IJE TRANSACTIONS B: Applications Vol. 28, No. 2, (February 2015) 224-233

International Journal of Engineering

Journal Homepage: www.ije.ir

Estimating the Time of a Step Change in Gamma Regression Profiles Using MLE Approach

F. Sogandi, A. Amiri*

Industrial Engineering Department, Shahed University, Tehran, Iran

PAPER INFO

Paper history:
Received 04 May 2014
Received in revised form 08 July 2014
Accepted 13 November 2014

Keywords:
Gamma Regression Profile
Change Point Estimation
Maximum Likelihood Estimator (MLE)
Statistical Process Control (SPC)
Phase II.

ABSTRACT

Sometimes the quality of a process or product is described by a functional relationship between a response variable and one or more explanatory variables referred to as profile. In most researches in this area the response variable is assumed to be normally distributed. However, occasionally in certain applications, the normality assumption is violated. In these cases, the Generalized Linear Models (GLM) such as Gamma regression models are used to characterize the profile. Also, in statistical process control finding the real time of change in process, called as change point, is necessary because it leads to saving time and cost in finding assignable cause(s). Therefore, in this paper we consider Gamma regression profile and use maximum likelihood to estimate the real time of a step change in Phase II. We evaluate accuracy and precision of the proposed change point estimator by simulation. The results show that the proposed change point estimator is effective in estimating the real time of step shifts in the process parameters of Gamma regression profiles. Also, a confidence set for the process change point based on the logarithm of the likelihood function is presented. Finally, the performance of the estimator is illustrated through a real case.

***doi**: 10.5829/idosi.ije.2015.28.02b.08*

1. INTRODUCTION

Control chart is an effectual tool to reduce variation of process and to monitor quality characteristics. Occasionally, quality of a product or performance of a process is described by a relationship between a response variable and one or more explanatory variables that known as profile. According to the type of this relationship, profiles are classified into categories such as simple linear profiles, multiple linear profiles, polynomial profiles, multivariate linear profiles, non-linear profiles, logistic profiles, and so on.

Control charts have been proven to be effective in detecting out-of-control signals. However, usually the time of the control chart signals is after the real time of a change. Identification of the exact time which in process has changed would simplify exploration and removing of the assignable cause. Consequently, having an estimate of the process change point would be very useful due to reduction of risk of misdiagnosing the control chart signals, which often leads to unnecessary and costly adjustments of the process. Change point problems are classified according to change types including step, drift and monotonic shifts. Generally, step shift happens when the parameter changes suddenly and remains constant until the assignable cause is detected and removed. To find the real time of a change, many authors have suggested several methods. See a comprehensive review on change point estimation methods for control chart post signal diagnostics by Amiri and Allahyari [1]. Perry and Pignatiello [2] showed that the performance of an MLE is better than the built-in EWMA and CUSUM estimator in identifying the change point of a normal and Poisson process, respectively. Amiri and Khosravi [3] proposed an MLE change point estimator in high quality processes under a drift in nonconforming proportion parameter. Amiri and Khosravi [4] proposed an MLE change point estimator under monotonic change for process fraction nonconforming in a high-quality process monitored by a cumulative count of conforming control chart. Ghazanfari et al. [5] suggested a

*Corresponding Author's Email: amiri@shahed.ac.ir (*A. Amiri*)

clustering approach to estimate a step change point. Zandi et al. [6] introduced an MLE of change point under a linear trend disturbance in the fraction nonconforming of a proccess. Noorossana and Heydari [7] proposed a change point estimator to find the real time of a monotonic change in the variance of a normal quality characteristic. Niaki and Khedmati [8] proposed a multi-attribute T^2 control chart to monitor the parameters vector of multi-attribute Poisson processes and then presented an MLE of change point for linear trend and step change disturbances. Moreover, Niaki and Khedmati [9] applied an MLE of change point in a high-yield process when linear trend disturbance occurs in the proportion nonconforming of the process.

Change point estimation after getting a signal from the control chart is also considered in the area of profile monitoring. Mahmoud et al. [10] and Zou et al. [11] proposed methods based on likelihood ratio test to estimate the step change point in simple linear profiles in Phases I and II, respectively. Kazemzadeh et al. [12] used the same method to estimate the change point in polynomial profiles under a step shift in Phase I. Fallahnezhad et al. [13] used a bayesian analysis to estimate the change-point in a sequence of independent random variables from exponential distributions. Keramatpour et al. [14] proposed a remedial measure to remove the effect of autocorrelation in monitoring of autocorrelated polynomial profiles. Then, after using some traditional methods in the literature to monitor the polynomial profiles, they estimated the real time of a step change in the parameters of the polynomial profiles.

In most of researches, distribution of the response variable is assumed to be normal, while in some problems in real world, response variable may follow other exponential family distributions such as Bernoulli, Poisson, Exponential, Gamma, and etc. Sharafi et al. [15] suggested an MLE method to find the exact time of a step change in monitoring of Binary profiles in Phase II. Sharafi et al. [16] investigated estimation of change point of Binary profiles with a linear trend disturbance as well. Recently, Sharafi et al. [17] proposed an MLE of change point method to identify the real time of a step change in Phase II monitoring of poisson regression profiles. To the best of our knowledge, there is no method for estimating the real time of a step change in Gamma regression profiles. Despite there are many real cases which can be characterized by a Gamma regression profile such as the amount of rainfall accumulated in reservoir under different levels of temperature. Hence, it is important to monitor Gamma regression profiles and estimate the real time of a change in the parameters of Gamma regression profile.

In this paper, we propose an MLE method to estimate step change point in Phase II monitoring of Gamma regression profiles. The structure of the paper is as follows: section 2 explains the Gamma regression model and its parameters estimation procedure. In section 3, the change point model is presented. The performance of the proposed model is investigated in section 4 through simulation studies. In section 5, a confidence set is defined and the set cardinality and coverage percentage criteria are computed to evaluate the performance of the change point estimator. In the next section, a real case presented. Finally, conclusions and some recommendations for future researches are provided.

