



Estimation of parameters and quantiles of the Weibull distribution

Alicja Jokił-Rokita¹ · Sylwester Piątek¹

Received: 10 March 2022 / Revised: 21 September 2022 / Accepted: 24 November 2022 /
Published online: 16 December 2022
© The Author(s) 2022

Abstract

We propose three new estimators of the Weibull distribution parameters which lead to three new plug-in estimators of quantiles. One of them is a modification of the maximum likelihood estimator and two of them are based on nonparametric estimators of the Gini coefficient. We also make some review of estimators of the Weibull distribution parameters and quantiles. We compare the small sample performance (in terms of bias and mean squared error) of the known and new estimators and extreme quantiles. Based on simulations, we obtain, among others, that the proposed modification of the maximum likelihood estimator of the shape parameter has a smaller bias and mean squared error than the maximum likelihood estimator, and is better or as good as known estimators when the sample size is not very small. Moreover, one of the proposed estimator, based on the nonparametric estimator of the Gini coefficient, leads to good extreme quantiles estimates (better than the maximum likelihood estimator) in the case of small sample sizes.

Keywords Extreme quantiles · Gini coefficient · Monte Carlo simulation · Point estimation

Mathematics Subject Classification 62F10 · 62P05

The second author contributed equally to this article is not required.

✉ Alicja Jokił-Rokita
alicia.jokił-rokita@pwr.edu.pl
Sylwester Piątek
sylwester.piatek@pwr.edu.pl

¹ Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, Wrocław 50-370, Poland

1 Introduction

The Weibull distribution appears in many fields of science such as reliability (Almeida 1999; Fok et al. 2001; Queshi and Sheikh 1997), survival analysis (Carroll 2003), hydrology (Heo et al. 2001), social sciences (Roed and Zhang 2002), wind energy industry (Kang et al. 2018), actuarial science (Bolancé and Guillen 2021), financial mathematics (Chen and Gerlach 2013; Gebizlioglu et al. 2011). In the last one, it may be used to describe the distribution of the return rates of certain investments.

This paper is concerned with estimation of the parameters and quantiles of the Weibull distribution on the basis of complete data. Let X_1, X_2, \dots, X_n be independent identically distributed (i.i.d.) random variables from the two-parameter Weibull distribution with the probability density function defined by

$$f(x) = \frac{\beta}{\sigma} \left(\frac{x}{\sigma}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\sigma}\right)^\beta\right], \quad x \geq 0, \beta > 0, \sigma > 0.$$

We propose three estimators of the shape parameter β of the Weibull distribution defined above, when the scale parameter σ is unknown. One of them is based on a modification of the equation from which the maximum likelihood estimator of the shape parameter is determined. The other two estimators are obtained on the basis of the non-parametric estimators of the Gini coefficient. The estimators proposed lead to new estimators of the scale parameter and new plug-in estimators of quantiles. The main purpose of this paper is to compare the performance of these estimators of the Weibull distribution parameters and plug-in estimators of Weibull distribution quantiles with those recommended in the literature. We use the bias and the mean squared error as criteria for comparison.

The paper is organized as follows. In Sect. 2 we present a review of known estimators of the Weibull distribution parameters and papers in which their accuracy were compared. In Sect. 3 we propose new estimators of the shape parameters of the Weibull distribution which lead to new estimators of the scale parameter and plug-in estimators of quantiles. Some comments on plug-in estimators of quantiles of the Weibull distribution are provided in Sect. 4. The estimators of the parameters and quantiles of the Weibull distribution are compared in the simulation study, and results of this study are presented in Sect. 5. Section 6 includes a real data analysis. Some concluding remarks are presented in Sect. 7.

2 A review of estimators in two-parameter Weibull distribution model

In the literature, over a dozen methods have been applied to estimation of the Weibull distribution parameters. These methods include: the method of moments (MM) (Cran 1988), the modified MM (Cohen and Whitten 1982), the method of probability weighted moments (Greenwood et al. 1979; Singh et al. 1990), the maximum likelihood estimation (ML) (Cohen 1965; Balakrishnan and Kateri 2008), the weighted maximum likelihood estimation (Jacquelin 1993), a few modifications of the ML estimation (Cohen and Whitten 1982), specially the Tiku's modified ML (TMML)

estimator (Gebizlioglu et al. 2011), method based on L -moments (LM) (Teimouri et al. 2013), the logarithmic moments (Johnson et al. 1994), the method of entropy (Singh 1987), the least squares (LS) (Bain and Antle 1967; Swain et al. 1988), the weighted least squares (WLS) (Hung 2001; Lu et al. 2004; Swain et al. 1988; Van Zyl and Schall 2012; Zhang et al. 2008; Kantar 2015), the generalized least squares (Engeman and Keefe 1982; Kantar 2015), the percentiles method (Hassanein 1971; Seki and Yokoyama 1993), a method based on U -statistics (U) (Sadani et al. 2019), the median rank regression method (Genschel and Meeker 2010; George 2014), the method based on second-kind statistics (log-cumulant estimator) (Sun and Han 2010), the generalized spacing method (GS) (Ghosh and Jammalamadaka 2001; Gebizlioglu et al. 2011). In parentheses, we gave abbreviations of estimators, which we will consider in the simulation study.

Closed form expressions of estimators are only given for L -moments estimators, estimators of logarithmic moment method, percentile estimators, Tiku's modified ML estimators, and the estimators based on U -statistics.

Comparisons of various estimators of the Weibull distribution parameters have been performed in the following papers (in chronological order): Gross and Lurie (1977), Gibbons and Vance (1981), Al-Baidhani and Sinclair (1987), Singh et al. (1990), Wu et al. (2006), Genschel and Meeker (2010), Gebizlioglu et al. (2011), Teimouri et al. (2013), Akram and Hayat (2014), George (2014), Pobocikova and Sedliackova (2014), Kantar (2015), and Sadani et al. (2019). However, the conclusions of these comparisons differ. For example, Al-Baidhani and Sinclair (1987) compared the generalised least squares, maximum likelihood, two estimators proposed in Bain and Antle (1967), and two mixed methods of estimating the parameters of the two-parameter Weibull distribution. The comparison was made using the observed relative efficiency of parameter estimates to summarize the results of 1000 simulated samples of sizes 10 and 25. The results were that: generalised least squares is the best method of estimating the shape parameter, the best method of estimating the scale parameter depends on the size of the shape parameter. Gebizlioglu et al. (2011) compared nine estimators and recommended ML estimator and TMML as the best for estimating both parameters. Teimouri et al. (2013) proposed LM and compared it with four other methods. They concluded that LM and ML estimators had the best performance. Pobocikova and Sedliackova (2014) generally recommended ML estimators, but for small samples ($n = 10$ or $n = 20$) suggested using WLS estimators. Wu et al. (2006) compared ML estimators with methods based on linear regression, considering different approximations to obtain them. They noticed that ML and MM estimators have tendency to overestimate rather than underestimate the parameters. Furthermore they showed that certain approximations are better than others in methods based on linear regression.

