Environmental Resources Research
Vol. 8, No. 2, 2020

GUASNR

Statistical inference for Birnbaum-Saunders and Weibull distributions fitted to grouped and ungrouped data

M. Teimouri*[1] and Quang V. Cao[2]
[1]Assistant Professor, Gonbad Kavous University, Gonbad Kavous, Iran
[2]Professor, School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, LA 70803

Received: Jun 2020 ; Accepted: September 2020

Abstract
For a given sample of grouped and ungrouped (raw) data, the maximum likelihood (ML) estimator is obtained using iterative algorithms such as Newton -Raphson (NR), which may not be converged always. Three-parameter Birnbaum-Saunders (BS) and Weibull distributions are frequently used in forestry and environmental sciences. In this study, we suggest using the expectation-maximization (EM) algorithm to estimate the parameters of BS and Weibull distributions when these models are fitted to grouped data. The EM algorithm is an iterative procedure that is used to obtain the ML estimator and always converges, whereas it is shown through simulation that the NR method may fail to converge. We demonstrate through three illustrations that the EM algorithm applied to the grouped data works efficiently. For the first illustration, the ML estimates of the grouped data exist and they are almost the same as the output of the EM algorithm. In the second and third real data examples that are of small sizes, the ML estimator does not exist for the ungrouped data but, we find it using the EM algorithm applied to the grouped data.

Keywords: Birnbaum-Saunders distribution, diameter modelling, expectation-maximization algorithm, Forest management, Grouped data, Maximum likelihood method, Weibull distribution.

Introduction

The application of the Weibull model for characterizing real phenomena has a long history. It was originally introduced by Waloddi Weibull for modeling the distribution of breaking strength of the materials. Since then it has applied in a wide range of applications in many fields of study. For example, it has received much attention in forestry (Bailey and Dell, 1973; Maltamo et al., 2000; Cao, 2004; Merganiand and Sterba, 2006; Zhang and Liu. 2006; Lei, 2008; Kuo-Chao et al., 2009; Stankova and Zlatanov, 2010; Schmidt et al., 2020), engineering (Johnson et al., 1994; Bartolucci et al., 1999; Murthy et al., 2004; Lai et al., 2006; Ambrozic and Vidovic, 2007; Teimouri et al., 2013a; Almalki and Nadarajah, 2014), and medicine (Zhu et al. 2011). Due to its popularity, statistical inference about the parameters of the Weibull model also received much attention in the literature. We refer readers to Green et al. (1994) (for Bayesian method), Cohen and Whitten (1982) (for modified ML) and modified moment method), Cran(1988) (for method of moment), Cousineau (2009) (for weighted ML method), Teimouri et al. (2013a) (for the linear combination of order statistics of three-parameter Weibull model), and Nagatsuka et al. (2013 (for ratio of order statistics). The Birnbaum-Saunders (BS) model was introduced by Birnbaum and Saunders (1969) for modelling the failures

*Corresponding author; teimouri@aut.ac.ir

due to cracks in materials. It has been used by researchers from different fields of study. For example, the BS distribution is used in forestry (Podlaski 2008), in reliability analysis (Leiva et al., 2007; Barros et al., 2008; Kundu et al., 2008; Lio and Park, 2008; Teimouri et al., 2013b) and biology (Hassani et al. 2019). For comprehensive information about the theory and applications of BS distribution, we refer readers to Balakrishnan and Kundu (2019) and Leiva (2015). Grouped data are frequently used in biological sciences. Instead of raw observations, data are grouped into discrete classes. Suppose that n observations of tree diameter at breast height (DBH) come independently from a family with distribution function (cdf) $G(.|\Theta)$, where $\Theta = (\theta_1, \dots, \theta_d)^T$ is is the family d-dimensional parameter vector. Furthermore, suppose data are classified into m distinct groups of the form $[r_{i-1}, r_i)$, for $i = 1, \dots, m$. So, the likelihood function of the grouped data is given by by

$$L(\Theta) = \frac{n!}{f_1! f_2! \dots f_m!} \prod_{i=1}^{m} [G(r_i|\Theta) - G(r_{i-1}|\Theta)]^{f_i}, \quad (1)$$

where $G(r_0|\Theta) = 0$, and f_i is frequency of the i-th group such that $n = \sum_{i=1}^{m} f_i$. The ML estimator of Θ, is obtained by maximizing Equation (1). Iterative methods such as Newton-Raphson (NR) are often used to numerically perform maximization. If the NR method encounters a local extremum, then this method cannot find the true solution. In another case, if the log-likelihood function is malformed, then the NR method could put itself into an infinite loop (Schworer and Hovey, 2004). Furthermore, when the initial values for the NR method are far away from the true parameter (which is where the log-likelihood function reaches its global maximum), or when the log-likelihood function at the initial values becomes large, then there is no guarantee that the NR method will converge. In this study, we derive the estimators of Θ using the expectation maximization (EM) algorithm. The EM algorithm is an iterative procedure and always converges (Little and Rubin, 2004; McLachlan and Krishnan, 2007). Each iteration of the EM algorithm guarantees increase of the log-likelihood value. Eventually, a local maximum of the log-likelihood function is reached. It is shown that the time for convergence is much longer than that of the NR method, (Little and Rubin, 2004). Because the EM method does not necessarily maximize the likelihood function, the estimates of Θ using the EM algorithm can be used as the initial values for the iterative NR method that should lead to the ML estimators. The probability density function (pdf) of Weibull and BS distributions are given,respectively, by the following.

