ISSN 0970-5226

JOURNAL OF NATIONAL ACADEMY OF MATHEMATICS INDIA

Vol. 29
(2015)

Published By
National Academy of Mathematics, India
D. D. U. Gorakhpur University
GORAKHPUR - 273 009, INDIA
e-mail : nationalacademyofmathematics@gmail.com

Officials of the National Academy of Mathematics, India

JOURNAL
OF
NATIONAL ACADEMY OF MATHEMATICS
INDIA

Vol. 29

(2015)

Published By :

National Academy of Mathematics, India

D.D.U. Gorakhpur University,

Gorakhpur - 273009

India

e-mail : nationalacademyofmathematics@gmail.com

J. Nat. Acad. Math.
Vol. 29 (2015)
Printed in India

CONTENTS

Printed at : Sinha Brother's, Indralok talkies Campus, Gorakhpur

J. Nat. Acad. Math.
Vol. 29 (2015), pp.99-108 ISSN 0970-5228

AN EXTENSION OF EXPONENTIAL DISTRIBUTION : THEORY AND APPLICATIONS

By

RAMESH KUMAR JOSHI

Abstract

In this paper a new extension of exponential distribution is introduced. The parameters are estimated using likelihood based inferential procedures. The maximum likelihood(ML) estimates and asymptotic confidence intervals based on ML estimates are obtained. All the computations are performed in R software. A real data set is analyzed for illustration and applications.

Keywords : Exponential distribution, Maximum likelihood estimation.
2010 AMS Subject Classification : 62F15, 65C05.

1. Introduction

The most exploited life test model is exponential distribution due to the existence of simple elegant closed form solutions to many life-testing problems. It can easily be justified under the assumption of constant failure rate but in the real world, the failure rates are not always constant. Hence, indiscriminate use of exponential lifetime model seems to be inappropriate and unrealistic. In recent years, new classes of models have been proposed based on modifications of the existing classical probability models, Marshall and Olkin(2007).

Recently, some attempts have been made to define new families of distributions to extend well known models and at the same time provide great flexibility in modeling data in practice. Several techniques could be employed to form a larger family from an existing distribution by incorporating extra parameters. So, several classes by adding one or more parameters to generate new models have been proposed in the statistical literature, Rinne(2009) and Pham and Lai(2007).

The continuous random variable X follows the exponential distribution with parameter θ if its cumulative distribution function(CDF) is given by

$$F(x;\theta) = 1 - e^{-\theta x} \quad ; \ \theta > 0, x > 0. \tag{1.1}$$

In order to provide some flexibility, alternative generalizations of the exponential distribution have been proposed.

Received : April 23, 2015; Accepted : July 22, 2015

There are several generalizations of exponential distribution. The five generalizations of the exponential distribution have received increased attention in the literature as compared to others, viz:

(i) Gupta and Kundu (1999) proposed the generalized-exponential (GE) distribution. The GE density can have decreasing and unimodal (right-skewed) shapes, and the hazard rate exhibits IFR and DFR shapes. Let X follow the GE random variable (rv), say $X \sim GE(\alpha, \lambda)$. The CDF of X is given by

$$F(x;\, \alpha,\, \lambda) = \left\{1 - e^{-\lambda x}\right\}^{\alpha} \quad ; \; x > 0, \tag{1.2}$$

where $\alpha > 0$ is a shape parameter and $\lambda > 0$ is a scale parameter. For $\alpha = 1$, the GE distribution reduces to the exponential distribution. This family has lots of properties which are quite similar to those of a Gamma distribution but it has an explicit expression of the survival function like a Weibull distribution. Gupta and Kundu(1999) provided a detailed review and some developments on the generalized exponential distribution.

