Using P-Values To Design Statistical Process Control Charts

Zhonghua Li[1], Peihua Qiu[2], Snigdhansu Chatterjee[2] and Zhaojun Wang[1]

[1]LPMC and School of Mathematical Sciences, Nankai University, Tianjin 300071, China

[2]School of Statistics, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

Conventional Phase II statistical process control (SPC) charts are designed using control limits; a chart gives a signal of process distributional shift when its charting statistic exceeds a properly chosen control limit. To do so, we only know whether a chart is out-of-control (OC) at a given time. It is therefore not informative enough about the likelihood of a potential distributional shift. In this paper, we suggest designing the SPC charts using p-values. By this approach, at each time point of Phase II process monitoring, the p-value of the observed charting statistic is computed, under the assumption that the process is in-control (IC). If the p-value is less than a pre-specified significance level, then a signal of distributional shift is delivered. This p-value approach has several benefits, compared to the conventional design using control limits. First, after a signal of distributional shift is delivered, we could know how strong the signal is. Second, even when the p-value at a given time point is larger than the significance level, it still provides us useful information about how stable the process performs at that time point. The second benefit is especially useful when we adopt a variable sampling scheme, by which the sampling time can be longer when we have more evidence that the process runs stably, supported by a larger p-value. To demonstrate the p-value approach, we consider univariate process monitoring by cumulative sum (CUSUM) control charts in various cases.

Key words: Bootstrap; Cumulative sum control charts; Process monitoring; Self-starting; Variable Sampling.

1 Introduction

Statistical process control (SPC) charts are used widely for monitoring the stability of different processes over time. Practical applications of SPC charts have now extended far beyond manufacturing industries to other industries such as biology, genetics, medicine, finance and so forth. Some widely used control charts include the Shewhart charts (Shewhart 1931), the cumulative sum (CUSUM) control charts (Page 1954), and the exponentially weighted moving average (EWMA) control charts (Roberts 1959). It is well known that Shewhart charts are effective in detecting isolated shifts or relatively large sustained shifts, while CUSUM and EWMA control charts are effective in detecting small to moderate sustained shifts. Comparing CUSUM with EWMA control charts, it has been demonstrated that they often have similar performance in terms of the average run length (ARL) (cf., Lucas and Saccucci 1990, Luo et al. 2009), which is the average number of observations needed for a control chart to signal a shift. Recently, process monitoring based on change-point detection also receives much attention (cf., Hawkins et al. 2003, Zhou et al. 2009). By the change-point approach, besides a signal of shift, the shift location can be obtained simultaneously. See Hawkins and Olwell (1998) and Montgomery (2004) for more complete discussion about theory and methodologies about SPC.

A conventional control chart gives a signal of process distributional shift when its charting statistic falls beyond the control limit(s). In practice, after a signal of shift is obtained, practitioners would also be interested in knowing how strong the signal is, so that appropriate subsequent actions can be taken accordingly. In cases when a variable sampling scheme is adopted (cf., Reynolds et al. 1990), even if a shift is not detected at a given time point, it would still be helpful to know the likelihood of a potential shift. The sampling time can be adjusted according to the likelihood as follows. It can be longer if the likelihood is smaller, and shorter otherwise. With a conventional design of control charts using control limits, such a quantitative measure of the shift likelihood is difficult to obtain.

In the context of hypothesis testing, early testing procedures make decisions using the concepts of rejection region and acception region (cf., Lehmann 1997). A null hypothesis would be rejected when the observed value of the related test statistic falls in the rejection region. This conventional way of hypothesis testing has been replaced by the p-value approach in recent text books, because the p-value approach can not only make a decision about the hypotheses, but also tell us how strong the evidence in the observed data is against the null hypothesis. Motivated by the p-value approach in hypothesis testing, in this paper, we suggest designing control charts using the p-value approach as well. By the p-value approach, for a given control chart, the in-control (IC) distribution of the charting statistic is first computed or estimated. Then, at a given time point, the p-value corresponding to the observed value of the charting statistic can be obtained. If the p-value is less than a pre-specified significance level, then the chart signals a process distributional shift. Compared to conventional control charts using control limits, this p-value approach has several benefits, including the following ones. First, at a given time point, even if a shift is not detected, the p-value can provide us a quantitative measure of the likelihood of a potential shift, so that the subsequent sampling interval can be adjusted properly. Second, conventional control charts may take different forms (e.g., the one-sided or two-sided charts) and their control limits are different in different situations. As a comparison, all control charts using the p-value approach have a same format, in the sense that the vertical axis is always in the range of $[0, 1]$, denoting the p-values, and there is only one control limit corresponding to the significance level. This makes the charts more convenient to use.

In the literature, p-value calculation of the charting statistic of a traditionally designed control chart has been discussed by several papers. In cases when the IC process distribution is assumed normal with a known variance, Grigg and Spiegelhalter (2008) provide an approximation formula for the IC distribution of the charting statistic of the conventional CUSUM chart. Li and Tsung (2009) study the false discovery rate in multistage process monitoring, in which p-value calculation of the charting statistics used in different stages of process monitoring is discussed. However, general discussion about how to design control charts using p-values in various cases is still lacking.

In this paper, we demonstrate the p-value approach in cases when we are interested in detecting mean shifts in Phase II SPC using a CUSUM control chart. Its applications in Phase I SPC, or its applications in Phase II SPC for detecting scale shifts and other process distributional shifts using other charting schemes can be discussed similarly. The rest part of the paper is organized as follows. In Section 2, our p-value approach is described in detail. Its numerical performance is evaluated in Section 3. We then demonstrate this approach using a real data example in Section 4. Finally, several remarks conclude the paper in Section 5.

2 Designing Phase II CUSUM Charts Using p-Values

In this section, we describe our proposed p-value approach in cases to design Phase II CUSUM charts for detecting mean shifts of univariate processes. In the literature, the IC process distribution is often assumed known in Phase II SPC. However, in practice, the IC process distribution is rarely known. Instead, it needs to be estimated from an IC dataset obtained at the end of Phase I analysis after the process has been adjusted properly so that it works stably. Hawkins (1987) proposes the self-starting CUSUM chart for Phase II SPC in cases when the IC process distribution is assumed normal but its mean and variance parameters need to be estimated. Chatterjee and Qiu (2009) discuss Phase II SPC in cases when the IC process distribution is estimated from an IC dataset using bootstrap.

To account for different cases, our description of the p-value approach is organized in four parts. In Section 2.1, the p-value approach is introduced in cases when the IC process distribution is completely known. The cases when the IC process distribution follows a parametric distribution with unknown parameters and when the IC process distribution is completely unknown are discussed in Sections 2.2 and 2.3, respectively. Then, CUSUM chart using the p-value approach together with the variable sampling scheme is discussed in Section 2.4.

2.1 Cases when the IC process distribution is known

Assume that $X_1, X_2, \ldots, X_t, \ldots$ are a sequence of independent observations obtained during Phase II process monitoring. Their cumulative distribution functions (cdfs) are the same to be F_0 up to an unknown time point τ, and change to another cdf F_1 after the time point τ. For simplicity, we further assume that F_0 and F_1 are the same except their means μ_0 and μ_1. Then, the process has a mean shift at τ, and the major goal of Phase II process monitoring is to detect the mean shift as soon as possible. To this end, the conventional CUSUM chart for detecting an upward mean shift uses the charting statistic

$$\begin{cases} C_0^+ = 0, \\ C_t^+ = \max(0, C_{t-1}^+ + X_t - \mu_0 - k), \ \text{for } t \geq 1, \end{cases} \tag{1}$$

where k is an allowance constant. The chart gives a signal of mean shift when

$$C_t^+ > h,$$

where h is a control limit chosen to achieve a pre-specified IC ARL (denoted as ARL_0) value.

