On the Joint Type-II Progressive Censoring Scheme

Shuvashree Mondal*, Debasis Kundu †

Abstract

Recently the progressive censoring scheme has been extended for two or more populations. In this article we consider the joint Type-II progressive censoring (JPC) scheme for two populations when the lifetime distributions of the experimental units of the two populations follow two-parameter generalized exponential distributions with the same scale parameter but different shape parameters. The maximum likelihood estimators of the unknown parameters cannot be obtained in explicit forms. We propose to use the expectation maximization (EM) algorithm to compute the maximum likelihood estimators. The observed information matrix based on missing value principles is derived. We study the Bayesian inference of the unknown parameters based on a beta-gamma prior for the shape parameters, and an independent gamma prior for the common scale parameter. The Bayes estimators with respect to the squared error loss function cannot be obtained in explicit form. We propose to use the importance sampling technique to compute the Bayes estimates and the associated credible intervals of the unknown parameters. Extensive simulation experiments have been performed to study the performances of the different methods. Finally a real data set has been analyzed for illustrative purposes.

Key Words and Phrases: Progressive censoring scheme; Joint progressive censoring scheme; Generalized Exponential distribution; EM Algorithm; Fisher Information; Bayes estimator; Importance sampling; Credible interval.

AMS Subject Classifications: 62N01, 62N02, 62F10.

*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.
†Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. Corresponding author. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

1 Introduction

The progressive censoring scheme has received a considerable amount of attention during the last ten to fifteen years. Briefly we describe the progressive Type-II censoring scheme as follows. Suppose n units are put on a life testing experiment and $k < n$ be the number of failures to be observed in the experiment. Let $R_1, \ldots, R_k$ be non-negative integers satisfying $\sum_{i}^{k}(R_i+1) = n$. In the progressive Type-II censoring scheme, at the time of the first failure, we withdraw R_1 units randomly from the remaining $n-1$ surviving units. Next at the time of the second failure, R_2 units are withdrawn randomly from the remaining $n-2-R_1$ surviving units. The test is continued until the k-th failure takes place and at the k-th failure the test stops with removal of the remaining R_k surviving units. An exhaustive collection of work on progressive censoring schemes can be found in Balakrishnan and Cramer (2014).

Recently Rasouli and Balakrishnan (2010) introduced the joint progressive censoring (JPC) scheme for a comparative study of two populations. It can be briefly described as follows. Suppose m units are drawn from Population-A (Pop-A) and n units are drawn from Population-B (Pop-B). Under the JPC scheme two samples are combined and put on a life testing experiment. Let k be the total number of failures to be observed in the experiment and $R_1, \ldots, R_k$ be the non-negative integers satisfying $\sum_{i=1}^{k}(R_i + 1) = m + n$. From the combined sample, at the time of the first failure W_1, R_1 units are randomly withdrawn from the remaining $m + n - 1$ surviving units. These R_1 units consist of S_1 units from the sample of Pop-A and T_1 units from the sample of Pop-B, where S_1, T_1 are random and $S_1 + T_1 = R_1$. Similarly, at the second failure time point W_2, R_2 units are withdrawn randomly from the remaining $m+n-2-R_1$ surviving units which consist of S_2 units from the sample of Pop-A and T_2 units from the sample of Pop-B where $S_2 + T_2 = R_2$. The test is terminated at the k-th failure time point W_k with removal of all the reaming surviving units from both the samples. Here we introduce another set of random variables $Z_1, \ldots, Z_k$, where $Z_i = 1$ or 0 if i-th failure comes from the sample of Pop-A or Pop-B, respectively. Therefore, the censored sample is of the form $((W_1, Z_1, S_1), \ldots, (W_k, Z_k, S_k))$. Let us denote $K_1 = \sum_{i=1}^{k} Z_i$ and $K_2 = \sum_{i=1}^{k}(1 - Z_i) = k - K_1$, the total number of failures from Pop-A and Pop-B, respectively, out of total k observed failures.

Rasouli and Balakrishnan (2010) provided the likelihood and Bayesian inference of two exponential distributions for the JPC scheme. Balakrishnan, Su and Liu (2015) extended the JPC scheme for more than two exponential populations and provided the likelihood and Bayesian inference. Parsi, Ganjali and Sanjari (2011) studied conditional maximum likelihood and interval estimation of the unknown parameters of two Weibull populations under the JPC scheme. Mondal and Kundu (2017) addressed the problem of point and interval estimation of the unknown parameters of two Weibull populations under the Bayesian frame-work. Doostparast, Ahmadi and Ahmadi (2013) obtained the Bayesian estimates of the parameters for a general class of distributions with respect to the squared error and the LINEX loss functions under the JPC scheme. Parsi and Bairamov (2009) determined the expected number of failures in life testing experiment under the JPC scheme for different parametric families of distributions.

In this paper we analyze the joint progressively censored data when the lifetime distributions of the experimental units of the two populations follow two-parameter generalized exponential distributions with the same scale parameter but different shape parameters. We study the likelihood as well as the Bayesian inference of the unknown model parameters. It is observed that the maximum likelihood estimators (MLEs) of the unknown parameters can be obtained by solving a three dimensional optimization problem. The standard Newton-Raphson method may be used to solve this problem. In this case one needs to compute the Hessian matrix, which may not be in a very convenient form. Moreover, it is also observed that for small effective sample size, the standard Newton-Raphson method may not converge. Nelder-Mead simplex algorithm can provide the MLEs. But in this method if the initial guesses are not so close to the MLEs, it may enter a region of local optimum. Due to this reason, we propose to use the Expectation Maximization (EM) algorithm based on missing value principle to compute the MLEs of the unknown parameters. In the proposed EM algorithm, at each 'E'-step the corresponding 'M'-step can be performed by solving a single one dimensional optimization problem. It is verified through extensive grid search that EM estimates indeed maximize the corresponding likelihood function. Also in these context EM algorithm is very robust to the initial guesses. We further compute the observed information matrix based on missing value principle and it can be used to compute the approximate confidence intervals of the unknown parameters.

To compute the Bayes estimates we have assumed a very flexible beta-gamma prior on the shape parameters, and an independent gamma prior on the common scale parameter. The beta-gamma distribution is a very flexible bivariate absolute continuous distribution, and it incorporates different dependency structure between the marginals, depending on the hyper-parameters. It can produce both positive and negative dependence among the marginals. It is observed that the Bayes estimates of the unknown parameters based on the squared error loss function cannot be obtained in closed form. We propose to use importance sampling technique to compute the Bayes estimates and the associated credible intervals. A rigorous Monte Carlo simulation experiment is performed to check the performance of the different methods. One real data set has been analyzed for illustrative purpose.

The rest of the paper is arranged as follows. In section 2 we provide the notations and derive the likelihood function. We present the EM algorithm in section 3. In section 4 we compute the observed information matrix based on missing value principle. The Bayesian analysis is performed in section 5. The simulation results and the analysis of a real data set are presented in section 6. Finally, in section 7 we conclude the paper.

2 Notation and Likelihood function

2.1 Notations

CRI : Credible Interval

MLE : Maximum likelihood estimator

EM : Expectation Maximization

HPD : Highest Posterior Density

i.i.d. : Independent and identically distributed.

PDF : Probability density function.

Beta(a, b) : Beta distribution with PDF:
$f_{Beta}(x, a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \ 0 < x < 1, a, b > 0.$

GA(a, b) : Gamma distribution with PDF:
$f_{GA}(x, a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, \ x > 0, a, b > 0.$

GE(α, λ) : Generalized exponential distribution with PDF:
$f_{GE}(x, \alpha, \lambda) = \alpha\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha-1}, \ x > 0, \alpha, \lambda > 0$
and the distribution function, $F_{GE}(x, \alpha, \lambda) = (1 - e^{-\lambda x})^{\alpha}$.

