



Article

Estimating Variances in Weighted Least-Squares Estimation of Distributional Parameters

Yeliz Mert Kantar

Department of Statistics, Faculty of Science, Anadolu University, 26470, Eskisehir, Turkey; ymert@anadolu.edu.tr

Academic Editor: Mehmet Pakdemirli

Received: 13 January 2016; Accepted: 14 March 2016; Published: 25 March 2016

Abstract: Many estimation methods have been proposed for the parameters of statistical distribution. The least squares estimation method, based on a regression model or probability plot, is frequently used by practitioners since its implementation procedure is extremely simple in complete and censoring data cases. However, in the procedure, heteroscedasticity is present in the used regression model and, thus, the weighted least squares estimation or alternative methods should be used. This study proposes an alternative method for the estimation of variance, based on a dependent variable generated via simulation, in order to estimate distributional parameters using the weighted least squares method. In the estimation procedure, the variances or weights are expressed as a function of the rank of the data point in the sample. The considered weighted estimation method is evaluated for the shape parameter of the log-logistic and Weibull distributions via a simulation study. It is found that the considered weighted estimation method shows better performance than the maximum likelihood, least-squares, and certain other alternative estimation approaches in terms of mean square error for most of the considered sample sizes. In addition, a real-life example from hydrology is provided to demonstrate the performance of the considered method.

Keywords: cumulative distribution function; variance estimation; heteroscedasticity; weighted least squares; simulation

1. Introduction

Statistical distributions have many applications in the areas of engineering, medical sciences, air quality determination, and so forth. Since the parameters of the considered distribution are used for inference results regarding topics of interest, their estimation methods have received great interest in the literature. The maximum likelihood estimation (MLE) method has good theoretical properties for large sample sizes ($n > 250$) and is often preferred [1]. However, it can show poor performance for small sample cases [1,2], and the MLE also requires an iterative numerical method, such as Newton–Raphson, for most distributions. Additionally, some research has been conducted to find the best linear unbiased estimators for distributions. However, the proposed estimators in the literature are generally computationally intensive and require numerical methods to obtain estimates [1]. On the other hand, a regression procedure based on a probability plot to estimate the parameters of statistical distributions is frequently used by practitioners since its implementation procedure is simple in the cases of complete and censoring data. This procedure is easily performed if a distribution function can be expressed as an explicit function. A linear regression model, whose dependent variable is the nonparametric estimate of the value of the distribution function at the ranked sample, is obtained. Thus, the least squares (LS) estimates of the parameters of the resulting regression model become the estimates of the parameters of the considered statistical distribution. In the literature, parameters of different statistical distributions have been estimated by the least squares estimation (LSE) method. For example, Altun [3] compares the LSE, the weighted least squares estimation (WLSE), and certain other estimation methods for the Weibull distribution. [4–9] have used the LSE method for different

statistical distributions. However, heteroscedasticity (non-constant variance) is present in the used regression model and, thus, LS estimates lose the efficiency property, even if most studies do not take into account this reality [9]. In such cases, the WLSE or alternative methods should be used [10–20]. It is well-known that when conducting the WLSE, the variances of the dependent variables are unknown and must be estimated to implement the WLSE procedure. Considering this reality, [11–16] discuss that variances should be found in order to perform the WLSE. They all generally consider different weights using large sample properties of the empirical distribution function or order statistics to stabilize the variance when conducting the WLSE. For example, Hung [11] proposes the WLSE for estimating the shape parameter of the Weibull distribution. His results from simulation studies illustrate that the mean-squared error of WLSE is smaller than competing procedures. Lu *et al.* [12] consider the LSE and WLSE for the Weibull distribution and compare their WLSE method with three existing WLSE methods. They found that their WLSE method is more precise and has smaller variance than Bergman’s WLS estimators. Lu and Tao [14] consider the LSE and WLSE for the Pareto distribution. Zyl [15] considers regression procedure for the parameters of the three-parameter generalized Pareto distribution and applies the WLSE with the Box–Cox procedure. Zhang *et al.* [13] discuss the WLSE methods for Weibull distribution and related works. They find that the WLSE is an efficient method that makes good use of small datasets. Kantar [20] introduces the generalized least squares and WLSE methods, based on an easily-calculated approximation of the covariance matrix, for distributional parameters. They all emphasize that a weight function should be used when performing the regression procedures.

In this article, a different approach is considered for the estimation of the weights for the WLSE. The variances are modeled as a function of the rank of the data point in the sample by using samples which are simulated from the dependent variable of the established regression model for the specified distribution. Thus, the estimation of variances or weights is expressed as a simple mathematical function of its rank. Next, the proposed WLSE is then applied to the estimation of the parameters of the log-logistic and Weibull distributions. The simulation results show that the considered WLSE for the shape parameter of the log-logistic and Weibull distributions provides better performance than the MLE, LSE, and WLSE (Zyl and Schall) in terms of bias and mean square errors for most of the considered sample cases.

The rest of the paper is organized as follows: the LSE for estimating distributional parameters is given in Section 2; Section 3 introduces an estimation of variances to use in the WLSE as an application of estimation of the parameters of the log-logistic and Weibull distributions; to show the performance of the proposed WLSE, a simulation study is presented in Section 4; an application from real-life is provided in Section 5; and, finally, the last section summarizes the conclusions of the study.

2. LSE Method for Log-Logistic and Weibull Distributions

The distribution function can be transformed to a linear regression model, $Y_i = \beta_0 + \beta_1 X_i$, if it can be written as an explicit function. Next, the LSE method can be used to calculate the parameters. In estimation, the sum of the squares of the errors, which is defined below, should be minimized;

$$\min_{\beta_0, \beta_1} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2 \tag{1}$$

The obtained LS estimates are:

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n Y_i X_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2} \tag{2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \tag{3}$$

If we consider the LSE for the distributional parameters, the following procedure should be followed for the Weibull distribution.

The cumulative distribution function (cdf) of the Weibull random variable, widely-used in engineering fields [1,21,22] is given as follows:

$$F(x, \lambda, \alpha) = 1 - e^{-(\lambda x)^\alpha}, \text{ for } x > 0 \tag{4}$$

where λ is the scale parameter and α is the shape parameter, which is of great importance to the Weibull since it determines the shape of distribution. α and λ are unknown parameters in real applications.

