Taylor & Francis
Taylor & Francis Group

Sequential Analysis
Design Methods and Applications

ISSN: 0747-4946 (Print) 1532-4176 (Online) Journal homepage: http://www.tandfonline.com/loi/lsqa20

A decision theoretic approach to change point estimation for binomial CUSUM control charts

Takeshi Emura & Yi-Ting Ho

SEQUENTIAL ANALYSIS
2016, VOL. 35, NO. 2, 238–253
http://dx.doi.org/10.1080/07474946.2016.1165543

Taylor & Francis
Taylor & Francis Group

A decision theoretic approach to change point estimation for binomial CUSUM control charts

Takeshi Emura and Yi-Ting Ho

Graduate Institute of Statistics, National Central University, Taoyuan, Taiwan

ABSTRACT
Detecting when the process has changed is a classical problem in sequential analysis and is an important practical issue in statistical process control. This article is concerned about the binomial cumulative sum (CUSUM) control chart, which is extensively applied to industrial process control, health care, public health surveillance, and other fields. For the binomial CUSUM, a maximum likelihood estimator has been proposed to estimate the change point. In our article, following a decision theoretic approach, we develop a new estimator that aims to improve the existing methods. For interval estimation, we propose a parametric bootstrap procedure to construct the confidence set of the change point. We compare our proposed method with the maximum likelihood estimator and Page's last zero estimator in terms of mean squared error by simulations. We find that the proposed method gives more unbiased and robust results than the existing procedures under various parameter designs. We analyze jewelry manufacturing data for illustration.

ARTICLE HISTORY
Received 21 July 2015
Revised 13 December 2015, 14 February 2016
Accepted 5 March 2016

KEYWORDS
Attribute control chart; *np*-chart; parametric bootstrap; sequential analysis; SPRT; statistical process control

MATHEMATICS SUBJECT CLASSIFICATION
62L0; 62C05; 62P30

1. Introduction

Statistical process control (SPC) is a branch of quality control that performs a sequential monitoring of manufactured products using control charts (Montgomery, 2009; Wetherill and Brown, 1991). SPC is applied in order to monitor and control manufacturing, service, and clinical processes, ensuring that they operate under a desired condition. With SPC, the process can safely produce as much conforming product as possible with a minimum of waste. One of the major goals of SPC is to detect the occurrence of irregular operating conditions that make nonconforming products at an unusually high rate.

In SPC, the *np*-chart (or *p*-chart) is one of the most widely used charts applied to monitor the number of nonconforming items for industrial manufactures. As explained by textbooks on SPC (e.g., Montgomery, 2009; Wetherill and Brown, 1991), the *np*-chart sequentially plots the number of nonconforming items over time. The *np*-chart even arises as a binary categorization of any continuous outcome (Yang et al., 2011). If the number of nonconforming items exceeds the control limits (typically 3-sigma limits), the *np*-chart issues a signal that some change occurs in the nonconforming rate. Despite its popularity, some conditions are necessary for the *np*-chart to work properly.

CONTACT T. Emura takeshiemura@gmail.com Graduate Institute of Statistics, National Central University, Jhongda Road, Jhongli City, Taoyuan, Taiwan.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsqa.
Recommended by Nitis Mukhopadhyay

First, the *np*-chart works well only if the sample size n is large enough and nonconforming rate p is not too small (Duran and Albin, 2009; Emura and Lin, 2015; Montgomery, 2009). However, small n is common in biostatistical contexts (e.g., Assareh et al., 2015), and small p is common in industrial practices (e.g., Wang, 2009). Second, the *np*-chart fails to detect a small change in p. The *np*-chart can only effectively detect the out-of-control signals if the magnitude of the change is large.

The binomial cumulative sum (CUSUM) control chart is a good alternative when small but consistent changes in p are expected. This well-known fact will be further explained by our real data analysis. In addition, there appears no requirement for sample size n because it does not rely on the normal approximation. The binomial CUSUM chart plots the cumulative sums of the number of nonconforming items over time.

Nowadays, the binomial CUSUM is extensively applied to a variety of fields, including not only industrial process control but also health care and public health surveillance. Recently, the binomial CUSUM has become increasingly popular in biostatistics. Rossi et al. (2014) and Assareh et al. (2015) used the binomial CUSUM chart in the evaluations of clinical programs. Their charts are called risk-adjusted CUSUM (RACUSUM) and adjust the risk of nonconforming defective rates over time using patients' profiles (Grigg et al., 2003). Some other applications of the binomial CUSUM include detecting the change in sensor networks (Fuh and Mei, 2008).

The CUSUM chart gives a way to estimate the change point of the process as first suggested by Page (1954) for continuous observations. If the CUSUM chart gives an out-of-control signal, then we suspect that there is a change point somewhere before the signal. In the binomial CUSUM chart, Page's change point estimator is defined as the last zero of the chart (Perry and Pignatiello, 2005). Alternatively, Pignatiello and Samuel (2001) and Perry and Pignatiello (2005) proposed a maximum likelihood estimator (MLE) for estimating the change point, which uses data more effectively. Perry and Pignatiello (2005) claimed that the MLE is usually more precise than Page's estimator. However, as we read their simulation results, Page's estimator can still be a better estimator if the design value in the CUSUM is correctly chosen. Perry et al. (2007) extended the MLE approach to more general monotonic change types. Perry and Pignatiello (2008) also developed the MLE under a general exponential family and constructed a confidence set for the change point. Overall, the MLE method to infer the change point is workable in various different models for change types. To the best of our knowledge, no other estimator has been considered beyond the MLE and Page's estimator.

In this article, we propose a new change point estimator that combines the MLE of Pignatiello and Samuel (2001) and Page's estimator. Our estimator is derived as a weighted sum of the two estimators, which is a common approach in statistical decision theory. The resultant estimator aims to utilize the information adaptively from both estimators and to be more unbiased and accurate for the true change point.

The article is organized as follows. Section 2 reviews the background. Section 3 introduces the proposed methods. Section 4 presents the simulation results, and Section 5 provides the real data analysis. Section 6 concludes the article.

2. Background

2.1. Binomial CUSUM control chart

In this article, we assume that the process is in control when independent observations $X_1, X_2, \ldots$ come from a binomial distribution with parameters n and p_0, namely, $Bin(n, p_0)$. The value p_0 is called "in-control fraction nonconforming," which is known a priori. We let $X_1, X_2, \ldots, X_\tau \sim Bin(n, p_0)$ for subgroups $i = 1, 2, \ldots, \tau$, where τ represents the last subgroup taken from the in-control process. Following an unknown subgroup τ, $i = \tau+1, \tau+2, \ldots, T$, where T is the most recent subgroup, the in-control fraction nonconforming p_0 changes to the out-of-control fraction nonconforming $p_a^{True} \neq p_0$, which is unknown. Accordingly, we let $X_{\tau+1}, X_{\tau+2}, \ldots, X_T \sim Bin(n, p_a^{True})$. Therefore,

$$P(X_i = x) = \begin{cases} \binom{n}{x} p_0^x (1-p_0)^{n-x}, & \text{if } i = 1, 2, \ldots, \tau, \\ \binom{n}{x} (p_a^{True})^x (1-p_a^{True})^{n-x}, & \text{if } i = \tau+1, \tau+2, \ldots, T. \end{cases}$$

We focus only on the increasing change $p_a^{True} > p_0$. Hence, the parameter space is

$$\Theta = \{(\tau, p_a^{True}) \mid \tau \in \{1, 2, \ldots, T\}, p_0 \le p_a^{True} \le 1\}$$

for a known value $T \in \{1, 2, \ldots\}$. Typically, T is the subgroup at which a control chart produces an out-of-control signal. The unknown value τ is called change point, which needs to be estimated from observations.

