

## A Bayesian Approach to a Process Control Problem

Özlem Sunar<sup>1</sup>, Gül Ergün<sup>2</sup>

<sup>1</sup>General Directorate of Mineral Research and Exploration, 06520 Çukurambar, Ankara, Türkiye

e-mail: ozlemsunar@mta.gov.tr

<sup>2</sup>Hacettepe University, Faculty of Science, Department of Statistics, 06800 Beytepe, Ankara, Türkiye

e-mail: gul@hacettepe.edu.tr

Received Date: January 5, 2010

Accepted Date: June 18, 2010

**Abstract.** Classical process control methods are not suitable when the number of observations is low. In this study, a Bayesian scheme is considered for a statistical process control problem where process mean is unknown and has jumps in both sides for normal observations. The aim of the study is to detect the changes in the mean for a short-run process. A random walk is defined to indicate the changes of the process mean in time. The posterior distribution of the process mean at time  $n$  is found to be a mixture of normal distributions.

**Key words:** Bayesian process control; random walk model; mixture models.  
*2000 Mathematics Subject Classification.* 62C10, 62F15, 62N05.

### 1. Introduction

During the World War II, the importance and usage of statistical quality control definitions becomes widespread, and it is understood how necessary the statistical techniques are for controlling and developing the product quality. Generally, statistical process control aims to investigate and prevent quality problems during the process without any delay that appeared due to unnatural causes and to provide maximum productivity (Montgomery, 2001). Control cards, first presented by Shewhart in 1931, are used to control the process statistically and to verify developments in the process. Usually, the assumptions are not hold for real life problems. One of the important problems in applications is to have unknown process parameters that have shifts. Statistical process control is interested in making decisions about the process parameters. As it is known that the main purpose in modeling is to predict the unknown parameters by using

existing information. Classical process control methods cannot be applied to workshop processes when the number of the data is few. Thus, Bayesian approach in statistical process control gives conceptually more effective results for short-run productions. A Bayesian framework can be an alternative solution to overcome the problems for that case. When the process mean is unknown, its structure should be represented by a system equation as in state-space model.

Then Kalman filter (1960) can be applied for estimating recursively the unknown quantity at each time. However, in this study the unknown process mean for normal observations is assumed to have three possible changes. Therefore, Kalman filter is not used here. This paper is inspired by the studies of Tsiamyrtzis and Hawkins (2005, : 2007) on a Bayesian process control problem. They considered the problem for only positive changes in unknown process mean. This study considers not only positive jumps but the negative jumps as well. A random walk and normal mixture models are used to detect the change of the unknown process mean in time. The predictions of the process mean is obtained by combining the prior knowledge defined for at each  $n$  time point and likelihood function via Bayesian theorem, and they are presented step by step in the study. In the second section of the paper, suggested model and a related theorem are presented with its proof. The model is implemented on an artificial data set by a program designed in Matlab 7.1 and the results are presented in Section 3. The last section is dedicated to conclusion.

## 2. A Bayesian Model to Detect Two Sided Jumps in Process Mean

In the study, a random walk model which contains jumps in both descending and ascending directions is evaluated for the unknown process mean and this structure is defined totally by the use of normal mixture model.

When  $\theta_{n-1}$  is process mean at the  $(n-1)^{th}$  step, random walk model for  $\theta$  at the  $n^{th}$  step is defined as follows:

$$(1) \quad \theta_n/\theta_{n-1} \sim \begin{cases} N(\theta_{n-1}, \sigma^2) & , \quad p_1 \\ N(\theta_{n-1} + \delta_1, \sigma^2) & , \quad p_2 \\ N(\theta_{n-1} - \delta_2, \sigma^2) & , \quad p_3 \end{cases} \quad , \quad p_1 + p_2 + p_3 = 1.$$