$$g(x|\Theta) = \frac{\gamma}{\beta}\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1}\exp\left\{-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right\}, \quad (2)$$

$$g(x|\Theta) = \frac{\sqrt{\frac{x-\alpha}{\beta}}+\sqrt{\frac{\beta}{x-\alpha}}}{2\gamma(x-\alpha)}\phi\left(\frac{\sqrt{\frac{x-\alpha}{\beta}}-\sqrt{\frac{\beta}{x-\alpha}}}{\gamma}\right), \quad (3)$$

where $\Theta = (\gamma, \beta, \alpha)^T$ with $\gamma > 0$, $\beta > 0$, $x > \alpha \geq 0$, and $\phi(.)$ denots the standard normal pdf. The cumulative distribution function (cdf) of the Weibull and BS distributions are given, respectively, by

$$G(x|\Theta) = 1 - \exp\left\{-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right\}, \quad (4)$$

$$G(x|\Theta) = \Phi\left[\frac{1}{\gamma}\left(\sqrt{\frac{x-\alpha}{\beta}} - \sqrt{\frac{\beta}{x-\alpha}}\right)\right]. \quad (5)$$

The aim of this paper is to derive estimators for the parameters of the Weibull and BS distributions when these models are fitted to grouped data.

Materials and Methods

Materials

We use three samples to demonstrate the performance of the EM algorithm compared with the ML method in estimating the parameters of the BS and Weibull distributions. In the first example the real data were originally grouped while for the second and third example, the observations were originally ungrouped but were treated them as grouped data in the evaluation. This study is about the tree's diameter at breast height (DBH). Characterizing the distribution of DBH received much attention in forestry for the purpose of modeling the forest vertical structure, forest dynamics, and comparing forest stands. Table 1 presents the observed DBH (in inch) studied by Gove and Fairweather (1989).

Table 1: Tree diameter data

Class number	Class (in inch)	frequency
1	(2-3.9]	9
2	(4-5.9]	60
3	(6-7.9]	89
4	(8-9.9]	90
5	(10-11.9]	77
6	(12-13.9]	41
7	(14-15.9]	19

In the second sample, we focus on the beach pollution data in South Wales reported by Steen and Stickler (1976). This set of data shows the pollution level (number of coliform per 100 ml) on 20 days over a 5-week period (Cheng and Amin, 1983) and is given in Table 2.

Table 2: Number of coliform per 100 ml on 20 days over a 5-week period.

Week 1	Week 2	Week 3	Week 4	Week 5
1364	2154	2236	2518	2527
2600	3009	3045	4109	5500
5800	7200	8400	8400	8900
11500	12700	15300	18300	20400

Finally, we analyze the maximum flood level (in millions of cubic feet per second) for the Susquehanna River of Harrisburg reported by Dumonceaux and Antle (1973). This set of data is given in Table 3.

Table 3: Maximum flood level data in millions of cubic feet per second.

0.654	0.613	0.315	0.449	0.297
0.402	0.379	0.423	0.379	0.3235
0.269	0.740	0.418	0.412	0.494
0.416	0.338	0.392	0.484	0.265

Methods

For grouped data, let $G(.|\Theta)$ denote the cdf of the family, the ML estimator is obtained by maximizing the logarithm of the likelihood function (1). For DBH data, let f_i indicate on the number of recorded DBHs in interval $(r_{i-1}, r_i]$, for $i = 1, \dots, m$. Then, the ML estimator of Θ is obtained by maximizing $l(\Theta) = \log L(\Theta)$ with respect to Θ. We have

$$l(\Theta) \propto \sum_{i=1}^{m} f_i \log[G(r_i|\Theta) - G(r_{i-1}|\Theta)].$$

The EM algorithm (Dempster et al., 1977) is known as the popular method for computing the ML estimators when we encounter the incomplete data problem. In other words, the use of the EM algorithm involves cases that we are dealing with latent variables, provided that the statistical model is formulated as a missing data problem. In what follows, we give a brief description of the EM algorithm. Let $\boldsymbol{\xi}$, $\boldsymbol{Z}$, and $\boldsymbol{\omega}$ denote the complete, unobservable variable, and observed data, respectively (complete data consists of observed values and unobservable variables, i.e., $\boldsymbol{\xi} = (\boldsymbol{Z}, \boldsymbol{\omega})$). The EM algorithm works by maximizing the conditional expectation $Q = Q(\Theta|\Theta^{(t)}) = E(l_c(\Theta; \boldsymbol{\xi})|\boldsymbol{\omega}, \Theta^{(t)})$ of the complete data log-likelihood function, given the observed data and a current estimate $\Theta^{(t)}$ of the parameter vector Θ where $l_c(\Theta; \boldsymbol{\xi})$ denotes the log-likelihood function for the complete data. Each iteration of the EM algorithm consists of two steps: E-step (i.e., computing Q at the t-th iteration) and M-step (maximizing Q with respect to Θ to get $\Theta^{(t+1)}$). The E-step and M-step are repeated until convergence occurs. Suppose n independent realizations come from a distribution with pdf $g(.|\Theta)$, in which Θ is the unknown parameter vector. Further, let X_{ij}, for $j = 1, \dots, f_i$, denote the independent and identically distributed (iid) random variables within subinterval $(r_{i-1}, r_i]$, for $i = 1, \dots, m$. The complete data log-likelihood function is

$$l_c(\Theta) \propto \sum_{i=1}^{m} \sum_{j=1}^{f_i} \log g(X_{ij}|\Theta), \quad (6)$$

where $\sum_{i=1}^{m} f_i = n$. In the EM algorithm framework, X_{ij} (missing data within subinterval $(r_{i-1}, r_i]$) is regarded as the j-th realization within i-th group. So, analyzing grouped data can be considered as solving an incomplete data problem in which $\boldsymbol{\omega} = (f_1, \dots, f_m)$ is vector of observed values and $\boldsymbol{Z} = (X_{i1}, \dots, X_{if_i})$ is vector of unobservable variables, for $i = 1, \dots, m$. The ML and EM methods are employed to estimate the parameters of the BS and Weibull models fitted to the grouped and ungrouped data. Also, we performed a simulation to check the robustness of the ML and EM methods with respect to initial values. For obtaining these estimators the package ForestFit (Teimouri et al., 2020) developed for R (R Core Team, 2018) environment is uded.

