Journal of the Chinese
Statistical Association
Vol. 52, (2014) 466–496

A CONTROL CHART USING COPULA-BASED MARKOV CHAIN MODELS

Ting-Hsuan Long and Takeshi Emura[†1]
Graduate Institute of Statistics, National Central University

ABSTRACT

Statistical process control is an important and convenient tool to stabilize the quality of manufactured goods and service operations. The traditional Shewhart control chart has been used extensively for process control, which is valid under the independence assumption of consecutive observations. In real world applications, there are many types of dependent observations in which the traditional control chart cannot be used. In this paper, we propose to apply a copula-based Markov chain to perform statistical process control for correlated observations. In particular, we consider three methods to obtain the estimates of upper control limit (UCL) and lower control limit (LCL) for the control chart. It is shown by simulations that Joe's parametric maximum likelihood method provides the most reliable estimates of the UCL and LCL compared to the other methods. We also propose simulation techniques to compute the average run length (ARL) of the proposed charts, which can be used to set the UCL and LCL for a given value of ARL. The piston rings data are analyzed for illustration.

Key words and phrases: Average run length, Clayton model, Correlated data, Kendall's tau, Markov chain.
JEL classification: C13, C15, C18. C22.

[†]Correspondence to: Takeshi Emura
E-mail:emura@stat.ncu.edu.tw

1. Introduction

With the promotion of industrial technologies, statistical process control (SPC) has been essential and convenient tools for manufacturers. Unavoidably, as factories produce items in mass production, they encounter some defective items. The basic idea of SPC is to keep the defective rate at some specified threshold (often at 0.27%). Consequently, the manufacturers can control the loss of their business profits.

In the traditional Shewhart charts, the process measurements on items are assumed to be independent. Unfortunately, the assumption of independence usually does not satisfy when the intermission between samples is short. For instance, manufacturing ill-conditioned items could cause machine's temperature to get higher than the normal condition. If the intermission is short, the chance of producing ill-conditioned product in the next item increases. Hence, very often in industrial practice, measurements are positively correlated.

A first order autoregressive AR(1), a first order moving average MA(1), and a first order integrated moving average IMA(1) model are typically used for SPC with correlated observations. A concise review of these models in the SPC literature is found in Box and Narasimhan (2010). The early work starts with the papers by Johnson and Bagshaw (1974), Bagshaw and Johnson (1975) and Vasilopoulos and Stamboulis (1978). After that, the problem of SPC with correlated observation has been widely studied. A comprehensive overview of this problem is found in Wieringa (1999), Knoth and Schmid (2004) and Psarakis and Papaleonida (2007). Although higher order models are available, the literature on SPC remains focused on the first order models (Wetherill and Brown 1991; Wardell, et al. 1994; Wieringa, 1999; Knoth and Schmid, 2004; Psarakis and Papaleonida 2007; Montgomery 2009a, b; Box and Narasimhan, 2010). In this paper, we also consider a first order (i.e., Markov) model, but the dependence is modeled via copulas, which has not been considered in the SPC literature.

2. Background

2.1 Copula-based Markov Chain Model

A copula is a bivariate distribution function with the two marginals being unif(0, 1). Copulas are useful to model the dependence between the two random variables that are transformed to unif(0, 1). Sklar (1959) showed that for any bivariate distribution function $H(\mathrm{y}_1, \mathrm{y}_2)$ with marginal distributions $G_1(\mathrm{y}_1)$ and $G_2(\mathrm{y}_2)$, there exists a copula $C : [0,1]^2 \to [0,1]$ such that

$$H(\mathrm{y}_1, \mathrm{y}_2) = C(G_1(\mathrm{y}_1), G_2(\mathrm{y}_2)).$$

More information on copulas can be found in the books of Joe (1997) and Nelsen (2006).

Darsow, Nguten and Olsen (1992) first introduced copula-based Markov chain models for serially correlated observations $\{Y_t : t = 1, \ldots, n\}$, where a copula defines the correlation between Y_{t-1} and Y_t . The resultant series become a stationary process with the stationary distribution $G_1 = G_2$ (Joe, 1997; Chen and Fan, 2006). The copula-based Markov model includes the 1st order autoregressive model, or AR(1), as a special case with a Gaussian copula and normal margin (p.260 of Joe, 1997).

This paper focuses on the one-parameter Clayton copula defined as:

$$C(u_1, u_2; \alpha) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha} \mathbf{I}(u_1^{-\alpha} + u_2^{-\alpha} - 1 > 0),$$

where $\alpha \epsilon (-1, \infty) \backslash \{0\}$ describe the correlation between Y_{t-1} and Y_t. If $\alpha \epsilon (-1, 0)$, Y_{t-1} and Y_t have negative correlation; when $\alpha \epsilon (0, \infty)$, Y_{t-1} and Y_t have positive correlation. It is well known that the correlation measure on the scale of [-1, 1] is represented by Kendall's tau $\tau = \alpha/(\alpha + 2)$. Figure 1 shows the plot of the first-order Markov series $\{Y_t : t = 1, \ldots, n\}$ under the Clayton copula with the marginal being the standard normal distribution. It is seen that when α increases, the serial correlations become strong.

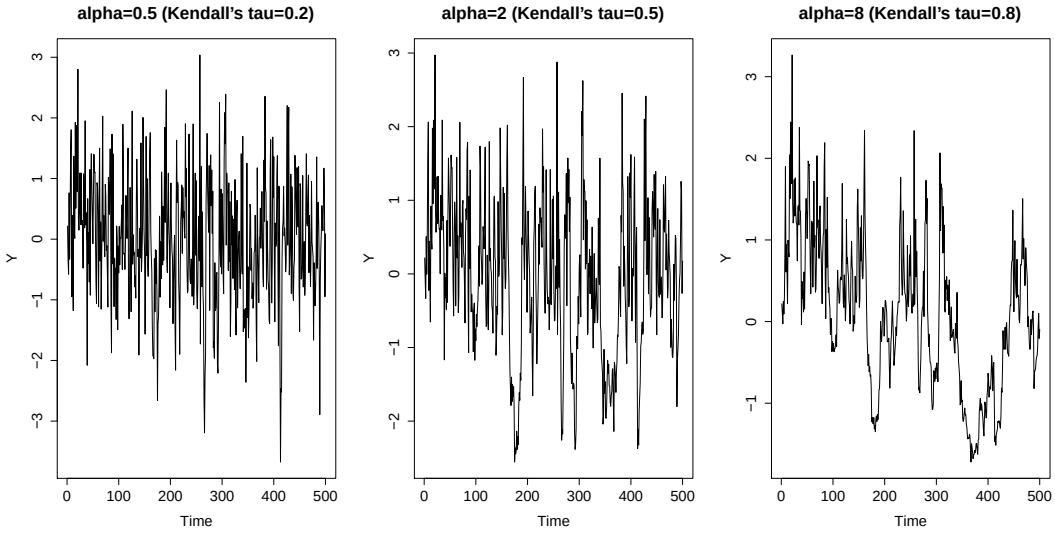

Fig.1. The plot of $\{Y_t : t = 1, \ldots, n\}$ under the Clayton copula with the marginal being the standard normal distribution, where n=500 .

In this article, we focus on the Clayton copula due to its popularity in applications. Some recent applications of the Clayton copula are referred to Sari et al. (2009) for industrial statistics and Emura and Chen (2014) for biostatistics.

2.2 Motivation and the Organization

In this paper, we assume that only one observation is available at each time, as usually assumed in the literature of SPC for autocorrelated data (Schmid, 1995; Wieringa, 1999; Kramer and Schmid, 2000; Knoth and Schmid, 2004; Psarakis and Papaleonida, 2007; Box and Narasimhan, 2010; Hryniewicz, 2012). Hence, we monitor individual observations $\{Y_t : t = 1, \ldots, n\}$ rather than subgroup averages. This is because the serial correlation reduces by taking subgroup averages (Wieringa, 1999).

The main theme of this paper is the application of the copula-based Markov chain models to SPC. In SPC, one often needs to estimate the upper control limit (UCL) and lower control limit (LCL) for $\{Y_t : t = 1, \ldots, n\}$. If the marginal mean μ and the marginal standard deviation σ of Y_t are known, one may set the three-sigma limits LCL=$\mu - 3\sigma$ and UCL=$\mu + 3\sigma$. For the aforementioned example of the Clayton model with the standard normal margins ($\mu = 0, \sigma = 1$), the UCL and LCL are +3 and -3, respectively.

In many real examples, such theoretical limits are unknown since the marginal distributions (μ and σ) are unknown. Therefore, the control limits must be estimated using the in-control data or Phase I data (p.230 of Montgomery 2009a). To the best of our knowledge, there is no paper discussing the estimation of the control limits under the copula-based time series models. Therefore, the primary objective of this paper is developing estimation procedures for ULC and LCL, which is detailed in Section 3. Then, Section 4 contains simulations that investigate the performance of the methods in Section 3.

