School Working Paper

Comparison of estimators of the Weibull Distribution

M. Akram and A. Hayat

www.**deakin**.edu.au

The difference is **Deakin University.**

Comparison of estimators of the Weibull Distribution

Muhammad Akram
Research Fellow
Department of Epidemiology & Preventive Medicine
School of Public Health & Preventive Medicine
Monash University, The Alfred, VIC 3800
Australia
Email: Muhammad.Akram@monash.edu

Aziz Hayat
School of Accounting, Economics and Finance
Deakin University, VIC 3125
Australia
Email: Aziz.Hayat@deakin.edu.au

28 February 2013

Comparison of estimators of the Weibull Distribution

Abstract:

We compare the small sample performance (in terms of bias and root mean squared error) of L-moment estimator of 3-parameter Weibull distribution with Maximum likelihood Estimation (MLE), Moment Estimation (MoE), Least squared estimation (LSE), the Modified MLE (MMLE), Modified MoE (MMoE), and the Maximum Product of Spacing (MPS). Overall, the LM method has the tendency to perform well as it is almost always close to the best method of estimation. The ML performance is remarkable even in small sample of size $n = 10$ when the shape parameter β lies in [1.5, 4] range. The MPS estimator dominates others when $0.5 \leq \beta < 1.5$. For Large $\beta \geq 6$, the MMLE outweighs others in large sample size of $n \geq 50$, whereas LM takes the lead in small samples.

1 Introduction

Maximum likelihood Estimation (MLE) is generally a starting point when it comes to estimating the 3-parameter Weibull distribution. However, there are other estimation methods developed over time for Weibull, which are based on different methodologies such as Moment Estimation (MoE), L-moment (LM) method, Probability Weighted Moment estimators (PWM) and Leas-squares estimation (LSE). Hosking (1990) proposed a class of moments called L-moments (LM), which are defined for Weibull by Goda et al. (2010) and applied by Lana et al. (2008). Following Jing et al. (1989), Bartoluccia et al. (1999) derived the PWM of the 3-parameter Weibull distribution.

The literature on Weibull estimation further developed in two ways: One is to improve upon the MLE and the other is to supply a solution for estimation when MLE fails. The method of estimation which attempts to improve estimation over MLE are the modified MLE (MMLE) and modified MoE (MMoE) of Cohen & Whitten (1982) and perhaps the method due to McCool (1998). The estimation methods, which provide solution when MLE fails to converge are the Bayesian procedure of Smith & Naylor (1987), and the maximum product of spacing (MPS) estimators of Cheng & Amin (1983).

Now that we have got a stream of estimation methods available for 3-parameter Weibull distribution, which are well researched on theoretical aspect of the estimation method individually. However, the small sample comparison in terms of bias and Root Mean Squared Error (RMSE) of L-moment estimator vis'-a-vis other popular estimators is not documented in the literature. We attempt to fill this gap up by comparing LM estimation method with that of other methods; namely, MLE, MMLE, LSE, MoE, MMoE, and the MPS. Thereby the main aim of the paper becomes developing a guideline of choosing the best estimation method for 3-Parameter Weilbull distribution, which we think would be of deep interest to applied statisticians.

The reason for leaving out PWM is based on Bartoluccia et al. (1999) who conclude that the best use of PWM is for initializing the MLE method, because the PWM estimators are not as accurate as the ML method. Likewise, Bayesian is more subjective approach as it requires specifying a suitable prior distribution, which is reported to be a difficult task apart from the lack of interpretability of the estimation results.

We see L-moments having the theoretical advantages over conventional moments of (i) being able to characterize a wider range of distributions, and (ii) when estimated from a sample, being more robust to the presence of outliers in the data. It would be in this sense, the parameter estimates obtained from L-moments may be more accurate than MoE and ML and/or their variants in small samples due to smaller small sample bias (SSB) and being less affected by the outliers. Even for the cases where ML parameters are optimal, SSB yet exists since the ML methodology is based on asymptotic (large sample) theory. These motivated us to compare L-moment estimators with that of others.

The rest of the paper is organized as follows. In the next section, we introduce the Weibull distribution. The estimation methods and procedures are covered in Section 3. Using Monte Carlo simulations, the bias and RMSE in the estimation methods are analyzed and the guideline of the best estimation method is laid down in Section 4. In Section 5, we provide examples of choosing the best estimation method using the guidelines developed in previous section. In the last section, we summarize the work presented in this paper.

2 Weibull distribution

Weibull (1951) derived the generalization of the exponential distribution that now bears his name.[1] Since that time, the Weibull distribution has proved to be a successful model for many product failure mechanisms because it is a flexible distribution (given that it can, for example, take the form of either the exponential distribution or the approximate normal distribution, and can be skewed either positively or negatively), with a wide variety of possible failure rate curve shapes. However, Lloyd (1967), Ku et al. (1972), Hammitt (2004), and McCool (1998), among others, have extended the use of the Weibull distribution to other branches of statistics, such as reliability, risks, and quality control work.

A distribution with a general probability density function (pdf) of

$$f(x) = \frac{\beta}{\eta} \left(\frac{x-\nu}{\eta} \right)^{\beta-1} \exp\left(-\left(\frac{x-\nu}{\eta} \right)^{\beta} \right) \qquad x > \nu, \beta > 0, \eta > 0 \qquad (1)$$

---

[1]The Weibull distribution was first published in 1939 (Weibull 1939). It was an attempt to explain what were (at that time) well known but unexplained facts about material strengths, namely that the relative strength of a specimen decreases with increasing dimensions, and that its bending strength is greater than its tensile strength.

is known as a 3-parameter Weibull distribution. Its cumulative distribution function $F(x)$ and quantile function $x(F)$ are defined as

$$F(x) = 1 - \exp\left(-\left(\frac{x-\nu}{\eta}\right)^{\beta}\right) \tag{2}$$

and

$$x(F) = \eta\left(-\ln(1 - F(x))\right)^{1/\beta} + \nu, \tag{3}$$

respectively. The parameters ν, β, and η determine the location, shape and scale of the distribution, respectively. For example, if we set the shape parameter $\beta = 1$ and the location parameter $\nu = 0$, the distribution changes its shape (say from skewed when $\beta = 1.5$ to exponential). The location for the two-parameter Weibull distribution is simply at the origin, i.e., $\nu = 0$.

3 Methods of estimation

In this section, we discuss the LM estimators in detail, followed MPS estimation. The MLE method and the MoE methods and their variants due to Cohen & Whitten (1982) are discussed briefly, each followed by the procedure adopted for the estimation.

3.1 L-Moment (LM)

It is standard statistical practice to summarize a probability distribution or an observed data set by its moments or cumulants. Moment based methods, although they have long been established in statistics, are not always satisfactory. It is sometimes difficult to reach a conclusion about the shape of the distribution based on its higher order moments. The alternative approach for describing the shape of a probability distribution can be based on quantiles, known as L-moments, and was introduced by Hosking (1990). Historically, it arose from modifications of the 'probability weighted moments' of Greenwood et al. (1979).

Let X be a real-valued random variable with cumulative distribution function $F(x)$ and quantile function $x(F)$, as defined in equations (2) and (3) respectively, and let $X_{1:n} \leq$

$X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics of a random sample of size n drawn from the distribution of X. The L-moments of X are then defined as the quantities

$$\lambda_r \equiv r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}), \text{ where } r = 1, 2, \ldots.$$

The L in 'L-moments' emphasizes the fact that λ_r is a linear function of the expected order statistics. The natural estimator of λ_r based on an observed sample of data is a linear combination of the ordered data values, i.e., L-statistics. Following David & Nagaraja (2003), the expectation of an order statistic can be written as

$$E(X_{j:r}) = \frac{r!}{(j-1)!(r-j)!} \int x(F(x))^{j-1}(1-F(x))^{r-j} dF(x). \tag{4}$$

The first few (population) L-moments $(\lambda_1, \lambda_2, \ldots)$ of a distribution are defined as

$$
\begin{aligned}
\lambda_1 &= E(X) = \int_0^1 x(F) dF, \\
\lambda_2 &= \frac{1}{2} E(X_{2:2} - X_{1:2}) = \int_0^1 x(F)(2F-1) dF, \\
\lambda_3 &= \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}) = \int_0^1 x(F)(6F^2 - 6F + 1) dF, \\
\lambda_4 &= \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) = \int_0^1 x(F)(20F^3 - 30F^2 + 12F - 1) dF.
\end{aligned}
\tag{5}
$$

Furthermore, the L-moment ratios are defined as $\tau_3 = \lambda_3/\lambda_2$ and $\tau_4 = \lambda_4/\lambda_2$, where τ_3 and τ_4 measure L-skewness and L-kurtosis respectively.

