IJSAR, 3(3), 2016; 51-59

International Journal of Sciences & Applied Research

www.ijsar.in

Parameter and confidence interval estimation of the extended exponential distribution from right censored survival time data

Kamil ALAKUŞ[1], Necati Alp ERİLLİ[2]*

[1]Department of Statistics, OndokuzMayisUniversity, Samsun, TURKEY.
[2]Department of Econometrics, Cumhuriyet University, Sivas, TURKEY.

Correspondence Address: *Necati Alp ERİLLİ, Department of Econometrics, Cumhuriyet University, Sivas, TURKEY.

Abstract

Introduction: Extended Exponential Distribution is a new expansion of Exponential distribution. It has a rich hazard family which can study out a wide application area in survival analysis or reliability theory. Extended Exponential Distribution has also increasing, decreasing and constant hazard properties like Weibull, Exponentiated Exponential and Gamma Distributions. It is a kind of distribution which has accelerated hazard ratio properties.

Case Report: In this study parameter estimation and asymptotic confidence intervals of parameters estimation in Extended Exponential Distribution for survival time analysis of right censored is discussed. A real time data is applied for illustrating to consider parameter estimation and asymptotic confidence intervals of parameters estimation in the study.

Conclusion: Results for Extended Exponential distribution for analysis in Modal Choice and Correlation Results for Lung Cancer Test Group Patients give the best results than other distributions. This shows that Extended Exponential distribution could find a large application area in survival analysis and theory.

Keywords: Extended Exponential Distribution, Point Estimation, Confidence Interval, Right Censoring

Introduction

Survival analysis is a discipline of statistics which deals with death in biological organisms and failure in mechanical systems. This topic is called reliability theory or reliability analysis in engineering and duration analysis or duration modeling in economics or even thistory analysis in sociology. Survival analysis attempts to answer questions such as fraction of a population which will survive in a certain time or what rate they will die or fail (Alakus et al., 2015).

In recent years there are two important distributions frequently studied in survival time analysis or reliability theory. One is Exponentiated Exponential distribution and the other is Extended Exponential Distribution. Both distributions have rich hazard family as Weibull distribution as. Gupta and Kundu (2001) have introduced the Exponentiated Exponential distribution. This distribution has lots of properties which are quite similar to Gamma distribution but it has an explicit expression of the survival function like Gamma or Weibull distributions. Gupta and Kundu (2007) provided a detailed review for Exponentiated Exponential distribution also. Extended Exponential Distribution is studied by Haghighi and Sadeghi (2009) and Nadarajah and Haghighi (2010) in recent years. Haghighi and Sadeghi (2009) is studied maximum likelihood estimators for the distribution. Nadarajah and Haghighi (2010) is studied some important moments, interval statistics and moments of likelihood estimators of extended exponential distribution. Both studies are used with uncensored data sets. In real life applications especially in health studies, many survival times have censored structure. It is hard to find any sample with uncensored survival times in real studies. For this reason, point and asymptotic confidence intervals for right censoring data for real applications is discussed in this study.

Materials and Methods

A probability distribution plays a central role not only in the analysis of survival data but also for other parametric counting process models. On the other hand the hazard function plays a central role in survival time data because it gives more information on risks.

The hazard plot technique, which was developed by Nelson (1969, 1972), provides a simple graphical procedure for both choosing an appropriate model (i.e., distribution) and approximate estimates of distribution parameters from sample data. These results may be used for all other parametric counting process models.

The hazard function must have an appropriate structure, from which may be deduced other properties such as cumulative hazard and survivor function. Some suitable parametric distributions considering in this study and their hazard and survival forms are introduced and summarized in Table.1. For other parametric distributions, hazard forms and details can be found in Alakuş (1997).

In Table.1, we can easily say that there are two types of hazard family: One is decreasing, increasing and constant type hazard family and the other is decreasing and uni-modal type hazard family. These groups can be summarizing as given in Figure 1 and Figure 2.

Table 1: Important Functions for Some Useful Probability Distributions.

