# Limit properties of EWMA charts for stationary processes

**Manuel Cabral Morais, Yarema Okhrin, Wolfgang Schmid**

**Angaben zur Veröffentlichung / Publication details:**

Morais, Manuel Cabral, Yarema Okhrin, and Wolfgang Schmid. 2012. “Limit properties of EWMA charts for stationary processes.” In *Frontiers in Statistical Quality Control 10*, edited by Hans-Joachim Lenz, Wolfgang Schmid, and Peter-Theodor Wilrich, 69–83. Heidelberg: Physica-Verlag. https://doi.org/10.1007/978-3-7908-2846-7_5.

# Limit Properties of EWMA Charts for Stationary Processes

**Manuel Cabral Morais, Yarema Okhrin, and Wolfgang Schmid**

**Abstract** In this paper we consider a general family of EWMA charts for an arbitrary parameter of the target process. Our assumptions on the target process are very weak and they are usually satisfied if it is stationary. We distinguish between the EWMA chart based on the exact variance and the EWMA scheme based on the asymptotic variance. In the case of the EWMA chart with exact variance the in-control variance of the EWMA recursion at time $t$ is used for the decision at time $t$ while in the case of the asymptotic variance at each time point the limit of the in-control variance of the EWMA chart for $t$ tending to infinity is applied. It is analyzed how the distributions of the corresponding run lengths behave if the smoothing parameter tends to zero. We show that the distribution of the run length of the EWMA chart based on the exact variance converges to the distribution of the run length of the repeated significance test while the limit of the EWMA scheme based on the asymptotic variance is degenerate. It is either 0 or 1. This result underlines the weakness of the schemes based on the asymptotic variance if the smoothing parameter is small. Moreover, several properties of the limit chart, i.e. the chart based on the repeated significance test, are presented as well.

**Keywords** EWMA chart • Stationary process • Statistical process control • Run length

---

M.C. Morais (✉)
Department of Mathematicss and CEMAT Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001, Lisboa, Portugal
e-mail: maj@math.ist.utl.pt

Y. Okhrin
Department of Statistics, University of Augsburg, Universitätsstr. 16, 86159, Augsburg, Germany
e-mail: yarema.okhrin@wiwi.uni-augsburg.de

W. Schmid
Department of Statistics, European University Viadrina, Grosse Scharrnstrasse 59, 15230, Frankfurt (Oder), Germany
e-mail: schmid@euv-frankfurt-o.de

# 1 Introduction

Since the introduction of the exponentially weighted moving average (EWMA) chart by Roberts (1959) this scheme has become one of the most discussed control charts in the literature on statistical process control. Because its decision rule is quite simple it has shown to be an attractive control chart for practitioners. Similar to the CUSUM chart of Page (1954) it depends on a further parameter, the smoothing parameter, which regulates the influence of past observations on the present decision. If the smoothing parameter is equal to 1 the chart coincides with the Shewhart chart (cf. Shewhart 1931). The smaller the value of the smoothing parameter the larger is the influence of the preceding observations. Recommendations about the choice of the smoothing parameter were given by several authors. Montgomery (2009) writes that values of $\lambda$ in the interval $0.05 \leq \lambda \leq 0.25$ work well in practice. The aim of Lucas and Saccucci (1990) was to find an optimal design procedure. Fixing the desired in-control average run length (ARL) of the EWMA chart they determined the value of the smoothing parameter which minimizes the out-of-control ARL for a given value of the expected shift. In their paper the minimum is determined for values of the smoothing parameter lying within the interval (0.03, 1.0). One reason why a lower bound for the smoothing parameter was chosen probably lies in the fact that the calculation of the average run length of the EWMA chart turns out to be quite complicate (see, e.g., Brook and Evans 1972; Crowder 1987) and numerically instable if the smoothing parameter is very small.

Most of the literature on EWMA charts is concerned with the monitoring of the mean of an independent random sample. A further great advantage of the EWMA approach consists in the fact that it can be easily extended to monitor other parameters of a process and that the structure of the underlying process can be quite general. EWMA charts for the standard deviation of an independent random process have been introduced by Crowder and Hamilton (1992). The extension of the EWMA chart to monitor the mean of a stationary time series process was given by Schmid (1997). EWMA charts for the variance of a stationary process were proposed in Schipper and Schmid (2001) while Rosołowski and Schmid (2003) discussed simultaneous EWMA schemes for the mean, the variances, and the autocovariances. These examples describe only a few applications of EWMA charts.

Many researchers followed the proposal of Lucas and Saccucci (1990) about the determination of the optimal smoothing parameter and they obtained it by minimizing the out-of-control ARL for a fixed in-control ARL. Due to the improvement of the computer power in the last years it is nowadays possible to calculate the ARL for smaller bounds than 0.03 as well. Doing this something interesting happens. It has turned in some cases that the smaller the lower bound the smaller the ARL and thus the optimal smoothing parameter is equal to the lower bound. This problem was described by Chan and Zhang (2000) and it was analyzed in more detail by Frisén and Sonesson (2006); both papers focus on the mean chart and independent samples.

A deeper analysis of the convergence of the distribution of the run length of the one-sided EWMA chart for the mean was provided by Morais et al. (2010). They proved that for a fixed control limit the in-control ARL of the EWMA scheme based on the asymptotic variance is a decreasing function in the smoothing parameter. Moreover, they analyzed the limit of the distribution of the run length as the smoothing parameter turns to zero. The resulting limit chart turns out to be equal to the repeated significance test. These authors also derived several properties of the limit chart.

