The Lehmann type II inverse Weibull distribution in the presence of censored data

Vera L. D. Tomazella[1], Pedro L. Ramos[2], Paulo H. Ferreira[3], Alex L. Mota [1,2] and Francisco Louzada[2]

[1] Federal University of São Carlos, São Carlos, SP, Brazil (vera@ufscar.br, alexlealmota@usp.br)

[2] University of São Paulo, São Carlos, SP, Brazil (pedrolramos@usp.br, louzada@icmc.usp.br)

[3] Federal University of Bahia, Salvador, BA, Brazil (paulohenri@ufba.br)

Abstract

In this paper, we investigate the mathematical properties of the Lehmann type II inverse Weibull distribution. We show that this model is a reparameterized version of the Kumaraswamy-inverse Weibull distribution without identifiability problems. Parameter estimation is discussed using maximum likelihood (ML) method under a right-censoring scheme. Furthermore, a bootstrap resampling approach is considered to reduce the bias of the ML estimates. In order to illustrate the proposed methodology, we consider a real data set related to the failure time of devices in an aircraft.

Keywords: Bootstrap-based bias correction; Maximum likelihood estimation; Reliability data; Right-censoring; Unimodal failure rate.

1 Introduction

Motivated by researches developed in recent years, many authors have proposed new classes of probability distributions, which are modifications of the baseline probability distribution functions that provide hazard rates contemplating various shapes. For instance, Mudholkar et al. (1995) presented the three-parameter exponentiated Weibull distribution, which, depending on the chosen parameters, can exhibit a failure rate function with non-monotone (i.e., unimodal or bathtub-shaped) behavior. Carrasco et al. (2008) derived a four-parameter distribution, named generalized modified Weibull distribution, with the ability to model monotone as well as non-monotone failure rates. Pescim et al. (2010) proposed distribution with four parameters called the beta generalized half-normal distribution, which can accommodate all the forms of the failure rate function. On the other hand, Kumaraswamy (1980) argued that the beta distribution does not faithfully fit hydrological random variables (as well as other random processes which are bounded both at the lower and upper ends, and which have a mode occurring between these two bounds), such as daily rainfall and daily streamflow. Moreover, as stated in Jones (2009), "the beta distribution is fairly tractable, but in some ways not fabulously so; in particular, its distribution function is an incomplete beta function ratio and its quantile function the inverse thereof."

Still little explored in the statistical literature, the Kumaraswamy distribution (Kumaraswamy, 1980) has a domain on $(0, 1)$. Such a feature allows the Kumaraswamy distribution to be merged with other models in order to generate new families of distributions. In this sense, Cordeiro and de Castro (2011) proposed to use the Kumaraswamy distribution to generalize other distributions and derived many critical mathematical properties of the new class of models.

Based on these ideas, Shahbaz et al. (2012) presented the Kumaraswamy-inverse Weibull (Kum-IW) distribution, which includes several well-known distributions used in survival and reliability data analysis (see Gusmão et al., 2017, Section 2.1). However, in this work, we prove that the Kum-IW distribution has identifiability problems, which is undesirable. Thus, we propose one way to correct this problem by considering a useful reparameterization of the model. The obtained (identifiable) distribution can be seen as the Lehmann type II inverse Weibull (LIW2) distribution, which first appeared in a letter presented by Gusmão et al. (2012). Although its cumulative distribution function was presented, no mathematical and inferential issues were discussed for this distribution. Therefore, we provide a comprehensive treatment of the mathematical properties of the LIW2 distribution. The parameter estimation of this model is carried out using the maximum likelihood (ML) method in the presence of right-censored observations. Further, we consider the bootstrap resampling method proposed by Efron (1992), aiming to reduce the bias of the ML estimates. This approach consists of generating pseudo-samples from the original data to estimate the bias of the ML estimates. Finally, in order to illustrate our proposed methodology, we consider a real data set related to the failure time of 194 devices in an aircraft.

The remainder of this paper is organized as follows. In Section 2, we present the Kum-IW distribution and solve its identifiability problem by considering the LIW2 distribution. In Section 3, we provide some mathematical properties of the LIW2 distribution. In Section 4, we discuss the ML estimation approach in the presence of right-censored data. A bootstrap-based bias correction procedure is also presented to reduce the bias associated with the ML estimation of the model parameters. In Section 5, we show a simulation study designed to verify the effectiveness of the proposed estimators. In Section 6, we illustrate the usefulness of the LIW2 distribution through an aircraft dataset. Finally, concluding remarks are provided in Section 7.

2 The LIW2 distribution

The inverse Weibull distribution has received some attention in the literature. Keller et al. (1982) studied the shapes of the probability density and failure rate functions for the basic inverse model.

Let T be a random variable with inverse Weibull distribution, i.e. $T \sim \mathrm{IW}(\alpha, \beta)$. Then, its cumulative distribution function (cdf) can be written as

$$G(t; \alpha, \beta) = \exp\left\{ -\left(\frac{\alpha}{t}\right)^{\beta} \right\}, \quad t > 0, \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ are, respectively, the scale and shape parameters, and its probability density function (pdf) is given by

$$g(t; \alpha, \beta) = \beta \alpha^{\beta} t^{-(\beta+1)} \exp\left\{ -\left(\frac{\alpha}{t}\right)^{\beta} \right\}.$$

On the other hand, the class of Kumaraswamy distributions (Kumaraswamy, 1980), denoted by $\mathrm{Kum}(\lambda, b)$, has cdf given by

$$F\left(x; \lambda, b\right) = 1 - \left(1 - x^{\lambda}\right)^{b}, \quad 0 < x < 1, \tag{2}$$

and its pdf is

$$f\left(x; \lambda, b\right) = \lambda b x^{\lambda - 1} \left(1 - x^{\lambda}\right)^{b-1},$$

where $\lambda > 0$ and $b > 0$ are the two shape parameters.

Considering that a random variable X has distribution $G(\cdot)$, Cordeiro and de Castro (2011) suggested to apply the Kumaraswamy distribution to $G(x)$. Note that, since $0 < G(x) < 1$ for any distribution G, then we have, by applying $G(x)$ to Equation (2), that

$$F_G(x; \lambda, b) = 1 - \left(1 - [G(x)]^{\lambda}\right)^{b}, \tag{3}$$

where $\lambda > 0$ and $b > 0$ are the new shape parameters. Hence, F_G is the cdf of the generalized Kumaraswamy-G distribution. The great advantage of the class of Kumaraswamy distributions is that it has a closed-form cdf.

Inserting Equation (1) into Equation (3), we get the cdf of the Kum-IW distribution (Shahbaz et al., 2012), which is given by

$$F_G(t; \lambda, b, \alpha, \beta) = 1 - \left(1 - \exp\left\{ -\lambda \left(\frac{\alpha}{t}\right)^{\beta} \right\}\right)^{b}, \quad t > 0. \tag{4}$$

Although the distribution above has a simple structure, its parameters are non-identifiable. A parameter θ for a family of distributions F such that $F = \{f : f(x; \theta),\, x \in \mathbb{R}, \theta \in \Theta\}$ is said to be identifiable if different values of θ correspond to different probability density or mass functions. That is to say, if $\theta \neq \theta'$ then $f(x; \theta)$ is not the same function as $f(x; \theta')$.

Let T be a positive random variable with cdf given by Equation (4). Also, let $\boldsymbol{\theta} = (\lambda, b, \alpha, \beta) \in \mathbb{R}^4_+$. Then, the Kum-IW model, as defined in Equation (4), is not identifiable. The proof is available in Appendix A. It is essential to clarify that identifiability is a property of the model and not of a particular estimation technique, but if a model is not identifiable, then the inference can be difficult.

In order to overcome the identifiability problem pointed out above, we isolate the quantity c in Equation (10), obtaining $c = \alpha\lambda^{\frac{1}{\beta}}$ (see Gusmão et al., 2017). The proof is available in Appendix B. Then, we rewrite Equation (4) as

$$F_G(t; b, c, \beta) = 1 - \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\right)^b, \quad t > 0, \tag{5}$$

where $b > 0$ and $\beta > 0$ are the shape parameters, and $c > 0$ is the scale parameter. Hence, the new parameterized (identifiable) Kum-IW distribution (also referred to as the LIW2 distribution) has three parameters, and its pdf is given by

$$f_G(t; b, c, \beta) = \beta b c^\beta t^{-(\beta+1)} \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\left(1 - \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\right)^{b-1}. \tag{6}$$

According to Cordeiro and Castro (2011), we have the following expansions for the cdf and pdf of the LIW2 distribution:

$$F_G(t; b, c, \beta) = 1 - b\sum_{r=0}^{\infty} \frac{(-1)^r\Gamma(b)}{(b-r)\Gamma(b-r)r!} \exp\left\{-r\left(\frac{c}{t}\right)^\beta\right\}$$

and

$$f_G(t; b, c, \beta) = \beta b c^\beta t^{-(\beta+1)} \sum_{r=0}^{\infty} \frac{(-1)^r\,\Gamma(b)}{\Gamma(b-r)\,r!} \exp\left\{-\left(\frac{c}{t}\right)^\beta\ (r+1)\right\}.$$

Notice that taking $b = 1$ in (6), we have the inverse Weibull distribution with pdf given by

$$g(t; c, \beta) = \beta c^\beta t^{-(\beta+1)} \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}.$$

3 Other properties of the LIW2 distribution

The reliability (or survival) and failure rate (or hazard) functions of the LIW2 distribution are given, respectively, by

$$S_G(t; b, c, \beta) = \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\right)^b \quad \text{and} \quad h_G(t; b, c, \beta) = \frac{\beta b c^\beta t^{-(\beta+1)} \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}}{1 - \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}}.$$

Figure 1 shows some examples of the shapes of the density and failure rate functions of the LIW2 distribution, considering different values of b, c, and β.