To the best of our knowledge, comparisons of various estimators of the Weibull distribution quantiles of order p have been performed only in the following papers: (Gibbons and Vance 1981) ($n = 10, n = 25$ and $p = 0.1, p = 0.9$), Al-Baidhani and Sinclair (1987) ($n = 10, n = 25$), Singh et al. (1990) ($n \in \{10, 50, 100, 1000\}$, $p \in \{0.9, 0.99, 0.999\}$). Gibbons and Vance (1981) concluded that for estimating quantile of order .9, the best are the ML estimator and the best linear invariant estimator, followed by the best linear unbiased estimator. The results of their simulations suggest

that estimator proposed by Bain and Antle (1967) and Gumbel (1958) are the best for estimating the quantile of order. 1. Al-Baidhani and Sinclair (1987) showed that ML estimator outperforms the least squares estimator and the generalised least squares estimator of quantiles of order 0.95 and 0.99.

3 New estimators of the shape parameter

3.1 Modified maximum likelihood estimators

Let x_1, \dots, x_n be the realizations of i.i.d. random variables X_1, \dots, X_n from the Weibull distribution with the scale parameter σ and the shape parameter β , and denote by L the likelihood function. The likelihood equations are

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n\beta}{\sigma} + \frac{\beta}{\sigma^{\beta+1}} \sum_{i=1}^n x_i^\beta = 0, \quad (1)$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} - n \ln \sigma + \sum_{i=1}^n \ln x_i + \frac{\ln \sigma}{\sigma^\beta} \sum_{i=1}^n x_i^\beta - \frac{1}{\sigma^\beta} \sum_{i=1}^n x_i^\beta \ln x_i = 0. \quad (2)$$

From Eq. (1) we have a closed-form expression for the ML estimator of σ , namely

$$\hat{\sigma}_{ML} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}_{ML}} \right)^{1/\hat{\beta}_{ML}}, \quad (3)$$

where the ML estimator $\hat{\beta}_{ML}$ of the parameter β is a solution to the equation

$$\sum_{i=1}^n x_i^\beta + \frac{1}{n} \left[\sum_{i=1}^n \ln(x_i^\beta) \right] \sum_{i=1}^n x_i^\beta - \sum_{i=1}^n x_i^\beta \ln(x_i^\beta) = 0. \quad (4)$$

Lemat 1 *The equation (4) is not an unbiased estimating equation, i.e.,*

$$E_{\sigma, \beta} \left\{ \sum_{i=1}^n X_i^\beta + \frac{1}{n} \left[\sum_{i=1}^n \ln(X_i^\beta) \right] \sum_{i=1}^n X_i^\beta - \sum_{i=1}^n X_i^\beta \ln(X_i^\beta) \right\} \neq 0. \quad (5)$$

Proof Let us denote $Y_i = X_i^\beta$. Then the left side of (5) is equal to

$$E_{\sigma, \beta} \left[\sum_{i=1}^n Y_i + \frac{1}{n} \left(\sum_{i=1}^n \ln Y_i \right) \sum_{i=1}^n Y_i - \sum_{i=1}^n Y_i \ln Y_i \right]. \quad (6)$$

Using the fact that if X has the Weibull distribution with the shape parameter β and the scale parameter σ then X^β has the exponential distribution with the scale parameter σ^β ,

we have that

$$E_{\sigma,\beta}(Y_i) = \sigma^\beta, \tag{7}$$

$$E_{\sigma,\beta}(\ln Y_i) = \psi(1) - \ln(\sigma^\beta) \tag{8}$$

and

$$E_{\sigma,\beta}(Y_i \ln Y_i) = \sigma^\beta [\psi(1) + 1 - \ln(\sigma^\beta)], \tag{9}$$

$i = 1, \dots, n$, where ψ is the digamma function. Thus, expression (6) is equal to

$$\sum_{i=1}^n E_{\sigma,\beta}(Y_i) + \frac{1}{n} \sum_{i=1}^n E_{\sigma,\beta}(Y_i \ln Y_i) + \frac{1}{n} \sum_{i \neq j} E_{\sigma,\beta}(Y_i \ln Y_j) - \sum_{i=1}^n E_{\sigma,\beta}(Y_i \ln Y_i) = \sigma^\beta.$$

□

From the proof of Lemma 1, it can be easily seen that the following equation

$$\frac{n-1}{n} \sum_{i=1}^n x_i^\beta + \frac{1}{n} \left[\sum_{i=1}^n \ln(x_i^\beta) \right] \sum_{i=1}^n x_i^\beta - \sum_{i=1}^n x_i^\beta \ln(x_i^\beta) = 0 \tag{10}$$

is an unbiased estimating equation for the parameter β .