In the present paper we consider a very general family of EWMA schemes which can be used to monitor an arbitrary real-valued parameter. It is assumed that the statistics to which the EWMA recursion is applied are governed by a stationary process. Because of its generality this approach covers most of the EWMA schemes discussed in literature. We distinguish between two types of decision rules. The first rule is based on a comparison of the deviation of the EWMA statistic and the target parameter with the in-control variance of the EWMA recursion. Because at each time point the variance changes it must be calculated in each step. This procedure is called the EWMA chart based on the exact variance has also been considered by Morais et al. (2010) as well. Because in each step the variance must be calculated most practitioners prefer to work with the asymptotic variance. In that case the distance between the EWMA recursion and the target value is compared with the asymptotic in-control variance of the EWMA recursion. This is our second procedure. It turns out to be much simpler and is used in most papers on EWMA charts.

In Section 2 we introduce the EWMA charts based on the asymptotic and exact variance. Moreover, we analyze the one-sided and the two-sided monitoring problem.

Section 3 deals with the limit distributions of the run lengths of the EWMA recursions. It is shown that the run lengths of the charts based on the exact variance converge to the distribution of the run length of the repeated significance test. This result holds for the in-control and out-of-control case as well. As special cases we consider the monitoring of the mean and the variance of a stationary process. In the second part of this section the limit of the charts based on the asymptotic variance is determined. It is shown to be degenerate, i.e. it is either 0 or 1. This is a clear hint that the charts based on the asymptotic variance should not be used with an extremely small smoothing parameter. In the most realistic case the probability of a signal up to a fixed time point converges to 1 as well in the in-control state as in the out-of-control state. This is a very unpleasant property. Moreover, it shows that the ARL converges to infinity if the smoothing parameter tends to zero.

Section 4 is devoted to the analysis of the limit scheme, i.e. the chart using the repeated significance test. Here we restrict ourselves to independent samples and a one-sided chart. Several properties of the limit scheme are presented which question the usefulness of the ARL as a performance measure for EWMA schemes.

In Section 5 we summarize the main conclusions of our paper.

## 2 EWMA Charts for Stationary Processes

In what follows $\{Y_t\}$ denotes the target process and $\{X_t\}$ stands for the observed process. Both processes are assumed to be the same up to time point $q-1$, i.e. $X_t = Y_t$ for $t = 1, \ldots, q-1$. At the time point $q$ it is assumed that $X_q \neq Y_q$. The unknown position $q$ is assumed to be a deterministic quantity taking values within the set $\mathbb{N} \cup \{\infty\}$. If $q < \infty$ then a sustained shift has occurred at time $t = q$. The process $\{X_t\}$ is called to be out-of-control. Needless to say, $\{X_t\}$ is said to be in-control if $q = \infty$.

We are interested to monitor a parameter $\theta$ of the target process by using the statistic $T_t$, which we assume is a function of past and present values of the observed process, i.e. $T_t = f_t(X_1, \ldots, X_t)$. The statistic $T_t$ can be interpreted as a point estimator for the parameter $\theta$. Suppose that $T_t$ is an unbiased estimator of $\theta$ in the in-control state. If $\theta$ is equal to the mean $\mu$ of the $\{Y_t\}$ then we can choose, e.g., $T_t = X_t$ (see Schmid 1997), $T_t = \sum_{v=1}^{k} X_{t+1-v}/k$ or $T_t = c_1 med\{X_{t+1-k}, .., X_t\}$. In case $\theta$ is the variance of $\{Y_t\}$ possible choices would be, e.g., $T_t = (X_t - \mu)^2$ (e.g., Schipper and Schmid 2001), $T_t = \sum_{v=1}^{k}(X_{t+1-v} - \mu)^2/k$ or $T_t = c_2 \sum_{v=1}^{k} |X_{t+1-v} - \mu|/k$.

The EWMA recursion applied to $T_t$ is given by

$$Z_t = \begin{cases} Z_0, & t = 0 \\ (1-\lambda)Z_{t-1} + \lambda T_t, & t = 1, 2, \ldots, \end{cases} \tag{1}$$

with an initial value $Z_0$. The parameter $\lambda$ is a smoothing parameter taking values within the set $(0, 1]$ and it corresponds to the weight given to the most recent observed value. Having in mind that $Z_t$ can be equivalently written as the following moving average

$$Z_t = \lambda \sum_{i=0}^{t-1} (1-\lambda)^i T_{t-i} + (1-\lambda)^t Z_0, \quad t = 1, 2, \ldots, \tag{2}$$

whose weights fall off geometrically, we immediately conclude that a value of $\lambda$ close to one leads to a short memory EWMA chart – in fact $\lambda = 1$ leads to nothing but a Shewhart chart –, whereas values of $\lambda$ close to zero lead to EWMA charts that give little importance to the most recent observations.

Since $T_t$ is an unbiased estimator of $\theta$, we can assert that in the in-control state,

$$E_\infty(Z_t) = \theta + (1-\lambda)^t [Z_0 - \theta] \, .$$

The index "$\infty$" means, throughout the remainder of this paper, that the quantity (an expectation, a variance, a covariance, a probability, etc.) is calculated with respect to the in-control situation.