Instead of comparing the charting statistic C_t^+ with the control limit value h, in this paper, we suggest computing the p-value corresponding to the value of C_t^+, and then comparing the p-value with a pre-specified significance level α for making a decision whether the process is out-of-control (OC). To this end, we need to find the IC distribution of C_t^+ first. Recently, Grigg and Spiegelhalter (2008) provide an approximation formula for this IC distribution in cases when the IC process distribution is normal with a known variance. The derivation of this formula is part theoretical and part empirical. Since we will discuss various other cases, including the ones when the IC process distribution is a t distribution or a chi-squared distribution that represents a symmetric distribution with heavy tails or a skewed distribution, the conditions required by the method of Grigg and Spiegelhalter (2008) are not satisfied in such cases. Therefore, we suggest computing p-values by simulation as follows.

Assume that the IC process distribution is completely known. Then, for a given time point $t \geq 1$ and a given allowance constant k, we generate Phase II observations $X_1, X_2, \ldots, X_t$ and compute the value of C_t^+ by (1). This process is then repeated many times (e.g., 1 million times), and the empirical distribution of C_t^+ can be determined by the computed C_t^+ values. For a given observed value of C_t^+, denoted as C_t^{+*}, the corresponding p-value is then computed by

$$p_{C_t^{+*}} = \mathrm{P}(C_t^+ > C_t^{+*}). \tag{2}$$

Figure 1 presents the p-values computed by (2) in cases when the IC distribution of C_t^+ is determined by 1 million replications, $k = 0.5$, $t = 1, 5, 10, 50$ and 100, and the IC process distribution is the normalized versions of the standard normal $N(0,1)$, t distribution with 4 degrees of freedom (denoted as t_4), chi-squared distribution with 1 degree of freedom (denoted as χ_1^2), and chi-squared distribution with 4 degrees of freedom (denoted as χ_4^2). From the plots of the figure, it can be seen that the p-values, which are the right-tail probabilities of the IC distribution of C_t^+, depend on t; but they are stable when $t \geq 50$, which is consistent with the well-known steady-state distribution of C_t^+ (cf., Hawkins and Olwell 1998). Therefore, in practice, to use the p-values to monitor the process, we only need to compute the probability distributions of C_t^+, for $t < 50$. From Figure 1, we can also see that the p-values depend slightly on the IC process distribution. For instance, p-values in the case of $N(0,1)$ are slightly different from the corresponding ones in the case of t_4.

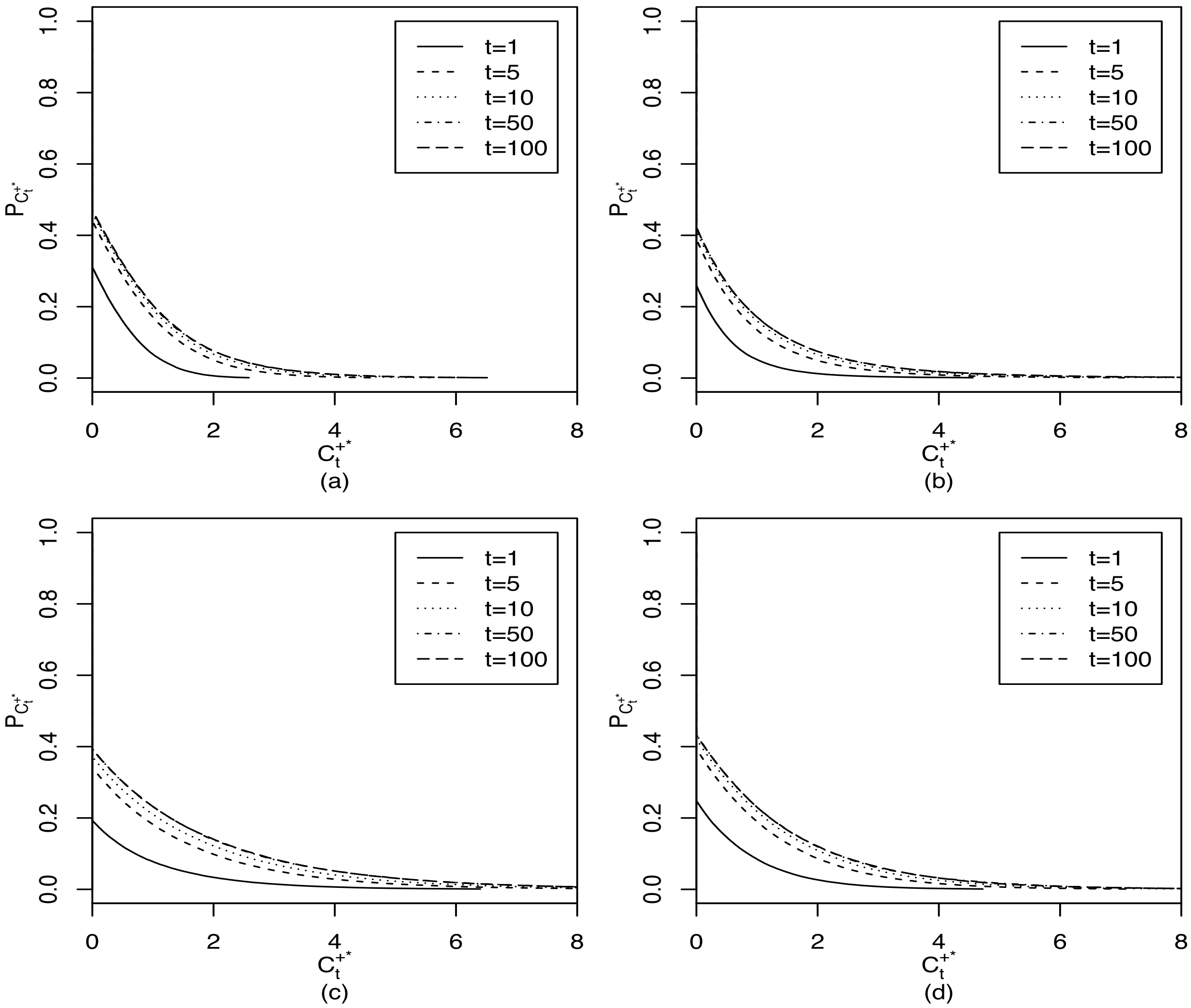

Figure 1: p-values computed by (2) with different t and IC process distributions. (a) $N(0,1)$, (b) t_4, (c) χ_1^2, (d) χ_4^2.

Next, for several commonly used significance levels α =0.01, 0.02, 0.05, and 0.10, we provide the corresponding critical values (CVs) of C_t^+ in Table 1, in cases when $k = 0.25$ or 0.5, and the IC process distribution is the normalized version of $N(0,1)$, t_4, χ_1^2, or χ_4^2. The corresponding ARL_0 values are also provided. These values are all computed based on 1 million simulation runs for C_t^+ when $t = 50$. From the table, it can be seen that the CV values and the ARL_0 values decrease

Table 1: Critical values (CVs) and the corresponding ARL_0 values for several commonly used significance levels α =0.01, 0.02, 0.05, and 0.10, and several IC process distributions.

