JOURNAL OF APPLIED STATISTICS
2024, VOL. 51, NO. 8, 1590–1608
https://doi.org/10.1080/02664763.2023.2229969

Taylor & Francis
Taylor & Francis Group

Robust explicit estimators using the power-weighted repeated medians

Chanseok Park [a], Xuehong Gao [b] and Min Wang [c,d]

[a] Applied Statistics Laboratory, Department of Industrial Engineering, Pusan National University, Busan, Republic of Korea; [b] Department of Safety Science and Engineering, University of Science and Technology Beijing, Beijing, People's Republic of China; [c] Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, TX, USA; [d] School of Data Science, The University of Texas at San Antonio, San Antonio, TX, USA

ABSTRACT

This paper consists of two parts. The first part of the paper is to propose an explicit robust estimation method for the regression coefficients in simple linear regression based on the power-weighted repeated medians technique that has a tuning constant for dealing with the trade-offs between efficiency and robustness. We then investigate the lower and upper bounds of the finite-sample breakdown point of the proposed method. The second part of the paper is to show that based on the linearization of the cumulative distribution function, the proposed method can be applied to obtain robust parameter estimators for the Weibull and Birnbaum-Saunders distributions that are commonly used in both reliability and survival analysis. Numerical studies demonstrate that the proposed method performs well in a manner that is approximately comparable with the ordinary least squares method, whereas it is far superior in the presence of data contamination that occurs frequently in practice.

ARTICLE HISTORY

Received 24 September 2022
Accepted 26 May 2023

KEYWORDS

Finite-sample breakdown point; linear model; repeated median; robustness; weighted median

MATHEMATICAL SUBJECT CLASSIFICATIONS

62F10; 62J05

1. Introduction

In an introductory statistics course, we commonly teach students how to determine the nature of a linear relationship between two quantitative random variables. Let (x_i, y_i) be an ordered pair from an experiment for $i = 1, 2, \ldots, n$. Knowing that the Pearson correlation coefficient is a measure of linear correlation between two sets of data from two random variables, we are interested in building a simple linear regression model that relates them, which is given by

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \tag{1}$$

where β_0 and β_1 are two unknown regression coefficients that stand for the intercept and slope terms in the simple linear regression model, respectively, and ε_i's are independent and identically distributed random variables with zero mean and finite variance.

To obtain estimates of the regression coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$, we usually adopt the least squares approach that chooses $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the residual sum of squares (RSS)

CONTACT Min Wang min.wang3@utsa.edu

defined as

$$\mathrm{RSS} = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2 .$$

By using some simple calculus, we can show that the minimizers, so-called the ordinary least squares (OLS) estimators of β_0 and β_1, are given by

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \text{and} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \tag{2}$$

respectively, where $\bar{x} = \sum_{i=1}^{n} x_i/n$ and $\bar{y} = \sum_{i=1}^{n} y_i/n$ are the sample means.

The OLS estimators in Equation (2) enjoy various desirable statistical properties; see, for example, Rousseeuw and Leroy [32], for more details. However, these estimators are highly affected by the presence of data contamination since they depend on the sample means whose finite-sample breakdown point is 0%. Here, data contamination implies that there may be one or more observations whose values seem extreme relative to the majority of the observations in a data set. The finite-sample breakdown point of an estimator is defined as the proportion of incorrect or contaminated observations (i.e. arbitrarily large or small observations) the estimate of a parameter can deal with before giving estimated values arbitrarily close to zero (implosion) or infinity (explosion). For more details on the finite-sample breakdown point, one can refer to Definition 2 of [27]. This finite-sample breakdown point is commonly based on the ϵ-replacement breakdown, as observations are replaced by contaminated values; see [5] and Subsection 1.6.1 of [12]. Thus, the OLS method suffers a lack of robustness towards outliers, indicating that a single extreme observation could have a large impact on the whole estimation. This motivates researchers to replace the OLS method with other methods that are robust against outliers; see, for example, the references in [3,10,31,33–37], to name just a few.

It is worth mentioning that by rewriting $\hat{\beta}_1$ as a weighted average, Siegel [34] suggested the repeated median (RM) estimator for β_1 given by

$$\tilde{\beta}_1 = \underset{1\le i\le n}{\mathrm{med}}\ \underset{j\neq i}{\mathrm{med}} \frac{y_i - y_j}{x_i - x_j}.$$

Siegel [34] showed when all x_i's are distinct, $\tilde{\beta}_1$ has a finite-sample breakdown point of $\lfloor n/2 \rfloor / n$ and an asymptotic breakdown point of 50%. Here, the asymptotic breakdown point is the limit of the finite-sample breakdown point as the sample size goes to infinity.

As will be shown, it is reasonable to use the weighted median [6] with powered weights instead of the inside conventional median in the RM estimator. Thus, we propose the power-weighted repeated medians (PWRM) estimator which uses the inside weighted median with powered weights. The weighted median with powered weights is easily computed by using the weighted empirical CDF provided in [7,8]. Using convex hull geometry, we also provide explicit lower and upper bounds for the finite-sample breakdown point of the proposed method. It is worth pointing out that the PWRM has a tuning parameter that deals with the trade-offs between relative efficiency and robustness red(breakdown point) [21]. The idea of this tuning parameter is similar in spirit to the tuning constants of Huber's M-estimation [15] and the generalized Kullback-Leibler divergence [25]. With a special choice of the tuning parameter, the PWRM becomes the conventional RM estimator with a breakdown point of 50%. Numerical results showed that with respect to the ratio of the generalized mean square errors (MSEs), the PWRM method performs well in a manner that is approximately comparable with the OLS method, whereas it is far superior when the data contain outliers. In addition, it is shown that by linearizing the CDF of the distribution, the PWRM method can be used to obtain robust parameter estimations for the Weibull and Birnbaum-Saunders distributions that are commonly used in both reliability and survival analysis.

The remainder of this paper is organized as follows. In Section 2, we begin by providing a brief overview of the RM estimator, propose robust explicit estimators for the regression coefficients in simple linear regression, and then conduct simulations to compare the performance of various estimators under consideration. In Section 3, we investigate the lower and upper bounds of the finite-sample breakdown point of the proposed method. We apply the proposed method for obtaining novel robust parameter estimations for the Weibull and Birnbaum-Saunders distributions in Section 4. Two real-data examples are provided for illustrative purposes in Section 5. We provide several concluding remarks in Section 6 with a main result of the finite-sample breakdown point deferred to Appendix.

2. Robust explicit parameter estimations

In this section, we first review the RM estimators for the regression coefficients in simple linear regression and then employ the power-weighted repeated medians technique to develop the PWRM as an alternative in Section 2.1. We carry out simulation studies to compare the finite-sample performance of various estimators in Section 2.2.

2.1. Power-weighted repeated medians

To estimate β_1, Siegel [34] considered the repeated median slope estimator given by

$$\tilde{\beta}_1 = \operatorname*{med}_{1\le i\le n} \operatorname*{med}_{j\neq i} \frac{y_i - y_j}{x_i - x_j}, \tag{3}$$

which behaves like a U-statistic [13] except that nested medians are used instead of the overall mean. We can also obtain the RM estimator for β_1 as follows. We observe that after tedious but simple algebra manipulations, the term S_{xy} in Equation (2) can be equivalently rewritten as

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)(y_i - y_j),$$

which shows that $\hat{\beta}_1$ can be re-expressed as

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)(y_i - y_j)}{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)^2} \\ &= \sum_{i=1}^{n}\sum_{j=1(j\neq i)}^{n}\left[\frac{(x_i - x_j)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_j)^2}\left(\frac{y_i - y_j}{x_i - x_j}\right)\right]\end{aligned}$$

$$= \sum_{i=1}^{n} \sum_{j=1(j\neq i)}^{n} w_{ij} \left(\frac{y_i - y_j}{x_i - x_j} \right), \tag{4}$$

where

$$w_{ij} = \frac{|x_i - x_j|^2}{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|^2} = \frac{|x_i - x_j|^2}{\sum_{i=1}^{n} \sum_{j=1(j\neq i)}^{n} |x_i - x_j|^2}.$$