Estimation methods for the Weibull distribution

The ML and EM methods described in the previous section are employed to estimate the parameters of the three-parameter Weibull distribution when this model is fitted to the grouped data. The log-likelihood function, $l(\Theta)$, of the three-parameter Weibull distribution is given by

$$l(\Theta) \propto \sum_{i=1}^{m} f_i \log\{G(r_i|\Theta) - (r_{i-1}|\Theta)\}, \quad (7)$$

where $G(.|\Theta)$ is given in (4) and $\Theta = (\gamma, \beta, \alpha)^T$, $\alpha \le r_0 < r_1 < \cdots < r_m$. Details for deriving the estimators of Θ using ML and EM methods are given in Appendices A and B, respectively.

Estimation methods for the BS distribution

In this section, we derive estimators for parameters of the BS distribution when this model is fitted to grouped data using ML and EM methods. Suppose n observations that follow the BS distribution with pdf given in (3) are discretized into m distinct intervals with end-points $r_1, \dots, r_m$. The ML estimator $\widehat{\Theta} = (\hat{\gamma}, \hat{\beta}, \hat{\alpha})^T$, of the parameter vector $\Theta = (\gamma, \beta, \alpha)^T$ is obtained by maximizing the log-likelihood function, $l(\Theta)$, of the three-parameter BS distribution given by

$$l(\Theta) \propto \sum_{i=1}^{m} f_i \log\{G(r_i|\Theta) - G(r_{i-1}|\Theta)\}, \quad (8)$$

where $G(.|\Theta)$ is given in (5) and $\alpha \le r_0 < r_1 < \cdots < r_m < \infty$. Detailes for deriveing the ML and EM estimators of Θ are not given here for space reasons.

Simulation study for checking convergence of th NR method

The reason for checking the convergence via simulation is that the NR method is sensitive to the initial values for maximizing the right-hand side of (8) or (7). Simulation is run 1000 times and in each run, the elements of the parameter vector $\Theta = (\gamma, \beta, \alpha)^T$ of the BS distribution come from uniform distributions of the ranges $(0.1,10)$, $(0.1,20)$, and $(0,30)$, respectively. The simulated observations in each run were grouped into separate classes of size 10, 20, 30, and 40.

Results

To obtain the ML estimator, the EM estimator is used as the initial values. The Anderson-Darling (AD) and Chi-square (Chi-sq) statistics were used to determine which of these distributions fitted the DBH data better. Details for computing the AD and Chi-sq statistics are given in Teimouri (2020). For sample one, estimated parameters as well as the goodness-of-fit (GOF) statistics are given in Table 4. The estimated parameters are almost the same of those obtained by Gove and Fairweather (1989) namely $\hat{\gamma} = 2.55$, $\hat{\beta} = 7.776$, and $\hat{\mu} = 2$.

Table 4: Estimated parameters and GOF statistics when Weibull distribution is fitted to sample 1 (DBH data).

Method	Estimator			GOF statistics	
	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\alpha}$	AD	Chi-sq
ML	2.610	7.771	1.999	29.635	5.165
EM	2.578	7.687	2.077	29.716	5.197

In the case of second sample, for the ungrouped data, the ML estimator of the parameters of a Weibull distribution fitted to the pollution data breaks down. The partially log-likelihood function, i.e., $\hat{L}(\alpha) = \log L(\Theta_{ML})$ where $\Theta_{ML} = (\gamma_{ML}, \beta_{ML}, \alpha)^T$ is shown in figure 1(a).

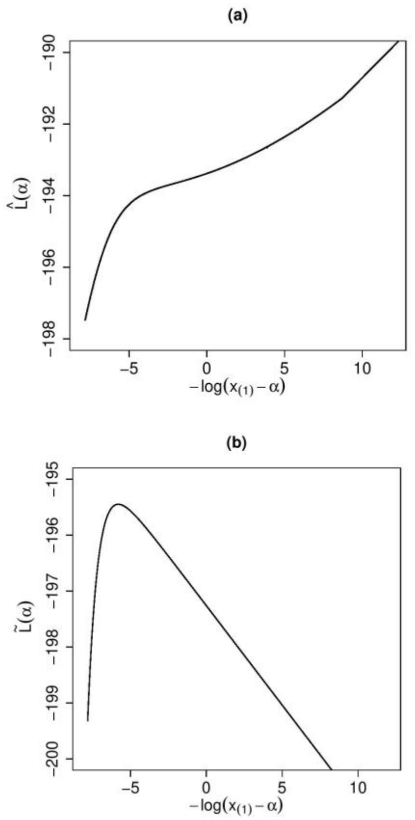

Figure 1: Partially log-likelihood functions for the beach pollution data. (a): the partially log-likelihood function $\hat{L}(\alpha)$ based on ML method applied to ungrouped data and (b) the partially log-likelihood function $\tilde{L}(\alpha)$ based on the EM algorithm applied to grouped data with 3 classes.

