Communications for Statistical Applications and Methods
2017, Vol. 24, No. 2, 129–142

https://doi.org/10.5351/CSAM.2017.24.2.129
Print ISSN 2287-7843 / Online ISSN 2383-4757

Closeness of Lindley distribution to Weibull and gamma distributions

Mohammad Z. Raqab[1,a,b], Reem A. Al-Jarallah[a], Dhaifallah K. Al-Mutairi[a]

[a]Department of Statistics and Operations Research, Kuwait University, Kuwait;
[b]Department of Mathematics, The University of Jordan, Jordan

Abstract

In this paper we consider the problem of the model selection/discrimination among three different positively skewed lifetime distributions. Lindley, Weibull, and gamma distributions have been used to effectively analyze positively skewed lifetime data. This paper assesses how much closer the Lindley distribution gets to Weibull and gamma distributions. We consider three techniques that involve the likelihood ratio test, asymptotic likelihood ratio test, and minimum Kolmogorov distance as optimality criteria to diagnose the appropriate fitting model among the three distributions for a given data set. Monte Carlo simulation study is performed for computing the probability of correct selection based on the considered optimality criteria among these families of distributions for various choices of sample sizes and shape parameters. It is observed that overall, the Lindley distribution is closer to Weibull distribution in the sense of likelihood ratio and Kolmogorov criteria. A real data set is presented and analyzed for illustrative purposes.

Keywords: likelihood ratio statistic, asymptotic distributions, Kolmogorov distance, Lindley distribution, probability of correct selection, Monte Carlo simulation

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a univariate lifetime data set and coming from a positively skewed distribution. Based on the distributional relationship among the Lindley (LI), Weibull (WE), and gamma (GA) distributions, it is quite natural that the three models may provide similar data fit in the sense the three models can be used effectively to analyze skewed data. It can be observed that for certain ranges of the parameter values, the probability density functions (PDFs) or the cumulative distribution functions (CDFs) of the three distributions are close to each other; however, these distributions are different in terms of other characteristics. Let us introduce the underlying distributions (LI, WE, GA) for a data set. The LI distribution was used by Lindley (1958) in the context of Fiducial and Bayesian statistics. Its PDF is

$$\begin{aligned} g(x;\lambda) &= \frac{\lambda^2}{\lambda+1}(1+x)e^{-\lambda x}, \\ &= p g_1(x;\lambda) + (1-p) g_2(x;\lambda), \quad x > 0,\ \lambda > 0, \end{aligned}$$

where

$$p = \frac{\lambda}{\lambda+1}, \qquad g_1(x;\lambda) = \lambda e^{-\lambda x},\ x > 0, \qquad g_2(x;\lambda) = \lambda^2 x e^{-\lambda x},\ x > 0.$$

[1] Corresponding author: Department of Statistics and Operations Research, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait. E-mail: mraqab@ju.edu.jo

Published 31 March 2017 / journal homepage: http://csam.or.kr

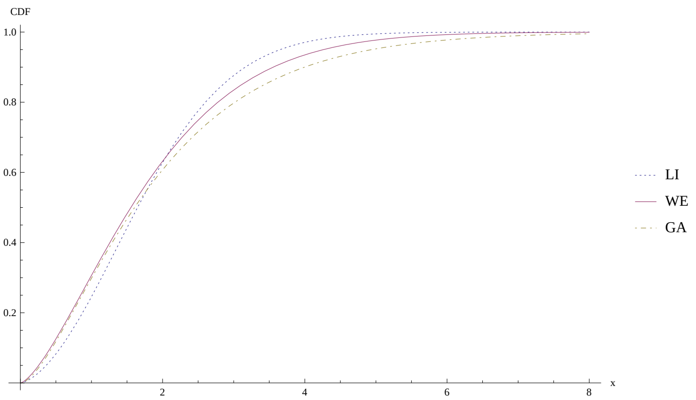


Figure 1: *The CDFs of LI*$(1.52, 0.6)$, *WE*$(1.45, 0.5)$, *and GA*$(1.7, 0.86)$. *CDF = cumulative distribution function; LI = Lindley; WE = Weibull; GA = gamma.*

The Lindley PDF can be written as a mixture of an exponential distribution with scale parameter λ and a GA $G(2, \lambda)$ with mixing proportion $p = \lambda/(\lambda + 1)$. This distribution received attention in the last few years. Ghitany *et al.* (2008) showed that the LI distribution is a better model than exponential distribution for some applications. Mazucheli and Achcar (2011) showed that the LI distribution can be used effectively in modelling strength data as well as for modelling general lifetime data. The LI distribution has been studied and extended by many authors in recent years. See, for example, Asgharzadeh *et al.* (2016), Asgharzadeh *et al.* (2017), Bakouch *et al.* (2012), Gomez-Deniz *et al.* (2014), Nadarajah *et al.* (2011), Okwuokenye and Peace (2016), and Zakerzadeh and Jafari (2009). Statistical inference for LI distribution based on complete and censored data has been discussed by several authors. See for example, Ali *et al.* (2013), Al-Mutairi *et al.* (2013), Gupta and Singh (2013), and Krishna and Kumar (2011). For $\lambda > 0, \alpha > 0$, the two-parameter LI distribution has the PDF

$$\begin{aligned} f(x; \alpha, \lambda) &= \frac{\alpha\lambda^2}{\lambda + 1} x^{\alpha-1}(1 + x^{\alpha}) e^{-\lambda x^{\alpha}}, \\ &= p g_3(x; \alpha, \lambda) + (1 - p) g_4(x; \alpha, \lambda), \quad x > 0,\ \alpha, \lambda > 0, \end{aligned}$$

where

$$\begin{aligned} g_3(x; \alpha, \lambda) &= \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}}, \qquad x > 0, \\ g_4(x; \alpha, \lambda) &= \alpha\lambda^2 x^{2\alpha-1} e^{-\lambda x^{\alpha}}, \quad x > 0. \end{aligned}$$

Here the LI(α, λ) is a two-component mixture of Weibull and a generalized gamma with shape α and scale λ parameters.