Thus, we can view $\hat{\beta}_1$ as the weighted mean of observations $(y_i - y_j)/(x_i - x_j)$'s with weights w_{ij}'s. In particular, by assuming that the weight w_{ij} is given by

$$w_{ij} = \frac{|x_i - x_j|^p}{\sum_{i=1}^{n} \sum_{j=1(j\neq i)}^{n} |x_i - x_j|^p},$$

we observe that when $p \approx 0$, the weights w_{ij} can be approximated by $1/\{n(n-1)\}$, resulting in the following approximated estimator for β [34] given by

$$\hat{\beta}_1 \approx \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1(j\neq i)}^{n} \left(\frac{y_i - y_j}{x_i - x_j} \right) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n-1} \sum_{j=1(j\neq i)}^{n} \left(\frac{y_i - y_j}{x_i - x_j} \right).$$

Using this approximation and simply replacing the mean parts $\left(\frac{1}{n}\sum_{i=1}^{n} \text{ and } \frac{1}{n-1}\sum_{j\neq i}\right)$ with the nested medians, we obtain the following RM estimator of β_1 given by

$$\tilde{\beta}_1 = \operatorname*{med}_{1\le i\le n} \operatorname*{med}_{j\neq i} \frac{y_i - y_j}{x_i - x_j}.$$

After $\hat{\beta}_1$ is estimated, one can estimate β_0 using the hierarchical structure given by

$$\tilde{\beta}_0 = \operatorname*{med}_{1\le i\le n} \left(y_i - \hat{\beta}_1 x_i \right). \tag{5}$$

One can also estimate β_0 using

$$\tilde{\beta}_0 = \operatorname*{med}_{1\le i\le n} \operatorname*{med}_{j\neq i} \frac{x_i y_j - x_j y_i}{x_i - x_j}. \tag{6}$$

The estimator in Equation (5) has the same asymptotic variance as the one in Equation (6), while it is computationally more economical than the latter. For a more detailed discussion, see, for example [14]. Thus, in this paper, we only use the estimator in Equation (5) for estimating β_0 in the simple linear regression model in Equation (1).

Considering the weights w_{ij} in (4) and the above relation with the RM estimator, we are motivated to use the weighted median instead of the conventional median in the inside of the RM method. In doing so, we have

$$\hat{\beta}_1 = \underset{1\le i\le n}{\text{med}}\ \underset{j\neq i}{\text{wmed}} \left(\frac{y_i - y_j}{x_i - x_j} \middle| \mathbf{w}_i^* \right), \tag{7}$$

where $\mathbf{w}_i^* = (w_{i1}, \ldots, w_{i,i-1}, w_{i,i+1} \ldots, w_{in})$. For each of $i = 1, 2, \ldots, n$, we may choose the weights w_{ij}'s, such that

$$w_{ij} = \frac{|x_i - x_j|^p}{\sum_{j=1(j\neq i)}^{n} |x_i - x_j|^p}, \tag{8}$$

where $j = 1, 2, \ldots, i-1, i+1, \ldots, n$. Since the denominator in Equation (8) is just a normalizing constant, it should satisfy

$$\sum_{j=1(j\neq i)}^{n} w_{ij} = 1$$

for each of $i = 1, 2, \ldots, n$. Without loss of generality, we denote (8) by

$$w_{ij} \propto |x_i - x_j|^p \tag{9}$$

for $j = 1, 2, \ldots, i-1, i+1, \ldots, n$. Note that the constant p in Equation (9) plays like a tuning constant that evaluates a trade-off between efficiency and robustness. Of particular note is that as $p \to 0$, the PWRM estimator in Equation (7) becomes the RM in Equation (3) having a breakdown point of 50%.

By following the work of [7,8], we could employ the weighted empirical CDF to calculate the weighted median in Equation (7). For brevity, we assume that there are n observations $\{x_1, x_2, \ldots, x_n\}$ with their corresponding weights, denoted by $\{w_1, w_2, \ldots, w_n\}$, respectively. Let $\mathbb{I}(A)$ be the indicator function of an event A, such that $\mathbb{I}(A) = 1$ when A is true and $\mathbb{I}(A) = 0$ when A is false. Then the weighted empirical CDF is given by

$$\hat{F}_w(t) = \sum_{j=1}^{n} w_j \mathbb{I}(x_j \le t),$$

where $w_j > 0$ and $\sum_{j=1}^{n} w_j = 1$. This weighted empirical CDF includes the conventional empirical CDF as a special case, when $w_j = 1/n$. If we denote $\hat{Q}_{wL}(\xi) = \inf\{u : \hat{F}_w(t) \ge \xi\}$ and $\hat{Q}_{wR}(\xi) = \inf\{u : \hat{F}_w(t) > \xi\}$, then the weighted median is obtained by $\hat{Q}_{wL}(1/2) = \inf\{u : \hat{F}_w(t) \ge 1/2\}$ or $\hat{Q}_{wR}(1/2) = \inf\{u : \hat{F}_w(t) > 1/2\}$ with $\hat{Q}_{wL}(1/2) \le \hat{Q}_{wR}(1/2)$. However, unlike the work of [8], who used $\hat{Q}_{wL}(1/2)$ as the weighted median, we suggest the average of the two medians given by $\{\hat{Q}_{wL}(1/2) + \hat{Q}_{wR}(1/2)\}/2$.

2.2. Simulation studies

For illustrative purposes, we here carry out a small simulation to compare the finite-sample performance of various estimators under consideration of the absence and presence of data contamination. We generate $y_i = 1 + 2x_i + \epsilon_i$, where the values of x_i are given by $x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$, and ϵ_i's are randomly generated from $N(0, 1)$. We iterate this experiment $I = 10^5$ times.

It is known that a direct comparison of two estimators (say, $\hat{\theta}_1$ and $\hat{\theta}_2$) is to compare their variances (Section 2.2 of [19]) given by

$$\mathrm{RE}(\hat{\theta}_2 \mid \hat{\theta}_1) = \frac{\mathrm{Var}(\hat{\theta}_1)}{\mathrm{Var}(\hat{\theta}_2)} \times 100\%,$$

which is often referred to as the relative efficiency (RE) of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$. Since the above RE compares the variances of the two univariate parameter estimators, it may be inappropriate to directly apply this RE to our estimation framework, in which we focus on the bivariate parameter estimations, namely the intercept and slope together.

When dealing with multivariate estimators, as is the case in our study, it is desirable to construct a single numerical measure that represents the variation of the multivariate estimators [28,29]. A common method is to use the determinant of the estimated variance-covariance matrix, so-called the generalized variance [17]. When there is no data contamination, the intercept and slope estimators are unbiased. However, when there is data contamination, they are not unbiased anymore. To tackle the biasedness issue of an estimator in the presence of data contamination, we may use the MSE as an alternative and then consider the determinant of the following matrix with the true intercept and slope being $\beta_0 = 1$ and $\beta_1 = 2$ given by

$$\mathbf{S}_M = \begin{pmatrix} \frac{1}{I}\sum_{k=1}^{I}(\hat{\beta}_{0,M,k} - 1)^2 & \frac{1}{I}\sum_{k=1}^{I}(\hat{\beta}_{0,M,k} - 1)(\hat{\beta}_{1,M,k} - 2) \\ \frac{1}{I}\sum_{k=1}^{I}(\hat{\beta}_{0,M,k} - 1)(\hat{\beta}_{1,M,k} - 2) & \frac{1}{I}\sum_{k=1}^{I}(\hat{\beta}_{1,M,k} - 2)^2 \end{pmatrix},$$

where M denotes the estimation method, k indicates the kth experiment, and I is the total number of iterations in the simulation. The determinant of $\mathbf{S}_M$ is often termed as the generalized MSE. Then the ratio of the generalized MSEs allows us to compare the performance of the two different estimation methods based on their variance and bias as well. Then the RE (%) using the method M with the corresponding sample is defined as

$$\mathrm{RE}(M \mid \mathrm{OLS}) = \frac{\det(\mathbf{S}_{\mathrm{OLS}}) \ \text{with no contamination}}{\det(\mathbf{S}_M) \ \text{with corresponding sample}} \times 100\%.$$

To further investigate the effect of data contamination, we artificially contaminate one response value by setting $y_1 = 100$. Simulation results are summarized in Table 1. As expected, the OLS method performs the best when there is no data contamination, whereas it is the worst in the presence of data contamination. We provide the scatter plots of the bivariate parameter estimates using the OLS, RM, PWRM ($p = 1$), and PWRM ($p = 2$) methods in Figure 1, where we only use two hundred resulting data points to avoid extreme cluttering with all the 10^5 observations. The circle denotes the parameter estimation with no contamination and the cross denotes that with contamination. We observe from this figure that the OLS method is seriously biased when there is data contamination, but other robust methods are not. We also observe that PWRM method generally outperforms the conventional RM and that the improvement is quite noticeable, especially when there exists data contamination. Finally, we note that the power p in the weights ($w_{ij} \propto |x_i - x_j|^p, j \neq i$) does not change the results much in this case for the values of p considered here.

