Discriminating among Weibull, log-normal and log-logistic distributions

M. Z. Raqab[a,b*], S. Al-Awadhi[a], Debasis Kundu[c]

[a] *Department of Statistics and OR, Kuwait University, Safat 13060, Kuwait*
[b] *Department of Mathematics, The University of Jordan, Amman 11942, Jordan*
[c] *Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA*

Abstract

In this paper we consider the problem of the model selection/ discrimination among three different positively skewed lifetime distributions. All these three distributions namely; the Weibull, log-normal and log-logistic, have been used quite effectively to analyze positively skewed lifetime data. In this paper we have used three different methods to discriminate among these three distributions. We have used the maximized likelihood method to choose the correct model and computed the asymptotic probability of correct selection. We have further obtained the Fisher information matrices of these three different distributions and compare them for complete and censored observations. These measures can be used to discriminate among these three distributions. We have also proposed to use the Kolmogorov-Smirnov distance to choose the correct model. Extensive simulations have been performed to compare the performances of the three different methods. It is observed that each method performs better than the other two for some distributions and for certain range of parameters. Further, the loss of information due to censoring are compared for these three distributions. The analysis of a real data set has been performed for illustrative purposes.

Keywords: Likelihood ratio method; Fisher information matrix; probability of correct selection; percentiles; model selection; Kolmogorov-Smirnov distance.

AMS Subject Classification (2010): 62E15; 62F10; 62E25.

1 Introduction

Among several right skewed distributions, the Weibull (WE), log-normal (LN) and log-logistic (LL) distributions have been used quite effectively in analyzing positively skewed lifetime data. These three distributions have several interesting distributional properties and their probability density functions also can take different

* *Corresponding author*
E-mail addresses: mraqab@ju.edu.jo (M. Raqab), alawadhidodo@yahoo.com (S. Al-Awadhi), kundu@iitk.ac.in (D. Kundu).

shapes. For example, the WE distribution can have a decreasing or an unimodal probability density function (PDF), and a decreasing, a constant and an increasing hazard function, depending on the shape parameter. Similarly, the PDF of a LN density function is always unimodal and it has an inverted bathtub shaped hazard function. Moreover, the LL distribution has either a reversed J shaped or an unimodal PDF, and the hazard function of the LL distribution is either a decreasing or an inverted bathtub shaped. For further details about the distributional behaviors of these distributions, one may refer to Johnson et al. (1995).

Let us consider the following problem. Suppose $\{x_1, \ldots, x_n\}$ is a random sample of size n from some unknown lifetime distribution function $F(\cdot)$, i.e. $F(0-) = 0$, and the preliminary data analysis suggests that it is coming from a positively skewed distribution. Hence, any one of the above three distributions can be used to analyze this data set. In this paper, we would like to explore among these WE, LN and LL distributions, which one fits the data 'best'. It can be observed that for certain ranges of the parameters, the corresponding PDFs or the cumulative distribution functions(CDFs) are very close to each other but can be quite different with respect to other characteristics. Before explaining this with an example let us introduce the following notations.

The WE distribution with the shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$ will be denoted by $WE(\alpha, \lambda)$. The corresponding PDF and CDF for $x > 0$, are

$$f_{WE}(x; \alpha, \lambda) = \alpha\lambda^{\alpha} x^{\alpha-1} e^{-(\lambda x)^{\alpha}} \quad \text{and} \quad F_{WE}(x; \alpha, \lambda) = 1 - e^{-(\lambda x)^{\alpha}},$$

respectively. The LN distribution is denoted by $LN(\sigma, \beta)$ with the shape parameter $\sigma > 0$ and scale parameter $\beta > 0$. The PDF and CDF of this distribution for $x > 0$, can be written as

$$f_{LN}(x; \sigma, \beta) = \frac{1}{\sqrt{2\pi}\ \sigma\ x} e^{-\frac{1}{2\sigma^2}(ln\ x - ln\ \beta)^2}$$

and

$$F_{LN}(x; \sigma, \beta) = \Phi(\frac{ln\ x - \ln\ \beta}{\sigma}) = \frac{1}{2} + \frac{1}{2}\ Erf(\frac{ln\ x - \ln\ \beta}{\sqrt{2}\ \sigma}),$$

respectively, where $\Phi(.)$ is the CDF of a standard normal distribution with $Erf(x) = 2\Phi(\sqrt{2}\ x) - 1$. The PDF and CDF of the LL distribution, denoted by $LL(\gamma, \xi)$, with the shape parameter $\gamma > 0$ and scale parameter $\xi > 0$, for $x > 0$, are

$$f_{LL}(x; \gamma, \xi) = \frac{1}{\gamma\ x} \frac{e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}{(1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}})^2} \quad \text{and} \quad F_{LL}(x; \gamma, \xi) = 1 - \frac{1}{1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}},$$

respectively.

In Figure 1, we have plotted the CDFs of WE(4.18,0.56), LN(0.27,1.60) and LL(0.16,1.60). It is clear from Figure 1 that all the CDFs are very close to each other. Therefore, if the data are coming from any one of these three distributions, the other two distributions can easily be used to analyze this data set. Although, these three CDFs are quite close to each other, the hazard functions of the above three distribution functions, see Figure 2, are completely different. Moreover, the

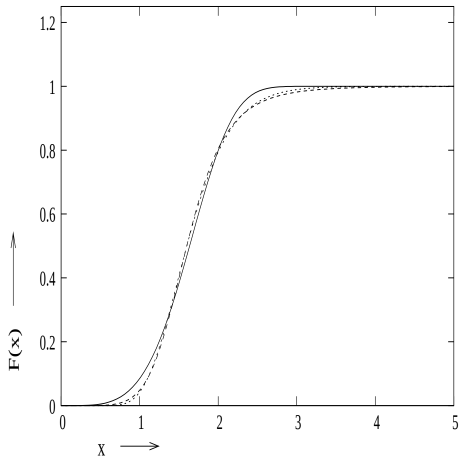

Figure 1: The CDFs of WE(4.18,0.56), LN(0.27,1.60) and LL(0.16,1.60).

5-th percentile points of WE(4.18,0.56), LN(0.27,1.60) and LL(0.16,1.60), are 0.8779, 1.1655 and 1.0053, respectively. Clearly, the 5-th percentile points of these three distributions are also significantly different. On the other hand if we choose the correct model based on the maximized likelihood ratio (the details will be explained later), and the data are obtained from a LL distribution, then the probability of correct selection (PCS) for different sample sizes are presented in Table 1. It can be seen that the PCS is as small as only 0.33 when the sample size is 20. Therefore, it is clear that choosing the correct model is a challenging problem, particularly when the sample size is small.

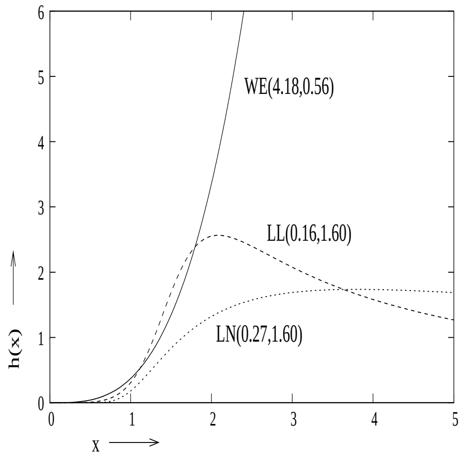

Figure 2: The hazard functions of WE(4.18,0.56), LN(0.27,1.60) and LL(0.16,1.60).

It is clear that choosing the correct model is an important problem as the effect of mis-classification can be quite severe as we have seen in case of the hazard

Table 1: PCS based on Monte Carlo simulations using ratio of maximized likelihood

Sample size	20	30	50	100	500
PCS	0.330	0.390	0.525	0.668	0.982

function or for the percentile points. This issue would be more crucial when the sample sizes are small or even moderate. Therefore, the discrimination problem between different distributions received a considerable attention in the last few decades.

Cox (1962) first addressed the problem of discriminating between the LN and the exponential distributions based on the likelihood function and derived the asymptotic distribution of the likelihood ratio test statistic. Since then extensive work has been done in discriminating among different distributions. Some of the recent work regarding discriminating between different lifetime distributions can be found in Alshunnar et al.(2010), Pakyari (2012), Elsherpieny et al. (2013), Sultan and Al-Moisheer (2013), Ahmad et al. (2016) and the references cited therein. Although, extensive work has been done in discriminating between two distributions not much work has been done when more than two distributions are present, except the work of Marshall et al. (2001) and Dey and Kundu (2009). Moreover, most of the work till today are based on the likelihood ratio test.

The aim of this paper is two fold. First of all we derive the Fisher information matrices of these three distributions and obtain different Fisher information measures both for the complete and censored samples. We also provide the loss of information due to truncation for these three different distributions. It is observed that the Fisher information measure can be used in discriminating purposes. Our second aim of this paper is to compare three different methods namely (i) the method based on the Fisher information measures, (ii) the method based on the likelihood ratio and (iii) the method based on the Kolmogorov-Smirnov distance, in discriminating among these three different distributions. We perform extensive simulation experiments to compare different methods for different sample sizes and for different parameter values. It is observed that the performance of each method depends on the true underlying distribution and the set of parameter values.

Rest of the paper is organized as follows. In Section 2, we derive the Fisher information measures for complete sample for all the three cases and show that how they can be used for discrimination purposes. In Section 3, we provide the discrimination procedure based on likelihood ratio statistics and derive their asymptotic properties. Monte Carlo simulation results are presented in Section 4. In Section 5, we provide the Fisher information measures for censored samples and the loss of information due to truncation. The analysis of a data set is presented in Section 6, and finally we conclude the paper in Section 7.