If observations are correlated, however, the above three-sigma limits $\mu \pm 3\sigma$ may not keep the average run length (ARL) at the desired level (often at ARL=370). In the case of dependent observation, one might alternatively determine the UCL and LCL such that the ARL is equal to a given value. This is done by selecting a constant c such that the limits $\mu \pm c\sigma$ achieve a given ARL (Schmid, 1995). Besides setting the control limits, the ARL is an important measure of the performance of a control chart. Therefore, the secondary objective of this paper is developing appropriate simulation techniques for calculating the ARL under a copula-based Markov model, which is detailed in Section 5. The choice of c will be discussed with the real data analysis in Section 6.

Chapter 7 concludes the paper. Detailed calculations are given in Appendices.

3. Estimation of Process Parameters

We introduce methods to estimate parameters that are useful for SPC. Such parameters include the center line, UCL, and LCL.

3.1 Model Assumptions

Following Joe (1997) and Chen and Fan (2006), we impose the following assumption throughout the paper:

Assumption 1
Let $\{Y_t : t = 1, \ldots, n\}$ be a sequence of random variables, representing a quality characteristics. The variables follow a stationary first-order Markov process with the transition probability determined by

$$P(Y_{t-1} \leq y_{t-1}, Y_t \leq y_t) = H^*(y_{t-1}, y_t) = C(G^*(y_{t-1}), G^*(y_t); \alpha^*),$$

where $G^*(\cdot)$ is continuous marginal (stationary) distribution and $C(\cdot, \cdot; \alpha^*)$ is the true parametric copula for an unknown value α^*. Assume that, the copula is also continuous, and is neither the Fréchet-Hoeffding upper nor lower bound.

Assumption 1 derives the conditional density of Y_t given Y_{t-1} via

$$g^*(y_t)c(G^*(y_{t-1}), G^*(y_t); \alpha^*),$$

where $c(\cdot, \cdot; \alpha^*)$ is the copula density of $C(\cdot, \cdot; \alpha^*)$, and $g^*(\cdot)$ is the density of the true marginal (stationary) distribution $G^*(\cdot)$.

Under Assumption 1, the transformed process, $\{U_t : U_t \equiv G^*(Y_t)\}$ is a stationary Markov process of order 1 in which the joint distribution of U_t and U_{t-1} is given by the copula $C(u_0, u_1; \alpha^*)$, and the conditional density of U_t given $U_{t-1} = u_0$ is $f_{U_t|U_{t-1}=u_0}(u) = c(u_0, u_1; \alpha^*)$. This property is shown to be useful for generating the data.

3.2 Joe's Method

We demonstrate how the likelihood estimator of Joe (Joe, 1997) can be used to estimate relevant parameters. In most quality control work, relevant parameters are $\mu = E(Y_t)$ and $\sigma = \sqrt{\mathrm{var}(\mathrm{Y_t})}$ to get the control limits $\mu \pm 3\sigma$. Hence, it is convenient to parameterize G^* in terms of (μ, σ). Here we propose to set $G^*(y) = \Phi\{(y - \mu)/\sigma\}$, where Φ is the distribution function of $N(0, 1)$.

The log-likelihood function given data $\{y_t : t = 1, \ldots, n\}$ is

$$L(\mu, \sigma, \alpha) = \frac{1}{n}\sum_{t=1}^{n} \log\{\frac{1}{\sigma}\phi(\frac{y_t - \mu}{\sigma})\} + \frac{1}{n}\sum_{t=2}^{n} \log\, c\{\Phi(\frac{y_{t-1} - \mu}{\sigma}), \Phi(\frac{y_t - \mu}{\sigma}); \alpha\}.$$

The formula of log-copula density $\log c(u_1, u_2; \alpha)$ is given in Appendix A.1. The maximum likelihood estimator (MLE) that maximizes the preceding formula is denoted by $(\hat{\mu}, \hat{\sigma}, \hat{\alpha})$. The resultant estimators of LCL and UCL are $\hat{\mu} - 3\hat{\sigma}$ and $\hat{\mu} + 3\hat{\sigma}$, respectively.

The log-likelihood function $L(\mu, \sigma, \alpha)$ is twice differentiable and the formulas of the first and second derivatives are given in Appendix A.2. The derivatives are quite complicated but they are useful for likelihood inference.

It is well-known that the Newton-Raphson algorithm is sensitive to the initial values, especially in estimating three or more parameters (see Section 5.7 of Knight (2000)). We also encounter the cases that the algorithm diverges due to a wrong initial value. Knight (2000) suggests trying several different initial values. Based on this suggestion and our own numerical experiences, we propose the following "randomized" Newton-Raphson algorithm:

Newton-Raphson algorithm with randomization
Step 1: Choose the initial value $(\mu_0, \sigma_0, \alpha_0)$, defined as

$$\mu_0 = \bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t \ , \ \sigma_0 = \sqrt{\sum_{t=1}^{n} Y_t^2/n - \bar{Y}^2} \ , \ \alpha_0 = -2\tau_0/(\tau_0 - 1) \ ,$$

where

$$\tau_0 = \begin{pmatrix} n \\ 2 \end{pmatrix}^{-1} \sum_{i<j} \{\mathrm{sgn}(Y_j - Y_i)\mathrm{sgn}(Y_{j+1} - Y_{i+1})\} \ ,$$

and where sgn(x)=-1 for $x < 0$, sgn(x) = 0 for $x = 0$ and sgn(x) = 1 for $x > 0$.
Step 2:

$$\text{Set} \begin{bmatrix} \mu_{k+1} \\ \sigma_{k+1} \\ \alpha_{k+1} \end{bmatrix} = \begin{bmatrix} \mu_k \\ \sigma_k \\ \alpha_k \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 L}{\partial\mu\partial\mu} & \frac{\partial^2 L}{\partial\mu\partial\sigma} & \frac{\partial^2 L}{\partial\mu\partial\alpha} \\ \frac{\partial^2 L}{\partial\sigma\partial\mu} & \frac{\partial^2 L}{\partial\sigma\partial\sigma} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} \\ \frac{\partial^2 L}{\partial\mu\partial\alpha} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} & \frac{\partial^2 L}{\partial\alpha\partial\alpha} \end{bmatrix}^{-1} \Bigg|_{(\mu_k, \sigma_k, \alpha_k)} \begin{bmatrix} \frac{\partial L}{\partial\mu} \\ \frac{\partial L}{\partial\sigma} \\ \frac{\partial L}{\partial\alpha} \end{bmatrix}_{(\mu_k, \sigma_k, \alpha_k)}$$

for $k = 0, 1, \ldots\ldots$, where the formulas for the derivatives are given in Appendix A.2.

- If $|\mu_{k+1}$ - $\mu_k| < 10^{-5}, |\sigma_{k+1}$ - $\sigma_k| < 10^{-5}$ and $|\alpha_{k+1}$ - $\alpha_k| < 10^{-5}$, stop the algorithm and set $(\hat{\mu}, \hat{\sigma}, \hat{\alpha}) = (\mu_{k+1}, \sigma_{k+1}, \alpha_{k+1})$.
- If $|\mu_{k+1}$ - $\mu_k| > 10^{20}, |\sigma_{k+1}$ - $\sigma_k| > 10^{20}$ or $|\alpha_{k+1}$ - $\alpha_k| > 10^{20}$, replace $(\mu_0, \sigma_0, \alpha_0)$ with $(\mu_0, \sigma_0, \alpha_0 + \mu)$, where $\mu \sim$ unif(-0.1, 0.1), and return to Step 1.

Recently, Hu (2014) successfully applied a similar randomized Newton-Raphson method to stabilize the computation of the MLE under double-truncation. The estimators for the LCL and UCL are $\hat{\mu} - 3\hat{\sigma}$ and $\hat{\mu} + 3\hat{\sigma}$, respectively.

3.3 Chen and Fan's Method

Chen and Fan (2006) proposed a copula-based Markov chain to describe the dependence structure for financial time-series data. In their paper, they considered a semi-parametric copula model with non-parametric marginal distributions. In this section, we discuss how to apply their method to estimate relevant parameters that are useful for SPC.