		α=0.01		α=0.02		α=0.05		α=0.10	
	k	CV	ARL_0	CV	ARL_0	CV	ARL_0	CV	ARL_0
$N(0,1)$	0.25	8.1841	996.238	6.9167	464.853	5.2237	172.234	3.9236	71.278
	0.50	4.0606	322.823	3.3483	169.538	2.4170	56.003	1.7237	25.425
t_4	0.25	8.8185	1262.625	7.2411	576.815	5.2305	195.717	3.7918	82.608
	0.50	4.9217	506.946	3.7781	218.192	2.5281	73.202	1.6415	33.620
χ_1^2	0.25	11.5085	1153.281	9.5924	582.298	6.9887	212.468	5.0404	86.091
	0.50	7.3315	481.008	5.8988	233.748	4.0530	82.886	2.6607	35.554
χ_4^2	0.25	9.9038	1048.514	8.3649	512.733	6.1924	183.958	4.5247	76.111
	0.50	5.6788	419.965	4.6678	203.735	3.3290	68.505	2.2905	28.742

when k increases or α increases. By the way, the CVs listed in Table 1 are the upper α-quantiles of C_t^+. To use the CVs in Table 1, one can have a rough idea about the p-value after he computes the value of the charting statistic. For instance, in cases when the IC process is $N(0,1)$ and $k = 0.25$, if the computed value of the charting statistic is 8.2, then from Table 1, we can know that the corresponding p-value would be less than 0.01 because 8.2 is larger than 8.1841 which is the upper 0.01-quantile in such cases.

2.2 Cases when the IC process distribution follows a parametric model with unknown parameters

The assumption that the IC distribution is completely known may not be reasonable for some applications. In this part, we consider a more general case when the IC process distribution follows a parametric model with one or more unknown parameters. One example of this scenario is when it is reasonable to assume that the IC process distribution is $N(\mu_0, \sigma^2)$, but the parameters μ_0 and σ are both unknown. In such cases, if there is an IC dataset, then μ_0 and σ can be estimated from the IC dataset beforehand. However, the sample size of the IC dataset should be large enough to guarantee that the resulting control chart performs reasonably well. Otherwise, control charts with estimated parameters would have a large bias in terms of the ARL_0 value, and they will lose some power in detecting process distributional shifts as well. See, for instance, Jensen et al. (2006). To overcome this difficulty, Hawkins (1987) proposes the self-starting method for constructing control charts in such cases. Construction of the self-starting CUSUM using p-values is described below.

Assume that $X_1, X_2, \ldots, X_t, \ldots$ are a sequence of i.i.d. observations with common distribution $N(\mu_0, \sigma^2)$. Let $m \geq 3$ be a fixed number. For $t \geq m$, define

$$\begin{aligned} T_t &= \frac{X_t - \overline{X}_{t-1}}{s_{t-1}}, \\ U_t &= \Phi^{-1}\left[G_{t-2}\left(T_t\sqrt{\frac{t-1}{t}}\right)\right], \end{aligned} \tag{3}$$

where $\overline{X}_{t-1}$ and s_{t-1} are the sample mean and sample standard deviation of the first $t-1$ observations, and $\Phi(\cdot)$ and $G_{t-2}(\cdot)$ are the cdf's of the standard normal and the t distribution with $t-2$ degrees of freedom, respectively. Then, it can be shown that the sequence $\{U_t, t \geq m\}$ are i.i.d. with the common distribution $N(0,1)$ (cf., Hawkins (1969) and Quesenberry (1991)). Therefore, we can now monitor the sequence $\{U_t, t \geq m\}$ for possible mean shifts using the control chart discussed in Section 2.1. Design of self-starting CUSUM charts when the IC process distribution is Gamma, binomial, and Poisson has been discussed in Hawkins and Olwell (1998).

By using the self-starting control chart based on (3), we need to have m IC observations collected before process monitoring. Otherwise, the sample standard deviation s_{t-1} is not well defined. Figure 2 presents the ARL_0 values at several commonly used α levels when m takes values of 3, 5, 7, 10 and ∞. Note that the case of $m = \infty$ actually denotes the case when the IC process distribution is completely known. From the plot, it can be seen that the ARL_0 values do not depend on the value of m much, which is appealing because it implies that we do not have to collect too many IC observations before using the self-starting control chart for process monitoring, in cases when we know the parametric form of the IC distribution. This result is not surprising because the self-starting control chart keeps updating the estimates of the IC parameters by using the first m IC observations and all subsequent observations as long as no signal of mean shifts is delivered. See related discussion in Hawkins and Olwell (1998, Section 7.2).

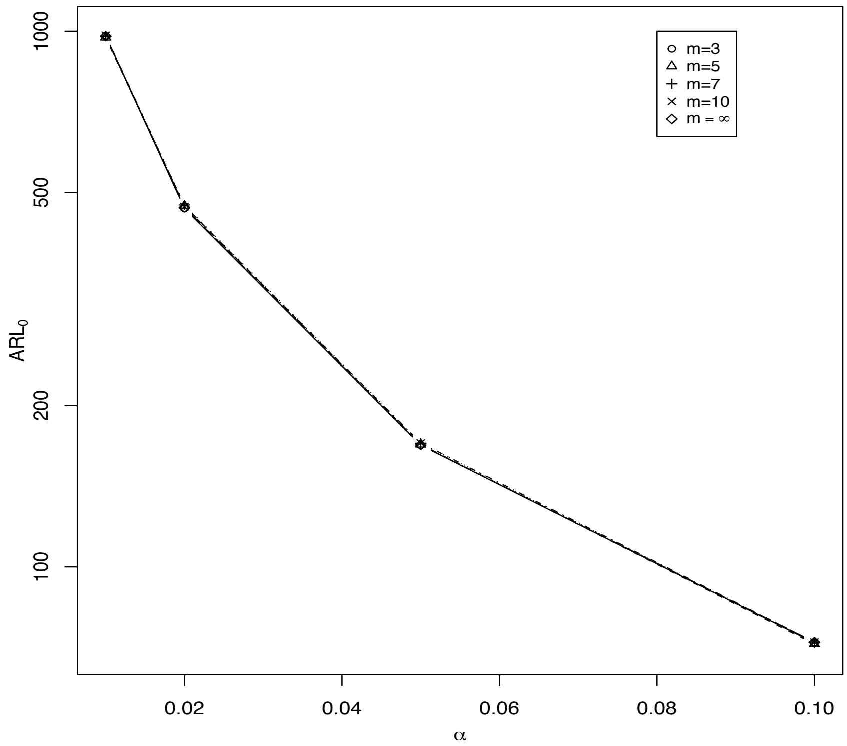

Figure 2: ARL_0 values at several commonly used α levels when m takes values of 3, 5, 7, 10 and ∞.

2.3 Cases when the IC process distribution is completely unknown

In this part, we discuss the case when the IC process distribution is completely unknown and when a set of IC data is available. In such cases, there are two possible ways to use the IC data for process monitoring. One way is to first estimate the IC process distribution from the IC data, and then monitor the process as usual using the estimated IC process distribution. The second approach is to estimate the IC distribution of the charting statistic using a bootstrap resampling technique from the IC data. Chatterjee and Qiu (2009) has demonstrated that the bootstrap approach is often more effective, which is also adopted here.

By the bootstrap method (cf., Efron 1979, Efron and Tibshirani 1993), we repeatedly draw bootstrap observations with replacement from the IC dataset, and the bootstrap observations are used as the Phase II observations. Then, we can compute the value of the charting statistic C_t^+. This process is then repeated B times, and the IC distribution of C_t^+ can be estimated from the B calculated values of C_t^+.

