Journal of Statistical Theory and Applications, Vol. 16, No. 2 (June 2017) 209–218

Comparing Two MLEs of The Change Point When an $\bar{X}$ Control Chart Is Used

Mohammad E. Dehghan Monfared

Department of Statistics, Faculty of Sciences, Persian Gulf University, Bushehr 75169, Iran
monfared2@gmail.com

Fazlollah Lak*

Department of Statistics, Faculty of Sciences, Persian Gulf University, Bushehr 75169, Iran
fazlollahlak@gmail.com

Received 29 February 2016

Accepted 22 October 2016

To find the change point of a normal process monitored by an $\bar{X}$ control chart, the first and most important step is to find a good estimate of the change point. Different from all the previous studies in which the MLE of the change point is computed from a conditional likelihood function given the signal time of the $\bar{X}$ control chart, in this paper we also derive the MLE of the change point from the complete likelihood function. Then, the two MLEs are compared through a series of simulations based on two criteria, MSE and a new proposed criterion. The new criterion is just the average number of points in time, ordered according to their likelihood, at which the process should be examined to find the change point. The results show that our estimator is superior to the usual MLE which was used in previous studies and this superiority is shown better by our new criterion, which is more rational, when a control chart is used.

Keywords: Change point; Maximum Likelihood estimation; Mean square error; $\bar{X}$ chart.

2000 Mathematics Subject Classification: 62F10, 62F30, 62F99

1. Introduction

Statistical process control (SPC) charts are commonly used for detecting the presence of disturbances in a process. The primary function of SPC charts is that an out-of-control signal will be triggered when the process disturbances have occurred in the process. In most situations, it is assumed that the quality characteristic of a product follows a normal distribution. In this article, we will consider the Shewhart $\bar{X}$ control chart, the most commonly used statistical process control chart in industry. The control limits for a Shewhart $\bar{X}$ control chart is as follows,

$$UCL = \mu_0 + k\sigma_0$$
$$LCL = \mu_0 - k\sigma_0,$$

*Corresponding author.

where μ_0 and σ_0 are the mean and standard deviation of the normal process, when it is under control. It is usually assumed that $k = 3$, although it can be identified in light of the specified type I error (i.e., α).

When the mean of the normal process which is monitored by an $\bar{X}$ control chart changes at an unknown time, referred to as a change point, it takes some time for control chart to detect the change and make a signal. The main goal of estimating the change point when a control chart is used, is to identify the time at which the disturbances are introduced into the process. Throughout this paper, we refer to this time as the true change point. Using this fact, a reasonable criterion for estimating the change point, can be based on the number of the points in time (NPT) at which the process should be examined to find the true change point. Obviously this criterion depends on the method which is used to search the process and find the true change point. The typical method (TM) is to search for the true change point starting from the initial signal time. If the root causes cannot be identified at the initial signal time T; the searching process may proceed at time $T-1$. This process should continue until the true change point is found.

Two other intuitive methods were proposed by Dehghan Monfared and Lak (2015) to find the true change point, which have different NPTs, and corresponding to each, a loss function is defined. The average of each one of these two loss functions, defines a risk function. It is shown that how this fact can be used to propose a new criterion for the efficiency of the MLE of the change point.

In their paper, Samuel, *et al.* (1998) addressed the issue of estimating the change point of a normal process. Pinatiello and Samuel (2001) used exponentially weighted moving average (EWMA) and Cusum charts and 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 in the case of the S chart and MLE are used in a gamma process. Later Shao, *et al.* (2006), used an $\bar{X}$ control chart and MLE to estimate the change point of a Gamma process, also, in Machine Learning (ML) approach for the change point problems, see Cheng and Cheng (2008) and Shao and Hsu (2009). Shao and Hou (2013a) applied an integrated approach of neural network and analysis of variance to identify a change point in an industrial process. Shao and Hou (2013b) used a two- stage hybrid scheme to estimate a change point for a multivariate process. Hou, *et al.* (2013) used a combined MLE and Generalized P chart approach to estimate the change point of a multinomial process.
All the above studies assume that the signal time of the control chart, denoted by T, is fixed and use a conditional likelihood function to derive the MLE of the change point. In addition to the usual MLE, which is computed from a conditional likelihood given the signal time T, here we also consider a case in which T is a random variable and based on a complete likelihood function, which contains the information of the signal time T about the change point, the MLE of the change point is derived. We also compare the two MLEs of the change point based in two criteria, the MSE and a new proposed criterion which is based on the NPT of a suitable method for finding the true change point. We show that, in this case, the MSE is not a good criterion while our proposed criterion performs quite well.