Theorem 1 *The solution to equation (10) exists and is unique.*

Proof The existence and uniqueness of the solution to equation (10) can be proved analogously to the existence and uniqueness of the ML estimator given in Balakrishnan and Kateri (2008). □

We take the unique solution to equation (10) as the modified maximum likelihood (MML) estimator $\hat{\beta}_{MML}$ of the shape parameter β . As the MML estimate of the parameter σ we take

$$\hat{\sigma}_{MML} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}_{MML}} \right)^{1/\hat{\beta}_{MML}}. \tag{11}$$

Theorem 2 *With probability 1, $\hat{\beta}_{MML} < \hat{\beta}_{ML}$, and $\hat{\sigma}_{MML} < \hat{\sigma}_{ML}$.*

Proof Equation (4) can be expressed alternatively as

$$1 = \frac{\sum_{i=1}^n x_i^\beta \ln(x_i^\beta)}{\sum_{i=1}^n x_i^\beta} - \frac{1}{n} \sum_{i=1}^n \ln(x_i^\beta) =: H_{\mathbf{x}}(\beta), \tag{12}$$

and equation (10) as

$$\frac{n-1}{n} = H_{\mathbf{x}}(\beta). \tag{13}$$

It can be shown that for each $\mathbf{x} = (x_1, \dots, x_n)$, the function $H_{\mathbf{x}}(\beta)$ of the argument β increases from 0 to infinity. Hence $\hat{\beta}_{MML} < \hat{\beta}_{ML}$ with probability 1. The second part of the theorem follows from the form of $\hat{\sigma}_{ML}$, $\hat{\sigma}_{MML}$, the relation between $\hat{\beta}_{ML}$ and $\hat{\beta}_{MML}$, and the power mean inequality. \square

Remark 1 It can be easily shown that $\hat{\beta}_{MML}(X_1, \dots, X_n) = \hat{\beta}_{MML}(cX_1, \dots, cX_n)$, and $\hat{\sigma}_{MML}(cX_1, \dots, cX_n) = c\hat{\sigma}_{MML}(X_1, \dots, X_n)$, for an arbitrary $c > 0$.

Remark 2 In Cohen and Whitten (1982) five modifications of the ML estimator of the three-dimensional Weibull distribution parameter are considered. They were obtained by replacing one of the likelihood equations (the one derived by differentiating the likelihood function with respect to the location parameter) with alternate functional relationships. Two of these modifications, under the assumption that the location parameter equals zero, are considered in Gebizlioglu et al. (2011) in the case of the two-parameter Weibull distribution. In Gebizlioglu et al. (2011) the TMML estimator is also considered which is based on linearizing the likelihood equations around the first two terms of the Taylor series. The proposed MML estimator is constructed by a slight modification of the equation used to determine the ML estimate of the shape parameter and differs from the modifications of the ML estimator considered in the literature.

3.2 Estimators based on the Gini index

Both equations (4) and (10) can only be solved numerically. Motivated by the need to determine an appropriate starting point for finding solutions to these equations, we look for a simple estimate of the shape parameter β . We took into account the fact that the Lorenz curve (see for example Sarabia 2008), and consequently the Gini index G , does not depend on the scale parameter (here σ). For the Weibull distribution $G = 1 - 2^{-1/\beta}$, so

$$\beta = [-\log_2(1 - G)]^{-1}. \quad (14)$$

One of the nonparametric estimators of G is

$$\hat{G} = \frac{1}{n^2 \bar{X}} \sum_{i=1}^n (2i - n - 1) X_{i:n},$$

where $X_{i:n}$ is the i -th order statistic of the sample (X_1, \dots, X_n) . The estimator \hat{G} is biased and in Davidson (2007) an approximated expression for bias of \hat{G} is given. Based on the result in Davidson (2007), in Mirzaei et al. (2017) the following bias-corrected estimator of G is proposed

$$\tilde{G} = \frac{n}{n-1} \hat{G}.$$

The estimator \tilde{G} is not unbiased, but its bias is of order $\frac{1}{n}$. We consider the following two simple estimators of the parameter β

$$\hat{\beta}_{\hat{G}} = [-\log_2(1 - \hat{G})]^{-1} \quad \text{and} \quad \hat{\beta}_{\tilde{G}} = [-\log_2(1 - \tilde{G})]^{-1}.$$

As the estimates of the parameter σ we take

$$\hat{\sigma}_{\hat{G}} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}_{\hat{G}}} \right)^{1/\hat{\beta}_{\hat{G}}} \quad \text{and} \quad \hat{\sigma}_{\tilde{G}} = \left(\frac{1}{n} \sum_{i=1}^n x_i^{\hat{\beta}_{\tilde{G}}} \right)^{1/\hat{\beta}_{\tilde{G}}}.$$

It can be easily shown that $\hat{\beta}_{\hat{G}} > \hat{\beta}_{\tilde{G}}$ with probability 1, and based on the power mean inequality, it implies that $\hat{\sigma}_{\hat{G}} > \hat{\sigma}_{\tilde{G}}$ with probability 1.

Fact 1 The estimator $\hat{\beta}_{\tilde{G}}$ has the same form as the L-moment estimator $\hat{\beta}_{LM}$ proposed by Teimouri et al. (2013).

Proof The estimator $\hat{\beta}_{LM}$ of the parameter β , proposed by Teimouri et al. (2013), is of the form

$$\hat{\beta}_{LM} = -\frac{\log 2}{\log(1 - m_2^0/m_1^0)},$$

where

$$m_1^0 = \bar{X} \quad \text{and} \quad m_2^0 = \frac{2}{n(n-1)} \sum_{i=1}^n (i-1)X_{i:n} - \bar{X}. \tag{15}$$

Hence,

$$\begin{aligned} \hat{\beta}_{LM} &= -\frac{1}{\log_2(1 - m_2^0/m_1^0)} = \left[-\log_2 \left(1 - \frac{1}{n(n-1)\bar{X}} \sum_{i=1}^n (2i-2)X_{i:n} - \bar{X} \right) \right]^{-1} \\ &= \left[-\log_2 \left(1 - \frac{1}{n(n-1)\bar{X}} \sum_{i=1}^n (2i-2)X_{i:n} - \frac{1}{n(n-1)\bar{X}} \sum_{i=1}^n (n-1)X_{i:n} \right) \right]^{-1} \\ &= \left[-\log_2 \left(1 - \frac{1}{n(n-1)\bar{X}} \sum_{i=1}^n (2i-n-1)X_{i:n} \right) \right]^{-1} = [-\log_2(1 - \tilde{G})]^{-1} = \hat{\beta}_{\tilde{G}}. \end{aligned}$$

□

Remark 3 Although $\hat{\beta}_{LM} = \hat{\beta}_{\tilde{G}}$, the estimator of the parameter σ proposed by Teimouri et al. (2013) is not of the same form as $\hat{\sigma}_{\tilde{G}}$.

