

An adaptive Shiryaev-Roberts procedure for signalling varying location shifts

Jiujun Zhang^a, Zhonghua Li^b, Qin Zhou^c, Zhaojun Wang^{b,*}

^a*Department of Mathematics, Liaoning University, Shenyang 110036, P.R.China*

^b*LPMC and Institute of Statistics, Nankai University, Tianjin 300071, P.R.China*

^c*School of Mathematical Sciences, Jiangsu Normal University, Xuzhou 221116, P.R.China*

Abstract

This paper proposes a new adaptive chart based on the Shiryaev-Roberts procedure, by updating the reference value in an adaptive way to achieve the aim of overall good performance over a range of future expected but unknown mean shifts. A two-dimensional Markov chain model is developed to analyze the run length performance. The design guidelines are given. The comparisons of run length performance of the proposed scheme and other charts show that the proposed chart provides quite effective detecting ability over a range of mean shift sizes. The implementation of the new chart is illustrated by a real data example.

Key words: Shiryaev-Roberts Procedure; EWMA; Statistical Process Control
1991 MSC: 62P30

1 Introduction

Statistical surveillance is used when regularly observed time series data are available, with the goal of giving a signal as soon as the data provide enough evidence of an important distributional change while maintaining the rate of false alarms at a low level. For this purpose, many alternative control chart approaches have been proposed in the literature. A statistical process control (SPC) chart is a procedure devised to monitor the stability of a process by

* Corresponding author.

Email addresses: jiujunzhang2000@yahoo.com.cn (Jiujun Zhang),
zli@nankai.edu.cn (Zhonghua Li), zhouqin8808@163.com (Qin Zhou),
zjwang@nankai.edu.cn (Zhaojun Wang).

plotting a sequence of statistics on a chart that involves a centerline and one or more statistically determined control limits and is used in evaluating whether a process is in control or out of control.

SPC charts originated in the late 1920s with the works of [Shewhart \(1926\)](#) and [Shewhart \(1931\)](#). While they are effective in detecting large shifts in a process parameter, Shewhart charts are not very effective in reacting to small parameter shifts. Thus, alternative control chart approaches that combine information over time have been proposed in the literature, e.g. the cumulative sum (CUSUM) chart, exponentially weighted moving average (EWMA) chart. These alternative control charts have been shown to be more effective than the Shewhart charts in detecting small, sustained shifts in a process parameter.

In 1961, for detecting a change in the drift of a Brownian motion, Shiryaev introduced an alternative change detection procedure, which is now usually referred to as the Shiryaev-Roberts (SR) procedure ([Shiryaev \(1961\)](#) and [Roberts \(1966\)](#)). The SR procedure has received some attention in the literature. [Pollak and Siegmund \(1985\)](#) compared the SR procedure with the CUSUM procedure for detecting a change in the drift of a Brownian motion process based on the conditional average delay time criterion. In addition, [Pollak and Siegmund \(1991\)](#) compared the performance of the CUSUM with the SR procedure for detecting a shift in a normal mean when the in-control value of the mean is unknown, but can be estimated from a training sample. [Kenett and Pollack \(1996\)](#) showed that the SR surveillance scheme has several advantages over classical CUSUM charts when a non-homogeneous Poisson process is considered. [Moustakides et al. \(2009\)](#) also offered a detailed comparative performance analysis of the SR procedure and the CUSUM chart, but the results are based on asymptotic approximations. In addition, [Moustakides et al. \(2011\)](#) developed integral equations and a concise numerical method to compute a number of performance metrics and evaluated the SR procedure's performance for various initialization strategies. Furthermore, [Polunchenko et.al \(2014\)](#) extended the framework of [Moustakides et al. \(2011\)](#). Recently, [Zhang et al. \(2011a\)](#) proposed a new single chart with SR procedure to monitor both mean and variance. It is shown that the new chart performs better than the other charts in most cases.

All of the properties mentioned above give the SR procedure an advantage over the other competing procedures. Moreover, the optimal performance of the SR charts heavily depends on the assumption that the shift magnitude, say δ , is known. In other words, optimal selection of the reference value relies on the target shift. Because we rarely know the exact shift value of a process before it is detected, it may be more important to look at a range of known or unknown mean shifts. Thus, a SR chart with a pre-specified reference value usually can not have an optimal performance for both small and large shifts. It is, therefore, desirable that a control chart performs well over a range of

mean shift sizes.

To provide good ARL performance over a range of mean shift sizes, an alternative approach is to consider a multiple model or a mixture of several control charts. In fact, [Lorden \(1971\)](#) has already considered and studied such a model. [Lucas \(1982\)](#) and [Hawkins and Zamba \(2003\)](#) proposed a Shewhart limit in conjunction with the CUSUM chart. [Lucas and Saccucci \(1990\)](#) recommended that a Shewhart limit be used in conjunction with the EWMA chart. This method can increase the efficiency of a chart for detecting large shifts. In addition, [Dragalin \(1997\)](#) investigated and studied a combination of several CUSUM charts with running extensive simulations to detect mean shifts in a range. They have shown the efficiency of the combined CUSUM charts, and provided various designs for these procedures, based on numerical simulations. This approach is very flexible, but with this comes greater complexity in its design.

Another effective alternative method is to use a single adaptive control chart. When a single chart is used, the design and operation of the monitoring scheme can be greatly simplified compared with the combined charts. Among them, [Lorden \(1971\)](#) first proposed a so-called generalized CUSUM procedure when the parameter is unknown. Later [Pollak and Siegmund \(1975\)](#) proposed a mixture type analog of CUSUM procedure. [Pollak \(1987\)](#) introduced the mixture-type analog of SR procedure and proved the second-order asymptotic minimaxity of this procedure in a certain sense. The main difficulties in implementing these procedures are the complexity of their statistics. [Dragalin \(1996\)](#) presented two adaptive procedures (VD chart) to estimate the process parameter with the CUSUM and SR chart when the parameter is unknown. The main drawback of this procedure is that the statistic is inappropriate since the control limit of the chart changes as estimation of δ , varies and using a unified control limit will result in imbalanced detection ability for different values of δ .