2. The Model

This study assumes the process is initially in control, and the sample observations come from a normal distribution with a known mean, μ_0 and a known standard deviation, σ_0. However, after an unknown point in time τ, a disturbance is introduced into the process and starting from the

point in time $\tau+1$ (known as the process change point) it changes the process mean from μ_0 to $\mu_1 = \mu_0 + \delta\sigma_0/\sqrt{n}$, where n is the subgroup size and δ is the unknown magnitude of the change. It is also assumed that once the parameter μ_0 changed, it remains at the new level of μ_1 until the root causes of the disturbance have been identified and removed.
Let X_{ij} denote the jth observation in subgroup i with normal distribution N(·,·). That is,

$$X_{ij} \overset{iid}{\sim} N(\mu_0, \sigma_0^2), \quad \begin{array}{l} i = 1,2,\dots,\tau \\ j = 1,2,\dots,n, \end{array}$$

and

$$X_{ij} \overset{iid}{\sim} N(\mu_1, \sigma_0^2), \quad \begin{array}{l} i = \tau+1,\dots,T \\ j = 1,2,\dots,n, \end{array}$$

where n is the subgroup sample size (the size of a sample drawn at time i) and T is the signal time that the subgroup mean exceeds one of the $\bar{X}$ control chart's limits.The notation $\overset{iid}{\sim}$ stands for independent and identically distributed.

In previous studies it is assumed that the value of T is known, but in application one takes a sample at time and investigate weather the process is under control or not. The sampling continues until a signal time happens and process out of control. Therefore, T is not known and random, rationaly.
Also, in past studies, conditional on T, they obtained the MLE for parameters in $\bar{X}$ control chart by, $\hat{\mu}_1 = \bar{\bar{X}}_\tau = (T-\tau)^{-1}\sum_{i=\tau+1}^{T}\bar{X}_i$, the average of the $T-\tau$ most recent subgroups and $\hat{\tau} = arg\max_\tau \left\{ (T-\tau)(\bar{\bar{x}}_\tau - \mu_0)^2 \right\}$, for example see, Dehghan Monfared and Lak (2013). In the next section, using this rational assuming, T is random, we estimate the unknown parameters.

2.1. *Unconditional MLE*

In Section 2, we describe that signal time is random and it is a rational assumption, therefore T is also a random variable. To find its distribution, first we find $\alpha(\mu_1)$, the probability of each subgroup mean being out of control limits, after the process mean changes to μ_1.

$$\begin{aligned}
\alpha(\mu_1) &= P_{\mu_1}\left(\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}} > 3\right) + P_{\mu_1}\left(\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}} < -3\right) \\
&= P_{\mu_1}\left(\frac{\bar{X}-\mu_1+\mu_1-\mu_0}{\sigma_0/\sqrt{n}} > 3\right) + P_{\mu_1}\left(\frac{\bar{X}-\mu_1+\mu_1-\mu_0}{\sigma_0/\sqrt{n}} < -3\right) \\
&= P\left(Z + \frac{\Delta\mu}{\sigma_0/\sqrt{n}} > 3\right) + P\left(Z + \frac{\Delta\mu}{\sigma_0/\sqrt{n}} < -3\right) \\
&= \Phi\left(-3 + \frac{\Delta\mu}{\sigma_0/\sqrt{n}}\right) + \Phi\left(-3 - \frac{\Delta\mu}{\sigma_0/\sqrt{n}}\right).
\end{aligned}$$

Given that a change in the process mean have been occurred at an unknown time $\tau+1$ and using the independence property of the subgroups, the probability mass function of T is,

$$h(t|\tau) = (1-\alpha(\mu_1))^{t-\tau-1}\alpha(\mu_1), \quad t = \tau+1, \tau+2, \cdots.$$