4 Estimators of quantiles of the Weibull distribution

Another issue we investigated is the estimation of the extreme quantiles of the Weibull distribution. We decided to investigate the extreme quantiles, as they are widely used for durability tests (Gibbons and Vance 1981), in hydrology (Smith 1987), and in financial mathematics, specifically for estimating risk measures (Gebizlioglu et al. 2011; Alemany et al. 2013).

The quantile of order p of the Weibull distribution is given by

$$Q(p) = \sigma[-\ln(1 - p)]^{1/\beta},$$

where $p \in (0, 1)$. The plug-in estimators \hat{Q} of the quantiles can be obtained substituting the unknown parameters by their estimators. For example, the MML estimator of the quantile of order p is

$$\hat{Q}_{MML}(p) = \hat{\sigma}_{MML}[-\ln(1 - p)]^{1/\hat{\beta}_{MML}} = \left[-\frac{1}{n} \ln(1 - p) \sum_{i=1}^n x_i^{\hat{\beta}_{MML}} \right]^{1/\hat{\beta}_{MML}}.$$

In the simulation study, presented in Sect. 5, we compare various plug-in estimators of quantiles obtained by using estimators of parameters which are recommended in the literature and those proposed in Sect. 3.

5 Simulation study

To investigate the performance of the new proposed estimators of the Weibull distribution parameters and quantiles relative to the estimators recommended in the literature we conducted a simulation study. The measures that we use to compare the estimators are the mean squared error (MSE) and bias. We performed the simulations using the R programming language.

5.1 Simulation design

For 26 values of shape parameter varying between 0.5 and 3 and 20 sample sizes varying between 10 and 200, we simulated 2000 samples from Weibull distribution. Without loss of generality we set scale parameter equal to 1. For each sample we calculated the values of the following parameter estimators recommended in the literature: MM, ML, WLS, LM, TMML, GS, LS, U and the new estimators proposed: MML, G1, G2. Then we used the estimates of the parameters to compute the plug-in estimates of quantiles of order .9, .95, .99 and .999.

Most of the estimators considered require numeric computation. To execute them we used functions *uniroot* from library *rootSolve* and *optim* from library *stats* in R package. Some of the estimators considered, namely MM, ML, GS, LM and U estimators, are implemented in the *ForestFit* (Teimouri et al. 2020) library in R. Determining the values of some estimators considered in the literature may be problematic due to

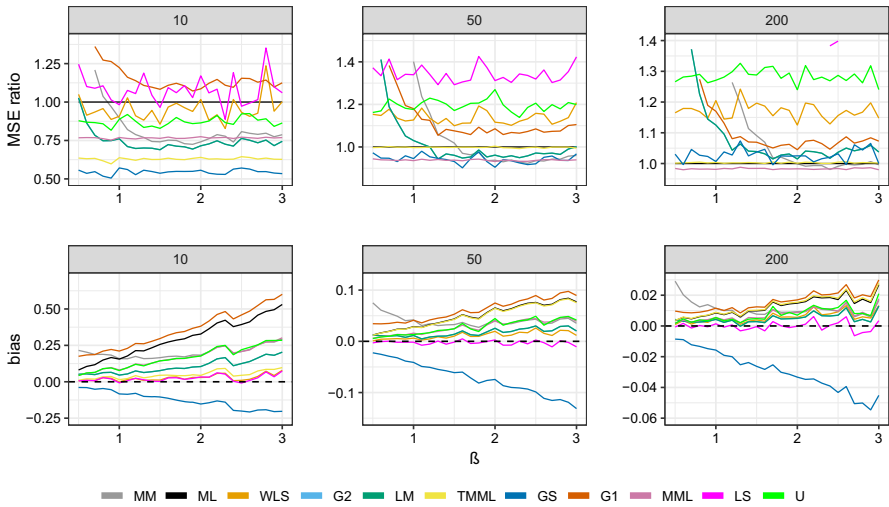


Fig. 1 Empirical MSE's ratios of selected estimators of β to ML estimator and empirical biases of the estimators of β for different values of β and $\sigma = 1$, for $n = 10$ (on the left), $n = 50$ (in the middle), $n = 200$ (on the right)

the lack of roots or having many roots, resulting in nonconvergence or convergence to incorrect values. This issue was discussed in Kantar and Şenoğlu (2008) (in the case of estimating the location and scale parameters with known shape parameter of the Weibull distribution) and Gebizlioglu et al. (2011). For example the equation used to determine the MML2 estimator of the shape parameter (proposed in Cohen and Whitten (1982), considered in Gebizlioglu et al. (2011) but not considered in this paper) often has no solution when the sample size and shape parameter are small. The estimators given by explicit formulas (i.e. G1, G2, TMML) and those whose existence and uniqueness has been theoretically proven (e.g. ML and MML) do not show such problems. During the simulation study, no problems with computing the values of other compared estimators (MM, WLS, LM, GS, LS, U) have been encountered.

5.2 Comparison of the estimators of the parameters

In Fig. 1 we present the empirical MSE's ratios of selected estimators of β with respect to the ML estimator and empirical biases of the estimators of β for different values of β and $\sigma = 1$. Lines denoted by G1 and G2 correspond to estimators based on \hat{G} and \tilde{G} , respectively. The values on the x axis has been customised to present explicitly the results of several best estimators. Consequently, when a certain estimator is significantly worse than others only a fragment of its plot is included in the chart (e.g. the plot corresponding to the LS estimator in the top right chart in Fig. 1) or it is not included at all (e.g. the plot corresponding to the LS estimator in the top right chart in Fig. 2).

Figure 2 contains a comparison of the scale parameter estimators analogous to the one presented in Fig. 1.