After algebraic manipulation, Equation (4) can be linearized as follows:

$$\ln[-\ln(1 - F(x))] = \alpha \ln \lambda + \alpha \ln x \tag{5}$$

For a sample of size n and $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$, the regression model is rewritten as:

$$\ln[-\ln(1 - F(x_{(i)}))] = \alpha \ln \lambda + \alpha \ln x_{(i)} \tag{6}$$

If we replace $\ln[-\ln(1 - F(x_{(i)}))]$ with Y , $\alpha \ln \lambda$ with β_0 , α with β_1 , and $\ln x_{(i)}$ with X , the regression model with error term occurs as:

$$Y_{(i)} = \beta_0 + \beta_1 x_{(i)} + \varepsilon_i \tag{7}$$

If we consider the log-logistic distribution, its cdf is given as follows:

$$F(x) = 1 - \left(1 + \left(\frac{x}{\gamma}\right)^\delta\right)^{-1}, \text{ } x \geq 0 \text{ and } \delta, \gamma > 0 \tag{8}$$

where γ is the scale parameter and δ is the shape parameter, the following linear model is obtained to estimate the parameters of log-logistic distribution:

$$\ln\left((1 - F(x))^{-1} - 1\right) = \delta \ln(x) - \delta \ln(\gamma) \tag{9}$$

For a sample of size n , let $x_{(1)}, \dots, x_{(n)}$ be the order statistics and thus, the regression model can be expressed as follows:

$$\ln\left(1 - \left(1 - F(x_{(i)})\right)^{-1}\right) = \delta \ln(x_{(i)}) - \delta \ln(\gamma) \tag{10}$$

$\frac{i-a}{(n+b)}$, ($0 \leq a \leq 0.5$, $0 \leq b \leq 1$) is used as estimate of $F(x_{(i)})$, where i is the rank of the data point in the sample in ascending order. Since $\frac{i-0.3}{(n+0.4)}$ shows better performance than the other considered alternatives, we use it in this study.

Replacing $F(x_{(i)})$ by its estimate, called \hat{F}_i , the Equations (6) and (10) with error terms yield the following equations for estimating parameters of the Weibull and log-logistic distributions:

$$\ln[-\ln(1 - \hat{F}_i)] = \alpha \ln \lambda + \alpha \ln x_{(i)} + \varepsilon_i \tag{11}$$

$$\ln(1 - (1 - \hat{F}_i)^{-1}) = \delta \ln(x_{(i)}) - \delta \ln(\gamma) + \varepsilon_i \tag{12}$$

On the other hand, it should be noted that the error terms of the models given in Equations (11) and (12) are not identically distributed as mentioned by [9,15–17,19,20], since $x_{(i)}$ is beta distributed with parameters i and $n - i + 1$ and, thus, the models (11) and (12) have a different variance for each i as earlier mentioned [9,19,20]. In other words, error terms of the previously mentioned models have no equal variance. This situation can negatively affect the LSE. In such cases, alternative estimation approaches to stabilize variances should be used.

It should be emphasized that, without considering non-equal variances on Equations (11)–(12), robust regression methods have also been examined in other research [3,8,9,21,23,24].

3. Estimating Weights in the Weighted Least-Squares Estimation for Parameters of the Log-Logistic and Weibull Distributions

If we take into account the models given in Equations (11) and (12) it can be seen that the estimate of $F(x_{(i)})$ is a function of $(i - a, n + b)$. Thus, the estimate of the variance of $\ln[-\ln(1 - F(x_{(i)}))]$ can be expressed as a function of $(i - a, n + b)$. That is, the following equation can be written for the variance of the dependent variable or error term of the model Equation (11):

$$\text{Var}(\ln[-\ln(1 - F(x_{(i)}))]) = G(i - a, n + b) + \varepsilon_i \tag{13}$$

where Var denotes the variance, G is a differentiable function and ε_i is an error term. If we replace $\text{Var}(\ln[-\ln(1 - F(x_{(i)}))])$ by Z_i , the model occurs as follows:

$$Z_i = G(i - a, n + b) + \varepsilon_i \tag{14}$$

Replacing Z_i by its estimate, \hat{Z}_i and considering \hat{G} as estimation of G , the following model yields:

$$\hat{Z}_i = \hat{G}(i - a, n + b) \tag{15}$$

It is known that cdf of any continuous random is distributed uniformly $(U(0,1))$ with zero and unit parameters, thus, $F(x_{(i)}) \sim U(0,1)$. By using this reality, the distribution of $(\ln[-\ln(1 - F(x_{(i)}))])$ for each, $i = 1, \dots, n$ can be generated via simulation and then variance can be estimated. Figure 1a–d shows scatter plots the variance of $\ln[-\ln(1 - F(x_{(i)}))]$, estimated from 5000 simulated samples, versus the rank $i, i = 1, \dots, n$, respectively. Thus, it can be seen from the figures that the variances of Z_i tend to decrease as the rank i increases. In other words, the weight is inversely related to the rank (the rank 1 indicates the highest weight, while rank n indicates the lowest weight). Thereby, the variance of $\ln[-\ln(1 - F(x_{(i)}))]$ can be modeled by a decreasing function of rank i .

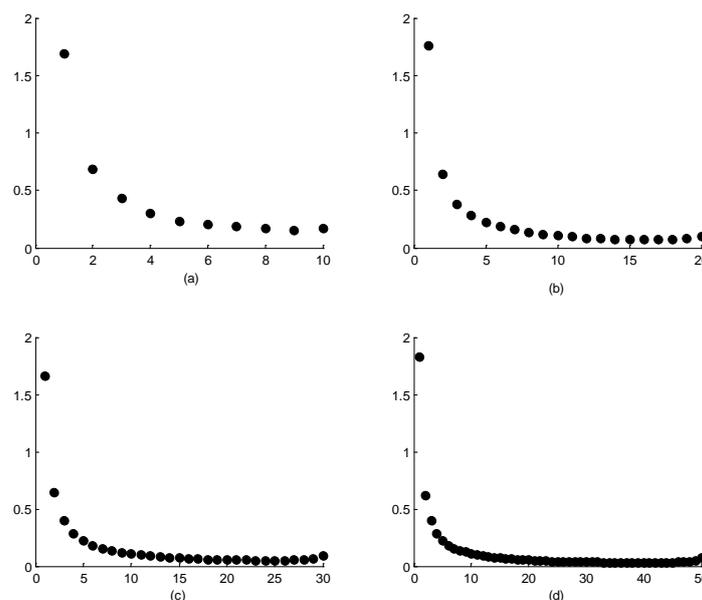


Figure 1. (a–d) show the estimated variance for the Weibull distribution versus $i = 1, \dots, n$ when the sample size is taken as $n = 10, 20, 30, 50$, respectively.