It is expected that the size m of the IC data would have an impact on the performance of our control chart. To see this, we compute the ARL_0 values at several commonly used α levels when m changes its value among 100, 500, 1000, 2000, and ∞. Again, the case when $m = \infty$ denotes the case when the IC process distribution is completely known. These ARL_0 values are shown in the plots of Figure 3. From the plots, we can see that, as long as $m \geq 1000$, the ARL_0 values are quite stable in various cases. Therefore, we suggest collecting 1000-2000 IC observations before Phase II process monitoring in cases when the IC process distribution is completely unknown. This requirement on the IC data size is similar to those given by Jones et al. (2001), Jones (2002) and Jones et al. (2004) in cases when certain IC parameters need to be estimated from an IC data with the traditional control charts.

2.4 CUSUM charts using variable sampling schemes

In recent years, variable sample rate (VSR) control charts have received much attention in the literature, by which the sampling rate changes as a function of the current and prior sample results. One major advantage of the VSR charts, compared to the fixed sampling rate (FSR) charts, is that VSR charts often provide faster detection of small to moderate process changes, for a given ARL_0 and a given IC average sampling rate. There are several possible approaches that can be used to vary the sampling rate, which include the variable sampling intervals (VSI), the variable sample size (VSS), and the variable sample size and sampling intervals (VSSI) schemes (e.g., Montgomery 2007). The sampling scheme of the conventional VSI charts is to use a longer sampling interval when the charting statistic value is far away from the control limits (i.e., within the so-called central region), and use a shorter sampling interval when the charting statistic value does not exceed the control limits but within the so-called warning region. If the charting statistic value exceeds the control limits (i.e., falls in the so-called action region), then the process is considered to be OC.

In this paper, we propose to monitor the p-values of the charting statistic. In this context, the variable sampling scheme can work as follows. We use a longer sampling interval as long as $P_{C_t^{+*}}$ is much larger than the significance level α, and a shorter sampling interval if $P_{C_t^{+*}}$ is larger than α but their values are close to α. The process is considered OC if $P_{C_t^{+*}} < \alpha$. Intuitively, it might be reasonable to set the sampling interval to be a continuously increasing function of $P_{C_t^{+*}}$.

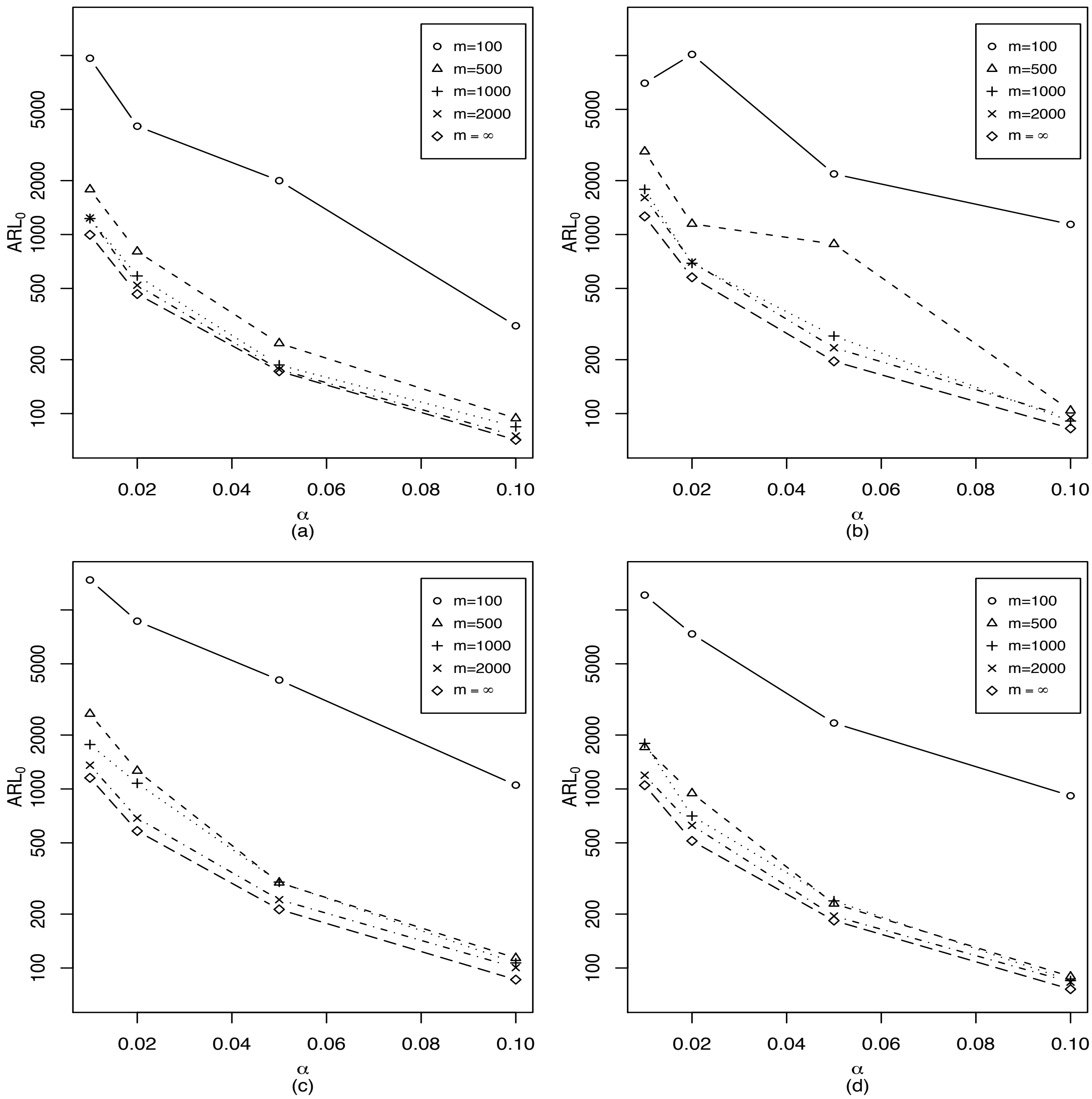

Figure 3: ARL_0 values at several commonly used α levels. (a) $N(0,1)$, (b) t_4, (c) χ^2_1, (d) χ^2_4.

However, previous research has demonstrated that it is sufficient to use only two possible values of the sampling interval for the control chart with a variable sampling scheme to achieve some good statistical properties (cf., Costa 1998, Luo et al. 2009, Reynolds and Arnold 2001, Wu et al. 2007). For this reason, we also use two values of the sampling interval here. Let $0 < d_1 < d_2$ be two possible sampling intervals. Then, the sampling interval function, denoted as $d(t)$, can be defined by

$$d(t) = \begin{cases} d_1, & \text{if } P_{C_t^{+*}} \in R_w, \\ d_2, & \text{if } P_{C_t^{+*}} \in R_c. \end{cases}$$

where R_w and R_c denote the warning region and central region, respectively.

Traditionally, the ARL value is commonly used as a performance measure of a control chart. However, when the sampling interval is variable, the time to signal may not be a constant multiple of the ARL. In such cases, widely used performance measures are the average time to signal (ATS), which is defined to be the expected value of the time from the start of the process to the time when the chart indicates an OC signal, and the adjusted average time to signal (AATS), which is defined to be the expected value of the time from the occurrence of an assignable cause of shifts to the time when the chart gives an OC signal. The AATS is also called the steady-state ATS (SSATS) if the assignable cause of shifts occurs at a late time. When the process is IC, the ATS value can be used to setup the chart so that the false alarm rate is controlled at a certain level. When the process is OC, the AATS value can be used to measure the performance of the chart. A chart with a smaller OC AATS value, denoted as $AATS_1$, performs better. Regarding d_1 and d_2, existing research (cf., Costa 1998, Luo et al. 2009, Reynolds and Arnold 2001, Wu et al. 2007) shows that d_1 should be small and d_2 should be large. Based on extensive numerical study, Reynolds et al. (1990) suggest using $d_1 = 0.1$ and $d_2 = 1.9$. These values of d_1 and d_2 are also used in this paper. It should be pointed out that the design of the above VSI control chart using p-values may not be optimal. As a matter of fact, even using the conventional control limits in its design, it is still an open problem how to design an optimal VSI control chart. The optimal design of the VSI control chart based on p-values will be left for our future research.

