



Change-Point-Based Statistical Process Controls

19

Partha Sarathi Mukherjee

Contents

19.1	Introduction	361
19.2	Traditionally Used SPC Charts	363
19.2.1	Shewhart Chart.....	363
19.2.2	CUSUM Chart.....	366
19.2.3	EWMA Chart.....	367
19.2.4	Control Charts by Change-Point Detection (CPD).....	368
19.3	The Proposed SPC Chart	369
19.3.1	A Computationally Efficient Approach to Estimate $\tau^*(t)$	370
19.3.2	Integration of Ansari-Bradley Test with Cramer-von Mises Test.....	370
19.3.3	Data Pruning Based on P-Values.....	371
19.3.4	The Algorithm of the Proposed Control Chart.....	371
19.3.5	Implementation.....	371
19.4	Numerical Studies	372
19.4.1	Location Changes.....	372
19.4.2	Scale Changes.....	373
19.4.3	Location-Scale Changes.....	374
19.4.4	General Distributional Changes.....	374
19.5	Analysis of Various Real-World Data	376
19.5.1	Climate Data on Minneapolis, USA.....	377
19.5.2	Blood Glucose Monitoring Data.....	379
19.6	Concluding Remarks	379
	References	380

first briefly discuss a few commonly used SPC charts along with relevant references and then present a new chart for univariate continuous processes. Unlike most SPC charts in the literature, it neither assumes any “in-control” probability distribution nor requires any “in-control” Phase I data, and it aims to detect arbitrary distributional change. This chart uses a computationally efficient method to find the possible change-point. Moreover, the proposed chart uses a p-value-based data pruning approach to further increase the efficiency, and it combines the strengths of two different tests of hypotheses, which has a potentially broad application. Numerical simulations and two real-data analyses show that the chart can be used in various monitoring problems when the nature of distributional change is unknown.

Keywords

Ansari-Bradley test · Cramer-von Mises test · Distributional change · Nonparametric SPC · Pruning · P-value

Abstract

In the current era of computers, statistical monitoring of sequential observations is an important research area. In problems such as monitoring the quality of industrial products, health variables, climatological variables, etc., we are often interested in detecting a change in the process distribution in general, not just in mean or variance. We

19.1 Introduction

Statistical process control (SPC) charts are widely used for monitoring the stability of certain sequential processes in various disciplines including manufacturing and healthcare systems. Typically, the SPC charts assume that there are two causes of variability in the process measurements: one is “common cause” which is due to unavoidable randomness and another is “special cause” when an undesirable variability intervenes, for example, mechanical defects, improper handling of machines, human errors, onset of certain medical conditions, etc. When the variability is only due to common causes, the process is said to be “in-control”. “In-control” process measurements can be considered as realizations of

P. S. Mukherjee (✉)
Interdisciplinary Statistical Research Unit (ISRU), Indian Statistical Institute (ISI), Kolkata, West Bengal, India
e-mail: psm@isical.ac.in

a random model, for example, independent and identically distributed (i.i.d.) observations from a cumulative distribution function (c.d.f.) F_1 . When a special cause interferes, the process measurements no longer appear as i.i.d. realizations of F_1 , and then the system is said to be “out-of-control”. Practitioners typically divide SPC into two phases. Initially, a set of process measurements are analyzed in Phase I. If any “unusual” patterns in the process measurements are found, they make necessary adjustments and fine-tuning of the system. After removing all such special causes, we have a clean set of process measurements data under stable operating conditions, and they are representative of the actual process performance. The major goal of Phase II SPC control charts is to detect any change in process distribution after an unknown time-point.

A change in the process distribution may not always be in location and scale only; it can be general, for example, changes in degrees of freedom of chi-squared distribution. Furthermore, changes can either be *isolated*, i.e., the system goes “out-of-control” for a short time and then returns to “in-control”, or *persistent*, i.e., once the system goes “out-of-control,” it remains “out-of-control” or even goes further away from control until the special causes are removed. Among existing SPC charts in the literature, Shewhart-type [1] charts are used to detect *isolated* changes and cumulative sum (CUSUM) type charts (e.g., [2]) are used to detect *persistent* changes. However, most SPC charts consider shifts in location and/or scale, because they are most common and often capture other departures. In real-world problems, shifts in skewness and kurtosis can happen without much change in location and scale. For example, the shape of the process distribution gradually changing over time without much change in mean or variance. If we fail to detect those changes and let the process run, it can eventually become worse and a shift in location and scale can creep in. Moreover, the special causes that initiated the change can cause more damage to the system, and it may become a challenge to fix. If we can detect such a change in skewness, we can avoid subsequent troubles. Therefore, it is imperative to develop an SPC chart that can detect changes in the process distribution in general. The proposed chart precisely focuses on this.

Various types of SPC control charts have been proposed in the literature including Shewhart-type charts [1], cumulative sum (CUSUM)-type charts (e.g., [2]), exponentially weighted moving average (EWMA chart) [3] (e.g., [4, 5]), etc. Many control charts assume that “in-control” process distribution F_1 is either known or has a known parametric form (e.g., normal). In real-life problems, this is usually not the case. It has been demonstrated in the literature that the SPC charts using prespecified distribution in their designs may not be reliable in such cases (e.g., [6, 7]). To address this, a number of non-parametric or distribution-free SPC charts have been proposed. For example [8–18], and so on. [19] provides

an overview of nonparametric SPC charts. Discussions for multivariate cases are provided by Qiu and Hawkins [20, 21] and Qiu [22]. Some SPC charts (e.g., [23–25]) monitor process mean and variance jointly. Moustakides [26] proposes a method to detect distributional changes when both “in-control” and “out-of-control” distributions are known. Ross and Adams [27] propose two nonparametric control charts to detect arbitrary distributional change under the change point detection (CPD) framework when both “in-control” and “out-of-control” distributions are unknown. Mukherjee [28] also proposes a Phase II nonparametric SPC chart for detecting arbitrary distributional change, but it requires a substantial amount of “in-control” Phase I data. A thorough literature review on SPC charts can be found in [29] as well as in [30].

Most existing SPC charts mentioned above focus to detect changes in process location and scale but do not consider an arbitrary distributional change. Moreover, some methods require multiple observations at each time-point; some others require “in-control” Phase I observations. All of these may not be reasonable in many real-life problems. The proposed chart focuses on univariate continuous processes and proposes a p-value-based nonparametric SPC chart to detect an arbitrary distributional change when we can assume that the observations we collect are independent, but “in-control” c.d.f. F_1 is unknown, and “in-control” Phase I data are unavailable. p-value-based SPC charts are new trends as some researchers have already started developing those for better interpretation, e.g., [31, 32], etc.

Additionally, the proposed chart uses the strengths of Cramer-von Mises test and Ansari-Bradley test and demonstrates a better performance in detecting an arbitrary distributional change early. It is demonstrated in the literature [27] that the power of two-sample Cramer-von Mises test to detect a small change in standard deviation or scale is rather weak. The proposed chart overcomes that by integrating this test with Ansari-Bradley test, a nonparametric test for detecting scale change. The integration process is quite general in nature, and hence similar integration techniques can be used to design better control charts in many scenarios.

The major steps of the proposed SPC chart are the following: First, we estimate the possible change-point that gives the minimum p-value of the relevant two-sample Cramer-von Mises test statistic. We use a computationally efficient method to do this. Using that change-point, we calculate the p-value of Ansari-Bradley test and pick the smallest of the two p-values as the effective p-value and use it as the charting statistic. Since it is demonstrated in the literature (e.g., [27]) that Cramer-von Mises test does not have strong power to detect scale changes, if there is indeed a scale change in the process distribution, Ansari-Bradley test gives a smaller p-value than Cramer-von Mises test. Therefore, to detect such changes, it is better to use the smaller of the p-values. If

the process distribution changes without affecting the scale parameter, Cramer-von Mises test gives a smaller p-value than Ansari-Bradley test. It is also better to use the smaller of the two p-values to detect such a change as well. If the p-value is large enough, and our effective sequence of observations is too long, we prune observations from distant past and we collect the next observation. The amount of pruning is a nondecreasing function of the effective p-value as the higher the p-value, the more likely it is that no distributional change has occurred, and it is therefore better to ignore information from distant past to speed up computation. If the p-value is small enough, the chart signals a distributional change.

The remainder of the chapter is organized as follows. At the end of this paragraph, a nomenclature subsection is provided so that this chapter can be read smoothly. Next, brief descriptions of traditionally used SPC charts are provided in Sect. 19.2. The proposed control chart is described in Sect. 19.3. Numerical studies to evaluate its performance in comparison with several existing control charts are presented in Sect. 19.4. A climatological data analysis and a blood sugar monitoring data analysis by the proposed chart and its competitors are presented in Sect. 19.5. A few remarks in Sect. 19.6 conclude this chapter.

19.2 Traditionally Used SPC Charts

Statistical process control of a production process is roughly divided into two phases: Phase I and Phase II. In Phase I, i.e., in the initial stage, we usually do not have enough information about the performance of the production process, and our major goal in this stage is to adjust the production process so that it can run in a stable manner. First, we usually let the process produce a given amount of products, and then the quality characteristics of these products are analyzed. If the statistical analysis of these data shows indication that the process is not running stably, we try to figure out the root causes for that and make adjustments of the process so that it can run stably. After the adjustments, another set of data is collected and analyzed, and the production process should be adjusted again, if necessary. The analysis and adjustment process is iterated several times until we are confident that the performance of the production process is stable. Once all “special causes” have been removed and the production process is “in-control,” we collect an “in-control” dataset from the products manufactured under the stable operating conditions, and the “in-control” data are used for estimating the “in-control” distribution of the quality characteristic(s) of interest. Based on the actual “in-control” distribution if known, or estimated “in-control” distribution otherwise, a Phase II SPC control chart is designed. It is typically used for online monitoring of the production process. When the chart detects a significant change in the distribution of the quality

characteristic(s) from “in-control” distribution, it gives a signal and the production process is stopped immediately for identification of the root cause for such a change and its removal. This online monitoring stage is often called Phase II SPC.

In both Phase I and Phase II of SPC, many statistical tools such as histograms, stem-and-leaf plots, regression, and design of experiments are very helpful. Among all these, control charts are especially useful since they are constructed specifically to detect “out-of-control” performance of the production process. A charting statistic should be chosen such that it contains as much of the information in the observed data about the distribution of the quality characteristic(s) as possible and be sensitive to any distributional change as well. In the literature, different types of control charts have been developed including the Shewhart charts, the cumulative sum (CUSUM) charts, the exponentially weighted moving average (EWMA) charts, charts based on change-point detection (CPD), and so on. Brief descriptions of some of these control charts are provided below.

19.2.1 Shewhart Chart

The first control chart was proposed by Shewhart [1] in 1931. The chart assumes that the quality variable X follows normal distribution $N(\mu_0, \sigma^2)$, and at each time-point we obtain m independent quality observations. Denote $(X_{n1}, X_{n2}, \dots, X_{nm})$ to be the n -th batch of observations, where the batch size is $m \geq 2$. Traditional z-test is used to check if the process observations are “in-control” at the n -th time-point. The process is considered “out-of-control” when

$$\bar{X}_n > \mu_0 + z_{1-\alpha/2} \frac{\sigma}{\sqrt{m}} \quad \text{or} \quad \bar{X}_n < \mu_0 - z_{1-\alpha/2} \frac{\sigma}{\sqrt{m}},$$

where \bar{X}_n is the sample mean of $(X_{n1}, X_{n2}, \dots, X_{nm})$ and $z_{1-\alpha/2}$ is $(1 - \alpha/2)$ -th quantile of the standard normal distribution. This version can be used when both μ_0 and σ are known. However, it is usually not the case in reality. In that case, they have to be estimated from a dataset known to be “in-control.” Suppose, $(X_{i1}^*, X_{i2}^*, \dots, X_{im}^*), i = 1, 2, \dots, M$ be an “in-control” dataset. Let \bar{X}_i^* and R_i^* be the sample mean and sample range of the i -th batch of “in-control” dataset and \bar{X}^* and \bar{R}^* be the sample means of $\{\bar{X}_i^*, i = 1, 2, \dots, M\}$ and $\{R_i^*, i = 1, 2, \dots, M\}$, respectively. It can be easily verified that \bar{X}^* is an unbiased estimator of μ_0 and $\bar{R}^*/d_1(m)$ is an unbiased estimator of σ , where $d_1(m) = E(R_i^*/\sigma)$ is a constant depending on m . When $m = 2$, $d_1(m) = 1.128$, when $m = 5$, $d_1(m) = 2.326$. $d_1(m)$ values for many other commonly used m are provided in Table 3.1 of [30]. Therefore, the Shewhart chart signals a shift in process mean if

$$\begin{aligned} \bar{X}_n &> \bar{X}^* + z_{1-\alpha/2} \frac{\bar{R}^*}{d_1(m)\sqrt{m}} \quad \text{or} \\ \bar{X}_n &< \bar{X}^* - z_{1-\alpha/2} \frac{\bar{R}^*}{d_1(m)\sqrt{m}}. \end{aligned} \tag{19.1}$$

Traditionally, the manufacturing industry uses $\alpha = 0.0027$, and hence $(1 - \alpha/2)$ -th quantile of $N(0, 1)$, i.e., $z_{1-\alpha/2}$ is 3. Therefore, the chart signals mean shift at time-point n if \bar{X}_n falls outside the interval of width six sigma centered at μ_0 and where sigma is the standard deviation of \bar{X}_n . Thus, the terminology ‘‘six sigma’’ originated in the domain of quality control.