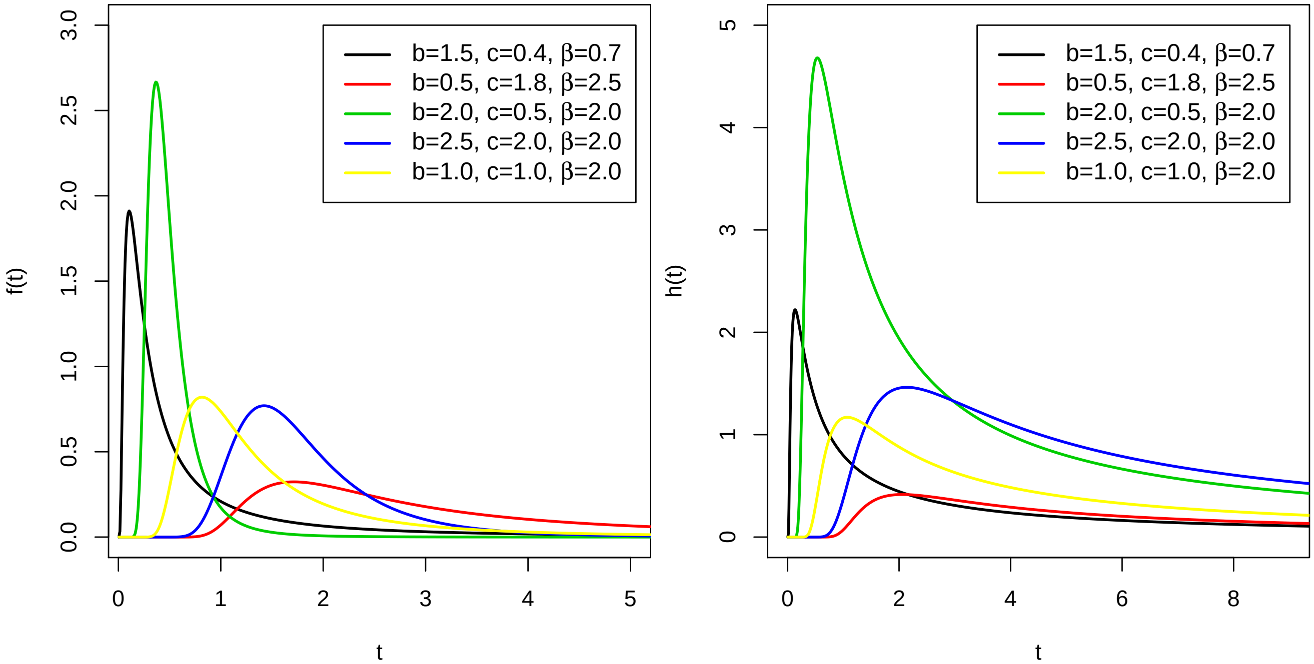

Figure 1: Density and failure rate functions of the LIW2 distribution for different values of b, c and β.

As can be seen in Figure 1, the failure rate function has a unimodal shape for different parameters' values. The proof that the failure rate function is only unimodal is not an easy task due to the complexity of the density and failure rate functions.

For $p \in (0,1)$, the quantile function of the LIW2 distribution is given by

$$Q_G(p;b,c,\beta) = F_G^{-1}(p;b,c,\beta) = c\left(-\log\left(1-(1-p)^{\frac{1}{b}}\right)\right)^{-\frac{1}{\beta}}.$$

Corollary 3.1 *Let T be a positive random variable with cdf given by Equation (5). Then, the k-th moment of T about the origin can be computed as*

$$E\left[T^k\right] = bc^k \sum_{r=0}^{\infty} \frac{(-1)^r\,\Gamma(b)}{\Gamma(b-r)\,r!}(r+1)^{\frac{k}{\beta}-1}\,\Gamma\left(1-\frac{k}{\beta}\right), \qquad \textit{for } \beta > k,$$

where $\Gamma(\cdot)$ is the gamma function.

Proof. The above result can be obtained as follows:

$$\begin{aligned}
E\left[T^k\right] &= \int_{-\infty}^{\infty} t^k f(t)dt = \int_0^{\infty} t^k \beta b c^\beta t^{-(\beta+1)} \exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\left(1-\exp\left\{-\left(\frac{c}{t}\right)^\beta\right\}\right)^{b-1} dt \\
&= \int_0^{\infty} t^k \beta b c^\beta t^{-(\beta+1)} \sum_{r=0}^{\infty} \frac{(-1)^r\,\Gamma(b)}{\Gamma(b-r)\,r!} \exp\left\{-\left(\frac{c}{t}\right)^\beta\,(r+1)\right\} dt \\
&= b\sum_{r=0}^{\infty} \frac{(-1)^r\,\Gamma(b)}{\Gamma(b-r)\,r!} \int_0^{\infty} \beta c^\beta t^{k-\beta-1} \exp\left\{-c^\beta t^{-\beta}(r+1)\right\} dt.
\end{aligned}$$

Then, making the changes of variable:

$$u = c^{\beta} t^{-\beta}(r+1)$$

and

$$du = -\beta c^{\beta} t^{-\beta-1}(r+1)dt,$$

we have $t^{-\beta} = uc^{-\beta}(r+1)^{-1}$ and $t^k = c^k (r+1)^{\frac{k}{\beta}} u^{-\frac{k}{\beta}}$. Consequently,

$$\begin{aligned} E\left[T^k\right] &= b\sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)}{\Gamma(b-r)\, r!} \int_0^{\infty} c^k u^{\left(1-\frac{k}{\beta}\right)-1} (r+1)^{\frac{k}{\beta}-1} \exp\{-u\} du \\ &= bc^k \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)}{\Gamma(b-r) r!} (r+1)^{\frac{k}{\beta}-1} \int_0^{\infty} u^{\left(1-\frac{k}{\beta}\right)-1} \exp\{-u\} du \\ &= bc^k \Gamma\left(1-\frac{k}{\beta}\right) \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)}{\Gamma(b-r)\, r!} (r+1)^{\frac{k}{\beta}-1}. \quad \blacksquare \end{aligned}$$

Similarly to the proof of Corollary 3.1, we obtain that the k-th negative and logarithmic moments of T are given by

$$E\left[T^{-k}\right] = \frac{kb}{\beta c^k} \Gamma\left(\frac{k}{\beta}\right) \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)}{\Gamma(b-r)\, r!} (r+1)^{\frac{-k}{\beta}-1}$$

and

$$E\left[\log(T^k)\right] = kb \sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)}{\Gamma(b-r)\, r!} (r+1)^{-1} \left[\frac{1}{\beta}\log(r+1) + \log(c) + \frac{\gamma}{\beta}\right],$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

Corollary 3.2 *Let $T_1, T_2, \ldots, T_n$ be independent and identically distributed random variables from a LIW2 distribution. Then, the density $f_{T_{(j)}}(t; b, c, \beta)$ of the j-th order statistic, for $j = 1, 2, \ldots, n$, is given by*

$$f_{T_{(j)}}(t; b, c, \beta) = \frac{\beta c^{\beta} t^{-(\beta+1)} \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}}{B(j, n-j+1)} \sum_{r=0}^{n-j} (-1)^r \binom{n-j}{r} \left[1 - \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^b\right]^{j+r-1},$$

where $B(.,.)$ denotes the beta function.