To use the VSI control chart, we still need to choose the warning and central regions R_w and R_c properly. To this end, let d_0 be the fixed sampling interval without employing VSI. Without loss of generality, we assume that $d_0 = 1$ in this paper. In existing control charts using variable sampling schemes, the warning and central regions are defined based on the charting statistic C_t^+ directly. Usually, the central region takes the form of $(-\infty, h_1]$, and the warning region takes the form of (h_1, h), where $0 < h_1 < h$ are two control limits and they are chosen such that (i) a given value of ARL_0 is achieved, and (ii) ARL_0 is about the same as the IC ATS value, denoted as ATS_0. Similarly, in the proposed chart using p-values, the warning and central regions can take the forms of $R_c = [\alpha_1, 1]$ and $R_w = (\alpha, \alpha_1)$, respectively, where $0 < \alpha < \alpha_1$ are two significance levels. They are chosen such that a fixed level of ARL_0 is achieved and $ARL_0 = ATS_0$. Next, we briefly explain that the requirement of $ARL_0 = ATS_0$ guarantees that the IC average inspection rates of the monitoring schemes based on p-values with and without the VSI are the same. To this end, let

$$p_1 = P[P_{C_t^{+*}} \in R_w | P_{C_t^{+*}} > \alpha],$$

$$p_2 = P[P_{C_t^{+*}} \in R_c | P_{C_t^{+*}} > \alpha],$$

n_1 and n_2 be the expected numbers of observations before C_t^{+*} falls into R_w and R_c, respectively. Then, $p_1 = n_1/ARL_0$ and $p_2 = n_2/ARL_0$. By Reynolds and Arnold (2001), ATS_0 can be written as

$$ATS_0 = d_1 n_1 + d_2 n_2.$$

Therefore, the requirement of $ARL_0 = ATS_0$ implies that

$$\frac{n_1}{ARL_0} d_1 + \frac{n_2}{ARL_0} d_2 = 1,$$

which is equivalent to

$$p_1 d_1 + p_2 d_2 = d_0.$$

The left-hand-side of the last equation is the IC average inspection rate of the monitoring scheme with the VSI, and the right-hand-side is the IC average inspection rate of the monitoring scheme without the VSI. Therefore, the two IC average inspection rates are the same.

To design the proposed CUSUM chart using p-values and VSI, we need to choose the parameters $(k, \alpha, \alpha_1, d_1, d_2)$ properly such that the pre-specified ARL_0 value is achieved. To this end, we can follow the several steps described below.

1. Select the reference value k to be the half of a target mean shift, as usual. The resulting

control chart will be good for detecting small mean shifts if k is chosen small, and it will be good for detecting large mean shifts if k is chosen large.

2. Determine the sampling intervals d_1 and d_2 properly. Usually, d_1 is chosen to be the shortest time to sample an item, and d_2 is chosen to be $2 - d_1$. In such cases, the average of d_1 and d_2 equals $d_0 = 1$.

3. If the value of ARL_0 is given beforehand, then choose α such that the CUSUM chart using p-values and a fixed sampling scheme achieves the pre-specified ARL_0 value. We can also specify the value of α in advance (e.g., $\alpha = 0.05$), as in the hypothesis testing setup. Then, the value of ARL_0 of the CUSUM chart using p-values and a fixed sampling scheme can be determined accordingly.

4. The parameter α_1 is searched such that the actual value of ATS_0 of the CUSUM chart using p-values and VSI equals the value of ARL_0.

3 Simulation Study

In this section, a comparative numerical study is conducted by Monte Carlo simulation to evaluate the performance of our proposed CUSUM charts using p-values. The control charts considered include the chart using p-values when the IC process distribution is assumed known to be normal with known parameters (denoted as N-CUSUM), the self-starting chart using p-values when the IC process distribution is assumed to follow a parametric distribution with unknown parameters (denoted as S-CUSUM), the chart using p-values and bootstrap when the IC process distribution is unknown but there is an IC dataset (denoted as B-CUSUM), the B-CUSUM chart used together with a variable sampling intervals scheme (denoted as V-CUSUM), the CUSUM chart using the traditional control limit, the fixed sampling scheme and bootstrap (denoted as NT-CUSUM), and the NT-CUSUM chart using a variable sampling intervals scheme (denoted as VT-CUSUM).

As in Figures 1 and 3, four different IC distributions, including the normalized versions with mean 0 and variance 1 of $N(0,1)$, t_4, χ_1^2 and χ_4^2, are considered. In all the CUSUM charts, the reference value k is chosen to be 0.25. Since the V-CUSUM chart and VT-CUSUM chart are involved, we use ATS as the performance measure in this numerical study, although ATS and ARL are the same for the charts N-CUSUM, S-CUSUM, B-CUSUM and NT-CUSUM who use a fixed sampling scheme. For all the control charts using p-values, the significant level is set to be $\alpha = 0.05$. From Table 1, the ARL_0 values of the CUSUM chart using p-values in cases when the IC process distribution is assumed known (which is called the ideal CUSUM here) to be the four distributions are respectively 172.234, 195.717, 212.468, and 183.958. For the S-CUSUM chart, we set $m = 10$. For the B-CUSUM chart, the IC sample size equals 2000. In the V-CUSUM chart, $d_1 = 0.1$, $d_2 = 1.9$, and α_1 is chosen so that $ATS_0 = ARL_0$, where the value of ARL_0 is the same as that of the ideal CUSUM chart. For all the charts using the traditional control limits, the control limits and warning limits are searched by simulation such that $ATS_0 = ARL_0$ and their ARL_0 values all equal to that of the ideal CUSUM chart. All ATS values presented in this section are computed from 1,000,000 replications. For charts B-CUSUM and V-CUSUM, these 1,000,000 replications are arranged as follows. We first generate 2,000 IC observations. Then, the IC distribution of the charting statistic is estimated by bootstrap from the IC data, and the ATS value is determined based on 10,000 replications of Phase II monitoring. This process is then repeated 100 times, with 100 different IC datasets being generated and used.

Note that, when the IC distribution is not normal, the normal-distribution-based chart N-CUSUM may not be appropriate to use, which has been well demonstrated in the literature (e.g., Qiu and Hawkins 2001, Qiu and Li 2011). Figure 4 further presents the nominal (dotted lines) and actual ATS_0 values of the chart N-CUSUM when the true IC distribution is t and χ^2 distributions with degrees of freedom changing from 1 to 20. We can see from Figure 4 that, when the degrees of freedom is small, the actual ATS_0 values are quite different from the nominal ATS_0 values. Therefore, it is misleading to use the N-CUSUM chart for detecting mean shifts in such cases. In order to make the comparison fair, when the IC distribution is not normal, we adjust the significance level of the N-CUSUM chart properly such that its actual ATS_0 value is about the same as those of other charts, and the adjusted N-CUSUM chart is denoted as N-CUSUM(A).

The OC ATS values of all the charts considered are presented in the four plots of Figure 5 for four different IC distributions. The S-CUSUM chart is not included in plot (b) when the IC distribution is t_4 because we have not found any existing discussion on how to construct such a chart in that case yet.

