Brazilian Journal of Probability and Statistics (2004), 18, pp. 151–162.

Conditional maximum likelihood estimation for control charts in the presence of correlation

Tzong-Ru Tsai, **Yi-Chen Chiang** and **Shuo-Jye Wu**
Tamkang University

Abstract: In practice, the observations are usually autocorrelated. The autocorrelation between successive observations has a large impact on control charts with the assumption of independence. It can decrease the in-control average run length which leads to a higher false alarm rate than in the case of independent process. This paper considers the problem of monitoring the mean of AR(1) process with a random error and provides a conditional maximum likelihood estimation method to improve the control chart performance when the sample size is small. Numerical result shows that the standard estimation method is very unstable when the sample size is small, and there is a large probability that the standard estimation method breaks down if the level of correlation between successive means is small-to-moderate. The new method given here overcomes this difficulty.

Key words: Autoregressive moving average model, exponentially weighted moving average control charts; first-order autoregressive model; maximum likelihood estimation; Shewhart control chart.

1 Introduction

The assumption usually made when applying control charts to monitor a process is that the process observations $Y_1, Y_2, \dots$ are independent, identically distributed (i.i.d.) variables with constant mean and variance. The process observations can be expressed as

$$Y_t = \mu + \epsilon_t, \quad t = 1, 2, \dots,$$

where μ is the process mean and random error terms ϵ_t's are i.i.d. normal random variables with mean 0 and variance σ_ϵ^2. When the process is in-control, the mean μ is often assumed to be the target value and can change to some other values when a special cause occurs. However, observations from a process are often autocorrelated, and this autocorrelation significantly affects the performance of control charts under the assumption of independence. If the process observations are autocorrelated, there are two general methods to construct control charts. The first method uses the Shewhart control chart but adjusts the control limits to account for the autocorrelation (*e.g.*, see Vasilopoulos and Stamboulis (1978) and

VanBrackle and Reynolds (1997)). The second method fits a time series model to the process data and then uses the residuals from this time series forecast to develop control charts. Control charts based on residuals have been considered by Abraham and Kartha (1979), Alwan (1991), Alwan and Roberts (1988), Harris and Ross (1991), Montgomery (2000) and Lu and Reynolds (1999a, 1999b, 2001).

Reynolds *et al.* (1996) studied the properties of fixed sampling interval and variable sampling interval $\overline{X}$ control charts for a first-order autoregressive (AR(1)) process with a random error. This process can be written as

$$Y_t \quad = \quad \mu_t + \epsilon_t \tag{1.1}$$

and

$$\mu_t \quad = \quad (1-\phi)\xi + \phi\mu_{t-1} + \alpha_t, \quad t = 1, 2, \ldots, \tag{1.2}$$

where ϕ is the autoregressive parameter. The α_t's are i.i.d. normal random variables with mean 0 and variance σ_α^2. It is assumed that the mean μ_t is not a fixed constant but rather continually wander over time. Here, the mean μ_t can be described as the AR(1) process. The process will be stationary if $|\phi| < 1$, however, most cases of practical interest involve positive autocorrelation. The model determined by equations (1.1) and (1.2) is a special case of a first-order autoregressive moving average (ARMA(1,1)) process. If $\phi = 0$, then the AR(1) process with a random error still holds and the means at different times are independent. The distribution of means μ_t's depends on the starting point μ_0. If the starting point μ_0 follows a normal distribution with mean ξ and variance $\sigma_\mu^2 = \frac{\sigma_\alpha^2}{1-\phi^2}$, then μ_t is normally distributed with mean ξ and variance σ_μ^2. Moreover, it is easy to show that Y_t is normally distributed with mean ξ and variance $\sigma_Y^2 = \sigma_\mu^2 + \sigma_\epsilon^2$, for $t = 1, 2, \ldots$. Let $\psi = \frac{\sigma_\mu^2}{\sigma_Y^2}$ denote the proportion of the process variance due to the AR(1) process. It can be shown that the correlation between Y_k and Y_{k-1} is $\rho = \phi\psi$. When $\psi = 1$, the sequence $Y_1, Y_2, \ldots$ is an AR(1) process. This model has been used frequently in practical applications, for example, by Harris and Ross (1991), Wardell *et al.* (1994) and Lu and Reynolds (1999a, 1999b, 2001).