Recently, [Sparks \(2000\)](#) proposed a new adaptive CUSUM (ACUSUM chart) procedure by dynamically adjusting the reference value of the conventional CUSUM chart. Simulation results showed that the proposed ACUSUM chart is robust at signaling varying location shifts. Further, [Shu and Jiang \(2006\)](#) developed a two-dimensional Markov chain model to analyze the performance of ACUSUM charts. [Jiang et al \(2008\)](#) also proposed an adaptive CUSUM chart with EWMA-based shift estimators (JSA chart). [Zhang et al. \(2011b\)](#) discussed an adaptive SR chart for monitoring the process variance over a range of dispersion shift sizes. To this end, in this paper, we develop a new adaptive SR procedure for monitoring the process mean. Our new adaptive chart is much simpler than the VD chart and it can be anticipated that the new chart will provide an overall good performance over a range of future mean shifts.

The remainder of the paper is organized as follows. In Section 2, the SR test is presented and the comparison between the SR and the CUSUM chart is briefly reviewed. Furthermore, an adaptive SR procedure is proposed and the run length performance study through a Markov chain method is discussed in Section 3. The effects of parameters and design guidelines are provided in Section 4. The performance comparison for detecting is presented in Section 5. Section 6 contains a real data example to illustrate the application of our proposed chart. The conclusion and discussion are given in the last section.

2 The motivation of the proposed chart

In this section, we firstly review the SR procedure for location briefly and then the comparison with the CUSUM procedure is given by numerical analysis in terms of ARL.

2.1 The derivation of SR chart for monitoring a normal process mean

Let $\mathbf{x}_t = (x_{t1}, \dots, x_{tn})$ denotes a sample of size $n \geq 1$ taken on a quality characteristic x . The monitoring problems with $n > 1$ and $n = 1$ are usually referred to as group observations case and individual observations case, respectively. In what follows, we assume that the \mathbf{x}_t for $t \geq 1$, the observations collected over time, come from the following process model

$$x_{ti} = \mu + \varepsilon_{ti}, \quad i = 1, \dots, n, \quad t = 1, 2, \dots,$$

where $\varepsilon_{t1}, \dots, \varepsilon_{tn}$ are identically and independently distributed (i.i.d) normal variables with mean 0 and standard deviation σ . When the process is in-control, $\mu = \mu_0$ and $\sigma = \sigma_0$. In this paper, we consider the Phase II case in which the in-control (IC) μ_0 and σ_0 , are assumed to be known, i.e., it is assumed that the IC data set used in Phase I is enough to estimate the parameters well.

When a process shift occurs, $\mu = \mu_1$, where $\mu_1 = \mu_0 + \delta\sigma_0$ and $\delta \neq 0$ and the probability density function (pdf) is denoted as $g(x)$. The values of δ is known before monitoring. Without loss of generality, we assume $\mu_0 = 0$ and $\sigma_0 = 1$ and the pdf is denoted as $f(x)$. If the in-control mean is not 0, or the in-control variance is not 1, one can transform the random variable such that the distribution of the transformed variable is $N(0, 1)$.

According to [Moustakides et al. \(2009\)](#), for $t \geq 1$, define $\Lambda_t = \frac{g(\mathbf{x}_t)}{f(\mathbf{x}_t)}$, to be the ‘‘instantaneous’’ likelihood ratio between the post-change and pre-change hypotheses. Then, under the assumption, we have

$$\Lambda_t = \frac{g(x_{t1}, x_{t2}, \dots, x_{tn})}{f(x_{t1}, x_{t2}, \dots, x_{tn})} = \exp\left\{\delta \sum_{j=1}^n x_{tj} - \frac{n\delta^2}{2}\right\} = \exp\left\{n\delta\bar{x}_t - \frac{n\delta^2}{2}\right\}, \quad (1)$$

where $\bar{x}_t = \frac{1}{n} \sum_{j=1}^n x_{tj}$ is the sample mean. The SR procedure stops and raises an alarm at $T_h^{SR} = \inf\{t \geq 1 : R_t \geq h\}$, where R_t is the SR detection statistic defined as

$$R_t = \sum_{l=1}^t \prod_{j=l}^t \Lambda_j, \quad (2)$$

and $h > 0$ is chosen to achieve a specified IC-ARL. It is easily verified that the SR statistic allows the following convenient recursive representation

$$R_t = (1 + R_{t-1})\Lambda_t, \quad (3)$$

where $R_0 = 0$. Next, we will compare the SR chart with the traditional CUSUM chart in terms of ARL.

2.2 Comparison between the SR and the CUSUM chart

The CUSUM chart proposed by Page (1954) is one of the most popular algorithms for accomplishing this, in part due to certain ARL optimality properties (Lorden, 1971; Pollak, 1985; Moustakides, 1986). The CUSUM works by accumulating the deviations from the target μ_0 that are above target with one statistic C^+ and that are below target with another statistic C^- . The statistics C^+ and C^- are called one sided upper and lower CUSUM respectively. They are calculated as

$$C_t^+ = \max(0, C_{t-1}^+ + (x_t - \mu_0) - k), \quad (4)$$

$$C_t^- = \max(0, C_{t-1}^- - (x_t - \mu_0) - k), \quad (5)$$

with the initial value $C_0^+ = C_0^- = 0$. The value of k is called the reference or allowable value and it is often chosen about halfway between the target μ_0 and the shift of mean that one is interested in detecting. Thus, $k = \frac{1}{2}|\mu_1 - \mu_0|$. If either of the two equations (4) and (5) exceeds the control limit h , the process is said to be out of control. In case of rational subgroups (of sizes $n > 1$), x_t is replaced by \bar{x}_t .

CUSUM chart is superior to the Shewhart chart at least for small and moderate shifts. Therefore, it is worthwhile to compare the performance of SR and CUSUM charts. Theoretical properties of the SR procedure have been developed and are well documented in the literature. However, most of the results

are based on asymptotic approximations found in sequential analysis or diffusion theory. The limited analytical results found in the literature are usually comparisons of the asymptotic approximations to results obtained through Monte Carlo simulations.

In the literature, run length performance is usually used to compare surveillance procedures, where the run length is the number of samples taken until the chart signals a change in the process parameter. The most commonly used efficiency criterion is the average run length (ARL), which is the average number of observations needed for the procedure to signal a change in the distribution. It is desirable that the ARL should be large if the process is in control and be small if the process is out-of-control. Quite often zero-state (ZS) and steady-state (SS) ARL measures are used for this purpose. The ZS-ARL is based on the assumption that the specified shift in the parameter occurs when the chart is at its initial condition. On the other hand, the SS-ARL is based on averaging these out-of-control values, but under the assumption that the process has been operating for a while and that the process mean stays on the target until the specified shift in the mean occurs. Based on the run length performance, the properties of the CUSUM chart have been thoroughly studied in the literatures.