The semi-parametric copula-based Markov chain model has unknown parameters (G^*, α^*). Chen and Fan (2006) proposed to estimate the unknown marginal (stationary) distribution G^* using $G_n(\cdot)$, the rescaled empirical distribution function defined as

$$G_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} \mathbf{I}\{Y_t \leq y\}.$$

Now, we estimate mean $\mu = E[Y_t]$ and standard deviation $\sigma = SD[Y_t]$ using $G_n(\cdot)$. One can use the Stieltjes integral to get the estimators

$$\hat{\mu} = \int y dG_n(y) = \frac{n}{n+1}\bar{Y},$$

where $\bar{Y} = \sum_{t=1}^{n} Y_t/n$, and

$$\hat{\sigma}^2 = \int y^2 dG_n(y) - [\int y dG_n(y)]^2 = \frac{1}{n+1}\sum_{t=1}^{n} Y_t^2 - (\frac{n}{n+1}\bar{Y})^2.$$

Hence,

$$\hat{\sigma} = \sqrt{\frac{1}{n+1}\sum_{t=1}^{n} Y_t^2 - (\frac{n}{n+1}\bar{Y})^2}.$$

If the marginal distribution $G^*(\cdot)$ is known, then the log-likelihood function is given by

$$L(\alpha) = \frac{1}{n}\sum_{t=1}^{n} \log\, g^*(Y_t) + \frac{1}{n}\sum_{t=2}^{n} \log\, c(G^*(Y_{t-1}), G^*(Y_t); \alpha).$$

Then, the unknown α^* is estimated by maximizing the above function with $G^*(\cdot)$ being replaced by $G_n(\cdot)$. The estimators for LCL and UCL are $\hat{\mu}-3\hat{\sigma}$ and $\hat{\mu}+3\hat{\sigma}$, respectively.

3.4 Standard Method

It is of interest to compare the above two methods with the standard estimators defined as

$$\hat{\mu} = \bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t,\ \hat{\sigma} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} Y_t^2 - (\bar{Y})^2}.$$

The corresponding estimators for the LCL and UCL are $\hat{\mu}-3\hat{\sigma}$ and $\hat{\mu}+3\hat{\sigma}$, respectively. Such estimators were considered in Kramer and Schmit (2000) under AR(1) models. The standard estimator is consistent but may incur the loss of efficiency by ignoring correlation.

4. Simulations

We have introduced three methods to estimate the process parameters. To know which method is most suitable for the SPC work, we compare the performance of the methods via simulations.

4.1 Simulation Methods

We develop the algorithm for generating $\{Y_t : t = 1, \ldots, n\}$ by extending the conditional approach for bivariate copula models (Frees and Valdez, 1998). Our simulations focus on the Clayton copula with $\alpha = 2(\tau = 0.5)$, $\alpha = 8(\tau = 0.8)$ and $\alpha = -1/3(\tau = -0.2)$. We choose the marginal (stationary) distribution to be the normal distribution $G^*(y) = \Phi\{(y-\mu)/\sigma\}$ with $(\mu, \sigma) = (1, 1)$. The algorithm is stated as follows:

Algorithm 1 (Data generation)

1. Generate a random number U_1, where $U_1 \sim$ unif(0,1). Then, set $Y_1 = \Psi^{-1}(U_1)$, where $\Psi(y) = \Phi\{(y-\mu)/\sigma\}$.
2. Set $Y_{t+1} = \Psi^{-1}\{[1 + (U_{t+1}^{-\alpha/(\alpha+1)} - 1)\Psi(Y_t)^{-\alpha}]^{-1/\alpha}\}$, where $U_{t+1} \sim$ unif(0,1), $t = 1, \ldots, n$.

After generating the data, we calculate parameter estimates using the three methods:

Method 1 (Joe's method; see Section 3.2)
Method 2 (Chen and Fan's method; see Section 3.3)
Method 3 (Standard method; see Section 3.4)

The MSE of an estimator $\hat{\theta}$ with respect to the unknown parameter θ is defined as MSE $(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$. We compare the three methods in terms of the MSE for $\theta = \mu,\ \sigma$, and $\mu + 3\sigma$. We also examine the bias, defined as Bias $(\hat{\theta}) = E(\hat{\theta}) - \theta$.

4.2 Simulation Results

The results based on 1000 repetitions are given in **Tables 1-3**. Generally speaking, the three methods give estimates $\hat{\theta}$ close to the true values of $\theta = \mu,\ \sigma$ and $\mu + 3\sigma$, respectively. Their MSE get close to zero as the sample sizes increase.

Under positive correlation ($\alpha > 0$), it is clear that the MSE for Joe's method is always smaller than other two methods (**Tables 1-3**). Under $\alpha = 2(\tau = 0.5)$, Joe's method reduces MSE($\hat{\sigma}$) and MSE($\hat{\mu} + 3\hat{\sigma}$) about by half the MSEs for the other two methods. When $\alpha = 8(\tau = 0.8)$, Joe's method gives remarkably superior MSE($\hat{\mu} + 3\hat{\sigma}$) to the other methods. The dominance of Joe's method over the other two becomes modest in terms of MSE($\hat{\mu}$). In SPC, however, the accuracy of MSE($\hat{\mu} + 3\hat{\sigma}$) is more important than MSE($\hat{\mu}$) since the out-of-control signals are decided by the UCL and LCL.

Under negative correlation ($\alpha < 0$), the three methods are quite comparable. The MSE of the three methods are very similar for all configurations. Overall, the MSE under the negative correlation is much smaller than that under positive correlation.

The efficiency of Joe's method is reasonable since it is performed under the correct assumptions on the Clayton copula and the normality. On the other hand, Chen and Fan's method and standard method do not rely on the distributional assumptions. To see the performance under a model misspecification, we generate heavy-tailed data $\{Y_t^* : t = 1, \ldots, n\}$ under the t-distribution by

$$Y_t^* = \mu + \sigma\sqrt{\frac{\nu - 2}{\nu}}\Psi^{-1}[\Phi\{(Y_t - \mu)/\sigma\}; \nu],$$

where $\Psi^{-1}[\cdot; \nu]$ is the quantile function of the t-distribution with degree of freedom $\nu = 10$. The performance of the three methods are comapred in **Table 4**. Although Joe's method is still the best for all configurations, its superiority becomes somewhat offset.

Therefore, as long as the true model is correctly specified or approximated well, Joe's method is most accurate in terms of MSE under positively correlated series. Since industrial settings typically faces with positively correlated series, Joe's method seems to be of great value. In practice, it is important to check the goodness-of-fit before using Joe's method.

5. Average Run Length

In this section, we develop simulation techniques to obtain the average run length (ARL). In particular, we choose the antithetic variables method to gain computational efficiency.

5.1 Calculation of ARL

The average run length (ARL) of a control chart is one way to determine the performance of control charts. The ARL is the average number of sample points that are plotted before a point is beyond the control limits. The ARL can help engineers know the performance of chart under study. For instance, if the process is in-control, engineers wish to keep the production process as long as possible. Hence, the chart that has a large ARL is preferred. The ARL is defined as follows:

Definition (ARL):

Let $\{Y_t, t = 1, 2, \ldots\}$ *be a sequence of random variables, representing a quality characteristics and* A=min$\{t : Y_t < \mu - 3\sigma$ *or* $Y_t > \mu + 3\sigma\}$ *be the run length. Then, the ARL is defined to be* $\mathrm{E}(A)$.

If $\{Y_t = 1, 2, \ldots\}$ are independent and identically distributed, the ARL is easily calculated as $\mathrm{E}(A) = 1/p$, where $p = P(Y_1 < \mu - 3\sigma$ or $Y_1 > \mu + 3\sigma\}$ [see p.37 of Wieringa (1999); p. 249 of Montgomery (2009a)]. However, for correlated observations, the ARL calculation is extremely difficult. Schmid (1995) proposed some analytical methods to calculate the ARL. However, his formula is complicated and does not give us practical way to calculate the ARL. Schmid (1995) and Hryniewicz (2012) used Monte Carlo simulations to calculate the ARL under the autoregressive model and copula-based chain model, respectively. In the following, we also suggest the Monte Carlo method to calculate the ARL under the copula-based models.

One can use Algorithm 1 to generate data until the data falls outside control limits and then obtain the value of run length. Repeating this step many times, we get the ARL. In this paper, we set m= 10000 repetitions. The algorithm is as follows:

Algorithm 2 (ARL with Monte Carlo)

1. Draw $Y_1 \sim N(\mu, \sigma)$.
2. Draw $U_{t+1} \sim$ unif$(0, 1)$, and then set $Y_{t+1} = \Psi^{-1}\{[1+(U_{t+1}^{-\alpha/(\alpha+1)} - 1)\Psi(Y_t)^{-\alpha}]^{-1/\alpha}\}$ for $t = 1, 2, \ldots\ldots$, where $\Psi(y) = \Phi\{(y - \mu)/\sigma\}$.
3. Calculate the run length $A = \min\{t : Y_t < \mu - 3\sigma$ or $Y_t > \mu + 3\sigma\}$.
4. Repeat Step 1$\sim$ Step 3 m times. The ARL is the average of the m run length.

5.2 Antithetic Variables

The calculation of the ARL requires a large number of Monte Carlo runs to get an accurate result. Some simulation techniques can help reduce the computational cost. The well-known techniques are common random number, antithetic variables, control variates, stratified sampling and important sampling (Chapter 9, Ross, 2013).

We introduce antithetic variables method, which is a simple method to reduce variable and computational cost.