(ii) Haghighi and Sadeghi(2009) and Nadarajah and Haghighi (2011) introduced a generalization of the exponential distribution and called it as an extension of exponential distribution. Its density can have decreasing and unimodal shapes, and the hazard rate exhibits increasing and decreasing shapes. Let X have the exponential extension distribution, say $X \sim NHE(\alpha, \lambda)$. The CDF of X is given by

$$F(x; \alpha, \lambda) = 1 - \exp\left\{1 - (1 + \lambda x)^{\alpha}\right\} \quad ; x \geqslant 0, \tag{1.3}$$

where $\lambda > 0$ is a scale parameter and $\alpha > 0$ is a shape parameter. The exponential distribution is a basic exemplar when $\alpha = 1$. Nadarajah and Haghighi (2011) proved that the NHE density allows decreasing and unimodal shapes and that the hazard rate shapes are increasing, decreasing and constant.

(iii) Lan and Leemis (2008) introduced the logistic-exponential distribution. Let X have the logistic-exponential (LE) distribution, say $X \sim LE(\alpha, \lambda)$. The CDF of X is given by

$$F(x; \alpha, \lambda) = \quad 1 - \frac{1}{1 + (e^{\lambda x} - 1)^{\alpha}} \quad ; x \geqslant 0,\ \alpha > 0,\ \lambda > 0. \tag{1.4}$$

For $\alpha = 1$, the LE distribution reduces to the exponential distribution. The logistic-exponential distribution has several useful probabilistic properties for lifetime modeling. It has increasing failure rate, decreasing failure rate, bathtub(BT)-shaped failure rate, and upside-down bathtub(UBT)-shaped failure rates for various parameter combinations.

Unlike most distributions in the BT and UBT classes, the LE distribution enjoys closed-form density, hazard, cumulative hazard, and survival functions. The moments are finite, although they cannot be expressed in closed form.

(iv) Adding parameters to a well-established family of distributions is a time honored device for obtaining more flexible new families of distributions. Marshall and Olkin (1997) introduced an inventive general method of adding a parameter to a family of distributions. Let X have the Marshall-Olkin Extended Exponential (MOEE) distribution, say $X \sim MOEE\,(\alpha, \lambda)$. The CDF of X is given by

$$F(x;\alpha,\lambda) = \frac{1-e^{-\lambda x}}{1-(1-\alpha)\,e^{-\lambda x}} \quad ; \; x>0. \tag{1.5}$$

Here $\alpha > 0$ and $\lambda > 0$ are the tilt and scale parameters respectively. For $\alpha = 1$, the MOEE distribution reduces to the exponential distribution.

(v) Gómez et al. (2014) introduced another extension of the exponential distribution. A random variable X is distributed according to the extended exponential distribution (EE) with parameters $\alpha > 0$ and $\lambda > 0$, say $X \sim EE\,(\alpha, \lambda)$ if its CDF is given by

$$F\,(x;\alpha,\lambda) = 1-\left(1+\frac{\alpha\lambda}{(\alpha+\lambda)}x\right)\;\exp\left(-\lambda x\right) \quad ; x>0. \tag{1.6}$$

For $\alpha = 1$, the EE distribution also reduces to the exponential distribution.

The rest of the article is arranged as follows. The new extension of exponential distribution is introduced and their properties are discussed in Section 2. The Section 3 deals with the maximum likelihood estimation procedure to estimate the model parameters and associated confidence intervals using the observed information matrix are discussed. A real data set has been analysed in Section 4 to illustrate the application and suitability of the proposed distribution . In this section, the ML estimators of the parameters, approximate confidence intervals are presented. Finally, Section 5 ends up with some general concluding remarks.