Proof. This proof is similar to that given in Casella and Berger (2002, p.299). $\blacksquare$

Now, from Corollary 3.2 we get that the minimum and maximum densities are, respectively, expressed as

$$f_{T_{(1)}}(t; b, c, \beta) = n\beta b c^{\beta} t^{-(\beta+1)} \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{b(n-1)} \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\} \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{b-1}$$

and

$$f_{T_{(n)}}(t; b, c, \beta) = n\beta b c^{\beta} t^{-(\beta+1)} \left[1 - \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^b\right]^{n-1} \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\} \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{b-1}.$$

Corollary 3.3 *The characteristic function of a random variable T with LIW2 distribution is given by*

$$\Psi_T(s) = E\left[\exp\{isT\}\right] = \sum_{k=0}^{\infty} \frac{b(isc)^k}{k!} \Gamma\left(1 - \frac{k}{\beta}\right) \sum_{r=0}^{\infty} \frac{(-1)^r\, \Gamma(b)}{\Gamma(b-r)\, r!} (r+1)^{\frac{k}{\beta}-1},$$

for $\beta > k$.

Proof. By expanding the exponential function in power series, we have, for random variable T,

$$\exp\{isT\} = \sum_{k=0}^{\infty} \frac{(isT)^k}{k!}.$$

Now, for $\beta > k$, we get

$$\Psi_T(s) = E\left[\exp\{isT\}\right] = \sum_{k=0}^{\infty} \frac{(is)^k E\left[T^k\right]}{k!}$$

and the result follows by using Corollary 3.1. ∎

It is well known in information theory that entropy is an important measure of uncertainty associated with a random variable. Perhaps the Shannon's entropy proposed by Shannon (1948) and defined by $H_{\mathrm{S}}(T) = E[-\log f(T)]$, with $f(.)$ being a particular pdf, is the most widely employed in applications; see Kapur (1994). However, in recent years, Rényi's entropy introduced by Rényi (1961) has substantially attracted several researchers. This is since the Rényi's entropy generalizes several other entropy measures, including Shannon's entropy; see, e.g., Csiszar and Korner (2011) and Jost (2006).

The Rényi's entropy of a positive random variable T with pdf $f(.)$ is defined by

$$H_{\mathrm{R}}(T) = \frac{1}{1-\alpha} \log\left\{\int_0^{\infty} [f(t)]^{\alpha}\, dt\right\},$$

where $\alpha > 0$ and $\alpha \neq 1$.

Proposition 3.4 *The Rényi's entropy of a random variable T with LIW2 distribution is given by*

$$H_{\mathrm{R}}(T) = \frac{1}{1-\alpha} \log\left\{\beta^{\alpha-1} b^{\alpha} c^{1-\alpha} \sum_{r=0}^{\infty} (-1)^r \binom{\alpha(b-1)}{r} (\alpha+r)^{\frac{1-\alpha(\beta+1)}{\beta}} \Gamma\left(\frac{\alpha(\beta+1)-1}{\beta}\right)\right\}, \tag{7}$$

where $\alpha > \frac{1}{\beta+1}$, for all $\beta > 0$.

Proof. In fact,

$$H_{\mathrm{R}}(T) = \frac{1}{1-\alpha} \log\left\{\beta^{\alpha} b^{\alpha} c^{\alpha\beta} \int_0^{\infty} t^{-\alpha(\beta+1)} \exp\left\{-\alpha\left(\frac{c}{t}\right)^{\beta}\right\} \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{\alpha(b-1)} dt\right\}.$$

By using the expansion $(1-x)^{b-1} = \sum_{r=0}^{\infty} (-1)^r \binom{b-1}{r} x^r$, for $-1 < x < 1$ and $b > 0$, we have

$$H_{\mathrm{R}}(T) = \frac{1}{1-\alpha} \log\left\{\beta^{\alpha} b^{\alpha} c^{\alpha\beta} \sum_{r=0}^{\infty} (-1)^r \binom{\alpha(b-1)}{r} \int_0^{\infty} t^{-\alpha(\beta+1)} \exp\left\{-(\alpha+r)\left(\frac{c}{t}\right)^{\beta}\right\} dt\right\}.$$

Then, making the change of variable:

$$u = c^{\beta} t^{-\beta}(r+1) \Rightarrow du = -\beta c^{\beta} t^{-\beta-1}(r+\alpha)dt,$$

and, hence, we have $t^{-\beta} = uc^{-\beta}(r+\alpha)^{-1}$. Thus, after some algebraic manipulations, we get (7). ■

From Proposition 3.4, taking $\alpha \to 1$, we obtain the Shannon's entropy expressed as

$$H_{\mathrm{S}}(T) = -\log(\beta b c^{\beta}) + b\sum_{r=0}^{\infty} \frac{(-1)^r \Gamma(b)(r+1)^{-1}}{\Gamma(b-r)\, r!} \left\{(\beta+1)\left[\frac{\log(r+1)}{\beta} + \log(c) + \frac{\gamma}{\beta}\right] + (r+1^{-1})\right\} + \frac{(b-1)}{b},$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

4 Maximum likelihood estimation for right-censored LIW2 data

A model is said to be identifiable when parameter values uniquely determine the probability distribution of data, and the probability distribution of data uniquely determines the parameter values. Therefore, when a model is unidentifiable (e.g., the Kum-IW distribution), various parameter values correspond to the same data distribution. This fact makes impossible the use of classical methods of inference, in which there is the need for an identifiable model (e.g., the LIW2 distribution) to be obtained, for example, ML estimators that are unique.

Let T_i be a random variable with a LIW2 distribution (5) indexed by a parameter vector $\boldsymbol{\theta} = (b, c, \beta)$. Since the data in survival and reliability analysis are generally censored, a straightforward random censoring mechanism, which is often realistic, is one in which each subject or device i is assumed to have a lifetime T_i and a censoring time C_i, where T_i and C_i are independent random variables. Let δ_i be an indicator of whether the actual lifetime of the i-th individual or item is observed or not, i.e. $\delta_i = 1$ if observed and $\delta_i = 0$ if not. Suppose that the data set consists of n independent observations $t_i = \min\{T_i, C_i\}$, for $i = 1, 2, \ldots, n$, where the distribution of C_i does not depend on the parameters of interest (the so-called non-informative censoring hypothesis). Classical parametric inference for such data is typically based on likelihood methods and their asymptotic theory.

The censored likelihood function for the LIW2 distribution can be written as

$$\begin{aligned} L(\boldsymbol{\theta}) =& \beta^r b^r c^{r\beta} \prod_{i=1}^{n} t_i^{-\delta_i(\beta+1)} \prod_{i=1}^{n} \left(1 - \exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right)^{\delta_i(b-1)} \exp\left\{-\sum_{i=1}^{n} \delta_i \left(\frac{c}{t_i}\right)^{\beta}\right\} \\ & \times \prod_{i=1}^{n} \left(1 - \exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right)^{b(1-\delta_i)}, \end{aligned}$$

where $r=\sum_{i=1}^{n}\delta_i$. The log-likelihood function is

$$
\begin{aligned}
\ell(\boldsymbol{\theta}) = {} & r\log\left(\beta b c^{\beta}\right) - (\beta+1)\sum_{i=1}^{n}\delta_i\log(t_i) + (b-1)\sum_{i=1}^{n}\delta_i\log\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right) \\
& + b\sum_{i=1}^{n}(1-\delta_i)\log\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right) - c^{\beta}\sum_{i=1}^{n}\delta_i\left(\frac{1}{t_i}\right)^{\beta},
\end{aligned} \tag{8}
$$

The score functions of the log-likelihood function (8) are obtained as follows:

$$
U_b(\boldsymbol{\theta}) = \frac{\partial\ell(\boldsymbol{\theta})}{\partial b} = \frac{r}{b} + \sum_{i=1}^{n}\delta_i\log\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right) + \sum_{i=1}^{n}(1-\delta_i)\log\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right),
$$

$$
\begin{aligned}
U_c(\boldsymbol{\theta}) = \frac{\partial\ell(\boldsymbol{\theta})}{\partial c} = {} & \frac{r\beta}{c} - \beta c^{\beta-1}\sum_{i=1}^{n}\delta_i\left(\frac{1}{t_i}\right)^{\beta} + \beta c^{\beta-1}(b-1)\sum_{i=1}^{n}\delta_i\frac{t_i^{-\beta}\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}}{1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}} \\
& + \beta b c^{\beta-1}\sum_{i=1}^{n}(1-\delta_i)\frac{t_i^{-\beta}\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}}{1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}}
\end{aligned}
$$

and

$$
\begin{aligned}
U_\beta(\boldsymbol{\theta}) = \frac{\partial\ell(\boldsymbol{\theta})}{\partial\beta} = {} & \frac{r}{\beta} + r\log(c) - c^{\beta}\log(c)\sum_{i=1}^{n}\delta_i t_i^{-\beta} + c^{\beta}\sum_{i=1}^{n}\delta_i t_i^{-\beta}\log(t_i) \\
& - \sum_{i=1}^{n}\delta_i\log(t_i) + c^{\beta}(b-1)\sum_{i=1}^{n}\delta_i t_i^{-\beta}\log\left(\frac{c}{t_i}\right)\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right)^{-1} \\
& + bc^{\beta}\sum_{i=1}^{n}(1-\delta_i)t_i^{-\beta}\log\left(\frac{c}{t_i}\right)\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\left(1-\exp\left\{-\left(\frac{c}{t_i}\right)^{\beta}\right\}\right)^{-1}.
\end{aligned}
$$

The maximum likelihood estimator (MLE) $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is then obtained by solving the nonlinear likelihood equations: $U_b(\boldsymbol{\theta})=0$, $U_c(\boldsymbol{\theta})=0$ and $U_\beta(\boldsymbol{\theta})=0$. Notice that these equations can not be solved analytically, but we can use, for instance, the *optim* or *maxLik* routines of R software (R Core Team, 2018) to find the parameter estimates numerically.

Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, $\widehat{\boldsymbol{\theta}}$ is asymptotically normally distributed with a joint trivariate normal distribution given by

$$
\sqrt{n}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \sim \mathrm{N}_3(\mathbf{0}, \mathbf{I}^{-1}(\boldsymbol{\theta})) \quad \text{for} \quad n\to\infty,
$$

where $\mathbf{I}(\boldsymbol{\theta})$ is the expected information matrix. This asymptotic behavior is still valid if $\mathbf{I}(\boldsymbol{\theta})$ is replaced by $\mathbf{J}(\widehat{\boldsymbol{\theta}})$, which is the observed information matrix evaluated at $\widehat{\boldsymbol{\theta}}$. Thus, the asymptotic trivariate normal distribution $\mathrm{N}_3(\mathbf{0}, \mathbf{J}^{-1}(\widehat{\boldsymbol{\theta}}))$ can be used to build approximate confidence intervals for the individual parameters, as well as for the failure rate and reliability functions. It is also useful for testing goodness-of-fit of the LIW2 distribution and for comparing this distribution with some of its special sub-models using one of the three well-known asymptotically equivalent test statistics, namely the likelihood ratio, Wald and Rao's score statistics.

It is well known that the MLEs $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_p\right)$ of the unknown parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ are usually biased for small samples, and this bias can be written as (see Cordeiro and Klein, 1994):

$$\text{Bias}(\hat{\theta}_m) = \sum_{j=1}^{p} s_{mj}(\boldsymbol{\theta}) \sum_{k=1}^{p} \sum_{l=1}^{p} s_{kl}(\boldsymbol{\theta}) \left(h_{jk}^{(l)}(\boldsymbol{\theta}) - 0.5 h_{jkl}(\boldsymbol{\theta}) \right) + O(n^{-2}),$$

where

$$h_{jkl}(\boldsymbol{\theta}) = E\left[\frac{\partial^3 \ell(\boldsymbol{\theta})}{\partial\theta_j \partial\theta_k \partial\theta_l}\right],$$

$$h_{jk}^{(l)}(\boldsymbol{\theta}) = \frac{\partial h_{jk}(\boldsymbol{\theta})}{\partial\theta_l} \quad \text{with} \quad h_{jk}(\boldsymbol{\theta}) = E\left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\theta_j \theta_k}\right],$$

and $s_{jk}(\boldsymbol{\theta})$ is the negative of the (j,k)-th element of the inverse of the Fisher information matrix for $\boldsymbol{\theta}$, that is, $\mathbf{I}^{-1}(\boldsymbol{\theta}) = \{-h_{jk}(\boldsymbol{\theta})\}^{-1} = \{-s_{jk}(\boldsymbol{\theta})\}$, for $j, k, l, m = 1, 2, \ldots, p$.

However, such terms can not be obtained since the elements of the Fisher information matrix do not have closed-form expressions. To overcome this problem, we can resort to bias correction using bootstrap techniques (Efron, 1992). In order to perform the non-parametric bootstrap method, let $\boldsymbol{t} = (t_1, \ldots, t_n)$ be a sample with n observations randomly drawn from a LIW2 distribution. The pseudo-samples (also referred to as the bootstrap samples) $\boldsymbol{t^*} = (t_1^*, \ldots, t_n^*)$ are obtained by resampling with replacement from the original sample $\boldsymbol{t}$. If B bootstrap samples $\{\boldsymbol{t}^{*(1)}, \ldots, \boldsymbol{t}^{*(B)}\}$ are generated independently from $\boldsymbol{t}$ and their respective estimates $\{\hat{\theta}^{*(1)}, \ldots, \hat{\theta}^{*(B)}\}$ are calculated using the ML method, then the bootstrap expectations are approximated by

$$\hat{\theta}^{*(\cdot)} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}^{*(b)}. \tag{9}$$

From (9), the bootstrap bias estimate is given by $\hat{B}_F(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - \hat{\theta}$, where $\hat{\theta}$ is the MLE of θ. The bias-corrected MLE obtained by the bootstrap resampling method is given by

$$\hat{\theta}^{\text{B}} = \hat{\theta} - \hat{B}_F(\hat{\theta}, \theta) = 2\hat{\theta} - \hat{\theta}^{*(\cdot)}.$$

Here, we have $\hat{\theta}^{\text{B}}$ denoted by $\hat{\boldsymbol{\theta}}^{\text{B}} = (\hat{b}^{\text{B}}, \hat{c}^{\text{B}}, \hat{\beta}^{\text{B}})$.

5 Simulation study

In this section, we perform a Monte Carlo simulation study to check the validity of the results discussed in the previous sections. All computations were carried out using the R software. The following approach was adopted:

1. Generate n values from the LIW2 distribution with parameters b, c and β;

2. Using the values obtained at step 1, calculate the estimates (MLEs and bootstrap bias-corrected MLEs) $\hat{b}$, $\hat{c}$ and $\hat{\beta}$ of the parameters b, c and β, respectively;

3. Repeat steps 1-2 M times;

4. Considering $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) = (\hat{b}, \hat{c}, \hat{\beta})$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (b, c, \beta)$, compute the mean relative estimate (MRE), $\frac{1}{M}\sum_{m=1}^{M} \frac{\hat{\theta}_{l,m}}{\theta_l}$, the mean square error (MSE), $\frac{1}{M}\sum_{m=1}^{M}(\hat{\theta}_{l,m} - \theta_l)^2$, where $\hat{\theta}_{l,m}$ denotes the estimate of θ_l obtained from sample m, for $l = 1, 2, 3$ and $m = 1, 2, \ldots, M$. Also, calculate the coverage probability (CP) of the asymptotic normal 95% confidence intervals for these parameters, $\frac{1}{M}\sum_{m=1}^{M} I\left(\hat{\theta}_{l,m} - 1.959964\ s_{\hat{\theta}_{l,m}} < \theta_l < \hat{\theta}_{l,m} + 1.959964\ s_{\hat{\theta}_{l,m}}\right)$, where $I(\cdot)$ is the indicator function and $s_{\hat{\theta}_{l,m}}$ denotes the estimated standard error of $\hat{\theta}_{l,m}$.

With this approach, it is expected that the bias and MSE will return values closer to zero, and the MRE will be closer to one. The 95% coverage probabilities were also computed, considering the 95% confidence interval. For a large number of experiments using a 95% confidence intervals, the frequencies of these intervals that covered the true values of $\boldsymbol{\theta}$ should be closer to 0.95. The coverage probabilities were calculated using the numeric observed information matrix obtained from the maxLik package results.