2 FI measure for complete sample

Let $X > 0$ be a continuous random variable with PDF and CDF as $f(x;\boldsymbol{\theta})$ and $F(x;\boldsymbol{\theta})$, respectively, where $\boldsymbol{\theta} = (\theta_1, \theta_2)$ is a vector parameter. Under the standard regularity conditions, see Lehmann(1991), the FI matrix for the parameter vector $\boldsymbol{\theta}$ is

$$I(\boldsymbol{\theta}) = E\left(\begin{bmatrix} \frac{\partial}{\partial\theta_1}\ln f(X;\boldsymbol{\theta}) \\ \\ \frac{\partial}{\partial\theta_2}\ln f(X;\boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial\theta_1}\ln f(X;\boldsymbol{\theta}) & \frac{\partial}{\partial\theta_2}\ln f(X;\boldsymbol{\theta}) \end{bmatrix}\right).$$

In this section, we present the FI measures for the WE, LN and LL distributions based on a complete data. The FI matrices of WE and LN (see, for example, Alshunnar et al. (2010) and Ahmad et al. (2016)) can be described as follows:

$$I_W(\alpha, \lambda) = \begin{pmatrix} f_{11W} & f_{12W} \\ f_{21W} & f_{22W} \end{pmatrix} \text{ and } I_N(\sigma^2, \beta) = \begin{pmatrix} f_{11N} & f_{12N} \\ f_{21N} & f_{22N} \end{pmatrix},$$

where

$$f_{11W} = \frac{1}{\alpha^2}\left(\psi'(1) + \psi^2(2)\right),\ f_{12W} = f_{21W} = \frac{1}{\lambda}\left(1 + \psi(1)\right),\ f_{22W} = \frac{\alpha^2}{\lambda^2},$$

and

$$f_{11N} = \frac{2}{\sigma^2},\ f_{12N} = f_{21N} = 0,\ f_{22N} = \frac{1}{\beta^2\,\sigma^2}.$$

Here $\psi(x) = \Gamma'(x)/\Gamma(x)$ and $\psi'(x)$ are the psi(or digamma) and tri-gamma functions, respectively, with $\Gamma(.)$ being the complete gamma function. The FI matrix for LL distribution based on a complete data in terms of the parameters γ and ξ is presented in Theorem 1 given below. The proof can be seen easily via differentiation techniques and straight-forward algebra.

Theorem 1: *The FI matrix for the LL distribution is*

$$I_L(\gamma, \xi) = \begin{pmatrix} f_{11L} & f_{12L} \\ f_{21L} & f_{22L} \end{pmatrix},$$

where

$$f_{11L} = \frac{1}{3\gamma^2}\left(1 + \frac{\pi^2}{3}\right),\ \ f_{22L} = \frac{1}{3\gamma^2\xi^2},\ \ \textit{and}\ \ f_{12L} = f_{21L} = 0.$$

Proof: See in the Appendix.

In the arguments similar to the E-optimality of the design of experiment problems, we consider the trace of the FI matrix as a measure of the total information in the data about the parameters involved in a specific model. For example, the trace of the FI matrix of the WE distribution is the sum of the information measure of α when λ is known and λ when α is known. The traces of the FI matrix of the LN and LL are defined similarly. In spite of the fact that the shape and scale parameters are essential tools in many distributional properties, these parameters do not characterize the same prominent distributional features of the corresponding distributions. For comparison purposes of distributional characteristics of the three distributions, we evaluate the asymptotic variances of the percentile estimators for these distributions. In our case, the the p-th $(0 < p < 1)$ percentiles of the WE, LN and LL distributions are respectively,

$$P_{WE}(\alpha,\lambda) = \frac{1}{\lambda}\left(-ln(1-p)\right)^{1/\alpha},\ P_{LN}(\beta,\sigma) = \beta e^{\sigma\,\Phi^{-1}(p)},\ P_{LL}(\gamma,\xi) = \xi\left(\frac{p}{1-p}\right)^{\gamma}.$$

Therefore, $Var_{WE}(p)$, $Var_{LN}(p)$,$Var_{LL}(p)$, the asymptotic variances of the logarithm of the p-th percentile estimators of the WE, LN and LL distributions, respectively, can be written as follows:

$$Var_{WE}(p) = \begin{bmatrix}\frac{\partial ln\,P_{WE}}{\partial\alpha} & \frac{\partial ln\,P_{WE}}{\partial\lambda}\end{bmatrix}\begin{bmatrix}f_{11W} & f_{12W}\\ f_{21W} & f_{22W}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial ln\,P_{WE}}{\partial\alpha}\\ \frac{\partial ln\,P_{WE}}{\partial\lambda}\end{bmatrix}, \tag{1}$$

$$Var_{LN}(p) = \begin{bmatrix}\frac{\partial ln\,P_{LN}}{\partial\sigma} & \frac{\partial ln\,P_{LN}}{\partial\beta}\end{bmatrix}\begin{bmatrix}f_{11N} & f_{12N}\\ f_{21N} & f_{22N}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial ln\,P_{LN}}{\partial\sigma}\\ \frac{\partial ln\,P_{LN}}{\partial\beta}\end{bmatrix}, \tag{2}$$

and

$$Var_{LL}(p) = \begin{bmatrix}\frac{\partial ln\,P_{LL}}{\partial\gamma} & \frac{\partial ln\,P_{LL}}{\partial\xi}\end{bmatrix}\begin{bmatrix}f_{11L} & f_{12L}\\ f_{21L} & f_{22L}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial ln\,P_{L}}{\partial\gamma}\\ \frac{\partial ln\,P_{L}}{\partial\xi}\end{bmatrix}. \tag{3}$$

The asymptotic variances for the median or the 95-th percentile estimators of the three distributions can be used for comparison purposes. Using the average asymptotic variance with respect to probability measure $W(.)$ proposed by Gupta and Kundu (2006), we compare the following measures:

$$AV_{WE} = \int_0^1 Var_{WE}(p)\,dW(p),\ AV_{LN} = \int_0^1 Var_{LN}(p)\,dW(p),$$

and

$$AV_{LL} = \int_0^1 Var_{LL}(p)\,dW(p),$$

where $W(.)$ is a weighted function such that $\int_0^1 W(p)\,dp = 1$. For more convenience, one may consider the average asymptotic variances of all percentile estimators, that is, $W(p) = 1, 0 < p < 1$.

To conduct a comparative study of the total information measure between any two distributions, we have to compute these measures at their closest values. One way to define the closeness (distance) between two distributions is to use the Kullback-Leibler (KL) distance (see, for example, White (1982, Theorem 1), Kundu and Manglick (2004)). For notational convenience, let $\tilde{\theta_1}$ and $\tilde{\theta_2}$ be the miss-classified parameters of F_1 distribution for given δ_1 and δ_2 of F_2 distribution so that $F_1(\tilde{\theta_1}, \tilde{\theta_2})$ is closest to $F_2(\delta_1, \delta_2)$ in terms of the Kullback-Leibler distance. Lemma 1 below provides the estimates of the parameters where any two distributions among WE, LN and LL are closest to each other. Further, the maximum likelihood estimates (MLEs) of θ_1 and θ_2 are denoted by $\hat{\theta_1}$ and $\hat{\theta_2}$.

Lemma 1 [Kundu and Manglick (2004)]:
(i) If the the underlying distribution is WE(α, λ), then the closest LN distribution in terms of KL distance is LN$(\tilde{\sigma}, \tilde{\beta})$, where

$$\tilde{\sigma} = \frac{\sqrt{\psi'(1)}}{\alpha}, \text{ and } \tilde{\beta} = \frac{1}{\lambda}\, e^{\frac{\psi(1)}{\alpha}}. \tag{4}$$

(ii) If LN(σ, β) is the valid distribution then the closest WE distribution in terms of KL distance is $WE(\tilde{\alpha}, \tilde{\lambda})$ such that

$$\tilde{\alpha} = \frac{1}{\sigma}, \text{ and } \tilde{\lambda} = \frac{1}{\beta}\, e^{\frac{-\sigma}{2}}. \tag{5}$$

Lemma 2 [Dey and Kundu (2009)]:
(i) Under the assumption that the data are coming from LN(σ, β) and for $n \to \infty$, we have $\hat{\gamma} \to \tilde{\gamma}$, a.s., and $\hat{\xi} \to \tilde{\xi}$, a.s., where

$$E_{LN}\left(ln\ f_{LL}(X; \tilde{\gamma}, \tilde{\xi})\right) = \max_{\gamma,\xi} E_{LN}\left(ln\ f_{LL}(X; \gamma, \xi)\right).$$

(ii)Under the assumption that the data are from LL(γ, ξ), we have the following results as $n \to \infty$, $\hat{\sigma} \to \tilde{\sigma}$, a.s., and $\hat{\beta} \to \tilde{\beta}$, a.s., where

$$E_{LL}\left(ln\ f_{LN}(X; \tilde{\sigma}, \tilde{\beta})\right) = \max_{\sigma,\beta} E_{LL}\left(ln\ f_{LN}(X; \sigma, \beta)\right).$$

In fact, Dey and Kundu (2009) have shown that for $\sigma = \beta = 1$, $E_{LN}(ln\ f_{LL}(X; \gamma, \xi)$ is maximized when $\tilde{\gamma} = 0.5718$ and $\tilde{\xi} = 1$, while when $\gamma = \xi = 1$, the maximization of $E_{LL}(ln\ f_{LN}(X; \sigma, \beta)$ occurs at $\tilde{\sigma} = \sqrt{3}$ and $\tilde{\beta} = 1$. In general if the data are coming from LN(σ, β), then $\tilde{\gamma}$ and $\tilde{\xi}$ can be obtained by maximizing

$$\begin{aligned} E_{LN}(ln\ f_{LL}(X; \gamma, \xi) &= E_{LN}\left\{-ln\ \gamma + \frac{lnX - ln\xi}{\gamma} - 2ln\left(1 + e^{\frac{lnX - ln\xi}{\gamma}}\right)\right\} \\ &= -ln\gamma + \frac{ln\beta - ln\xi}{\gamma} - 2E_N\left[ln\left(1 + (\frac{\beta}{\xi})^{\frac{1}{\gamma}}\ e^{\frac{\sigma}{\gamma}X}\right)\right], \end{aligned}$$

where $E_N(.)$ stands for the expectation under standard normal distribution. They do not have explicit forms and they need to be obtained numerically.