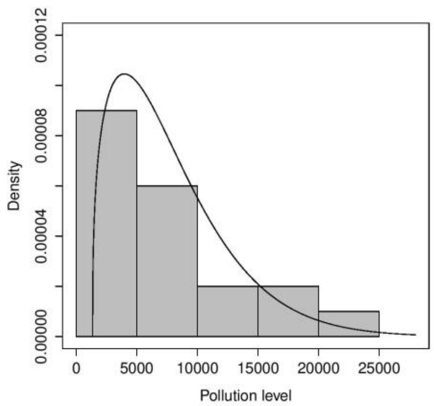

Figure 2: Histogram of the beach pollution data. Superimposed are the fitted Weibull pdfs whose parameters are estimated through EM algorithm applied to grouped beach pollution data into 5 classes. It should be noted that both pdfs coincide with each others.

As it is seen from Figure 1(a), the ML estimator does not exist since the partially maximized log-likelihood function has no local maximum (Cheng and Amin, 1983). If data are grouped into 3 classes, $\tilde{L}(\alpha) = \log L(\Theta_{EM})$ where $\Theta_{EM} = (\gamma_{EM}, \beta_{EM}, \alpha)^T$, as shown in Figure 1(b), has a global maximum. The EM and ML estimators are $\widehat{\Theta} = (1.3555{,}7020.8200{,}1364)^T$ and $\widehat{\Theta} = (1.3548{,}7018.3900{,}1363.8800)^T$, rspectively. For this, the EM estimator are used as the initial values to obtain the ML estimator. It is clear that the ML estimator of the three-parameter Weibull model breaks down and so the EM algorithm can find the ML estimator when data if data are grouped. The fitted Weibull distribution whose parameters are estimated via the EM and ML methods applied to the grouped beach pollution data into 5 classes are shown in Figure 2.

For the third sample of ungrouped maximum flood level data, the ML estimator for the BS distribution fitted to the breaks down. The partially log-likelihood function, i.e., $\hat{L}(\alpha) = \log L(\Theta_{ML})$ where $\Theta_{ML} = (\gamma_{ML}, \beta_{ML}, \alpha)^T$ is shown in Figure 3(a). As it is seen, the ML estimator does not exist since the partially maximized log-likelihood function has no local maximum. Based on data grouped into 3 classes, the graph of $\tilde{L}(\alpha) = \log L(\Theta_{EM})$ has global maximum as shown in figure 3(b) where $\Theta_{EM} = (\gamma_{EM}, \beta_{EM}, \alpha)^T$. Here, γ_{EM} and β_{EM} are estimators of γ and β via the EM algorithm. The EM and ML estimators are $\widehat{\Theta} = (0.5719{,}0.1689{,}0.2149)^T$ and $\widehat{\Theta} = (0.9028{,}0.1054{,}0.2571)^T$, respectively. Also, the EM estimator are used as the initial values in order to obtain the ML estimator. This example shows that if the ML estimator of the three-parameter Weibull distribution breaks down, then the EM algorithm can find the ML estimator when data are treated as grouped. This example reveals that if ML estimator for BS distribution breaks own, then by grouping the originally ungrouped data, we can always find the EM and maybe the ML estimator of the parameters. The fitted BS distribution whose parameters are estimated via the EM and ML methods applied to the grouped maximum flood level data into 3 classes are shown in Figure 4.

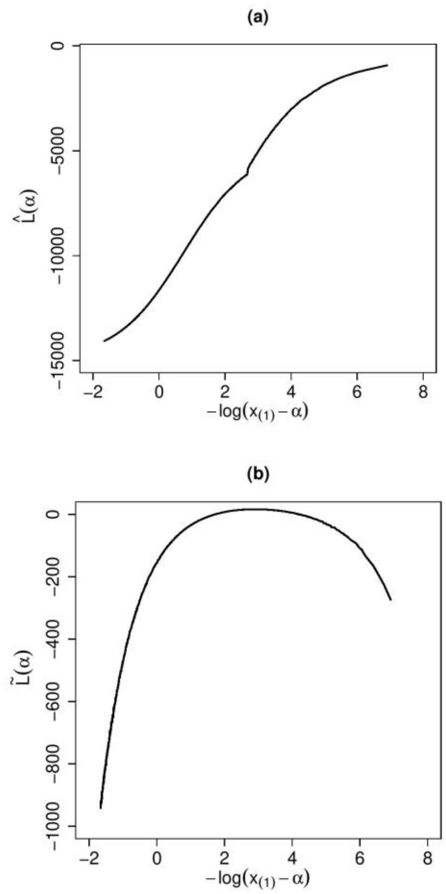

Figure 3: Partially log-likelihood function for the maximum flood level data. (a): the partially log-likelihood function $\hat{L}(\alpha)$ based on ML method applied to ungrouped data and (b) the partially log-likelihood function $\tilde{L}(\alpha)$ based on the EM algorithm applied to grouped data with 3 classes.

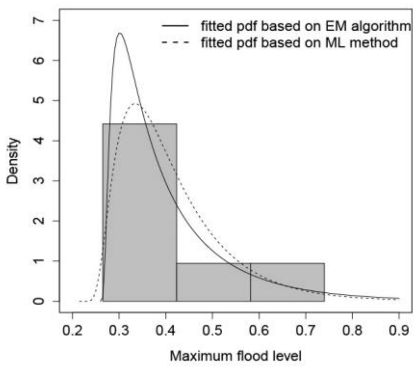

Figure 4: Histogram of the maximum flood level data. Superimposed are fitted BS pdfs whose parameters are estimated through EM algorithm and ML method applied to maximum flood level data grouped into 3 classes.