2. The new extended exponential model:

The CDF of the new extended exponential distribution, say $X \sim EEN\,(\alpha, \lambda)$ is defined as

$$F\,(x;\alpha,\lambda) = 1-\exp\left\{-\alpha\, x\, e^{-\lambda/x}\right\} \quad ; \; x>0,\; \alpha>0,\; \lambda>0 \tag{2.1}$$

The probability density function(PDF) is given by

$$f(x;\alpha,\lambda)=\alpha\left(1+\frac{\lambda}{x}\right)e^{-\lambda/x}\exp\left\{-\alpha\, x\, e^{-\lambda/x}\right\}\quad ;\, x>0. \tag{2.2}$$

The reliability/survival function is

$$R(x)=\exp\left\{-\alpha\, x\, e^{-\lambda/x}\right\}\quad ;\, x>0 \tag{2.3}$$

The hazard rate function(HRF) is

$$h(x)=\alpha\left(1+\frac{\lambda}{x}\right)e^{-\lambda/x}\quad ;\, x>0. \tag{2.4}$$

The quantile function is one of the ways of specifying the distribution of a random variable and it is also an alternative to the PDF and CDF. The quantile function is usually used to obtain the statistical measures such as the median, skewness, kurtosis and it is also used to generate the random variables. The definition of the p^{th} quantile is the real solution of the following equation

$$F(Q(p))=p$$

follows easily by inverting (2.1) as

$$\frac{\lambda}{Q(p)}-\log(Q(p))+\log\left\{-\frac{1}{\alpha}(1-p)\right\}=0,\;\; 0<p<1. \tag{2.5}$$

Quantiles of interest can be obtained from (2.5) by substituting appropriate values for p. In particular, the median of $p=0.5$ is

Here, we consider simulating values of a random variable X with the CDF (2.1). Let U denote a uniform random variable on the interval $(0,1)$. One way to simulate values of X is to set

$$\frac{\lambda}{x}-\log(x)+\log\left\{-\frac{1}{\alpha}(1-u)\right\}=0 \tag{2.6}$$

and solve for X.

The skewness and kurtosis measures are used in statistical analyses to characterize a distribution or a data set. The Bowley's skewness measure based on quartiles is given by

$$S_k=\frac{Q(0.75)+Q(0.25)-2Q(0.5)}{Q(0.75)-Q(0.25)},$$

and the Moors's kurtosis measure based on octiles (Moors (1988)) is given by

$$K_u=\frac{Q(0.875)-Q(0.625)+Q(0.375)-Q(0.125)}{Q(0.75)-Q(0.25)},$$

where the $Q(.)$ is the quantile function. The skewness and kurtosis measures based on quantiles like Bowley's skewness and Moors's kurtosis have a number of advantages compared to the classical measures of skewness and kurtosis, e.g. they are less sensitive to outliers and they exist for the distributions even without defined the moments.

Plots of the PDF and hazard rate function of the $EEN(\alpha, \lambda)$ some choices of values of the parameters are displayed in Figure 1.

