On Nonparametric Statistical Process Control Of Univariate Processes

Peihua Qiu[1] and Zhonghua Li[1,2]

[1]School of Statistics, University of Minnesota, USA

[2]LPMC and Department of Statistics, Nankai University, China

Abstract

This paper considers statistical process control (SPC) of univariate processes when the parametric form of the process distribution is unavailable. Most existing SPC procedures are based on the assumption that a parametric form (e.g., normal) of the process distribution can be specified beforehand. In the literature, it has been demonstrated that their performance is unreliable in cases when the pre-specified process distribution is invalid. To overcome this limitation, some nonparametric (or distribution-free) SPC control charts have been proposed, most of which are based on the ordering information of the observed data. This paper tries to make two major contributions to the nonparametric SPC literature. First, we propose an alternative framework for the construction of nonparametric control charts, by first categorizing observed data and then applying methods of categorical data analysis to SPC. Under this framework, some new nonparametric control charts are proposed. Second, we compare our proposed control charts with several representative existing control charts in various cases. Some empirical guidelines are provided for practitioners to choose a proper nonparametric control chart for a specific application.

Key Words: Categorical data; Categorization; Discrete data; Distribution-free; Non-Gaussian data; Nonparametric procedures; Ranks.

1 Introduction

Statistical process control (SPC) charts are widely used in industry for monitoring the stability of certain sequential processes (e.g., manufacturing processes, health care systems, internet traffic flow, and so forth). In practice, SPC is generally divided into two phases. In Phase I, a set of process data is gathered and analyzed. Any unusual "patterns" in the data lead to adjustments and fine tuning of the process. Once all such assignable causes are accounted for, we are left with a

clean set of data, gathered under stable operating conditions and illustrative of the actual process performance. This set is then used for estimating the in-control (IC) distribution of the process response. In Phase II, the estimated IC distribution of the response from a Phase I data is used, and the major goal of this phase is to detect changes in the response distribution after an unknown time point. Performance of a Phase II SPC procedure is often measured by the average run length (ARL), which is the average number of observations needed for the procedure to signal a change in the response distribution. The IC ARL value of the procedure is often controlled at some specific level. Then, the procedure performs better if its out-of-control (OC) ARL is shorter, when detecting a given distributional change. See, e.g., Chen et al. (2001), Chen and Zhang (2005), Hawkins and Olwell (1998), Montgomery (2001), Woodall (2000), and Yeh et al. (2004) for related discussion. This paper focuses on Phase II monitoring of univariate processes in cases when a parametric form (e.g., normal) of the process response distribution is unavailable.

In the literature, many Phase II SPC charts have been proposed, including different versions of the Shewhart chart, the cumulative sum (CUSUM) chart, the exponentially weighted moving average (EWMA) chart, and the chart based on change-point detection (cf., e.g., Hawkins and Olwell 1998, Hawkins et al. 2003, Montgomery 2001). Design and implementation of these control charts usually requires the assumption that the process response distribution follows a parametric form (e.g., normal). In practice, however, process observations may not follow the pre-specified parametric distribution. In such cases, it has been demonstrated in the literature that results from these charts using the pre-specified distribution in their design may not be reliable (cf., Amin et al. 1995, Hackl and Ledolter 1992, Lucas and Crosier 1982). As a demonstration, Figure 1 shows the actual IC ARL values of the conventional CUSUM chart based on the assumption that the IC response distribution is $N(0, 1)$, in cases when the allowance constant of the chart (see introduction in Section 3.1) is 0.5, the assumed IC ARL value equals 500 and the true response distribution is the standardized version (with mean 0 and variance 1) of the chi-square (plot (a)) or t (plot (b)) distribution with degrees of freedom (df) changing from 1 to 60 in plot (a) and from 3 to 60 in plot (b). From the plots, it can be seen that the actual IC ARL values of the conventional CUSUM chart are much smaller than the nominal IC ARL value when the df is small, which implies that the related process would be stopped too often by the control chart and consequently a considerable amount of time and resource would be wasted in such cases.

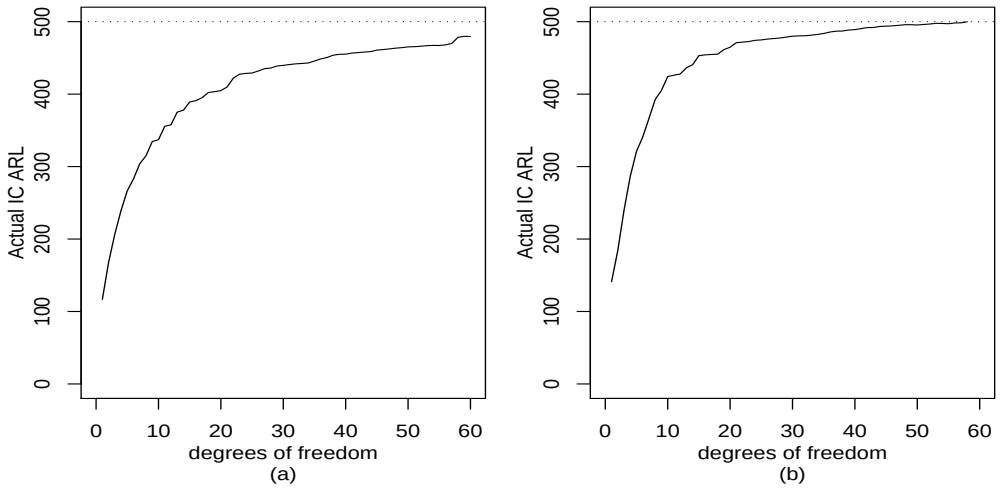

Figure 1: Actual IC ARL values of the conventional CUSUM chart in cases when the nominal IC ARL value is 500, and the true response distribution is the standardized version (with mean 0 and variance 1) of the chi-square (plot (a)) or t (plot (b)) distribution with degrees of freedom changing from 1 to 60 in plot (a) and from 3 to 60 in plot (b).

From Figure 1, it seems necessary to develop appropriate control charts that do not require specifying the parametric form of the response distribution beforehand. To this end, a number of distribution-free or nonparametric control charts have been developed. See, for instance, Albers and Kallenberg (2004, 2009), Albers et al. (2006), Amin et al. (1995), Amin and Searcy (1991), Amin and Widmaier (1999), Bakir(2004, 2006), Bakir and Reynolds (1979), Chakraborti et al. (2000, 2004, 2009), and Chakraborti and Eryilmaz(2007). Chakraborti et al. (2001) gives a thorough overview on existing research in the area of univariate distribution-free SPC. In multivariate cases, see Qiu and Hawkins (2001, 2003) and Qiu (2008) for related discussion.

Most existing nonparametric SPC charts mentioned above are based on the ordering or ranking information of the observations obtained at the same or different time points. Some of them require multiple observations at each time point (i.e., cases with batch or subgrouped data). In this paper, we suggest an alternative method to handle the univariate SPC problem with unknown response distribution. By the proposed method, observed data are first categorized, and then certain statistical procedures for categorical data analysis are used for constructing nonparametric SPC charts. Major considerations behind this method are as follows. First, statistical tools for describing and analyzing non-normal numerical data are limited; but, there are a considerable amount of statistical tools for handling categorical data (cf., Agresti 2002). Second, while both

data ranking and data categorization result in a loss of information, the amount of lost information due to categorization can be controlled by the number of categories used. The bigger the number of categories used, the less information would be lost. As a comparison, the information loss due to ranking can hardly be controlled. Furthermore, the proposed method does not require multiple observations at a single time point. It does not require observations to be numerical either. Methods based on ranking are generally difficult to use in such cases.

For a given application with unknown response distribution, an important decision to make is to choose an appropriate control chart from all the existing ones. To address this issue, a numerical study is performed to compare our proposed control charts with certain representative existing control charts in various cases. Some practical guidelines are provided for users to choose a proper control chart for a specific application. This comparative study shows that our proposed control charts perform favorably in the cases considered.

Phase II monitoring of categorized (or grouped) data has been discussed in the literature by Steiner et al. (1996) and Steiner (1998). In their papers, Steiner and co-authors consider the case when the IC distribution has a known parametric form (e.g., normal), individual observations may not be completely known, instead, it is known that they belong to certain given intervals. In such cases, a CUSUM chart based on the likelihood ratio formulation and an EWMA chart have been proposed. Obviously, the motivation of the current research is different from that in Steiner et al. (1996) and Steiner (1998). In the current research, we consider the case when the parametric form of the IC distribution is unavailable; thus, the regular likelihood ratio formulation is unavailable either. Categorization of the observed numerical data is for overcoming this difficulty, in order to use existing statistical methods for categorical data analysis.

The rest part of the paper is organized as follows. Our proposed nonparametric control charts are described in Section 2. A numerical study to evaluate their performance in comparison with several existing control charts is presented in Section 3. An application is discussed in Section 4 to demonstrate the use of the proposed methods in a real world setting. Some remarks conclude the article in Section 5.

2 Proposed Nonparametric Control Charts

As described in Section 1, our proposed nonparametric control charts can monitor processes with both categorical and numerical observations. For simplicity, we focus on cases with continuous observations first. Cases with discrete or categorical observations can be handled similarly, and will be briefly discussed at the end of this section. Like some existing nonparametric control charts (e.g., Chakraborti et al. 2009), our charts do not assume that the IC process response cumulative distribution function (cdf) F_0 is known. Instead, we assume that an IC dataset has been collected in the Phase I analysis, and it can be used for estimating certain IC parameters. Note that, in practice, it might still be an issue how to do Phase I analysis in cases when F_0 is nonparametric and unknown, although Phase I analysis in such cases is not the focus of the current research.

Let $\mathbf{X}(n) = (X_1(n), X_2(n), \ldots, X_m(n))'$ be m observations obtained at the nth time point during Phase II process monitoring. In the literature, such data are often called subgrouped or batch data, and m is the batch size. When $m = 1$, the data are sometimes called single-observation data. The first step of our proposed method to construct nonparametric control charts is to categorize the observed data as follows. Let $I_1 = (-\infty, q_1], I_2 = (q_1, q_2], \ldots, I_p = (q_{p-1}, \infty)$ be a partition of the real line, where $-\infty < q_1 < q_2 < \cdots < q_{p-1} < \infty$ are $p-1$ boundary points of the partitioning intervals. Define

$$Y_{jl}(n) = I(X_j(n) \in I_l), \text{ for } j = 1, 2, \ldots, m, \ l = 1, 2, \ldots, p, \tag{1}$$

and $\mathbf{Y}_j(n) = (Y_{j1}(n), Y_{j2}(n), \ldots, Y_{jp}(n))'$, where $I(a) = 1$ if $a =$"True" and 0 otherwise. Then, $\mathbf{Y}_j(n)$ has one and only one component being 1, the index of the component being 1 is the index of the partitioning interval that $X_j(n)$ belongs to, and this index has a discrete distribution with probabilities $f_l = P(X_j(n) \in I_l)$, for $l = 1, 2, \ldots, p$. For simplicity, $\mathbf{f} = (f_1, f_2, \ldots, f_p)'$ is also called the distribution of $\mathbf{Y}_j(n)$ in this paper. Let $\mathbf{f}^{(0)} = (f_1^{(0)}, f_2^{(0)}, \ldots, f_p^{(0)})'$ be the IC distribution of $\mathbf{Y}_j(n)$, and $\mathbf{f}^{(1)} = (f_1^{(1)}, f_2^{(1)}, \ldots, f_p^{(1)})'$ be its OC distribution. For the distribution of $\mathbf{Y}_j(n)$, it can be checked that, if the support of the IC distribution F_0 of the process observations contains at least one of the boundary points $\{q_1, q_2, \ldots, q_{p-1}\}$ and $p \geq 2$, then any mean shift in $\mathbf{X}(n)$ would result in changes in the distribution $\mathbf{f}$. It should be pointed out that the conditions used in this result are weak. For instance, the supports of most commonly used continuous numerical distributions (e.g., normal, t, χ^2, uniform, exponential distributions) are connected intervals on the number line. For

these distributions, any reasonable set of boundary points $\{q_1, q_2, \ldots, q_{p-1}\}$ should have at least one point contained in their supports. Otherwise, there must exist one partitioning interval that contains all observations, which is contradictory to the purpose of categorization. By the above result, under some mild regularity conditions, $\mathbf{f}^{(1)}$ would be different from $\mathbf{f}^{(0)}$ if there is a mean shift in $\mathbf{X}(n)$. Therefore, detection of a mean shift in $\mathbf{X}(n)$ is equivalent to detection of a change in the distribution of $\mathbf{Y}_j(n)$ in such cases.

To define $\mathbf{Y}_j(n)$ by (1), we need to choose the boundary points $\{q_1, q_2, \ldots, q_{p-1}\}$ beforehand. For that purpose, existing research in categorical data analysis has demonstrated that it would help detect shifts from $\mathbf{f}^{(0)}$ if the boundary points are chosen such that the expected counts under F_0 in the partitioning intervals are roughly the same (cf., e.g., Agresti 2002, Section 1.5). Therefore, we suggest choosing q_l to be the l/p-th quantile of F_0, for $l = 1, 2, \ldots, p-1$. In such cases, $\mathbf{f}^{(0)} = (1/p, 1/p, \ldots, 1/p)'$. In practice, q_l can be estimated by the l/p-th sample quantile of the IC dataset, and the resulting estimator of $\mathbf{f}^{(0)}$, denoted as $\widehat{\mathbf{f}}^{(0)}$, would be a statistically consistent estimator of $\mathbf{f}^{(0)}$.