Now suppose that in the in-control state $\{T_t\}$ is a (weakly) stationary process with mean $\theta$ and autocovariance function $\{\gamma_{T,h}\}$. Then

$$\begin{aligned}\mathrm{Var}_\infty(Z_t) &= \lambda^2 \sum_{i,j=0}^{t-1} (1-\lambda)^{i+j} \gamma_{T,|i-j|} \qquad (3)\\ &= \frac{\lambda}{2-\lambda}\left[[1-(1-\lambda)^{2t}]\gamma_{T,0} + 2\sum_{v=1}^{t-1}(1-\lambda)^v[1-(1-\lambda)^{2(t-v)}]\gamma_{T,v}\right].\end{aligned}$$

Upward shifts in $\theta$ can be detected by upper one-sided EWMA charts which give a signal at the sampling period $t \geq 1$, suggesting that the parameter $\theta$ increased, if

$$Z_t > \theta + c\sqrt{\mathrm{Var}_\infty(Z_t)},$$

for some fixed constant critical value $c$ that defines the range of these exact control limits. Note that we make use of the asymptotic mean $\theta$ instead of the exact one $E_\infty(Z_t)$. This signal is a valid one, in case the process is out-of-control, and it is called a false alarm, otherwise.

In order to detect an upward or a downward shift, we have to make use of two-sided EWMA charts which trigger a signal whenever

$$|Z_t - \theta| > c\sqrt{\mathrm{Var}_\infty(Z_t)},$$

with $c > 0$.

To address both types of EWMA charts, we define $\mathcal{R}$ as the rejection area and thus a signal is given if

$$\frac{Z_t - \theta}{\sqrt{\mathrm{Var}_\infty(Z_t)}} \in \mathcal{R},$$

where $\mathcal{R} = (c, \infty)$ in the upper one-sided case, and $\mathcal{R} = (-\infty, -c) \cup (c, \infty)$ in the two-sided case.

In practice the asymptotic variance is frequently used instead of the exact one. In that case a signal is given if

$$\frac{Z_t - \theta}{\sqrt{\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t)}} \in \mathcal{R}\ .$$

Note that

$$\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t) = \lambda^2 \sum_{i,j=0}^{\infty} (1-\lambda)^{i+j} \gamma_{T,|i-j|}$$

if $\{T_t\}$ has an absolutely summable autocovariance function $\{\gamma_{T,h}\}$. Another representation of the asymptotic variance based on the spectral density function is given in Schmid (1997).

In what follows, $\theta$, $\text{Var}_\infty(Z_t)$, and $\lim_{t\to\infty} \text{Var}_\infty(Z_t)$ are assumed to be known quantities. We shall not discuss the influence of parameter estimation.

# 3 Limit Behavior for Stationary Processes

Let $N(\lambda, \mathcal{R})$ denote the run length of control chart based on the exact variance, i.e.

$$N(\lambda, \mathcal{R}) = \inf\left\{ t \in \mathbb{N} : \frac{Z_t - \theta}{\sqrt{\text{Var}_\infty(Z_t)}} \in \mathcal{R} \right\},$$

and $N_{asymp}(\lambda, \mathcal{R})$ be the run length of the scheme with the asymptotic variance

$$N_a(\lambda, \mathcal{R}) = \inf\left\{ t \in \mathbb{N} : \frac{Z_t - \theta}{\sqrt{\lim_{t\to\infty} \text{Var}_\infty(Z_t)}} \in \mathcal{R} \right\}.$$

We denote the run length of the repeated significance test by

$$N(\mathcal{R}) = \inf\left\{ t \in \mathbb{N} : \frac{\bar{T}_t - \theta}{\sqrt{\text{Var}_\infty(\bar{T}_t)}} \in \mathcal{R} \right\},$$

where $\bar{T}_t = \sum_{v=1}^{t} T_v/t$. It is worth mentioning that

$$t\,\text{Var}_\infty(\bar{T}_t) = \gamma_{T,0} + 2\sum_{v=1}^{t-1}(1 - v/t)\gamma_{T,v}, \tag{4}$$

whose limit, when $t$ tends to infinity, exists as long as $\sum_{v=1}^{\infty} |\gamma_{T,v}| < \infty$.

## *3.1 The EWMA Scheme Based on the Exact Variance*

First, we consider the probability of getting no signal up to a fixed time point for the present EWMA scheme if $\lambda$ tends to zero. It is shown that the limit is equal to the probability of no false signal for the repeated significance test, i.e. the sequential application of the corresponding significance test using the critical value $c$.

**Theorem 1.** *Assume that the k-dimensional random vector $(X_1, \ldots, X_k)$ is continuous or discrete and that its distribution does not depend on $\lambda$. If $Z_0 = \theta$ then*

$$\lim_{\lambda\to 0+} P[N(\lambda, \mathcal{R}) > k] = P[N(\mathcal{R}) > k],$$

*for any fixed $k = 1, 2, \ldots$ and any set $\mathcal{R} \subset I\!R$ which does not depend on $\lambda$.*