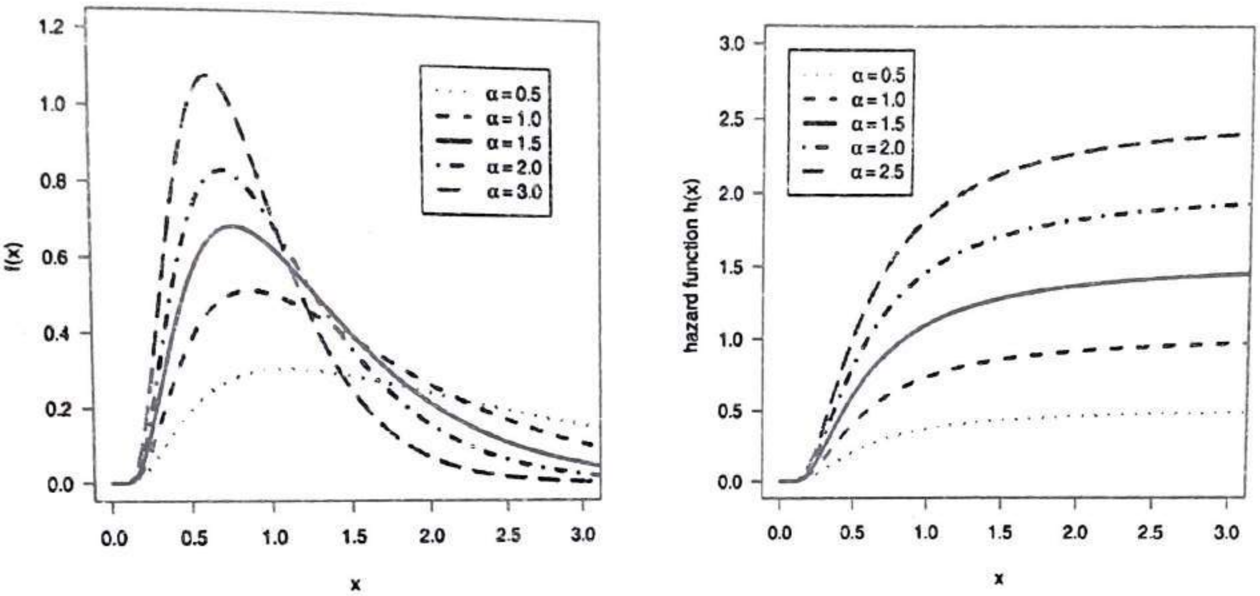


Figure 1. Plots of the probability density function(left panel) and hazard function (right panel), for λ=1 and different values of α.

3. Maximum Likelihood Estimation

In this section, we determine the maximum likelihood estimates of the model parameters and asymptotic confidence intervals. Let $\underline{x} = (x_1, \ldots, x_n)$ be a sample from a distribution with probability density function (2.1). The likelihood function of the parameter $L(\alpha, \lambda)$ is given by

$$L(\alpha, \lambda|\underline{x}) = \prod_{i=1}^{n} \alpha \left(1 + \frac{\lambda}{x_i}\right) e^{-\lambda/x_i} \exp\left\{-\alpha\, x_i\, e^{-\lambda/x_i}\right\} \tag{3.1}$$

The log likelihood function of the parameter $\ell(\alpha, \lambda)$ is given by

$$\ell(\alpha, \lambda|\underline{x}) = n \log(\alpha) + \sum_{i=1}^{n} \log\left(1 + \frac{\lambda}{x_i}\right) - \sum_{i=1}^{n} \frac{\lambda}{x_i} - \alpha \sum_{i=1}^{n} x_i\, e^{-\lambda/x_i} \tag{3.2}$$

The maximum likelihood estimators of the parameters have obtained by differentiating (3.2) w.r.t.to parameters and equating to zero, we have

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^{n} x_i\, e^{-\lambda/x_i} = 0 \\
\frac{\partial \ell}{\partial \lambda} &= -n \sum_{i=1}^{n} \frac{\lambda}{x_i\,(\lambda + x_i)} + \alpha \sum_{i=1}^{n} e^{-\lambda/x_i} = 0
\end{aligned} \tag{3.3}$$

The maximum likelihood estimator (MLE) $\hat{\delta} = (\hat{\alpha}, \hat{\lambda})^{\top}$ of $\delta = (\alpha, \lambda)^{\top}$ is obtained by solving simultaneously the nonlinear equations (3.3). These equations cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as a Newton-Raphson type algorithm to calculate the estimate $\hat{\delta}$. For example, *optim*() function in R software can be used to compute $\hat{\delta}$ numerically.

The $(1-\gamma)\,\%$ confidence intervals for α, β and λ can be obtained from the usual asymptotic normality of the maximum likelihood estimators with $var\,(\hat{\alpha})$, $var\left(\hat{\beta}\right)$ and $var\left(\hat{\lambda}\right)$ estimated from the inverse of the matrix of second derivatives of the log-likelihood function locally at $\hat{\alpha}$ and $\hat{\lambda}$, Casella and Berger(2002).

