

Bayesian Change Point Estimation Based on Masked Data in Exponential Distribution Parallel System

Yuejun Liu, Huaikou Miao

Abstract—Change point reflects a qualitative change in things. It has gained some applications in the field of reliability. In order to estimate the position parameters of the change point, a Bayesian change point model based on masked data and Gibbs sampling was proposed. By filling in missing lifetime data and introducing latent variables, the simple likelihood function is obtained for exponential distribution parallel system under censored data. This paper describes the probability distributions and random generation methods of the missing lifetime variables and latent variables, and obtains the full conditional distributions of the change point position parameters and other unknown parameters. By Gibbs sampling and estimation of unknown parameters, the estimates of the mean, median, and quantile of the parameter posterior distribution are obtained. The specific steps of Gibbs sampling are introduced in detail. The convergence of Gibbs sampling is also diagnosed. Random simulation results show that the estimations are fairly accurate.

Keywords—parallel system, masked data, change point, Gibbs sample, censored data.

I. INTRODUCTION

Exponential distribution is a very important lifetime distribution. In many cases, it is assumed that the product lifetime follows an exponential distribution, such as the lifetime of an electronic component, the length of talk time on a telephone and the service time of a random service system. Exponential distribution has a wide range of applications in reliability. The change point model is a very important probabilistic statistical model. It is widely used in industrial quality control, economics and other fields.

When the reliability of the system is statistically analyzed, the set of components causing the system failure is deterministic, but the specific component cannot be identified. That is to say, the failure reasons are masked. Such lifetime data are called masked data. Usher and Hodgson first gave the definition of masking, and divided the observations of the

system failure mechanism into three categories: precise, partial and missing. The corresponding data types were competitive failure data, partially masked data and completely masked data. Similarly, Reiser and Guttman obtained the analytical form of Bayesian estimation using Bayesian analysis method [1]. Berger and Sun proposed a new Bayesian method, which introduces the auxiliary variable to describe the failure mechanism of the masking, and finally obtained the Bayesian estimation of the parameters by Gibbs sampling. In recent years, some achievements have been made in the study of masked data and Bayesian estimates. [2-4] conducted statistical analysis of masked data in accelerated life tests. [5-6] studied Bayesian analysis of masked series system series system. [7-9] conducted parameter inference in a hybrid system with masked data.

However, the lifetime distribution problems in above studies have not involved change point. The lifetime distribution study of change point under masked data has not been seen yet. Gibbs sampling as a general method for Bayesian calculation can simplify a change point model. For example, the full conditional distributions of unknown parameters can be transformed into a posterior distribution without change points. This paper mainly discusses Bayesian parameter estimation of change point model based on censored data and masked data in an exponential distribution parallel system. By adding the missing lifetime data and latent variables, a relatively simple likelihood function is obtained. The parameter estimation is obtained based on Gibbs samples which are obtained via Gibbs sampling methods under the full conditional distributions. Random simulation results show that the estimations are fairly accurate.

The rest of the paper is organized as follows: Section II and section III introduce parallel system and the masked data, Section IV describes two-component parallel system with exponential distribution under masked data, Section V discusses Bayesian estimation of change point model, Section VI gives an example of random simulation, and finally, Section VII concludes the study and points out the application value of this method.

This work was partly supported by National Natural Science Foundation of China (No.61572306), Science and Technology Plan Project of Anyang (No.59) and Soft Science Projects of Henan Province (132400410249).

Yuejun Liu is with School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China. (corresponding author; e-mail: aylyj@qq.com).

Huaikou Miao is with School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China.

II. PARALLEL SYSTEM

Assuming that the parallel system is composed of J independent components, the lifetime of the j th component be X_j , and the density function and distribution function are respectively $f_j(t|\theta_j)$ and $F_j(t|\theta_j)$. The lifetime of the parallel system is $T = \max(X_1, \dots, X_J)$. Let K denote the coding of the component that causes the system to fail. If $K = j$, it denotes that the j th component causes the system to fail.

Let $\theta = (\theta_1, \dots, \theta_J)$. Based on the observational data $(T = t, K = j)$, the likelihood function of the parallel system is

$$L(\theta|T = t, K = j) = f_j(t|\theta_j) \prod_{k \neq j} F_k(t|\theta_k).$$

If the observational data t is censored data, and K is not visible, then the likelihood function is

$$\begin{aligned} L(\theta|t) &= P(T > t) = 1 - P(T \leq t) \\ &= 1 - \prod_{k=1}^J F_k(t|\theta_k) \triangleq R_T(t|\theta) \end{aligned}$$

Let δ be an indicator variable. $\delta = 1$ denotes that the observational data are failure data. $\delta = 0$ denotes that the observational data are censored data.

