مجلّه‌ی پژوهش‌های آماری ایران
پژوهشکده‌ی آمار
دوره‌ی ۴، شماره‌ی ۲، پاییز و زمستان ۱۳۸۶، صص ۲۰۳-۲۱۶

J. Statist. Res. Iran **4** (2007): 203-216

Bayesian Estimation of the Multiple Change Points in a Gamma Process Using $\bar{X}$ Chart

M. Esmail Dehghan Monfared* and Mohammad Reza Meshkani

Shahid Beheshti University

Abstract. The process personnel always seek the opportunity to improve the processes. One of the essential steps for process improvement is to quickly recognize the starting time or the change point of a process disturbance. Different from the traditional normally distributed assumption for a process, this study considers a process which follows a gamma process. In addition, we consider the possibility of the existence of more than one change point. The proposed approach combines the commonly used $\bar{X}$ control chart with the Bayesian estimation technique using reversible jump Markov chain Monte Carlo method (RJMCMC) to obtain Bayes estimates. The efficiency of our proposed method is evaluated through a series of simulations. The results show that in many cases if there exist more than one change point, our proposed method is able to estimate the true model. Consequently, if there exist more than one change point in the process we have some chance to estimate the true model which will be helpful to determine and remove the root causes introduced into the process. This method is more flexible than the case we assumed that there is just one change point in the process.

Keywords. Bayesian estimation; gamma process; multiple change points; RJMCMC; X-bar chart.

* Corresponding author

1 Introduction

Nowadays, the statistical process control (SPC) charts are widely used in industries. Typically it is assumed that, the quality characteristic of a process would follow a certain probability distribution. One can derive the upper control limit (UCL) and the lower control limit (LCL) in light of the specified type I error. When the sample observations are located within the control limits, we may conclude that the process is in a state of statistical control. On the the other hand, an out-of-control could occur when the sample observations fall outside the control limits. The major reason for the occurrence of out-of-control is that the disturbances have been introduced into the process. These disturbances change the probability distribution of the quality characteristic of the process, and this is detected by the SPC charts after some periods of time. There is a difference between the change point and the out-of-control signal time which is triggered by the SPC charts. The starting point of a process disturbance is called, change point. If we can accurately estimate the change point, we should have a good chance to easily and quickly determine the root causes of the problem and bring the process back to the in-control state.

SPC applications usually assume that the observations are normally distributed. If the sample size is moderate, according to the central limit theorem, this assumption is not far from reality. However, due to the shortage of sampling personnel, the cost of sample products, there are certain applications in which it may not be possible to have such moderate sample size. The semiconductor industry, for example, is a familiar case with nonnormal and skewed distributions for impurities and particle counts (Levinson, 1997; Levinson and Polny, 1999). In their papers, Samuel et al. (1998a, 1998b) addressed the issue of estimating the change point of a normal process. Pignatiello and Samuel (2001) used the EWMA and CUMSUM charts and the MLE to estimate the change point of a process. Shao and Hou (2004) provided some statistical properties for the change point estimators. In addition, Shao and Hou (2006) derived the change point estimators under the case where the X-bar chart and the MLE are used in a gamma process. Different from the Shao and Hou's (2006) concern, in this paper we assume after the change point occurrence, the parameter λ may change several times before the out-of-control signal is triggered by the SPC chart. In addition, we use the Bayesian approach to estimate the parameters. The multiple change point occurs when several disturbances are introduced into the process at different times. Knowing the number of the change points and estimating them is helpful to determine the root causes of the disturbances.

Below, we present the model in section 2. Section 3 is devoted to estimation of the parameters. In section 4, we illustrate our proposed method by various

examples. Section 5 concludes the paper.

2 The Model

It is assumed that we have a gamma process which is initially in control and the sample observations come from a gamma distribution with a pair of known parameters α_0 and λ_0. After an unknown time period τ_1, the process parameter λ_0 changes to an unknown value λ_1 and remains at this new level until the unknown time τ_2, again at this time the parameter λ_1 changes to λ_2 and remains at the new level λ_2 until the unknown time τ_3. This process is repeated k times where k is also an unknown parameter. After time τ_k the parameter λ_k remains unchanged until time T, the trigger time of the SPC chart. Thus, we have k change points $\tau_1, \tau_2, \ldots, \tau_k$. Among these change points, τ_1 may be more important, because it is the first time some disturbance is introduced into the process. In many cases it may be the main reason for the process going out of control. Let X_{ij} denote the jth observation in subgroup i with gamma distribution $G(\cdot,\cdot)$. That is,