To obtain $\tilde{\sigma}$ and $\tilde{\beta}$ such that $LL(\gamma,\xi)$ is closest to $LN(\tilde{\sigma},\tilde{\beta})$, we have to maximize

$$E_{LL}(ln\ f_{LN}(X;\sigma,\beta) = -\frac{ln(2\pi)}{2} - ln\ \sigma - E_{LL}(ln\ X) - \frac{1}{2\sigma^2}\ E_{LL}\left(ln\ X - ln\ \beta\right)^2.$$

It is easily seen that $E_{LL}(lnX) = ln\xi$. By applying Taylor's series $-ln(1-z) = \sum_{j=1}^{\infty} z^j/j$, it readily follows

$$\int_0^1 lnz\ ln(1-z) = \sum_{j=1}^{\infty}\frac{1}{j(j+1)^2} = 2 - \frac{\pi^2}{6}.$$

Summing up, the most simplified function to be maximized is

$$E_{LL}(ln\ f_{LN}(X;\sigma,\beta) = -\frac{ln(2\pi)}{2} - ln\ \sigma - ln\ \xi - \frac{1}{2\sigma^2}\left\{(ln\xi - ln\beta)^2 + \frac{\pi^2\gamma^2}{3}\right\}.$$

It can be easily verified that $\tilde{\beta} = \xi$ and $\tilde{\sigma} = \pi\gamma/\sqrt{3}$ maximize $E_{LL}(ln\ f_{LN}(X;\sigma,\beta)$. Similarly as in Lemma 1 and Lemma 2, we have the following Lemma related to the WE and LL distributions.

Lemma 3:
(i) Suppose the data come from $WE(\alpha,\lambda)$ then as $n\to\infty$, $\hat{\gamma}\to\tilde{\gamma}$, a.s. and $\hat{\xi}\to\tilde{\xi}$, a.s., where

$$E_{WE}\left(ln\ f_{LL}(X;\tilde{\gamma},\tilde{\xi})\right) = \max_{\gamma,\xi} E_{WE}\left[ln\ f_{LL}(X;\gamma,\xi)\right].$$

(ii) If the data come from $LL(\gamma,\xi)$, then as $n\to\infty$, we have $\hat{\alpha}\to\tilde{\alpha}$, a.s. and $\hat{\lambda}\to\tilde{\lambda}$, a.s., where

$$E_{LL}\left(ln\ f_{WE}(X;\tilde{\alpha},\tilde{\lambda})\right) = \max_{\alpha,\lambda} E_{LL}\left[ln\ f_{WE}(X;\alpha,\lambda)\right].$$

Proof: The proof is followed along the lines of Dey and Kundu (2009). ∎
(i) To find $\tilde{\gamma}$ and $\tilde{\xi}$, let us define

$$\begin{aligned} g(\gamma,\xi) &= E_{WE}\left(ln\ f_{LL}(X;\gamma,\xi)\right) \\ &= -ln\ \gamma - \left(1-\frac{1}{\gamma}\right)\ E_{WE}(ln\ X) - \frac{ln\ \xi}{\gamma} - 2E_{WE}\ ln\left(1+\left(\frac{X}{\xi}\right)^{\frac{1}{\gamma}}\right). \end{aligned} \tag{6}$$

The second term in (6) is evaluated to be $E_{WE}(ln\ X) = -ln\ \lambda + \dfrac{\psi(1)}{\alpha}$, where $C = -\psi(1) = 0.57721...$, is the Euler's constant, while the last term can be rewritten as

$$E_{WE}\ ln\left(1+\left(\frac{X}{\xi}\right)^{\frac{1}{\gamma}}\right) = \int_0^{\infty} ln\left(1+\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}\right)e^{-y}\ dy.$$

Consequently, Eq. (6) can take the following simplified form

$$\begin{aligned} g(\gamma,\xi) &= -ln\gamma - \left(1-\frac{1}{\gamma}\right)\left(-ln\lambda + \frac{\psi(1)}{\alpha}\right) - \frac{ln\xi}{\gamma} - \\ &\qquad 2\int_0^{\infty} ln\left(1+\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}\right) e^{-y}dy. \end{aligned} \tag{7}$$

For $\tilde{\xi}$, it can be obtained directly by differentiating (7) with respect to ξ and equating the resulting expression to 0. Then, $\tilde{\xi}$ can be found by solving the following equation numerically,

$$\int_0^{\infty} \frac{\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}}{1+\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}} e^{-y}dy = \frac{1}{2}. \tag{8}$$

Upon differentiating (7) with respect to γ, using (8) and equating the resulting expression to 0, we compute $\tilde{\gamma}$ as a numerical solution to

$$\frac{2}{\alpha}\int_0^{\infty} \frac{lny\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}}{1+\left(\frac{y^{\frac{1}{\alpha}}}{\lambda\xi}\right)^{\frac{1}{\gamma}}} e^{-y}dy - \gamma - \frac{\psi(1)}{\alpha} = 0, \tag{9}$$

(ii) For given γ and ξ, we maximize $E_{LL}[ln\ f_{WE}(X;\alpha,\lambda)]$, with respect to α and λ. Let

$$\begin{aligned} h(\alpha,\lambda) &= E_{LL}\left[ln\ f_{WE}(X;\alpha,\lambda)\right] \\ &= ln\ \alpha + \alpha ln\ \lambda + (\alpha-1)E_{LL}(lnX) - \lambda^{\alpha}E_{LL}(X^{\alpha}). \end{aligned}$$

By using the facts, $E_{LL}(lnX) = ln\ \xi$ and $E_{LL}(X^{\alpha}) = \xi^{\alpha}B(1-\alpha\gamma, 1+\alpha\gamma)$, $\alpha\gamma < 1$, we have

$$h(\alpha,\lambda) = ln\ \alpha + \alpha ln\lambda + (\alpha-1)ln\ \xi - \lambda^{\alpha}\ \xi^{\alpha}\ B(1-\alpha\gamma, 1+\alpha\gamma),$$

where $B(a,b) = \int_0^1 w^{a-1}(1-w)^{b-1}\ dw = \Gamma[a]\Gamma[b]/\Gamma[a+b]$ is beta function. It can be checked that

$$\frac{\partial}{\partial\alpha}B(1-\alpha\gamma,\alpha\gamma+1) = \gamma B(1-\alpha\gamma,\alpha\gamma+1)\left[\psi(1+\alpha\gamma)-\psi(1-\alpha\gamma)\right].$$

Arguments similar to those in (i), we find $\tilde{\alpha}$ and $\tilde{\lambda}$ for which $LL(\gamma,\xi)$ is closest to $WE(\tilde{\alpha},\tilde{\lambda})$ by solving the following normal equations:

$$\tilde{\lambda} = \frac{1}{\xi\left[B(1-\alpha\gamma, 1+\alpha\ \gamma)\right]^{1/\alpha}}, \quad \text{and} \quad \psi(1+\alpha\ \gamma) - \psi(1-\alpha\ \gamma) = \frac{1}{\gamma\alpha}. \tag{10}$$

Table 2: The TI and TV of FI matrices of $LN(\tilde{\sigma}, \tilde{\beta})$, $LL(\tilde{\gamma}, \tilde{\xi})$ and $WE(\alpha, 1)$ for different α

$\alpha \downarrow$	$\tilde{\sigma}$	$\tilde{\beta}$	$\tilde{\gamma}$	$\tilde{\xi}$	TI_{WE}	TI_{LN}	TI_{LL}	TV_{WE}	TV_{LN}	TV_{LL}
0.5	2.5651	0.3152	1.3918	0.3965	7.5447	1.8337	1.8328	4.5866	3.9436	2.2683
0.8	1.6032	0.4860	0.8699	0.5609	3.4895	2.4254	3.2898	2.1214	1.8922	1.2434
1.0	1.2826	0.5615	0.6959	0.6297	2.8237	3.1438	4.6886	1.7166	1.3412	0.9147
1.2	1.0688	0.6182	0.5799	0.6802	2.7064	4.0414	6.3946	1.6453	1.0077	0.7019
1.4	0.9161	0.6621	0.4971	0.7187	2.8905	5.1012	8.3983	1.7572	0.7875	0.5557
1.6	0.8016	0.6971	0.4349	0.7490	3.2724	6.3151	10.7019	1.9894	0.6335	0.4506
1.8	0.7125	0.7257	0.3866	0.7734	3.8029	7.6799	13.2961	2.3119	0.5212	0.3727
2.0	0.6413	0.7493	0.3480	0.7935	4.4559	9.1938	16.1791	2.7089	0.4365	0.3134

Here, we consider the trace of the information matrix to discriminate between any two distributions. In this case the information content is the sum of the FI measure of the shape parameter assuming the scale parameter to be known and the FI measure of the scale parameter assuming the shape parameter to be known. Another related measure is the the trace of the inverse of the FI information matrix, which is also known in the statistical literature as the sum of the asymptotic variances of the MLEs of the shape and scale parameters. To conduct a comparison between the total information (TI) and total variance (TV) measures between any two distributions, it is appropriate to do so at the points where both distributions are closest to each other. The TI and TV for all the three distributions when the data come from WE, LN and LL are reported in Tables 2, 3 and 4, respectively. It is clearly observed from these tables that the corresponding FI contents are quite different even the distribution functions are closest in terms of the Kullback-Leibler distance. It is obvious that the parametric values are essential in estimating the FI quantities computed in these tables. We will point out in Section 4 how this affects on choosing the correct model.