*Proof.* The run length of the EWMA scheme with exact variance exceeds $k$ with probability

$$P[N(\lambda,\mathcal{R})>k] = P\left[\frac{Z_t-\theta}{\sqrt{\mathrm{Var}_\infty(Z_t)}} \notin \mathcal{R} \quad \forall\, t=1,\ldots,k\right]$$
$$= \int_{A(\lambda)} \cdots \int f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)dx_1\ldots dx_k,$$

where $A(\lambda)=\cap_{t=1}^{k} A_t(\lambda)$ with

$$A_t(\lambda)=\left\{(x_1,\ldots,x_t):\frac{Z_t-\theta}{\sqrt{\mathrm{Var}_\infty(Z_t)}}\notin\mathcal{R}\right\}.$$

Because

$$\frac{Z_t-\theta}{\sqrt{\mathrm{Var}_\infty(Z_t)}}=\frac{\sum_{i=0}^{t-1}(1-\lambda)^i(T_{t-i}-\theta)}{\sqrt{\mathrm{Var}_\infty(Z_t)/\lambda}} \tag{5}$$

and

$$\lim_{\lambda\to 0+}\frac{\mathrm{Var}_\infty(Z_t)}{\lambda^2}$$
$$=\frac{1}{2}\lim_{\lambda\to 0+}\frac{\left[[1-(1-\lambda)^{2t}]\gamma_{T,0}+2\sum_{v=1}^{t-1}(1-\lambda)^v[1-(1-\lambda)^{2(t-v)}]\gamma_{T,v}\right]}{\lambda}$$
$$=t^2\mathrm{Var}_\infty(\bar{T}_t),$$

the use of Eq. 4 leads to

$$\lim_{\lambda\to 0+}A_t(\lambda)=\left\{(x_1,..,x_t):\frac{\bar{T}_t(x_1,\ldots,x_t)-\theta}{\sqrt{\mathrm{Var}_\infty(\bar{T}_t)}}\notin\mathcal{R}\right\}$$

and

$$\lim_{\lambda\to 0+}\bigcap_{t=1}^{k}A_t(\lambda)=\bigcap_{t=1}^{k}\left\{(x_1,..,x_t):\frac{\bar{T}_t(x_1,\ldots,x_t)-\theta}{\sqrt{\mathrm{Var}_\infty(\bar{T}_t)}}\notin\mathcal{R}\right\}. \tag{6}$$

Moreover, if we define the set on the right side of Eq. 6 by $A$, then

$$\lim_{\lambda\to 0+}\int_{A(\lambda)}\cdots\int f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)dx_1\ldots dx_k$$
$$=\int_A\cdots\int f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)dx_1\ldots dx_k,$$

thus proving the result.

Remarkably, the theorem holds both in the in-control and in the out-of-control states. In addition, if we combine the fact that $E[N] = \sum_{k=0}^{\infty} P[N > k]$ and Theorem 1, we can conclude that

$$\lim_{\lambda \to 0+} E[N(\lambda, \mathcal{R})] = E[N(\mathcal{R})].$$

This means that if $\lambda$ converges to zero then the ARL of the EWMA chart for monitoring $\theta$ converges to the ARL of the repeated significance test. We ought to mention that in some cases $E[N(\mathcal{R})] = \infty$; this point is analyzed in more detail in the next section.

Note that the above result is quite general and valid for any EWMA chart whose input statistics $\{T_t\}$ are governed by a stationary process.

*Example 1.* (a) Monitoring the mean of a stationary process

In this case we have $\theta = \mu$. Following Schmid (1997) we choose $T_t = X_t$. If $\{Y_t\}$ is stationary then in the in-control state $\{T_t\}$ is stationary as well. In the special case that $\{Y_t\}$ is a causal ARMA process (cf. Brockwell and Davis 1991) the autocovariances of $\{Y_t\}$ can be determined recursively by making use of the Yule-Walker equations (e.g., Brockwell and Davis 1991, Chap. 3). If $\{Y_t\}$ is a stationary GARCH process (cf. Tsay 2005, Chap. 3) then the determination of the variance of the EWMA recursion is easier and it holds that

$$\mathrm{Var}_{\infty}(Z_t) = \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2t}]\gamma_{T,0}$$

(b) Monitoring the variance of a stationary process

Suppose that $\mu = 0$, $\theta = \gamma_0$, $T_t = X_t^2$ (e.g., Schipper and Schmid 2001) and $Y_t = \sum_{i=-\infty}^{\infty} a_i \varepsilon_{t-i}$, where $\{a_i\}$ is absolutely summable. Let $\{\varepsilon_t\}$ be independent and normally distributed with $E(\varepsilon_t) = 0$ and $\mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ then $\gamma_{T,h} = \gamma_0^2 + 2\gamma_h^2$ (cf. Brockwell and Davis 1991, p. 227), where $\gamma_h$ stands for the autocovariance function of $\{Y_t\}$. Thus, we get

$$\begin{aligned}\mathrm{Var}_{\infty}(Z_t) = \frac{\lambda}{2-\lambda}\Bigg[\gamma_0^2\Bigg[&[3(1-(1-\lambda)^{2t}]\\ &+\frac{2(1-\lambda)}{\lambda}[1-(1-\lambda)^t][1-(1-\lambda)^{t-1}]\Bigg]\\ &+2\sum_{i=1}^{t-1}(1-\lambda)^i[1-(1-\lambda)^{2(t-i)}]\gamma_i^2\Bigg].\end{aligned}$$