If n parallel systems are independently tested for the lifetime, the likelihood function based on the observational data $\{(t_i, K_i, \delta_i) : i = 1, \dots, n\}$ is

$$\begin{aligned} L(\theta|(t_i, K_i, \delta_i), i = 1, \dots, n) \\ = \prod_{i=1}^n \left\{ [f_{K_i}(t_i|\theta_{K_i}) \prod_{k \neq K_i} F_k(t_i|\theta_k)]^{\delta_i} [R_T(t_i|\theta)]^{1-\delta_i} \right\}. \end{aligned}$$

III. THE MASKED DATA OF PARALLEL SYSTEM

In many cases, the possible failure mechanism set M of the parallel system may be observed, but the exact failure mechanism K is not clear. That is to say, we only know that K belongs to the set $M \subseteq \{1, \dots, J\}$, but do not know the exact value of K . So, the observational data (T, M) are called masked data. If M is a single point set, the observational data are the usual data (T, K) described previously; If $M = \{1, \dots, J\}$, the observational data (T, M) are called completely masked data.

The likelihood function based on failure data of a single point $(T = t, M = M_0)$ is as follows

$$\begin{aligned} L(\theta|T = t, M = M_0) \\ &= \sum_{j \in M_0} P(T = t, M = M_0, K = j) \\ &= \sum_{j \in M_0} P(T = t, K = j) P(M = M_0 | T = t, K = j) \\ &= \sum_{j \in M_0} P(M = M_0 | T = t, K = j) f_j(t|\theta_j) \prod_{k \neq j} F_k(t|\theta_k), \end{aligned}$$

where $P(M = M_0 | T = t, K = j)$ is called masking probability.

The likelihood function based on masked data $\{(t_i, M_i, \delta_i) : i = 1, \dots, n\}$ of n independent parallel systems is

$$\begin{aligned} L(\theta|(t_i, K_i, \delta_i), i = 1, \dots, n) \\ = \prod_{i=1}^n \left\{ \left[\sum_{j \in M_i} P(M = M_i | T = t_i, K = j) f_j(t_i|\theta_j) \right. \right. \\ \left. \left. \prod_{k \neq j} F_k(t_i|\theta_k) \right]^{\delta_i} [R_T(t_i|\theta)]^{1-\delta_i} \right\}. \end{aligned}$$

It is assumed that the masking probability has nothing to do with the failure time but something to do with the failure mechanism, i.e.,

$$\begin{aligned} P(M = M_0 | T = t, K = j) \\ = P(M = M_0 | K = j) \equiv p_j(M_0). \end{aligned} \quad (1)$$

At this time, the following likelihood function can be obtained

$$\begin{aligned} L(\theta|(t_i, K_i, \delta_i), i = 1, \dots, n) \\ = \prod_{i=1}^n \left\{ \left[\sum_{j \in M_i} p_j(M_i) f_j(t_i|\theta_j) \prod_{k \neq j} F_k(t_i|\theta_k) \right]^{\delta_i} [R_T(t_i|\theta)]^{1-\delta_i} \right\}. \end{aligned} \quad (2)$$

IV. TWO-COMPONENT PARALLEL SYSTEM WITH EXPONENTIAL DISTRIBUTION UNDER MASKED DATA

If the probability density function of X is $f(x|\lambda) = \lambda e^{-\lambda x}$ ($x \geq 0, \lambda > 0$), X follows exponential distribution. It is denoted as $X \sim \text{Exp}(\lambda)$. Now the distribution function of X is $F(x|\lambda) = 1 - e^{-\lambda x}$. In two-component parallel system, if the j th component lifetime X_j follows $\text{Exp}(\lambda_j)$ ($j = 1, 2$), we can get $\theta_1 = \lambda_1, \theta_2 = \lambda_2$. The following contents are the lifetime test of two-component parallel systems.

The expressions of $F_j(t|\theta_j)$ and $R_T(t|\theta)$ are polynomials. Next, adding the missing lifetime value of X_j simplifies the formula (2):

When $\delta_i = 1$ and $k \neq j$, we add the missing lifetime value $X_k = Z_{ki} = z_{ki}$, where Z_{ki} follows the truncated exponential distribution $\text{Exp}(\lambda_k)$ on the interval $(0, t_i]$.

When $\delta_i = 0$, we add the missing lifetime value $X_1 = Y_{1i} = y_{1i}$ and $X_2 = Y_{2i} = y_{2i}$.