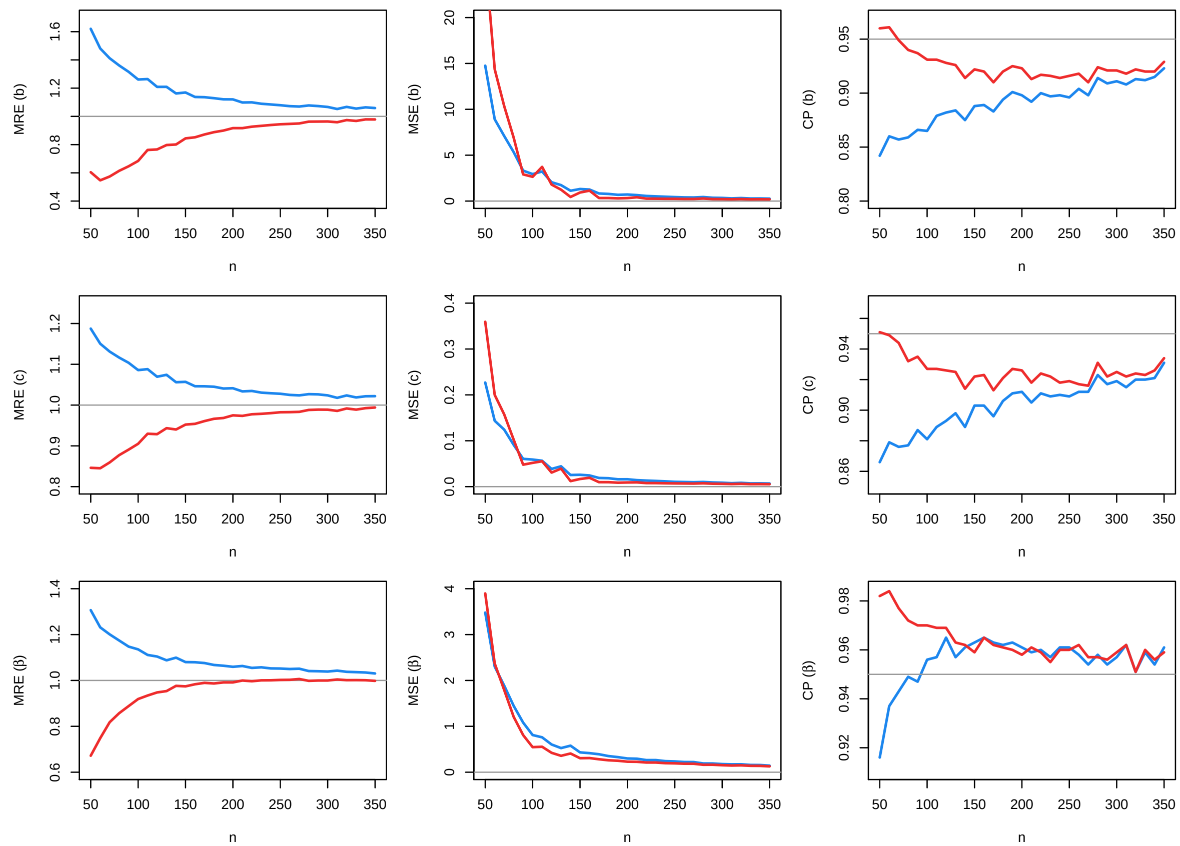

Figure 2: MRE, MSE and CP for the MLEs (solid blue line) and bootstrap bias-corrected MLEs (solid red line) of $b = 1.3$, $c = 0.5$ and $\beta = 2$, based on $M = 5,000$ generated samples of size n with 30% of censoring.

The seed used by the pseudo-random number generators was 2018. We set $M = 5,000$, $n = \{50, 60, \dots, 350\}$ and $\boldsymbol{\theta} = (1.3, 0.5, 2)$, as well as two different censoring rates: 30% and 50%. In order to generate randomly censored data, we used the same procedures as in Goodman et al. (2006). Here, we considered $B = 500$ for the bootstrap method. The MLEs were computed using the log-likelihood function (8) with the R routine *maxBFGS* from the "maxLik" package (Henningsen and Toomet, 2011), which was able to locate the maximum of the log-likelihood surface for a wide range of starting values. In this case, the solution for the maximum was unique for all starting values.

Tables 4-5 (see Appendix C) display the MRE, MSE and CP for the MLEs of b, c and β, considering $M = 5,000$ simulated samples with different sizes and different percentages of censored observations. Figures 2-3 summarize these results. From these tables and figures, we can observe that, for all parameters, the MSE tends to zero and the MRE tends to one as n increases, i.e., the MLEs are asymptotically unbiased and consistent. However, the bootstrap bias-corrected MLEs (BCMLEs, whose results are also shown in the tables mentioned above and figures) returned more precise estimates when compared with the MLEs. Additionally, for all parameters, the estimated CP of the asymptotic normal confidence interval is closer to the nominal value (0.95) using the bootstrap approach. Based on these results, we can conclude that the ML estimation method with the bootstrap resampling technique works well to find the estimates for the LIW2 distribution parameters in the presence of right-censored data.

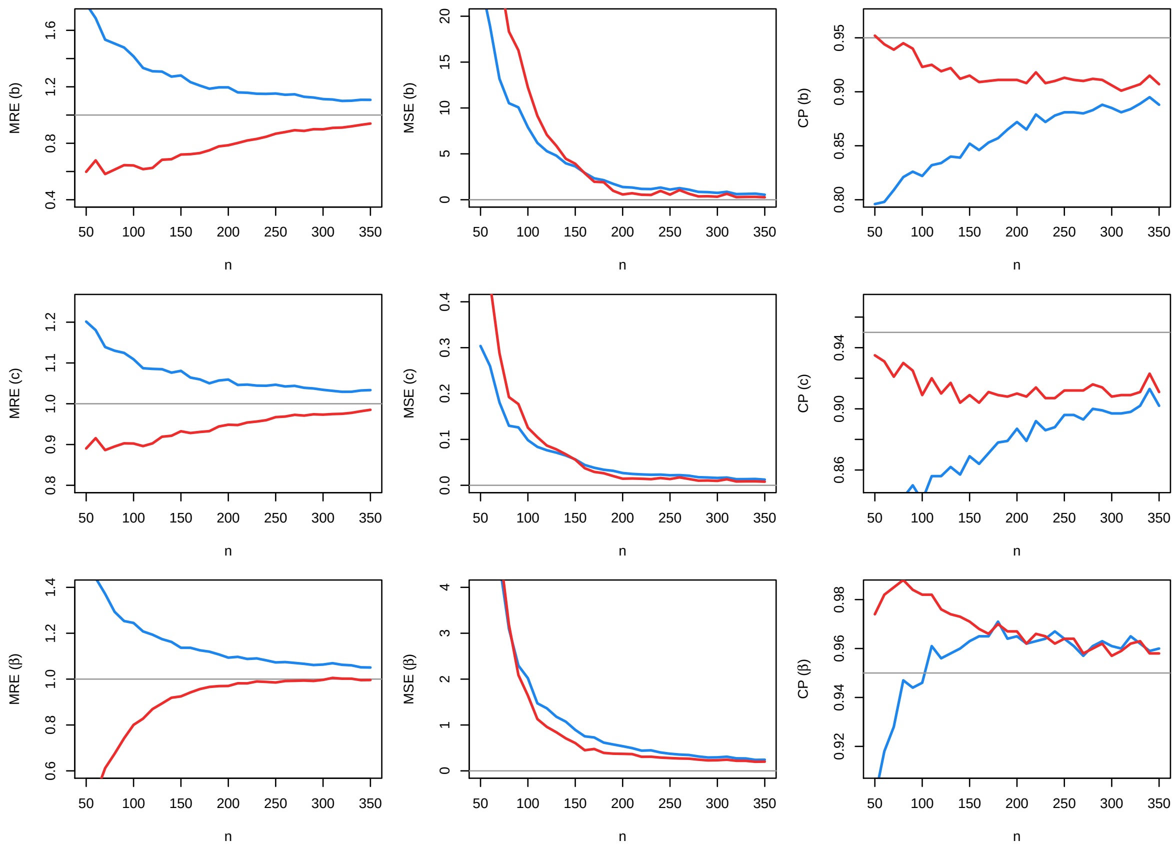

Figure 3: MRE, MSE and CP for the MLEs (solid blue line) and bootstrap bias-corrected MLEs (solid red line) of $b = 1.3$, $c = 0.5$ and $\beta = 2$, based on $M = 5,000$ generated samples of size n with 50% of censoring.

6 Application

In this section, we recall the real data set related to the failure time (in days) of 194 devices in an aircraft, first presented in Ramos et al. (2018). All data are available in Table 1, where + indicates the presence of censorship. The choice of the distribution that better fits these data is essential to avoid higher costs for the airline company.

The results obtained using the LIW2 distribution were compared to the corresponding ones achieved

Table 1: Data set related to the failure time (in days) of 194 devices in an aircraft.

43	29	37	88	5	14	9	43+	1	78	1	77	17	100
3	119+	22	3	8	80	1	19	157+	65	34	13	62+	2
1	1	2	3	6	1	2	5	7	6	1	1	4	1
1	1	2	7	2	1	1	2	1	1	7	1	1	4
1	4	2	4	5	5	4	3	2	2	2	3	3	9
1	6	9	2	5	7	4	2	1	2	2	3	11	8
3	1	2	2	2	2	2	1	3	20+	8	8	197	20
14	7	29	7	16	34	25	10	80	42	32	1	3	1
12	7	7	39+	60	53	32	9	8	1	1	27	2	4
8	13	7	7	1	19	7	12	19	5	18	1	4	18
20	9	14	13	70	18	3	7	20	3	11	10	3	38+
278	13	79	145+	19	2	18	2	65	14	31	10	19	5
9	45	13	5	1	1	31	35	34	4	3	5	12	140+
106	5	40	130+	21	19	7	10	91	193	64	85+		

with the use of the three-parameter extended Poisson-Weibull (EPW) (Ramos et al., 2020), generalized extended exponential-Poisson (GE2P) (Ramos et al., 2020) and generalized gamma (GG) (Stacy, 1962) distributions and four-parameter Kum-IW distribution given in Equation (4) due to Shahbaz et al. (2012), in addition to the non-parametric reliability curve estimated using the Kaplan-Meier method (Kaplan and Meier, 1958). First, in order to identify the behavior of the empirical failure rate function, we considered the total-time-on-test (TTT) plot introduced by Barlow and Campo (1975). The TTT plot is obtained by plotting $(r/n, G(r/n))$, for $r = 1, 2, \ldots, n$, where $G(r/n) = \left(\sum_{i=1}^{r} T_{(i)} + (n-r)T_{(r)}\right) / \sum_{i=1}^{n} T_{(i)}$ and $T_{(i)}$ is the order statistic of the sample T_i $(i = 1, 2, \ldots, n)$. If the TTT plot is concave, convex, first convex then concave and first concave then convex, the shape of the corresponding failure rate function is increasing, decreasing, bathtub and unimodal, respectively.