A CUSUM procedure involves a cumulative sum of the deviations of X_is with respect to some reference value (Hawkins and Olwell, 1998). Define the binomial CUSUM recursively by

$$S_0 = 0, \quad S_i = \max\{0, X_i - nk + S_{i-1}\}, \quad i = 1, 2, \ldots,$$

where

$$k = -\log\left(\frac{1-p_a^{as}}{1-p_0}\right) \Big/ \log\left\{\frac{p_a^{as}(1-p_0)}{p_0(1-p_a^{as})}\right\},$$

and where $p_a^{as} > p_0$ is a design value, which is prespecified by engineers. Here, "*as*" stands for "assumed value." The formula of k is derived by using Wald's (1947) sequential probability ratio test for testing

$$H_0: p = p_0 \quad \text{v.s.} \quad H_1: p = p_a^{as}.$$

When S_i exceeds a decision interval $h > 0$, the CUSUM detects a change in the fraction nonconforming. The CUSUM for the decreasing change $p_a^{\mathrm{True}} < p_0$ can be considered similarly as in Perry and Pignatiello (2005).

With this CUSUM scheme, one can define T as the first subgroup that yields the out-of-control signal. Formally, $T = \inf\{i; S_i > h\}$ is called the run length. The average run length (ARL) is the expected value of T, defined as $\mathrm{ARL} = E[T]$. The decision interval h is usually defined such that ARL equals a specified value (e.g., 150 or 370). This is done either by simulations or by utilizing some theoretical analyses on the ARL as previously developed by Khan and Khan (2004) and Khan (2008).

2.2. Maximum likelihood estimator of change point

We introduce the MLE considered by Pignatiello and Samuel (2001) and Perry and Pignatiello (2005) under the binomial CUSUM.

Under the CUSUM scheme of Section 2.1, the likelihood function is given by

$$L(\tau, p_a | p_0) = \prod_{i=1}^{\tau} \binom{n}{x_i} p_0^{x_i}(1-p_0)^{n-x_i} \prod_{i=\tau+1}^{T} \binom{n}{x_i} p_a^{x_i}(1-p_a)^{n-x_i},$$

where $(\tau, p_a) \in \Theta$. The log-likelihood is

$$\begin{aligned}\log L(\tau, p_a|p_0) = C + \log(p_0)\sum_{i=1}^{\tau} x_i + \log(1-p_0)\sum_{i=1}^{\tau}(n-x_i) \\ + \log(p_a)\sum_{i=\tau+1}^{T} x_i + \log(1-p_a)\sum_{i=\tau+1}^{T}(n-x_i),\end{aligned}$$

where C is a constant. As in Perry and Pignatiello (2005), we rewrite

$$\log L(\tau, p_a|p_0) = C^* + \log\left(\frac{p_a}{p_0}\right)\sum_{i=\tau+1}^{T} x_i + \log\left(\frac{1-p_a}{1-p_0}\right)\sum_{i=\tau+1}^{T}(n-x_i),$$

where $C^* = C + \log(p_0)\sum_{i=1}^{T} x_i + \log(1-p_0)\sum_{i=1}^{T}(n-x_i)$ is a constant. If τ is given, the log-likelihood equation of p_a becomes

$$\frac{\partial}{\partial p_a}\log L(\tau, p_a|p_0, \mathrm{X}) = \sum_{i=\tau+1}^{T} x_i \Big/ p_a - \sum_{i=\tau+1}^{T}(n-x_i)\Big/(1-p_a).$$

Setting the above equation to 0, one can easily estimate the value of p_a by

$$\hat{p}_a(\tau) = \sum_{i=\tau+1}^{T} x_i \Big/ \sum_{i=\tau+1}^{T} n.$$

This is indeed the maximum of the log-likelihood because

$$\frac{\partial^2}{\partial(p_a)^2}\log L(\tau, p_a|p_0, \mathrm{X}) = -\sum_{i=\tau+1}^{T} x_i \Big/ p_a^2 - \sum_{i=\tau+1}^{T}(n-x_i)\Big/(1-p_a)^2 < 0, \quad 0 < p_a < 1.$$

Putting $\hat{p}_a(\tau)$ into $\log L(\tau, p_a|p_0)$, one has a profile log-likelihood for τ as

$$\log L(\tau, \hat{p}_a(\tau)|p_0) = C^* + \log\left\{\frac{\hat{p}_a(\tau)}{p_0}\right\}\sum_{i=\tau+1}^{T} x_i + \log\left\{\frac{1-\hat{p}_a(\tau)}{1-p_0}\right\}\sum_{i=\tau+1}^{T}(n-x_i).$$

Therefore, the MLE for the change point becomes

$$\hat{\tau}_{\mathrm{MLE}} = \operatorname*{arg\,max}_{\tau\in\{1,2,\ldots,T\}}\left[\log\left\{\frac{\hat{p}_a(\tau)}{p_0}\right\}\sum_{i=\tau+1}^{T} x_i + \log\left\{\frac{1-\hat{p}_a(\tau)}{1-p_0}\right\}\sum_{i=\tau+1}^{T}(n-x_i)\right].$$

The numerical value of $\hat{\tau}_{\mathrm{MLE}}$ is obtained by plotting the profile log-likelihood for all $\tau \in \{1, 2, \ldots, T\}$ and then finding the maximizing point.

2.3. Page's last zero estimator

The CUSUM control chart has an automatic change point estimator, as suggested by Page (1954). After an out-of-control signal of the binomial CUSUM control chart is found, one can use Page's last zero estimator to estimate the true change point. If the CUSUM control chart signals that an increase in the process fraction nonconforming has occurred, then an estimate of the change point is given by the last zero $\hat{\tau}_{\text{CUSUM}} = \max\{i : S_i = 0\}$. Note that this estimator is biased unless the design value p_a^{as} is correctly specified. Because $p_a^{as} \neq p_a^{\text{True}}$ in general, $\hat{\tau}_{\text{CUSUM}}$ is a biased estimator.

3. Proposed method

3.1. Idea of statistical decision theory

The idea of combing two estimators to reduce the risk or the mean squared error (MSE) is popular in statistical decision theory (Casella and Berger, 2002; Emura et al., 2014; James and Stein, 1961; Khan, 1968; Laheetharan and Wijekoon, 2010; Wencheko and Wijekoon, 2005). The resultant estimator has a weighted sum of two estimators, where the weight is chosen to minimize the MSE. In the Bayesian framework, a frequentist estimator and a prior value are combined to form a Bayes estimator with the weight chosen to minimize the Bayes risk. The idea is also useful to reduce some deficiency of the MLE under high-dimensional regression parameters (Emura et al., 2012).