Then the joint density functions of Y_{1i} and Y_{2i} is

$$\begin{aligned} H(y_{1i}, y_{2i} | \lambda_1, \lambda_2, t_i) \\ = \frac{f_1(y_{1i} | \lambda_1) f_2(y_{2i} | \lambda_2)}{R_T(t_i | \theta)}, \max(y_{1i}, y_{2i}) > t_i. \end{aligned}$$

y_{1i}, y_{2i} can be obtained by random method. The steps are as follows:

(1) Generate randomly numbers x_1 of $\text{Exp}(\lambda_1)$ and x_2 of $\text{Exp}(\lambda_2)$ respectively;

(2) If $\max(x_1, x_2) > t_i$, let $y_{1i} = x_1, y_{2i} = x_2$, and stop;

(3) If $\max(x_1, x_2) \leq t_i$, return to step (1).

Therefore, the Likelihood function is

$$L(\theta|(t_i, K_i, \delta_i), i = 1, \dots, n) \\ = \prod_{i=1}^n \{ [\sum_{j \in M_i} p_j (M_i) f_j(t_i | \lambda_j) \prod_{k \neq j} f_k(z_{2i} | \lambda_k)]^{\delta_i} \\ [f_1(y_{1i} | \lambda_1) f_2(y_{2i} | \lambda_2)]^{1-\delta_i} \}.$$

Assume that the masking probability follows the formula (1). Let $P(M = \{1\} | T = t, K = 1) = p_1$ and $P(M = \{2\} | T = t, K = 2) = p_2$, then the following two formulas hold:

$$P(M = \{1, 2\} | T = t, K = 1) = 1 - p_1 \triangleq q_1,$$

$$P(M = \{1, 2\} | T = t, K = 2) = 1 - p_2 \triangleq q_2.$$

To simplify the writing, we introduce the variables α_i and β_i . The following three classifications are available:

(1) If $M_i = \{1\}$, then $\alpha_i = 1, \beta_i = 0$;

(2) If $M_i = \{2\}$, then $\alpha_i = 0, \beta_i = 1$;

(3) If $M_i = \{1, 2\}$, then $\alpha_i = 0, \beta_i = 0$.

Let $p = (p_1, p_2)$, the likelihood function is

$$L(\theta, p) = \prod_{i=1}^n \{ [p_1 f_1(t_i | \lambda_1) f_2(z_{2i} | \lambda_2)]^{\alpha_i} [p_2 f_2(t_i | \lambda_2) f_1(z_{1i} | \lambda_1)]^{\beta_i} \\ [q_1 f_1(t_i | \lambda_1) f_2(z_{2i} | \lambda_2) + q_2 f_2(t_i | \lambda_2) f_1(z_{1i} | \lambda_1)]^{1-\alpha_i-\beta_i} \}^{\delta_i} \\ [f_1(y_{1i} | \theta_1) f_2(y_{2i} | \theta_2)]^{1-\delta_i}.$$

When $\delta_i = 1, \alpha_i = 0, \beta_i = 0$, i.e., the observational data are $(t_i, M_i = \{1, 2\})$, the distribution rule of latent variable v_i is as follows:

$$P(v_i = 1) = P(K = 1 | T = t_i, M_i = \{1, 2\}) \\ = \frac{P(K = 1, T = t_i, M_i = \{1, 2\})}{P(T = t_i, M_i = \{1, 2\})} \\ = \frac{P(M_i = \{1, 2\} | T = t_i, K = 1) P(T = t_i, K = 1)}{\sum_{j=1}^2 P(M_i = \{1, 2\} | T = t_i, K = j) P(T = t_i, K = j)} \\ = \frac{q_1 f_1(t_i | \lambda_1) f_2(z_{2i} | \lambda_2)}{q_1 f_1(t_i | \lambda_1) f_2(z_{2i} | \lambda_2) + q_2 f_2(t_i | \lambda_2) f_1(z_{1i} | \lambda_1)} \\ = \frac{q_1 \lambda_1 e^{-\lambda_1 t_i} \lambda_2 e^{-\lambda_2 z_{2i}}}{q_1 \lambda_1 e^{-\lambda_1 t_i} \lambda_2 e^{-\lambda_2 z_{2i}} + q_2 \lambda_2 e^{-\lambda_2 t_i} \lambda_1 e^{-\lambda_1 z_{1i}}} \triangleq g \\ = g(\lambda_1, \lambda_2, p_1, p_2, z_{1i}, z_{2i}, t_i)$$

where $v_i = 1$ means that the system failure is caused by the first component, and $v_i = 0$ means that the system failure is caused by the second component.