Of course, the L-moments defined above are for a probability distribution, while in practice they are often estimated from a finite sample. Estimation is based on a sample of size n, arranged in order of ascending magnitude. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the ordered sample, which can be characterized better by the estimator of the probability weighted moment β_r. Following Hosking & Wallis (1997), an unbiased estimator of β_r based on

ordered samples is defined as

$$b_r = n^{-1} \binom{n-1}{r}^{-1} \sum_{j=r+1}^{n} \binom{j-1}{r} X_{j:n}. \tag{6}$$

The corresponding sample L-moments $(\ell_1, \ell_2, \dots)$ can then be defined from their population counterparts in equation (5). They are found to appear as

$$\ell_1 = b_0,$$
$$\ell_2 = 2b_1 - b_0,$$
$$\ell_3 = 6b_2 - 6b_1 + b_0,$$
$$\ell_4 = 20b_3 - 30b_2 + 12b_1 - b_0,$$

and in general

$$\ell_{r+1} = \sum_{k=0}^{r} p_{r,k} b_k, \qquad \text{for } r = 0, 1, \dots, n-1,$$

where

$$p_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+1}{k}.$$

The first three population L-moments of the 3-parameter Weibull distribution, as per equation (5), are found to be

$$\lambda_1 = \eta \Gamma\left(\frac{1}{\beta} + 1\right) + \nu,$$
$$\lambda_2 = \eta \Gamma\left(\frac{1}{\beta} + 1\right)\left(1 - \frac{1}{2^{1/\beta}}\right), \text{ and}$$
$$\lambda_3 = \eta \Gamma\left(\frac{1}{\beta} + 1\right)\left(1 - \frac{3}{2^{1/\beta}} + \frac{2}{3^{1/\beta}}\right),$$

where Γ above stands for gamma functions.

Analogously to the usual method of moments, the L-moment method also consists of equating the first few population L-moments (λ_r) to the corresponding sample L-moments (ℓ_r), thus obtaining as many equations as are needed to solve for the unknown population parameters, i.e.

$$\lambda_r = \ell_r, \qquad r = 1, 2, \ldots, p,$$

for the p parameters.

Since we aim to estimate the 3-parameter Weibull distribution, we require three equations for this purpose. As was mentioned above, the equations can be obtained by equating the first three population L-moments to the corresponding first three sample L-moments, i.e., $\lambda_1 = \ell_1$, $\lambda_2 = \ell_2$, and $\lambda_3 = \ell_3$. Substituting the values of the population L-moments λ_1, λ_2, and λ_3 for the sample L-moments, we get

$$\ell_1 = \eta \Gamma \left(\frac{1}{\beta} + 1 \right) + \nu \tag{7}$$

$$\ell_2 = \eta \Gamma \left(\frac{1}{\beta} + 1 \right) \left(1 - \frac{1}{2^{1/\beta}} \right) \tag{8}$$

$$\ell_3 = \eta \Gamma \left(\frac{1}{\beta} + 1 \right) \left(1 - \frac{3}{2^{1/\beta}} + \frac{2}{3^{1/\beta}} \right). \tag{9}$$

Solving these equations do not yield explicit solution for the estimates of parameters. Following Goda et al. (2010), we used the L-skewness measure to estimate β:

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = \frac{1 - \frac{3}{2^{1/\beta}} + \frac{2}{3^{1/\beta}}}{1 - \frac{1}{2^{1/\beta}}}. \tag{10}$$

Goda et al. (2010) estimates the parameters of Weibull distribution by fitting a polynomial functional relation of τ_3 with β. The fitting error in estimating the β was set to 0.3% for $0.6 < \beta < 3.0$. The estimate of β using equation (10) is very similar to the Goda's estimate of β. The L-moment estimate of the β is actually the iterative solution of the above non-linear equation, where τ_3 is replaced by its sample counterpart, $\hat{\tau}_3 = \hat{\ell}_3 / \hat{\ell}_2$. Also note that the Weibull distribution can be rewritten as a generalized extreme-value distribution, and approximations for the Weibull β can be obtained from those for expressions for the GEV

shape parameter given by Hosking et al. (1985) and Donaldson (1996).

After estimating β (say $\hat{\beta}$) iteratively, the estimators of η and v can easily be obtained by substituting the value of $\hat{\beta}$ in equations (7) and (8) respectively, to give:

$$\hat{\eta} = \frac{\ell_2}{\Gamma\left(\frac{1}{\hat{\beta}}+1\right)\left(1-\frac{1}{2^{1/\hat{\beta}}}\right)} \tag{11}$$

and

$$\hat{v} = \ell_1 - \hat{\eta}\Gamma\left(\frac{1}{\hat{\beta}}+1\right). \tag{12}$$

It is of interest to compare the method of L-moments with the asymptotically optimal method of maximum likelihood. The method of L-moments is usually computationally more tractable than the method of maximum likelihood, and requires less frequent recourse to iterative procedures. Following Hosking (1990), the asymptotic standard errors of L-moment estimators usually show the method of L-moments to be reasonably efficient, compared to maximum likelihood estimators.

3.2 Maximum Likelihood Estimators (MLE)

The ML estimators are inconsistent and inefficient when $\beta \leq 1$. The ML method gives estimators that are asymptotically normal and efficient when $\beta > 2$, and is usually preferred for estimating all three of the parameters when $\beta > 1$ (see for instance Cheng & Amin 1983, Johnson & Haskell 1983, Smith 1985). Using the density of x, as defined in equation (1), the log-likelihood function for the Weibull distributed independent data observations $x_1, x_2, \ldots, x_n$ can be written as

$$L(x_1, x_2, \ldots, x_n; v, \beta, \eta) = n\left(\log\beta - \beta\log\eta\right) + (\beta-1)\sum_{i=1}^{n}\log(x_i - v) - \frac{1}{\eta^\beta}\sum_{i=1}^{n}(x_i - v)^\beta, \tag{13}$$

The function in equation (13) can be maximized with respect to the parameters v, β, and η, using a routine optimization criteria such as, for example, the free quasi-Newton algorithm. However, the convergence issues with the MLE (where the function in equation (13) is flat

at the optimal parameter values) are still there and are well documented; see for instance Zanakis (1979), for details. In certain cases, the ML method may not have a solution; that is, the likelihood function in (13) may have no stationary point for estimates to converge to. In other cases, the estimates of the likelihood may converge to a saddle point, rather than the maximum of the likelihood function (see Rockette et al. 1974, for details).

3.3 Maximum Product of Spacing (MPS)

The MPS estimation technique is based on maximizing the geometric mean of identically distributed spacing of the sample y_i, where $y_i = F(x_i) = 1 - \exp\left(-\left(\frac{x_i - \nu}{\eta}\right)^{\beta}\right)$ for Weibull distributed ordered sample $x_1 < x_2 < \cdots < x_n$. The method is documented in Cheng & Amin (1983). The spacing is defined as:

$$D_i = y_i - y_{i-1} = \int_{x_{i-1}}^{x_i} f(x, \Theta) dx \quad (i = 1, 2, \ldots, n+1)$$

where $x_0 = F^{-1}(0)$, $x_{n+1} = F^{-1}(1)$, and the $\sum D_i = 1$. The x_0 and x_{n+1} can be the part of the parameter space Θ, which includes β, η, and ν. The x_0 and x_{n+1} are the endpoints of the distribution (in fact they can be any values such that $F(x_0) = 0$ and $F(x_{n+1}) = 1$, so they could be defined as $x_0 = -\infty$ and $x_{n+1} = \infty$. The MPS method is to choose Θ to maximize the geometric mean of the spacings

$$G = \left\{ \prod_{i=1}^{n+1} D_i \right\}^{1/(n+1)} \tag{14}$$

Aa per Cheng & Amin, we estimated the Weibull parameters by maximizing the logarithm of above equation.

