Weighted Maximum Likelihood Approach for Robust Estimation: Weibull Model

Mahfuzur Rahman Khokan, Wasimul Bari[*] and Jafar A. Khan

Department of Statistics, Biostatistics and Informatics, Dhaka University, Dhaka- 1000, Bangladesh

Abstract

A robust weighted maximum likelihood (WML) estimation approach is proposed for the Weibull model in the presence of outliers. Extensive simulation studies are conducted to examine the performance of the proposed approach and simulated results reveals that the WML approach provides better results compared to the usual maximum likelihood (ML) approach in the presence of outliers.

I. Introduction

When an experiment is conducted, there is always a chance of occurring outlying observation in the data set. An outlying observation or outlier is an observation that is markedly different from the bulk of the observations. It may arise because of generating from different mechanisms or assumptions[1]. In statistical literature, it is well established that the usual maximum likelihood (ML) estimation approach does not provide consistent as well as efficient estimators for the parameters of interest in the presence of outliers. Ejaz et al.[2] indicated this problem of estimating the parameter of exponential model when a portion of observations are outlier. To overcome this problem, they suggested using the robust weighted maximum likelihood (WML) approach to estimate the parameters of the model. Note that WML approach is the generalization of ML approach where weights are estimated using the data as proposed by Tibshirani and Hastie[3], Staniswalis[4], Fan et al.[5], Hu[6], Hu and Zidek[7], Hu and Zidek[8]. The weighted likelihood theory can also be used as a simple alternative to the empirical Bayesian approach for many complex problems. The main advantage of this principle is that it enables a bias-precision trade of to be made without relying on a Bayesian approach.

The Weibull distribution is widely used model in both industrial and biomedical applications. This distribution appears to be the most appropriate choice in describing the lifetimes. In the presence of outliers in the data, the traditional maximum likelihood estimation approach does not provide reliable estimators for the scale and shape parameters involved in the Weibull distribution. To handle the outliers, Bayesian approaches were proposed by Liangappaiah[9] and Dixit[10]. He and Fung[11] proposed the method of median for the robust estimation of Weibull distribution. Boudt et al.[12] proposed three robust and explicit Weibull parameter estimators. These are the quantile least squares, the repeated median and the median/Q_n. In this paper, we apply the WML approach to the Weibull model for the robust estimation of shape and scale parameters when the data set is contaminated with outliers. This is done by assigning zero weights to the observations with small likelihoods. Extensive simulation studies are conducted to examine the performance of the WML approach and results are compared with the results obtained

from ML approach. The paper is organized as follows. Section 2 describes the WML approach and Weibull model. In Section 3, we describe how we conduct the simulation study. The simulation results are also discussed in the same section. An example is given in Section 4. We conclude the paper in Section 5.

II. The WML Approach and Weibull Model

Let $f(x, \theta)$ be the probability distribution function for a random variable X, where θ be the vector of parameters. Under the WML approach the weighted likelihood for a random sample of size n can be written as $L^*(\theta) = \prod_{i=1}^{n} f^{\delta_i}(x_i, \theta)$, where δ_i indicates whether observation x_i is an outlying observation or not. Note that δ_i takes value 1 if the observation is not an outlying observation, otherwise it takes value 0. This likelihood function reduces to the usual likelihood function $L(\theta)$ if $\delta_i = 1, \forall i$. Let $\hat{\theta}$ be the usual maximum likelihood estimator (MLE) of θ obtained from $L(\theta)$. Ejaz et al.[3] proposed to define δ_i as

$$\delta_i = \begin{cases} 1 & if\ f(x_i, \hat{\theta}) > k \\ 0, & otherwise \end{cases},$$

which is based on the robust estimation theory and maximum likelihood estimation method to reject the outlying observation. The choice of tuning constant k depends on the form of the probability distribution function $f(x, \theta)$. For an exponential distribution with mean θ, an observation is an outlier if $\hat{\theta}^{-1} \exp(-\hat{\theta}^{-1} x_i) < k$, i.e. $x_i > -\hat{\theta} ln(\tau)$, where $\tau = k\hat{\theta}$. Ejaz et al.[3] obtained the value of k as follows. The largest observation in the dataset is considered as an outlier with a pre-assigned small probability α. That is,

$$\alpha = Pr.\left(\max X_i > -\theta ln(\tau)\right) \approx 1 - \prod_{i=1}^{n} Pr.\left(X_i \leq -\theta ln(\tau)\right) = 1 - [1 - \tau]^n.$$

