Inference and Optimal Censoring Schemes for Progressively Censored Birnbaum-Saunders Distribution

Biswabrata Pradhan [*]& Debasis Kundu [†]

Abstract

The aim of this article is two fold. First we discuss the maximum likelihood estimators of the unknown parameters of a two-parameter Birnbaum-Saunders distribution when the data are progressively Type-II censored. The maximum likelihood estimators are obtained using the EM algorithm by exploiting the property that the Birnbaum-Saunders distribution can be expressed as an equal mixture of an inverse Gaussian distribution and its reciprocal. From the proposed EM algorithm, the observed information matrix can be obtained quite easily, which can be used to construct the asymptotic confidence intervals. We perform the analysis of two real and one simulated data sets for illustrative purposes, and the performances are quite satisfactory. We further propose the use of different criteria to compare two different sampling schemes, and then find the optimal sampling scheme for a given criterion. It is observed that finding the optimal censoring scheme is a discrete optimization problem, and it is quite a computer intensive process. We examine one sub-optimal censoring scheme by restricting the choice of censoring schemes to one-step censoring schemes as suggested by Balakrishnan (2007), which can be obtained quite easily. We compare the performances of the sub-optimal censoring schemes with the optimal ones, and observe that the loss of information is quite insignificant.

Kew Words and Phrases; Maximum likelihood estimation; EM algorithm; Progressive Censoring Scheme; Fisher information matrix; Inverse Gaussian distribution.

[*]SQC & OR Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata, Pin 700108, India

[†]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. Corresponding author, e-mail: kundu@iitk.ac.in

1 Introduction

Birnbaum and Saunders (1969a, 1969b) introduced a failure time distribution for fatigue failure which is caused due to periodic stress. A more general derivation was provided by Desmond (1985) based on a biological model. He also strengthened the physical justification for the use of this distribution by relaxing the assumptions originally made by Birnbaum and Saunders (1969a). Although the Weibull distribution may be the most widely used failure time distribution, recently the Birnbaum-Saunders distribution has received considerable attention mainly due to the different shapes of the probability density function (PDF) and due to the non-monotonicity property of the hazard function.

The cumulative distribution function (CDF) of a two-parameter Birnbaum-Saunders random variable T is of the form;

$$F(t;\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left(\left\{\frac{t}{\beta}\right\}^{\frac{1}{2}} - \left\{\frac{\beta}{t}\right\}^{\frac{1}{2}}\right)\right]; \quad t > 0. \tag{1}$$

Here $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively, and $\Phi(\cdot)$ is the CDF of a standard normal distribution. From here on, a Birnbaum-Saunders distribution with shape and scale parameters α and β, respectively, as in (1), will be denoted by $\mathrm{BS}(\alpha,\beta)$. It can be easily seen that extensive work has been done on different aspects of Birnbaum-Saunders distribution since its introduction. A comprehensive review on Birnbaum-Saunders distribution since the mid 1990's has been given by Johnson *et al.* (1995). For some later developments, the reader is referred to Dupis and Mills (1998), Rieck (1995, 1999), From and Li (2006), Owen (2006), Xie and Wei (2007), Ahmed *et al.* (2008), Balakrishnan *et al.* (2011) and see the references cited therein.

Since the CDF of a Birnbaum-Saunders random variable can be written in terms of $\Phi(\cdot)$, it can be used quite effectively even for censored data. Traditionally, Type-I and Type-II were the most popular censoring schemes which have been used in practice. Unfortunately, none of these allow the removal of any experimental units during the experiment. Type-I and Type-II progressive censoring allow removal of the experimental unit during the experiment. Because of this flexibility, progressive censoring scheme became quite popular and has received considerable attention in the applied statistics literature during the last 10-12 years. The recent review article by Balakrishnan (2007) provided details on the different estimation procedures in the case of progressive censoring and its different applications. Recently, some attention has been given to construct the optimal progressive censoring scheme. This proves to be a very practical and challenging problem.

The aim of this article is two fold. First we consider the maximum likelihood estimators (MLEs) of the unknown parameters of the Birnbaum-Saunders distribution when the data are Type-II progressively censored. It is well known that even in the complete sample case, the MLEs cannot be obtained in closed form. Also the standard Newton-Raphson algorithm may not always converge.

Jorgensen *et al.* (1991) observed that the two-parameter Birnbaum-Saunders distribution can be expressed as an equal mixture of an inverse Gaussian distribution and its reciprocal. It thus makes it very natural to use the EM algorithm to compute the MLEs. In each expectation-step ('E'-step) of the proposed EM algorithm the maximization of the pseudo log-likelihood function ('M'-step) can be performed very easily, and they have explicit solutions. Therefore, the implementation of the proposed EM algorithm is very simple. Some initial values of the unknown parameters are also proposed. Moreover, at the last step of the EM algorithm, using the idea of Louis (1982), the observed Fisher information matrix can be easily computed, which can be used to construct the asymptotic confidence intervals of the unknown parameters. We have analyzed three data sets, two real data sets and one simulated one, mainly for illustrative purpose. It is observed that the performances of the proposed EM algorithm is quite satisfactory.

Our second aim of this paper is to construct the optimal progressive censoring schemes. Different criteria are available in the literature to compare two different censoring schemes in terms of their information content. The most popular ones are the determinant or trace of the expected Fisher information matrix. Zhang and Meeker (2005) uses the variance of the p-th percentile point as an effective criterion. A more general criterion as suggested by Kundu (2008) can also be adopted. We provide different optimal censoring schemes in terms of the different optimality criteria. In this case finding the optimal censoring scheme turns out to be a discrete optimization problem, and is quite computer intensive. We have examined one sub-optimal censoring scheme by restricting the choice of schemes to one-step censoring schemes, as suggested by Balakrishnan (2007), and compare its performance with the optimal one. The performances of the sub-optimal censoring schemes are quite satisfactory, in the sense that the loss of information of the sub-optimal censoring scheme is not very significant.

The rest of the paper is organized as follows. In section 2, we briefly describe the Birnbaum-Saunders distribution and the procedure of progressive censoring. In section 3, we describe the aforementioned EM algorithm. Observed and expected Fisher information matrices are provided in Section 4. The analysis of three data sets are presented in Section 5. We present the optimal censoring schemes in Section 6, and finally conclude the paper in section 7.