The antithetic variables method aims to reduce the variance by introducing correlation in the series of Monte Carlo runs. In Algorithm 2, the ARL is written as $A = h\{U_t; t = 1, 2, \ldots\}$. It is important to notice that $B = h\{1 - U_t; t = 1, 2, \ldots\}$ has the same distribution as A. This implies that $(A + B)/2$ is unbiased for the ARL. Furthermore, if $\mathrm{cov}(A, B) < 0$, then

$$\mathrm{var}[(A + B)/2] < \mathrm{var}[A]/2$$

which becomed smaller than the variance of the average of two independent sequences. The following algorithm shows how to use the antithetic variables method:

Algorithm 3 (ARL with antithetic variables)

1. Draw $Y_1 \sim N(\mu, \sigma)$.
2. Draw $U_1 \sim \mathrm{unif}(0, 1)$ and then set $Y_{1,1} = \Psi^{-1}(U_1)$ and $Y_{2,1} = \Psi^{-1}(1 - U_1)$, where $\Psi(y) = \Phi\{(y - \mu)/\sigma\}$.
3. Draw $U_{t+1} \sim \mathrm{unif}(0, 1)$ and set $Y_{1,t+1} = \Psi^{-1}\{[1 + (U_{t+1}^{-\alpha/(\alpha+1)} - 1)\Psi(Y_{1,t})^{-\alpha}]^{-1/\alpha}\}$, $t = 1, 2, \ldots\ldots$
4. Calculate $A = \min\{t : Y_{1,t} < \mu - 3\sigma \text{ or } Y_{1,t} > \mu + 3\sigma\}$.
5. Set $Y_{2,t+1} = \Psi^{-1}\{[1 - U_{t+1})^{-\alpha/(\alpha+1)} - 1)\Psi(Y_{2,t})^{-\alpha}]^{-1/\alpha}\}$, $t = 1, 2, \ldots\ldots$.
6. Calculate $B = \min\{t : Y_{2,t} < \mu - 3\sigma \text{ or } Y_{2,t} > \mu + 3\sigma\}$.
7. Repeat step 1 $\sim$ step 6 m times, and get two sequences $(A_1, \ldots, A_m)$ and $(B_1, \ldots, B_m)$. The ARL is $\sum_{i=1}^{m}(A_i + B_i)/2m$.

Remark: In Step 3 and Step 5, we use common uniform random variables. In this way, we save the number of generating uniform random numbers by half, compared with Algorithm 2.

5.3 Simulation Results

We compare the calculation of the ARL between the Monte Carlo method (Algorithm 2) and antithetic variables method (Algorithm 3) under the same simulation settings as Section 4.1. To check whether the use of antithetic variables reduces the variance or not, we compare the standard deviation (SD) of the antithetic variables method with that of the usual Monte Carlo method. For the two algorithms to be comparable, the ARL for the Monte Carlo is $\sum_{i=1}^{2m} A_i/(2m)$ and for the anntithetic variables method is $\sum_{i=1}^{m}(A_i + B_i)/(2m)$, where $m = 10000$. Thus, the SD for the Monte Carlo method is $\sqrt{\sum_{i=1}^{2m} A_i^2/(2m) - \{\sum_{i=1}^{2m} A_i/(2m)\}^2}$ and for the antithetic variables method is $\sqrt{\sum_{i=1}^{m}(A_i^2 + B_i^2)/(2m) - \{\sum_{i=1}^{m}(A_i + B_i)/(2m)\}^2}$. Smaller SD

corresponds to better computational efficiency. We also calculate the sample correlation between the two sequences of the antithetic variables, denoted by $\text{cor}(A, B)$. If $\text{cor}(A, B) < 0$, we expect that the antithetic variables method reduces the SD.

The results are given in **Table 5**. The Monte Carlo and antithetic variables methods produce similar values for the ARL, which means that the two methods give a good approximation to the true ARL. Also, the SD of the two methods is quite similar. This implies that the variance reduction using the antithetic variables is quite modest. This result agrees with the fact that $\text{cor}(A, B)$ is very close to zero. However, it should be noted that the antithetic variables method reduces by half the number of generating random numbers.

We also examine the case of the one-sided control limit in which the ARL is the average number of sample points that are plotted before a point is beyond the UCL only, i.e. $A = \min\{t : Y_t > \mu + 3\sigma\}$. The results are summarized **Table 6**. In this case, $\text{cor}(A, B) < 0$ occurs in all configurations. The reason is that, if the sequence A reaches the UCL, the alternative sequence B gets close to the LCL. However, the effect of negative correlation is modest and there is no apparent efficiency gain. In conclusion, the antithetic variables method saves the number of random samples but does not improve efficiency.

We display the properties of both in-control and out-of-control ARL under various Kendall's tau in **Table 7**. It is seen that the ARL increases as Kendall's tau increases. This kind of the increase of the ARL with the correlation is well known (e.g., Schmid, 1995; Wieringa, 1999; Konth and Schmid 2004), and is in accordance with the simulation results of Hryniewicz (2012). The out-of-control ARL (1 σ -shift or 2 σ -shift) is substantially smaller than the in-control ARL, showing the good performance in detecting the out-of-control state. However, when the correlation is large, there is some delay in detecting the out-of-control signals. Under the case that Kendall's tau=0.0001, the ARL values agree with the well-known ARL values of the Shewhart chart for independent observations (ARL=370 under in-control; ARL=43.96 under 1 σ -shift; ARL=6.30 under 2 σ -shift).

To keep the in-control ARL at desired level (e.g., 370), one can select constant c such that the limits $\mu \pm c\sigma$ achieve a given ARL (Schmid, 1995). To do this, one can try many different values of c to calculate the ARL using either Algorithm 2 or 3. Then, the appropriate value of c is the one that is closest to the desired ARL. This procedure will be explained in the subsequent real data analysis. Obviously, this is a computationally intensive procedure. As the future work, we wish to reduce the computational cost by importance sampling [see Chen, Fuh, and Teng (2013) and reference therein]. However, this is challenging since the definition of the ARL involves infinitely many random variables.

5.4 ARL Under Estimated Parameters

If the marginal mean μ and the marginal standard deviation σ are unknown, one needs to estimate them from data under in-control status, which is often done in Phase I trial (Montgomery, 2009a). These estimates are used to calculate the UCL and LCL to set the control limit of Phase II. Here, the primary interest is to find a good estimator that have small deviation from the specified in-control ARL (true ARL). We conduct simulations to investigate the influence of parameter estimation on the ARL of Phase II. Such simulation designs have been considered in Kramer and Schmid (2000) and Hryniewicz (2012).

Table 8 compares the ARL under the estimated parameters and the true ARL, where the true ARL represents the case of the known UCL and LCL. For estimated parameters, the UCL and LCL have random variation due to estimation in Phase I. We use the three methods (Joe's method, Chen and Fan's method and standard method) to estimate the UCL and LCL. **Table 8** shows that the standard method leads to the ARL that are somewhat different from the true ARL. This is because the standard method provides less accurate estimate of the UCL and LCL, especially for strongly correlated cases. Although this is pointed out by Hryniewicz (2012), the paper does not offer the solution. Similarly, Chen and Fan (2006) also performed poorly **(Table 8)**. However, Joe's method provides the most unbiased ARL. Especially, when the correlation is high (Kendall's tau = 0.8), only Joe's method give reasonable approximation to the true ARL.

Based on the fact that Joe's method relies on the normality assumption, we also conduct simulations under model misspecifications. **Table 8** shows the results under misspecification in Phase I, based on the t-distribution with degree of freedom $\nu = 10$ (as considered in Section 4.2). Although Joe's method still provides the best approximation to the true ARL, the advantage is reduced. Therefore, as long as the normality assumption is approximated well, Joe's method would be recommended.

6. Data Analysis

We demonstrate the proposed copula-based control chart using the data on diameter measurements of piston rings (Montgomery, 2009b). We download the data available from R qcc package (Luca, 2014), and obtained the diameter measurements $\{Y_t; t = 1, 2, \ldots, 200\}$ for the 200 samples.