It is well known that both schemes enjoy specific optimality properties under different optimality criteria. More precisely, it follows from [Moustakides \(1986\)](#) that the CUSUM procedure is (min-max) optimal with respect to the [Lorden \(1971\)](#) detection measure. On the other hand, it follows from [Pollak and Tartakovsky \(2009\)](#) that the SR procedure is optimal with respect to the relative integral average detection delay (RIADD) measure. The exact analytical characterization of the two performance measures was recently made possible by [Moustakides et al. \(2009\)](#) through a set of integral equations. It is shown that the CUSUM procedure outperforms the SR procedure with respect to Lorden's performance measure, i.e., ZS-ARL, while the SR procedure is superior with respect to the RIADD measure, i.e., SS-ARL.

In order to make it clear, we made a similar comparison between the SR procedure and the CUSUM procedure in terms of SS-ARL for different values of δ (the smallest standardized shift considered important to be detected quickly) and under different possible shift sizes in the process mean, μ . The IC-ARL is chosen to be 200 and the sample size are set to be 1 and 5, respectively. All ARL values are estimated using 100,000 simulations of data taken as normally distributed. The results are summarized in Table 1. It can be seen that the SS-ARLs for the SR procedure are always smaller than the SS ARLs of CUSUM chart only if $\mu \leq \delta$. Similar conclusions were also given in [Mahmouda et al. \(2008\)](#), which were based on only the case of $n = 1$.

[Insert Table 1 about here]

It should be noted that δ is the smallest standardized shift size that a surveillance method is designed to detect optimally. In practice, however, the true shift can be different from the shift for which the surveillance method is optimal. A drawback of the SR scheme is that it is designed to detect a particular mean shift size, say δ , and may perform far away for detecting shifts of sizes smaller or larger than δ . To this end, in the next section, we propose an adaptive SR (ASR chart) chart for monitoring process mean by integrating the idea of the adaptive CUSUM for process mean problem proposed by Sparks (2000) with the SR chart. It can be easily implemented and provides a balanced protection against a broad range of shift sizes by properly setting parameters.

3 An adaptive SR chart for the process mean

As discussed above, the SR chart can only be optimized in terms of SS-ARL if we have accurate information on the value of δ . In practice, δ is unknown and in certain cases its magnitude even varies over time due to causes such as gradual deterioration of equipment, waste accumulation or human causes. In order to enhance the detection ability for a range of shift sizes, we propose an adaptive charting method based on SR procedure.

3.1 An adaptive SR chart for the process mean

Assume that we are only interested in detecting process increase, i.e., $\delta > 0$. Let δ_t standard for the process mean at time t . It is natural to use this value in the SR statistic by setting $\delta = \delta_t$, say,

$$R_t^U = (1 + R_{t-1}^U)\Lambda_t^*, \quad (6)$$

where

$$\Lambda_t^* = \exp\left\{n\delta_t\bar{x}_t - \frac{n}{2}\delta_t^2\right\}. \quad (7)$$

A signal is given when R_t^U exceeds a specified control limit h . It might be expected to be nearly optimal in detecting the mean change in such situations. However, this statistic is inappropriate since the control limit of R_t^U changes as δ_t varies and using a unified control limit will result in imbalanced detection ability for different values of δ (see Sparks (2000) for a related discussion). Alternatively, we proposed to use $R_t^U/h_U(\delta_t)$, where $h_U(\delta_t)$ is an operating function that denotes the control limit for the upper-sided ASR chart. Using

operating function is one of the main differences between our chart and the VD chart.

In this paper, we come up with a regression fit approximating this function and use it as a convenient formula for practitioners. We use EXCEL to fit these functions. Table 2 tabulates the approximated $h_U(\cdot)$ for $n = 1$ and $n = 5$ and IC ARL=100, 200, 300, 400 and 500 by six-order polynomial fits based on forty points $(\delta_i, h_U(\delta_i))$ with δ_i equally spaced in $[0.05, 2.0]$.

[Insert Table 2 about here]

Certainly, it is required to estimate δ_t in applications. Moreover, the EWMA scheme has been used in Sparks (2000) and Shu and Jiang (2006) to estimate the current process mean level. Although some other methods can be used, we still use the EWMA scheme due to its simplicity and efficiency in the current paper. To be specific, the conventional EWMA statistic for the process mean can be obtained recursively as

$$Q_t = (1 - \lambda)Q_{t-1} + \lambda\bar{x}_t, \quad (8)$$

where λ is a smoothing parameter with $0 < \lambda < 1$. Note that this estimator is such simpler than its counterpart VD chart.

Practically speaking, when detecting upward shift, there is always a minimum magnitude of interest of $\delta > 0$ for early detection, say δ_{\min} . In order to maintain the advantages of ASR charts for detecting small shifts while improving the sensitivity for detecting large shifts, δ_{\min} usually takes small values in the ASR procedures to guarantee the sensitivity to small shifts.

Therefore, for the purpose of improving the efficiency in detecting shifts larger than δ_{\min} , we suggest to use

$$Q_t^+ = \max\{\delta_{\min}, (1 - \lambda)Q_{t-1}^+ + \lambda\bar{x}_t\}, \quad (9)$$

instead, where the starting value of Q_0^+ can be set to δ_{\min} or some other values.

Finally, by using (5), the upper-sided ASR charting statistic is defined as

$$R_t^U = (1 + R_{t-1}^U)\Lambda_t^*, \quad (10)$$

where

$$\Lambda_t^* = \exp\left\{nQ_t^+\bar{x}_t - \frac{n}{2}Q_t^{+2}\right\}. \quad (11)$$

The ASR chart triggers a signal when $R_t^U/h_U(Q_t^+) > h_c$, where $h_c > 0$ is a control limit chosen to achieve a specified IC-ARL.

3.2 Markov chain estimates of ARL for the ASR chart

The Markov chain approach, integral equation approach and Monte Carlo simulations have been widely used to study the run length performance of control charts (Li et al, 2014). Due to the simple format of the estimator of δ in equation (8), the run length performance of our ASR charts, as described below, can be investigated using a two-dimensional Markov chain model. This allows one to apply a two-dimensional extension of the univariate Markov chain approach to calculate the ARL. For the VD chart, the run length performance was evaluated by simulation and this is the advantage of our ASR chart compared with the VD chart.

