Estimating the change point of the cumulative count of a conforming control chart under a drift

Amirhossein Amiri *, Ramezan Khosravi

Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, P.O. Box 18151-159, Iran

Received 8 May 2011; revised 21 November 2011; accepted 18 December 2011

KEYWORDS
Drift;
Change point;
CCC control chart;
Statistical process control;
Maximum likelihood estimator.

Abstract In a high quality process, the fraction of nonconforming is very low. In this area, standard Shewhart control charts are no longer useful. The Cumulative Count of Conforming (CCC) control charts, which enumerates the number of conforming items between the occurrences of two nonconforming ones, have been shown to be effective in the monitoring of high quality processes. When the CCC control chart signals an out-of-control condition, the process engineers should search for the source of the assignable causes. Knowing the exact time of the process change would help them to reduce the time for identification of the assignable causes. This paper provides a maximum likelihood estimator for the change point of the nonconforming level of the high quality process with a linear trend. Then, a Monte Carlo simulation is applied to evaluate the performance of the proposed estimator. In addition, the proposed estimator is compared with the MLE of the process fraction nonconforming, derived under a single step change. The results show that the proposed estimator outperforms the MLE designed for step change, when a linear trend disturbance is present in the process.

1. Introduction

Control charts are useful tools used to monitor the process change by distinguishing between the assignable causes and the common causes of variation. When a control chart signals an out-of-control alarm, the process engineers initiate the search for assignable causes of variation. For this purpose, the more experience and knowledge they have about the process, the faster and more accurately they can identify and eliminate the assignable causes. However, if they can estimate the time when the process disturbance first manifested into the observations, valuable time can be saved, because they concentrate on the smaller range of observations, which helps them to quickly find the possible source of variations. This time is considered the change point, and there are different change types considered in the literature. One potential type of change is a step change. Estimating the real time of a step change in the parameters of various distributions has been addressed by different researchers. Samuel et al. [1,2] considered step changes in a normal distribution with mean \( \mu \) and variance \( \sigma^2 \) parameters, respectively. Samuel and Pignatiello [3] estimated the change point in a normal process mean in SPC application, and compared the performance of a maximum likelihood estimator with built-in estimators of CUSUM and EMWA. Pignatiello and Simpson [4] studied normal processes, and proposed a magnitude-robust control chart for monitoring and estimating a step change in the mean. Nedumaran et al. [5] investigated the time of the step change in \( \chi^2 \) control charts. Samuel and Pignatiello [6] estimated the real time of a step change in the parameter of the Poisson process, \( \lambda \). Pignatiello and Samuel [7] identified the time of a step change in the process fraction nonconforming. Their proposed change point estimator was a maximum likelihood estimator used when a step change occurred in the fraction nonconforming of a binomial process.

Another potential change type is a drift in which the parameter of a process begins to change linearly from its control value in the period of time, until one point lies at out-of-control limits. These types of changes, addressed by different researchers, can occur as a result of tool wear, worker's fatigue, filters

* Corresponding author.
E-mail addresses: amiri@shahed.ac.ir (A. Amiri), r.khosravi@shahed.ac.ir (R. Khosravi).
Peer review under responsibility of Sharif University of Technology.
that become dirty over time or any other time related factors.
Perry et al. [8] proposed methods for estimating the change point of a Poisson rate parameter with linear trend disturbance. Perry and Pigantiello [9] proposed a change point estimator to find the real time of a drift in the mean of the process level, following a signal issued by $X$ control charts. Fahmy and Elsayed [10] proposed MLE for estimating the change point in Shewhart control charts under linear trend disturbance. Perry [11] developed a MLE change point estimator with a drift in the mean of auto correlated processes. Zandi et al. [12] proposed a model for the change point problem in which a maximum likelihood estimator is applied when a linear trend disturbance in the process fraction nonconforming is present.

There is another change type, called monotonic change, which is more general, because it encompasses both step changes and drifts. In monotonic changes, the type of change is unknown, but it is assumed that the direction of shifts is the same, increasing or decreasing. Perry et al. [13] estimated the time of a monotonic change in the parameter of the Poisson process, $\lambda$. Perry et al. [14] proposed a maximum likelihood estimator for estimating the change point of the process fraction nonconforming with the monotonic change. More information about change points can be found in a review paper provided by Amiri and Allahyari [15].

