مجلة التجارة والتمويل

https://caf.journals.ekb.eg/

كلية التجارة – جامعة طنطا

العدد : الثالث

سبتمبر 2022

Robust Estimation for Weibull Regression Model

Zeinab M. Abd El-Raoof

Assistant Lecturer of Statistics
Department of Statistics
Faculty of Commerce
Al-Azhar University- Girls' Branch
Tafahna Al-Ashraf, Egypt

Email: zeinab.abdelroof@azhar.edu.eg

Prof. Mervat M. El-Gohary

Professor of Statistics
Department of Statistics
Faculty of Commerce
Al-Azhar University- Girls' Branch
Cairo, Egypt

Asso. Prof. Enas G. Yehia

Associate Professor of Statistics
Department of Statistics
Faculty of Commerce
Al-Azhar University- Girls' Branch
Tafahna Al-Ashraf, Egypt

Robust Estimation for Weibull Regression Model

Abstract

Weibull regression model is one of the most prevalently used parametric regression models. Estimation of the parameters of this model is affected by the presence of outliers. Outliers may be caused by actual rare events or by measurement, coding, or data entry errors, which cause a serious problem in parameter estimation. Therefore, robust estimation methods are used to overcome this problem since they are insensitive to perturbations. This paper discusses the efficiency of using two robust estimation methods, the M-estimation method and the MM-estimation method for estimating the parameters of Weibull regression model. Also, a Monte Carlo simulation study was conducted to compare the performance of robust M-estimation and MM-estimation methods with the maximum likelihood method for estimating the parameters of Weibull regression model in the presence of outliers. The simulation results showed that the robust MM-estimation method gives better performance than the maximum likelihood method and the M-estimation method.

Keywords: *Weibull regression model, Outliers, Robust estimation, M-estimation, MM-estimation.*

1. Introduction

Weibull regression is one of the most popularly used parametric models in reliability, lifetime, environmental, and survival data analysis in which the values of the dependent variable are duration observations. The *maximum likelihood* (ML) estimation method is usually used for estimating the regression coefficients of Weibull regression model since it does not impose any constraints on the properties of the dependent variable. However, the *maximum likelihood estimator*

(MLE) is affected by the presence of outliers in the data [Ahmed and Cheng (2020)].

Outliers in the regression are defined by Rousseeuw and Leroy (1987) as cases for which an observation deviates from the linear relation followed by the majority of the data. It may appear in explanatory variables space, dependent variable space, or both. One outlier can destroy estimation methods, resulting in parameter estimates that do not provide useful information [Chatterjee and Hadi (1986), Ullah *et al*. (2006), and Elgohary (2008)].

Generally, robust estimation is an alternative to the ML when outliers exist in the data since it is insensitive to perturbations. Therefore, many robust estimation methods have been proposed, such as the M-estimation (Huber, 1964), *least median of squares* (LMS) (Rousseeuw, 1984), *least trimmed squares* (LTS) (Rousseeuw, 1984), S-estimation (Rousseeuw and Yohai, 1984), and MM-estimation (Yohai, 1987).

Based on robust estimation, many studies have been presented. By Alma (2011), four robust methods: the M-estimation, MM-estimation, LTS, and S-estimation were comparatively evaluated against the *ordinary least squares* (OLS) method in multiple regression model. Gad and Qura (2016) reviewed some robust estimation methods: *least absolute deviation* (LAD), M-estimation, MM-estimation, τ-estimation, and *robust efficiency weighted least squares estimation* (REWLSE) to overcome the presence of outliers in linear regression model. Almongy and Almetwaly (2017) presented comparisons between the LAD, LMS, *least quantile of squares* (LQS), LTS, M-estimation, and S-estimation to determine a suitable regression in the presence of outliers. Yu and Yao (2017) reviewed and described some robust methods including M-estimatiom, MM-estimation, LMS, LTS, S, generalized M-estimation, R-estimation, LAD, and REWLSE in linear regression model. Almetwally and Almongy (2018) studied six robust estimation methods against the OLS in linear regression model. Tinungki (2018) used the MM-estimation method based on the objective function of Tukey Bisquare to overcome the existence of extreme outlier data.

Also, Abonazel (2020) studied some handling methods of outliers and missing data in linear regression model. Nugroho *et al.* (2020) studied the M-estimation for analyzing malnutrition data in East Java with outliers at some significance levels. Piradl *et al.* (2021) considered the M-estimation for the linear regression model with correlated error terms and outliers. A comparison of M-estimation, S-estimation with MM-estimation was proposed by Singgih and Fauzan (2022) to get the best estimation of robust regression in criminal cases in Indonesia.

For Weibull regression model, few studies have been presented to deal with outliers, for example, Mann (1982) proposed optimal outliers tests in Weibull model, Khokan *et al.* (2013) presented the weighted maximum likelihood method for robust estimation in Weibull model, and Shu *et al.* (2018) used Bayesian method to estimate the parameters of Weibull regression with outliers.

This paper aims to compare the performance of classical estimation method, the ML and two robust estimation methods, the M, and MM for estimating the parameters of Weibull regression model in the presence of outliers.

The organization of this paper is as follows. In Section 2, the specification of the Weibull regression model is presented. In Section 3, the ML estimation method is provided to estimate the parameters of the Weibull regression model. In Section 4, two robust estimation methods, the M-estimation and MM-estimation are proposed for estimating the parameters of the Weibull regression model. In Section 5, a Monte Carlo simulation study is conducted to investigate the performance of the ML, M-estimation, and MM-estimation methods in Weibull regression model with outliers. In Section 6, the concluding remarks of this paper are provided.