As a simple model for $G(\cdot)$, we take $a = 0$ and $b = 0$. Table 1 shows the considered models and the obtained R^2 from these models.

Table 1. The considered linear models for $G(\cdot)$ for the Weibull distribution.

$G(\cdot)$	α/i	$(1/i)^\alpha$	$(\alpha/i) + (\alpha/i^2)$	$exp(\alpha i)$
R^2	0.98	0.98	0.99	0.95

It can be deduced from Table 1 that $G(i, n) = \alpha/i$ with the determination of coefficient 0.98 and the estimated $\alpha = 1.51$ may be an alternative function to model variance. Since the weight is the inverse of the variance, $w_i = i/\alpha, i = 1, 2, \dots, n$. Hence, weights are expressed as a simple mathematical function or as a model of its rank.

For the log-logistic distribution case, Figure 2a–d shows scatter plots of the variance of $\ln\left(1 - \left(1 - F(x_{(i)})\right)^{-1}\right)$, which is estimated from 5000 simulated samples, versus rank $i, i = 1, \dots, n$. It can be seen observed from the figures that the variances, Z_i , tend to decrease and increase, while rank i increases to $\frac{n}{2}$ and n , respectively. Therefore, we can use the given functions in Table 1 to model variance of the $\ln\left(1 - \left(1 - F(x_{(i)})\right)^{-1}\right)$ for $i = 1, \dots, (\frac{n}{2})$, and the determined function can be used for $i = (\frac{n}{2}) + 1, \dots, n$. Alternatively, U shaped functions, which are more expensive than those given in Table 2, may be used.

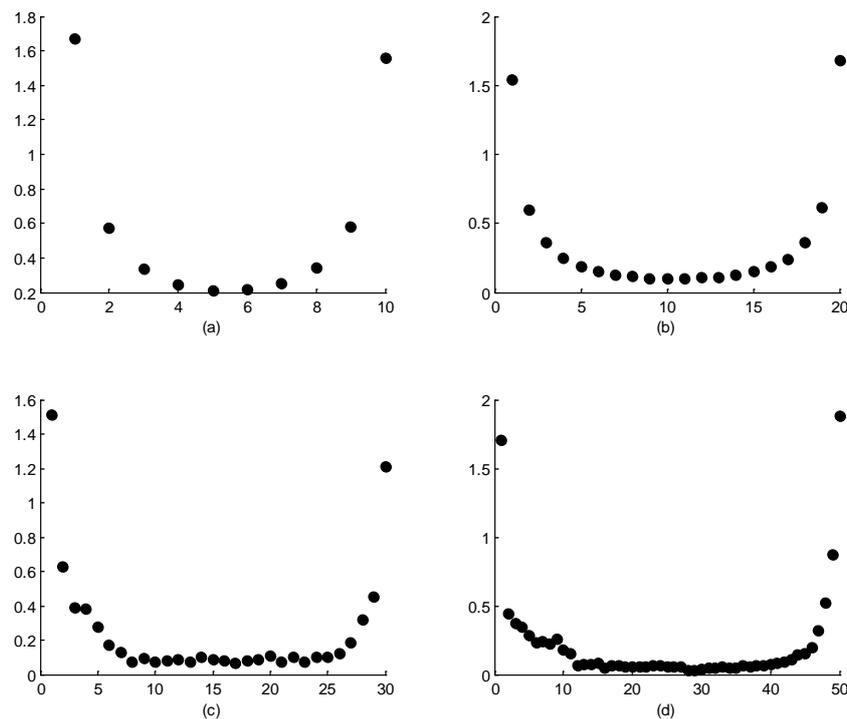


Figure 2. (a–d) show the estimated variance for the log-logistic distribution versus $i = 1, \dots, n$ when the sample size is taken as $n = 10, 20, 30, 50$, respectively.

Table 2. The considered linear models for $G(\cdot)$ for the log-logistic distribution.

$G(\cdot)$	α/i	$(1/i)^\alpha$	$(\alpha/i) + (\alpha/i^2)$	$exp(\alpha i)$
R^2	0.98	0.98	0.99	0.95

Parallel to the Weibull distribution case, it can be concluded from Table 2 that $G(i, n) = \alpha/i$ with the determination of coefficient 0.98, and the estimated $\alpha = 1.43$ can be a plausible model for the variance and, thus, $w_i = i/\alpha$.

Therefore, we suggest a simple formula to calculate the weights to be used in WLS for estimating the parameters of the Weibull and log-logistic distributions. Such formulae for weights of the Pareto, Logistic, Gumbel, Burr, and such like random variables, whose cumulative distribution functions have explicit functional form, can be easily obtained by means of the simulation study mentioned above.

The WLSE procedure for the Weibull and log-logistic distributions can be carried out by minimizing the weighted sum of squares with respect to the unknown shape and scale parameters, thus:

$$\sum w_i (\ln[-\ln(1 - \hat{F}_i)] - \alpha \ln \lambda - \alpha \ln(x_{(i)}))^2 \tag{16}$$

$$\sum w_i (\ln(1 - (1 - \hat{F}_i)^{-1}) - \delta \ln(x_{(i)}) + \delta \ln(\gamma))^2 \tag{17}$$

With matrix notations for the Weibull distribution, $y = (\ln[-\ln(1 - \hat{F}_1)], \dots, \ln[-\ln(1 - \hat{F}_n)])'$,

$$X = \begin{pmatrix} 1 & \log(x_{(1)}) \\ \vdots & \vdots \\ 1 & \log(x_{(n)}) \end{pmatrix} \text{ and } = \begin{pmatrix} w_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_n \end{pmatrix}, \beta = (\beta_0, \beta_1), \text{ where } \beta_0 = \alpha \ln \lambda \text{ and } \beta_1 = \alpha,$$

the solution of the WLSE is obtained as follows:

$$\hat{\beta} = (X'WX)^{-1} X'Wy \tag{18}$$

where $\hat{\alpha} = \beta_1$ and $\hat{\lambda} = \exp\left(\frac{\beta_0}{\beta_1}\right)$.