Because of the low fraction nonconforming items in high quality processes, the traditional Shewhart control charts cannot be used. This has led to development of a new type of control chart, based on the cumulative count of conforming items, called CCC control charts. In this chart, the number of items is counted until observation of a nonconforming one. Since these counts follow a geometric distribution, the control chart based on this distribution was designed to monitor the process fraction nonconforming in high quality processes. These charts were first developed by Calvin [16], and were then further studied by several researchers, such as Goh [17], Kaminsky et al. [18], Nelson [19], Quezennery [20], Xie and Goh [21], McCool and Joyner-Moltery [22] and Xie et al. [23].

The problem of change point in high quality processes was considered too. Noorossana et al. [24] identified the period of step change following a signal from a CCC control chart with a Maximum Likelihood Estimator (MLE) in a high quality process. When the process fraction nonconforming, $p$, is changed, the CCC control chart signals an out-of-control condition. They assumed that the change in the process fraction nonconforming is a simple step change, and derived the likelihood function of geometric distribution in order to estimate the change point. They applied Monte Carlo simulation and showed that their proposed model performs well in estimating change point when a single step change occurs in the fraction nonconforming in a high quality process.

However, as discussed previously, sometimes a drift can occur in the process parameters, such as a fraction of nonconforming, in a high yield process. This paper provides a maximum likelihood estimator for the period when change with a linear trend in the process nonconforming level, $p$, occurs in high quality processes. The remainder of the paper is outlined as follows: The proposed MLE approach is presented in Section 2. An illustrative example is given in Section 3 to show the proposed model in more detail. In Section 4, Monte Carlo simulation is applied to evaluate the performance of this estimator, and then the proposed model is compared with the MLE of the process fraction nonconforming derived under a simple step change. Our concluding remarks are given in the final Section.

2. Proposed change point estimator

In high quality processes, the fraction of nonconforming, $p$, is in the range of part per million. In this situation, the traditional $p$ control chart cannot be used. Many researchers, such as Xie and Goh [25] and Woodall [26], have warned about this problem. The reason is that when the fraction of nonconforming in a process is very low, the signals for an out-of-control depend heavily on the choice of sample size, and a relatively large sample size is required to detect, on time, shifts in the process mean. To overcome this problem, a new type of control chart has been developed, which is based on the Cumulative Count of Conforming (CCC) items. In the CCC control charts, the number of items until the occurrence of a nonconforming one is considered as a geometric random variable and a control chart based on this distribution is applied to monitor the process nonconforming level (Goh [17]). With the process nonconforming fraction, $p$, the P.D.F and C.D.F of the CCC control chart are:

\[
g(x, p) = p(1-p)^{x-1}, \quad x \geq 1.
\]

\[
x \mapsto \ln g(x, p) = \ln p + (x-1) \ln(1-p).
\]

Equation (1) can be written in terms of a geometric distribution in order to estimate the changepoint, the process fraction of nonconforming changes to $\lambda$, after an unknown point in period $\tau$, referred to as a change point, the process fraction of nonconforming changes to $p_1 = p_0 + \beta(j - \tau)$ for observations $j = \tau + 1, \ldots$, and it is supposed that it remains in this situation until the assignable cause of variation is identified and eliminated. $\beta$ is the slope of linear trend disturbance, $X_j$ is the number of items sampled up to observation of the first nonconforming item at period $j$. Period $j$ is the time between the observation of the $(j-1)$st and $j$th nonconforming items. It is obvious that $X$ follows a geometric distribution. The generation of $X$ from an out-of-control process, with the proportion, $p_1$, continues until period $T$ at which the CCC control chart signals an out-of-control situation. In other words, $LC_L < X_t < UCL_c$ for $j < T$, and $X_t < LC_L$ or $X_t > UCL_c$ for $j = T$. After the control chart issues a signal, we should look for assignable causes of variations. It is desired for us to identify the change point, $\tau$, in order to constrict the search range.