It can be found that v_i follows the binomial distribution $b(1, g)$.

After introducing the latent variable v_i , the likelihood function is

$$L(\theta, p) = \prod_{i=1}^n \{ (p_1 \lambda_1 e^{-\lambda_1 t_i} \lambda_2 e^{-\lambda_2 z_{2i}})^{\alpha_i} (p_2 \lambda_2 e^{-\lambda_2 t_i} \lambda_1 e^{-\lambda_1 z_{1i}})^{\beta_i} \\ [(q_1 \lambda_1 e^{-\lambda_1 t_i} \lambda_2 e^{-\lambda_2 z_{2i}})^{v_i} (q_2 \lambda_2 e^{-\lambda_2 t_i} \lambda_1 e^{-\lambda_1 z_{1i}})^{1-v_i}]^{1-\alpha_i-\beta_i} \}^{\delta_i} \\ (\lambda_1 e^{-\lambda_1 y_{1i}} \lambda_2 e^{-\lambda_2 y_{2i}})^{1-\delta_i} \\ \propto (\lambda_1^l e^{-\lambda_1 s_1}) (\lambda_2^l e^{-\lambda_2 s_2}) (p_1^{n_1} q_1^{r_1}) (p_2^{n_2} q_2^{r_2}),$$

where

$$s_1 = s_1(D) \\ = \sum_{i \in D} \delta_i \{ \alpha_i t_i + \beta_i z_{1i} + (1 - \alpha_i - \beta_i) [v_i t_i + (1 - v_i) z_{1i}] \} + (1 - \delta_i) y_{1i}, \\ s_2 = s_2(D) \\ = \sum_{i \in D} \delta_i \{ \beta_i t_i + \alpha_i z_{2i} + (1 - \alpha_i - \beta_i) [(1 - v_i) t_i + v_i z_{2i}] \} + (1 - \delta_i) y_{2i}, \\ n_1 = n_1(D) \\ = \sum_{i \in D} \delta_i \alpha_i, r_1 = r_1(D) = \sum_{i \in D} \delta_i (1 - \alpha_i - \beta_i) v_i, \\ n_2 = n_2(D) \\ = \sum_{i \in D} \delta_i \beta_i, r_2 = r_2(D) = \sum_{i \in D} \delta_i (1 - \alpha_i - \beta_i) (1 - v_i), \\ l = l(D) = \sum_{i \in D} 1 = n.$$

Note that s_i is the sum of the lifetime of all i -th components in n parallel systems. For all the i -th components of n parallel systems which cause the system to fail, n_i and r_i respectively indicate the number of unmasked components and the number of masked components.

V. BAYESIAN ESTIMATION OF CHANGE POINT MODEL

Let $D_1 = \{1, 2, \dots, k\}$ and $D_2 = \{k+1, k+2, \dots, n\}$.

For the i -th parallel system, when $i \in D_m$, we have $(p_1, p_2, \lambda_1, \lambda_2) = (p_{1m}, p_{2m}, \lambda_{1m}, \lambda_{2m})$ ($m = 1, 2$). If $p_{11} \neq p_{12}, p_{21} \neq p_{22}, \lambda_{11} \neq \lambda_{12}$ and $\lambda_{21} \neq \lambda_{22}$, the model is called a change point model of two-component parallel system with exponential distribution under the masked data. k is called the change point position parameter.

If the change points exist in the above model, the Bayesian estimation of the unknown parameters is as follows.

Let $\gamma = (k, \lambda_{11}, \lambda_{21}, p_{11}, p_{21}, \lambda_{12}, \lambda_{22}, p_{12}, p_{22})$.

The vectors $\mathbf{T}, \mathbf{M}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{y}_1, \mathbf{y}_2, \boldsymbol{\delta}, \mathbf{v}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ are composed of $t_i, M_i, z_{1i}, z_{2i}, y_{1i}, y_{2i}, \delta_i, v_i, \alpha_i, \beta_i$ respectively. The likelihood function of the change point model is

$$L(\mathbf{T}, \mathbf{M}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{y}_1, \mathbf{y}_2, \boldsymbol{\delta}, \mathbf{v}, \boldsymbol{\alpha}, \boldsymbol{\beta} | \boldsymbol{\gamma}) \\ \propto \prod_{m=1}^2 (\lambda_{1m}^{l_m} e^{-\lambda_{1m} s_{1m}}) (\lambda_{2m}^{l_m} e^{-\lambda_{2m} s_{2m}}) (p_{1m}^{n_{1m}} q_{1m}^{r_{1m}}) (p_{2m}^{n_{2m}} q_{2m}^{r_{2m}}),$$

where

$$l_m = l(D_m), \quad s_{1m} = s_1(D_m), \quad s_{2m} = s_2(D_m), \\ n_{1m} = n_1(D_m), \quad r_{1m} = r_1(D_m), \quad r_{2m} = r_2(D_m), \\ (m = 1, 2).$$