Distribution	Survival Function	Hazard Function	Hazard Type
Standard Exp.	$exp(-\lambda t)$	λ	Constant
Weibull	$exp\{-(\lambda t)^{\alpha}\}$	$\alpha\lambda(\lambda t)^{\alpha-1}$	For α<1 Decreasing For α=1 Constant For α>1 Increasing
Log-Normal	$1-\Phi\{(logt-\mu)/\sigma\}$	$f(t;\sigma,\mu)/S(t;\sigma,\mu)$	For $\sigma \geq$ e Decreasing For $\sigma < e$ Uni-modal
Log-Logistic	$[1+(\lambda t)^{\alpha}]^{-1}$	$\alpha\lambda(\lambda t)^{\alpha-1}\times[1+(\lambda t)^{\alpha}]^{-1}$	For α≤1 Decreasing For α>1 Uni-modal
Expo-Expo	$1-[1-exp(-\lambda t)]^{\alpha}$	$\alpha\lambda exp(-\lambda t)[1-exp(-\lambda t)]^{\alpha-1}$ $\times\{1-[1-exp(-\lambda t)]^{\alpha}\}^{-1}$	For α<1 Decreasing For α=1 Constant For α>1 Increasing
Extended Expo.	$exp[1-(1+\lambda t)^{\alpha}]$	$\alpha\lambda(1+\lambda t)^{\alpha-1}$	For α<1 Decreasing For α=1 Constant For α>1 Increasing

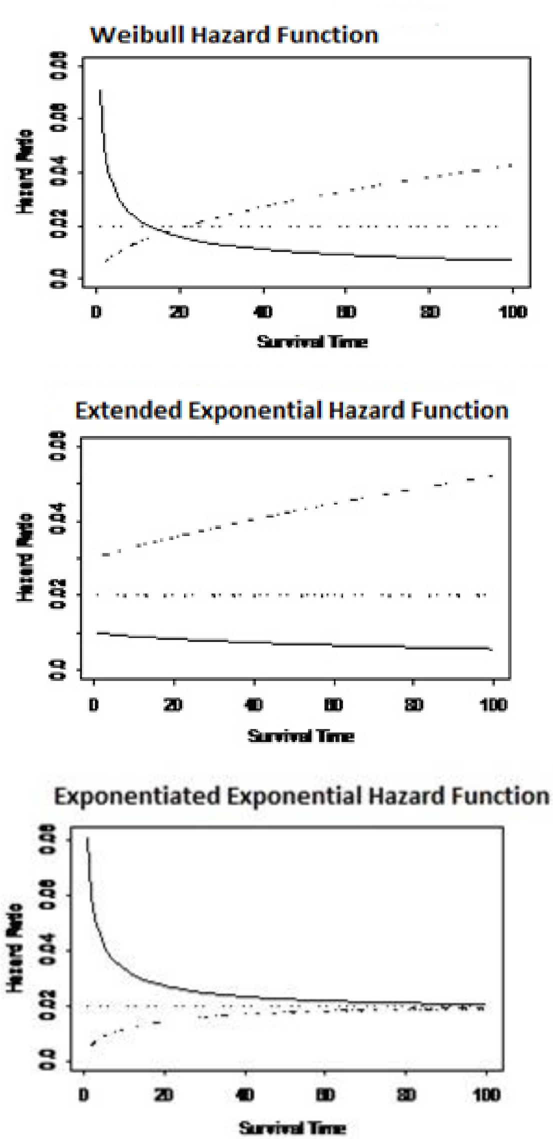

Fig. 1: Monotonic Decreasing, Increasing and Constant Type Hazard Probability Distribution Functions.

It can be easily conclude that; the extended exponential distribution has a rich hazard family as well as Weibull distribution.

Parameter Estimations: Right Censoring Case

Extended Exponential Distribution

Probability Density Function of Extended Exponential Distribution can be written like as above:

$$f(t) = h(t)S(t) = \begin{cases} \alpha\lambda(1+\lambda t)^{\alpha-1} e^{1-(1+\lambda t)^{\alpha}} & ; \quad t > 0 \\ 0 & ; \quad .... \end{cases} \quad (1)$$

where λ is scale parameter and α is shape parameter. Corresponding to eq.(1) Hazard Function can written as given in equation (2):