Different model discrimination criteria were also considered, such as the AIC (Akaike Information Criterion), AICc (Corrected Akaike Information Criterion), HQIC (Hannan-Quinn Information Criterion) and CAIC (Consistent Akaike Information Criterion), which are calculated as follows: $\text{AIC} = -2\ell(\hat{\boldsymbol{\theta}}) + 2p$, $\text{AICc} = \text{AIC} + [2\,p\,(p+1)/(n-p-1)]$, $\text{HQIC} = -2\,\ell(\hat{\boldsymbol{\theta}}) + 2\,p\,\log(\log(n))$ and $\text{CAIC} = -2\ell(\hat{\boldsymbol{\theta}}) + p(\log(n) + 1)$, where p is the number of parameters in the model and $\hat{\boldsymbol{\theta}}$ is the bootstrap bias-corrected MLE of $\boldsymbol{\theta}$.

The best fitted model is the one that provides the minimum values of these criteria.

Figure 4 presents the TTT plot, the reliability function adjusted by different distributions, and the Kaplan-Meier estimate, as well as the failure rate function adjusted by the LIW2 distribution. Moreover, Table 2 displays the AIC, AICc, HQIC, and CAIC values for the different probability distributions.

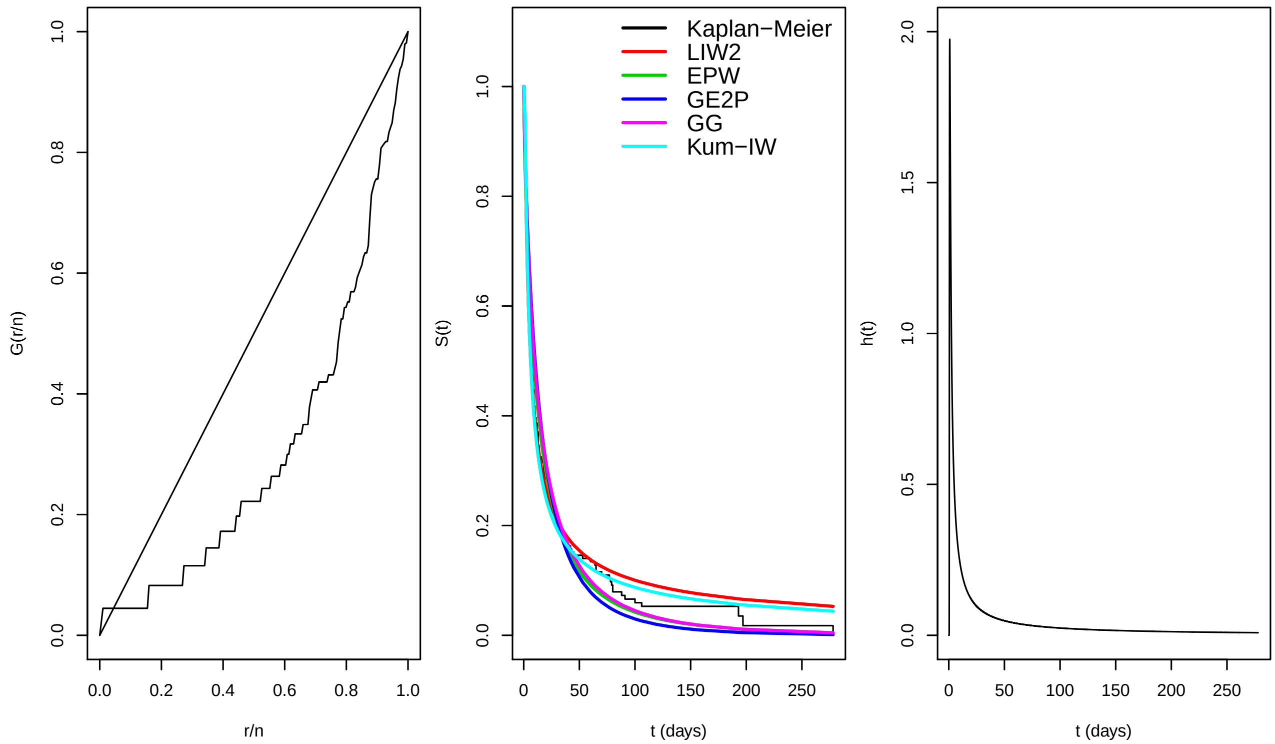

Figure 4: Left panel: TTT plot. Middle panel: reliability function adjusted by different probability distributions and the Kaplan-Meier estimate. Right panel: failure rate function adjusted by the LIW2 distribution, considering the data set related to the failure time of 194 devices in an aircraft

.

Table 2: AIC, AICc, HQIC and CAIC values for different probability distributions, considering the data set related to the failure time of 194 devices in an aircraft.

Criterion	LIW2	EPW	GE2P	GG	Kum-IW
AIC	**1396.026**	1437.686	1451.593	1430.244	1396.373
AICc	**1396.153**	1437.812	1451.719	1430.371	1396.585
HQIC	**1399.996**	1441.656	1455.562	1434.214	1401.666
CAIC	**1408.830**	1450.490	1464.396	1443.048	1413.444

Based on the TTT plot, there is an indication that the failure rate function is unimodal (or upsidedown bathtub-shaped). By comparing the Kaplan-Meier curve with the fitted curves drawn from the five candidate parametric models, we can see that the LIW2 distribution gives a better fit to the aircraft data set. This finding is corroborated by the AIC, AICc, HQIC, and CAIC criteria values, which provide evidence in favor of the LIW2 distribution.

Table 3 shows the BCMLEs, standard errors (SE) and asymptotic normal 95% confidence intervals (95% CI) for the parameters of the LIW2 distribution.

Table 3: BCMLEs, SE and 95% CI for the parameters of the LIW2 distribution, considering the aircraft data set.

Parameter	BCMLE	SE	95% CI
b	0.744	0.248	(0.259; 1.223)
c	2.808	1.066	(0.718; 4.898)
β	0.859	0.213	(0.446; 1.272)

In order to verify the goodness of fit of the LIW2 distribution to the aircraft data, we calculate the Cox-Snell residuals (Cox and Snell, 1968) defined by

$$e_i = -\log\left(\hat{S}(t_i)\right), \quad i = 1, 2, \ldots, n,$$

where $\hat{S}(t_i)$ is the fitted LIW2 reliability function of the i-th lifetime. If the LIW2 distribution is correctly specified, then Cox-Snell residuals e_i's are a censored random sample from the standard exponential distribution, that is, $e_i \sim \text{Exp}(1)$, for $i = 1, 2, \ldots, n$.

Figure 5 presents the graph of Kaplan-Meier *versus* standard exponential survival, both fitted to the Cox-Snell residuals. Observe that most of the points are over the line, showing the goodness of fit for the proposed data set using the LIW2 distribution. Therefore, from the proposed methodology, the data set related to the failure time of 194 devices in an aircraft can be satisfactorily described by the LIW2 distribution.