The method presented here is based on the MLE approach. When the CCC control chart signals an out-of-control state, this estimator can be applied to determine at which point the process parameter, $p_0$, has changed. We denote the MLE of the change point as $\hat{\tau}$. The likelihood function is:

\[
L(\tau, \beta|x) = \prod_{j=1}^{T} p_0(1-p_0)^{y_j-1} \prod_{j=\tau+1}^{T} (1-p_0 - \beta(j-\tau)).
\]

(2)

$\hat{\tau}$ is the value of $\tau$ that maximizes this likelihood function or equivalently its logarithm. The logarithm of the likelihood function is:

\[
\ln L(\tau, \beta|x) = \tau \ln p_0 + \sum_{j=1}^{\tau} x_j - \tau \right) \ln(1-p_0)
\]

\[
\ln L(\tau, \beta|x) = \tau \ln p_0 + \sum_{j=1}^{\tau} x_j - \tau \right) \ln(1-p_0)
\]

\[
\times (p_0 + \beta(j-\tau)).
\]

(3)
In the logarithm of the likelihood function, two parameters, \( \tau \) and \( \beta \), are unknown. Hence, we first obtain an approximation for \( \beta \) in terms of \( \tau \), which maximizes Eq. (3) and denotes it as \( \hat{\beta} \). Taking the partial derivative of Eq. (3), with respect to \( \beta \), we have:

\[
\frac{\partial \ln L}{\partial \beta} = \sum_{j=1}^{T} \left[ \frac{(x_j - \tau) - (j - \tau)}{1 - p_0 - \hat{\beta} (j - \tau)} + \frac{j - \tau}{p_0 + \hat{\beta} (j - \tau)} \right] = 0.
\]  

(4)

As can be seen in Eq. (4), there is no closed-form solution for \( \beta \). So, we use a numerical calculation to solve for \( \beta \) at each potential change point. Newton’s method is a derivative based algorithm that uses a linear approximation for finding the roots of an equation (see [27]).

Since \( \tau \) is unknown, we should perform the algorithm for every potential change point, \( \tau \), so that the result for each change point will be \( \hat{\beta}_{\tau_k} \), i.e. for every potential \( \tau \), there is a corresponding \( \hat{\beta}_{\tau_k} \).

\[
\hat{\beta}_{\tau_{k+1}} = \hat{\beta}_{\tau_k} - \sum_{j=1}^{T} \left[ \frac{(x_j - \tau) - (j - \tau)}{1 - p_0 - \hat{\beta}_{\tau_k} (j - \tau)} + \frac{j - \tau}{p_0 + \hat{\beta}_{\tau_k} (j - \tau)} \right] - \sum_{j=1}^{T} \left[ \frac{-(j - \tau)^2 (x_j - \tau)}{1 - p_0 - \hat{\beta}_{\tau_k} (j - \tau)} - \frac{(j - \tau)^2}{p_0 + \hat{\beta}_{\tau_k} (j - \tau)} \right].
\]  

(5)

It is necessary to notice that the fraction of nonconforming, \( p \), is always greater than zero. Therefore, \( \beta \) must be greater than \(-p_0/(j - \tau)\) for any given \( j \). For increasing rates, the procedure will work without any problem, because \( \beta \) has no upper bound. But, in cases of decreasing rates, Newton’s method will fail, because in this state, \( \beta \) is no longer unconstrained, and the linear trend would eventually produce a negative value for \( p \), which is impossible. So, in this procedure, only the positive values of \( \beta \) are considered.

Finally, substituting \( \hat{\beta} \), for \( \beta \) in Eq. (3) and evaluating Eq. (3) among all potential change points in search of \( \tau \) leads to:

\[
\hat{\tau} = \arg \max_{0 \leq \tau < t} \left\{ t \ln p_0 + \left( \sum_{j=1}^{T} x_j - \tau \right) \ln (1 - p_0) + \sum_{j=1}^{T} \ln \left[ \frac{(p_0 + \hat{\beta} (j - \tau))^{n-1}}{(1 - p_0 - \hat{\beta} (j - \tau))} \right] \right\}.
\]  

(6)

### 3. Numerical example

In order to show the procedure more clearly, a numerical example is presented in this section. A set of nonconforming data from a high quality process, with the known fraction nonconforming of \( p_0 \) equal to 0.0005 (500 ppm), is given in [28]. The following control limits were obtained for this process quality characteristic, assuming the probability of Type I error being equal to 0.0027:

\[
\text{UCL} = \frac{\ln(\alpha/2)}{\ln(1 - p_0)} = 13211.99.
\]

And:

\[
\text{LCL} = \frac{\ln(1 - \alpha/2)}{\ln(1 - p_0)} = 2.70.
\]