We demonstrate the idea of statistical decision theory by a well-known example of Khan (1968), which also appears in the textbook of Casella and Berger (2002). Let $X_1, X_2, \ldots, X_n \sim N(\theta, a\theta^2)$, where $\theta > 0$ is unknown and $a > 0$ is known. In this setting, two estimators of θ are available, which are given by

$$d_1 = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

and

$$d_2 = c_n \left\{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right\}^{\frac{1}{2}} = c_n\sqrt{n-1}S_n,$$

where $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ and $c_n = \frac{1}{\sqrt{2a}}\Gamma\left(\frac{n-1}{2}\right)/\Gamma\left(\frac{n}{2}\right)$. The two estimators are unbiased for θ so that $E_\theta(d_1) = E_\theta(d_2) = \theta$. In addition, $\text{Var}_\theta(d_1) = a\theta^2/n$, $\text{Var}_\theta(d_2) = \{c_n^2 a(n-1) - 1\}\theta^2$, and $\text{Cov}_\theta(d_1, d_2) = 0$.

Khan (1968) considered a class of unbiased estimators of θ as

$$d(\alpha) = \alpha d_2 + (1-\alpha)d_1, \quad 0 \le \alpha \le 1,$$

where α is a weight. The optimal weight that minimizes the MSE is shown to be

$$\alpha^* = \frac{\text{Var}_\theta(d_1)}{\text{Var}_\theta(d_1) + \text{Var}_\theta(d_2)} = \frac{a}{a + n\{c_n^2 a(n-1) - 1\}}.$$

In the change point estimation, both $\hat{\tau}_{\text{CUSUM}}$ and $\hat{\tau}_{\text{MLE}}$ are biased for τ. Combination of the two estimators with a weight also leads to a biased estimator of the change point. Nevertheless, if the signs of the biases are different between $\hat{\tau}_{\text{CUSUM}}$ and $\hat{\tau}_{\text{MLE}}$, their combination may reduce the bias. To reduce the bias (and hence the MSE as well) of the change point estimator, a special weight is needed. Hereafter, we will present a new estimator with a special weight.

3.2. Proposed estimator for change point

We propose a new change point estimator that combines $\hat{\tau}_{\text{MLE}}$ (Section 2.2) and $\hat{\tau}_{\text{CUSUM}}$ (Section 2.3) in a decision theoretic manner. Let $p_a^{\text{True}} \geq p_0$ be the true fraction nonconforming, which is unknown. Also, we let $p_a^{as} \geq p_0$ be a design value, which is known for or set by engineers. We propose a change point estimator

$$\hat{\tau}_{\text{NEW}}(w) = w\hat{\tau}_{\text{CUSUM}} + (1-w)\hat{\tau}_{MLE}, \quad 0 \leq w \leq 1,$$

where w is a weight.

In order to reduce the MSE and the bias of the proposed estimator, we require that the weight w satisfies certain conditions. If p_a^{True}=p_a^{as}, then one should set $w = 1$, which implies that $\hat{\tau}_{\text{NEW}}(w) = \hat{\tau}_{\text{CUSUM}}$. This is because $\hat{\tau}_{\text{CUSUM}}$ is much more accurate than $\hat{\tau}_{\text{MLE}}$ under $p_a^{\text{True}} = p_a^{as}$, which can be observed from Perry and Pignatiello (2005) and our simulation results. If p_a^{True} is far from p_a^{as}, then $\hat{\tau}_{\text{MLE}}$ is more precise than $\hat{\tau}_{\text{CUSUM}}$ because the latter is highly biased.

With the above considerations, we propose the following weight function:

$$w \equiv w(p_a^{\text{True}} |\, p_a^{as}) = \begin{cases} \left(\dfrac{p_a^{\text{True}} - p_0}{p_a^{as} - p_0}\right)^{\left(\frac{p_a^{True}}{p_0}\right)}, & \text{if } p_a^{\text{True}} \leq p_a^{as}, \\ \left(\dfrac{p_a^{as} - p_0}{p_a^{\text{True}} - p_0}\right)^{\left(\frac{p_a^{True}}{p_0}\right)}, & \text{if } p_a^{\text{True}} > p_a^{as}. \end{cases}$$

The value of w decreases from 1 to 0 when p_a^{True} deviates from p_a^{as} (Figure 1). In addition, $w = 1$ if and only if $p_a^{\text{True}} = p_a^{as}$. Hence, the weight satisfies our requirements.

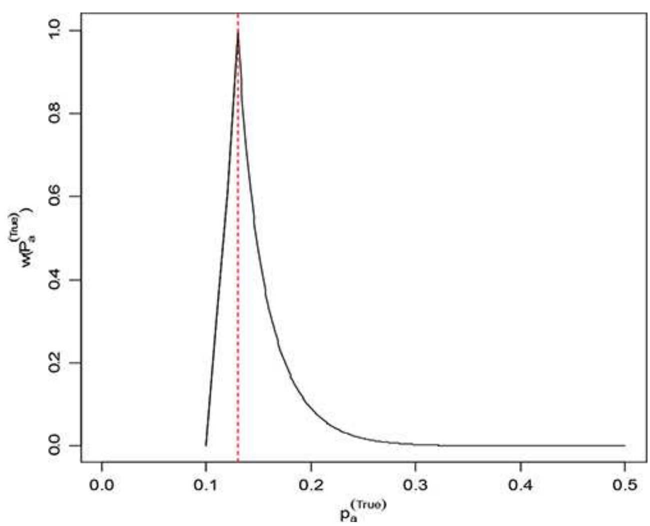

Figure 1. Plot of the weight function $w(p_a^{\text{True}}|p_a^{as} = 0.13)$ under $p_0 = 0.1$.

The preceding weight w involves an unknown parameter p_a^{True}, which needs to be estimated in practical applications. We propose to use the MLE given by $\hat{p}_a(\hat{\tau}_{MLE})$, where $\hat{p}_a(\tau)$ is defined in Section 2.2. Another choice is $\hat{p}_a(\hat{\tau}_{\text{CUSUM}})$. Our numerical results reveal that $\hat{p}_a(\hat{\tau}_{\text{MLE}})$ is more accurate than $\hat{p}_a(\hat{\tau}_{\text{CUSUM}})$ (The results are not shown but are available upon request.) Hence, we propose the following estimator of the weight

$$\hat{w} \equiv w(\hat{p}_a(\hat{\tau}_{\text{MLE}})|p_a^{as}) = \begin{cases} \left(\frac{\hat{p}_a(\hat{\tau}_{\text{MLE}})-p_0}{p_a^{as}-p_0}\right)^{\left(\frac{\hat{p}_a(\hat{\tau}_{\text{MLE}})}{p_0}\right)}, & \text{if } \hat{p}_a(\hat{\tau}_{\text{MLE}}) \leq p_a^{as}, \\ \left(\frac{p_a^{as}-p_0}{\hat{p}_a(\hat{\tau}_{\text{MLE}})-p_0}\right)^{\left(\frac{\hat{p}_a(\hat{\tau}_{\text{MLE}})}{p_0}\right)}, & \text{if } \hat{p}_a(\hat{\tau}_{\text{MLE}}) > p_a^{as}. \end{cases}$$

Accordingly, we define $\hat{\tau}_{\text{NEW}}(\hat{w})$, the proposed estimator with the weight being estimated.