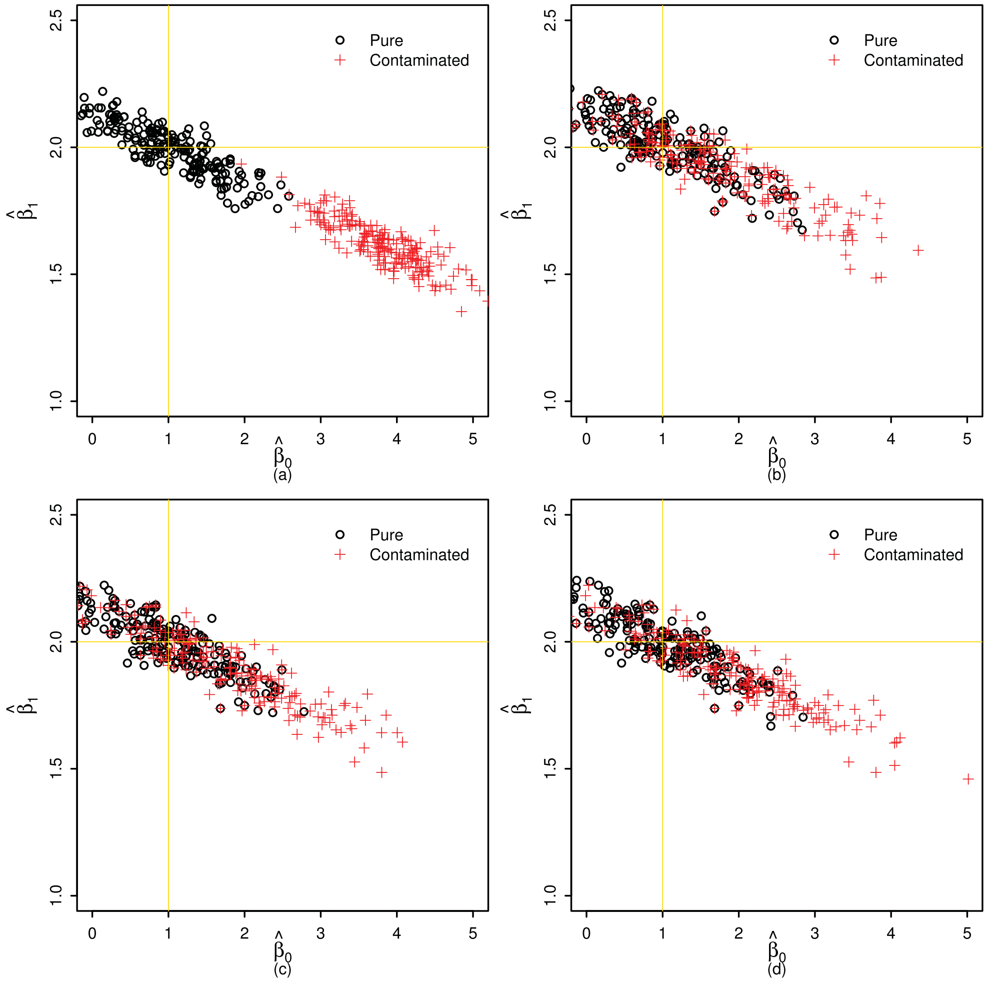

Figure 1. The parameter estimates obtained using (a) OLS, (b) RM, (c) PWRM ($p = 1$), and (d) PWRM ($p = 2$).

Table 1. The values of the generalized MSEs and REs.

	OLS	RM	PWRM ($p = 1$)	PWRM ($p = 2$)	PWRM ($p = 1/2$)
No contamination					
MSE	0.00120	0.00304	0.00230	0.00225	0.00235
RE(%)	100.00	39.45	52.25	53.48	51.16
Contaminated sample					
MSE	0.02124	0.00675	0.00571	0.00651	0.00550
RE(%)	5.66	17.80	21.04	18.43	21.83

3. Finite-sample breakdown point

In this section, we provide the lower and upper bounds of the finite-sample breakdown point for the PWRM estimators for the regression coefficients in simple linear regression. Unlike the conventional RM, its finite-sample breakdown point depends on the weights and it has lower and upper bounds. For more details on the finite-sample breakdown points, one can refer to Definition 2 of [27]. Since geometry plays an important role in obtaining these bounds of the breakdown points, we briefly introduce the geometry originally provided by [7].

For notational convenience, we assume that there are n observations $\{x_1, x_2, \ldots, x_n\}$ with their corresponding weights $\{w_1, w_2, \ldots, w_n\}$, respectively. First, we consider the upper bound as follows. Let $w_{(j)}$'s be the values of the order statistics, such that $w_{(1)} \leq w_{(2)} \leq \cdots \leq w_{(n)}$. We assume that the observations, x_i's, are contaminated one by one starting from the one with the smallest weight in a pattern, such that the observation corresponding to its weight $w_{(1)}$ is contaminated first, the one with $w_{(2)}$ next, and so on. Let $b_k = \sum_{j=1}^{k} w_{(j)}$. Then, considering that the conventional RM is resistant when there is less than 50% contamination, the weighted median is resistant while $b_k < 1/2$. Thus, the upper bound is obtained as $\epsilon_U = \sum_{k=1}^{n} \mathbb{I}(b_k < 1/2)/n$. Next, we consider the lower bound as follows. We assume that the observations, x_i's, are contaminated one by one starting from the one with the largest weight. Then the observation with its weight $w_{(n)}$ is contaminated first, the one with $w_{(n-1)}$ next, and so on. Let $a_k = \sum_{j=1}^{k} w_{(n+1-j)}$. Similar to the case of the upper bound, the minimum breakdown point is obtained as $\epsilon_L = \sum_{k=1}^{n} \mathbb{I}(a_k < 1/2)/n$.

Now suppose that we have weights w_{ij}'s as in Equation (9). Then the lower and upper bounds of the finite-sample breakdown point are given by

$$\epsilon_L = \frac{\min_{1\leq i\leq n} \sum_{k=1}^{n-1} \mathbb{I}(\sum_{\ell=1}^{k} w_{i(n-\ell)} < 1/2)}{n-1}$$

and

$$\epsilon_U = \frac{\max_{1\leq i\leq n} \sum_{k=1}^{n-1} \mathbb{I}(\sum_{\ell=1}^{k} w_{i(\ell)} < 1/2)}{n-1},$$

respectively, where $w_{i(\ell)}$ be the values of the order statistics such that $0 < w_{i(1)} \leq w_{i(2)} \leq \ldots \leq w_{i(n-1)}$ for each of $i = 1, 2, \ldots, n$. For a more detailed proof, we refer the interested reader to Theorem A.2.

4. Practical applications

In this section, we apply the PWRM method for obtaining novel robust estimators for the Weibull distribution in Section 4.1 and the Birnbaum-Saunders distribution in Section 4.2 by linearizing their respective CDFs.

4.1. Weibull distribution

A random variable X is said to follow the two-parameter Weibull distribution if its CDF can be written as

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{\theta}\right)^{\alpha}\right\},$$

where $\theta > 0$ and $\alpha > 0$ represent the scale and shape parameters, respectively. It can be easily shown that the CDF of a Weibull random variable can be linearized as

$$\log\left(-\log(1-p_i)\right) = -\alpha\log\theta + \alpha\log x_{(i)}, \quad i = 1, \ldots, n, \tag{10}$$

Table 2. The values of the generalized MSEs and REs for the Weibull distribution.