$$h(t) = \alpha\lambda(1+\lambda t)^{\alpha-1}, t > 0 \quad (2)$$

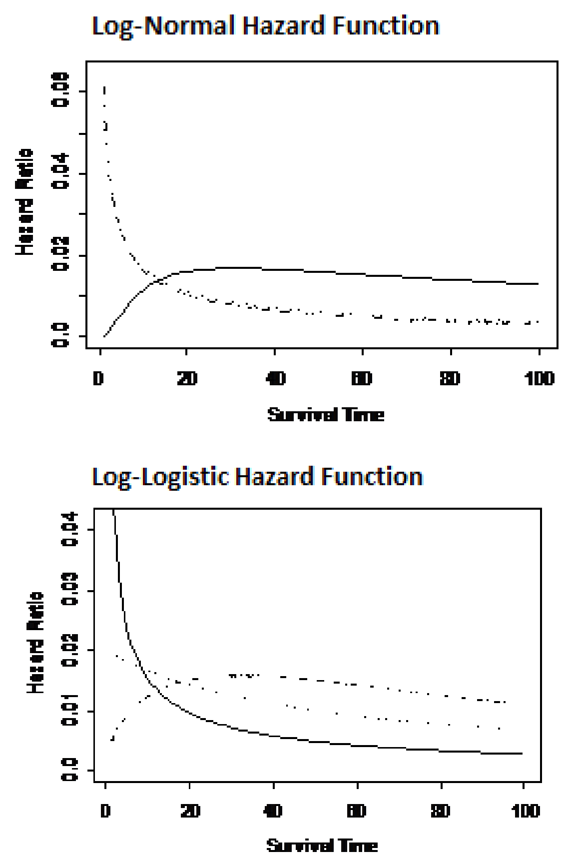

Fig. 2: Monotonic Decreasing and Uni-Modal Type Hazard Probability Distribution Functions.

Cumulative Hazard Function can be found by integrated with the Hazard function. Than cumulative Hazard Function can found as given in equation (3):

$$H(t) = \int_0^t h(s)\,ds = \alpha\lambda \int_0^t (1+\lambda s)^{\alpha-1} ds \quad (3)$$

If we write $1 + \lambda s = y$ than $\lambda ds = dy$ and limits must change as;

$s = 0 \Longrightarrow y = 1$ and $s = t \Longrightarrow y = 1 + \lambda t$.

So Hazard Function can be tidy up as given:

$$H(t) = \alpha \int_1^{1+\lambda t} y^{\alpha-1} ds = y^{\alpha}|_1^{1+\lambda t} = (1+\lambda t)^{\alpha} - 1 \quad (4)$$

Finally, corresponding Survival Function can be found as below:

$$S(t) = e^{-H(t)} = e^{1-(1+\lambda t)^{\alpha}}, \; t > 0 \quad (5)$$

Point Estimation: Studying with censored data in real life time applications is inevitable for the reason of survival time cannot be observed exactly. For instance in searching HIV deaths in a location, patients can die with heart attack or cold virus or other diseases during the study. Sometimes patients are still alive when study is totally over. Such like this kind of censoring is common problem in survival analysis studies. In this study point and confidence intervals are studied under right censoring data.

Extended Exponential Likelihood Function with two parameters under right censoring is given in eq. (6).

$$L(t;\alpha,\lambda) = \prod_{i=1}^{n}[f(t_i;\alpha,\lambda)]^{w_i} \prod_{i=1}^{n}[S(t_i;\alpha,\lambda)]^{1-w_i}$$

$$\alpha^d \lambda^d \prod_{i=1}^{n}[(1+\lambda t_i)^{\alpha-1}]^{w_i} \prod_{i=1}^{n} exp[1-(1+\lambda t_i)^{\alpha}] \quad (6)$$

where d means uncensored observation number and w_i means dummy variable which 0 is censoring and 1 is uncensored data. Taking logarithm of eq.(6), Log-likelihood function can be given as in equation (7):

$$\ell L(.) = dlog\alpha + dlog\lambda + (\alpha-1)\sum_{i=1}^{n} log(1+\lambda t_i) + n - \sum_{i=1}^{n}(1+\lambda t_i)^{\alpha}$$

(7)

Maximum likelihood estimates can be found by simultaneous solutions of the equations (8) and (9).

