 = 2.75$

(n,m)	$\hat{\beta}_2$	Bias $\hat{\beta}_2$	MSE $\hat{\beta}_2$
(15,15)	3.002631681	0.252631681	0.9558988122
(20,20)	2.932442797	0.182442797	0.6090520559
(25,25)	2.891735451	0.141735451	0.4601051183
(15,20)	3.002309839	0.252309839	0.9969076983
(20,15)	2.900321238	0.150321238	0.5981276207
(15,25)	2.969478335	0.219478335	0.8879485964
(25,15)	2.863795246	0.113795246	0.4227174191
(20,25)	2.935928610	0.185928610	0.6328966532
(25,20)	2.871473618	0.121473618	0.4015675626

References

[1] Ahmad, K.E., Fakhry, M.E. and Jaheen, Z.F., Empirical Bayes estimation of $P(Y < X)$ and charaterizations of Burr-type X model, *Journal of Statistical Planning and Inference* **64** (1997), 297-308.

[2] Awad, A.M., Azzam, M.M. and Hamadan, M.A., Some inference results in $P(Y < X)$ bivariate expoential model, *Communications in Statistics-Theroy and Methods* **10**(1981), 2515–2524.

[3] Baklizi, A. and Dayyh, W.A., Shrinkage estimation of $P(Y < X)$ in the exponential case, *Comm. in Statistics, Theory and Methods* **32**(1)(2003), 31–42.

[4] Church, J.D. and Harris, B., The estimation of reliablity from stress strength relationships, *Technometrics* **12**(1970), 49–54.

[5] Constantine, K. and Karson, M., The estimation of $P(Y < X)$ in gamma case, *Comm. in Statistics-Computations and Simulations* **15**(1986), 365–388.

[6] Downtown, F., The estimation of $P(Y < X)$ in normal case, *Technometrics* **15**(1973), 551–558.

[7] Dixit, U.J., Estimation of parameters of the Gamma distribution in the presence of outliers, *Comm. in Statistics, Theory and Methods* **18**(1989), 3071–3085.

[8] Dixit, U.J. and Nasiri, P., Estimation of parameters of the exponential distribution in the presence of outliers generated from uniform distribution, *Metron* **49**(3-4)(2001), 187–198.

[9] Govidarajulu, Z., Two sided confidence limits for $P(X > Y)$ based on normal samples of X and Y, *Sankhya B* **29** (1967), 35-40.

[10] Gupta, R.D. and Gupta, R.C., Estimation of $P(aX > bY)$ in the multivarite normal case, *Statistics* **1**(1990), 91–97.

[11] Gupta, R.D. and Kundu, D., Generalied exponential distributions, *Australian and New zealand Journal of Stastistics* **41**(1999), 173–188.

[12] Gupta, R.D. and Kundu, D., Generalied exponential distributions; Different Method of Esimations, *Journal of Stastistics Computation and Simulation* **69**(2001a), 315–338.

[13] Gupta, R.D. and Kundu, D., Generalied exponential distributions; An alternative to gamma or Weibull distribution, *Biometrical Journal* **43**(2001b), 117–130.

[14] Gupta, R.D. and Kundu, D., Generalied exponential distributions; Statistical Inferences, *Journal of Statistical Theory and Applications* **1**(2002), 101–118.

[15] Gupta, R.D. and Kundu, D., Closeness between the gamma and generalized exponentioal distributions, *Communications in Statistcs-Theory and Methods* **32**(2003a), 705–722.

[16] Gupta, R.D. and Kundu, D., Discriminating between the Weibull and Generalied exponential distributions, *Computational Statistics and Data Analysis* **43**(2003b), 179–196.

[17] Gupta, R.D. and Kundu, D., Discriminating between gamma and the Generalied exponential distributions, *Journal of Statistical Computation and Simulation* **74**(2004), 107–122.

[18] Kelley, G.D., Kelley, J.A. and Schucany, W.R., Efficient estimation of $P(Y < X)$ in expoential case, *Technometrics* **15**(1976), 359–360.

[19] Kundu, D. and Gupta R.D., Estimation of $P[Y < X]$ for generalied exponentioal distribution, *Metrika* **61**(3)(2005), 291–308.

[20] Kundu, D. and Gupta R.D., Estimation of $P[Y < X]$ for Weibull distribution, *IEEE Transactions on Reliability* **55**(2)(2006), 270-280.

[21] Kundu, D., Gupta, R.D. and Manglick, A., Discriminating between lognormal and generalized exponential distribution,*Journal of Statistical Planning and Inference* **127** (2005), 213–227.

[22] Mudholkar, G.S. and Srivastava, D.K., Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability* **42**(1993), 299–302.

[23] Mudholkar, G.S. and Hutson, A.D., The exponentiated Weibull family: some properties and a flood data applications, *Communications in Statistics-Theory and Methods* **25**(1996), 3059–3083.

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

[25] Owen, D.B., Craswell, K.J. and Hanson, D.L., Non-parametric upper confidence bounds for $P[Y < X]$ and confidence limits for $P(Y < X)$ when X

and Y are normal, *Journal of the American Statistical Association* **59**(1977), 906–924.

[26] Ragab, M.Z., Inference for generalized exponential distribution based on record statistics, *Journal of Statistical Planning and Inference* **104**(2002), 339–350.

[27] Ragab, M.Z. and Ahsanullah, M., Estimation of the location and parameters of the generalized exponential distribution based on order statistics, *Journal of Statistical Computation and Simulation* **69**(2001), 109–124.

[28] Sathe, Y.S. and Dixit, U.J., Estimation of $(X \leq Y)$ in negative binomial distribution, *Journal of Statistical Planning and Inference* **03**(2001), 83–92.

[29] Sathe, Y.S. and Shah, S.P., On estimation $P(Y < X)$ for the exponential distribution, *Communications in Statistics-Theory and Methods* **10**(1981), 39–47.

[30] Surles, J.G. and Padgett, W.J., Inference for reliablitiy and stress-strength for a scaled Burr-type X distribution, *Lifetime Data Analysis* **7**(2001), 187–200.

[31] Tong, H.A., A note on the estimation of $P(Y < X)$ in the exponential case, *Technometrics* **17**, 625. (Errara 1975, Vol. **17**(1975), 395).

[32] Tong, H.A., On the of $P(Y < X)$ for exponential families, *IEEE Trunsactions on Reliablility* **26**(1977), 54–56.

[33] Woodward, W.A. and Kelley, G.D., Minimum variance unbiased estimation of $P[Y < X]$ in the normal case, *Technometrics* **19**(1977), 95–98.

[34] Zheng, G., On the Fisher information matrix in type-II censored data from the exponentiaed exponetial family, *Biometrical Journal* **44**(2002), 353–357.