2.2 Likelihood Function

Suppose m items from Pop-A say $X_1, \ldots, X_m$ are i.i.d. $GE(\alpha_1, \lambda)$ and n items from Pop-B say $Y_1, \ldots, Y_n$ are i.i.d. $GE(\alpha_2, \lambda)$. For a given JPC scheme $(R_1, \ldots, R_k)$, the observed data are of the form $\{(w_1, z_1, s_1), \ldots, (w_k, z_k, s_k)\}$. Therefore, the likelihood function without the normalizing constant can be written as

$$\begin{aligned} L(\alpha_1, \alpha_2, \lambda | data) &= \alpha_1^{k_1} \alpha_2^{k_2} \lambda^k e^{-\lambda \sum\limits_{i=1}^{k} w_i} \times \prod_{i=1}^{k} ((1 - e^{-\lambda w_i})^{\alpha_1 - 1})^{z_i} ((1 - e^{-\lambda w_i})^{\alpha_2 - 1})^{1 - z_i} \\ &\times \prod_{i=1}^{k} (1 - (1 - e^{-\lambda w_i})^{\alpha_1})^{s_i} (1 - (1 - e^{-\lambda w_i})^{\alpha_2})^{t_i}, \end{aligned} \tag{1}$$

where $k_1 = \sum\limits_{i=1}^{k} z_i$, $k_2 = \sum\limits_{i=1}^{k}(1 - z_i) = k - k_1$ and $t_i = R_i - s_i$ for $i = 1, \ldots k$. When $k_1 = 0$, we have $k_2 = k$, $z_i = 0$ for all $1 \leq i \leq k$. Thus, the likelihood function (1) becomes,

$$\begin{aligned} L(\alpha_1, \alpha_2, \lambda | data) &= \alpha_2^k \lambda^k e^{-\lambda \sum_{i=1}^k w_i} \prod_{i=1}^{k} (1 - e^{-\lambda \sum_{i=1}^k w_i})^{\alpha_2 - 1} \\ &\times \prod_{i=1}^{k} (1 - (1 - e^{-\lambda \sum_{i=1}^k w_i})^{\alpha_1})^{s_i} (1 - (1 - e^{-\lambda \sum_{i=1}^k w_i})^{\alpha_2})^{t_i}. \end{aligned}$$

For $s_i = 0$, $(1 - (1 - e^{-\lambda \sum_{i=1}^k w_i})^{\alpha_1})^{s_i} = 1$ and for $s_i \neq 0$, $(1 - (1 - e^{-\lambda \sum_{i=1}^k w_i})^{\alpha_1})^{s_i}$ is a strictly increasing function of α_1 for fixed λ, and it increases to 1. Therefore, when $k_1 = 0$, for fixed λ and α_2, $L(\alpha_1, \alpha_2, \lambda | data)$ is a strictly increasing function of α_1 on the parameter space $(0, \infty)$. Similarly when $k_2 = 0$, for fixed λ and α_1, $L(\alpha_1, \alpha_2, \lambda | data)$ is a strictly increasing function of α_2 on the parameter space $(0, \infty)$. Therefore, the MLEs do not exist when $k_1 = 0$ or $k_2 = 0$. For existence of the MLEs, it is assumed that $k_1, k_2 > 0$.

Note that, the MLEs can be obtained by maximizing the likelihood equation (1). It involves solving a three dimensional optimization problem. The standard Newton-Raphson algorithm may be used for that purpose. It has its own difficulties. To avoid that we propose to use EM algorithm to compute the MLEs, and it will be explained in the next section.

3 Expectation Maximization (EM) Algorithm

In this section we apply the EM algorithm by treating the problem as a missing value problem. The missing values in this case are the lifetimes of the censored units at each failure time point.

We denote U_{ij} as the lifetime of the j-th censored unit from Pop-A at the i-th failure time-point W_i, for $j = 1, \ldots, S_i$ and $V_{ij'}$ as the lifetime of j'-th censored unit from Pop-B at i-th failure time point W_i, for $j' = 1, \ldots, T_i = R_i - S_i$ and $i = 1, \ldots, k$. The observed data are $((w_1, z_1, s_1), \ldots, (w_k, z_k, s_k))$ and the missing data are $\mathscr{U} = \Big((u_{11}, \ldots, u_{1s_1}), \ldots, (u_{k1}, \ldots u_{ks_k})\Big)$ and $\mathscr{V} = ((v_{11}, \ldots, v_{1t_1}), \ldots, (v_{k1}, \ldots, v_{kt_k}))$. Combining the observed and missing data forms the complete data, $\Big((w_1, z_1, s_1), \ldots, (w_k, z_k, s_k), \mathscr{U}, \mathscr{V}\Big) = data^*$ (say). Based on the complete data the log-likelihood function without the normalizing constant can be obtained as

$$\begin{aligned}
\ln L_c(\alpha_1, \alpha_2, \lambda | data^*) &= m \ln \alpha_1 + n \ln \alpha_2 + (m+n) \ln \lambda \\
&\quad -\lambda \Big(\sum_{i=1}^{k} w_i + \sum_{i=1}^{k} \sum_{j=1}^{s_i} u_{ij} + \sum_{i=1}^{k} \sum_{j'=1}^{t_i} v_{ij'} \Big) \\
&\quad +(\alpha_1 - 1) \Big(\sum_{i=1}^{k} z_i \ln(1 - e^{-\lambda w_i}) + \sum_{i=1}^{k} \sum_{j=1}^{s_i} \ln(1 - e^{\lambda u_{ij}}) \Big) \\
&\quad +(\alpha_2 - 1) \Big(\sum_{i=1}^{k} (1 - z_i) \ln(1 - e^{-\lambda w_i}) + \sum_{i=1}^{k} \sum_{j'=1}^{t_i} \ln(1 - e^{\lambda v_{ij'}}) \Big).
\end{aligned}$$

At the 'E'-step of the EM algorithm the pseudo log-likelihood function can be obtained as

$$\begin{aligned}
l_s(\alpha_1, \alpha_2, \lambda) &= m \ln \alpha_1 + n \ln \alpha_2 + (m+n) \ln \lambda \\
&\quad -\lambda \Big(\sum_{i=1}^{k} w_i + \sum_{i=1}^{k} \sum_{j=1}^{s_i} E(U_{ij} | U_{ij} > w_i) + \sum_{i=1}^{k} \sum_{j'=1}^{t_i} E(V_{ij'} | V_{ij'} > w_i) \Big) \\
&\quad +(\alpha_1 - 1) \Big(\sum_{i=1}^{k} z_i \ln(1 - e^{-\lambda w_i}) + \sum_{i=1}^{k} \sum_{j=1}^{s_i} E(\ln(1 - e^{\lambda U_{ij}}) | U_{ij} > w_i) \Big) \\
&\quad +(\alpha_2 - 1) \Big(\sum_{i=1}^{k} (1 - z_i) \ln(1 - e^{-\lambda w_i}) + \sum_{i=1}^{k} \sum_{j'=1}^{t_i} E(\ln(1 - e^{\lambda V_{ij'}}) | V_{ij'} > w_i) \Big).
\end{aligned} \tag{2}$$

The following result is needed to compute the necessary conditional expectations as required in (2).

Result 1: Given $(W_1 = w_1, Z_1 = z_1, S_1 = s_1), \ldots, (W_i = w_i, Z_i = z_i, S_i = s_i)$ the conditional PDF of U_{ij} is given by

$$f_{U_{ij}|(W_1,Z_1,S_1),\ldots,(W_i,Z_i,S_i)}(u_{ij}|(w_1,z_1,s_1),\ldots,(w_i,z_i,s_i)) = f_{U_{ij}|W_i}(u_{ij}|w_i) = \frac{f_{GE}(u_{ij},\alpha_1,\lambda)}{1-F_{GE}(w_i,\alpha_1,\lambda)},$$

and the conditional PDF of $V_{ij'}$ is given by

$$f_{V_{ij'}|(W_1,Z_1,S_1),\ldots,(W_i,Z_i,S_i)}(v_{ij'}|(w_1,z_1,s_1),\ldots,(w_i,z_i,s_i)) = f_{V_{ij'}|W_i}(v_{ij'}|w_i) = \frac{f_{GE}(v_{ij'},\alpha_2,\lambda)}{1-F_{GE}(w_i,\alpha_2,\lambda)},$$

for $i = 1, \ldots, k$.

Proof: It can be obtained similar as in the progressive Type-II censoring case discussed in Chapter 9 of Balakrishnan and Cramer (2014).