Assuming $\mathbf{x_i} = (x_{i1}, \ldots, x_{in}),\quad i = 1, \ldots, T$, the joint distribution of the observations, i.e. $\mathbf{x}_i$ s and T is,

$$\begin{aligned} f(\mathbf{x}_1, \ldots, \mathbf{x}_t, t|\tau, \mu_1) &= \prod_{i=1}^{\tau}\prod_{j=1}^{n} f(x_{ij}) \prod_{i=\tau+1}^{t}\prod_{j=1}^{n} f(x_{ij}|\mu_1) h(t|\tau) \\ &= (2\pi\sigma_0^2)^{-\frac{nt}{2}} \exp\Bigg\{ -\frac{n}{2\sigma_0^2}\Bigg(\sum_{i=1}^{t} \bar{x}_i^2 - 2\mu_0 \sum_{i=1}^{\tau} \bar{x}_i + \tau\mu_0^2 \\ &- (t-\tau)\bar{\bar{x}}_\tau^2 + (t-\tau)\times(\mu_1 - \bar{\bar{x}}_\tau)^2 \Bigg) \Bigg\} \\ &\times \big(1-\alpha(\mu_1)\big)^{(t-\tau-1)} \alpha(\mu_1), \end{aligned}$$

where in the last expression $\bar{\bar{x}}_\tau = \frac{1}{t-\tau} \sum\limits_{i=\tau+1}^{t} \bar{x}_i$. Throughout the paper, if $\tau = 0$, the values of the product $\prod\limits_{i=1}^{\tau}$ and the sum $\sum\limits_{i=1}^{\tau}$ are defined as 1 and 0, respectively.

To find the MLE of the parameters τ and μ_1 it suffices to maximize the logarithm of the likelihood function with respect to these parameters. The logarithm of the likelihood function (apart from a constant) is

$$\begin{aligned} l(\tau, \mu_1) = \log L(\tau, \mu_1|\mathbf{x}) = -\frac{n}{2\sigma_0^2}\Bigg\{ -2\mu_0 \sum_{i=1}^{\tau} \bar{x}_i + \tau\mu_0^2 - (t-\tau)\bar{\bar{x}}_\tau{}^2 \\ + (t-\tau)(\mu_1 - \bar{\bar{x}}_\tau)^2 \Bigg\} + (t-\tau-1)\log\big(1-\alpha(\mu_1)\big) + \log\big(\alpha(\mu_1)\big). \end{aligned} \tag{2.1}$$

There are two unknown parameters in the log-likelihood function: τ and μ_1. If the change point τ were known, the MLE of μ_1, i.e $\hat{\mu}$, is obtained by maximizing $l(\tau, \mu_1)$ as a function of μ_1. To do this, we set the first partial derivative of $l(\tau, \mu_1)$ with respect to μ_1 equal to zero; that is,

$$\begin{aligned} l'(\tau, \mu_1) = \frac{\partial l(\tau, \mu_1)}{\partial \mu_1} &= -\frac{n}{\sigma_0^2}(t-\tau)(\mu_1 - \bar{\bar{x}}_\tau) - (t-\tau-1)\frac{\alpha'(\mu_1)}{1-\alpha(\mu_1)} \\ &+ \frac{\alpha'(\mu_1)}{\alpha(\mu_1)} = 0. \end{aligned}$$

Solving this equation can't be done analytically, so the Newton Raphson iterative algorithm is used. For our equation the algorithm is as follows,

$$\mu_1^{[i]} = \mu_1^{[i-1]} - \frac{l'(\tau, \mu_1^{[i-1]})}{l''(\tau, \mu_1^{[i-1]})}, \quad i = 1, 2, \cdots.$$

The algorithm starts with arbitrary initial value $\mu_1^{[0]}$ and proceeds until a sufficiently accurate value is reached. Using this method, for each τ, the MLE of the μ_1, shown as $\hat{\mu}_1(\tau)$ is computed for each τ.