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha^2} &= -\frac{n}{\alpha^2} \\
\frac{\partial^2 \ell}{\partial \alpha\, \partial \lambda} &= \sum_{i=1}^{n} e^{-\lambda/x_i} \\
\frac{\partial^2 \ell}{\partial \lambda^2} &= -\alpha \sum_{i=1}^{n} (1/x_i)\, e^{-\lambda/x_i} - \sum_{i=1}^{n} \frac{1}{(\lambda + x_i)^2}
\end{aligned}$$

Hence, from the asymptotic normality of MLEs, approximate $(1-\gamma)\,\%$ confidence intervals for α and λ can be constructed as

$$\hat{\alpha} \pm z_{\gamma/2}\sqrt{var(\hat{\alpha})};\ \text{ and }\ \hat{\lambda} \pm z_{\gamma/2}\sqrt{var(\hat{\lambda})} \tag{3.4}$$

where $z_{\gamma/2}$ is the upper percentile of standard normal variate.

4. Data Analysis: Application

In this section we present the analysis of one real data set for illustration of the proposed methodology. The following data represent the number of million revolutions before failing for each of 23 ball bearings in a life test, Lawless(2003).

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 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 computed the ML estimates by maximizing the log-likelihood function given in equation (3.1) directly using *optim*() function in R software (R Development Core Team (2015) and Rizzo(2008)). We obtain $\hat{\alpha} = 0.0347$ and $\hat{\lambda} = 75.1948$, and the corresponding log-likelihood value is -112.9694. The contour plot and fitted CDF with empirical distribution function(EDF)are presented in Figure 2, Kumar and Ligges(2011).

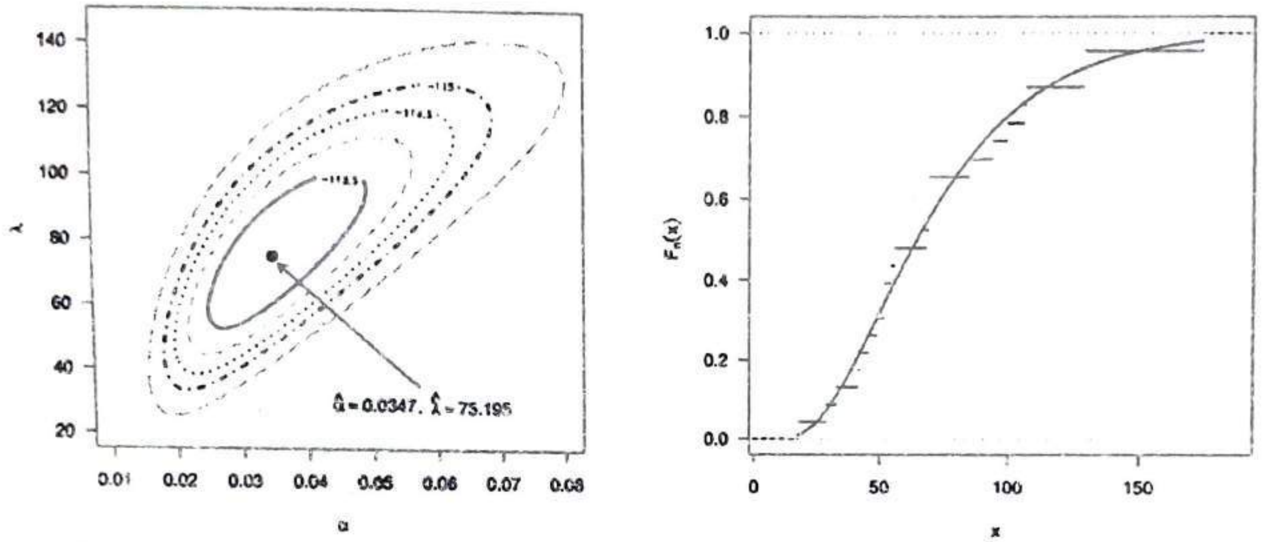


Figure 2. Contour plot(left panel) and the fitted CDF with empirical distribution function (right panel).