2 Preliminaries

2.1 Birnbaum-Saunders Distribution

If $T \sim \mathrm{BS}(\alpha, \beta)$, it has the CDF (1), and the corresponding PDF for $t > 0$ becomes

$$f(t;\alpha,\beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right] exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right]. \quad (2)$$

The shape of the PDF of a two-parameter Birnbaum-Saunders distribution is always uni-modal and it is a right skewed distribution. The mean, variance and other higher order moments can be expressed in explicit forms. For example

$$E(T) = \beta\left(1+\frac{1}{2}\alpha^2\right), \quad V(T) = (\alpha\beta)^2\left(1+\frac{5}{4}\alpha^2\right), \tag{3}$$

$$\beta_1(T) = \frac{16\alpha^2(11\alpha^2+6)}{(5\alpha^2+4)^3}, \quad \beta_2(T) = 3 + \frac{6\alpha^2(93\alpha^2+41)}{(5\alpha^2+4)^2}, \tag{4}$$

where $E(T)$, $V(T)$, $\beta_1(T)$ and $\beta_2(T)$ are the expected value, variance, coefficient of skewness, and the coefficient of kurtosis, respectively. It is clear that as α increases to ∞, $\beta_1(T)$ gradually decreases to zero. It implies that the PDF becomes symmetric as α becomes very large. Note that if we make the transformation

$$X = \frac{1}{2}\left[\left(\frac{T}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{T}\right)^{\frac{1}{2}}\right] \tag{5}$$

or

$$T = \beta\left[1 + 2X^2 + 2X(1+X^2)^{\frac{1}{2}}\right], \tag{6}$$

then it easily follows that X is normally distributed with mean zero and variance $\frac{\alpha^2}{4}$. The transformation (6) is very useful in determining many properties of the Birnbaum-Saunders distribution. For example, establishing different moments as described in (3) and (4), the transformation (6) may be used very effectively. Moreover, generating random sample from the Birnbaum-Saunders distribution can be performed using the normal random variable. It can also be seen, using the above transformation, that if $T \sim \text{BS}(\alpha,\beta)$, then $T^{-1} \sim \text{BS}(\alpha,\beta^{-1})$.

Another interesting relation between Birnbaum-Saunders distribution and inverse Gaussian distribution has been established by Jorgensen *et al.* (1991). Note that a random variable X_1 follows an inverse Gaussian distribution with parameters $\mu > 0$ and $\lambda > 0$, if it has the PDF

$$f_{X_1}(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}; \tag{7}$$

for $x > 0$ and 0 otherwise. From now on, an inverse Gaussian distribution with the PDF (7) will be denoted by $\text{IG}(\mu,\lambda)$. Moreover, suppose X_2 is a random variable such that $X_2^{-1} \sim$

$\mathrm{IG}(\mu^{-1}, \lambda/\mu^2)$. Consider a new random variable X, such that

$$X = \begin{cases} X_1 & \text{with probability } \frac{1}{2} \\ \\ X_2 & \text{with probability } \frac{1}{2}. \end{cases} \tag{8}$$

Clearly, X is a mixture of X_1 and X_2, and the PDF of X is

$$f_X(x;\mu,\lambda) = \frac{1}{2} f_{X_1}(x;\mu,\lambda) + \frac{1}{2} f_{X_2}(x;\mu,\lambda), \tag{9}$$

where $f_{X_1}(\cdot)$ and $f_{X_2}(\cdot)$ are the PDFs of X_1 and X_2 respectively. It may be observed that, see Gupta and Akman (1995), $f_{X_2}(\cdot)$ can be written as

$$f_{X_2}(x;\mu,\lambda) = \frac{x}{\mu} f_{X_1}(x;\mu,\lambda). \tag{10}$$

It can be easily verified that (9) coincides with (2) when $\mu = \beta$ and $\lambda = \beta/\alpha^2$. Therefore, a two-parameter Birnbaum-Saunders distribution can be expressed as an equal mixture of an inverse Gaussian distribution and its reciprocal.

2.2 Type-II Progressive Censoring

Type-II progressive censoring was first introduced by Cohen (1963). A Type-II progressively censored life test is conducted as follows. Put n identical units in a test. Unlike Type-I or Type-II censoring schemes, an item can be removed during the experiment, even if it does not fail. At the time of the first failure, say at $t_{(1)}$, R_1 units from the remaining $n-1$ survival items are removed. Similarly, at the time of the second failure, say at $t_{(2)}$, R_2 units from the remaining $n - R_1 - 1$ items are removed, and so on. Finally, at the time of the m-th failure, say at $t_{(m)}$, the remaining survival units, say R_m, are removed, and the experiment stops at $t_{(m)}$. Clearly, R_m satisfies the following;

$$R_m = n - R_1 - \cdots - R_{m-1} - m. \tag{11}$$

In Type-II progressive censoring experiment, m and $\{R_1, \cdots, R_m\}$ are pre-specified. Therefore, in Type-II progressive censoring experiment, the data will be of the form $\{(t_{(1)}, R_1), \cdots,$

$(t_{(m)}, R_m)\}$. Note that if we put $R_1 = \cdots = R_{m-1} = 0$, then $R_m = n - m$, and it becomes the classical Type-II censoring. In the next section we discuss the maximum likelihood estimation of the unknown parameters, when the lifetime distribution of the experimental units follow two-parameter Birnbaum-Saunders distribution.

3 Maximum Likelihood Estimation

Based on the progressively censored observations $\{(t_{(1)}, R_1), \cdots, (t_{(m)}, R_m)\}$, without the additive constant, the log-likelihood function of the observed data is

$$l(\alpha, \beta | data) = \sum_{i=1}^{m} \{\ln f(t_{(i)}; \alpha, \beta) + R_i \ln(\Phi(-g(t_{(i)}; \alpha, \beta)))\}, \tag{12}$$

where $f(\cdot)$ is the PDF of the Birnbaum-Saunders distribution as defined in (2), and

$$g(t; \alpha, \beta) = \frac{1}{\alpha} \left\{ \left(\frac{t}{\beta} \right)^{\frac{1}{2}} - \left(\frac{\beta}{t} \right)^{\frac{1}{2}} \right\}. \tag{13}$$

We propose the EM algorithm to compute the MLEs, and we use the mixture representation as given in (9). For convenience we consider the re-parameterization (μ, λ) of (α, β), where $\mu = \beta$ and $\lambda = \beta/\alpha^2$, as defined before. The progressive censored data can be viewed as an incomplete data problem (See Ng *et al.*, 2002). The complete observations can be written as follows;

$$\{(T_{(1)}, Z_1), \cdots, (T_{(m)}, Z_m), (Y_{ij}, W_{ij}; j = 1, \ldots, R_i; i = 1, \ldots, m)\}, \tag{14}$$

where $T_{(1)}, \ldots, T_{(m)}$ are observed data, $\{Z_i, Y_{ij}, Wij; j = 1, \ldots, R_i; i = 1, \ldots, m\}$ are not observable. Moreover, Z_i (W_{ij}) is an indicator variable taking value 0 if the observation $T_{(i)}$ (Y_{ij}) comes from X_1 or 1, if it comes from X_2. Due to (9), clearly $P(Z = 0) = P(Z = 1) = \frac{1}{2}$ and $P(W = 0) = P(W = 1) = \frac{1}{2}$. Then the complete data likelihood, involving only the unknown parameters, can be written as

$$L_c(\mu, \lambda) \quad = \quad \prod_{i=1}^{m} \left[f_T(t_{(i)}; \mu, \lambda) \prod_{j=1}^{R_i} f_T(y_{ij}; \mu, \lambda) \right]$$

$$
\begin{aligned}
&= \prod_{i=1}^{m}\left[\left\{f_{X_1}(t_{(i)};\mu,\lambda)\right\}^{1-z_i}\left\{f_{X_2}(t_{(i)};\mu,\lambda)\right\}^{z_i}\right.\\
&\qquad \left.\times\prod_{j=1}^{R_i}\left\{f_{X_1}(y_{ij};\mu,\lambda)\right\}^{1-w_{ij}}\left\{f_{X_2}(y_{ij};\mu,\lambda)\right\}^{w_{ij}}\right]. \qquad (15)
\end{aligned}
$$