A similar process can be applied to estimate the parameters of the log-logistic distribution. In conclusion, the considered WLSE is explicit functions of sample observations and is, therefore, easy to compute and does not have computational complexities like the MLE [1].

4. Monte Carlo Simulation

Different estimation methods may result in different estimates. Thus, it is important to have objective criteria to inform the chosen method over other alternatives. A common approach to select the best method is the Monte Carlo simulation by using appropriate criteria: bias and mean squared error (MSE) [25]. In this section, the performance of the WLSE with the proposed weights is evaluated by means of a simulation study. The considered WLSE is compared with the MLE, LSE, and WLSE (Zyl and Schall) for the parameters of the log-logistic and Weibull distributions. In the simulation study, the replication number is taken as 5000. All computations for simulation are performed using Matlab 10.1. Sample sizes are taken as $n = 10, 20, 30, 50, 100,$ and 250 and the shape parameters are considered as $\alpha = 1, 2, 3$ and 6 as parallel to previous studies concerning the log-logistic and Weibull distributions. Without any loss of generality, the scale parameter is taken to be equal to 1.

From the simulation results presented in Table 3 for the shape parameter of the log-logistic distribution, the following conclusions may be summarized. The MSE and bias values of the proposed WLSE decrease when sample sizes increase. Thus, we show that the proposed WLSE provides consistent estimates. According to the MSE criterion, the proposed WLSE apparently shows better performance than the MLE, LSE and WLSE (Zyl and Schall) for all considered shape parameters when sample sizes are $n = 10, 20, 30,$ and 50 . Since the MLE is asymptotically the best, it provides the best performance for $n = 100$ and 250 , as expected. It can be seen from analysis that WLSE has a better performance than the LSE and the WLSE (Zyl and Schall) for estimation of the shape parameter of the log-logistic distribution in terms of MSE.

Table 3. Bias and MSE of the estimated shape parameters of the log-logistic distribution.

Shape Parameter	1		2		3		6	
<i>n</i> = 10								
Methods	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	0.15170	−0.15212	0.61690	−0.31629	1.37293	−0.44112	5.67004	−0.98941
LSE	0.11440	−0.09667	0.54134	−0.20337	1.10040	−0.28495	3.99536	−0.64275
WLSE(Zyl and Schall)	0.19999	−0.21543	0.81328	−0.44568	1.78391	−0.62753	7.43891	−1.37432
Proposed WLSE	0.10839	−0.02486	0.43330	−0.06128	1.00824	−0.06953	3.98816	−0.21838
<i>n</i> = 20								
MLE	0.05213	−0.06977	0.20253	−0.13622	0.45077	−0.20854	1.79207	−0.40466
LSE	0.05036	−0.03032	0.19174	−0.05249	0.42394	−0.08664	1.68610	−0.17020
WLSE(Zyl and Schall)	0.07190	−0.11669	0.28007	−0.22935	0.63664	−0.35112	2.46200	−0.68556
Proposed WLSE	0.04278	0.02031	0.16383	0.04718	0.36171	0.06327	1.45807	0.14500
<i>n</i> = 30								
MLE	0.02778	−0.04194	0.11692	−0.09339	0.26793	−0.12972	1.15141	−0.29157
LSE	0.02773	−0.00947	0.11498	−0.02627	0.26564	−0.03313	1.12986	−0.09620
WLSE(Zyl and Schall)	0.03941	−0.07720	0.16531	−0.15971	0.37430	−0.23364	1.59306	−0.50179
Proposed WLSE	0.02509	0.02770	0.10194	0.04546	0.23752	0.08039	0.99113	0.12622
<i>n</i> = 50								
MLE	0.01566	−0.02583	0.03070	−0.02512	0.14066	−0.07853	0.57653	−0.13840
LSE	0.01633	0.00006	0.03428	0.01186	0.14630	0.00037	0.60268	0.01698
WLSE(Zyl and Schall)	0.02169	−0.04636	0.04053	−0.04015	0.19377	−0.13955	0.79926	−0.27090
Proposed WLSE	0.01521	0.02195	0.03136	0.02328	0.13589	0.06200	0.56752	0.13359
<i>n</i> = 100								
MLE	0.00737	−0.01237	0.02925	−0.02169	0.06983	−0.04588	0.26815	−0.07447
LSE	0.00819	0.00447	0.03224	0.01293	0.07529	0.00617	0.29422	0.02979
WLSE(Zyl and Schall)	0.00993	−0.02040	0.03966	−0.03845	0.09062	−0.07264	0.34757	−0.12550
Proposed WLSE	0.00749	0.01454	0.03025	0.03137	0.07067	0.03276	0.27823	0.08239
<i>n</i> = 250								
MLE	0.00288	−0.00549	0.01132	−0.00937	0.02607	−0.01758	0.10527	−0.02345
LSE	0.00325	0.00419	0.01289	0.01079	0.02891	0.01211	0.12167	0.03412
WLSE(Zyl and Schall)	0.00375	−0.00810	0.01484	−0.01423	0.03342	−0.02428	0.13741	−0.03586
Proposed WLSE	0.00312	0.00624	0.01228	0.01308	0.02845	0.01610	0.11304	0.04641

Note: Scale is taken as 1/λ.

If we evaluate the WLSE in terms of bias criterion, it is observed from Table 3 that the proposed WLSE is best for *n* = 10, 20 and it is the best performer next to the LSE for *n* = 30, 50. In addition, the WLSE provides less bias than the WLSE (Zyl and Schall) for *n* = 100 and 250. As a result, the considered WLSE with the proposed weights can be a good alternative estimation method for the shape parameter of the log-logistic distribution. In particular, it is useful in dealing with small sample sizes.