$$
\begin{aligned}
X_{ij} &\overset{iid}{\sim} G(\alpha_0, \lambda_0), && i = 1, 2, \ldots, \tau_1 - 1, \quad j = 1, 2, \ldots, n_i \\
X_{ij} &\overset{iid}{\sim} G(\alpha_0, \lambda_1), && i = \tau_1, \tau_1 + 1, \ldots, \tau_2 - 1, \quad j = 1, 2, \ldots, n_i \\
X_{ij} &\overset{iid}{\sim} G(\alpha_0, \lambda_2), && i = \tau_2, \tau_2 + 1, \ldots, \tau_3 - 1, \quad j = 1, 2, \ldots, n_i \\
&\vdots \\
X_{ij} &\overset{iid}{\sim} G(\alpha_0, \lambda_k), && i = \tau_k, \tau_k + 1, \ldots, T, \quad j = 1, 2, \ldots, n_i
\end{aligned}
$$

where n_i are the sample sizes, T is the signal time when the subgroup sample mean $\bar{X}$ crosses one of the $\bar{X}$ control chart's limits. For convenience, one can set $n_i = n$ for $i = 1, \ldots, T$. For fixed T and $\mathbf{x_i} = (x_{i1}, \ldots, x_{in})$, $i = 1, \ldots, T$, the joint distribution of the observations is

$$
\begin{aligned}
f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T|\theta) &= \prod_{i=1}^{\tau_1 - 1} \prod_{j=1}^{n} f(x_{ij}) \prod_{i=\tau_1}^{\tau_2 - 1} \prod_{j=1}^{n} f(x_{ij}) \cdots \prod_{i=\tau_k}^{T} \prod_{j=1}^{n} f(x_{ij}) \\
&= \prod_{i=1}^{\tau_1 - 1} \prod_{j=1}^{n} \frac{\lambda_0^{\alpha_0}}{\Gamma(\alpha_0)} x_{ij}^{\alpha_0 - 1} e^{-\lambda_0 x_{ij}} \prod_{i=\tau_1}^{\tau_2 - 1} \prod_{j=1}^{n} \frac{\lambda_1^{\alpha_0}}{\Gamma(\alpha_0)} x_{ij}^{\alpha_0 - 1} e^{-\lambda_1 x_{ij}}
\end{aligned}
$$

$$\times \cdots \times \prod_{i=\tau_k}^{T} \prod_{i=1}^{n} \frac{\lambda_k^{\alpha_0}}{\Gamma(\alpha_0)} x_{ij}^{\alpha_0-1} e^{-\lambda_k x_{ij}}$$

$$= \frac{\prod_{i=1}^{T} (\prod_{j=1}^{n} x_{ij})^{\alpha_0-1}}{(\Gamma(\alpha_0))^{nT}} \lambda_0^{\alpha_0(\tau_1-1)} \lambda_1^{\alpha_0(\tau_2-\tau_1)} \cdots \lambda_k^{\alpha_0(T-\tau_k+1)}$$

$$\times e^{-\lambda_0 \sum_{i=1}^{\tau_1-1} \sum_{j=1}^{n} x_{ij} - \lambda_1 \sum_{i=\tau_1}^{\tau_2-1} \sum_{j=1}^{n} x_{ij} - \cdots - \lambda_k \sum_{i=\tau_k}^{T} \sum_{j=1}^{n} x_{ij}},$$

where $\theta = (k, \tau_1, \tau_2, \ldots, \tau_k, \lambda_1, \lambda_2, \ldots, \lambda_k)$. By factorization theorem, $(Y_1, Y_2, \ldots, Y_T)$ is a sufficient statistic for θ, where $Y_r = \sum_{j=1}^{n} X_{rj}$, $r = 1, 2, \ldots, T$. Thus one can make inference about θ, using the joint distribution of $(Y_1, Y_2, \ldots, Y_T)$. By reproductive property of gamma distribution,

$$f(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T|\theta) = \prod_{i=1}^{\tau_1-1} \frac{\lambda_0^{n\alpha_0}}{\Gamma(n\alpha_0)} y_i^{n\alpha_0-1} e^{-\lambda_0 y_i} \prod_{i=\tau_1}^{\tau_2-1} \frac{\lambda_1^{n\alpha_0}}{\gamma(n\alpha_0)} y_i^{n\alpha_0-1} e^{-\lambda_1 y_i}$$

$$\times \cdots \times \prod_{i=\tau_k}^{T} \frac{\lambda_k^{n\alpha_0}}{\Gamma(n\alpha_0)} y_i^{n\alpha_0-1} e^{-\lambda_k y_i}$$

$$= \frac{\prod_{i=1}^{T} y_i^{n\alpha_0-1}}{\{\Gamma(n\alpha_0)\}^T} \lambda_0^{n\alpha_0(\tau_1-1)} \lambda_1^{n\alpha_0(\tau_2-\tau_1)} \cdots \lambda_k^{n\alpha_0(T-\tau_k+1)}$$

$$\times e^{-\lambda_0 \sum_{i=1}^{\tau_1-1} y_i - \lambda_1 \sum_{i=\tau_1}^{\tau_2-1} y_i - \cdots - \lambda_k \sum_{i=\tau_k}^{T} y_i}.$$

Below, this distribution is considered as the likelihood function of the parameters. Assuming some priors, it will be used to derive the posterior distribution of the parameters.