Box *et al.* (1994) showed that the AR(1) process with a random error is equivalent to an ARMA(1,1) process. The ARMA(1,1) process can be written as

$$Y_t = (1-\phi)\xi + \phi Y_{t-1} + \gamma_t - \theta\gamma_{t-1}, \quad t = 1, 2, \ldots, \tag{1.3}$$

where the γ_t's are i.i.d. normal random variables with mean 0 and variance σ_γ^2. If $0 < \phi < 1$ and $\sigma_\epsilon^2 > 0$, Reynolds *et al.* (1996) gave equations for expressing the ARMA(1,1) parameters in terms of the parameters in the AR(1) process with a random error as

$$\theta = \frac{\sigma_\alpha^2 + (1+\phi^2)\sigma_\epsilon^2}{2\phi\sigma_\epsilon^2} - \frac{1}{2}\sqrt{\left(\frac{\sigma_\alpha^2 + (1+\phi^2)\sigma_\epsilon^2}{\phi\sigma_\epsilon^2}\right)^2 - 4} \tag{1.4}$$

and

$$\sigma_\gamma^2 = \frac{\phi}{\theta}\sigma_\epsilon^2. \tag{1.5}$$

Conversely, if the parameters in the ARMA(1,1) process have values $0 \leq \theta \leq \phi < 1$ and $\sigma_\gamma^2 > 0$, then the parameters in the AR(1) process with a random error can be expressed in terms of the ARMA(1,1) parameters as

$$\sigma_\alpha^2 = \frac{(\phi-\theta)(1-\phi\theta)}{\phi}\sigma_\gamma^2 \tag{1.6}$$

and

$$\sigma_\epsilon^2 = \frac{\theta}{\phi}\sigma_\gamma^2. \tag{1.7}$$

When observations are taken from the AR(1) process with a random error, it is more convenient to think of these observations as taking from an ARMA(1,1) process. Standard time series estimation techniques can be used to estimate the parameters, and then equations (1.6) and (1.7) can be used to determine the estimates of the corresponding parameters in the AR(1) process with a random error as long as the ARMA(1,1) estimates satisfy $0 \leq \hat{\theta} \leq \hat{\phi} < 1$ and $\hat{\sigma}_\gamma^2 > 0$. This result depends on a precise parameter estimation in the ARMA(1,1) process. However, performance of the parameter estimation in the ARMA(1,1) process depends on the sample size. Lu and Reynolds (1999a) pointed out that the estimation of ϕ and θ in an ARMA(1,1) process is unstable when the sample size is not large enough. In their simulation for the ARMA(1,1) process, the parameter estimation for ϕ and θ are stable when the sample sizes $n = 500$ and $n = 1000$. But the parameter estimation become very unstable when only $n = 100$ data points are taken. In their numerical examples, they considered the case of $\phi = 0.6$, $\sigma_\epsilon^2 = 0.1$, and $\sigma_Y^2 = 1$ for an AR(1) process with a random error. The series contains 100, 200, 500 and 1000 data points, respectively. They translated this process to an ARMA(1,1) process first by equations (1.4) and (1.5) and then think of these data points as taking from an ARMA(1,1) process with $\phi = 0.6$ and $\theta = 0.085$. Through the standard time series estimation method, they found that there is a large probability to produce negative estimates of ϕ and θ when only 100 or 200 data points are taken. Incorrect estimates of parameters result in a large impact on developing control charts.

In this paper, a conditional maximum likelihood (ML) estimation method is provided in Section 2 to overcome this difficulty for the case of small sample size. Some numerical results are provided in Section 3 to evaluate the performance of the method. Conclusions are made in Section 4.

2 Conditional maximum likelihood estimation method

If observations are taken from an AR(1) process with a random error, it is more convenient to consider that they are taken from an ARMA(1,1) process. Assume that $0 \leq \theta \leq \phi < 1$ and $\sigma_\gamma^2 > 0$ for the ARMA(1,1) process, and equations (1.6) and (1.7) are used to determine the estimates of the corresponding parameters in the AR(1) process with a random error. The standard estimation method breaks down easily if the sample size is small. One situation of that is the parameters ϕ and θ in the ARMA(1,1) process are estimated as negative values.