2. Weibull Regression Model

Suppose the random variable T has a Weibull distribution with *probability density function* (pdf) as follows:

$$f_T(t;\alpha,\theta) = \frac{\theta}{\alpha}\left(\frac{t}{\alpha}\right)^{\theta-1} e^{-\left(\frac{t}{\alpha}\right)^{\theta}}, t > 0, \quad (1)$$

where $\theta > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter. The *cumulative distribution function* (cdf) of Weibull distribution is given as

$$F_T(t;\ \alpha,\theta) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\theta}}, \tag{2}$$

The Weibull regression model for the i^{th} lifetime T_i, of a sample of n observations can be obtained from the following expression:

$$log(T_i) = Y_i = X_i'\beta + \sigma\, Z_i, \tag{3}$$

where $X_i' = (1, x_{i1}, x_{i2}, \dots, x_{ip})$ is a $p+1$ vector of covariates, $\beta' = (\beta_0, \beta_1, \dots, \beta_p)$ are $p+1$ unknown regression coefficients, σ is a scale parameter, and Z_i is the independent and identically random errors distributed according to standard extreme value distribution with pdf

$$f_Z(z) = e^{z-e^z}, \tag{4}$$

and the cdf is

$$F_Z(z) = 1 - e^{-e^z}. \tag{5}$$

Then, the pdf of T can be obtained from (3) as follows:

$$T = e^{X'\beta+\sigma Z}, \tag{6}$$

$$Z = g^{-1}(T) = \frac{log(T) - X'\beta}{\sigma}, \tag{7}$$

$$|J| = \left|\frac{d[g^{-1}(T)]}{dT}\right| = \frac{1}{\sigma T}. \tag{8}$$

By substituting (7), (8) in (4), then,

$$f_T(t) = f_Z\left(g^{-1}(t)\right)|J|$$

$$= e^{\left(\frac{log(t)-X'\beta}{\sigma}\right) - e^{\left(\frac{log(t)-X'\beta}{\sigma}\right)}} \cdot \frac{1}{\sigma t} \tag{9}$$

$$= \left(\frac{t}{e^{X'\beta}}\right)^{\frac{1}{\sigma}} . e^{-\left(\frac{t}{e^{X'\beta}}\right)^{\frac{1}{\sigma}}} . \frac{1}{\sigma t}$$

$$= \frac{1/\sigma}{e^{X'\beta}} \left[\frac{t}{e^{X'\beta}}\right]^{\frac{1}{\sigma}-1} e^{-\left(\frac{t}{e^{X'\beta}}\right)^{\frac{1}{\sigma}}}, \qquad t \geq 0. \tag{10}$$

From (10), it can be observed that T has a Weibull distribution with the shape parameter $\frac{1}{\sigma}$ and scale parameter $e^{X'\beta}$, *i.e.*, $T \sim Weibull\left(\frac{1}{\sigma}, e^{X'\beta}\right)$.

The mean and variance of the random variable T are given respectively as follows:

$$E(T) = e^{X'\beta}\Gamma\left(1 + \frac{1}{1/\sigma}\right), \tag{11}$$

and

$$Var(T) = (e^{X'\beta})^2\left[\Gamma\left(1 + \frac{2}{1/\sigma}\right) - \left(\Gamma\left(1 + \frac{1}{1/\sigma}\right)\right)^2\right]. \tag{12}$$

3. Maximum Likelihood Estimation

The ML method is the most commonly used methods of estimation of regression parameters. The aim of ML method in regression model is to estimate the unknown parameters by maximizing the likelihood function. The parameters of Weibull regression model as in (3) can be estimated by the ML method as follows:

The likelihood function is given by

$$L(\sigma, \beta) = \prod_{i=1}^{n} \frac{1}{\sigma} e^{\left(\frac{y_i - X_i'\beta}{\sigma}\right) - e^{\left(\frac{y_i - X_i'\beta}{\sigma}\right)}}. \tag{13}$$

The maximum likelihood estimates of $\sigma, \beta_0, \beta_1, \dots, \beta_p$ can be obtained by differentiating (13) with respect to σ and β and equating the results with zero.

The log-likelihood function will be written as

$$l = ln\, L(\sigma, \beta) = -nln\sigma + \sum_{i=1}^{n} \frac{y_i - X_i'\beta}{\sigma} - \sum_{i=1}^{n} e^{\frac{y_i - X_i'\beta}{\sigma}}. \tag{14}$$

Then, the first partial derivatives of l with respect to σ and β are given respectively as follows:

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} - \frac{1}{\sigma}\sum_{i=1}^{n}\frac{y_i - X_i'\beta}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n}\frac{y_i - X_i'\beta}{\sigma} e^{\frac{y_i - X_i'\beta}{\sigma}}, \quad (15)$$

$$\frac{\partial l}{\partial \beta_j} = -\frac{1}{\sigma}\sum_{i=1}^{n} X_{ij} + \frac{1}{\sigma}\sum_{i=1}^{n} X_{ij} e^{\frac{y_i - X_i'\beta}{\sigma}}, j = 0, \dots, p. \quad (16)$$

The MLE of $\hat{\sigma}$, $\hat{\beta}$ are given by setting (15) and (16) equal to zero to get the solution. It is well known that it is difficult to solve these equations exactly. So, the MLE of $\hat{\sigma}$, $\hat{\beta}$ can be obtained by solving (15) and (16) numerically using Newton-Raphson iterative method [Al-Adilee and Mohamed (2014)].