In the 'M'-step, we need to maximize the pseudo log-likelihood function with respect to α_1, α_2 and λ. If at the r-th stage of the iteration the estimate of $(\alpha_1, \alpha_2, \lambda)$ is denoted by $(\alpha_1^{(r)}, \alpha_2^{(r)}, \lambda^{(r)})$, then $(\alpha_1^{(r+1)}, \alpha_2^{(r+1)}, \lambda^{(r+1)})$ can be obtained by maximizing $g(\alpha_1, \alpha_2, \lambda)$ with respect to $\alpha_1, \alpha_2, \lambda$, where

$$\begin{aligned} g(\alpha_1,\alpha_2,\lambda) &= m\ln\alpha_1 + n\ln\alpha_2 + (m+n)\ln\lambda \\ &\quad -\lambda\Big(\sum_{i=1}^k w_i + \sum_{i=1}^k s_i E_{(\alpha_1^{(r)},\lambda^{(r)})}(U|U>w_i) + \sum_{i=1}^k t_i E_{(\alpha_2^{(r)},\lambda^{(r)})}(V|V>w_i)\Big) \\ &\quad +(\alpha_1-1)\Big(\sum_{i=1}^k z_i\ln(1-e^{-\lambda w_i}) + \sum_{i=1}^k s_i E_{(\alpha_1^{(r)},\lambda^{(r)})}(\ln(1-e^{-\lambda U})|U>w_i)\Big) \\ &\quad +(\alpha_2-1)\Big(\sum_{i=1}^k (1-z_i)\ln(1-e^{-\lambda w_i}) + \sum_{i=1}^k t_i E_{(\alpha_2^{(r)},\lambda^{(r)})}(\ln(1-e^{-\lambda V})|V>w_i)\Big). \end{aligned} \tag{3}$$

Here for any function $h(\cdot, \alpha_1, \alpha_2, \lambda)$,

$$\begin{aligned} E_{(\alpha_1^{(r)},\lambda^{(r)})}(h(U,\alpha_1,\alpha_2,\lambda)|U>w_i) &= \int_{w_i}^{\infty} h(u,\alpha_1,\alpha_2,\lambda)\frac{f_{GE}(u,\alpha_1^{(r)},\lambda^{(r)})}{1-F_{GE}(w_i,\alpha_1^{(r)},\lambda^{(r)})}\,du \\ E_{(\alpha_2^{(r)},\lambda^{(r)})}(h(V,\alpha_1,\alpha_2,\lambda)|V>w_i) &= \int_{w_i}^{\infty} h(v,\alpha_1,\alpha_2,\lambda)\frac{f_{GE}(v,\alpha_2^{(r)},\lambda^{(r)})}{1-F_{GE}(w_i,\alpha_2^{(r)},\lambda^{(r)})}\,dv. \end{aligned}$$

For any fixed λ, $\alpha_1 = \widehat{\alpha}_1(\lambda)$ and $\alpha_2 = \widehat{\alpha}_2(\lambda)$ maximize $g(\alpha_1, \alpha_2, \lambda)$, where,

$$\begin{aligned}
\widehat{\alpha}_1(\lambda) &= -\frac{m}{\sum_{i=1}^{k} z_i \ln(1-e^{-\lambda w_i}) + \sum_{i=1}^{k} s_i E_{(\alpha_1^{(r)},\lambda^{(r)})}(\ln(1-e^{-\lambda U})|U > w_i)}, \\
\widehat{\alpha}_2(\lambda) &= -\frac{n}{\sum_{i=1}^{k} (1-z_i) \ln(1-e^{-\lambda w_i}) + \sum_{i=1}^{k} t_i E_{(\alpha_2^{(r)},\lambda^{(r)})}(\ln(1-e^{-\lambda V})|V > w_i)}.
\end{aligned}$$

Therefore, at the $(r+1)$-th stage, $\lambda^{(r+1)}$ can be obtained by maximizing $g(\widehat{\alpha}_1(\lambda), \widehat{\alpha}_2(\lambda), \lambda)$ with respect to λ. Once $\lambda^{(r+1)}$ is obtained, we compute $\alpha_1^{(r+1)}$ and $\alpha_2^{(r+1)}$ as $\widehat{\alpha}_1(\lambda^{(r+1)}), \widehat{\alpha}_2(\lambda^{(r+1)})$, respectively. We stop the iteration when the convergence takes place.

Therefore, it has been observed that at the 'M'-step of the EM algorithm, the optimization becomes a one dimensional optimization problem. Moreover, in our extensive simulation experiments the EM algorithm has always converged.

4 Observed Information Matrix

In this section following the idea of Louis (1982), we compute the observed information matrix based on the missing value principles. The observed information can be obtained as: Observed Information = Complete Information - Missing Information. We denote the observed information matrix by $I_{obs}(\alpha_1, \alpha_2, \lambda)$, complete information matrix by $I_c(\alpha_1, \alpha_2, \lambda)$ and missing information matrix by $I_{missing}(\alpha_1, \alpha_2, \lambda)$. They are presented below.

$$I_c(\alpha_1, \alpha_2, \lambda) = mI_1(\alpha_1, \alpha_2, \lambda) + nI_2(\alpha_1, \alpha_2, \lambda),$$

where

$$I_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{and} \quad I_2 = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}.$$

Here, $a_{11} = -E\left(\frac{\partial^2 \ln f_{GE}(x,\alpha_1,\lambda)}{\partial \alpha_1^2}\right)$, $a_{12} = a_{21} = a_{22} = a_{23} = a_{32} = 0$,
$a_{13} = a_{31} = -E\left(\frac{\partial^2 \ln f_{GE}(x,\alpha_1,\lambda)}{\partial \alpha_1 \partial \lambda}\right)$, $a_{33} = -E\left(\frac{\partial^2 \ln f_{GE}(x,\alpha_1,\lambda)}{\partial \lambda^2}\right)$

and $b_{11}=b_{12}=b_{13}=b_{21}=b_{31}=0, b_{22}=-E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha_2,\lambda)}{\partial \alpha_2^2}\Big)$,
$b_{23}=b_{32}=-E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha_2,\lambda)}{\partial \alpha_2 \partial \lambda}\Big), b_{33}=-E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha_2,\lambda)}{\partial \lambda^2}\Big)$.

The missing observation matrix can be obtained as follows.

$$I_{missing}(\alpha_1,\alpha_2,\lambda)=\sum_{i=1}^{k}\sum_{j=1}^{s_i} I_{U_{ij}|W_i}(\alpha_1,\alpha_2,\lambda)+\sum_{i=1}^{k}\sum_{j'=1}^{t_i} I_{V_{ij'}|W_i}(\alpha_1,\alpha_2,\lambda),$$

where

$$I_{U_{ij}|W_i}(\alpha_1,\alpha_2,\lambda)=\begin{bmatrix} c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ c_{31} & c_{32} & c_{33}\end{bmatrix} \quad \text{and} \quad I_{V_{ij'}|W_i}(\alpha_1,\alpha_2,\lambda)=\begin{bmatrix} d_{11} & d_{12} & d_{13}\\ d_{21} & d_{22} & d_{23}\\ d_{31} & d_{32} & d_{33}\end{bmatrix}.$$

Here, $c_{11}=-E(\frac{\partial^2 \ln f_{U_{i1}|W_i}(u_{ij}|w_i)}{\partial \alpha_1^2}), c_{12}=c_{21}=c_{22}=c_{23}=c_{32}=0$,
$c_{13}=c_{31}=-E(\frac{\partial^2 \ln f_{U_{i1}|W_i}(u_{ij}|w_i)}{\partial \alpha_1 \partial \lambda}), c_{33}=-E(\frac{\partial^2 \ln f_{U_{i1}|W_i}(u_{ij}|w_i)}{\partial \lambda^2})$,
and $d_{11}=d_{12}=d_{13}=d_{21}=d_{31}=0, d_{22}=-E(\frac{\partial^2 \ln f_{V_{i1}|W_i}(v_{ij}|w_i)}{\partial \alpha_2^2})$,
$d_{23}=d_{32}=-E(\frac{\partial^2 \ln f_{V_{i1}|W_i}(v_{ij}|w_i)}{\partial \alpha_2 \lambda}), d_{33}=-E(\frac{\partial^2 \ln f_{V_{i1}|W_i}(v_{ij}|w_i)}{\partial \lambda^2})$.

Expressions of all the expectations are given in Appendix.

5 Bayesian Inference

In this section we provide the Bayesian inference of the unknown parameters. To provide the Bayesian inference we need to assume some priors on the unknown parameters. In this case for unknown α_1, α_2 and λ, no conjugate prior exists. It is assumed that α_1 and α_2 jointly follow a beta-gamma (BG) distribution. It is well known that four-parameter bivariate beta-gamma distribution is a very flexible absolute continuous distribution. For $a_0, b_0, a, b > 0$, the joint PDF of α_1, α_2 can be written as

$$\pi(\alpha_1,\alpha_2|a_0,b_0,a,b)=\frac{\Gamma(a+b)}{\Gamma(a_0)\Gamma(a)\Gamma(b)} b_0^{a_0}\alpha_1^{a-1}\alpha_2^{b-1}(\alpha_1+\alpha_2)^{a_0-a-b}e^{-b_0(\alpha_1+\alpha_2)}; \qquad (4)$$

for $0<\alpha_1,\alpha_2<\infty$ and zero, otherwise. A beta-gamma distribution with the joint PDF (4) will be denoted by BG(a_0,b_0,a,b). The following results will be used to generate samples from a beta-gamma distribution.