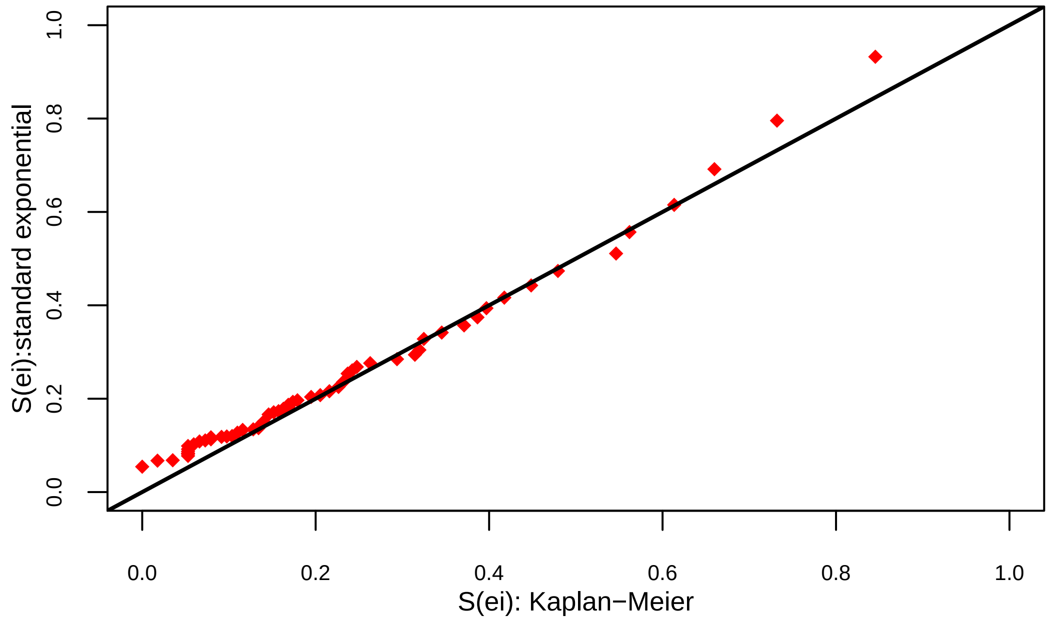

Figure 5: Kaplan-Meier vs standard exponential survival, both fitted to the Cox-Snell residuals.

7 Conclusions

In this paper, we revisited the LIW2 distribution and presented several mathematical properties of this distribution, which can be used in situations where the data present a unimodal failure rate. Initially, we proved that such a model is the reparameterized version of the Kum-IW model without identifiability problems. Then, the LIW2 model parameters estimation was discussed under a random right-censored data scheme, which is often realistic to describe survival and reliability data. Since the MLEs are biased for small sample sizes, we proposed a bias correction approach using bootstrap techniques.

An extensive numerical simulation study was conducted to assess the performance of our proposed methodology, which revealed that the bias-corrective approach should be used to achieve good estimates for the parameters of the LIW2 distribution. These results are of great practical interest since they will enable the use of the LIW2 distribution in various application issues.

Many possible extensions of the current work can be further considered. The presence of covariates, as well as of long-term survivals, is ubiquitous in practice. Our approach should be investigated in both contexts; see, e.g., Perdoná and Louzada-Neto (2011).

Acknowledgements

The research of Francisco Louzada is supported by the Brazilian Governmental Agencies, CNPq and FAPESP. Pedro L. Ramos acknowledges support from the São Paulo State Research Foundation (FAPESP Proc. 2017/25971-0).

References

Barlow, R. E., and Campo, R. A. (1975). Total Time on Test Processes and Applications to Failure Data Analysis. In: Reliability and Fault Tree Analysis, pp. 451-481, R. E. Barlow, J. Fussell and N. D. Singpurwalla (eds.). SIAM, Philadelphia.

Carrasco, J. M. F., Ortega, E. M. M., and Cordeiro, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis*, 53(2):450–462.

Casella, G., and Berger, R. L. (2002). *Statistical Inference*. 2nd Edition. Pacific Grove, CA: Duxbury.

Cordeiro, G. M., and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation & Simulation*, 81(7):883–898.

Cordeiro, G. M., and Klein, R. (1994). Bias correction in ARMA models. *Statistics & Probability Letters*, 19(3):169–176.

Cox, D. R., and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society*: Series B (Methodological), 30(2), 248-265.

Csiszar, I., and Korner, J. (2011). Information theory: coding theorems for discrete memoryless systems. *Cambridge University Press.*

Efron, B. (1992). Bootstrap Methods: Another Look at the Jackknife. In: Kotz S., Johnson N.L. (eds.) Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY.

Goodman, M. S., Li, Y., and Tiwari, R. C. (2006). Survival analysis with change point hazard functions. Harvard University Biostatistics Working Paper Series, Paper 40.

Gusmão, F. R. S., Ortega, E. M. M., and Cordeiro, G. M. (2012). Reply to the “Letter to the Editor” of M. C. Jones. *Statistical Papers*, 53(2):253–254.

Gusmão, F. R. S., Tomazella, V. L. D., and Ehlers, R. S. (2017). Bayesian estimation of the Kumaraswamy Inverse Weibull distribution. *Journal of Statistical Theory and Applications*, 16(2):248–260.

Henningsen, A., and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, 26(3):443–458.

Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(1):70–81.

Jost, L. (2006). Entropy and diversity. *Oikos*, 113(2), 363-375.

Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Kapur, J. N. (1994). Measures of information and their applications. *Wiley-Interscience*.

Keller, A. Z., Kamath, A. R. R., and Perera, U. D. (1982). Reliability analysis of CNC machine tools. *Reliability Engineering*, 3(6):449–473.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1–2):79–88.

Mudholkar, G. S., Srivastava, D. K., and Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4):436–445.

Perdoná, G. C., and Louzada-Neto, F. (2011). A general hazard model for lifetime data in the presence of cure rate. *Journal of Applied Statistics*, 38(7):1395–1405.

Pescim, R. R., Demétrio, C. G. B., Cordeiro, G. M., Ortega, E. M. M., and Urbano, M. R. (2010). The beta generalized half-normal distribution. *Computational Statistics & Data Analysis*, 54(4):945–957.

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0.

Ramos, P. L., Dey, D. K., Louzada, F., and Lachos, V. H. (2020). An extended poisson family of life distribution: A unified approach in competitive and complementary risks. *Journal of Applied Statistics*, 47(2), 306-322.

Ramos, P. L., Louzada, F., Shimizu, T. K. O., and Luiz, A. O. (2018). The inverse weighted Lindley distribution: Properties, estimation and an application on a failure time data. *Communications in Statistics - Theory and Methods*, 48(10):2372-2389.

Rényi, A. (1961). On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. *The Regents of the University of California.*

Shahbaz, M. Q., Shahbaz, S., and Butt, N. S. (2012). The Kumaraswamy-Inverse Weibull Distribution. *Pakistan Journal of Statistics & Operation Research*, 8(3):479–489.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379-423.

Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 33(3), 1187-1192.

Appendix A: On the non-identifiability of the Kum-IW distribution

In order to prove the identifiability problems of the Kum-IW distribution, note that, it is known that if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ implies $F_G(t;\boldsymbol{\theta}_1) \neq F_G(t;\boldsymbol{\theta}_2)$, then a parameter vector $\boldsymbol{\theta}$ for a family of distributions F_G is identifiable. Therefore, the model is identifiable.

Let $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ be parameter vectors in $\mathbb{R}_+^4$ such that $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$. Also, let $\boldsymbol{\theta}_i = (\lambda_i, b, \alpha_i, \beta)$ and consider the following relation:

$$\lambda_i = \left(\frac{c}{\alpha_i}\right)^{\beta}, \tag{10}$$

for $i = 1, 2$ and $c > 0$ fixed.

Considering the cdf F_G, given by Equation (4), as well as the parameter vector $\boldsymbol{\theta}_1 = (\lambda_1, b, \alpha_1, \beta)$, we have

$$F_G(t;\boldsymbol{\theta}_1) = 1 - \left(1 - \exp\left\{-\lambda_1\left(\frac{\alpha_1}{t}\right)^{\beta}\right\}\right)^{b}. \tag{11}$$

Substituting equation (10) into (11), we have

$$F_G(t;\boldsymbol{\theta}_1) = 1 - \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{b}.$$

Now, using Equation (4) and the parameter vector $\boldsymbol{\theta}_2 = (\lambda_2, b, \alpha_2, \beta)$, we have

$$F_G(t; \boldsymbol{\theta}_2) = 1 - \left(1 - \exp\left\{-\lambda_2 \left(\frac{\alpha_2}{t}\right)^{\beta}\right\}\right)^{b}. \tag{12}$$

Finally, considering Equations (10) and (12), we have

$$F_G(t; \boldsymbol{\theta}_2) = 1 - \left(1 - \exp\left\{-\left(\frac{c}{t}\right)^{\beta}\right\}\right)^{b}.$$

Thus, $F_G(t; \boldsymbol{\theta}_1) = F_G(t; \boldsymbol{\theta}_2)$ for all $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ where $\boldsymbol{\theta}_i = (\lambda_i, b, \alpha_i, \beta)$, with $\lambda_i = \left(\frac{c}{\alpha_i}\right)^{\beta}$ and $c > 0$ fixed, for $i = 1, 2$. Thus, a parameter vector $\boldsymbol{\theta}$ for a family of distributions F_G, given by Equation (4), is not identifiable, that is, the model is not identifiable.