4. Robust Estimation Methods

Robust estimation, a type of regression analysis intended to get beyond some of the drawbacks of conventional approaches, is a crucial technique for analyzing data that have been contaminated by outliers.

In the following subsections, two robust estimation methods: the M-estimation and MM-estimation are proposed to estimate the parameters of the Weibull regression model.

4.1. M-estimation

Huber (1964) introduced the M-etimation method as an extension of the maximum likelihood method of estimation by minimizing the function of the residuals instead of the squared residuals.

For estimating the parameters of the Weibull regression model as referred in (3) by the M-estimation method, it is required to minimize the following objective function

$$\hat{\beta}_{\mathrm{M}} = min[\sum_{i=1}^{n} \rho(e_i)]$$

$$= min\left[\sum_{i=1}^{n} \rho(\frac{e_i}{\hat{\sigma}})\right]$$

$$= min[\sum_{i=1}^{n} \rho(u_i)], \quad (17)$$

where ρ is symmetric, continuous, and nonnegative objective function, $e_i = \frac{y_i - X_i' \beta}{\hat{\sigma}}$ are the residuals, $\hat{\sigma}$ is the median absolute deviation of the residuals as an estimate of scale parameter by $\hat{\sigma} = \frac{median|e_i - median(e_i)|}{0.6745}$, and $u_i = \frac{e_i}{\hat{\sigma}}$.

By setting the first partial derivatives of Equation (17) with respect to the coefficients equals to zero, the following $p + 1$ nonlinear equations are obtained:

$$\sum_{i=1}^{n} x_{ij} \psi(u_i) = 0, \; j = 0, \dots, p, \quad (18)$$

where $\psi(.) = \rho'(.)$is the first partial derivative of the function ρ, known as the influence function.

For the ρ function, the Tukey's Bisquare objective function by Beaton and Tukey (1974) can be used as

$$\rho = \begin{cases} \frac{c^2}{6}(1 - (\frac{u_i}{c})^2)^3 \,, & |u_i| \le c \\ \frac{c^2}{6} \quad , & |u_i| > c \,, \end{cases} \quad (19)$$

where c is called the tuning constant chosen as $c = 4.685$.

To get the solution of Equation (18), a weighted function is considered as

$$w_i(M) = \frac{\psi(u_i)}{u_i}. \quad (20)$$

By using the Tukey's Bisquare weight function, then,

$$w_i(M) = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2, & |u_i| \leq c \\ 0, & |u_i| > c, \end{cases} \quad (21)$$

where $c = 4.685$.

Consequently, Equation (18) can be written as follows:

$$\sum_{i=1}^{n} x_{ij} w_i(M) u_i = 0, \; j = 0, \ldots, p. \quad (22)$$

The *iterative reweighted least squares* (IRLS) can be used to solve Equation (22) numerically. In the final step of iteration, the M-estimator is obtained as follows:

$$\hat{\beta}_{\mathrm{M}} = (X' W^M X)^{-1} X' W^M Y, \quad (23)$$

where W^M is a $n \times n$ diagonal matrix of weights $w_i(M)$.

4.2. MM-estimation

The MM method is a modified type of M method that is developed by Yohai (1987) by combining a method of high breakdown value and an efficient estimation method to achieve the breakdown point up to 50% and asymptotic efficiency as close to one as desired.

For estimating the parameters of the Weibull regression model as in (3) using the MM-estimation method, the following stages are considered:

- Estimating the scale of the residuals e_i by using the S-estimator as follows:
  1. Estimate the regression coefficients by the OLS method.
  2. Calculate the residuals.
  3. Use the following expression:

$$\hat{\sigma}_s = \begin{cases} \frac{median|e_i - median(e_i)|}{0.6745} & ; iteration = 1 \\ \sqrt{\frac{1}{nk}\sum_{i=1}^{n} w_i(S) e_i^2} & ; iteration > 1 \,, k = 0.199, \end{cases} \tag{24}$$

where $w_i(S)$ is given by the Tukey's Bisquare function as

$$w_i(S) = \begin{cases} \left[1 - \left(\frac{v_i}{c}\right)^2\right]^2 & , |v_i| \le c \\ 0 \quad , & |v_i| > c \;, \end{cases} \tag{25}$$

where $v_i = \frac{e_i}{\hat{\sigma}_s}$ and $c = 4.685$.

- Minimizing the following objective function:

$$min \sum_{i=1}^{n} \rho(v_i) \tag{26}$$

1. By differentiating (26) with respect to β , then,

$$\sum_{i=1}^{n} x_{ij} \psi(v_i) = 0, \; j = 0, \dots, p, \tag{27}$$

where $\psi(v_i)$ is the first derivative of $\rho(v_i)$.

2. Defining the weights:

$$w_i(MM) = \frac{\psi(v_i)}{v_i}. \tag{28}$$

3. Using the IRLS to solve the following equation:

$$\sum_{i=1}^{n} x_{ij} w_i(MM) v_i = 0, \; j = 0, \dots, p. \tag{29}$$

- Calculating $\hat{\beta}_{\text{MM}}$ after convergence is obtained as follows:

$$\hat{\beta}_{\text{MM}} = (X' W^{MM} X)^{-1} X' W^{MM} Y, \tag{30}$$

where W^{MM} is a $n \times n$ diagonal matrix of weights $w_i(MM)$.