RESULT: $(X,Y) \sim \text{BG}(a_0, b_0, a, b)$ if and only if $U \sim \text{GA}(a_0, b_0)$, $V \sim \text{Beta}(a,b)$ and U, V are independently distributed. Here $\sim$ means follows, $U = (X+Y)$ and $V = \dfrac{X}{X+Y}$.

PROOF: Using the transformation of variable it can be shown easily. ∎

Also

$$\begin{aligned}
E(\alpha_1) &= \frac{a_0}{b_0}\frac{a}{(a+b)}, \quad E(\alpha_2) = \frac{a_0}{b_0}\frac{b}{(a+b)}, \\
E(\alpha_1^2) &= \frac{a_0(a_0+1)}{b_0^2}\frac{a(a+1)}{(a+b)(a+b+1)}, \quad E(\alpha_2^2) = \frac{a_0(a_0+1)}{b_0^2}\frac{b(b+1)}{(a+b)(a+b+1)}, \\
E(\alpha_1\alpha_2) &= \frac{a_0(a_0+1)}{b_0^2}\frac{ab}{(a+b)(a+b+1)}, \quad Cov(\alpha_1,\alpha_2) = \frac{a_0 ab(a+b-a_0)}{b_0^2(a+b)^2(a+b+1)}.
\end{aligned}$$

The beta-gamma prior incorporates different dependency structure between the two shape parameters. When $a_0 > a+b$, α_1, α_2 are positively correlated. For $a_0 < a+b$ they are negatively correlated and independent when $a_0 = a+b$.

As in Raqab and Madi (2005) it is assumed that for $c > 0, d > 0$, the scale parameter $\lambda \sim \text{GA}(c,d)$. It is further assumed that (α_1, α_2) and λ are independently distributed. Based on the observed data the joint posterior density can be obtained as

$$\begin{aligned}
\pi(\alpha_1, \alpha_2, \lambda|data) \;\propto\; & \alpha_1^{a+k_1-1}\alpha_2^{b+k_2-1}\lambda^{c+k-1}\, e^{-\lambda(d+\sum_{i=1}^{k} w_i)} e^{-b_0(\alpha_1+\alpha_2)}(\alpha_1+\alpha_2)^{a_0-a-b} \\
& \times \prod_{i=1}^{k}\left((1-e^{-\lambda w_i})^{\alpha_1-1}\right)^{z_i}\left((1-e^{-\lambda w_i})^{\alpha_2-1}\right)^{1-z_i} \\
& \times \prod_{i=1}^{k}\left(1-(1-e^{-\lambda w_i})^{\alpha_1}\right)^{s_i}\left(1-(1-e^{-\lambda w_i})^{\alpha_2}\right)^{t_i}. \qquad (5)
\end{aligned}$$

The Bayes estimator of any function of $\alpha_1, \alpha_2, \lambda$, say $g(\alpha_1, \alpha_2, \lambda)$ with respect to the squared error loss function can be obtained as

$$E(g(\alpha_1,\alpha_2,\lambda)|data) = \int_0^\infty \int_0^\infty \int_0^\infty g(\alpha_1,\alpha_2,\lambda)\pi(\alpha_1,\alpha_2,\lambda)\; d\alpha_1\; d\alpha_2\; d\lambda, \qquad (6)$$

provided it exists. It is immediate that the Bayes estimators cannot be obtained in explicit forms. We would like to use the importance sampling technique for that purpose.

For further development, (5) can be re-written as

$$\begin{aligned}\pi(\alpha_1,\alpha_2,\lambda|data) \;\propto\; & \alpha_1^{a+k_1-1}\alpha_2^{b+k_2-1}\lambda^{c+k-1}e^{-\lambda(d+\sum_{i=1}^{k}w_i)} \\ & \times(\alpha_1+\alpha_2)^{a_0-a-b}e^{-(\alpha_1+\alpha_2)(b_0-A(\lambda))}e^{-\alpha_1(A(\lambda)-A_1(\lambda))}e^{-\alpha_2(A(\lambda)-A_2(\lambda))} \\ & \times e^{\left(\sum_{i=1}^{k}s_i\ln(1-(1-e^{-\lambda w_i})^{\alpha_1})+\sum_{i=1}^{k}t_i\ln(1-(1-e^{-\lambda w_i})^{\alpha_2})-\sum_{i=1}^{k}\ln(1-e^{\lambda w_i})\right)},\end{aligned}$$

where

$$\begin{aligned}A_1(\lambda) &= \sum_{i=1}^{k}z_i\ln(1-e^{-\lambda w_i}), \\ A_2(\lambda) &= \sum_{i=1}^{k}(1-z_i)\ln(1-e^{-\lambda w_i}), \quad \text{and} \\ A(\lambda) &= \min(A_1(\lambda),A_2(\lambda)).\end{aligned}$$

We rewrite $\pi(\alpha_1,\alpha_2,\lambda|data)$ as

$$\pi(\alpha_1,\alpha_2,\lambda|data) \;\propto\; \pi_1^*(\alpha_1,\alpha_2|\lambda,data)\times\pi_2^*(\lambda|data)\times h(\alpha_1,\alpha_2,\lambda),$$

where

$$\begin{aligned}\pi_1^*(\alpha_1,\alpha_2|\lambda,data) &\sim \text{BG}(a_0+k,b_0-A(\lambda),a+k_1,b+k_2), \\ \pi_2^*(\lambda|data) &\sim \text{GA}(c+k,d+\sum_{i=1}^{k}w_i), \\ h(\alpha_1,\alpha_2,\lambda) &= e^{-\alpha_1(A(\lambda)-A_1(\lambda))}e^{-\alpha_2(A(\lambda)-A_2(\lambda))} \\ &\quad\times e^{\left(\sum_{i=1}^{k}s_i\ln(1-(1-e^{-\lambda w_i})^{\alpha_1})+\sum_{i=1}^{k}t_i\ln(1-(1-e^{-\lambda w_i})^{\alpha_2})-\sum_{i=1}^{k}\ln(1-e^{\lambda w_i})\right)} \\ &\quad\times\frac{1}{(b_0-A(\lambda))^{a_0+k}}.\end{aligned}$$

The following algorithm can be used to compute a simulation consistent estimate of (6).

Algorithm:

Step 1: Given data, generated λ from $\pi_2^*(\lambda|data)$.

Step 2: For a given λ, generate α_1, α_2 from $\pi_1^*(\alpha_1, \alpha_2|\lambda, data)$.

Step 3: Repeat the process say N times to generate $((\alpha_{11}, \alpha_{21}, \lambda_1), \ldots, (\alpha_{1N}, \alpha_{2N}, \lambda_N))$.

Step 4: To compute Bayes estimate of $g(\alpha_1, \alpha_2, \lambda)$ compute $(g_1, \ldots, g_N)$ and $(h_1, \ldots, h_N)$, where $g_i = g(\alpha_{1i}, \alpha_{2i}, \lambda_i)$ and $h_i = h(\alpha_{1i}, \alpha_{2i}, \lambda_i)$.

Step 5: A simulation consistent estimate of $g(\alpha_1, \alpha_2, \lambda)$ can be obtained

$$\frac{\sum_{i=1}^N h_i g_i}{\sum_{j=1}^N h_j} = \sum_{i=1}^N v_i g_i, \quad \text{where} \quad v_i = \frac{h_i}{\sum_{j=1}^N h_j}.$$

Step 6 To compute $100(1-\gamma)\%$ CRI of $g(\alpha_1, \alpha_2, \lambda)$, arrange g_i's in ascending order to obtain $(g_{(1)}, \ldots g_{(N)})$ and record the corresponding v_i's as $(v_{(1)}, \ldots, v_{(N)})$. A $100(1-\gamma)\%$ CRI can be obtained as $(g_{(j_1)}, g_{(j_2)})$ where j_1, j_2 such that

$$j_1 < j_2, \quad j_1, j_2 \in \{1, \ldots, N\} \quad \text{and} \quad \sum_{i=j_1}^{j_2} v_i \le 1-\gamma < \sum_{i=j_1}^{j_2+1} v_i. \tag{7}$$

The $100(1-\gamma)\%$ symmetric credible interval is obtained as $(g_{([N\frac{\gamma}{2}])}, g_{([N(1-\frac{\gamma}{2})])})$. The $100(1-\gamma)\%$ highest posterior density (HPD) CRI can be obtained as $(g_{(j_1^*)}, g_{(j_2^*)})$, where j_1^*, j_2^* such that $g_{(j_2^*)} - g_{(j_1^*)} \le g_{(j_2)} - g_{(j_1)}$ and j_1^*, j_2^* satisfying (7) for all j_1, j_2 satisfying (7). Here $[x]$ is the greatest integer less than or equal to x.