Now we would like to use the Fisher information for discriminating among these three distributions. Let $\{x_1, \ldots, x_n\}$ be a random sample from one of these three distribution functions and the observed total information measures for the WE, LN and LL be denoted by $TI_{WE}(\hat{\alpha}, \hat{\lambda})$, $TI_{LN}(\hat{\sigma}, \hat{\beta})$ and $TI_{LL}(\hat{\gamma}, \hat{\xi})$, respectively. Consider the following statistics:

$$\begin{aligned} FD_1 &= TI_{WE}(\hat{\alpha}, \hat{\lambda}) - TI_{LN}(\hat{\sigma}, \hat{\beta}), \quad FD_2 = TI_{WE}(\hat{\alpha}, \hat{\lambda}) - TI_{LL}(\hat{\gamma}, \hat{\xi}), \\ FD_3 &= TI_{LN}(\hat{\sigma}, \hat{\beta}) - TI_{LL}(\hat{\gamma}, \hat{\xi}). \end{aligned} \tag{11}$$

Based on (11), we choose WE if $FD_1 > 0$ and $FD_2 > 0$, LN if $FD_1 < 0$ and

Table 3: The TI and TV of FI matrices of $WE(\tilde{\alpha}, \tilde{\lambda})$, $LL(\tilde{\gamma}, \tilde{\xi})$ and $LN(\sigma, 1)$ for different σ

$\sigma^2 \downarrow$	$\tilde{\alpha}$	$\tilde{\lambda}$	$\tilde{\gamma}$	$\tilde{\xi}$	TI_{LN}	TI_{WE}	TI_{LL}	TV_{LN}	TV_{WE}	TV_{LL}
0.5	1.4142	0.7022	0.4043	1.0000	6.0000	4.9679	10.7874	0.7500	1.4892	0.6047
0.8	1.1180	0.6394	0.5115	1.0000	3.7500	4.5163	6.7396	1.2000	1.1225	0.9679
1.0	1.0000	0.6065	0.5718	1.0000	3.0000	4.5422	5.3931	1.5000	1.0157	1.2095
1.2	0.9129	0.5783	0.6264	1.0000	2.5000	4.6802	4.4939	1.8000	0.9515	1.4515
1.4	0.8452	0.5534	0.6766	1.0000	2.1429	4.8855	3.8518	2.1000	0.9096	1.6935
1.6	0.7906	0.5313	0.7233	1.0000	1.8750	5.1320	3.3704	2.4000	0.8807	1.9354
1.8	0.7454	0.5113	0.7672	1.0000	1.6667	5.4076	2.9958	2.7000	0.8594	2.1774
2.0	0.7071	0.4931	0.8087	1.0000	1.5000	5.7038	2.6962	3.0000	0.8431	2.4193

Table 4: The TI and TV of FI matrices of $WE(\tilde{\alpha}, \tilde{\lambda})$, $LN(\tilde{\sigma}, \tilde{\beta})$ and $LL(\gamma, 1)$ for different γ

$\gamma \downarrow$	$\tilde{\alpha}$	$\tilde{\lambda}$	$\tilde{\sigma}$	$\tilde{\beta}$	TI_{LL}	TI_{WE}	TI_{LN}	TV_{LL}	TV_{WE}	TV_{LN}
0.1	5.0000	0.9136	0.1814	1.0000	176.3290	30.0251	91.1689	0.0370	15.2352	0.0494
0.3	1.6667	0.7627	0.5441	1.0000	19.5921	5.4319	10.1336	0.3329	1.9209	0.4441
0.5	1.0000	0.6366	0.9069	1.0000	7.0532	4.2912	3.6476	0.9248	1.0572	1.2337
0.8	0.6250	0.4855	1.4510	1.0000	2.7551	6.3259	1.4249	2.3676	0.9065	3.1580
1.0	0.5000	0.4053	1.8138	1.0000	1.7633	8.8166	0.9119	3.6993	0.8805	4.9348
1.2	0.4167	0.3383	2.1766	1.0000	1.2245	12.0199	0.6332	5.3270	0.8363	7.1064
1.4	0.3571	0.2824	2.5393	1.0000	0.8996	15.9001	0.4653	7.2507	0.7709	9.6721
1.6	0.3125	0.2357	2.9021	1.0000	0.6888	20.4323	0.3562	9.4703	0.6901	12.6333
1.8	0.2778	0.1968	3.2648	1.0000	0.5442	25.6237	0.2815	11.9858	0.6033	15.9884
2.0	0.2500	0.1643	3.6276	1.0000	0.4408	31.4942	0.2280	14.7973	0.5168	19.7392

$FD_3 > 0$, LL if $FD_2 < 0$ and $FD_3 < 0$, as the preferred distribution. The respective PCSs are defined as follows;

$$
\begin{aligned}
PCS_{WE}^{FI} &= P(FD_1 > 0, FD_2 > 0 | \text{data follow WE}), \\
PCS_{LN}^{FI} &= P(FD_2 < 0, FD_3 > 0 | \text{data follow LN}), \\
PCS_{LL}^{FI} &= P(FD_2 < 0, FD_3 < 0 | \text{data follow LL}).
\end{aligned}
$$

3 Ratio of Maximized Likelihood Method

In this section we describe the discrimination procedure based on the ratio of maximized likelihood method. It is assumed that the sample $\{x_1, \ldots, x_n\}$ has been obtained from one of these three distributions; namely WE(α, λ), LN(σ, β) or LL(γ, ξ), and the corresponding likelihood functions are

$$
\begin{aligned}
L_{WE}(\alpha, \lambda) &= \prod_{i=1}^{n} f_{WE}(x_i; \alpha, \lambda), \qquad L_{LN}(\sigma, \beta) = \prod_{i=1}^{n} f_{LN}(x_i; \sigma, \beta), \qquad \text{and} \\
L_{LL}(\gamma, \xi) &= \prod_{i=1}^{n} f_{LL}(x_i; \gamma, \xi),
\end{aligned}
$$

respectively. Let us consider the following test statistics

$$
Q_1 = ln\left(\frac{L_{WE}(\hat{\alpha}, \hat{\lambda})}{L_{LN}(\hat{\sigma}, \hat{\beta})}\right), \quad Q_2 = ln\left(\frac{L_{WE}(\hat{\alpha}, \hat{\lambda})}{L_{LL}(\hat{\gamma}, \hat{\xi})}\right), \quad Q_3 = ln\left(\frac{L_{LN}(\hat{\sigma}, \hat{\beta})}{L_{LL}(\hat{\gamma}, \hat{\xi})}\right),
$$

The statistics Q_1, Q_2 and Q_3 can be written as follows:

$$
\begin{aligned}
Q_1 &= n\left[ln\left(\sqrt{2\pi}\,\hat{\alpha}\,\hat{\sigma}(\hat{\lambda})^{\hat{\alpha}}\right) - \frac{1}{2}\right], \\
Q_2 &= n\left[ln\left(\hat{\alpha}\,\hat{\gamma}(\hat{\lambda})^{\hat{\alpha}}(\hat{\xi})^{\frac{1}{\hat{\gamma}}}\right) - 1\right] + (\hat{\alpha} - \frac{1}{\hat{\gamma}})\sum_{i=1}^{n} lnX_i + 2\sum_{i=1}^{n} ln\left(1 + (\frac{X_i}{\hat{\xi}})^{\frac{1}{\hat{\gamma}}}\right), \\
Q_3 &= n\left[ln\left(\frac{1}{\sqrt{2\pi}}\frac{\hat{\gamma}(\hat{\xi})^{\frac{1}{\hat{\gamma}}}}{\hat{\sigma}}\right) - \frac{1}{2}\right] - \frac{1}{\hat{\gamma}}\sum_{i=1}^{n} lnX_i + 2\sum_{i=1}^{n} ln\left(1 + (\frac{X_i}{\hat{\xi}})^{\frac{1}{\hat{\gamma}}}\right).
\end{aligned}
$$

Now we choose WE if $Q_1 > 0, Q_2 > 0$, LN if $Q_1 < 0, Q_3 > 0$ and LL if $Q_2 < 0, Q_3 < 0$. Hence, the respective PCS can be defined as follows:

$$PCS_{WE}^{LR} = P(Q_1 > 0, Q_2 > 0 | \text{data follow WE}).$$

$$PCS_{LN}^{LR} = P(Q_1 < 0, Q_3 > 0 | \text{data follow LN}),$$

$$PCS_{LL}^{LR} = P(Q_2 < 0, Q_3 < 0 | \text{data follow LL}).$$

Now we have the following results.

Theorem 2:
(i) *Under the assumptions that the data are from WE*(α, λ)*,* (Q_1, Q_2) *is asymptotically bivariate normally distributed with the mean vector* $(E_{WE}(Q_1), E_{WE}(Q_2))$*, and the dispersion matrix*

$$\Sigma_{WE} = \begin{bmatrix} Var_{WE}(Q_1) & Cov_{WE}(Q_1, Q_2) \\ Cov_{WE}(Q_1, Q_2) & Var_{WE}(Q_2) \end{bmatrix}$$

(ii) *Under the assumptions that the data are from LN*(σ, β)*,* (Q_1, Q_3) *is asymptotically bivariate normally distributed with the mean vector* $(E_{WE}(Q_1), E_{WE}(Q_3))$*, and the dispersion matrix*

$$\Sigma_{LN} = \begin{bmatrix} Var_{LN}(Q_1) & Cov_{LN}(Q_1, Q_3) \\ Cov_{WE}(Q_1, Q_3) & Var_{WE}(Q_3) \end{bmatrix}$$

(iii) *Under the assumptions that the data are from LL*(γ, ξ)*,* (Q_2, Q_3) *is asymptotically bivariate normally distributed with the mean vector* $(E_{LL}(Q_2), E_{LL}(Q_3))$*, and the dispersion matrix*

$$\Sigma_{LL} = \begin{bmatrix} Var_{LL}(Q_2) & Cov_{LL}(Q_2, Q_3) \\ Cov_{LL}(Q_2, Q_3) & Var_{LL}(Q_3) \end{bmatrix}$$

Proof: The proof can be obtained along the same line as the Theorem 1 of Dey and Kundu (2009), the details are avoided.

4 Numerical Comparisons

In this section we performed some simulation experiments to compare different methods to discriminate among these three distributions. We have used the method based on Fisher information (FI), the ratio of maximized likelihood (RML) method and the method based on Kolmogorov-Smirnov (MK) distance. The method based on the KS distance (KSD) can be described as follows. Based on the observed sample we compute the KS distances between the (i) empirical distribution function (EDF) and WE$(\hat{\alpha}, \hat{\lambda})$, (ii) EDF and LN$(\hat{\sigma}, \hat{\beta})$, (iii) EDF and LL$(\hat{\gamma}, \hat{\xi})$. Which ever gives the minimum we choose that particular distribution as the best fitted one. The PCS can be defined along the same manner. It can be proved theoretically that the PCS based on RML does not depend on the parameter values. In all the cases we have taken the scale parameters to be 1.

To compare the performances of the different methods for different sample sizes and for different parameter values, we have generated samples from the three different distributions and compute the PCS for each method based on 1000 replications. The results are reported in Table 5. Some of the points are quite clear from these experimental results. It is clear that as the sample size increases, the PCS increases in each case. It indicates the consistency properties of each method. In case of FI, the PCS decreases as the shape parameter increases in all the three cases.

Although no such pattern exists in case of KSD. It is further observed that the performance of the RML is very poor when the data are obtained from the LL distribution. It seems, although we could not prove it theoretically that PCS remains constant in case of KSD, when the data are obtained from the LL distribution. If the shape parameter is less than 1, then the method based on FI out performs the RML method. Although, nothing can be said when the shape parameter is greater than 1.