The prior distribution of parameters can be determined by the following methods:

(1) Let the prior distribution of k be a discrete uniform distribution $\pi(k) = 1/(n-1)$, $1 \leq k \leq n-1$.

(2) Let the prior distribution of λ_{im} be a gamma distribution $Ga(a_m, b_m)$:

$$\pi(\lambda_{im}) \propto \lambda_{im}^{a_m-1} e^{-b_m \lambda_{im}}, \lambda_{im} > 0, i = 1, 2; m = 1, 2.$$

(3) Let the prior distribution of p_{im} be a beta distribution $Be(c_m, d_m)$: $\pi(p_{im}) \propto p_{im}^{c_m-1} (1-p_{im})^{d_m-1}$, $0 < p_{im} < 1$.

Assuming that k, λ_{im} and p_{im} are independent of each other, where $i = 1, 2$ and $m = 1, 2$, the following formula holds:

$$L(\gamma | T, M, z_1, z_2, y_1, y_2, \delta, v, \alpha, \beta) \\ \propto L(T, M, z_1, z_2, y_1, y_2, \delta, v, \alpha, \beta | \gamma) \pi(k) \prod_{m=1}^2 \prod_{i=1}^2 \pi(p_{im}) \pi(\lambda_{im}) \\ \propto \prod_{m=1}^2 \left[\lambda_{1m}^{l_m+a_{1m}-1} e^{-\lambda_{1m}(s_{1m}+b_{1m})} \lambda_{2m}^{l_m+a_{2m}-1} e^{-\lambda_{2m}(s_{2m}+b_{2m})} \right. \\ \left. p_{1m}^{n_{1m}+c_{1m}-1} q_{1m}^{r_{1m}+d_{1m}-1} p_{2m}^{n_{2m}+c_{2m}-1} q_{2m}^{r_{2m}+d_{2m}-1} \right].$$

If $\delta_i = 1, \alpha_i = 0, \beta_i = 0$, i.e., the observational data are $T = t_i, M_i = \{1, 2\}$, we have

$$\pi(v_i | T, M, z_1, z_2, y_1, y_2, \delta, v_{-i}, \alpha, \beta) \\ \propto b(1, g_{im}), i \in D_m, m = 1, 2,$$

where $v_{-i} = \{v_j : j \neq i\}$, $g_{im} = g(\lambda_{1m}, \lambda_{2m}, p_{1m}, p_{2m}, z_{1i}, z_{2i}, t_i)$.

If $\delta_i = 1, \alpha_i = 1, \beta_i = 0$, i.e., the observational data are $T = t_i, M_i = \{1\}$ or $\delta_i = 1, \alpha_i = 0, \beta_i = 0, v_i = 1$, i.e., the observational data are $T = t_i, M_i = \{1, 2\}$ and the system failure is caused by the first component, we have $\pi(z_{2i} | T, M, z_1, z_{-2i}, y_1, y_2, \delta, v, \alpha, \beta)$, which follows the truncated exponential distribution $Exp(\lambda_{2m}), i \in D_m (m = 1, 2)$ on the interval $(0, t_i]$.

If $\delta_i = 1, \alpha_i = 0, \beta_i = 1$, i.e., the observational data are $T = t_i, M_i = \{2\}$ or $\delta_i = 1, \alpha_i = 0, \beta_i = 0, v_i = 0$, i.e., the observational data are $T = t_i, M_i = \{1, 2\}$ and the system failure is caused by the second component, we have $\pi(z_{1i} | T, M, z_{-1i}, z_2, y_1, y_2, \delta, v, \alpha, \beta)$, which follows the truncated exponential distribution $Exp(\lambda_{1m}), i \in D_m, m = 1, 2$ on the interval $(0, t_i]$.

If $\delta_i = 0$, i.e., the observational data is $T = t_i$, we have

$$\pi(y_{1i}, y_{2i} | T, M, z_1, z_2, y_{-1i}, y_{-2i}, \delta, v, \alpha, \beta) \\ \propto H(y_{1i}, y_{2i} | \lambda_{1m}, \lambda_{2m}, t_i), i \in D_m, m = 1, 2.$$

The full conditional distribution of λ_{11} can be abbreviated as $\pi(\lambda_{11} | \cdot)$.