$$\frac{\partial \ell L(.)}{\partial \alpha} = 0 \to \frac{d}{\alpha} + \sum_{i=1}^{n} w_i \log(1+\lambda t_i) - \sum_{i=1}^{n}(1+\lambda t_i)^{\alpha} \lg(1+\lambda t_i) = 0$$

(8)

$$\frac{\partial \ell L(.)}{\partial \lambda} = 0 \to \frac{d}{\lambda} + (\alpha-1)\sum_{i=1}^{n} w_i t_i (1+\lambda t_i)^{-1} - \alpha \sum_{i=1}^{n} t_i (1+\lambda t_i)^{\alpha-1} = 0$$

(9)

Interval Estimations: Second partial deviations of the function (7) are given in equation (10).

$$\frac{\partial^2 \ell L(.)}{\partial \alpha^2} = 0 \Rightarrow -\frac{d}{\alpha^2} - \sum_{i=1}^{n}(1+\lambda t_i)^{\alpha} \lg^2(1+\lambda t_i) = 0$$

$$\frac{\partial^2 \ell L(.)}{\partial \lambda^2} = 0 \Rightarrow -\frac{d}{\lambda^2} - (\alpha-1)\sum_{i=1}^{n} w_i t_i^2 (1+\lambda t_i)^{-2}$$

$$-\alpha(\alpha-1)\sum_{i=1}^{n} t_i^2(1+\lambda t_i)^{\alpha-2} = 0$$

$$\frac{\partial^2 \ell L(.)}{\partial \alpha \partial \lambda} = 0 \Rightarrow \sum_{i=1}^{n} w_i t_i (1+\lambda t_i)^{-1} - \sum_{i=1}^{n} t_i (1+\lambda t_i)^{\alpha-1}$$

$$-\alpha \sum_{i=1}^{n} t_i (1+\lambda t_i)^{\alpha-1} log(1+\lambda t_i) = 0 \quad (10)$$

The MLE of the vector of unknown parameters $\theta = (\alpha, \lambda)$ is not obtained in closed form. It is not possible to derive the exact distribution of the MLE. Here, we derive the confidence intervals of the parameters based on the asymptotic distributions of the MLE of the parameters. Miller (1981) gives the asymptotic distribution of the MLE as given:

$$(\hat{\theta} - \theta) \sim N_2(0, \Sigma_{\theta})$$

Thus, the Hessian Matrix can be written as in equation (11):

$$H(\Theta) = \begin{bmatrix} \frac{\partial^2 \ell L(.)}{\partial \alpha^2} & \frac{\partial^2 \ell L(.)}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 \ell L(.)}{\partial \lambda \partial \alpha} & \frac{\partial^2 \ell L(.)}{\partial \lambda^2} \end{bmatrix} \quad (11)$$

So, Fischer Information matrix can be calculated as given;

$$I(\theta) = -H(\theta) = \begin{bmatrix} -\frac{\partial^2 \ell L(.)}{\partial \alpha^2} & -\frac{\partial^2 \ell L(.)}{\partial \alpha \partial \lambda} \\ -\frac{\partial^2 \ell L(.)}{\partial \lambda \partial \alpha} & -\frac{\partial^2 \ell L(.)}{\partial \lambda^2} \end{bmatrix} \quad (12)$$

Therefore Variance-Covariance matrix is given by;

$$\Sigma_{\theta} = \frac{1}{|I(\theta)|} Adj.(I(\theta)) =$$

$$\frac{1}{\det(I(\theta))} \begin{bmatrix} -\frac{\partial^2 \ell L(.)}{\partial \lambda^2} & \frac{\partial^2 \ell L(.)}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 \ell L(.)}{\partial \lambda \partial \alpha} & -\frac{\partial^2 \ell L(.)}{\partial \alpha^2} \end{bmatrix} \quad (13)$$

We can get variance of α and λ as given in equation (14) and (15):

$$V(\hat{\alpha}) = \frac{\left(-\frac{\partial^2 \ell L(.)}{\partial \lambda^2}\right)}{|I(\theta)|} \quad (14)$$

$$V(\hat{\lambda}) = \frac{\left(-\frac{\partial^2 \ell L(.)}{\partial \alpha^2}\right)}{|I(\theta)|} \quad (15)$$