Once w is estimated by $\hat{w}$ in the preceding formula, $\hat{w} < 0$ occurs by chance. In particular, if $\hat{p}_a(\hat{\tau}_{\text{MLE}}) < p_0$, the weight value will be negative. In this case, we set $\hat{w} = 0$. This is a natural restriction to meet $0 \leq w \leq 1$. However, the occurrence is very minor because it is found in only a few cases during a large number of simulation repetitions.

3.3. Confidence set for change point

A confidence set for the change point is the random set of τs that contains the true change point with a specified probability, say $1-\alpha$. Though the MLE $\hat{\tau}_{\text{MLE}}$ allows the use of the profile likelihood to obtain the confidence set (Perry and Pignatiello, 2008), this approach cannot be used to the proposed estimator $\hat{\tau}_{\text{NEW}}(\hat{w})$.

We propose the percentile confidence set with the parametric bootstrap (Efron and Tibshirani, 1993). Recall that the estimator $\hat{\tau}_{\text{NEW}}(\hat{w})$ is calculated from in-control data $X_1, X_2, \ldots, X_\tau \sim Bin(n, p_0)$ and out-of-control data $X_{\tau+1}, X_{\tau+2}, \ldots, X_T \sim Bin(n, p_a^{\text{True}})$, where $T = \inf\{i; S_i > h\}$ and S_i is the CUSUM based on X_is. The bootstrap algorithm of obtaining the confidence set (interval) is stated as follows:

Confidence set (interval) for τ with confidence level $(1 - \alpha) \times 100\%$:

Step 0: Estimate the population parameters (τ, p_a) by $(\hat{\tau}_{\text{NEW}}(\hat{w}), \hat{p}_a(\hat{\tau}_{\text{MLE}}))$.

Step 1: Generate bootstrap in-control data $X_1^*, X_2^*, \ldots, X_{\hat{\tau}_{NEW}(\hat{w})}^* \sim Bin(n, p_0)$ and bootstrap out-of-control data $X_{\hat{\tau}_{NEW}(\hat{w})+1}^*, X_{\hat{\tau}_{NEW}(\hat{w})+2}^*, \ldots, X_{T*}^* \sim Bin(n, \hat{p}_a(\hat{\tau}_{MLE}))$, where $T^* = \inf\{\, i;\ S_i^* > h \,\}$ and S_i^* is the bootstrap CUSUM based on X_i^*s.

Step 2: Obtain the bootstrap estimate $\hat{\tau}_{\text{NEW}}^*(\hat{w}^*)$ using the bootstrap data in Step 1.

Step 3: Repeat Steps 1–2 to obtain the bootstrap estimates $\hat{\tau}_{\text{NEW}}^*(\hat{w}^*)^{(b)}$, $b = 1, 2, \ldots, B$, where B is the large number of bootstraps replications.

Step 4: The bootstrap confidence interval consists of the lower percentile $[B \times \alpha/2$th value in the ordered list of $\hat{\tau}_{\text{NEW}}^*(\hat{w}^*)^{(b)}, b = 1, 2, \ldots, B]$ and the upper percentile $[B \times (1 - \alpha/2)$th value in the ordered list of $\hat{\tau}_{\text{NEW}}^*(\hat{w}^*)^{(b)}, b = 1, 2, \ldots, B]$.

In Step 1, the number of data in the bootstrap samples vary with the bootstrap replication (b) and hence one can write $T^{*(b)}, b = 1, 2, \ldots, B$. This is different from the usual bootstrap samples for independent and identically distributed data. The distinction is necessary because each bootstrap data must issue the out-of-control signal at time $T^{*(b)}, b = 1, 2, \ldots, B$, as in the original data.

In Step 3, using $B \geq 1,000$ would usually be suggested by Efron and Tibshirani (1993). We use $B = 1,000$ for our data analysis, but in our simulations, we use $B = 500$ to save the computation time.

In Step 4, we define the confidence set as the percentile interval. We have checked in simulations that the usual properties of the bootstrap samples are met; the average of $\hat{\tau}^*_{\text{NEW}}(\hat{w}^*_{\text{NEW}})^{(b)}$, $b = 1, 2, \ldots, B$, is always very close to the original estimate $\hat{\tau}_{\text{NEW}}(\hat{w})$; the percentile interval always (i.e., with probability one) covers $\hat{\tau}_{\text{NEW}}(\hat{w})$. By its definition, it is required that the confidence interval covers the true τ with probability $1-\alpha$. This requirement is met only when the estimate $\hat{\tau}_{\text{NEW}}(\hat{w})$ is a reasonably good estimate of the true τ. We will demonstrate the performance of the coverage probability by simulations.

4. Simulations

We use Monte Carlo simulations to compare the change point estimators $\hat{\tau}_{\text{CUSUM}}$, $\hat{\tau}_{\text{MLE}}$, $\hat{\tau}_{\text{NEW}}(w)$, and $\hat{\tau}_{\text{NEW}}(\hat{w})$ in term of the unbiasedness and accuracy. To do a fair comparison, we follow the published design of Perry and Pignatiello (2005).

4.1. Simulation designs

We set the true change point $\tau = 100$ and in-control fraction nonconforming $p_0 = 0.1$. Observations $X_1, X_2, \ldots, X_{100}$ follow a binomial distribution with $p_0 = 0.1$ and $n = 50$ for the first $\tau = 100$ subgroups. After subgroup 100, observations $X_{101}, X_{102}, \ldots$ follow a binomial distribution with $p_a^{\text{True}} \in \{0.11, 0.12, \ldots, 0.25, 0.30\}$ and $n = 50$ until the CUSUM chart issues a signal. We design the CUSUM charts to detect a 30% increase in p_a^{True} by setting $p_a^{as} = 1.3 \times p_0 = 0.13$. Then, the reference value of X_i is nk, where

$$k = -\ln\left(\frac{1-0.13}{1-0.1}\right) \Big/ \ln\left\{\frac{0.13(1-0.1)}{0.1(1-0.13)}\right\} = 0.1144295.$$

If the chart yields a signal at $i \leq \tau$, we restart by setting $S_i \equiv 0$. Then, $T > \tau$ is the subgroup at which the chart issues a signal. We choose the decision interval $h = 6.57$ or 11.42 such that the in-control ARL is 150 or 370, respectively. We do 1,000 simulations to run the CUSUM charts and calculate estimators. Then, we evaluate the performance of the new estimator in terms of the unbiasedness and the MSE defined as

$$\text{MSE}\{\hat{\tau}_{\text{NEW}}(\hat{w})\} = E\{\hat{\tau}_{\text{NEW}}(\hat{w}) - \tau\}^2.$$

4.2. Simulation results

We first examine the unbiasedness of the new estimators. Table 1 compares $E(\hat{\tau}_{\text{MLE}})$, $E(\hat{\tau}_{\text{CUSUM}})$, $E\{\hat{\tau}_{\text{NEW}}(w)\}$, and $E\{\hat{\tau}_{\text{NEW}}(\hat{w})\}$, where w is known and $\hat{w}$ is estimated. Table 1 shows that $E(\hat{\tau}_{\text{CUSUM}})$ is clearly closer to the true $\tau = 100$ than $E(\hat{\tau}_{\text{MLE}})$ when $p_a^{\text{True}} = p_a^{as} = 0.13$ (i.e., the design value is correctly specified). When $p_a^{\text{True}} > p_a^{as} = 0.13$, we see that $E(\hat{\tau}_{\text{CUSUM}})$ underestimates and $E(\hat{\tau}_{\text{MLE}})$ slightly overestimates the true $\tau = 100$. The proposed estimator $\hat{\tau}_{\text{NEW}}(w)$ is an intermediate between $\hat{\tau}_{\text{CUSUM}}$ and $\hat{\tau}_{\text{MLE}}$, and produces most unbiased results in majority of cases. Note that the performance of $\hat{\tau}_{\text{NEW}}(\hat{w})$ depends on how well w is estimated by $\hat{w}$. Table 1 reveals that the performance of $\hat{\tau}_{\text{NEW}}(\hat{w})$ is fairly competitive to that of $\hat{\tau}_{\text{NEW}}(w)$. This implies that the effect of estimating w is modest.