The full conditional distributions of parameters can be determined as follows:

$$\pi(\lambda_{im} | \cdot) \propto \lambda_{im}^{l_m+a_m-1} e^{-\lambda_{im}(s_{im}+b_m)} \propto Ga(l_m + a_m, s_{im} + b_m), \\ i = 1, 2; m = 1, 2;$$

$$\pi(p_{im} | \cdot) \propto p_{im}^{n_{im}+c_{im}-1} (1-p_{im})^{r_{im}+d_{im}-1} \propto Be(n_{im} + c_{im}, r_{im} + d_{im});$$

$$\pi(k | \cdot) \propto \prod_{m=1}^2 (\lambda_{1m}^{l_m} e^{-\lambda_{1m} s_{1m}}) (\lambda_{2m}^{l_m} e^{-\lambda_{2m} s_{2m}}) (p_{1m}^{n_{1m}} q_{1m}^{r_{1m}}) (p_{2m}^{n_{2m}} q_{2m}^{r_{2m}}) \\ \propto [\lambda_{11} \lambda_{21} (\lambda_{12} \lambda_{22})^{-1}]^k e^{-[(\lambda_{11} - \lambda_{12}) s_{11} + (\lambda_{21} - \lambda_{22}) s_{21}]} \\ (p_{11} p_{12})^{n_{11}} (p_{21} p_{22})^{n_{21}} (q_{11} q_{12})^{r_{11}} (q_{21} q_{22})^{r_{21}}.$$

z_{1i}, z_{2i} can be obtained by using the inverse transform method to sample the truncated exponential distribution; (y_{1i}, y_{2i}) can be obtained by random method; k can be obtained by inverse transform method. The full conditional distributions of $v_i, \lambda_{im}, p_{im}$ can be obtained by direct Gibbs sampling.

The steps of the Gibbs sampling are as follows:

Let the initial value

$$\gamma^{(0)} = (k^{(0)}, \lambda_{11}^{(0)}, \lambda_{21}^{(0)}, p_{11}^{(0)}, p_{21}^{(0)}, \lambda_{12}^{(0)}, \lambda_{22}^{(0)}, p_{12}^{(0)}, p_{22}^{(0)})$$

and the estimate value at the beginning of the t -th iteration be $\gamma^{(t-1)}$. Then the t -th iteration is split into seven steps:

(1) If $i \in D_m, \delta_i = 1, \alpha_i = 0, \beta_i = 0$, v_i is updated by $v_i^{(t)}$ which is obtained by the binomial distribution function $b(1, g_{im})$ sampling.

(2) If $i \in D_m, \delta_i = 1, \alpha_i = 1, \beta_i = 0$ or $i \in D_m, \delta_i = 1, \alpha_i = 0, \beta_i = 0, v_i = 1$, z_{2i} is updated by $z_{2i}^{(t)}$ which is obtained by the truncated exponential distribution function $Exp(\lambda_{2m})$ on the interval $(0, t_i]$ sampling.

(3) If $i \in D_m, \delta_i = 1, \alpha_i = 0, \beta_i = 1$ or $i \in D_m, \delta_i = 1, \alpha_i = 0, \beta_i = 0, v_i = 0$, z_{1i} is updated by $z_{1i}^{(t)}$ which is obtained by the truncated exponential distribution function $Exp(\lambda_{1m})$ on the interval $(0, t_i]$ sampling.

(4) If $i \in D_m$ and $\delta_i = 0$, (y_{1i}, y_{2i}) is updated by $(y_{1i}^{(t)}, y_{2i}^{(t)})$ which is obtained by the joint density function $H(y_{1i}, y_{2i} | \lambda_{1m}, \lambda_{2m}, t_i)$ sampling.

(5) If $i \in D_m$, λ_{im} is updated by $\lambda_{im}^{(t)}$ which is obtained by the gamma distribution function $Ga(l_m + a_m, s_{im} + b_m)$ sampling.

(6) If $i \in D_m$, p_{im} is updated by $p_{im}^{(t)}$ which is obtained by the beta distribution function $Be(n_{im} + c_{im}, r_{im} + d_{im})$ sampling.

(7) k is updated by $k^{(t)}$ which is obtained by the full conditional distribution function $\pi(k | \cdot)$ sampling. $\gamma^{(t)}$ is called Gibbs samples.