The approximate $100(1-\Upsilon)\%$ two sided confidence interval of α and λ can be written as, respectively;

$$Pr\left\{\hat{\alpha} - Se(\hat{\alpha}) z_{\left(\frac{\gamma}{2}\right)} \le \alpha \le \hat{\alpha} + Se(\hat{\alpha}) z_{1-\left(\frac{\gamma}{2}\right)}\right\} = 1-\gamma \quad (16)$$

and

$$Pr\left\{\hat{\lambda} - Se(\hat{\lambda}) z_{\left(\frac{\gamma}{2}\right)} \le \lambda \le \hat{\lambda} + Se(\hat{\lambda}) z_{1-\left(\frac{\gamma}{2}\right)}\right\} = 1-\gamma \quad (17)$$

Here $Se(\hat{\alpha})$ and $Se(\hat{\lambda})$ denotes standard error of α and λ respectively. $z_{\left(\frac{\gamma}{2}\right)}$ is the upper $\left(\frac{\gamma}{2}\right)$th percentile of a standard normal distribution.

Special Properties: Central Tendency Measures and measures of variability for Extended Exponential Distribution with two parameters can be given as follows. Median can be found by equation (18);

$$S(t_{;}\alpha,\lambda) = \frac{1}{2} \Rightarrow \tilde{t} = Med(T) = \frac{(ln2+1)^{1/\alpha} - 1}{\lambda} \quad (18)$$

Mean can be found by his formula:

$$E(T) = \int_0^\infty t f(t) dt = \int_0^\infty \alpha\lambda t (1+\lambda t)^{\alpha-1} e^{1-(1+\lambda t)^\alpha} dt \quad (19)$$

If we write $(1+\lambda t)^\alpha = y$ than $\alpha\lambda(1+\lambda t)^{\alpha-1} dt = dy$. For this, we can write *t* as;

$t = \frac{y^{1/\alpha}-1}{\lambda}$ and limits must change as: for $t = 0 \rightarrow y = 1$ and for $t = \infty \rightarrow y = \infty$.

$$E(T) = \frac{1}{\lambda}\int_1^\infty \left(y^{1/\alpha} - 1\right) e^{1-y} dy$$

$$= \frac{1}{\lambda} e \left[\int_1^\infty \left(y^{1/\alpha} e^{-y}\right) dy - \int_1^\infty e^{-y} dy\right]$$

$$= \frac{e}{\lambda}\left[\Gamma\left(1+\frac{1}{\alpha};1\right) + e^{-y}\Big|_1^\infty\right]$$

$$E(T) = \frac{e}{\lambda}\left[\Gamma\left(1+\frac{1}{\alpha};1\right) - 1\right] \quad (20)$$

Variance and standard deviation are found as given;

$$E(T^2) = \int_0^\infty t^2 f(t)\, dt$$

$$= \frac{1}{\lambda^2}\int_1^\infty \left(y^{1/\alpha} - 1\right)^2 e^{1-y} dy$$

$$= \frac{1}{\lambda^2}\int_1^\infty \left(y^{2/\alpha} - 2y^{1/\alpha} + 1\right) e^{1-y} dy$$

$$= \frac{1}{\lambda^2}\left[e\int_1^\infty y^{2/\alpha} e^{-y} dy - 2e\int_1^\infty y^{1/\alpha} e^{-y} dy + e\int_1^\infty e^{-y} dy\right]$$

$$= \frac{1}{\lambda^2}\left[e\Gamma\left(1+\frac{2}{\alpha};1\right) - 2e\Gamma\left(1+\frac{1}{\alpha};1\right) - ee^{-y}\Big|_1^\infty\right]$$

$$= \frac{1}{\lambda^2}\left[e\left\{\Gamma\left(1+\frac{2}{\alpha};1\right) - 2e\Gamma\left(1+\frac{1}{\alpha};1\right)\right\} + 1\right]$$

$$V(T) = E(T^2) - [E(T)]^2$$

$$V(T) = \frac{1}{\lambda^2}\left[e\left\{\Gamma\left(1+\frac{2}{\alpha};1\right) - 2e\Gamma\left(1+\frac{1}{\alpha};1\right)\right\} + 1\right] - \left\{\frac{e\Gamma\left(1+\frac{1}{\alpha};1\right) - 1}{\lambda}\right\}^2$$