Similar to Zhang et al. (2011b), we develop a Markov chain model for an upper ASR chart. It can be seen that the random vector $(\frac{R_t^U}{h(Q_t^+)}, Q_t^+)$ can be modeled as a Markov chain. First, we choose a large control limit L for the process of Q_t^+ so that the IC region can be partitioned within a two-dimensional rectangle $[0, h_c] \times [\delta_{\min}, L]$ to a discretized Markov chain, i.e., $\{0 \leq \frac{R_t^U}{h(Q_t^+)} \leq h_c, \delta_{\min} \leq Q_t^+ \leq L\}$. Assume that the number of states along the axis $\frac{R_t^U}{h(Q_t^+)}$ over the range $[0, h_c]$ is m_1 , then the width of each segment is $\omega = 2h_c/(2m_1 - 1)$, except that the width of the first segment is $\frac{\omega}{2}$. Similarly, the axis Q_t^+ over the interval $[\delta_{\min}, L]$ is segmented into m_2 states such that the width of each segment is $\Delta = 2(L - \delta_{\min})/(2m_2 - 1)$, except that the width of the first segment is $\Delta/2$. The states along the axis $R_t^U/h(Q_t^+)$ and Q_t^+ are labeled by $i = 0, 1, \dots, (m_1 - 1)$ and $j = 0, 1, \dots, (m_2 - 1)$, respectively. The center point of state i along the axis $R_t^U/h(Q_t^+)$ is $i\omega$, and the center point of state j along the axis Q_t^+ is $\delta_{\min} + j\Delta$. Therefore, the IC region is divided into a number of $N = m_1 \times m_2$ two-dimensional rectangles.

Define:

$$\begin{aligned} a_1 &= \frac{1}{n(\delta_{\min} + l\Delta)} \left\{ \log \left[\frac{(k - 0.5)\omega h(\delta_{\min} + l\Delta)}{1 + i\omega h(\delta_{\min} + j\Delta)} \right] + \frac{n(\delta_{\min} + l\Delta)^2}{2} \right\}, \\ a_2 &= \frac{1}{n(\delta_{\min} + l\Delta)} \left\{ \log \left[\frac{(k + 0.5)\omega h(\delta_{\min} + l\Delta)}{1 + i\omega h(\delta_{\min} + j\Delta)} \right] + \frac{n(\delta_{\min} + l\Delta)^2}{2} \right\}, \\ b_1 &= \delta_{\min} + [l - 0.5 - (1 - \lambda)j] \frac{\Delta}{\lambda}, \quad b_2 = \delta_{\min} + [l + 0.5 - (1 - \lambda)j] \frac{\Delta}{\lambda}, \\ c &= \min[b_2, \max(a_1, b_1)], \quad d = \max[b_1, \min(a_2, b_2)], \quad e = \max[a_1, \min(a_2, b_2)]. \end{aligned}$$

Let $P_{(i,j)(k,l)}$ be the transition probability of $(\frac{R_t^U}{h(Q_t^+)}, Q_t^+)$ from state (i, j) to state (k, l) . Then, when $k \neq 0$ and $l \neq 0$, the transition probability $P_{(i,j)(k,l)}$ can be evaluated by

$$\begin{aligned}
P_{(i,j)(k,l)} &= Pr\left\{\frac{R_t^U}{h(Q_t^+)} \text{ instate } k, Q_t^+ \text{ instate } i \mid \frac{R_{t-1}^U}{h(Q_{t-1}^+)} \text{ instate } i, Q_{t-1}^+ \text{ instate } j\right\} \\
&= Pr\left\{(k - 0.5)\omega < R_t/h(Q_t^+) < (k + 0.5)\omega, \right. \\
&\quad \left. \delta_{\min} + (l - 0.5)\Delta < (1 - \lambda)Q_{t-1}^+ + \lambda\bar{x}_t < \delta_{\min} + (l + 0.5)\Delta \right. \\
&\quad \left. \mid \frac{R_{t-1}^U}{h(Q_{t-1}^+)} = i\omega, Q_{t-1}^+ = \delta_{\min} + j\Delta\right\} \\
&= Pr\left\{(k - 0.5)\omega < \frac{1 + R_{t-1}^U}{h(Q_t^+)}\Lambda_t^* < (k + 0.5)\omega, b_1 < \bar{x}_t < b_2\right\} \\
&= Pr\left\{\frac{(k - 0.5)\omega h(Q_t^+)}{1 + i\omega h(\delta_{\min} + j\Delta)} < \Lambda_t^* < \frac{(k + 0.5)\omega h(Q_t^+)}{1 + i\omega h(\delta_{\min} + j\Delta)}, b_1 < \bar{x}_t < b_2\right\} \\
&= Pr\{a_1 < \bar{x}_t < a_2, b_1 < \bar{x}_t < b_2\},
\end{aligned}$$

then, when the process is IC, the probability can be presented as

$$\begin{aligned}
P_{(i,j)(k,l)} &= Pr\{c < \bar{x}_t < d\} \\
&= \Phi(\sqrt{nd}) - \Phi(\sqrt{nc}),
\end{aligned}$$

where $\phi(\cdot)$ is the probability distribution function of the standard normal distribution.

Similarly, when $k \neq 0$ and $l = 0$,

$$\begin{aligned}
P_{(i,j)(k,0)} &= Pr\{a_1 < \bar{x}_t < a_2, \bar{x}_t < b_2\} \\
&= Pr\{a_1 < \bar{x}_t < \max[a_1, \min(a_2, b_2)]\} \\
&= \Phi(\sqrt{ne}) - \Phi(\sqrt{na_1}).
\end{aligned}$$

When $k = 0$ and $l \neq 0$,

$$\begin{aligned}
P_{(i,j)(0,l)} &= Pr\{\bar{x}_t < a_2, b_1 < \bar{x}_t < b_2\} \\
&= Pr\{b_1 < \bar{x}_t < \max[b_1, \min(a_2, b_2)]\} \\
&= \Phi(\sqrt{nd}) - \Phi(\sqrt{nb_1}).
\end{aligned}$$