It is an extended case of Tsiamyrtzis and Hawkins models published in 2005 and 2007. Beside the positive jump, a negative jump with size of  $\delta_2$  is included into the model. Here  $\sigma^2$  is variance definition and presents random drift of the mean. It is seen in Equation (1) that the process mean may have a positive shift with the probability “ $p_2$ ” and negative shift with the probability “ $p_3$ ”. According to the definition, the value of  $\theta_n$  at each time point is centered to  $\theta_{n-1}, : \theta_{n-1} + \delta_1, : \theta_{n-1} - \delta_2$ . Then, the conditional distribution of  $\theta_n/\theta_{n-1}$  is presented by the following normal mixture model.

$$(2) \quad f(\theta_n/\theta_{n-1}) \sim p_1 N(\theta_{n-1}, \sigma^2) + p_2 N(\theta_{n-1} + \delta_1, \sigma^2) + p_3 N(\theta_{n-1} - \delta_2, \sigma^2).$$

It is another definition of Equation (1). When the observations have normal distributions then the observation equation is defined as below:

$$(3) \quad f(x_n/\theta_n) \sim N(\theta_n, \tau^2).$$

Here  $\tau^2$  is the variance of the normal model. The process mean  $\theta$  is updated by obtaining the posterior distribution of  $\theta_n$  after observing  $x_n$  at time  $n$ . Here  $x_n$  can either be a single data or it presents a data set. The parameters  $\Omega = (\sigma, \tau, p_1, p_2, p_3, \delta_1, \delta_2)$  in the models are assumed as nuisance parameters to avoid the probable analytic difficulties in the prediction process. The each component of  $\Omega$  is assumed to be known in advance or predicted by using the previous data set.

Since the Bayesian model, evaluated for a statistical process control problem to detect the change in unknown mean for normal observations, contains no jump or jump in upper threshold or jump in lower threshold at time  $n$ , the posterior distribution after  $n$  successive times will have a normal mixture distribution with  $3^n$  components. When the nuisance parameters are omitted in the study, the prior distribution for  $\theta$  at the  $n^{th}$  time point is also a mixture of the normal distributions with  $3^n$  component. When  $x_n$  presents all the observations up to  $n^{th}$  time point, the posterior distribution of  $\theta_n/x_n$  is then obtained by combining likelihood function and prior distribution. In the study, the theorem and definition related to the prediction of unknown process mean  $\theta$  is presented below:

**Theorem 2.1.** Since the possible changes of unknown process mean at time  $n$  are defined by a random walk containing positive jumps, negative jumps and no jump, then the posterior distribution of  $\theta_n/x_n$  is found to have a normal mixture model with  $3^n$  component as below.

$$(4) \quad f(\theta_n/x_n) = \sum_{i=0}^{3^n-1} \alpha_i^{(n)} N(\theta_i^{(n)}, \hat{\sigma}_n^2).$$

The variance, posterior weights and means in Equation (4) are obtained as follows:

$$\begin{aligned} \hat{\sigma}_n^2 &= (1 - K_n) \tau^2 = K_n (\sigma^2 + \hat{\sigma}_{n-1}^2), \\ \alpha_{3i}^{(n)} &= \frac{p_1 \alpha_i^{(n-1)} f_{1i}(x_n)}{NC}, \text{ for } i = 0, 1, \dots, 3^{n-1} - 1, \\ \alpha_{3i+1}^{(n)} &= \frac{p_2 \alpha_i^{(n-1)} f_{2i}(x_n)}{NC}, \\ \alpha_{3i-1}^{(n)} &= \frac{p_3 \alpha_i^{(n-1)} f_{3i}(x_n)}{NC}, \\ \theta_{3i}^{(n)} &= K_n \theta_i^{(n-1)} + (1 - K_n) x_n, \end{aligned}$$

$$\begin{aligned}\theta_{3i+1}^{(n)} &= K_n \left( \theta_i^{(n-1)} + \delta_1 \right) + (1 - K_n) x_n, \\ \theta_{3i-1}^{(n)} &= K_n \left( \theta_i^{(n-1)} - \delta_2 \right) + (1 - K_n) x_n\end{aligned}$$