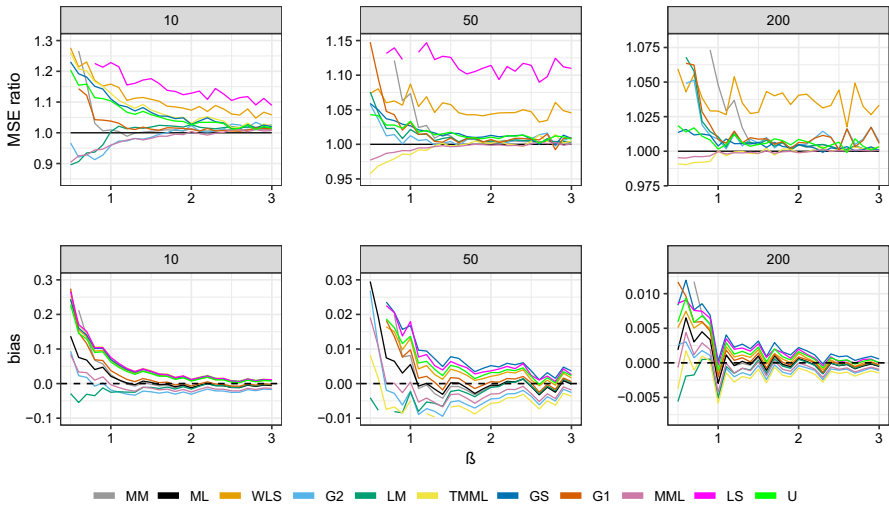


Fig. 2 Empirical MSE's ratios of selected estimators of σ to ML estimator and empirical biases of the estimators of σ for different values of β and $\sigma = 1$, for $n = 10$ (on the left), $n = 50$ (in the middle), $n = 200$ (on the right)

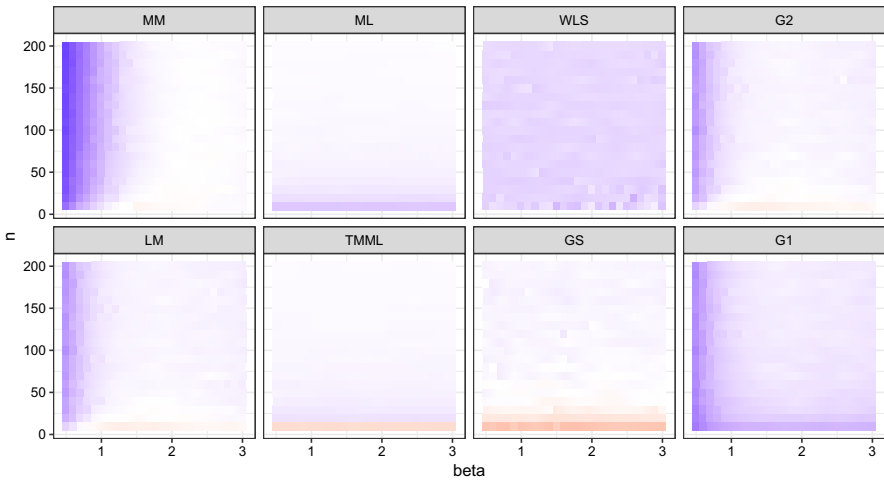


Fig. 3 Comparison of empirical MSE's of MML estimator and selected estimators of the shape parameter when $\sigma = 1$

Figures 3 and 4 include heatmap plots related to the empirical MSE's of the shape and scale parameter estimators, respectively. The closer the color is to blue, the better the MML estimator is compared to the selected estimator. The closer the color is to red, the worse the MML estimator is compared to the selected estimator. If the color is white, the compared estimators have comparable empirical MSE's.

Our simulations concerning the parameter β show that:

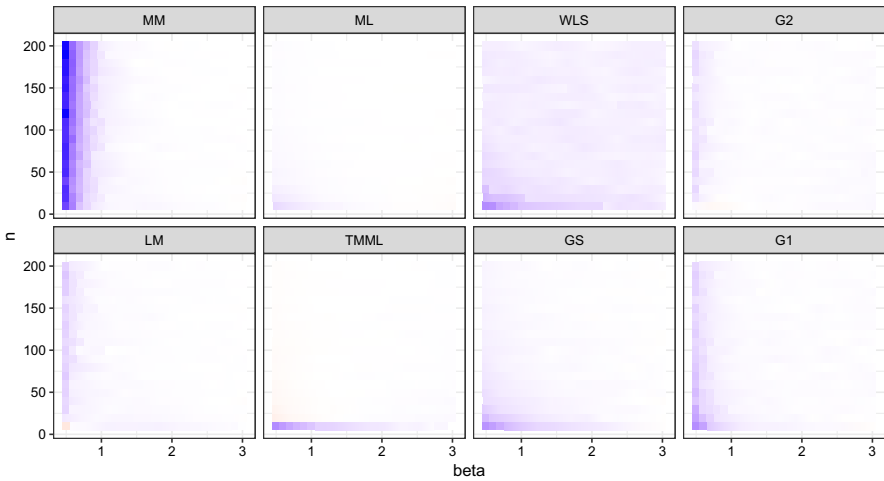


Fig. 4 Comparison of empirical MSE's of MML estimator and selected estimators of the scale parameter when $\sigma = 1$

- the MML estimator has the smallest empirical MSE's for samples of sizes $n \geq 50$ (see Fig. 3);
- the MML estimator has the smaller empirical MSE's than ML estimator for all cases considered; for $n = 10$ the ratio of these errors is approximately 3/4 and increases to 1 as n increases;
- for all considered cases, the empirical bias of the MML estimator is approximately half that of the ML estimator;
- the G2, LM, TMML, GS estimators have smaller empirical MSE's than MML estimator when $n = 10$ (see Fig. 3);
- the LS estimator has the smallest empirical bias, close to zero (see Fig. 1);
- most of the estimators tend to overestimate the shape parameter, while GS estimator underestimates it (see Fig. 1).

5.3 Comparison of the estimators of extreme quantiles

Figure 5 contains a comparison of MSE's and biases of selected estimators of quantile $q = Q(p)$ of order $p = .99$. The values displayed in the plots are the ratios of MSE's (on top) (or biases — bottom) of quantile estimator considered and ML estimator of q . This means, that if the curve corresponding to the estimator, say \hat{q} , is below 1, the estimator \hat{q} is better (in terms of MSE) than ML estimator of q .