If observations X_{ij}, $i = 1, \ldots, T$, $j = 1, \ldots, n$ were independent and under control, then all of them would have the same gamma distribution, $G(\alpha_0, \lambda_0)$. It can be derived that the subgroup sample mean $\bar{X}_i$, $(i = 1, \ldots, T)$ has a gamma distribution $G(n\alpha_0, n\lambda_0)$. The upper $(100 \times \alpha)$th percentile of a gamma distribution $G(n\alpha_0, n\lambda_0)$ can be easily obtained and expressed as $\chi^2_\alpha(2n\alpha_0)/(2n\lambda_0)$, where $\chi^2_\alpha(d)$ denotes the upper $(100 \times \alpha)$th percentile of a chi-square distribution with d degrees of freedom. If the process is in-control, then $100(1-\alpha)$ percent of the $\bar{X}$ values would fall between $\chi^2_{1-\alpha}(2n\alpha_0)/(2n\lambda_0)$ and $\chi^2_\alpha(2n\alpha_0)/(2n\lambda_0)$.

Obviously, almost all of the X values would fall between $\chi^2_{0.99865}(2n\alpha_0)/(2n\lambda_0)$ and $\chi^2_{0.000135}(2n\alpha_0)/(2n\lambda_0)$, unless the system had actually changed. Accordingly, if α_0 and λ_0 are known, we can use $\chi^2_{0.99865}(2n\alpha_0)/(2n\lambda_0)$ and $\chi^2_{0.00135}(2n\alpha_0)/(2n\lambda_0)$ as the lower and upper control limits.

Now, we assume that there have been some change points which were indicated by the trigger time of SPC chart and we intend to estimate their number and locations via a Bayesian procedure.

3 Bayesian Estimation of the Parameters

To estimate the parameters using Bayesian approach, first we should determine the prior distributions for these parameters. In choosing the prior distributions for the parameters, we should consider their nature. We are dealing with a positive discrete variable k which cannot be too large. First we assume that the maximum number of change points is a known integer value $k_{\max}$, and k is distributed uniformly on $\{1, 2, \ldots, k_{\max}\}$. In absence of any explicit information about k, we can use,

$$\pi_1(k) = \frac{1}{k_{\max}}, \qquad k = 1, 2, \ldots, k_{\max}$$

Given k, we assume that the parameters λ_i, $i = 1, 2, \ldots, k$ are independent and distributed as $G(\alpha^\star, \beta)$ since it is a positive quantity and its smaller values are more likely than its larger values. Therefore,

$$\pi_2(\lambda_i|k) = \frac{\beta^{\alpha^\star}}{\Gamma(\alpha^\star)}\lambda_i^{\alpha^\star - 1}e^{-\beta\lambda_i}, \qquad i = 1, 2, \ldots, k$$

seems a sensible choice. Finally, due to lack of sufficient knowledge on change points, we may only imagine that $\tau_1, \tau_2, \ldots, \tau_k$ are the order statistics of k values drawn without replacement from the set $\{2, 3, \ldots, T\}$. Then we can set

$$\pi_3(\tau_1, \tau_2, \ldots, \tau_k|k) = \frac{1}{\binom{T-1}{k}}.$$

Then the joint posterior of the parameters is

$$\pi(k, \tau_1, \ldots, \tau_k, \lambda_1, \ldots, \lambda_k|\mathbf{y}) = \frac{f(\mathbf{y}|\boldsymbol{\theta})\pi_3(\boldsymbol{\theta}_1|k)\pi(\boldsymbol{\theta}_2|k)\pi_1(k)}{\sum_{k=1}^{k_{\max}}\sum_{\tau_1}\cdots\sum_{\tau_k}\int_0^\infty\cdots\int_0^\infty f(\mathbf{y}|\boldsymbol{\theta})\pi_3(\boldsymbol{\theta}_1|k)\pi(\boldsymbol{\theta}_2|k)\pi_1(k)d\lambda_1\cdots d\lambda_k}$$

and the marginal posterior of $(k, \tau_1, \ldots, \tau_k)$ is given as

$$\pi(k, \tau_1, \ldots, \tau_k|\mathbf{y}) = \frac{\int_0^\infty\cdots\int_0^\infty f(\mathbf{y}|\boldsymbol{\theta})\pi_3(\boldsymbol{\theta}_1|k)\pi(\boldsymbol{\theta}_2|k)\pi_1(k)d\lambda_1\cdots d\lambda_k}{\sum_{k=1}^{k_{\max}}\sum_{\tau_1}\cdots\sum_{\tau_k}\int_0^\infty\cdots\int_0^\infty f(\mathbf{y}|\boldsymbol{\theta})\pi_3(\boldsymbol{\theta}_1|k)\pi(\boldsymbol{\theta}_2|k)\pi_1(k)d\lambda_1\cdots d\lambda_k},$$