5. Monte Carlo Simulation Study

In this section, a Monte Carlo simulation study is conducted in order to compare the performance of two robust estimation methods: the M-estimation and MM-estimation versus the ML method. The performance of these estimators can be evaluated using the bias and the *mean square error* (MSE) which are given by

$$Bias(\hat{\xi}_r) = \frac{1}{R}\sum_{r=1}^{R}\left|\hat{\xi}_r - \xi\right|, \qquad (31)$$

and

$$MSE(\hat{\xi}_r) = \frac{1}{R}\sum_{r=1}^{R}\left(\hat{\xi}_r - \xi\right)^2, \qquad (32)$$

where $\hat{\xi}_r$ is the estimated value of the considered parameter ξ at the rth repetition out of R replicates.

The computation of the simulation study is developed using R programming (version 3.6.3).

The simulation design is as follows:

- Generate two random variables x_1, and x_2 from uniform distribution (0,1) for different sample sizes $n = 50, 100,$ and 200.
- The regression parameters are assumed to be $\beta_0 = 0.5$, $\beta_1 = 1$, and $\beta_2 = 1$.
- The random errors, Z_i are generated from standard normal distribution.
- Generate the random variable, T_i from $Weibull\ (\frac{1}{\sigma}, e^{X_i'\beta})$ where the value of σ is taken as $\sigma = 0.5$ and 1.
- Generate randomly different percentages of outliers such that $\tau\%$ =10% and 30% in the response variable space to investigate the robustness of the proposed methods in the presence of outliers.
- All simulation results are repeated R =1000 times and all the results of all separate experiments are obtained by the same series of random numbers.

The results of the estimated mean, bias, and MSE of the estimators, MLE, M, and MM under different levels of outliers $\tau\% = 0\%,\ 10\%,$ and 30%, different sample size $n = 50, 100,$ and 200, and $\sigma = 0.5$ and 1 are reported in Tables 1-6. Also, Figures 1-4 show the estimated MSE of the estimators, MLE, M, and MM with various values of n, σ, and $\tau\%$.

Table 1: Estimated Mean, bias, and MSE values when $n = 50$ and $\sigma = 0.5$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.7021	1.0114	1.0118	0.4491
		Bias	0.2021	0.0114	0.0118	0.0509
		MSE	0.1760	0.0825	0.0755	0.0785
	M	Mean	0.6632	1.0051	1.0107	0.4653
		Bias	0.1632	0.0051	0.0107	0.0347
		MSE	0.1345	0.0403	0.0492	0.0546
	MM	Mean	0.6101	1.0039	1.0077	0.4751
		Bias	0.1101	0.0039	0.0077	0.0249
		MSE	0.0981	0.0208	0.0226	0.0327
$\tau = 10\%$	ML	Mean	0.7703	0.8389	0.8422	0.4006
		Bias	0.2703	0.1611	0.1578	0.0994
		MSE	0.2440	0.1518	0.1547	0.1342
	M	Mean	0.7103	0.8876	0.8923	0.4221
		Bias	0.2103	0.1124	0.1077	0.0779
		MSE	0.1694	0.1055	0.1057	0.1018
	MM	Mean	0.6585	0.9336	0.9310	0.4255
		Bias	0.1585	0.0664	0.0690	0.0745
		MSE	0.1435	0.0749	0.0802	0.0866
$\tau = 30\%$	ML	Mean	0.8769	1.2679	1.2856	0.3527
		Bias	0.3769	0.2679	0.2856	0.1473
		MSE	0.3867	0.2545	0.2615	0.1662
	M	Mean	0.8016	1.1469	1.1553	0.4028
		Bias	0.3016	0.1469	0.1553	0.0972
		MSE	0.2755	0.1794	0.1864	0.1248
	MM	Mean	0.7733	1.1035	1.0912	0.4174
		Bias	0.2733	0.1035	0.0912	0.0826
		MSE	0.2444	0.1496	0.1374	0.1114

Table 2: Estimated Mean, bias, and MSE values when $n = 50$ and $\sigma = 1$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.7343	1.0281	1.0218	0.9243
		Bias	0.2343	0.0281	0.0218	0.0757
		MSE	0.2012	0.1103	0.1158	0.1004
	M	Mean	0.7052	1.0125	1.0154	0.9496
		Bias	0.2052	0.0125	0.0154	0.0504
		MSE	0.1644	0.0953	0.0863	0.0805
	MM	Mean	0.6691	1.0093	1.0069	0.9605
		Bias	0.1691	0.0093	0.0069	0.0395
		MSE	0.1323	0.0565	0.0501	0.0767
$\tau = 10\%$	ML	Mean	0.8138	1.2294	1.2021	0.8626
		Bias	0.3138	0.2294	0.2021	0.1374
		MSE	0.3033	0.2342	0.2121	0.1735
	M	Mean	0.7482	1.1318	1.1551	0.8965
		Bias	0.2482	0.1318	0.1551	0.1035
		MSE	0.2339	0.1702	0.1548	0.1322
	MM	Mean	0.7193	1.1113	1.1432	0.9096
		Bias	0.2193	0.1113	0.1432	0.0904
		MSE	0.1822	0.1332	0.1477	0.1146
$\tau = 30\%$	ML	Mean	0.9037	1.3854	1.3751	0.8202
		Bias	0.4037	0.3854	0.3751	0.1798
		MSE	0.4308	0.3564	0.3468	0.2351
	M	Mean	0.8511	1.2769	1.2533	0.8674
		Bias	0.3511	0.2769	0.2533	0.1326
		MSE	0.3282	0.2651	0.2464	0.1717
	MM	Mean	0.8111	1.1975	1.1827	0.8757
		Bias	0.3111	0.1975	0.1827	0.1243
		MSE	0.2712	0.2075	0.1979	0.1529