Here,

$$\begin{aligned}K_n &= \frac{\tau^2}{\tau^2 + \sigma^2 + \hat{\sigma}_{n-1}^2}, \\ f_{1i}(x_n) &= \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{(x_n - \theta_i^{(n-1)})^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\}, \\ f_{2i}(x_n) &= \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{[x_n - (\theta_i^{(n-1)} + \delta_1)]^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\}, \\ f_{3i}(x_n) &= \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{[x_n - (\theta_i^{(n-1)} - \delta_2)]^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\}\end{aligned}$$

and

$$NC = \sum_{i=0}^{3^n-1} \left[ p_1 \alpha_i^{(n-1)} f_{1i}(x_n) + p_2 \alpha_i^{(n-1)} f_{2i}(x_n) + p_3 \alpha_i^{(n-1)} f_{3i}(x_n) \right].$$

**Proof.** Like Tsiamyrtzis and Hawkins study, the inductive method is used here for the proof of the mixture model that is offered for the posterior distribution. It is assumed in the study that the theorem is hold for  $(n-1)^{th}$  stages of the process and the distribution of  $\theta_{n-1}/x_{n-1}$  is defined as follows:

$$(5) \quad f(\theta_{n-1}/x_{n-1}) = \sum_{i=0}^{3^{n-1}-1} \alpha_i^{(n-1)} N(\theta_i^{(n-1)}, \hat{\sigma}_{n-1}^2).$$

The distribution of  $\theta_n/\theta_{n-1}$  from Equation (2) and the posterior distribution of  $\theta_{n-1}/x_{n-1}$  from Equation (5) are both used in the following integral.

$$\begin{aligned}(6) \quad f(\theta_n/x_{n-1}) &= \int_{\theta} f(\theta_n/\theta_{n-1}) f(\theta_{n-1}/x_{n-1}) d\theta_{n-1} \\ &= \int_{\theta} \left[ \frac{p_1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta_n - \theta_{n-1})^2}{2\sigma^2}} + \frac{p_2}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta_n - \theta_{n-1} - \delta_1)^2}{2\sigma^2}} + \right. \\ &\quad \left. + \frac{p_3}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta_n - \theta_{n-1} + \delta_2)^2}{2\sigma^2}} \right] \times \sum_{i=0}^{3^{n-1}-1} \alpha_i^{(n-1)} \frac{1}{\sqrt{2\pi\hat{\sigma}_{n-1}^2}} e^{-\frac{(\theta_{n-1} - \theta_i^{(n-1)})^2}{2\hat{\sigma}_{n-1}^2}} d\theta_{n-1}\end{aligned}$$

Integrating the each component separately and completing squares in Equation (6), the prior distribution of  $\theta_n/x_{n-1}$  at time n is obtained as follows:

$$f(\theta_n/x_{n-1}) = \sum_{i=0}^{3^{n-1}-1} [p_1 \alpha_i^{(n-1)} N(\theta_i^{(n-1)}, \sigma^2 + \hat{\sigma}_{n-1}^2) + p_2 \alpha_i^{(n-1)} N(\theta_i^{(n-1)} + \delta_1, \sigma^2 + \hat{\sigma}_{n-1}^2) + p_3 \alpha_i^{(n-1)} N(\theta_i^{(n-1)} - \delta_2, \sigma^2 + \hat{\sigma}_{n-1}^2)]$$

$$f(\theta_n/x_{n-1}) = \sum_{i=0}^{3^{n-1}-1} [p_1 \alpha_i^{(n-1)} f_{1i}(\theta_n) + p_2 \alpha_i^{(n-1)} f_{2i}(\theta_n) + p_3 \alpha_i^{(n-1)} f_{3i}(\theta_n)]$$

The purpose of the study is to apply Bayesian approach for detecting the changes in the process mean at each time. Therefore, the following step in this study is to obtain the posterior distribution of the process mean. For this purpose, posterior distribution is obtained by using Bayesian theorem and total probability formulation given as follows:

$$f(\theta_n/x_n) = \frac{f(x_n/\theta_n) f(\theta_n/x_{n-1})}{\int f(x_n/\theta_n) f(\theta_n/x_{n-1}) d\theta_n}$$