6 Simulation results and Data Analysis

6.1 Simulation results

In this section we perform some simulation experiments to study the performances of the different methods. The simulation study is conducted for different JPC schemes. We have considered $m = 25, n = 20$, effective sample sizes $k = 20$ and 25, and $\alpha_1 = 1.5, \alpha_2 = 2, \lambda = 2$. The MLEs are calculated using the EM algorithm as discussed before. In simulation experiments we set the real values of the parameters as the initial guesses in EM algorithm. In all the cases considered here, it is observed that the estimators obtained using EM algorithm maximize the likelihood function. The Bayes estimates with respect to the squared error loss function are computed based on both informative and non-informative priors. For the informative prior, we have considered the hyper parameters as $a_0 = 3.5$, $b_0 = 1$, $a = 1$, $b = 1.3$, $c = 2$, $d = 1$. These hyper parameters are chosen so that the prior expectations of the two populations match with the corresponding true expected values. Following the idea of Congdon (2014) for the non-informative prior we have taken the hyper parameters as $a_0 = b_0 = a = b = c = d = 10^{-5}$, which are close to zero. Based on maximum likelihood estimates computed by EM algorithm we compute percentile bootstrap confidence intervals. Bayesian symmetric credible intervals are constructed for comparison purposes.

In Table 1 we report average estimates (AE) and the associated mean squared errors (MSE) of the MLEs for different JPC schemes. All these results are obtained based on 1000 replications. In Table 2 we record the average Bayes estimates (BE) and associated MSE both for the informative and non-informative priors. In this case also all the results are based on 1000 replications. Here the notation $\mathscr{R}$=($0_{(2)}$,25,$0_{(17)}$) indicates $R_1 = R_2 = 0, R_3 = 25, R_4 = \cdots = R_{20} = 0$. In Table 3 we record 90% percentile bootstrap confidence intervals along with 90% symmetric credible intervals based on informative and non-informative priors. The average length (AL) and the coverage percentage (CP) of these intervals are computed based on 1000 replications. In bootstrap confidence intervals, in each replication, intervals are computed based on 1000 bootstrap samples.

From Table 1 it is clear that as the effective sample size increases the MLEs are performing better in terms of the MSEs. In all the cases considered the MLEs over-estimate the true parameter values. Another point is evident that the MSEs associated with α_2 are always much bigger compared to the other two MSEs associated with α_1 and λ. From Table 2 it is observed that similarly as the MLEs, the Bayes estimates perform better as the effective sample size k increases, for both the priors. As expected the Bayes estimates perform better for the informative prior than the non-informative prior. For both the priors, the Bayes estimators always over-estimate λ and α_1, while under-estimate α_2. Comparing the MLEs and Bayes estimators it is clear that the Bayes estimators with non-informative priors perform better than the MLEs. Hence, we recommend to use the Bayes estimators with non-informative priors in this case if we do not have any prior information, otherwise informative priors should be preferred. From Table 3, it is clear that the symmetric credible intervals are providing shorter average length the the bootstrap CI. Again the symmetric confidence intervals based on informative prior is performing better than the other two methods.

Table 1: AE and MSE of maximum likelihood estimates based on 1000 simulations with $m = 25, n = 20, \alpha_1 = 1.5, \alpha_2 = 2, \lambda = 2$ for different JPC schemes

Censoring Scheme	Parameter	AE	MSE
k=20,$\mathscr{R}$=(25,$0_{(19)}$)	α_1	1.781	0.572
	α_2	2.885	4.851
	λ	2.291	0.491
k=20,$\mathscr{R}$=($0_{(5)}$,25,$0_{(14)}$)	α_1	1.776	0.475
	α_2	2.569	2.481
	λ	2.294	0.527
k=20,$\mathscr{R}$=($0_{(10)}$,25,$0_{(9)}$)	α_1	1.817	0.578
	α_2	2.539	1.546
	λ	2.355	0.639
k=20,$\mathscr{R}$=($0_{(14)}$,25,$0_{(5)}$)	α_1	1.812	0.590
	α_2	2.559	1.574
	λ	2.335	0.637
k=20,$\mathscr{R}$=($0_{(19)}$,25)	α_1	1.843	0.626
	α_2	2.559	1.574
	λ	2.384	0.726
k=20,$\mathscr{R}$=(10,$0_{(18)}$,15)	α_1	1.850	0.626
	α_2	2.617	1.902
	λ	2.356	0.610
k=20,$\mathscr{R}$=($2_{(12)}$,1,$0_{(7)}$)	α_1	1.772	0.448
	α_2	2.510	1.605
	λ	2.292	0.515
k=20,$\mathscr{R}$=($0_{(7)}$,$2_{(12)}$,1)	α_1	1.866	0.670
	α_2	2.614	2.139
	λ	2.388	0.760
k=25,$\mathscr{R}$=(20,$0_{(24)}$)	α_1	1.767	0.531
	α_2	2.509	1.608
	λ	2.213	0.341
k=25,$\mathscr{R}$=($0_{(5)}$,20,$0_{(19)}$)	α_1	1.700	0.312
	α_2	2.469	1.396
	λ	2.205	0.320
k=25,$\mathscr{R}$=($0_{(10)}$,20,$0_{(14)}$)	α_1	1.748	0.401
	α_2	2.405	1.124
	λ	2.244	0.387
k=25,$\mathscr{R}$=($0_{(15)}$,20,$0_{(9)}$)	α_1	1.805	0.498
	α_2	2.419	1.087
	λ	2.268	0.442
k=25,$\mathscr{R}$=($0_{(20)}$,20,$0_{(4)}$)	α_1	1.773	0.441
	α_2	2.444	1.238
	λ	2.258	0.436
k=25,$\mathscr{R}$=($0_{(24)}$,20)	α_1	1.836	0.561
	α_2	2.535	1.400
	λ	2.330	0.502
k=25,$\mathscr{R}$=(10,$0_{(23)}$,10)	α_1	1.790	0.546
	α_2	2.560	1.750
	λ	2.292	0.453
k=25,$\mathscr{R}$=($2_{(10)}$,$0_{(15)}$)	α_1	1.756	0.471
	α_2	2.546	3.339
	λ	2.246	0.384
k=25,$\mathscr{R}$=($0_{(15)}$,$2_{(10)}$)	α_1	1.789	0.471
	α_2	2.515	1.315
	λ	2.320	0.496

Table 2: BE and MSE of Bayes estimates for informative and non-informative prior based on 1000 simulations with $m = 25, n = 20, \alpha_1 = 1.5, \alpha_2 = 2, \lambda = 2$ for different JPC schemes