Table 3: Estimated Mean, bias, and MSE values when $n = 100$ and $\sigma = 0.5$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.6811	1.0091	1.0087	0.4582
		Bias	0.1810	0.0091	0.0087	0.0418
		MSE	0.1604	0.0583	0.0675	0.0552
	M	Mean	0.6446	1.0024	1.0071	0.4831
		Bias	0.1446	0.0024	0.0071	0.0169
		MSE	0.1153	0.0333	0.0314	0.0425
	MM	Mean	0.5838	1.0016	1.0031	0.4927
		Bias	0.0838	0.0016	0.0031	0.0073
		MSE	0.0774	0.0195	0.0195	0.0218
$\tau = 10\%$	ML	Mean	0.7413	0.8582	0.8661	0.4181
		Bias	0.2413	0.1418	0.1342	0.0819
		MSE	0.2234	0.1286	0.1266	0.1088
	M	Mean	0.6838	0.9099	0.9135	0.4342
		Bias	0.1838	0.0901	0.0865	0.0662
		MSE	0.1430	0.0844	0.0845	0.0779
	MM	Mean	0.6311	0.9572	0.9529	0.4302
		Bias	0.1311	0.0432	0.0471	0.0608
		MSE	0.1119	0.0612	0.0568	0.0711
$\tau = 30\%$	ML	Mean	0.8483	1.1939	1.2224	0.3733
		Bias	0.3483	0.1939	0.2224	0.1267
		MSE	0.3609	0.2019	0.2191	0.1385
	M	Mean	0.7935	1.1174	1.1255	0.4151
		Bias	0.2935	0.1174	0.1255	0.0849
		MSE	0.2661	0.1492	0.1543	0.1171
	MM	Mean	0.7464	1.0911	1.0882	0.4291
		Bias	0.2464	0.0911	0.0882	0.0711
		MSE	0.1732	0.1381	0.1287	0.0931

Table 4: Estimated Mean, bias, and MSE values when $n = 100$ and $\sigma = 1$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.7121	1.0173	1.0135	0.9362
		Bias	0.2121	0.0173	0.0135	0.0638
		MSE	0.1854	0.1001	0.1100	0.0854
	M	Mean	0.6713	1.0095	1.0093	0.9661
		Bias	0.1713	0.0095	0.0093	0.0339
		MSE	0.1396	0.0867	0.0753	0.0706
	MM	Mean	0.6477	1.0054	1.0043	0.9772
		Bias	0.1477	0.0054	0.0043	0.0228
		MSE	0.1136	0.0636	0.0536	0.0636
$\tau = 10\%$	ML	Mean	0.7938	1.1763	1.1632	0.8818
		Bias	0.2938	0.1763	0.1632	0.1182
		MSE	0.2811	0.2006	0.1919	0.1612
	M	Mean	0.7282	1.1278	1.1404	0.9182
		Bias	0.2282	0.1278	0.1404	0.0818
		MSE	0.1959	0.1554	0.156	0.1102
	MM	Mean	0.6993	1.0974	1.1193	0.9245
		Bias	0.1993	0.0974	0.1193	0.0755
		MSE	0.1566	0.1168	0.1156	0.1021
$\tau = 30\%$	ML	Mean	0.8634	1.3053	1.2716	0.8343
		Bias	0.3634	0.3053	0.2716	0.1657
		MSE	0.3838	0.2785	0.2566	0.2269
	M	Mean	0.8312	1.2008	1.1849	0.8999
		Bias	0.3312	0.2008	0.1849	0.1001
		MSE	0.2929	0.2202	0.2041	0.1446
	MM	Mean	0.7942	1.1252	1.1145	0.9136
		Bias	0.2942	0.1252	0.1145	0.0864
		MSE	0.2363	0.1503	0.1454	0.1311

Table 5: Estimated Mean, bias, and MSE values when $n = 200$ and $\sigma = 0.5$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.6439	1.0098	1.0094	0.4748
		Bias	0.1439	0.0098	0.0094	0.0252
		MSE	0.1339	0.0438	0.0539	0.0427
	M	Mean	0.6063	1.0055	1.0052	0.4917
		Bias	0.1063	0.0055	0.0052	0.0083
		MSE	0.0897	0.0296	0.0265	0.0318
	MM	Mean	0.5578	1.0011	1.0025	0.5017
		Bias	0.0578	0.0011	0.0025	0.0017
		MSE	0.0626	0.0181	0.0166	0.0113
$\tau = 10\%$	ML	Mean	0.7244	0.8861	0.8934	0.4282
		Bias	0.2244	0.1139	0.1066	0.0718
		MSE	0.1934	0.0943	0.0926	0.0863
	M	Mean	0.6718	0.9249	0.9293	0.4422
		Bias	0.1718	0.0751	0.0707	0.0578
		MSE	0.0916	0.0568	0.0632	0.0661
	MM	Mean	0.6112	0.9655	0.9640	0.4589
		Bias	0.1112	0.0345	0.0360	0.0411
		MSE	0.0867	0.0356	0.0326	0.0435
$\tau = 30\%$	ML	Mean	0.8149	1.1693	1.1758	0.4056
		Bias	0.3149	0.1693	0.1758	0.0944
		MSE	0.2996	0.1727	0.1727	0.1093
	M	Mean	0.7783	1.0831	1.1294	0.4216
		Bias	0.2783	0.0832	0.1294	0.0784
		MSE	0.2233	0.1201	0.1072	0.0989
	MM	Mean	0.7364	1.0537	1.0934	0.4528
		Bias	0.2364	0.0537	0.0934	0.0472
		MSE	0.1468	0.0968	0.0912	0.0781