Suppose that observations are taken from an AR(1) process with a random error. Assume that parameters $\phi = 0.3(0.1)0.8$, $\sigma_\epsilon^2 = 0.25$, $\sigma_Y^2 = 1$, and $\psi = 0.75$. This process implies that 75% of the process variability is due to the AR(1) process and the correlation between adjacent observations is $\rho = 0.75\phi$. Using equations (1.4) and (1.5), we can think of these observations as coming directly from an ARMA(1,1) process and without reference to the original process. For each combination of the specified values of ϕ, θ and σ_γ^2, n ($n = 100$ and 300) observations are generated from the corresponding ARMA(1,1) process until 10000 runs are obtained; moreover, the ARMA(1,1) parameters are estimated based on the standard ML estimation. Some criteria for evaluating the performance of estimators are considered. They are: (1) biases of estimators $\hat{\phi}$ and $\hat{\theta}$ of parameters ϕ and θ, respectively; (2) mean squared error (MSE) of estimators $\hat{\phi}$ and $\hat{\theta}$ of parameters ϕ and θ, respectively; (3) proportion of negative estimates of ϕ and θ when $0 < \theta < \phi < 1$, say Pr-1; (4) proportion of $0 < \hat{\phi} < \hat{\theta} < 1$ when $0 < \theta < \phi < 1$, say Pr-2. Table 1 and Table 2 present the simulated results. It can be seen that the biases, MSEs and Pr-2 are not very large. However, the values of Pr-1 are large for small-to-moderate values of ϕ. Even though the initial sample size reaches 300, the probability Pr-1 is still large and, hence equations (1.6) and (1.7) break down for these cases when ϕ is small-to-moderate. Therefore, a conditional ML estimation method is provided here to overcome this difficulty when the sample size is not large enough.

Let $\tilde{\mu}_t = \mu_t - \xi$ for $t = 0, 1, 2, \ldots, m$, where m is the initial sample size. In this article, we choose $\mu_0 = \xi$ to be the starting point. Hence, $\tilde{\mu}_0 = \mu_0 - \xi = 0$ and equation (1.2) can be rewritten as

$$\tilde{\mu}_t = \phi\tilde{\mu}_{t-1} + \alpha_t, \quad t = 1, 2, \ldots, m. \tag{2.8}$$

From the derivation in Appendix A, we have

$$\tilde{\mu}_t = \sum_{j=1}^{t} \phi^{t-j}\alpha_j, \quad t = 1, 2, \ldots, m. \tag{2.9}$$

Let $\tilde{\boldsymbol{\mu}}^T = (\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_m)$. Then $E(\tilde{\boldsymbol{\mu}}) = \mathbf{0}$ and the variance-covariance matrix $Cov(\tilde{\boldsymbol{\mu}}) = \sigma_\alpha^2\boldsymbol{V}$, where $\boldsymbol{V}$ is an $m \times m$ symmetric matrix with entry $v_{ij} = \sum_{h=1}^{i}\phi^{i+j-2h}$, for $i \le j$. Moreover, we can show that $\tilde{\boldsymbol{\mu}}$ follows a normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $\sigma_\alpha^2\boldsymbol{V}$. Let $\boldsymbol{\mu}^T = (\mu_1, \mu_2, \ldots, \mu_m) = \tilde{\boldsymbol{\mu}}^T + \xi\mathbf{1}^T$ and $\boldsymbol{\epsilon}^T = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m)$, where $\mathbf{1}$ is an $m \times 1$ vector with entries 1. Then, $\boldsymbol{\mu}$ follows a normal distribution with mean $\xi\mathbf{1}$ and variance-covariance matrix $\sigma_\alpha^2\boldsymbol{V}$, and $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\epsilon}$ follows a normal distribution with mean $\xi\mathbf{1}$ and variance-covariance matrix $\sigma_\alpha^2\boldsymbol{V} + \sigma_\epsilon^2\boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix of order m. Let $\delta = \frac{\sigma_\epsilon^2}{\sigma_\alpha^2}$ and $\boldsymbol{P}(\delta) = \boldsymbol{V} + \delta\boldsymbol{I}$. Then $\boldsymbol{Y}$ follows a normal distribution with mean $\xi\mathbf{1}$ and variance-covariance matrix $\sigma_\alpha^2\boldsymbol{P}(\delta)$. Suppose that parameters ϕ, ψ and δ