where $\boldsymbol{\theta} = (k, \tau_1, \ldots, \tau_k, \lambda_1, \ldots, \lambda_k)$, $\boldsymbol{\theta}_1 = (\tau_1, \ldots, \tau_k)$, $\boldsymbol{\theta}_2 = (\lambda_1, \ldots, \lambda_k)$, and $\pi(\boldsymbol{\theta}_2|k) = \prod_{i=1}^k \pi_2(\lambda_i|k)$. The summations on τ_i's are taken for all $(\tau_1, \ldots, \tau_k)$ in the sets

$$A_k = \{(\tau_1, \ldots, \tau_k)|\ 2 \leqslant \tau_1 < \tau_2 < \cdots < \tau_k \leqslant T,\ \tau_i \in N,\ i = 1, \ldots, k\}.$$

Regarding the unknown dimension of the parameters when $k_{\max}$ is large, computation of the posterior distributions of the parameters is difficult. Therefore we use the reversible jump Markov chain Monte Carlo (RJMCMC)(see Green(1995) or Givens and Hoeting (2005) for details) to compute the Bayesian estimates of the parameters. RJMCMC is an extension of Metropolis-Hastings method, and applies to various varying-dimension problems.

To apply this method, one starts using an initial value for the parameters. If the value of the jth random draw is $x_j = (k, \theta^{(k)})$, a proposal value for x_{j+1} is $x^\star = (k', \theta^{(k')})$. The relation between $\theta^{(k)}$ and $\theta^{(k')}$ is as follows: If the dimension of $\theta^{(k)}$ and $\theta^{(k')}$ are not equal, the random vectors $u^{(1)}$ and $u^{(2)}$ are generated such that the dimension of $(\theta^{(k')}, u^{(2)})$ and $(\theta^{(k)}, u^{(1)})$ would be equal. Then we set $\theta^{(k')}$ to be some deterministic function of $\theta^{(k)}$ and $u^{(1)}$, and $\theta^{(k)}$ to be some function of $\theta^{(k')}$ and $u^{(2)}$.

Now x_{j+1} is set to be $x^\star$ with probability r and it is equal to x_j with probability $1 - r$, where

$$r = \min\left\{1, \frac{p(k', \theta^{(k')}|y)j(k', \theta^{(k')})q_2(u^{(2)})}{p(k, \theta^{(k)}|y)j(k, \theta^{(k)})q_1(u^{(1)})} \cdot \left|\frac{\partial(\theta^{(k')}, u^{(2)})}{\partial(\theta^{(k)}, u^{(1)})}\right|\right\},$$

in which $j(x)$ is the probability of choosing this move, $q_1(\cdot)$ and $q_2(\cdot)$ are the pdf's of $u^{(1)}$ and $u^{(2)}$ respectively, and the last term is the Jacobian.

In practice, we do not have to generate both $u^{(1)}$ and $u^{(2)}$. For example, if the dimension of $\theta^{(k')}$ is m and that of $\theta^{(k)}$ is n, such that $m > n$ then it is enough to generate random vector $u^{(1)}$ with dimension n_2 where $m = n + n_2$. In this case $\theta^{(k')}$ will be some function of $(\theta^{(k)}, u^{(1)})$, and the acceptance probability becomes

$$r = \min\left\{1, \frac{p(k', \theta^{(k')}|y)j(k', \theta^{(k')})}{p(k, \theta^{(k)}|y)j(k, \theta^{(k)})q_1(u^{(1)})} \cdot \left|\frac{\partial(\theta^{(k')})}{\partial(\theta^{(k)}, u^{(1)})}\right|\right\}. \tag{1}$$

3.1 Using RJMCMC for Change Points

In developing a RJMCMC sampler for the change point problem, we are guided by intuition in designing appropriate moves, coupled with the requirements that the dimensions can be balanced properly, that the moves can be simulated conveniently, and that the acceptance ratio can be computed economically.

In using this method, some possible transitions are (a) change to the λ of a randomly chosen change point, (b) a change to the position of a randomly chosen change point, (c) *birth* of a new change point at a randomly chosen point from the set $\{2, 3, \ldots, T\} - \{\tau_1, \tau_2, \ldots, \tau_k\}$, and (d) *death* of a randomly chosen change point. We denote these transitions with $(L, P, k+1, k-1)$. Here

L means a λ change, P a position change, $k-1$ denotes a death of a change point that reduces the number of change points from k to $k-1$, and $k+1$ denotes a birth of a change point that increases the number of change points from k to $k+1$.