To test whether the distribution of $\mathbf{Y}_j(n)$ equals its IC distribution $\mathbf{f}^{(0)}$, the Pearson's chi-square test is a natural choice. Let

$$g_l(n) = \sum_{j=1}^{m} Y_{jl}(n), \text{ for } l = 1, 2, \ldots, p, \ n = 1, 2, \ldots, \tag{2}$$

and $\mathbf{g}(n) = (g_1(n), g_2(n), \ldots, g_p(n))'$. Then, $g_l(n)$, for $l = 1, 2, \ldots, p$, denotes the number of observations at the n-th time point that fall into the l-th interval I_l. The Pearson's chi-square test statistic at the n-th time point is defined by

$$\widetilde{X}^2(n) = \sum_{l=1}^{p} \frac{\left[g_l(n) - m f_l^{(0)}\right]^2}{m f_l^{(0)}},$$

which measures the discrepancy between the observed counts $\mathbf{g}(n)$ and the expected counts $m\mathbf{f}^{(0)}$ in the p partitioning intervals at the n-th time point. However, this statistic uses observations at a single time point only; it may not be effective to detect persistent changes in the distribution of $\mathbf{Y}_j(n)$, as demonstrated in the literature (cf., Hawkins and Olwell 1998). To overcome this limitation, we suggest the following CUSUM chart, which adopts the structure of Crosier's (1988) multivariate CUSUM scheme, although the two charts are for different purposes. Let $\mathbf{S}_0^{obs} = \mathbf{S}_0^{exp} =$

$\mathbf{0}$ be two $p \times 1$ column vectors, and

$$\begin{cases} \mathbf{S}_n^{obs} = \mathbf{0}, & \text{if } C_n \le k_P \\ \mathbf{S}_n^{exp} = \mathbf{0}, & \\ \mathbf{S}_n^{obs} = (\mathbf{S}_{n-1}^{obs} + \mathbf{g}(n))(C_n - k_P)/C_n, & \text{otherwise} \\ \mathbf{S}_n^{exp} = (\mathbf{S}_{n-1}^{exp} + m\mathbf{f}^{(0)})(C_n - k_P)/C_n, & \end{cases}$$

where

$$C_n = \left\{ (\mathbf{S}_{n-1}^{obs} - \mathbf{S}_{n-1}^{exp}) + \left(\mathbf{g}(n) - m\mathbf{f}^{(0)} \right) \right\}' \left(diag(\mathbf{S}_{n-1}^{exp} + m\mathbf{f}^{(0)}) \right)^{-1} \left\{ (\mathbf{S}_{n-1}^{obs} - \mathbf{S}_{n-1}^{exp}) + \left(\mathbf{g}(n) - m\mathbf{f}^{(0)} \right) \right\},$$

$k_P \ge 0$ is the so-called allowance parameter, $diag(A)$ denotes a diagonal matrix with its diagonal elements equal to the corresponding elements of the vector A, and the superscripts "obs" and "exp" denote observed and expected counts, respectively. Define

$$u_{n,P} = \left(\mathbf{S}_n^{obs} - \mathbf{S}_n^{exp} \right)' (diag(\mathbf{S}_n^{exp}))^{-1} \left(\mathbf{S}_n^{obs} - \mathbf{S}_n^{exp} \right).$$

Then, a mean shift in $\mathbf{X}(n)$ is signaled if

$$u_{n,P} > h_P, \tag{3}$$

where $h_P > 0$ is a control limit chosen to achieve a given IC ARL level. Chart (3) will be called P-CUSUM chart hereafter, to reflect the fact that it is constructed based on the Pearson's chi-square test.

When $k_P = 0$, it is not difficult to check that $\mathbf{S}_n^{obs}$ is a frequency vector with its l-th element denoting the cumulative observed count of observations in the l-th interval I_l as of the time point n, for $l = 1, 2, \ldots, p$, and $\mathbf{S}_n^{exp}$ equals $nm\mathbf{f}^{(0)}$ which is the vector of the corresponding cumulative expected counts. Therefore, in such cases, $u_{n,P}$ is the conventional Pearson's chi-square test statistic that measures the difference between the cumulative observed and expected counts as of the time point n. Further, it can be checked (cf., Appendix C, Qiu and Hawkins 2001) that

$$u_{n,P} = \max(0, C_n - k_P).$$

Therefore, the charting statistic $u_{n,P}$ is defined in the way that the CUSUM chart can be repeatedly restarted when there is little evidence of distributional shift in $\mathbf{Y}_j(n)$.

For the P-CUSUM chart (3), the control limit h_P can be determined easily by a numerical algorithm as follows. First, choose an initial value for h_P. Then, compute the IC ARL value

of the P-CUSUM chart based on a large number (e.g., 10000) of replicated simulation runs in which the IC multinomial observations $\mathbf{Y}_j(n)$ are sequentially generated from the IC distribution $\mathbf{f}^{(0)} = (1/p, 1/p, \ldots, 1/p)'$. If the computed IC ARL value is smaller than the nominal IC ARL value, then increase the value of h_P. Otherwise, choose a smaller h_P value. The above process is repeated until the nominal IC ARL value is reached within a desired precision. In this process, some numerical searching schemes, such as the bisection search, can be applied (cf., Qiu 2008). From the above description, it can be seen that determination of h_P does not require any information about the IC distribution F_0 of the process response. Instead, it only depends on the nominal IC ARL value, the allowance constant k_P, the batch size m, and the number of categories p. In this sense, the P-CUSUM chart is distribution-free. For some commonly used values of the nominal IC ARL, k_P, m, and p, we compute the h_P values and the results are presented in Table 2 in Section 3.3.

As a remark, due to the fact that the charting statistic $u_{n,P}$ takes discrete values on the positive number line, certain pre-specified nominal IC ARL values cannot be reached within a desired precision. This phenomenon is common when handling discrete data (cf., Section 5.2, Hawkins and Olwell 1998). In actual implementation of the proposed control chart, it should be okay to use an IC ARL value that the chart can reach. However, when comparing different control charts, it is often desirable to let all of them have a common IC ARL value, which may not be achieved easily in our method. To overcome this technical difficulty, in this paper, we propose the following simple but efficient modification of the charting statistic $u_{n,P}$. Let $\mathbf{b}_j(n)$ be a sequence of i.i.d. random vectors generated from the distribution $\mathbf{N}_p(\mathbf{0}, s^2\mathbf{I}_{p\times p})$, where s^2 is a small positive number and $\mathbf{I}_{p\times p}$ is the $p \times p$ identity matrix. Then, when computing $u_{n,P}$, we suggest replacing $\mathbf{Y}_j(n)$ by $\mathbf{Y}_j(n) + \mathbf{b}_j(n)$. Namely, we add a small random number to each component of $\mathbf{Y}_j(n)$ to alleviate the discreteness of the charting statistic $u_{n,P}$. Based on our numerical study (not reported here but available from the authors), as long as s is chosen small, the OC behavior of the P-CUSUM chart is little affected by this modification. However, most nominal IC ARL values can be reached within a desired precision after the modification, which can be seen from some numerical results reported in Section 3 (cf., Table 1). By the way, in all the numerical examples presented in this paper, we use $s = 0.01$.

To detect changes in the distribution of $\mathbf{Y}_j(n)$, besides the Pearson's chi-square test (cf., $\widetilde{X}^2(n)$

defined above), another popular test is based on the likelihood ratio statistic

$$\widetilde{G}^2(n) = 2\sum_{l=1}^{p} g_l(n) \log\left(\frac{g_l(n)}{m f_l^{(0)}}\right).$$

Then, a CUSUM chart can be constructed similarly to (3) as follows. Let $\widetilde{\mathbf{S}}_n^{obs}$ and $\widetilde{\mathbf{S}}_n^{exp}$ be quantities defined in the same way as $\mathbf{S}_n^{obs}$ and $\mathbf{S}_n^{exp}$ used in chart (3), except that the allowance parameter k_P is replaced by another constant k_L, and C_n is replaced by

$$\widetilde{C}_n = 2\left(\widetilde{\mathbf{S}}_{n-1}^{obs} + \mathbf{g}(n)\right)' \log\left(\frac{\widetilde{\mathbf{S}}_{n-1}^{obs} + \mathbf{g}(n)}{\widetilde{\mathbf{S}}_{n-1}^{exp} + m\mathbf{f}^{(0)}}\right),$$

where $\mathbf{a}/\mathbf{b}$ denotes a vector obtained by component-wise division of the vector $\mathbf{a}$ divided by the vector $\mathbf{b}$, and $\log(\mathbf{a}/\mathbf{b})$ denotes a vector obtained by applying the log operator to each component of $\mathbf{a}/\mathbf{b}$. Define

$$u_{n,L} = 2\left(\widetilde{\mathbf{S}}_n^{obs}\right)' \log\left(\frac{\widetilde{\mathbf{S}}_n^{obs}}{m\mathbf{f}^{(0)}}\right).$$

Then a mean shift in $\mathbf{X}(n)$ is signaled if

$$u_{n,L} > h_L, \tag{4}$$

where the control limit $h_L > 0$ can be chosen similarly to h_P to achieve a given IC ARL level. The control chart (4), that is based on the likelihood ratio test, is called L-CUSUM chart in this paper.

In categorical data analysis, there is much discussion about the relationship between the two tests based on $\widetilde{X}^2(n)$ and $\widetilde{G}^2(n)$, respectively (cf., Agresti 2002, Section 1.5.5). Basically, the two tests are asymptotically equivalent, and the distribution of $\widetilde{X}^2(n)$ converges to a chi-square distribution faster than that of $\widetilde{G}^2(n)$, when the observed counts in $\mathbf{g}(n)$ increase. In SPC, however, asymptotic properties are usually irrelevant because the related control chart should be stopped immediately after a signal of mean shift is obtained. Therefore, relationship between the P-CUSUM and L-CUSUM charts still needs to be studied, which will be discussed in Section 3.2.

From the construction of the charts (3) and (4), we notice that the partitioning intervals $\{I_1, I_2, \ldots, I_p\}$ are ordered on the number line; but, both the charting statistics $u_{n,P}$ and $u_{n,L}$ ignore such ordering information contained in the observed counts $\{g_1(n), g_2(n), \ldots, g_p(n)\}$ completely. In the nonparametric statistics literature, a popular test that takes into account the ordering information among observations when testing a potential change in the observation distribution is

the Kolmogorov-Smirnov goodness-of-fit test (cf., Chakravarti et al. 1967). To apply this test here for detecting potential shifts from the IC process response distribution F_0, the test statistic can be written as

$$D(n) = \max_{1\le j\le m}\left(\widehat{F}_0\left(X_{(j)}(n)\right) - \frac{j-1}{m},\ \frac{j}{m} - \widehat{F}_0\left(X_{(j)}(n)\right)\right),$$

where $\{X_{(j)}(n), j = 1, 2, \ldots, m\}$ are order statistics of $\{X_j(n), j = 1, 2, \ldots, m\}$, and $\widehat{F}_0$ is the empirical IC cdf constructed from the IC data. Based on $D(n)$, a CUSUM chart can be constructed as follows. Let $u_{0,K} = 0$, and for $n \ge 1$,

$$u_{n,K} = \max\left(0, u_{n-1,K} + (D(n) - D_0) - k_K\right),$$

where D_0 denotes the IC mean of $D(n)$ which can be estimated from the IC data, and k_K is an allowance constant. Then, a mean shift in $\mathbf{X}(n)$ is signaled if

$$u_{n,K} > h_K, \tag{5}$$

where $h_K > 0$ is the control limit. Chart (5) is called K-CUSUM chart in the remaining part of the paper. It should be pointed out that the K-CUSUM chart (5) is constructed from the original observations $\mathbf{X}(n)$, instead of from their categorized version $\{\mathbf{Y}_j(n), j = 1, 2, \ldots, m\}$. Although this chart can also be constructed from $\{\mathbf{Y}_j(n), j = 1, 2, \ldots, m\}$, this approach is not recommended due to the reasons that (i) chart (5) based on $\mathbf{X}(n)$ is already nonparametric, and (ii) chart (5) based on $\{\mathbf{Y}_j(n), j = 1, 2, \ldots, m\}$ would lose certain efficiency when detecting potential mean shifts in $\mathbf{X}(n)$.

The above discussion focuses on cases when process observations are continuous. Situations when observations are discrete or categorical can be handled similarly. For instance, when process observations are categorical, then the categorization step (cf., equation (1)) can be skipped when constructing the proposed control charts P-CUSUM and L-CUSUM (note: chart K-CUSUM can only be used in cases with numerical observations). When process observations are discrete and the number of different observation values is quite small, we can use each observation value as a single category, and construct the three charts as usual. In cases when process observations are discrete but the number of different observation values is quite large, proper combination of certain observation values might be necessary when defining the categories. To this end, practical guidelines provided in the paragraph immediately before the one containing equation (2) and the guidelines given in Section 3.3 regarding construction of the partitioning intervals should be helpful.

3 Numerical Study

In this section, we present some numerical examples to evaluate the performance of the charts P-CUSUM, L-CUSUM and K-CUSUM, in comparison with certain representative existing control charts. The existing control charts considered here are briefly introduced in Section 3.1. Then, the related control charts are compared in various cases in Section 3.2. From the comparison, we can see that the P-CUSUM chart has a favorable performance in different cases. To use this chart, the number of categories p should be properly chosen beforehand. This and other issues of the P-CUSUM chart are discussed in Section 3.3.

3.1 Some representative existing control charts

The traditional CUSUM chart is a standard tool for monitoring univariate processes in practice. Its charting statistics of the two-sided version are defined by

$$\begin{aligned} u^+_{n,N} &= \max\left(0, u^+_{n-1,N} + \overline{X}(n) - k_N\right), \\ u^-_{n,N} &= \min\left(0, u^-_{n-1,N} + \overline{X}(n) + k_N\right), \text{ for } n \geq 1, \end{aligned}$$

where $u^+_{0,N} = u^-_{0,N} = 0$, k_N is an allowance constant, and $\overline{X}(n) = \frac{1}{m}\sum_{j=1}^{m} X_j(n)$. Then, a mean shift in $\mathbf{X}(n)$ is signaled if

$$u^+_{n,N} > h^+_N \quad \text{or} \quad u^-_{n,N} < -h^-_N \tag{6}$$

where the control limits $h^+_N, h^-_N > 0$ are chosen to achieve a given IC ARL level. When the IC process distribution is symmetric (e.g., normal or t), then we can set $h^+_N = h^-_N$. For skewed IC process distributions, h^+_N and h^-_N can be chosen different as follows. As demonstrated by Hawkins and Olwell (1998), when n is large enough (e.g., $n \geq 50$), the distributions of the charting statistics $u^+_{n,N}$ and $u^-_{n,N}$ would become stable (i.e., they do not change much when n increases), which is the so-called steady-state phenomenon. So, in cases when the IC process distribution is known, we can determine the steady-state distributions of $u^+_{n,N}$ and $u^-_{n,N}$ by simulation and choose h^+_N and h^-_N such that $P(u^+_{n,N} > h^+_N) = P(u^-_{n,N} < -h^-_N)$ and the pre-specified IC ARL value is achieved. When the IC process distribution is unknown but an IC dataset is available, the steady-state distributions of $u^+_{n,N}$ and $u^-_{n,N}$ can be estimated by a bootstrap procedure using bootstrap samples drawn from the IC data, as done in Chatterjee and Qiu (2009), and then h^+_N and h^-_N can be determined in a

similar way to that when the IC process distribution is known, as described above. The chart (6) is called N-CUSUM chart hereafter.

When the process distribution is not normal, Borror et al. (1999) showed that a properly designed EWMA chart is robust to departures from normality. More specifically, the EWMA charting statistic is defined by

$$v_n = \lambda \overline{X}(n) + (1-\lambda) v_{n-1},$$

where $v_0 = 0$, $\lambda \in [0,1]$ is a weighting parameter, and $\overline{X}(n) = \frac{1}{m}\sum_{j=1}^{m} X_j(n)$. Then, a mean shift in $\mathbf{X}(n)$ is signaled if

$$|v_n| \geq h_R,$$

where $h_R > 0$ is a control limit chosen to achieve a pre-specified IC ARL level under the normality assumption. Borror et al. (1999) showed that this chart when $\lambda = 0.05$ performed reasonably well in various cases when the IC process distribution was actually non-normal. This EWMA chart with $\lambda = 0.05$ is called R-EWMA chart hereafter.