2. GAMMA REGRESSION MODEL

In this paper, we concentrate on estimating the time of step shifts in the Gamma regression profiles in Phase II. Gamma distribution is a distribution that arises naturally in processes for which the waiting times between events are relevant. Thus, there are many real cases in which the response variable follows Gamma distribution. It can be used in a range of disciplines including queuing models, climatology, and financial services. Examples of events that may be modeled by Gamma distribution include:

- The size of loan defaults or aggregate insurance claims.
- The flow of items through manufacturing and distribution processes.
- The load on web servers.
- Waiting time between Poisson distributed events.

The aforementioned Gamma quality characteristics can be related to an explanatory variable and describe a Gamma regression profile. For example, the last quality characteristic and the type of tools or materials used can characterize a Gamma regression profile.

Gamma distribution belongs to a larger class of distributions called the exponential family. Other distributions belonging to the exponential family are the normal, poisson, exponential, and binomial distributions. There are three components that comprise GLM: (i) a random component, the random component is the outcome (Y) and follows a distribution from the exponential family. (ii) a systematic component which needs the **x**'s to be combined in the model as a linear function. (iii) the link function that relates the mean of response variable to linear combination of explanatory variables. Let assume there are P predictor variables for any of n independent experimental sets, which are shown by $\mathbf{x_j} = (x_{j1}, x_{j2}, \ldots, x_{jp})^T$ in which " j=1, 2, ..., n." Assume " y_j " as a Gamma distribution with parameters "(m_j, λ_j)". Then, the Gamma regression profile is modeled by relationship between mean of a Gamma random variable $(\frac{m_j}{\lambda_j})$ and P predictor variables through a log link function, "Often, noncanonical links such as the identity, $m = \mathbf{x\beta}$, and the log, $\log(m) = \mathbf{x\beta}$, are used with Gamma distributed data [18]. The identity link requires restrictions on $\boldsymbol{\beta}$; the log link does not. "The log link is probably the most commonly used for Gamma regression" [19]. Hence, the log link function is used in this paper as follows:

$$log(\frac{m_j}{\lambda_j}) = \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jp}, \quad (1)$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^T$, is the regression parameters vector. We consider $x_{j1} \equiv 1$ for β_1 as the intercept of the model. The alternative equation which directly specify λ_j is as follows:

$$\lambda_j = \frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta})}, \quad (2)$$

In the field, Albert and Anderson [20] used the following likelihood function to approximate the model parameters:

$$l(\mathbf{y}, \boldsymbol{\lambda}) = \prod_{j=1}^{n} \frac{\lambda_j^{m_j}}{\Gamma(m_j)} y_j^{m_j - 1} e^{-\lambda_j y_j}. \quad (3)$$

On the other hand, Equation (4) can be concluded from Equation (2):

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n) = (\frac{m_1}{\exp(\mathbf{x_1}^{\mathbf{T}}\boldsymbol{\beta})}, \frac{m_2}{\exp(\mathbf{x_2}^{\mathbf{T}}\boldsymbol{\beta})}, \ldots, \frac{m_n}{\exp(\mathbf{x_n}^{\mathbf{T}}\boldsymbol{\beta})})^T, \quad (4)$$

and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. Thus, Equation (5) is obtained by replacing Equations (4) in Equation (3):

$$l(\boldsymbol{\lambda}, \mathbf{y}) = \prod_{j=1}^{n} \frac{(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta})})^{m_j}}{\Gamma(m_j)} y_j^{m_j - 1} e^{-(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta})}) y_j}. \quad (5)$$

3. PROPOSED MLE STEP CHANGE POINT

It is supposed that the process performs in a state of statistical control with samples that coming from a Gamma distribution with the known parameters $\boldsymbol{\beta} = \boldsymbol{\beta_0}$ that is a P-dimensional vector in Phase II. Thus, the mass probability function is as follows:

$$f(y_{ij}) = \frac{e^{-(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})}) y_{ij}} (\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})})^{m_j} y_{ij}^{m_j - 1}}{(m_j - 1)!}, \quad (6)$$

where y_{ij} is the value taken by the response variable for the j^{th} value of the predictor variable in the i^{th} profile. After an indefinite amount of time passes, in an unknown profile in τ, which called as the process change point, the parameteres of the process change to an unknown out-of-control state. The parameters after τ can be denoted by $\boldsymbol{\beta_1} = \boldsymbol{\beta_0} + \boldsymbol{\delta}$. Since we consider a step shift in Phase II, the parameters remain at the new level until the source of the assignable cause is identified and omitted. Hence, in the likelihood function for $i = 1, 2, \ldots, \tau$, the process parameter λ is equal to it's known in-control value λ_0. Similarly, for profiles $i = \tau + 1, \tau + 2, \ldots, T$, it become equal to some unknown parameter λ_1 where T is the last profile sampled, in which, unknown parameters in the model are τ and λ_1, which represent the last profile taken from an in-control process and the out-of-control process parameter, respectively. To estimate these unknown parameters along with the change point, we use the MLE approach. The proposed change point estimator is denoted as $\hat{\tau}$. We describe the level of shift $\boldsymbol{\delta}$ in $\boldsymbol{\beta_0}$ and then assess the performance of change point estimator by shifts in the parameters. Based on the aforementioned explanations, the likelihood function for Gamma regression profile is given by:

$$l(\tau, y_{ij}) = \frac{\prod_{i=1}^{\tau}\prod_{j=1}^{n}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})})^{m_j} \prod_{i=1}^{\tau}\prod_{j=1}^{n} y_{ij}^{m_j - 1} e^{-\sum_{i=1}^{\tau}\sum_{j=1}^{n}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})}) y_{ij}}}{\prod_{i=1}^{\tau}\prod_{j=1}^{n}(m_j - 1)!} \times \frac{\prod_{i=\tau+1}^{T}\prod_{j=1}^{n}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_1})})^{m_j} \prod_{i=\tau+1}^{T}\prod_{j=1}^{n} y_{ij}^{m_j - 1} e^{-\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_1})}) y_{ij}}}{\prod_{i=\tau+1}^{T}\prod_{j=1}^{n}(m_j - 1)!}. \quad (7)$$