$$V(T) = \frac{e}{\lambda^2}\left[\Gamma\left(1+\frac{2}{\alpha};1\right) - e\Gamma^2\left(1+\frac{1}{\alpha};1\right)\right] \quad (21)$$

where $\Gamma(.;.)$ denotes the complementary incomplete gamma function. Sometimes it is known upper limit of incomplete gamma integral. In this special case this function is calculated using $\Gamma(c;1) = \int_1^\infty t^{c-1} e^{-t} dt$, for $= 1,2,\ldots, c = \left(1+\frac{k}{\alpha}\right)$. Standard deviation is just root square of $V(T)$.

Results

Data are taken from Kalbfleisch and Prentice (1981). There are 137 survival times and eight covariates which are observed as days. Variables are performance status, disease duration (months), age (years), prior therapy, squamous cell type, small cell type, adeno cell type and treatment. There are two main groups of patients. One is standard group and the other is test groups. 69 patients appear in standard group and 68 appear in test group. Five of standard group and four of test group have censored survival times. In this study we analyses only test group patients' survival

times. For details of data it is suggested to look up Kalbfleisch and Prentice (1981).
Kolmogorov-Smirnov (K.S.) goodness-of-fit test statistic has been modified by many authors to the right censored survival time data (e.g., Fleming et al., 1980; Pepe and Fleming, 1989). When some survival times are censored the Modified K.S. test statistic is given as below:

$$D_n = \sup_{0 \le t \le t_m} |S_0(t) - S_S(t)|, \text{ for all } t, \text{ and } m \le n \quad (22)$$

In equation (22); $S_0(t)$ represent nonparametric estimate of the observed survival times and $S_S(t)$ also represent the considered distribution survival function estimate of the same survival times. Critical values can be used in making decision with ordinary one sample K.S. statistical table.
In this study we also tested performance of proposed model with log-likelihood value and besides with AIC and BIC model selection criterions. The Akaike information criterion (AIC) is a measure of the relative goodness of fit of a statistical model. It was developed by Akaike (1974). Bayesian Information Criterion (BIC) is a criterion for model selection among a class of parametric models with different numbers of parameters. It has been introduced by Schwarz (1978). The AIC and BIC is given as follows respectively;

$$AIC = -2loglik + 2p \quad (22a)$$

When data has small samples size, corrected AIC can be given;

$$AICc = -2loglik + \frac{2pn}{(n-p-1)} \quad (22b)$$

$$BIC = -2loglik + plog(n) \quad (23)$$

Where p is defined as number of free parameters.
Given a set of candidate models for the data, the preferred model is the one with the smallest value of AIC and BIC, and largest value of Log Likelihood. Reference books on model selection include Linhart and Zucchini (1986), Burnham and Anderson (2002), Miller (2002), Claeskens and Hjort (2008).

Estimation Results: In application, it is analyzed Exponential family distributions with Kolmogorov-Smirnov goodness of fit test for Lung Cancer Test Group. Results are given in Table.2.

Table 2: K.S. Type Goodness of Fit Test for Lung Cancer Test Group Results.

Name of Distribution	Test Statistics	d.f.	Important Level
Standard Exp.	0.1764	68	Not Convenient
Weibull	0.0812	68	Convenient
Log-Normal	0.0511	68	Convenient
Log-Logistic	0.0405	68	Convenient
Expo-Expo	0.1126	68	Convenient
Extended Expo	0.0566	68	Convenient

Looking to the results in Table.2, we can simply say that data is not suitable for only standard exponential distribution. Other distributions are very convenient for exponential. Secondly, modal choice results are given in Table.3. In Table.3, it can be seen that the best result for data is extended exponential distribution. AIC and BIC values are the minimum ones against other distributions and correspondingly Log-Lik and correlation values are the biggest ones.

Table 3: Modal Choice and Correlation Results for Lung Cancer Test Group Patients.