Censoring Scheme	Parameter	IP		NIP	
		BE	MSE	BE	MSE
k=20,$\mathscr{R}$=(25,$0_{(19)}$)	α_1	1.615	0.208	1.662	0.392
	α_2	1.928	0.214	1.927	0.319
	λ	2.068	0.199	2.078	0.266
k=20,$\mathscr{R}$=($0_{(5)}$,25,$0_{(14)}$)	α_1	1.562	0.191	1.588	0.290
	α_2	1.723	0.246	1.736	0.324
	λ	2.070	0.189	2.071	0.300
k=20,$\mathscr{R}$=($0_{(10)}$,25,$0_{(9)}$)	α_1	1.552	0.207	1.574	0.309
	α_2	1.679	0.285	1.668	0.406
	λ	2.067	0.221	2.162	0.387
k=20,$\mathscr{R}$=($0_{(14)}$,25,$0_{(5)}$)	α_1	1.563	0.214	1.606	0.371
	α_2	1.661	0.313	1.656	0.456
	λ	2.200	0.311	2.282	0.487
k=20,$\mathscr{R}$=($0_{(19)}$,25)	α_1	1.589	0.238	1.686	0.457
	α_2	1.614	0.373	1.622	0.507
	λ	2.433	0.452	2.607	0.824
k=20,$\mathscr{R}$=($2_{(12)}$,1,$0_{(7)}$)	α_1	1.555	0.198	1.569	0.305
	α_2	1.741	0.244	1.706	0.365
	λ	2.038	0.213	2.089	0.338
k=20,$\mathscr{R}$=($0_{(7)}$,$2_{(14)}$,1)	α_1	1.576	0.211	1.645	0.377
	α_2	1.652	0.335	1.650	0.460
	λ	2.259	0.321	2.390	0.551
k=20,$\mathscr{R}$=(10,$0_{(18)}$,15)	α_1	1.620	0.238	1.727	0.459
	α_2	1.758	0.291	1.797	0.446
	λ	2.227	0.295	2.344	0.469
k=25,$\mathscr{R}$=(20,$0_{(24)}$)	α_1	1.591	0.203	1.620	0.309
	α_2	1.921	0.210	1.887	0.270
	λ	2.026	0.154	2.007	0.189
k=25,$\mathscr{R}$=($0_{(5)}$,20,$0_{(19)}$)	α_1	1.550	0.196	1.560	0.246
	α_2	1.778	0.196	1.752	0.270
	λ	2.003	0.177	2.012	0.219
k=25,$\mathscr{R}$=($0_{(10)}$,20,$0_{(14)}$)	α_1	1.533	0.197	1.551	0.291
	α_2	1.734	0.226	1.689	0.317
	λ	2.020	0.182	2.014	0.248
k=25,$\mathscr{R}$=($0_{(15)}$,20,$0_{(9)}$)	α_1	1.529	0.174	1.569	0.298
	α_2	1.694	0.245	1.675	0.337
	λ	2.049	0.184	2.070	0.289
k=25,$\mathscr{R}$=($0_{(20)}$,20,$0_{(4)}$)	α_1	1.538	0.175	1.609	0.309
	α_2	1.659	0.263	1.664	0.396
	λ	2.106	0.183	2.222	0.325
k=25,$\mathscr{R}$=($0_{(24)}$,20)	α_1	1.582	0.191	1.688	0.374
	α_2	1.669	0.301	1.714	0.409
	λ	2.267	0.274	2.355	0.425
k=25,$\mathscr{R}$=($2_{(10)}$,$0_{(15)}$)	α_1	1.555	0.176	1.582	0.294
	α_2	1.817	0.195	1.797	0.284
	λ	2.006	0.147	2.020	0.229
k=25,$\mathscr{R}$=($0_{(15)}$,$2_{(10)}$)	α_1	1.570	0.175	1.673	0.374
	α_2	1.688	0.277	1.711	0.395
	λ	2.216	0.242	2.305	0.393
k=25,$\mathscr{R}$=(10,$0_{(23)}$,10)	α_1	1.628	0.215	1.671	0.391
	α_2	1.822	0.238	1.848	0.372
	λ	2.131	0.190	2.169	0.320

Table 3: AL and CP of 90% Bootstrap CI and Bayesian 90% symmetric CRI for informative and non-informative prior based on 1000 simulations with $m = 25, n = 20, \alpha_1 = 1.5, \alpha_2 = 2, \lambda = 2$ for different JPC schemes

Censoring Scheme	Parameter	Bootstrap CI		Symmetric CRI (IP)		Symmetric CRI (NIP)	
		AL	CP	AL	CP	AL	CP
k=20,$\mathscr{R}$=(25,$0_{(19)}$)	α_1	2.783	82.3%	1.370	87.1%	1.474	83.9%
	α_2	7.864	77.5%	1.493	83.0%	1.552	77.2%
	λ	2.155	81.6%	1.318	86.5%	1.386	83.0%
k=20,$\mathscr{R}$=($0_{(5)}$,25,$0_{(14)}$)	α_1	2.351	82.1%	1.142	80.2%	1.197	75.5%
	α_2	5.796	79.4%	1.118	65.7%	1.157	58.6%
	λ	2.303	79.2%	1.310	84.3%	1.391	79.3%
k=20,$\mathscr{R}$=($0_{(10)}$,25,$0_{(9)}$)	α_1	2.581	79.7%	1.094	76.5%	1.140	70.4%
	α_2	5.100	78.8%	1.071	58.2%	1.100	50.3%
	λ	2.503	77.2%	1.314	82.6%	1.393	77.2%
k=20,$\mathscr{R}$=($0_{(14)}$,25,$0_{(5)}$)	α_1	2.728	78.5%	1.052	71.6%	1.150	68.1%
	α_2	4.832	78.7%	1.020	53.0%	1.089	48.0%
	λ	2.600	77.9%	1.265	78.1%	1.376	72.5%
k=20,$\mathscr{R}$=($0_{(19)}$,25)	α_1	3.002	74.7%	0.990	71.1%	1.157	63.0%
	α_2	5.436	76.9%	0.979	50.4%	1.062	48.0%
	λ	2.870	74.3%	1.164	60.1%	1.243	50.0%
k=25,$\mathscr{R}$=(20,$0_{(24)}$)	α_1	2.436	82.8%	1.300	89.2%	1.366	83.2%
	α_2	5.548	78.0%	1.382	85.4%	1.408	78.8%
	λ	1.864	77.0%	1.167	88.4%	1.204	85.2%
k=25,$\mathscr{R}$=($0_{(5)}$,20,$0_{(19)}$)	α_1	2.077	83.6%	1.112	80.7%	1.197	79.5%
	α_2	4.509	78.9%	1.112	68.5%	1.141	60.7%
	λ	1.888	82.2%	1.162	85.4%	1.237	82.1%
k=25,$\mathscr{R}$=($0_{(10)}$,20,$0_{(14)}$)	α_1	2.149	80.5%	1.070	78.6%	1.131	73.7%
	α_2	3.885	82.2%	1.044	61.5%	1.073	55.6%
	λ	1.968	81.9%	1.197	84.3%	1.238	79.4%
k=25,$\mathscr{R}$=($0_{(15)}$,20,$0_{(9)}$)	α_1	2.310	80.2%	1.076	78.4%	1.147	74.9%
	α_2	4.020	79.8%	1.043	59.3%	1.113	57.0%
	λ	2.108	78.1%	1.194	83.5%	1.263	77.9%
k=25,$\mathscr{R}$=($0_{(20)}$,20,$0_{(4)}$)	α_1	2.408	79.6%	1.080	77.2%	1.173	73.1%
	α_2	4.272	77.9%	1.052	58.2%	1.129	55.1%
	λ	2.172	78.9%	1.175	80.2%	1.267	76.8%
k=25,$\mathscr{R}$=($0_{(24)}$,20)	α_1	2.560	76.1%	1.009	77.4%	1.177	71.9%
	α_2	4.556	76.5%	1.030	57.0%	1.170	56.7%
	λ	2.237	75.5%	1.112	70.9%	1.221	66.0%

6.2 Data Analysis

The real data sets have been taken from Xia et al. (2009) which represent the breaking strength of jute fiber of gauge length 10 mm and 20 mm. Each group of fibers contains 30 similar fibers. The data are presented for easy references.

Data set 1 (Gauge length 10 mm.): 43.93, 50.16, 101.15 ,108.94, 123.06, 141.38, 151.48, 163.40, 177.25, 183.16, 212.13, 257.44, 262.90, 291.27, 303.90, 323.83, 353.24, 376.42, 383.43, 422.11, 506.60, 530.55, 590.48, 637.66, 671.49, 693.73, 700.74, 704.66, 727.23, 778.17.

Data set 2 (Gauge length 20 mm.): 36.75, 45.58, 48.01, 71.46, 83.55, 99.72, 113.85, 116.99, 119.86, 145.96, 166.49, 187.13, 187.85, 200.16, 244.53, 284.64, 350.70, 375.81, 419.02, 456.60, 547.44, 578.62, 581.60, 585.57, 594.29, 662.66, 688.16, 707.36, 756.70, 765.14.

We divide both the data sets by 1000, and it will not affect the inference procedure. For each data set we fit a two parameter generalized exponential distribution. The MLEs and the Kolmogorov-Smirnov (K-S) distance between empirical distribution functions and fitted distributions along with the corresponding p values are provided in Table 4 for both the data sets.

Table 4: MLEs and K-S distance

Data set	MLE from complete sample		K-S distance	p value
	shape parameter	scale parameter		
Data Set 1	$\alpha_1 = 2.224$	$\lambda_1 = 4.311$	0.100	0.921
Data Set 2	$\alpha_2 = 1.605$	$\lambda_2 = 3.893$	0.149	0.514

To test $H_0 : \lambda_1 = \lambda_2$, we have performed a likelihood-ratio test and the corresponding p value is 0.722. Hence, we cannot reject the null hypothesis. Based on this assumptions we have obtained the MLEs of the unknown parameters as; $\widehat{\alpha}_1 = 2.112$, $\widehat{\alpha}_2 = 1.680$ and $\widehat{\lambda} = 4.103$. From the above data sets we have generated two data sets based on two JPC schemes; Scheme 1: $k = 25$ with $\mathscr{R}$=($2_{(17)}$,1,$0_{(7)}$) and Scheme 2: $k = 30$ with $\mathscr{R}$=($2_{(15)}$,$0_{(15)}$).