If the simulation results for the scale parameter of the log-logistic distribution are evaluated, it can be concluded that the proposed WLSE is the best method, when sample sizes are taken as 10, 20, 50, and 100 in terms of MSE for most of the shape parameter cases except for the cases of shape parameters 3 and 6. Similar to the results of the shape parameter, the MLE naturally provides the least MSE for *n* = 250. It can be seen from Table 3 that the WLSE and WLSE (Zyl and Schall) perform very similarly and they outperform the LSE for almost all cases. According to bias, the WLSE provides less bias than the LSE and WLSE (Zyl and Schall) in almost all cases as seen in Table 4.

Based on the results presented in Table 5 for the shape parameter of the Weibull distribution, the following conclusions may be summarized. The proposed WLSE shows the best performance for all sample sizes except *n* = 250, which favors the MLE. The WLSE outperforms the LSE and WLSE (Zyl and Schall) for all cases. According to bias criterion, while the LSE outperforms others for *n* = 10 and 20, the proposed WLSE is the best performer next to the LSE. For other sample sizes, the proposed WLSE gives satisfactory results.

Table 4. Bias and MSE of the estimated scale parameters of the log-logistic distribution.

Shape Parameter	1		2		3		6	
<i>n</i> = 10								
Methods	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	0.53291	−0.16983	0.08620	−0.03690	0.03530	−0.01621	0.0086	−0.00188
LSE	0.28823	0.28573	0.09542	0.19277	0.04874	0.14074	0.01472	0.07957
WLSE(Zyl and Schall)	0.27738	0.20752	0.08143	0.14670	0.04022	0.10846	0.01173	0.06139
Proposed WLSE	0.27505	0.19492	0.07949	0.13951	0.03900	0.10382	0.01127	0.05852
<i>n</i> = 20								
MLE	0.20817	−0.08162	0.04173	−0.02409	0.01733	−0.00745	0.00442	−0.00346
LSE	0.16078	0.19233	0.04737	0.11886	0.02340	0.08807	0.00654	0.04575
WLSE(Zyl and Schall)	0.14875	0.11459	0.03943	0.07457	0.01863	0.05758	0.00509	0.02970
Proposed WLSE	0.14925	0.10117	0.03864	0.06670	0.01819	0.05303	0.00491	0.02705
<i>n</i> = 30								
MLE	0.11511	−0.04886	0.02753	−0.01467	0.01128	−0.00538	0.00276	−0.00181
LSE	0.10550	0.15271	0.03210	0.08778	0.01482	0.06448	0.00399	0.03366
WLSE(Zyl and Schall)	0.09359	0.08221	0.02668	0.05080	0.01187	0.03837	0.00308	0.02013
Proposed WLSE	0.09378	0.07111	0.02630	0.04494	0.01161	0.03451	0.00299	0.01821
<i>n</i> = 50								
MLE	0.06937	−0.03995	0.00766	−0.00371	0.00696	−0.00388	0.00173	−0.00055
LSE	0.06566	0.09608	0.00896	0.03468	0.00870	0.04251	0.00234	0.02330
WLSE(Zyl and Schall)	0.05982	0.03974	0.00765	0.01630	0.00714	0.02252	0.00185	0.01269
Proposed WLSE	0.05973	0.03129	0.00745	0.01355	0.00702	0.01960	0.00182	0.01126
<i>n</i> = 100								
MLE	0.03096	−0.01622	0.00757	−0.00265	0.00327	−0.00273	0.00085	−0.00016
LSE	0.03227	0.05936	0.00890	0.03609	0.00395	0.02353	0.00107	0.01298
WLSE(Zyl and Schall)	0.02912	0.02344	0.00763	0.01716	0.00332	0.01043	0.00088	0.00646
Proposed WLSE	0.02901	0.01845	0.00750	0.01454	0.00328	0.00867	0.00088	0.00559
<i>n</i> = 250								
MLE	0.01215	−0.00633	0.00296	−0.00003	0.00137	−0.00053	0.00034	0.00010
LSE	0.01320	0.02975	0.00345	0.01802	0.00156	0.01156	0.00041	0.00610
WLSE(Zyl and Schall)	0.01181	0.00944	0.00300	0.00800	0.00138	0.00477	0.00035	0.00277
Proposed WLSE	0.01175	0.00720	0.00298	0.00667	0.00138	0.00399	0.00035	0.00236

Table 5. Bias and MSE of the estimated shape parameters of the Weibull distribution.

Shape parameter	1		2		3		6	
<i>n</i> = 10								
Methods	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	0.16290	−0.17684	0.61002	−0.34672	1.39320	−0.48813	5.30808	−1.00735
LSE	0.10553	0.02617	0.40657	0.05791	0.92350	0.10020	3.63395	0.18918
WLSE(Zyl and Schall)	0.09019	0.04832	0.33633	0.10347	0.80193	0.17179	3.01066	0.32833
Proposed WLSE	0.08970	0.04528	0.33353	0.09746	0.79808	0.16480	2.98083	0.31209
<i>n</i> = 20								
MLE	0.05051	−0.07862	0.20273	−0.15801	0.43237	−0.22142	1.76956	−0.44148
LSE	0.05122	0.03667	0.20062	0.07636	0.45081	0.12566	1.81754	0.24342
WLSE(Zyl and Schall)	0.03805	0.04307	0.15174	0.08599	0.33618	0.13991	1.38091	0.28552
Proposed WLSE	0.03713	0.04112	0.14802	0.08202	0.32646	0.13488	1.34091	0.27419
<i>n</i> = 30								
MLE	0.02709	−0.04567	0.10771	−0.09483	0.23843	−0.14551	0.96885	−0.30166
LSE	0.03418	0.04049	0.13882	0.07656	0.29428	0.11197	1.23346	0.22298
WLSE(Zyl and Schall)	0.02383	0.03874	0.09409	0.07331	0.20710	0.10699	0.83766	0.20695
Proposed WLSE	0.02288	0.03733	0.08996	0.07127	0.19936	0.10311	0.79907	0.19891
<i>n</i> = 50								
MLE	0.01465	−0.0233	0.06141	−0.0627	0.13222	−0.0875	0.52323	−0.1647
LSE	0.02333	0.06806	0.09433	0.14283	0.21502	0.21366	0.86411	0.42794
WLSE(Zyl and Schall)	0.01490	0.04802	0.05899	0.08928	0.13479	0.14028	0.55811	0.28503
Proposed WLSE	0.01440	0.04733	0.05896	0.08917	0.13236	0.10311	0.5300	0.28891

Table 5. Cont.