The corresponding CCC control chart was depicted, showing that all points are within the control limits and the process is in control. But, in order to show the application of the proposed change point estimator, we used the first 10 points of those data exactly and after it, we induce a drift to process fraction nonconforming with a rate of \( \beta = 0.005 \). The CCC control chart signaled an out-of-control situation at period, \( T = 19 \). To find the change point by using Eq. (5), we computed the values of \( \hat{\beta} \) for period time between 0 and 18. Then, the values of the logarithm for each point have been computed from Eq. (3). As shown in Table 1, the largest value of likelihood function corresponds to period 10, which indicates that the change in the process nonconforming level has most likely occurred at this observation. So, to find assignable causes, we should check the records around the period time of 9.

<table>
<thead>
<tr>
<th>Period no. (j)</th>
<th>( t )</th>
<th>( X_j )</th>
<th>( \hat{\beta}_j )</th>
<th>( \ln L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>227</td>
<td>3.48E-05</td>
<td>-159.827</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2269</td>
<td>4.51E-05</td>
<td>-159.335</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1193</td>
<td>6.07E-05</td>
<td>-158.657</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>4106</td>
<td>8.7E-05</td>
<td>-157.579</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>154</td>
<td>0.000134</td>
<td>-156.083</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>12198</td>
<td>0.00255</td>
<td>-152.934</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>201</td>
<td>0.00436</td>
<td>-150.288</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>9612</td>
<td>0.01365</td>
<td>-142.391</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>4045</td>
<td>0.04113</td>
<td>-134.967</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>678</td>
<td>0.006462</td>
<td>-134.508</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>37</td>
<td>0.007557</td>
<td>-137.101</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>9</td>
<td>0.009168</td>
<td>-139.63</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>132</td>
<td>0.012849</td>
<td>-141.705</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>4</td>
<td>0.015119</td>
<td>-144.836</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>17</td>
<td>0.02147</td>
<td>-147.526</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>75</td>
<td>0.048814</td>
<td>-149.241</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>35</td>
<td>0.159189</td>
<td>-151.562</td>
</tr>
<tr>
<td>18</td>
<td>17</td>
<td>14</td>
<td>0.226457</td>
<td>-157.398</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>1</td>
<td>0.312841</td>
<td>-161.915</td>
</tr>
</tbody>
</table>

### 4. Performance evaluation

For evaluating the performance of the proposed estimator, we performed a Monte Carlo simulation study. First, by using a geometric distribution, we generated 100 observations from an in-control process, with \( p = p_0 \). It should be noted that during the generation of these observations, it is possible to have some observations with \( X_j \) values, exceeding either of the control limits. These observations are considered a false alarm, since we know that the process is in-control and observations come from an in-control process, with \( p = p_0 \). So, if a false alarm is occurred at period \( t < \tau \), it is rational to exclude such observations from the computation and restart the control chart, i.e. in general, the first 100 observations have to be in control limits.

Starting with period 101, we induced a drift in the process nonconforming level that changes \( p_0 \) to \( p_1 = p_0 + \beta (t - \tau) \). Then, we generated enough observations until the CCC control chart signals an out-of-control condition. At this point, \( T \), the process is stopped and the proposed MLE estimator is applied to estimate the change point \( \hat{\tau}_k \), which should be close to 100. Furthermore, we recorded the number at which the
control chart has signaled to compute the expected number of samples taken until the first alarm is given, named $\hat{E}(T)$, which relates to ARL with $\hat{E}(T) = \text{ARL} + \tau$. To evaluate and compare the precision of the proposed change point estimator with the estimator provided by Noorossana et al. [24], i.e. $\hat{\tau}_{sc}$, 10,000 simulation runs are applied. The average and standard deviations of the proposed change point estimator, as well as the one by Noorossana et al. [24], along with estimation of $\hat{E}(T)$, are computed and the results are shown in Table 2.

Table 2: Accuracy performances for two MLEs of the change point for $p_0 = 0.0005$ and different values of $\beta$ following a signal from a CCC control chart when a linear trend change is present ($\tau = 100$).