Scheme 1: The generated data applying Scheme 1 are presented as:

(36.75,0,1), (43.93,1,1), (45.58,0,1) , (48.01,0,2), (50.16,1,1), (83.55,1,2), (99.72,1,0), (101.15,1,2), (108.94,1,2), (116.99,0,1), (119.86,0,1), (151.48,1,1), (183.16,1,0), (187.13,0,0), (187.85,0,0), (200.16,0,2), (257.44,1,0), (291.27,1,0), (376.42,1,0), (383.43,1,0), (422.11,1,0), (530.55,1,0), (578.62,0,0), (671.49,1,0), (765.14,0,0).

To find out the initial guesses of the EM algorithm, it is assumed that we have two independent complete samples of size k_1 and k_2. i.e. the first sample consists of those w_i's for which z_i's are 1 and the other sample consists of those w_i's for which z_i's are 0. Here $k_1 = 15, k_2 = 10$ and the samples are

$(0.04393, 0.05016, 0.10894, 0.12306, 0.15148, 0.16340, 0.18316, 0.32383, 0.35324, 0.37581,$
$0.50660, 0.67149, 0.69373, 0.70074, 0.77817)$

and

$(0.03675, 0.04558, 0.04801, 0.08355, 0.09972, 0.11385, 0.11699, 0.16649, 0.18785, 0.20016,$
$0.24453, 0.37581, 0.54744, 0.58160, 0.68816, 0.70736).$

Based on these two samples we compute the MLEs of the parameters λ, α_1, α_2 and consider these estimates as the initial guesses in the EM algorithm. For Scheme 1 the initial guesses are 1.771, 1.333, 5.484 for α_1, α_2 and λ, respectively.

In Table 5 we record the maximum likelihood estimates and the Bayes estimates based on the above data sets. The maximum likelihood estimates are calculated using the EM algorithm and in this case the number of iterations to converge the EM algorithm is 21. As theoretically it is difficult to prove the pseudo profile log likelihood function $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ in EM algorithm is uni-modal, we plot $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ for different iterations of EM algorithm. In Fig 1 we plot $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ for the first four iterations of the EM algorithm for Scheme 1 and it is observed that each time it is uni-modal. The Bayes estimates are computed based on non-informative priors. Both the maximum likelihood estimates and the Bayes are slightly higher than the MLEs based on complete data sets. In Table 6 we compute the 90% percentile bootstrap confidence intervals along with the 90% Bayesian symmetric credible intervals. The bootstrap confidence intervals are wider than the symmetric credible intervals for all the parameters.

Table 5: Maximum likelihood estimate and Bayes estimate for Scheme 1

Parameter	maximum likelihood estimate	Bayes estimate
α_1	2.184	2.291
α_2	1.824	1.769
λ	4.516	4.412

Table 6: 90% Bootstrap CI and Bayesian 90% symmetric CRI for Scheme 1

Parameter	Bootstrap CI		symmetric CRI	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
α_1	1.415	3.746	1.072	3.394
α_2	1.435	3.843	1.312	2.455
λ	3.412	7.447	2.783	5.768

Scheme 2: The data generated applying Scheme 2 are presented as: (36.75, 0, 1),(43.93, 1, 0),(45.58, 0, 1), (48.01, 0 ,1),(50.16 , 1, 2),(83.55, 0 ,2), (99.72 , 0, 1),(108.94, 1 ,2),(113.85, 0, 1),(116.99 , 0, 2), (123.06 , 1, 1),(151.48 , 1, 2),(163.40 , 1, 0), (166.49 , 0, 0),(183.16 ,1, 0),(187.85 , 0, 0), (200.16 , 0 ,0),(244.53 , 0 ,0),(323.83 , 1 ,0), (353.24 , 1, 0),(375.81 , 0 ,0),(506.60 , 1, 0, (547.44 , 0 ,0),(581.60, 0, 0),(671.49 , 1, 0), (688.16, 0, 0),(693.73 , 1, 0),(700.74 , 1, 0),(707.36 , 0, 0), (778.17, 1 ,0).

The initial guesses of EM algorithm are computed as described in Scheme 1. The maximum likelihood estimates and the Bayes estimates are recorded in Table 7. Here the number of iterations for EM algorithm to converge is 14. In Fig 2 we plot $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ for different iterations of EM algorithm for Scheme 2. Also EM algorithm produces estimates very close to the estimates based on complete data sets. The Bayes estimates are performing quite well. In Table 8 we compute the 90% bootstrap confidence intervals along with the 90% symmetric Bayesian credible intervals. In this scheme also bootstrap confidence intervals are wider than the symmetric credible intervals.

Table 7: Maximum likelihood estimate and Bayes estimate for Scheme 2

Parameter	maximum likelihood estimate	Bayes estimate
α_1	2.080	2.113
α_2	1.660	1.637
λ	4.071	4.192

Table 8: 90% Bootstrap CI and Bayesian 90% symmetric CRI for Scheme 2

Parameter	Bootstrap CI		symmetric CRI	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
α_1	1.499	3.797	1.290	2.508
α_2	1.199	2.976	1.061	2.334
λ	3.083	6.385	2.891	5.557

Comparing Scheme 1 and Scheme 2 we can conclude that as effective sample size increases the EM algorithm produces estimates very close to the estimates based on complete data and the number of iteration to converge is reducing. Similar conclusion will go with Bayes estimates too.

7 Conclusion

In this article we study the joint progressive censoring scheme (JPC) applied on two generalized exponential populations with same scale parameter and different shape parameters. In likelihood inference the maximum likelihood estimators cannot be obtained in closed form. We apply expectation maximization (EM) algorithm based on missing value principle to derive maximum likelihood estimates of unknown parameters. EM method reduces the likelihood inference into a one dimensional optimization problem. In this context EM algorithm always converge to maximum likelihood estimates. Observed information matrix is derived using missing value principle. Bayesian analysis is performed assuming a beta-Gamma prior for shape parameters and a gamma prior for scale parameter. Beta-gamma prior incorporates different dependency structure between two shape parameters. As the joint prior is not coming out as a conjugate prior, the Bayes estimator out of squared error loss function can not be obtained in closed form. We reply on importance sampling technique to compute the Bayes estimates and the associated credible intervals. In this article we study two generalized exponential populations which can be extended to general number of populations. More work is needed along this direction.

Acknowledgements

The authors would like to thank two unknown reviewers and the Associate Editor for their constructive suggestions which have helped to improve the paper significantly.

Appendix

$$
\begin{aligned}
E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha,\lambda)}{\partial \alpha^2}\Big) &= -\frac{1}{\alpha^2}, \\
E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha,\lambda)}{\partial \alpha \partial \lambda}\Big) &= \begin{cases} \frac{1}{\lambda}[\frac{\alpha}{\alpha-1}(\Psi(\alpha)-\Psi(1)) - (\Psi(\alpha+1)-\Psi(1))], & \text{if } \alpha > 2, \\ \frac{\alpha}{\lambda}\int_0^\infty x e^{-2x}(1-e^{-x})^{\alpha-2}\, dx, & \text{if } 0 < \alpha \le 2. \end{cases} \\
E\Big(\frac{\partial^2 \ln f_{GE}(x,\alpha,\lambda)}{\partial \lambda^2}\Big) &= \begin{cases} -\frac{1}{\lambda^2}[1+\frac{\alpha(\alpha-1)}{\alpha-2}(\Psi'(1)-\Psi'(\alpha-1)+(\Psi(\alpha-1)-\Psi(1))^2)] & \\ -\frac{\alpha}{\lambda^2}[\Psi'(1)-\Psi(\alpha)+(\Psi(\alpha)-\Psi(1))^2], & \text{if } \alpha > 2, \\ -\frac{1}{\lambda^2} - \frac{\alpha(\alpha-1)}{\lambda^2}\int_0^\infty x^2 e^{-2x}(1-e^{-x})^{\alpha-3}\, dx & \text{if } 0 < \alpha \le 2. \end{cases}
\end{aligned}
$$

Here $\Psi()$ and $\Psi'()$ are *digamma* and *trigamma* functions. Readers may refer to Abramowitz and Stegun (1964).