5 Fisher Information Matrix: Censored Sample

In engineering, economics, actuarial studies and medical research, censoring is a condition in which an observation is only partially known. With censoring, observations result is either lying within an interval or taking other exact values outside the interval. For example, suppose a study is performed to study the effect of a treatment on heart stroke. In such a study, we consider patients in a clinical trial to study the effect of treatments on stroke occurrence during time period (T_1, T_2). Those patients who had no strokes during (T_1, T_2) are censored. Right truncation occurs when the study patients have already stroke experienced after a fixed time (say T_2). Left truncation occurs when the subjects have been at risk before entering the study. Let X be a right and left censored at fixed times T_1 and T_2, respectively. Therefore, we observe the random variable Y as follows:

$$Y = \begin{cases} X, & \text{if } T_1 < X < T_2, \\ T_1, & \text{if } X < T_1, \\ T_2, & \text{if } X > T_2. \end{cases}$$

The FI about $\boldsymbol{\theta}$ in the observation Y can be written as

$$I_C(\boldsymbol{\theta}; T_1, T_2) = I_M(\boldsymbol{\theta}; T_1, T_2) + I_R(\boldsymbol{\theta}; T_2) + I_L(\boldsymbol{\theta}; T_1), \tag{12}$$

where,

$$\begin{aligned} I_M(\boldsymbol{\theta}; T_1, T_2) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ I_R(\boldsymbol{\theta}; T_2) &= \frac{1}{\overline{F}(T_2, \boldsymbol{\theta})} \begin{bmatrix} \frac{\partial}{\partial \theta_1} \overline{F}(T_2, \boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_2} \overline{F}(T_2, \boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial \theta_1} \overline{F}(T_2, \boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} \overline{F}(T_2, \boldsymbol{\theta}) \end{bmatrix} \\ I_L(\boldsymbol{\theta}; T_1) &= \frac{1}{F(T_1, \boldsymbol{\theta})} \begin{bmatrix} \frac{\partial}{\partial \theta_1} F(T_1, \boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_2} F(T_1, \boldsymbol{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial \theta_1} F(T_1, \boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} F(T_1, \boldsymbol{\theta}) \end{bmatrix}, \end{aligned}$$

$\overline{F}(T_2, \boldsymbol{\theta}) = 1 - F(T_2, \boldsymbol{\theta})$, and

$$a_{ij} = \int_{T_1}^{T_2} \left(\frac{\partial}{\partial \theta_i} \ln f(x; \boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \theta_j} \ln f(x; \boldsymbol{\theta}) \right) f(x; \boldsymbol{\theta}) dx,$$

Table 5: The PCSs under WE, LN and LL distributions for different sample sizes and for different parameter values .

$WE(\alpha, 1)$									
	$n = 20$			$n = 30$			$n = 50$		
$\alpha \downarrow$	RML	FI	K-S	RML	FI	K-S	RML	FI	K-S
0.3	0.750	0.905	0.400	0.825	0.945	0.510	0.890	0.970	0.650
0.6	0.760	0.890	0.420	0.795	0.905	0.485	0.895	0.980	0.655
0.8	0.770	0.850	0.448	0.815	0.870	0.480	0.890	0.975	0.660
1.0	0.764	0.695	0.440	0.800	0.726	0.485	0.902	0.782	0.630
1.2	0.770	0.620	0.460	0.820	0.680	0.560	0.915	0.740	0.645
1.5	0.755	0.478	0.490	0.795	0.496	0.615	0.910	0.556	0.780
$LN(\sigma, 1)$									
	$n = 20$			$n = 30$			$n = 50$		
$\sigma^2 \downarrow$	RML	FI	K-S	RML	FI	K-S	RML	FI	K-S
0.3	0.775	0.840	0.775	0.800	0.905	0.825	0.905	0.945	0.900
0.5	0.780	0.800	0.730	0.810	0.935	0.885	0.900	0.950	0.920
0.8	0.770	0.780	0.715	0.800	0.845	0.875	0.905	0.930	0.885
1.0	0.775	0.735	0.800	0.820	0.855	0.860	0.900	0.875	0.905
1.2	0.785	0.685	0.585	0.830	0.745	0.590	0.910	0.810	0.775
1.5	0.795	0.565	0.354	0.825	0.590	0.385	0.890	0.652	0.740
$LL(\gamma, 1)$									
	$n = 20$			$n = 30$			$n = 50$		
$\gamma \downarrow$	RML	FI	K-S	RML	FI	K-S	RML	FI	K-S
0.2	0.330	1.000	0.615	0.395	1.000	0.650	0.520	1.000	0.705
0.4	0.350	0.980	0.625	0.390	0.990	0.665	0.525	1.000	0.685
0.6	0.328	0.566	0.626	0.376	0.570	0.628	0.535	0.620	0.745
0.8	0.340	0.525	0.620	0.425	0.550	0.670	0.530	0.610	0.735
1.0	0.350	0.480	0.645	0.395	0.520	0.650	0.530	0.550	0.725
1.5	0.330	0.420	0.680	0.410	0.480	0.685	0.520	0.530	0.745

for $i, j = 1,2$. Therefore, the FI for complete sample or for fixed right censored (at time T_2) sample or for fixed left censored (at time T_1) sample with vector of parameters $\boldsymbol{\theta}$ can be obtained as

$$I_M(\boldsymbol{\theta}; 0, \infty), I_M(\boldsymbol{\theta}; 0, T_2) + I_R(\boldsymbol{\theta}; T_2) \text{ and } I_M(\boldsymbol{\theta}; T_1, \infty) + I_L(\boldsymbol{\theta}; T_1),$$

respectively, by observing the fact that $I_R(\boldsymbol{\theta}; \infty) = 0$ and $I_L(\boldsymbol{\theta}; 0) = 0$. In this section first we provide the FI matrices for the WE, LN and LL distributions.

Theorem 3: *Let $p_1 = F_{WE}(T_1), p_2 = F_{WE}(T_2)$, or $p_1 = F_{LN}(T_1), p_2 = F_{LN}(T_2)$ or $p_1 = F_{LL}(T_1), p_2 = F_{LL}(T_2)$. Then the FI matrices of the WE, LN and LL distributions for censored sample are obtained, respectively, to be as follows:*
(i)

$$I_{MW}(\alpha, \lambda, T_1, T_2) = \begin{pmatrix} a_{11W} & a_{12W} \\ a_{21W} & a_{22W} \end{pmatrix}, \; I_{RW}(\alpha, \lambda, T_2) = \begin{pmatrix} b_{11W} & b_{12W} \\ b_{21W} & b_{22W} \end{pmatrix},$$

$$I_{LW}(\alpha, \lambda, T_1) = \begin{pmatrix} c_{11W} & c_{12W} \\ c_{21W} & c_{22W} \end{pmatrix},$$

where

$$a_{11W} = \frac{1}{\alpha^2} \int_{-\ln(1-p_1)}^{-\ln(1-p_2)} (1 + \ln u - u \ln u)^2 e^{-u} \, du, \quad a_{22W} = \frac{\alpha^2}{\lambda^2} \int_{-\ln(1-p_1)}^{-\ln(1-p_2)} (1-u)^2 \, e^{-u} \, du,$$

$$a_{12W} = a_{21W} = \frac{1}{\lambda} \int_{-\ln(1-p_1)}^{-\ln(1-p_2)} (1 + \ln u - u \ln u)(1-u) \, e^{-u} \, du,$$

$$b_{11W} = \frac{(1-p_2) \, ln^2(1-p_2) \, ln^2[-ln(1-p_2)]}{\alpha^2}, \; b_{22W} = \frac{\alpha^2 \, (1-p_2) ln^2(1-p_2)}{\lambda^2},$$

$$b_{12W} = b_{21W} = \frac{(1-p_2) ln^2(1-p_2) \, ln[-ln(1-p_2)]}{\lambda},$$

$$c_{11W} = \frac{(1-p_1)^2 \, ln^2(1-p_1) \, ln^2[-ln(1-p_1)]}{\alpha^2 \, p_1}, \; c_{22W} = \frac{\alpha^2 \, (1-p_1)^2 ln^2(1-p_1)}{\lambda^2 \, p_1},$$

$$c_{12W} = c_{21W} = \frac{(1-p_1)^2 \, ln^2(1-p_1) \, ln[-ln(1-p_1)]}{\lambda \, p_1}.$$

(ii)

$$I_{MN}(\alpha, \lambda, T_1, T_2) = \begin{pmatrix} a_{11N} & a_{12N} \\ a_{21N} & a_{22N} \end{pmatrix}, \; I_{RN}(\alpha, \lambda, T_2) = \begin{pmatrix} b_{11N} & b_{12N} \\ b_{21N} & b_{22N} \end{pmatrix},$$

$$I_{LN}(\alpha, \lambda, T_1) = \begin{pmatrix} c_{11N} & c_{12N} \\ c_{21N} & c_{22N} \end{pmatrix},$$

where

$$
\begin{aligned}
a_{11N} &= \frac{1}{\sigma^2}\int_{\Phi^{-1}(p_1)}^{\Phi^{-1}(p_2)} (1-y^2)^2\phi(y)dy, \quad a_{22N} = \frac{1}{\beta^2\sigma^2}\int_{\Phi^{-1}(p_1)}^{\Phi^{-1}(p_2)} y^2\phi(y)dy, \\
a_{12N} &= a_{21L} = \frac{1}{\beta\sigma^2}\int_{\Phi^{-1}(p_1)}^{\Phi^{-1}(p_2)} (y^2-1)y\phi(y)dy, \\
b_{11N} &= \frac{1}{(1-p_2)\sigma^2}\left[\phi\left(\Phi^{-1}(p_2)\right)\right]^2\left(\Phi^{-1}(p_2)\right)^2, \quad b_{22L} = \frac{1}{\beta^2(1-p_2)\sigma^2}\left[\phi\left(\Phi^{-1}(p_2)\right)\right]^2 \\
b_{12N} &= b_{21L} = \frac{1}{\beta(1-p_2)\sigma^2}\left[\phi\left(\Phi^{-1}(p_2)\right)\right]^2\left(\Phi^{-1}(p_2)\right) \\
c_{11N} &= \frac{1}{p_1\sigma^2}\left[\phi\left(\Phi^{-1}(p_1)\right)\right]^2\left(\Phi^{-1}(p_1)\right)^2, \quad c_{22L} = \frac{1}{\beta^2 p_1\sigma^2}\left[\phi\left(\Phi^{-1}(p_1)\right)\right]^2 \\
c_{12N} &= c_{21N} = \frac{1}{\beta p_1\sigma^2}\left[\phi\left(\Phi^{-1}(p_1)\right)\right]^2\left(\Phi^{-1}(p_1)\right).
\end{aligned}
$$