Since the denominator given in the right hand side of the above equation is the normalizing coefficient (NC), it is omitted at the beginning of operations and the posterior distribution is described approximately as follows:

$$f(\theta_n/x_n) \propto \sum_{i=0}^{3^{n-1}-1} [p_1 \alpha_i^{(n-1)} f(x_n/\theta_n) f_{1i}(\theta_n) + p_2 \alpha_i^{(n-1)} f(x_n/\theta_n) f_{2i}(\theta_n) + p_3 \alpha_i^{(n-1)} f(x_n/\theta_n) f_{3i}(\theta_n)]$$

As in the previous step for obtaining the prior distribution, the equation is evaluated as three parts and each result related to each part is obtained one by one. After all intermediate operations, the posterior distribution is approximately found as follows:

$$f(\theta_n/x_n) \propto \sum_{i=0}^{3^{n-1}-1} [p_1 \alpha_i^{(n-1)} f_{1i}(x_n) f_{1i}(\theta_n/x_n) + p_2 \alpha_i^{(n-1)} f_{2i}(x_n) f_{2i}(\theta_n/x_n) + p_3 \alpha_i^{(n-1)} f_{3i}(x_n) f_{3i}(\theta_n/x_n)],$$

where

$$f_{1i}(x_n) = \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{(x_n - \theta_i^{(n-1)})^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\},$$

$$f_{2i}(x_n) = \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{[x_n - (\theta_i^{(n-1)} + \delta_1)]^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\}$$

and

$$f_{3i}(x_n) = \frac{1}{\sqrt{2\pi(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)}} \exp \left\{ -\frac{[x_n - (\theta_i^{(n-1)} - \delta_2)]^2}{2(\sigma^2 + \hat{\sigma}_{n-1}^2 + \tau^2)} \right\}.$$

The posterior distributions of  $f_{1i}(\theta_n/x_n)$ ,  $f_{2i}(\theta_n/x_n)$  and  $f_{3i}(\theta_n/x_n)$  are normally distributed. The normalizing coefficient, which is omitted at the beginning of the operations for practical reasons, is placed in, the following equation.

$$(7) \quad f(\theta_n/x_n) = \sum_{i=0}^{3^{n-1}-1} \left[ \underbrace{\left( \frac{p_1 \alpha_i^{(n-1)} f_{1i}(x_n)}{NC} \right)}_{\alpha_{3i}^{(n)}} f_{1i}(\theta_n/x_n) + \underbrace{\left( \frac{p_2 \alpha_i^{(n-1)} f_{2i}(x_n)}{NC} \right)}_{\alpha_{3i+1}^{(n)}} f_{2i}(\theta_n/x_n) + \underbrace{\left( \frac{p_3 \alpha_i^{(n-1)} f_{3i}(x_n)}{NC} \right)}_{\alpha_{3i-1}^{(n)}} f_{3i}(\theta_n/x_n) \right]$$

Although the posterior distribution has three components in this study, the model structure is similar to the result of Tsiamyrtzis and Hawkins's (2005) study for two components. The exact form of the posterior distribution after analytic procedure is given as follows:

$$f(\theta_n/x_n) = \sum_{i=0}^{3^{n-1}-1} \left[ \alpha_{3i}^{(n)} N(\theta_{3i}^{(n)}, \hat{\sigma}_n^2) + \alpha_{3i+1}^{(n)} N(\theta_{3i+1}^{(n)}, \hat{\sigma}_n^2) + \alpha_{3i-1}^{(n)} N(\theta_{3i-1}^{(n)}, \hat{\sigma}_n^2) \right]$$

It is clear that the posterior distribution is a normal mixture model with three components like the prior distribution. We assume three situations in the model: the jump is positive ( $\delta_1 > 0$ ) or negative ( $\delta_2 > 0$ ) or equal to zero. In the study, the upper and lower threshold values are set for detecting the changes in process mean and an estimate which exceeds any of those values indicates a possible change in the mean. An algorithm considered as an extension of Tsiamyrtzis and Hawkins algorithm is developed to implement the proposed Bayesian framework and an artificial data set is used to detect the changes in the mean for short run process.