Now, we describe these transitions in more details. At each transition, an independent random choice is made between attempting each of the at most four available move types. We have probability η_k for L, π_k for P, b_k for $k+1$, and d_k for $k-1$, depending only on the current number of change points k, and satisfying $\eta_k + \pi_k + b_k + d_k = 1$. Naturally, $d_1 = 0$ and $b_{k_{\max}} = 0$. If a move of type L or P is chosen, the remaining details are straightforward. A change to a λ_j is attempted by first choosing one of $\lambda_1, \lambda_2, \ldots, \lambda_k$ at random, obtaining λ_j say, then proposing a change to λ'_l such that $u = \log(\lambda'_j/\lambda_j)$ is uniformly distributed on the interval $(-\frac{1}{2}, \frac{1}{2})$. This choice is made for convenience, so that the proposal density ratio takes a simple form. Then we have

$$\frac{\partial \lambda'_j}{\partial \lambda_j} = \exp(u) = \frac{\lambda'_j}{\lambda_j},$$

and the acceptance probability for this move is found to be

$$\min\left\{1, (\text{likelihood ratio}) \times \left(\frac{\lambda'_j}{\lambda_j}\right)^{\alpha^\star} e^{-\beta(\lambda'_j - \lambda_j)}\right\}$$

in the usual way. Hereafter, the *likelihood ratio* means $p(y|x^\star)/p(y|x)$, where x and $x^\star$ stand for the current and proposed new values of all parameters.

For a position change move, one of $\tau_1, \tau_2, \ldots, \tau_k$ is drawn at random, obtaining say τ_j. The proposed replacement value is τ'_j, drawn uniformly from $\{\tau | \tau \in N,\ \tau_{j-1} < \tau < \tau_{j+1}\}$, and the acceptance probability turns out to be

$$\min\{1, (\text{likelihood ratio})\}.$$

The details for a birth of a change point are more complicated, as follows. We first choose an integer value $\tau^\star$ for the proposed new change point, uniformly distributed on $\{\tau_1, \tau_1 + 1, \ldots, T\} - \{\tau_1, \tau_2, \ldots, \tau_k\}$. The reason why we choose $\tau^\star$ greater than τ_1 is that τ_1 is the first change point, and there is not any change point before τ_1. The parameter $\tau^\star$ must lie, with probability 1 within an existing interval (τ_j, τ_{j+1}), say. If accepted, τ'_{j+1} will be set to $\tau^\star$, and $\tau_{j+1}, \tau_{j+2}, \ldots, \tau_k$ will be relabeled as $\tau'_{j+2}, \tau'_{j+3}, \ldots, \tau'_{k+1}$ with corresponding changes to the labeling of the corresponding λ's. We wish to propose new λ's, λ'_j, and λ'_{j+1} for the distribution of the observations on the subintervals $[\tau_j, \tau^\star)$ and $[\tau^\star, \tau_{j+1})$ which recognize that the current λ on the union of these two intervals is typically well-supported in the posterior distribution, and should

therefore not be completely discarded. Thus the new λ's, λ'_j, and λ'_{j+1} should be perturbed in either direction from λ_j in such a way that λ_j is a compromise between them. To preserve positivity and maintain simplicity in the acceptance ratio calculations, we use a weighted geometric mean for this compromise, so that $(\tau^\star - \tau_j)\log(\lambda'_j) + (\tau_{j+1} - \tau^\star)\log(\lambda'_{j+1}) = (\tau_{j+1} - \tau_j)\log(\lambda_j)$ and define the perturbation to be such that

$$\frac{\lambda'_{j+1}}{\lambda'_j} = \frac{1-u}{u}, \tag{2}$$

with u drawn uniformly from $[0,1]$. Details for a death are as follows. The τ_{j+1} that is proposed for removal is simply drawn at random from $\tau_1, \tau_2, \ldots, \tau_k$. If τ_{j+1} is removed, the new λ over the interval $(\tau'_j, \tau'_{j+1}) = (\tau_j, \tau_{j+2})$ is λ'_j, the weighted geometric mean satisfying

$$(\tau_{j+1} - \tau_j)\log(\lambda_j) + (\tau_{j+2} - \tau_{j+1})\log(\lambda_{j+1}) = (\tau'_{j+1} - \tau'_j)\log(\lambda'_j).$$