3.4 Modified Maximum Likelihood Estimators (MMLE)

Cohen & Whitten (1982) documents the modified maximum likelihood estimators, which aim to produce satisfactory estimates of Wiebull's parameters when MLE fails to do so. They are based on the modified log-likelihood function and the modified moments conditions and hence they referred the new estimators as the modified estimators. We discuss their various modified estimators briefly here.

3.4.1 MMLE1

In this case, the log-likelihood function in equation (13) is maximized with respect to the parameters v, β, and η by imposing the constraint

$$E\big[F(x_r)\big] = F(x_r),$$

where $E(.)$ is the usual expectation, x_r is the rth order statistic in a random sample of size n, and $F(x_r)$ is the associated value of the cumulative distribution. Since $E[F(x_r)] = r/(n+1)$, the above equation becomes

$$1 - \exp\left(-\left(\frac{x_r - v}{\eta}\right)^{\beta}\right) = r/(n+1).$$

As per Cohen & Whitten (1982) restricting $r = 1$ reduces the above equation to

$$-ln\left(\frac{n}{n+1}\right) = \left(\frac{x_1 - v}{\eta}\right)^{\beta}. \tag{15}$$

Maximizing the log-likelihood function in equation (13) by imposing the above constraint is the basic methodology of Cohen and Whitten's computational procedure.

3.4.2 MMLE2

In this method, the constraint given in equation (15) is replaced by

$$E(x_1) = x_1,$$

which becomes

$$v + \left(\frac{\eta}{n}\right)^{1/\beta} \Gamma\left(\frac{1}{\beta} + 1\right) = x_1. \tag{16}$$

3.4.3 MMLE3

Likewise, the constraint given in equation (15) is replaced by

$$E(x) = \bar{x}$$

which thus becomes

$$\nu + \eta\Gamma\left(\frac{1}{\beta} + 1\right) = \bar{x}. \tag{17}$$

3.5 Moment Estimators (MoE)

The Weibull parameters cannot be expressed explicitly as functions of the conventional raw moments. However, the parameters can easily be derived by using central moments. The mean (μ) and the second and third central moments (μ_2 and μ_3) as functions of the Weibull parameters are given below.

$$\mu = \nu + \eta\,\Gamma\left(1 + \frac{1}{\beta}\right) \tag{18}$$

$$\mu_2 = \eta^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right] \tag{19}$$

$$\mu_3 = \eta^3\left[\Gamma\left(1 + \frac{3}{\beta}\right) - 3\Gamma\left(1 + \frac{1}{\beta}\right)\Gamma\left(1 + \frac{2}{\beta}\right) + 2\Gamma^3\left(1 + \frac{1}{\beta}\right)\right] \tag{20}$$

Interestingly, the coefficient of skewness (sk), given below, is invariant to location and scale transformation of the data. As can be seen, it does not depend on the location parameter ν or the scale parameter η.

$$sk = \sqrt{\frac{\mu_3^2}{\mu_2^3}} = \sqrt{\frac{\left[\Gamma\left(1 + \frac{3}{\beta}\right) - 3\Gamma\left(1 + \frac{1}{\beta}\right)\Gamma\left(1 + \frac{2}{\beta}\right) + 2\Gamma^3\left(1 + \frac{1}{\beta}\right)\right]^2}{\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right]^3}} \tag{21}$$

The shape parameter β can be estimated by finding the roots of equation (21) by replacing sk with its sample counterpart. Once the parameter β is known (estimated), the other parameters can easily be estimated using equations (18) and (19). Moment estimators are easy to compute but are less efficient than MLE (see Bartoluccia et al. 1999).

3.6 Modified Moments Estimators (MMoE)

3.6.1 MMoE1

Following Cohen & Whitten (1982), the equation (21) is replaced by

$$E\Big[F(x_r)\Big] = F(x_r)$$

As in the case of MMLE1, r is restricted to 1, and the above equation becomes

$$-ln\left(\frac{n}{n+1}\right) = \left(\frac{x_1 - v}{\eta}\right)^\beta \tag{22}$$

Solving equations (18), (19) and (22), we obtain

$$\frac{\mu_2}{(\mu - x_1)^2} = \frac{\Gamma_2 - \Gamma_1^2}{\left[\Gamma_1 - (-ln(n/(n+1)))^{1/\beta}\right]^2}, \tag{23}$$

where $\Gamma_k = \Gamma\left(1 + \frac{k}{\beta}\right)$.

The β can be estimated by solving equation (23). Once the parameter β is estimated, the other parameters follow from (18) and (19).

3.6.2 MMoE2

In this case, the equation (21) is replaced by

$$E[x_1] = x_1,$$

which reduces to

$$\nu + \frac{\eta \Gamma_1}{n^{1/\beta}} = x_1. \tag{24}$$

Solving equations (18), (19) and (24), we obtain

$$\frac{\mu_2}{(\mu - x_1)^2} = \frac{\Gamma_2 - \Gamma_1^2}{\left[\Gamma_1(1 - n^{-1/\beta})\right]^2}. \tag{25}$$

Equation (25) solves for β by replacing μ and μ_2 with their sample counterparts. The other parameters, ν and η, follow from equations (18) and (19).

3.6.3 MMoE3

In this modification, equation (21) is replaced by

$$Median = x_{me}$$

where x_{me} is a sample median, which reduces to

$$\nu + \eta (ln2)^{1/\beta} = x_{me} \tag{26}$$

Solving equations (18), (19) and (26), we obtain

$$\frac{\mu_2}{(\mu - x_{me})^2} = \frac{\Gamma_2 - \Gamma_1^2}{\left[\Gamma_1 - (ln2)^{1/\beta}\right]^2}. \tag{27}$$

Equation (27) can solve for β after its left-hand parameters replaced by their sample counterparts. Remaining parameters (ν and η) follow from equations (18) and (19).

4 Monte Carlo simulations

To establish the small sample performance of various estimators, we conduct a simulation study. We chose parameter values which were consistent with most of the

applications in various fields including quality control, engineering, actuarial sciences and finance. The parameter values used in simulations are reported in the simulation design of Table 1. The effect of various shape parameter values on the distribution is depicted in Figure 1. We opted for eight values of the shape parameter, $\beta = 0.5$, 0.75,

Table 1: *True parameters for simulation design.*

		Parameters		
Application	Distribution shape	β	η	ν
Quality control	Positively skewed	0.5	100	90
Quality control	Positively skewed	0.75	100	90
Quality control	Positively skewed	1.0	100	90
Quality control	Positively skewed	1.5	100	90
Finance	Positively skewed	3	100	90
Finance	Negatively skewed	4	100	90
Finance	Negatively skewed	6	100	90
Finance	Negatively skewed	9	100	90

For each of the parameter combinations above, we computed the bias and root mean squared errors (RMSE) for sample sizes of $n = 10, 20, 50$ and 100. The results are based on 5000 simulations.

1, 1.5, 3, 4, 6 and 9, to cover the cases where the distribution is positively skewed (for $\beta \leq 3$, generally in both the engineering and finance cases, see for example Corrado (1996), Corrado & Su (1996) and Corrado & Su (1997)), and negatively skewed for $\beta \geq 4$.