The log-likelihood involving only the unknown parameters, becomes

$$
\begin{aligned}
l_c(\mu,\lambda) &= \frac{n}{2}\ln\lambda+\frac{n\lambda}{\mu}-\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}t_{(i)}-\frac{\lambda}{2}\sum_{i=1}^{m}\frac{1}{t_{(i)}}-\ln\mu\sum_{i=1}^{m}z_i-\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}\sum_{j=1}^{R_i}y_{ij}\\
&\quad -\frac{\lambda}{2}\sum_{i=1}^{m}\sum_{j=1}^{R_i}\frac{1}{y_{ij}}-\ln\mu\sum_{i=1}^{m}\sum_{j=1}^{R_i}w_{ij}. \qquad (16)
\end{aligned}
$$

Therefore, at the k-th stage of the EM algorithm, if the values of μ and λ are $\mu^{(k)}$ and $\lambda^{(k)}$ respectively, then at the k-th stage the pseudo log-likelihood function (E-step) becomes

$$
\begin{aligned}
l_c^*(\mu,\lambda) &= \frac{n}{2}\ln\lambda+\frac{n\lambda}{\mu}-\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}t_{(i)}-\frac{\lambda}{2}\sum_{i=1}^{m}\frac{1}{t_{(i)}}-\ln\mu\sum_{i=1}^{m}E\left[Z_i|t_{(i)},\mu^{(k)},\lambda^{(k)}\right]\\
&\quad -\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}\sum_{j=1}^{R_i}E\left[Y_{ij}|Y_{ij}>t_{(i)},\mu^{(k)},\lambda^{(k)}\right]-\frac{\lambda}{2}\sum_{i=1}^{m}\sum_{j=1}^{R_i}E\left[\frac{1}{Y_{ij}}|Y_{ij}>t_{(i)},\mu^{(k)},\lambda^{(k)}\right]\\
&\quad -\ln\mu\sum_{i=1}^{m}\sum_{j=1}^{R_i}E\left[W_{ij}|t_{(i)},\mu^{(k)},\lambda^{(k)}\right]\\
&= \frac{n}{2}\ln\lambda+\frac{n\lambda}{\mu}-\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}t_{(i)}-\frac{\lambda}{2}\sum_{i=1}^{m}\frac{1}{t_{(i)}}-\ln\mu\sum_{i=1}^{m}a_i^{(k)}-\frac{\lambda}{2\mu^2}\sum_{i=1}^{m}R_ib_{1i}^{(k)}\\
&\quad -\frac{\lambda}{2}\sum_{i=1}^{m}R_ib_{2i}^{(k)}-\ln\mu\sum_{i=1}^{m}R_ia_i^{(k)}, \qquad (17)
\end{aligned}
$$

where

$$
\begin{aligned}
b_{1i}^{(k)} &= E\left[X_1|X_1>t_{(i)};\mu^{(k)},\lambda^{(k)}\right]=\frac{1}{C(t_{(i)})}\int_{t_{(i)}}^{\infty}x\left(\frac{\lambda^{(k)}}{2\pi x^3}\right)^{\frac{1}{2}}e^{-\frac{\lambda^{(k)}(x-\mu^{(k)})^2}{2\mu^{(k)2}x}}dx, \qquad (18)\\
b_{2i}^{(k)} &= E\left[\frac{1}{X_1}|X_1>t_{(i)};\mu^{(k)},\lambda^{(k)}\right]=\frac{1}{C(t_{(i)})}\int_{t_{(i)}}^{\infty}\frac{1}{x}\left(\frac{\lambda^{(k)}}{2\pi x^3}\right)^{\frac{1}{2}}e^{-\frac{\lambda^{(k)}(x-\mu^{(k)})^2}{2\mu^{(k)2}x}}dx, \qquad (19)\\
C(t_{(i)}) &= P[X_1>t_{(i)}]=1-\Phi\left(\sqrt{\frac{\lambda}{t_{(i)}}}\left(\frac{t_{(i)}}{\mu}-1\right)\right)-\exp\left(\frac{2\lambda}{\mu}\right)\Phi\left(-\sqrt{\frac{\lambda}{t_{(i)}}}\left(\frac{t_{(i)}}{\mu}+1\right)\right)
\end{aligned}
$$

and

$$
\begin{aligned}
a_i^{(k)} &= E(Z_i | t_{(i)}, \mu^{(k)}, \lambda^{(k)}) \\
&= \frac{\frac{1}{2} f_{X_2}(t_{(i)}; \mu^{(k)}, \lambda^{(k)})}{\frac{1}{2} f_{X_1}(t_{(i)}; \mu^{(k)}, \lambda^{(k)}) + \frac{1}{2} f_{X_2}(t_{(i)}; \mu^{(k)}, \lambda^{(k)})} \\
&= \frac{t_{(i)}}{t_{(i)} + \mu^{(k)}}. \qquad (20)
\end{aligned}
$$

The maximization (M-step) of (17) with respect to μ and λ can be performed very easily we obtain $\mu^{(k+1)}$ and $\lambda^{(k+1)}$ as follows;

$$
\mu^{(k+1)} = \frac{(2\bar{z}^{(k)} - 1) + \sqrt{(2\bar{z}^{(k)} - 1)^2 + 4\bar{z}^{(k)}(1 - \bar{z}^{(k)})\bar{t}^{(k)} s_1^{(k)}}}{2\bar{z}^{(k)} s_1^{(k)}} \qquad (21)
$$

and

$$
\lambda^{(k+1)} = \frac{(\mu^{(k+1)})^2}{\bar{t}^{(k)} - 2\mu^{(k+1)} + (\mu^{(k+1)})^2 s_1^{(k)}}, \qquad (22)
$$

where

$$
\bar{t}^{(k)} = \frac{1}{n} \sum_{i=1}^{m} (t_i + R_i b_{1i}^{(k)}), \quad s_1^{(k)} = \frac{1}{n} \sum_{i=1}^{m} \left[\frac{1}{t_i} + R_i b_{2i}^{(k)} \right] \quad \text{and} \quad \bar{z}^{(k)} = \frac{1}{n} \sum_{i=1}^{m} (a_i^{(k)} + R_i a_i^{(k)}).
$$