Using Joe's method, the estimates are obtained as $\hat{\mu} = 74.0036$, $\hat{\sigma} = 0.0115$, and $\hat{\alpha} = 0.1422$ that corresponds to Kendall's tau = 0.0664. Therefore, the data exhibit weak positive dependence. In the last step of the Newton-Raphson algorithm, we examine the gradient

$$\left[\frac{\partial L}{\partial \mu} \quad \frac{\partial L}{\partial \sigma} \quad \frac{\partial L}{\partial \alpha}\right]\Bigg|_{(\hat{\mu},\hat{\sigma},\hat{\alpha})} = \left[-1.517216\times10^{-9} \quad 2.072281\times10^{-9} \quad 2.248266\times10^{-11}\right],$$

and the Hessian matrix

$$\left[\begin{array}{ccc} \frac{\partial^2 L}{\partial\mu\partial\mu} & \frac{\partial^2 L}{\partial\mu\partial\sigma} & \frac{\partial^2 L}{\partial\mu\partial\alpha} \\ \frac{\partial^2 L}{\partial\sigma\partial\mu} & \frac{\partial^2 L}{\partial\sigma\partial\sigma} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} \\ \frac{\partial^2 L}{\partial\mu\partial\alpha} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} & \frac{\partial^2 L}{\partial\alpha\partial\alpha} \end{array}\right]\Bigg|_{(\hat{\mu},\hat{\sigma},\hat{\alpha})} = \left[\begin{array}{ccc} -6108.555329 & -646.07069 & -3.2773394 \\ -646.070688 & -15025.21851 & 26.6075763 \\ -3.277339 & 26.60758 & -0.4012899 \end{array}\right].$$

Since all the eigenvalues is negative, the Hessian matrix is negative definite. Hence, $(\hat{\mu}, \hat{\sigma}, \hat{\alpha})$ is a local maxima. We have confirmed the global uniqueness of the MLE by drawing the likelihood functions. **Figure 2** shows that $(\hat{\mu}, \hat{\sigma}, \hat{\alpha})$ attains the maximum.

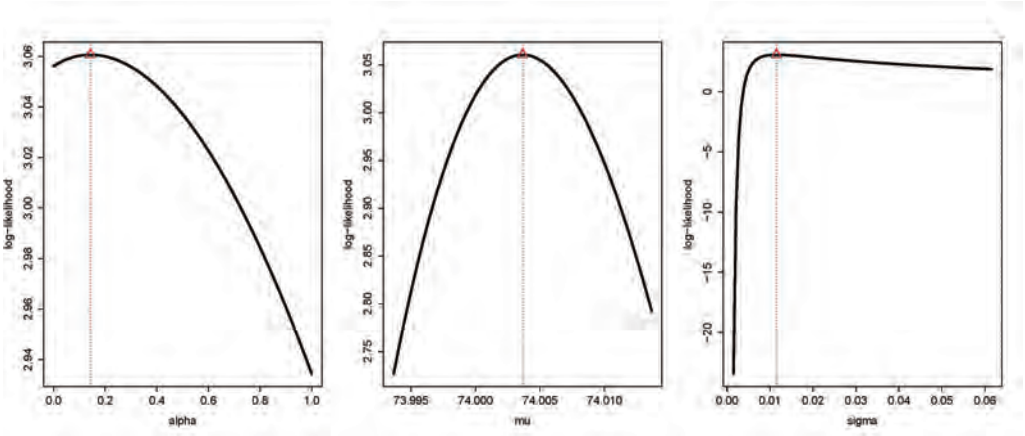

Fig. 2. The likelihood function for the piston rings data (Montgomery 2009b). The vertical line signifies the MLE $\hat{\alpha}$=0.1422, $\hat{\mu}$=74.0036, and $\hat{\sigma}$=0.0115.

The resultant control chart is displayed in **Figure 3**. The MLE produces UCL $(\hat{\mu} + 3\hat{\sigma})$=74.0381 and LCL $(\hat{\mu} + 3\hat{\sigma})$=73.9691. Only one point, corresponding to the 67^{th} observation, gives the out-of-control signal, and all the others fall between LCL and UCL. Therefore, some assignable cause that changes the mean should be investigated for the 67^{th} sample.

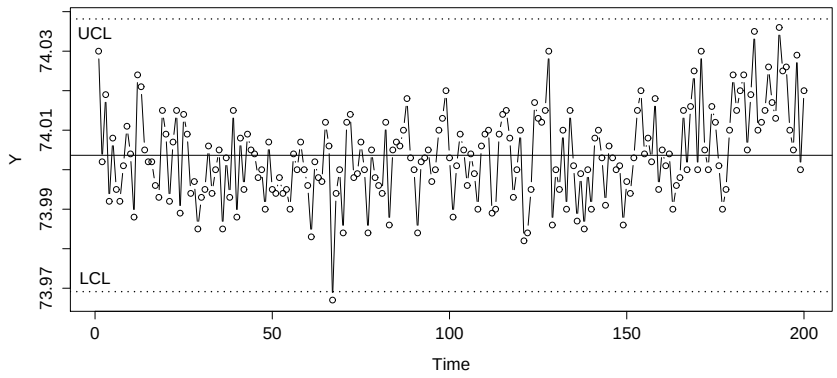

Fig. 3. The control chart using the piston rings data (Montgomery 2009b). The center line represents the estimated mean $\hat{\mu}$, and the other two straight lines are UCL ($\hat{\mu} + 3\hat{\sigma}$) and LCL ($\hat{\mu} + 3\hat{\sigma}$), which are obtained by Joe's method under the Clayton copula.

Under the estimated parameters, we calculate the ARL under the Clayton copula with α=0.1535 and the normal distribution with μ=74.0036 and σ=0.0115. Here, the 3-sigma limits are UCL ($\mu+3\sigma$)=74.0381 and LCL ($\mu+3\sigma$)=73.9691. Using Algorithm 2 (Monte Carlo) with $m = 10000$, we obtain the ARL = 382.442 (se = 3.885). Suppose that one wishes to have a control chart with ARL = 370. Accordingly, we reduce the coefficient from 3 to 2.99. Then, the choice UCL ($\mu + 2.99\sigma$)=74.0380 and LCL ($\mu - 2.99\sigma$) = 73.9693 achieves the desired ARL = 371.155 (se = 3.767).

7. Conclusion and Discussion

This paper provides a framework for performing statistical process control using copula-based Markov chain models. Although the copula-based Markov models have been utilized to many different fields, the application to statistical process control has not been considered in the literature. In particular, we demonstrate how to apply Joe's method, Chen and Fan's method, and the standard method for calculating the control limits, and then compare their performance via simulations. The results show that Joe's method performs best in terms of accuracy of the estimated control limits and the average run length with estimated parameters, when the model assumptions are adequate. Hence we propose to use Joe's method for the application to statistical process control. For illustration, we demonstrate the usage of Joe's method for diameter measurements of piston rings data.

We also propose simulation techniques to calculate the average run length of the proposed control charts. The Monte Carlo method and antithetic variables method are presented, where the latter reduces by half the number of generating uniform random numbers. It is demonstrated through the data analysis that the algorithms are useful when one wishes to set copula-based control limits for a given value of the average run length.

Although we have applied a copula-based Markov model for a serially correlated data, there are many cases where two series of correlated data are available. Specifically, suppose that one observe two quality characteristics, say $\{X_t : t = 1, \ldots, n\}$ and $\{X_t : t = 1, \ldots, n\}$. If there is no serial correlation within each series, the correlation between the two series is modeled by a bivariate normal distribution. Then, the simultaneous monitoring of the two series is performed by a control ellipse or Hotelling T^2-chart (Chap.11 of Montgomery, 2009a). These approaches must be modified to take into account serial correlation in the two series. In the presence of two series, one is not only interested in monitoring the process mean, but also the association between the two series. For instance, the inner diameter X_t may change the outer diameter Y_t of some parts. Such monitoring schemes have not been considered in the SPC context, but there are a rich literature studying on the causal relationship between two series. We only mention that many methods to study the causal relationship between the series have been proposed (Hung and Tseng, 2012 and references therein).

The extension of the copula-based process control to the discrete variables is an important direction for future research. Due to Assumption I, the method presented in this paper is only applicable to continuous margins. However, the well-known np-control chart and c-control chart assume that the observations follow independent binomial distribution and Poisson distribution, respectively (Wetherill and Brown 1991; Montgomery 2009a, b). The copula approach to incorporate the dependence is a challenging but interesting topic since estimation under the copula models with discrete margins are relatively new. The difficulty comes from the fact that the correlation parameter for copula models may be affected by the marginal distributions [see Nešlehová (2007) for binomial margins and Genest, Nešlehová and Rémillard (2013) for Poisson margins]. Another challenge in the np-control chart is that it requires large n and not too small p (Emura and Lin, 2013). The copula approach to the discrete cases will be a promising topic for research.

One of important issues that we did not discuss in this paper is the goodness-of-fit of a given copula. We choose the Clayton copula for its popularity in applications and mathematical tractability. Obviously, there are many other choices, such as Frank, Gumbel, Gaussian copulas (Nelsen, 2006). Many of the goodness-of-fit methods for parametric models use distance statistics, such as the Kolmogorov-Smirnov statistic and Cramér-von Mises statistic. The asymptotic distribution of such statistics under the null model typically requires the empirical process techniques (Genest, Rémillard, and Beaudoin, 2009; Emura and Konno, 2012), which needs further study.

Acknowledgments

We would like to thank the editor, associate editor and two anonymous reviewers for their helpful comments that greatly improved the manuscript. This work was financially supported by the National Science Council of Taiwan (NSC101-2118-M008-002-MY2).