Table 6: Estimated Mean, bias, and MSE values when $n = 200$ and $\sigma = 1$

Outliers	Method		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}$
$\tau = 0\%$	ML	Mean	0.6878	1.0118	1.0098	0.9574
		Bias	0.1878	0.0118	0.0098	0.0426
		MSE	0.1382	0.0854	0.0853	0.0713
	M	Mean	0.6537	1.0082	1.0045	0.9718
		Bias	0.1537	0.0082	0.0045	0.0282
		MSE	0.1139	0.0622	0.0562	0.0584
	MM	Mean	0.6244	1.0035	1.0041	0.9904
		Bias	0.1244	0.0035	0.0041	0.0096
		MSE	0.0973	0.0314	0.0306	0.0433
$\tau = 10\%$	ML	Mean	0.7687	1.1511	1.1365	0.9089
		Bias	0.2687	0.1511	0.1365	0.0911
		MSE	0.2456	0.1531	0.1632	0.1309
	M	Mean	0.7075	1.1061	1.1243	0.9234
		Bias	0.2075	0.1061	0.1243	0.0766
		MSE	0.1495	0.1137	0.1396	0.0816
	MM	Mean	0.6684	1.0891	1.0961	0.9315
		Bias	0.1684	0.0891	0.0961	0.0685
		MSE	0.1225	0.091	0.0928	0.0782
$\tau = 30\%$	ML	Mean	0.8591	1.2026	1.1902	0.8683
		Bias	0.3591	0.2026	0.1902	0.1317
		MSE	0.3288	0.2143	0.1992	0.1689
	M	Mean	0.7912	1.1092	1.1591	0.9019
		Bias	0.2912	0.1092	0.1591	0.0981
		MSE	0.2418	0.1555	0.1637	0.1031
	MM	Mean	0.7508	1.0478	1.1297	0.9112
		Bias	0.2508	0.0478	0.1297	0.0888
		MSE	0.1607	0.1198	0.1359	0.0901

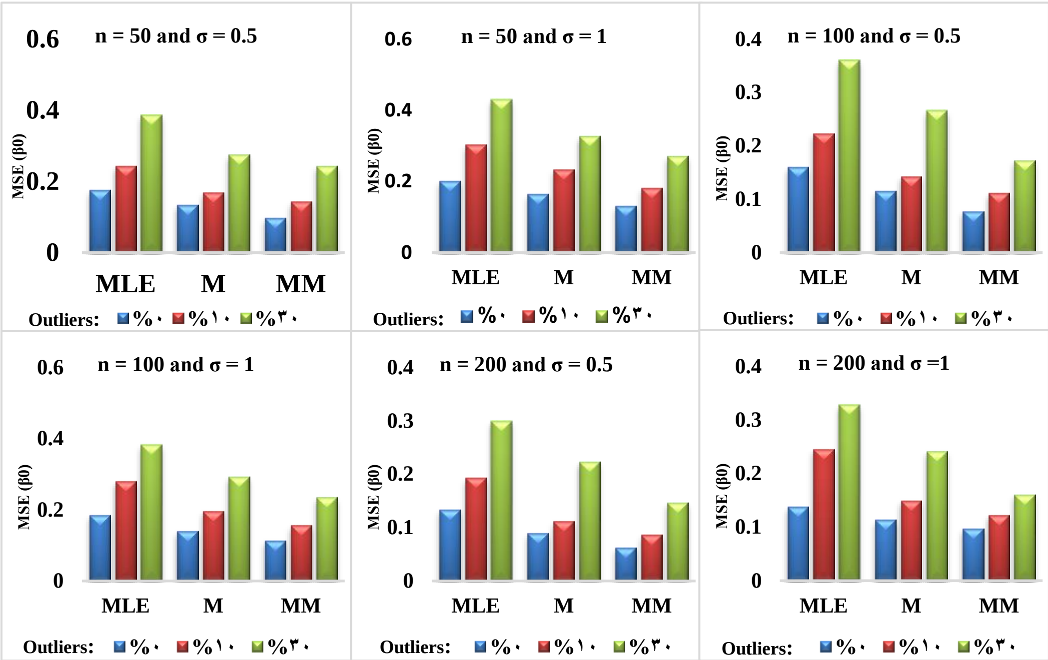

Figure 1: The MSE of different estimators for β_0 with different percentages of outliers