	OLS	RM	PWRM ($p=1$)	PWRM ($p=1/2$)	PWRM ($p=1/4$)
No contamination					
MSE	0.00246	0.00471	0.00272	0.00303	0.00354
RE(%)	100.00	52.39	90.61	81.24	69.61
With contamination ($\delta = 10$)					
MSE	0.02628	0.00515	0.00436	0.00406	0.00432
RE(%)	9.38	47.88	56.48	60.74	57.018867

where $x_{(i)}$'s are the values of the order statistics, such that $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$, and $p_i = F(x_{(i)})$, which can be easily estimated by using the plotting position, an increasing step function jumping at $x_{(i)}$. There are several different versions of the plotting positions in the literature. One of the most popular plotting positions is

$$p_i = \begin{cases} \dfrac{i-3/8}{n+1/4} & \text{for } n \le 10 \\ \dfrac{i-1/2}{n} & \text{for } n \ge 11 \end{cases}, \tag{11}$$

which is due to [2,38]. For more details of a comparison of these plotting positions, we refer the interested reader to [22].

By denoting $y_i^* = \log(-\log(1-p_i))$, $x_i^* = \log x_{(i)}$, $\beta_0^* = -\alpha \log \theta$, and $\beta_1^* = \alpha$, we can rewrite Equation (10) as

$$y_i^* = \beta_0^* + \beta_1^* x_i^*, \quad i = 1, \ldots, n.$$

Thereafter, based on observations $\{(x_1^*, y_2^*), \ldots (x_n^*, y_n^*)\}$, we can easily calculate the estimate of β_1^*, denoted by $\hat{\beta}_1^*$, by using the RM estimator in Equation (3) and the PWRM estimator in Equation (7). After $\hat{\beta}_1^*$ is obtained, we can estimate $\hat{\beta}_0^*$ easily using Equation (5). After $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ are obtained, we obtain the original parameter estimators by reparametrizing as

$$\hat{\alpha} = \hat{\beta}_1^* \quad \text{and} \quad \hat{\theta} = \mathrm{e}^{-\hat{\beta}_0^*/\hat{\beta}_1^*}.$$

We carry out a Monte Carlo simulation to compare the performance of the OLS, RM and PWRM methods. For the PWRM method, we consider $p = \{1, 1/2, 1/4\}$. We generate $n = 20$ observations from the Weibull distribution with $\theta = 1$ and $\alpha = 2$. We repeat this experiment 10^5 times and summarize the results in Table 2. It is known that some estimators can break down under a single bad observation. For more details, one can refer to Section 11.2 of Huber and Ronchetti [16]. Thus, in this paper, to investigate the sensitivity of the considered methods in the presence of data contamination, we replace the first observation with a large outlier (say, 10 in this simulation) and rerun the above simulation again. Numerical results are also reported in Table 2. We observe that the PWRM method always outperforms the RM method for all the values of $p = \{1, 1/2, 1/4\}$. Of particular note is that when $p = 1$, the RE is 90.61% with no contaminated data and is 56.48% with contaminated data. However, the OLS only achieves the RE of 9.38%, due to its lack of robustness. Thus, we prefer the use of the PWRM method with $p = 1$ in this Weibull case.

Table 3. The values of the generalized MSEs and REs for the Birnbaum-Saunders distribution.

	OLS	RM	PWRM ($p = 1/2$)	PWRM ($p = 1/4$)	PWRM ($p = 1/8$)
No contamination					
MSE	0.01178	0.03319	0.02703	0.02734	0.02931
RE(%)	100.00	35.49	43.57	43.09	40.20
With contamination ($\delta = 100$)					
MSE	0.33387	0.06092	0.16732	0.07466	0.06358
RE(%)	3.53	19.34	7.04	15.78	18.53

4.2. Birnbaum-Saunders distribution

A random variable X is said to follow the two-parameter Birnbaum-Saunders distribution if its CDF can be written as

$$F(x) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right],$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution, and $\alpha > 0$ and $\beta > 0$ represent the shape and scale parameters, respectively. It can be easily shown that the CDF of the Birnbaum-Saunders distribution can be linearized as

$$\sqrt{x_{(i)}}\,\Phi^{-1}(p_i) = -\frac{\sqrt{\beta}}{\alpha} + \frac{1}{\alpha\sqrt{\beta}} \cdot x_{(i)}, \tag{12}$$

where $x_{(i)}$'s are the values of the order statistics. By letting $y_i^* = \sqrt{x_{(i)}}\,\Phi^{-1}(p_i)$, $x_i^* = x_{(i)}$, $\beta_0^* = -\sqrt{\beta}/\alpha$, and $\beta_1^* = 1/(\alpha\sqrt{\beta})$, we can rewrite Equation (12) as

$$y_i^* = \beta_0^* + \beta_1^* x_i^*.$$

In a similar manner as done for the Weibull case, we first obtain the PWRM estimators of the regression coefficients above, denoted by $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$, respectively. We then obtain the original parameter estimators of the Birnbaum-Saunders distribution by reparametrizing as

$$\hat{\alpha} = \frac{1}{\sqrt{-\hat{\beta}_0^* \hat{\beta}_1^*}} \quad \text{and} \quad \hat{\beta} = -\frac{\hat{\beta}_0^*}{\hat{\beta}_1^*}.$$

We again conduct a Monte Carlo simulation to compare the performance of the considered methods. For the PWRM method, we consider $p = \{1/2, 1/4, 1/8\}$. We generate $n = 20$ observations from the Birnbaum-Saunders distribution with $\alpha = 2$ and $\beta = 1$. We repeat this experiment 10^5 times and summarize the results in Table 3. To investigate the effect of a contaminated value, we replace the first observation with 100 and rerun the above simulation. Numerical results are also provided in Table 3. We observe that the PWRM method outperforms the RM method in the absence of data contamination and that the PWRM method with $p = 1/8$ behaves similarly to the RM method in the presence of data contamination. Consequently, we suggest the use of the PWRM method with $p = 1/8$ for the Birnbaum-Saunders case.

Table 4. The data set from Example 8.1 of [18].

17.88	28.92	33.00	41.52	42.12	45.60	48.48	51.84
51.96	54.12	55.56	67.80	68.64	68.64	68.88	84.12
93.12	98.64	105.12	105.84	127.92	128.04	173.40	

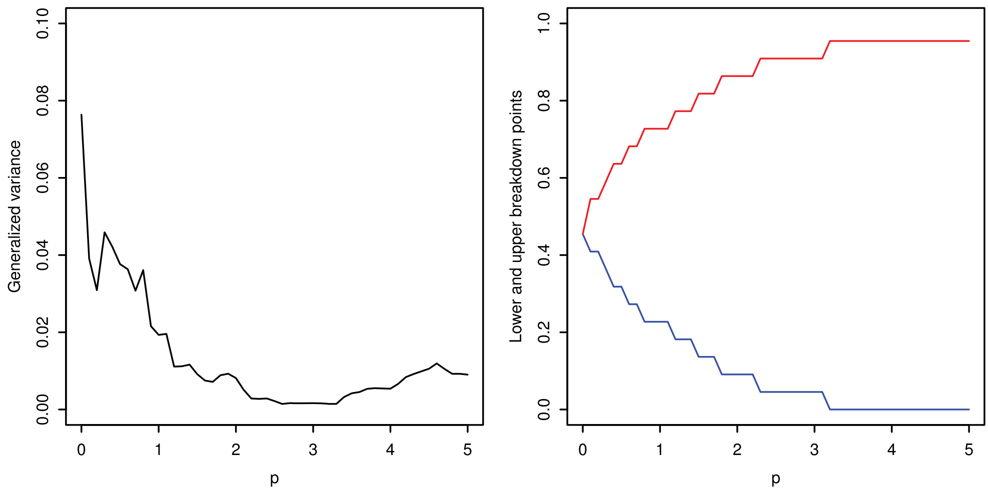

Figure 2. (a) Generalized variance from LOO cross-validation and (b) Lower and upper breakdown points with the Weibull Data.

5. Illustrative examples

In this section, we illustrate the practical applications of the PWRM method for estimating the unknown parameters of the Weibull distribution in Section 5.1 and the Birnbaum Saunders distribution in Section 5.2 through two-real data examples as discussed below.

5.1. Weibull data

We consider a real-data example originally provided by [20]. This data set is used frequently for the Weibull model analysis; see, for example, Example 8.1 of [18,24]. Twenty-three ball bearings were under endurance tests and the numbers to failures (in millions of revolutions) from these tests are provided in Table 4.