Figure 1: *Weibull density at various values of shape parameter β with $\eta = 100$ and $\nu = 90$.*
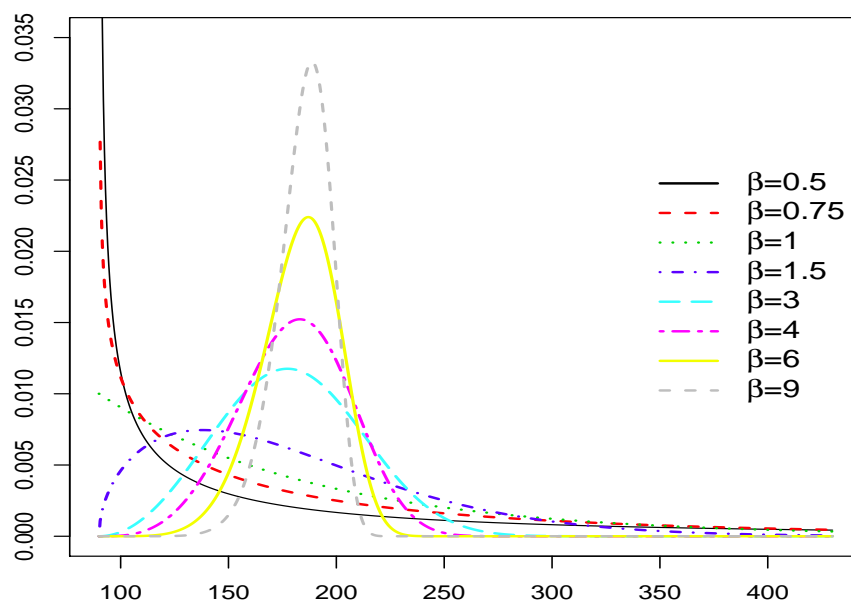
We considered $v = 90$ and kept the scale parameter η constant at 100, for consistency. The sample sizes (n) considered are 10, 20, 50, and 100, in order to cover the small, medium, and large sample sizes encountered in real life data.

In MLE, there is no guarantee that the location parameter, v, will remain non-negative (see for example, Green et al. (1994) and Smith & Naylor (1987)). Green et al. (1994) examined profile of the log-likelihood function to investigate why ML yield negative estimates of v. Smith & Naylor (1987) imposed the constraint of $0 < v \leq X_{1:n}$ in their estimation procedure. We imposed the same constraint on v for MLE and as well for other estimation methods for comparison purpose. We used estimates of MoM as starting values for ML method, and used MLE estimates as starting values for the modified versions of ML method. For the MPS method we used both MLE and MoE estimates as starting values; the results did not depend on the choice of starting values.

4.1 Simulation design

For each sample size (n) and the specified values of the parameters (see Table 1), 5000 data sets are generated from the Weibull distribution, as per equation (1). From each data set, the estimates of the parameters (β, η, v) are obtained by the methods listed in Section 3. For comparing the small sample performance of the estimators, as discussed earlier we considered, bias and the RMSE. We compute the bias of an estimator as the difference between an estimator's expectation and the true value of the parameter being estimated. Since the bias is the difference between the estimated and true values of the parameter, it does not take the variation in estimates into account, which is measured by the standard deviation of an estimator. If an estimator is biased, then the minimum variance of an estimator does not make it the best estimator. Thus, in such a situation, the RMSE is used to compare the different estimators, rather than using the unbiasedness and minimum variance. The RMSE of an estimator is one way of quantifying the difference between an estimator and the true value of the quantity being estimated. We take the RMSE as the square-root of the sum of the variance and the squared bias of an estimator, i.e., $\text{RMSE} = \sqrt{\text{Var} + \text{Bias}^2}$, and thus assesses the quality of an estimator in terms of its variation and unbiasedness together. It incorporates both the variance of an estimator and its bias. For an unbiased estimator, the RMSE is simply its standard deviation. The RMSE

will only be small when both the variance and the bias of an estimator are small.

In simulations, we considered only those cases where ML converged and the sample L-skewness was $\geq$ –0.1699. The reason is, in practice, such a sample shall not be estimated either by ML or LM. One choice would be the MPS method of Cheng & Amin, which converges when ML fails and asymptotically equivalent to the MLE. This provides a fair go of comparing the small sample performance among various estimators.

4.2 Findings

We initially considered twelve methods of estimation. Of all the methods, LSM, MMoE2, MMoE3, and MML3 showed the worst small sample performance. So, we removed these methods from our presentation. The methods (along with the simulation results) reported in Tables 2–9 are: ML, LM, MML1, MML2, MoE, MMoE1, and the MPS. The smallest the RMSE/bias the best the estimator is. The choice of the estimator in finite sample is largely depended on estimating the shape parameter (β) accurately and precisely (jointly measured by the RMSE) as much as possible. When the RMSE of two methods is close, we preferred the method with the least bias between the two. Keeping this in view, we analyzed the results. The key findings are as follows:

- $\beta = 0.5$: MPS estimator is the best choice among others. Interestingly, LM's η estimator dominates its performance in all sample sizes. Therefore, in applications where variation parameter is of deep interest, LM has the tendency to meet the analyst expectations with actually quite a reasonable estimate of β.

- $\beta = 0.75$: This time round MPS estimator surpasses all other estimators for all parameters.

- $\beta = 1$: The choice of estimator leans more towards MPS estimation method in terms of bias only but for small sample of size n=10 where MLE seemingly snatching its domination.

- $\beta = 1.5$: The MLE performs the best among all estimation methods in terms of bias even in sample size as small as n=10. However, its performance is very close with MPS in terms of RMSE of β. Since other MLE estimators (of η and ν) have smaller bias/RMSE while lower bias in β, we rank MLE on top of MPS at this contest.

- $\beta = 3$: With β getting closer to symmetric shape, as expected MLE beats all other estimation methods in bias. Interestingly, MoE offers the least RMSE, however, with almost doubling the β bias.

- $\beta = 4$: MLE clearly surpasses all other estimation methods in bias. MoE still offers the least RMSE but with bias many times fold this time round.

- $\beta = 6$: We find the positive role of modified MLE when β is large. In large sample of size n$\geq$50, MMLE1 outweighs other estimation methods. While LM performance is the best in small sample of sizes n=10, 20.

- $\beta = 9$: When β is really large, the story is virtually the same as for $\beta = 6$. LM performance snatches back its domination from others in small sample of sizes n=10,20. The MMLE1 takes the lead in large sample of sizes n=50, 100.

What we find is that as long as $0.5 \leq \beta < 1.5$, MPS is the right estimators in the sense of offering the least bias/MSE of the parameters. The reason may lie on the fact that it works (converges) when MLE fails in this range of β. For $1.5 \leq \beta \leq 4$, MLE is the way forward for estimating the 3-parameter Weibull distribution. For large values of shape parameter such as $\beta = 6$ & 9 and in the sample of large size, MMLE1 is the right choice, and in the small sample, LM estimator shall be used.

For the wide range of β values, the ML performance is quite notable even in small sample sizes. At the outset, we thought that the MLE is a starting choice (among other estimators) but now find that it can come on top of other methods for estimating the 3-parameter Weibull distribution as long as its estimates see convergence. However, the main issue with the MLE is the non-convergence of its estimates to true parameter values for small values of β and/or n. For instance, with $\beta = 1.5$ for n=10 the non-convergence rate of MLE was about 44%. The good news is the MPS estimator rises when MLE drops in performance (especially towards convergence).

For researchers less concerned with the mighty accuracy of estimators and especially after the reasonable estimator without being toggled between MLE and MPS, LM is the right choice, which we find is not the best but always close (in terms of bias/RMSE) to its competitor whether MLE, or MPS, or any other. So, we think the LM estimators for Weibull do deserve its due credit. Just to get the meaning of non-convergence mark of 44% in above

for MLE, there were only 2% cases in which the sample L-skewness <-0.1699 (the lowest theoretical bound for Skewness of Weibull distributed random variable) in L-moment estimation. So, LM-bound researchers will almost always remain close to the target without being too much worried about convergence and/or achieving theoretical bounds/limits.