The PDFs of WE and GA distributions are, respectively, of the forms:

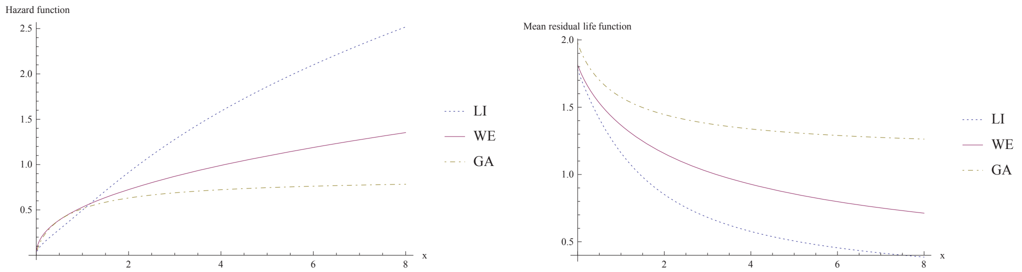
$$\begin{aligned} f_{\mathrm{WE}}(x; \beta, \theta) &= \beta\theta^{\beta} x^{\beta-1} e^{-(\theta x)^{\beta}}, \quad \text{for } x > 0, \beta, \theta > 0, \\ f_{\mathrm{GA}}(x; \nu, \tau) &= \frac{\tau^{\nu}}{\Gamma(\nu)} x^{\nu-1} e^{-\tau x}, \qquad \text{for } x > 0, \nu, \tau > 0. \end{aligned}$$

The WE and GA distributions are two of the most popular distributions used for analyzing skewed lifetime data. They have many applications that include queuing systems, reliability assessment, and

Table 1: The PCS based on the maximized likelihood when the model LI(1.5, 1)

n	20	30	50	80	100	500
PCS	0.392	0.445	0.594	0.655	0.691	0.945

PCS = probability of correct selection.



(a) (b)

Figure 2: *The hazard (a) and mean residual life (b) functions of LI*$(1.52, 0.6)$*, WE*$(1.45, 0.5)$*, and GA*$(1.7, 0.86)$*. LI = Lindley; WE = Weibull; GA = gamma.*

insurance claims. A detailed discussion on WE distribution and its related issues can be found in Kim (2016) and Park *et al.* (2016). A WE distribution with shape and scale parameters will be denoted by WE(β, θ) and similarly a GA distribution with shape parameter ν and inverted scale parameter τ will be denoted as GA(ν, τ).

Bain and Engelhardt (1980) proposed the use of the likelihood ratio statistic when choosing between a Weibull or gamma model and the values of the probability of correct selection (PCS) are obtained by Monte Carlo (MC) simulation. Marshall *et al.* (2001) considered the discrimination problem between the geometric extreme exponential model and other families of distributions that included Weibull, gamma and lognormal families as alternatives. Wiens (1999) discussed the problem of discriminating between lognormal and GA distributions. Dumonceaux and Antle (1973) and Pascual (2005) used maximum likelihood estimation (MLE) to study the miss-specification of lognormal and WE distributions.

Now we briefly describe the importance of choosing the best correct model, assuming that all the three models fit the data reasonable well. Let us consider the CDFs of LI($1.52, 0.6$), WE($1.45, 0.5$), and GA($1.7, 0.86$) presented in Figure 1. Data can be also fitted well by the other two distributions if it follows any one of the three distributions. This means that the separation between the three distributions are of great interest because they allow practitioners to choose the best choice among the distributions. The problem of choosing the correct model is not an easy task if the sample size n is small. For example, if the data are coming from LI($1.5, 1$), then the PCSs based on MC simulations based on the maximized likelihood method are computed as in Table 1. It is observed from Table 1 that the PCS can be only 0.392 when the sample size is 20. The mis-specification problem would lead to a serious effect. Let us look at the hazard functions or the mean residual life function of the three distribution functions displayed in Figure 2. The three CDFs are quite close to each other; however, it is evident that other characteristics such as the hazard functions or mean residual life functions of these distribution are clearly separated. Significant separation may occur for other characteristics as a variable sampling plan (Schneider, 1989).

This paper is to help the experimenter decide which distribution best fit his real data. The most appropriate fitting model allows us to develop valid model of random processes and make better decisions. In practice, the probability distributions are crucial for making efficient inferences in diverse fields as risk analysis, actuarial studies, reliability engineering and medicine. The problem of discrimination among some specific lifetime distributions has received a considerable attention in the literature due to increasing applications of lifetime distributions. We use the maximized likelihood ratio test (LRT), its asymptotic distribution and the minimum Kolmogorov distance (KD) procedures as optimal criteria to discriminate among the three competitive models: LI, WE, and GA distributions. We use the asymptotic distributions of the maximized likelihood functions since the exact distributions of the maximized likelihood functions are not available for finite samples. By using these procedures, PCSs are computed to discriminate among these models and determine the correct model by extensive computer simulations for different sample sizes.

The rest of the paper is organized as follows. We briefly describe the maximized LRT and minimum KD procedures in Section 2. In Section 3, the asymptotic distributions of the ratio of maximized likelihoods (RMLs) to discriminate between all pairs of the distributions are obtained. MC simulation experiments of PCS values using the LRT and KD procedures as well as the asymptotic results are presented in Section 4. Analysis of real data is also performed in Section 5. Finally, we conclude the paper in Section 6 with a look towards future work.

2. Likelihood ratio test and Kolmogorov distance procedures

Here in this section, we use the maximum LRT and the minimum KD criteria to choose the desirable fitted model among LI, WE, and GA distributions. Assume that the random sample $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is known to come from either a LI distribution, $X \sim \mathrm{LI}(\alpha, \lambda)$, or a WE distribution, $X \sim \mathrm{WE}(\beta, \theta)$ or a GA distribution, $X \sim \mathrm{GA}(\nu, \tau)$. Based on the observed sample, the corresponding likelihood functions are $L_L(\alpha, \lambda; \boldsymbol{x})$, $L_W(\beta, \theta; \boldsymbol{x})$, $L_G(\nu, \tau; \boldsymbol{x})$, respectively. For LI observed data with given α, the MLE of λ is

$$\hat{\lambda}(\alpha) = \frac{n - \sum_{i=1}^n x_i^\alpha + \sqrt{\left(n - \sum_{i=1}^n x_i^\alpha\right)^2 + 8n \sum_{i=1}^n x_i^\alpha}}{2 \sum_{i=1}^n x_i^\alpha}. \tag{2.1}$$

Now, to obtain the MLE of α, we need to maximize the profile log-likelihood function with respect to α. It can be shown that the maximum of $\log L(\alpha, \hat{\lambda}(\alpha))$ can be found as a fixed point solution of the following non-linear equation $g(\alpha) = \alpha$, where

$$g(\alpha) = \left[\frac{\hat{\lambda}(\alpha)}{n} \sum_{i=1}^n x_i^\alpha \ln x_i - \frac{1}{n} \sum_{i=1}^n \ln x_i - \frac{1}{n} \sum_{i=1}^n \frac{x_i^\alpha \ln x_i}{1 + x_i^\alpha}\right]^{-1}.$$

Very simple iterative procedure $g(\alpha^{(k)}) = \alpha^{(k+1)}$ can be used, where $\alpha^{(k)}$ is the k^{th} iterate. Once the MLE of α is computed, the MLE of λ can be computed immediately from (2.1).