This quantity can be calculated recursively for a stationary ARMA process as described in (a). For the special case of an ARMA(1,1) process $Y_t = \alpha Y_{t-1} + \varepsilon_t + \beta\varepsilon_{t-1}$, we get

$$\begin{aligned}\mathrm{Var}_{\infty}(Z_t) = \frac{\lambda}{2-\lambda}\Bigg[\gamma_0^2\Bigg[&[3(1-(1-\lambda)^{2t}]\\ &+\frac{2(1-\lambda)}{\lambda}[1-(1-\lambda)^t][1-(1-\lambda)^{t-1}]\Bigg]\\ &+2(1-\lambda)\gamma_1^2\left[\frac{1-[\alpha^2(1-\lambda)]^{t-1}}{1-\alpha^2(1-\lambda)}-(1-\lambda)^t\frac{(1-\lambda)^{t-1}-\alpha^{2(t-1)}}{1-\lambda-\alpha^2}\right]\Bigg].\end{aligned}$$

since $\gamma_i = \alpha^{i-1}\gamma_1$, for $i \geq 1$, with

$$\gamma_0 = \sigma_\varepsilon^2\frac{1+2\alpha\beta+\beta^2}{1-\alpha^2} \quad \text{and} \quad \gamma_1 = \sigma_\varepsilon^2\frac{(1+\alpha\beta)(\alpha+\beta)}{1-\alpha^2}.$$

Note that a stronger result for the mean chart for independent normal variables was shown by Morais et al. (2010). They proved that in the in-control state the probability of a false signal is a decreasing function in $\lambda \in (0, 1]$. This result was obtained by using monotonicity results for the multivariate normal distribution (cf. Tong 1990).

## 3.2 *The EWMA Scheme Based on the Asymptotic Variance*

Next we analyze the run length of the scheme based on the asymptotic variance in the in-control state.

**Theorem 2.** *Assume that the k-dimensional random vector $(X_1, \ldots, X_k)$ is continuous or discrete and that its distribution does not depend on $\lambda$. Let $\mathcal{R}$ be an arbitrary subset of IR which does not depend on $\lambda$ and suppose that $Z_0 = \theta$.*

*(a) If $0 \notin \mathcal{R}$ then*

$$\lim_{\lambda\to 0+} P[N_{asymp}(\lambda, \mathcal{R}) > k] = 1,\ k \in \mathbb{N}.$$

*(b) If $0 \in \mathcal{R}$ then*

$$\lim_{\lambda\to 0+} P[N_{asymp}(\lambda, \mathcal{R}) > k] = 0,\ k \in \mathbb{N}.$$

*Proof.* We only prove part (a). The proof of part (b) follows immediately.

First, note that

$$\begin{aligned}P[N_{asymp}(\lambda, \mathcal{R}) > k] &= P\left[\frac{Z_t-\theta}{\sqrt{\lim_{t\to\infty}\mathrm{Var}_{\infty}(Z_t)}} \notin \mathcal{R}, \quad \forall\, t = 1, \ldots, k\right]\\ &= \int_{A_a(\lambda)}\cdots\int f_{X_1,\ldots,X_k}(x_1, \ldots, x_k)dx_1\ldots dx_k,\end{aligned}$$

where $A_a(\lambda) = \cap_{t=1}^{k} A_{a,t}(\lambda)$ with

$$A_{a,t}(\lambda) = \{(x_1, \ldots, x_t) : \frac{Z_t - \theta}{\sqrt{\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t)}} \notin \mathcal{R}\}.$$

Because

$$\frac{Z_t - \theta}{\sqrt{\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t)}} = \frac{\sum_{i=0}^{t-1}(1-\lambda)^i (T_{t-i} - \theta)}{\sqrt{\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t)/\lambda}} \tag{7}$$

and

$$\lim_{\lambda\to 0+} \frac{\lim_{t\to\infty} \mathrm{Var}_\infty(Z_t)}{\lambda^2} = \frac{1}{2} \lim_{\lambda\to 0+} \frac{\left[\gamma_{T,0} + 2\sum_{v=1}^{\infty}(1-\lambda)^v \gamma_{T,v}\right]}{\lambda} = \infty,$$

we successively get, for case (a),

$$\lim_{\lambda\to 0+} A_t(\lambda) = \Omega$$
$$\lim_{\lambda\to 0+} \bigcap_{t=1}^{k} A_t(\lambda) = \Omega,$$

and

$$\lim_{\lambda\to 0+} \int_{A(\lambda)} \cdots \int f_{X_1,\ldots,X_k}(x_1, \ldots, x_k) dx_1 \ldots dx_k$$
$$= \int_{\Omega} \cdots \int f_{X_1,\ldots,X_k}(x_1, \ldots, x_k) dx_1 \ldots dx_k = 1,$$

thus proving the result for case (a).

Note that case (a) is the one we usually deal with. It arises if we choose $c > 0$ and take $\mathcal{R} = (c, \infty)$, in the one-sided case, or consider $\mathcal{R} = (-\infty, -c) \cup (c, \infty)$, in the two-sided case. Choosing $c < 0$ means that $0 \in \mathcal{R}$ and that we are dealing with a negative control limit, in the one-sided case, which is rather strange to most practitioners. However, we shall address this case later on.