The performance of a control chart is traditionally measured by average run length (ARL). Since the charts use control limits for making decision on process performance, an ‘‘in-control’’ process sometimes give false signals of distributional shift. This phenomenon is analogous to having type I error in hypothesis testing. The number of samples or batches collected from the initial time-point of consideration to the occurrence of first false ‘‘out-of-control’’ signal when the process remains ‘‘in-control’’ is called ‘‘in-control’’ run length. The mean of such a run length is called ‘‘in-control’’ average run length denoted as ARL_0 . On the other hand, the number of samples or batches collected from the time-point when the shift actually occurs to the time-point of signal of shift is called ‘‘out-of-control’’ run length. Its mean is called ‘‘out-of-control’’ average run length denoted as ARL_1 . The ideal situation is that, for a control chart, ARL_0 value is large and ARL_1 value is small. However, similar to type I and type II error probabilities in hypothesis testing, it is difficult to achieve. Usually, when ARL_0 is large, ARL_1 is also relatively large and vice versa. In SPC literature, we usually fix the ARL_0 value at a given level and compare the performances of the control charts by comparing how small their ARL_1 values are. In the \bar{X} Shewhart chart as described above, the distribution of ‘‘in-control’’ run length is clearly geometric with parameter α . $\alpha = 0.0027$ makes $ARL_0 = 1/\alpha = 370.37$. ARL_1 can also be computed easily as a function of the shifted mean.

In the literature, there are many versions of Shewhart chart. One of them uses sample standard deviation to estimate σ rather than sample range. In this case, the chart gives a signal for mean shift when

$$\begin{aligned} \bar{X}_n &> \bar{X}^* + z_{1-\alpha/2} \frac{\bar{S}^*}{d_3(m)\sqrt{m}} \\ \text{or } \bar{X}_n &< \bar{X}^* - z_{1-\alpha/2} \frac{\bar{S}^*}{d_3(m)\sqrt{m}}, \end{aligned} \tag{19.2}$$

where \bar{S}^* is the sample mean of the batch-wise sample standard deviations $\{S_i^*, i = 1, 2, \dots, M\}$, i.e.,

$$S_i^* = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (X_{ij} - \bar{X}_i^*)^2} \quad \text{for } 1, 2, \dots, M,$$

and $d_3(m) = E(S_i^*/\sigma)$ is a constant depending on the value of m . Under the same setup, Shewhart charts were constructed to monitor process variability. Defining $d_2(m) = \sqrt{\text{Var}\left(\frac{R_i^*}{\sigma}\right)}$, and using $d_1(m) = E(R_i^*/\sigma)$, we can estimate $\sigma_{R_i^*}$ by $\frac{d_2(m)}{d_1(m)}\bar{R}^*$. Therefore, this version of Shewhart chart gives a signal for a change in variability if

$$R_n > \bar{R}^* + z_{1-\alpha/2} \frac{d_2(m)}{d_1(m)}\bar{R}^* \quad \text{or} \quad R_n < \bar{R}^* - z_{1-\alpha/2} \frac{d_2(m)}{d_1(m)}\bar{R}^*. \tag{19.3}$$

Similarly, another version of Shewhart chart was constructed using sample standard deviation instead of sample range. Using the result $\sigma_{S_i^*} = \sigma\sqrt{1 - d_3^2(m)}$ proved by Kenney and Keeping [34], $\sigma_{S_i^*}$ can be estimated by $\frac{\bar{S}^*}{d_3(m)}\sqrt{1 - d_3^2(m)}$. Therefore, the chart gives a signal of a change in variability if S_n , the sample standard deviation of $(X_{n1}, X_{n2}, \dots, X_{nm})$, satisfies the following condition:

$$\begin{aligned} S_n &> \bar{S}^* + z_{1-\alpha/2} \frac{\sqrt{1 - d_3^2(m)}}{d_3(m)}\bar{S}^* \\ \text{or } S_n &< \bar{S}^* - z_{1-\alpha/2} \frac{\sqrt{1 - d_3^2(m)}}{d_3(m)}\bar{S}^*. \end{aligned} \tag{19.4}$$

Using the result that $\frac{(m-1)S_i^{*2}}{(\sigma^2)} \sim \chi_{m-1}^2$, another chart was constructed using sample variance instead of sample standard deviation. This chart gives a signal for a change in process variance when

$$S_n^2 > \bar{S}^{*2} \frac{\chi_{1-\alpha/2, m-1}^2}{m-1} \quad \text{or} \quad S_n^2 < \bar{S}^{*2} \frac{\chi_{\alpha/2, m-1}^2}{m-1}, \tag{19.5}$$

where \bar{S}^{*2} is the sample average of $\{S_i^{*2}, i = 1, 2, \dots, M\}$.

Next, we discuss \bar{X} Shewhart chart for monitoring individual observations rather than batched observations in Phase I SPC. The idea is to artificially create grouped data by grouping consecutive observations. First, we fix the size of each group $\tilde{m} > 1$. Then, the first \tilde{m} observations form the first group, the next \tilde{m} observations form the second group, and so on. Next, we can apply \bar{X} Shewhart chart (19.1) on the

grouped data. However, one problem here is that consecutive groups are \tilde{m} time-points apart. Hence, it is difficult to know the process behavior at each time-point. To overcome this limitation, most people adopt the idea of moving windows. We artificially create grouped data as follows: Group 1 ($X_1, X_2, \dots, X_{\tilde{m}}$), Group 2 ($X_2, X_3, \dots, X_{\tilde{m}+1}$), and so on until Group $(n - \tilde{m} + 1)$ ($X_{n-\tilde{m}+1}, X_{n-\tilde{m}+2}, \dots, X_n$). Denote $MR_1, MR_2, \dots, MR_{n-\tilde{m}+1}$ to be the sample ranges of the $(n - \tilde{m} + 1)$ groups of data and \overline{MR} to be their sample mean. From the definition of $d_1(\tilde{m})$, we can estimate σ by $\overline{MR}/d_1(\tilde{m})$. Therefore, the \bar{X} Shewhart chart for monitoring individual observations gives a signal of mean shift when

$$X_n > \bar{X} + z_{1-\alpha/2} \frac{\overline{MR}}{d_1(\tilde{m})} \quad \text{or} \quad X_n < \bar{X} - z_{1-\alpha/2} \frac{\overline{MR}}{d_1(\tilde{m})}. \tag{19.6}$$

Similarly, R Shewhart chart for monitoring individual observations gives a signal of a variability shift when

$$MR_i > \overline{MR} + z_{1-\alpha/2} \frac{d_2(\tilde{m})}{d_1(\tilde{m})} \overline{MR} \\ \text{or} \quad MR_i < \overline{MR} - z_{1-\alpha/2} \frac{d_2(\tilde{m})}{d_1(\tilde{m})} \overline{MR}. \tag{19.7}$$

Here, we should check at all time-points that belong to the i -th group, i.e., from the i -th to the $(i + \tilde{m} - 1)$ -th time-points. Other Shewhart charts for monitoring individual observations were constructed similarly.

In many applications, quality characteristics are categorical. Now, we describe some Shewhart charts for monitoring such characteristics. After certain products are randomly chosen for monitoring purposes, they are classified in conforming and nonconforming products based on requirements on the quality characteristics. Now, we monitor the proportion of nonconforming products over time. We assume that when a production process is “in-control,” the true proportion of nonconforming products is π_0 , and we obtain a random sample of m products at each time-point. Let Y be the number of nonconforming products obtained at a given time-point. Therefore, $Y \sim \text{Binomial}(m, \pi_0)$ when the process is “in-control.” Let $p = Y/m$ be the sample proportion of nonconforming products at the time-point. When m is large, the probability distribution of p can be well approximated by $N(\pi_0, \pi_0(1 - \pi_0)/m)$ by the central limit theorem when the process is “in-control.” Hence, the process can be called “out-of-control” if

$$p > \pi_0 + z_{1-\alpha/2} \sqrt{\frac{\pi_0(1 - \pi_0)}{m}} \\ \text{or} \quad p < \pi_0 - z_{1-\alpha/2} \sqrt{\frac{\pi_0(1 - \pi_0)}{m}}.$$

In practice, π_0 is often unknown and should be estimated from collected Phase I data, just like in the original version of \bar{X} Shewhart chart use estimated μ_0 and σ . As before, we assume that we have M batches of Phase I data. Let p_i^* be the sample proportion of nonconforming products in the i -th batch of Phase I data for $i = 1, 2, \dots, M$ and \bar{p}^* be their sample mean. Therefore, we can estimate π_0 by \bar{p}^* , and hence the p Shewhart chart gives a signal for a change in the proportion of nonconforming products at n -th time-point when

$$p_n > \bar{p}^* + z_{1-\alpha/2} \sqrt{\frac{\bar{p}^*(1 - \bar{p}^*)}{m}} \\ \text{or} \quad p_n < \bar{p}^* - z_{1-\alpha/2} \sqrt{\frac{\bar{p}^*(1 - \bar{p}^*)}{m}}. \tag{19.8}$$

There are other versions of Shewhart chart in the literature for monitoring count processes having distributions such as Poisson. Detailed descriptions of such charts are provided in Chapter 3 of [30].

Shewhart charts are good at detecting relatively large isolated shifts, but not so efficient in detecting relatively small but persistent shifts. This is because Shewhart charts evaluate the process performance based on the observed data collected at each individual time-point and ignore observations collected previously. Therefore, Shewhart charts are popular in Phase I SPC where large and isolated shifts are common but less commonly used in Phase II SPC.

As we mentioned before, the \bar{X} , R , and S Shewhart charts are appropriate to use only in cases when the process distribution is normal and the observations are independent of each other. When the process distribution is not normal, the probability of type I error, i.e., the probability of false “out-of-control” signal when the process is actually “in-control,” can substantially differ from the prefixed value of α . If the type I probability of a Shewhart chart is larger than α , then the chart will give false “out-of-control” signal more often than expected. Consequently, much time and resource will be wasted to find the cause of such signals and adjusting the related production process. On the other hand, if the type I probability of a Shewhart chart is smaller than α , then real shifts will be missed more often than expected and hence many nonconforming products could be manufactured. However, when the distribution of X_{ij} is non-normal but the batch size m is large, the issue described above will not be serious, because the distribution of \bar{X}_n can be well approximated by normal distribution due to the central limit theorem. In cases when the distribution of X_{ij} is non-normal and the batch size m is small, mainly two approaches are usually taken. One approach is to transform the non-normal data to normal and then use the conventional Shewhart charts to the transformed data [35,36], and the other approach is to use Shewhart charts that are constructed to monitor non-normal data.

If the performances of the Shewhart charts are evaluated by “in-control” average run length ARL_0 and “out-of-control” average run length ARL_1 , then the performance measures will be not accurate if the observations are correlated [37–39]. In these cases, correlations have to be handled properly.

19.2.2 CUSUM Chart

Shewhart chart makes a decision whether a process is “in-control” or not at a time-point using only the observations obtained at that time-point and ignoring all previous observations. Therefore, it is not very effective in Phase II monitoring in most cases, because previous observations contain helpful information about process performance at present. Page [2] suggested the first cumulative sum (CUSUM) chart to overcome this limitation. Let us first describe the basic CUSUM chart for detecting mean shift of a process following normal distribution. Again, the chart assumes that the process observations X_1, X_2, X_3, \dots follow $N(\mu_0, \sigma^2)$ and are independent. The CUSUM charting statistics are given by

$$C_n^+ = \max(0, C_{n-1}^+ ((X_n - \mu_0)/\sigma) - k), \quad (19.9)$$

$$C_n^- = \min(0, C_{n-1}^- ((X_n - \mu_0)/\sigma) + k) \quad (19.10)$$

where $C_0^+ = C_0^- = 0$, and $k > 0$ is an allowance constant. The CUSUM chart gives a signal of mean shift if

$$C_n^+ > \rho_c \quad \text{or} \quad C_n^- < -\rho_c \quad (19.11)$$

where $\rho_c > 0$ is a control limit. The allowance constant k is prespecified, and the value of ρ_c is chosen so that the average run length when the process is “in-control,” denoted as ARL_0 , equals a given number, say, 200, 370, 500, etc. Table 4.1 of [30] provides the values of ρ_c for various values of allowance constant k and ARL_0 . We can easily see that the charting statistics C_n^+ and C_n^- make use of all available data before the n -th time-point, and they restart from 0 when the cumulative charting statistics suggest no significant evidence for mean shift in the sense that $C_{n-1}^+ ((X_n - \mu_0)/\sigma) < k$ and $C_{n-1}^- ((X_n - \mu_0)/\sigma) > -k$. Because of this restarting mechanism, the CUSUM chart enjoys a good theoretical property that it has smallest “out-of-control” ARL, denoted as ARL_1 , among all control charts having the same value of ARL_0 . Moustakides [26] proved that the CUSUM chart with an allowance constant k has the shortest ARL_1 value among all charts with a fixed ARL_0 value for detecting a persistent shift of size $\delta = 2k$.

To be able to use the above CUSUM chart, the process observations should be independent before and after a potential shift, both “in-control” and “out-of-control” distributions have to be normal, and the parameters μ_0 and σ of the “in-control” distribution have to be known. However, in practice these assumptions may not be reasonable.