Table 1 *Bias, MSE, Pr−1 and Pr−2 when $n = 100$*

ϕ	θ	$bias(\hat{\phi}, \phi)$	$bias(\hat{\theta}, \theta)$	$MSE(\hat{\phi})$	$MSE(\hat{\theta})$	Pr−1	Pr−2
0.3	0.07902	-0.11864	-0.11589	0.15814	0.16059	0.5251	0.0101
		(0.19144)	(0.20508)	(0.07040)	(0.08277)		
0.4	0.11001	-0.09901	-0.09281	0.11046	0.11414	0.4465	0.0017
		(0.11806)	(0.13977)	(0.03109)	(0.05169)		
0.5	0.14590	-0.07520	-0.07212	0.07504	0.08272	0.3647	0.0003
		(0.07436)	(0.09492)	(0.02529)	(0.03537)		
0.6	0.18950	-0.04745	-0.04717	0.04705	0.05729	0.2416	0
		(0.03787)	(0.05327)	(0.01786)	(0.02582)		
0.7	0.24579	0.00782	-0.00749	0.02692	0.03716	0.1001	0
		(0.04178)	(0.03404)	(0.01653)	(0.02322)		
0.8	0.32523	0.10729	0.02643	0.01946	0.01840	0.1070	0
		(0.11098)	(0.03115)	(0.01871)	(0.01670)		

Note: The values in parentheses are computed only for the cases of $0 < \hat{\theta} < \hat{\phi} < 1$.

are known. The log−likelihood function is maximized for σ_α^2 by

$$\hat{\sigma}_\alpha^2 = \frac{1}{m}(\boldsymbol{Y} - \xi\mathbf{1})^T \boldsymbol{P}^{-1}(\delta)(\boldsymbol{Y} - \xi\mathbf{1}). \tag{2.10}$$

In practical applications, the values of ϕ, ξ and δ are usually unknown. We can get these estimates from the corresponding ARMA(1,1) model based on the standard ML procedure (*e.g.*, see Box *et al.* (1994)) first and then plug these estimates into equation (2.10) to obtain a conditional estimate $\hat{\sigma}_\alpha^2$. Assume that observations are taken from an AR(1) process with a random error with parameters $\phi > 0$, $\sigma_\alpha^2 > 0$ and $\sigma_\epsilon^2 > 0$. The corresponding ARMA(1,1) process and equations (1.6) and (1.7) can be used to determine the estimates of the parameters in the original process. When $\hat{\phi} < 0$, $\hat{\theta} < 0$ or $0 < \hat{\phi} < \hat{\theta} < 1$, a conditional ML estimation procedure is considered. The steps are as follows:

Step 0: Compute the estimates $\hat{\phi}$ and $\hat{\sigma_\gamma^2}$ of the parameters ϕ and σ_γ^2 in an AR(1) model, respectively. Let $\sigma_\gamma^{2*} = \hat{\sigma}_\gamma^2$. Set the initial values $\phi^{(0)} = \hat{\phi}$ and $\theta^{(0)} = 0$. From equation (1.7), we have $\sigma_\epsilon^{2(0)} = 0$. Using equation (1.6), we can obtain that $\sigma_\alpha^{2(0)} = \sigma_\gamma^{2*}$, $\sigma_Y^{2(0)} = \frac{\sigma_\gamma^{2*}}{1-\phi^{(0)2}}$, and $\sigma_\mu^{2(0)} = \sigma_Y^{2(0)}$. Hence, $\psi^{(0)} = \frac{\sigma_\mu^{2(0)}}{\sigma_Y^{2(0)}}$=1 and $\delta^{(0)} = \frac{\sigma_\epsilon^{2(0)}}{\sigma_\alpha^{2(0)}} = 0$.

Table 2 *Bias, MSE, Pr−1 and Pr−2 when* $n = 300$

ϕ	θ	$bias(\hat{\phi},\phi)$	$bias(\hat{\theta},\theta)$	$MSE(\hat{\phi})$	$MSE(\hat{\theta})$	Pr−1	Pr−2
0.3	0.07902	-0.04258	-0.04124	0.06167	0.06457	0.4327	0
		(0.12217)	(0.13186)	(0.03436)	(0.04026)		
0.4	0.11001	-0.03432	-0.03237	0.03667	0.04020	0.3332	0
		(0.06552)	(0.07654)	(0.01720)	(0.02154)		
0.5	0.14589	-0.02357	-0.02257	0.02057	0.02534	0.2068	0
		(0.02665)	(0.03522)	(0.01065)	(0.01434)		
0.6	0.18949	-0.00906	-0.00783	0.01165	0.01640	0.0805	0
		(0.00915)	(0.01455)	(0.00781)	(0.01136)		
0.7	0.24579	0.00865	0.00562	0.00673	0.01118	0.0131	0
		(0.01181)	(0.00963)	(0.00601)	(0.01012)		
0.8	0.32522	0.06048	0.03717	0.00738	0.00941	0.0003	0
		(0.06057)	(0.03730)	(0.00736)	(0.00937)		

Note: The values in parentheses are computed only for the cases of $0 < \hat{\theta} < \hat{\phi} < 1$.