For WE distribution with given β, the MLE $\hat{\theta}$ of θ is obtained as

$$\hat{\theta} = \left(\frac{n}{\sum_{i=1}^n x_i^\beta}\right)^{\frac{1}{\beta}}.$$

By substituting $\hat{\theta}$ in the respective log-likelihood, we obtain the MLE of β as a fixed point solution to

$h(\beta) = \beta$, where

$$h(\beta) = \left[\frac{\sum_{i=1}^{n} x_i^{\beta} \ln x_i}{\sum_{i=1}^{n} x_i^{\beta}} - \frac{\sum_{i=1}^{n} \ln x_i}{n} \right]^{-1}.$$

Under the assumption that the observations follow GA distribution, the MLEs of ν and τ for $G(\nu, \tau)$ are obtained by solving the nonlinear equation:

$$\ln \nu - \psi(\nu) + \frac{1}{n} \sum_{i=1}^{n} \ln x_i - \ln \overline{X} = 0,$$

where $\overline{X}$ is the sample mean. If $\hat{\nu}$ is obtained, the MLE of τ is computed to be $\hat{\tau} = \hat{\nu}/\overline{X}$.

2.1. Likelihood ratio test procedure

Let $X_1, \ldots, X_n$ be independent and identically distributed (iid) random variables (r.v.'s) coming from either a LI(α, λ) or a WE(β, θ). The log-likelihood ratio statistic T_{LW}, is defined as the logarithm of RML functions (see for example, Cox, 1961):

$$\begin{aligned} T_{\text{LW}} &= \ln \left(\frac{L_L\left(\hat{\alpha}, \hat{\lambda}; \boldsymbol{x}\right)}{L_W\left(\hat{\beta}, \hat{\theta}; \boldsymbol{x}\right)} \right) \\ &= n \ln \left(\frac{\hat{\alpha}\hat{\lambda}^2}{\left(\hat{\lambda}+1\right)\hat{\beta}\hat{\theta}^{\hat{\beta}}} \right) + \left(\hat{\alpha} - \hat{\beta}\right) \sum_{i=1}^{n} \ln x_i + \sum_{i=1}^{n} \ln\left(1 + x_i^{\hat{\alpha}}\right) - \hat{\lambda} \sum_{i=1}^{n} x_i^{\hat{\alpha}} + n, \end{aligned} \tag{2.2}$$

where $\hat{\alpha}$ and $\hat{\lambda}$ are the MLEs of α and λ, respectively, under the LI distribution and $\hat{\beta}$ and $\hat{\theta}$ are the MLEs of β and θ based on WE distribution. Now, the log-likelihood ratio statistic T_{LG} based on LI(α, λ) and GA(ν, τ) is also defined as

$$\begin{aligned} T_{\text{LG}} &= \ln \left(\frac{L_L(\hat{\alpha}, \hat{\lambda}; \boldsymbol{x})}{L_G(\hat{\nu}, \hat{\tau}; \boldsymbol{x})} \right) \\ &= n \ln \left(\frac{\hat{\alpha}\hat{\lambda}^2\Gamma(\hat{\nu})}{\hat{\tau}^{\hat{\nu}}\left(\hat{\lambda}+1\right)} \right) + (\hat{\alpha} - \hat{\nu}) \sum_{i=1}^{n} \ln x_i + \sum_{i=1}^{n} \ln\left(1 + x_i^{\hat{\alpha}}\right) + \hat{\tau} \sum_{i=1}^{n} x_i - \hat{\lambda} \sum_{i=1}^{n} x_i^{\hat{\alpha}}. \end{aligned} \tag{2.3}$$

The decision rule for discriminating between the LI and WE distributions is to choose the LI if $T_{\text{LW}} > 0$, and to reject the LI in favor of the WE, otherwise. Similarly, for discriminating between any LI and GA, we choose LI if $T_{\text{LG}} > 0$, and GA otherwise.

2.2. Kolmogorov distance procedure

The KD between two distributions is a tool to measure how similar two probability distributions are. It is the usually preferred way of measuring similarity of distributions. Let $F_L(\hat{\alpha}, \hat{\lambda})$, $F_W(\hat{\beta}, \hat{\theta})$ and $F_G(\hat{\nu}, \hat{\tau})$ be the CDFs computed at the MLEs of the parameters of LI, WE, and GA distributions, respectively. Also, let $F_n(x)$ be the empirically observed CDF computed based on the available data. KDs associated with the three models are defined as:

$$\text{KD}_L = \sup_{-\infty<x<\infty} \left| F_L\left(\hat{\alpha}, \hat{\lambda}\right) - F_n(x) \right|, \qquad \text{KD}_W = \sup_{-\infty<x<\infty} \left| F_W\left(\hat{\beta}, \hat{\theta}\right) - F_n(x) \right|,$$

and

$$\mathrm{KD}_G = \sup_{-\infty<x<\infty} |F_G(\hat{\nu}, \hat{\tau}) - F_n(x)|, \tag{2.4}$$

respectively. The decision rule here for discriminating among the three distributions is to choose the model with the minimum KD.

3. Asymptotic results of the ratio of maximized likelihood

In this section, we provide the asymptotic distribution of the logarithm of RML statistics L_{LW} and L_{LG} under null hypotheses in two different cases for each statistic. The results are concluded using the Central Limit Theorem (CLT) and White (1982). For notational convenience, let $\phi(V)$ be any Borel measurable function. The mean and variance of $\phi(V)$ under the assumption that V follows $\mathrm{LI}(\cdot,\cdot)$ are denoted by $E_L(\phi(V))$ and $\mathrm{Var}_L(\phi(V))$, respectively. Similarly, $E_W(\phi(V)), \mathrm{Var}_W(\phi(V))$ and $E_G(\phi(V)), \mathrm{Var}_G(\phi(V))$ stand for the mean and variance of $\mathrm{WE}(\cdot,\cdot)$ and $\mathrm{GA}(\cdot,\cdot)$, respectively. We also denote the almost sure convergence by a.s.

3.1. Asymptotic distribution of Lindley against Weibull

Case 1: The data come from a LI distribution $\mathrm{LI}(\alpha,\lambda)$ and competing distribution is a WE distribution $\mathrm{WE}(\beta,\theta)$. In this case, let us present the following lemma which is quite useful for the asymptotic results in Theorem 1 given below.