A simulation study was conducted on the Weibull distribution, with parameters γ, β, and α following uniform distributions of ranges $(0.15,15)$, $(0.1,10)$, and $(0,30)$, respectively. The results are given in Table 5. It is worth noting that since the EM method always converged, we do not consider it in simulation study. In what follows, the inferential methods for BS and Weibull distributions will be given and then these methds are used to model three sets of real data described in Material section. The simulation results, i.e., the percentage of repetitions that the NR method converged truly, are given in Table 4.

Table 4: Percentage of repetitions that NR method truly converged for the BS distribution. The dash "-", sign in each cell indicates that the simulation was not performed for that combination of number of classes and sample size.

	Sample size			
Number of classes	50	100	250	500
5	81%	77%	80%	75%
10	84%	73%	76%	72%
20	82%	74%	74%	70%
40	-	-	-	71%

Table 5: Percentage of repetitions that NR method truly converged for the Weibull distribution. The dash "-", sign in each cell indicates that the simulation was not performed for that combination of number of classes and sample size.

	Sample size			
Number of classes	50	100	250	500
5	39%	36%	34%	33%
10	38%	35%	35%	32%
20	35%	33%	32%	30%
40	-	-	-	33%

Discussion

Here, we give methods for computing the ML estimators of the parameters of the three-parameter BS and Weibull distributions fitted to the ungrouped and grouped data. Suppose $x_1, \dots, x_n$ denote a sample of n independent realizations coming from either three-parameter BS or Weibull distributions. We suggest the algorithm given by the following for estimating the parameters of these families.

1. Classify the data into m separate groups with lower bounds $r_0, r_1, \dots, r_m$ where $r_0 = \min\{x_1, \dots, x_n\}$, $r_m = \max\{x_1, \dots, x_n\}$, and frequency of each group is f_i, for $i = 1, \dots, m$. A schematic representation of the grouping procedure is shown in Table 7;
2. We suggest that $m = 3$ for $n \leq 20$, and $3 \leq m \leq n/8$ otherwise;
3. Apply the EM algorithm as described in Appendices B and D for estimating the parameters of the the three-parameter BS and Weibull distributions, respectively;

Further study shows that grouping the pollution and maximum flood level data into 4, 5, or more classes yields still the $\tilde{L}(\alpha)$ unimodal but estimations based on EM algorithm are slightly different. The above algorithm is very useful when a location family of statistical distributions that does not satisfy the regularity conditions (i.e., the dependency the support of random variable on the parameter) is fitted to the ungrouped data. If the ML estimator breaks down, we use the estimator obtained through the EM algorithm as described by the above, otherwise we use the EM estimator as the initial values for the NR method to obtain the ML estimator whose methodology is described in Appendices A and C for three-parameter BS and Weibull distributions, respectively. When data are given in grouped form as shown in Table 7, we suggest to use first the EM algorithm to estimate the parameters and then use the estimated parameters as the initial values for NR method for computing the ML estimator.

Conclusion

We have carried out a simulation study and discovered that percentage of failed attempts to reach convergence through the Newton-Raphson (NR) method is considerable (say, on the average 25% for the Birnbaum-Saunders (BS) distribution and 65% for the Weibull distribution) when these models fitted to the raw or ungrouped data. Also, under some situations, the ML estimator may break down. In both cases, either NR method does not converge or maximum likelihood (ML) estimator breaks down, we suggest using the expectation-maximization (EM) algorithm or ML method for estimating parameters of the three-parameter BS and Weibull distributions when these models are fitted to the grouped data. If ML estimator exists, the estimated parameters through EM algorithm or ML method applied to the grouped data can be used as the initial values for the NR method to obtain the ML estimator. Otherwise, the EM estimator of the grouped data can be considered as the desired estimators. We have demonstrated the performance of the EM algorithm by three real examples including tree's diameter at breast height (DBH), water pollution level, and maximum flood level. For the first example, the developed EM algorithm for grouped data revealed that the Weibull model outperforms the BS distribution for DBH modelling. In the case of the second and the third examples, the ML estimator breaks down but the EM algorithm works efficiently. The number of classes for turning the ungrouped data into grouped data can be considered as an interesting problem for future work. However, as a rule of thumb, we suggested using $m = 3$ for $n \leq 20$, and $3 \leq m \leq n/8$ otherwise. The EM algorithm presented in this work can be applied to obtain the ML estimators for the parameters of a distribution fitted to ungrouped data when the ML estimators break down.

Appendix A: Estimation parameters of the Weibull distribution fitted to the grouped data using the ML method

The log-likelihood function of the three-parameter Weibull distribution with cdf given in (4) is given by

$$l(\Theta) \propto \sum_{i=1}^{m} f_i \log\left[\exp\left\{-\left(\frac{r_{i-1}-\alpha}{\beta}\right)^{\gamma}\right\} - \exp\left\{-\left(\frac{r_i-\alpha}{\beta}\right)^{\gamma}\right\}\right]. \quad (9)$$