The pair of birth and death moves thus defined satisfy the dimension-matching requirement. The birth increases the dimensionality from $2k$ to $2k+2$, the difference is 2 which is the number of variables we need to separate λ'_j and λ'_{j+1} namely $\tau^\star$ and u. In deriving an expression for the acceptance probability of the birth proposal, it is helpful to re-write (1) in the form

$$\min\{1, (\text{likelihood ratio}) \times (\text{prior ratio}) \times (\text{proposal ratio}) \times (\text{Jacobian})\},$$

noting that $p(x|y) = p(y|x)p(x)/p(y)$. Then the prior ratio which previously was $p(k', \theta^{(k')})/p(k, \theta^{(k)})$ becomes

$$\frac{\binom{T-1}{k}}{\binom{T-1}{k+1}} \cdot \frac{\beta^{\alpha^\star}}{\Gamma(\alpha^\star)} \left(\frac{\lambda'_j \lambda'_{j+1}}{\lambda_j}\right)^{\alpha^\star - 1} e^{-\beta(\lambda'_j + \lambda'_{j+1} - \lambda_j)}$$

and the proposal ratio which was $j(k', \theta^{(k')})/j(k, \theta^{(k)})q_1(u^{(1)})$ becomes

$$\frac{d_{k+1}(T - \tau_1 - k + 1)}{b_k k}.$$

Now, we compute the Jacobian. We have $\frac{\lambda'_{j+1}}{\lambda'_j} = \frac{1-u}{u} = \frac{1}{u} - 1$, then

$$(\tau^\star - \tau_j)\log(\lambda'_j) + (\tau_{j+1} - \tau^\star)\log\left\{\left(\frac{1}{u} - 1\right)\lambda'_j\right\} = (\tau_{j+1} - \tau_j)\log(\lambda_j). \tag{3}$$

Taking derivative with respect to u from both sides, we have

$$(\tau^\star - \tau_j)\frac{\frac{\partial \lambda'_j}{u}}{\lambda'_j} + (\tau_{j+1} - \tau^\star)\left(\frac{\frac{-1}{u^2}}{\frac{1-u}{u}} + \frac{\frac{\partial \lambda'_j}{\partial u}}{\lambda'_j}\right) = 0.$$

Then we have

$$(\tau_{j+1}-\tau_j)\frac{\frac{\partial\lambda'_j}{\partial u}}{\lambda'_j}=\frac{\tau_{j+1}-\tau^\star}{u(1-u)}.$$

Thus,

$$\frac{\partial\lambda'_j}{\partial u}=\frac{(\tau_{j+1}-\tau^\star)\lambda'_j}{(\tau_{j+1}-\tau_j)u(1-u)}.$$

Using (2) we have

$$\begin{aligned}\frac{\partial}{\partial u}(\lambda'_{j+1}) &= \frac{\partial}{\partial u}\left(\frac{1-u}{u}\lambda'_j\right)=-\frac{1}{u^2}\lambda'_j+\frac{1-u}{u}\left(\frac{\tau_{j+1}-\tau^\star}{\tau_{j+1}-\tau_j}\right)\frac{\lambda'_j}{u(1-u)}\\ &= \frac{(\tau_j-\tau^\star)}{u^2(\lambda_{j+1}-\lambda_j)}\lambda'_j.\end{aligned}$$

Taking derivative with respect to λ_j from both sides of (3) we have

$$(\tau^\star-\tau_j)\frac{\frac{\partial\lambda'_j}{\partial\lambda_j}}{\lambda'_j}+(\tau_{j+1}-\tau^\star)\frac{\frac{\partial\lambda'_j}{\partial\lambda_j}}{\lambda'_j}=\frac{\tau_{j+1}-\tau_j}{\lambda_j}.$$

Then

$$(\tau_{j+1}-\tau_j)\frac{\frac{\partial\lambda'_j}{\partial\lambda_j}}{\lambda'_j}=\frac{\tau_{j+1}-\tau_j}{\lambda_j}.$$

That is,

$$\frac{\partial\lambda'_j}{\partial\lambda_j}=\frac{\lambda'_j}{\lambda_j}.$$

Using (2) we have

$$\frac{\partial\lambda'_{j+1}}{\partial\lambda_j}=\frac{\partial}{\partial\lambda_j}\left(\frac{1-u}{u}\lambda'_j\right)=\left(\frac{1-u}{u}\right)\frac{\lambda'_j}{\lambda_j}.$$

Thus we have

$$J=\begin{vmatrix}\frac{\partial\lambda'_j}{\partial\lambda_j} & \frac{\partial\lambda'_{j+1}}{\partial\lambda_j}\\ \frac{\partial\lambda'_j}{\partial u} & \frac{\partial\lambda'_{j+1}}{\partial u}\end{vmatrix}=-\left(\frac{\lambda'^2_j}{\lambda_j}\right)\frac{1}{u^2}=-\frac{(\lambda'_j+\lambda'_{j+1})^2}{\lambda_j}.$$

The last equality is obtained using the fact that

$$\frac{1-u}{u}=\frac{\lambda'_{j+1}}{\lambda'_j}.$$

Then, we have

$$\frac{1}{u^2} = \frac{(\lambda'_j + \lambda'_{j+1})^2}{\lambda'^2_j}.$$

At last we have

$$|J| = \frac{(\lambda'_j + \lambda'_{j+1})^2}{\lambda_j}.$$

The acceptance probability for the corresponding death step has the same form with the appropriate changes of labeling of the variables, and the ratio terms inverted.