Lemma 1. *Under the assumption that the data are sampling from LI(α,λ), we have the following results as $n \to \infty$,*

(a) $\hat{\alpha} \to \alpha$ a.s., and $\hat{\lambda} \to \lambda$ a.s., where

$$E_L(\ln f_L(X;\alpha,\lambda)) = \max_{\bar{\alpha},\bar{\lambda}} E_L\left(\ln f_L\left(X;\bar{\alpha},\bar{\lambda}\right)\right).$$

(b) $\hat{\beta} \to \tilde{\beta}$ a.s., and $\hat{\theta} \to \tilde{\theta}$ a.s., where

$$E_L\left(\ln f_W\left(X;\tilde{\beta},\tilde{\theta}\right)\right) = \max_{\beta,\theta} E_L(\ln f_W(X;\beta,\theta)).$$

(c) If $T^*_{LW} = \ln[(L_L(\alpha,\lambda;\boldsymbol{x}))/(L_W(\tilde{\beta},\tilde{\theta};\boldsymbol{x}))]$, then $n^{-1/2}(T_{LW} - E_L(T_{LW}))$ is asymptotically equivalent to $n^{-1/2}(T^*_{LW} - E_L(T^*_{LW}))$.

Proof: Using arguments similar to those in White (1982, Theorem 1), we conclude the results. □

Theorem 1. *Under the assumption that the data are coming from LI(α,λ), then the distribution of T_{LW} as defined in (2.2) is approximately normally distributed with mean $E_L(T_{LW})$ and $\mathrm{Var}_L(T_{LW})$ as $n \to \infty$.*

Proof: It follows from CLT and part (b) of Lemma 1 that $n^{-1/2}(T^*_{\mathrm{LW}} - E_L(T^*_{\mathrm{LW}}))$ is asymptotically normally distributed with mean 0 and variance $\mathrm{Var}_L(T^*_{\mathrm{LW}}) = \mathrm{Var}_L(T_{\mathrm{LW}})$. Then by using (c) of Lemma 1, $n^{-1/2}(T_{\mathrm{LW}} - E_L(T_{\mathrm{LW}}))$ also approximately normally distributed as $n \to \infty$. Then for large n, T_{LW} converges in distribution to normal distribution with mean $E_L(T_{\mathrm{LW}})$ and variance $\mathrm{Var}_L(T_{\mathrm{LW}})$. □

Now, the mis-specified WE parameters when the data are coming from LI distribution, $\tilde{\beta}$ and $\tilde{\theta}$ are computed based on maximizing $\pi_1(\beta,\theta) = E_L[\ln f_W(X;\beta,\theta)]$. Equivalently, $\tilde{\beta}$ and $\tilde{\theta}$ are, respectively, solutions to

$$\frac{1}{\beta} + E_L(\ln X) - \frac{E_L(X^\beta \ln X)}{E_L\left(X^\beta\right)} = 0, \tag{3.1}$$

$$\theta = \left(\frac{1}{E_L\left(X^\beta\right)}\right)^{\frac{1}{\beta}}, \tag{3.2}$$

where

$$E_L\left(X^\beta\right) = \frac{\lambda\Gamma\left(\frac{\beta}{\alpha}+1\right)}{\lambda^{\frac{\beta}{\alpha}}(\lambda+1)}\left\{\frac{1}{\lambda}\left(\frac{\beta}{\alpha}+1\right)+1\right\}, \qquad E_L(\ln X) = \frac{1}{\alpha}\left\{\psi(1) + \frac{1}{\lambda+1} - \ln\lambda\right\},$$

$$E_L\left(X^\beta \ln X\right) = \frac{\lambda\Gamma\left(\frac{\beta}{\alpha}+1\right)}{\lambda^{\frac{\beta}{\alpha}}(\lambda+1)\alpha}\left\{\left[\psi\left(\frac{\beta}{\alpha}+1\right)-\ln\lambda\right]\left[1+\frac{1}{\lambda}\left(\frac{\beta}{\alpha}+1\right)\right]+\frac{1}{\lambda}\right\}.$$

Consequently, for large n, T_{LW} is asymptotically normal distribution with mean $E_L(T_{\text{LW}})$ and $\text{Var}_L(T_{\text{LW}})$ and variance, where

$$\text{AM}_{\text{LW}}^L = \lim_{n\to\infty}\frac{E_L(T_{\text{LW}})}{n} = E_L\left[\ln f_L(X;\alpha,\lambda) - \ln f_W\left(X;\tilde{\beta},\tilde{\theta}\right)\right], \tag{3.3}$$

$$\text{AV}_{\text{LW}}^L = \lim_{n\to\infty}\frac{\text{Var}_L(T_{\text{LW}})}{n} = \text{Var}_L\left[\ln f_L(\alpha,\lambda) - \ln f_W\left(\tilde{\beta},\tilde{\theta}\right)\right], \tag{3.4}$$

provided that AM_{LW}^L and AV_{LW}^L exist. The asymptotic PCS given that the true sampled distribution is LI and that the competing distribution is WE, is given by

$$\text{PCS}_L = P(T_{\text{LW}} > 0) \approx 1 - \Phi\left(\frac{-n\text{AM}_{\text{LW}}^L}{\sqrt{n\text{AV}_{\text{LW}}^L}}\right),$$

where $\Phi(.)$ is the distribution function of the standard normal distribution.

Case 2: The data come from a WE distribution $\text{WE}(\beta,\theta)$ and competing distribution is a $\text{LI}(\alpha,\lambda)$. Using similar procedures used in Case 1, it can be shown that as $n \to \infty$, T_{LW} is normally distributed with mean $E_W(T_{\text{LW}})$ and variance $\text{Var}_W(T_{\text{LW}})$, where

$$\text{AM}_{\text{LW}}^W = \lim_{n\to\infty}\frac{E_W(T_{\text{LW}})}{n} = E_W\left[\ln f_L\left(X;\tilde{\alpha},\tilde{\lambda}\right) - \ln f_W(X;\beta,\theta)\right], \tag{3.5}$$

$$\text{AV}_{\text{LW}}^W = \lim_{n\to\infty}\frac{\text{Var}_W(T_{\text{LW}})}{n} = \text{Var}_W\left[\ln f_L\left(X;\tilde{\alpha},\tilde{\lambda}\right) - \ln f_W(X;\beta,\theta)\right], \tag{3.6}$$