If the observations are autocorrelated, then the actual value of ARL_0 will be different from the specified value for which the control limit ρ_c is determined in case of i.i.d. (independent and identically distributed) normal observations as provided by Table 4.1 of [30]. For example, if the production process is autoregressive of order 1, and the autocorrelation is negative, then the actual value of ARL_0 will be larger than the specified value for which ρ_c is determined. Consequently, the chart will not be sensitive enough to mean shifts and a lot of nonconforming products can be manufactured. On the other hand, if the production process is autoregressive of order 1, with positive autocorrelation, then the actual value of ARL_0 will be smaller than the specified value for which ρ_c is determined. That means the chart will give too many false signals of mean shift than expected and the production process has to be stopped too many times unnecessarily, and hence many resources will be wasted. One commonly used approach to accommodate possible autocorrelation among observed data is to group neighboring observations into batches and then apply conventional CUSUM charts for independent data to the batch means. One major reason behind this idea is that possible autocorrelation in the original data will be mostly eliminated in the new batch means [40]. Because of autocorrelation, the standard deviation of standardized group means may not be 1, and it can differ very much from 1. Therefore, the actual ARL_0 may be far away from the specified ARL_0 value. To overcome this limitation of the grouping approach, the group means need to be scaled properly which is difficult to do unless we know the nature of correlation in the original data. For related descriptions, see [40, 41], and [42]. Another disadvantage of the grouping approach is that the control chart cannot detect a mean shift promptly as it has to wait until all observations within a group are obtained. An alternative approach to the grouping idea is to describe the correlation by a statistical model such as autoregressive moving average (ARMA) model. In many practical applications, appropriate special cases of the ARMA model, such as first-order autoregressive model, can be used; otherwise, an appropriate model can be selected by a model selection procedure. After a time-series model is chosen and fitted by a certain routine procedure in time-series analysis, we can calculate the residuals. If the chosen time-series model describes the observed “in-control” data adequately, and the production process is “in-control” until the given time-point, the residuals should approximately be independent with a zero-mean common normal distribution whose variance can be estimated by an appropriate estimator. Then, we can apply the conventional CUSUM chart to the calculated residuals. However, we should be careful that an “out-of-control” signal may not always be due to a shift in mean; it can be due to a change in correlation structure in the observations as well. Related discussions on model-based control charts for monitoring autocorrelated processes can be found in [43–45], and so on.

In cases when the observations are i.i.d. normal but the “in-control” mean μ_0 and variance σ^2 are unknown, a common approach is to estimate those from a large “in-control” dataset. However, even a small randomness of these estimated values affects the performance of the CUSUM charts in a significant way. Hawkins [46] explored this research question in detail. Since in many applications we cannot have an extremely large “in-control” dataset, Hawkins [46] proposed the self-starting CUSUM charts in which “in-control” parameters are estimated from the observations collected in Phase II SPC. Assume that no “out-of-control” signal is given until the $(n-1)$ -th time-point. A new observation is collected at the n -th time-point and we want to make a decision whether a signal of mean shift should be given at this time-point. Since no signal for mean shift is given at $(n-1)$ -th time-point, we can consider all observations collected at that time-point and before, i.e., $X_{n-1}, X_{n-2}, \dots, X_1$, as “in-control” observations. Therefore, μ_0 and σ^2 can be estimated by their sample mean \bar{X}_{n-1} and sample variance S_{n-1}^2 , respectively, as long as $n \geq 3$. Therefore, for constructing a CUSUM chart, it is natural to replace $(X_n - \mu_0)/\sigma$ in (19.9) and (19.10) by $T_n = (X_n - \bar{X}_{n-1})/S_{n-1}$. When X_1, X_2, \dots, X_n are i.i.d. $N(\mu_0, \sigma^2)$ which is the case when the process is “in-control” up to the n -th time-point, we can easily check that $(\sqrt{(n-1)/n})T_n \sim t_{n-2}$. As proved in [47], T_1, T_2, \dots, T_n are independent of each other when the process is “in-control” up to the n -th time-point. Therefore, in that case,

$$Z_n = \Phi^{-1} \left[\Upsilon_{n-2} \left(\sqrt{\frac{n-1}{n}} T_n \right) \right]$$

follow i.i.d. $N(0, 1)$ distribution, where Φ and Υ_{n-2} are cumulative distribution functions (c.d.f.) of $N(0, 1)$ and t_{n-2} , respectively. Since Φ^{-1} and Υ_{n-2} are increasing functions, a mean shift in the original observations X_i , $1 \leq i \leq n$ indicates a mean shift in the transformed observations Z_i , $1 \leq i \leq n$ and vice versa. Therefore, detection of mean shift in Phase II monitoring can be accomplished by using transformed observations Z_n , $n \geq 1$ in place of $(X_n - \mu_0)/\sigma$ in (19.9) and (19.10). It has been demonstrated in the literature that if a persistent mean shift occurs within the first few observations in Phase II monitoring, the self-starting CUSUM chart as described above has a weak power to detect it. Therefore, in practice, at least a dozen or more “in-control” observations should be collected in Phase II before using the self-starting CUSUM chart. Self-starting control charts are now popular in the literature ([48, 49], and so on).

The traditional CUSUM chart (19.9–19.10) has an allowance parameter k , which should be set as $\delta/2$ where δ is the size of potential mean shift. In practice, δ is often unknown at the time when we design the CUSUM chart and hence choice of k is not straightforward. Sparks [50] proposed two approaches to solve this issue. Sparks’ first approach is to use several CUSUM charts with different k

values simultaneously so that these charts target to detect mean shifts of different sizes. Such a joint control scheme gives an “out-of-control” signal of mean shift if at least one of the CUSUM charts detects a mean shift. Of course, the ARL_0 values of these CUSUM charts have to be the same prefixed number. If we have prior information about the potential mean shift, we should incorporate the information while determining the k values. Sparks’ second approach is to estimate the size of mean shift δ recursively at each time-point and updating the value of k accordingly. The control limit should also be updated at each time-point so that we can maintain the prespecified ARL_0 value. These are called adaptive CUSUM charts.

When the “in-control” distribution of the production process is not normal, then the traditional CUSUM charts should not be used. If we know that the “in-control” distribution is in exponential family such as gamma and Weibull distribution, then we can similarly construct CUSUM charts by using sequential probability ratio test. However, if the “in-control” distribution is completely unknown, we can use nonparametric control charts [11–18] that do not assume any “in-control” distribution.

The versions of CUSUM chart mentioned above were designed for detecting a step shift in process mean. However, in many applications, the process mean and/or variance changes gradually with or without a known parametric pattern, after the process becomes “out-of-control.” Such changes are called drifts. Gan [51], Davis and Woodall [52], and many other researchers proposed CUSUM charts for detecting linear drifts.

Recently, CUSUM charts with variable sampling rate become popular ([53, 54] and many others). In this type of CUSUM chart, the sampling rate varies over time based on all observed data. There are many different types of sampling rate such as variable sampling intervals, variable sample sizes, etc. One major advantage of variable sampling rate CUSUM charts compared to fixed sampling rate CUSUM charts is faster detection of small to moderate shift in process mean. Recently, Li and Qiu [32] suggested implementing a CUSUM chart using statistical p-values and proposed the concept of dynamic sampling.

In the literature, researchers have constructed CUSUM charts for monitoring the variance of the process distribution as well.

19.2.3 EWMA Chart

In spite of having good theoretical properties, CUSUM charts were difficult to use in the 1950s when there was no computers. A simpler chart, called exponentially weighted moving average (EWMA) chart, was proposed by Roberts [55] in 1959. Under the same assumptions and notations of CUSUM chart, the EWMA charting statistic is defined as

$$E_n = \lambda X_n + (1 - \lambda)E_{n-1} \quad (19.12)$$

where $E_0 = \mu_0$ and $0 < \lambda \leq 1$ is a weighting parameter. We can easily check that

$$E_n = \lambda \sum_{i=1}^n (1 - \lambda)^{n-i} X_i + (1 - \lambda)^n \mu_0, \quad (19.13)$$

and when the process is “in-control,”

$$E_n \sim N\left(\mu_0, \frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2n}] \sigma^2\right).$$

That means E_n is a weighted average of μ_0 and all observations up to time-point n , and the weight received by X_i decays exponentially fast when i moves away from n . Therefore, it becomes easy to study the properties of E_n when the process is “in-control.” From the probability distribution of E_n , EWMA chart gives a signal for mean shift when

$$|E_n - \mu_0| > \rho_e \sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2n}]}, \quad (19.14)$$

where $\rho_e > 0$ is a control limit. $\lambda > 0$ is chosen beforehand, and the value of ρ_e is determined such that a specified value of ARL_0 is achieved.

To be able to use the EWMA chart in practice, we need to choose λ values properly and then the value of ρ_e so that the prespecified ARL_0 value is achieved. Just like the CUSUM charts, we need to specify a target shift size first and then we can search for a λ value and the corresponding ρ_e value such that the prespecified ARL_0 value is achieved and the ARL_1 value for detecting the mean shift of target size is minimized. As a general guideline, small λ values are good for detecting relatively small mean shifts, and large λ values are good for detecting relatively large means shifts. Crowder [56] provides a discussion on this issue. While we assume that the “in-control” process distribution is normal, some researchers have demonstrated that the EWMA chart is quite robust to normality assumption [3].

In case of autocorrelated observations, similar approaches as we described in the CUSUM chart can be implemented in case of EWMA charts as well. However, some researchers suggest applying the EWMA charts directly to the original data and then adjusting the control limits to reflect the impact of autocorrelations [57].

For using Shewhart, CUSUM, or EWMA chart as described before, μ_0 and σ has to be known or estimated before the monitoring starts. As we have discussed before, this is not convenient in many applications, and hence we should use self-starting control charts. Just like the self-starting CUSUM control charts, we first transform the original data to Z_i , $i \geq 1$ and then apply the traditional EWMA chart to the

transformed data. In the literature, there are several methods such as [58] where λ is chosen adaptively depending on the size of potential mean shift. These are called adaptive EWMA charts.

In the literature, EWMA charts have been developed when the process distribution follows other parametric forms. Some researchers [59, 60] constructed EWMA charts for monitoring processes following Weibull distributions. Borrer et al. [61], Gan [62], and some others discussed process monitoring when the process distribution is Poisson. Perry and Pignatiello [63], Sparks et al. [64], and many others discussed EWMA process monitoring with binomial or negative binomial distributions.

All versions EWMA charts that we discussed so far are designed for detecting step shift in the process mean. In some practical situations, when the process becomes “out-of-control,” its mean departs gradually from the “in-control” level. It is important that we can detect such gradual departures, called drifts, as early as possible. In the literature, some researchers have modified EWMA charts to detect drifts efficiently ([65] and others).

Just like CUSUM charts, researchers have constructed EWMA charts to monitor the variance of the production processes as well.

19.2.4 Control Charts by Change-Point Detection (CPD)

In change-point detection (CPD), the distribution of the first part of a sequence of random variables is assumed to be the same, the distribution of the remaining part of the random variables of that sequence is also assumed to be the same, but the distributions of the two parts are assumed to be different. The specific position in the sequence at which the distribution of the random variables changes from one to the next is called a change-point. Our major goal is to estimate the position of the change-point. Gombay [66], Hinkley [67], and many others in the literature provide detailed description of this topic. In change-point detection problems, the sample size is usually fixed. In Phase I SPC, the sample size is usually fixed, and then the change-point methods can be applied directly. In Phase II SPC, observations are obtained sequentially over time. Therefore, change-point methods must be applied appropriately in such cases. Recently, change-point methods have been modified and applied to the SPC problems ([4, 5], and others) as well. Change-point-based control charts are good at detecting small and persistent shifts and can estimate the position of change-point efficiently.

Let us describe the change-point based control chart proposed by Hawkins et al. [4]. It assumes that the observations X_1, X_2, \dots, X_n follow this change-point model

$$\begin{aligned}
 X_i &= \mu_0 + \varepsilon_i & \text{for } 1 \leq i \leq r \\
 X_i &= \mu_1 + \varepsilon_i & \text{for } (r + 1) \leq i \leq n
 \end{aligned}$$

where r is the change-point and ε_i , $1 \leq i \leq n$ is a sequence of i.i.d. random variables having the common distribution $N(0, \sigma^2)$. For testing the existence of the change-point, the likelihood ration test statistic is

$$T_{max,n} = \max_{1 \leq j \leq n-1} \sqrt{\frac{j(n-j)}{n}} \left| \bar{X}_j - \bar{X}'_j \right| / \tilde{S}_j \quad (19.15)$$

where \bar{X}_j and \bar{X}'_j are sample means of first j and the remaining $(n-j)$ observations, and $\tilde{S}_j^2 = \sum_{i=1}^j (X_i - \bar{X}_j)^2 + \sum_{i=j+1}^{n-j} (X_i - \bar{X}'_j)^2$. The change-point-based chart gives an “out-of-control” signal of mean shift when

$$T_{max,n} > \rho_n \quad (19.16)$$

where $\rho_n > 0$ is a control limit. If we have an “out-of-control” signal, then the position of change-point r is estimated by the maximizer in (19.15). Hawkins et al. [4] provided formulas to calculate the values of ρ_n for commonly used prespecified values of ARL_0 .

Just like other control charts, change-point-based charts were also constructed to accommodate other situations such as if the “in-control” process distribution is not normal, but other known parametric distribution. Change-point-based charts were also developed to monitor process variance as well.

Although the change-point-based control charts have their advantages over Shewhart, CUSUM, and EWMA charts, their computation is still relatively complex. This is because the estimator of the position of change-point has to be recalculated each time a new observation is obtained which involve a search in a sequence of n time-points. A systematic comparison between the performances of self-starting and adaptive traditional (CUSUM and EWMA) control charts and the change-point-based control charts is currently lacking in the literature.

In practice, many variables do not follow normal distribution. For example, many economic indices, rainfall distribution, lifetime distribution of many products, etc. are usually skewed to the right. In most instances it is difficult to find an appropriate parametric distribution for their modeling. Researchers use nonparametric methods to describe their distributions. When the normality assumption is not reasonable, several researchers have pointed out that the traditional control charts would be unreliable for process monitoring ([6,7,16], and many others). In such cases, nonparametric statistical methods based on the ranking or ordering information of the observations can be considered for making inferences about the underlying process distribution. Another approach is based on data categorization. Clearly, both approaches

have the limitation of losing useful information during ranking and categorization. However, the methods based on data categorization seem to be more efficient in process monitoring as use some information about observation magnitudes. Nonparametric versions of conventional charts such as Shewhart, CUSUM, EWMA, and change-point-based have been developed in the literature. Most nonparametric control charts in the literature are for Phase II SPC. However, there is only a limited discussion on Phase I SPC when the process distribution does not follow any common parametric form [68,69].