4.3 Discussion

In some cases it is hard to know a priori whether the location parameter ν will be zero or not — especially when the minimum observation of the observed data set is close to zero, and/or when the standard deviation is large relative to the mean of the data. We provide an example in the application section (from a published study) where the location parameter ν was assumed to be zero where it seems not to be zero. In such cases, it may be safer to apply the 3-parameter Weibull distribution than the 2-parameter Weibull distribution. Thus, it is important to investigate which of the various estimation methods are good enough at distinguishing the cases where ν is zero or close to zero when we fit the 3-parameter Weibull distribution. We therefore considered $\nu = 0, 10$ in order to cover such cases with the same β and η values appeared in Table 1. The methodology for Monte Carlo simulations is the same as adopted in previous subsection. The results are not shown here for $\nu = 0, 10$ and can be provided on request.

Moreover, it is almost a general practice to assume the location parameter of the Weibull distribution to be positive (rather then zero) as long as $X_{1:n} > 0$ (no matter how close the first data observation is to zero); for instance, see Bartoluccia et al. (1999). A priori believe of ν to be zero will simply mean the use of 2-Parameter Weibull distribution. And the results of the best estimator in the previous section may sill be used. However in situations, where a priori believe on ν cannot be established and one would be relying on the 3-Parameter Weibull distribution then the results of best estimators in previous section may not hold. We therefore provide a comparison of the performance of previously listed methods under such scenarios of interest, i.e., when ν is actually zero or close to zero and the researcher still persists to use 3-Parameter Weibull distribution.

Interestingly, the previous results (of subsection 4.2) do not hold. The reason is this time MMoE1 comes out on top both in terms of bias and the RMSE for $\beta \geq 3$. This may be on account of MMoE1 estimates location parameter most accurately than others when it

is zero or close to zero. The message is clear, if the true location parameter is zero or close to zero, the best method of estimation in 3-Parameter Weibull distribution for $\beta \geq 3$ is MMoE1. Moreover for $\beta < 3$, LM has performed overall well. However, MPS did not perform that bad either. It is next to MMoE1 or LM in performance. The call of MMoE1 over LM was a tight one though.

5 Application

In this section we apply MLE, LM, MoE, MMLE1, MMLE2, MMoE1, and the MPS estimators to the following two examples. The data are taken from published works.

Dataset 1 in Example 1 (see Table 10) is taken from McCool (1998). The data set 1 in Table 10 looks at the rolling contact fatigue, and was first used by Ku et al. (1972). In their study, ten different test machines were employed to evaluate the effects of two lubricants on the fatigue life. One lubricant met the specifications of military standard Mil-1-7808, while the other met standard Mil-1-23699. Dataset 1 is the sorted lives in hours achieved using the Mil-1-23699.

The data in Example 2 have been taken from Thoman et al. (1969), who used it to estimate 2-parameter Weibull distribution. Their MLE estimate of β reported after adjusting for the small sample bias was 1.976. The data are the results of the tests, in millions of revolutions, of 23 ball bearings. Lieblein & Zelen (1956) originally gave the results of tests of the endurance of nearly 5000 deep-groove ball bearings. Thoman et al. (1969) basically used the sample data for size 23, given on page 286 of Lieblein & Zelen (1956).

The attempt is made to use our findings of section 4.2 to get the best estimator of β (or the true value of β) from real data in Table 10. The competition boils down to just two–MLE and MPS as others can easily be ruled out due to β expected to be not greater than 2 in these examples. In Example 1, we believe that β is less than or equal to 1 and the MPS is the right estimator to use here. The reason is when $\beta \leq 1$ (see Tables 2–4), the LM, MoE, MMoE1 has overestimated β and MLE and MMLE2 underestimated β, and MMLE1 estimate is far off true β on average. The same pattern follows in this sample (Example 1) if we believe true β is less than or equal to 1. Nevertheless, MLE is not the ideal choice when $\beta \leq 1$ as the estimator is inconsistent and inefficient and MLE fails to converge. We

can get an idea of it from v estimate, which is failed to converge and is replaced by first data observation. So, we believe MPS is the right estimation method and that its estimates are more reliable than any others.

In Example 2, the choice is little difficult as it is the value of β (between 1.0 and 2) at which both MLE and MPS are going to perform very closely. We leans towards MLE for the reasons (i) it has converged and (ii) as per our simulation's findings, its performance is best when β is in the range of $[1.5, 4]$ even for sample as small as n=10, on average. The MLE estimate of v (14.87) appears to provide the evidence of it as it is not as far off the first observation of the dataset (17.88) as the same MPS estimate (8.65). In both examples, the performance of LM is not bad either, that could be considered next best to the MPS or ML method.

6 Conclusion

In this paper, we compare the small sample performance (in terms of bias and root mean squared error) of L-moment (LM) estimation method with other estimation methods in a simulation study for various shape parameter values of 3-parameter Weibull distribution. The other estimation methods considered are Least Squared Estimation (LSE), Maximum likelihood Estimation (MLE), Moment Estimation (MoE), Modified MLE (MMLE), Modified MoE (MMoE), and the Maximum Product of Spacing (MPS). In simulations, we considered only those cases in which MLE converged and the sample L-skewness $\geq$ –0.1699 along with location parameter constraining to $0 < v < X_{1:n}$ for all methods. The reason is, in practice, such a sample shall not be estimated either by MLE or LM. This provides a fair basis of comparing the small sample performance of LM estimators among others.

We found that MPS performs well when the shape parameter β is less than 1.5, the case where MLE is inconsistent and inefficient and may not exist at all. The MLE performs relatively well when β is between 1.5 and 4, inclusive. This is because when $\beta > 2$ the Fisher information matrix is finite, and the classical properties (consistency and efficiency) hold. For relatively larger values of β, such as 6 and 9, LM method takes the lead in small samples (n=10,20), and MMLE outweighs other estimation methods in large sample size (n$\geq$50). When v is zero or close to zero, the MMoE for 3-Parameter Weibull performs the

best as compared to others, whereas MPS estimator stands next behind it among many others—although the call of MMoE over LM was a tough one.

Acknowledgement

References

Bartoluccia, A. A., Singha, K. P., Bartoluccib, A. D. & Baea, S. (1999), 'Applying medical survival data to estimate the three-parameter Weibull distribution by the method of probability-weighted moments', *Mathematics and Computers in Simulation* **48**, 385–392.

Cheng, R. C. H. & Amin, N. A. K. (1983), 'Estimating parameters in continuous univariate distributions with a shifted origin', *Journal of the Royal Statistical Society, Series B (Methodological)* **45**(3), 394–403.

Cohen, A. C. & Whitten, B. (1982), 'Modified maximum likelihood and modified moment estimators for the three-parameter weibull distribution', *Communications in Statistics– Theory and Methods* **11**.

Corrado, C. J. (1996), 'S&P 500 index option tests of Jarrow and Rudd's approximate option valuation formula', *The Journal of Futures Markets* **16**, 611–630.

Corrado, C. J. & Su, T. (1996), 'Skewness and kurtosis in S&P 500 index returns implied by option prices', *The Journal of Financial Research* **19**, 175–192.

Corrado, C. J. & Su, T. (1997), 'Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices', *The Journal of Derivatives* **4**, 8–19.

David, H. A. & Nagaraja, H. N., eds (2003), *Order statistics*, John Wiley and Sons.

Donaldson, R. (1996), 'Calculating inverse cv, skew and pwm functions for pearson–3, log-normal, extreme-value and log-logistic distributions', *Communications in Statistics— Simulation and Computation* **25**, 741–747.

Goda, Y., Kudaka, M. & Kawai, H. (2010), Incorporation of weibull distribution in L-moments method for regional frequency analysis of peaks-over-threshold wave heights, *in* 'In Proceendings of the International Conference on Coastal Engineering'. Avaliable at http://journals.tdl.org/ICCE/article/viewFile/1154/pdf_42.