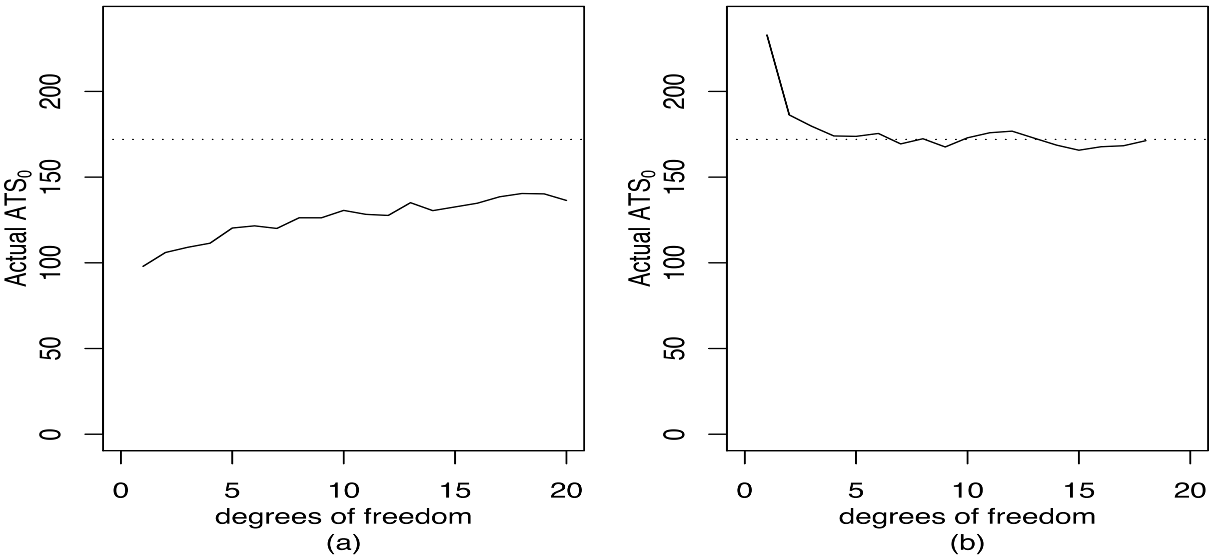

Figure 4: (a) Actual ATS_0 values of the chart N-CUSUM when the true IC distribution is χ^2 with different degrees of freedom. (b) Actual ATS_0 values of the chart N-CUSUM when the true IC distribution is t. In each plot, the dotted line denotes the nominal ARL_0 value.

From Figure 5, we can have the following conclusions. First, it seems that the charts V-CUSUM and VT-CUSUM with variable sampling schemes perform the best in all cases considered, except certain cases when the IC distribution is very skewed (cf., plot (c)). Therefore, in practice, it might be a good idea to adopt a variable sampling scheme, if it is practically convenient. Second, the two charts V-CUSUM and VT-CUSUM perform similarly, because their only difference is that the former uses a significance level α while the latter uses a control limit h in their decision rules and the two charts are designed to have the same ARL_0 and ATS_0 values. The same conclusion can also be made for the charts N-CUSUM and NT-CUSUM. Third, when the true IC distribution is known to be $N(0,1)$, the charts N-CUSUM, B-CUSUM and NT-CUSUM all perform reasonably well, while the chart S-CUSUM performs slightly worse because it does not assume the IC parameters to be known and it does not use many IC data either. Fourth, when the true IC distribution is t_4 which is symmetric with heavy tails, the charts N-CUSUM, B-CUSUM and NT-CUSUM perform similarly well. Fifth, when the true IC distribution is χ^2_4 which is skewed with a relatively small skewness, the self-starting chart S-CUSUM performs slightly better than the charts N-CUSUM, B-CUSUM and NT-CUSUM. Sixth, when the true IC distribution is χ^2_1 which is very skewed, the self-starting chart S-CUSUM performs the best when the shift size is smaller than 0.6.

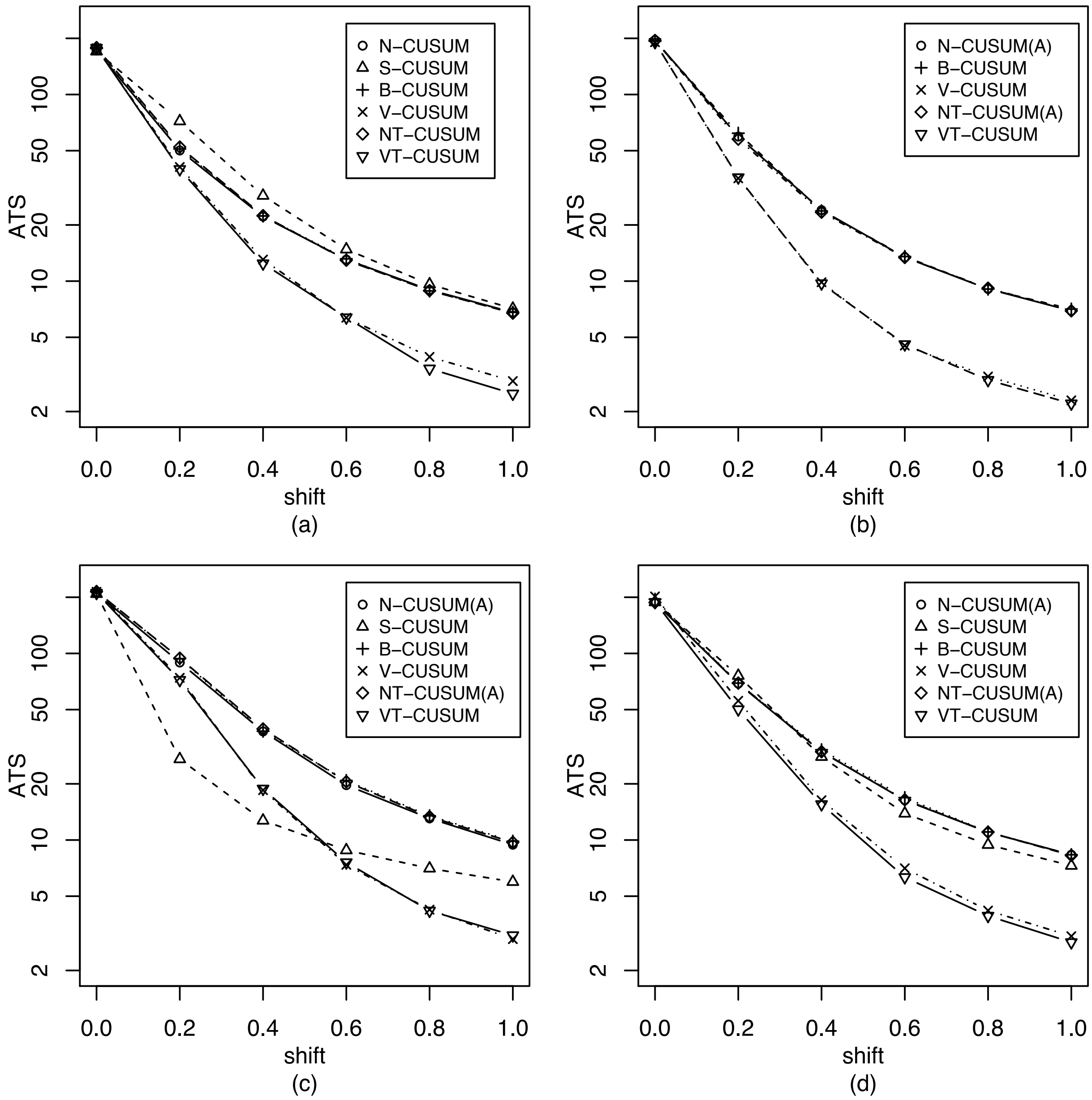

Figure 5: ATS values when the mean shift changes from 0 to 1. (a) $N(0,1)$, (b) t_4, (c) χ_1^2, (d) χ_4^2.