Appendices

A.1 Log-density for the Clayton copula

The density of the Clayton copula is given by

$$c(u_1, u_2; \alpha) = \partial^2 C(u_1, u_2; \alpha)/\partial u_1 \partial u_2 = (1+\alpha) u_1^{-(1+\alpha)} u_2^{-(1+\alpha)} [\, u_1^{-\alpha} + u_2^{-\alpha} - 1\,]^{-(\alpha^{-1}+2)} \quad , \quad \alpha > 0 .$$

Hence, the log-copula density is:

$$l(u_1, u_2; \alpha) = \log(1+\alpha) - (1+\alpha)\log u_1 - (1+\alpha)\log u_2 - \left(\frac{1}{\alpha} + 2\right)\log(u_1^{-\alpha} + u_2^{-\alpha} - 1).$$

A.2 Likelihood function and its first and second derivatives

Let $u_{t-1} = \phi\{(Y_{t-1} - \mu)/\sigma\}$ and $U_{t-1} = \Phi\{(Y_{t-1} - \mu)/\sigma\}$. Using the formulas of Appendix A.1, the likelihood function of Section 3.2 is rewritten as

$$L(\mu, \sigma, \alpha) = -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{1}{n}\sum_{t=1}^{n}\frac{(Y_t - \mu)^2}{2\sigma^2} + \frac{1}{n}\sum_{t=2}^{n}\left[\, \log(1+\alpha) - (1+\alpha)\log U_{t-1} - (1+\alpha)\log U_t - \left(\frac{1}{\alpha} + 2\right)\log(U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1)\right].$$

Hence, its derivatives are

$$\frac{\partial L(\mu, \sigma, \alpha)}{\partial \mu} = \frac{1}{n}\sum_{t=1}^{n}\frac{Y_t - \mu}{\sigma^2} + \frac{1}{n}\sum_{t=2}^{n}\left[\, \frac{1+\alpha}{\sigma}\left(\frac{u_{t-1}}{U_{t-1}} + \frac{u_t}{U_t}\right) - \frac{2\alpha+1}{\sigma}\left(\frac{U_{t-1}^{-(\alpha+1)} u_{t-1} + U_t^{-(\alpha+1)} u_t}{U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1}\right)\right],$$

$$\frac{\partial L(\mu, \sigma, \alpha)}{\partial \sigma} = \frac{1}{n}\sum_{t=1}^{n}\left[\, \frac{(Y_t - \mu)^2}{\sigma^3} - \frac{1}{\sigma}\right] + \frac{1}{n}\sum_{t=2}^{n}\left\{ (1+\alpha)\left[\left(\frac{Y_{t-1} - \mu}{\sigma^2}\right)\frac{u_{t-1}}{U_{t-1}} + \left(\frac{Y_t - \mu}{\sigma^2}\right)\frac{u_t}{U_t}\right]\right\} + \frac{1}{n}\sum_{t=2}^{n}\left\{ -(1+2\alpha)\frac{\{(Y_{t-1} - \mu)/\sigma^2\} U_{t-1}^{-(1+\alpha)} u_{t-1} + \{(Y_t - \mu)/\sigma^2\} U_t^{-(1+\alpha)} u_t}{U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1}\right\},$$

$$\frac{\partial L(\mu, \sigma, \alpha)}{\partial \alpha} = \frac{1}{n}\sum_{t=2}^{n}\left\{ \frac{1}{1+\alpha} - \log(U_{t-1} U_t) + \frac{\log(U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1)}{\alpha^2}\right\} + \frac{1}{n}\sum_{t=2}^{n}\left\{ \left(2 + \frac{1}{\alpha}\right)\frac{U_{t-1}^{-\alpha}\log U_{t-1} + U_t^{-\alpha}\log U_t}{U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1}\right\},$$

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\mu^2}=-\frac{1}{\sigma^2}+\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{1+\alpha}{\sigma}\left[\frac{\{(Y_{t-1}-\mu)/\sigma^2\}U_{t-1}u_{t-1}+u_{t-1}^2/\sigma}{U_{t-1}^{\ 2}}+\frac{\{(Y_t-\mu)/\sigma^2\}U_tu_t+u_t^2/\sigma}{U_t^2}\right]\right\}$$
$$-\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{1+2\alpha}{\sigma(U_{t-1}^{\ -\alpha}+U_t^{-\alpha}-1)^2}\left[H_1(U_{t-1}^{\ -\alpha}+U_t^{-\alpha}-1)-\frac{\alpha}{\sigma}(U_{t-1}^{\ -(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t)^2\right]\right\},$$

where

$$H_1=\frac{1+\alpha}{\sigma}U_{t-1}^{\ -(2+\alpha)}u_{t-1}^{\ 2}+\left(\frac{Y_{t-1}-\mu}{\sigma^2}\right)U_{t-1}^{-(1+\alpha)}u_{t-1}+\frac{1+\alpha}{\sigma}U_t^{\ -(2+\alpha)}u_t^{\ 2}+\left(\frac{Y_t-\mu}{\sigma^2}\right)U_t^{-(1+\alpha)}u_t.$$

In addition,

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\sigma^2}=\frac{1}{n}\sum_{t=1}^{n}\left\{-3\frac{(Y_t-\mu)^2}{\sigma^4}+\frac{1}{\sigma^2}\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{(1+\alpha)\left(\frac{Y_{t-1}-\mu}{\sigma^3}\right)\left(\frac{u_{t-1}}{U_{t-1}}\right)\left(-2+\frac{(Y_{t-1}-\mu)^2}{\sigma^2}+\left(\frac{Y_{t-1}-\mu}{\sigma}\right)\left(\frac{u_{t-1}}{U_{t-1}}\right)\right)\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{(1+\alpha)\left(\frac{Y_t-\mu}{\sigma^3}\right)\left(\frac{u_t}{U_t}\right)\left(-2+\frac{(Y_t-\mu)^2}{\sigma^2}+\left(\frac{Y_t-\mu}{\sigma}\right)\left(\frac{u_t}{U_t}\right)\right)\right\}$$
$$-\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{(1+2\alpha)}{(U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1)^2}K_1(U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1)-\alpha K_2\right\},$$

where

$$K_1=\frac{Y_{t-1}-\mu}{\sigma^3}U_{t-1}^{-(1+\alpha)}u_{t-1}\left[-2+(1+\alpha)\left(\frac{u_{t-1}}{U_{t-1}}\right)\frac{Y_{t-1}-\mu}{\sigma}+\left(\frac{Y_{t-1}-\mu}{\sigma}\right)^2\right]$$
$$+\frac{Y_t-\mu}{\sigma^3}U_t^{-(1+\alpha)}u_t\left[-2+(1+\alpha)\left(\frac{u_t}{U_t}\right)\frac{Y_t-\mu}{\sigma}+\left(\frac{Y_t-\mu}{\sigma}\right)^2\right],$$

and $K_2=\{(Y_{t-1}-\mu)/\sigma^2\}\{U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t\}$. Finaly,

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\alpha^2}=\frac{1}{n}\sum_{t=2}^{n}\left\{-\frac{1}{(1+\alpha)^2}-\frac{2}{\alpha^3}\log(U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1)-\frac{2(U_{t-1}^{-\alpha}\log U_{t-1}+U_t^{-\alpha}\log U_t)}{\alpha^2(U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1)}\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{\left(2+\frac{1}{\alpha}\right)\left(\frac{(U_{t-1}^{-\alpha}\log U_{t-1}+U_t^{-\alpha}\log U_t)^2}{U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1}\right)-\frac{U_{t-1}^{-\alpha}(\log U_{t-1})^2+U_t^{-\alpha}(\log U_t)^2}{U_{t-1}^{\ -\alpha}+U_t^{\ -\alpha}-1}\right\},$$

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\mu\partial\sigma} = -\frac{2}{n}\sum_{t=1}^{n}\frac{Y_t-\mu}{\sigma^3}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{-\frac{\alpha+1}{\sigma^2}\left(\frac{u_{t-1}}{U_{t-1}}+\frac{u_t}{U_t}\right)+\frac{\alpha+1}{\sigma}\left[\frac{(Y_{t-1}-\mu)^2u_{t-1}U_{t-1}/\sigma^3+(Y_{t-1}-\mu)u_{t-1}^2/\sigma^2}{U_{t-1}^2}\right]\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{\alpha+1}{\sigma}\left[\frac{(Y_t-\mu)^2u_tU_t/\sigma^3+(Y_{t-1}-\mu)u_t^2/\sigma^2}{U_t^2}\right]+\frac{2\alpha+1}{\sigma^2}\left(\frac{U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right)\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{-\frac{2\alpha+1}{\sigma(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)}\left[(\alpha+1)U_{t-1}^{-(2+\alpha)}(Y_{t-1}-\mu)u_{t-1}^2/\sigma^2+U_{t-1}^{-(1+\alpha)}(Y_{t-1}-\mu)^2u_{t-1}/\sigma^3\right]\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{-\frac{2\alpha+1}{\sigma(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)}\left[(\alpha+1)U_t^{-(2+\alpha)}\frac{Y_t-\mu}{\sigma^2}u_t^2+U_t^{-(1+\alpha)}\frac{(Y_t-\mu)^2}{\sigma^3}u_t\right]\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{2\alpha+1}{\sigma(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)^2}\left[\alpha(U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t)\left(\frac{Y_{t-1}-\mu}{\sigma^2}U_{t-1}^{-(1+\alpha)}u_{t-1}\right)\right]\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{\frac{2\alpha+1}{\sigma(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)^2}\left[\alpha(U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t)\left(\frac{Y_t-\mu}{\sigma^2}U_t^{-(1+\alpha)}u_t\right)\right]\right\},$$