Distribution	Log-Lik	AIC	BIC	Correlation
Standard Exp.	-378.51	759.025	761.245	0.9760
Weibull	-373.84	751.682	756.121	0.9846
Log-Normal	-373.02	750.044	754.483	0.9827
Log-Logistic	-372.56	749.137	753.576	0.9918
Expo-Expo	-375.51	755.034	759.473	0.9750
Extended Expo	-372.17*	748.34*	752.779*	0.9926*

(*) The best model is extended exponential distribution for all criterions.

The results are also supported with QQ Plot given in Figure.3. Moreover correlation values given in Table.3 are calculated with QQ Plot transforms. Looking for all results, we can simply say that the best result for the data is Extended Exponential distribution.

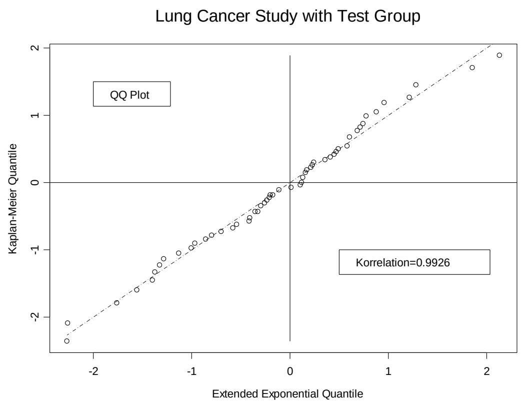

Fig. 3: QQ Plot for Lung Cancer Study with Test Group.

Therefore we estimated the parameters for point and interval estimations. Maximum Likelihood Estimates are given in Table.4. Results for parameters are very significant at level 95%.

Table 4: Results of Extended Exponential Distribution for The Lung Cancer Data Test Group.

Parameter	Value	Std. Err.	z Test	p value
$\hat{\alpha}$	0.50000	0.088938	5.622	0.000
$\hat{\lambda}$	0.02845	0.011322	2.513	0.006

Variance-Covariance matrix of Extended Exponential distribution is given as below:

$$\hat{\Sigma} = \frac{1}{|\hat{I}|} Adj(\hat{I}) = \begin{bmatrix} 0.007910018 & 0.000907702 \\ 0.000907702 & 0.000128193 \end{bmatrix}$$

95% Asymptotic Confidence Interval Estimates are given below:

$$Se(\hat{\alpha}) = 0{,}088938 \rightarrow 95\% Conf.Int.: 0{,}32568; 0{,}67431$$
$$Se(\hat{\lambda}) = 0{,}011322 \rightarrow 95\% Conf.Int.: 0{,}00625; 0{,}05064$$

As it can see, intervals are very tight and results are very satisfactory. Median, Mean and Variance of Extended Exponential distribution are found as given:

$$Mean = \hat{\mu}_T = \frac{e}{\hat{\lambda}}\left[\Gamma\left(1+\frac{1}{\hat{\alpha}};1\right)-1\right] = \frac{1}{0.02845}[e\Gamma(1+2;1)-1]$$
$$= \frac{1}{0.02845}[e(1.83939721)-1] = \frac{4}{0.02845} = 140.5975$$
$$V(T) = \frac{e}{\lambda^2}\left[\Gamma\left(1+\frac{2}{\alpha};1\right) - e\Gamma^2\left(1+\frac{1}{\alpha};1\right)\right]$$
$$= \frac{e}{\lambda^2}\left[\Gamma\left(1+\frac{2}{0.5};1\right) - e\Gamma^2\left(1+\frac{1}{0.5};1\right)\right] = \frac{40}{(0.02845^2)}$$
$$= 49416.17 \Longrightarrow \sqrt{\hat{V}(T)} = 222.3040$$

Considering hazard functions for all distribution which analyzed in this study; median and parameter estimation results are given in Table.5.

Table 5: Results of Lung Cancer Test Group Patients.