The MLE of τ is the value of τ that maximizes the likelihood function in Equation (7) or, equivalently, its logarithm. Hence, it is better to take the logarithm of the likelihood function which is shown in Equation (9).

$$\begin{aligned} \ln l(\tau, y_{ij}) = & \sum_{i=1}^{\tau}\sum_{j=1}^{n} m_j \ln(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})}) - \sum_{i=1}^{\tau}\sum_{j=1}^{n} \ln((m_j - 1)!) \\ & + \sum_{i=1}^{\tau}\sum_{j=1}^{n}(m_j - 1)\ln(y_{ij}) - \sum_{i=1}^{\tau}\sum_{j=1}^{n} y_{ij}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_0})}) \\ & + \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}(m_j - 1)\ln(y_{ij}) + \sum_{i=\tau+1}^{T}\sum_{j=1}^{n} m_j \ln(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_1})}) \\ & - \sum_{i=\tau+1}^{T}\sum_{j=1}^{n} y_{ij}(\frac{m_j}{\exp(\mathbf{x_j}^{\mathbf{T}}\boldsymbol{\beta_1})}) - \sum_{i=\tau+1}^{T}\sum_{j=1}^{n} \ln((m_j - 1)!). \end{aligned} \quad (8)$$

To determine the unknown elements of vector $\boldsymbol{\beta}_1$ in Equation (8), we should take the partial derivatives from the vector $\boldsymbol{\beta}_1$ respect to its elements β_{01} and β_{11}, then solve the equations to find the MLE of the parameters β_{01} and β_{11}. Hence, $\hat{\beta}_{01}$ is computed by using Equation (9) as follows:

$$\hat{\beta}_{01} = \ln\left(\frac{\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\frac{m_j y_{ij}}{\exp(x_j\beta_{11})}}{\sum_{j=1}^{n} m_j(T-\tau)}\right). \quad (9)$$

by obtaining $\hat{\beta}_{01}$ and replacing it in partial derivative function respect to β_{11}, Equation (10) is obtained where there is no closed-form solution for β_{11}. Hence, we use Newton's method to solve β_{11} in Equation (10) at each potential change point value. This provides an estimate of β_{11} for each τ without requiring an explicit closed-form expression.

$$\frac{\partial l(\tau, y_{ij})}{\partial \beta_{11}} = \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{x_j y_{ij} n m_j (T-\tau)}{\left(\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{m_j y_{ij}}{\exp(\beta_{11}x_j)}\right)\right)\times\exp(\beta_{11}x_j)}\right) - \sum_{j=1}^{n}(T-\tau)x_j. \quad (10)$$

Newton's method is a derivative-based root finding algorithm that uses the linear approximation. If τ was known, the Newton's method could be used to solve for β_{11} in Equation (10). The $\hat{\beta}_{11,\tau}$ is computed through an iterative algorithm using Equation (11). In this equation, the initial value for $\hat{\beta}_{11,\tau,k}$ is set equal to zero:

$$\hat{\beta}_{11,\tau,k+1} = \hat{\beta}_{11,\tau,k} - \frac{f(\hat{\beta}_{11,\tau,k})}{f'(\hat{\beta}_{11,\tau,k})}, \quad (11)$$

and we have:

$$f(\hat{\beta}_{11,\tau,k}) = \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{x_j y_{ij} n m_j (T-\tau)}{\left(\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{m_j y_{ij}}{\exp(\hat{\beta}_{11,\tau,k}x_j)}\right)\right)\times\exp(\hat{\beta}_{11,\tau,k}x_j)}\right) - \sum_{j=1}^{n}(T-\tau)x_j, \quad (12)$$

and

$$f'(\hat{\beta}_{11,\tau,k}) = \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{-\exp(\hat{\beta}_{11,\tau,k}x_j)x_j y_{ij} n m_j (T-\tau)\times A}{\left(\left(\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{m_j y_{ij}}{\exp(\hat{\beta}_{11,\tau,k}x_j)}\right)\right)\times\exp(\hat{\beta}_{11,\tau,k}x_j)\right)^2}\right),$$
$$A = \left(\hat{\beta}_{11,\tau,k}\times\sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{m_j y_{ij}}{\exp(\hat{\beta}_{11,\tau,k}x_j)}\right) - \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}\left(\frac{\hat{\beta}_{11,\tau,k}m_j y_{ij}}{\exp(\hat{\beta}_{11,\tau,k}x_j)}\right)\right). \quad (13)$$

After computing $\hat{\beta}_{11,\tau}$ based on Newton's algorithm and by using Equation (13), we can compute $\hat{\beta}_{01,\tau}$ for each τ. Then, we replace $\hat{\beta}_{01,\tau}$ and $\hat{\beta}_{11,\tau}$ in the vector of $\hat{\boldsymbol{\beta}}_{1\tau}$ in the logarithm of likelihood function for all possible change point values. The MLE of the change point τ is the value which maximizes the expression in Equation (8). Hence, the estimator of the change point by using the MLE approach is shown as follows:

$$\hat{\tau} = \arg\max\Big[\sum_{i=1}^{\tau}\sum_{j=1}^{n} m_j \ln\left(\frac{m_j}{\exp(\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_0)}\right) + \sum_{i=1}^{\tau}\sum_{j=1}^{n}(m_j-1)\ln(y_{ij}) - \sum_{i=1}^{\tau}\sum_{j=1}^{n} y_{ij}\left(\frac{m_j}{\exp(\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_0)}\right) + \sum_{i=\tau+1}^{T}\sum_{j=1}^{n}(m_j-1)\ln(y_{ij})$$
$$+\sum_{i=\tau+1}^{T}\sum_{j=1}^{n} m_j \ln\left(\frac{m_j}{\exp(\mathbf{x}_j^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{1\tau})}\right) - \sum_{i=\tau+1}^{T}\sum_{j=1}^{n} y_{ij}\left(\frac{m_j}{\exp(\mathbf{x}_j^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{1\tau})}\right)\Big]. \quad (14)$$