Green, E. J., Roesch, F. A., Smith, A. F. M. & Strawderman, W. E. (1994), 'Bayesian estimation for the three-parameter weibull distribution with tree diameter data', *Biometrics* **50**, 254–269.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C. & Wallis, J. R. (1979), 'Probability

weighted moments: Definition and relation to parameters of several distributions expressable in inverse form', *Water Resources Research* **15**, 1049–1054.

Hammitt, J. K. (2004), 'Effects of disease type and latency on the value of mortality risk', *The Journal of Risk and Uncertainty* **28**, 73–95.

Hosking, J. R. M. (1990), 'L-moments: analysis and estimation of distributions using linear combinations of order statistics', *Journal of the Royal Statistical Society, Series B (Methodological)* **52**, 105–124.

Hosking, J. R. M. & Wallis, J. R. (1997), *Regional Frequency Analysis: An Approach Based on L-Moments*, Cambridge University Press, Cambridge, UK.

Hosking, J. R. M., Wallis, J. R. & Wood, E. F. (1985), 'Estimation of the generalized extreme-value distribution by the method of probability-weighted moments', *Technometrics* **27**(3), pp. 251–261.

Jing, D., Dedun, S., Ronfu, Y. & Yu, H. (1989), 'Expressions relating probability weighted moments to parameters of several distributions inexpressible in inverse form', *Journal of Hydrology* **110**, 259–270.

Johnson, R. A. & Haskell, J. H. (1983), 'Sampling properties of estimators of a Weibull distribution of use in the lumber industry', *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* **11**(2), 155–169.

Ku, P. M., Anderson, E. L. & Carper, H. J. (1972), 'Some considerations in rolling fatigue evaluation', *ASLE Transactions* **15**, 113–129.

Lana, X., Martinez, M. D., Burgueno, A. & Sorraa, C. (2008), 'Return period maps of dry spells for Catalonia', *Hydrological Sciences Journal* **53**.

Lieblein, J. & Zelen, M. (1956), 'Statistical investigation of the fatigue life of deep-groove ball bearings', *Journal of Research of the National Bureau of Standards* **47**, 273–316.

Lloyd, N. S. (1967), 'Weibull probability paper', *Industrial Quality Control* **23**, 452–453.

McCool, J. I. (1998), 'Inference on the Weibull location parameter', *Journal of Quality Technology* **30**, 119–126.

Rockette, H., Antle, C. & Klimko, L. A. (1974), 'Maximum likelihood estimation with the Weibull model', *Journal of the American Statistical Association* **69**(345), 246–249.

Smith, R. L. (1985), 'Maximum likelihood estimation in a class of nonregular cases', *Biometrika* **72**(1), 67–90.

Smith, R. L. & Naylor, J. C. (1987), 'A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution', *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **36**, 358–369.

Thoman, D. R., Bain, L. J. & Antle, C. E. (1969), 'Inference on the parameters of the Weibull distribution', *Technometrics* **11**, 445–460.

Weibull, W. (1939), 'A statistical theory of the strength of material', Swedish Royal Institute for Engineering Research, 151.

Weibull, W. (1951), 'A statistical distribution function of wide applicability', *AMSE Journal of Applied Mechanics* **18**, 292–297.

Zanakis, S. H. (1979), 'Extended pattern search with transformations for the three-parameter Weibull MLE problem', *Management Science* **25**, 1149–1161.

Table 2: *True parameters are $\beta = 0.5$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.17	0.38	0.09	0.20	0.04	0.11	0.02	0.07
$\hat{\eta}$	23.39	107.94	17.40	70.61	10.03	39.44	6.77	27.06
$\hat{v}$	-14.22	27.14	-8.12	14.51	-4.24	6.56	-2.79	4.05
MLE								
$\hat{\beta}$	-0.13	0.28	-0.07	0.16	-0.01	0.09	0.01	0.06
$\hat{\eta}$	89.91	148.96	91.67	126.94	81.59	100.72	69.43	86.48
$\hat{v}$	1.62	5.43	0.50	1.18	0.08	0.17	0.02	0.05
MMLE1								
$\hat{\beta}$	0.88	3.47	0.60	2.55	0.51	1.38	0.66	1.32
$\hat{\eta}$	93.37	150.73	95.05	129.37	84.91	103.63	72.73	89.47
$\hat{v}$	-1.22	10.21	-1.73	6.55	-1.88	3.40	-2.29	3.62
MMLE2								
$\hat{\beta}$	-0.12	0.25	-0.06	0.16	0.01	0.08	0.07	0.09
$\hat{\eta}$	-86.95	103.97	-94.33	95.50	-72.52	81.34	-12.68	59.45
$\hat{v}$	1.56	5.30	0.49	1.18	0.06	0.17	-0.02	0.06
MoE								
$\hat{\beta}$	0.63	0.74	0.42	0.49	0.26	0.31	0.19	0.23
$\hat{\eta}$	149.48	185.75	140.18	161.54	117.42	131.30	95.72	109.18
$\hat{v}$	-78.26	80.47	-77.38	79.48	-71.91	74.66	-62.36	66.33
MMoE1								
$\hat{\beta}$	0.28	0.46	0.16	0.21	0.08	0.13	0.05	0.09
$\hat{\eta}$	51.20	117.03	36.23	75.75	22.81	45.66	15.63	32.63
$\hat{v}$	-7.63	15.84	-2.60	3.81	-2.38	3.13	-2.49	3.08
MPS								
$\hat{\beta}$	-0.01	0.24	-0.03	0.11	-0.02	0.06	-0.01	0.04
$\hat{\eta}$	43.22	110.30	43.12	90.91	53.51	88.19	52.52	80.10
$\hat{v}$	-0.64	9.12	0.15	1.13	0.03	0.15	0.01	0.05

Table 3: *True parameters are $\beta = 0.75$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.29	0.72	0.12	0.30	0.05	0.14	0.03	0.09
$\hat{\eta}$	16.58	70.46	10.16	43.59	5.10	24.76	4.04	17.07
$\hat{v}$	-14.81	30.49	-7.15	14.92	-3.43	6.42	-2.37	4.01
MLE								
$\hat{\beta}$	-0.20	0.64	-0.21	0.31	-0.09	0.14	-0.03	0.08
$\hat{\eta}$	32.37	83.03	34.25	65.43	23.76	39.00	19.18	30.11
$\hat{v}$	2.37	13.84	2.09	3.73	0.63	1.07	0.24	0.40
MMLE1								
$\hat{\beta}$	1.32	4.40	1.71	4.62	2.05	4.17	2.20	3.96
$\hat{\eta}$	36.28	84.19	38.83	68.16	25.39	41.21	19.06	32.39
$\hat{v}$	-2.85	18.30	-4.32	12.92	-6.46	11.06	-7.52	10.73
MMLE2								
$\hat{\beta}$	-0.25	0.53	-0.22	0.29	-0.07	0.14	-0.02	0.07
$\hat{\eta}$	-80.30	94.93	-93.59	96.23	-96.79	96.98	-93.74	93.94
$\hat{v}$	2.35	13.38	2.09	3.73	0.62	1.07	0.23	0.39
MoE								
$\hat{\beta}$	0.82	0.98	0.51	0.61	0.25	0.34	0.16	0.23
$\hat{\eta}$	96.57	120.09	77.21	96.36	46.00	61.35	30.98	43.24
$\hat{v}$	-66.19	70.73	-51.60	57.50	-30.66	36.30	-20.47	25.22
MMoE1								
$\hat{\beta}$	0.43	0.89	0.14	0.27	0.05	0.13	0.03	0.09
$\hat{\eta}$	25.36	69.87	9.33	37.40	3.39	21.90	2.51	15.66
$\hat{v}$	-14.28	30.00	-2.44	6.24	-0.65	1.25	-0.70	0.92
MPS								
$\hat{\beta}$	0.07	0.58	-0.03	0.19	-0.03	0.09	-0.02	0.06
$\hat{\eta}$	15.86	63.18	6.93	36.80	2.84	21.74	2.10	15.50
$\hat{v}$	-5.19	21.40	-0.06	4.36	0.12	0.92	0.05	0.33