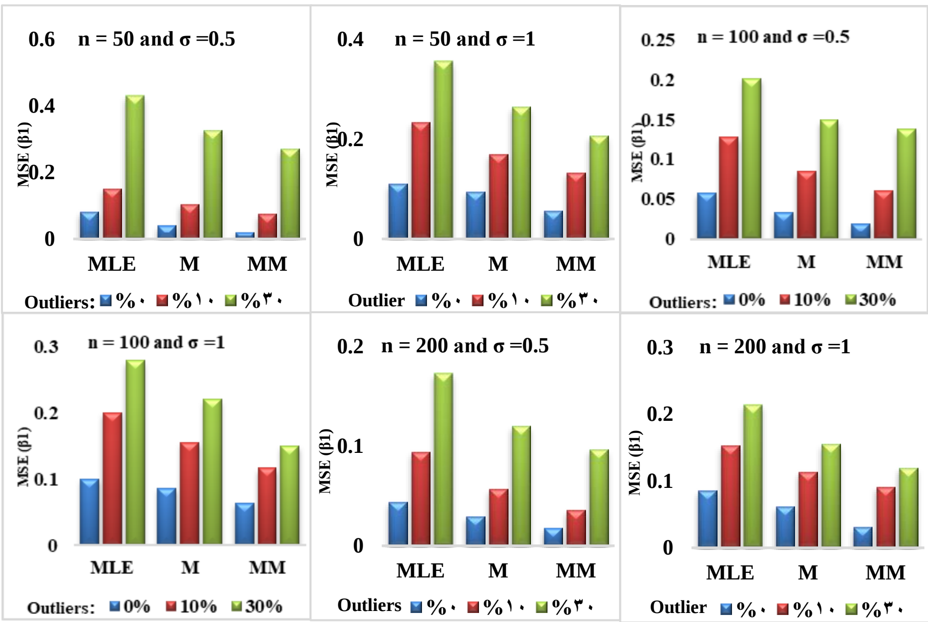

Figure 2: The MSE of different estimators for β_1 with different percentages of outliers

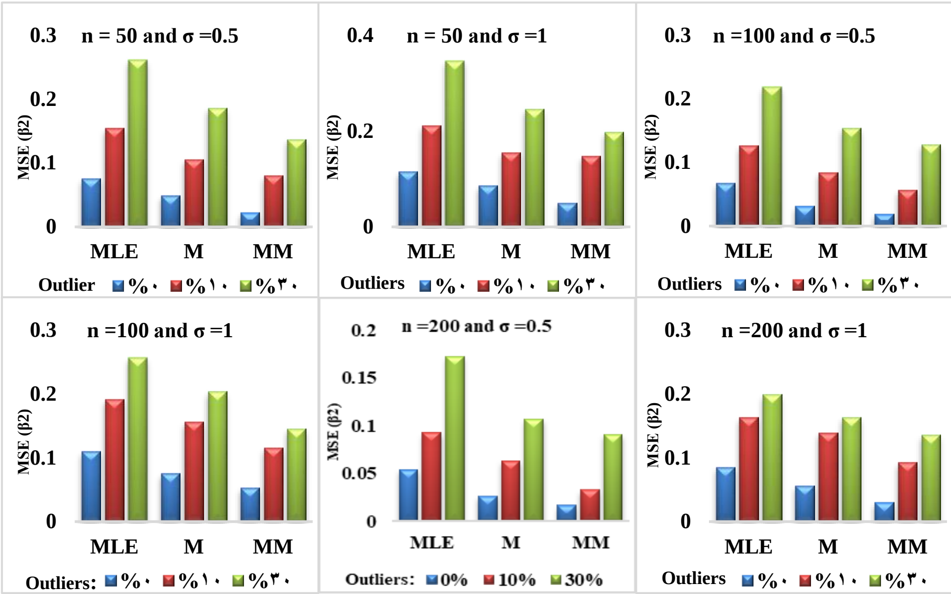

Figure 3: The MSE of different estimators for β_2 with different percentages of outliers

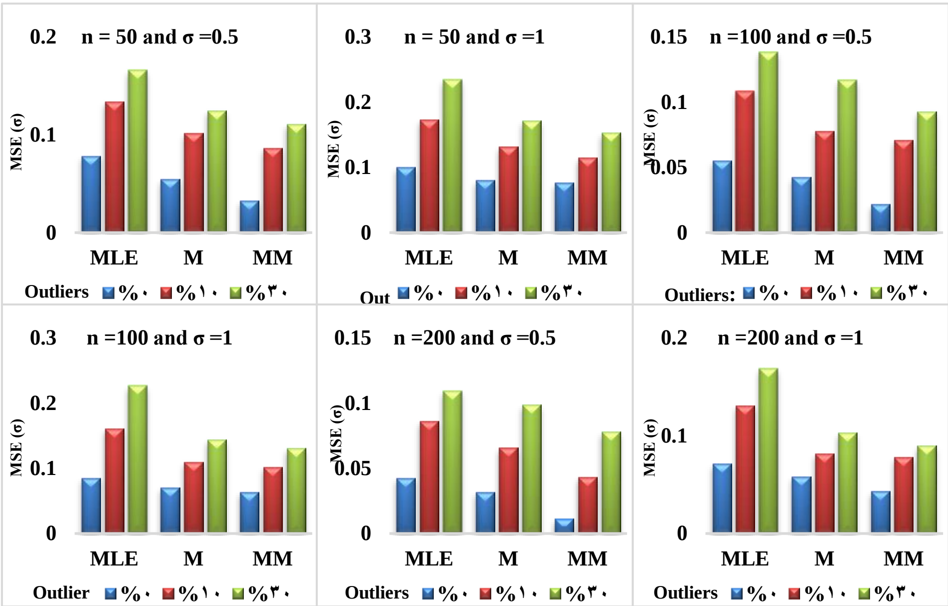

Figure 4: The MSE of different estimators for σ with different percentages of outliers

From Tables 1-6, it can be seen that when the sample size n increases, the estimated values of the bias and MSE decrease for all estimates of the MLE, M, and MM in all cases. However, as σ increases, the estimated bias and MSE increase for all estimates in all cases. In addition, as the percentage values of outliers increase, the estimated bias and MSE of all estimators are increasing. It is also seen from Tables 1-6 and Figures 1-4 that the estimates of M and MM are better than MLE in all cases. Furthermore, the MM-estimator gives the best performance compared to the other estimators since it has the smallest values of bias and MSE for all levels of outliers corresponding to different values of n and σ.