Using the method described in Section 3, we can construct the approximate confidence intervals(ACI) based on MLE's. Table 1 shows the MLE's with their standard errors(SE) and 95% confidence intervals for α and λ.

Table 1

MLE, SE and 95% confidence intervals

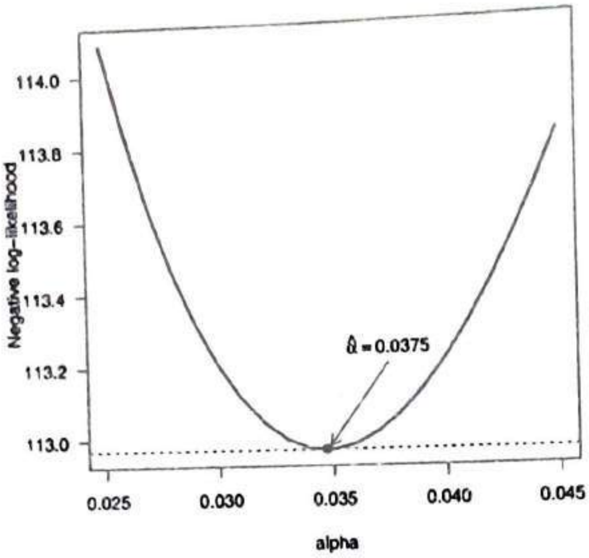
Parameter	MLE	SE	95% ACI
alpha	0.0347	0.00808	(0.0189, 0.0505)
lambda	75.1948	9.3599	(56.849, 93.540)

From Figure 3 it is evident that MLE's are unique. We have considered five alternative models, namely NHE(Nadarajah and Haghighi (2011)), EE(Gómez et al. (2014)), MOEE(Marshall and Olkin (1997)), LE(Lan and Leemis (2008)) and GE(Gupta and Kundu (1999)). The models are compared via the Akaike Information Criterion (AIC), the Corrected Akaike Information criterion (AICC), and Bayesian information criterion (BIC) which are used to select the best model among several models. The definitions of AIC, BIC and AICC are given below:

$AIC = -2\ell\left(\hat{\theta}\right) + 2k$; $\quad BIC = -2\ell\left(\hat{\theta}\right) + k\log(n)$, and

$AICC = AIC + \frac{2k(k+1)}{n-k-1}$

where k is the number of parameters in the model under consideration



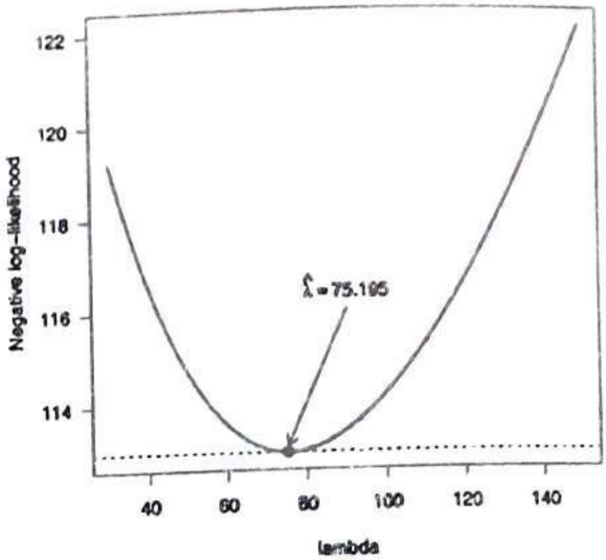


Figure 3. Profile log-likelihood functions of α and λ.

The maximum likelihood estimates, the log-likelihood value, the AIC, BIC and AICC are reported in Tables 2, we conclude that the EEN distribution provides a better fit to this data than other models.