Therefore, $\tau = 1 - (1 - \alpha)^{\frac{1}{n}} \approx \frac{\alpha}{n}$. That is, $k = \alpha(n\ \hat{\theta})^{-1}$. Note that the choice of value for α is completely user-dependent and it can be treated as the desired level of robustness.

The Weibull distribution is the most commonly used model in reliability and life testing analysis. In this paper, we consider the WML approach for the robust estimation of shape and scale parameters of Weibull distribution when the

---

*Author for correspondence, *email :*w _bari@yahoo.com

data set is contaminated with outliers. The probability density function of Weibull distribution is given by

$$f(x,\theta) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} exp\left\{-\left(\frac{x}{\lambda}\right)^{\beta}\right\}, x \geq 0, \beta, \lambda > 0; \theta = (\beta,\lambda)',$$

where β and λ are shape and scale parameters, respectively. Note that the Weibull distribution reduces to the exponential distribution when shape parameter $\beta = 1$. Regarding the robust estimation for Weibull distribution, following exponential distribution, we modify the tuning constant k by replacing $\hat{\theta}$ with MLE of mean of Weibull distribution. Thus, one may obtain $k = \alpha(n\,\hat{\lambda} * \Gamma(1 + (1/\hat{\beta})))^{-1}$, where $\hat{\lambda}$ and $\hat{\beta}$ are the usual MLEs.

Let $\hat{\theta}^* = (\hat{\beta}^*, \hat{\lambda}^*)'$ be the weighted maximum likelihood estimator (WMLE) of θ obtained from using the weighted likelihood function $L^*(\theta)$. To obtain this WMLE, one may need to solve the estimating equation $U(\hat{\theta}^*) = 0$, where $U(\theta)$ is the score function defined as

$$U(\theta) = \left[\frac{\partial}{\partial\beta}l^*(\theta), \frac{\partial}{\partial\lambda}l^*(\theta)\right]' = [U_1(\theta), U_2(\theta)]'$$

with

$$l^*(\theta) = lnL^*(\theta) = \sum_{i=1}^{n} \delta_i\, lnf(x_i,\theta)$$

$$= \sum_{i=1}^{n}\delta_i\,ln\beta - \beta\sum_{i=1}^{n}\delta_i\,ln\lambda + (\beta - 1)\sum_{i=1}^{n}\delta_i\,lnx_i - \frac{1}{\lambda^\beta}\sum_{i=1}^{n}\delta_i\,x_i{}^\beta.$$

Therefore, $\quad U_1(\theta) = \frac{1}{\beta}\sum_{i=1}^{n}\delta_i - \sum_{i=1}^{n}\delta_i\,ln\lambda +$

$$\sum_{i=1}^{n}\delta_i\,lnx_i - \sum_{i=1}^{n}\delta_i\left(\frac{x_i}{\lambda}\right)^\beta ln\left(\frac{x_i}{\lambda}\right) \text{ and}$$

$$U_2(\theta) = -\frac{\beta}{\lambda}\sum_{i=1}^{n}\delta_i + \frac{\beta}{\lambda^{\beta+1}}\sum_{i=1}^{n}\delta_i\,x_i{}^\beta.$$