Substituting this back into equation (1), we get

$$l^{\star}(\tau) = l(\tau,\hat{\mu}_1(\tau)) = -\frac{n}{2\sigma_0^2}\Bigg\{-2\mu_0\sum_{i=1}^{\tau}\bar{x}_i + \tau\mu_0^2 - (t-\tau)\bar{\bar{x}}_\tau{}^2 \\ + (t-\tau)(\hat{\mu}_1(\tau)-\bar{\bar{x}}_\tau)^2\Bigg\} + (t-\tau-1)\log\big(1-\alpha(\hat{\mu}_1(\tau))\big) + \log\big(\alpha(\hat{\mu}_1(\tau))\big).$$

It then follows that the value of τ that maximizes the log-likelihood function is,

$$\hat{\tau} = arg\max_{\tau}\{l^{\star}(\tau)\}.$$

That is, $\hat{\tau}$ is the value of τ in the range $\tau = 0,1,\ldots,T-1$ which maximizes $l^{\star}(\tau)$. Due to complicated nature of this function, it has to be maximized in τ numerically, as we have done in Section 4.

3. A new Criterion for the efficiency of the MLE of the change point

Using an estimate of the change point at hand, starting from the estimated value, a search for finding the true change point should be initiated. In fact the main purpose of estimating the change point is to determine the time at which the disturbances effect the process. A common criterion for examining estimators is the Mean Square Error (MSE). This criterion is a risk function which is the expectation of quadratic loss function. This loss function is based on the distance between the unknown parameter and its estimate. But, in our problem a more natural loss function is the number of the points in time at which the process should be examined to find the true change point.

The following example shows that it is possible a closer estimate to the true change point may have a weaker performance based on this new loss function.

Example. Suppose that at time $T = 104$ the control chart triggers a signal and the true change point is $\tau+1 = 100$. If we have two estimates $\hat{\tau}_1+1 = 97$ and $\hat{\tau}_2+1 = 104$. Although $\hat{\tau}_1+1$ is closer to $\tau+1 = 100$ than $\hat{\tau}_2+1$, but starting from $\hat{\tau}_1+1 = 97$, we may examine the process at points in time: $97,98,96,99,95,100$ or $97,96,98,95,99,94,100$ to find the true change point. In other words in average 6.5 points should be checked to get the true change point $\tau+1 = 100$. While starting from $\hat{\tau}_2+1 = 104$, we may examine the process at the points in time: $104,103,102,101,100$ to find 100 as the true change point. In this case the process should only be checked at 5 points in time.

To find the true time of the change in the process mean or the change point, two algorithms were proposed by Dehghan Monfared and Lak (2015). The first one, namely Chronological-ordered method (COM), can be applied to any estimate of the change point while the second algorithm called Likelihood-Ordered Method (LOM), can be employed exclusively for MLE. These methods are described briefly in the following subsections.

3.1. *Chronologically-Ordered Method*

The first method is based on the intuitive strategy the more the point is close to the estimate of the change point the more likely to be the true change point. Note that, unless $\hat{\tau}+1 = \tau+1$, the change point is overestimated or underestimated. When it is overestimated ($\hat{\tau}+1 > \tau+1$), to find the true change point, the process should be examined at the points in time: $\hat{\tau}+1, \hat{\tau}, \hat{\tau}-1, \ldots$ until the special cause of the process disturbance is found at $\tau+1$, and in the case of underestimation, this is done by examining the process at the points in time: $\hat{\tau}+1, \hat{\tau}+2, \hat{\tau}+3, \ldots$. But in practice it is unknown that weather the change point is overestimated or underestimated. Thus, a reasonable strategy for finding the true change point would be examination of the process at the points in time: $\hat{\tau}+1, \hat{\tau}, \hat{\tau}+2, \hat{\tau}-1, \ldots$ until the special cause of the process disturbance is found at $\tau+1$.

Note that the points which are the same distance from $\hat{\tau}+1$, like $\hat{\tau}$ and $\hat{\tau}+2$, are exchangeable in this algorithm. In other words, for finding the true change point, we should first examine the process at the estimated change point time, then at points in time which are one time unit away from the estimated change point and so on.