Shape parameter	1		2		3		6	
<i>n</i> = 100								
MLE	0.00677	-0.01440	0.02660	-0.03043	0.05868	-0.04105	0.24495	-0.09501
LSE	0.01114	0.02299	0.04404	0.04507	0.09679	0.07418	0.38372	0.13333
WLSE(Zyl and Schall)	0.00711	0.01278	0.02742	0.02316	0.06246	0.04011	0.24995	0.06395
Proposed WLSE	0.00663	0.01205	0.02589	0.02420	0.05872	0.04100	0.24438	0.06505
<i>n</i> = 250								
MLE	0.00242	-0.00611	0.00982	-0.01081	0.02268	-0.02154	0.08933	-0.02439
LSE	0.00434	0.01466	0.01728	0.03187	0.03825	0.03658	0.16016	0.09904
WLSE(Zyl and Schall)	0.00270	0.00405	0.01112	0.01105	0.02511	0.00906	0.10153	0.04266
Proposed WLSE	0.00249	0.00404	0.01024	0.01169	0.02329	0.01205	0.09355	0.04241

The results of the simulation for the scale parameter of the Weibull distribution are given in Table 6. While, according to the MSE criterion, the proposed WLSE, LSE, and WLSE (Zyl and Schall) yield similar results, the proposed WLSE provides comparable efficiency in comparison with LSE in terms of bias. Moreover, the proposed WLSE is compared to two alternative WLSE (Hung [11], Lu *et al.* [12]) and also PE [2] for the shape parameter of the Weibull distribution. It can be seen that the proposed WLSE demonstrates satisfactory results in terms of MSE.

Table 6. Bias and MSE of the estimated scale parameters of the Weibull distribution.

Shape Parameter	1		2		3		6	
<i>n</i> = 10								
Methods	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias
MLE	0.11141	-0.01587	0.02711	0.00626	0.01246	0.00561	0.00312	0.00412
LSE	0.15382	-0.04226	0.03203	-0.00566	0.01343	0.00074	0.00326	0.00258
WLSE(Zyl and Schall)	0.17017	-0.08331	0.03342	-0.02544	0.01375	-0.01216	0.00328	-0.00398
Proposed WLSE	0.17033	-0.08574	0.03400	-0.03128	0.01372	-0.01586	0.00328	-0.00590
<i>n</i> = 20								
MLE	0.05694	-0.00975	0.01438	0.00157	0.00613	0.00245	0.00162	0.00248
LSE	0.06613	-0.00934	0.01561	0.00333	0.00660	0.00446	0.00168	0.00229
WLSE(Zyl and Schall)	0.07133	-0.03953	0.01603	-0.01178	0.00661	-0.00566	0.00168	-0.00258
Proposed WLSE	0.07289	-0.04828	0.01619	-0.01601	0.00661	-0.00840	0.00168	-0.00297
<i>n</i> = 30								
MLE	0.03774	-0.00943	0.00924	0.00119	0.00415	0.00154	0.00103	0.00096
LSE	0.04201	0.00142	0.01003	0.00484	0.00440	0.00435	0.00110	0.00328
WLSE(Zyl and Schall)	0.04407	-0.02275	0.01004	-0.00763	0.00438	-0.00371	0.00107	-0.00091
Proposed WLSE	0.04455	-0.02932	0.01003	-0.01076	0.00437	-0.00478	0.00106	-0.00119
<i>n</i> = 50								
MLE	0.02187	-0.00347	0.00558	0.00231	0.00248	0.00129	0.00062	0.00103
LSE	0.02378	0.00483	0.00599	0.00366	0.00264	0.00385	0.00067	0.00229
WLSE(Zyl and Schall)	0.02425	-0.01454	0.00597	-0.00594	0.00261	-0.00256	0.00064	-0.00097
Proposed WLSE	0.02423	-0.01881	0.00597	-0.00813	0.00259	-0.00302	0.00064	-0.00108
<i>n</i> = 100								
MLE	0.01100	-0.00012	0.00276	0.00007	0.00122	0.00199	0.00031	0.00101
LSE	0.01167	0.00325	0.00297	0.00447	0.00132	0.00169	0.00033	0.00112
WLSE(Zyl and Schall)	0.01168	-0.00876	0.00291	-0.00193	0.00129	-0.00251	0.00033	-0.00102
Proposed WLSE	0.01162	-0.01009	0.00287	-0.00300	0.00127	-0.00307	0.00032	-0.00112
<i>n</i> = 250								
MLE	0.00432	0.00057	0.00107	0.00078	0.00048	0.00056	0.00012	0.00050
LSE	0.00457	0.00327	0.00113	0.00227	0.00052	0.00153	0.00013	0.00072
WLSE(Zyl and Schall)	0.00454	-0.00396	0.00112	-0.00122	0.00051	-0.00083	0.00012	-0.00042
Proposed WLSE	0.00436	-0.00379	0.00110	-0.00112	0.00050	-0.00109	0.00012	-0.00043

To sum up, the proposed WLSE for the shape parameter of the Weibull distribution, which characterizes failure rate and has a significant impact on the accuracy of wind power estimation, provides better performance than the MLE, LSE, and three alternative WLSE in terms of MSE for the most of the considered sample (*n* < 100) and shape parameter cases. For the log-logistic distribution, it

may be concluded that the proposed WLSE for the shape parameter is a good alternative to the MLE, LSE, and WLSE.

5. Real-Life Application from Hydrology

In this study, a flood frequency analysis is performed for one station of the Aji River basin in Iran as an illustrative example. Data measured at the SofiChai station is used. An analysis of the considered flood data is studied in detail in [26] with most statistical distributions including the log-logistic and Weibull distributions. We also research whether data follows the Weibull and log-logistic distributions using statistical tests, such as Kolmogorov–Smirnov, Anderson–Darling, and Chi-squared. In the results of the considered tests and a q-q plot, we conclude that the Weibull and log-logistic distributions may be used as alternative models for model flood data.