We use a shewhart T^2 control chart to monitor a Gamma regression profile and estimate change point in Phase II. It should be noted that Yeh et al. [21] introduced five Hotelling T^2 control charts to monitor Binary profiles in Phase I that any of these T^2 charts uses a different method to estimate the mean vector and covariance matrix. They showed that the T^2 control chart, which estimates the covariance matrix by averaging the covariance estimates of each given sample, is more effective in detecting both step and drift shifts. This control chart is applied in Phase II with the assumption that the mean vector and covariance matrix are known. The T^2 statistic for sample $i(i=1,2,\ldots,T)$ in Phase II is defined as:

$$T_i^2 = (\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_0)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_0), \quad (15)$$

where $\boldsymbol{\beta}_0$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix of Gamma regression parameters, respectively that $\boldsymbol{\Sigma}$ is defined by Equation (16). When the process is in-control, the upper control limit for the proposed control chart is equal to $\chi^2_{2,\alpha}$ which is the α percentile of the chi-square distribution with 2 degrees of freedom and $\mathbf{W} = \mathrm{diag}(\mathrm{var}(y_{i1}), \mathrm{var}(y_{i2}), \ldots, \mathrm{var}(y_{in}))$ is a $P\times P$ diagonal matrix.

$$\Sigma = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}. \quad (16)$$

when the T^2 control chart is employed to monitor a process, as long as the plotted points fall below the upper control limit, the process is assumed in-control. However, when a point exceeds the upper control limit, the control chart signals a change in the parameters of the process and the process is assumed to be out-of-control. In these situations, the most important problem is that there is usually a considerable time lag between the signaling time and the real time at which the change has happened. Thus, whenever the T^2 control chart signals an out-of-control state, the real time of a change can be estimated via Equation (14).

4. PERFORMANCE OF THE MLE ESTIMATOR

In this section, the performance of the proposed estimator is examined using the Monte Carlo simulation with an example. We assume that $\boldsymbol{\beta}_0$ is a 2-dimensional vector for ease representation of formulation and β_{00} and β_{10} are the in-control intercept and the slope of the regression function, respectively, shown by vector $\boldsymbol{\beta}_0 = (\beta_{00}, \beta_{10})^T$. Moreover, we set the design matrix $\mathbf{x}$ as:

$$\mathbf{x} = \begin{bmatrix} 1 & 1 & & 1 & 1 \\ \log(10) & \log(15) & \cdots & \log(45) & \log(50) \end{bmatrix}.$$

It is assumed that the in-control $\boldsymbol{\beta}$ vector is $\boldsymbol{\beta}_0 = (-4,\ 2)^T$, which comes from the historical dataset in Phase I. In addition, the covariance matrix of the Gamma regression parameters Σ in Phase II is computed by the following equation:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 2.0472 & -0.5433 \\ -0.5433 & 0.1446 \end{pmatrix}.$$

Assuming α equal to 0.005, the upper control limit for the T^2 control chart is equal to $\chi^2_{2,0.005} = 10.59$, n=9 and m_j=30 for all j=1, . . ., 9. Now, assume the parameters of Gamma regression model have increasing shifts and the vector $\boldsymbol{\beta}_0$ changes to $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_0 + \boldsymbol{\delta}$ where $\boldsymbol{\delta} = (\delta_1\sigma_1, \delta_2\sigma_2)$ and δ_1, δ_2 are constant coefficients of shifts in the intercept and slope of the Gamma regression profile, respectively, and σ_1 and σ_2 are equal to 1.4308 and 0.38026, respectively. Also, the convergence threshold in the Newton's method to estimate β_{11} is considered equal to 0.0005. A Monte Carlo simulation study is performed to test the performance of the estimator of the step change point. In this study, the process change point is considered at $\tau = 50$. During the generating of profiles $i = 1, 2, ..., 50$, the process parameter is equal to its known in-control value of $\boldsymbol{\beta}_0$. Therefore, for these profiles, the dependent observations are randomly generated from a Gamma regression with parameter vector $\boldsymbol{\beta}_0 = (-4, 2)^T$. It is assumed from profile 51, observations are generated from the out-of-control process with parameter vector $\boldsymbol{\beta}_1$ until the T^2 control chart signals an out-of control state. At this time, the change point estimator in Equation (14) is used and the real time of the process change is determined. This procedure is repeated 10,000 times for different step shifts considered in the paper.

The simulation results are demonstrated in Tables 1, 2, 3 and 4. The mean and the corresponding standard error of simultaneous and individual shifts in paramters of Gamma regression profile are summarized in Tables 1 and 2, respectively. However, the precision performance of the estimator under the mentioned shifts is illustrated in Tables 3 and 4. In each simulation run, E(T) is the expected value of the number of samples taken until the first alarm happens; so, E(T)=ARL+50. Table 1 shows the average change point estimator, the standard error of the change point estimator and E(T) under different magnitudes of step shifts that are considered in this paper. For example, we conclude from the results of simulation for shift equal to (0.01,0.02), the expected number of samples taken until the signal is E(T)=185.98. For this case, the average of the change point estimates is 50.58, which is close to the actual change point of τ = 50 as shown in Table 1. Moreover, the standard error of the change point estimator is 0.12. As another example from the Table 2, for shift equal to (0.03,0), the expected number of samples taken until the signal is E(T)=180.06. For this case, the average of the change point estimates is 50.53 which is close to the actual change point and the standard error of the change point estimator is 0.08. Hence, the proposed change point estimator is suitable for all types of shifts even in small shifts in both simultaneous and individual changes. Furthermore, as the magnitude of the step change increases, the performance of the proposed estimator improve significantly. In other words, the proposed method works well and provides adequately accurate and reliable estimates of the real change point. In order to illustrate the benefit of the proposed change point estimator, E($\hat{\tau}$) is compared to E(T) in Figure 1. It can be easily seen from Figure 1 that if one only relies on E(T) and searches for the special cause around it, most probably, one will not be able to find the cause.