3.2. *Likelihood-Ordered Method*

Another method to find the true change point is solely can be used for MLE of the change point. In this method the points in time are examined in order of their likelihood. That is, a point which has a larger likelihood has more priority to be examined as a possible change point. In this case, the loss function is,

$$L(\hat{\tau}, \tau) = \sum_{i=1}^{T} iI(t_{(i)} = \tau), \tag{3.1}$$

where $t_{(1)}, t_{(2)}, \cdots, t_{(T)}$ are the points $0, 1, \cdots, T-1$, decreasingly ordered by their likelihoods and $I(\cdot)$ is the indicator function, note that $\hat{\tau} = t_{(1)}$.

In fact each of these two methods, when it is used with the MLE, identifies a path which starts from the MLE of the change point and passes through the true change point. Obviously, the sooner a path get us to the true change point the better the path is. Dehghan Monfared and Lak (2015) showed that for all values of δ, LOM has smaller risk than COM. Thus another criteria for the efficiency of the MLE of the change point can be proposed as the risk function of LOM (the average length of the shorter path which is identified by LOM). In the next section this new criteria for the efficiency of the MLE of the change point along with MSE criteria is applied to compare the efficiency of conditional and unconditional MLE of the change point.

4. Comparing Two MLEs

In this section, the efficiency of two MLEs of the change point, i.e. conditional and unconditional MLE are evaluated, through a series of simulations. When the $\bar{X}$ control chart signals which the mean of the process has changed, both of these MLEs are computed. A Monte Carlo simulation study was conducted to study the performance of these two estimators. Suppose $n=4$, sample observations are randomly generated from standard Normal distribution for subgroups $1, 2, \ldots, \tau$. Then, starting with subgroup $\tau+1$, observations were randomly generated from $N(\delta, 1)$ until the $\bar{X}$ control chart triggers a signal. For each of the values of $\tau = 100, 200, 400$ and $\delta = 0.5, 1, 1.5, 2, 2.5, 3$, this procedure is repeated a total of 1000 times. For each simulation run, both estimators and their respective LOM loss functions are computed. The mean of the LOM loss function of each MLE of τ is just the estimated risk function of the MLE of τ. The average values of $\hat{\tau}^{(ML)}$, the MLE of τ and the signal time T, denoted by $\bar{\hat{\tau}}^{(ML)}$ and $\bar{T}$ respectively, along with LOM the estimated risk functions of conditional and unconditional MLEs of τ, are given in table 1.

We consider that for each fixed δ, for $\tau = 100, 200, 400$, while conditional MLE has non constant bias and standard deviation, the bias and standard deviation of unconditional MLE are almost constant. In addition the LOM risk of unconditional MLE, except for small δ, is almost constant but for unconditional MLE is non constant. In addition, based on the MSE (Figure 1):

For $\tau = 100, 200, 400$, conditional MLE outperforms unconditional MLE for small values of δ while for other values conditional MLE is outperforms by unconditional MLE.

Based on LOM the estimated risk function, for $\tau = 100, 200, 400$, unconditional MLE outperforms conditional MLE for all values of δ.

Table 1. The mean of the signal time T, unconditional MLE (UNMLE) and conditional MLE (CONMLE) of the change point along with the estimated risk function of each of two MLEs, shown as R^{un} and R^{con} respectively, and their standard deviations for different values of δ. (subgroup size$n = 4$.)