To solve the estimating equation for $\hat{\theta}^*$, one may use the Newton-Raphson iterative procedure. The estimator obtained at the mth iteration is given by

$$\hat{\theta}^{*(m)} = \hat{\theta}^{*(m-1)} + I^*(\theta)^{-1}\Big|_{\theta=\hat{\theta}^{*(m-1)}} \times$$

$$\begin{bmatrix} U_1(\theta) \\ U_2(\theta) \end{bmatrix}_{\theta=\hat{\theta}^{*(m-1)}}; m = 1,2,3,\dots,$$

where $I^*(\theta)$ is a 2×2 matrix with

$I_{11}{}^*(\theta) = -\frac{\partial}{\partial\beta}U_1(\theta), I_{22}{}^*(\theta) = -\frac{\partial}{\partial\lambda}U_2(\theta)$, and $I_{12}{}^*(\theta) = -\frac{\partial}{\partial\lambda}U_1(\theta) = -\frac{\partial}{\partial\beta}U_2(\theta) = I_{21}{}^*(\theta)$. These components for Weibull distribution under WML approach are

$$I_{11}{}^*(\theta) = \frac{1}{\beta^2}\sum_{i=1}^{n}\delta_i + \sum_{i=1}^{n}\delta_i\left(\frac{x_i}{\lambda}\right)^\beta\{ln\left(\frac{x_i}{\lambda}\right)\}^2,$$

$$I_{12}{}^*(\theta) = I_{21}{}^*(\theta) =$$

$$\frac{1}{\lambda}\sum_{i=1}^{n}\delta_i - \frac{1}{\lambda}\sum_{i=1}^{n}\delta_i\left(\frac{x_i}{\lambda}\right)^\beta - \frac{\beta}{\lambda}\sum_{i=1}^{n}\delta_i\left(\frac{x_i}{\lambda}\right)^\beta ln\left(\frac{x_i}{\lambda}\right),$$

and $I_{22}{}^*(\theta) = -\frac{\beta}{\lambda^2}\left[\sum_{i=1}^{n}\delta_i - (\beta+1)\sum_{i=1}^{n}\delta_i\left(\frac{x_i}{\lambda}\right)^\beta\right].$

Asymptotic properties of weighted maximum likelihood estimators are discussed by Wang et al.[13]. They established that under some appropriate conditions the WMLE is weakly consistent and follows a normal distribution for large sample.

III. Simulation Study

Simulation studies have been conducted to examine the performance of WML approach in the presence of outliers in the data set. The performance is examined by comparing the bias and risk of the WMLEs with those of MLEs obtained from 5000 replicates in a simulation study. For this purpose, we compute simulated mean (SM) and simulated variance (SV) for the parameter of interest and compute the bias and risk as *Bias=SM-True parameter value* and *Risk=SV+Bias²*, respectively. We use different sample sizes for the simulation study such as $n = 50, 100, and\ 200$. But, we only report the results obtained from $n = 100$. For the robust estimation, we consider a = 0.01, 0.03, 0.05, 0.07, 0.09.

In this paper, we use the WML approach for the robust estimation of parameters of the Weibull distribution in the presence of outliers. Since the mean of the Weibull distribution is $\lambda\Gamma(1 + \frac{1}{\beta})$, the outlier may arise in the data through the scale parameter λ as well as through the shape parameter β. We consider both cases in our simulation studies. To generate the outlying observations in a data set, one may use the obstructing distribution defined as

$$G(x) = (1 - \varepsilon)F(x,\theta) + \varepsilon F(x,\theta_1); \theta \neq \theta_1,$$

where F is the cumulative distribution function associated with the probability distribution function of interest f and ε is the percentage of outliers in the data set. In our simulation study, we consider different values for ε such as $\varepsilon = 0.01, 0.03, 0.05, 0.07, and\ 0.09$. Note that θ is the true parameter to be estimated and outliers arise using the parameter θ_1.

Outlier through Scale Parameter

Recall that $\theta = (\beta,\lambda)'$ denotes the vector of parameters for Weibull distribution. Keeping shape parameter β fixed, we define $\theta_1 = (\beta,\lambda(1 + \Delta))', \Delta > 0$ to generate the outliers in the data set using $G(x)$, where

$$G(x) = 1 - exp\left\{-\left(\frac{x}{\lambda}\right)^\beta\right\} - \varepsilon\left[exp\left\{-\left(\frac{x}{\lambda(1+\Delta)}\right)^\beta\right\} - exp\left\{-\left(\frac{x}{\lambda}\right)^\beta\right\}\right].$$

We use different values for Δ in our simulation studies, but we report results only for $\Delta = 1$. The biases and risks obtained under this set up are given in Table 1.