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\hat{E}(T)$</th>
<th>$\hat{\tau}_{nt}$</th>
<th>$\text{SE}(\hat{\tau}_{nt})$</th>
<th>$\hat{\tau}_{sc}$</th>
<th>$\text{SE}(\hat{\tau}_{sc})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000005</td>
<td>402.4105</td>
<td>122.623</td>
<td>126.033</td>
<td>152.212</td>
<td>39.277</td>
</tr>
<tr>
<td>0.00001</td>
<td>299.9815</td>
<td>107.9</td>
<td>65.1</td>
<td>135.835</td>
<td>27.247</td>
</tr>
<tr>
<td>0.00002</td>
<td>254.7</td>
<td>102.787</td>
<td>32.209</td>
<td>125.029</td>
<td>18.834</td>
</tr>
<tr>
<td>0.00003</td>
<td>218.6618</td>
<td>101.669</td>
<td>25.749</td>
<td>114.75</td>
<td>11.892</td>
</tr>
<tr>
<td>0.00005</td>
<td>179.3</td>
<td>100.7</td>
<td>15.4</td>
<td>101.065</td>
<td>6.179</td>
</tr>
<tr>
<td>0.001</td>
<td>157.7</td>
<td>99.998</td>
<td>8.6</td>
<td>94.743</td>
<td>4.528</td>
</tr>
<tr>
<td>0.005</td>
<td>126.9718</td>
<td>99.979</td>
<td>4.5</td>
<td>102.483</td>
<td>10.107</td>
</tr>
<tr>
<td>0.005</td>
<td>108.7</td>
<td>99.8</td>
<td>4.1</td>
<td>101.097</td>
<td>8.518</td>
</tr>
</tbody>
</table>

Table 2 reveals that a drift in the fraction of nonconforming level, with a rate of 0.00001, would be detected by the CCC control chart at an average of 199.99 periods after the occurrence of the actual change. In this case, the proposed MLE provides an average of 107.9 for the change point, which is reasonably close to the actual change point of 100, and which provides a better estimation, in comparison to $\hat{\tau}_{sc}$, which provides an average of 135.84 for the change point. Table 2 indicates that as the slope of the change in the process nonconforming level increases, the standard error becomes very small and the value of $\hat{\tau}$ gets closer to the true value, but more accurate estimates of the change point for different values of $\beta$ are obtained using the proposed estimator.

The probability of the change point estimate lying within a certain difference from the true change point, under $p_0 = 0.00005$ and different values of $\beta$, for both mentioned estimators, is reported in Table 3. In this table, the precision estimates of $\hat{\tau}_{sc}$ are shown in parentheses.

According to the results in Table 3, the $\hat{\tau}_{nt}$ estimator provides more or at least equal precision, in comparison with $\hat{\tau}_{sc}$, if the changes in the parameter of fraction nonconforming follow a linear trend model. For example, when the magnitude of the drift is 0.0001, the $\hat{\tau}_{nt}$ estimates the change point correctly 12% of the time, while the $\hat{\tau}_{nt}$ only in 1% of the time finds the true change point. Based on the results in Table 3, the precision estimates of the two estimators are plotted in Figures 1–4, where they show the precision of estimators versus different slopes of change trend for specified tolerances (the scale of the slope is $10^{-6}$).

It can be seen from the figures that the $\hat{\tau}_{nt}$ estimator outperforms the $\hat{\tau}_{sc}$ estimator. If obtaining the true change point with minimum possible tolerance is especially crucial for process engineers, applying the proposed model is completely essential, as can be seen in Figures 1 and 2.

5. Conclusions

In this paper, we proposed an MLE estimator that finds the real time of a change in the high yield process nonconforming level, under a linear trend disturbance, following a signal given by a CCC control chart. Then, a Monte Carlo simulation is applied to evaluate and compare the precision of this estimator with the method developed by Noorossana et al. [24] for single step change, when the linear trend disturbance is present. The results show the priority of the proposed MLE estimator for

References


Amirhossein Amiri is Assistant Professor at Shahed University, in Iran. He holds B.S., M.S., and Ph.D. degrees in Industrial Engineering from Khajeh Nasir University of Technology, Iran University of Science and Technology, and Tarbiat Modares University, respectively, and is a member of the Iranian Statistical Association. His research interests are statistical quality control, profile monitoring and Six Sigma.

Ramezan Khosravi is an M.S. degree student in Industrial Engineering at Shahed University in Iran. He holds a B.S. degree in Industrial Engineering (2009) from Bojnourd University in Iran. His research interests include statistical process control especially in high quality processes and change-point analysis.