The result of part (a) is highly desirable in the in-control state because it means that the probability of a false alarm converges to zero as $\lambda$ tends to zero. However, the result also holds in the out-of-control state, that is, the probability of a signal within the first $k$ samples converges to 1. As a consequence the EWMA scheme with asymptotic variance behaves quite chaotically if $\lambda$ reaches 0. It also implies that $\lim_{\lambda\to 0+} E[N_{asymp}(\lambda, c)] = \infty$, therefore the behavior of the EWMA control charts with different values of $\lambda$ cannot be compared by means of the average run length. Let us remind the reader that the in-control ARL is used to determine the control limits but in the present case this cannot be done since this quantity does not exist if $\lambda$ converges to zero.

For case (b) the scheme shows similar undesirable properties. For instance, the probability of a correct signal converges to 0 in the in-control state when $\lambda$ converges to zero.

In sum, if $\lambda$ takes values close to zero the control chart based on the exact variance must be favored: it has at least a reasonable limit behavior – the repeated significance test. For a better understanding of this limit behavior, it is necessary, however, to analyze the properties of the repeated significance test.

## 4 Some Properties of the Limit Chart

In this section we present some results of the limit chart. We focus on the detection of a change in the mean and choose $T_t = X_t$. Moreover, we restrict ourselves to the one-sided problem, i.e. we choose $\mathcal{R} = (c, \infty)$. Then the limit chart, i.e. the repeated significance test, has run length given by

$$N(c) = \inf\{t \in \mathbb{N} : \frac{\sum_{i=1}^{t}(X_i - \mu)}{\sqrt{t[\gamma_0 + 2\sum_{i=1}^{t-1}(1 - i/t)\gamma_i]}} > c\}.$$

In order to illustrate how the limit scheme may behave, we focus on the case of independent random variables. A detailed analysis of the repeated significance test is given in Morais et al. (2010). In that case we have that

$$N(c) = \inf\{t \in \mathbb{N} : \frac{\sum_{i=1}^{t}(X_i - \mu)}{\sqrt{t\gamma_0}} > c\}.$$

We make use of the following change point model

$$X_t = \begin{cases} Y_t & \text{for } t < q \\ Y_t + a\sqrt{\gamma_0} & \text{for } t \geq q. \end{cases} \tag{8}$$

In what follows we use the symbols $P_{a,q}$, $E_{a,q}$, etc. to denote a probability, expectation, etc. taken with respect to model Eq. 8.

First, we discuss the behavior of the limit scheme in the in-control state.

**Theorem 3.** *Assume that the random variables $\{Y_t\}$ are independent and identically distributed with mean $\mu$ and variance $\gamma_0$.*

*(a) If $c \geq 0$ then $E_\infty[N(c)] = \infty$.*

**Table 1** The fraction (in %) of the run lengths falling into the respective interval as a function of $c$. The computations are based on $10^9$ replications of iid Gaussian observations

| $c$ | $E[N(c)]$ | $N(c) \leq 10$ (in %) | $10 < N(c) \leq 10^4$ (in %) | $10^4 < N(c) \leq 10^7$ (in %) | $10^7 < N(c)$ (in %) |
|---|---|---|---|---|---|
| −1.5 | 1.266 | 99.64 | 0.36 | 3.84e−05 | 0 |
| −1.4 | 1.406 | 99.50 | 0.51 | 8.97e−05 | 0 |
| −1.3 | 1.614 | 99.28 | 0.71 | 2.05e−04 | 1e−07 |
| −1.2 | 2.078 | 99.02 | 0.98 | 4.35e−04 | 7e−07 |
| −1.1 | 2.707 | 98.66 | 1.34 | 8.74e−04 | 4e−07 |
| −1.0 | 4.342 | 98.20 | 1.80 | 1.83e−03 | 1.5e−06 |
| −0.9 | 8.284 | 97.61 | 2.38 | 3.62e−03 | 6.7e−06 |
| −0.8 | 21.81 | 96.87 | 3.12 | 7.05e−03 | 1.53e−05 |
| −0.7 | 59.14 | 95.95 | 4.03 | 1.34e−02 | 4.27e−05 |
| −0.6 | 140.9 | 94.84 | 5.15 | 2.47e−02 | 1.16e−04 |

(b) *Suppose that $P[Y_t = 0] < 1$ and that the variables $\{Y_t\}$ are symmetric around $\mu$. If $c < 0$ then $E_\infty[N(c)] < \infty$ and $Var_\infty[N(c)] < \infty$.*

This result is remarkable. It implies that the in-control ARL of the limit chart is equal to infinity if the control limit is nonnegative. Thus, together with the results of Section 3, it follows that $\lim_{\lambda \to 0+} E_\infty[N(\lambda, c)] = \infty$. Note that practitioners choose a positive control limit and thus the in-control ARL is not finite in the most popular case. Theorem 3 also has an important consequence on the comparison of EWMA charts because it states that EWMA charts should not be compared by means of the average run length, at least when $\lambda$ is very small.

Table 1 illustrates the divergence of the run length if the critical value tends to zero from the left. Despite of a small in-control ARL, there is a substantial number of run lengths exceeding $10^7$. It is assumed for simplicity that the observations are iid and follow Gaussian distribution.

Now let us assume that the control limit $c$ is chosen as a solution of $E_\infty[N(\lambda, c)] = \xi$, with $\xi > 1$, denote $c = c(\lambda, \xi)$ and discuss the behavior of $c(\lambda, \xi)$ as $\lambda$ tends to 0.

**Theorem 4.** *Assume that the random variables $\{Y_t\}$ are independent and identically distributed to $\mathcal{N}(\mu, \gamma_0)$. Then* $\lim_{\lambda \to 0+} c(\lambda, \xi) < 0$.