Table 2

MLE, log-likelihood, AIC, BIC and AICC

Model	$\hat{\alpha}$	$\hat{\lambda}$	$\ell(\hat{\theta})$	AIC	BIC	AICC
NHE	33.2019	0.0003	-117.2509	238.5019	240.7729	239.0473
EE	0.0277	12.11	-115.5449	235.0897	237.3607	235.6352
MOEE	17.9213	0.0435	-114.3503	232.7006	234.9716	233.2461
LE	2.3675	0.0106	-113.2403	230.4806	232.7516	231.026
GE	5.2832	0.0323	-112.9762	229.9524	232.2234	230.4979
EEN	0.0347	75.1947	-112.9694	229.9388	232.2098	230.4842

Moreover, perfection of competing models is also tested via the Kolmogrov-Simnorov(K-S), the Anderson-Darling (A^2) and the Cramer-Von Mises (W) statistics. The mathematical expressions for the statistics above are given below

$$KS = \max_{1\leqslant i\leqslant n}\left(z_i - \frac{i-1}{n}, \frac{i}{n} - z_i\right)$$

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\ln z_i + \ln\left(1 - z_{n+1-i}\right)\right\}$$

$$W = \frac{1}{12n} + \sum_{i=1}^{n}\left\{\frac{(2i-1)}{2n} - z_i\right\}^2$$

where $z_i = CDF(x_i)$; the x_i's being the ordered observations, D'Agostino and Stephens(1986).

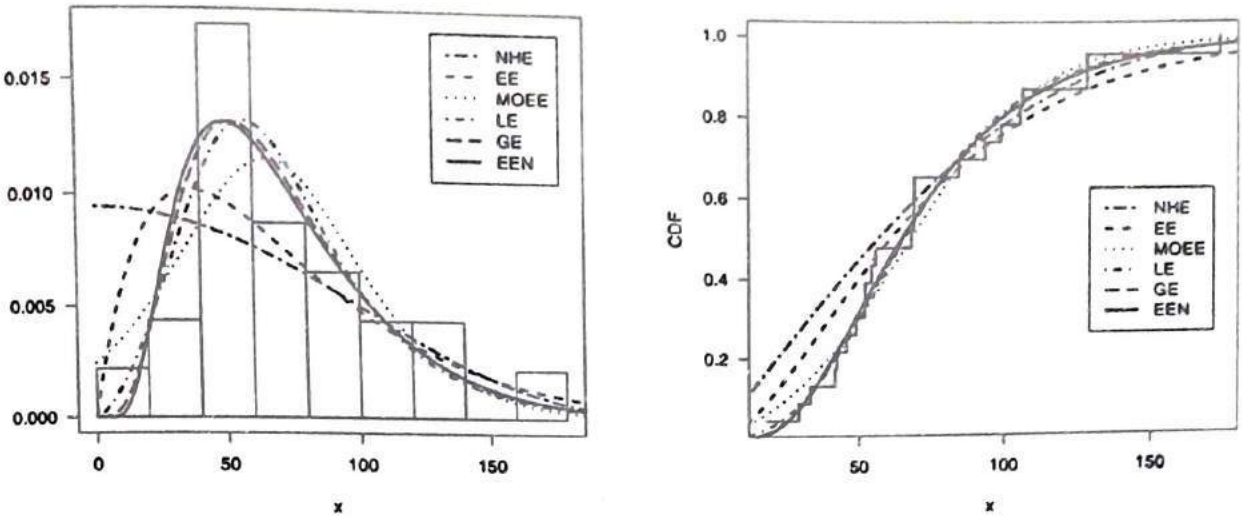


Figure 4. The Histogram and the PDF of fitted distributions (left panel); Empirical CDF with estimated CDF(right panel).