Table 1. Bias and Risk for shape and scale parameters of Weibull distribution obtained from a contaminated (contamination through scale parameter) sample of size n = 100, $\beta = 1$, Δ=1 [Risks are given in parentheses]

Parameters	ε=0.01		ε=0.03		ε=0.05		ε=0.07		ε=0.09	
	shape	scale	Shape	scale	shape	scale	shape	scale	Shape	scale
ML Approach	.0094 (.0065)	.0204 (.0446)	.0027 (.0065)	.0533 (.0486)	.0036 (.0065)	.0868 (.0549)	.0096 (.0066)	.1207 (.0637)	.0148 (.0067)	.1546 (.0747)
α =.01	.0103 (.0066)	.0197 (.0447)	.0042 (.0066)	.0522 (.0486)	.0018 (.0066)	.0852 (.0546)	.0074 (.0067)	.1185 (.0631)	.0121 (.0067)	.1517 (.0738)
α =.03	.0117 (.0068)	.0182 (.0447)	.0061 (.0068)	.0499 (.0485)	.0005 (.0067)	.0824 (.0544)	.0049 (.0068)	.1156 (.0626)	.0094 (.0068)	.1485 (.0729)
α =.05	.0129 (.0068)	.0166 (.0447)	.0078 (.0068)	.0478 (.0484)	.0025 (.0068)	.0799 (.0541)	.0026 (.0068)	.1125 (.0620)	.0069 (.0068)	.1451 (.0720)
α =.07	.0142 (.0069)	.0151 (.0448)	.0090 (.0069)	.0461 (.0483)	.0041 (.0068)	.0776 (.0538)	.0009 (.0068)	.1101 (.0616)	.0051 (.0068)	.1422 (.0713)
α =.09	.0154 (.0069)	.0134 (.0445)	.0103 (.0069)	.0443 (.0482)	.0056 (.0068)	.0753 (.0534)	.0007 (.0068)	.1075 (.0611)	.0035 (.0068)	.1397 (.0707)

It is clear from Table 1 that the ML approach provides better results for the shape parameter in term of bias compared to the WML approach if the percentage of outliers in the data set is smaller, say $\varepsilon = 0.01, 0.03, 0.05$. For example, when $\varepsilon = 0.01$ and $\alpha = 0.05$, biases are 0.0094 and 0.0129 with risks 0.0065 and 0.0068 under the ML and WML approaches, respectively. For large values of ε, say $\varepsilon = 0.07, 0.09$, for the shape parameter, the WML approach performs better than the ML approach in term of bias. For example, for $\varepsilon = 0.07$ and $\alpha = 0.05$, biases are 0.0096 and 0.0026 with risks 0.0066 and 0.0068 under the ML and WML approaches, respectively. The WML approach estimates the scale parameter with smaller biases regardless of values for ε. For example, when $\varepsilon = 0.05$ and $\alpha = 0.07$, biases for scale parameter are 0.0868 and 0.0776 with risks 0.0549 and 0.0538 under the ML and WML approaches,

respectively. It is apparent from the above table that in most of the cases, bias decreases as the level of robustness, α, increases.

Outlier through Shape Parameter

To generate outliers through only shape parameter, keeping the scale parameter fixed, we choose $\theta_1 = (\beta(1-\Delta), \lambda)^{/}, 0 < \Delta < 1$ and thus get the $G(x)$ given by

$$G(x) = 1 - \exp\left\{-\left(\frac{x}{\lambda}\right)^{\beta}\right\} - \varepsilon\left[\exp\left\{-\left(\frac{x}{\lambda}\right)^{\beta(1-\Delta)}\right\} - \exp\left\{-\left(\frac{x}{\lambda}\right)^{\beta}\right\}\right].$$

We generated data using $G(x)$ for different values of Δ, but in this paper we report the results in Table 2 obtained from $\Delta = 0.75$.