As aforementioned, the PWRM has the tuning parameter p which deals with the trade-offs between efficiency and robustness [21]. To determine the tuning parameter p, we can perform the leave-one-out cross-validation (LOOCV) method as follows. In this example, we have $n = 23$ training sets of size 22. For each of n training sets with p given, we first estimate the shape and scale parameters and then calculate the generalized variance [17] based on the n pairs of shape and scale parameter estimates. We repeated the above LOOCV method for each of $p = 0(0.1)5$. The results are summarized in Figure 2(a). As will be detailed, in Figure 2(b), we also provide the lower and upper breakdown points which are calculated based on Theorem A.2 for each of $p = 0(0.1)5$. We observe from Figure 2

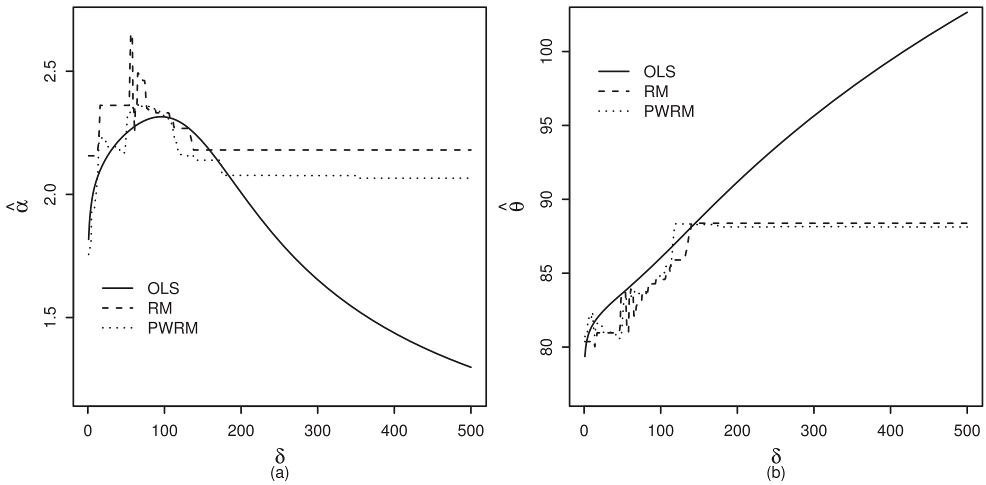

Figure 3. Estimates of the Weibull parameters: (a) shape and (b) scale with $p = 1$.

that the optimal value of p based on the value of generalized variance is around three. However, when $p = 3$, the lower breakdown point is too small. When p is between one and two, there is a small increase in the value of the generalized variance but the value of the lower breakdown point increase significantly. Thus, we recommend the tuning parameter p between one and two.

To investigate the sensitivity to an outlier or data contamination, one can consider the influence function [4,11] or ϵ-influence function [25], etc. It is well known that an estimator does not break down as an observation in the sample becomes arbitrarily large when an influence function is bounded. But, in this case, it may be extremely difficult or impossible to obtain the conventional influence function. Thus, we consider an empirical approach analogous to the underlying idea of the influence function by investigating how the parameter estimates are affected by the contamination level. For more details on the implementation of this idea, see Figure 2 of [28], Figures 7 and 8 of [25], 5 of [26], Figures 1 and 2 of [30] Figure 3 of [27]. To this end, we replace the first observation (17.88) with δ, whose value ranges from 1 to 500 in increments of 1. We draw the plot of the parameter estimates versus δ in Figure 3 ($p = 1$) and Figure 4 ($p = 2$). Note that each convex hull is constructed using $a_{ik} = \sum_{\ell=1}^{k} w_{i(n-\ell)}$ and $b_{ik} = \sum_{\ell=1}^{k} w_{i(\ell)}$ for $k = 1, 2, \ldots, n-1$ as seen in Figure A1, where $w_{i(\ell)}$ are the values of the order statistics of w_{ij} such that $0 < w_{i(1)} \leq w_{i(2)} \leq \cdots \leq w_{i(n-1)}$. Thus, for $i = 1, 2, \ldots, 23$, we can draw the twenty-three convex hulls as in Figure 5. Using Theorem A.2 (the largest convex hull), we can obtain the lower and upper bounds for the finite-sample breakdown point of the PWRM method. In particular, when $p = 1$, we have 5/22 (lower bound) and 16/22 (upper bound), indicating that its finite-sample breakdown point should be between 22.7% and 72.7%. With $p = 2$, we have 2/22 (lower bound) and 19/22 (upper bound). Thus, the finite-sample breakdown point should be between 9.1% and 86.4%. With $p = 3$, we have 1/22 (lower bound) and 20/22 (upper bound) which give 4.5% and 90.9%.

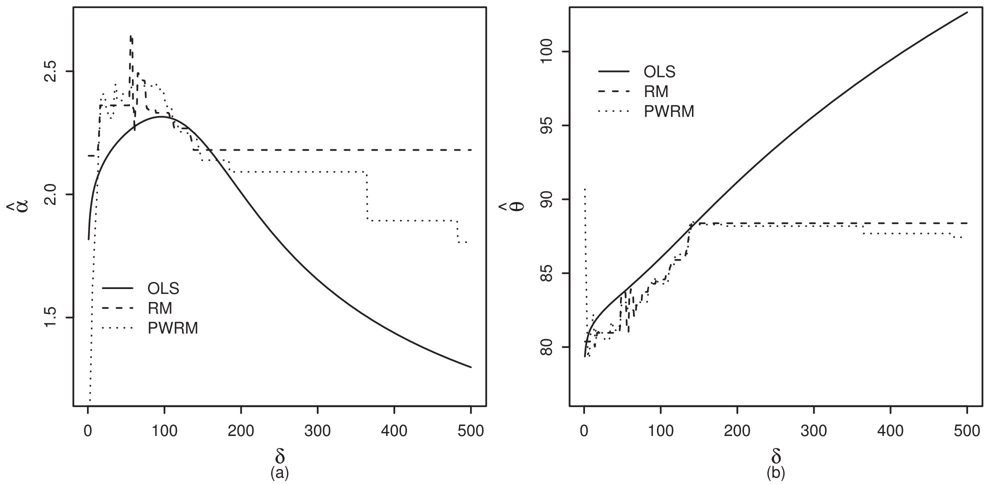

Figure 4. Estimates of the Weibull parameters: (a) shape and (b) scale with $p = 2$.

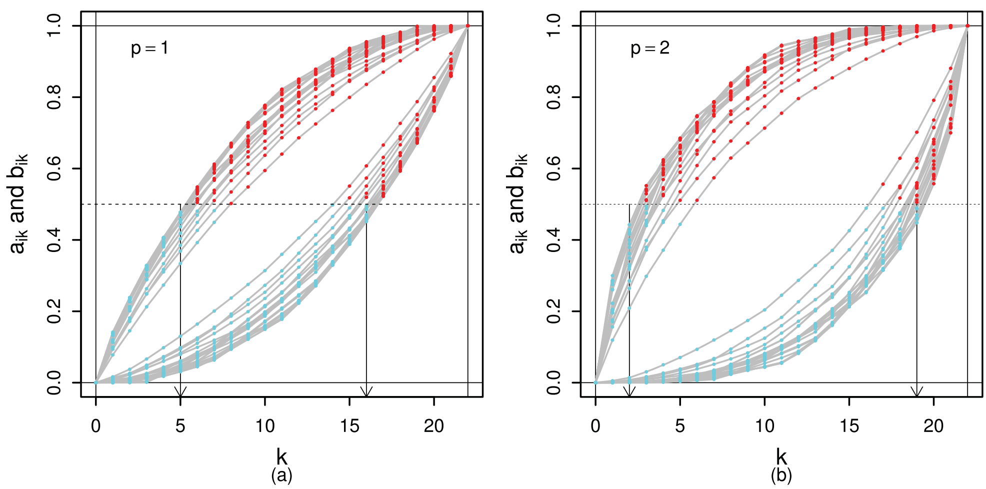

Figure 5. Convex hulls using the Weibull data set: (a) $p = 1$ and (b) $p = 2$.

5.2. Birnbaum-Saunders data

We consider the data set from [1] about fatigue life (in 10^5 cycles) to failure of aluminum coupons. This data set is often used for the Birnbaum-Saunders model; see, for example, Table I of [23]. These data are also provided in Table 5.