However, the change point estimator $\hat{\tau}$, on average, directs one accurately to the actual change point and enables one to find the cause more effectively. The results in Table 2 are similar to Table l. However, as discussed before, shitfs in the parameters are separetely. As shown in Tables 1 and 2, our proposed change point estimator performs satisfactory for all types of shifts. By comparing the results of simultaneous and individual changes in the regression parameters, we understand that accuracy of the change point estimator for simultaneous shifts is better than the individual shifts in the intercept or slope. Also, accuracy of the change point estimator for shift in the intercept of Gamma regression model is better than the shift in the slope.

TABLE 1. The averages and standard errors of the change point estimator under different step shifts in the parameters (β_1, β_2) simultaneously with 10,000 simulations runs when P =2 and τ =50.

(δ_1, δ_2)	E(T)	$\bar{\hat{\tau}}$	se($\bar{\hat{\tau}}$)
(0.01,0.01)	244.18	52.89	0.23
(0.01,0.02)	185.98	50.58	0.12
(0.01,0.04)	107.18	49.78	0.06
(0.01,0.06)	86.18	50.04	0.03
(0.02,0.01)	183.36	50.38	0.09
(0.02,0.02)	105.45	49.75	0.06
(0.02,0.04)	88.02	49.82	0.04
(0.02,0.06)	67.93	49.98	0.03
(0.03,0.01)	129.26	49.69	0. 05
(0.03,0.02)	107.89	49.77	0.03
(0.03,0.04)	77.75	49.96	0.02
(0.04,0.01)	110.05	49.75	0.06
(0.04,0.02)	90.42	49.79	0.05
(0.05,0.01)	88.19	49.85	0.04
(0.05,0.02)	78.47	49.97	0.03

TABLE 2. The averages and standard errors of the change point estimator under different step shifts in the parameters (β_1, β_2) individually with 10,000 simulations runs when P =2 and τ =50.

(δ_1, δ_2)	E(T)	$\bar{\hat{\tau}}$	se($\bar{\hat{\tau}}$)
(0,0.01)	265.41	53.71	0. 3
(0,0.02)	199.15	52.26	0.15
(0,0.03)	191.54	50.97	0.09
(0,0.04)	114.76	50.34	0.09
(0,0.05)	107.45	50.08	0.03
(0,0.06)	99.94	50.06	0.02
(0.01,0)	258.9	52.97	0.28
(0.02,0)	189.9	51.95	0.11
(0.03,0)	180.06	50.53	0. 08
(0.04,0)	111.05	50.29	0.07
(0.05,0)	95.1	50.26	0.03
(0.06,0)	85.54	49.95	0.02

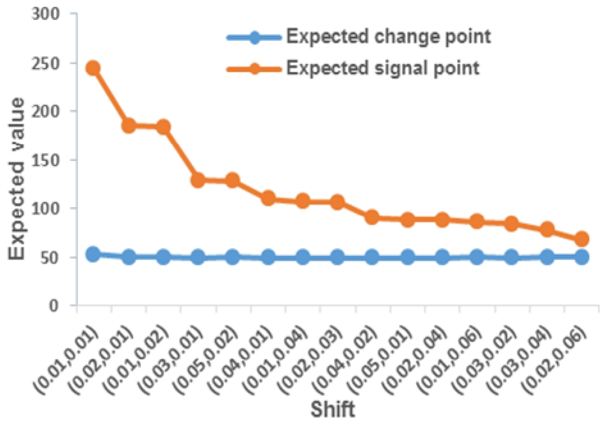

Figure 1. Expected time of a signal with T^2 control chart, average of change point estimates E($\hat{\tau}$) by MLE in 10,000 replicated simulation for τ = 50.

The results of the proportion of 10,000 simulation runs in Tables 3 and 4 demonstrate that the estimator lies within a specified tolerance of the real change point value under different shifts. Assume that precision i under the given shifts are percent of results which distance of the change point estimator from exact change point is i or less than i .

In other words, a measure of the precision of the change point estimator can be examined by constructing a frequency distribution for $P(|\tau - \bar{\hat{\tau}}| \le i)$ where i is equal to 0, 1, . . . , 7, so the results demonstrate that the estimator lies within a specified tolerance. For example in Table 3, under shift equal to (0.02, 0.02), in 58% of simulation runs, distance of change point estimator from exact change point is less than 2 and probability that $\hat{\tau}$ lies within tolerance of 2 or less from the real change point is 0.58 that named as precision 2. Moreover, in this case in 35% of the simulation runs, the estimator correctly identify the real time of the change. Also, we notice that the percentage of those simulation trials identifying the change point correctly are 8%, 14%, 30%, 45%, 21%, 35%, 44% , 58%, 67% and 36%, 48%, 71%, 88%, 90% and 91% for the magnitude of shifts (δ_1, δ_2) = (0.01,0.01), (0.01,0.02), (0.01,0.04), (0.01,0.06), (0.02,0.01), (0.02,0.02), (0.02,0.04), (0.02,0.06), (0.03,0.01) and (0.03,0.02), (0.03, 0.04), (0.04,0.01), (0.04,0.02), (0.05,0.01) and (0.05,0.02) respectively. Clearly, the probability of exact estimation increases as the magnitude of the step shift increases. Hence, as the magnitude of the step change increases, the performance of the estimator improves significantly. Also, the precision of the change point estimator under individual shifts in the intercept and slope are summarized in Table 4. Precision analysis of the change point estimator in this case is similar to the simultaneously shifts in Table 3.

The results show the suitable precision of the proposed estimator in estimating the real time of a step change in the parameters intercept and slope separately. By comparison of Tables 3 and 4, we conclude that the precision of the change point estimator for simultaneous shifts in parameters of Gamma regression model is better than the precision estimator under shifts in the intercept or slope shifts individually. Furthermore, we can conclude that precision of the change point estimator for given shifts in the intercept of Gamma regression model is better than shifts in the slope.