A distribution-free control chart by Chakraborti and Eryilmaz (2007) is also considered in our numerical study. The chart originally proposed by Chakraborti and Eryilmaz (2007) is a Shewhart-type chart, and it is constructed based on the statistic

$$\psi(n) = 2W_n^+ - \frac{m(m+1)}{2}, \text{ for } n \geq 1,$$

where W_n^+ is the Wilcoxon signed-rank statistic of $\mathbf{X}(n)$, defined to be the sum of the ranks of $\{|X_j(n) - \theta_0|, j = 1, 2, \ldots, m\}$ over all positive components of $\{X_j(n) - \theta_0, j = 1, 2, \ldots, m\}$, and θ_0 is the IC median of the process distribution which can be estimated from the IC data. As well demonstrated in the literature (e.g., Hawkins and Olwell 1998), CUSUM charts are more favorable for detecting persistent shifts, compared to Shewhart-type charts. For that reason, we construct a CUSUM chart based on $\psi(n)$ as follows. Let $u_{0,S}^+ = u_{0,S}^- = 0$, and

$$\begin{aligned} u_{n,S}^+ &= \max\left(0, u_{n-1,S}^+ + (\psi(n) - \psi_0) - k_S\right), \\ u_{n,S}^- &= \min\left(0, u_{n-1,S}^- + (\psi(n) - \psi_0) + k_S\right), \text{ for } n \geq 1, \end{aligned}$$

where k_S is an allowance constant, and ψ_0 is the IC mean of $\psi(n)$ which can be estimated from the IC data. Then, the CUSUM chart signals a mean shift in $\mathbf{X}(n)$ if

$$u_{n,S}^+ > h_S \quad \text{or} \quad u_{n,S}^- < -h_S, \tag{7}$$

where the control limit h_S is chosen to achieve a given IC ARL level. The chart (7) is called S-CUSUM chart hereafter.

As described in Section 1, Steiner et al. (1996) proposed a CUSUM chart for grouped data. To use this control chart, we need to assume that the process response distribution has a parametric cdf $F(x, \theta)$, where θ is a parameter. When the process is IC, $\theta = \theta_0$. After the process becomes OC, θ shifts to θ_1. Assume that the data are grouped into s intervals: $(-\infty, t_1], (t_1, t_2], \ldots, (t_{s-1}, \infty)$. For $j = 1, 2, \ldots, s$, let

$$\pi_j^+(\theta) = F(t_j, \theta) - F(t_{j-1}, \theta), \qquad l_j^+ = \log\left(\pi_j^+(\theta_1)/\pi_j^+(\theta_0)\right), \qquad w_j^+ = round\left(Q^+ l_j^+\right),$$

where $t_0 = -\infty$, $t_s = \infty$, $round(a)$ is a rounding function to round a to its nearest integer, and $Q^+ = 50/[\max(l_1^+, \ldots, l_s^+) - \min(l_1^+, \ldots, l_s^+)]$. Then, to detect the shift in θ from θ_0 to θ_1, the charting statistic is defined to be

$$T_n^+ = \max\left(0, T_{n-1}^+ + \sum_{j=1}^{s} m_j^+(n) l_j^+\right),$$

where $T_0^+ = 0$, and $m_j^+(n)$ denotes the number of observations that belong to the interval $(t_{j-1}, t_j]$ at the n-th time point, for $j = 1, 2, \ldots, s$. Then, the chart gives a signal of shift if

$$T_n^+ > h_{St},$$

where $h_{St} > 0$ is a control limit chosen to achieve a given IC ARL value. In this paper, we choose F to be the cdf of $N(\theta, 1)$, $\theta_0 = 0$, and $s = 5$. As suggested by Steiner et al. (1996), $\{t_j, j = 1, 2, \ldots, s\}$ are chosen to be the optimal CUSUM gauge points listed in Table 2 of their paper that are scaled by θ_1. From the above description, we can see that the control chart using T_n^+ is one-sided; but, all other control charts considered in the paper are two-sided. To make the related control charts comparable, we modify it to a two-sided control chart as follows. For $j = 1, 2, \ldots, s$, let

$$\pi_j^-(\theta) = F(-t_{j-1}, \theta) - F(-t_j, \theta), \qquad l_j^- = \log\left(\pi_j^-(-\theta_1)/\pi_j^-(\theta_0)\right), \qquad w_j^- = round\left(Q^- l_j^-\right),$$

and

$$T_n^- = \min\left(0, T_{n-1}^- - \sum_{j=1}^{s} m_j^-(n) l_j^-\right),$$

where $Q^- = 50/[\max(l_1^-, \ldots, l_s^-) - \min(l_1^-, \ldots, l_s^-)]$, and $m_j^-(n)$ denotes the number of observations that belong to the interval $(-t_j, -t_{j-1}]$ at the n-th time point, for $j = 1, 2, \ldots, s$. Then, the two-sided control chart gives a signal of shift if

$$T_n^+ > h_{St} \quad \text{or} \quad T_n^- < -h_{St},$$

where h_{St} is chosen to achieve a given IC ARL value. This chart is called St-CUSUM chart hereafter.

In cases when the IC process response distribution is normal, it is known that the CUSUM chart (6) when $h_N^+ = h_N^- = h_N$ and k_N chosen to be half of a given mean shift has certain optimality properties for detecting the given mean shift (Page 1954). This chart is also considered here as a golden standard when comparing different control charts, and is denoted as I-CUSUM where "I" represents the "ideal" CUSUM chart. When the IC process response distribution is $t(d)$, the charting statistics of the I-CUSUM (cf., Hawkins and Olwell 1998, Section 6.2.9) for detecting a mean shift from $\mu = 0$ to $\mu = \mu_1$ is

$$\begin{aligned} Z_n^+ &= \max\left(0, Z_{n-1}^+ + \frac{d+1}{2}\log\left(\frac{d + X_n^2}{d + (X_n - \mu_1)^2}\right)\right) \\ Z_n^- &= \min\left(0, Z_{n-1}^- - \frac{d+1}{2}\log\left(\frac{d + X_n^2}{d + (X_n + \mu_1)^2}\right)\right), \end{aligned}$$

and a mean shift is signaled if

$$Z_n^+ > h_I \text{ or } Z_n^- < -h_I,$$

where $h_I > 0$ is a control limit chosen to achieve a given IC ARL value.

3.2 Numerical comparison of the control charts

In this part, we compare the two control charts P-CUSUM and L-CUSUM that are both based on categorization of the process observations with six alternative charts K-CUSUM, N-CUSUM, R-EWMA, S-CUSUM, St-CUSUM, and I-CUSUM. In the comparison, the IC distribution is chosen to be the standardized version with mean 0 and standard deviation 1 of one of the following four distributions: $N(0,1)$, $t(4)$, $\chi^2(1)$ and $\chi^2(4)$. Among these distributions, $t(4)$ represents symmetric distributions with heavy tails, and $\chi^2(1)$ and $\chi^2(4)$ represent skewed distributions with different skewness. For the chart I-CUSUM, it is considered only in cases when the IC distribution is $N(0,1)$

or the standardized version of $t(4)$, because we do not know how to construct an appropriate I-CUSUM chart when the IC distribution is chi-square. It is also assumed the pre-specified IC ARL value is 500, and the batch size of Phase II observations at each time point is $m = 5$. In this comparison, we fix $p = 5$ in the P-CUSUM and L-CUSUM charts.

First, we compute the actual IC ARL values of the eight control charts, based on 10,000 replicated simulations, in cases with the four actual IC process distributions described above. For the charts P-CUSUM, L-CUSUM, K-CUSUM, N-CUSUM, and S-CUSUM, their control limits are determined based on 500 IC observations when their allowance constants are chosen to be 0.1. The control limit of the chart R-EWMA is chosen when $\lambda = 0.05$ and the IC distribution is assumed normal. For the chart St-CUSUM, it is assumed that the IC process response distribution is normal in all cases, as described in Section 3.1, and $\theta_1 = 0.6$. For the I-CUSUM chart, the IC distribution is assumed known, $k_N = 0.1$ when the IC distribution is $N(0, 1)$, and $\mu_1 = 0.2$ when the IC distribution is the standardized version of $t(4)$. The results are shown in Table 1. From the table, it can be seen that the actual IC ARL values of the charts P-CUSUM, L-CUSUM, and K-CUSUM are very close to 500 in all cases, as expected. The actual IC ARL values of the charts N-CUSUM and R-EWMA are quite different from 500 when the actual IC process distribution is not normal. For the chart S-CUSUM, its actual IC ARL values are moderately different from 500 due to the discreteness of its charting statistic, as shown in Chakraborti and Eryilmaz (2007). Because the chart St-CUSUM is constructed using the likelihood of normal distributions, its actual IC ARL value is close to 500 when the actual IC process distribution is $N(0, 1)$; but, its actual IC ARL values are very different from 500 in the other three cases when the actual IC distribution is not normal. The actual IC ARL values of the chart I-CUSUM are close to 500 in the two cases considered, as expected.

Next, we compare the OC performance of the related control charts in cases when the IC sample size $M = 200$ or 500. In order to make the comparison more meaningful, we intentionally adjust the parameters of the charts N-CUSUM, R-EWMA, and St-CUSUM so that their actual IC ARL values equal 500 in all cases considered. For the chart S-CUSUM, we use the same modification procedure as the one described in Section 2 for the P-CUSUM chart to overcome the difficulty caused by the discreteness of their charting statistics, and the actual IC ARL value of its modified version can reach 500 in all cases considered. In this study, 10 mean shifts ranging from -1.0 to 1.0

Table 1: The actual IC ARL values and their standard errors (in parentheses) of the eight control charts when the nominal IC ARL values are fixed at 500 and the actual IC process distribution is the standardized version of $N(0,1)$, $t(4)$, $\chi^2(1)$ and $\chi^2(4)$.

	$N(0,1)$	$t(4)$	$\chi^2(1)$	$\chi^2(4)$
P-CUSUM	501.9 (5.51)	503.3 (5.55)	504.8 (5.46)	501.1 (5.45)
L-CUSUM	498.1 (4.96)	495.3 (5.02)	504.2 (5.13)	497.4 (5.06)
K-CUSUM	499.0 (4.48)	499.7 (4.47)	504.4 (4.61)	496.4 (4.46)
N-CUSUM	498.9 (4.35)	156.0 (1.13)	321.4 (2.63)	371.5 (3.27)
R-EWMA	502.2 (6.24)	578.2 (20.87)	605.5 (8.65)	533.7 (6.57)
S-CUSUM	497.3 (5.22)	532.2 (5.57)	544.5 (5.88)	518.3 (5.63)
St-CUSUM	499.5 (4.87)	3037.7 (27.33)	9316.6 (20.46)	1862.2 (18.08)
I-CUSUM	499.8 (4.84)	499.8 (4.92)	–	–

with step 0.2 are considered, representing small, medium and large shifts. Due to the facts that different control charts have different parameters (e.g., k_P in the chart P-CUSUM, k_N in the chart N-CUSUM, and λ in the chart R-EWMA), and that the performance of different charts may not be comparable if their parameters are pre-specified, we use the following two approaches to set up their parameters. One is that all their parameters are chosen to be the optimal ones for detecting a given shift (e.g., the one of size 0.6), by minimizing the OC ARL values of the charts for detecting that shift while their IC ARL values are all fixed at 500, and the chosen parameters are used in all other cases as well. The second approach is to compare the optimal performance of all the charts when detecting each shift, by selecting their parameters to minimize the OC ARL values for detecting each individual shift, while their IC ARL values are all fixed at 500. The second approach for comparing different control charts has been used in the literature. See, for instance, Qiu and Hawkins (2001). Based on 10,000 replications, in the case when $M = 500$, the OC ARL values of the related control charts by the first approach when the procedure parameters are chosen to be the optimal ones for detecting the shift of 0.6 are shown in Figures 2 and 3. The results are shown in two figures, instead of one, to better demonstrate the difference among different control charts. Also, the OC ARL values of the P-CUSUM chart are plotted in both figures for convenience of comparing other charts to it. The corresponding results when $M = 200$ are shown in Figures 4 and 5. When reading the plots in these figures, readers are reminded that the scale on the y-axis is in natural logarithm, to better demonstrate the difference among different control charts when detecting relatively large shifts. The corresponding results when the procedure parameters are chosen by the second approach are included in a supplementary file, to save some space here.

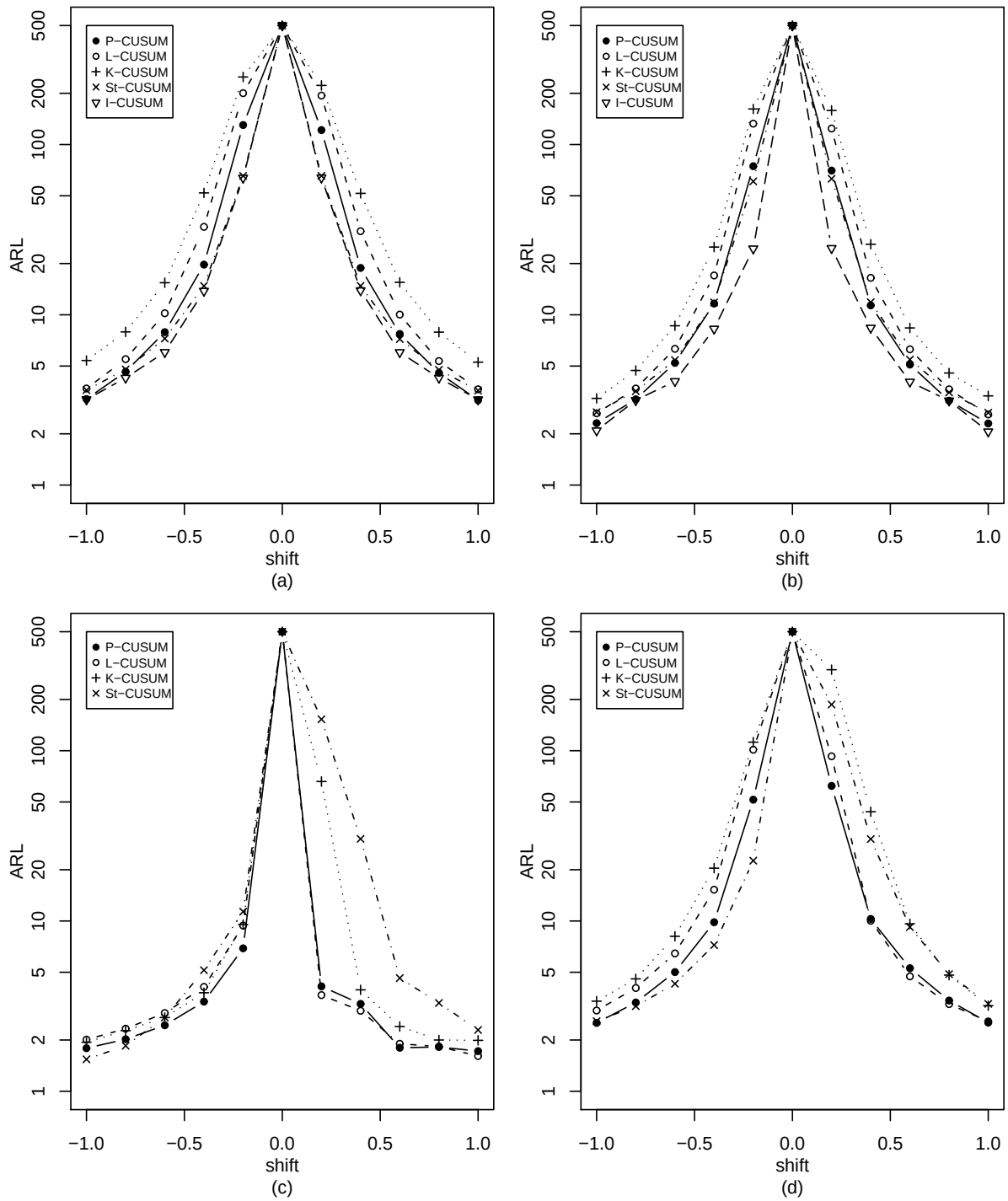

Figure 2: OC ARL values of five control charts when the IC ARL is 500, $M = 500$, $m = 5$, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the y-axis is in natural logarithm.