Many researchers have developed control charts in the literature to jointly monitor process mean and variance as well. We skip such details in this chapter. We also omit descriptions of multivariate control charts as well. Interested readers should go through [30].

19.3 The Proposed SPC Chart

In this section, we describe our proposed Phase II chart for detecting persistent distributional change in univariate continuous processes. We assume that “in-control” probability distribution of the process is unknown, “in-control” Phase I data are unavailable, and the observations we collect are independent of each other. Let $Y_1, Y_2, \dots, Y_t, \dots$ be a sequence of independent observations during Phase II process monitoring. We start performing statistical monitoring right after time-point $S_0 \geq 4$, because we need to need have at least a few observations so that statistical tests have enough power. This is a common practice as discussed in [4,27], etc. At each time $t \geq S_0$, we consider the following change-point framework. Assuming τ , $2 \leq \tau \leq t - 2$ to be a possible point of distributional change, we consider two samples $\{Y_1, Y_2, \dots, Y_\tau\}$ and $\{Y_{\tau+1}, Y_{\tau+2}, \dots, Y_t\}$, and perform Cramer-von Mises test to check if they are coming from the same unknown continuous cumulative distribution function (cdf). For statistically testing whether two samples are from the same continuous distribution, two tests are commonly used: Kolmogorov-Smirnov test and Cramer-von Mises test. Kolmogorov-Smirnov test does not work well if there are ties in the samples. Also, Cramer-von Mises test usually have more power than Kolmogorov-Smirnov test in many situations. Therefore, we use Cramer-von Mises test ahead of Kolmogorov-Smirnov test. The two-sample Cramer-von Mises test statistic is given by

$$\begin{aligned}
 C_\tau(t) &= \frac{\tau(t-\tau)}{t^2} \left(\sum_{i=1}^{\tau} (F^*(Y_i) - G^*(Y_i))^2 \right. \\
 &\quad \left. + \sum_{j=\tau+1}^t (F^*(Y_j) - G^*(Y_j))^2 \right), \quad (19.17)
 \end{aligned}$$

where F^* and G^* are the empirical distributions associated with the samples $\{Y_1, Y_2, \dots, Y_\tau\}$ and $\{Y_{\tau+1}, Y_{\tau+2}, \dots, Y_t\}$. Large values of C_τ means $\{Y_1, Y_2, \dots, Y_\tau\}$ and $\{Y_{\tau+1}, Y_{\tau+2}, \dots, Y_t\}$ are possibly coming from different cdfs. Suppose $P_\tau^{(CvM)}(t)$ is the p-value of the two-sample Cramer-von Mises test. The details of computing $P_\tau^{(CvM)}(t)$ are provided in Sect. 19.3.5 along with other computational technicalities of the proposed chart. A *natural approach* in this framework is to (i) calculate the p-values of two-sample Cramer-von Mises test for all possible values of τ , i.e., $\tau = 2, 3, \dots, t-2$, (ii) determine $\tau^*(t) = \arg \min_{\tau \in \{2, 3, \dots, (t-2)\}} P_\tau^{(CvM)}(t)$, and (iii) use $P_{\tau^*}^{(CvM)}(t)$ as the charting statistic.

However, there are *two major drawbacks* of this approach. *The first drawback* is that the determination of $\tau^*(t)$ can be computationally expensive because we have to compute $P_\tau^{(CvM)}(t)$ for all possible values of τ , i.e., $2, 3, \dots, (t-2)$. Moreover, we need to execute similar procedures for all values of $t \geq S_0$ until we get a signal. We run into this issue in the change-point framework wherever we need to compute the relevant statistic for each value of the possible change-point. To reduce computation, we consider the following two techniques and both can be applied to the proposed SPC chart. The first technique is to estimate $\tau^*(t)$ efficiently by the method proposed in Sect. 19.2.1 which can be applied to other change-point problems of similar nature. The second technique is the pruning of data from distant past based on p-values as proposed by Mukherjee [28]. *The second drawback* is the weakness of the power of Cramer-von Mises test to detect a change in standard deviation or scale. The proposed chart addresses this issue by integrating Cramer-von Mises test with Ansari-Bradley test which is a nonparametric test to detect scale difference. We describe these approaches below.

19.3.1 A Computationally Efficient Approach to Estimate $\tau^*(t)$

We note that the values of $P_\tau^{(CvM)}(t)$ and $P_{\tau'}^{(CvM)}(t)$ should not be too different as long as τ and τ' are close. In other words, for each t , the values of $P_\tau^{(CvM)}(t)$ for different τ are strongly autocorrelated. Therefore, we can make the procedure faster by first calculating $P_\tau^{(CvM)}(t)$ for τ -values that are multiples of $\lfloor \sqrt{t} \rfloor$ instead of all possible values of τ . For a demonstration purpose, consider the case when $t = 100$. Calculate $P_\tau^{(CvM)}(t)$ for $\tau = 10, 20, \dots, 90$ instead of $\tau = 2, 3, 4, \dots, 98$. Then, select the value of τ for which $P_\tau^{(CvM)}(t)$ is minimum, and calculate $P_\tau^{(CvM)}(t)$ for τ -values in the two adjoining intervals. In other words, if $\tau = 30$ gives the smallest $P_\tau^{(CvM)}(t)$ -value among $\tau = 10, 20, 30, \dots, 90$, compute $P_\tau^{(CvM)}(t)$ for $\tau = 21, 22, \dots, 29, 30, 31, \dots, 39$ and select the τ -value for which $P_\tau^{(CvM)}(t)$ is the smallest. The detailed procedure to estimate $\tau^*(t)$ is described below.

For each t , instead of calculating $P_\tau^{(CvM)}(t)$ for possible values of τ , i.e., $2, 3, \dots, (t-2)$, calculate the statistic for $\tau = i \cdot \lfloor \sqrt{t} \rfloor$, where possible values of i are $1, 2, \dots, I(t)$, where $I(t)$ is the largest integer for which $I(t) \cdot \lfloor \sqrt{t} \rfloor \leq (t-2)$. Here, $\lfloor \sqrt{t} \rfloor$ is the largest integer smaller than or equal to t . Since we start monitoring when the time-point $t \geq S_0 \geq 4$, we always have at least one positive integer i for which we can compute $P_\tau^{(CvM)}(t)$. For example, when $t = S_0 = 4$, we have to compute $P_\tau^{(CvM)}(t)$ for only one value of τ , i.e., when $\tau = 2$. Next, we find $\tilde{\tau} = \arg \min_{i \in \{1, 2, \dots, I(t)\}} P_{i \cdot \lfloor \sqrt{t} \rfloor}^{(CvM)}$. Since the $\tau^*(t)$ should be close to $\tilde{\tau}$, we calculate $P_\tau^{(CvM)}(t)$ for all integer values of τ within $[\max\{\tilde{\tau} - \lfloor \sqrt{t} \rfloor + 1, 2\}, \min\{\tilde{\tau} + \lfloor \sqrt{t} \rfloor - 1, (t-2)\}]$ and pick the integer within that interval for which $P_\tau^{(CvM)}(t)$ is minimum. This is our estimated $\tau^*(t)$ and we call it $\tilde{\tau}^*(t)$. The method is summarized as follows:

- (i) Calculate $P_\tau^{(CvM)}(t)$ for $\tau = i \cdot \lfloor \sqrt{t} \rfloor$, where possible values of i are $1, 2, \dots, I(t)$, where $I(t)$ is the largest integer for which $I(t) \cdot \lfloor \sqrt{t} \rfloor \leq (t-2)$.
- (ii) Find $\tilde{\tau} = \arg \min_{i \in \{1, 2, \dots, I(t)\}} P_{i \cdot \lfloor \sqrt{t} \rfloor}^{(CvM)}$.
- (iii) Calculate $P_\tau^{(CvM)}(t)$ for all integer values of τ within the interval:

$$[\max\{\tilde{\tau} - \lfloor \sqrt{t} \rfloor + 1, 2\}, \min\{\tilde{\tau} + \lfloor \sqrt{t} \rfloor - 1, (t-2)\}].$$

- (iv) Estimate $\tau^*(t)$ by the integer within the interval in (iii) for which $P_\tau^{(CvM)}(t)$ is minimum.

19.3.2 Integration of Ansari-Bradley Test with Cramer-von Mises Test

It is well documented in the literature (e.g., [27]) that Cramer-von Mises test does not have high power to detect changes in scale parameters. However, it has high power to detect changes in location parameters. Because of this weakness, we integrate Cramer-von Mises test with Ansari-Bradley test [70] in the proposed control chart. Ansari-Bradley test is a nonparametric test based on rank sum to detect differences in scale parameters. The integration procedure is as follows. For each t , once $\tau^*(t)$ is estimated by the procedure in Sect. 19.2.1:

- (i) Record the p-value of two-sample Cramer-von Mises test for checking if $\{Y_1, Y_2, \dots, Y_{\tilde{\tau}^*}\}$ and $\{Y_{\tilde{\tau}^*+1}, Y_{\tilde{\tau}^*+2}, \dots, Y_t\}$ are realizations from the same continuous cdf. Obviously, it is $P_{\tilde{\tau}^*}^{(CvM)}(t)$.
- (ii) Perform Ansari-Bradley two-sample test for checking if the scale parameters of $\{Y_1, Y_2, \dots, Y_{\tilde{\tau}^*}\}$ and $\{Y_{\tilde{\tau}^*+1}, Y_{\tilde{\tau}^*+2}, \dots, Y_t\}$ are same. Call the p-value $P_{\tilde{\tau}^*(t)}^{(AB)}$.

- (iii) Calculate $p^{(E)}(t) = \min\{P_{\hat{\tau}^*(t)}^{(CvM)}(t), P_{\hat{\tau}^*(t)}^{(AB)}(t)\}$, and use this as charting statistic.

When we integrate two different tests in a change-point-based chart, it is natural to perform change-point analysis for both tests, because the estimated change-points can be different for different tests. However, to avoid extra computation, we do not consider this approach. Numerical studies show that the proposed chart based on the procedures as described performs well in detecting changes in scale parameters. When there is a small change in scale parameters and the p-value of Cramer-von Mises test is not small enough to give a signal, it is still small and hence the change-points based on both tests should be close. The proposed chart just detects such change by integrating a more powerful test appropriately.

19.3.3 Data Pruning Based on P-Values

In case the proposed chart does not detect any distributional change for a long time, the sequence of “in-control” observations will be long, and hence calculation of $\hat{\tau}^*(t)$ as described in Sect. 19.2.1 can be time-consuming as well. If the influx of observations is rapid, then this chart in its present form cannot be used to monitor such processes. To solve this problem, consider datapruning from distant past. If $p^{(E)}(t)$, calculated by the procedure described in Sect. 19.2.2 is large, it is very unlikely that a distributional change has taken place. Therefore, we prune a few observations from distant past and focus on more recent observations. In this way, we make sure that the sequence of “in-control” observations does not become too long. We provide the description of the method below.

When $\tilde{t} \geq T_0$, where \tilde{t} is the current length of the sequence of observations, and T_0 is a threshold parameter of the SPC chart, we consider the possibility of data pruning. If $p^{(E)}(t) \geq P_0$, we prune the oldest $C(\tilde{t}, p^{(E)}(t), P_0) = \left\lceil \tilde{t} \cdot \min\left(0.2, \left(\frac{p^{(E)}(t) - P_0}{1 - P_0}\right)\right) \right\rceil$ number of observations. Here, we make sure not to prune too much in one step, by specifying that we cannot prune more than 20% of the current length of the sequence at one step. Here, we set the maximum amount of pruning at one step to be 20% based on the performance of the chart in our simulation experiments in terms of how fast it can compute and how early it can give a signal when a process goes “out-of-control.” While one can introduce a parameter for the maximum percentage we can prune at one step and select its value based on a reasonable criterion, we set it equal to 20% for simplicity. Once we prune, we can estimate $\tau^*(t)$ faster in the next step, i.e., after the arrival of next observation. The summary of data pruning is as follows:

- (i) When $p^{(E)}(t) < P_0$, give a signal for distributional change and stop process monitoring; otherwise, go to Step (ii).
- (ii) If $\tilde{t} \geq T_0$, prune the oldest $C(\tilde{t}, p^{(E)}(t), P_0) = \left\lceil \tilde{t} \cdot \min\left(0.2, \left(\frac{p^{(E)}(t) - P_0}{1 - P_0}\right)\right) \right\rceil$ number of observations, and collect the next observation. Otherwise, go to Step (iii).
- (iii) Do not prune and collect the next observation.

19.3.4 The Algorithm of the Proposed Control Chart

The summary of the procedures to run the proposed control chart is as follows:

1. When $\tilde{t} < S_0$, collect the next observation. Otherwise, go to Step 2.
2. Calculate $\hat{\tau}^*(t)$, an estimate of $\tau^*(t)$ by the method described in Sect. 19.2.1. Go to Step 3.
3. Calculate $P_{\hat{\tau}^*(t)}^{(CvM)}(t)$, $P_{\hat{\tau}^*(t)}^{(AB)}(t)$, and $p^{(E)}(t)$ by the method described in Sect. 19.2.2. If $p^{(E)}(t) < P_0$, give a signal for distributional change and stop process monitoring; otherwise, go to Step 4.
4. Perform the data pruning procedure as described in Sect. 19.2.3. Go to Step 1.