Step 1: Substitute $\delta^{(0)}$, $\phi^{(0)}$ and $\hat{\xi}$ into equation (2.10) to get $\sigma_\alpha^{2(1)}$. Then, update the following estimates: $\sigma_\mu^{2(1)} = \frac{\sigma_\alpha^{2(1)}}{1-\phi^{(0)2}}$, $\sigma_\epsilon^{2(1)} = \sigma_Y^{2(0)} - \sigma_\mu^{2(1)}$, $\psi^{(1)} = \frac{\sigma_\mu^{2(1)}}{\sigma_Y^{2(0)}}$, $\phi^{(1)} = \frac{\phi^{(0)}}{\psi^{(1)}}$, $\delta^{(1)} = \frac{\sigma_\epsilon^{2(1)}}{\sigma_\alpha^{2(1)}}$,

$$\theta^{(1)} = \frac{\sigma_\alpha^{2(1)} + (1+\phi^{(1)2})\sigma_\epsilon^{2(1)}}{2\phi^{(1)}\sigma_\epsilon^{2(1)}} - \frac{1}{2}\sqrt{\left(\frac{\sigma_\alpha^{2(1)} + (1+\phi^{(1)2})\sigma_\epsilon^{2(1)}}{\phi^{(1)}\sigma_\epsilon^{2(1)}}\right)^2 - 4},$$

and $\sigma_Y^{2(1)} = \sigma_Y^{2(0)}$.

Step 2: For integer $\ell \geq 2$, compute $\sigma_\alpha^{2(\ell)}$ by equation (2.10) with $\delta^{(\ell-1)}$, $\phi^{(\ell-1)}$ and $\hat{\xi}$, and update the following estimates: $\sigma_\epsilon^{2(\ell)} = \frac{\theta^{(\ell-1)}}{\phi^{(\ell-1)}}\sigma_\gamma^{2*}$, $\sigma_\mu^{2(\ell)} = \frac{\sigma_\alpha^{2(\ell)}}{1-\phi^{(\ell-1)2}}$, $\sigma_Y^{2(\ell)} = \sigma_\mu^{2(\ell)} + \sigma_\epsilon^{2(\ell)}$, $\psi^{(\ell)} = \frac{\sigma_\mu^{2(\ell)}}{\sigma_Y^{2(\ell)}}$, $\phi^{(\ell)} = \frac{\phi^{(\ell-1)}}{\psi^{(\ell)}}$, $\delta^{(\ell)} = \frac{\sigma_\epsilon^{2(\ell)}}{\sigma_\alpha^{2(\ell)}}$ and

$$\theta^{(\ell)} = \frac{\sigma_\alpha^{2(\ell)} + (1+\phi^{(\ell)2})\sigma_\epsilon^{2(\ell)}}{2\phi^{(\ell)}\sigma_\epsilon^{2(\ell)}} - \frac{1}{2}\sqrt{\left(\frac{\sigma_\alpha^{2(\ell)} + (1+\phi^{(\ell)2})\sigma_\epsilon^{2(\ell)}}{\phi^{(\ell)}\sigma_\epsilon^{2(\ell)}}\right)^2 - 4}.$$

Repeat Step 2 until the minimum $\sigma_\alpha^{2(k)}$ is obtained, where k is a positive integer.

Step 3: The estimate $\theta^{(k)}$ is taken as the new initial estimate of θ. Moreover, $\theta^{(k)}$, $\phi^{(0)}$, equation (1.6) and equation (1.7) are used to get new initial estimates of the parameters σ_α^2 and σ_ϵ^2, respectively. Repeat Step 1 and Step 2 again through these new initial estimates until the minimum $\sigma_\alpha^{2(h)}$ is obtained, where h is a positive integer.

The estimates $\phi^{(h)}$, $\sigma_Y^{2(h)}$ and $\psi^{(h)}$ can be used to develop control charts and help us to understand the autocorrelated process.