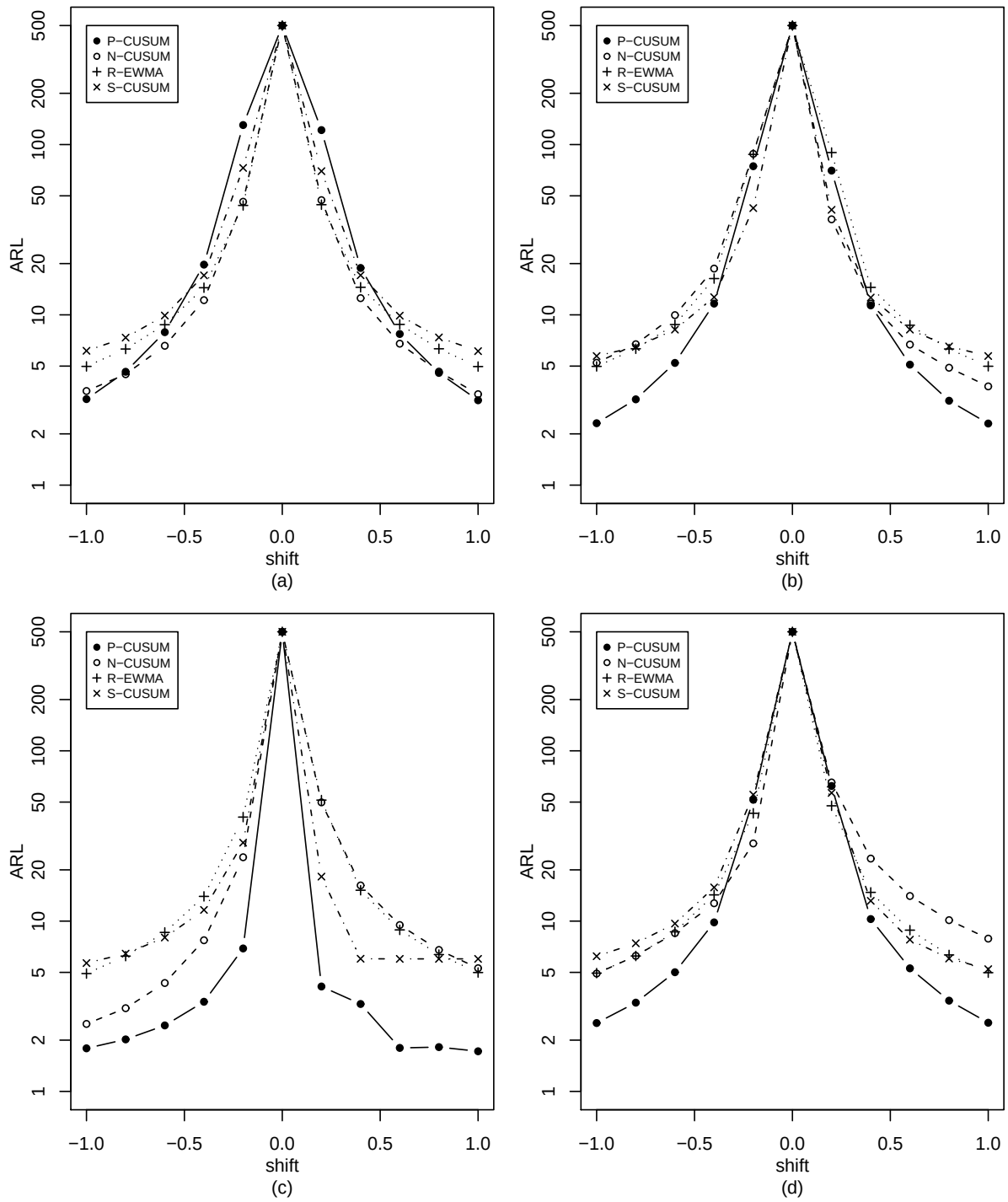

Figure 3: OC ARL values of four control charts when the IC ARL is 500, $M = 500$, $m = 5$, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the y-axis is in natural logarithm.

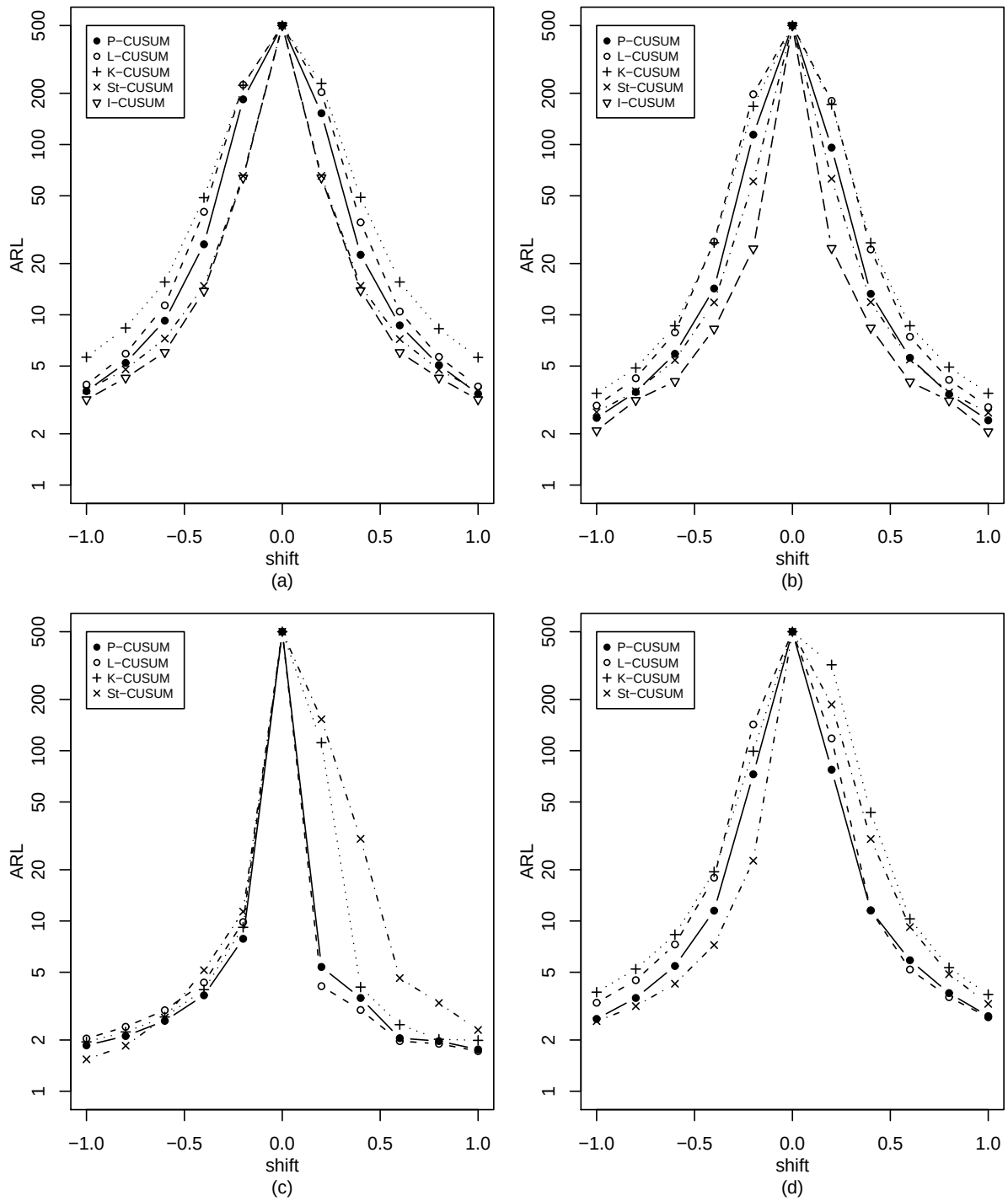

Figure 4: OC ARL values of five control charts when the IC ARL is 500, $M = 200$, $m = 5$, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the y-axis is in natural logarithm.

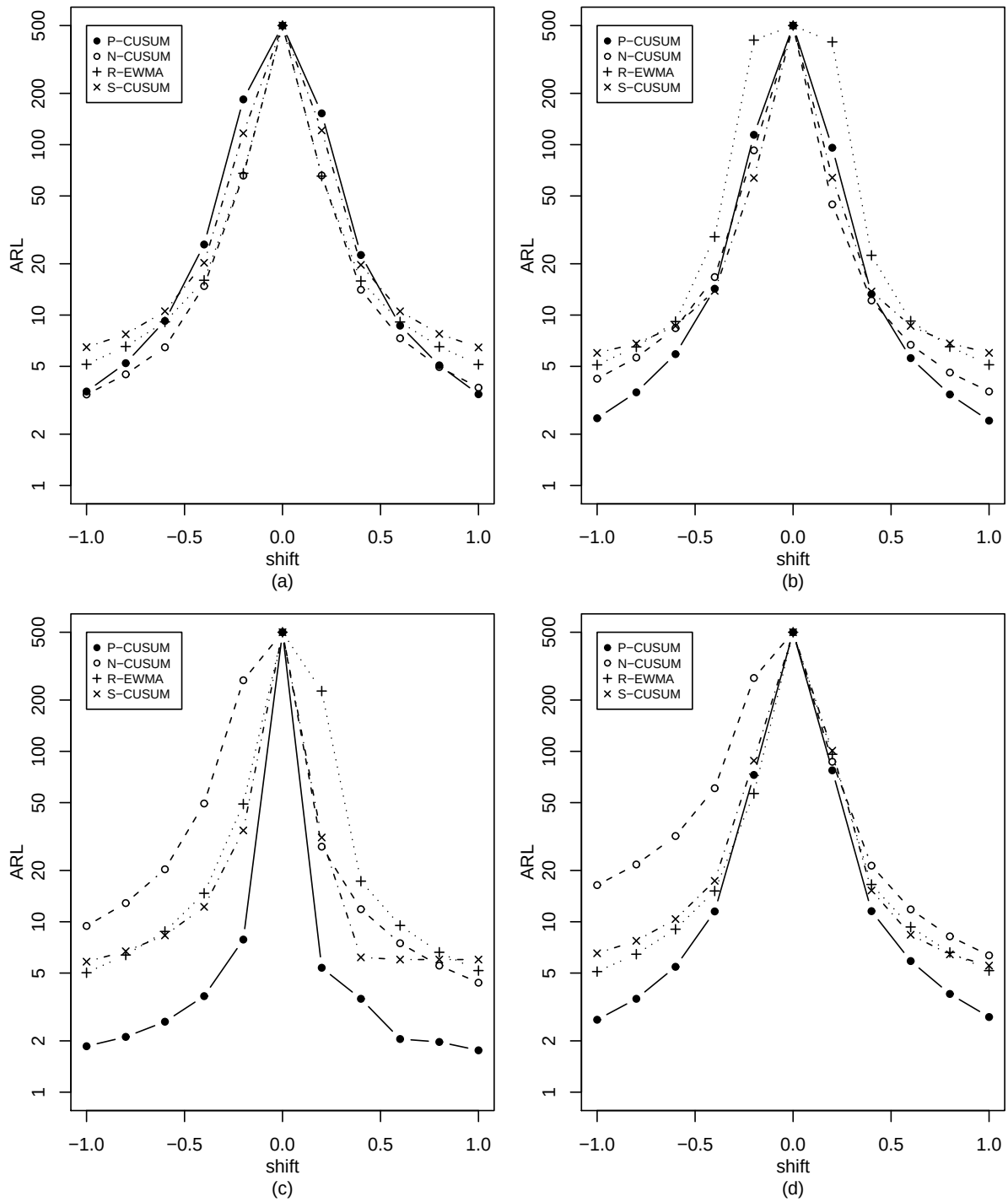

Figure 5: OC ARL values of four control charts when the IC ARL is 500, $M = 200$, $m = 5$, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Procedure parameters of the control charts are chosen to be the ones that minimize their OC ARL values when detecting the shift of 0.6. Scale on the y-axis is in natural logarithm.

From Figures 2(a) and 3(a), we can see that when the actual process distribution is normal, the I-CUSUM and N-CUSUM charts perform the best, as expected, the St-CUSUM chart is slightly worse, the P-CUSUM chart performs worse than the St-CUSUM chart when the shift is small and compatible or slightly better when the shift is large, the K-CUSUM and L-CUSUM charts is uniformly worse than the P-CUSUM chart, and the R-EWMA and S-CUSUM charts perform well for detecting small shifts, and become worse when detecting relatively large shifts. When reading the plots, readers are reminded that all the control charts are tuned optimal for detecting the shift of 0.6 in Figures 2 and 3. This example shows that, when the normality assumption is valid, the I-CUSUM and N-CUSUM charts are the best to use, because of the optimality of the conventional CUSUM chart (cf., Moustakides 1986, Ritov 1990). The St-CUSUM and P-CUSUM charts would lose certain power in detecting mean shifts because of grouping in the St-CUSUM chart and the categorization in the P-CUSUM chart; but the loss in power is small, which is especially true for the St-CUSUM chart. The R-EWMA and S-CUSUM charts perform well for detecting small shifts only. It seems that the K-CUSUM and L-CUSUM charts are not appropriate to use in this scenario.