19.3.5 Implementation

Calculation of p-values for two-sample Cramer-von Mises tests is computationally expensive and time-consuming especially when the sample sizes are large [71]. However, [72] provide the asymptotic distribution of two-sample Cramer-von Mises criterion. We also note that the convergence rate is rapid. Therefore, in our implementation, we use asymptotic p-values rather than the exact ones. Using the software R (<https://www.r-project.org/>) and the R-package *CvM2SL2Test* developed by Xiao [73], we extend the table provided by Anderson and Darling [72] to calculate p-values of two-sample Cramer-von Mises tests when they are smaller than 0.01. We use the same R-package *CvM2SL2Test* to calculate the two-sample Cramer-von Mises test statistics given as in (1) and then calculate their asymptotic p-values. For calculating (1), the R-package uses a C++ program developed by Xiao et al. [74]. We use R-function *ansari.test()* to perform Ansari-Bradley tests. It should be noted that approximate p-values are calculated even for small samples in presence of ties. The proposed method is designed to monitor univariate continuous

processes and hence cannot handle even a moderate number of ties in the observation sequence. However, this chart can still perform well in presence of a small number of ties.

19.4 Numerical Studies

We perform several numerical studies to evaluate the performance of the proposed method in comparison with a number of state-of-the-art change-point-based control charts. Our main goal is to find a chart that shows overall better performance to detect arbitrary distributional changes. We also study how the performances are affected by the number of “in-control” observations before the distributional change. Like most research articles in the literature, we also evaluate the performances of the control charts by the average number of observations, called “run length” after the first observation from the changed distribution is obtained. We call it ARL_1 . The smaller the value of ARL_1 , the better the performance. Of course, we set the control limits of all control charts such that the average number of “in-control” observations required to give a false signal when no distributional change takes place is a prefixed large number. We call it ARL_0 . In all our numerical simulations, we set $ARL_0 = 500$ unless mentioned otherwise.

We compare the proposed method with four other change-point-based methods. The first method is proposed by Hawkins and Deng [13]. It is based on Mann-Whitney test and it aims to detect location shift. We call this method MW. The next competing control chart, proposed by Ross et al. [33], is based on Mood test [75]. This chart aims to detect scale change in the process distribution. We call it Mood. The third competing chart, also proposed by Ross et al. [33], is based on Lepage test [76]. It aims to detect changes in both location and scale of the process distribution, and we call it Lepage. Our fourth competing chart, proposed by Ross and Adams [27], aims to detect arbitrary distributional change using Cramer-von Mises test. We call this chart CvM.

Initially, we consider three “in-control” process distributions: standard normal, $N(0, 1)$; standardized t-distribution with 2.5 degrees of freedom, $ST(2.5)$; and standardized log-normal distribution with parameters 1.0 and 0.5, $SLN(1.0, 0.5)$. Note that the mean and standard deviation of log-normal distribution with parameters 1.0 and 0.5 are 3.08 and 1.64, respectively. However, we approximate those by 3 and 1.6 while standardizing. We first consider shifts in location only, scale only, and both location and scale simultaneously. Finally, we consider arbitrary changes of various distributions. To compare the performances of various methods, we consider two cases: when the distribution change occurs early, right after

time-point $\tau = 50$, and when the change occurs late, right after time-point $\tau = 300$. It should be noted that if a false signal occurs before the actual distributional change, we disregard that sequence in our simulation, as it is a reasonable practice.

We use the R-package *cpm* to run the competing charts. For the proposed method, we need to select S_0 and T_0 . For all methods, we select the startup time to be 20, and for the proposed method, we select $S_0 = 20$ and $T_0 = ARL_0 = 500$. For comparing the performances of the methods when the distributional change is arbitrary, we run four versions of the proposed chart called P_{250} , P_{500} , P_{1000} , and P_∞ when $T_0 = 250, 500, 1000$, and ∞ , respectively. From Table 19.6, we see that $T_0 = 500$ is a reasonable choice considering the fact that larger values of T_0 requires more time for computation. It should be noted that the proposed chart that is designed to detect an arbitrary distributional change may not outperform the charts designed to detect a specific type of distributional changes when the actual distributional change is of that particular type. For example, MW should perform better than the proposed chart when the distributional change involves location change only. However, our goal is to find a chart that performs well to detect all types of distributional changes so that it can be used in various applications where the natures of changes are unknown.

19.4.1 Location Changes

First, we focus on changes in location only. We consider four different amounts of shift, $\delta_L = 0.25, 0.5, 1.0$, and 2.0 . For each of three distributions $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$, we generate 50,000 sequences of observations where the post-change observations right after time-point τ are calculated by adding δ_L . From Table 19.1, we observe that

- MW and CvM are slightly better than the proposed method when δ_L is small or moderate. When δ_L is large, all methods except Mood perform well. The reason is that MW is designed to detect location changes only; therefore, it has better power when detecting location changes compared to the methods designed to detect arbitrary distributional changes. Mood is designed to detect scale changes; therefore, it cannot perform well in this case.
- Detections of location changes are faster when the “in-control” distribution is $ST(2.5)$ or $SLN(1.0, 0.5)$ compared to the $N(0, 1)$ case.
- Changes occurring right after time-point $\tau = 300$ are easier to detect compared to the changes occurring right after $\tau = 50$.

Table 19.1 Mean delay in detection of various location shifts δ_L occurring right after time τ . “In-control” distributions considered are $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$. The results are based on 50,000 random simulations when $ARL_0 = 500$

	$\tau = 50$				$\tau = 300$			
	$\delta_L =$				$\delta_L =$			
$N(\delta_L, 1)$	0.25	0.50	1.00	2.00	0.25	0.50	1.00	2.00
MW	377.1	134.3	13.8	4.9	150.8	35.1	10.2	3.9
Mood	502.6	506.3	310.5	32.9	479.3	399.0	71.2	4.3
Lepage	434.4	225.6	20.0	3.7	228.6	49.5	11.5	3.1
CvM	387.7	153.9	14.8	4.3	168.3	37.8	10.7	3.9
Proposed	400.0	169.6	16.0	4.8	190.5	41.5	12.3	5.3
$ST(2.5) + \delta_L$								
MW	199.8	21.4	6.0	3.5	46.8	13.3	5.2	3.1
Mood	492.1	346.5	72.1	11.5	446.0	180.0	7.9	2.5
Lepage	299.2	38.3	6.4	2.4	74.9	16.9	4.5	2.1
CvM	195.5	20.6	5.7	3.2	46.3	13.1	4.9	3.0
Proposed	212.1	22.3	6.2	3.6	50.3	14.6	6.4	4.2
$SLN(1.0, 0.5) + \delta_L$								
MW	368.4	96.0	10.9	4.3	113.8	26.4	8.4	3.8
Mood	527.3	272.4	71.5	45.9	406.2	171.3	139.3	5.6
Lepage	411.3	93.9	15.9	4.4	143.3	37.6	11.9	3.3
CvM	378.1	103.3	10.7	4.0	122.3	27.0	8.3	3.6
Proposed	356.9	93.0	11.2	4.4	120.8	28.2	9.6	5.0

Table 19.2 Mean delay in detection of various scale changes δ_S occurring right after time τ . “In-control” distributions considered are $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$. The results are based on 50,000 random simulations when $ARL_0 = 500$

	$\tau = 50$				$\tau = 300$			
	$\delta_S =$				$\delta_S =$			
$N(0, \delta_S^2)$	1.50	2.00	0.50	0.25	1.50	2.00	0.50	0.25
MW	320.2	226.4	753.0	843.5	154.0	73.7	1470.1	1929.3
Mood	94.6	17.1	37.2	11.2	26.9	10.6	21.3	10.1
Lepage	144.6	25.8	63.3	15.4	34.1	13.2	31.0	14.8
CvM	320.5	196.6	563.9	69.4	131.6	47.2	102.8	30.4
Proposed	195.2	37.8	84.6	13.2	59.1	24.4	31.9	13.8
$ST(2.5).\delta_S$								
MW	359.4	273.6	704.5	816.5	205.1	107.8	1272.2	1798.7
Mood	205.0	49.4	79.8	14.0	49.2	17.3	31.0	12.3
Lepage	246.8	85.3	141.3	19.7	64.0	21.3	45.6	17.9
CvM	348.9	252.5	613.3	140.6	182.4	68.6	161.4	38.7
Proposed	275.4	80.1	144.9	17.2	89.4	33.0	42.5	16.1
$SLN(1.0, 0.5).\delta_S$								
MW	274.5	167.1	671.4	660.7	102.7	47.1	721.6	485.9
Mood	66.2	13.8	28.8	12.0	20.4	8.8	20.1	11.3
Lepage	95.9	17.3	42.6	15.3	25.0	10.7	27.3	15.3
CvM	268.0	123.5	421.5	47.3	83.9	33.8	65.4	24.2
Proposed	142.1	27.3	55.7	12.6	45.1	20.3	26.7	13.5

19.4.2 Scale Changes

Now, we focus on changes in scale only. We consider four scale changes, namely, $\delta_S = 1.50, 0.50, 2.00$, and 0.25 . For each of three distributions, $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$, we generate 50,000 sequences of observations where the post-change observations are calculated by multiplying δ_S . From Table 19.2, we observe that

- Mood and Lepage are slightly better than the proposed method in some cases. The reason is that Mood is designed

to detect scale changes only and Lepage is designed to detect both location and scale change. Therefore, they have better powers when detecting scale changes compared to the methods designed to detect arbitrary distributional changes. MW is designed to detect location changes; therefore, it cannot perform well in this case. The proposed method performs much better than CvM. The reason is the incorporation of Ansari-Bradley test with Cramer-von Mises test.

- MW performs fairly well when $\delta_S > 1.00$ but performs very poorly when $\delta_S < 1.00$. One explanation of this is

when $\delta_S > 1.00$, extremely large or small numbers are more likely, and MW interprets those as location changes. This phenomenon is commented in [27] and in [24].

- Detections of scale changes are faster when the “in-control” distribution is $ST(2.5)$ and $SLN(1.0, 0.5)$ compared to the $N(0, 1)$ case.
- In this case also, changes occurring right after time-point $\tau = 300$ are easier to detect compared to the changes occurring right after $\tau = 50$. MW, however, cannot detect a decrease in scale at all.

- Lepage performs the best, as it should because it is designed to detect changes in location and scale simultaneously. Mood works well in some cases. The proposed method works well, better than CvM in many cases.
- In these case also, detections of changes in location and scale simultaneously are faster when the “in-control” distribution is $ST(2.5)$ and $SLN(1.0, 0.5)$ compared to $N(0, 1)$ case.
- Changes occurring right after time-point $\tau = 300$ are easier to detect compared to the changes occurring right after $\tau = 50$. MW, however, cannot detect a decrease in scale at all.

19.4.3 Location-Scale Changes

Now, we focus on changes in location and scale simultaneously. We consider 8 changes in location and scale simultaneously, namely, $(\delta_L, \delta_S) = (0.50, 1.50), (0.50, 0.50), (0.50, 2.00), (0.50, 0.25), (1.00, 1.50), (1.00, 0.50), (1.00, 2.00),$ and $(1.00, 0.25)$. For each of the three distributions $N(0, 1), ST(2.5),$ and $SLN(1.0, 0.5)$, we generate 50,000 sequences of observations where the post-change observations are calculated by multiplying δ_S and then adding δ_L . From Tables 19.3, 19.4, and 19.5, we observe that

19.4.4 General Distributional Changes

Finally, we consider general changes of various “in-control” distributions. The corresponding changes in the pair of mean and standard deviations are large in some cases (e.g., Weibull(1) to Weibull(3) and vice versa, Gamma(2,2) to Gamma(3,2) and vice versa) and small or zero in some other cases (e.g., $N(0, 1)$ to $ST(2.5)$ and vice versa, $N(0, 1)$ to $SLN(1.0, 0.5)$ and vice versa). From Table 19.6, we observe that

Table 19.3 Mean delay in detection of various amounts of changes in location and scale simultaneously, i.e., (δ_L, δ_S) , occurring right after time τ . “In-control” distribution considered here is $N(0, 1)$. The results are based on 50,000 random simulations when $ARL_0 = 500$

$N(\delta_L, \delta_S^2)$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 50$ and $(\delta_L, \delta_S) =$								
MW	137.1	125.4	127.2	129.0	20.2	9.4	28.3	8.3
Mood	116.6	45.6	18.2	20.4	168.0	62.1	23.3	44.5
Lepage	90.4	31.2	20.7	16.8	19.7	14.0	12.5	12.0
CvM	131.0	66.9	94.6	19.9	20.9	8.7	24.0	6.6
Proposed	116.4	33.1	33.5	14.0	21.7	9.2	19.0	7.1
$\tau = 300$ and $(\delta_L, \delta_S) =$								
$N(\delta_L, \delta_S^2)$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
MW	38.7	26.9	37.1	23.7	13.6	7.5	16.4	6.6
Mood	22.5	52.0	9.9	26.9	14.0	148.4	8.3	136.8
Lepage	21.1	27.7	11.3	18.9	10.2	10.7	8.3	9.9
CvM	37.9	20.7	30.4	13.1	13.9	7.0	15.6	5.6
Proposed	33.9	20.8	20.8	13.8	15.1	8.5	14.6	7.0

Table 19.4 Mean delay in detection of various amounts of changes in location and scale simultaneously, i.e., (δ_L, δ_S) , occurring right after time τ . “In-control” distribution considered here is $ST(2.5)$. The results are based on 50,000 random simulations when $ARL_0 = 500$

$ST(2.5) \cdot \delta_S + \delta_L$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 50$ and $(\delta_L, \delta_S) =$								
MW	33.9	13.4	47.4	10.8	7.7	4.7	10.0	4.3
Mood	241.9	66.4	70.1	41.8	110.2	52.5	87.0	46.3
Lepage	57.1	18.5	34.1	14.2	7.4	5.1	7.9	4.7
CvM	31.4	11.0	36.7	7.9	7.4	4.3	9.3	3.8
Proposed	33.3	11.6	30.6	8.3	7.9	4.7	9.3	4.2
$\tau = 300$ and $(\delta_L, \delta_S) =$								
MW	17.3	9.4	20.5	7.9	6.4	4.1	7.9	3.8
Mood	25.0	166.7	12.5	120.7	7.6	6.9	6.8	6.0
Lepage	14.7	14.6	11.6	12.7	5.1	3.7	5.4	3.4
CvM	16.8	8.5	18.7	6.5	6.2	3.8	7.5	3.4
Proposed	17.8	9.9	17.3	7.9	7.7	5.3	8.7	4.8