Appendix B: On the identifiability of the LIW2 distribution

An important result is that the LIW2 model, as defined in Equation (5), is identifiable. In order to prove that, we must show that, for all $\boldsymbol{\theta}_i \neq \boldsymbol{\theta}_j$ with $i \neq j$ and $i, j \geq 1$, we have $F_G(t; \boldsymbol{\theta}_i) \neq F_G(t; \boldsymbol{\theta}_j)$. So, let $\boldsymbol{\theta}_i = (b_i, c_i, \beta_i)$ and $\boldsymbol{\theta}_j = (b_j, c_j, \beta_j)$, with $b_i \neq b_j$, $c_i \neq c_j$ and $\beta_i \neq \beta_j$, for all $i \neq j$ and $i, j \geq 1$. Suppose that $F_G(t; \boldsymbol{\theta}_i) = F_G(t; \boldsymbol{\theta}_j)$ for any $\boldsymbol{\theta}_i \neq \boldsymbol{\theta}_j$ and for all $t > 0$, with $i \neq j$ and $i, j \geq 1$. Hence,

$$F_G(t; \boldsymbol{\theta}_i) = F_G(t; \boldsymbol{\theta}_j) \quad \Rightarrow \quad 1 - \left(1 - \exp\left\{-\left(\frac{c_i}{t}\right)^{\beta_i}\right\}\right)^{b_i} = 1 - \left(1 - \exp\left\{-\left(\frac{c_j}{t}\right)^{\beta_j}\right\}\right)^{b_j},$$

for all $i \neq j$ and $i, j \geq 1$ and for all $t > 0$.

Let $t = 1$ and $c_{i_0}^{\beta_{i_0}} = c_{j_0}^{\beta_{j_0}}$, where $i_0 \in \{1, 2, 3, ...\}$, $j_0 \in \{1, 2, 3, ...\}$ and $i_0 \neq j_0$. Thus,

$$\left(1 - \exp\left\{-\left(\frac{c_{i_0}}{1}\right)^{\beta_{i_0}}\right\}\right)^{b_{i_0}} = \left(1 - \exp\left\{-\left(\frac{c_{j_0}}{1}\right)^{\beta_{j_0}}\right\}\right)^{b_{j_0}} \quad \Rightarrow \quad b_{i_0} = b_{j_0}.$$

By hypothesis, $b_{i_0} \neq b_{j_0}$ for all $i \neq j$ and $i, j \geq 1$ and for all $t > 0$. Hence, we have a contradiction. Then, a parameter vector $\boldsymbol{\theta}$ for a family of distributions F_G, given by Equation (5), is identifiable, i.e. the model is identifiable.

Appendix C: Simulation tables

Table 4: MRE, MSE and CP for the MLEs and BCMLEs of $b = 1.3$, $c = 0.5$ and $\beta = 2$, based on $M = 5,000$ generated samples of size n with 30% of censoring.

		MLE									BCMLE								
		b			c			β			b			c			β		
Cens.p	n	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**
30%	50	1.620	14.755	0.842	1.188	0.227	0.866	1.306	3.479	0.916	0.604	26.726	0.960	0.846	0.359	0.951	0.672	3.897	0.982
	70	1.412	7.095	0.857	1.131	0.124	0.876	1.201	1.888	0.943	0.574	10.356	0.949	0.859	0.158	0.944	0.818	1.793	0.977
	90	1.315	3.318	0.866	1.104	0.061	0.887	1.148	1.082	0.947	0.646	2.910	0.937	0.891	0.048	0.935	0.888	0.809	0.970
	110	1.264	3.223	0.879	1.088	0.056	0.889	1.111	0.761	0.957	0.762	3.728	0.931	0.930	0.055	0.927	0.934	0.555	0.969
	130	1.209	1.739	0.884	1.074	0.044	0.898	1.088	0.526	0.957	0.797	1.242	0.926	0.943	0.039	0.925	0.954	0.358	0.963
	150	1.169	1.316	0.888	1.057	0.026	0.903	1.080	0.432	0.963	0.844	0.933	0.922	0.952	0.017	0.922	0.974	0.305	0.959
	170	1.136	0.828	0.883	1.046	0.019	0.896	1.076	0.392	0.963	0.872	0.334	0.910	0.961	0.010	0.913	0.989	0.282	0.962
	190	1.121	0.684	0.901	1.041	0.016	0.911	1.064	0.329	0.963	0.900	0.294	0.925	0.968	0.009	0.927	0.991	0.250	0.960
	210	1.098	0.653	0.892	1.034	0.014	0.905	1.063	0.295	0.959	0.917	0.415	0.913	0.974	0.009	0.918	1.000	0.227	0.961
	230	1.089	0.517	0.897	1.031	0.012	0.909	1.057	0.265	0.957	0.933	0.261	0.916	0.979	0.008	0.922	1.000	0.211	0.955
	250	1.079	0.435	0.896	1.028	0.011	0.909	1.052	0.235	0.961	0.944	0.249	0.916	0.982	0.007	0.919	1.002	0.193	0.960
	270	1.070	0.393	0.898	1.024	0.010	0.912	1.051	0.221	0.954	0.950	0.233	0.910	0.983	0.007	0.916	1.006	0.185	0.957
	290	1.073	0.362	0.909	1.026	0.009	0.917	1.040	0.190	0.954	0.963	0.221	0.921	0.989	0.006	0.922	1.000	0.163	0.956
	310	1.052	0.295	0.908	1.018	0.008	0.915	1.043	0.172	0.962	0.959	0.197	0.918	0.986	0.006	0.922	1.004	0.146	0.962
	330	1.055	0.282	0.912	1.019	0.007	0.920	1.036	0.159	0.959	0.968	0.192	0.920	0.989	0.005	0.923	1.002	0.140	0.960
	350	1.059	0.269	0.923	1.022	0.007	0.931	1.030	0.144	0.961	0.978	0.187	0.929	0.994	0.005	0.934	0.998	0.127	0.959

Table 5: MRE, MSE and CP for the MLEs and BCMLEs of $b = 1.3$, $c = 0.5$ and $\beta = 2$, based on $M = 5,000$ generated samples of size n with 50% of censoring.

		MLE									**BCMLE**								
		b			c			β			b			c			β		
Cens.p	n	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**	**MRE**	**MSE**	**CP**
	50	1.786	23.819	0.796	1.201	0.304	0.823	1.545	9.080	0.899	0.598	47.884	0.952	0.890	0.518	0.935	0.361	11.757	0.974
	70	1.534	13.181	0.809	1.139	0.181	0.832	1.372	4.519	0.928	0.582	24.598	0.939	0.886	0.288	0.921	0.612	4.866	0.985
	90	1.478	10.068	0.826	1.125	0.126	0.850	1.253	2.295	0.944	0.645	16.283	0.940	0.903	0.177	0.925	0.742	2.088	0.984
	110	1.334	6.197	0.832	1.087	0.084	0.856	1.208	1.472	0.961	0.617	9.147	0.925	0.896	0.105	0.920	0.828	1.129	0.982
	130	1.308	4.806	0.840	1.085	0.071	0.862	1.174	1.180	0.958	0.683	5.883	0.922	0.919	0.078	0.917	0.894	0.842	0.974
	150	1.280	3.639	0.852	1.081	0.057	0.869	1.137	0.894	0.963	0.720	3.921	0.915	0.933	0.056	0.909	0.925	0.606	0.971
	170	1.209	2.331	0.853	1.060	0.038	0.871	1.126	0.728	0.965	0.730	1.962	0.910	0.931	0.029	0.911	0.957	0.477	0.966
	190	1.196	1.734	0.865	1.057	0.032	0.879	1.108	0.576	0.964	0.778	0.976	0.911	0.944	0.020	0.908	0.970	0.376	0.967
50%	210	1.160	1.329	0.865	1.046	0.025	0.879	1.097	0.496	0.962	0.802	0.707	0.908	0.948	0.015	0.908	0.982	0.366	0.962
	230	1.151	1.165	0.872	1.044	0.023	0.886	1.090	0.447	0.964	0.831	0.522	0.908	0.956	0.013	0.907	0.990	0.308	0.965
	250	1.153	1.109	0.881	1.047	0.022	0.896	1.073	0.373	0.964	0.868	0.558	0.913	0.967	0.014	0.912	0.985	0.279	0.964
	270	1.147	1.099	0.880	1.044	0.021	0.893	1.071	0.346	0.957	0.893	0.650	0.910	0.973	0.014	0.912	0.992	0.266	0.958
	290	1.124	0.826	0.888	1.037	0.017	0.899	1.061	0.292	0.963	0.900	0.381	0.911	0.974	0.010	0.914	0.992	0.229	0.962
	310	1.110	0.863	0.881	1.032	0.017	0.897	1.069	0.309	0.960	0.909	0.648	0.901	0.975	0.013	0.909	1.005	0.244	0.959
	330	1.102	0.637	0.889	1.029	0.014	0.902	1.060	0.271	0.962	0.920	0.306	0.907	0.978	0.009	0.911	1.002	0.218	0.963
	350	1.108	0.547	0.888	1.033	0.012	0.902	1.051	0.243	0.960	0.940	0.269	0.907	0.985	0.008	0.911	0.996	0.201	0.958