The values of Kolmogorov-Smirnov(KS), Anderson-Darling(A^2) and Cramer-Von Mises (W) statistic with their respective p-value of different models are reported in Table 3. As we can see in Table 3, the proposed model has the minimum values of the test statistics and higher p-value. Figure 4 (left panel) displays the histogram and the fitted density functions, which support the results in Tables 2 and 3. Also, Figure 4(right panel) which compares the distribution functions for the different models with the empirical distribution function reveals the same. Therefore, for the given data set shows the proposed distribution gets better fit and more reliable solutions from other alternatives.

Table 3

The KS, A^2 and W statistics with p-value

Model	KS(p-value)	A^2(p-value)	W(p-value)
NHE	0.2484(0.1170)	1.6365(0.1473)	0.2970(0.1376)
EE	0.1887(0.3858)	0.8803(0.4254)	0.1408(0.4215)
MOEE	0.1383(0.7714)	0.3795(0.8675)	0.0589(0.8255)
LE	0.1100(0.9437)	0.2193(0.9843)	0.0390(0.9422)
GE	0.1058(0.9589)	0.1871(0.9937)	0.0322(0.9711)
EEN	0.1003(0.9748)	0.1876(0.9936)	0.0303(0.9776)

8. Conclusion

In this study we have proposed a new distribution called an exponential extension new(EEN). We derived the maximum likelihood estimates of the parameters and their associated confidence intervals. A real data application has demonstrated that the proposed model is quite useful for dealing with real data and behaves better in terms of fitting than other models such as the NHE, EE, MOEE, LE and GE models. It is also worth mentioning that we have not only confine our interest in the information criterion and statistics but also adopted the graphical displays, so that the reader can gain a perspective of the various meanings and associated interpretations.

Trichandra Multiple Campus
Tribhuvan University, Kathmandu, Nepal
email : rmshjoshi@gmail.com

References

1. **Casella, G. and Berger, R. :** Statistical Inference, Second Edition. Belmont: Duxbury Press, (2002).
2. **D'Agostino, R.B. and Stephens, M.A. :** Goodness-of-fit techniques. New York, Marcel Dekker; (1986).
3. **Gomez, Y. M., Bolfarine, H. and Gomez, H. W. :** A New Extension of Exponential Distribution. Rev. Colombiana de Estadistica, **37(1)**(2014), 25-34.
4. **Gupta, R. D. and Kundu, D. :** Generalized exponential distributions, Australian and New Zealand Journal of Statistics, **41(2)**(1999), 173 - 188.
5. **Haghighi,F. and Sadeghi, S. :** An Exponential Extension, 41émes Journées de Statistique, https://hal.inria.fr/inria-00386755 (2009).
6. **Kumar, V. and Ligges, U. :** reliaR : A package for some probability distributions, http://cran.r-project.org/web/packages/reliaR/index.html, (2011).
7. **Lan, Y. and Leemis, L. M. :** The Logistic-Exponential Survival Distribution. Naval Research Logistics, **55** (2008), 252-264.
8. **Marshall, A. W. and Olkin, I. :** A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, **84(3)**, (1997), 641?652.
9. **Marshall, A. W. and Olkin, I. :** Life Distributions: Structure of Nonparametric, Semiparametric and Parametric Families, Springer, New York, (2007).
10. **Moors, J. :** A quantile alternative for kurtosis. The Statistician, **37** (1988), 25-32.
11. **Nadarajah, S. and Haghighi, F. :** An extension of the exponential distribution, Statistics, **45(6)** (2011), 543-558.
12. **Pham, H. and Lai, C.D. :** On recent generalizations of the Weibull distribution, IEEE Trans. on Reliability, **56(3)** (2007), 454-458.
13. **R Development Core Team :** An Introduction to R, ISBN 3-900051-12-7 (2015), URL http://www.r-project.org.
14. **Rinne, H. :** The Weibull Distribution : A Handbook, CRC, (2009).
15. **Rizzo, M. L. :** Statistical computing with R. Chapman & Hall/CRC, (2008).