Figure 3 shows the weights calculated for sample size, $n = 34$ for Weibull and log-logistic distributions in order to perform the WLSE.

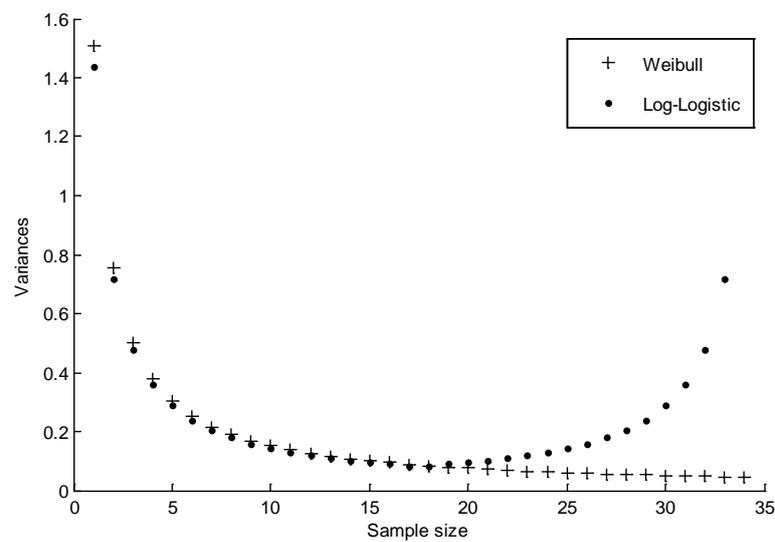


Figure 3. The estimated variances corresponding to Weibull and log-logistic distributions.

We now calculate the estimates of the scale and shape parameters of the log-logistic and Weibull distributions using the previously mentioned estimation methods in this study; see Table 7.

Table 7. Parameter estimates for log-logistic and Weibull distributions.

Parameters	Log-logistic Distribution						
	MLE	LSE	WLSE (Zyl and Schall)	Proposed WLSE	LAD	WLSE (Hung)	WLSE (Lu et al.)
Scale	30.0184	27.0712	28.8336	29.4025	28.5428	-	-
Shape	3.94881	3.58588	4.47555	3.48969	3.82828	-	-
Weibull Distribution							
Scale	35.21229	35.5353	35.3532	35.2858	35.3344	35.0005	35.3607
Shape	2.80503	2.548301	2.634572	2.638023	2.659215	2.738903	2.63333

Figure 4 provides a histogram and the fitted Weibull probability density functions (pdfs) for flood data to evaluate the results of the MLE, LSE, and WLSE. It is observed that the Weibull pdfs capture most features of the observed frequency of the histogram of the observed data. Additionally, the performance of WLSE is evaluated based on root mean square error (RMSE) between the area of the histogram and the total area under the Weibull pdfs curve. It is clear that the values of RMSE are very

close to each other for all methods. 0.0532, 0.0543, and 0.0538 are obtained for the MLE, LSE, and WLSE, respectively. It is seen from this example that the RMSE of the WLSE is less than the LSE.

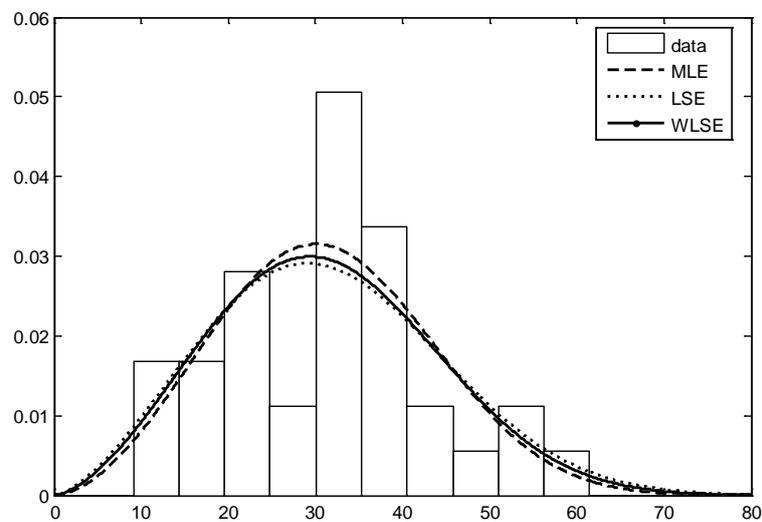


Figure 4. Weibull pdfs.

The log-logistic pdfs and the histogram are given in Figure 5. The RMSE of the MLE, LSE, and the considered WLSE are 0.0601, 0.0674, and 0.0610 respectively. As the Weibull case, the RMSE of WLSE is less than the LSE. Thus, the simulation results are supported by the results of application.

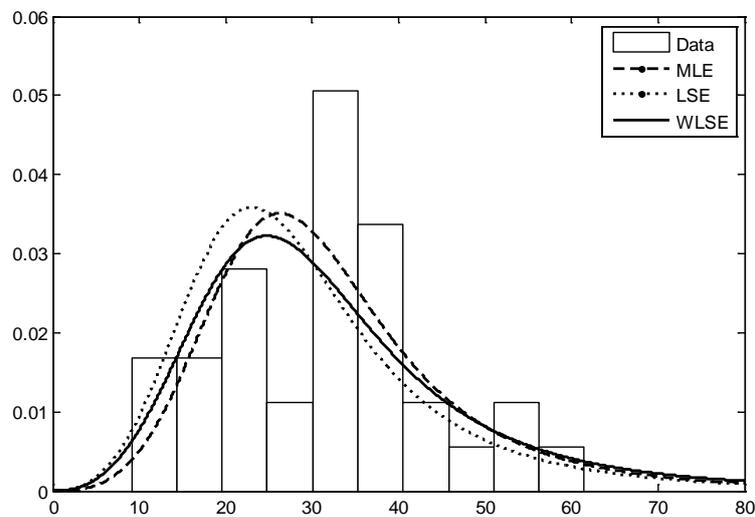


Figure 5. Log-logistic pdfs.