3.2 Estimating the Hyperparameters $\alpha^\star$ and β

Here we use the emperical Bayes method to estimate the hyperparameters $\alpha^\star$ and β. To do so, first we should compute the likelihood function $L(\alpha^\star, \beta|y_1, y_2, \ldots, y_T)$. We have

$$L(\alpha^\star, \beta|y_1, y_2, \ldots, y_T) = \sum_{k=1}^{k_{\max}} \sum_{2\leqslant\tau_1<\tau_2<\cdots<\tau_k\leqslant T-1} \sum \cdots \sum \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(y_1, \ldots, y_T|\theta) \times \pi_1(k) \left(\prod_{i=1}^{k} \pi_2(\lambda_i|k) \right) \pi_3(\tau_1, \ldots, \tau_k|k) d\lambda_1 \cdots d\lambda_k,$$

where $\theta = (k, \tau_1, \tau_2, \ldots, \tau_k, \lambda_1, \lambda_2, \ldots, \lambda_k)$.

It is noticed that computation of the above expression is difficult, but can be approximated by Monte Carlo integration method. The values of $\alpha^\star$ and β which maximize $L(\alpha^\star, \beta|y_1, \ldots, y_T)$ are the estimates of $\alpha^\star$ and β, which can be found numerically.

4 Examples

Suppose that a gamma process is monitored by an $\bar{X}$ chart. When a signal is triggered by the $\bar{X}$ chart, the RJMCMC is then applied to data to obtain Bayes estimates of the number of change points, k, and their corresponding τ's. To do this, first we assume that $\lambda_i \sim G(\alpha^\star, \beta)$, $i = 1, 2, \ldots, k$, and the method mentioned in subsection 3.2 is used to estimate $\alpha^\star$ and β. In addition we assume that $k_{\max} = 10$, the probabilities of changes in the parameter values are $b_i = d_i = 0.45$ for $i = 2, \ldots, 9$, $b_{10} = d_1 = 0$, $d_{10} = b_1 = 0.9$, and $\pi_i = \eta_i = 0.05$ for $i = 1, \ldots, 10$. For this setting, we have run the Monte Carlo simulation for 40000 updates. We use the posterior mode as an estimate of k, while the

nearest integers to posterior means are used to estimate $\tau_1, \tau_2 \ldots, \tau_k$. To evaluate the effectiveness of our proposed approach for estimating the number of change points and their values, a series of simulations were conducted. Sample observations were produced randomly in subgroups of size 10 from a gamma distribution with parameters α and λ. In the following simulation examples, 100 runs were performed for each example, then using the proposed method, the average of the estimates of $k, \tau_1, \tau_2 \ldots, \tau_k$ for those 100 runs were recorded with their standard errors for each estimates of k. The results for Example 1 are given in Tables 1 and 2. For Example 2, they are given in Tables 3 and 4, and for the last example in Tables 5 and 6. The examples are as follows.

Example 1. Suppose for the first 99 subgroups, the observations followed $G(1,1)$; but starting from subgroup 100, sample observations followed $G(1,0.8)$.

In this example, the parameters are, $k = 1$, $\lambda_0 = 1$, $\lambda_1 = 0.8$, and $\tau_1 = 100$. According to Table 1, we see that in 95% of the times our method estimats the true model ($k = 1$). The estimates of τ for each estimated value of k are given in Table 2.

Example 2. Suppose for the first 49 subgroups, the observations followed $G(1,1)$; but starting from subgroup 50 to 99, the observations followed $G(1,1.2)$. In addition, we assume that all these observations are located within the control limits. Finally, suppose starting from subgroup 100, sample observations followed $G(1,0.8)$.

In Example 2, the parameters are $k = 2$, $\lambda_0 = 1$, $\lambda_1 = 1.2$, $\lambda_2 = 0.8$, $\tau_1 = 50$, $\tau_2 = 100$. According to Table 3, we see that in 75% of the times our method estimate the true model. The estimates of τ for each estimated value of k are given in Table 4.

Example 3. Suppose for the first 69 subgroups, the observations followed $G(1,1)$; but starting from subgroup 70 to 129 the observations followed $G(1,0.8)$, while starting from subgroup 130 to 179 the observations followed $G(1,1.1)$. In addition, we assume that all these observations are located within control limits. At last, starting from observation 180 the observations followed $G(1,0.9)$.