It is assumed that M times Gibbs samples are performed and the sampling converges after the B -th iteration. The parameters can be estimated according to the last $M-B$ iterative values. For example, the mean and the median of the last $M-B$ iterative values can be used as the posterior expectation estimation and the posterior median estimation, respectively.

VI. RANDOM SIMULATION

Let the number of the parallel system be $n=300$, and the initial value of $(k, \lambda_{11}, \lambda_{21}, p_{11}, p_{21}, \lambda_{12}, \lambda_{22}, p_{12}, p_{22})$ be $(130, 3, 5, 0.35, 0.8, 4.5, 2.5, 0.7, 0.5)$. Let's take the random censored lifetime test, and take the censored variable as $Y_i \sim Exp(2.8)$. The simulation data can be generated as follows :

(1) The k groups data x_{1i}, x_{2i}, w_i ($i=1, 2, \dots, k$) are independently generated from $Exp(\lambda_{11}), Exp(\lambda_{21}), Exp(2.8)$. If $\max(x_{1i}, x_{2i}) > w_i$, then w_i is truncated data. If $\max(x_{1i}, x_{2i}) \leq w_i$, then $\max(x_{1i}, x_{2i})$ is the failure data and is marked from component 1 or component 2.

(2) For the data generated in step (1): if the data are marked as coming from component 1, the new data are randomly selected by $100(1-p_{11})\%$ to mask; if the data are marked from component 2, the new data are randomly selected by $100(1-p_{21})\%$ to mask.

(3) Replace $(k, \lambda_{11}, \lambda_{21}, p_{11}, p_{21})$ with $(n-k, \lambda_{12}, \lambda_{22}, p_{12}, p_{22})$ in step (1), and then repeat step (1) and step (2).

Let the prior distributions of $\lambda_{11}, \lambda_{21}, \lambda_{12}, \lambda_{22}$ as $Ga(7.5, 2), Ga(0.6, 0.15), Ga(22, 5), Ga(2.8, 1.3)$ respectively.

Let the prior distributions of $p_{11}, p_{21}, p_{12}, p_{22}$ as $Be(4, 7), Be(17, 5), Be(8.5, 7), Be(12, 11)$ respectively.

The change point position parameter k is estimated and analyzed, when $B = 10000, M = 20000$, the estimated values of each parameter are shown in Table 1.

Next, the correlation of Gibbs sample is tested. If Gibbs sample are autocorrelated, Gibbs sampling will converge slowly and be less efficient. The autocorrelation plot of the Gibbs sample of parameter k is shown in Figure 1. Whether or not traversal mean converges can be used to determine whether the Gibbs sample converges. To ensure that the variables that calculate the mean are approximately independent, one sample is usually extracted every other segment in the chain sampled by Gibbs. The traversal mean of the Gibbs sample of k is shown in Figure 2.

It can be seen from Table 1. When the mean value of the Gibbs sample is used as the parameter value for Bayesian estimation, the relative errors of the parameter k and other parameters are less than 4.5% and less than 10%, respectively, and the Monte Carlo error is also small. The estimation accuracy is higher. The 0.95 confidence interval for each parameter can be approximated as [2.5% percentile, 97.5% percentile]. The length of the approximate confidence interval is very short, so the effect of interval estimation is also good. It can be seen from Figure 1 and Figure 2: The autocorrelation coefficient of the Gibbs sample decays to 0 quickly, which indicates that the convergence of Gibbs sampling is fast; when the iteration reaches 10,000 times, the travel mean has become stable, i.e., the Gibbs sampling has converged. In summary, the random simulation works well.

Table 1. Random Simulation Results

Parameter Value	Mean	Relative error	Monte Carlo error	92.5% percentile	Median	97.5% percentile	
k	130	135.789700	0.044536	0.055636	123	137	147
l_{11}	3	3.11171	0.037237	0.003624	2.51844	3.106337	3.704439
l_{21}	5	5.201003	0.040201	0.006003	4.214427	5.201701	6.18801
l_{12}	4.5	4.387944	0.024901	0.005066	3.546424	4.398168	5.21506
l_{22}	2.5	2.643648	0.057459	0.003091	2.14041	2.646967	3.14993
p_{11}	0.35	0.323423	0.075934	0.000373	0.262435	0.322739	0.385228
p_{21}	0.8	0.747434	0.065708	0.000859	0.605844	0.747836	0.889692
p_{12}	0.7	0.637267	0.089619	0.000736	0.515933	0.637879	0.75752
p_{22}	0.5	0.545179	0.090357	0.000633	0.441414	0.546213	0.649042