In the scenario of Figures 2(b) and 3(b), the IC process distribution is $t(4)$, which is still symmetric but has heavier tails, compared to the normal distribution. In such cases, the I-CUSUM chart is still the best, as expected. Among other control charts, we can notice the following changes in the results, compared to the results obtained from Figures 2(a) and 2(b). (i) The P-CUSUM chart performs the best among all control charts other than the I-CUSUM chart in most cases of this scenario. (ii) The N-CUSUM chart is not among the best any more when the shift is large, although it still performs well when the shift is small.

From Figures 2(c), 2(d), 3(c) and 3(d), it can be seen that, when the actual IC distribution is skewed, the charts P-CUSUM and L-CUSUM perform reasonable well. The chart St-CUSUM performs well only in cases when detecting downward shifts. The K-CUSUM chart performs well when the skewness is relatively large. The other three charts do not perform well in such cases.

From Figures 4 and 5, we can see that the corresponding results when $M = 200$ are similar to those in Figures 2 and 3 when $M = 500$, although most OC ARL values presented in Figures 2 and 3 are slightly smaller than the corresponding ones presented in Figures 4 and 5, which is intuitively reasonable. The results presented in the supplementary file when all procedure parameters are

chosen to be the optimal ones for detecting each individual shift are overall better than those presented in Figures 2–5, in the sense that most OC ARL values in that scenario are smaller than the corresponding ones in Figures 2–5. However, the relative performance of the eight control charts is similar to that observed from Figures 2–5.

3.3 More numerical results about the P-CUSUM chart

From Section 3.2, we can see that the P-CUSUM chart performs well in most cases considered. In this part, we present more numerical results to further investigate its performance. First, we fix the value of its allowance constant k_P at 0.001, 0.005, 0.01, 0.05, or 0.1, the nominal IC ARL value (denoted as ARL_0) at 500, $M = 500$, and compute its OC ARL values based on 10,000 replications when the mean shift in the original process observations, the IC process distribution, and other parameters are in the same setup as those in Figures 2 and 3. The results are shown in Figure 6. From the four plots in the figure, it can be seen that, the OC ARL values when $k_P = 0.1$ are uniformly larger than those in cases when k_P is chosen smaller in all scenarios considered. Therefore, to use the P-CUSUM chart, it seems that k_P should not be chosen larger than 0.05. An intuitive explanation of this result is that k_P works as a threshold value of the statistic C_n in the P-CUSUM chart, and C_n measures the relative difference between certain observed counts and the corresponding expected counts, which usually takes small values (cf., the discussion given in the paragraph immediately after the one containing expression (3) in Section 2). Consequently, a good value for k_P should also be relatively small. If we check the plots carefully, then we can also find that results when $k_P = 0.01$ are good only when the shift is large. When the shift is small, k_P should be a smaller number.

As discussed in Section 2, selection of the control limit h_P depends only on the number of categories p used in categorizing the Phase II data, the allowance constant k_P, the batch size m of the Phase II data, and the nominal IC ARL value ARL_0. Next, in the cases when $ARL_0 = 200$, 300, 500, and 1000, $p = 2$, 3, 5, 10, 15, and 20, $k_P = 0.001$, 0.005, 0.01, and 0.05, $m = 1$, and 5, and $M = 500$, we compute the h_P values based on 10,000 replications, and the results are presented in Table 2. In the cases considered, the actual IC ARL values are also recorded, and they are all within one standard error of ARL_0, as in the example of Table 1. Therefore, to save some space,

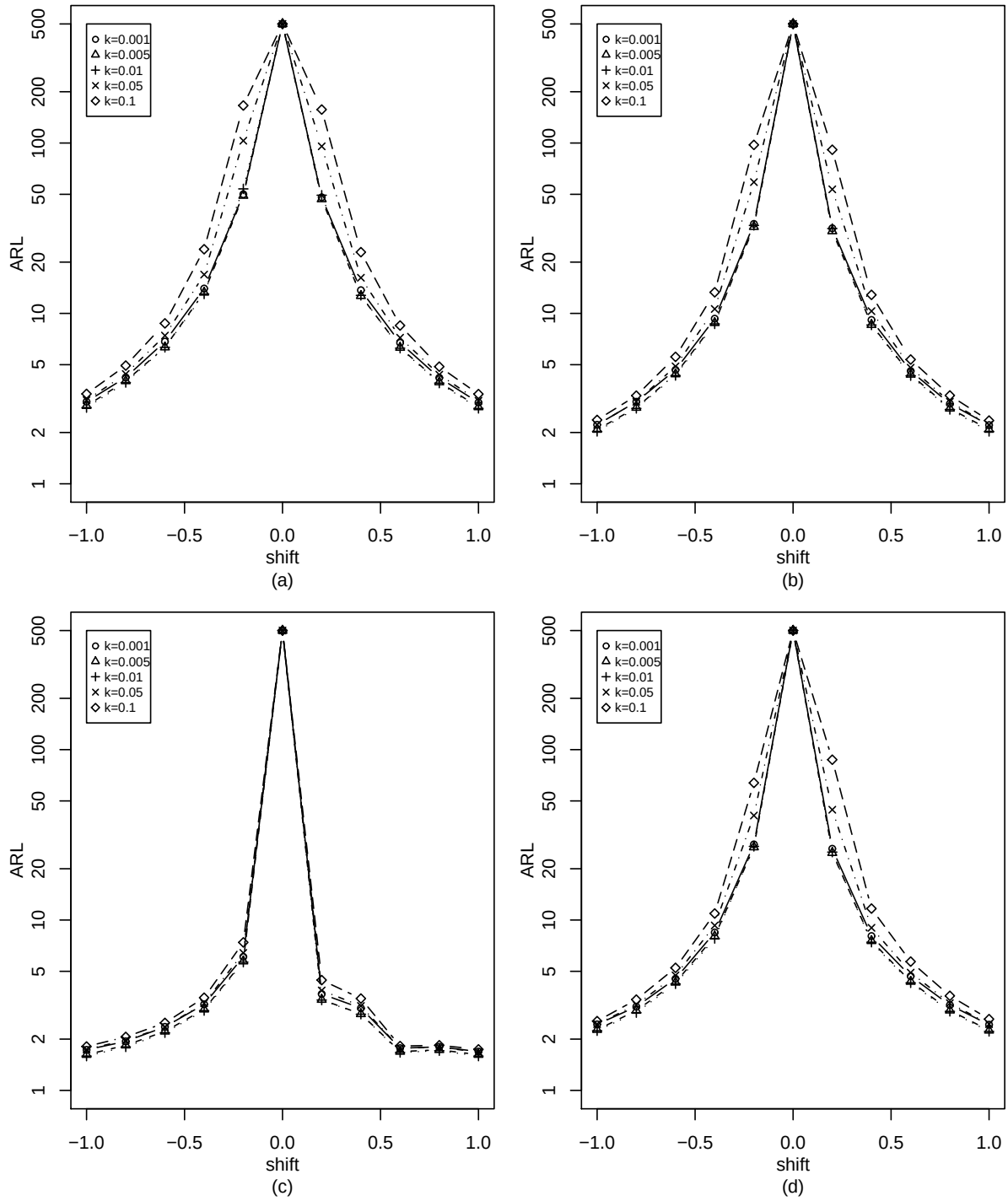

Figure 6: OC ARL values of the P-CUSUM chart when $k_P = 0.001$, 0.005, 0.01, 0.05, or 0.1, the IC ARL is 500, $M = 500$, $m = 5$, $p = 5$, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Scale on the y-axis is in natural logarithm.

they are omitted here. From the table, it can be seen that h_P increases with ARL_0, k_P, and p, and decreases with m.

To use the P-CUSUM chart, an IC dataset of size M is required to estimate the IC quantiles $\{q_l, l = 1, 2, \ldots, p-1\}$ such that categorization of the Phase II data can be proceeded (cf., expression (1)). Therefore, performance of the P-CUSUM chart may depend on the IC sample size M. To study this potential dependence, we compute the optimal OC ARL values of the P-CUSUM chart when M changes its values among 100, 200, 500, and 1000, and the optimal OC ARL values are presented in Figure 7. From the figure, it can be seen that (i) the P-CUSUM chart performs better when M is larger, (ii) its OC performance is already quite stable when $M \geq 100$, and (iii) its optimal OC ARL values are stable when $M \geq 200$, and almost identical when $M \geq 500$ in most cases considered. Intuitively, the value of M may also have an impact on the variability of the run length distribution of the P-CUSUM chart. To investigate this issue, we compute the sample standard deviation of the 10,000 run length values (denoted as SDRL) obtained from the 10,000 replications in each case considered, and the results are presented in Figure 8. From the figure, it can be seen that (i) the SDRL value tends to be smaller when M increases, and (ii) its values are stable when $M \geq 200$ and are almost identical when $M \geq 500$ in most cases considered.

From the construction of the P-CUSUM chart discussed in Section 2, it can be seen that the number of categories p should be chosen properly beforehand, which is studied in the following examples. We first assume that the actual IC distribution is the standardized version with mean 0 and standard deviation 1 of the $t(4)$ distribution, $M = 500$, and the shift ranges from -1.0 to 1.0 with step 0.2, as in Figure 2(b). The optimal OC ARL values of the P-CUSUM chart based on 10,000 replications are presented in the four plots of Figure 9, respectively, when the batch size m changes its value among 1, 5, 10, and 20. In all four plots, we let p change among 2, 3, 5, 10, and 15. From Figure 10, it can be seen that (i) when $m = 1$ (i.e., for single-observation data), it seems that p should be chosen relatively large, although the performance of the P-CUSUM chart does not change much when $p \geq 5$, and (ii) when m gets larger, it seems that p can be chosen smaller. For instance, when $m = 20$ (cf., Figure 9(d)), it seems that $p = 2$ is already good enough.

Next, we repeat the above simulation, after changing the actual IC distribution to the standardized version of $\chi^2(1)$. The corresponding results are presented in Figure 10. Similar conclusions

Table 2: Computed h_P values based on 10,000 replications when the nominal IC ARL value $ARL_0 =$ 200, 300, 500, and 1000, $p = 2$, 3, 5, 10, 15, and 20, $k_P = 0.001$, 0.005, 0.01, and 0.05, $m = 1$, and 5, and $M = 500$.

$m = 1$							
k_P	ARL_0	$p = 2$	3	5	10	15	20
0.001	200	4.144	4.429	5.654	10.783	15.777	20.883
	300	4.898	4.614	5.887	11.015	16.230	21.422
	500	5.445	4.833	6.265	11.305	16.570	21.827
	1000	6.402	5.552	6.843	11.960	16.979	22.155
0.005	200	5.089	5.148	6.215	11.180	16.142	21.297
	300	5.848	5.481	6.650	11.531	16.407	21.589
	500	6.497	6.099	7.251	12.006	16.992	22.106
	1000	7.735	7.146	8.056	12.758	17.655	22.677
0.01	200	5.529	5.555	6.665	11.377	16.400	21.448
	300	6.123	6.061	7.209	11.842	16.739	21.856
	500	6.954	6.899	7.929	12.343	17.307	22.312
	1000	8.078	7.979	8.600	13.205	18.056	22.993
0.05	200	5.918	6.860	7.957	12.171	16.981	21.931
	300	6.656	7.685	8.438	12.841	17.550	22.397
	500	7.595	8.492	9.269	13.616	18.181	22.958
	1000	8.559	9.773	10.614	14.461	19.149	23.986
$m = 5$							
k_P	ARL_0	$p = 2$	3	5	10	15	20
0.001	200	0.899	0.983	1.318	2.430	3.588	4.753
	300	1.080	1.065	1.379	2.496	3.647	4.814
	500	1.265	1.183	1.485	2.598	3.734	4.916
	1000	1.469	1.384	1.636	2.718	3.879	5.054
0.005	200	1.214	1.245	1.512	2.575	3.732	4.912
	300	1.326	1.375	1.618	2.678	3.824	4.987
	500	1.522	1.556	1.759	2.799	3.948	5.115
	1000	1.778	1.815	1.966	2.981	4.119	5.254
0.01	200	1.276	1.389	1.635	2.686	3.826	5.024
	300	1.412	1.530	1.750	2.777	3.909	5.084
	500	1.600	1.725	1.911	2.911	4.032	5.214
	1000	1.889	2.007	2.152	3.100	4.228	5.396
0.05	200	1.348	1.704	1.994	2.927	4.071	5.255
	300	1.488	1.898	2.149	3.047	4.187	5.365
	500	1.718	2.040	2.362	3.220	4.355	5.528
	1000	1.924	2.272	2.628	3.459	4.577	5.745

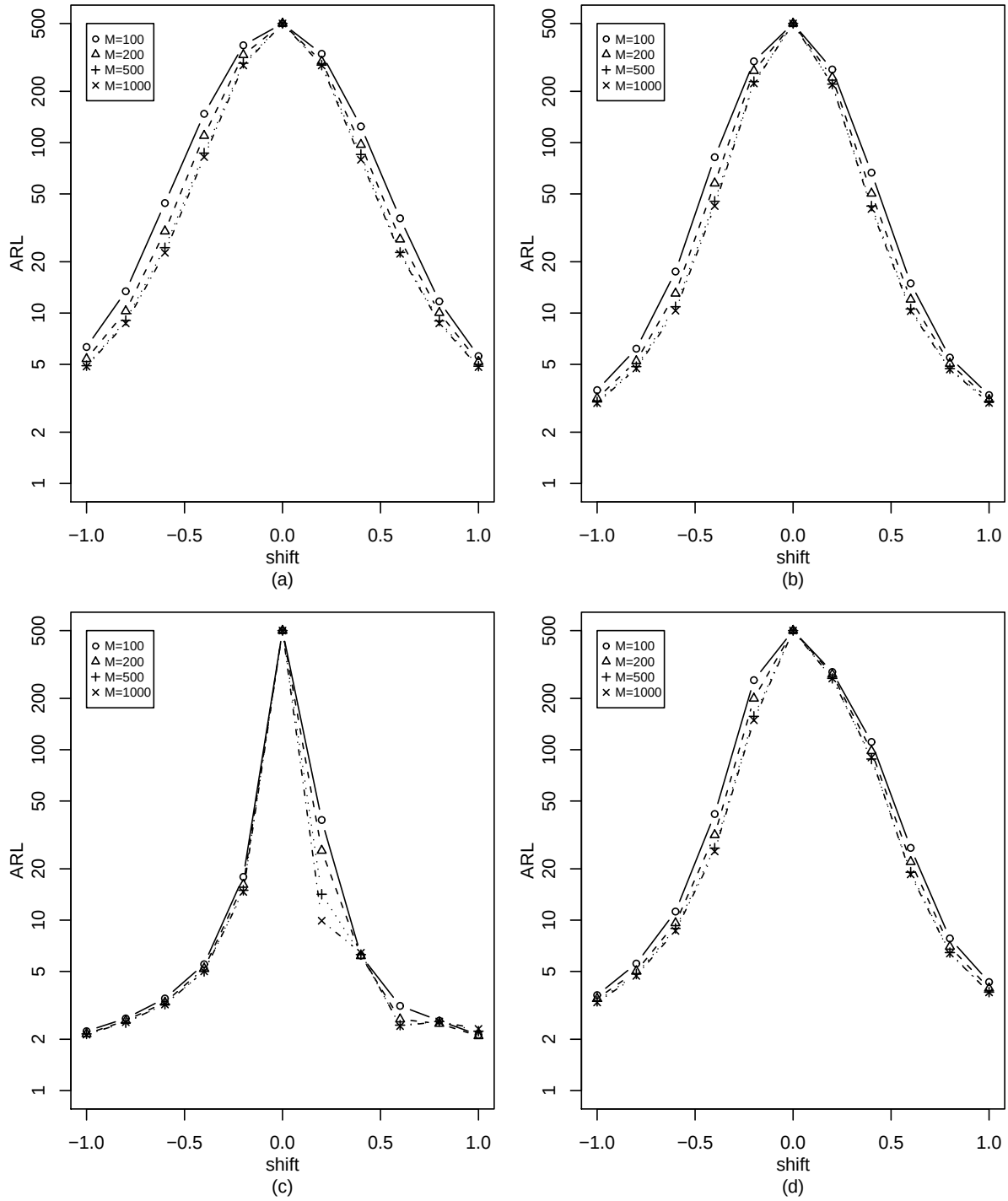

Figure 7: Optimal OC ARL values of the P-CUSUM chart when the IC ARL is 500, $m = 5$, the IC sample size $M = 100$, 200, 500, and 1000, and the actual IC process distribution is the standardized version of $N(0, 1)$ (plot (a)), $t(4)$ (plot (b)), $\chi^2(1)$ (plot (c)), and $\chi^2(4)$ (plot (d)). Scale on the y-axis is in natural logarithm.