### 3. Application

An artificial data set is used for the problem of detecting changes in the mean of a short-run process. Since the posterior distribution of the mean is found in the form of  $3^n$  mixture model, then the size of  $n$  is taken as 9 in the application for getting rid of complexity. The artificial data is given in Table 1.

Time	1	2	3	4	5	6	7	8	9
Observation	15	10	10	10	10	10	10	10	5

**Table 1.** An artificial data set

The prior distribution of process mean ( $\theta$ ) at starting point is defined as  $\theta_0 \sim N(10, 4)$ . Where initial mean of  $\theta$  at the  $0^{th}$  stage is  $\zeta = 10$  and initial variance of  $\theta$  is  $\sigma_0^2 = 4$ . The likelihood function of the observations is considered as  $x_n/\theta_n \sim N(\theta_n, 2)$ . Then the random walk model for the application is stated as follows:

$$\theta_n/\theta_{n-1} \sim \begin{cases} N(\theta_{n-1}, 2) & , \quad p_1 = 0.8 \\ N(\theta_{n-1} + \delta_1, 2) & , \quad p_2 = 0.1 \\ N(\theta_{n-1} - \delta_2, 2) & , \quad p_3 = 0.1 \end{cases}$$

The size of a jump in positive direction is taken as  $\delta_1 = 3\sigma$  and size of a jump in negative direction is taken as  $\delta_2 = 2\sigma$  here. Finally, the lower and the upper threshold values are taken as  $T_{\text{lower}} = 8$  and  $T_{\text{upper}} = 13$  respectively. An algorithm designed in Matlab 7.1 which is an extended version of Tsiamyrtzis and Hawkins's (2005), is used to apply the model to the artificial data. The posterior estimates of the means and posterior probabilities are obtained and presented in the following table.

Time	1	2	3	4	5	6	7	8	9
Posterior mean	14.1285	11.1203	10.3231	10.0797	10.0016	9.9764	9.9682	9.9655	6.3789
Posterior probability	0.1993	0.9319	0.9631	0.9553	0.9509	0.9493	0.9487	0.9485	0.0954

**Table 2.** Posterior mean and probability predictions

It is seen that the posterior probabilities which represent the  $P(T_{\text{lower}} < \theta < T_{\text{upper}})$  are close to 1 for the  $2^{nd}$  to  $8^{th}$  observations. It means there is no change in the process mean. But it is not valid for the first and the last values of the application. According to posterior probabilities which are close to zero, the jumps in data indicate the real changes in the process mean.

### 4. Conclusion and Suggestions

In the study, a Bayesian framework is proposed for detecting the change in the unknown mean for a short run process. Since the dynamic structure of  $\theta$

is defined with a random walk containing jumps in both directions at time  $n$ , the prior and posterior distributions of are found to be in a form of  $3^n$  normal mixture model. Existence of the three possible changes at each time point emphasizes that the number of components in the posterior distribution will increase exponentially with  $n$ . Therefore, this model is applicable for only short run processes. The process mean has no change when the point estimate of the unknown mean lies within the upper and lower thresholds and the posterior probability is close to 1. In case of violation, the process mean is assumed to have a change in either positive or negative direction. Such change should be checked carefully and necessary interventions should be applied for the real life process control problem.

## References

1. Kalman, R. E., A new approach to linear filtering and prediction problems, Journal of Basic Engineering, 82, 35-45, 1960.
2. Montgomery, D. C., Introduction to Statistical Quality Control, 2001.
3. Tsiamyrtzis, P. and Hawkins, D. M., A Bayesian scheme to detect changes in the mean of a short run process, Technometrics, 47, 446-456, 2005.
4. Tsiamyrtzis, P. and Hawkins, D. M., A Bayesian Approach to Statistical Process Control, in Bayesian Process Monitoring, Control and Optimization edited by Colosimo, B.M. and Castillo, E.D., 2007.