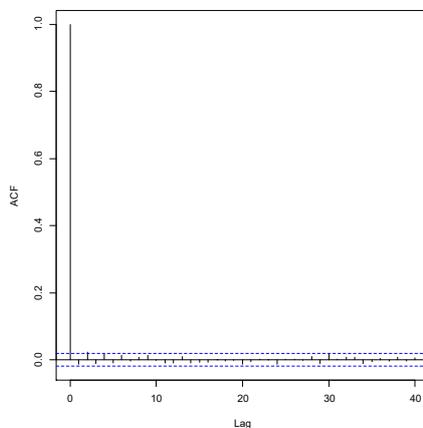


Figure 1. Autocorrelation of the Gibbs sample

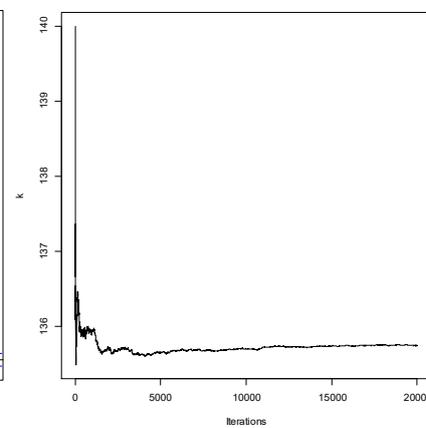


Figure 2. Traversal mean of the Gibbs sample

VII. CONCLUSION

In this paper, we consider Bayesian change point estimation for exponential distribution parallel system under masked data. We fill in all the missing data of interest and introduce latent variables by sampling methods. We study the probability distributions and random generation methods of the missing lifetime data and latent variables. The full conditional distributions of all parameters are discussed. Bayesian estimations of parameters are studied by Gibbs sampling of MCMC methods. Our simulation results show that Bayesian parameter estimations are fairly accurate and the effect of simulation is good using MCMC methods. In this paper, we make statistical inference on the base of Gibbs sampling, so Bayesian approach is common and suitable for change point problem of other life distributions and has its value in popularization.

REFERENCES

- [1] Reiser B , et al., "Bayesian Inference for Masked System Lifetime Data", *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, vol.44, no.1, pp. 79–90, 1995.
- [2] Xu, Ancha, Sanjib Basu, and Yincai Tang, "A Full Bayesian Approach for Masked Data in Step-Stress Accelerated Life Testing", *IEEE Transactions on Reliability*, vol.63, no.3, pp. 798-806, 2014.
- [3] Liu, Bin , et al., "Reliability Analysis of Masked Data in Adaptive Step-Stress Partially Accelerated Lifetime Tests with Progressive Removal", *Communications in Statistics - Theory and Methods*, vol.46, no.12, pp. 6174-6191, 2017.
- [4] Cai, Jing, Yimin Shi, and Hongbo Yue, "Accelerated life tests for log-normal series system with dependent masked data under Type-I progressive hybrid censoring", *Communications in Statistics-Simulation and Computation*, vol.46, no.2, pp. 1628-1646, 2017.
- [5] Tomer, Sanjeev K., Ashok K. Singh, and M. S. Panwar, "Bayesian analysis of masked series system lifetime data from a family of lifetime distributions", *International Journal of System Assurance Engineering and Management*, vol.6, no.4, pp. 495-502, 2014.
- [6] Cai, Jing, Yimin Shi, and Bin Liu, "Inference for a series system with dependent masked data under progressive interval censoring", *Journal of Applied Statistics*, vol.44, no.1, pp. 2-15, 2017.
- [7] Wang, Ronghua, et al., "Parameter inference in a hybrid system with masked data", *IEEE Transactions on Reliability*, vol.64, no.2, pp.636-644, 2015.
- [8] Cai, Jing, Yimin Shi, and Xuchao Bai, "Statistical analysis of masked data in a hybrid system based on copula theory under progressive hybrid censoring", *Sequential Analysis*, vol.36, no.2, pp. 240-250, 2017.
- [9] Liu, Bin, et al., "Nonparametric bayesian analysis for masked data from hybrid systems in accelerated lifetime tests", *IEEE Transactions on Reliability*, vol. 66, no.3, pp. 662-676, 2017.

Yuejun Liu was born on July 7, 1976. He is a Ph.D. candidate at school of Computer Engineering and Science of Shanghai University and is a researcher (associate professor) at Anyang Normal University, China. His major research interests include formal methods and software reliability.

Huaikou Miao is currently a professor in computer engineering and science with Shanghai University, China. His research interests include formal methods and software engineering.