The results are similar for all four investigated orders (.9, .95, .99 and .999) of quantiles, hence we only present results for one order $p = .99$ of quantile. The conclusions are as follows:

- when $n \geq 50$ and $\beta \geq 1.5$ ML, MML and TMML estimators have almost identical values of MSE's (see Figs. 5, 6);

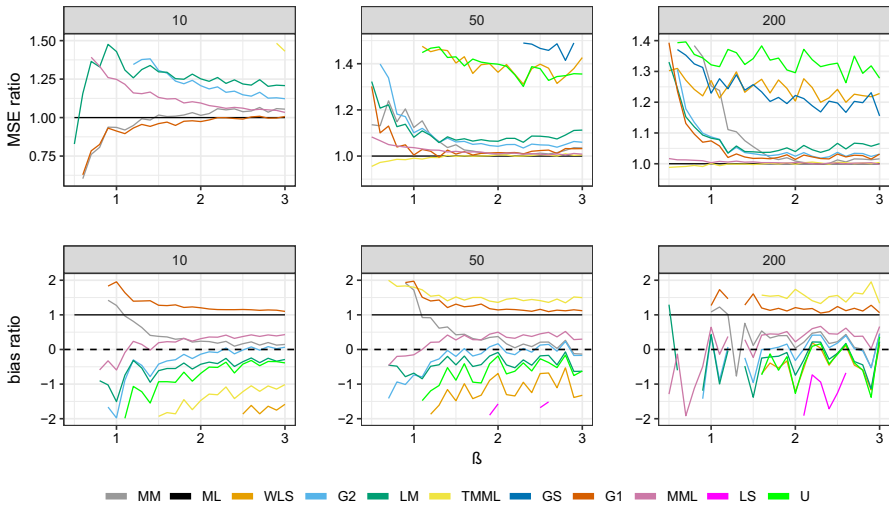


Fig. 5 Empirical MSE's ratios and empirical biases ratios of quantiles of order .99 estimators to ML estimator for different values of β and $\sigma = 1$, for $n = 10$ (on the left), $n = 50$ (in the middle), $n = 200$ (on the right)

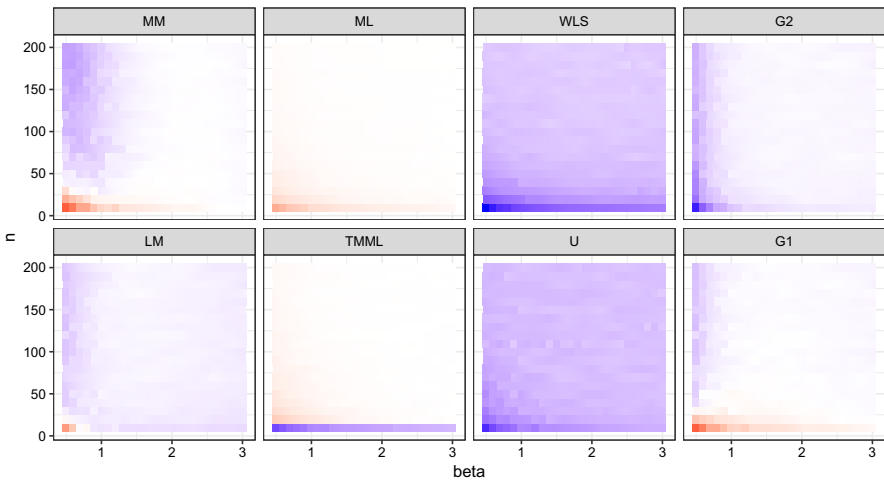


Fig. 6 Comparison of empirical MSE's of the MML estimator and selected estimators of the quantile of order .99 when $\sigma = 1$

- when $\beta > 0.8$ MML estimator has much smaller, and TMML estimator has much larger bias than ML estimator (see Fig. 5);
- for small sample sizes ($n = 10$ and $n = 20$) the G1 estimator has the smallest MSE (see Figs. 5, 6);
- the LM, G2, U estimators and for larger sample sizes the WLSE have much smaller bias than ML, MML and TMML estimators, although they have significantly larger MSE's (see Fig. 5);

Table 1 The average CPU times (in milliseconds) per sample for selected sample sizes and estimators when $\beta = 2$

	10	20	50	100	200
MM	0.258	0.252	0.250	0.223	0.260
ML	0.477	0.772	1.374	2.596	4.686
LS	3.670	4.141	4.735	5.665	7.901
WLS	3.841	4.105	4.953	5.899	8.386
G1	0.175	0.159	0.189	0.228	0.253
G2	0.157	0.153	0.174	0.214	0.281
GS	2.384	2.914	4.246	6.142	10.499
MML	0.398	0.578	1.093	2.001	3.828
LM	0.166	0.144	0.147	0.174	0.203
TMML	0.179	0.151	0.184	0.208	0.312
U	1.346	2.158	6.447	20.172	76.534

- for all cases considered, the MML estimator has slightly larger MSE's than ML estimator;
- for $\beta > 0.75$ the MML estimator has much smaller bias (50%) than ML estimator;
- the G2 estimator has the smallest bias when $\beta > 1.5$.

5.4 Comparison of the CPU times of the estimators

We consider the central processing unit (CPU) times to compare the computational complexity of the considered estimators. Some estimators, namely G1, G2, LM, and TMML, are given by explicit formulas, while others require iterative methods to find the root of the corresponding equation. For most estimators (except MM, G1, G2, LM, and TMML), the time required increases significantly with sample size. Table 1 shows the comparison of average CPU times (in milliseconds) when shape parameter is equal to 2. Similar results have been achieved with several different values of this parameter. CPU times of the estimators given by explicit formulas (i.e. G1, G2, TMML) and the MM estimator are much shorter than those of other estimators, especially when the sample size is large. The shape and scale estimators based on U-statistic, which outperform several other estimators in the sense of bias, when the sample size is large is the slowest one in the experiment and it might be extremely slow (see (Sadani et al. 2019)). In case when two estimators have similar efficiency, the one which requires less time to be computed may be chosen.

6 Real data analysis

The data, given in Table 2, has been taken from Murthy et al. (2004) (data set 4.1). It consists of 20 observations of the time till failure. The data is complete, which means that it has not been censored.