4 A Real Data Example

In this section, we demonstrate our proposed p-value approach using a real-data obtained from a chemical process. The data set contains 149 readings of triglyceride content of chemical products, which is described in more detail in Chapter 3 of Hawkins and Olwell (1998). The data can be downloaded from the web page http://www.stat.umn.edu/cusum/data.htm.

Hawkins and Olwell (1998) have shown that the process mean appears to be well within its allowable range in the early part of the data set. Based on that result, we use the first 75 observations as an IC data, and the remaining observations are used for testing. Since the observations in this example are collected at equally spaced time points, only the three control charts N-CUSUM, S-CUSUM and B-CUSUM are considered here. In all three proposed control charts, we use $k = 0.25$, and $\alpha = 0.05$. In chart S-CUSUM, we choose $m = 10$. Then, the three charts are shown in Figure 6, where the horizontal dashed lines denote the significance level $\alpha = 0.05$. From Figure 6, we can see that all the three charts give a signal of mean shift at the 123rd time point and the subsequent p-values are all well below the significant level. Therefore, these signals are convincing enough, which are also consistent with the findings in Hawkins and Olwell (1998).

At the end of this section, we would like to use this example to illustrate the use of p-values in designing the V-CUSUM chart with variable sampling intervals. Note that, in the original data, observations are collected at equally spaced time points. However, for the illustration purpose, we assume that they are collected with variable sampling intervals specified by our V-CUSUM chart. As in the simulation examples in Section 3, we choose $d_1 = 0.1$, $d_2 = 1.9$, $\alpha_1 = 0.5$, and $\alpha = 0.05$. In such cases, $ATS_0 = ARL_0 \approx 200$. For the testing data (i.e., from the 76th to 149th observations), the charting statistic values C_t^{+*}, the corresponding p-values $P_{C_t^{+*}}$, and the sampling intervals $d(t)$ of the V-CUSUM chart are presented in Table 2. From the table, after the 76th observation is obtained, the value of $P_{C_t^{+*}}$ is computed to be 0.086, which is between (α, α_1). Therefore, we collect the next observation (i.e., the 77th observation) after $d_1 = 0.1$ units of the time. This monitoring process continues until the 123rd observation at which we obtain a signal of mean shift.

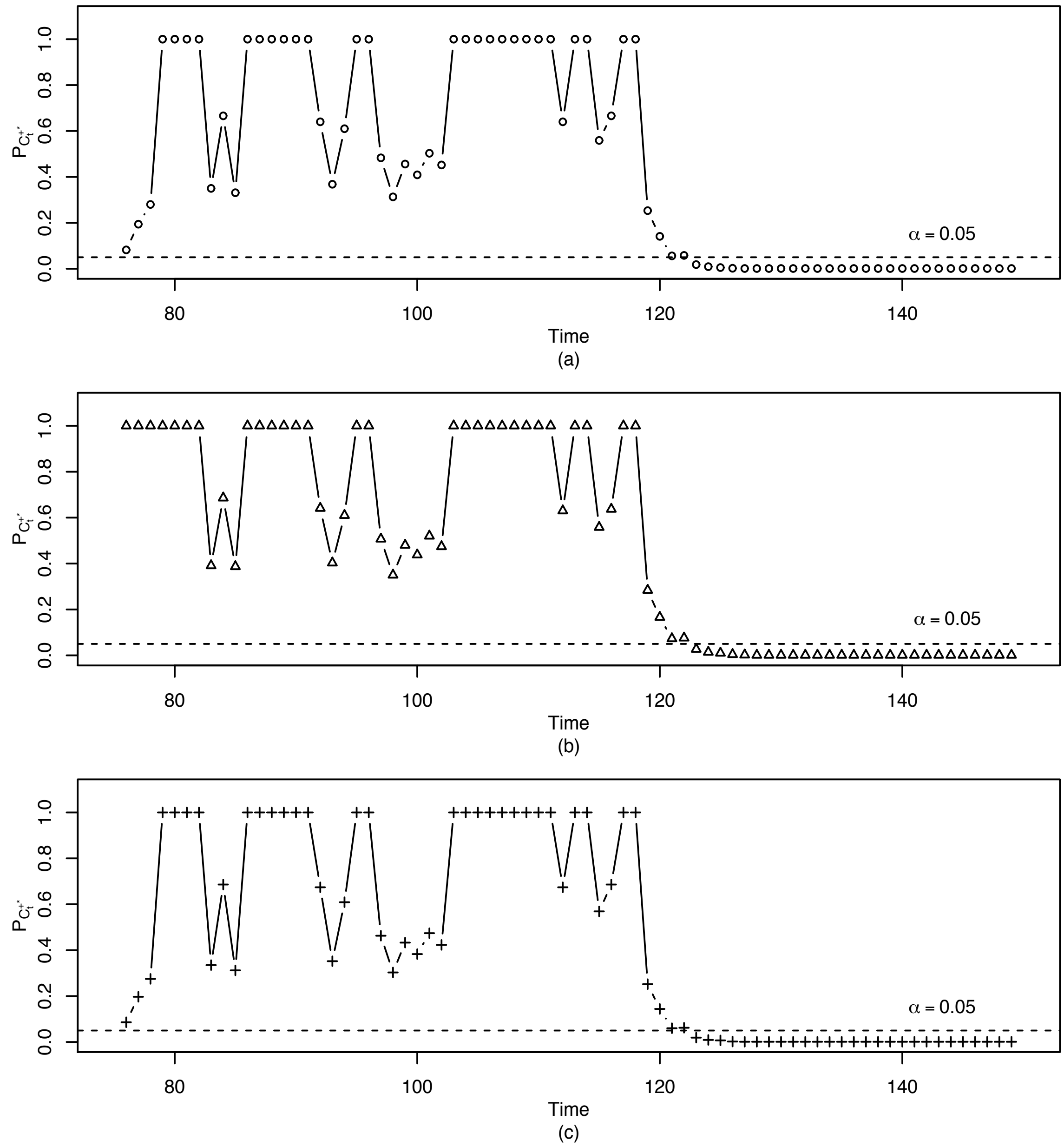

Figure 6: Control charts using p-values for monitoring the chemical process concerning triglyceride content. (a) N-CUSUM, (b) S-CUSUM, (c) B-CUSUM.