where $\tilde{\alpha}$ and $\tilde{\lambda}$ are mis-specified LI parameters when the data are coming from WE distribution. Here $\tilde{\alpha}$ and $\tilde{\lambda}$ are the values that maximizes $\pi_2(\alpha,\lambda) = E_W[\ln f_L(X;\alpha,\lambda)]$. Therefore $\tilde{\alpha}$ and $\tilde{\lambda}$ are solutions to the following equations:

$$\frac{1}{\alpha} + E_W(\ln X) + E_W\left(\frac{X^\alpha \ln X}{1+X^\alpha}\right) - \tilde{\lambda}(\alpha)E_W(X^\alpha \ln X) = 0,$$

$$\tilde{\lambda}(\alpha) = \frac{1 - E_W(X^\alpha) + \sqrt{(1-E_W(X^\alpha))^2 + 8E_W(X^\alpha)}}{2E_W(X^\alpha)},$$

where

$$E_W(X^\alpha) = \Gamma\left(\frac{\alpha}{\beta} + 1\right), \qquad E_W(\ln X) = \left(\frac{\psi(1)}{\beta} - \ln\theta\right),$$

$$E_W(X^\alpha \ln X) = \frac{\Gamma\left(\frac{\alpha}{\beta}+1\right)}{\theta^\alpha}\left(\frac{\psi\left(\frac{\alpha}{\beta}+1\right)}{\beta} - \ln\theta\right),$$

$$E_W\left(\frac{X^\alpha \ln X}{1+X^\alpha}\right) = \frac{1}{\beta}\int_0^\infty \frac{x^{\frac{\alpha}{\beta}}}{1+x^{\frac{\alpha}{\beta}}} \ln x\, e^{-x} dx.$$

3.2. Asymptotic distribution of Lindley against gamma

Case 1: The data come from a LI distribution $\mathrm{LI}(\alpha,\lambda)$ and competing distribution is a GA distribution $\mathrm{GA}(\nu,\tau)$. T_{LG} is approximately normally distributed as $n \to \infty$, with mean $E_L(T_{\mathrm{LG}})$ and variance $\mathrm{Var}_L(T_{\mathrm{LG}})$, where

$$\mathrm{AM}^L_{\mathrm{LG}} = \lim_{n\to\infty} \frac{E_L(T_{\mathrm{LG}})}{n} = E_L\left[\ln f_L(X;\alpha,\lambda) - \ln f_G\left(X;\tilde{\nu},\tilde{\tau}\right)\right], \tag{3.7}$$

$$\mathrm{AV}^L_{\mathrm{LG}} = \lim_{n\to\infty} \frac{\mathrm{Var}_W(T_{\mathrm{LG}})}{n} = \mathrm{Var}_L\left[\ln f_L(\alpha,\lambda) - \ln f_G\left(\tilde{\nu},\tilde{\tau}\right)\right], \tag{3.8}$$

where $\tilde{\nu}$ and $\tilde{\tau}$ are mis-specified GA parameters when the data are coming from LI distribution. Under LI, the mis-specified parameters $\tilde{\nu}$ and $\tilde{\tau}$ are derived based on maximizing $\pi_3(\nu,\tau) = E_L[\ln f_G(X;\tilde{\nu},\tilde{\tau})]$. Precisely, $\tilde{\nu}$ is the solution to

$$\psi(\nu) - \ln\nu + E_L(X) - E_L(\ln X) = 0, \tag{3.9}$$

$$\tilde{\tau} = \frac{\tilde{\nu}}{E_L(X)}. \tag{3.10}$$

Note that

$$E_L(X) = \frac{\Gamma\left(\frac{1}{\alpha}\right)\left(\frac{1}{\alpha}+\lambda+1\right)}{\alpha\lambda^{\frac{1}{\alpha}}(\lambda+1)}, \quad \text{and} \quad E_L(\ln X) = \frac{(1+\lambda)(\psi(1)-\ln\lambda)+1}{\alpha(\lambda+1)}.$$

Case 2: The data come from a GA distribution $\mathrm{GA}(\nu,\tau)$ and competing distribution is a LI distribution $\mathrm{LI}(\alpha,\lambda)$. In this case, T_{LG} is approximately normally distributed as $n \to \infty$, with mean $E_G(T_{\mathrm{LG}})$ and variance $\mathrm{Var}_G(T_{\mathrm{LG}})$, where

$$\mathrm{AM}^G_{\mathrm{LG}} = \lim_{n\to\infty} \frac{E_G(T_{\mathrm{LG}})}{n} = E_G\left[\ln f_L\left(X;\tilde{\alpha},\tilde{\lambda}\right) - \ln f_G(X;\nu,\tau)\right], \tag{3.11}$$

$$\mathrm{AV}^G_{\mathrm{LG}} = \lim_{n\to\infty} \frac{\mathrm{Var}_G(T_{\mathrm{LG}})}{n} = \mathrm{Var}_G\left[\ln f_L\left(\tilde{\alpha},\tilde{\lambda}\right) - \ln f_G(\nu,\tau)\right], \tag{3.12}$$

where $\tilde{\alpha}$ and $\tilde{\lambda}$ are mis-specified LI parameters when the data are coming from GA distribution. Here, under LI, the mis-specified parameters $\tilde{\alpha}$ and $\tilde{\lambda}$ are derived based on maximizing $\pi_4(\alpha,\lambda) = E_G[\ln f_L(X;\tilde{\alpha},\tilde{\lambda})]$. It is easily checked that $\tilde{\alpha}$ is the solution to

$$\frac{1}{\alpha} + E_G(\ln X) + E_G\left(\frac{X^\alpha \ln X}{1+X^\alpha}\right) - \tilde{\lambda}(\alpha) E_G\left(X^\alpha \ln X\right) = 0,$$

$$\tilde{\lambda}(\alpha) = \frac{\left(1 - E_G\left(X^{\tilde{\alpha}}\right)\right) + \sqrt{\left(1 - E_G\left(X^{\tilde{\alpha}}\right)\right)^2 + 8E_G\left(X^{\tilde{\alpha}}\right)}}{2E_G\left(X^{\tilde{\alpha}}\right)},$$

where

$$E_G(X^\alpha) = \frac{\Gamma(\alpha+\nu)}{\Gamma(\nu)\,\tau^\alpha}, \qquad E_G(\ln X) = \psi(\nu) - \ln\tau,$$
$$E_G(X^\alpha \ln X) = \frac{\Gamma(\alpha+\nu)}{\tau^\alpha\Gamma(\nu)}\left(\psi(\alpha+\nu) - \ln\tau\right),$$
$$E_G\left(\frac{X^\alpha \ln X}{1+X^\alpha}\right) = \frac{\tau^\nu}{\Gamma(\nu)}\int_0^\infty \frac{x^{\alpha+\nu-1}}{1+x^\alpha}\ln x\, e^{-\tau x}dx.$$

4. Numerical simulation study

Here in this section, we use the MC simulation to see which model among LI, WE, and GA distributions is preferred for fitting the data under the considered optimality criteria. In fact, we study the behavior of the PCS of the LI, WE, and GA distributions based on LRT, KD, and asymptotic LRT procedures. Let us consider the case where the null distribution is LI$(1,\lambda)$ and the alternative is WE(β,θ). In this case we consider $n = 20, 40, 60, 80$ and $\lambda = 0.1, 0.3, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0$. The MC samples are generated using the following algorithm:

Step 1: Generate data set from a LI distribution of size n.