We applied the Anderson-Darling test (implemented in *gofest* package) to verify the null hypothesis that the observed times to failure are realizations of the random

Table 2 Real data consisting of 20 observations of the time till failure

11.24	1.92	12.74	22.48	9.60	11.50	8.86	7.75	5.73	9.37
30.42	9.17	10.20	5.52	5.85	38.14	2.99	16.58	18.92	13.36

Table 3 Estimates of the parameters and extreme quantiles

	$\hat{\beta}$	$\hat{\sigma}$	q = .9	q = .99	q = .999
MM	1.42	13.87	24.99	40.77	54.28
ML	1.55	14.12	24.20	37.86	49.19
LS	1.73	12.98	21.02	31.37	39.64
WLS	1.61	13.40	22.48	34.56	44.44
G1	1.58	14.20	24.09	37.38	48.33
G2	1.48	13.93	24.48	39.11	51.45
MML	1.51	14.00	24.36	38.60	50.53
LM	1.48	13.95	24.52	39.18	51.54
TMML	1.57	13.83	23.55	36.67	47.50
U	1.72	14.07	22.87	34.25	43.37

variables from the Weibull distribution. The obtained p value in this test is equal to 0.8257, so we do not reject it on the significance level 0.05.

Table 3 includes the estimates of the shape and scale parameters and quantiles.

The estimated values of the shape parameter β are between 1.42 and 1.73. According to the results of our simulations, the MM and G1 estimators of quantiles are recommended for such combination of the sample size $n = 20$ and the shape parameter β . The MM estimator though gives significantly greater values of quantiles, so we recommend using the G1 estimator of quantiles.

7 Concluding remarks

We proposed three Weibull distribution shape parameter estimators, which lead to three scale parameter estimators and three new quantile estimators of this distribution. One of the proposed estimators (the MML) is a modification of the ML estimator. Based on the simulations performed, we can conclude that it has both a smaller mean squared error and a smaller bias in relation to the ML estimator. However, the plug-in estimators of the Weibull distribution quantiles based on the proposed modification of the ML estimator of the shape parameter turned out to be worse (due to the mean squared error) compared to the plug-in estimators based on ML estimators of the parameters. We also proposed two estimators of the shape parameter which have closed form expressions. The main aim of this proposal was to obtain a good starting point for the numerical computation of the ML and MML estimators. They are based on the nonparametric estimators of the Gini coefficient, thus we did not expect that they would have smaller errors than known estimators, in particular the ML estimator. It turned out, however,

that one of the proposed estimators leads to good extreme quantile estimates (better than the ML) in the case of small sample sizes.

In general, the plug-in estimator of extreme quantiles based on the ML estimators of the parameters is the best or as good as other estimators (due to the mean squared error), except the case of small samples (up to 30), where we recommend G1 and MM estimators. For parameters estimation we recommend using MML estimator.

The approach presented in this paper can also be applied to censored data. For example, in case of Type-I censoring, Type-II censoring or progressive Type-II censoring, the ML estimator of the scale parameter σ can be expressed as a function of the ML estimator of the parameter β (see e.g. Cohen (1965), Balakrishnan and Kateri (2008)), which is the solution to an equation (depending on the censoring type) which can only be solved numerically. However, analogously to the case of complete data, relationship (14) between this parameter and the Gini index can be used. As the estimator of the Gini index we can take one of the non-parametric estimators of this index with right censored data considered, for example, in Gigliarano and Muliere (2013), Lv et al. (2017), Hong et al. (2018), and recently (Kattumannil et al. 2021). To the best of our knowledge, this approach has not been used to estimate the Weibull distribution parameters based on censored data. Applying the approach analogous to that presented in Sect. 3.1, especially in the case of Type-I censored data, seems to be a more difficult task. This is because it results in having a sum of a random number of functions of order statistics in an equation for the ML estimator of the parameter β . Comparison of estimators of the parameters and quantiles of the Weibull distribution based on censored data, obtained using the methods sketched above with the estimators known from the literature, e.g. Gibbons and Vance (1981), Zhang et al. (2008), Balakrishnan and Kateri (2008), Genschel and Meeker (2010), Guure and Ibrahim (2013), Yu and Peng (2013), Jia (2020), Jiang (2022), is an interesting task which we intend to undertake.

Acknowledgements We would like to thank the Referees very much for their valuable comments and suggestions.

Declarations

Conflict of interest We certify that we have no affiliations with or involvement in any organization or entity with any financial interest, or non-financial interest that are directly or indirectly related in the subject matter or material discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akram M, Hayat A (2014) Comparison of estimators of the Weibull distribution. *J Stat Theory Pract* 8(2):238–259
- Al-Baidhani PA, Sinclair CD (1987) Comparison of methods of estimation of parameters of the Weibull distribution. *Commun Stat Simul Comput* 16:373–384
- Almeida JB (1999) Application of the Weibull statistics to the failure of coatings. *J Mater Technol* 93:259–263
- Alemay R, Bolancé C, Guillén M (2013) A nonparametric approach to calculating value-at-risk. *Insurance Math Econom* 52(2):255–262
- Bain LJ, Antle CE (1967) Estimation of parameters in the Weibull distribution. *Technometrics* 9:621–627
- Balakrishnan N, Kateri M (2008) On the maximum likelihood estimation of Weibull distribution based on complete and censored data. *Stat Probab Lett* 78:2971–2975
- Bolancé C, Guillén M (2021) Nonparametric estimation of extreme quantiles with an application to longevity risk. *Risks* 9(4):77
- Carroll KJ (2003) On the use and utility of the Weibull model in the analysis of survival data. *Control Clin Trials* 24:682–701
- Chen Q, Gerlach RH (2013) The two-sided Weibull distribution and forecasting financial tail risk. *Int J Forecast* 29
- Cohen AC (1965) Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics* 7(4):579–588
- Cohen CA, Whitten B (1982) Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. *Commun Stat Theory Methods* 11(23):2631–2656
- Cran GW (1988) Moment estimators for the 3-parameter Weibull distribution. *IEEE Trans Reliab* 37:360–363
- Davidson R (2007) Reliable inference for the Gini index. *J Econom* 150:30–40
- Engeman RM, Keefe TJ (1982) On generalized least squares estimation of the Weibull distribution. *Commun Stat Theory Methods* 11(19):2181–2193
- Fok SL, Mitchel BC, Smart J, Marsden BJ (2001) A numerical study on the application of the Weibull theory to brittle materials. *Eng Fract Mech* 68:1171–1179
- Gebizlioglu OL, Şenoğlu B, Kantar YM (2011) Comparison of certain value-at-risk estimation methods for the two-parameter Weibull loss distribution. *J Comput Appl Math* 235:3304–3314
- Genschel U, Meeker W (2010) A comparison of maximum likelihood and median-rank regression for Weibull estimation. *Qual Eng* 22:236–255
- George F (2014) A comparison of shape and scale estimators of the two-parameter Weibull distribution. *J Mod Appl Stat Methods* 13(1):23–35
- Ghosh K, Jammalamadaka SR (2001) A general estimation method using spacings. *J Stat Plan Inference* 93:71–82
- Gibbons DI, Vance LC (1981) A simulation study of estimators for the 2-parameter Weibull distribution. *IEEE Trans Reliab* 30(1):61–66
- Gigliarano C, Muliere P (2013) Estimating the Lorenz curve and Gini index with right censored data: a Polya tree approach. *Metron* 71:105–122
- Greenwood JA, Landwehr JM, Matalas JM, Wallis JR (1979) Probability weighted moment: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res* 15:1049–1054
- Gross AJ, Lurie D (1977) Monte Carlo comparison of parameter estimators of the 2-parameter Weibull distribution. *IEEE Trans Reliab* 26(5):356–358
- Gumbel EJ (1958) *Statistics of extremes*. Columbia University Press, New York
- Guure C, Ibrahim N (2013) Methods for estimating the 2-parameter Weibull distribution with Type-I censored data. *Res J Appl Sci Eng Technol* 5:689–694
- Hassanein KM (1971) Percentile estimators for the parameters of the Weibull distribution. *Biometrika* 58:673–676
- Heo JH, Boes DC, Salas JD (2001) Regional flood frequency analysis based on a Weibull model: part 1. estimation, asymptotic variances. *J Hydrol* 242:157–170
- Hong L, Alfani G, Gigliarano C, Bonetti M (2018) giniinc: a Stata package for measuring inequality from incomplete income and survival data. *Stand Genomic Sci* 18(3):692–715