Let $q(x,\alpha,\lambda|x>c)=\frac{f_{GE}(x,\alpha,\lambda)}{1-F_{GE}(c,\alpha,\lambda)}$.

$$E(\frac{\partial^2 \ln q(x,\alpha,\lambda|x>c)}{\partial\alpha^2}) = -\frac{1}{\alpha^2}+[\ln(1-e^{-\lambda c})]^2\frac{(1-e^{-\lambda c})^\alpha}{(1-(1-e^{-\lambda c})^\alpha)^2},$$

$$\begin{aligned} E(\frac{\partial^2 \ln q(x,\alpha,\lambda|x>c)}{\partial\alpha\partial\lambda}) &= h_2(c,\alpha,\lambda)-\frac{ce^{-\lambda c}(1-e^{-\lambda c})^{\alpha-1}}{(1-(1-e^{-\lambda c})^\alpha)^2}\Big[1+\alpha\ln(1-e^{-\lambda c})-(1-e^{-\lambda c})^\alpha\Big], \\ E(\frac{\partial^2 \ln q(x,\alpha,\lambda|x>c)}{\partial\lambda^2}) &= -\frac{1}{\lambda^2}-(\alpha-1)h_1(c,\alpha,\lambda)+\frac{\alpha c^2e^{-\lambda c}(1-e^{-\lambda c})^{\alpha-2}}{(1-(1-e^{-\lambda c})^\alpha)^2}\Big[\alpha e^{-\lambda c}-1 \\ &\quad +(1-e^{-\lambda c})^\alpha\Big], \end{aligned}$$

$$\begin{aligned} \text{where}\quad h_1(c,\alpha,\lambda) &= \frac{1}{\lambda^2(1-(1-e^{-\lambda c})^\alpha)}\int_{(1-e^{-\lambda c})^\alpha}^{1}(\ln(1-x^{\frac{1}{\alpha}}))^2(1-x^{\frac{1}{\alpha}})x^{-\frac{2}{\alpha}}dx, \\ h_2(c,\alpha,\lambda) &= \frac{1}{\lambda(1-(1-e^{-\lambda c})^\alpha)}\int_{(1-e^{-\lambda c})^\alpha}^{1}(-\ln(1-x^{\frac{1}{\alpha}}))(1-x^{\frac{1}{\alpha}})x^{-\frac{1}{\alpha}}dx. \end{aligned}$$

References

[1] Abramowitz, M., and I. A. Stegun. 1964. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. New York:Dover.

[2] Balakrishnan, N., M. Burkschat, E. Cramer, and G. Hofmann. 2008. Fisher information based progressive censoring plans. *Computational Statistics & Data Analysis* 53(2):366-380.

[3] Balakrishnan, N., and E. Cramer. 2014. *The Art of Progressive Censoring: Applications to Reliability and Quality*. New York: Springer.

[4] Balakrishnan, N., F. Su, and K. Y. Liu. 2015. Exact likelihood inference for k exponential populations under joint progressive Type-II censoring. *Communications in Statistics-Simulation and Computation* 44(4):902-923.

[5] Congdon, P. 2014. *Applied Bayesian Modelling* . John Wiley & Sons.

[6] Doostparast, M., M. V. Ahmadi, and J. Ahmadi. 2013. Bayes Estimation Based on Joint Progressive Type II Censored Data Under LINEX Loss Function. *Communications in Statistics-Simulation and Computation* 42(8):1865-1886.

[7] Gupta, R. D., and D. Kundu. 1999. Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics* 41(2):173-188.

[8] Gupta, R. D., and D. Kundu. 2001. Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical journal* 43(1):117-130.

[9] Kundu, D., and R. D. Gupta. 2008. Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis* 52(4):1873-1883.

[10] Kundu, D., and B. Pradhan. 2009. Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring. *Communications in StatisticsTheory and Methods* 38(12):2030-2041.

[11] Louis, T. A. 1982. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 44(2):226-233.

[12] Mondal, S., and D. Kundu. 2017. Point and Interval Estimation of Weibull Parameters Based on Joint Progressively Censored Data. *Sankhya B.* doi:10.1007/s13571-017-0134-1.

[13] Ng, H. K. T., P. S. Chan, and N. Balakrishnan. 2004. Optimal progressive censoring plans for the Weibull distribution. *Technometrics* 46(4):470-481.

[14] Parsi, S., and I. Bairamov. 2009. Expected values of the number of failures for two populations under joint Type-II progressive censoring. *Computational Statistics & Data Analysis* 53(10):3560-3570.

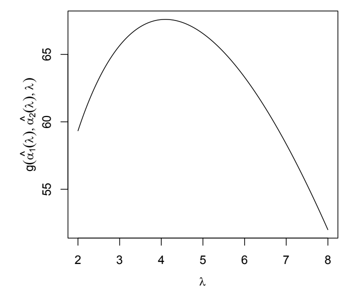
[15] Parsi, S., M. Ganjali, and N. Sanjari Farsipour. 2011. Conditional maximum likelihood and interval estimation for two Weibull populations under joint Type-II progressive censoring. *Communications in Statistics-theory and Methods* 40(12): 2117-2135.

[16] Pradhan, B., and D. Kundu. 2009. On progressively censored generalized exponential distribution. *Test* 18(3):497-515.

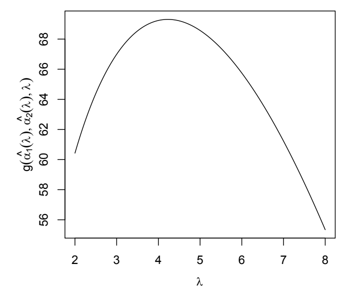
[17] Rasouli, A., and N. Balakrishnan. 2010. Exact likelihood inference for two exponential populations under joint progressive type-II censoring. *Communications in Statistics - Theory and Methods* 39(12):2172-2191.

[18] Raqab, M. Z., and M. T. Madi. 2005. Bayesian inference for the generalized exponential distribution. *Journal of Statistical computation and Simulation* 75(10):841-852.

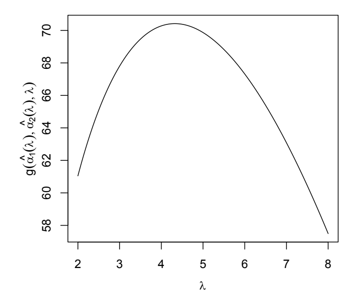
[19] Xia, Z. P., J. Y. Yu, L. D. Cheng, L. F. Liu, and W. M. Wang. 2009. Study on the breaking strength of jute fibers using modified Weibull distribution. *Composites Part A: Applied Science and Manufacturing* 40(1):54-59.



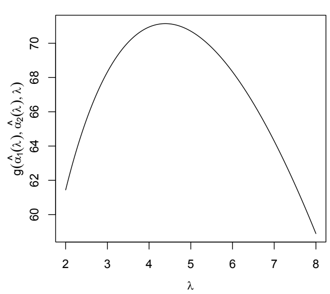
(a) pseudo profile likelihood for 1st iteration of EM algorithm



(b) pseudo profile likelihood for 2nd iteration of EM algorithm

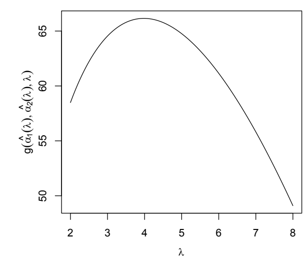


(c) pseudo profile likelihood for 3rd iteration of EM algorithm

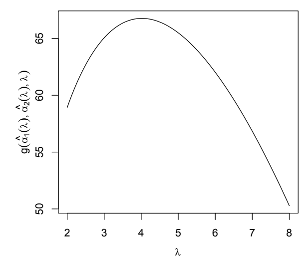


(d) pseudo profile likelihood for 4th iteration of EM algorithm

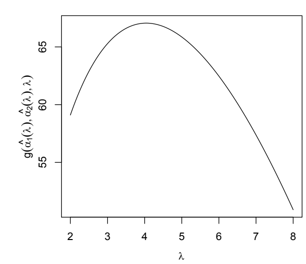
Figure 1: Plot of pseudo profile log likelihood function $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ for different iteration in EM algorithm for Scheme 1.



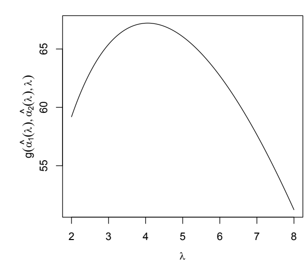
(a) pseudo profile likelihood for 1st iteration of EM algorithm



(b) pseudo profile likelihood for 2nd iteration of EM algorithm



(c) pseudo profile likelihood for 3rd iteration of EM algorithm



(d) pseudo profile likelihood for 4th iteration of EM algorithm

Figure 2: Plot of pseudo profile log likelihood function $g(\hat{\alpha}_1(\lambda), \hat{\alpha}_2(\lambda), \lambda)$ for different iteration in EM algorithm for Scheme 2.