When $k = 0$ and $l = 0$,

$$\begin{aligned}
P_{(i,j)(0,0)} &= Pr\{\bar{x}_t < a_2, \bar{x}_t < b_2\} \\
&= Pr\{\bar{x}_t < \min(a_2, b_2)\}
\end{aligned}$$

$$= \Phi(\sqrt{n} \min(a_2, b_2)).$$

Then, we have

$$P_{(i,j)(k,l)} = \begin{cases} Pr\{c < \bar{x}_t < d\}, & \text{if } k \neq 0 \text{ and } l \neq 0 \\ Pr\{a_1 < \bar{x}_t < e\}, & \text{if } k \neq 0 \text{ and } l = 0 \\ Pr\{b_1 < \bar{x}_t < d\}, & \text{if } k = 0 \text{ and } l \neq 0 \\ Pr\{\bar{x}_t < \min(a_2, b_2)\}, & \text{if } k = 0 \text{ and } l = 0, \end{cases}$$

The ZS-ARL of the upper ASR chart can then be obtained by the following equation:

$$ARL = \mathbf{p}_{ini}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1},$$

where \mathbf{p}_{ini} is any initial probability vector of states and the submatrix Q contains the probabilities of going from one transient state to another. Transient states are referred to as in-control states, and the absorbing state is referred to as the out-of-control state.

In order to calibrate the SS-ARL, according to [Lucas and Saccucci \(1990\)](#), define

$$\mathbf{P}^* = \begin{pmatrix} \mathbf{Q} & (\mathbf{I} - \mathbf{Q})\mathbf{1} \\ 0 \dots 1 \dots 0 & 0 \end{pmatrix}.$$

Let \mathbf{p}_{ss} be the cyclical steady-state probability vector, then \mathbf{p}_{ss} can be obtained by solving simultaneously $\mathbf{p} = \mathbf{P}^{*T} \mathbf{p}$ subject to $\mathbf{p}\mathbf{1} = 1$, where $\mathbf{1}$ is a vector of ones. The cyclical SS-ARL is given by

$$ARL = \mathbf{p}_{ss}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}.$$

4 Effects of parameters and design guidelines

From the description of Section 3, the performance of the ASR chart is related to the initial value Q_0^+ , the smoothing parameter λ , and the minimum

magnitude δ_{\min} . First, we investigate the effect of the initial value of Q_0^+ on charts' performance.

It can be expected that Q_0^+ does not affect charts' steady-state performance and thus we only need to consider the zero-state cases. Table 3 tabulates the OC-ARL values of ASR charts with different values of $Q_0^+ = 1.0, 1.5, 2.0, 2.5$ under different possible shift sizes in the process mean, μ , when $\lambda = 0.1, \delta_{\min} = 1.0, n = 1$. Q_0^+ has a small effect on the OC-ARL and less effect on the IC-ARL. In general, a large value of Q_0^+ slightly improves the sensitivity of ASR charts to large shifts but reduces the sensitivity to small shifts, and vice versa for a small value of Q_0^+ . Thus, we recommend setting $Q_0^+ = 2.0$ and this choice will be used in the following comparisons.

[Insert Table 3 about here]

It should be noted that the control limit $h = 1.5$ provides nearly an identical IC-ARL for all ASR charts with different Q_0^+ . [Shu and Jiang \(2006\)](#) pointed out that this insensitivity of the IC-ARL to Q_0^+ indicates a very good property because the OC performance can be adjusted by changing Q_0^+ without changing the control limit. This is similar to the effect of a head start on conventional CUSUM and EWMA charts ([Lucas \(1982\)](#)).

The effects of the parameters (δ_{\min}, λ) are also investigated. Table 4 presents the OC-ARL values of ASR charts with different values of $\delta_{\min} = 0.25, 0.5, 0.75, 1.0$ for various mean shifts when $\lambda = 0.1, Q_0^+ = 2.0, n = 5$. Generally, a large value of δ_{\min} improves the sensitivity of ASR charts to large shifts but reduces the sensitivity to small shifts, and vice versa for a small value of δ_{\min} . Moreover, the value of δ_{\min} improves the detection ability of ASR when $\delta \geq \delta_{\min}$ but reduces the efficiency when $\delta < \delta_{\min}$. Overall, the ASR chart with $\delta_{\min} = 0.5$ seems to be always robust in detecting various mean changes on the whole, and thus it is a reasonable choice in practice.

[Insert Table 4 about here]

We can also observe that, once the parameters δ_{\min} and Q_0^+ are fixed, λ can be tuned to minimize the OC-ARL for any particular shift. However, this optimal value of λ degrades the sensitivity to other shifts and finding the optimal value of λ is too complicated for practical use. To this end, we can select $\lambda \in [0.05, 0.2]$, as is often suggested in literature.

From the performance of the ASR charts shown above, the following design procedure is recommended when the ASR chart is used: 1) Specify the IC-ARL; 2) If there is no prior knowledge of the dispersion shift size, $\delta_{\min} = 0.5$ and $\lambda = 0.1$ are suggested, which are useful choices in practice in order to

detect small and moderate mean shifts quickly; 3) If there is some knowledge of the mean shift size and the shift size is small, small δ_{\min} and λ are suggested, and vice versa; 4) After specifying (δ_{\min}, λ) , determine the control limit h to achieve a desired IC-ARL. A Fortran program to find the control limit h is available from the authors upon request.

5 Comparison with other charts

In this section, we compare our ASR chart with the SR chart, JSA chart and the VD chart. The JSA chart performs better than the ACUSUM chart, so, we will not consider the ACUSUM chart in this comparison. Some other works, such as those of [Lorden and Pollak \(2005\)](#) and [Lorden and Pollak \(2008\)](#), also address this problem via direct estimation. However, their works are based on mathematical theory, which we will not consider, either.

Suppose $[\delta_1, \delta_2]$ is the potential mean shift range for detection. In this comparison, the IC-ARL is taken as 400 and $n = 1$ is considered. In order to be consistent with [Jiang et al \(2008\)](#), the ASR parameters are chosen as $\lambda = 0.3$, $\delta_{\min} = 1.0$, the JSA parameters are chosen as $\lambda = 0.3$, $\delta_{\min} = 1.0$, $\gamma = 1.5$ and the VD parameter is chosen as $\theta_0 = 1$ (see [Dragalin \(1996\)](#) and [Jiang et al \(2008\)](#) for more details). For the other two SR charts, $\delta = \delta_1 = 0.25$ and $\delta = \delta_2 = 2$ (denoted as SR₁ and SR₂ chart) are considered. This range is considered to be small and moderate shift size.