(iii)

$$
I_{ML}(\alpha,\lambda,T_1,T_2) = \begin{pmatrix} a_{11L} & a_{12L} \\ a_{21L} & a_{22L}\end{pmatrix}, \; I_{RL}(\alpha,\lambda,T_2) = \begin{pmatrix} b_{11L} & b_{12L} \\ b_{21L} & b_{22L}\end{pmatrix},
$$

$$
I_{LL}(\alpha,\lambda,T_1) = \begin{pmatrix} c_{11L} & c_{12L} \\ c_{21L} & c_{22L}\end{pmatrix},
$$

$$
\begin{aligned}
a_{11L} &= \frac{1}{\gamma^2}\int_{p_1/(1-p_1)}^{p_2/(1-p_2)} \left[1+\ln u - 2\frac{u\ln u}{1+u}\right]^2 \frac{1}{(1+u)^2}\,du, \quad a_{22L} = \frac{1}{\gamma^2\,\xi^2}\int_{p_1/(1-p_1)}^{p_2/(1-p_2)} \frac{(1-u)^2}{(1+u)^4}\,du, \\
a_{12L} &= a_{21L} = \frac{1}{\gamma^2\,\xi}\int_{p_1/(1-p_1)}^{p_2/(1-p_2)} \left[1+\ln u - 2\frac{u\ln u}{1+u}\right]\frac{1-u}{(1+u)^3}\,du, \\
b_{11L} &= \frac{(1-p_2)p_2^2\,ln^2\left(\frac{p_2}{1-p_2}\right)}{\gamma^2}, \; b_{22L} = \frac{p_2^2(1-p_2)}{\gamma^2\xi^2}, \; b_{12L} = b_{21L} = \frac{p_2^2(1-p_2)\,ln^2\left(\frac{p_2}{1-p_2}\right)}{\gamma^2\xi}, \\
c_{11L} &= \frac{p_1(1-p_1)^2\,ln^2\left(\frac{p_1}{1-p_1}\right)}{\gamma^2}, \; c_{22L} = \frac{p_1(1-p_1)^2}{\gamma^2\xi^2}, \; c_{12L} = c_{21L} = \frac{p_1(1-p_1)^2 ln^2\left(\frac{p_1}{1-p_1}\right)}{\gamma^2\xi}.
\end{aligned}
$$

It may be observed that as $p_1 \downarrow 0$ and $p_2 \uparrow 1$, then $a_{ijW} \to f_{ijW}, a_{ijN} \to f_{ijN}$ and $a_{ijL} \to f_{ijL}$. To compare between the FI and its respective variance of the three distributions, we consider three different censoring schemes; Scheme 1: $p_1 = 0, p_2 = 0.75$; Scheme 2: $p_1 = 0.10, p_2 = 0.85$ and Scheme 3: $p_1 = 0.25, p_2 = 1$. These schemes represent right censoring, interval censoring and left censoring, respectively. The results are displayed in Table 6 under the assumption that the parent distribution is $WE(\alpha, 1)$ with $\alpha = 0.5, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0$. It is also observed from Table 6 that the FI for the right and left tails for LN are exactly same in this case since $p_{1L} = 1 - p_{2R}$. It can be easily observed from the expressions

Table 6: Loss of FI of the shape and scale parameters for Weibull distribution

Scheme→ Parameter ↓	Scheme 1	Scheme 2	Scheme 3
Shape	46%	78%	38%
Scale	25%	78%	9%

of the Fisher information matrix that if $p_{1L} = 1 - p_{2R}$, where $p_{1L} = F_{LN}(T_1)$ and $p_{2R} = F_{LN}(T_2)$ corresponding to left and right censoring cases, respectively, then both left and right censored data for LN distribution have the same FI about both the parameters. The same is true in case of LL distribution also. It seems it is due to the fact both normal (ln LN) and logistic (ln LL) distributions are symmetric distributions, although we could not prove theoretically. In case of WE distribution, the FI gets higher values towards the right tail than the left tail.

Now one important question is which of the two parameters has more impact. For this, we would like to discuss the loss of information due to truncation in one parameter, when the other parameter is known. Suppose the WE distribution is the underlying distribution with fixed truncation points. If the scale (shape) parameter is known, the loss of information of the shape (scale) parameter for WE distribution is

$$\begin{aligned} Loss_{WE}(\alpha) &= 1 - \frac{a_{11W} + b_{11W} + c_{11W}}{f_{11}} \\ &= 1 - \frac{1}{\psi'(1) + \psi^2(2)} \left[\int_{-ln(1-p_1)}^{-ln(1-p_2)} (1 + \ln u - u \ln u)^2 e^{-u}\, du + \right. \\ &\quad \left. (1-p_2)\ln^2(1-p_2)\ln^2(-ln(1-p_2)) + \frac{(1-p_1)^2 \ln^2(1-p_1)\ln^2(-ln(1-p_1))}{p_1} \right], \end{aligned}$$

and

$$\begin{aligned} Loss_{WE}(\lambda) &= 1 - \frac{a_{22W} + b_{22W} + c_{22W}}{f_{22}} \\ &= 1 - \left[\int_{-ln(1-p_1)}^{-ln(1-p_2)} (1-u)^2 e^{-u}\, du + (1-p_2)\ln^2(1-p_2) + (1-p_1)^2 \ln^2(1-p_1) \right]. \end{aligned}$$

Clearly both losses are free of any parameter and depend only on the truncation parameters. The loss of FI of the shape and scale parameters due to Schemes 1, 2 and 3 are presented in Table 6. It is easily seen that the information due to the interval censoring are similar in both cases while the information due to the right and left censoring are different. It is of interest to see also that the last portion of the data contains higher information of the shape/or scale parameter for the WE distribution.

Now for LN distribution, if the scale (shape) parameter is known, the loss of

Table 7: Loss of FI of the shape and scale parameters for the log-normal distribution

Scheme→ Parameter ↓	Scheme 1	Scheme 2	Scheme 3
Shape	31%	80%	51%
Scale	6%	72%	13%

information of the shape (scale) parameter is

$$Loss_{LN}(\sigma) = 1 - \frac{1}{2}\left[\int_{\Phi^{-1}(p_1)}^{\Phi^{-1}(p_2)} (1-u^2)^2\phi(u)\,du + \frac{[\phi(\Phi^{-1}(p_2))]^2(\Phi^{-1}(p_2))^2}{1-p_2} + \frac{[\phi(\Phi^{-1}(p_1))]^2(\Phi^{-1}(p_1))^2}{p_1}\right],$$

and

$$Loss_{LN}(\beta) = 1 - \left[\int_{\Phi^{-1}(p_1)}^{\Phi^{-1}(p_2)} u^2\phi(u)\,du + \frac{[\phi(\Phi^{-1}(p_2))]^2}{1-p_2} + \frac{[\phi(\Phi^{-1}(p_1))]^2}{p_1}\right].$$

For LN distribution, the losses of FI of the shape and scale parameters under different schemes are presented in Table 7. It is clear that the maximum information of the shape and scale parameters of LN distribution is occurred in the initial portion of the data.

Proceeding similarly, the losses of FI of the shape and scale parameters of LL are

$$Loss_{LL}(\gamma) = 1 - \frac{1}{3(1+\frac{\pi^2}{3})}\left[\int_{\frac{p_1}{1-p_1}}^{\frac{p_2}{1-p_2}} [1+\ln u - \frac{2u\ln u}{1+u}]^2 + (1-p_2)p_2^2\ln^2(\frac{p_2}{1-p_2}) + p_1(1-p_1)^2\ln^2(\frac{p_1}{1-p_1})\right],$$

and

$$Loss_{LL}(\xi) = 1 - \left[\int_{\frac{p_1}{1-p_1}}^{\frac{p_2}{1-p_2}} \frac{(1-u)^2}{(1+u)^4} + p_2^2\,(1-p_2) + p_1\,(1-p_1)^2\right],$$

respectively. The loss of information for the shape and scale parameters for different censoring schemes are presented in Table 8. It is clearly observed that for both the LN and LL, the initial portion of the data has more information than the right tail, where as for the WE distribution it is the other way.

Table 8: Loss of FI of the shape and scale parameters for the log-logistic distribution

Scheme→ Parameter ↓	Scheme 1	Scheme 2	Scheme 3
Shape	26%	74%	51%
Scale	2%	57%	14%

6 Data analysis

Here we discuss the analysis of real life data representing the fatigue life(rounded to the nearest thousand cycles) for 67 specimens of Alloy $T7987$ that failed before having accumulated 300 thousand cycles of testing. Their recorded values (in hundreds) are

```
0.94 1.18 1.39 1.59 1.71 1.89 2.27 0.96 1.21 1.40 1.59 1.72 1.90 2.56
0.99 1.21 1.41 1.59 1.73 1.96 2.57 0.99 1.23 1.41 1.59 1.76 1.97 2.69
1.04 1.29 1.43 1.62 1.77 2.03 2.71 1.08 1.31 1.44 1.68 1.80 2.05 2.74
1.12 1.33 1.49 1.68 1.80 2.11 2.91 1.14 1.35 1.49 1.69 1.84 2.13 1.17
1.36 1.52 1.70 1.87 2.24 1.17 1.39 1.53 1.70 1.88 2.26.
```

First we provide some preliminary data analysis results. The mean, median, standard deviation and the coefficient of skewness are 1.6608, 1.5900, 0.4672, 0.7488, respectively. The histogram of the above data set is presented in Figure 3 From

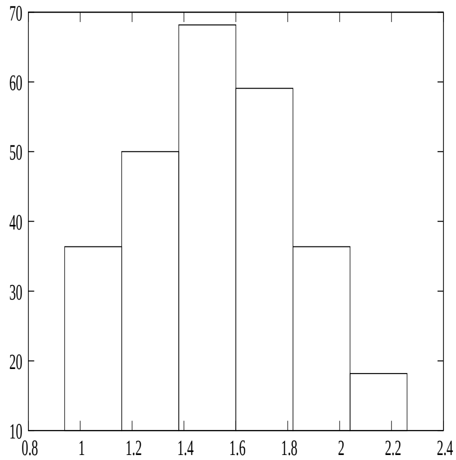

Figure 3: Histogram of the fatigue data.