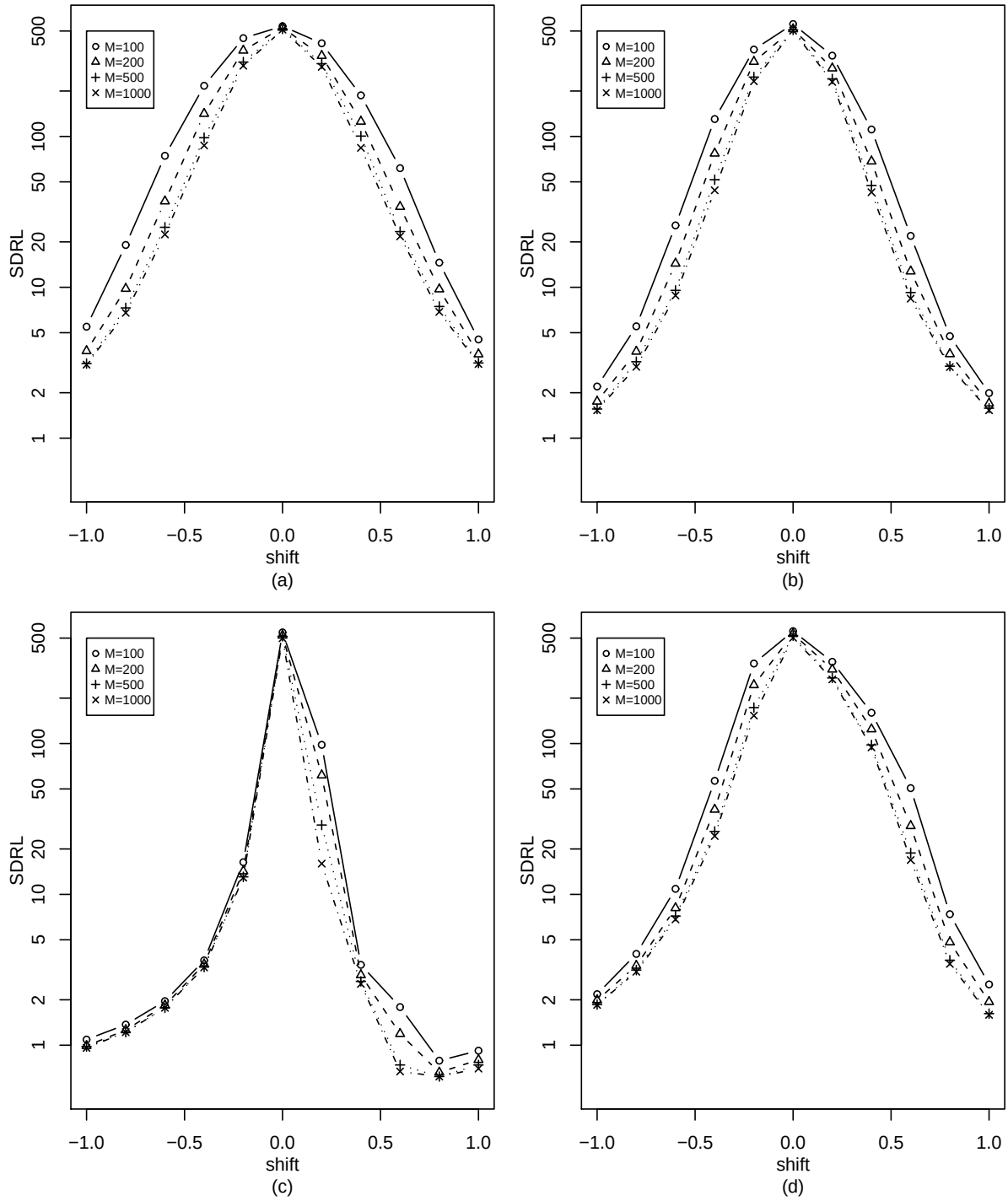

Figure 8: Sample standard deviation values of the run length distribution of the P-CUSUM chart, computed from 10,000 replicated simulations, in cases considered in Figure 7. Scale on the y-axis is in natural logarithm.

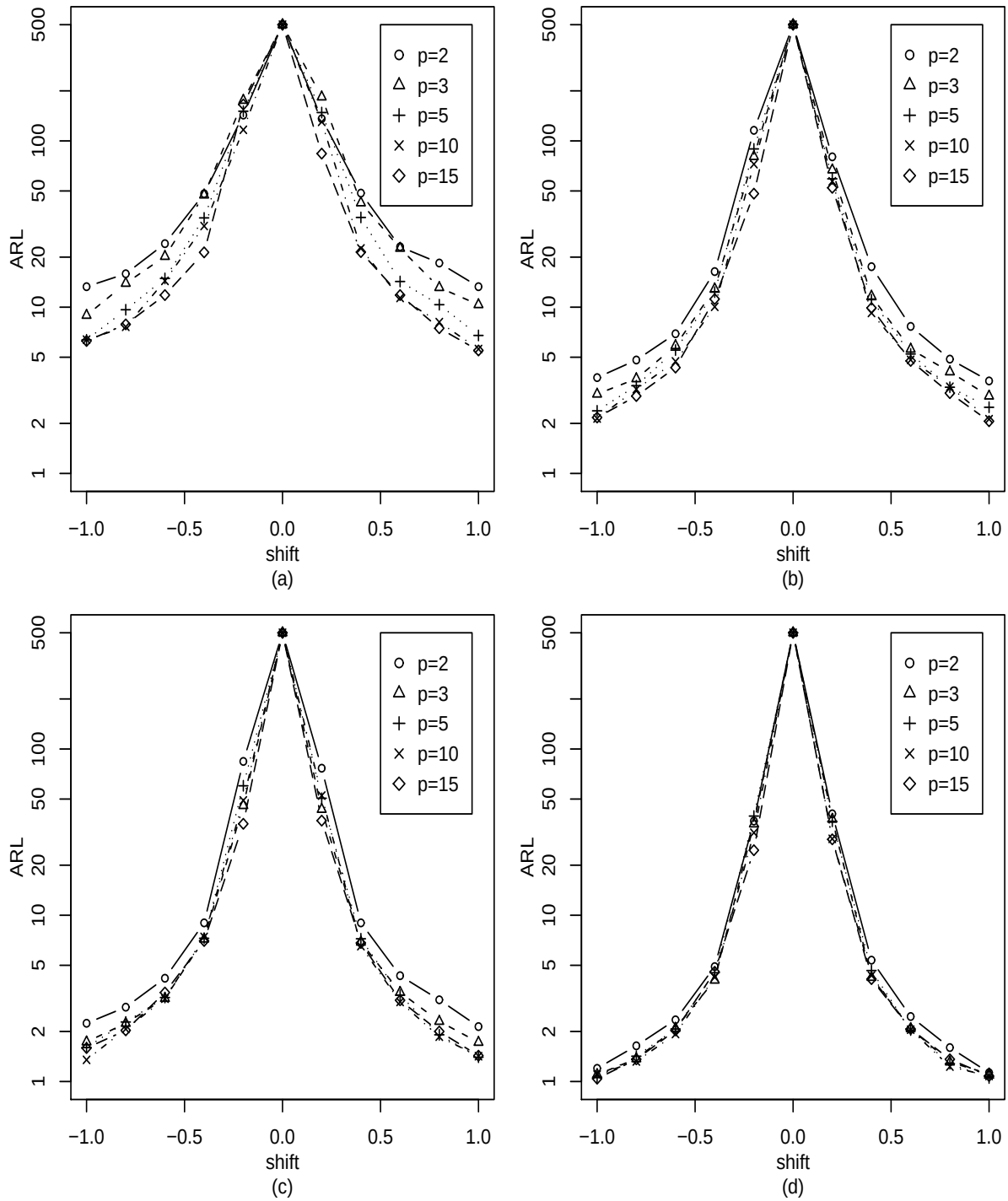

Figure 9: Optimal OC ARL values of the P-CUSUM chart when the actual IC distribution is the standardized version of $t(4)$, $M = 500$, $p = 2, 3, 5, 10$ and 15, and the batch size is $m = 1$ (plot (a)), $m = 5$ (plot (b)), $m = 10$ (plot (c)), and $m = 20$ (plot (d)). Scale on the y-axis is in natural logarithm.

to those in the previous example can be made here for positive shifts. When the shift is negative and relatively small, it seems that p should be at least 10 to have reasonably good results. When the shift is negative and large, p can be chosen smaller. We also performed simulations in cases when the actual IC distribution is $N(0,1)$ or the standardized version of $\chi^2(4)$. Results in these two cases are similar to those in Figures 9 and 10, respectively. So, they are omitted here, and are available from the authors upon request.

Based on the results shown in Figures 9 and 10, we provide the following guidelines for selecting p. (i) p can be chosen smaller when m is larger. (ii) In cases when we do not have any prior information about the process distribution, then we can choose $p = 10$. In such cases, the P-CUSUM chart should perform reasonably well. (iii) If we know that the process distribution is quite symmetric, or that it is skewed but the potential shift is in the direction of the longer tail, then p can be chosen as small as 5. The above guideline (i) can be explained intuitively as follows. When the batch size m is larger, the observed count vector $\mathbf{g}(n)$ (cf., expression (2)) carries more information about a potential distributional shift in the categorized data $\mathbf{Y}_j(n)$, allowing p to be smaller without losing much efficiency of the P-CUSUM chart. More specifically, assume that the process is OC at the n-th time point. Then, it can be checked that

$$E_p\left(\widetilde{X}^2(n)\right) = \sum_{l=1}^{p} \frac{f_l^{(1)}(1-f_l^{(1)}) + m(f_l^{(1)} - f_l^{(0)})^2}{f_l^{(0)}}, \tag{8}$$

where E_p denotes the expectation when p categories are used in the P-CUSUM chart, $\widetilde{X}^2(n)$ is the Pearson's chi-square statistic on which the P-CUSUM chart is based (cf., its definition in Section 2), and $\mathbf{f}^{(0)} = (f_1^{(0)}, f_2^{(0)}, \ldots, f_p^{(0)})'$ and $\mathbf{f}^{(1)} = (f_1^{(1)}, f_2^{(1)}, \ldots, f_p^{(1)})'$ are the IC and OC distributions of $\mathbf{Y}_j(n)$. From (8), when m is larger, there is indeed more information in $\widetilde{X}^2(n)$ about the mean shift. In the case when the IC distribution is $N(0,1)$ and the mean shift is 0.4, $\mathbf{f}^{(0)} = (0.5, 0.5)'$ and $\mathbf{f}^{(1)} = (0.345, 0.655)'$ when $p = 2$, and $\mathbf{f}^{(0)} = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)'$ and $\mathbf{f}^{(1)} = (0.046, 0.061, 0.071, 0.079, 0.088, 0.097, 0.108, 0.121, 0.140, 0.189)'$ when $p = 10$. In such cases, when m changes among 1, 5, 10, and 20, $E_2(\widetilde{X}^2(n))$, $E_{10}(\widetilde{X}^2(n))$, and their relative difference $RD = (E_{10}(\widetilde{X}^2(n)) - E_2(\widetilde{X}^2(n)))/E_2(\widetilde{X}^2(n)))$ are listed in Table 3. From the table, it can be seen that, when m increases, the relative difference between $E_2(\widetilde{X}^2(n))$ and $E_{10}(\widetilde{X}^2(n))$ gets smaller. This explains why the OC ARL values of the P-CUSUM chart when $p = 2$ get closer to those when $p = 10$ when m increases, as seen in Figures 9 and 10. The guidelines (ii) and (iii) imply that