In this Table, δ is the smallest standardized shift considered important to be detected quickly and μ is the possible shift sizes in the process mean. The ARL values of the SR and ASR charts are obtained using the Markov chain method and the ARL values of the JSA chart are selected from Table 2 of [Jiang et al \(2008\)](#). Both of the ZS-ARL and SS-ARL are considered in this comparison. The same values of control limits were used for computing both zero-state and steady-state OC-ARL values. The ZS-ARL and the SS-ARL comparison results are summarized in Table 5. From this table we observed the following results:

- The comparison between the ASR and SR charts:

It can be seen from this Table that a single SR chart can only signal either small or large shifts quickly once the corresponding δ is chosen. For example, for detecting the small shift, e.g., $\mu = 0.5$, the SS-ARL of the SR₁ chart is 21.93, but for the SR₂ chart, the SS-ARL increases to 43.77. However, for detecting the large shift, e.g., $\mu = 2$, the SS-ARL of the SR₁ chart is 5.31, but for the SR₂ chart, the corresponding value deduces to 2.83. Note that the difference between the ZS-ARL and SS-ARL of the SR₁ chart is very significant. Also we can see that the ASR chart performs better than the

SR₁ chart when $\mu > 0.50$ and almost always outperforms the SR₂, except for very large shift.

- The comparison between the ASR and JSA charts:

Apparently, for detecting small and moderate shift sizes, say, $\mu < 1$, the ASR chart is more sensitive than the JSA chart. On the other hand, for detecting large shift sizes, e.g., $\mu \geq 1$, the JSA chart performs a little better than the ASR chart.

- The comparison between the ASR and VD charts:

It can be seen that the VD chart is more sensitive than the ASR chart for detecting small shifts ,e.g., $\mu \leq 0.5$, but less sensitive for detecting large shifts.

Considering the satisfactory and robust performance for different sizes in detecting the process shifts, we believe it is worth taking the effort in designing and applying our proposed new chart. We will show the application of our proposed chart by a real data example in the next section.

[Insert Table 5 about here]

6 An example

To illustrate how the ASR procedure is implemented, in this section, the application of our proposed ASR chart is illustrated by a real data example which contains a data set consisting of measurements of the inside diameter of the cylinder bores in an engine block. [Chen et al. \(2001\)](#) used this data set to show the implementation of their MaxEWMA chart for monitoring the process mean and variance shifts. [Zhang et al. \(2011b\)](#) also used this data set to show the implementation of their ASR chart for monitoring the process variance shifts. The original data set can be found in Table 2 in [Chen et al. \(2001\)](#).

First, we use the grand average \bar{x} of the preliminary data to estimate the process mean μ and use \bar{S}/c_n to estimate the process variance σ , where c_n is a constant that depends only on the subgroup size n , $\bar{S} = (S_1 + \dots + S_m)/m$ is the average of the sample standard deviations, and $S_i^2 = \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 / (n - 1)$ is the i -th sample variance. In this data set, c_4 is the known constant and $c_4 = 0.94$ is used.

The process mean and variance are estimated from the data set and we have $\mu = 200.251$ and $\sigma = 3.306$. To use our ASR chart to monitor process mean, all the original data are standardized by the process mean and variance. In this example, $\delta_{min} = 0.25$, $Q_0 = 2.0$, $\lambda = 0.1$ is used and the IC-ARL is chosen

to be 200. The corresponding control limit is 2.30. The chart is shown as a plot in Fig. 1-(a). We can see that the curve has a suddenly decrease from sample 1 and this point exceeds the control limit, which is related to the process mean. According to [Chen et al. \(2001\)](#), sample 1 corresponds to the time when the regular operator is absent and a relief operator is in charge of the production. This result is also consistent with [Zhang et al. \(2011b\)](#). When this point is removed, we obtain $\mu = 200.123$ and $\sigma = 3.346$, and our second ASR chart is given as a plot in Fig. 1-(b). The OC signal is triggered at sample 11 and this sample is also related to the process mean, which is consistent with the results of the MaxEWMA chart when $k = 1$. When these two samples are removed, we obtain $\mu = 199.98$ and $\sigma = 3.378$, and our third ASR chart is given as a plot in Fig. 1-(c). It shows that there is no point falling outside the control limit.

It should be noted that there are two other OC points, point 6 and point 16 in [Chen et al. \(2001\)](#) and [Zhang et al. \(2011b\)](#), which are still in control in our chart. From the original analysis we know that these two points are related to the process variance. So, it can be seen that our ASR chart is not sensitive to the variance shift. This, again, shows that the ASR chart is quite a useful tool for practitioners to monitor the process mean.

[Insert Figure 1 about here]

7 Conclusions and considerations

In this paper, an adaptive SR chart (ASR chart) based on the Shiryaev-Roberts procedure is proposed for monitoring the process mean of a normal process that can be efficient in detecting a broader range of shifts, by dynamically adjusting its reference values according to current process information. A two-dimensional Markov chain model to analyze the run-length distribution of the ASR chart is proposed. Moreover, a simplified operating function is derived by six-order polynomial fits and this function is very accurate in the range $[0.05, 2.0]$. Since the ASR chart performs more robustly than other procedures for detecting the process mean shift, we recommend its use in practice. This paper considers the ASR charts performance with known parameters. When the process parameters, mean and standard deviation are estimated, the performance of the ASR chart will be affected. This will be further investigated in the future paper.

In addition, as pointed out by [Jiang et al \(2008\)](#), it is known that the EWMA scheme is inefficient in capturing abrupt process mean changes of moderate and large magnitudes. The best estimating procedures, according to [Yashchin \(1995\)](#), can not be linear if they are to adapt to changes of large magnitude

because the inertia (a long term measurement of estimation error) increases as the magnitude of the shift increases. To alleviate this limitation, [Yashchin \(1995\)](#) suggests using a non-linear estimator called the EWMA-C estimator, which is a generalization of the EWMA statistic. Thus, it is expected that the ASR chart using an EWMA-C mean estimator improves the same chart using an EWMA estimator. This will be our interest in the future research.

Acknowledgements

The authors are grateful to the editor and three anonymous referees for their valuable comments that have greatly improved the paper. This paper is supported by NSF of China Grant 11101198, 11201246, 11371202, 11131002, RFDP of China Grant 20110031110002, PAPD of Jiangsu Higher Education Institutions and the Research Fund for the Doctoral Program of Liaoning University.