To determine the tuning parameter p, we perform the LOOCV method as done before. In this example, we have $n = 101$ training sets of size 100. Similar to the case of the Weibull data, we repeat this LOOCV method for each of $p = 0(0.1)5$ and calculate the generalized variance. The results are summarized in Figure 6. We observe from Figure 6 that the optimal value of p based on the value of generalized variance is around one. Thus, we recommend the tuning parameter p between half and two.

Table 5. The data set from Table I of [23].

3.70	7.06	7.16	7.46	7.85	7.97	8.44	8.55	8.58	8.86	8.86
9.30	9.60	9.88	9.90	10.00	10.10	10.16	10.18	10.20	10.55	10.85
11.02	11.02	11.08	11.15	11.20	11.34	11.40	11.99	12.00	12.00	12.03
12.22	12.35	12.38	12.52	12.58	12.62	12.69	12.70	12.90	12.93	13.00
13.10	13.13	13.15	13.30	13.55	13.90	14.16	14.19	14.20	14.20	14.50
14.52	14.75	14.78	14.81	14.85	15.02	15.05	15.13	15.22	15.22	15.30
15.40	15.60	15.67	15.78	15.94	16.02	16.04	16.08	16.30	16.42	16.74
17.30	17.50	17.50	17.63	17.68	17.81	17.82	17.92	18.20	18.68	18.81
18.90	18.93	18.95	19.10	19.23	19.40	19.45	20.23	21.00	21.30	22.15
22.68	24.40									

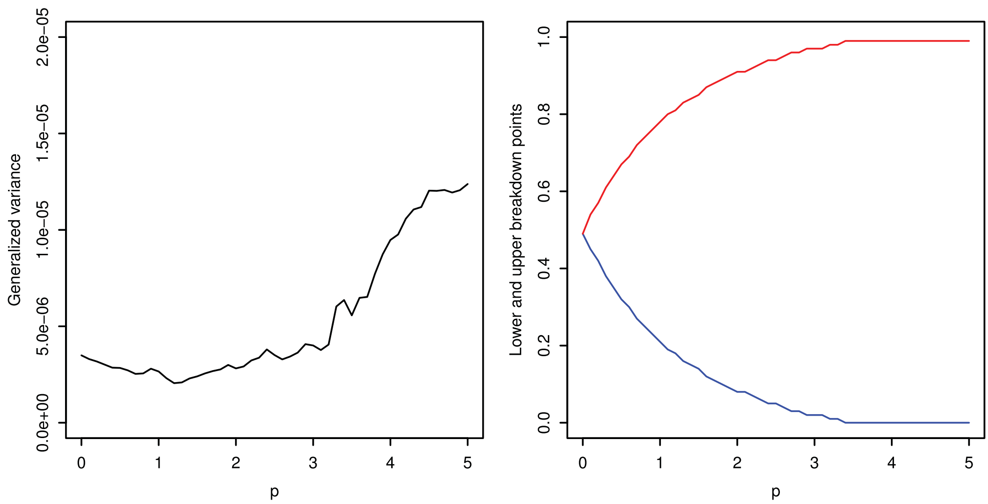

Figure 6. (a) Generalized variance from LOO cross-validation and (b) Lower and upper breakdown points with the Birnbaum-Saunders Data.

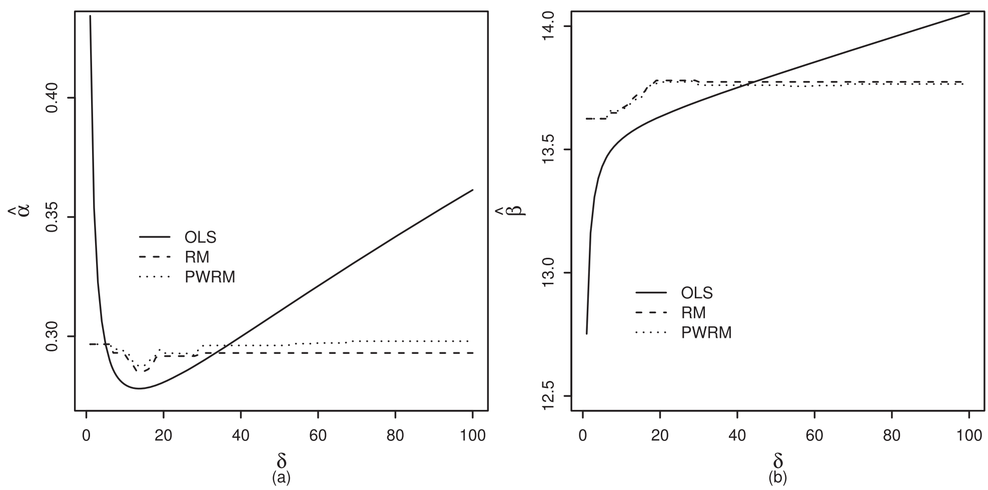

Figure 7. Estimates of the Birnbaum-Saunders parameters: (a) shape and (b) scale with $p = 1/2$.

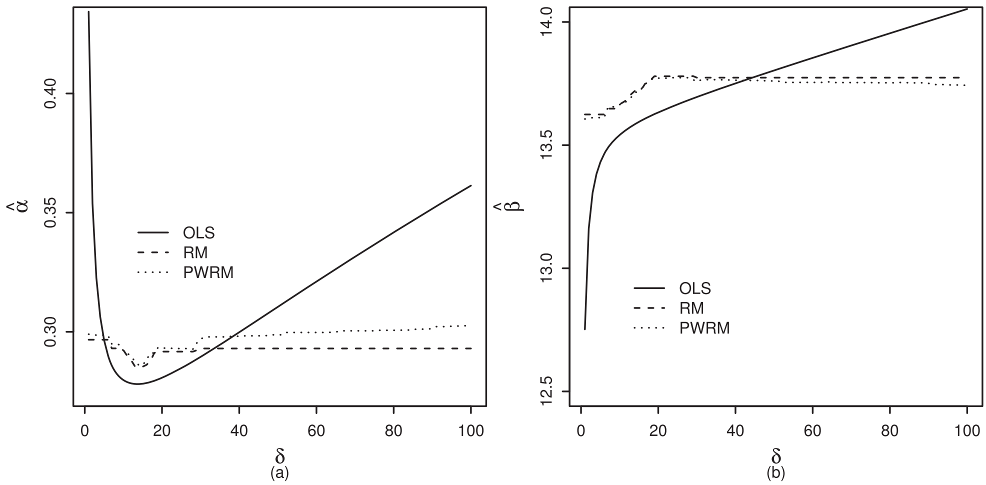

Figure 8. Estimates of the Birnbaum-Saunders parameters: (a) shape and (b) scale with $p = 1$.

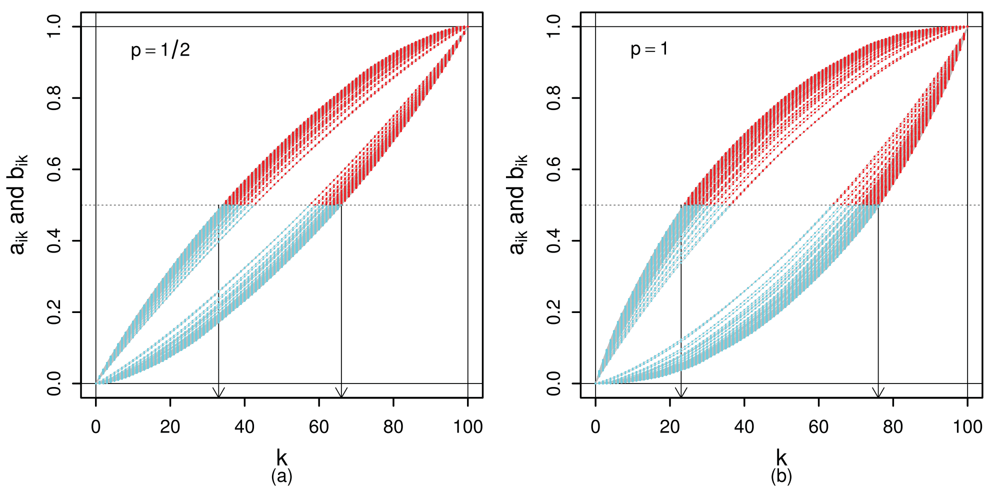

Figure 9. Convex hulls using the Birnbaum-Saunders data set: (a) $p = 1/2$ and (b) $p = 1$.