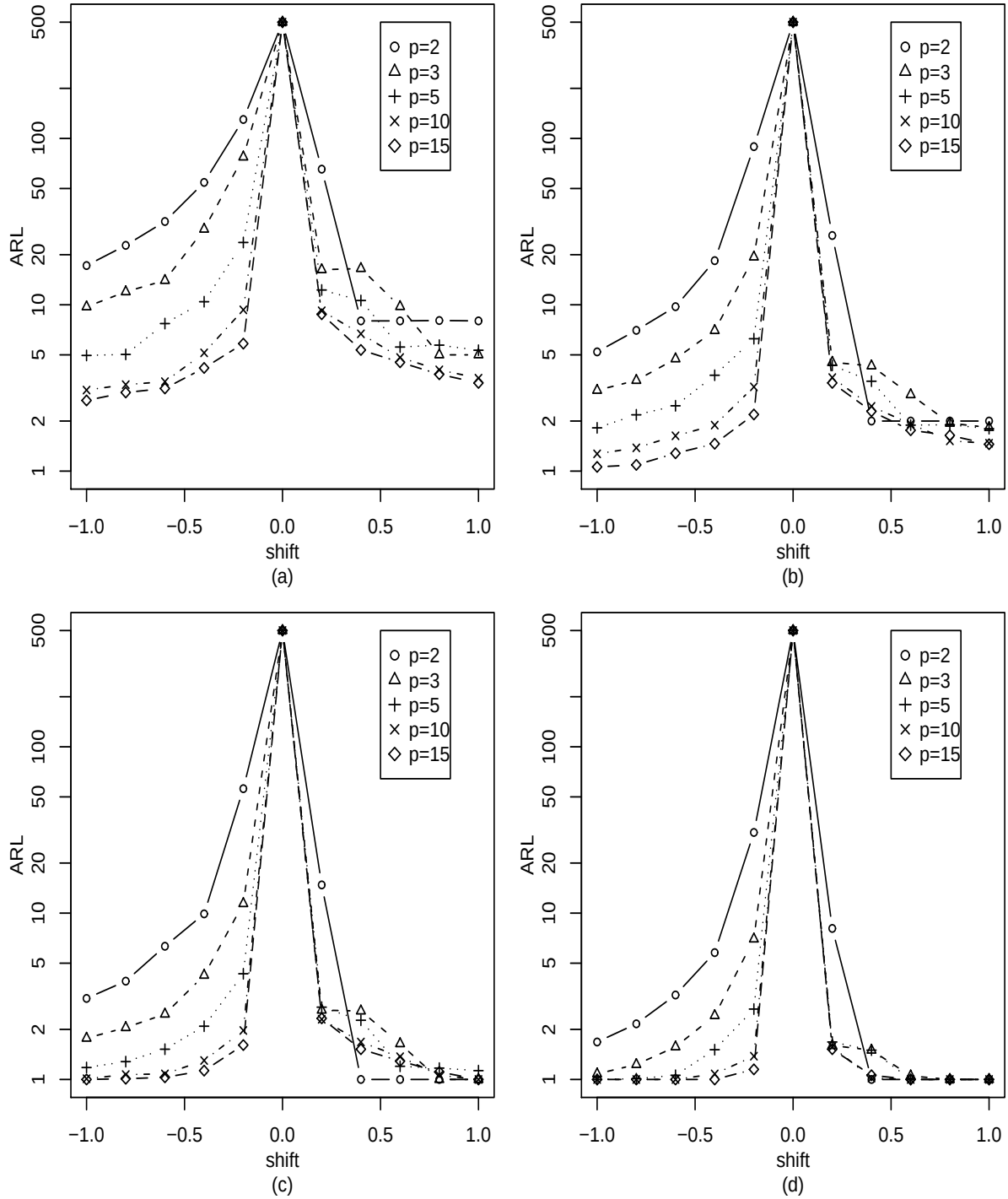

Figure 10: Optimal OC ARL values of the P-CUSUM chart when the actual IC distribution is the standardized version of $\chi^2(1)$, $M = 500$, $p = 2, 3, 5, 10$ and 15, and the batch size is $m = 1$ (plot (a)), $m = 5$ (plot (b)), $m = 10$ (plot (c)), and $m = 20$ (plot (d)). Scale on the y-axis is in natural logarithm.

Table 3: Values of $E_2(\widetilde{X}^2(n))$, $E_{10}(\widetilde{X}^2(n))$, and $RD = (E_{10}(\widetilde{X}^2(n)) - E_2(\widetilde{X}^2(n)))/E_2(\widetilde{X}^2(n)))$ in cases when the IC distribution is $N(0,1)$, the mean shift is 0.4, $p = 2$ or 10, and m changes among 1, 5, 10, and 20.

m	$E_2(\widetilde{X}^2(n))$	$E_{10}(\widetilde{X}^2(n))$	RD
1	1.000	9.000	8.000
5	1.386	9.638	5.954
10	1.870	10.434	4.580
20	2.836	12.028	3.241

selection of p also depends on the shape of the process distribution, which can be demonstrated by the calculation shown in Table 4. In the table, we consider two IC distributions: $N(0,1)$ and the standardized version of $\chi^2(1)$, as in Figures 9 and 10. We also assume that the mean shift in the original process observations $\mathbf{X}(n)$ is 0.2, and p is 2 or 5. The table presents the IC distribution $\mathbf{f}^{(0)}$, the OC distribution $\mathbf{f}^{(1)}$, and the relative distance between $\mathbf{f}^{(0)}$ and $\mathbf{f}^{(1)}$, defined by $Q = \sum_{l=1}^{p}(f_l^{(1)} - f_l^{(0)})^2/f_l^{(0)}$, of the categorized data $\mathbf{Y}_j(n)$. It can be seen from the table that, when the IC distribution is $N(0,1)$ which is symmetric, Q changes its value from 0.025 to 0.036 when p increases from 2 to 5. However, when the IC distribution is the standardized version of $\chi^2(1)$ which is skewed, Q changes its value from 0.127 to 0.843. The increase in the Q value is about 7 times in the latter case, which is much more dramatic compared to the increase in the former case. Therefore, the effect of p on the P-CUSUM chart does depend on the shape of the IC process distribution. By similar arguments, we can also explain why the P-CUSUM chart with $p = 2$ performs better than the chart with $p = 3$ in some cases shown in Figure 10. For instance, in the case when the IC process distribution is the standardized version of $\chi^2(1)$ (cf., Figure 10(c)), $p = 2$, and the mean shift in the original process observations $\mathbf{X}(n)$ is 0.4, the IC distribution $\mathbf{f}^{(0)}$ and the OC distribution $\mathbf{f}^{(1)}$ of the categorized data $\mathbf{Y}_j(n)$ are $(0.5, 0.5)'$ and $(0,1)'$, respectively, and the resulting Q value is 1.0. After p increases from 2 to 3 and other parameters remain unchanged, $\mathbf{f}^{(0)}$ and $\mathbf{f}^{(1)}$ become $(1/3, 1/3, 1/3)'$ and $(0, 0.457, 0.543)'$. The resulting Q value is 0.511 in this case, which is much smaller than the Q value when $p = 2$. Therefore, the shift would be relatively easier to detect by the P-CUSUM chart with $p = 2$. By the way, the above justifications with Tables 3 and 4 are all based on theoretical results. In applications, observed values of $\widetilde{X}^2(n)$ would be different from its mean values presented in Table 3. Similarly, although $\mathbf{f}^{(0)}$ presented in Table 4 have equal probabilities for each group, its estimator $\widehat{\mathbf{f}}^{(0)}$ obtained from an IC data would not have

Table 4: The IC distribution $\mathbf{f}^{(0)}$, the OC distribution $\mathbf{f}^{(1)}$, and the relative distance between them, defined by $Q = \sum_{l=1}^{p}(f_l^{(1)} - f_l^{(0)})^2/f_l^{(0)}$, of the categorized data $\mathbf{Y}_j(n)$ when the IC distribution is $N(0,1)$ or the standardized version of $\chi^2(1)$, and p is 2 or 5.

IC Distribution	p	$\mathbf{f}^{(0)}$	$\mathbf{f}^{(1)}$	Q
$N(0,1)$	2	(0.5,0.5)'	(0.421,0.579)'	0.025
	5	(0.2,0.2,0.2,0.2,0.2)'	(0.149,0.716,0.196,0.218,0.261)'	0.036
$\chi^2(1)$	2	(0.5,0.5)'	(0.322,0.678)'	0.127
	5	(0.2,0.2,0.2,0.2,0.2)'	(0,0,0.486,0.271,0.243)'	0.843

this property due to estimation error.

From the construction of the P-CUSUM chart, we can see that it should be able to detect any distributional shift of the original observations $\mathbf{X}(n)$ that results in a shift in the distribution of the corresponding categorized data $\mathbf{Y}_j(n)$. Therefore, the P-CUSUM chart should have certain ability to detect shifts in dispersion, or in both location and dispersion, of the distribution of the original observations $\mathbf{X}(n)$. To demonstrate this, next we consider the cases when the IC process distribution is the standardized version with mean 0 and standard deviation 1 of $\chi^2(1)$ or $\chi^2(4)$, and the OC distribution has a (location, dispersion) shift from $(0,1)$ to (μ_1, σ_1^2), where (μ_1, σ_1^2) takes the value of $(0,1.5), (0,2.0), (0.5,1.5)$, or $(0.5,2.0)$. When $M = 500$, $m = 5$, and $p = 5$, the computed optimal OC ARL values of the P-CUSUM chart are presented in Table 5. From the table, it can be seen that the P-CUSUM chart does have certain ability to detect dispersion shifts, and its computed optimal OC ARL values are much smaller in cases when both the location and dispersion are shifted, compared to cases when only the dispersion is shifted, which is reasonable because the resulting shifts in the distribution of the categorized data $\mathbf{Y}_j(n)$ are bigger in the former cases. For a similar reason, it can also be seen that the P-CUSUM chart is more sensitive to dispersion shifts when the IC process distribution is more skewed. However, based on the P-CUSUM chart alone, we cannot distinguish location shifts from dispersion shifts, after a signal of shift is given. It is left for our future research to design a control chart based on the categorized data that is powerful in detecting both location and dispersion shifts and is able to make a distinction between the two types of shifts.

Table 5: Computed optimal OC ARL values of the P-CUSUM chart when the IC observation distribution is the standardized version with mean 0 and standard deviation 1 of $\chi^2(1)$ or $\chi^2(4)$, and the OC distribution has a (location, dispersion) shift to (μ_1, σ_1^2).

IC Distribution	(μ_1, σ_1^2)=(0,1.5)	(0,2.0)	(0.5,1.5)	(0.5,2.0)
$\chi^2(1)$	51.90	17.30	8.13	15.87
$\chi^2(4)$	279.24	117.69	86.96	79.94

4 An Application

We illustrate the proposed methods using a real-data example about daily exchange rates between Korean Won and US currency Dollar between March 28, 1997 and December 02, 1997. During this period of time, the daily exchange rates were quite stable early on and became unstable starting from early August. This can be seen from Figure 11(a) in which 162 daily exchange rates (Won/Dollar) observed in that period are shown. As a side note, the world financial market experienced a serious crisis in the winter of 2007, and Korean Won was seriously affected by the crisis.

Like many other Phase II SPC procedures, our proposed procedure assumes that observations at different time points are independent of each other. However, for the exchange rate data, we found that observations at different time points are substantially correlated. Following the suggestions in Qiu and Hawkins (2001), we first pre-whiten the data using the following auto-regression model that can be accomplished by the R function ar.yw():

$$\begin{aligned} x(i) - 919.01 \;=\; & 0.86(x(i-1) - 919.01) + 0.05(x(i-2) - 919.01) - 0.09(x(i-3) - 919.01) + \\ & 0.12(x(i-4) - 919.01) - 0.06(x(i-5) - 919.01) + 0.02(x(i-6) - 919.01) + \\ & 0.25(x(i-7) - 919.01) - 0.40(x(i-8) - 919.01) + 0.16(x(i-9) - 919.01) + \epsilon(i), \\ & \text{for } i = 10, 11, \ldots, 162, \end{aligned}$$

where $x(i)$ denotes the i-th observation of the exchange rate, $\epsilon(i)$ is a white noise process with zero mean, and the order of the model is determined by the default Akaike's Information Criterion (AIC) in ar.yw(). The pre-whitened data are shown in Figure 11(b).

We then try to apply the related control charts considered in the previous section to the pre-whitened data. To this end, the first 96 residuals are used as an IC data, which correspond to the first 105 original observations, and the remaining residuals are used for testing. Based on

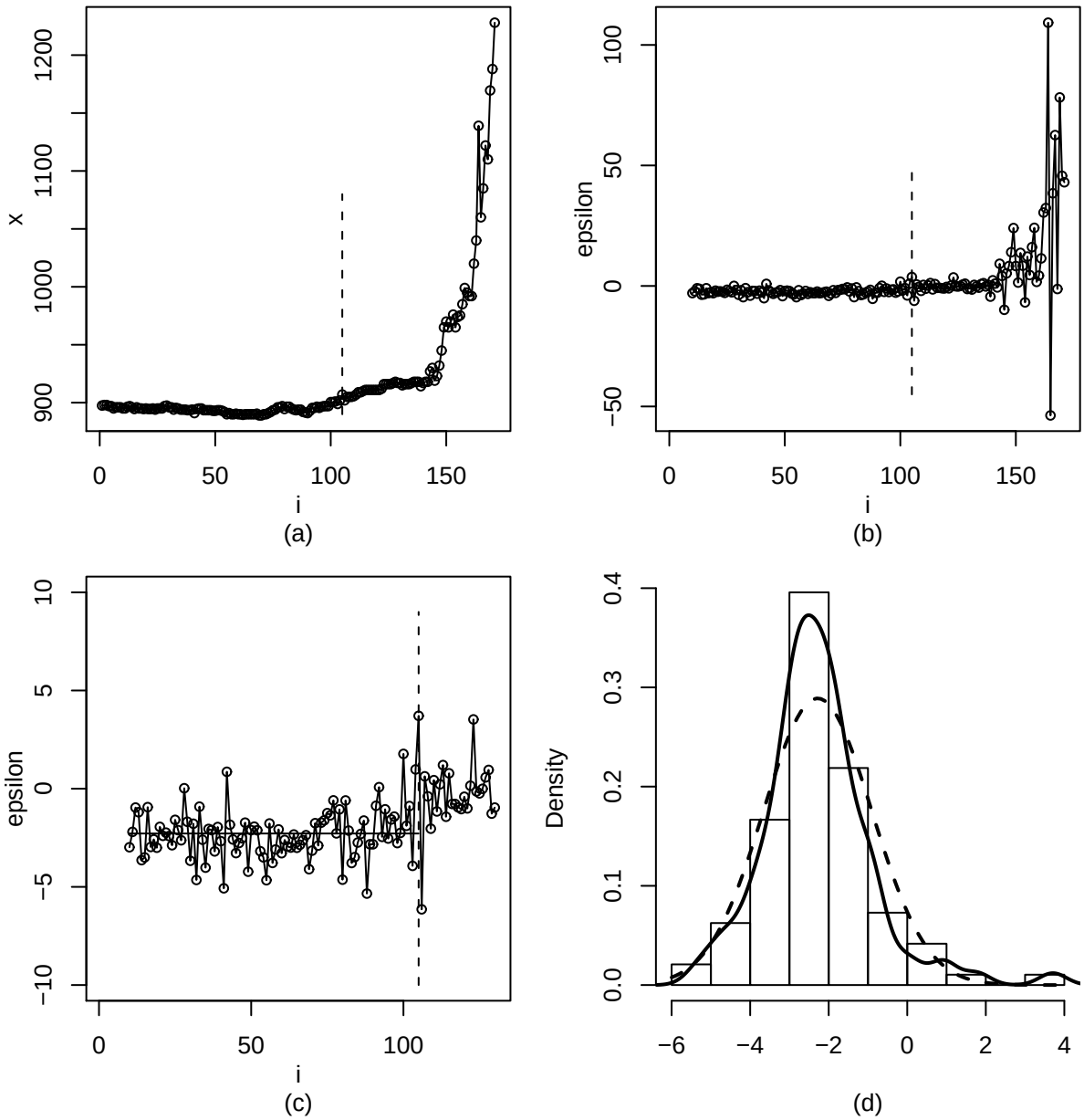

Figure 11: (a) Original observations of the exchange rates between Korean currency Won and US currency Dollar between March 28, 1997 and December 02, 1997. (b) Pre-whitened values of the original observations. (c) The first 121 pre-whitened values. (d) Density histogram, estimated density curve (solid) of the first 96 pre-whitened values (i.e., IC data), and the density curve of a normal distribution (dashed) with the same mean and variance as those of the IC data. In plots (a)–(c), the dashed vertical line separates the IC and testing data. In plot (c), the solid horizontal line denotes the sample mean of the IC data.

the simulation results shown in Figure 7, the OC performance of the P-CUSUM chart should be quite stable with this amount of IC data. In Figure 11(a)-(b), the training and testing data are separated by a dashed vertical line. To take a closer look at the IC data and at the first several testing observations as well, the first 121 residuals are presented in Figure 11(c) again, in which the solid horizontal line denotes the sample mean of the IC data and the dashed vertical line separates the IC and testing data. From plot (c), it can be seen that there is an upward mean shift starting from the very beginning of the test data.