		δ					
τ		.5	1	1.5	2	2.5	3
100	$\bar{T}$	255.47	144.39	114.98	106.12	103.20	101.98
		(4.87)	(1.34)	(0.46)	(0.18)	(0.08)	(0.04)
	$\bar{\hat{\tau}}^{(UNMLE)}$	114.68	101.98	100.84	100.42	100.18	100.07
		(1.00)	(0.22)	(0.11)	(0.05)	(0.03)	(0.02)
	$\bar{\hat{\tau}}^{(CONMLE)}$	103.8	99.99	99.93	99.92	99.54	99.6
		(0.73)	(0.26)	(0.15)	(0.09)	(0.15)	(0.15)
	R^{un}	17.79	4.91	2.87	1.80	1.42	1.23
		(0.62)	(0.18)	(0.12)	(0.04)	(0.03)	(0.02)
	R^{con}	17.91	5.14	3.01	1.96	1.54	1.34
		(0.67)	(0.23)	(0.16)	(0.1)	(0.08)	(0.05)
200	$\bar{T}$	359.71	244.15	215.07	206.22	203.3	202.03
		(5.18)	(1.34)	(0.46)	(0.18)	(0.09)	(0.05)
	$\bar{\hat{\tau}}^{(UNMLE)}$	215.13	202.08	200.86	200.43	200.15	200.12
		(0.98)	(0.22)	(0.10)	(0.05)	(0.03)	(0.02)
	$\bar{\hat{\tau}}^{(CONMLE)}$	203.35	200.2	199.75	199.83	199.42	199.63
		(0.80)	(0.20)	(0.16)	(0.19)	(0.22)	(0.12)
	R^{un}	19.11	5.03	2.78	1.80	1.47	1.22
		(0.73)	(0.18)	(0.08)	(0.04)	(0.04)	(0.02)
	R^{con}	20.83	5.4	2.93	1.84	1.6	1.33
		(0.89)	(0.26)	(0.14)	(0.05)	(0.08)	(0.05)
400	$\bar{T}$	557.39	441.47	415.32	406.36	403.23	401.97
		(5.23)	(1.24)	(0.45)	(0.18)	(0.08)	(0.04)
	$\bar{\hat{\tau}}^{(UNMLE)}$	416.44	402.66	400.83	400.42	400.17	400.06
		(1.09)	(0.23)	(0.10)	(0.05)	(0.03)	(0.02)
	$\bar{\hat{\tau}}^{(CONMLE)}$	402.11	400.29	399.08	399.75	399.54	398.92
		(1.06)	(0.31)	(0.46)	(0.21)	(0.14)	(0.46)
	R^{un}	21.90	5.57	2.74	1.83	1.44	1.22
		(1.18)	(0.23)	(0.09)	(0.05)	(0.03)	(0.02)
	R^{con}	26.20	6.84	3.46	1.99	1.54	1.61
		(1.59)	(0.56)	(0.4)	(0.10)	(0.06)	(0.22)

5. Conclusion

In this study, two MLEs of the change point, when implementing an $\bar{X}$ control chart, were computed. The first MLE of the change point was computed from the conditional likelihood function of the change point given the signal time T. Almost all the previous papers used this method to