the preliminary data analysis, it is clear that a skewed distribution may be used to analyze this data set. Barreto-Souza et al. (2010) fitted the Weibull-geometric, extended exponential geometric and WE models to this real data set. Graphically, Meeker and Escobar(1998, p. 149) showed that the LN distribution provides a much better fit than WE distribution. Before progressing further we provide the plot of

Table 9: The traces and variances of the FI matrices of $LN(\tilde{\sigma}, \tilde{\beta})$, $LL(\tilde{\gamma}, \tilde{\xi})$ and $WE(\alpha, 1)$ for three different censoring schemes

$\alpha \downarrow$	$\tilde{\sigma}$	$\tilde{\beta}$	$\tilde{\gamma}$	$\tilde{\xi}$	TI_{WE}	TI_{LN}	TI_{LL}	TV_{WE}	TV_{LN}	TV_{LL}
0.5	2.5651	0.3152	1.3918	0.3965	4.1134	1.6458	1.6252	5.5938	5.6498	2.7661
					1.6913	0.4951	0.6623	18.8675	18.6383	7.3245
					6.2757	1.6458	1.6252	4.7996	5.6498	2.7661
0.8	1.6032	0.4860	0.8699	0.5609	2.0136	2.0815	2.7803	2.7382	2.5921	1.4452
					0.7800	0.6242	1.0962	8.7020	8.5259	3.7021
					2.9928	2.0815	2.7803	2.2889	2.5921	1.4452
1.0	1.2826	0.5615	0.6959	0.6297	1.7315	2.6455	3.9000	2.3546	1.8015	1.0463
					0.6292	0.7923	1.5197	7.0191	5.9175	2.6492
					2.5048	2.6455	3.9000	1.9157	1.8015	1.0463
1.2	1.0688	0.6182	0.5799	0.6802	1.7616	3.3532	5.2638	2.3956	1.3346	0.7946
					0.6011	1.0033	2.0356	6.7053	4.3797	1.9967
					2.4837	3.3532	5.2638	1.8996	1.3346	0.7946
1.4	0.9161	0.6621	0.4971	0.7187	1.9708	4.1886	6.8642	2.6800	1.0322	0.6247
					0.6402	1.2524	2.6403	7.1415	3.3846	1.5611
					2.7253	4.1886	6.8642	2.0843	1.0322	0.6247
1.6	0.8016	0.6971	0.4349	0.7490	2.3034	5.1441	8.7018	3.1324	0.8237	0.5039
					0.7233	1.5372	3.3339	8.0692	2.6996	1.2543
					3.1440	5.1441	8.7018	2.4046	0.8237	0.5039
1.8	0.7125	0.7257	0.3866	0.7734	2.7329	6.2169	10.7689	3.7165	0.6734	0.4152
					0.8395	1.8569	4.1134	9.3650	2.2060	1.0303
					3.6994	6.2169	10.7689	2.8293	0.6734	0.4152
2.0	0.6413	0.7493	0.3480	0.7935	3.2454	7.4050	13.0637	4.4134	0.5612	0.3481
					0.9828	2.2110	4.9782	10.9638	1.8378	0.8618
					4.3698	7.4050	13.0637	3.3420	0.5612	0.3481

the scaled total time on test (TTT) transform. It is well known that the scaled TTT transform plot provides an indication about the shape of the hazard function. For example, if the plot is a concave function then it indicates that the hazard function is an increasing function, or if the plot is first concave and then convex, it indicates that the hazard function is an upside down function etc. We provide the plot of the scaled TTT transform in Figure 4. Although, it is not very clear, it has an indication that at the beginning it is concave and then it is convex. It indicates that the hazard function is an upside down function.

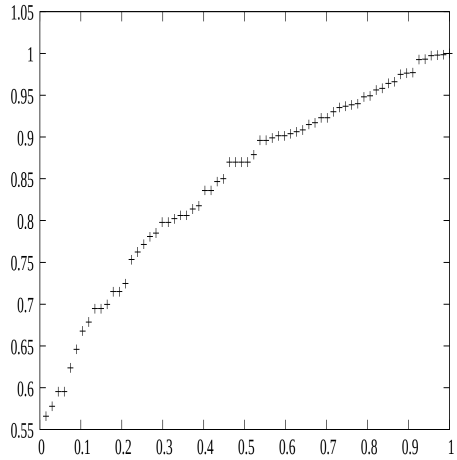

Figure 4: Scaled TTT transform plot.

Now we fit all the three distributions WE, LN and LL to this data set. The MLEs of the models parameters are computed numerically using Newton-Raphson (NR) method. The MLEs, the Kolmogorov-Smirnov (K-S) distances between the fitted and the empirical distribution functions and the corresponding p-values (between parentheses) are presented in Table 10. The empirical and fitted distribution functions are presented in Figure 5. The CDFs of LN and LL are very close to each other, and the CDF of WE is quite different than the other two.

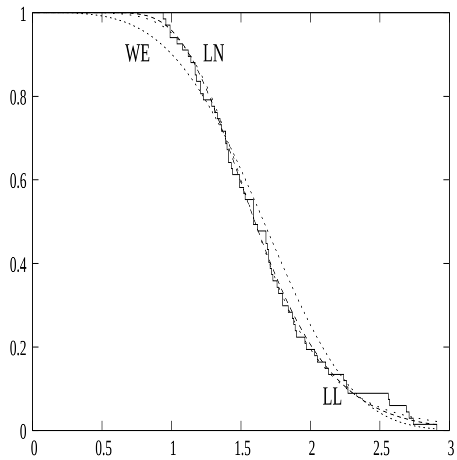

Figure 5: Empirical and fitted distribution functions.

Table 10: MLEs, K-S statistics and the corresponding p-values of the data set

	WE model	
α	λ	$K-S$
3.7257	0.5446	0.0973(0.5497)
	LN model	
σ	β	$K-S$
0.2722	1.5998	0.0418(0.9998)
	LL model	
γ	ξ	$K-S$
0.1576	1.5998	.0502(0.9959)

The maximized log-likelihood (MLL) values for WE, LN and WE distributions are, respectively, $-44.746, -39.381$ and -40.640. Therefore, based on MLL and K-S statistics it is observed that both LN and LL fit the data equally well, although LN provides a slightly better fit than LL.

Now let us compute the estimated FI matrices for the WE, LN and LL models and they are as follows:

$$I_{WE}(\hat{\alpha}, \hat{\lambda}) = \begin{pmatrix} 0.1314 & 0.7763 \\ 0.7763 & 46.8016 \end{pmatrix}, \; I_{LN}(\hat{\sigma}, \hat{\beta}) = \begin{pmatrix} 26.9840 & 0 \\ 0 & 5.2734 \end{pmatrix},$$

and

$$I_{LL}(\hat{\gamma}, \hat{\xi}) = \begin{pmatrix} 57.5573 & 0 \\ 0 & 5.2437 \end{pmatrix}.$$

Now if we consider the total asymptotic variances of the three cases, they are as follows:

$$TV_{WE} = 8.4622, \; TV_{LN} = 0.2267, \; TV_{LL} = 0.2081.$$

Hence, based on all these we conclude that LN is the most preferred among these three distributions for this particular data set.

Next, let us assess the variance of the p-th percentile estimators of the three distributions for various choices of p. Figure 3 shows the asymptotic variance of the p-th percentile estimators for complete and censored samples. For complete sample, it is clear that the WE distribution has higher variance than LN and LL distributions for $p < 0.7$. It is also evident from Figure Figure 6 that asymptotic variances of the p-th percentile estimators for censored samples for LN and LL are different while their respective curves tend to be identical for complete sample case. This concludes that both distributions LN and LL can be discriminated easily when the censoring data set is available. By taking a censoring observation on $[T_1, T_2]$ with $p_1 = 0.15$ and $p_2 = 0.90$, the FI matrices for WE, LN and LL distributions are computed to be

$$I_{WE}(\hat{\alpha}, \hat{\lambda}) = \begin{pmatrix} 0.0488 & 0.8918 \\ 0.8918 & 35.1012. \end{pmatrix}, I_{LN}(\hat{\beta}, \hat{\theta}) = \begin{pmatrix} 5.4529 & 0.1678 \\ 0.1678 & 1.4947 \end{pmatrix}$$

and

$$I_{LL}(\hat{\beta}, \hat{\theta}) = \begin{pmatrix} 28.5881 & 5.7614 \\ 5.7614 & 3.2150 \end{pmatrix}$$

respectively. Based on the FI matrices for complete and censored data sets, it is observed the loss of information due to truncation for LN distribution is much more than WE and LL distributions with respect to both parameters, while the loss of information of the shape parameter due truncation for WE and LL is much more than that of the scale parameters for the same distributions.

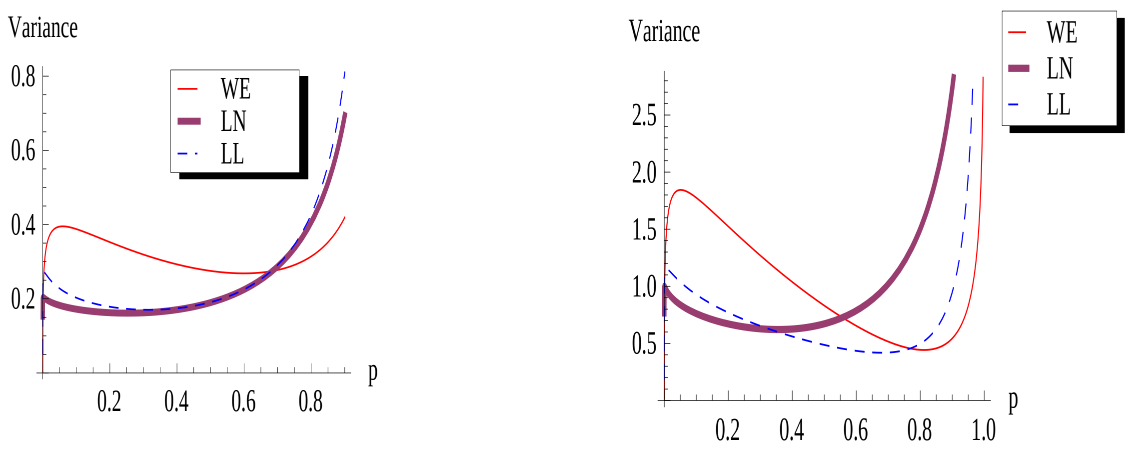

Figure 6: The variances of the pth percentile estimators for WE, LN and LL distributions for complete (left) and censored (right) samples.