The ML estimators of the γ, β, and α are obtained by solving simultaneously the first derivatives of right-hand side of (9) with respect to γ, β, and α, respectively. It follows that

$$\frac{\partial l(\Theta)}{\partial\gamma} = \sum_{i=1}^{m} \frac{f_i}{dG_i(\Theta)} [(\frac{r_i-\alpha}{\beta})^\gamma \log\left(\frac{r_i-\alpha}{\beta}\right)) \times \exp\{(\frac{r_{i-1}-\alpha}{\beta})^\gamma\} - (\frac{r_{i-1}-\alpha}{\beta})^\gamma \times \log(\frac{r_{i-1}-\alpha}{\beta})\exp\{(\frac{r_i-\alpha}{\beta})^\gamma\}] = 0, \quad (10)$$

$$\frac{\partial l(\Theta)}{\partial\beta} = \frac{\gamma}{\beta}\sum_{i=1}^{m} \frac{f_i}{dG_i(\Theta)} [(\frac{r_{i-1}-\alpha}{\beta})^\gamma \exp\{(\frac{r_i-\alpha}{\beta})^\gamma\} -(\frac{r_i-\alpha}{\beta})^\gamma \exp\{(\frac{r_{i-1}-\alpha}{\beta})^\gamma\}] = 0, \quad (11)$$

$$\frac{\partial l(\Theta)}{\partial\alpha} = \gamma\sum_{i=1}^{m} \frac{f_i}{dG_i(\Theta)} [\frac{1}{(r_{i-1}-\alpha)}(\frac{r_{i-1}-\alpha}{\beta})^\gamma \times \exp\{(\frac{r_i-\alpha}{\beta})^\gamma\} - \frac{1}{(r_i-\alpha)}(\frac{r_i-\alpha}{\beta})^\gamma \times \exp\{(\frac{r_{i-1}-\alpha}{\beta})^\gamma\}] = 0, \quad (12)$$

where

$$dG_i(\Theta) = \exp\{(\frac{r_i-\alpha}{\beta})^\gamma\} - \exp\{(\frac{r_{i-1}-\alpha}{\beta})^\gamma\}.$$

Solving simultaneously Equations (10), (11), and (12), yields the ML estimators of the parameters γ, β, and α. We use the optim (,) command in R environment for this aim.

Appendix B: Estimation parameters of the Weibull distribution fitted to the grouped data using the EM algorithm

The complete data log-likelihood function is given by

$$l_c(\Theta) \propto \sum_{i=1}^{m}\sum_{j=1}^{f_i} \log g(x_{ij}|\Theta),$$

where $g(.|\Theta)$ is given by (2). It follows that

$$l_c(\Theta) \propto n\log\left(\frac{\gamma}{\beta}\right) + (\gamma-1) \times \sum_{i=1}^{m}\sum_{j=1}^{f_i}\left(\log\left(\frac{x_{ij}-\alpha}{\beta}\right) - (\frac{x_{ij}-\alpha}{\beta})^\gamma\right). \quad (13)$$

Differentiating right-hand side of (13) with respect to γ and β, we have

$$\frac{\partial l_c(\Theta)}{\partial\gamma} = \frac{n}{\gamma} + \sum_{i=1}^{m}\sum_{j=1}^{f_i} \log\left(\frac{x_{ij}-\alpha}{\beta}\right) - \sum_{i=1}^{m}\sum_{j=1}^{f_i} \log(\frac{x_{ij}-\alpha}{\beta})(\frac{x_{ij}-\alpha}{\beta})^\gamma,$$

$$\frac{\partial l_c(\Theta)}{\partial\beta} = -\frac{n}{\beta} - \frac{n(\gamma-1)}{\beta} + \gamma\beta^{-\gamma-1}\sum_{i=1}^{m}\sum_{j=1}^{f_i}(x_{ij}-\alpha)^\gamma.$$

Assume that we are at t-th iteration of the EM algorithm. By equating the expected values of both sides of partial derivatives to zero, we obtain

$$\gamma = \frac{n}{\sum_{i=1}^{m} f_i E_{1i}^{(t)} - \sum_{i=1}^{m} f_i E_{2i}^{(t)}},$$

$$\beta = (\frac{\sum_{i=1}^{m} f_i E_{3i}^{(t)}}{n})^{\frac{1}{\gamma}},$$

where

$$E_{1i}^{(t)} = E(\log(\frac{X_{ij}-\alpha}{\beta})|X_{ij},\Theta^{(t)}) = \frac{1}{D_i}\int_{r_{i-1}}^{r_i} \log(\frac{x-\alpha^{(t)}}{\beta^{(t)}}) g(x|\Theta^{(t)})dx, \quad (14)$$

$$E_{2i}^{(t)} = E(\log(\frac{X_{ij}-\alpha}{\beta})(\frac{X_{ij}-\alpha}{\beta})^\gamma|X_{ij},\Theta^{(t)}) = \frac{1}{D_i}\int_{r_{i-1}}^{r_i} (\frac{x-\alpha^{(t)}}{\beta^{(t)}})^{\gamma^{(t)}} \log(\frac{x-\alpha^{(t)}}{\beta^{(t)}}) g(x|\Theta^{(t)})dx, \quad (15)$$

$$E_{3i}^{(t)} = E((X_{ij}-\alpha)^\gamma|X_{ij},\Theta^{(t)}) = \frac{1}{D_i}\int_{r_{i-1}}^{r_i}(x-\alpha^{(t)})^{\gamma^{(t)}} g(x|\Theta^{(t)})dx, \quad (16)$$

in which $X_{ij} \in (r_{i-1}, r_i\,]$, $D_i = G(r_i|\Theta^{(t)}) - G(r_{i-1}|\Theta^{(t)})$, for $i = 1, \dots, m$, $G(r_0|\Theta^{(t)}) = 0$, and $G(.|\Theta^{(t)})$ is defined as (4). The location parameter can be updated through conditionally maximization (CM) step. So, at t-th iteration of the algorithm, the location parameter is updated as