6. Conclusions

The key of WLSE is to calculate the weights. In this study, we propose estimating weights as a function of the rank of the data point in the sample based on dependent variable generated via simulation. The considered weighted least squares estimation, which is computationally easy, is then applied to the estimation of the parameters of the log-logistic and Weibull distributions. Considering the results of the Monte Carlo simulation and a real application, it is shown that the proposed WLSE for the shape parameters of the Weibull and the log-logistic distributions shows better performance than other considered methods for most of the considered parameters and sample cases.

Acknowledgments: This study was supported by Anadolu University Scientific Research Project Commission under the grant no: 1505F311.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Kantar, Y.M.; Şenoğlu, B. A comparative study for the location and scale parameters of the Weibull distribution with a given shape parameter. *Comput. Geosci.* **2008**, *34*, 1900–1909. [[CrossRef](#)]
2. Marks, N.B. Estimation of Weibull parameters from common percentiles. *J. Appl. Statist.* **2005**, *1*, 17–24. [[CrossRef](#)]
3. Altun, A. Estimation of the shape parameter of the Weibull distribution using linear regression methods: Non-censored samples. *Qual. Reliab. Eng. Int.* **2012**, *29*, 1207–1219.
4. Tiryakioglu, M.; Hudak, D. On estimating Weibull modulus by the linear regression method. *Mater. Sci.* **2007**, *42*, 10173–10179. [[CrossRef](#)]
5. Zhang, L.F.; Xie, M.; Tang, L.C. A study of two estimation approaches for parameters of Weibull distribution based on WPP. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 360–368. [[CrossRef](#)]
6. Genschel, U.; Meeker, W.Q. A Comparison of maximum likelihood and median-rank regression for Weibull estimation. *Qual. Eng.* **2010**, *22*, 236–255. [[CrossRef](#)]
7. Hossain, A.; Howlader, H.A. Unweighted least squares estimation of Weibull parameters. *J. Statist. Comput. Sim.* **1996**, *54*, 265–271. [[CrossRef](#)]
8. Kantar, Y.M.; Arik, I. M-estimation of log-logistic distribution parameters with outliers. *Int. J. Agric. Stat. Sci.* **2014**, *10*, 7–16.
9. Boudt, K.; Caliskan, D.; Croux, C. Robust explicit estimators of Weibull parameters. *Metrika* **2011**, *73*, 187–209. [[CrossRef](#)]
10. Bergman, B. Estimation of Weibull parameters using a weight function. *J. Mater. Sci.* **1986**, *5*, 611–614. [[CrossRef](#)]
11. Hung, W.L. Weighted least squares estimation of the shape parameter of the Weibull distribution. *Qual. Reliab. Eng. Int.* **2001**, *17*, 467–469. [[CrossRef](#)]
12. Lu, H.L.; Chen, C.H.; Wu, J.W. A Note on weighted least-squares estimation of the shape parameter of the Weibull distribution. *Qual. Reliab. Eng. Int.* **2004**, *20*, 579–586. [[CrossRef](#)]
13. Zhang, L.F.; Xie, M.; Tang, L.C. On weighted least squares estimation for the parameters of Weibull distribution. In *Recent Advances in Reliability and Quality in Design*, 1st ed.; Pham, H., Ed.; Springer Series in Reliability Engineering: London, UK, 2008; pp. 57–84.
14. Lu, H.L.; Tao, S.H. The Estimation of Pareto distribution by a weighted least square method. *Qual. Quant.* **2007**, *41*, 913–926. [[CrossRef](#)]
15. Zyl, J.M.V. A median regression model to estimate the parameters of the three-parameter generalized Pareto distribution. *Commun. Stat.-Simul.* **2012**, *41*, 544–553.
16. Zyl, J.M.V.; Schall, R. Parameter Estimation through weighted least-squares rank regression with specific reference to the Weibull and Gumbel distributions. *Commun. Stat. Simul.* **2012**, *41*, 1654–1666.
17. Kantar, Y.M.; Arik, A. The use of the data transformation techniques in estimating the shape parameter of the Weibull distribution. *IJEQE* **2014**, *3*, 20–33. [[CrossRef](#)]
18. Kantar, Y.M.; Arik, A.; Yenilmez, İ.; Usta, İ. Comparison of some estimation methods of the two parameter Weibull distribution for unusual wind speed data cases. *Int. J. Inform. Technol.* **2016**. in press.
19. Engeman, R.M.; Keefe, T.J. On generalized least squares estimation of the Weibull distribution. *Commun. Stat. Theory Methods* **1982**, *19*, 2181–2193. [[CrossRef](#)]
20. Kantar, Y.M. Generalized least squares and weighted least squares estimation methods for distributional parameters. *REVSTAT—Stat. J.* **2015**, *13*, 263–282.
21. Zhang, L.F.; Xie, M.; Tang, L.C. Robust regression using probability plots for estimating the Weibull shape parameter. *Qual. Reliab. Eng. Int.* **2006**, *22*, 905–917. [[CrossRef](#)]
22. Kantar, Y.M.; Usta, I. Analysis of wind speed distributions: Wind distribution function derived from minimum cross entropy principles as better alternative to Weibull function. *Energy Convers. Manag.* **2008**, *49*, 962–973. [[CrossRef](#)]

23. Acıtaş, Ş.; Kasap, P.; Şenoğlu, B.; Arslan, O. Robust estimation with the skew 2t distribution. *Pak. J. Stat.* **2013**, *29*, 409–430.
24. Arik, I.; Kantar, Y.M.; Yenilmez, I. The evaluation of robust and efficient estimators for log-logistic distribution for censored data with/without outliers. *JSRAD* **2015**, *2*, 24–32.
25. Kantar, Y.M.; Usta, I.; Acitas, S. A Monte Carlo Simulation Study on Partially Adaptive Estimators of linear Regression Models. *J. Appl. Stat.* **2011**, *38*, 1681–1699. [[CrossRef](#)]
26. Ghorbani, M.A.; Ruskepa, H.; Singh, V.P.; Sivakumar, B. Flood frequency analysis using mathematica. *Turk. J. Eng. Environ. Sci.* **2010**, *34*, 171–188.



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).