5. CARDINALITY AND COVERAGE PERCENTAGE OF CONFIDENCE SET ESTIMATOR

We consider constructing confidence set on the process change point. Such a set would provide a window of possible change points that covers the true process change point with a given level of confidence and enhances the identification chance of special cause. Box and Cox [22] suggested constructing confidence regions on parameter estimates using the likelihood function by a confidence set of the form:

$$CS = \{t : \log_e L(t) > \log_e L(\hat{\tau}) - D\}, \quad (17)$$

where, $\log_e L(\hat{\tau})$ is the maximum of the log likelihood function evaluated over all possible change points t. If the value of the log likelihood function at time t, $\log_e L(t)$, exceeds the maximum of the log likelihood function minus a reference value D, then t is included in the confidence set.

We use critical values of D between 1 and 8 and δ vectors given in Figure 1 to compute the cardinality and coverage percentage of confidence set estimator. Figure 3 provides a surface plot showing the relationship among cardinality, coverage, δ and D for the confidence set estimator. For example, if δ =(0.04,0.01) and D = 8, the confidence set obtained using Equation (17) will yield an expected cardinality of approximately 37.3. In addition, 47 percent of possible change points contain the change point estimator. Also, the result of coverage and cardinality regarding to different small values of D and δ are reported in Table 5 for more clarification.

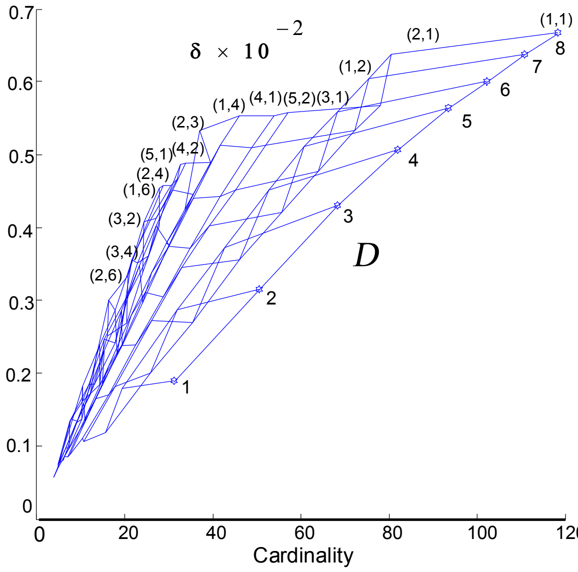

Figure 2. Surface plot obtained from confidence set estimator showing estimated relationships between set cardinality, coverage, and δ value D. Estimated surface obtained from 10,000 independent simulation runs and $\tau = 50$.

TABLE 3. Estimated precision performances under shifts in parameters of β_0 simultaneously with 10,000 simulations runs when P =2 and τ =50.

(δ_1, δ_2)	Precision 0	Precision 1	Precision 2	Precision 3	Precision 4	Precision 5	Precision 6	Precision 7
(0.01,0.01)	0.08	0.17	0.23	0.3	0.35	0.4	0.45	0.56
(0.01,0.02)	0.14	0.28	0.39	0.47	0.53	0.58	0.69	0.72
(0.01,0.04)	0.3	0.52	0.64	0.73	0.79	0. 83	0.95	1
(0.01,0.06)	0.45	0.69	0.77	0.83	0.97	1		
(0.02,0.01)	0.21	0.32	0.49	0.52	0.69	0.89	0.95	1
(0.02,0.02)	0.35	0.4	0.58	0.65	0.78	0.9	0.98	1
(0.02,0.04)	0.44	0.63	0.76	0.85	0.94	0.97	1	
(0.02,0.06)	0.58	0.7	0.89	0.95	1			
(0.03,0.01)	0.67	0.79	0.85	0.96	1			
(0.03,0.02)	0.36	0.43	0.66	0.74	0.8	0.97	1	
(0.03,0.04)	0.48	0.71	0.86	0.98	0.89			
(0.04,0.01)	0.71	0.85	0.96	1	1			
(0.04,0.02)	0.88	0.95	0.98	1				
(0.05,0.01)	0.9	0.96	1					
(0.05,0.02)	0.91	0.98	1					

TABLE 4. Estimated precision performances under shifts in parameters of β_0 individually with 10,000 simulations runs when P =2 and τ =50.

(δ_1,δ_2)	Precision 0	Precision 1	Precision 2	Precision 3	Precision 4	Precision 5	Precision 6	Precision 7
(0,0.01)	0.05	0.11	0.22	0.26	0.35	0.39	0.42	0.5
(0,0.02)	0.12	0.25	0.34	0.41	0.5	0.55	0.57	0.69
(0,0.03)	0.26	0.33	0.42	0.54	0.65	0.71	0.86	0.95
(0,0.04)	0.29	0.45	0.55	0.67	0.7	0.79	0.86	0.98
(0,0.05)	0.31	0.5	0.59	0.75	0. 81	0.95	0.99	1
(0,0.06)	0.42	0.57	0.87	0.9	0.96	1		
(0.01,0)	0.06	0.13	0.2	0.29	0.31	0.45	0.5	0.67
(0.02,0)	0.19	0.3	0.41	0.5	0.63	0.79	0.91	0.98
(0.03,0)	0.35	0.45	0.54	0. 6	0.79	0.86	0.95	1
(0.04,0)	0.41	0.54	0.74	0.88	0.91	0.95	0.98	1
(0.05,0)	0.56	0.65	0.76	0.89	0.92	0.94	1	
(0.06,0)	0.7	0.85	0.94	1				

6. A REAL CASE

In this section, data set from Miller and Wu [23] is used to show the applicability of the Gamma regression profile and evaluate the performance of the change point estimator.