Table 2 gives the performances of $\hat{p}_a^{\text{True}}$ for estimating p_a^{True}. This is of interest because the performance of $\hat{\tau}_{\text{NEW}}(\hat{w})$ depends on the accuracy of $\hat{p}_a^{\text{True}}$. It is seen that $\hat{p}_a^{\text{True}}$ is slightly

Table 1. Simulation results for estimating the change point $\tau = 100$ under $n = 50$, $p_0 = 0.1$, and $p_a^{as} = 0.13$ based on 1,000 simulation runs[a].

$h = 6.57$	p_a^{True}	$\hat{ARL}$	$E(\hat{\tau}_{MLE})$	$E(\hat{\tau}_{CUSUM})$	$E\{\hat{\tau}_{NEW}(w)\}$	$E\{\hat{\tau}_{NEW}(\hat{w})\}$
	0.11	121.75	118.90	115.98	118.03	117.61
	0.12	111.50	108.50	105.51	106.66	107.32
	$0.13(=p_a^{as})$	107.13	104.10	101.33	101.33	103.08
	0.14	105.16	102.53	99.84	100.73	101.70
	0.15	104.14	101.54	99.11	100.41	100.87
	0.16	103.34	101.13	98.78	100.35	100.52
	0.17	102.81	100.63	98.47	100.11	100.17
	0.18	102.50	100.55	98.54	100.21	100.15
	0.19	102.23	100.16	98.45	99.95	99.88
	0.20	101.95	100.26	98.42	100.09	99.96
	0.25	101.42	100.11	98.34	100.08	99.88
	0.30	101.17	100.07	98.44	100.07	99.91
$h = 11.42$	p_a^{True}	$\hat{ARL}$	$E(\hat{\tau}_{MLE})$	$E(\hat{\tau}_{CUSUM})$	$E\{\hat{\tau}_{NEW}(w)\}$	$E\{\hat{\tau}_{NEW}(\hat{w})\}$
	0.11	160.99	153.07	147.96	151.54	150.93
	0.12	121.98	115.53	108.84	111.42	113.33
	$0.13(=p_a^{as})$	112.65	107.14	100.97	100.97	105.20
	0.14	108.48	103.79	98.63	100.34	102.25
	0.15	106.47	102.63	97.93	100.45	101.38
	0.16	105.03	101.77	97.75	100.44	100.81
	0.17	104.28	101.50	97.23	100.49	100.64
	0.18	103.76	101.13	97.26	100.46	100.44
	0.19	103.33	100.88	97.19	100.42	100.22
	0.20	102.94	100.62	97.02	100.30	100.11
	0.25	102.05	100.30	97.24	100.25	100.03
	0.30	101.63	100.14	97.30	100.13	99.98

[a] $\hat{ARL}$ = Average run length based on 1000 runs. The decision interval $h = 6.57$ or 11.42 is chosen such that the in-control ARL is 150 or 370, respectively.

Table 2. Simulation results for estimating p_a^{True} under $n = 50$, $p_0 = 0.1$, and $p_a^{as} = 0.13$ based on 1,000 simulation runs.[a]

$h = 6.57$	p_a^{True}	$E(\hat{p}_a^{True})$	$MSE(\hat{p}_a^{True})$
	0.11	0.153	0.0022
	0.12	0.155	0.0016
	$0.13(=p_a^{as})$	0.156	0.0011
	0.14	0.159	0.0008
	0.15	0.163	0.0007
	0.16	0.168	0.0007
	0.17	0.170	0.0006
	0.18	0.174	0.0007
	0.19	0.177	0.0010
	0.20	0.182	0.0011
	0.25	0.198	0.0037
	0.30	0.212	0.0090
$h = 11.42$	p_a^{True}	$E(\hat{p}_a^{True})$	$MSE(\hat{p}_a^{True})$
	0.11	0.154	0.0023
	0.12	0.155	0.0017
	$0.13(=p_a^{as})$	0.159	0.0014
	0.14	0.162	0.0010
	0.15	0.167	0.0009
	0.16	0.173	0.0008
	0.17	0.179	0.0007
	0.18	0.183	0.0007
	0.19	0.187	0.0008
	0.20	0.192	0.0008
	0.25	0.216	0.0023
	0.30	0.234	0.0056

[a] The decision interval $h = 6.57$ or 11.42 is chosen such that the in-control ARL is 150 or 370, respectively.

Table 3. Simulation results for the MSE of the change point $\tau = 100$ under $n = 50$, $p_0 = 0.1$, and $p_a^{as} = 0.13$ based on 1,000 simulation runs.[a]

$h = 6.57$	p_a^{True}	MSE($\hat{\tau}_{\text{MLE}}$)	MSE($\hat{\tau}_{\text{CUSUM}}$)	MSE$\{\hat{\tau}_{\text{NEW}}(w)\}$	MSE$\{\hat{\tau}_{\text{NEW}}(\hat{w})\}$
	0.11	814.11	669.15	759.25	735.72
	0.12	214.27	127.46	145.85	158.91
	$0.13 (= p_a^{as})$	96.686	30.862	30.862	50.698
	0.14	48.880	16.757	16.976	22.809
	0.15	43.172	10.166	15.404	16.536
	0.16	26.780	9.2780	13.807	10.237
	0.17	17.391	11.063	10.945	6.4579
	0.18	10.060	9.7160	7.2139	3.8668
	0.19	24.216	9.3510	18.894	7.1571
	0.20	9.6440	9.4110	8.0442	2.8040
	0.25	2.2600	8.2400	2.1866	1.0691
	0.30	0.2660	8.1890	0.2642	0.3753
$h = 11.42$	p_a^{True}	MSE($\hat{\tau}_{\text{MLE}}$)	MSE($\hat{\tau}_{\text{CUSUM}}$)	MSE$\{\hat{\tau}_{\text{NEW}}(w)\}$	MSE$\{\hat{\tau}_{\text{NEW}}(\hat{w})\}$
	0.11	6,119.14	5,743.80	5,835.6	5,758.0
	0.12	639.60	374.17	423.99	495.05
	$0.13 (= p_a^{as})$	226.71	60.876	60.876	122.26
	0.14	108.40	36.832	32.666	49.398
	0.15	53.433	31.370	24.371	28.747
	0.16	20.426	24.352	11.983	13.635
	0.17	12.091	30.333	7.8838	8.5666
	0.18	12.338	27.753	8.9755	6.8706
	0.19	4.5920	30.488	3.6264	4.4526
	0.20	5.6930	29.692	4.9255	4.3549
	0.25	0.9280	26.806	0.8860	1.5221
	0.30	0.4600	27.475	0.4562	0.6258

[a] The decision interval $h = 6.57$ or 11.42 is chosen such that the in-control ARL is 150 or 370, respectively.

biased for estimating p_a^{True}. However, the MSE of $\hat{p}_a^{\text{True}}$ is still close to zero and $\hat{p}_a^{\text{True}}$ captures the magnitude of p_a^{True} well.