Table 2. Bias and Risk for shape and scale parameters of Weibull distribution obtained from a contaminated (contamination through shape parameter) sample of size n = 100, $\lambda = 2$, Δ=0.75 [Risks are given in parentheses]

Parameters	ε=0.01		ε=0.03		ε=0.05		ε=0.07		ε=0.09	
	shape	scale	shape	Scale	shape	scale	shape	scale	shape	scale
ML Approach	.1016 (.0754)	.0110 (.0143)	.2988 (.1805)	.0313 (.0211)	.4506 (.2954)	.0509 (.0281)	.5719 (.4105)	.0099 (.0358)	.6668 (.5161)	.0858 (.0429)
α =.01	.0167 (.0281)	.0071 (.0113)	.1212 (.0489)	.0145 (.0131)	.2296 (.0915)	.0171 (.0156)	.3368 (.1562)	.0151 (.0184)	.4359 (.2338)	.0107 (.0213)
α =.03	.0130 (.0277)	.0077 (.0113)	.1087 (.0439)	.0172 (.0129)	.2115 (.0806)	.0222 (.0151)	.3127 (.1366)	.0235 (.0175)	.4092 (.2076)	.0216 (.0203)
α =.05	.0108 (.0275)	.0079 (.0112)	.1034 (.0422)	.0184 (.0128)	.2025 (.0758)	.0245 (.0149)	.3019 (.1282)	.0272 (.0173)	.3958 (.1947)	.0271 (.0197)
α =.07	.0094 (.0275)	.0081 (.0113)	.0992 (.0407)	.0190 (.0127)	.1968 (.0728)	.0259 (.0148)	.2931 (.1221)	.0298 (.0172)	.3876 (.1869)	.0303 (.0196)
α =.09	.0085 (.0276)	.0082 (.0112)	.0967 (.0399)	.0194 (.0127)	.1925 (.0706)	.0269 (.0147)	.2873 (.1177)	.0316 (.0170)	.3793 (.1796)	.0335 (.0194)

Table 2 reveals that the WML approach performs better in estimating both shape and scale parameters compared to the ML approach. When $\varepsilon = 0.05$ and level of robustness $\alpha = 0.05$, the biases (risks) for the shape and scale parameters are 0.4506 (0.2954) and 0.0509 (0.0281) under the ML approach, respectively; but these are 0.2025 (0.0758) and 0.0245 (0.0149), respectively under the WML approach. Though the ML approach provides smaller bias for the scale parameter than the WML approach, when $\varepsilon = 0.07$, the risk is higher under the ML approach than the WML approach. For example, when $\varepsilon = 0.07$ and $\alpha = 0.01$, the bias is 0.0099 with risk 0.0358; whereas the bias is 0.0151 with risk 0.0184 under the WML approach. In many cases, it is found that risk of estimator decreases as the level of robustness increases.

Note that we also consider the scenario when there is no outlier in the data set and apply both the ML and WML approaches to estimate the shape and scale parameters. We found that ML approach performs slightly better than the WML approach in estimating parameters of Weibull distribution. These results are not shown in this paper.

IV. An Example

We consider a data set on lifetimes of six mice with brain cancer[14]. The lifetimes are 3, 4, 6, 5, 8, and 10 months. To test whether this data set fits the Weibull distribution or not, we use the equation $lnx = ln\lambda + \frac{1}{\beta}\ln[-\ln\{S(x)\}]$, where $S(x)$ is the survival function defined as $S(x) = 1 - F(x)$. To check the goodness of fit, we plot lnx against $\ln\left[-\ln\hat{S}(x)\right]$, where $\hat{S}(x)$ is the product- limit estimator of the survival function. We found an approximate straight line with some intercept and slope values. This suggests that the data set follows the Weibull distribution. The ML estimators of the shape and scale parameters are 2.7494 and 6.7698 with standard deviation 0.8736 and 1.0643, respectively.