Step 2: For each fixed value of λ, we compute the MLEs of all parameters involved in the alternatives: WE and GA distributions. In addition, we compute the mis-specified parameters $\tilde{\beta}$ and $\tilde{\theta}$ of WE distribution as well as the mis-specified parameters $\tilde{\nu}$ and $\tilde{\tau}$ of GA distribution using Equations (3.1), (3.2), (3.9), and (3.10).

Step 3: From Step 2, compute the LRT statistics T_{LW} and T_{LG} in (2.2) and (2.3), respectively. Also, compute KD_L, KD_W, and KD_G in (2.4).

Step 4: Repeat Step 1–3 M times and obtain MC samples $T^i_{\text{LW}}, L^i_{\text{LG}}, \text{KD}^i_L, \text{KD}^i_W$, and KD^i_G, $i = 1, 2, \ldots, M$.

Step 5: Under the assumption that the true distribution is LI$(1,\lambda)$, the approximate PCS based on LRT and KD can be described as:

$$\text{PCS}^{\text{LRT}}_{\text{LW}} \approx \frac{\#\text{ of } T_{\text{LW}} \text{ values in Step 3} > 0}{M},$$
$$\text{PCS}^{\text{KD}}_{\text{LW}} = \frac{\#\text{ of times } \text{KD}_L \text{ is minimum with respect } \text{KD}_W}{M}.$$

The $\text{PCS}^{\text{LRT}}_{\text{LG}}$ and $\text{PCS}^{\text{KD}}_{\text{LG}}$ can be described similarly.

The approximate PCSs based on the normal distribution of T_{LW} and T_{LG} as $n \to \infty$ (using Equations (3.3), (3.4), (3.7), and (3.8) are

$$\text{PCS}_{\text{LW}} \approx 1 - \Phi\left(\frac{-n\text{AM}^L_{\text{LW}}}{\sqrt{n\text{AV}^L_{\text{LW}}}}\right), \quad \text{and} \quad \text{PCS}_{\text{LG}} \approx 1 - \Phi\left(\frac{-n\text{AM}^L_{\text{LG}}}{\sqrt{n\text{AV}^L_{\text{LG}}}}\right).$$

The PCS results obtained from MC simulations (carried out using Mathematica package) considering LRT, asymptotic approximations and KD are presented in Tables 2 and 3 for various sample sizes and different parameters values. Arguments similar to ones in the above algorithm can be conducted when the parent distribution is WE or GA. The simulation study is conducted under WE and GA true distributions and the respective results are presented in Tables 4 and 5, respectively. From Tables 2–5, we conclude the following comments:

- Clearly, Tables 2 and 4 show some moderate values of PCS based on LRT, asymptotic LRT and KD procedures. This shows that these procedures are not clearly discriminate between LI and WE distributions. Towards this end, LI is close to WE distribution in terms of these criteria.
- When compared to Tables 2 and 4, Tables 3 and 5 present high PCS values. That is, when the underlying distribution is LI (GA) and the competing GA (LI), the LI distribution can be discriminated well against GA. When discriminating between LI and GA, the PCS based on LRT, and KD procedures gets high as α or ν moves away from 1. This implies that the discrimination procedures perform well in discriminating between LI and GA when α or ν goes away from 1.
- It is easily checked that for fixed sample size, the KD procedure is not intrinsically sensitive to the choices of parametric values to discriminate between LI and WE distributions. This conclusion is reversed while discriminating between LI and GA distributions where one can observe significant variations for various parametric changes. This confirms apparent distinction between LI and GA fitting distributions and some difficulties in obtaining an obvious discrimination between LI and WE distributions.
- Generally, the asymptotic PCS and the simulated LRT are close in the most considered cases. This indicates that asymptotic LRT works reasonably well in discriminating between these distributions even when the sample size is small or moderate. For large sample sizes, PCS values computed based on LRT tends to the ones based on the simulated LRT.

5. Data analysis

In this section, we apply the procedures developed in this paper on a real life data set. We present a data analysis of the insulating fluid data presented by Nelson (1982). The data represent the times to breakdown of an insulating fluid between electrodes recorded at voltage of 34 kV. As indicated in Nelson (1982), the times to breakdown insulating fluid at these voltages are an exponential distribution. Awwad *et al.* (2014) analyzed the data assuming the WE distribution. The times to breakdown are recorded as

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67
4.85	6.50	7.35	8.01	8.27	12.06	31.75	32.52
33.91	36.71	72.89					

Here, we aim at discriminating among LI, WE, and GA and choosing the best preferred model for fitting this data set. For comparison purposes, we plot the three fitted distribution functions for data set in Figure 3. It is clear that the three fitted distribution functions are very close in the sense of fitting probability model; therefore, a tool is required to discriminate among them. Here, we show that the logarithm of RML plays an important role in discriminating among overlapping families. For this, it is worthy to decide which one of them represents the data better. Now, the MLEs of the unknown parameters are $\hat{\alpha} = 0.6141$, $\hat{\lambda} = 0.3979$, $\hat{\beta} = 0.7708$, $\hat{\theta} = 0.0818$, $\hat{\nu} = 0.6898$, and $\hat{\tau} = 0.0480$. The log-likelihood values corresponds to LI, WE, and GA distributions are -63.75 and -68.39 and -68.62. Therefore, $T_{\mathrm{LW}} = 4.64 > 0, T_{\mathrm{LG}} = 4.87 > 0$ and this indicates that the LI distribution is the preferred

Table 2: The probability of correct selection based on LRT, ALRT, and KD procedures when the true model is Lindley and the competing distribution is Weibull model