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\mu\partial\alpha} = \frac{1}{n}\sum_{t=2}^{n}\left\{\frac{1}{\sigma}\left(\frac{u_{t-1}}{U_{t-1}}+\frac{u_t}{U_t}\right)-\frac{2}{\sigma}\left(\frac{U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right)\right\}$$
$$-\frac{2\alpha+1}{n\sigma}\sum_{t=2}^{n}\left\{\frac{-U_{t-1}^{-(1+\alpha)}\log(U_{t-1})u_{t-1}-U_t^{-(1+\alpha)}\log(U_t)u_t}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right\}$$
$$-\frac{2\alpha+1}{n\sigma}\sum_{t=2}^{n}\left\{\frac{[-U_{t-1}^{-\alpha}\log(U_{t-1})-U_t^{-\alpha}\log(U_t)](U_{t-1}^{-(1+\alpha)}u_{t-1}+U_t^{-(1+\alpha)}u_t)}{(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)^2}\right\},$$

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\sigma\partial\alpha} = \frac{1}{n}\sum_{t=2}^{n}\left\{\frac{Y_{t-1}-\mu}{\sigma^2}\left(\frac{u_{t-1}}{U_{t-1}}\right)+\frac{Y_t-\mu}{\sigma^2}\left(\frac{u_t}{U_t}\right)\right\}$$
$$+\frac{1}{n}\sum_{t=2}^{n}\left\{-2\frac{(Y_{t-1}-\mu)U_{t-1}^{-(1+\alpha)}u_{t-1}/\sigma^2+(Y_t-\mu)U_t^{-(1+\alpha)}u_t/\sigma^2}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right\}$$
$$-\frac{2\alpha+1}{n}\sum_{t=2}^{n}\left\{\frac{-(Y_{t-1}-\mu)U_{t-1}^{-(1+\alpha)}u_{t-1}\log(U_{t-1})/\sigma^2}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right\}$$
$$-\frac{2\alpha+1}{n}\sum_{t=2}^{n}\left\{\frac{-(Y_t-\mu)U_t^{-(1+\alpha)}u_t\log(U_t)}{U_{t-1}^{-\alpha}+U_t^{-\alpha}-1}\right\}$$
$$-\frac{2\alpha+1}{n}\sum_{t=2}^{n}\left\{\frac{[(Y_{t-1}-\mu)U_{t-1}^{-(1+\alpha)}u_{t-1}/\sigma^2][U_{t-1}^{-\alpha}\log(U_{t-1})+U_t^{-\alpha}\log(U_t)]}{(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)^2}\right\}$$
$$-\frac{2\alpha+1}{n}\sum_{t=2}^{n}\left\{\frac{[(Y_t-\mu)U_t^{-(1+\alpha)}u_t/\sigma^2][U_{t-1}^{-\alpha}\log(U_{t-1})+U_t^{-\alpha}\log(U_t)]}{(U_{t-1}^{-\alpha}+U_t^{-\alpha}-1)^2}\right\}.$$

Table 1 Simulation results for $\hat{\mu}$ based on 1000 repetitions.

		Method 1 (Joe)	Method 2 (Chen & Fan)	Method 3 (Standard)
$\mu=1,\ \sigma=1,\ \alpha=2\ (\tau=0.5)$				
$n=300$	E($\hat{\mu}$)	0.9899	0.9911	0.9944
	Bias($\hat{\mu}$)	-0.0101	-0.0089	-0.0056
	MSE($\hat{\mu}$)	0.0271	0.0299	0.0301
$n=600$	E($\hat{\mu}$)	0.9929	0.9946	0.9963
	Bias($\hat{\mu}$)	-0.0071	-0.0054	-0.0037
	MSE($\hat{\mu}$)	0.0114	0.0126	0.0126
$n=1000$	E($\hat{\mu}$)	1.0000	1.0000	1.0010
	Bias($\hat{\mu}$)	0.0000	0.0000	0.0010
	MSE($\hat{\mu}$)	0.0079	0.0091	0.0091
$\mu=1,\ \sigma=1,\ \alpha=8\ (\tau=0.8)$				
$n=300$	E($\hat{\mu}$)	1.0161	1.0272	1.0306
	Bias($\hat{\mu}$)	0.0161	0.0272	0.0306
	MSE($\hat{\mu}$)	0.1761	0.2283	0.2300
$n=600$	E($\hat{\mu}$)	1.0055	1.0201	1.0218
	Bias($\hat{\mu}$)	0.0055	0.0201	0.0218
	MSE($\hat{\mu}$)	0.0783	0.1024	0.1028
$n=1000$	E($\hat{\mu}$)	0.9619	0.9702	0.9712
	Bias($\hat{\mu}$)	-0.0381	-0.0298	-0.0288
	MSE($\hat{\mu}$)	0.0676	0.0909	0.0910
$\mu=1,\ \sigma=1,\ \alpha=-1/3\ (\tau=-0.2)$				
$n=300$	E($\hat{\mu}$)	0.9983	0.9961	0.9994
	Bias($\hat{\mu}$)	-0.0017	-0.0039	-0.0006
	MSE($\hat{\mu}$)	0.0016	0.0016	0.0015
$n=600$	E($\hat{\mu}$)	1.0027	1.0016	1.0033
	Bias($\hat{\mu}$)	0.0027	0.0016	0.0033
	MSE($\hat{\mu}$)	0.0008	0.0008	0.0008
$n=1000$	E($\hat{\mu}$)	0.9989	0.9984	0.9994
	Bias($\hat{\mu}$)	-0.0011	-0.0016	-0.0006
	MSE($\hat{\mu}$)	0.0005	0.0005	0.0005

Table 2 Simulation results for $\hat{\sigma}$ based on 1000 repetitions.

		Method 1 (Joe)	Method 2 (Chen & Fan)	Method 3 (Standard)
$\mu=1,\ \sigma=1,\ \alpha=2\ (\tau=0.5)$				
$n=300$	E($\hat{\sigma}$)	0.9982	0.9842	0.9841
	Bias($\hat{\sigma}$)	-0.0018	-0.0158	-0.0159
	MSE($\hat{\sigma}$)	0.0060	0.0098	0.0100
$n=600$	E($\hat{\sigma}$)	0.9948	0.9882	0.9882
	Bias($\hat{\sigma}$)	-0.0052	-0.0118	-0.0118
	MSE($\hat{\sigma}$)	0.0027	0.0051	0.0051
$n=1000$	E($\hat{\sigma}$)	0.9955	0.9908	0.9907
	Bias($\hat{\sigma}$)	-0.0045	-0.0092	-0.0093
	MSE($\hat{\sigma}$)	0.0018	0.0033	0.0033
$\mu=1,\ \sigma=1,\ \alpha=8\ (\tau=0.8)$				
$n=300$	E($\hat{\sigma}$)	0.9371	0.8521	0.8508
	Bias($\hat{\sigma}$)	-0.0629	-0.1479	-0.1492
	MSE($\hat{\sigma}$)	0.0242	0.0444	0.0451
$n=600$	E($\hat{\sigma}$)	0.9857	0.9186	0.9182
	Bias($\hat{\sigma}$)	-0.0143	-0.0814	-0.0818
	MSE($\hat{\sigma}$)	0.0179	0.0271	0.0273
$n=1000$	E($\hat{\sigma}$)	1.0098	0.9537	0.9536
	Bias($\hat{\sigma}$)	0.0098	-0.0463	-0.0464
	MSE($\hat{\sigma}$)	0.0115	0.0212	0.0213
$\mu=1,\ \sigma=1,\ \alpha=-1/3\ (\tau=-0.2)$				
$n=300$	E($\hat{\sigma}$)	1.0042	1.0008	1.0009
	Bias($\hat{\sigma}$)	0.0042	0.0008	0.0009
	MSE($\hat{\sigma}$)	0.0019	0.0019	0.0019
$n=600$	E($\hat{\sigma}$)	1.0058	1.0046	1.0046
	Bias($\hat{\sigma}$)	0.0058	0.0046	0.0046
	MSE($\hat{\sigma}$)	0.0012	0.0011	0.0011
$n=1000$	E($\hat{\sigma}$)	1.0046	1.0039	1.0039
	Bias($\hat{\sigma}$)	0.0046	0.0039	0.0039
	MSE($\hat{\sigma}$)	0.0007	0.0006	0.0006

Table 3 Simulation results for UCL= $\hat{\mu}+3\hat{\sigma}$ based on 1000 repetitions.