In this case, the weight of the mold is the response variable and is measured over eight levels of a high injection pressure factor. It is known that the amount of material injected could by affected by this factor. Therefore, the relationship between the weights of the mold and pressure could be characterized by a Gamma regression profile. High injection pressure is varied over the range of 650-1000 psi. Miller and Wu [23] presented the data of this real case as 32 runs of an experimental design. We used the data set of run 5 with four replications. First, we estimated parameters of the Gamma regression models for four profiles and computed the average of the intercept and the slope estimates of the four profiles equal to 6.2543 and 0.0003, respectively. Then, we estimated the m_j (j=1,2,…,8) based on the observations in each level of the explanatory variable and then computed the $\bar{m}$ equal to 219265. After that, using the $\bar{m}$ and Equation (2), the scale parameter of the Gamma distribution in each level of the explanatory variable is computed as follows:

$$\lambda = [346.7 \quad 341.6 \quad 336.5 \quad 331.5 \quad 326.6 \quad 321.7 \quad 316.9 \quad 312.2].$$

Therfore, the parameters of the Gamma regression profile are estimated and known. To check the performance of the proposed estimator, we generated new data based on a shift with magnitude of 0.8 in the intercept parameter of the Gamma regression profile and continued generating the data until a signal is taken by the T^2 control chart. Then, we applied the MLE change point estimator and estimated the change point. The T^2 control chart on the real samples as well as simulated ones is illustrated in Figure 3. Also, the real change point which is 4 th point and the estimated change point which is 3^{th} point are shown in this Figure as well. This shows the suitable performance of the proposed estimator in real application.

TABLE 5. The relationships between set cardinality, coverage, and δ value D that obtained from 10,000 independent simulations runs and τ = 50 .

	D / δ	1	2	3	4	5
Coverage(%)	(0.01,0.01)	5.68	9.67	13.16	17	22.56
	(0.02,0.01)	6.97	9.68	13.77	18.39	25.05
	(0.01,0.02)	7.37	12.36	16.15	20	28.64
	(0.03,0.01)	7.53	13.37	18.16	22.9	30.09
	(0.05,0.02)	7.71	13.67	18.41	24.65	30.39
Cardinality	(0.05,0.02)	3.91	6.65	9.13	12.32	16.23
	(0.03,0.01)	5.01	7.33	10.41	13.42	16.56
	(0.01,0.02)	5.02	7.71	10.55	14.5	17.89
	(0.02,0.01)	5.1	7.78	10.59	14.52	18.42
	(0.01,0.01)	5.23	8.02	12.12	15.39	20.56

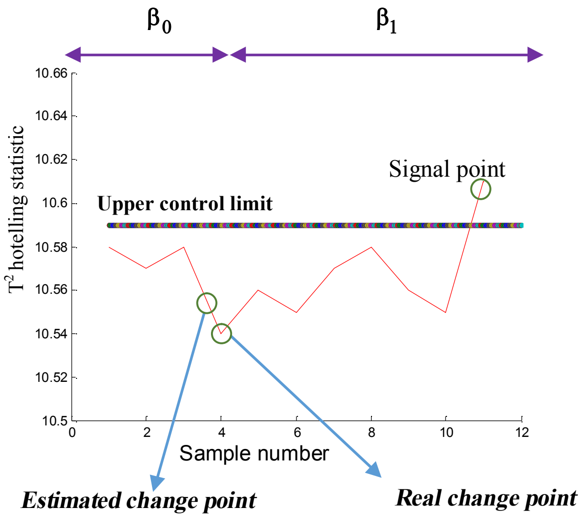

Figure 3. T^2 control chart with a step change in the intercept of Gamma regression profile.

7. CONCLUSION AND FUTURE RESEARCHES

There are many real cases in which the response variable of the profiles does not follow Normal distribution. In these situations, the Generalized Linear Models such as Gamma regression models are used to describe the profile. In this paper, an MLE approach was proposed to identify the time of a step shift in Phase II monitoring of Gamma regression profiles. Then, the performance of the proposed change point estimator was evaluated through simulation studies. Our results coming from Monte Carlo simulation revealed that the change point estimator performs satisfactory under different shifts. Also, the precision and accuracy of the change point estimator for simultaneous shifts is better than the accuracy and precision of the estimator under shifts in the intercept or the slope of the Gamma regression model individually. Moreover, the accuracy and precision of the change point estimator only for shifts in the intercept of Gamma regression model is better than shifts in the slope. Finally, cardinality and coverage percent of a confidence set estimator is analyzed.

Developing a change point estimator for the other change types such as drift and monotonic could be a fruitful area for researches. Also, investigating the other methods such as clustering and artificial neural network for step change point estimation could be considered as future researches. One also may consider the effect of other link functions such as reciprocal link function on the change point estimator and its performance as well. Finally, the effect of missing data (Ashuri and Amiri [24]) on change point estimates can be investigated as future researches.

8. ACKNOWLEDGEMENT

The authors are grateful to the respectful reviewers for precious comments which led to improvement in the paper.