- Hung W-L (2001) Weighted least-squares estimation of the shape parameter of the Weibull distribution. *Qual Reliab Eng Int* 17(6):467–469
- Jacquelin J (1993) Generalization of the method of maximum likelihood. *IEEE Trans Electr Insul* 28(1):65–72
- Jia X (2020) Reliability analysis for Weibull distribution with homogeneous heavily censored data based on Bayesian and least-squares methods. *Appl Math Model* 83:169–188
- Jiang R (2022) A novel parameter estimation method for the Weibull distribution on heavily censored data. *Proc Inst Mech Eng Part O* 236(2):307–316
- Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, vol 1. Wiley, New York
- Kang D, Ko K, Huh J (2018) Comparative study of different methods for estimating Weibull parameters: a case study on Jeju Island, South Korea. *Energies* 11(2)
- Kantar MY (2015) Generalized least squares and weighted least squares estimation methods for distributional parameters. *REVSTAT* 13(3):263–282
- Kantar YM, Şenoğlu B (2008) A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter. *Comput Geosci* 34:1900–1909
- Kattumannil SK, Dewan I, Sreelakshmi N (2021) Non-parametric estimation of Gini index with right censored observations. *Stat Probab Lett* 175:109113
- Lu H-L, Chen C-H, Wu J-W (2004) A note on weighted least-squares estimation of the shape parameter of the Weibull distribution. *Qual Reliab Eng Int* 20(6):579–586
- Lv X, Zhang G, Ren G (2017) Gini index estimation for lifetime data. *Lifetime Data Anal* 23:275–304
- Mirzaei S, Mohtashami Borzadaran GR, Dehak M (2017) A comparative study of the Gini coefficient estimators based on the linearization and U-statistics methods. *Revista Colombiana de Estadística* 40:205–221
- Murthy D, Xie M, Jiang R (2004) Weibull models. Wiley series in probability and statistics. Wiley, Hoboken
- Pobocikova I, Sedliackova Z (2014) Comparison of four methods for estimating the Weibull distribution parameters. *Appl Math Sci* 8(83):4137–4149
- Queeshi FS, Sheikh AK (1997) Probabilistic characterization of adhesive wear in metals. *IEEE Trans Reliab* 46:38–44
- Roed K, Zhang T (2002) A note on the Weibull distribution and time aggregation bias. *Appl Econ Lett* 9:469–472
- Sadani S, Abdollahnezhad K, Teimouri M, Ranjbar V (2019). A new estimator for Weibull distribution parameters: comprehensive comparative study for Weibull distribution. [arXiv:1902.05658](https://arxiv.org/abs/1902.05658)
- Sarabia JM (2008) Parametric Lorenz Curves: models and applications. Springer, New York
- Seki T, Yokoyama S (1993) Simple and robust estimation of the Weibull parameters. *Microelectron Reliab* 33:45–52
- Singh VP (1987) On application of the Weibull distribution in hydrology. *Water Resour Manag* 1:33–43
- Singh VP, Cruise JF, Ma M (1990) A comparative evaluation of estimators of the Weibull distribution by Monte Carlo simulation. *J Stat Comput Simul* 36(4):229–241
- Smith JA (1987) Estimating the upper tail of flood frequency distributions. *Water Resour Res* 23(8):1657–1666
- Sun Z, Han C (2010) Parameter estimation of Weibull distribution based on second-kind statistics. *J Commun Softw Syst* 6(3):109–114
- Swain JJ, Venkatraman S, Wilson JR (1988) Least-squares estimation of distribution functions in Johnson's translation system. *J Stat Comput Simul* 29(4):271–297
- Teimouri M, Hoseini SM, Nadarajah S (2013) Comparison of estimation methods for the Weibull distribution. *Statistics* 47(1):93–109
- Teimouri M, Doser JW, Finley AO (2020) Forestfit: an R package for modeling plant size distributions. *Environ Model Softw* 131:104668
- Van Zyl JM, Schall R (2012) Parameter estimation through weighted least-squares rank regression with specific reference to the Weibull and Gumbel distributions. *Commun Stat Simul Comput* 41(9):1654–1666
- Wu D, Zhou J, Li Y (2006) Methods for estimating Weibull parameters for brittle materials. *J Mater Sci* 41:5630–5638
- Yu H-F, Peng C-Y (2013) Estimation for Weibull distribution with Type-II highly censored data. *Qual Technol Quant Manag* 10(2):193–202
- Zhang L, Xie M, Tang L (2008) Recent advances in reliability and quality in design, chapter on weighted least squares estimation for the parameters of Weibull distribution. Springer, London, pp 57–84

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.