The control limit of the upper one-sided EWMA chart is, by definition, positive. Theorem 4 shows, however, that there is no positive solution for the limit chart. Thus, it is impossible to compare the behavior of the EWMA chart with that of the limit chart if the control limit is chosen to be equal to a specified constant. Figure 1 shows the behaviour of the control limit $c$ with $\xi = 50$ both for the asymptotic and the exact variances as a function of $\lambda$.

Now we investigate the out-of-control behavior of the limit scheme. We ought to begin by noting that contrary to the in-control ARL, the out-of-control ARL is always finite, as stated in the next theorem.

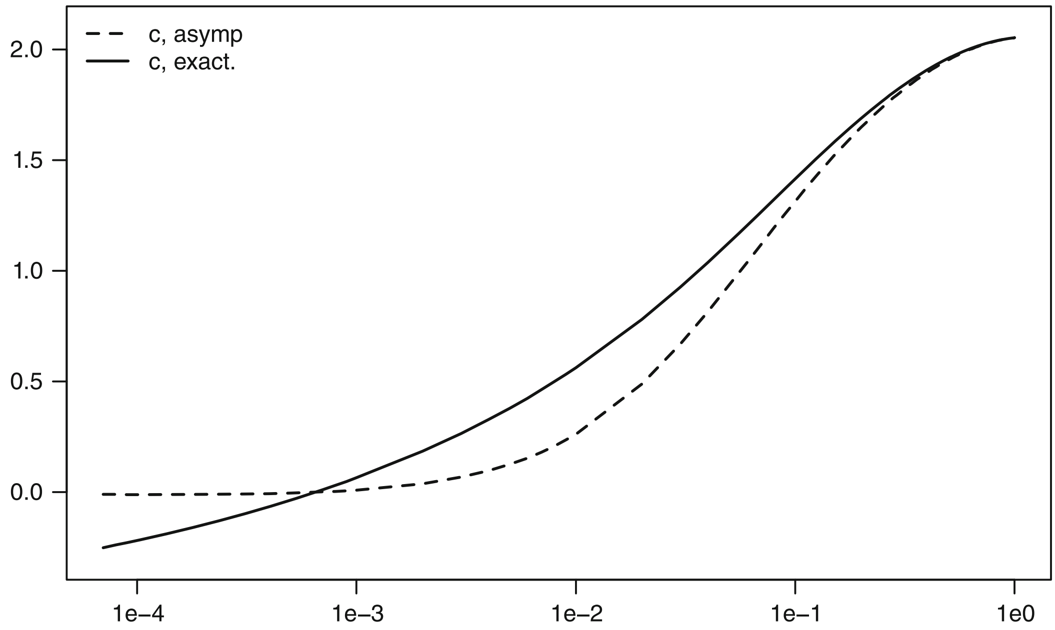


**Fig. 1** The critical value $c$ as a function of $\lambda$ for the EWMA schemes with exact and asymptotic variance, for $\xi = 50$ and iid Gaussian observations. The results are based on a Monte-Carlo study with $10^9$ replications

**Theorem 5.** *Assume that the random variables* $\{Y_t\}$ *are independent and identically distributed to* $\mathcal{N}(\mu, \gamma_0)$. *Then* $E_{a,q}[N(c)] < \infty$, *for all* $c \in IR$, $q \in \mathbb{N}$, *and shifts with magnitude* $a > 0$.

Figure 2 illustrates this result. In fact, it shows that the out-of-control ARL converges to a finite value as $\lambda \to 0+$ for several upward shifts with magnitude $a$.

We shall now discuss the average delay of the limit scheme, another frequently used performance measure. Recall that, when dealing with control chart with run length $N$, the average delay is given by

$$E_{a,q}(N(c) - q + 1|N(c) \geq q) = \sum_{k=q}^{\infty} \frac{P_{a,q}[N(c) \geq k]}{P_{a,q}[N(c) \geq q]}.$$

Note that Theorem 5 allows us to assert that the average delay exists for $q < \infty$ and $a > 0$. The next theorem refers to the behavior of the average delay when the magnitude of the shift, $a$, tends to zero.

**Theorem 6.** *Assume that the random variables* $\{Y_t\}$ *are independent and identically distributed to* $\mathcal{N}(\mu, \gamma_0)$ *and let* $q \in \mathbb{N}$.

*(a)* *If* $c > 0$ *then* $\lim_{a \to 0+} E_{a,q}[N(c) - q + 1|N(c) \geq q] = \infty$.
*(b)* *If* $c < 0$ *then* $\lim_{a \to 0+} E_{a,q}[N(c) - q + 1|N(c) \geq q] < \infty$.