Table 19.5 Mean delay in detection of various amounts of changes in location and scale simultaneously, i.e., (δ_L, δ_S) , occurring right after time τ . “In-control” distribution considered here is $SLN(1.0, 0.5)$. The results are based on 50,000 random simulations when $ARL_0 = 500$

$(SLN(1, .5)).\delta_S + \delta_L$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 50$ and $(\delta_L, \delta_S) =$								
MW	201.5	40.7	216.0	33.9	19.8	7.9	44.2	7.4
Mood	273.6	32.2	28.5	24.9	384.2	47.4	140.2	42.5
Lepage	217.9	19.1	45.0	15.2	33.3	11.7	48.1	10.7
CvM	206.1	17.9	185.3	11.1	21.6	6.6	42.2	5.6
Proposed	224.4	14.8	67.7	10.5	23.8	7.0	44.7	6.1
$\tau = 300$ and $(\delta_L, \delta_S) =$								
MW	53.3	15.1	65.2	13.0	12.9	6.4	20.2	5.9
Mood	69.0	55.8	13.3	48.3	154.7	137.8	18.9	133.5
Lepage	44.7	19.8	16.1	17.5	15.7	9.3	14.6	8.6
CvM	54.4	11.6	47.8	8.6	13.6	5.5	20.1	4.9
Proposed	55.7	12.6	28.4	9.8	15.3	7.0	20.6	6.3

Table 19.6 Mean delay in detection of various general changes in distribution. The results are based on 50,000 random simulations when $ARL_0 = 500$

	MW	Mood	Lepage	CvM	P_{250}	P_{500}	P_{1000}	P_∞
$\tau = 50$								
$N(0, 1) \rightarrow ST(2.5)$	685.7	109.4	192.2	617.5	173.1	178.1	173.4	173.1
$ST(2.5) \rightarrow N(0, 1)$	299.6	101.5	149.7	272.3	125.8	109.9	100.8	96.1
$N(0, 1) \rightarrow SLN(1, 0.5)$	555.5	539.1	544.9	516.2	455.8	478.0	487.8	499.1
$SLN(1, 0.5) \rightarrow N(0, 1)$	430.7	376.9	366.8	416.7	423.0	418.8	412.5	412.5
$\text{Gamma}(2,2) \rightarrow \text{Gamma}(3,2)$	46.8	466.0	86.3	55.0	63.0	60.2	60.5	62.2
$\text{Gamma}(3,2) \rightarrow \text{Gamma}(2,2)$	39.9	464.5	76.8	46.4	56.6	52.9	51.4	52.2
$\text{Weibull}(1) \rightarrow \text{Weibull}(3)$	657.5	18.3	23.2	177.7	25.5	24.9	24.0	23.6
$\text{Weibull}(3) \rightarrow \text{Weibull}(1)$	117.8	8.0	9.7	52.7	15.7	15.2	14.5	14.2
$\text{Uniform}(0,1) \rightarrow \text{Beta}(5,5)$	794.2	19.5	28.5	385.3	34.5	34.4	32.5	32.4
$\text{Beta}(5,5) \rightarrow \text{Uniform}(0,1)$	178.5	9.7	12.5	119.6	20.1	19.4	18.5	17.9
$\tau = 300$								
$N(0, 1) \rightarrow ST(2.5)$	1168.9	37.4	55.9	190.5	51.3	47.2	46.2	45.7
$ST(2.5) \rightarrow N(0, 1)$	129.9	23.0	28.7	79.4	40.6	37.4	36.3	35.9
$N(0, 1) \rightarrow SLN(1, 0.5)$	643.8	462.1	458.5	449.7	356.8	346.7	373.0	419.4
$SLN(1, 0.5) \rightarrow N(0, 1)$	316.1	168.9	178.0	265.0	322.5	277.9	269.6	288.0
$\text{Gamma}(2,2) \rightarrow \text{Gamma}(3,2)$	19.7	316.8	27.4	21.0	23.8	22.9	22.4	22.2
$\text{Gamma}(3,2) \rightarrow \text{Gamma}(2,2)$	18.7	155.3	21.7	20.2	22.9	22.2	21.8	21.5
$\text{Weibull}(1) \rightarrow \text{Weibull}(3)$	512.2	15.7	20.7	37.9	19.7	20.3	19.8	19.6
$\text{Weibull}(3) \rightarrow \text{Weibull}(1)$	31.1	5.9	7.1	21.8	13.4	14.2	14.0	13.8
$\text{Uniform}(0,1) \rightarrow \text{Beta}(5,5)$	1670.0	15.2	22.1	60.1	21.3	22.1	21.7	21.4
$\text{Beta}(5,5) \rightarrow \text{Uniform}(0,1)$	48.9	7.0	8.7	30.2	16.5	17.3	16.8	16.6

- No one method is uniformly best. P_{500} works well in all cases. Note that P_{500} is the proposed method when $T_0 = 500$ as defined in the fourth paragraph of Sect. 19.3.
- In cases where another method is the best, P_{500} is not far behind.
- The performance of P_{500} is not far from the best choice of T_0 for the proposed method. The larger the value of T_0 , the more the computing time. Therefore, P_{500} is a good balance between computing time and performance. Hence, we suggest using $T_0 = ARL_0$ in most applications.
- In cases where location change is large, MW works well, and when scale change is large, Mood works well, as expected.
- Distributional change from $N(0, 1)$ to $SLN(1.0, 0.5)$ which does not alter the mean and standard deviation is detected by the proposed method earlier than other charts.
- When we are not sure about the nature of possible distributional change, the proposed method is a good choice.
- Changes occurring right after time-point $\tau = 300$ are easier to detect compared to the changes occurring right after $\tau = 50$. MW cannot detect an arbitrary distributional change well if the median does not change by much.

Table 19.7 Mean delay in detection of various location shifts δ_L occurring right after time τ . “In-control” distributions considered are $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$. The results are based on 10,000 random simulations when $ARL_0 = 200$

	$\tau = 30$				$\tau = 150$			
	$\delta_L =$				$\delta_L =$			
$N(\delta_L, 1)$	0.25	0.50	1.00	2.00	0.25	0.50	1.00	2.00
MW	165.4	85.4	13.3	3.8	102.5	29.2	8.5	3.2
Mood	195.3	198.3	156.4	36.9	185.2	161.5	54.4	2.6
Lepage	150.6	100.4	18.9	3.1	112.7	39.0	9.2	2.3
CvM	166.4	95.1	15.2	3.7	106.6	32.5	8.9	3.1
Proposed	166.5	103.0	16.8	4.3	116.0	37.6	10.2	4.0
<i>ST(2.5) + δ_L</i>								
MW	110.0	24.3	5.2	3.0	40.0	11.1	4.2	2.7
Mood	191.0	154.4	52.6	17.4	178.8	105.4	8.8	2.0
Lepage	118.5	35.7	6.1	1.9	57.5	14.2	3.8	1.5
CvM	110.5	22.8	5.0	2.6	40.2	10.9	4.0	2.3
Proposed	119.1	26.1	5.6	3.2	46.6	12.2	4.9	3.1
<i>SLN(1.0, 0.5) + δ_L</i>								
MW	167.9	78.9	11.0	3.7	94.0	22.9	7.0	3.1
Mood	205.1	150.6	52.7	34.8	202.9	106.1	76.1	6.0
Lepage	153.0	69.9	13.6	4.1	101.8	30.1	9.8	2.7
CvM	165.8	82.5	10.6	3.4	95.0	23.5	7.0	2.9
Proposed	158.7	72.8	11.2	4.0	89.5	24.8	7.9	3.7

Table 19.8 Mean delay in detection of various scale changes δ_S occurring right after time τ . “In-control” distributions considered are $N(0, 1)$, $ST(2.5)$, and $SLN(1.0, 0.5)$. The results are based on 10,000 random simulations when $ARL_0 = 200$

	$\tau = 30$				$\tau = 150$			
	$\delta_S =$				$\delta_S =$			
$N(0, \delta_S^2)$	1.50	2.00	0.50	0.25	1.50	2.00	0.50	0.25
MW	129.2	93.5	299.8	340.8	73.4	39.8	561.6	739.2
Mood	60.7	16.8	39.8	9.6	20.1	8.0	17.5	8.1
Lepage	61.8	20.4	61.8	13.1	22.7	9.2	25.8	11.9
CvM	128.0	87.7	253.5	97.2	68.2	31.5	106.2	26.9
Proposed	94.3	31.6	70.2	13.7	45.7	18.9	26.9	10.9
<i>ST(2.5).δ_S</i>								
MW	145.2	111.4	284.3	332.1	92.2	54.1	481.6	685.7
Mood	103.4	41.3	68.7	12.4	37.6	13.8	25.3	9.9
Lepage	90.7	46.1	97.4	17.3	41.9	16.0	38.4	14.3
CvM	144.2	105.1	256.2	151.4	88.0	44.1	173.5	34.5
Proposed	119.5	55.5	95.9	18.3	68.0	26.0	35.3	13.2
<i>SLN(1.0, 0.5).δ_S</i>								
MW	113.8	75.2	281.6	296.5	55.7	29.0	381.0	357.7
Mood	47.6	13.9	31.9	10.1	15.6	6.6	16.4	9.1
Lepage	46.3	13.9	43.8	12.8	16.9	7.4	22.2	12.2
CvM	110.3	65.5	217.7	69.3	51.9	22.7	64.0	21.0
Proposed	76.1	22.8	53.8	12.3	34.8	15.5	22.5	10.6

Numerical simulations when $ARL_0 = 200$ are provided in Tables 19.7, 19.8, 19.9, 19.10, 19.11, and 19.12. Similar conclusions as provided above can be drawn from these tables as well.

From these simulation studies, we see that the proposed method works well in most applications. When we are trying to detect a specific type of distributional change, its performance is slightly worse than the chart that is designed to detect that particular type of changes. However, in most cases, the differences are not much. Therefore,

when the nature of distributional change is unknown, the proposed chart can be used with anticipation of a good performance.

19.5 Analysis of Various Real-World Data

Now, we focus on applications of the proposed chart on real-world problems. We consider two datasets: a climate data and a blood glucose monitoring data.

Table 19.9 Mean delay in detection of various changes in location and scale simultaneously, i.e., (δ_L, δ_S) occurring right after time τ . “In-control” distribution considered here is $N(0, 1)$. The results are based on 10,000 random simulations when $ARL_0 = 200$

$N(\delta_L, \delta_S^2)$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 30$ and $(\delta_L, \delta_S) =$								
MW	76.2	99.7	68.0	103.7	19.4	9.5	24.4	8.7
Mood	71.2	41.9	18.4	15.9	93.2	42.9	24.4	29.9
Lepage	45.7	29.2	17.2	13.3	17.8	12.0	11.0	9.8
CvM	75.8	71.1	59.2	26.4	20.2	8.4	22.2	6.2
Proposed	70.3	33.0	28.2	11.9	22.3	9.0	17.9	6.6
$\tau = 150$ and $(\delta_L, \delta_S) =$								
MW	29.3	24.7	25.0	22.2	10.9	6.2	12.6	5.5
Mood	17.5	37.6	7.5	20.0	11.4	82.8	6.5	76.1
Lepage	15.2	21.8	8.1	14.7	7.8	8.8	6.0	8.3
CvM	28.8	18.3	21.7	11.4	11.1	5.8	11.8	4.7
Proposed	27.0	17.6	16.3	11.3	12.2	6.9	11.4	5.5

Table 19.10 Mean delay in detection of various changes in location and scale simultaneously, i.e., (δ_L, δ_S) occurring right after time τ . “In-control” distribution considered here is $ST(2.5)$. The results are based on 10,000 random simulations when $ARL_0 = 200$

$ST(2.5).\delta_S + \delta_L$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 30$ and $(\delta_L, \delta_S) =$								
MW	32.4	15.1	38.8	12.7	6.9	4.1	8.9	3.8
Mood	111.1	47.4	51.9	28.6	74.8	36.9	59.3	32.0
Lepage	42.1	15.4	28.2	11.3	7.1	4.9	7.9	4.4
CvM	30.4	11.8	32.8	7.5	6.5	3.7	8.2	3.3
Proposed	34.5	11.5	29.5	7.7	7.2	4.3	8.4	3.8
$\tau = 150$ and $(\delta_L, \delta_S) =$								
MW	14.0	7.9	15.8	6.7	5.4	3.4	6.4	3.1
Mood	24.2	91.7	10.6	68.9	6.9	10.6	5.8	10.1
Lepage	12.0	12.1	9.2	10.5	4.3	3.1	4.3	2.8
CvM	13.4	7.1	14.4	5.4	5.1	3.1	6.2	2.8
Proposed	14.4	8.2	13.7	6.3	6.0	3.9	6.8	3.5

Table 19.11 Mean delay in detection of various changes in location and scale simultaneously, i.e., (δ_L, δ_S) occurring right after time τ . “In-control” distribution considered here is $SLN(1.0, 0.5)$. The results are based on 10,000 random simulations when $ARL_0 = 200$

$SLN(1, 0.5).\delta_S + \delta_L$	(.5, 1.5)	(.5, .5)	(.5, 2)	(.5, .25)	(1, 1.5)	(1, .5)	(1, 2)	(1, .25)
$\tau = 30$ and $(\delta_L, \delta_S) =$								
MW	101.6	49.1	96.5	43.4	20.7	8.0	38.1	7.3
Mood	119.9	24.2	26.0	18.1	176.4	31.9	76.3	27.8
Lepage	85.2	15.5	28.5	11.8	30.1	9.7	34.1	8.9
CvM	105.9	24.1	84.9	11.4	23.1	6.2	37.7	5.1
Proposed	110.3	13.6	47.4	9.0	27.6	6.5	40.9	5.6
$\tau = 150$ and $(\delta_L, \delta_S) =$								
MW	40.6	13.2	38.6	11.3	10.6	5.3	16.2	4.9
Mood	47.9	37.0	10.2	32.1	91.3	76.2	15.6	74.7
Lepage	31.8	15.6	11.0	13.6	12.8	7.8	11.2	7.2
CvM	42.5	9.9	32.5	7.3	11.3	4.6	16.2	4.0
Proposed	47.6	10.5	21.9	8.1	13.0	5.5	16.5	4.8