Now we discuss how to choose the initial values. The following two facts will be useful to find the initial values of the algorithm of μ and λ. Note that if $\{x_1, \cdots, x_n\}$ is a random sample from X_1, then the MLEs of μ and λ are

$$
\bar{x} \quad \text{and} \quad \left(\frac{1}{n} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{x_i \bar{x}^2} \right)^{-1}, \qquad (23)
$$

respectively. Similarly, if $\{x_1, \cdots, x_n\}$ is a random sample from X_2, then the MLEs of μ and λ are

$$
\frac{1}{s_1} \quad \text{and} \quad \left(\frac{1}{n} \sum_{i=1}^{n} \left(x_i^{-1} - s_1 \right)^2 x_i \right)^{-1}, \qquad (24)
$$

respectively, where $s_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{x_i}$. Unfortunately, in this case, not all the x_i's are known. Some of the x_i's are truncated at $t_{(j)}$. If x_i is truncated at $t_{(j)}$, we substitute the corresponding

x_i by $t_{(j)}$ into (23) and (24). We suggest to use the simple averages of these estimates as initial values of μ and λ, *i.e.*

$$\mu^{(0)} = \frac{1}{2}\left[\bar{x} + \frac{1}{s_1}\right] \text{ and } \lambda^{(0)} = \frac{1}{2}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{x_i\bar{x}^2}\right)^{-1} + \left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i^{-1}-s_1\right)^2 x_i\right)^{-1}\right]. \quad (25)$$

Therefore, we propose the following algorithm to compute the MLEs of the parameters of the Birnbaum-Saunders distribution.

ALGORITHM:

Step 1: Compute $\mu^{(0)}$ and $\lambda^{(0)}$ using (25).

Step 2: Compute $b_{1i}^{(0)}$, $b_{2i}^{(0)}$ numerically using (18), (19), respectively and $a_i^{(0)}$ using (20) for $i = 1, \cdots, m$.

Step 3: Compute $\mu^{(1)}$ and $\lambda^{(1)}$ using (21) and (22), respectively.

Step 4: Check the convergence, if not satisfied, go back to Step 2.

4 Fisher Information Matrices

In this section we provide both the observed and expected Fisher information matrices. We will be using the missing value principle of Louis (1982); see for example Ng *et al.* (2002, 2004), Balakrishnan *et al.* (2008) also in this connection. The observed information matrix will be useful to construct the asymptotic confidence intervals of the unknown parameters, where as the expected Fisher information matrix will be used to determine the optimal censoring scheme.

4.1 Observed Fisher Information Matrix

Louis (1982) provided the general methodology to compute the observed Fisher information matrix from the last step of the EM algorithm. For convenience, we are presenting the observed Fisher information matrix obtained from the EM algorithm. Using the same notation as in Louis (1982), the observed Fisher information matrix can be written as

$$\boldsymbol{F}_{obs} = \boldsymbol{B} - \boldsymbol{S}\boldsymbol{S}^T,$$

where $\boldsymbol{B}$ is the negative of the second derivative of the log-likelihood function (17) and $\boldsymbol{S}$ is the first derivative vector. We provide the elements of the matrix $\boldsymbol{B} = ((B(i,j)))$, the vector $\boldsymbol{S} = (S(j))$ and it is assumed that the EM algorithm stops at the k-step. We then have

$$B(1,1) = \frac{n\lambda^{(k)}}{(\mu^{(k)})^2}\left[\frac{3\bar{t}^{(k)}}{(\mu^{(k)})^2} - \frac{2}{\mu^{(k)}}\right] - \frac{n\bar{z}^{(k)}}{(\mu^{(k)})^2}, \qquad B(2,2) = \frac{n}{2(\lambda^{(k)})^2},$$

$$B(1,2) = B(2,1) = -\frac{n}{(\mu^{(k)})^3}\left(\bar{t}^{(k)} - \mu^{(k)}\right),$$

$$S(1) = -\frac{n\bar{z}^{(k)}}{\mu^{(k)}} + \frac{n\lambda^{(k)}}{(\mu^{(k)})^3}\left(\bar{t}^{(k)} - \mu^{(k)}\right)$$

and $$S(2) = \frac{n}{2\lambda^{(k)}} + \frac{n}{\mu^{(k)}} - \frac{n}{2(\mu^{(k)})^2}\bar{t}^{(k)} - \frac{n}{2}s_1^{(k)}.$$

4.2 Expected Fisher Information Matrix

We now provide the expected Fisher information matrix for the progressively censored data. We need the following result for further development. The PDF of $T_{(j)}$, for $j = 1, \cdots, m$, (see Balakrishnan, 2007), is

$$f_{T_{(j)}}(t) = c_{j-1}\sum_{i=1}^{j} a_{i,j}\Phi(-g(t))^{r_i-1}f(t;\alpha,\beta), \tag{26}$$

for $t > 0$, and 0 otherwise. Here

$$r_j = m - j + 1 + \sum_{i=j}^{m} R_i, \quad c_{j-1} = \prod_{i=1}^{j} r_i; \quad j = 1, \cdots, m, \tag{27}$$

and

$$a_{1,1} = 1, \quad a_{i,j} = \prod_{k=1,k\neq i}^{j} \frac{1}{r_k - r_i};$$

see for example Balakrishnan (2007). Based on (26), the expected Fisher information matrix, $\boldsymbol{F}$, can be obtained. Let us denote the 2×2 matrix $\boldsymbol{F}$ as follows;

$$\boldsymbol{F} = \begin{bmatrix} F_{11} & F_{12} \\ & \\ F_{21} & F_{22} \end{bmatrix} = -E \begin{bmatrix} \frac{\partial^2 \ln l(\alpha,\beta)}{\partial \alpha^2} & \frac{\partial^2 \ln l(\alpha,\beta)}{\partial \alpha \ \partial \beta} \\ & \\ \frac{\partial^2 \ln l(\alpha,\beta)}{\partial \alpha^2} & \frac{\partial^2 \ln l(\alpha,\beta)}{\partial \alpha^2} \end{bmatrix}. \tag{28}$$