After the parameters ϕ, σ_Y^2 and ψ are estimated, two control charts for monitoring the mean of AR(1) process with a random error are suggested. The first chart is the exponentially weighted moving average (EWMA) control chart based on the observations (see Lu and Reynolds (1999a)). This chart plots the control statistics

$$X_t = (1-\lambda)X_{t-1} + \lambda Y_t, \quad t = 1, 2, \ldots, \tag{2.11}$$

where λ is a smoothing constant satisfying $0 < \lambda \leq 1$ and X_0 is the starting value. If the target value of process is ξ_0, we can set $X_0 = \xi_0$. The control limits for the EWMA control chart are

$$\xi_0 \pm c\sqrt{\frac{\lambda}{2-\lambda}}\sigma_Y. \tag{2.12}$$

Lu and Reynolds (1999a) provided some tables for selecting adequate values of c and λ. In their tables, it seems that a relatively small value of λ, such as $\lambda = 0.2$, would work well across a range of shifts for both the EWMA chart of observations and the EWMA chart of residuals. For example, if $\psi = 0.5$, $\phi = 0.4$ and we hope to detect a shift in mean $(\xi - \xi_0)/\sigma_Y = 1.5$, we can select $\lambda = 0.2$ and $c = 3.391$.

The second control chart is the Shewhart control chart in the presence of autocorrelation. The control limits are

$$\xi_0 \pm h\sigma_Y. \tag{2.13}$$

The Table 7 of Reynolds *et al.* (1996) displays some adequate values of h. For example, if $\psi = 0.5$, $\phi = 0.4$, we can select $h = 2.998$.

3 Numerical results

We simulated 120 observations from an AR(1) process with a random error with parameters $\phi = 0.3$, $\sigma_\epsilon^2 = 0.25$, $\sigma_\mu^2 = 0.75$, $\sigma_Y^2 = 1$ and $\psi = 0.75$. The 120 observations have the target value of $\xi_0 = 1.25$, but a constant 1.25 is added to the last 20 observations to create a shift in the mean of size $\frac{\xi-\xi_0}{\sigma_Y^2} = 1.25$. Figure 1 shows the time plot of the data. Using SAS to estimate the corresponding ARMA(1,1) parameters for the first 100 observations, we get $\hat{\phi} = -0.95344$ and $\hat{\theta} = -0.91529$ and, hence the standard method based on equations (1.6) and (1.7) breaks down. Fitting these observations to an AR(1) model, we can get estimates $\hat{\phi} = 0.01679$, $\hat{\sigma}_\gamma^2 = 0.84924$ and $\hat{\psi} = 1$. These estimates seriously overestimate the

Table 3 *Estimation results of the example in Case 1 based on different methods*

	parameters					
	ϕ	θ	σ_ϵ^2	σ_μ^2	σ_Y^2	ψ
true value	0.30	0.07902	0.25	0.75	1	0.75
ARMA(1,1)	-0.95344	-0.91529	NA*	NA	NA	NA
AR(1)	0.01679	0	0	0.84948	0.84948	1
new method	0.37206	0.13315	0.30951	0.56002	0.86953	0.64404

Note: 'NA' denotes not available

value of ψ due to $\theta = 0$. The AR(1) model always gives $\hat{\psi} = 1$ whatever the data set is. Table 3 shows the estimation result based on the conditional ML estimation method. The estimates are much close to the true values of parameters. Hence, the control charts based on equations (2.12) and (2.13) can be constructed more adequately.

Consider the Shewhart chart with the control limits as in equation (2.13) first. From Table 3, we have $\hat{\sigma}_Y^2 = 0.86953$ based on the new method. Using an interpolation to the Table 7 in Reynolds *et al.* (1996), we have $h = 2.996$. Hence, we can get the control limits $1.25 \pm 2.996\sqrt{0.86953} = -1.54373$ and 4.04372, respectively. Figure 1 shows this chart and indicates an out of control signal at observation 113. Next, we consider the EWMA chart. The EWMA estimates based on equation (2.11) are plotted in Figure 2. Using $\lambda = 0.2$, $\hat{\phi} = 0.37206$ and $\hat{\psi} = 0.64404$, the Figure 1 in Lu and Reynolds (1999a) shows that $c = 3.5$. Substitute these estimates into equation (2.12), we can get the control limits are 0.16210 and 2.33790, respectively. The EWMA chart shows an out of control signal at observation 106 which exceeds the upper control limit and indicates a potential upward trend for the following estimates of the EWMA control statistics. The EWMA chart is faster to detect an out of control signal than the Shewhart chart does.