Table 4: *True parameters are $\beta = 1$, $\eta = 100, \nu = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.46	1.00	0.20	0.49	0.10	0.23	0.06	0.15
$\hat{\eta}$	18.04	60.39	9.28	37.61	5.97	20.77	4.26	14.15
$\hat{\nu}$	-17.40	36.76	-8.27	19.89	-3.96	8.61	-2.60	5.30
MLE								
$\hat{\beta}$	-0.10	0.97	-0.21	0.47	-0.10	0.23	-0.04	0.13
$\hat{\eta}$	10.44	62.17	3.97	40.01	1.42	21.40	0.55	13.48
$\hat{\nu}$	2.07	21.70	3.50	7.04	1.70	2.59	0.90	1.31
MMLE1								
$\hat{\beta}$	1.40	4.67	1.96	5.45	1.71	4.07	1.28	2.14
$\hat{\eta}$	14.87	62.75	8.50	41.72	4.18	23.19	4.79	15.94
$\hat{\nu}$	-4.29	24.55	-5.60	16.49	-9.38	17.19	-10.73	14.14
MMLE2								
$\hat{\beta}$	-0.30	0.78	-0.38	0.47	-0.30	0.38	-0.17	0.25
$\hat{\eta}$	-56.79	84.95	-64.61	82.61	-65.46	81.96	-73.16	86.15
$\hat{\nu}$	2.11	20.91	3.57	7.06	1.76	2.62	0.91	1.31
MoE								
$\hat{\beta}$	0.93	1.14	0.55	0.74	0.28	0.40	0.17	0.26
$\hat{\eta}$	65.55	86.04	43.84	61.92	25.55	37.85	16.78	26.26
$\hat{\nu}$	-50.72	57.55	-33.57	41.48	-18.34	24.12	-11.79	16.27
MMoE1								
$\hat{\beta}$	0.78	1.40	0.22	0.45	0.08	0.19	0.04	0.12
$\hat{\eta}$	32.83	68.13	7.93	32.18	2.77	16.89	1.37	11.49
$\hat{\nu}$	-27.22	46.06	-5.84	13.05	-1.17	2.94	-0.41	1.21
MPS								
$\hat{\beta}$	0.25	0.85	0.04	0.35	0.00	0.15	0.00	0.10
$\hat{\eta}$	19.56	56.90	6.16	30.34	3.05	16.62	1.74	11.35
$\hat{\nu}$	-12.77	32.45	-2.54	10.58	-0.43	2.52	-0.13	1.10

Table 5: *True parameters are $\beta = 1.5$, $\eta = 100, \nu = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.71	1.40	0.33	0.81	0.10	0.36	0.06	0.22
$\hat{\eta}$	23.95	59.00	12.20	37.52	4.65	19.15	2.78	12.27
$\hat{\nu}$	-22.76	46.74	-10.89	28.70	-4.25	13.26	-2.47	7.94
MLE								
$\hat{\beta}$	0.20	1.37	-0.03	0.66	-0.06	0.25	-0.04	0.16
$\hat{\eta}$	-0.95	47.21	-6.65	27.73	-5.34	13.57	-3.30	8.86
$\hat{\nu}$	-0.63	33.10	4.61	16.48	3.58	6.85	2.26	4.04
MMLE1								
$\hat{\beta}$	0.90	4.23	0.78	4.27	-0.17	1.27	-0.20	0.32
$\hat{\eta}$	3.38	46.82	-2.64	27.99	-2.02	13.12	0.23	8.77
$\hat{\nu}$	-6.21	32.79	-1.30	18.39	1.75	7.47	1.11	3.33
MMLE2								
$\hat{\beta}$	-0.27	1.12	-0.50	0.81	-0.51	0.62	-0.48	0.58
$\hat{\eta}$	-15.80	62.48	-13.25	42.66	-3.26	17.37	-0.30	11.19
$\hat{\nu}$	-0.66	31.87	4.77	16.47	3.82	6.96	2.61	4.27
MoE								
$\hat{\beta}$	0.90	1.11	0.57	0.83	0.26	0.47	0.14	0.29
$\hat{\eta}$	36.31	55.22	25.93	42.03	13.64	25.37	8.13	16.67
$\hat{\nu}$	-31.17	42.11	-21.94	32.32	-11.65	19.14	-6.86	12.29
MMoE1								
$\hat{\beta}$	1.50	2.14	0.48	0.96	0.09	0.29	0.03	0.17
$\hat{\eta}$	53.80	78.11	17.70	39.12	2.98	14.49	1.09	8.91
$\hat{\nu}$	-49.28	65.86	-15.33	30.04	-2.65	7.89	-1.00	4.00
MPS								
$\hat{\beta}$	0.52	1.22	0.14	0.64	-0.02	0.26	-0.02	0.16
$\hat{\eta}$	29.32	59.80	10.93	32.10	2.43	14.01	1.15	8.86
$\hat{\nu}$	-24.17	46.35	-7.58	22.55	-1.34	7.29	-0.57	3.86

Table 6: *True parameters are $\beta = 3$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.68	2.17	0.52	1.80	0.32	1.22	0.17	0.80
$\hat{\eta}$	16.85	58.07	12.49	47.59	7.93	32.88	4.23	22.19
$\hat{v}$	-17.64	54.47	-12.83	44.83	-7.95	30.92	-4.15	20.80
MLE								
$\hat{\beta}$	0.13	2.25	0.13	1.77	0.04	1.03	-0.02	0.62
$\hat{\eta}$	-7.63	52.26	-4.15	42.60	-2.44	26.13	-2.22	16.38
$\hat{v}$	4.58	47.76	2.42	39.38	1.79	24.11	1.94	14.98
MMLE1								
$\hat{\beta}$	-0.27	2.83	-0.33	1.88	-0.24	0.99	-0.12	0.79
$\hat{\eta}$	-3.68	50.08	0.35	41.04	3.09	26.53	3.64	18.46
$\hat{v}$	0.96	43.67	-0.53	36.46	-1.86	23.74	-2.48	16.94
MMLE2								
$\hat{\beta}$	-0.54	1.99	-0.39	1.62	-0.33	1.10	-0.34	0.78
$\hat{\eta}$	-6.94	51.49	-1.35	42.10	1.91	26.62	2.20	16.93
$\hat{v}$	4.12	45.22	2.04	37.60	1.87	23.64	2.40	14.77
MoE								
$\hat{\beta}$	-0.25	0.57	-0.25	0.57	-0.11	0.45	-0.04	0.39
$\hat{\eta}$	-9.20	27.19	-8.48	21.86	-4.02	15.70	-1.49	12.47
$\hat{v}$	8.09	25.01	7.65	20.29	3.64	14.37	1.39	11.40
MMoE1								
$\hat{\beta}$	2.67	3.49	1.62	2.72	0.93	1.87	0.49	1.16
$\hat{\eta}$	61.40	79.69	38.15	65.10	22.68	45.90	12.36	29.01
$\hat{v}$	-60.57	76.69	-37.70	62.53	-22.26	43.98	-11.97	27.45
MPS								
$\hat{\beta}$	0.49	2.01	0.37	1.71	0.24	1.13	0.11	0.69
$\hat{\eta}$	24.50	60.72	16.66	49.28	10.14	31.73	5.53	19.26
$\hat{v}$	-23.68	57.21	-16.11	46.54	-9.66	29.80	-5.12	17.84