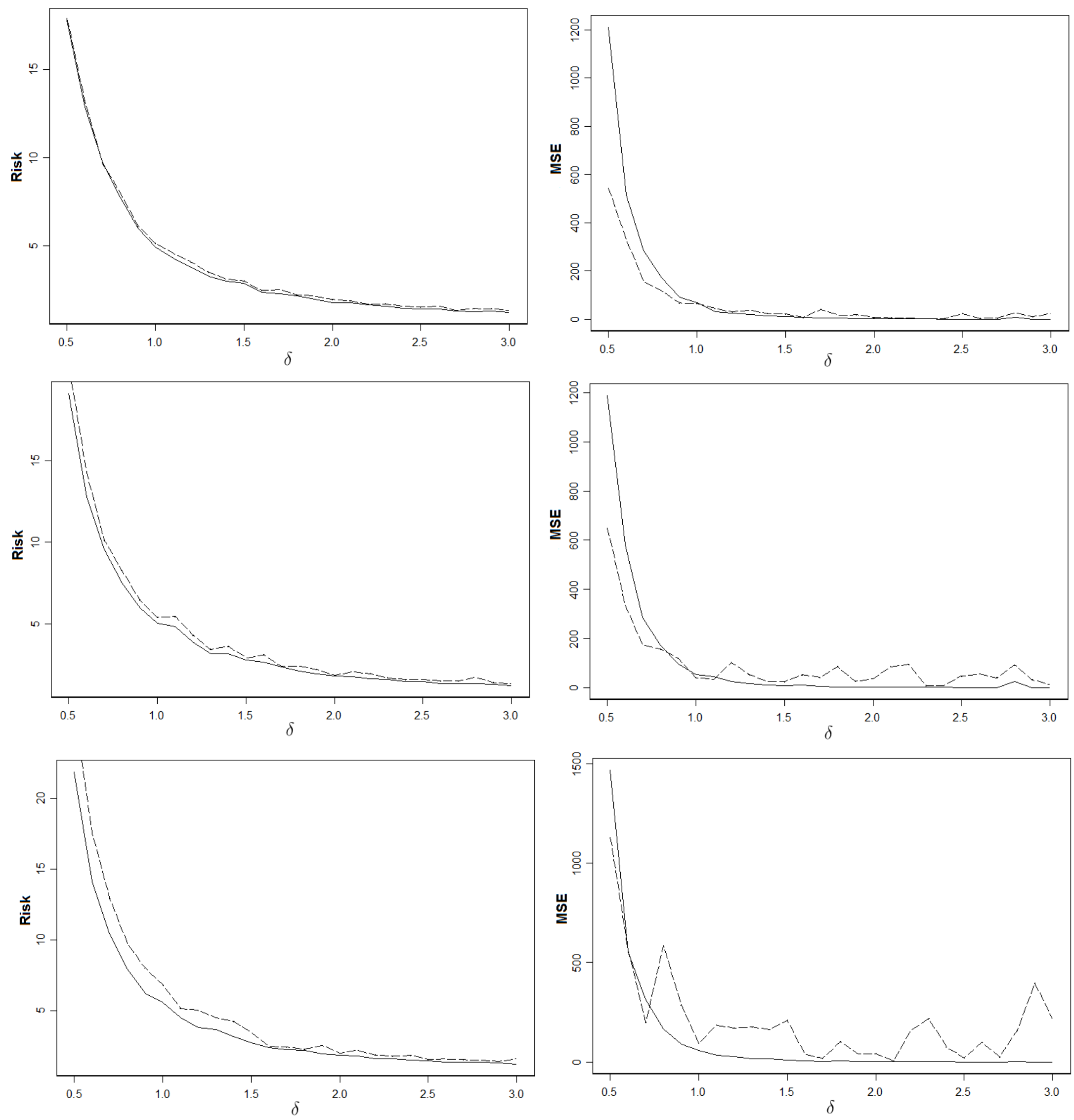

Fig. 1. Left: The estimated risk functions of UNMLE (solid curve) and CONMLE (broken curve). Right: The MSE of UNMLE (solid curve) and CONMLE (broken curve).(From above to the bottom: true change point $\tau = 100, 200, 400$ and subgroup size $n = 4$.)

compute the MLE of the change point. The second MLE of the change point is computed by maximizing complete (unconditional) likelihood function of the change point. It is expected the second MLE of the change point be more efficient than the first one because for computing the second MLE the extra information of the signal time T is also used. But paradoxically, when the MSE criterion is applied to compare the two MLEs, for small δ (shift in the mean of the process) the CONMLE of the change point outperforms the UNMLE of the change point. We introduced and applied a new criterion called LOM risk function, the paradox is resolved and based in this new criterion the unconditional MLE outperforms conditional MLE for all values of δ. Thus, using the MSE criterion in such a problem may be misleading. The method we used here to compare conditional and unconditional MLE estimates of the change point can be used with other control charts. In this paper, we used the $\bar{X}$ control chart to monitor a normal process. There are other types of control charts (for example, CUSUM and EWMA charts) which could be considered and the comparison be done. In addition, the study can be done for non-normal processes like gamma process.