9. REFERENCES

1. Amiri, A. and Allahyari, S., "Change point estimation methods for control chart postsignal diagnostics: A literature review", ***Quality and Reliability Engineering International***, Vol. 28, No. 7, (2012), 673-685.
2. Perry, M.B., Pignatiello, J.J. and Simpson, J.R., "Estimating the change point of a poisson rate parameter with a linear trend disturbance", ***Quality and Reliability Engineering International***, Vol. 22, No. 4, (2006), 371-384.
3. Amiri, A. and Khosravi, R., "Estimating the change point of the cumulative count of a conforming control chart under a drift", ***Scientia Iranica***, Vol. 19, No. 3, (2012), 856-861.
4. Amiri, A. and Khosravi, R., "Identifying time of a monotonic change in the fraction nonconforming of a high-quality process", ***The International Journal of Advanced Manufacturing Technology***, Vol. 68, No. 1-4, (2013), 547-555.
5. Ghazanfari, M., Alaeddini, A., Niaki, S.T.A. and Aryanezhad, M.B., "A clustering approach to identify the time of a step change in shewhart control charts", ***Quality and Reliability Engineering International***, Vol. 24, No. 7, (2008), 765-778.
6. Zandi, F., Niaki, S., Nayeri, M. and Fathi, M., "Change-point estimation of the process fraction non-conforming with a linear trend in statistical process control", ***International Journal of Computer Integrated Manufacturing***, Vol. 24, No. 10, (2011).
7. Noorossana, R. and Heydari, M., "Change point estimation of a normal process variance with monotonic change", ***Scientia Iranica***, Vol. 19, No. 3, (2012), 885-894.
8. Niaki, S.T.A. and Khedmati, M., "Estimating the change point of the parameter vector of multivariate poisson processes monitored by a multi-attribute t 2 control chart", ***The International Journal of Advanced Manufacturing Technology***, Vol. 64, No. 9-12, (2013), 1625-1642.
9. Niaki, S.T.A. and Khedmati, M., "Change point estimation of high-yield processes with a linear trend disturbance", ***The International Journal of Advanced Manufacturing Technology***, Vol. 69, No. 1-4, (2013), 491-497.
10. Mahmoud, M.A., Parker, P.A., Woodall, W.H. and Hawkins, D.M., "A change point method for linear profile data", ***Quality and Reliability Engineering International***, Vol. 23, No. 2, (2007), 247-268.
11. Shang, Y., Tsung, F. and Zou, C., "Profile monitoring with binary data and random predictors", ***Journal of Quality Technology***, Vol. 43, No. 3, (2011), 196-208.
12. Kazemzadeh, R., Noorossana, R. and Amiri, A., "Phase i monitoring of polynomial profiles", ***Communications in Statistics—Theory and Methods***, Vol. 37, No. 10, (2008), 1671-1686.
13. Fallahnezhad, M.S., Rasti, B. and Abooie, MH., , "Improving the performance of bayesian estimation methods in estimations of shift point and computations of shift point and comparison with mle approach", ***International Journal of Engineering,***, Vol. 27, No. 6, (2013), 921-932.
14. Keramatpour, M., Niaki, S., Khedmati, M. and Soleymanian, M., "Monitoring and change point estimation of ar (1) autocorrelated polynomial profiles", ***International Journal of***

Engineering-Transactions C: Aspects, Vol. 26, No. 9, (2013), 933-942.

15. Sharafi, A., Aminnayeri, M. and Amiri, A., "Identifying the time of step change in binary profiles", ***The International Journal of Advanced Manufacturing Technology***, Vol. 63, No. 1-4, (2012), 209-214.
16. Sharafi, A., Aminnayeri, M., Amiri, A. and Rasouli, M., "Estimating the change point of binary profiles with a linear trend disturbance", ***International Journal of Industrial Engineering***, Vol. 24, No. 2, (2013), 123-129.
17. Sharafi, A., Aminnayeri, M. and Amiri, A., "An mle approach for estimating the time of step changes in poisson regression profiles", ***Scientia Iranica***, Vol. 20, No. 3, (2013), 855-860.
18. McCullagh, P. and Nelder, J.A., "Generalized linear models", (1989).
19. Bedrick, E.J., Christensen, R. and Johnson, W., "A new perspective on priors for generalized linear models", ***Journal of the American Statistical Association***, Vol. 91, No. 436, (1996), 1450-1460.
20. Albert, A. and Anderson, J., "On the existence of maximum likelihood estimates in logistic regression models", ***Biometrika***, Vol. 71, No. 1, (1984), 1-10.
21. Yeh, A.B., Huwang, L. and Li, Y.-M., "Profile monitoring for a binary response", ***IIE Transactions***, Vol. 41, No. 11, (2009), 931-941.
22. Box, G.E. and Cox, D.R., "An analysis of transformations", ***Journal of the Royal Statistical Society. Series B (Methodological)***, (1964), 211-252.
23. Miller, A. and Wu, C., "Parameter design for signal-response systems: A different look at taguchi's dynamic parameter design", ***Statistical Science***, Vol. 11, No. 2, (1996), 122-136.
24. Ashuri, A.a.A., A., , " Evaluating estimation methods of missing data on a multivariate process capability index", ***International Journal of Engineering***, Vol. 28, No. 1, (2014), 88-96.

Estimating the Time of a Step Change in Gamma Regression Profiles Using MLE Approach

F. Sogandi, A. Amiri

Industrial Engineering Department, Shahed University, Tehran, Iran

P A P E R I N F O

Paper history:
Received 04 May 2014
Received in revised form 08 July 2014
Accepted 13 November 2014

Keywords:
Gamma Regression
Profile
Change Point Estimation
Maximum Likelihood
Estimator (Mle)
Statistical Process
Control (Spc)
Phase Ii.

چکیده

گاهی اوقات کیفیت یک محصول یا فرآیند بوسیله یک رابطه تابعی بین یک متغیر پاسخ و یک یا چند متغیر مستقل تحت عنوان پروفایل توصیف می‌شود. در اکثر تحقیقات در این حوزه توزیع متغیر پاسخ نرمال فرض می‌شود، در حالیکه گاهی در کاربردهای خاصی فرض نرمال بودن نقض می‌شود. در این موارد از مدل های خطی تعمیم یافته مثل مدل رگرسیونی گاما برای توصیف پروفایل استفاده می‌شود. همچنین در کنترل فرآیند آماری پیدا کردن زمان واقعی تغییردر فرآیند که نقطه تغییر نام دارد ضروری است زیرا منجر به صرفه جویی زمان وهزینه در شناسایی دلایل خاص می‌شود. بنابراین در این مقاله پروفایل رگرسیونی گاما در نظر گرفته می‌شود و از برآوردکننده حداکثر درست نمایی برای تخمین زمان واقعی یک تغییرپله ای در پارامترهای گاما در فاز ۲ استفاده می‌شود. به علاوه دقت وصحت برآورد کننده نقطه تغییر ارائه شده بوسیله شبیه سازی مونت کارلو ارزیابی می‌شود. نتایج شبیه‌سازی نشان می‌دهد که برآورد کننده نقطه تغییر ارائه شده در تخمین نقطه واقعی تغییرات پله ای در پارامترهای پروفایل رگرسیونی از عملکرد مناسبی برخوردار است. همچنین یک مجموعه اطمینان برای نقطه تغییر فرآیند براساس لگاریتم تابع درست نمایی ارائه شده است. نهایتاً عملکرد برآورد کننده با یک مثال واقعی نشان داده شده است.

.***doi***: *10.5829/idosi.ije.2015.28.02b.08*