19.5.1 Climate Data on Minneapolis, USA

Monitoring for the changes of patterns of various climatological measurements such as maximum and minimum temperatures on a daily, monthly, and yearly basis, amounts of rainfall, snow, measurements of snow depth, number of rainy or snowy days is an emerging research area. In the literature, various statistical methods are demonstrated with the capacity of analyzing such data. Modern statistical process control (SPC) charts also deserve a chance to monitor such climatological variables. In this context, we consider

mean daily maximum temperature in Fahrenheit in the month of January in Minneapolis, USA. We collect the data from http://www.dnr.state.mn.us/climate/twin_cities/listings.html. The data are from 1873 to 2017 with no missing value and are presented in Fig. 19.1. At first, before applying the control charts, we check if our assumption of independence of the observations is reasonable. Durbin-Watson test (R function: *dwtest*, R package: *lmtest*) for two-sided alternative gives a high p-value 0.8359 showing lack of autocorrelation. Therefore, we can assume that the observations are independent. Now, we apply the

Table 19.12 Mean delay in detection of various general changes in distribution. The results are based on 10,000 random simulations when $ARL_0 = 200$

	MW	Mood	Lepage	CvM	P_{100}	P_{200}	P_{400}	P_{∞}
$\tau = 30$								
$N(0, 1) \rightarrow ST(2.5)$	276.8	79.9	107.6	256.3	104.8	105.6	109.4	113.2
$ST(2.5) \rightarrow N(0, 1)$	121.0	67.0	70.2	114.0	78.5	69.1	63.5	62.1
$N(0, 1) \rightarrow SLN(1, 0.5)$	218.2	210.8	176.9	206.8	173.4	183.4	186.7	192.0
$SLN(1, 0.5) \rightarrow N(0, 1)$	174.1	155.1	121.3	166.3	161.7	160.5	159.0	157.2
$\text{Gamma}(2,2) \rightarrow \text{Gamma}(3,2)$	45.2	191.7	59.2	50.2	58.4	56.7	58.6	55.9
$\text{Gamma}(3,2) \rightarrow \text{Gamma}(2,2)$	38.4	184.8	51.3	42.4	54.6	50.3	49.0	48.8
$\text{Weibull}(1) \rightarrow \text{Weibull}(3)$	286.1	16.9	21.6	153.6	27.3	26.4	25.9	26.0
$\text{Weibull}(3) \rightarrow \text{Weibull}(1)$	56.8	7.7	7.6	41.2	14.0	12.7	12.4	12.0
$\text{Uniform}(0,1) \rightarrow \text{Beta}(5,5)$	322.6	19.7	29.3	219.5	38.8	37.7	39.4	41.4
$\text{Beta}(5,5) \rightarrow \text{Uniform}(0,1)$	77.9	9.2	9.7	62.8	17.9	16.4	15.4	15.0
$\tau = 150$								
$N(0, 1) \rightarrow ST(2.5)$	444.9	30.4	45.0	194.6	60.6	40.7	40.7	39.7
$ST(2.5) \rightarrow N(0, 1)$	61.9	19.0	22.0	51.6	44.6	29.0	28.4	28.2
$N(0, 1) \rightarrow SLN(1, 0.5)$	252.4	202.7	168.9	209.6	156.4	157.2	174.9	196.9
$SLN(1, 0.5) \rightarrow N(0, 1)$	137.9	85.4	69.6	123.6	141.3	127.8	129.7	131.0
$\text{Gamma}(2,2) \rightarrow \text{Gamma}(3,2)$	16.4	151.0	22.2	17.7	23.7	20.0	19.3	18.8
$\text{Gamma}(3,2) \rightarrow \text{Gamma}(2,2)$	15.5	86.6	17.0	16.6	22.1	19.1	18.3	17.9
$\text{Weibull}(1) \rightarrow \text{Weibull}(3)$	340.9	12.7	16.7	34.2	17.6	16.6	16.0	16.1
$\text{Weibull}(3) \rightarrow \text{Weibull}(1)$	20.8	4.5	5.0	15.4	11.5	11.0	10.7	10.4
$\text{Uniform}(0,1) \rightarrow \text{Beta}(5,5)$	640.5	12.4	18.0	56.2	21.6	18.5	18.2	18.0
$\text{Beta}(5,5) \rightarrow \text{Uniform}(0,1)$	29.3	5.4	6.2	20.4	13.9	13.2	12.8	12.6

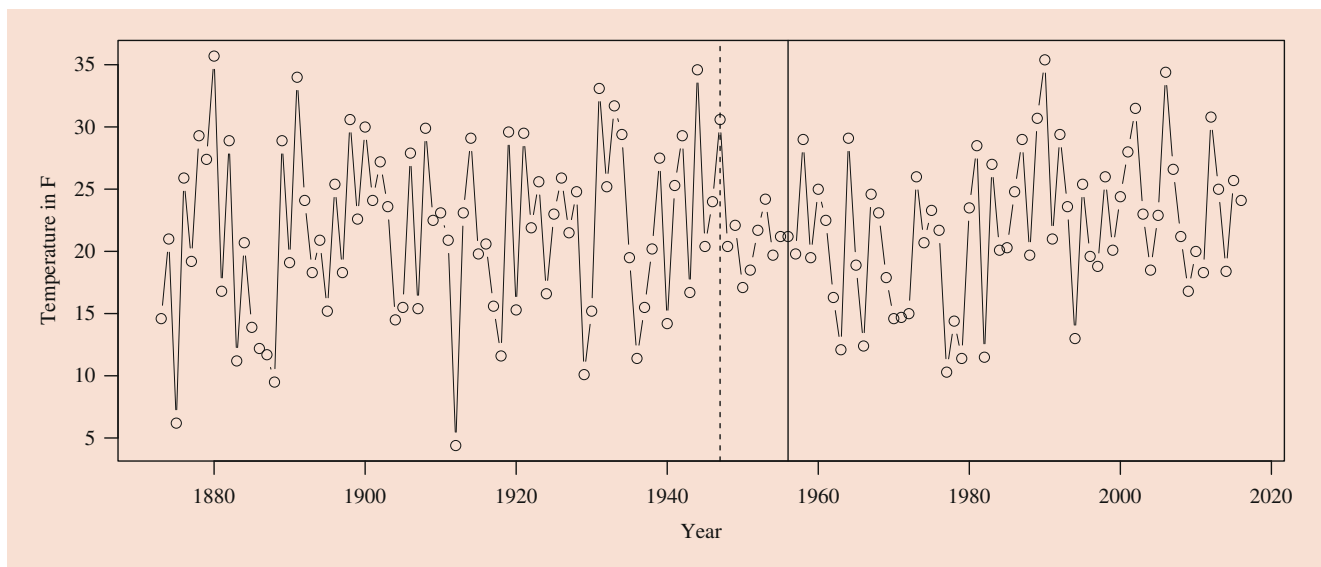


Fig. 19.1 Mean daily maximum temperature in the month of January in Minneapolis, USA. Setting $ARL_0 = 500$, the proposed chart detects a distributional change in the year 1956 estimating the change-point to be just after 1947. Other charts do not detect any distributional change

proposed method along with other competing methods when we set $ARL_0 = 500$. For the proposed method, we set $T_0 = ARL_0 = 500$ as per our suggestion mentioned before. The proposed method detects a distributional change at 1956 while the estimated change-point is 1947. All other methods, i.e., MW, Mood, Lepage, and CvM, do not detect any distributional change. We also run all SPC charts when we set $ARL_0 = 200$. In this case also, the proposed method detects a distributional change at 1956 while the estimated change-point is 1947. Mood also detects distributional change but at

a later time at 1960 while the estimated change point is 1947. MW, Lepage, and CvM still do not detect any distributional change. Now, we study the “in-control” distribution and the estimated “out-of-control” distribution. Table 19.13 shows the first four sample moments of “in-control” observations from 1873 to 1947 and “out-of-control” observations from 1948 to 1956. From Table 19.13, we see that the second moment changed a lot, but not the other three moments. We carry out the calculations using R-package *moments*.

19.5.2 Blood Glucose Monitoring Data

Monitoring blood glucose level on a daily basis is essential for advanced diabetic patients. It gives information on whether the particular lifestyle change and the treatment procedure including the medicine with its administered dosage are working well for the patient. Monitoring such measurements is complicated because we are not just focusing on the mean and the standard deviation only; we need to monitor the stability of its probability distribution as well. We collect a data from UCI Machine Learning Repository [77]. This directory contains a dataset prepared for the use of participants for the 1994 AAAI Spring Symposium on Artificial Intelligence in Medicine. For our analysis, we pick the data from the first patient. We choose to monitor pre-breakfast blood glucose level on a daily basis. We prefer this over blood glucose measurements at other times because they depend too much on many variables like the type of food, amount of food, etc. The data from the first patient contains daily pre-breakfast observations from 21 April 1991 to 3 September 1991 with the number on 18 August 1991 missing. The data are presented in Fig. 19.2 ignoring the only one missing value. As always, before applying the control chart, we check if our assumption of independence of the observations is reasonable. Durbin-Watson test for two-sided

alternative gives a high p-value 0.7727 showing lack of autocorrelation. Therefore, our assumption of independence is reasonable. Now, we apply the proposed method along with other competing methods when we set $ARL_0 = 500$. For the proposed method, we set $T_0 = ARL_0 = 500$ like before. The proposed method detects a distributional change on 11 June 1991 while the estimated change-point is on 5 June 1991. Other competing charts do not detect any distributional change. Running the charts for $ARL_0 = 200$ gives similar results except that the proposed chart detects distributional change one day earlier on 10 June 1991 while the estimated change-point being the same as before. Table 19.14 shows first four moments before and after the distributional change. The standard deviation decreases a lot showing a more stable fasting blood glucose numbers, and also the skewness appears to have decreased considerably. It is to be noted that the phrases “in-control” and “out-of-control” we use are in the sense of standard terminologies in SPC literature, not in the sense of blood glucose control.

Table 19.13 First four sample moments of the observations from 1873 to 1947 and 1948 to 1956

	Mean	St. deviation	Skewness	Kurtosis
1873–1947, “in-control”	21.7067	7.0842	−0.1898	2.3666
1948–1956, “out-of-control”	20.6778	2.0795	−0.1350	2.6056

19.6 Concluding Remarks

This chapter first describes a few commonly used traditional statistical process control (SPC) charts such as Shewhart, CUSUM, EWMA, and change-point-based (CPD) control charts and discusses a number of situations where these charts should be appropriately modified for practical use. Next, this chapter proposes a change-point based nonparametric statistical process control chart for detecting arbitrary distributional changes when the process distribution is univariate continuous. There are two specifically important

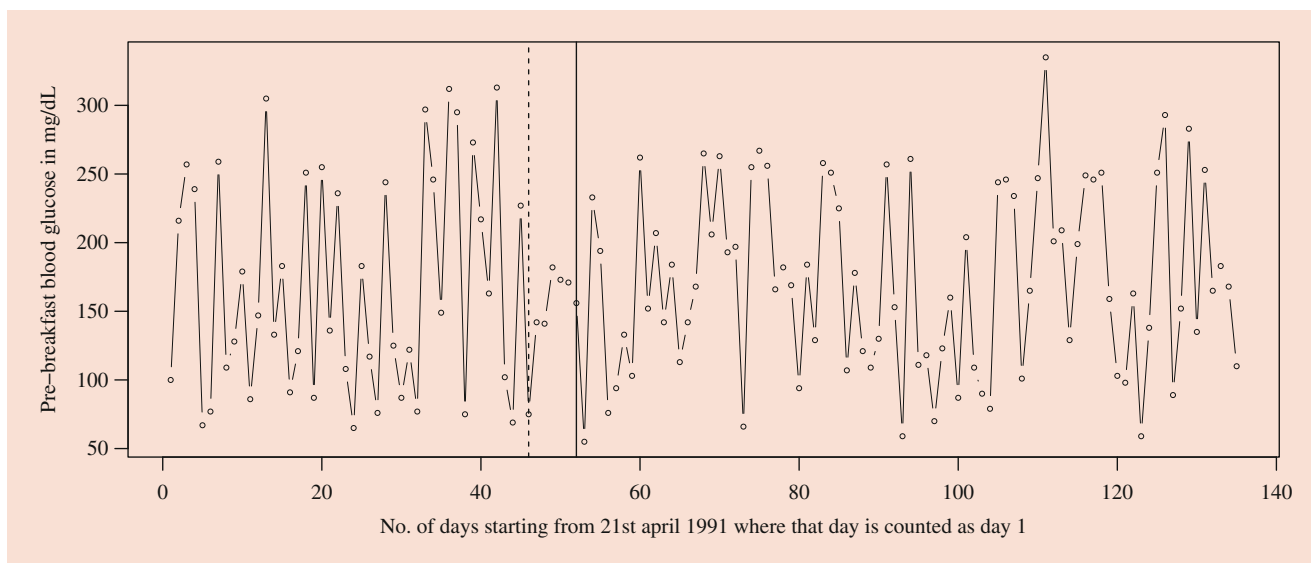


Fig. 19.2 Daily pre-breakfast blood glucose measurements of a selected patient. Setting $ARL_0 = 500$, the proposed chart detects a distributional change at the 52nd day estimating the change-point to be the 46th day. Other charts do not detect any distributional change

Table 19.14 First four sample moments of the observations from 21 April 1991 to 5–6 June 1991 to 11 June 1991

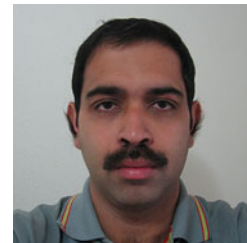
	Mean	St. deviation	Skewness	Kurtosis
21st April to 5th June, “in-control”	166.9	80.8	0.3824	1.7221
6th June to 11th June, “out-of-control”	160.8	17.2	−0.0981	1.4267

contributions of the proposed chart. The first one is the combination of the strengths of two statistical tests: Cramer-von Mises test and Ansari-Bradley test. The second one is the introduction of a numerically efficient technique to estimate the possible change-point without sacrificing the accuracy by much. Both these contributions are quite general in nature and have broad applications well beyond the numerical examples and real-world data analyses shown in this chapter including monitoring fast data streams. Another important aspect of this control chart is that runtime distribution is not geometric even when the process is “in-control.” The reason is that the probability of getting a signal is a function of current runtime. However, it appears that such charts can still be used in many applications. However, further research is required to fully understand the pros and cons of such charts.