Table 3 compares the precision of the new estimators in terms of MSE. Table 3 indicates that $\hat{\tau}_{\text{CUSUM}}$ provides the most precise estimator when the true value of the out-of-control process fraction nonconforming is near the specified value for which the CUSUM was designed (i.e., $p_a^{\text{True}} \approx p_a^{as} = 0.13$). However, we observe poor performance of all the estimators when $p_a^{\text{True}} < p_a^{as} = 0.13$. This is because the CUSUM chart cannot immediately detect the process change because p_a^{True} is close to p_0. Although some of the MSEs of $\hat{\tau}_{\text{NEW}}(\hat{w})$ are not the smallest, $\hat{\tau}_{\text{NEW}}(\hat{w})$ provides very precise estimator of the change point under various different values of p_a^{True}.

Table 4 shows the performance of the 95% bootstrap confidence interval. The results show some undercoverage with the settings of $p_a^{as} \le 0.14$ as there is systematic bias of $\hat{\tau}_{\text{NEW}}(\hat{w})$ for the true $\tau = 100$. In the worst case of $p_a^{as} = 0.11$, the average lower confidence interval exceeds the true $\tau = 100$, which leads to gross undercoverage. This problem is due to the lack of information to estimate the true change point (as mentioned in the previous paragraph) but is not due to the performance of the bootstrap. On the other hand, with the settings of $p_a^{as} \ge 0.17$, the coverage probabilities generally agree with the nominal 95% level, or slight overcoverage. In these settings, the lower and upper confidence limits successfully bracket the true change point $\tau = 100$.

Table 4. Simulation results for the 95% confidence interval of the change point $\tau = 100$ under $n = 50$, $p_0 = 0.1$, and $p_a^{as} = 0.13$ based on 1,000 simulation runs.[a]

$h = 6.57$	p_a^{True}	95% Lower confidence limit	95% Upper confidence limit	Coverage probability	$E\{\hat{\tau}_{\text{NEW}}(\hat{w})\}$
	0.11	111.77	125.18	0.34	117.61
	0.12	100.87	115.21	0.55	107.32
	$0.13(= p_a^{as})$	97.62	111.08	0.72	103.08
	0.14	96.04	108.48	0.88	101.70
	0.15	95.63	107.36	0.91	100.87
	0.16	95.17	106.14	0.96	100.52
	0.17	95.23	105.28	0.98	100.17
	0.18	95.15	104.86	0.99	100.15
	0.19	95.31	104.37	0.98	99.88
	0.20	95.22	104.19	0.99	99.96
	0.25	95.70	102.59	0.97	99.88
	0.30	96.05	101.97	0.91	99.91
$h = 11.42$	p_a^{True}	95% Lower confidence limit	95% Upper confidence limit	Coverage probability	$E\{\hat{\tau}_{\text{NEW}}(\hat{w})\}$
	0.11	141.99	163.84	0.18	150.93
	0.12	106.49	126.72	0.45	113.33
	$0.13(= p_a^{as})$	98.06	116.79	0.63	105.20
	0.14	96.21	112.19	0.77	102.25
	0.15	94.87	110.03	0.84	101.38
	0.16	94.54	108.25	0.88	100.81
	0.17	94.42	106.70	0.94	100.64
	0.18	94.37	105.88	0.97	100.44
	0.19	94.26	104.90	0.98	100.22
	0.20	94.61	104.59	0.98	100.11
	0.25	95.52	102.77	0.99	100.03
	0.30	96.06	102.08	0.98	99.98

[a] The decision interval $h = 6.57$ or 11.42 is chosen such that the in-control ARL is 150 or 370, respectively.

5. Data analysis

We analyze the jewelry manufacturing data obtained from Burr (1979). This data set counts the number of defective pieces in the jewelry manufacturing process. Each subgroup contains $n = 50$ beads, so we let

$$X_i = \text{the number of defective pieces in 50 beads, for } i = 1, 2, \ldots, T,$$

where the number of subgroups $T = 54$ is chosen as the first subgroup that yields an out-of-control signal under the setting of Burr (1979). By counting the observed number of defectives, we set the in-control fraction nonconforming $p_0 = 229/2,700 = 0.085$.

First, we use the *np*-chart that plots the number of defective pieces in each subgroup (Figure 2). We see that the *np*-chart shows a consistently increasing trend after the subgroup 40. However, all subgroups are still within the 3σ-control limits, so the process is declared to be in control. This is a typical example where the CUSUM is more appropriate than the *np*-chart to detect the small signals.

We compute the binomal CUSUM designed to detect a 30% increase in p_a^{True} by setting $p_a^{as} = 1.3 \times p_0 = 1.3 \times 0.085 = 0.11$, and

$$k = -\ln\left(\frac{1 - 0.11}{1 - 0.085}\right) \Big/ \ln\left\{\frac{0.11(1 - 0.085)}{0.085(1 - 0.11)}\right\} = 0.097.$$

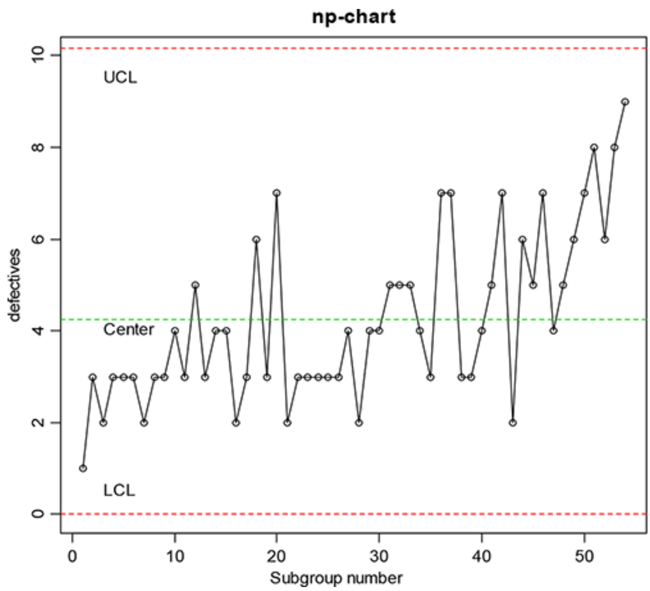

Figure 2. *np*-Chart for 54 subgroups of $n = 50$ pieces with 3σ limits. Here, Center = 4.25, lower control limit (LCL) = 0, and upper control limit (UCL) = 10.165.