6. Concluding Remarks

This paper presents two robust estimation methods, the M-estimation and MM-estimation to estimate the parameters of Weibull regression model in the presence of outliers. A simulation study is carried out to compare the performance of the M-estimation and MM-estimation methods against the ML method in various cases of outliers. The results showed that the MM-estimation method has better performance than the M-estimation and ML methods.

References

1. Abonazel, M. R. (2020). Handling Outliers and Missing Data in Regression Models Using R: Simulation Examples. *Academic Journal of Applied Mathematical Sciences*, 6(8): 187-203.
2. Ahmed, I. A. I. and Cheng, W. (2020). The Performance of Robust Methods in Logistic Regression Model. *Open Journal of Statistics*, 10: 127-138.
3. Al-Adilee, A., and Mohamed, D. K., (2014). Linear Regression Model Related to Gumbel Distribution. *Journal of Kerbala University*, 12(4): 70-78.

4. Alma, Ö. G. (2011). Comparison of Robust Regression Methods in Linear Regression. *International Journal of Contemporary Mathematical Sciences*, 6(9): 409-421.
5. Almetwally, E. M., and Almongy, H. M. (2018). Comparison between M-Estimation, S-Estimation, and MM Estimation Methods of Robust Estimation with Application and Simulation. *International Journal of Mathematical Archive*, 9(11): 55-63.
6. Almongy, H. M., and Almetwaly, E. M. (2017). Comparison between Methods of Robust Estimation for Reducing the Effect of Outliers. *The Egyptian Journal for Commercial Studies*, 41(4): 1-23.
7. Beaton, A. E., and Tukey, J. W. (1974). The Firing of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data. *Technometrics*, 16: 147-185.
8. Chatterjee, S., and Hadi, A. S. (1986). Influential Observations, High Leverage Points, and Outliers in Linear Regression. *Statistical Science*, 1(3): 379-393.
9. Elgohary, M. M. (2008). Robust Estimation in the Linear Model. Unpublished paper presented to the promotion committee, Al-Azhar University.
10. Gad, A. M., and Qura, M. E. (2016). Regression Estimation in the Presence of Outliers: A Comparative Study. *International Journal of Probability and Statistics*, 5(3): 65-72.
11. Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1): 73-101.
12. Khokan, M. R., Bari, W., and Khan, J. (2013). Weighted Maximum Likelihood Approach for Robust Estimation: Weibull Model. *Dhaka University Journal of Science*, 61(2): 153-156.
13. Mann, N. R. (1982). Optimal Outlier Tests for a Weibull Model to Identify Process Changes or Predict Failure Times. *Studies in the Management Sciences*, 19: 261-270.
14. Nugroho, W. H., Wardhani, N. W. S., Fernandes, A. A. R., and Solimun. (2020). Robust Regression Analysis Study for Data with Outliers at Some Significance Levels. *Mathematics and Statistics*, 8(4): 373-381.
15. Piradl, S., Shadrokh, A., and Yarmohammadi, M. (2021). A Robust Estimation Method for the Linear Regression Model Parameters with

Correlated Error Terms and Outliers. *Journal of Applied Statistics*, 49(7): 1663-1676.

16. Rousseeuw, P. J. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*, 79(388): 871-880.
17. Rousseeuw, P. J. and Leroy, A. M. (1987). *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.
18. Rousseeuw, P. J., and Yohai, V. J. (1984). Robust Regression by Means of S-Estimators. *Robust and Nonlinear Time Series Analysis*, J. Frand, W. Hardle, and R. D. Martin (eds.), Lecture Notes in Statistics (26): 256-272, Springer-Verlag, New York.
19. Shu, C., Qin, T., Chen, X., and Yin, P. (2018). Study on Outlier Detection Method in Survival Analysis: Weibull Regression Outlier Model. *Journal of Biometrics & Biostatistics*, 9(4): 1-7.
20. Singgih, M. N. A., and Fauzan, A. (2022). Comparison of M Estimation, with MM Estimation to Get the Best Estimation of Robust Regression in Criminal Cases in Indonesia. *Jurnal Matematika, Statistika & Komputasi*, 18(2): 251-260.
21. Tinungki, G. M. (2018). Overcoming the Existence of Extreme Outlier Data by Using Robust MM Method Based on the Objective Function of Tukey Bisquare. *Science Nature*, 1(1): 022-032.
22. Ullah, I., Qadir, M. F., and Ali, A. (2006). Insha's Redescending M-estimator for Robust Regression: A Comparative Study. *Pakistan Journal of Statistics and Operation Research*, 2(2): 135-144.
23. Yohai, V. J. (1987). High Breakdown Point and High Efficiency Robust Estimates for Regression. *The Annals of Statistics*, 15(2): 624-656.
24. Yu, C., and Yao, W. (2017). Robust Linear Regression: A Review and Comparison. *Communications in Statistics-Simulation and Computation*, 46(8): 6261-6282.