To create an outlier in the data set, we consider the third observation 6 as 600. We then apply the ML approach and obtain the MLEs of shape and scale parameters as 0.4533 and 32.3880 with standard deviation 0.1297 and 30.7308, respectively. Therefore, it is clear that in the presence of outlier in the data, the usual ML approach does not yield the better estimates for the parameters of interest. To obtain the better estimates, one needs a robust estimation approach. For the contaminated data set, we employ the robust WML approach and obtain the estimates for shape and scale parameters as 2.5175 and 6.7986 with standard deviation 0.8872 and 1.279, respectively. These results are close to the MLEs obtained from the data set without any outlier.

V. Conclusion

It is common in practice that the data set is contaminated with outliers. In the presence of outlier, the usual maximum likelihood (ML) approach does not provide consistent as well as efficient estimators for the parameters of interest. To overcome this problem, various robust estimation approaches have been proposed in the recent past. In this paper, we have suggested a robust weighted maximum

likelihood (WML) approach to estimate the parameters in the presence of outliers. It is well known that the Weibull distribution plays an important role in reliability and life testing theory. Therefore, one may need to estimate the parameters of this distribution when the data set will be contaminated with outliers. For this purpose, we use the WML approach for the robust estimation of shape and scale parameters. To examine the performance of the WML approach, we have conducted extensive simulation studies for small as well as large samples and found that the WML performs better that the ML approach in terms of biases and risks.

The one important feature of the suggested WML approach is that this can also be used for the exponential distribution when the shape parameter will be fixed at 1. One can extend this approach to the robust estimation for the log-normal and log-logistic distribution of lifetime random variables. This approach can also be extended to the case when the data will contain the censored observations.

References

1. Hawkins, D., 1980. *Identification of Outliers*, Chapman Hall. p. 58.

2. Ejaz, S. A., I. V, Andrei, and A. H, Abdulkadir, 2005. Robust Weighted Likelihood Estimation of Exponential Parameters. *IEEE Transactions on Reliability*, **54(3),** 389-395.

3. Tibshirani, R. and T, Hastie, 1987. Local Likelihood Estimation. *Journal of American Statistical Association*, **82**, 559-567.

4. Staniswalis, J. G., 1989. The Kernel Estimate of a Regression Function in Likelihood-Based Model. *Journal of American Statistical Association*, **84**, 276-283.

5. Fan, J., N. E., Hecman, and W. P., Wand, 1995. Local Polynomial Kernel Regression for Generalized Linear Models and Quasi-likelihood Functions. *Journal of American Statistical Association*, **90**, 826-838.

6. Hu, F., 1997. Asymptotic Properties of Relevance Weighted Likelihood Estimations. *Canadian Journal of Statistics,* **25**, 45-60.

7. Hu. F. and J. V., Zidek, 2001. *The Relevance Likelihood with Applications*, S. E. Ahmed and N. Reid, Eds; Springer Verlag, 25, Empirical Bayes and Likelihood Inference. p.234.

8. Hu. F. and J. V., Zidek, 2002. The Weighted Likelihood. *Canadian Journal of Statistics,* **30**, 347-372.

9. Lingappaiah, G., 1976. Effect of Outliers on the Estimation of Parameters. *Metrika*, **23**, 27–30.

10. Dixit, U. J., 1994. Bayesian Approach to Prediction in the Presence of Outliers for Weibull Distribution. *Metrika,* 41, 127-136.

11. He, X. and Fung, W. K., 1999. Method of Medians for Lifetime Data with Weibull Models. *Stat Med*, **18**, 1993–2009.

12. Boudt, K, D., Caliskan, and C., Croux, 2011. Robust Explicit Estimators of Weibull Parameters. *Metrika,* **73**, 187-209.

13. Wang, Xiaogang, Eeden, C. V and Zidek, J. V., 2004. Asymptotic Properties of Maximum Weighted Likelihood Estimators. *Journal of Statistical Planning and Inference*, **119**, 37-54.

14. Lee, E. T., 2003. *Statistical Methods for Survival Data Analysis*, Wiley and Sons, Inc., 3rd Edition, pp 206.