Table 5. CUSUM calculated on the 54 subgroups for the jewelry manufacturing data from Burr (1979).[a]

Subgroup number	Defectives X_i	Fraction defective	S_i	Subgroup number	Defectives X_i	Fraction defective	S_i
1	1	0.02	0	28	2	0.04	0
2	3	0.06	0	29	4	0.08	0
3	2	0.04	0	30	4	0.08	0
4	3	0.06	0	31	5	0.10	0.1489
5	3	0.06	0	32	5	0.10	0.2978
6	3	0.06	0	33	5	0.10	0.4468
7	2	0.04	0	34	4	0.08	0
8	3	0.06	0	35	3	0.06	0
9	3	0.06	0	36	7	0.14	2.1489
10	4	0.08	0	37	7	0.14	4.2978
11	3	0.06	0	38	3	0.06	2.4468
12	5	0.10	0.1489	39	3	0.06	0.5957
13	3	0.06	0	40	4	0.08	0
14	4	0.08	0	41	5	0.10	0.1489
15	4	0.08	0	42	7	0.14	2.2978
16	2	0.04	0	43	2	0.04	0
17	3	0.06	0	44	6	0.12	1.1489
18	6	0.12	1.1489	45	5	0.10	1.2978
19	3	0.06	0	46	7	0.14	3.4468
20	7	0.14	2.1489	47	4	0.08	2.5957
21	2	0.04	0	48	5	0.10	2.7447
22	3	0.06	0	49	6	0.12	3.8936
23	3	0.06	0	50	7	0.14	6.0426
24	3	0.06	0	51	8	0.16	9.1915
25	3	0.06	0	52	6	0.12	10.340
26	3	0.06	0	53	8	0.16	13.489
27	4	0.08	0	54	9	0.18	17.638

[a] The fraction nonconforming is $p_0 = 229/2700 = 0.085$. Each subgroup contains $n = 50$ beads.

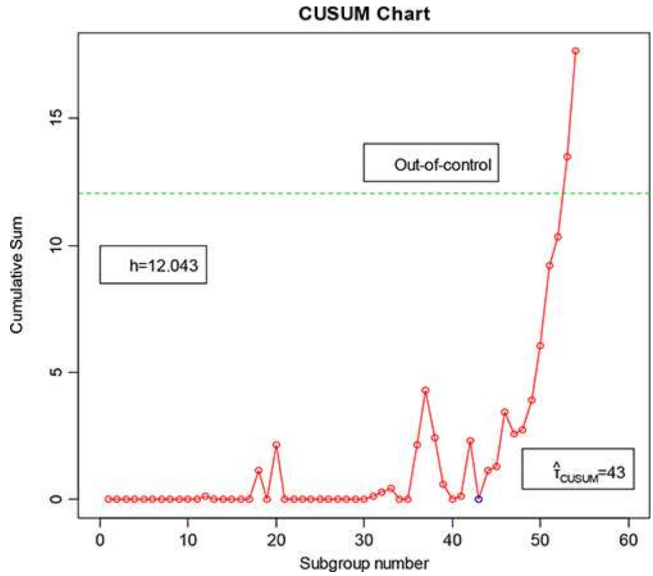

Figure 3. Binomial CUSUM chart from the jewelry manufacturing data for $p_0 = 0.085$, $h = 12.043$, and $p_a^{as} = 0.11$.

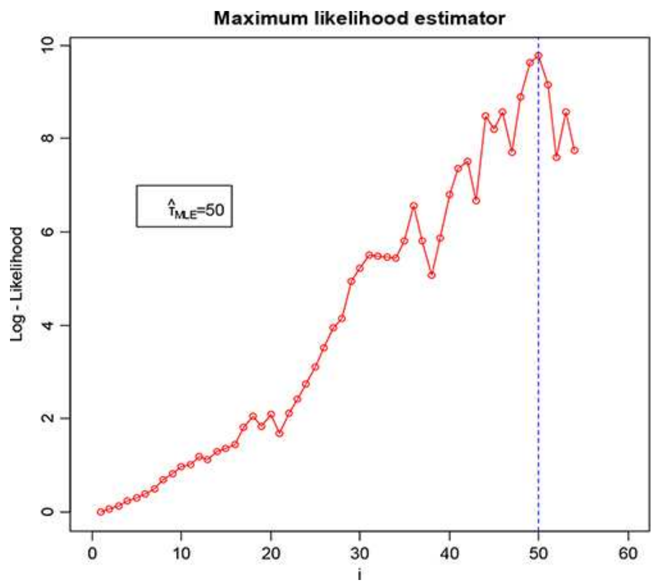

Figure 4. Results of maximum likelihood estimation from the jewelry manufacturing data for $p_0 = 0.085$, $h = 12.043$, and $p_a^{as} = 0.11$.

Then, the reference value become $nk = 50 \times 0.097 = 4.85$. The CUSUM is computed recursively. For instance,

$$S_1 = \max\{0, X_1 - nk + S_0\} = \max\{0, 1 - 4.85 + 0\} = \max\{0, -3.85\} = 0,$$

$$S_{49} = \max\{0, X_{49} - nk + S_{48}\} = \max\{0, 6 - 4.85 + 2.74\} = \max\{0, 3.89\} = 3.89.$$

All values of S_i are given in Table 5. We choose $h = 12.043$ such that the in-control ARL is close to 370. Then, the CUSUM is comparable to the np-chart.

Figure 3 shows that the out-of-control in the CUSUM is found as $S_{53} = 13.4894 > h$. Thus, the fraction nonconforming is changed. Figure 3 displays Page's change point estimator at $\hat{\tau}_{\text{CUSUM}} = 43$. Figure 4 shows that the MLE of the change point is $\hat{\tau}_{\text{MLE}} = 50$.

To obtain the proposed estimator $\hat{\tau}_{\text{NEW}}(\hat{w})$, we first need to estimate the weight $\hat{w} = w(\hat{p}_a(\hat{\tau}_{\text{MLE}})|p_a^{as})$. Because $\hat{\tau}_{\text{MLE}} = 50$,

$$\hat{p}_a(\hat{\tau}_{\text{MLE}}) = \frac{x_{51} + x_{52} + x_{53} + x_{54}}{50 + 50 + 50 + 50} = \frac{8 + 6 + 8 + 9}{200} = 0.155.$$

Because $\hat{p}_a(\hat{\tau}_{\text{MLE}}) > p_a^{as}$,

$$\hat{w} = \left(\frac{0.11 - 0.085}{0.155 - 0.085}\right)^{\left(\frac{0.155}{0.085}\right)} = 0.1529.$$

Therefore, the proposed change point estimator is

$$\begin{aligned}\hat{\tau}_{\text{NEW}}(w) &= w\hat{\tau}_{\text{CUSUM}} + (1 - w)\hat{\tau}_{\text{MLE}} \\ &= 0.1529 \times 43 + 0.8470 \times 50 \\ &= 48.9292.\end{aligned}$$

We know from the simulations that $\hat{\tau}_{\text{CUSUM}}$ typically underestimates whereas $\hat{\tau}_{\text{MLE}}$ slightly overestimates the true change point. This tendency is also suspected in the data analysis because $\hat{\tau}_{\text{CUSUM}} = 43 < \hat{\tau}_{\text{MLE}} = 50$. The results indicate that our proposed estimator may provide a more unbiased estimator by taking an intermediate between $\hat{\tau}_{\text{CUSUM}}$ and $\hat{\tau}_{\text{MLE}}$.