Part (a) from Theorem 6 reads as follows: the average delay of the limit chart has an undesirable behavior when a positive control limit is at use. Moreover, since this

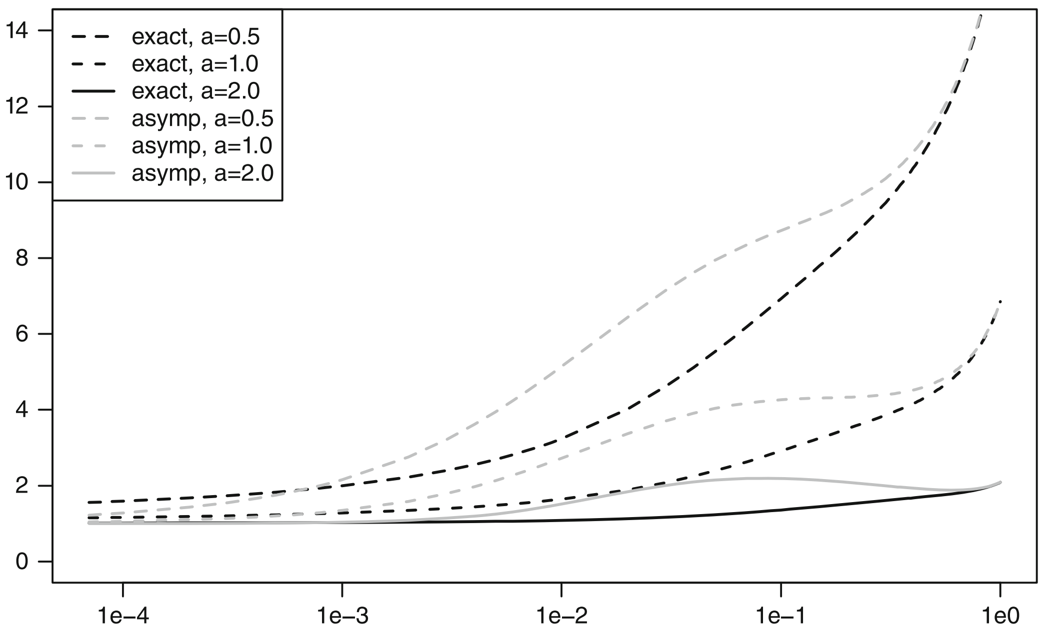


**Fig. 2** The out-of-control ARL as a function of $\lambda$ for the EWMA schemes based on the exact and the asymptotic variance, for $\xi = 50$ and iid Gaussian observations. The results are based on a Monte-Carlo study with $10^9$ replications

result implies that the EWMA scheme behaves similar as $\lambda$ tends to zero, we can conclude that all criteria based on the first moment of the run length are not suitable to assess or compare the performance EWMA charts with small values of $\lambda$.

## 5 Concluding Remarks

This paper essentially provides a thorough study on the behaviour of the run length of EWMA charts with exact and asymptotic control limits when $\lambda$ converges to zero. We ought to stress that the results are quite general and refer to the control of any parameter of a stationary process. For instance, we proved that, when the smoothing parameter $\lambda$ tends to zero:

- The run length of EWMA charts based on the exact variance has the same behavior as the run length of a chart based on a repeated significance test, what we called the limit chart;
- The out-of-control run length of EWMA charts based on the asymptotic variance is infinite if the rejection area $\mathcal{R}$ includes the origin.

Finally, this study also brought to light a few useful results concerning the run length of the limit chart, namely for the mean independent and identically distributed processes:

- Its in-control ARL is not finite if the control limit $c$ in nonnegative.

In summary, our results permit the following conclusions:

1. The EWMA scheme with exact variance must be preferred over the EWMA scheme with asymptotic variance.
2. The EWMA scheme with asymptotic variance should not be applied if the smoothing parameter is small (about $\lambda \leq 0.1$).
3. The average run length is no suitable performance measure for EWMA charts if the smoothing parameter is very small.
4. One-sided EWMA charts should make use of a reflecting boundary.

## References


Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods*. New York: Springer.

Brook, D., & Evans, D. A. (1972). An approach to the probability distributions of CUSUM run length. *Biometrika, 3*, 539–549.

Chan, L. K., & Zhang, J. (2000). Some issues in the design of EWMA charts. *Communications in Statistics – Simulation and Computation, 29*, 207–217.

Crowder, S. V. (1987). A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics, 29*, 401–407.

Crowder, S. V., & Hamilton, M. D. (1992). EWMA for monitoring a process standard deviation. *Journal of Quality Technology, 24*, 12–21.

Frisén, M., & Sonesson, C. (2006). Optimal surveillance based on exonentially weighted moving averages. *Sequential Analysis, 25*, 379–403.

Lucas, J. M., & Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics, 32*, 1–12.

Montgomery, D. C. (2009). *Introduction to statistical quality control* (6th ed.). New York: Wiley.

Morais, M. C., Okhrin, Y., & Schmid, W. (2009). *On the limiting behaviour of EWMA charts with exact control limits*. Discussion Paper 272 of the Faculty of Business Administration and Economics, EUV, Frankfurt (Oder).

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

Roberts, S.W. (1959). Control charts tests based on geometric moving averages. *Technometrics, 1*, 239–250.

Rosołowski, M., & Schmid, W. (2003). EWMA charts for monitoring the mean and the autocovariances of stationary Gaussian processes. *Sequential Analysis, 22*, 257–285.

Schmid, W. (1997). On EWMA charts for time series. In H. J. Lenz & P.-Th. Wilrich (Eds.), *Frontiers of statistical quality control* (Vol. 5, pp. 115–137). Heidelberg: Physica.

Schipper, S., & Schmid, W. (2001). Sequential methods for detecting changes in the variance of economic time series. *Sequential Analysis, 20*, 235–262.

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. Princeton: Van Notrand.

Tong, Y. L. (1990). *The multivariate normal distribution*. New York: Springer.

Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). Hoboken, NJ: Wiley.