Distribution	Median	Scale Param.	Shape Param.
Standard Exp.	94.4196	$\hat{\lambda}$=0.007341	None
Weibull	72.5175	$\hat{\lambda}$=0.0085582	$\hat{\alpha}$=0.76831
Log-Normal	58.7148	$\hat{\mu}$ =4.072692	$\hat{\sigma}$ =1.46916
Expo-Expo	81.9840	$\hat{\lambda}$=0.005603	$\hat{\alpha}$=0.694
Log-Logistic	60.8061	$\hat{\lambda}$=0.016445	$\hat{\alpha}$=1.21849
Extended Expo	65.6150	$\hat{\lambda}$=0.02845	$\hat{\alpha}$=0.5

Results and Discussion

Survivalanalysis is one of the important discipline of statistics which deals with death in biological organisms and failure in mechanical systems. In the survival analysis, one of the important functions is the survival function. For this reason, the estimation of a survival function is also very important case for survival literature. Both point estimation and confidence interval estimation of the survival function may be achieved by fitting parametric distributions. In this study we investigate, parameter and confidence interval estimation of the extended exponential distribution from right censored survival time data.

Firstly; in the analysis of survival times, hazard function has very important issue. In general, hazard function has three main structures in survival analysis: Increasing, decreasing and constant or combination of all types. In respect to hazard function of Extended Exponential distribution has all type of structures; it will be useful and practical distribution in survival analysis as Weibull distribution which found a large application area in survival literature.

Secondly, in point of Extended Exponential distribution has accelerated proportional hazard, it can be used in regression analysis as well. On the other hand; because of Extended Exponential distribution has no nice representation of parameter estimators, it may limit usefulness of this distribution in regression analysis. Thirdly, it is a useful distribution for researchers to study in survival analysis or reliability theory.

In Table.3 we conclude that the survival times of test group have decreasing hazard (extended exponential) or uni-modal hazard (log-logistic) ratios. Further, looking to Table.2 it is easily seen that the best fit for this data is extended exponential distribution. As a result, decreasing hazard ratio is considerable for the survival times of test group.

Extended Exponential distribution results for Lung Cancer Data gave the best result against other exponential family distributions which analyzed in this study. This shows that Extended Exponential distribution can find a large application area in survival analysis or reliability theory.

References

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control19 (6): 716–723.

Alakus, K., 1997. Contributions to Statistical Estimation and Testing in Survival Analysis Problems. Unpublished PhD Thesis, University of Strathclyde, The UK.

Alakuş, K.,Erilli, N. A., Öner Y., 2015. Confidenceintervalsestimationforsurvivalfunction in weibuldistributionbased on censoredsurvival time data. Erciyes UniversityJournal of ScienceandTechonology, No:31-2, pp. 489-493.

Burnham, P. K., Anderson, D. R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd edition). Springer-Verlag, New York.

Claeskens, G., Hjort, N. L., 2008. Model Selection and Model Averaging. Cambridge University Press, Cambridge.

Fleming, T. R., O'Fallon, J. R., O'Brien, P. C., Harrington, D. P., 1980. Modified Kolmogorov-Smirnov Test Procedures with Application to Arbitrarily Right-Censored Data.Biometrics 36, 607-625.

Gupta R.D., Kundu D., 2001.Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions. Bimetrical Journal 43, 117-130.

Gupta R.D., Kundu D., 2007.Generalized Exponential Distribution: Existing Results and Some Recent Development.Journal of Statistical Planning and Inference. 137, 3537-3547.

Haghighi F., Sadeghi S., 2009. An Exponential Extension. 41èmes Journées de Statistique

Kalbfleisch, J.D., Prentice, R.L., 1981. The Statistical Analysis of Failure Time Data, Wiley, New York.

Linhart, H., Zucchini, W., 1986. Model Selection.Wiley, New York.

Miller, A. J., 2002. Subset Selection in Regression (2nd edition).Chapman and Hall/CRC, Boca Raton, Florida.

Nadarajah S., Haghighi F., 2010. An Extension of the Exponential Distribution. Statistics: A Journal of Theoretical and Applied Statistics.2010, 1-10.

Nelson, W., 1969.Hazard Plotting for Incomplete Failure Data.J.Qual. Technol. 1,27-52.

Nelson, W., 1972.Theory and Applications of Hazard Plotting for Censored Failure Data.Technometrics 14, 945-965.

Pepe, M. S., Fleming, T. R., 1989. Weighted Kaplan-Meier Statistics: A Class of Distance Tests for Censored Survival Data. Biometrics 45, 497-507.

Schwarz, G. E., 1978. Estimating the Dimension of a Model. Annals of Statistics6 (2): 461–464.