Table 2: Values of C_t^{+*}, $P_{C_t^{+*}}$ and $d(t)$ of the V-CUSUM chart.

t	C_t^{+*}	$P_{C_t^{+*}}$	$d(t)$	t	C_t^{+*}	$P_{C_t^{+*}}$	$d(t)$
76	4.4142	0.086	0.1	113	0.0000	1.000	1.9
77	2.6969	0.197	0.1	114	0.0000	1.000	1.9
78	1.9478	0.275	0.1	115	0.5417	0.569	1.9
79	0.0000	1.000	1.9	116	0.1153	0.686	1.9
80	0.0000	1.000	1.9	117	0.0000	1.000	1.9
81	0.0000	1.000	1.9	118	0.0000	1.000	1.9
82	0.0000	1.000	1.9	119	2.1552	0.252	0.1
83	1.5098	0.335	0.1	120	3.3424	0.144	0.1
84	0.1153	0.686	1.9	121	5.1749	0.060	0.1
85	1.6251	0.312	0.1	122	5.0712	0.062	0.1
86	0.0000	1.000	1.9	123	7.5491	0.019	0.1
87	0.0000	1.000	1.9	124	9.0589	0.009	0.1
88	0.0000	1.000	1.9	125	9.9234	0.007	0.1
89	0.0000	1.000	1.9	126	11.7559	0.002	0.1
90	0.0000	1.000	1.9	127	14.2338	0.001	0.1
91	0.0000	1.000	1.9	128	15.4209	0.001	0.1
92	0.2190	0.674	1.9	129	17.5762	0.001	0.1
93	1.4061	0.352	0.1	130	19.4087	0.001	0.1
94	0.3343	0.609	1.9	131	20.9185	0.001	0.1
95	0.0000	1.000	1.9	132	23.0738	0.001	0.1
96	0.0000	1.000	1.9	133	24.2609	0.001	0.1
97	0.8644	0.463	0.1	134	26.0934	0.001	0.1
98	1.7288	0.303	0.1	135	25.0216	0.001	0.1
99	0.9797	0.433	0.1	136	26.8541	0.001	0.1
100	1.1987	0.383	0.1	137	27.0731	0.001	0.1
101	0.7723	0.474	0.1	138	29.8738	0.001	0.1
102	0.9913	0.423	0.1	139	34.2880	0.001	0.1
103	0.0000	1.000	1.9	140	34.8297	0.001	0.1
104	0.0000	1.000	1.9	141	40.2120	0.001	0.1
105	0.0000	1.000	1.9	142	40.4310	0.001	0.1
106	0.0000	1.000	1.9	143	43.5543	0.001	0.1
107	0.0000	1.000	1.9	144	42.8052	0.001	0.1
108	0.0000	1.000	1.9	145	45.9286	0.001	0.1
109	0.0000	1.000	1.9	146	47.1157	0.001	0.1
110	0.0000	1.000	1.9	147	49.9163	0.001	0.1
111	0.0000	1.000	1.9	148	52.0716	0.001	0.1
112	0.2190	0.674	1.9	149	53.2587	0.001	0.1

5 Concluding Remarks

In this paper, we propose designing control charts using p-values and a significance level. Compared to conventional control charts, a chart using p-values would have several advantages. First, at the time when it gives a signal of process distributional shift, it also provides a measure of the likelihood of the shift, so that subsequent actions can be taken properly. Second, even when the chart does not give a signal of shift at a given time, it still provides a measure regarding the likelihood of a potential shift, which is helpful especially when a variable sampling scheme is adopted. Third, a chart using p-values is easier to interpret. Finally, it has a unified form with the y-axis ranging between 0 and 1 in all cases. Although we demonstrate the proposed method only in cases when upward process mean shifts are our concerns, it is a quite general method and can be used in most other cases.

Acknowledgments

The authors are grateful to the editor and two anonymous referees for their valuable comments that have greatly improved the paper. This research is finished during Li's visit to School of Statistics at The University of Minnesota, whose hospitality is appreciated. The research is supported in part by the NSF grants DMS-0706082 and SES-0851705, the Natural Sciences Foundation of China grants 11071128 and 11131002, the RFDP of China Grant 20110031110002, the Fundamental Research Funds for the Central Universities 65012231, and the Office of International Programs at University of Minnesota.

References

Chatterjee, S. and Qiu, P., 2009. Distribution-free cumulative sum control charts using bootstrap-based control limits. The Annals of Applied Statistics, 3(1), 349-369.

Costa, A. F. B., 1998. Joint $\overline{X}$ and R charts with variable parameters. IIE Transactions, 30, 505-514.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. The Annals of Statistics, 7, 1-26.

Efron, B. and Tibshirani, R.J., 1993. *An introduction to the bootstrap.* Chapman and Hall, Boca Raton, FL.

Grigg, O. A. and Spiegelhalter, D. J., 2008. An empirical approximation to the null unbounded steady-state distribution of the cumulative sum statistic. Technometrics, 50, 501-511.

Hawkins, D. M., 1969. On the distribution and power of a test for a single outlier. South African Statistical Journal, 3, 9–15.

Hawkins, D. M., 1987. Self-starting CUSUM charts for location and scale. The Statistician, 36 (4), 299–316.

Hawkins, D. M. and Olwell, D. H., 1998. *Cumulative sum charts and charting for quality improvement.* Springer, Berlin.

Hawkins, D. M., Qiu, P. and Kang, C. W., 2003. The change point model for statistical process control. Journal of Quality Technology, 35 (4), 355-366.

Jensen, W. A., Jones-Farmer, L. A., Champ, C. W. and Woodall, W. H., 2006. Effects of parameter estimation on control chart properties: a literature review. Journal of Quality Technology, 38 (4), 349-364.

Jones, L. A., 2002. The statistical design of EWMA control charts with estimated parameters. Journal of Quality Technology, 34 (3), 277-288.

Jones, L. A., Champ, C. W. and Rigdon, S. E., 2001. The performance of exponentially weighted moving average charts with estimated parameters. Technometrics, 43 (2), 156-167.

Jones, L. A., Champ, C. W. and Rigdon, S. E., 2004. The run length distribution of the CUSUM with estimated parameters. Journal of Quality Technology, 36 (1), 95-108.

Lehmann, E.L. (1997), *Testing statistical hypotheses*, New York: Springer.

Li, Y., and Tsung, F. 2009. False discovery rate-adjusted charting schemes for multistage process monitoring and fault identification. Technometrics, 51, 186-205.

Lucas, J.M. and Saccucci, M.S., 1990. Exponentially weighted moving average control schemes: Properties and enhancements. Technometrics, 32, 1-12.

Luo, Y., Li, Z., Wang, Z., 2009. Adaptive CUSUM control chart with variable sampling intervals. Computational Statistics and Data Analysis, 53, 2693-2701.

Montgomery, D.C., 2004. *Introduction to statistical quality control*, fifth ed. John Wiley & Sons, New York.

Montgomery, D. C., 2007. SPC research–current trends. Quality and Reliability Engineering International, 23, 515-516.

Page, E. S., 1954. Continuous Inspection Schemes. Biometrika, 42, 243-254.

Qiu, P. and Hawkins, D. M. 2001. A rank based multivariate CUSUM procedure. Technometrics, 43, 120-132.

Qiu, P. and Li, Z. 2011. On nonparametric statistical process control of univariate processes. Technometrics, 53, 390-405.

Quesenberry, C. P. 1991. SPC Q charts for start-up processes and short or long runs. Journal of Quality Technology, 23 (3), 213-224.

Reynolds, M. R., Jr., Amin, R. W. and Arnold, J. C., 1990. CUSUM charts with variable sampling intervals. Technometrics, 32 (4), 371-384.

Reynolds, M. R., Jr. and Arnold, J. C., 2001. EWMA control charts with variable sample sizes and variable sampling intervals. IIE Transactions, 33, 511-530.

Roberts, G. C., 1959. Control Chart Tests Based on Geometric Moving Average. Technometrics, 1, 239-250.

Shewhart, W. A., 1931. *Economic control of quality of manufactured product.* New York: Van Nostrand.

Wu, Z., Zhang, S. and Wang, P., 2007. A CUSUM scheme with variable sample sizes and sampling intervals for monitoring the process mean and variance. Quality and Reliability Engineering International, 23, 157-170.

Zhou, C., Zou, C., Zhang, Y., and Wang Z., 2009. Nonparametric control chart based on change-point model. Statistical Papers, 50, 13-28.