Then

$$F_{11} = -\frac{m}{\alpha^2} + \frac{3}{\alpha^2}\sum_{j=1}^{m} h_{1j} - \frac{1}{\alpha^2}\sum_{j=1}^{m} R_j h_{2j} + \frac{2}{\alpha^2}\sum_{j=1}^{m} R_j h_{3j} + \frac{1}{\alpha^2}\sum_{j=1}^{m} R_j h_{4j}, \tag{29}$$

where

$$\begin{aligned}
h_{1j} &= E\left[\frac{1}{\alpha^2}\left(\frac{T_{(j)}}{\beta} + \frac{\beta}{T_{(j)}} - 2\right)\right] \\
h_{2j} &= E\left[\frac{g^3(T_{(j)})}{\Phi(-g(T_{(j)}))}\phi(g(T_{(j)}))\right] \\
h_{3j} &= E\left[\frac{g(T_{(j)})\phi(g(T_{(j)}))}{\Phi(-g(T_{(j)}))}\right] \\
h_{4j} &= E\left[\frac{g^2(T_{(j)})\phi^2(g(T_{(j)}))}{\left(\Phi(-g(T_{(j)}))\right)^2}\right], \quad \text{and} \\
g(t) &= \frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right].
\end{aligned}$$

Similarly,

$$F_{22} = -\frac{m}{\beta^2} - \sum_{j=1}^{m} h_{5j} + \sum_{j=1}^{m} h_{6j} - \sum_{j=1}^{m} R_j \ h_{7j}, \tag{30}$$

where

$$\begin{aligned}
h_{5j} &= E\left[\frac{h(T_{(j)})h_{b2}(T_{(j)}) - h_{b1}^2(T_{(j)})}{h^2(T_{(j)})}\right] \\
h_{6j} &= E\left[g_{b1}^2(T_{(j)}) + g(T_{(j)})g_{b2}(T_{(j)})\right] \\
h_{7j} &= E\left[\frac{\Phi(-g(T_{(j)}))P_{b2}(T_{(j)}) - P_{b1}^2(T_{(j)})}{\Phi^2(-g(T_{(j)}))}\right]
\end{aligned}$$

and

$$\begin{aligned}
h(t) &= \left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}} \\
h_{b1}(t) &= \frac{\partial}{\partial\beta}h(t) = \frac{1}{2\sqrt{\beta t}} + \frac{3}{2}\frac{\sqrt{\beta}}{t^{3/2}} \\
h_{b2}(t) &= \frac{\partial^2}{\partial\beta^2}h(t) = \frac{1}{4\beta}\left[-\frac{1}{\sqrt{\beta t}} + 3\frac{\sqrt{\beta}}{t^{3/2}}\right] \\
g_{b1}(t) &= \frac{\partial}{\partial\beta}g(t) = -\frac{1}{2\alpha\beta}\left[\left(\frac{t}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right] \\
g_{b2}(t) &= \frac{\partial}{\partial\beta}g_{b1}(t) = -\frac{1}{4\alpha\beta^2}\left[3\left(\frac{t}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right] \\
p_{b1}(t) &= -\phi(g(t))g_{b1}(t) \\
p_{b2}(t) &= \phi(g(t))g(t)g_{b1}^2(t) - \phi(g(t))g_{b2}(t).
\end{aligned}$$

Lastly, we have

$$F_{12} = F_{21} = -\frac{2}{\alpha}\sum_{j=1}^{m} h_{8j} + \frac{1}{\alpha}\sum_{j=1}^{m} R_j h_{9j}, \tag{31}$$

where

$$\begin{aligned}
h_{8j} &= E\left(g(T_{(j)})g_{b1}(T_{(j)})\right) \\
h_{9j} &= E\left[\frac{\phi(g(T_{(j)}))g_{b1}(T_{(j)})}{\Phi(-g(T_{(j)}))}\left\{g^2(T_{(j)}) - 1 - \frac{\phi(g(T_{(j)}))g(T_{(j)})}{\Phi(-g(T_{(j)}))}\right\}\right].
\end{aligned}$$

It is clear that the above elements need to be evaluated numerically.

5 Data Analysis

In this section, for illustrative purposes, we analyze three data sets; two real and one simulated. We want to show how the proposed algorithm works in practice. We employ the initial suggested in Section 3. Starting with the above initial estimates we compute the MLEs of α and β by using the EM algorithm as proposed in Section 3. In each case, the EM algorithm stops when $|\alpha_i - \alpha_{i-1}| + |\beta_i - \beta_{i-1}| \le 10^{-9}$. Here α_i and β_i denote the estimates of α and

β, respectively, at the i-th step. We further compute the asymptotic confidence intervals of the unknown parameters at the last step using the method of Louis (1982).

DATA SET 1: We take a very large data set with $n = 101$. The data set is given by Birnbaum and Saunders (1969b) on the fatigue life of 6061-T6 minimum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. The data set consists of the observations with maximum stress per cycle 31000 psi. The data are presented in Table 1.

Table 1: Fatigue lifetime data presented by Birnbaum Saunders (1969).

70	90	96	97	99	100	103	104	104	105	107	108	108	108	109
109	112	112	113	114	114	114	116	119	120	120	120	121	121	123
124	124	124	124	124	125	125	129	129	130	130	130	131	131	131
131	131	132	132	132	133	134	134	134	134	134	136	136	137	138
138	138	139	139	141	141	142	142	142	142	142	142	144	144	145
146	148	148	149	151	151	152	155	156	157	157	157	157	158	159
162	163	163	164	166	166	168	170	174	196	212				

Birnbaum and Saunders (1969b) and also Ng *et al.* (2006) analyzed this data set using Birnbaum-Saunders distribution. Based on the complete sample, the MLEs of α and β are 0.1704 and 131.8188 respectively. The corresponding standard errors, are respectively, 0.0120 and 2.2267. The 95% asymptotic confidence intervals of α and β are (0.1497,0.1976) and (127.5944, 136.3325) respectively. It is observed in both these cases that Birnbaum-Saunders distribution provides a good fit to the above data set.

We have generated different progressively censored samples from the complete data set for the different censoring schemes given below.

BCS-1: $m = 40$, $R_1 = 61$, $R_2 = \cdots = R_{40} = 0$.

BCS-2: $m = 40$, $R_1 = \cdots = R_{39} = 0$, $R_{40} = 61$.

BCS-3: $m = 50$, $R_1 = 51$, $R_2 = \cdots = R_{50} = 0$.

BCS-4: $m = 50$, $R_1 = \cdots = R_{49} = 0$, $R_{50} = 51$.

BCS-5: $m = 60$, $R_1 = 20$, $R_2 = 21$, $R_3 = \cdots = R_{60} = 0$.

BCS-6: $m = 60$, $R_1 = \cdots = R_{58} = 0$, $R_{59} = 20$, $R_{60} = 21$.

BCS-7: $m = 60$, $R_1 = 10$, $R_2 = 10$, $R_3 = 10$, $R_4 = 11$, $R_5 = \cdots = R_{60} = 0$.

BCS-8: $m = 70$, $R_1 = 15$, $R_2 = \cdots = R_{69} = 0$, $R_{70} = 16$.