$$\alpha^{(t+1)} = \underset{\alpha}{argmax} \sum_{i=1}^{m} f_i \log[G(r_i|\overline{\Theta}) - G(r_{i-1}|\overline{\Theta})],$$

where $\overline{\Theta} = (\gamma^{(t+1)}, \beta^{(t+1)}, \alpha)^T$, $r_0 = \min\{x_1, \dots, x_n\}$, and $G(r_0|\overline{\Theta}) = 0$. The M-step is complete. Fortunately, all three quantities $E_{1i}^{(t)}$, $E_{2i}^{(t)}$, and $E_{3i}^{(t)}$ given, respectively, in (14), (15), and (16) have closed-form expressions. Suppose $\gamma^{(t+1)}$, $\beta^{(t+1)}$, and $\alpha^{(t+1)}$ are updated values of the parameters γ, β, and α, respectively, at $(t+1)$-th iteration. We stop the EM algorithm if

$$\max\{|\gamma^{(t+1)}-\gamma^{(t)}|, |\beta^{(t+1)}-\beta^{(t)}|, |\alpha^{(t+1)}-\alpha^{(t)}|\} < 10^{-5}.$$

The initial values for starting the EM algorithm are obtained through the method of moments. Consider the transformation $Y = X - \alpha$, where X follows the Weibull distribution with pdf (2). It follows that

$$E(Y) = \beta\Gamma(1+\frac{1}{\gamma}),$$

$$E(Y^2) = \beta^2\Gamma(1+\frac{2}{\gamma}).$$

Equating the sample moment to the corresponding population moment up to the second order, the equations given by the following are used to find the moment-based estimators for the grouped data.

$$\frac{\beta}{\gamma}\Gamma\left(\frac{1}{\gamma}\right) = \frac{1}{n}\sum_{i=1}^{m} f_i E(Y_{ij}|Y_{ij}, \Theta^{(t)}), \quad (17)$$

$$\frac{2\beta^2}{\gamma}\Gamma\left(\frac{2}{\gamma}\right) = \frac{1}{n}\sum_{i=1}^{m} f_i E\left(Y_{ij}^2|Y_{ij}, \Theta^{(t)}\right), \quad (18)$$

where $Y_{ij} \in (r'_{i-1}, r'_i]$, for $i = 1, \dots, m$. Both expectations given in the right-hand side of (17) and (18) are expressed in terms of incomplete gamma function. To avoid computational complexity, we use the mid-point approximation of these expressions. It follows that

$$\beta\Gamma(1+\frac{1}{\gamma}) = \frac{1}{n}\sum_{i=1}^{m} f_i(\frac{r'_{i-1}+r'_i}{2}), \quad (19)$$

$$\beta^2\Gamma(1+\frac{2}{\gamma}) = \frac{1}{n}\sum_{i=1}^{m} f_i(\frac{r'_{i-1}+r'_i}{2})^2, \quad (20)$$

where $r'_0 = \min\{y_1, \dots, y_n\}$ and $r'_i = r_i - \alpha$, for $i = 1, \dots, m$. Solving simultaneously Equations (19) and (20), the initial values $\gamma^{(0)}$ and $\beta^{(0)}$ are obtained. For the location parameter, we set the minimum of original observations as the initial value, i.e., $\alpha^{(0)} = \min\{x_1, \dots, x_n\}$ when $x_1, \dots, x_n$ are independent realizations form the three-parameter Weibull distribution.

Acknowledgements

The authors would like to thank two anonymous referees for their constructive comments that greatly improved the quality of this paper.

References

Almalki, S.J. and Nadarajah, S. 2014. Modifications of the Weibull distribution: A review, Reliability Engineering and System Safety. 124, 32-55.

Ambrozic, M. and Vidovic, K. 2007. Computation of the parameters of the Weibull distribution for estimating the bending strength of corrugated roofing sheets. Materiali in Tehnologije. 41, 179-184.

Bailey, R.L. and Dell, T. 1973. Quantifying diameter distributions with the Weibull function. Forest Science. 19, 97-104.

Balakrishnan, N. and Kundu, D. 2019. Birnbaum-Saunders distribution: A review of models, analysis, and applications. Applied Stochastic Models in Business and Industry. 35, 4-49.

Barros, M., Paula, G.A. and Leiva, V. 2008. A new class of survival regression models with heavy-tailed errors: Robustness and Diagnostics, Lifetime Data Analysis. 14, 316-332.

Bartolucci, A.A., Singh, K.P., Bartolucci, A.D. and Bae, 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.

Birnbaum, Z.W. and Saunders, S.C. 1969. A new family of life distributions. Journal of Applied Probability. 6, 319-327.

Cao, Q.V. 2004. Predicting parameters of a Weibull function for modeling diameter distribution. Forest science. 50, 682-685.

Celeux, G. 1985. The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. Computational Statistics Quarterly. 2, 73-82.

Cheng, R. and Amin, N. 1983. Estimating parameters in continuous univariate distributions with a shifted origin, Journal of the Royal Statistical Society: Series B (Methodological). 45, 394-403.

Cohen, C.A. and Whitten, B. 1982. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. Communications in Statistics-Theory and Methods. 11, 2631-2656.

Cousineau, D. 2009. Nearly unbiased estimators for the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method. British Journal of Mathematical and Statistical Psychology. 62, 167-191.

Cran, G. 1988. Moment estimators for the 3-parameter Weibull distribution. IEEE Transactions on Reliability. 37, 360-363.