Appendix: Comparing methods of finding the true change point

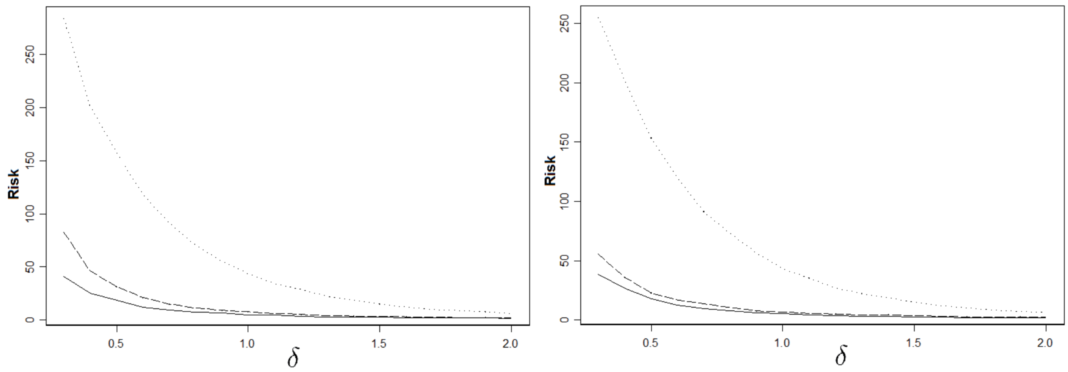

Fig. 2. Left: The estimated risk functions of TM (dotted curve), COM (broken curve) and LOM (solid curve) when unconditional likelihood is used. Right: The estimated risk functions of TM (dotted curve), COM (broken curve) and LOM (solid curve) when conditional likelihood is used. $\tau = 100$ and subgroup size $n = 4$.)

Figure 2 shows a comparison between different methods for finding true change point as it is done by Dehghan Monfared and Lak (2015). As seen in figure 2, for both cases, i.e. conditional and unconditional likelihood, LOM outperforms the two other methods.

Acknowledgments

The authors are grateful to the editor and anonymous referees for their valuable comments which will improve the presentation of this paper .

References

[1] C. S. Cheng, H. P. Cheng, (2008). Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines. *Expert Syst. Appl.* **35**: 198-206.

[2] M. E. Dehghan Monfared and F. Lak (2015), "Finding True Change Point Using $\bar{X}$ Control Chart and MLE", *Submitted*.

[3] M. E. Dehghan Monfared and F. Lak (2013), "Bayesian Estimation of the Change Point Using $\bar{X}$ Control Chart", *Communications in Statistics-Theory and Methods*, **42**: 1572-1582.

[4] C.D. Hou, Y.E. Shao and S. Huang (2013), "A combined MLE and generalized P chart approach to estimate the change point of a multinomial process", *Applied Mathematics & Information Sciences,* vol. 7, no. 4, pp. 1487-1493.

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

[6] W. Levinson (1997), ”Watch out for non normal distributions of impurities,” *Chemical Engineering Progress,* 93(5),70-76.
[7] D. C. Montgomery (2009), *”Introduction to Statistical Quality Control”*, 6th ed., Wiley, New York.
[8] S. Pal(2005), ”Evaluation of nonnormal process capability indices using generalized lambda distribution” *Quality Engineering,* **17(1)**,77-85.
[9] J.J. Pignatiello and T.R. Samuel (2001), ”Estimation of the change point of a normal process mean in SPC applications,” *Journal of Quality Technology,* **33(1)**, 82-95.
[10] T.R. Samuel, J.J. Pignatiello and J.A. Calvin (1998), ”Identifying the time of a step change with $\bar{X}$ control charts,” *Quality Engineering,* **10(3)**,521-527.
[11] Y.E. Shao and C.D. Hou (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(4)**, 349-357.
[12] Y.E. Shao and C.D. Hou (2006), ”Estimation of the starting time of a step change disturbance in a gamma process,” *Journal of the Chinese Institute of Industrial Engineers,* **23(4)**,319-327.
[13] Y.E. Shao, C.D. Hou and H.J. Wang (2006),”Estimation of the change point of a gamma process by using the S control chart and MLE,” *Journal of the Chinese Institute of Industrial Engineers,* **23(3)**,207-214.
[14] Y.E. Shao and C.D. Hou (2013a), ”Fault identification in industrial processes using an integrated approach of neural network and analysis of variance”, *Mathematical Problems in Engineering,* vol. 2013, article ID 516760, pp. 1-7.
[15] Y.E. Shao and C.D. Hou (2013b), ”Change point determination for a multivariate process using a two-stage hybrid scheme”, *Applied Soft Computing,* vol. **13**, no.3 pp. 1520-1527.
[16] Y. E. Shao and B. S. Hsu (2009). ”Determining the contributers for a multivariate SPC chart signal using artificial neural networks and support vector machine”, *Int. J. INNovative Comput. Inform. Cont.* **5(12B)**: 4899-4906 .