References

- Chen, G., Cheng, S. W., Xie, H. (2001). Monitoring process mean and variability with one EWMA chart. *Journal of Quality Technology* 33: 223-233.
- Dragalin, V. (1996). Adaptive procedures for detecting a change in distribution. In *Proceedings of the 4th Umea-Wrzburg Conference in Statistics*, pages 87-103.
- Dragalin, V. (1997). The design and analysis of 2-CUSUM procedure. *Communications in Statistics-Simulation and Computation* 26: 67-81.
- Hawkins, D. M., Zamba, K. D. (2003). On small shifts in quality control. *Quality & Engineering* 16(4), 143-149.
- Jiang, W., Shu, L., Apley, D. W. (2008). Adaptive CUSUM procedures with EWMA-based shift estimators. *IIE Transactions* 40: 992-1003.
- Kenett, R. S. and Pollack, M. (1996). Data-analytic aspects of the Shiriyayev-Roberts control chart: surveillance of a non-homogeneous poisson process. *Journal of Applied Statistics* 23: 125-137.
- Li, Z., Zou, C., Gong, Z., Wang, Z. (2014). The computation of average run length and average time to signal: an overview. *Journal of Statistical Computation and Simulation* to appear (DOI:10.1080/00949655.2013.766737).
- Lorden, G. (1971). Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics* 42: 1897-1908.
- Lorden, G., Pollak, M. (2005). Nonanticipating estimation applied to sequential analysis and changepoint detection. *Annals of Statistics* 33: 1422-1454.
- Lorden, G., Pollak, M. (2008). Sequential change-point detection procedures

- that are nearly optimal and computationally simple. *Sequential Analysis* 27(04): 476-512.
- Lucas, J. M. (1982). Combined Shewhart-CUSUM quality control schemes. *Journal of Quality Technology* 14: 51-59.
- Lucas, J. M., Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: properties and enhancements.(with discussion). *Technometrics* 32: 1-29.
- Mahmouda, M. A., Woodall, W. H., Davis, R. E. (2008). Performance comparison of some likelihood ratio-based statistical surveillance methods. *Journal of Applied Statistics* 35: 783-798.
- Moustakides, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Annals of Statistics* 14: 1379-1387.
- Moustakides, G. V., Polunchenko, A. S., Tartakovsky, A. G. (2009). Numerical comparison of CUSUM and Shiryaev-Roberts procedures for detecting changes in distributions. *Communications in Statistics-Theory and Methods* 38: 3225-3239.
- Moustakides, G. V., Polunchenko, A. S., Tartakovsky, A. G. (2011). A numerical approach to performance analysis of quickest change-point detection procedures. *Statistica Sinica* 21: 571-596.
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika* 41: 100-114.
- Pollak, M. (1985). Optimal detection of a change in distribution. *Annals of Statistics* 13: 206-227.
- Pollak, M. (1987). Average run length of an optimal method of detecting a change in distribution. *Annals of Statistics* 15: 749-779.
- Pollak, M., Siegmund, D. (1975). Approximations to the expected sample size of certain sequential tests. *Annals of Statistics* 3: 1267-1282.
- Pollak, M., Siegmund, D. (1985). A diffusion process and its applications to detecting a change in the drift of Brownian motion. *Biometrika*. 72: 267-280
- Pollak, M., Siegmund, D. (1991). Sequential detection of a change in a normal mean when the initial value is unknown. *Annals of Statistics* 19: 394-416.
- Pollak, M., Tartakovsky, A. G. (2009). Optimality properties of the Shiryaev-Roberts procedure. *Statistica Sinica* 19: 1729-1739.
- Polunchenko, A. S., Sokolov, G., Du, W. (2014). An accurate method for determining the pre change Run Length distribution of the Generalized Shiryaev-Roberts detection procedure. *Sequential Analysis* 33: 112-134.
- Roberts, S. W. (1966). A comparison of some control chart procedures. *Technometrics* 8: 411-430.
- Shewhart, W. A. (1926). Quality control charts. *Bell System Technical Journal* 593-603.
- Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. Princeton: Van Nostrand Reinhold.
- Shiryaev, A. N. (1961). The problem of the most rapid detection of a disturbance in a stationary process. *Soviet Mathematics-Doklady* 2: 795-799.
- Shu, L., Jiang, W. (2006). A markov chain model for the adaptive CUSUM control chart. *Journal of Quality Technology* 2: 135-147.

- Sparks, R. S. (2000). CUSUM charts for signalling varying location shifts. *Journal of Quality Technology* 32: 157-171.
- Yashchin, E. (1995). Estimating the current mean of a process subject to abrupt changes. *Technometrics* 37: 311-323.
- Zhang, J., Zou, C., Wang, Z. (2011a). A new chart for detecting the process mean and variability. *Communications in Statistics-Simulation and Computation* 40: 728-743.
- Zhang, J., Zou, C., Wang, Z. (2011b). An adaptive Shiryaev-Roberts procedure for monitoring dispersion. *Computers & Industrial Engineering* 61: 1166-1172.

Table 1

Comparison of SS-ARL between SR and CUSUM charts with different δ

n	μ	SR						CUSUM					
		δ						δ					
		0.25	0.50	0.75	1.00	1.50	2.0	0.25	0.50	0.75	1.00	1.5	2.0
1	0.00	200	200	200	200	200	200	200	200	200	200	200	200
	0.25	33.25	37.50	42.67	48.35	60.20	71.04	37.14	42.23	47.69	53.63	65.83	76.73
	0.50	16.53	16.27	16.94	18.67	23.36	29.19	16.68	17.08	18.43	20.55	26.34	32.98
	0.75	10.93	9.96	9.79	9.91	11.61	14.33	10.52	10.13	9.91	10.69	12.78	15.98
	1.00	8.21	7.11	6.75	6.60	7.09	8.21	7.67	6.95	6.81	6.77	7.52	8.99
	1.50	5.51	4.68	4.21	3.92	3.72	3.91	5.05	4.39	4.03	3.88	3.80	4.08
	2.00	4.21	3.52	3.10	2.81	2.58	2.50	3.81	3.31	2.93	2.74	2.54	2.52
5	0.25	14.21	16.45	20.98	27.21	39.15	44.32	14.81	18.08	23.73	30.46	41.57	45.49
	0.50	6.22	5.68	6.08	7.19	10.55	13.22	6.01	5.79	6.47	7.84	11.48	13.63
	0.75	4.03	3.37	3.21	3.34	4.16	5.11	3.80	3.31	3.26	3.47	4.44	5.27
	1.00	3.05	2.45	2.20	2.14	2.30	2.60	2.83	2.35	2.18	2.16	2.37	2.68
	1.50	2.12	1.66	1.41	1.30	1.25	1.27	1.97	1.57	1.35	1.28	1.25	1.27
	2.00	1.70	1.27	1.09	1.05	1.03	1.03	1.59	1.18	1.07	1.04	1.03	1.03