7 Conclusions

In this article, we have considered the problem of discrimination among the WE, LN and LL distributions using three different methods. The asymptotic variance of the percentile estimators is also compared for these three distributions. It is observed that although the three distributions can be chosen as appropriate fitting models for a specific data set, the total information of Fisher information matrix as well as the asymptotic variance of the percentile estimators can be quite different. An extensive simulation experiment has been carried out to compute the PCS by different methods, and it is observed that the method based on the Fisher information measure competes with the other existing methods well, especially for certain ranges of parameter values.

Acknowledgement

The authors would like to thank the unknown reviewers for their constructive comments which have helped to improve the manuscript significantly.

Appendix

Proof of Theorem 1:

Taking the natural logarithm for the PDF of the LL distribution, we get

$$\ln f_{LL}(x;\gamma,\xi) \propto -ln\ \gamma + \frac{1}{\gamma}(ln\ x - ln\ \xi) - 2\ ln(1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}). \tag{13}$$

Differentiating both sides of (13) with respect to γ and ξ, respectively, we have

$$\frac{\partial \ln f_{LL}(x;\gamma,\xi)}{\partial \gamma} = -\frac{1}{\gamma}[1 + \frac{1}{\gamma}(ln\ x - ln\ \xi) - 2\ \frac{e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}{1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}\ \frac{(ln\ x - ln\ \xi)}{\gamma}],$$

and

$$\frac{\partial \ln f_{LL}(x;\gamma,\xi)}{\partial \xi} = -\frac{1}{\gamma\ \xi}(1 - 2\ \frac{e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}{1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}).$$

Then

$$f_{11L} = \frac{1}{\gamma^2}\int_0^\infty [1 + \frac{(ln\ x - ln\ \xi)}{\gamma} - 2\frac{(ln\ x - ln\ \xi)}{\gamma}\ \frac{e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}{1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}]^2 f_{LL}(x;\gamma,\xi)\ dx,$$

$$f_{22L} = \frac{1}{\gamma^2\ \xi^2}\int_0^\infty [2\ \frac{e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}}{1 + e^{\frac{(ln\ x - ln\ \xi)}{\gamma}}} - 1]^2 f_{LL}(x;\gamma,\xi)\ dx.$$

By simple transformation techniques, we may readily obtain

$$f_{11L} = \frac{1}{\gamma^2}\int_0^1 [1 + ln(\frac{1-z}{z}) - 2\ ln(\frac{1-z}{z}(1-z)]^2\ dz, \ \text{ and } f_{22L} = \frac{1}{3\gamma^2\xi^2}.$$

A more simplified expression of f_{11L} can be obtained by using the second derivative approach as follows:

$$\begin{aligned} f_{11L} &= -E\,\frac{\partial^2 ln f(x;\gamma,\xi)}{\partial\gamma^2} \\ &= -\frac{1}{\gamma^2} - \frac{2}{\gamma^2}\,E\left(\frac{ln\ x - ln\ \xi}{\gamma}\right) + \frac{4}{\gamma^2}\,E\left[\frac{ln\ x - ln\ \xi}{\gamma}\,\frac{e^{\frac{ln\ x - ln\ \xi}{\gamma}}}{1+e^{\frac{ln\ x - ln\ \xi}{\gamma}}}\right] \\ &+ \frac{2}{\gamma^2}\,E[(\frac{ln\ x - ln\ \xi}{\gamma})^2\frac{e^{\frac{ln\ x - ln\ \xi}{\gamma}}}{\left(1+e^{\frac{ln\ x - ln\ \xi}{\gamma}}\right)^2}]. \end{aligned} \tag{14}$$

Since the integrand function involved in the second term of the right hand side of (14) is odd function, its integral becomes 0. By using the substitution arguments, we con readily obtain the following identities:

$$E\left[\frac{ln\ x - ln\ \xi}{\gamma}\,\frac{e^{\frac{ln\ x - ln\ \xi}{\gamma}}}{1+e^{\frac{ln\ x - ln\ \xi}{\gamma}}}\right] = \int_0^1 [ln(1-z) - ln\ z]\,(1-z)\,dz = \frac{1}{2},$$

and

$$\begin{aligned} E\left[(\frac{ln\ x - ln\ \xi}{\gamma})^2\frac{e^{\frac{ln\ x - ln\ \xi}{\gamma}}}{(1+e^{\frac{ln\ x - ln\ \xi}{\gamma}})^2}\right] &= \int_0^1 [ln(1-z) - ln\ z]^2\,z(1-z)\,dz \\ &= \frac{19}{54} - 2\int_0^1 ln\ z\ ln(1-z)\ z(1-z)\ dz. \end{aligned}$$

This leads to

$$f_{11L} = \frac{1}{\gamma^2}\left\{\frac{46}{27} - 4\int_0^1 ln\ z\ ln(1-z)\ z(1-z)\ dz\right\}. \tag{15}$$

By using the Taylor's series expansion of $-ln(1-z) = \sum_{j=0}^{\infty} z^j/j$ and the identity (see, Gradshteyn and Ryzhik (1994), §4.272.16, p.548), we have

$$\begin{aligned} \int_0^1 (-ln\ z)z^{j+1}(1-z)\ dz &= \sum_{j=0}^{\infty}\sum_{k=0}^{1}\binom{1}{k}\frac{(-1)^k}{(j+k+2)^2} \\ &= \sum_{j=1}^{\infty}\frac{1}{j}\left[\frac{1}{(j+2)^2} - \frac{1}{(j+3)^2}\right]. \end{aligned}$$

The series appeared on the right hand side of the above identity can be computed easily by a straightforward algebra of partial fractions decomposition, telescoping series arguments and using the Euler-Riemann zeta function $\zeta(2) = \sum_{i=1}^{\infty} 1/i^2 = \pi^2/6$. This gives

$$\int_0^1 (-ln\ z)z^{j+1}(1-z)\ dz = \frac{74-6\pi^2}{216},$$

and consequently we simplify (15) as follows:

$$f_{11L} = \frac{1}{3\gamma^2}\left(1 + \frac{\pi^2}{3}\right).$$

Similarly, by differentiating (13) with respect to ξ and γ, respectively, and then we have

$$f_{12L} = f_{21L} = -\frac{1}{\gamma^2\xi} + \frac{2}{\gamma^2\xi}\int_{-\infty}^{\infty}\left[1 + y + e^y\right]\frac{e^{2y}}{1 + e^{2y}}\,dy,$$

and this in turn $f_{12L} = f_{21L} = 0$ by straightforward integration techniques.

References

[1] Ahmad, M. A., Raqab, M. Z. and Kundu, D. (2017). Discriminating between the generalized Rayleigh and Weibull distributions: Some comparative studies, *Communications in Statistics-Simulation & Computation*, In Press, Online Link: http://dx.doi.org/10.1080/03610918.2015.1136415.

[2] Alshunnar, F.S., Raqab, M.Z. and Kundu, D. (2010). On the comparison of the Fisher information of the log-normal and Weibull distributions, *Journal of Applied Statistics* 37, 391 - 404.

[3] Barreto-Souza,W., de Morais, A. L. and Cordeiro, G.M. (2010). The Weibull-geometric distribution, *J*ournal of Statistical Computation and Simulation 81(5), 645-657.

[4] Cox, D. R. (1962). Further results on tests of separate families of hypotheses, *Journal of the Royal Statistical Society*, Series B, Vol. 24, pp. 406-424.

[5] Dey, A. K. and Kundu, D. (2009). Discriminating between the log-normal and log-logistic distributions, *Communications in Statistics-Theory & Methods* 39(2), 280-292.

[6] Elsherpieny, E. A., Ibrahim, S. A., Radwan, N. U. (2013). Discrimination between Weibull and log-logistic distributions, *International Journal of Innovative Research in Science, Engineering and Technology* 2(8), 3358-3371.

[7] Gradshteyn, I.S. and Ryzhik, I.M. (1994), In: A. Jeffrey (Ed), Table of Integrals Series and Products, 5th edn, Academic Press, San Diego, USA.

[8] Gupta, R.D. and Kundu, D, (2006). On the comparison of Fisher information of the Weibull and GE distributions, *Journal of Statistical Planning & Inference* 136, 3130-3144.

[9] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distribution*, Vol. 1, 2nd Ed., New York, Wiley.

[10] Kundu, D. and Manglick, A.(2004). Discriminating between the Weibull and log- normal distributions, *Naval Research Logistics* 51, 893 - 905.

[11] Lehmann, E. L. (1991). *Theory of Point Estimation*, Wiley, New York.

[12] Marshall, A. W., Meza, J. C., and Olkin, I.(2001). Can data recognize its parent distribution?, *Journal of Computational and Graphical Statistics* 10 (2001) 555-580.

[13] Meeker, W. Q. and Escobar, L. A.(1998). *Statistical Methods for Reliability Data*, New York: Wiley & Sons.

[14] Pakyari, R. (2010). Discriminating between generalized exponential, geometric extreme exponential and Weibull distributions, *Journal of Statistical Computation and Simulation* 80(12), 1403-1412.

[15] Raqab, M. Z.(2013). Discriminating between the generalized Rayleigh and Weibull distributions, *Journal of Applied Statistics* 40(7), 1480-1493.

[16] Sultan K. S. and Al-Moisheer, A. S. (2013). Estimation of a discriminant function from a mixture of two inverse Weibull distributions, *Journal of Statistical Computation and Simulation* 83(3), 405-416.

[17] White, H. (1982). Regularity conditions for Cox's test of non-nested hypotheses, *Journal of Econometrics* 19, 301-318.

[18] Yu, H. F. (2007). Mis-specification analysis between normal and extreme value distributions for a linear regression model, *Communications in Statistics-Theory & Methods* 36, 499-521.