References

- Shewhart, W.A.: *Economic Control of Quality of Manufactured Product*. Van Nostrand, New York (1931)
- Page, E.S.: Continuous inspection schemes. *Biometrika* **41**, 100–114 (1954)
- Borror, C.M., Montgomery, D.C., Runger, G.C.: Robustness of the EWMA control chart to non-normality. *J. Qual. Technol.* **31**, 309–316 (1999)
- Hawkins, D.M., P. Qiu, Kang, C.W.: The changepoint model for statistical process control. *J. Qual. Technol.* **35**, 355–366 (2003)
- Zhou, C., Zou, C., Zhang, Y., Wang, Z.: Nonparametric control chart based on change-point model. *Stat. Pap.* **50**, 13–28 (2009)
- Amin, R., M.R. Reynolds Jr., Bakir, S.T.: Nonparametric quality control charts based on the sign statistic. *Commun. Stat. Theory Methods* **24**, 1597–1623 (1995)
- Hackl, P., Ledolter, J.: A new nonparametric quality control technique. *Commun. Stat. Simul. Comput.* **21**, 423–443 (1992)
- Albers, W., Kallenberg, W.C.M.: Empirical nonparametric control charts: estimation effects and corrections. *J. Appl. Stat.* **31**, 345–360 (2004)
- Albers, W., Kallenberg, W.C.M.: CUMIN charts. *Metrika* **70**, 111–130 (2009)
- Chakraborti, S., van der Laan, P., van de Wiel, M.A.: A class of distribution-free control charts. *J. R. Stat. Soc. Ser. C Appl. Stat.* **53**, 443–462 (2004)
- Chakraborti, S., Eryilmaz, S., Human, S.W.: A phase II nonparametric control chart based on precedence statistics with runs-type signaling rules. *Comput. Stat. Data Anal.* **53**, 1054–1065 (2009)
- Chakraborti, S., Eryilmaz, S.: A nonparametric Shewhart-type signed-rank control chart based on runs. *Commun. Stat. Simul. Comput.* **36**, 335–356 (2007)
- Hawkins, D.M., Deng, Q.: A nonparametric change-point control chart. *J. Qual. Technol.* **42**, 165–173 (2010)
- Liu, L., Tsung, F., Zhang, J.: Adaptive nonparametric CUSUM scheme for detecting unknown shifts in location. *Int. J. Prod. Res.* **52**, 1592–1606 (2014)
- Mukherjee, A., Chakraborti, S.: A distribution-free control chart for the joint monitoring of location and scale. *Qual. Reliab. Eng. Int.* **28**(3), 335–352 (2012)
- Qiu, P., Li, Z.: On nonparametric statistical process control of univariate processes. *Technometrics* **53**, 390–405 (2011)
- Qiu, P., Li, Z.: Distribution-free monitoring of univariate processes. *Stat. Prob. Lett.* **81**, 1833–1840 (2011)
- Qiu, P., Zhang, J.: On phase II SPC in cases when normality is invalid. *Qual. Reliab. Eng. Int.* **31**, 27–35 (2015)
- Qiu, P.: Some perspectives on nonparametric statistical process control. *J. Qual. Technol.* **50**, 49–65 (2018)
- Qiu, P., Hawkins, D.M.: A rank based multivariate CUSUM procedure. *Technometrics* **43**, 120–132 (2001)
- Qiu, P., Hawkins, D.M.: A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions. *J. R. Stat. Soc. Ser. D Stat.* **52**, 151–164 (2003)
- Qiu, P.: Distribution-free multivariate process control based on log-linear modeling. *IIE Trans.* **40**, 664–677 (2008)
- Yeh, A.B., Lin, D.K.J., Venkatramani, C.: Unified CUSUM charts for monitoring process mean and variability. *Qual. Technol. Quant. Manag.* **1**, 65–86 (2004)
- Hawkins, D.M., Zamba, K.D.: Statistical process control for shifts in mean or variance using a changepoint formulation. *Technometrics* **47**, 164–173 (2005)
- Zhang, J., Li, E., Li, Z.: A Cramer-von Mises test-based distribution-free control chart for joint monitoring of location and scale. *Comput. Indust. Eng.* **110**, 484–497 (2017)
- Moustakides, G.V.: Optimal stopping times for detecting changes in distributions. *Ann. Stat.* **14**, 1379–1387 (1986)
- Ross, G.J., Adams, N.M.: Two nonparametric control charts for detecting arbitrary distribution changes. *J. Qual. Technol.* **44**, 102–116 (2012)
- Mukherjee, P.S.: On phase II monitoring of the probability distributions of univariate continuous processes. *Stat. Pap.* **57**(2), 539–562 (2016)
- Hawkins, D.M., Olwell, D.H.: *Cumulative Sum Charts and Charting for Quality Improvement*. Springer, New York (1998)
- Qiu, P.: *Introduction to Statistical Process Control*. CRC Press, Taylor & Francis Group, New York (2013)
- Li, Z., Qiu, P., Chatterjee, S., Wang, Z.: Using p-values to design statistical process control charts. *Stat. Pap.* **54**, 523–539 (2013)
- Li, Z., Qiu, P.: Statistical process control using dynamic sampling. *Technometrics* **56**, 325–335 (2014)
- Ross, G.J., Tasoulis, D.K., Adams, N.M.: Nonparametric monitoring of data streams for changes in location and scale. *Technometrics* **53**(4), 379–389 (2011)
- Kenney, J.F., Keeping, E.S.: *Mathematics of Statistics, Part Twol*. Van Nostrand Company Inc., Princeton (1951)
- Chou, Y.M., Polansky, A.M., Mason, R.L.: Transforming non-normal data to normality in statistical process control. *J. Qual. Technol.* **30**, 133–141 (1998)
- Yourstone, S.A., Zimmer, W.J.: Non-normality and the design of control charts for averages. *Decis. Sci.* **23**, 1099–1113 (1992)
- Alwan, L.C.: Effects of autocorrelation on control chart performance. *Commun. Stat. Theory Methods* **21**, 1025–1049 (1992)
- Black, R., Smith, J., Wells, S.: The impact of Weibull data and autocorrelation on the performance of the Shewhart and exponentially weighted moving average control charts. *Int. J. Ind. Eng. Comput.* **2**, 575–582 (2011)

39. English, J.R., Lee, S.C., Martin, T.W., Tilmon, C.: Detecting changes in autoregressive processes with X and EWMA charts. *IIE Tras.* **32**, 1103–1113 (2000)
40. Runger, G.C., Willemains, T.R.: Batch means control charts for autocorrelated data. *IIE Tras.* **28**, 483–487 (1996)
41. Kim, S.H., Alexopoulos, C., Tsui, K.L., Wilson, J.R.: A distribution-free tabular CUSUM chart for autocorrelated data. *IIE Tras.* **39**, 317–330 (2007)
42. Runger, G.C., Willemains, T.R.: Model-based and model-free control of autocorrelated processes. *J. Qual. Technol.* **27**, 283–292 (1995)
43. Johnson, R.A., Bagshaw, M.: The effect of serial correlation on the performance of CUSUM tests. *Technometrics* **16**, 103–112 (1974)
44. Montgomery, D.C., Mastrangelo, C.M.: Some statistical process control methods for autocorrelated data. *J. Qual. Technol.* **23**, 179–204 (1991)
45. Jiang, W., Tsui, K.L., Woodall, W.: A new SPC monitoring method: the ARMA chart. *Technometrics* **42**, 399–410 (2000)
46. Hawkins, D.M.: Self-starting cusums for location and scale. *Statistician* **36**, 299–315 (1987)
47. Hawkins, D.M.: On the distribution and power of a test for a single outlier. *S. Afr. Stat. J.* **3**, 9–15 (1969)
48. Chatterjee, S., Qiu, P.: Distribution-free cumulative sum control charts using bootstrap-based control limits. *Ann. Appl. Stat.* **3**, 349–369 (2009)
49. Hawkins, D.M., Maboudou-Tchao, E.M.: Self-starting multivariate exponentially weighted moving average control charting. *Technometrics* **49**, 199–209 (2007)
50. Sparks, R.S.: CUSUM charts for signalling varying location shifts. *J. Qual. Technol.* **32**, 157–171 (2000)
51. Gan, F.F.: CUSUM control charts under linear drift. *Statistician* **41**, 71–84 (1992)
52. Davis, R.B., Woodall, W.H.: Performance of the control chart trend rule under linear drift. *J. Qual. Technol.* **20**, 260–262 (1988)
53. Costa, A.F.B.: Joint \bar{X} and R charts with variable parameters. *IIE Trans.* **30**, 505–514 (1998)
54. Luo, Y., Li, Z., Wang, Z.: AdaptiveCUSUM control chart with variable sampling intervals. *Comput. Stat. Data Anal.* **53**, 2693–2701 (2009)
55. Roberts, S.V.: Control chart tests based on geometric moving averages. *Technometrics* **1**, 239–250 (1959)
56. Crowder, S.V.: Design of exponentially weighted moving average schemes. *J. Qual. Technol.* **21**, 155–162 (1989)
57. Zhang, N.F.: A statistical control chart for stationary process data. *Technometrics* **40**, 24–38 (1998)
58. Capizzi, G., Masarotto, G.: An adaptive exponentially weighted moving average control chart. *Technometrics* **45**, 199–207 (2003)
59. Pascual, F.: EWMA charts for the weibull shape parameter. *J. Qual. Technol.* **42**, 400–416 (2010)
60. Zhang, L., Chen, G.: EWMA charts for monitoring the mean of censored Weibull lifetimes. *J. Qual. Technol.* **36**, 321–328 (2004)
61. Borror, C.M., Champ, C.W., Ridgon, S.E.: Poisson EWMA control charts. *J. Qual. Technol.* **30**, 352–361 (1998)
62. Gan, F.F.: Monitoring Poisson observations using modified exponentially weighted moving average control charts. *Commun. Stat. Simul. Comput.* **19**, 103–124 (1990)
63. Perry, M.B., Pignatiello Jr. J.J.: Estimation of the change point of the process fraction nonconforming in SPC applications. *Int. J. Reliab. Qual. Saf. Eng.* **12**, 95–110 (2005)
64. Sparks, R.S., Keighley, T., Muscatello, D.: Optimal exponentially weighted moving average (EWMA) plans for detecting seasonal epidemics when faced with non-homogeneous negative binomial counts. *J. Appl. Stat.* **38**, 2165–2181 (2011)
65. Gan, F.F.: EWMA control chart under linear drift. *J. Stat. Comput. Simul.* **38**, 181–200 (1991)
66. Gombay, E.: Sequential change-point detection and estimation. *Seq. Anal.* **22**, 203–222 (2003)
67. Hinkley, D.V.: Inference about the change-point in a sequence of random variables. *Biometrika* **57**, 1–17 (1970)
68. Grahamy, M.A., Human, S.W., Chakraborti, S.: A Phase I nonparametric Shewhart-type control chart based on median. *J. Appl. Stat.* **37**, 1795–1813 (2010)
69. Jones-Farmer, L.A., Jordan, V., Champ, C.W.: Distribution-free phase I control charts for subgroup location. *J. Qual. Technol.* **41**, 304–317 (2009)
70. Ansari, A.R., Bradley, R.A.: Rank-sum tests for dispersions. *Ann. Math. Stat.* **31**(4), 1174–1189 (1960)
71. Anderson, T.W.: On the distribution of the two-sample Cramer-von Mises criterion. *Ann. Math. Stat.* **33**(3), 1148–1159 (1962)
72. Anderson, T.W., Darling, D.A.: Asymptotic theory of certain ‘goodness of fit’ criteria based on stochastic processes. *Ann. Math. Stat.* **23**, 193–212 (1952)
73. Xiao, Y.: (2012). <https://cran.r-project.org/src/contrib/Archive/CvM2SL2Test/>
74. Xiao, Y., Gordon, A., Yakovlev, A.: A C++ program for the Cramer-von Mises two sample test. *J. Stat. Softw.* **17**(8) (2007)
75. Mood, A.: On the asymptotic efficiency of certain nonparametric two-sample tests. *Ann. Math. Stat.* **25**, 514–533 (1954)
76. Lepage, Y.: Combination of Wilcoxiens and Ansari–Bradley Statistics. *Biometrika* **58**, 213–217 (1971)
77. Lichman, M.: UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine (2013). <http://archive.ics.uci.edu/ml>



Partha Sarathi Mukherjee received his PhD in Statistics from University of Minnesota at Twin Cities, USA in 2011. He worked as a faculty at Boise State University, USA from 2012 to 2019. Recently, he joined Indian Statistical Institute, Kolkata, India as a faculty. He works on statistical image restoration, statistical process control, and application of statistical tools in various disciplines.