α	n = 20			n = 40			n = 60			n = 80		
	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD
0.1	0.47	0.54	0.39	0.52	0.55	0.40	0.52	0.56	0.42	0.52	0.57	0.42
0.3	0.45	0.54	0.40	0.49	0.55	0.40	0.53	0.57	0.43	0.54	0.57	0.43
0.5	0.45	0.54	0.40	0.50	0.55	0.40	0.52	0.56	0.42	0.55	0.57	0.45
1.0	0.46	0.54	0.41	0.50	0.55	0.40	0.51	0.56	0.41	0.54	0.57	0.44
1.5	0.45	0.54	0.39	0.50	0.55	0.39	0.52	0.56	0.42	0.54	0.57	0.44
2.0	0.45	0.54	0.39	0.49	0.55	0.39	0.52	0.56	0.42	0.54	0.57	0.45
2.5	0.46	0.54	0.40	0.50	0.55	0.40	0.51	0.56	0.42	0.54	0.57	0.44
3.0	0.47	0.54	0.40	0.49	0.55	0.40	0.54	0.56	0.44	0.54	0.57	0.43

LRT = likelihood ratio test; ALRT = asymptotic likelihood ratio test; KD = Kolmogorov distance.

Table 3: The probability of correct selection based on LRT, ALRT, and KD procedures when the true model is Lindley and the competing distribution is gamma model

α	n = 20			n = 40			n = 60			n = 80		
	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD
0.1	0.70	0.54	0.92	0.88	0.56	0.98	0.95	0.57	0.99	0.98	0.58	1.00
0.3	0.61	0.69	0.82	0.74	0.76	0.88	0.83	0.81	0.92	0.88	0.84	0.95
0.5	0.50	0.62	0.75	0.60	0.67	0.79	0.65	0.70	0.81	0.69	0.73	0.82
1.0	0.49	0.56	0.41	0.56	0.59	0.44	0.60	0.61	0.50	0.61	0.63	0.51
1.5	0.63	0.66	0.38	0.69	0.72	0.51	0.74	0.76	0.57	0.78	0.79	0.64
2.0	0.66	0.70	0.41	0.76	0.77	0.55	0.81	0.82	0.65	0.85	0.86	0.72
2.5	0.70	0.73	0.43	0.80	0.81	0.60	0.85	0.85	0.70	0.88	0.89	0.76
3.0	0.71	0.74	0.45	0.82	0.82	0.63	0.86	0.87	0.71	0.90	0.91	0.79

LRT = likelihood ratio test; ALRT = asymptotic likelihood ratio test; KD = Kolmogorov distance.

Table 4: The probability of correct selection based on LRT, ALRT, and KD procedures when the true model is Weibull and the competing distribution is Lindley model

α	n = 20			n = 40			n = 60			n = 80		
	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD
0.1	0.57	0.58	0.65	0.56	0.61	0.66	0.57	0.64	0.65	0.59	0.66	0.66
0.3	0.55	0.53	0.66	0.56	0.55	0.66	0.57	0.56	0.65	0.58	0.56	0.66
0.5	0.55	0.53	0.65	0.58	0.55	0.66	0.56	0.56	0.66	0.57	0.56	0.64
1.0	0.55	0.53	0.65	0.57	0.55	0.65	0.58	0.56	0.65	0.59	0.56	0.65
1.5	0.56	0.53	0.67	0.56	0.55	0.66	0.57	0.56	0.65	0.58	0.56	0.64
2.0	0.57	0.53	0.66	0.56	0.55	0.66	0.58	0.56	0.65	0.60	0.56	0.66
2.5	0.55	0.53	0.66	0.57	0.55	0.66	0.57	0.56	0.65	0.58	0.56	0.65
3.0	0.57	0.53	0.65	0.57	0.55	0.66	0.58	0.56	0.64	0.59	0.56	0.65

LRT = likelihood ratio test; ALRT = asymptotic likelihood ratio test; KD = Kolmogorov distance.

Table 5: The probability of correct selection based on LRT, ALRT, and KD procedures when the true model is gamma and the competing distribution is Lindley model

α	n = 20			n = 40			n = 60			n = 80		
	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD	LRT	ALRT	KD
0.1	0.87	0.86	0.52	0.94	0.94	0.66	0.97	0.97	0.74	0.99	0.99	0.80
0.3	0.74	0.71	0.44	0.80	0.78	0.54	0.83	0.83	0.61	0.88	0.87	0.66
0.5	0.63	0.61	0.37	0.67	0.65	0.41	0.71	0.68	0.46	0.72	0.71	0.49
1.0	0.55	0.53	0.54	0.55	0.55	0.49	0.56	0.56	0.48	0.56	0.56	0.47
1.5	0.60	0.60	0.64	0.64	0.64	0.66	0.68	0.67	0.69	0.70	0.70	0.68
2.0	0.65	0.64	0.70	0.70	0.69	0.71	0.74	0.73	0.73	0.76	0.76	0.74
2.5	0.66	0.66	0.71	0.71	0.72	0.73	0.77	0.76	0.76	0.79	0.79	0.77
3.0	0.68	0.67	0.73	0.73	0.73	0.74	0.78	0.77	0.76	0.82	0.81	0.79

LRT = likelihood ratio test; ALRT = asymptotic likelihood ratio test; KD = Kolmogorov distance.

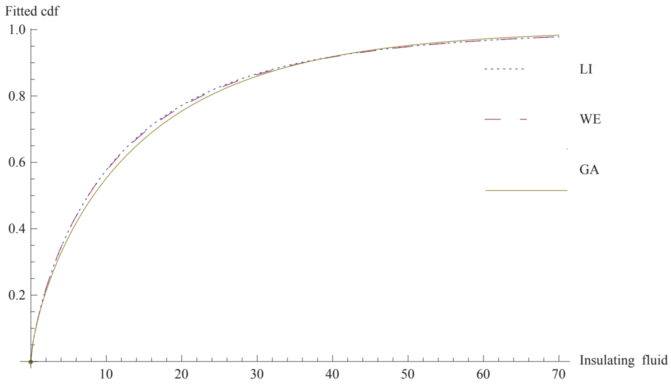


Figure 3: *The three fitted distribution (LI, WE, GA) functions for the insulating fluid data set. CDF = cumulative distribution function; LI = Lindley; WE = Weibull; GA = gamma.*

fitting model. The KD distances between the empirical distribution function and the fitted distributions and the associated p-values are ($\text{KD}_L = 0.1602, p\text{-value} = 0.7142$), ($\text{KD}_W = 0.1614, p\text{-value} = 0.7055$), and ($\text{KD}_G = 0.1849, p\text{-value} = 0.5346$). It is clear that the three distributions fit the data reasonably well but LI is the preferred one. The WE distribution is very competitive fitting distribution and is a good alternative. The GA distribution is fitting the data, but it is a worse one in terms of KD criterion. In terms of LRT criterion, LI distribution is relatively closer to WE than GA distribution.