Table 4 displays the average values of the estimates of parameters based on the new method. The value of n_1 denotes the total number of the simulated data sets with estimates $\hat{\phi} < 0$, $\hat{\theta} < 0$ or $0 < \hat{\phi} < \hat{\theta} < 1$ based on an ARMA(1,1) model. Within these n_1 data sets, the value of n_2 denotes the total number negative estimates for the autoregressive parameter based on an AR(1) model. In these n_2 data sets, the new method still breaks down. Hence, the average values of the estimates in Table 4 are computed based on only the $n_1 - n_2$ simulated data sets. From Table 4 we can see that the mean estimate of ψ is overestimated due to both the values of θ and σ_ϵ^2 are underestimated. But, the mean estimate of ψ is larger than 0.5 that keeps the same direction of the true value of ψ. The other estimates $\hat{\phi}$ and $\hat{\sigma}_Y^2$ are reasonably close to the true value and, hence can be used to construct the control charts adequately.

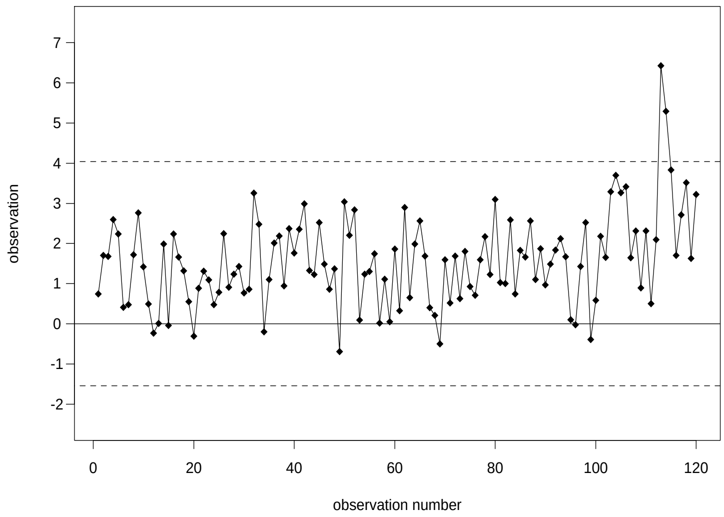

Figure 1 *Shewhart chart of the observations with nominal in control ARL=370.4.*

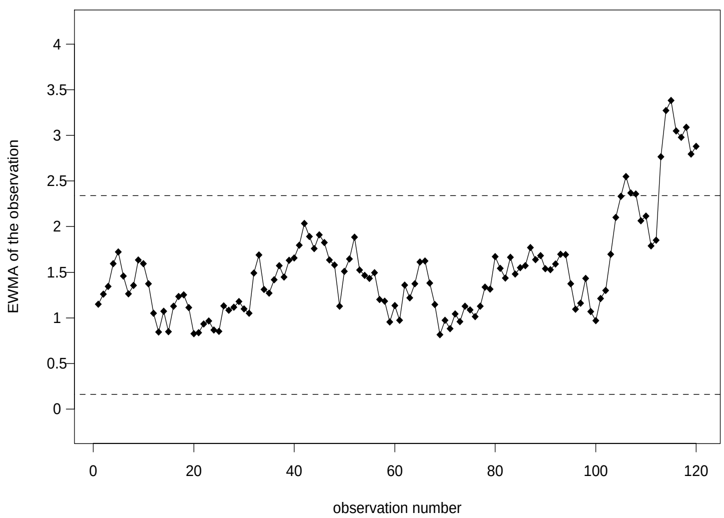

Figure 2 *EWMA of the observations.*

Table 4 *Estimation results of the examples in Case 1 based on new method*

	parameters							
	ϕ	σ_γ^2	σ_ϵ^2	σ_μ^2	σ_Y^2	ψ	n_1	n_2
true value	0.30	0.94907	0.25	0.75	1	0.75	5352	141
estimates	0.24164	0.94414	0.04524	0.95215	0.99739	0.95463		
true value	0.40	0.90899	0.25	0.75	1	0.75	4482	16
estimates	0.30323	0.90278	0.04132	0.94954	0.99087	0.95829		
true value	0.50	0.85676	0.25	0.75	1	0.75	3650	5
estimates	0.37308	0.84753	0.03603	0.94799	0.98402	0.96337		
true value	0.60	0.79157	0.25	0.75	1	0.75	2416	0
estimates	0.44227	0.78598	0.03027	0.95357	0.98385	0.96922		

4 Conclusions

If observations are taken from an AR(1) process with a random error, it is more convenient to think of these observations are taken from an ARMA(1,1) process. This paper provides a conditional ML estimation procedure to overcome the difficulty of that the standard ML estimation fails for the ARMA(1,1) process. If the sample size is small or we wish to chart the process early, then the proposed estimation method can be conducted even when the the standard ML estimation fails. But the proposed estimation method still fails if the AR(1) model can not provide a positive estimate for the parameter ϕ as an adequate initial value. Fortunately, this situation appears rarely in practice. The new estimation method can be used in the early stage and aid us to chart the process until enough observations are obtained, and then the standard ML estimation is used to construct the control charts.