Note that BCS-2 and BCS-4 are the usual Type-II censoring schemes, and BCS-1 and BCS-3 are just the opposite of Type-II censoring schemes, and they are known as the Type-III censoring scheme, see Pradhan and Kundu (2009). It may be noted that for fixed n and m, among different progressive censoring schemes, the expected duration of the experiment is minimum for Type-II censoring scheme and maximum for Type-III censoring scheme. The expected duration of the experiments for other censoring schemes will be in between these two extremes.

The MLEs of α and β, along with standard errors, and 95% confidence intervals are presented in Table 2. The iteration numbers required for the convergence of EM algorithm in the 8 censoring schemes are 67, 78, 49, 59, 29, 43, 29 and 29, respectively. As expected, the number of iteration decreases with an increase effective sample size. In all cases, the EM algorithm performs well.

Example 2: This is a very small data set with $n = 10$ observations. This data set (McCool, 1974) is on fatigue life in hours of bearings of a certain type. The data set is presented in Table 3.

We generate progressively censored data from the complete data set for the following censoring schemes.

MCS-1: $m = 6$, $R_1 = 4$, $R_2 = \cdots = R_6 = 0$.

MCS-2: $m = 6$, $R_1 = \cdots = R_5 = 0$, $R_6 = 4$.

Table 2: Estimates of α and β along with standard error and 95% confidence intervals for different censoring schemes.

Censoring	α			β		
schemes	Estimate	s.e.	95% CI	Estimate	s.e.	95% CI
BCS-1	0.1765	0.0124	[0.1521, 0.2008]	125.4913	2.1872	[121.2044, 129.7782]
BCS-2	0.1925	0.0135	[0.1660, 0.2191]	135.1504	2.5654	[130.1221, 140.1787]
BCS-3	0.1758	0.0124	[0.1516, 0.2001]	126.5748	2.1983	[122.2661, 130.8834]
BCS-4	0.1755	0.0124	[0.1513, 0.1997]	132.3971	2.2947	[127.8994, 136.8946]
BCS-5	0.1769	0.0125	[0.1525, 0.2013]	129.1470	2.2560	[124.7253, 133.5610]
BCS-6	0.1753	0.0123	[0.1512, 0.1995]	132.3745	2.2920	[127.8821, 136.8669]
BCS-7	0.1756	0.0123	[0.1514, 0.1998]	129.6502	2.2478	[125.2445, 134.0559]
BCS-8	0.1746	0.0123	[0.1505, 0.1987]	130.5407	2.2512	[126.1283, 134.9530]

Table 3: Fatigue lifetime data presented by McCool (1974).

152.7	172.0	172.5	173.3	193.0
204.7	216.5	234.9	262.6	422.6

MCS-3: $m = 7$, $R_1 = 2$, $R_2 = 1$, $R_3 \cdots = R_7 = 0$.

The estimate α and β, along with standard errors, and 95% confidence intervals are given in Table 4. It is observed that even when the sample size is very small, the proposed algorithm is working quite well. The iteration numbers needed for the convergence of EM algorithm for three schemes are 30, 44 and 24 respectively.

Table 4: Estimates of α and β along with standard error and 95% confidence intervals for the data set of McCool (1974)

Censoring	α			β		
schemes	Estimate	s.e.	95% CI	Estimate	s.e.	95% CI
MCS-1	0.1639	0.0367	[0.0921, 0.2358]	194.0795	9.9935	[174.4922, 213.6669]
MCS-2	0.1484	0.0332	[0.0833, 0.2134]	195.4253	9.1186	[177.5528, 213.2978]
MCS-3	0.1570	0.0351	[0.0882, 0.2258]	195.8228	9.6617	[176.8859, 214.7597]

Data Set 3: Finally we analyze one simulated data set. The progressively censored data set has been generated from a Birnbaum-Saunders distribution with parameters $\alpha = 1$, $\beta = 1$, the sample size $n = 30$, the effective sample size $m = 15$ and $R = (R_1 = 5, R_2 = 5, R_3 = 5, R_4 = \cdots = R_{15} = 0)$. The observations are presented below:

0.08528956	0.18289674	0.23667950	0.36458839	0.39295772
0.41571990	0.53403278	0.79707860	0.85808242	1.10885831
1.27737016	1.92730265	2.35140494	4.30906003	7.71967056

The estimate of α and β, with standard errors in parentheses, are 1.1705 (0.1554) and 0.8561 (0.1449), respectively. The EM algorithm stops after 25 iterations. The 95% confidence interval for α and β are [0.8660, 1.4750] and [0.5721, 1.1402], respectively. It is clear that the EM algorithm works very well in this case also.

6 Optimal Progressive Censoring Scheme

Up to this point we have assumed that n, m and the progressive censoring scheme $\{R_1, \cdots, R_m\}$ such that $R_1 + \cdots + R_m = n - m$, are fixed in advance and we have developed the inference procedure of the unknown parameters. In practice, it is very important to choose a 'optimal' censoring scheme so that it provides the maximum information of the unknown parameters. It is very clear that unless we fixed n and m, the problem may not make much sense. Intuitively it is very apparent that if we choose $n = m$ and make n larger it should provide more information of the unknown parameters. Moreover, in practice most of the time, the sample size n and the effective sample size m are fixed in advance. Therefore, the natural question is whether we should choose the progressive censoring scheme $\{R_1, \cdots, R_m\}$ based on convenience or based on some scientific basis.

In the last few years finding the 'optimal' censoring scheme has received considerable attention in the statistical literature; see for example the Chapter 10 of Balakrishnan and

Aggarwala (2000), Ng *et al.* (2004), Balasooriya and Balakrishnan (2000), Balasooriya *et al.* (2000), Burkchat *et al.* (2007, 2008), Burkchat (2008), Pradhan and Kundu (2009) and see the references cited therein. Here, possible censoring schemes means, for fixed n and m, all possible choices of $R_1, \cdots, R_m$, such that

$$\sum_{i=1}^{m} R_i + m = n,$$

and choosing the 'optimal' sampling scheme means, among all possible progressive censoring schemes find that particular progressive censoring scheme, which provides the maximum 'information' of the unknown parameters. Naturally, the first question that arises is how to define information measures of the unknown parameters based on particular progressive censoring data, and the second question is how to compare two different information measures based on two different progressive censoring schemes. Therefore, two important issues are involved in dealing with finding the optimal censoring scheme, namely (a) find a proper criterion and (ii) with respect to that criterion, find the best censoring scheme. Both points are quite important and neither is an trivial issue in this case.

In case of a single parameter, the problem is quite straightforward. In this case, the expected Fisher information measure can be used as an information measure. But when more than one parameter is present, it is not an easy problem anymore. Some of the existing choices are to consider the trace or determinant of the expected Fisher information matrix. Unfortunately, in presence of the shape parameter, it can be shown that, see Gupta and Kundu (2006), the trace or determinant are not scale invariant. In fact, it may happen that for a particular scheme, its determinant or trace of the expected Fisher information matrix is more than the other scheme, but if we change the unit (multiply the data by a positive constant), the inequality becomes reversed. This is not very desirable.