The results developed in the previous sections are used to examine that LI is the preferred model over WE and GA distributions. Suppose the distribution of the above data is LI with $\alpha = 0.6141 = \hat{\alpha}, \lambda = 0.3979 = \tilde{\lambda}$. In this case, the PCS using KD criteria is computed based on a simulation (10,000 runs) to be PCS = 0.60 if the competing distribution is WE and PCS = 0.72 if the competing distribution is GA distribution. This implies that we are confident more than 70% that LI distribution is preferred against GA distribution and just 60% that LI is preferred against WE distribution. In the case, we wrongly choose the GA with $\nu = 0.6898 = \hat{\nu}, \tau = 0.048 = \hat{\tau}$ as the underlying distribution of this data, then based on a simulation of 10,000 runs, we obtain PCS using LRT as 0.70.

6. Conclusions

In this paper, we study the problem of closeness of the LI distribution to the WE and GA distributions using statistics based on maximized RML and minimized KD. However, we obtain asymptotic distributions of the logarithm of RML statistics between LI and WE distributions and LI and GA distributions under two different conditions. We compare PCSs using the LRT, asymptotic LRT and KD criteria via MC simulations for various choices of sample sizes. The conclusion of this paper is mainly two fold. First, it is observed that overall, the LI distribution is closer to WE distribution in the sense of LRT and KD criteria while these criteria can discriminate the LI distribution well from the GA distribution. The second conclusion of this paper is that the asymptotic LRT works adequately even for small or moderate sample sizes. Clearly, the PCS depends on the model parameters of the parent distribution as well as sample sizes. Based on the first conclusion, one may question if we can approximate the LI distribution by WE distribution; subsequently, more work is required in this area. The results developed in the paper can be extended for Type I censored data. We have considered the closeness among LI, WE, and GA distributions; however, our procedures can be extended to check the closeness and separation among other related distributions.

Acknowledgement

We would like to thank the editor for his encouragement and the referees for their valuable suggestions and comments.

References

Abu Awwad RR, Raqab MZ, and Al-Mudahakha IM (2014). Statistical inference based on progressively type II censored data from Weibull model, *Communications in Statistics - Simulation and Computation*, **44**, 2654–2670.

Ali S, Aslam M, and Kazmi SMA (2013). A study of the effect of the loss function on Bayes estimate, posterior risk and hazard function for Lindley distribution, *Applied Mathematical Modelling*, **37**, 6068–6078.

Al-Mutairi DK, Ghitany ME, and Kundu D (2013). Inferences on stress-strength reliability from Lindley distributions, *Communications in Statistics - Theory and Methods*, **42**, 1443–1463.

Asgharzadeh A, Bakouch HS, Nadarajah S, and Sharafi F (2016). A new weighted Lindley distribution with application, *Brazilian Journal of Probability and Statistics*, **30**, 1–27.

Asgharzadeh A, Nadarajah S, and Sharafi F (2017). Generalized inverse Lindley distribution with application to Danish fire insurance data, *Communications in Statistics - Theory and Methods*, **46**, 5001–5021.

Bain LJ and Engelhardt M (1980). Probability of correct selection of Weibull versus gamma based on likelihood ratio, *Communications in Statistics - Simulation and Computation*, **9**, 375–381.

Bakouch HS, Al-Zahrani BM, Al-Shomrani AA, Marchi VA, and Louzada F (2012). An extended Lindley distribution, *Journal of the Korean Statistical Society*, **41**, 75–85.

Cox DR (1961). Tests of separate families of hypotheses. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 105–123.

Dumonceaux R and Antle CE (1973). Discrimination between the log-normal and the Weibull distributions, *Technometrics*, **15**, 923–926.

Ghitany ME, Atieh B, and Nadarajah S (2008). Lindley distribution and its application, *Mathematics and Computers in Simulations*, **78**, 493–506.

Gomez-Deniz E, Sordo MA, and Calderin-Ojeda E (2014). The log-Lindley distribution as an alternative to the beta regression model with applications in insurance, *Insurance: Mathematics and Economics*, **54**, 49–57.

Gupta PK and Singh B (2013). Parameter estimation of Lindley distribution with hybrid censored data, *International Journal of System Assurance Engineering and Management*, **4**, 378–385.

Kim N (2016). On the new maximum likelihood estimators for parameters of a Weibull distribution under random censoring, *Communications for Statistical Applications and Methods*, **23**, 241–250.

Krishna H and Kumar K (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample, *Mathematics and Computers in Simulation*, **82**, 281–294.

Lindley DV (1958). Fiducial distributions and Bayes' theorem, *Journal of the Royal Statistical Society Series B (Methodological)*, **20**, 102–107.

Marshall AW, Meza JC, and Olkin I (2001). Can data recognize its parent distribution?, *Journal of Computational and Graphical Statistics*, **10**, 555–580.

Mazucheli J and Achcar JA (2011). The Lindley distribution applied to competing risks lifetime data, *Computer Methods and Programs in Biomedicine*, **104**, 188–192.

Nadarajah S, Bakouch HS, and Tahmasbi R (2011). A generalized Lindley distribution, *Sankhya B*, **73**, 331–359.

Nelson W (1982). *Applied Life Data Analysis*, John Wiley & Sons, New York.

Okwuokenye M and Peace KE (2016). A comparison of inverse transform and composition methods of data simuladon from the Lindley distribution, *Communications for Statistical Applications and Methods*, **23**, 517–529.

Park S, Park J, and Choi Y (2016). A new flexible Weibull distribution, *Communications for Statistical Applications and Methods*, **23**, 399–409.

Pascual FG (2005). Maximum likelihood estimation under misspecified lognormal and Weibull distributions, *Communications in Statistics - Simulation and Computations*, **34**, 503–524.

Schneider H (1989). Failure-censored variables-sampling plans for lognormal and Weibull distributions, *Technometrics*, **31**, 199–206.

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

Wiens BL (1999). When log-normal and gamma models give different results: a case study, *The American Statistician*, **53**, 89–93.

Zakerzadeh H and Jafari AA (2009). Inference on the parameters of two Weibull distributions based on record values, *Statistical Methods and Applications*, **24**, 25–40.

Received October 21, 2016; Revised January 5, 2017; Accepted February 18, 2017