To investigate the sensitivity to an outlier or data contamination, we replace the first observation (3.70) with δ, whose value ranges from 1 to 100 in increments of 1. We draw the plot of the parameter estimates versus δ in Figure 7 ($p = 1/2$) and Figure 8 ($p = 1$) and the convex hulls in Figure 9 for each of $i = 1, 2, \ldots, 101$. We observe that when $p = 1/2$, we have 33/100 (lower bound) and 66/100 (upper bound), indicating that the finite-sample breakdown point should be between 33% and 66%. Also, with $p = 1$, we have 23/100 (lower bound) and 76/100 (upper bound), indicating that the finite-sample breakdown point should be between 23% and 76%.

6. Concluding remarks

In this paper, we have proposed a new method of robust parameter estimation of the regression coefficients in simple linear regression based on the power-weighted repeated medians technique. The proposed method not only has a single tuning constant that allows us to deal with the trade-offs between relative efficiency and robustness measured by breakdown points but also includes the repeated medians as a special case with the tuning parameter being zero. When the power parameter is zero, the lower breakdown point is highest and the generalized variance tends to have a larger value which reduces the relative efficiency. Thus, we recommend the use of a desired p based on the cross validation along with the breakdown point which corresponds to p.

In addition, we have derived the lower and upper bounds of the finite-sample breakdown point of the proposed method. It has been shown that through linearizing the cumulative distribution function technique, the proposed method can be generalized to derive novel robust parameter estimations for the Weibull and Birnbaum–Saunders distributions that are commonly used in both reliability and survival analysis. Numerical results from simulation studies and real data examples demonstrated the proposed method performs well in a manner that is approximately comparable with existing methods, whereas it is far superior when the data contain outliers that occur frequently in practice.

Finally, the R codes used for this research can be found on the web below:

https://github.com/AppliedStat/R-code/tree/master/2024a

Acknowledgments

The authors truly appreciate the valuable comments from two anonymous referees which led to a great improvement in our work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work of Professor Park was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A2C1091319). The work of Professor Gao was supported by National Science Foundation of China Grant No. 72104020 and China Postdoctoral Science Foundation, No. 15 Special Fund (In-Station), Grant No. 2022T150044. The work of Dr. Min Wang was partially supported by the Internal Research Awards (INTRA), UTSA program from the Vice President for Research, Economic Development, and Knowledge Enterprise at the University of Texas at San Antonio.

ORCID

Chanseok Park http://orcid.org/0000-0002-2208-3498
Min Wang http://orcid.org/0000-0002-9233-7844

References

[1] Z.W. Birnbaum and S.C. Saunders, *Estimation for a family of life distributions with applications to fatigue*, J. Appl. Probab. 6 (1969), pp. 328–347.
[2] G. Blom, *Statistical Estimates and Transformed Beta Variates*, Wiley, New York, 1958.
[3] K. Boudt, D. Caliskan, and C. Croux, *Robust explicit estimators of Weibull parameters*, Metrika 73 (2011), pp. 187–209.

[4] R.D. Cook and S. Weisberg, *Characterizations of an empirical influence function for detecting influential cases in regression*, Technometrics 22 (1980), pp. 495–508.
[5] D.L. Donoho and P.J. Huber, *The notion of breakdown point*, in *A Festschrift for Erich L. Lehmann*, P. Bickel, K. Doksum, and J.L. Hodges, Jr., eds., Wadsworth, Belmont, CA, 1983, pp. 157–184.
[6] F.Y. Edgeworth, *On a new method of reducing observations relating to several quantities*, Lond. Edinb. Dublin Philos. Mag. J. Sci. 25 (1888), pp. 184–191.
[7] X. Gao, *Determination of the optimal facility location based on the minimum distance approach*, Ph.D. diss., Pusan National University, Busan, Korean, 2020.
[8] X. Gao and C. Cui, *A note on the warehouse location problem with data contamination*, RAIRO–Oper. Res. 55 (2021), pp. 1113–1135.
[9] R.L. Graham and F. Frances Yao, *Finding the convex hull of a simple polygon*, J. Algorithms 4 (1983), pp. 324–331.
[10] F.R. Hampel, *The influence curve and its role in robust estimation*, J. Am. Stat. Assoc. 69 (1974), pp. 383–393.
[11] F.R. Hampel, A. Marazzi, E. Ronchetti, P.J. Rousseeuw, W.A. Stahel, and R.E. Welsch, *Handouts for the instructional meeting on robust statistical methods*, in *The 15th European Meeting of Statisticians*, Palermo, Italy. 1982.
[12] T.P. Hettmansperger and J.W. McKean, *Robust Nonparametric Statistical Methods*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL, 2010.
[13] W. Hoeffding, *A class of statistics with asymptotically normal distribution*, Ann. Math. Stat.19 (1948), pp. 293–325.
[14] O. Hossjer, P. Rousseeuw, and I. Ruts, *The repeated median intercept estimator: Influence function and asymptotic normality*, J. Multivar. Anal. 52 (1995), pp. 45–72.
[15] P.J. Huber, *Robust estimation of a location parameter*, Ann. Math. Stat. 35 (1964), pp. 73–101.
[16] P.J. Huber and E.M. Ronchetti, *Robust Statistics*, 2nd ed., John Wiley & Sons, New York, 2009.
[17] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, 6th ed., Prentice-Hall Inc, Englewood Cliffs, NJ, 2007.
[18] L.M. Leemis, *Reliability: Probabilistic Models and Statistical Methods*, 2nd ed., Lawrence M. Leemis, Williamsburg, VA, 2009.
[19] E.L. Lehmann, *Elements of Large-Sample Theory*, Springer, New York, 1999.
[20] J. Lieblein and M. Zelen, *Statistical investigation of the fatigue life of deep groove ball bearings*, J. Res. Bur. Standards 57 (1956), pp. 273–316.
[21] B.G. Lindsay, *Efficiency versus robustness: The case for minimum Hellinger distance and related methods*, Ann. Stat. 22 (1994), pp. 1081–1114.
[22] S.W. Looney and T.R. Gulledge, Jr., *Use of the correlation coefficient with normal probability plots*, Am. Stat. 39 (1985), pp. 75–79.
[23] W.J. Owen and W.J. Padgett, *A Birnbaum-Saunders accelerated life model*, IEEE Trans. Reliab. 49 (2000), pp. 224–229.
[24] C. Park, *Weibullness test and parameter estimation of the three-parameter Weibull model using the sample correlation coefficient*, Int. J. Ind. Eng. Theory Appl. Pract. 24 (2017), pp. 376–391.
[25] C. Park and A. Basu, *The generalized Kullback-Leibler divergence and robust inference*, J. Stat. Comput. Simul. 73 (2003), pp. 311–332.
[26] C. Park and A. Basu, *Minimum disparity inference based on tangent disparities*, Int. J. Inf. Manag. Sci. 22 (2011), pp. 1–25.
[27] C. Park, H. Kim, and M. Wang, *Investigation of finite-sample properties of robust location and scale estimators*, Commun. Stat. Simul. Comput. 51 (2022), pp. 2619–2645.
[28] C. Park and M. Leeds, *A highly efficient robust design under data contamination*, Comput. Ind. Eng.93 (2016), pp. 131–142.
[29] C. Park, L. Ouyang, J.H. Byun, and M. Leeds, *Robust design under normal model departure*, Comput. Ind. Eng. 113 (2017), pp. 206–220.
[30] C. Park, M. Wang, and W.Y. Hwang, *A study on robustness of the paired sample tests*, Ind. Eng. Manag. Syst. 19 (2020), pp. 386–397.
[31] P.J. Rousseeuw, *Least median of squares regression*, J. Am. Stat. Assoc. 79 (1984), pp. 871–880.