Appendix: Expression of the components of $\tilde{\boldsymbol{\mu}}$

By equation (2.8), we can show that

$$
\begin{aligned}
\tilde{\mu}_1 &= \phi\tilde{\mu}_0 + \alpha_1, \\
\tilde{\mu}_2 &= \phi\tilde{\mu}_1 + \alpha_2 = \phi^2\tilde{\mu}_0 + \phi\alpha_1 + \alpha_2, \\
&\vdots \\
\tilde{\mu}_m &= \phi\tilde{\mu}_{m-1} + \alpha_m = \phi^m\tilde{\mu}_0 + \phi^{m-1}\alpha_1 + \phi^{m-2}\alpha_2 + \cdots + \phi\alpha_{m-1} + \alpha_m.
\end{aligned}
$$

Hence, $E(\tilde{\mu}_t) = 0$, for all $t = 1, 2, \ldots, m$. Moreover, for $i \leq j$,

$$\begin{aligned} Cov(\tilde{\mu}_i, \tilde{\mu}_j) &= Cov\left(\sum_{h=1}^{i} \phi^{i-h}\alpha_h, \quad \sum_{s=1}^{j} \phi^{j-s}\alpha_s\right) \\ &= \sum_{h=1}^{i}\sum_{s=1}^{j} \phi^{i-h}\phi^{j-s} Cov(\alpha_h, \alpha_s) \\ &= \sum_{h=1}^{i} \phi^{i+j-2h}\sigma_\alpha^2. \end{aligned}$$

Thus, $v_{ij} = \sum_{h=1}^{i} \phi^{i+j-2h}$.

(Received March, 2004. Accepted August, 2004.)

References

Abraham, B. and Kartha, C. P. (1979). Forecast stability and control charts. *ASQC Technical Conference Transactions*, American Society for Quality Control, Milwaukee, 675-680.

Alwan, L. C. (1991). Autocorrelation: fixed versus variable control limits. *Quality Engineering*, **4**, 167-188.

Alwan, L. C. and Roberts, H. V. (1988). Time-series modeling for statistical process control. *Journal of Business and Economic Statistics*, **6**, 87-95.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*. 3rd edition. Prentice-Hall, Englewood Cliffs.

Harris, T. J. and Ross, W. H. (1991). Statistical process control procedures for correlated observations. *The Canadian Journal of Chemical Engineering*, **69**, 48-57.

Lu, C.-W. and Reynolds, M. R., JR. (1999a). EWMA control charts for monitoring the mean of autocorrelated process. *Journal of Quality Technology*, **31**, 166-188.

Lu, C.-W. and Reynolds, M. R., JR. (1999b). Control charts for monitoring the mean and variance of autocorrelated process. *Journal of Quality Technology*, **31**, 259-274.

Lu, C.-W. and Reynolds, M. R., JR. (2001). CUSUM charts for monitoring an autocorrelated process. *Journal of Quality Technology*, **33**, 316-334.

Montgomery, D. C. (2000). *Introduction to Statistical Quality Control*. 4th edition. New York: John Wiley and Sons.

Reynolds, M. R., JR., Arnold, J. C. and Baik, J. W. (1996). Variable sampling interval $\overline{X}$ charts in the presence of correlation. *Journal of Quality Technology*, **28**, 12-30.

Vanbrackle, L. N. and Reynolds, M. R., JR. (1997). EWMA and CUSUM control charts in the presence of correlation. *Communications in Statistics – Simulation and Computation*, **26**, 979-1008.

Vasilopoulos, A. V. and Stamboulis, A. P.(1978). Modification of control limits in the presence of data correlation. *Journal of Quality Technology*, **10**, 20-30.

Wardell, D. G., Moskowitz, H. and Plante, R. D. (1994). Run length distributions of special-cause control charts for correlated processes. *Technometrics*, **36**, 3-17.

Tzong-Ru Tsai, **Yi-Chen Chiang** and **Shuo-Jye Wu**
Department of Statistics
Tamkang University
Tamsui, Taipei, Taiwan 251, ROC.
E-mail: shuo@stat.tku.edu.tw