To avoid this problem, based on Zhang and Meeker (2005) and Kundu (2008), we use the following criteria for information measures, see also Pradhan and Kundu (2009). Consider the p-th quantile of the Birnbaum-Saunders distribution:

$$T_p = \frac{\beta}{4}\left[\alpha\Phi^{1}(p) + \sqrt{\{\alpha\Phi^{-1}(p)\}^2 + 4}\right]^2.$$

Using the idea of Gupta and Kundu (2006), the following information measure for a given censoring scheme $\{R_1, \cdots, R_m\}$ has been used

$$I_W(\{R_1, \cdots, R_m\}) = \int_0^1 V(\{R_1, \cdots, R_m\})_p w(p) dp,$$

where $V(\{R_1, \cdots, R_m\})_p$ denotes the asymptotic variance of $\widehat{T}_p$, the MLE of T_p, based on the censoring scheme $\{R_1, \cdots, R_m\}$. Moreover, $w(\cdot)$ is a non-negative weight function such that

$$\int_0^1 w(p) dp = 1.$$

In this case, $V(\{R_1, \cdots, R_m\})_p$ can be expressed as

$$T_p^2 \left[\frac{2\Phi^{-1}(p)T_p}{\sqrt{(\alpha\Phi^{-1}(p))^2 + 4}}, \frac{1}{\beta}\right] \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}^{-1} \begin{bmatrix} \frac{2\Phi^{-1}(p)T_p}{\sqrt{(\alpha\Phi^{-1}(p))^2+4}} \\ \frac{1}{\beta} \end{bmatrix}.$$

As it has been mentioned in Pradhan and Kundu (2009), it may be observed that the criterion $I_W(\{R_1, \cdots, R_m\})$ is very flexible. By the proper choice of the weight function $w(\cdot)$, the criteria of Zhang and Meeker (2005) and Ng *et al.* (2004) can be obtained as special cases.

For illustrative purposes, we have provided a small table indicating the optimal censoring scheme with respect to different criteria. We have used the following six criteria:

Criterion 1: Minimum trace.

Criterion 2: Minimum variance.

Criterion 3: Minimum variance of the p-th percentile estimator, $p = 0.5$

Criterion 4: Minimum variance of the p-th percentile estimator, $p = 0.9$

Criterion 5: Minimum variance of the p-th percentile estimator, $p = 0.999$

Criterion 6: Minimum variance of $I_W(\cdot)$, with $w(p) = 1$, for $0 < p < 1$.

The results are presented in Table 5. Note that in each case, the minimization has to be performed numerically. Interestingly, it is observed that Criterion 1 and Criterion 6 provide very similar results, as it was observed in Pradhan and Kundu (2009) in case of the generalized exponential distribution. Moreover, for Criterion 5, the usual Type-II censoring are the optimal schemes in all the cases considered.

Table 5: The optimal censoring scheme for different criteria when $\alpha = 2$, $\beta = 1$ and $n = 10$, 15, 20, 25 and $m = 5$. Against each criterion, the first row, second row third row and fourth row represent the optimal censoring scheme with respect to that criterion for $n = 10$, 15, 20 and 25 respectively.

Criterion	R_1	R_2	R_3	R_4	R_5
1	0	0	0	0	5
	0	0	0	0	10
	0	0	0	1	14
	0	0	0	1	19
2	5	0	0	0	0
	10	0	0	0	0
	12	3	0	0	0
	16	4	0	0	0
3	5	0	0	0	0
	12	3	0	0	0
	7	8	0	0	0
	0	0	12	8	0
4	0	0	0	0	5
	0	0	0	0	10
	0	0	0	2	13
	0	0	0	2	18
5	0	0	0	0	5
	0	0	0	0	10
	0	0	0	0	15
	0	0	0	0	20
6	0	0	0	1	4
	0	0	0	2	8
	0	0	1	2	12
	0	0	1	2	17

It may be mentioned that although the total number of censoring schemes is finite, but it can be quite large. For fixed n and m a total $\binom{n-1}{m-1}$ possible censoring schemes are available. For example when $n = 25$ and $m = 12$, the possible number of censoring schemes becomes $\binom{24}{11} =$ 2,496,144, which is quite large. We need to search all these cases to find the optimal censoring scheme. We do not have any efficient method to find the optimal scheme without searching all the possible schemes. We have provided the sub-optimal scheme for each criterion as suggested by Pradhan and Kundu (2009) which involves searching only among m schemes of the form $(n-m, 0, \cdots, 0)$, $\cdots, (0, \cdots, 0, n-m)$. The results are presented in Table 6, and we also present the relative efficiency (RE) of the sub-optimal scheme. It is clear that when the relative efficiency is 1, then the sub-optimal scheme is the optimal scheme also. In all the cases considered, the relative efficiencies of the sub-optimal schemes are quite high.

7 Conclusions

In this paper we have considered the estimation of the unknown parameters of the Birnbaum-Saunders distribution when the data are progressively Type-II censored. The MLEs of the unknown parameters cannot be obtained in explicit forms. We have proposed to use the EM algorithm to compute the unknown parameters, and also numerically estimate the asymptotic variance-covariance matrix.

We have provided the optimal censoring scheme based on different information measures. Since the computation of the optimal censoring scheme is quite computer intensive and the number of possible schemes grows fast, we have examined the sub-optimal censoring schemes also. It is observed in all the cases considered that the relative efficiencies of the sub-optimal censoring schemes are quite high. Therefore, sub-optimal censoring schemes can be used quite effectively in practice.

Acknowledgements:

The authors would like to thank the associate editor and two reviewers for their valuable comments which had improved the earlier version of the manuscript.

References

[1] Ahmed, S.E., Budsaba, K, Lisawadi, S. and Volodin, A. I. (2008), "Parametric estimation for the Birnbaum-Saunders lifetime distribution based on a new parameterization", *Thailand Statistician*, 6, 213-240.

[2] Balakrishnan, N. (2007), "Progressive Censoring Methodology: An Appraisal", (with discussion)", *TEST*, vol. 16, no. 2, 211 - 296.

[3] Balakrishnan, N. and Aggarwala, R. (2000), *Progressive censoring, theory, methods and applications*, Birkhauser, Boston.

[4] Balakrishnan, N., Burkschat, M., Cramer, E. and Hoffman, G. (2008), "Fisher information based progressive censoring plans", *Computational Statistics and Data Analysis*, vol. 53, 366 - 380.

[5] Balakrishnan, N., Gupta, R.C., Kundu, D., Leiva, V. and Sanhueza, A. (2011), "On some mixture models based on the Birnbaum-Saunders distribution and associated inference", *Journal of Statistical Planning and Inference*, vol. 141, 2175 - 2190.