In Example 3, the parameters are $k = 3$, $\lambda_0 = 1$, $\lambda_1 = 0.8$, $\lambda_2 = 1.1$, $\lambda_3 = 0.9$, $\tau_1 = 70$, $\tau_2 = 130$, $\tau_3 = 180$. According to Table 5, we see that in 48% of the times our method estimates the true model. The estimates of τ for each estimated value of k are given in Table 6.

Table 1. The estimate of posterior probability distribution of k under the conditions mentioned in Example 1.

k	1	2
$\pi(k\|\mathbf{x})$	0.95	0.05

Table 2. The estimates of k and average change point estimates with their standard errors for each estimated k under the conditions mentioned in Example 1.

$\hat{k}=1$	$\hat{k}=2$	
$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$
99.11	66.45	98.93
(12.32)	(30.62)	(22.62)

Table 3. The estimate of posterior probability distribution of k under the conditions mentioned in Example 2.

k	1	2	3
$\pi(k\|\mathbf{x})$	0.11	0.75	0.14

Table 4. The estimates of k and average change point estimates with their standard for each estimated k errors under the conditions mentioned in Example 2.

$\hat{k}=1$	$\hat{k}=2$		$\hat{k}=3$		
$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\tau}}_3$
103.17	49.93	100.58	54.27	93.75	157.33
(7.08)	(10.97)	(7.77)	(17.07)	(6.57)	(69.78)

Table 5. The estimate of posterior probability distribution of k under the conditions mentioned in Example 3.

k	1	2	3	4
$\pi(k\|\mathbf{x})$	0.35	0.04	0.48	0.13

Table 6. The estimates of k and average change point estimates with their standard errors for each estimated k under the conditions mentioned in Example 3.

$\hat{k}=1$	$\hat{k}=2$		$\hat{k}=3$			$\hat{k}=4$			
$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\tau}}_3$	$\bar{\hat{\tau}}_1$	$\bar{\hat{\tau}}_2$	$\bar{\hat{\tau}}_3$	$\bar{\hat{\tau}}_4$
118.73	94.47	157.72	75.03	129.19	198.7	72.96	123.05	170.59	306.85
(56.2)	(29.09)	(28.75)	(12.9)	(6.4)	(34.1)	(8.99)	(4.72)	(10.45)	(119.16)

5 Conclusion

To identify the change point of a gamma process, monitored by $\bar{X}$ control chart, there is no reason to constrain ourselves to one change point. In fact, some disturbances may be introduced into the process in different times before the $\bar{X}$ chart triggers a signal. Knowing the number of times when the disturbances were introduced into the process, and estimating those times will be helpful to determine the root causes of the disturbances and for removing them. This study proposes a method to estimate the the number of change points and their values in a gamma process monitored by $\bar{X}$ control chart, using RJMCMC algorithm. Application of the proposed method and the efficiency of its performance are illustrated through a series of simulations. In this paper, we used the $\bar{X}$ control chart to monitor a gamma process. There are other types of control charts (for example, S, EWMA, and Cusum control charts) to investigate. Those who are interested in this issue may apply the same concept to study the other control charts.

References

Givens, G.H. and Hoeting, J.A. (2005). *Computational Statistics*, Wiley, New York.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711-732.

Levinson, W. (1997). Watch out for nonnormal distributions of impurities. *Chemical Engineering Progress* **93**, 70-76.

Levinson, W. and Polny, A. (1999). SPC for tool particle counts. *Semiconductor International*, June, 117-122.

Pignatiello, J.J. and Samuel, T.R. (2001). Estimation of the change point of a normal process mean in SPC applications. *Journal of Quality Technology* **33**, 82-95.

Samuel, T.R., Pignatiello, J.J., and Calvin, J.A. (1998a). Identifying the time of a step change with $\bar{X}$ control charts *Quality Engineering* **10**, 521-527.

Samuel, T.R., Pignatiello, J.J., and Calvin, J.A. (1998b). Identifying the time of a step change in a normal process variance. *Quality Engineering* **10**, 529-538.

Shao, Y.E. and Hou, C.D. (2004). A study on identifying the time of a step change disturbance with S control charts and MLE method. *Journal of the Chinese Institute of Industrial Engineers* **21**, 349-357.

Shao, Y.E. and Hou, C.D. (2006). Estimation of the starting time of a step change disturbance in a gamma process. *Journal of the Chinese Institute of Industrial Engineers* **23**, 319-327.

M. Esmail Dehghan Monfared
Department of Statistics,
Faculty of Mathematical Sciences,
Shahid Beheshti University,
Tehran, Iran.
e-mail: *monfared2@gmail.com*

Mohammad Reza Meshkani
Department of Statistics,
Faculty of Mathematical Sciences,
Shahid Beheshti University,
Tehran, Iran.
e-mail: *mrmeshkani@gmail.com*