The change point estimate $\hat{\tau}_{\text{NEW}}(w) = 48.9292$ is better supplemented by interval estimates. We perform the parametric bootstrap (Section 3.3) and obtained the 95% confidence interval as [41.0113, 53.7573]. This suggests that, with 95% confidence, the true change point is one of {41, 42, 43, …, 51, 52, 53}.

6. Conclusion

This article proposes a new change point estimator $\hat{\tau}_{\text{NEW}}(\hat{w})$ under the binomial CUSUM control chart. Our estimator is not totally new but is a sensible combination of Page's estimator $\hat{\tau}_{\text{CUSUM}}$ (Page, 1954) and the maximum likelihood estimator $\hat{\tau}_{\text{MLE}}$ (Perry and Pignatiello, 2005; Pignatiello and Samuel, 2001). We derive our estimator such that the following principle holds: If the binomial CUSUM chart is correctly designed (i.e., $p_a^{\text{True}} = p_a^{as}$), one should use $\hat{\tau}_{\text{NEW}}(\hat{w}) = \hat{\tau}_{\text{CUSUM}}$. In this case, the chart can immediately detect the change and hence the last zero estimator $\hat{\tau}_{\text{CUSUM}}$ is very close to the true value. In the usual case (i.e., $p_a^{\text{True}} \neq p_a^{as}$), the proposed estimator is a weighed sum of $\hat{\tau}_{\text{MLE}}$ and $\hat{\tau}_{\text{CUSUM}}$. This scheme is done by using the proposed weight function $w \equiv w(p_a^{\text{True}}|\ p_a^{as})$ in Section 3.2. Summing up two existing estimators to improve accuracy is a well-known technique in the statistical decision theory (Section 3.1).

Our simulations suggest that the performance of the proposed estimator $\hat{\tau}_{\text{NEW}}(\hat{w})$ is less biased compared to $\hat{\tau}_{\text{CUSUM}}$ and $\hat{\tau}_{\text{MLE}}$ under various parameter settings and under various design values p_a^{as}. In terms of the mean squared error, the proposed method is not always the best but never becomes the worst among all competing estimators. This implies that our proposed estimator is more robust than the existing ones and reliably applied to many different parameter settings.

Acknowledgments

We thank the editor, associate editor, and referee for their careful reading of our article. We are also thankful to Professor Cheng-Der Fuh, Professor Ying-Chao Hung, and Professor Yi-Hua Tina Wang for their comments on an earlier version of our article.

Funding

Our work is supported by a research grant funded by the Ministry of Science and Technology of Taiwan (MOST 103-2118-M-008-MY2).

References

Assareh, H., Smith, I., and Mengersen, K. (2015). Change Point Detection in Risk Adjusted Control Charts, *Statistical Methods in Medical Reasearch* 24: 747–768.
Burr, W. I. (1979). *Elementary Statistical Quality Control*, New York: Dekker.
Casella, G. and Berger, R. L. (2002). *Statistical Inference*, Pacific Grove: Duxbury.
Duran, R. I. and Albin, S. L. (2009). Monitoring a Fraction with Easy and Reliable Settings of the False Alarm Rate, *Quality and Reliability Engineering International* 25: 1026–1043.
Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, London: Chapman & Hall.
Emura, T., Chen, Y. H., and Chen, H. Y. (2012). Survival Prediction Based on Compound Covariate Under Cox Proportional Hazard Models, *PLoS ONE* 7: e47627. doi:10.1371/journal.pone.0047627.
Emura, T., Kao, F. S., and Michimae, H. (2014). An Improved Nonparametric Estimator of Sub-Distribution Function for Bivariate Competing Risk Models, *Journal of Multivariate Analysis* 132: 229–241.
Emura, T. and Lin, Y. S. (2015). A Comparison of Normal Approximation Rules for Attribute Control Charts, *Quality and Reliability Enginerring International* 31: 411–418.
Fuh, C. D. and Mei, Y. (2008). Optimal Stationary Binary Quantizer for Decentralized Quickest Change Detection in Hidden Markov Models, in *Proceedings of 11th International Conference on IEEE Information Fusion*, Cologne, France.
Grigg, O. A., Farewell, V. T., and Spiegelhalter, D. J. (2003). Use of Risk-Adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts, *Statistical Methods in Medical Research* 12: 147–170.
Hawkins, D. M. and Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Wiley.
James, W. and Stein, C. (1961). Estimation with Quadratic Loss, in *Proceedings of Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1*, pp. 361–379.
Khan, R. A. (1968). A Note on Estimating the Mean of a Normal Distribution with Known Coefficient of Variation, *Journal of American Statistical Association* 63: 1039–1041.
Khan, R. A. (2008). Distributional Properties of CUSUM Stopping Times, *Sequential Analysis*, 27: 420–434.
Khan, R. A. and Khan, M. K. (2004). On the Use of the SPRT in Determining the Properties of Some CUSUM Procedures, *Sequential Analysis* 23: 355–378.
Laheetharan, A. and Wijekoon, P. (2010). Improved Estimation of the Population Parameters When Some Additional Information Is Available, *Statistical Papers* 51: 889–914.
Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*, New York: Wiley.
Page, E. S. (1954). Continuous Inspection Schemes, *Biometrika* 41: 100–114.
Perry, M. B. and Pignatiello, J. J., Jr. (2005). Estimation of the Change Point of the Process Fraction Nonconforming in SPC Applications, *Interational Journal of Reliability Quality and Safety Engineering* 12: 95–110.
Perry, M. B. and Pignatiello, J. J., Jr. (2008). A Change Point Model for the Location Parameter of Exponential Family Densities, *IIE Transactions* 40: 947–956.

Perry, M. B., Pignatiello, J. J., Jr., and Simpson, J. R. (2007). Estimating the Change Point of the Process Fraction Nonconforming with a Monotonic Change Disturbance in SPC, *Quality and Reliability Engineering Inernational* 23: 327–339.

Pignatiello, J. J., Jr. and Samuel, T. R. (2001). Identifying the Time of a Step-Change in the Process Fraction Nonconforming, *Quality Engineering* 13: 357–365.

Rossi, G., Sarto, S. D., and Marchi, M. A. (2014). New Risk-Adjusted Bernoulli Cumulative Sum Chart for Monitoring Binary Health Data. *Statistical Methods in Medical Reasearch*. doi:10.1177/0962280214530883.

Wald, A. (1947). *Sequential Analysis*, New York: Wiley.

Wang, H. (2009). Comparison of P Control Charts for Low Defective Rate, *Computational Statistics & Data Analysis* 53: 4210–4220.

Wencheko, E. and Wijekoon, P. (2005). Improved Estimation of the Mean in One-Parameter Exponential Families with Known Coefficient of Variation, *Statistical Papers* 46: 101–115.

Wetherill, G. B. and Brown, D. B. (1991). *Statistical Process Control, Theory and Practice*, London: Chapman and Hall.

Yang, S. F., Cheng, T. C., Hung, Y. C., and Cheng, S. W. (2011). A New Chart for Monitoring Service Process Mean, *Quality and Reliability Engineering International* 28: 377–386.