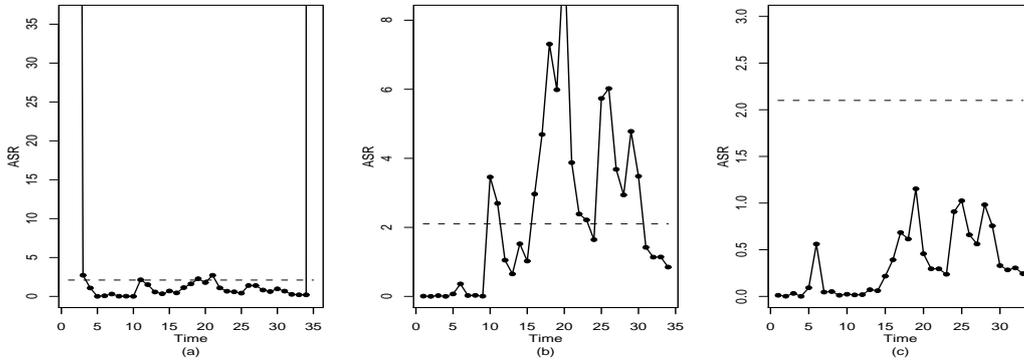


Fig. 1. The ASR chart for the cylinder data with the dashed horizontal line indicating its control limit.

Table 2

Model for the upper control limit as a function of the mean shift of δ for the upper-sided SR charts.

n	ARL	Operating models $h_U(\delta)$ for the SR chart
1	100	$99.158-51.376\delta-7.6842\delta^2+33.541\delta^3-26.256\delta^4+9.4566\delta^5-1.3297\delta^6$
	200	$199.28-109.01\delta+5.6921\delta^2+38.149\delta^3-32.423\delta^4+11.709\delta^5-1.5861\delta^6$
	300	$298.13-158.41\delta-5.4713\delta^2+78.355\delta^3-66.439\delta^4+25.232\delta^5-3.6687\delta^6$
	400	$400.54-233.25\delta+53.839\delta^2+22.583\delta^3-31.800\delta^4+14.061\delta^5-2.2433\delta^6$
	500	$499.54-285.38\delta+63.552\delta^2+11.349\delta^3-11.237\delta^4+1.4082\delta^5+0.3305\delta^6$
5	100	$100.01-126.05\delta+48.962\delta^2+54.341\delta^3-82.945\delta^4+39.084\delta^5-6.2576\delta^6$
	200	$198.04-237.10\delta+64.389\delta^2+133.23\delta^3-158.00\delta^4+64.789\delta^5-9.1826\delta^6$
	300	$296.69-348.09\delta+75.842\delta^2+212.44\delta^3-226.06\delta^4+85.396\delta^5-11.096\delta^6$
	400	$399.12-489.29\delta+179.11\delta^2+159.76\delta^3-199.05\delta^4+72.973\delta^5-8.6252\delta^6$
	500	$498.87-635.25\delta+342.45\delta^2-29.059\delta^3-40.623\delta^4+2.8779\delta^5+3.2956\delta^6$

Table 3

ZS-ARL values of ASR charts for different Q_0 when $\lambda = 0.1, \delta_{\min} = 1.0$ and $n = 1$

μ	Q_0			
	1.0	1.5	2.0	2.5
0.00	400	400	400	400
0.25	79.66	80.02	80.59	81.65
0.50	27.32	27.48	28.29	29.48
0.75	14.19	14.18	14.82	15.79
1.00	9.28	9.09	9.43	10.06
1.25	6.87	6.54	6.59	6.95
1.50	5.42	5.02	4.93	5.09
1.75	4.47	4.04	3.87	3.89
2.00	3.80	3.38	3.17	3.10
h	1.498	1.500	1.505	1.505

Table 4
 SS-ARL performance of ASR charts with different δ_{min} when $\lambda = 0.1, Q_0 = 2.0$

μ	n=1				n=5			
	δ_{min}				δ_{min}			
	0.25	0.5	0.75	1.0	0.25	0.5	0.75	1.0
0.00	400	400	400	400	400	400	400	400
0.25	53.35	53.61	64.58	76.12	17.93	21.77	29.84	40.76
0.50	25.43	21.04	22.13	24.91	7.18	6.72	7.35	8.93
0.75	15.52	12.33	12.10	12.46	4.35	3.81	3.69	3.88
1.00	10.71	8.53	8.04	7.94	3.10	2.68	2.46	2.40
1.25	8.06	6.41	5.97	5.76	2.44	2.08	1.88	1.77
1.50	6.40	5.11	4.72	4.49	2.02	1.73	1.53	1.41
1.75	5.30	4.25	3.90	3.68	1.75	1.47	1.29	1.20
2.00	4.52	3.63	3.34	3.14	1.54	1.27	1.13	1.08
h	65.8	3.8	2.00	1.50	2.99	1.38	1.138	1.05

Table 5
 ARL comparison between ASR, SR, JSA and VD charts when $n = 1$

μ	ASR		SR ₁		SR ₂		JSA		VD	
	ZS	SS	ZS	SS	ZS	SS	ZS	SS	ZS	SS
	0.00	400	400	400	400	400	400	400	400	400
0.25	81.95	78.85	70.02	47.52	122.7	121.7	92.82	88.28	73.17	68.97
0.50	28.10	25.69	37.02	21.93	44.59	43.77	30.52	26.14	27.71	23.66
0.75	14.56	11.08	25.71	14.17	19.73	19.16	14.71	11.28	15.47	12.75
1.00	9.38	7.92	19.98	10.54	10.63	10.22	9.09	6.57	10.34	8.48
1.50	5.13	4.49	14.03	7.03	4.82	4.54	4.89	3.52	5.74	4.82
2.00	3.38	3.05	10.96	5.31	3.03	2.83	3.23	2.44	3.74	3.26