The Shapiro test for checking the normality of the IC data gives a p-value of 1.323×10^{-4}, which implies that the IC data are significantly non-normal. To demonstrate this, the density histogram of the IC data is shown in Figure 11(d), along with its estimated density curve (solid) and the density curve of a normal distribution (dashed) with the same mean and standard deviation. We then consider to apply the related control charts to this dataset. Because the S-CUSUM chart can only be used for batch data and the current exchange rate data have a single observation at each time point, it is not considered here. The I-CUSUM and St-CUSUM charts are not considered here either because they require a parametric form of the IC distribution which is unavailable in this example. Among the remaining 5 charts, based on the simulation results shown in Table 1, the three nonparametric charts P-CUSUM, L-CUSUM, and K-CUSUM might be more appropriate to use in this case, compared to the charts N-CUSUM and R-EWMA. Because the IC process distribution is quite symmetric, as shown by Figure 11(d), we choose $p = 5$ in the P-CUSUM chart, by the practical guidelines given in Section 3.3. When the nominal IC ARL value is fixed at 200, and the allowance constants of the four CUSUM charts are all chosen to be 0.1 (remember that $\lambda = 0.05$ in the R-EWMA chart), the five control charts are shown in Figure 12, in which the dashed horizontal lines denote the control limits of the related control charts. From the five plots in the figure, it can be seen that the P-CUSUM, L-CUSUM, K-CUSUM, N-CUSUM, and R-EWMA charts give signals of shift at $i = 111$, 112, 115, 115, and 113, respectively. Therefore, the P-CUSUM chart gives an earliest signal in this example. By these results, the control charts confirm that the exchange rates between Korean currency Won and US currency Dollar started to become unstable at the very beginning of the Phase II monitoring, as demonstrated in Figure 11(c). From Figure 12(a), the P-CUSUM chart increases almost linearly over time, which can be explained as follows. Based on the IC data (i.e., from the 9th to the 105th residuals shown in

Figure 11(c)), the five intervals for categorization of the normalized data (i.e., $(\epsilon(i) - \widehat{\mu}_0)/\widehat{\sigma}$ where $\widehat{\mu}_0$ is the sample mean of the IC data, $\widehat{\sigma}$ is their sample standard deviation, and $\epsilon(i)$ is the ith residual) are $I_1 = (-\infty, -0.661], I_2 = (-0.661, -0.250], I_3 = (-0.250, 0.120], I_4 = (0.120, 0.503]$, and $I_5 = (0.503, \infty)$. From the 106th to the 162nd residuals, all of them, except the 106th, 109th, 139th, 145th and 154th residuals, belong to the last interval, which explains why the P-CUSUM chart looks straight and decreases only slightly at the five specific occasions.

5 Concluding Remarks

We have presented a new framework to construct nonparametric control charts for univariate SPC in cases when the process distribution cannot be specified beforehand. By this framework, original observations are first categorized, and then some statistical tools for categorical data analysis are used for constructing nonparametric control charts. Compared to existing control charts that are based on ranking information of the process observations, the proposed control charts have the advantages that (i) the information loss due to categorization can be controlled by adjusting the number of categories used, (ii) they can be used in cases with both numerical and categorical observations, and (iii) they can be used in cases with single-observation data. Numerical studies show that the P-CUSUM chart performs well in all cases considered, compared to its peers. Therefore, this is the nonparametric chart that we recommend to use in practice.

The comparative study presented in this paper is empirical. Much future research is required to confirm the conclusions presented here by mathematically more rigorous arguments. In this paper, we focus on Phase II online monitoring of the process mean. In some applications, simultaneous online monitoring of the process mean and variance would be our interest, which also requires much future research. Further, our proposed P-CUSUM chart does not require the specification of the IC distribution. But, we assume that an IC dataset is available for estimating certain IC parameters. As pointed out at the beginning of Section 2, it still requires much future research on how to do the corresponding Phase I analysis properly in cases when the IC distribution is unknown and how to collect the IC data at the end of Phase I SPC.

Acknowledgments: We thank the editor, the associate editor, and two referees for many constructive comments and suggestions, which greatly improved the quality of the paper.

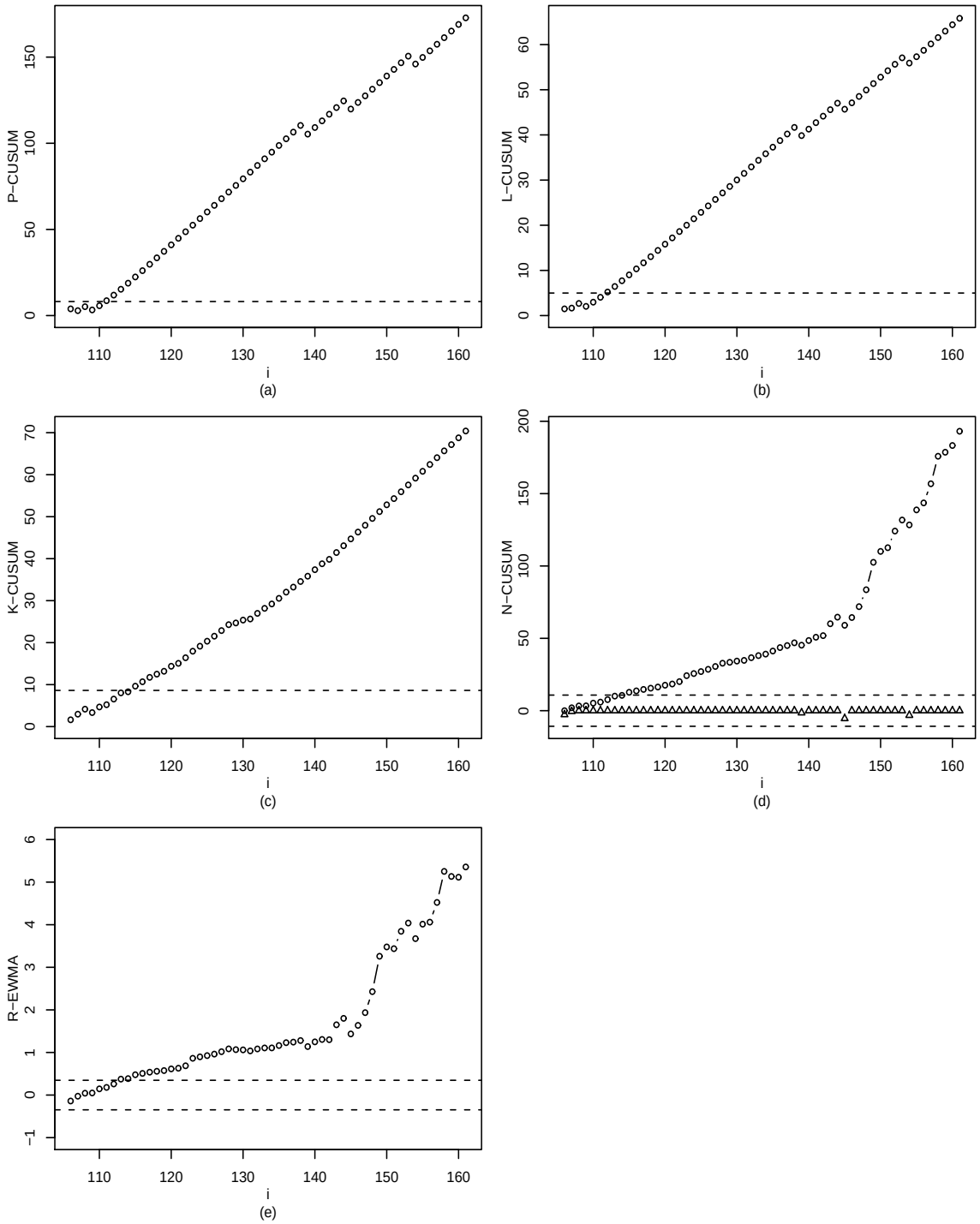

Figure 12: Control charts P-CUSUM, L-CUSUM, K-CUSUM, N-CUSUM, and R-EWMA when they are applied to the exchange rate data. In each plot, the horizontal dashed line(s) denotes the control limit(s). The N-CUSUM chart is two-sided; the little circles and triangles in plot (d) denote their upward and downward charting statistics, respectively.

References

Agresti, A. (2002), *Categorical data analysis (2nd edition)*, New York: John Wiley & Sons.

Albers, W., and Kallenberg, W.C.M. (2004), "Empirical nonparametric control charts: estimation effects and corrections," *Journal of Applied Statistics*, **31**, 345–360.

Albers, W., and Kallenberg W.C.M. (2009), "CUMIN charts," *Metrika*, **70**, 111–130.

Albers, W., Kallenberg, W.C.M., and Nurdiati, S. (2006), "Data driven choice of control charts," *Journal of Statistical Planning and Inference*, **136**, 909–941.

Amin, R., Reynolds, M. R. Jr., and Bakir, S. T. (1995), "Nonparametric quality control charts based on the sign statistic," *Communications in Statistics-Theory and Methods*, **24**, 1597–1623.

Amin, R.W., and Searcy, A.J. (1991), "A nonparametric exponentially weighted Moving Average Control Scheme," *Communications in Statistics -Simulation*, **20**, 1049–1072.

Amin, R.W., and Widmaier, O. (1999), "Sign control charts with variable sampling intervals," *Communications in Statistics: Theory and Methods*, **28**, 1961–1985.

Bakir, S. T. (2004), "A distribution-free Shewhart quality control chart based on signed-ranks," *Quality Engineering*, **16**, 613–623.

Bakir, S. T. (2006), "Distribution-free quality control charts based on signed-rank-like statistics," *Communications in Statistics-Theory and Methods*, **35**, 743–757.

Bakir, S.T., and Reynolds, M.R. Jr. (1979), "A nonparametric procedure for process control based on within group ranking," *Technometrics*, **21**, 175–183.

Borror, C.M., Montgomery, D.C., and Runger, G.C. (1999), "Robustness of the EWMA control chart to non-normality," *Journal of Quality Technology*, **31**, 309–316.

Chakraborti, S., and Eryilmaz, S. (2007), "A nonparametric Shewhart-type signed-rank control chart based on runs," *Communications in Statistics-Simulation and Computation*, **36**, 335–356.

Chakraborti, S., Eryilmaz, S., and Human, S. W. (2009), "A phase II nonparametric control chart based on precedence statistics with runs-type signaling rules," *Computational Statistics and Data Analysis*, **53**, 1054–1065.

Chakraborti, S., and van der Laan, P. (2000), "Precedence probability and Prediction Intervals," *Journal of the Royal Statistical Society (Series D)- The Statistician*, **49**, 219–228.

Chakraborti, S., van der Laan, P., and Bakir, S. T. (2001), "Nonparametric control charts: an overview and some results," *Journal of Quality Technology*, **33**, 304–315.

Chakraborti, S., van der Laan, P., and van de Wiel, M.A. (2004), "A class of distribution-free control charts," *Journal of the Royal Statistical society (Series C)-Applied Statistics*, **53**, 443–462.

Chakravarti, I.M., Laha, R.G., and Roy, J. (1967), *Handbook of Methods of Applied Statistics (Volume I)*, New York: John Wiley & Sons.

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

Chen, G., Cheng, S., and Xie, H. (2001), "Monitoring process mean and variability with one EWMA chart," *Journal of Quality Technology*, **33**, 223–233.

Chen, G., and Zhang, L. (2005), "An Extended EWMA Mean Chart," *Quality Technology and Quantitative Management*, **2**, 39–52.

Crosier, R. B. (1988), "Multivariate generalizations of cumulative sum quality-control schemes," *Technometrics*, **30**, 291–303.

Hackl, P., and Ledolter, J. (1992), "A new nonparametric quality control technique," *Communications in Statistics-Simulation and Computation*, **21**, 423–443.

Hawkins, D.M., and Olwell, D.H. (1998), *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag.

Hawkins, D.M., Qiu, P., and Chang Wook Kang (2003), "The changepoint model for statistical process control," *Journal of Quality Technology*, **35**, 355–366.

Lucas, J. M., and Crosier, R. B. (1982), "Robust CUSUM: a robust study for CUSUM quality control schemes," *Communications in Statistics-Theory and Methods*, **11**, 2669–2687.

Montgomery, D. C. (2001), *Introduction To Statistical Quality Control (4th edition)*, New York: John Wiley & Sons.

Moustakides, G.V. (1986), "Optimal stopping times for detecting changes in distributions," *The Annals of Statistics*, **14**, 1379–1387.

Page, E.S. (1954), "Continuous inspection schemes," *Biometrika*, **41**, 100-114.

Qiu, P. (2008), "Distribution-free multivariate process control based on log-linear modeling," *IIE Transactions*, **40**, 664–677.

Qiu, P., and Hawkins, D. M. (2001), "A rank based multivariate CUSUM procedure," *Technometrics*, **43**, 120–132.

Qiu, P., and Hawkins, D. M. (2003), "A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions," *Journal of Royal Statistical Society (Series D) - The Statistician*, **52**, 151–164.

Ritov, Y. (1990), "Decision theoretic optimality of the CUSUM procedure," *The Annals of Statistics*, **18**, 1464–1469.

Steiner, S.H. (1998), "Grouped data exponentially weighted moving average control charts," *Applied Statistics*, **47**, 203–216.

Steiner, S.H., Geyer, P.L., and Wesolowsky, G.O. (1996), "Grouped data-sequential probability ratio tests and cumulative sum control charts," *Technometrics*, **38**, 230–237.

Woodall, W.H. (2000), "Controversies and contradictions in statistical process control," *Journal of Quality Technology*, **32**, 341–350.

Yeh, A.B., Lin, D.K.-J., and Venkataramani, C., (2004), "Unified CUSUM control charts for monitoring process mean and variability," *Quality Technology and Quantitative Management*, **1**, 65–85.