[6] Balasooriya, U. and Balakrishnan, N. (2000), "Reliability sampling plans for log-normal distribution, based on progressively-censored samples", *IEEE Transactions on Reliability*, vol. 49, 199 - 203.

[7] Balasooriya, U., Saw, S.L.C. and Gadag, V. (2000), "Progressively censored reliability sampling plans for the Weibull distribution", *Technometrics*, vol. 42, 160 - 167.

[8] Birnbaum, Z. W. and Saunders, S. C., (1969a). "A New Family of Life Distribution", *Journal of Applied Probability*, vol. 6, 319-327.

[9] Birnbaum, Z. W. and Saunders, S. C., (1969b). "Estimation for a Family of Life Distributions With Applications to Fatigue", *Journal of Applied Probability*, vol. 6, 328-347.

[10] Burkschat, M. (2008), "On optimality of extremal schemes in progressive type II censoring", *Journal of Statistical Planning and Inference*, vol. 138, 1647 - 1659.

[11] Burkschat, M., Cramer, E. and Kamps, U. (2006), "On optimal schemes in progressive censoring", *Statistics and Probability Letters*, vol. 76, 1032 - 1036.

[12] Burkschat, M., Cramer, E. and Kamps, U. (2007), "Optimality criteria and optimal schemes in progressive censoring", *Communications in Statistics - Theory and Methods*, vol. 36, 1419 - 1431.

[13] Cohen, A. C., (1963). "Progressively censored samples in life testing" *Technometrics*, vol. 5, 327 - 329

[14] Desmond, A. F., (1985). "Stochastic Models of Failure in Random Environments", *Canadian Journal of Statistics,* vol. 13, 171-183.

[15] Dupuis, D. J. and Mills, J. E., (1998). "Robust Estimation of the Birnbaum–Saunders Distribution", *IEEE Transactions on Reliability*, vol. 47, 88-95.

[16] From, S.G., Li, L.X. (2006), "Estimation of the parameters of the Birnbaum-Saunders distribution", *Communications in Statistics Theory and Methods*, vol. 35, 2157-2169.

[17] Gupta, R.C. and Akman, O. (1995), "On the reliability studies of a weighted inverse Gaussian model", *Journal of Statistical Planning and Inference*, vol. 48, 69 - 83.

[18] Gupta, R.D. and Kundu, D. (2006), "Comparison of Fisher information matrices of the Weibull and generalized exponential distributions", *Journal of Statistical Planning and Inference*, vol. 136, 3130 - 3144.

[19] Johnson, N. L., Kotz, S. and Balakrishnan, N., 1995. *Continuous Univariate Distributions – Vol. 2*, Second edition. John Wiley & Sons, New York.

[20] Jorgensen, B., Seshadri, V. and Whitmore, G. A., (1991), "On the mixture of the inverse Gaussian distribution with its complementary reciprocal", *Scandinavian Journal of Statistics*, vol. 18, 77-89.

[21] Kundu, D. (2008), "Bayesian inference and life testing plan for Weibull distribution in presence of progressive censoring", *Technometrics*, vol. 50, no. 2, 144 - 154.

[22] Louis, T. A. (1982), "Finding the observed information matrix using the EM algorithm", *Journal of the Royal Statistical Society*, Ser B, vol. 44, 226 - 233.

[23] McCool, J.I. (1974), "Inferential Techniques for Weibull Populations", Aerospace Research Laboratories Report ARL TR74-0180,Wright-Patterson Air Force Base, Dayton, Ohio.

[24] McLachlan, G. J. and Krishnan, T. (1992), *The EM Algorithm and its Extensions*, John Wiley & Sons, New York.

[25] Ng, H. K. T., Chan, C. S. and Balakrishnan, N. (2002), " Estimation of parameters from progressively censored data using EM algorithm", *Computational Statistics and Data Analysis* vol. 39, 371 - 386

[26] Ng, H. K. T., Chan, C.S. and Balakrishnan, N. (2004), "Optimal progressive censoring plans for the Weibull distribution", *Technometrics*, vol. 46, 470 - 481.

[27] Ng, H. K. T., Kundu, D. and Balakrishnan, N. (2006), "Point and interval estimation for the two-parameter Birnbaum-Saunders distribution based on Type-II censored samples", *Computational Statistics & Data Analysis*, vol 50, 3222-3242.

[28] Owen, W.J. (2006). "A new three-parameter extension to the Birnbaum-Saunders distribution" *IEEE Transactions on Reliability*, vol. 55, 475-479.

[29] Pradhan, B. and Kundu, D. (2009), "On progressively censored generalized exponential distribution", *Test*, vol. 18, 497 - 515.

[30] Rieck, J. R., (1995). "Parametric Estimation for the Birnbaum–Saunders Distribution Based on Symmetrically Censored Samples", *Communications in Statistics – Theory and Methods*, vol. 24, 1721-1736.

[31] Rieck, J. R., (1999). "A Moment-generating Function with Application to the Birnbaum–Saunders Distribution", *Communications in Statistics – Theory and Methods*, vol. 28, 2213-2222.

[32] Xie, F.C. and Wei, B.C. (2007), "Diagnostics analysis for log-Birnbaum-Saunders regression models", *Computational Statistics & Data Analysis*, vol. 51, 4692-4706.

[33] Zhang, Y. and Meeker, W.Q. (2005), "Bayesian life test planning for the Weibull distribution with given shape parameter", *Metrika*, vol. 61, 237 - 249.

Table 6: The sub-optimal censoring scheme for different criteria when $\alpha = 2$, $\beta = 1$ and $n = 10, 15, 20, 25$ and $m = 5$. Against each criterion, the first row, second row third row and fourth row represent the optimal censoring scheme with respect to that criterion for $n = 10, 15, 20$ and 25 respectively. We have also provided the relative efficiency (RE) of the sub-optimal censoring scheme.

Criterion	R_1	R_2	R_3	R_4	R_5	RE
1	0	0	0	0	5	1.0
	0	0	0	0	10	1.0
	0	0	0	0	15	0.97
	0	0	0	0	20	0.98
2	5	0	0	0	0	1.0
	10	0	0	0	0	1.0
	15	0	0	0	0	0.93
	20	0	0	0	0	0.92
3	5	0	0	0	0	1.0
	15	0	0	0	0	0.90
	0	15	0	0	0	0.88
	0	0	20	0	0	0.86
4	0	0	0	0	5	1.0
	0	0	0	0	10	1.0
	0	0	0	0	15	0.90
	0	0	0	0	20	0.91
5	0	0	0	0	5	1.0
	0	0	0	0	10	1.0
	0	0	0	0	15	1.0
	0	0	0	0	20	1.0
6	0	0	0	0	9	0.99
	0	0	0	0	10	0.97
	0	0	0	0	15	0.94
	0	0	0	0	20	0.94