Table 7: *True parameters are $\beta = 4$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	0.49	2.82	0.53	2.44	0.56	1.95	0.38	1.46
$\hat{\eta}$	9.39	58.14	10.39	51.59	11.70	42.07	7.76	31.53
$\hat{v}$	-10.59	55.85	-10.96	49.81	-11.74	40.85	-7.78	30.54
MLE								
$\hat{\beta}$	-0.20	2.93	0.09	2.49	0.25	1.82	0.10	1.21
$\hat{\eta}$	-14.42	55.51	-5.10	48.77	1.96	36.83	0.46	25.13
$\hat{v}$	11.47	52.26	3.53	46.60	-2.40	35.56	-0.73	24.17
MMLE1								
$\hat{\beta}$	-0.78	2.73	-0.59	2.05	-0.26	1.49	-0.16	1.17
$\hat{\eta}$	-10.20	52.42	-0.68	46.32	6.93	36.53	6.00	26.54
$\hat{v}$	7.56	47.39	0.65	42.79	-4.99	33.62	-4.08	24.06
MMLE2								
$\hat{\beta}$	-0.85	2.81	-0.50	2.20	-0.19	1.70	-0.23	1.29
$\hat{\eta}$	-12.58	53.81	-2.11	47.16	6.48	37.20	5.29	26.40
$\hat{v}$	10.38	49.78	3.12	44.20	-2.37	34.29	-0.45	23.86
MoE								
$\hat{\beta}$	-1.22	1.31	-1.17	1.25	-1.00	1.06	-0.88	0.92
$\hat{\eta}$	-27.76	34.45	-26.45	30.71	-21.97	24.44	-19.19	20.87
$\hat{v}$	25.79	32.85	24.88	29.29	20.93	23.44	18.30	19.99
MMoE1								
$\hat{\beta}$	2.93	3.86	2.30	3.82	1.70	2.98	1.19	2.34
$\hat{\eta}$	53.25	73.40	41.96	70.48	33.19	58.97	23.41	46.49
$\hat{v}$	-53.35	72.01	-42.03	69.03	-32.93	57.73	-23.21	45.50
MPS								
$\hat{\beta}$	0.30	2.62	0.43	2.37	0.56	1.91	0.35	1.36
$\hat{\eta}$	17.77	61.24	16.39	54.58	16.43	43.30	10.27	30.20
$\hat{v}$	-17.88	59.23	-16.37	52.90	-16.11	42.07	-10.05	29.25

Table 8: *True parameters are $\beta = 6$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	-0.25	4.13	0.17	3.68	0.68	3.27	0.73	2.82
$\hat{\eta}$	-5.66	58.86	2.24	55.63	9.95	49.39	10.85	42.83
$\hat{v}$	4.73	57.44	-2.73	54.58	-10.16	48.66	-10.92	42.29
MLE								
$\hat{\beta}$	-1.17	4.44	-0.41	3.81	0.31	3.20	0.41	2.60
$\hat{\eta}$	-27.21	60.64	-11.79	55.04	1.76	46.56	4.55	38.39
$\hat{v}$	25.01	58.38	10.68	53.64	-2.19	45.79	-4.73	37.82
MMLE1								
$\hat{\beta}$	-1.96	3.66	-1.40	2.95	-0.54	2.39	-0.22	2.02
$\hat{\eta}$	-21.45	55.76	-6.26	51.08	7.39	44.41	10.70	37.91
$\hat{v}$	20.25	52.48	7.20	48.77	-5.13	42.05	-7.76	35.19
MMLE2								
$\hat{\beta}$	-1.33	5.14	-1.03	3.62	-0.37	2.83	-0.16	2.42
$\hat{\eta}$	-23.81	57.06	-7.90	52.12	5.36	45.70	8.67	38.87
$\hat{v}$	22.57	55.26	9.54	50.81	-1.75	43.80	-4.05	36.66
MoE								
$\hat{\beta}$	-3.19	3.22	-3.17	3.20	-2.99	3.01	-2.82	2.84
$\hat{\eta}$	-49.82	51.82	-48.53	49.67	-44.98	45.61	-42.00	42.58
$\hat{v}$	47.99	50.31	46.90	48.23	43.56	44.24	40.80	41.37
MMoE1								
$\hat{\beta}$	2.45	4.06	3.35	5.92	2.82	4.89	2.47	4.34
$\hat{\eta}$	26.19	50.05	41.89	75.96	38.21	67.79	34.34	61.15
$\hat{v}$	-26.48	49.50	-42.12	75.12	-38.27	67.21	-34.28	60.61
MPS								
$\hat{\beta}$	-0.21	4.11	0.23	3.68	0.84	3.26	0.89	2.79
$\hat{\eta}$	6.36	63.63	11.46	60.16	17.61	52.35	16.51	44.00
$\hat{v}$	-6.62	62.48	-11.58	59.15	-17.59	51.68	-16.41	43.45

Table 9: *True parameters are $\beta = 9$, $\eta = 100$, $v = 90$*

	n=10		n=20		n=50		n=100	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
LM								
$\hat{\beta}$	-1.74	6.20	-1.23	5.62	0.38	5.20	1.06	4.85
$\hat{\eta}$	-20.03	62.23	-13.45	58.56	3.64	54.50	10.97	50.92
$\hat{v}$	19.21	61.34	12.79	57.86	-3.83	54.08	-11.02	50.60
MLE								
$\hat{\beta}$	-2.96	6.91	-1.88	5.96	-0.06	5.20	0.72	4.72
$\hat{\eta}$	-39.19	67.34	-24.46	60.86	-3.40	53.04	5.91	48.56
$\hat{v}$	37.49	65.72	23.38	59.86	3.08	52.58	-6.02	48.23
MMLE1								
$\hat{\beta}$	-3.58	6.51	-3.09	5.10	-1.45	4.02	-0.68	3.42
$\hat{\eta}$	-32.02	60.75	-17.75	55.53	2.04	49.81	10.88	46.37
$\hat{v}$	31.98	59.74	19.17	54.49	0.36	48.10	-7.85	44.13
MMLE2								
$\hat{\beta}$	-0.95	8.39	-1.49	7.25	-0.69	5.29	-0.15	4.29
$\hat{\eta}$	-36.74	62.74	-20.84	56.24	0.42	51.32	9.96	47.97
$\hat{v}$	33.00	61.02	20.61	57.06	2.29	51.21	-5.98	46.95
MoE								
$\hat{\beta}$	-6.18	6.19	-6.19	6.20	-5.96	5.97	-5.63	5.65
$\hat{\eta}$	-65.71	66.43	-65.01	65.40	-61.54	61.88	-57.49	57.93
$\hat{v}$	64.12	65.01	63.41	63.89	60.33	60.68	56.52	56.95
MMoE1								
$\hat{\beta}$	1.46	5.57	4.99	9.23	4.15	7.68	3.95	7.06
$\hat{\eta}$	2.75	38.67	39.96	79.00	37.77	72.64	37.91	69.00
$\hat{v}$	-3.13	38.82	-40.45	78.69	-37.89	72.32	-37.92	68.69
MPS								
$\hat{\beta}$	-1.32	6.41	-0.73	5.72	0.86	5.22	1.53	4.92
$\hat{\eta}$	-5.53	67.51	-0.58	63.15	13.74	57.61	19.23	53.75
$\hat{v}$	5.12	66.91	0.16	62.60	-13.78	57.25	-19.18	53.42

Table 10: *Parameters estimates under various methods.*

	Example 1			Example 2		
	β	η	ν	β	η	ν
MLE	0.75	37.20	90.40	1.59	63.89	14.87
LM	0.95	37.96	85.34	1.44	60.13	17.65
MoE	1.86	73.22	59.20	1.68	68.72	10.87
MMLE1	20.34	53.64	42.61	0.99	67.26	15.09
MMLE2	0.59	41.79	89.11	0.97	66.94	15.20
MMoE1	1.05	38.56	86.33	1.94	78.76	2.38
MPS	0.80	37.11	88.44	1.61	72.87	8.65
Dataset 1:	90.4, 94.2, 97.8, 101.8, 104.6, 113, 118, 154.9, 181.3, 186.2					
Dataset 2:	17.88, 28.92, 33, 41.52, 42.12, 45.6, 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					

Datasets 1 and 2 are used for parameter estimation in Examples 1 and 2 respectively.