Dempster, A.P., Laird, N.M. and Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (methodological). 1-38.

Dumonceaux, R. and Antle, C.E. 1973. Discrimination between the log-normal and the Weibull distributions. Technometrics. 15, 923-926.

Gove, J.H. and Fairweather, S.E. 1989. Maximum-likelihood estimation of Weibull function parameters using a general interactive optimizer and grouped data. Forest Ecology and Management. 28, 61-69.

Green, E.J., Roesch, J.F.A., Smith, A.F. and Strawderman, W.E. 1994. Bayesian estimation for the three-parameter Weibull distribution with tree diameter data. Biometrics. 50, 254-269.

Hassani, H., Kalantari, M. and Ghodsi, Z. 2019. Spatial gradient of bicoid is well explained by Birnbaum-Saunders distribution. Medical Hypotheses. 122, 73-81.

Johnson, N.L., Kotz, S. and Balakrishnan, N. 1994. Continuous Univariate Distributions, John Wiley & Sons.

Kundu, D., Kannan, N. and Balakrishnan, N. 2008. On the hazard function of Birnbaum-Saunders distribution and associated inference, Computational Statistics and Data Analysis. 52, 2692- 2702.

Kuo-Chao, L., Keng-Tung, W. and Chien-Song, C. 2009. A new study on combustion behavior of pine sawdust characterized by the Weibull distribution. Chinese Journal of Chemical Engineering. 17, 860-868.

Lai, C.D., Murthy, D. and Xie, M. 2006. Weibull Distributions and Their Applications, Springer. Handbook of Engineering Statistics. 63-78.

Lei, Y. 2008. Evaluation of three methods for estimating the Weibull distribution parameters of Chinese pine (*Pinus tabulaeformis*). Journal of Forest Science. 54, 566-71.

Leiva, V. 2015. The Birnbaum-Saunders Distribution. Academic Press.

Leiva, V. Barros, M., Paula, G.A. and Galea, M. 2007. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. Computational Statistics and Data Analysis. 51, 5694-5707.

Lio, Y.L. and Park, C. 2008. A bootstrap control chart for Birnbaum-Saunders percentiles. Quality and Reliability Engineering International. 24, 585-600.

Little, R.J. and Rubin, D.B. 2004. Incomplete data. Encyclopedia of Statistical Sciences. 5.

Liu, C., and Rubin, D.B. 1994. The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. Biometrika. 81, 633-648.

Maltamo, M., Kangas, A., Uuttera, J., Torniainen, T. and Saramäki, J. 2000. Comparison of percentile based prediction methods and the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. Forest Ecology and Management. 133, 263-274.

McLachlan, G. and Krishnan, T. 2007. The EM Algorithm and Extensions, volume 382. John Wiley and Sons.

Meng, X.L. and Rubin, D.B. 1993. Maximum likelihood estimation via the ECM algorithm: A general framework, Biometrika. 80, 267-278.

Merganič, J. and Sterba, H. 2006. Characterisation of diameter distribution using the Weibull function: method of moments. European Journal of Forest Research. 125, 427-439.

Murthy, D.P., Xie, M. and Jiang, R. 2004. Weibull Models, volume 505. John Wiley and Sons.

Nagatsuka, H., Kamakura, T. and Balakrishnan, N. 2013. A consistent method of estimation for the three-parameter Weibull distribution. Computational Statistics and Data Analysis. 58, 210- 226.

Podlaski, R. 2008. Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. Canadian Journal of Forest Research. 38, 518-527.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: https://www.R-project.org/.

Schmidt, L.N., Sanquetta, M.N.I., McTague, J.P., da Silva, G.F., Fraga Filho, C.V., Sanquetta, C.R., Soares Scolforo, J.R. 2020. On the use of the Weibull distribution in modeling and describing diameter distributions of clonal eucalypt stands. Canadian Journal of Forest Research.

Schworer, A. and Hovey, P. 2004. Newton-Raphson versus Fisher scoring algorithms in calculating maximum likelihood estimates. Dayton.

Stankova, T.V. and Zlatanov, T.M. 2010. Modeling diameter distribution of Austrian black pine (*Pinus nigra* arn.) plantations: a comparison of the Weibull frequency distribution function and percentile-based projection methods, European Journal of Forest Research. 129, 1169-1179.

Steen, P. and Stickler, D. 1976. A sewage pollution study of beaches from Cardiff to Ogmore. UWIST, Dept. of Applied Biology Report, Cardiff.

Teimouri, M., Doser, J.W. and Finley, A.O. 2020. Forestfit: An R package for modeling plant size distributions. Environmental Modelling and Software. 104668.

Teimouri, M., Hoseini, S.M. and Nadarajah, S. 2013 a. Comparison of estimation methods for the Weibull distribution. Statistics. 47, 93-109.

Teimouri, M., Hosseini, S.M. and Nadarajah, S., 2013 b. Ratios of Birnbaum-Saunders random variables, Quality Technology and Quantitative Management. 10, 457-481.

Teimouri, M. 2020. Bias corrected maximum likelihood estimators under progressive type-I interval censoring scheme, https://doi.org/10.1080/03610918.2020.1819320.

Zhang, L. and Liu, C. 2006. Fitting irregular diameter distributions of forest stands by Weibull, modified Weibull, and mixture Weibull models. Journal of Forest Research. 11, 369-372.

Zhu, H.P., Xia, X., Chuan, H.Y., Adnan, A., Liu, S.F. and Du, Y.K., 2011. Application of Weibull model for survival of patients with gastric cancer. BMC Gastroenterology. 11, 1.