[32] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, 2nd ed., Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 2003. Available at https://doi.org/10.1002/0471725382.
[33] P.K. Sen, *Estimates of the regression coefficient based on Kendall's tau*, J. Am. Stat. Assoc. 63 (1968), pp. 1379–1389.
[34] A.F. Siegel, *Robust regression using repeated medians*, Biometrika 69 (1982), pp. 242–244.
[35] H. Theil, *A rank-invariant method of linear and polynomial regression analysis. I.*, Nederl. Akad. Wetensch. Proc. 53 (1950), pp. 386–392.
[36] M. Wang, C. Park, and X. Sun, *Simple robust parameter estimation for the Birnbaum-Saunders distribution*, J. Stat. Distrib. Appl. 2 (2015), pp. 1–11.
[37] M. Wang, J. Zhao, X. Sun, and C. Park, *Robust explicit estimation of the two-parameter Birnbaum-Saunders distribution*, J. Appl. Stat. 40 (2013), pp. 2259–2274.
[38] M.B. Wilk and R. Gnanadesikan, *Probability plotting methods for the analysis of data*, Biometrika 55 (1968), pp. 1–17.
[39] S. Wu and L. Debnath, *Inequalities for convex sequences and their applications*, Comput. Math. Appl. 54 (2007), pp. 525–534.

Appendix

We assume that there are n observations, denoted by $x_1, x_2, \dots, x_n$. Suppose that we estimate the slope parameter using Equation (7) with the corresponding weights given by Equation (8). For the weights, we use w_{ij} in Equation (9). Let $w_{i(\ell)}$'s be the values of the order statistics of w_{ij} such that $0 < w_{i(1)} \le w_{i(2)} \le \cdots \le w_{i(n-1)}$ for each of $i = 1, 2, \dots, n$. In the following lemma, we show that the sequences $a_{ik} = \sum_{\ell=1}^{k} w_{i(n-\ell)}$ and $b_{ik} = \sum_{\ell=1}^{k} w_{i(\ell)}$ for $k = 1, 2, \dots, n-1$ are concave and convex, respectively, for each of $i = 1, 2, \dots, n$. These properties play an important role in understanding the geometry of the lower and upper bounds of the finite-sample breakdown point of the proposed estimator.

Lemma A.1: *The sequences $a_{ik} = \sum_{\ell=1}^{k} w_{i(n-\ell)}$ and $b_{ik} = \sum_{\ell=1}^{k} w_{i(\ell)}$ are both strictly increasing for each of $i = 1, 2, \dots, n$. In addition, $\{(k, a_{ik}) : k = 0, 1, 2, \dots, n-1\}$ with $a_0 = 0$ and $\{(k, b_{ik}) : k = 0, 1, 2, \dots, n-1\}$ with $b_0 = 0$ construct the convex hull of the polygon.*

Proof: This proof is based on the results of [7]. Since $a_{ik} - a_{i,k-1} = w_{i(n-k)} > 0$ and $b_{ik} - b_{i,k-1} = w_{i(k)} > 0$, it is easily seen that a_{ik} and b_{ik} are strictly increasing for each of $i = 1, 2, \dots, n$. Since $2a_{ik} - (a_{i,k-1} + a_{i,k+1}) = w_{i(n-k)} - w_{i(n-k-1)} \ge 0$, it is immediate from Definition 1 of [39] that a_{ik} is concave. Similarly, b_{ik} is convex since $2b_{ik} - (b_{i,k-1} + b_{i,k+1}) = w_{i(k)} - w_{i(k+1)} \le 0$. Thus, (k, a_{ik}) and (k, b_{ik}) for $k = 1, 2, \dots, n-1$ construct the upper and lower hulls, respectively, for each of $i = 1, 2, \dots, n$. Then, using both (k, a_{ik}) and (k, b_{ik}), we can easily construct the convex hull of the polygon [9], which completes the proof. ∎

Let $\mathcal{W}_{ik} = \{\sum_{m \in \mathcal{I}_k} w_{im} : \text{forall } \mathcal{I}_k \subseteq \mathcal{J}_n\}$ for $k = 1, 2, \dots, n-1$, where $\mathcal{J}_n = \{1, 2, \dots, n-1\}$ and $\mathcal{I}_k$ is a k-element subset of $\mathcal{J}_n$. Then, since $a_{ik} = \sum_{\ell=1}^{k} w_{i(n-\ell)}$ is the sum of the k largest values of w_{ij} and that $b_{ik} = \sum_{\ell=1}^{k} w_{i(\ell)}$ is the sum of the k smallest values, the set $\mathcal{W}_{ik}$ is bounded by a_{ik} (upper) and b_{ik} (lower) for each of $i = 1, 2, \dots, n$. Thus, all the points in $\cup_{k=1}^{n-1}(\{k\} \times \mathcal{W}_{ik})$ are within the convex hull of the polygon constructed by Lemma A.1.

Theorem A.2: *Suppose that we have weights w_{ij} as in Equation* (9). *Then the lower and upper bounds of the finite-sample breakdown point are given by*

$$\epsilon_L = \frac{\min_{1 \le i \le n} \sum_{k=1}^{n-1} \mathbb{I}(\sum_{\ell=1}^{k} w_{i(n-\ell)} < 1/2)}{n-1}$$

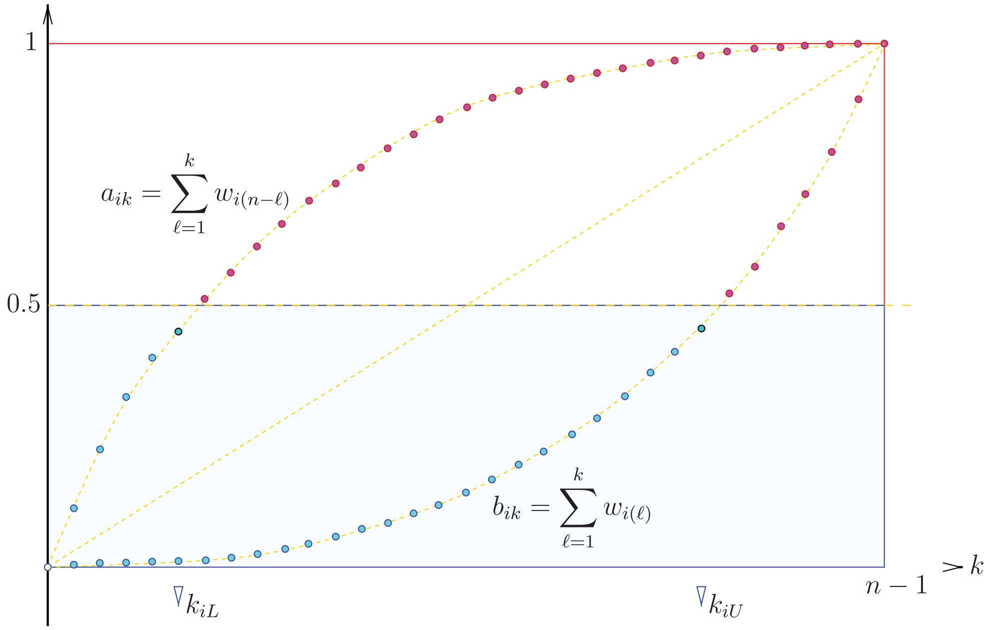

Figure A1. An illustration of the ith convex hull constructed by a_{ik} and b_{ik}.

and

$$\epsilon_U = \frac{\max_{1\le i\le n} \sum_{k=1}^{n-1} \mathbb{I}(\sum_{\ell=1}^{k} w_{i(\ell)} < 1/2)}{n-1}.$$

Proof: Let $w_{i(\ell)}$'s be the values of the order statistics of w_{ij}, such that $0 < w_{i(1)} \le w_{i(2)} \le \ldots \le w_{i(n-1)}$ for each of $i = 1, 2, \ldots, n$. Using Lemma A.1, we can construct the convex hull as seen in Figure A1. Thus, for each i, we have $k_{iL} = \sum_{k=1}^{n-1} \mathbb{I}(a_{ik} < 1/2)$ and $k_{iU} = \sum_{k=1}^{n-1} \mathbb{I}(b_{ik} < 1/2)$, where $a_{ik} = \sum_{\ell=1}^{k} w_{i(n-\ell)}$ and $b_{ik} = \sum_{\ell=1}^{k} w_{i(\ell)}$. Since

$$\epsilon_L = \frac{\min_{1\le i\le n} k_{iL}}{n-1} \quad \text{and} \quad \epsilon_U = \frac{\max_{1\le i\le n} k_{iU}}{n-1},$$

we have the results. This completes the proof of Theorem A.2. ∎

As seen in Figure A1, we construct the ith convex hull. For each of $i = 1, 2, \ldots, n$, we can construct the convex hull repeatedly, which results in the n convex hulls as seen in Figures 5 and 9 in the illustrative examples. Thus, we can obtain the lower and upper bounds of the finite sample breakdown point using the above theorem.