		Method 1 (Joe)	Method 2 (Chen & Fan)	Method 3 (Standard)
$\mu=1,\ \sigma=1,\ \alpha=2\ (\tau=0.5),$ UCL=4				
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	3.9845	3.9438	3.9467
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0155	-0.0562	-0.0533
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0320	0.0585	0.0585
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	3.9773	3.9592	3.9608
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0227	-0.0408	-0.0392
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0152	0.0304	0.0304
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	3.9863	3.9723	3.9732
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0137	-0.0277	-0.0268
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0092	0.0184	0.0184
$\mu=1,\ \sigma=1,\ \alpha=8\ (\tau=0.8),$ UCL=4				
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	3.8275	3.5834	3.5830
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.1725	-0.4166	-0.4170
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.3294	0.5220	0.5241
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	3.9627	3.7758	3.7765
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0373	-0.2242	-0.2235
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0789	0.1740	0.1738
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	3.9914	3.8314	3.8321
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0086	-0.1686	-0.1679
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0186	0.1083	0.1082
$\mu=1,\ \sigma=1,\ \alpha=-1/3\ (\tau=-0.2),$ UCL=4				
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	4.0108	3.9987	4.0020
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0108	-0.0013	0.0020
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0202	0.0199	0.0200
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	4.0200	4.0153	4.0169
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0200	0.0153	0.0169
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0125	0.0121	0.0121
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	4.0126	4.0102	4.0112
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0126	0.0102	0.0112
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0073	0.0070	0.0070

Table 4 Simulations under a misspecified model for UCL= $\hat{\mu}+3\hat{\sigma}$ based on 1000 repetitions.

		Method 1 (Joe)	Method 2 (Chen & Fan)	Method 3 (Standard)
$\mu=1,\ \sigma=1,\ \alpha=2\ (\tau=0.5),$		UCL= $\mu+3\sigma$ =4		
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	4.0287	3.9233	3.9261
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0287	-0.0767	-0.0739
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0608	0.0860	0.0861
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	4.0599	3.9726	3.9741
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0599	-0.0274	-0.0259
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0453	0.0581	0.0582
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	4.0479	3.9804	3.9814
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0479	-0.0196	-0.0186
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0202	0.0305	0.0305
$\mu=1,\ \sigma=1,\ \alpha=8\ (\tau=0.8),$		UCL= $\mu+3\sigma$ =4		
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	3.7583	3.5208	3.5206
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.2417	-0.4792	-0.4794
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.3948	0.5474	0.5478
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	3.9351	3.7161	3.7167
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0649	-0.2839	-0.2833
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.1772	0.1962	0.1961
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	3.9611	3.7946	3.7952
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0389	-0.2054	-0.2048
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0929	0.1507	0.1506
$\mu=1,\ \sigma=1,\ \alpha=-1/3\ (\tau=-0.2),$		UCL=4		
$n=300$	E($\hat{\mu}+3\hat{\sigma}$)	3.9880	3.9826	3.9859
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0120	-0.0174	-0.0141
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0280	0.0278	0.0279
$n=600$	E($\hat{\mu}+3\hat{\sigma}$)	4.0021	4.0060	4.0077
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0021	0.0060	0.0077
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0175	0.0169	0.0170
$n=1000$	E($\hat{\mu}+3\hat{\sigma}$)	3.9963	4.0021	4.0031
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0037	0.0021	0.0031
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0100	0.0092	0.0092

Table 5 Calculation of ARL based on Monte Carlo and antithetic variable methods with $2m = 20000$ repetitions.

(μ, σ)		Monte Carlo		Antithetic		cor(A, B)	Ratio[a]
		ARL	SD	ARL	SD		
$(0, 1)$	$\alpha = 2$[b]	620.930	632.505	616.383	627.166	0.0714	1.008
	$\alpha = 8$[c]	763.152	772.725	753.803	755.466	-0.0515	1.022
$(0, 3)$	$\alpha = 2$	621.307	616.196	615.454	613.850	0.0508	1.003
	$\alpha = 8$	764.426	771.462	754.856	753.365	-0.0536	1.024
$(1, 1)$	$\alpha = 2$	616.695	615.549	605.901	606.064	0.0532	1.015
	$\alpha = 8$	759.815	748.315	760.938	757.564	-0.0515	0.987
$(1, 3)$	$\alpha = 2$	613.520	622.567	615.391	615.253	0.0445	1.011
	$\alpha = 8$	762.030	761.094	753.432	751.918	-0.0786	1.012
$(2, 1)$	$\alpha = 2$	623.782	638.164	629.492	633.717	0.0577	1.007
	$\alpha = 8$	762.654	767.709	764.261	770.023	-0.0423	0.997
$(2, 3)$	$\alpha = 2$	616.641	613.566	620.629	616.973	0.0394	0.994
	$\alpha = 8$	766.410	758.899	781.132	772.514	-0.0430	0.982

[a] Ratio > 1 corresponds to better performance of the antithetic variables method;

Ratio = (SD for Monte Carlo) / (SD for Antithetic variables).

[b] $\alpha = 2$ corresponds to $\tau = 0.5$.

[c] $\alpha = 8$ corresponds to $\tau = 0.8$.

Table 6 Calculation of ARL based on Monte Carlo and antithetic variable methods with $2m$ = 20000 repetitions. One-sided case.

(μ, σ)		Monte Carlo		Antithetic		cor(A, B)	Ratio[a]
		ARL	SD	ARL	SD		
$(0, 1)$	$\alpha = 2$[b]	748.477	750.8413	744.3434	742.4753	-0.0219	1.0112
	$\alpha = 8$[c]	786.569	793.4528	796.9079	815.5159	-0.0867	0.972
$(0, 3)$	$\alpha = 2$	743.888	732.300	738.777	736.080	-0.0201	0.994
	$\alpha = 8$	783.669	773.330	776.231	777.225	-0.0995	0.994
$(1, 1)$	$\alpha = 2$	746.228	743.185	745.731	746.343	-0.0020	0.995
	$\alpha = 8$	786.765	783.535	787.907	784.702	-0.1082	0.998
$(1, 3)$	$\alpha = 2$	752.232	759.681	751.409	770.731	-0.0229	0.985
	$\alpha = 8$	784.119	787.141	791.712	786.279	-0.0731	1.001
$(2, 1)$	$\alpha = 2$	742.697	735.749	749.081	744.173	-0.0129	0.988
	$\alpha = 8$	784.096	790.829	785.912	779.862	-0.0856	1.014
$(2, 3)$	$\alpha = 2$	744.671	743.750	740.998	736.727	-0.0477	1.009
	$\alpha = 8$	777.763	793.051	777.577	793.779	-0.0771	0.999

[a] Ratio > 1 corresponds to better performance of the antithetic variables method;

Ratio = (SD for Monte Carlo) / (SD for Antithetic variables).

[b] $\alpha = 2$ corresponds to $\tau = 0.5$.

[c] $\alpha = 8$ corresponds to $\tau = 0.8$.

Table 7 Monte Carlo values for the ARL under the Clayton copula with the marginal being $N(\mu, \sigma = 1)$. The UCL = +3 and LCL = -3 are fixed. The in-control state is $\mu = 0$ while the out-of-control states are $\mu = 1$ (1σ shift) and $\mu = 2$ (2σ shift).

Kendall's tau $\tau = \alpha/(\alpha + 2)$	In-control ARL No shift: $\mu = 0$	Out-of-control ARL Shift: $\mu = 1$	Out-of-control ARL Shift: $\mu = 2$
0.9 ($\alpha = 18$)	934.598	255.900	184.529
0.8 ($\alpha = 8$)	766.300	91.150	45.126
0.5 ($\alpha = 2$)	632.918	49.151	10.107
0.3 ($\alpha = 6/7$)	505.197	45.168	7.520
0.1 ($\alpha = 2/9$)	390.536	44.386	6.589
0.0001($\alpha = 0.0002$)	373.174	44.106	6.356

Table 8 Monte Carlo values for the ARL under the Clayton copula, where UCL=+3 and LCL=-3 are known or they are estimated by the three methods.

Kendall's tau	known	Estimated by Joe's method	Estimated by Chen & Fan's method	Estimated by standard method
The true model				
0.8 ($\alpha = 8$)	766.656	748.436	501.773	502.175
0.5 ($\alpha = 2$)	627.335	690.920	736.589	738.521
0.3 ($\alpha = 6/7$)	500.518	545.310	579.958	581.683
0.1 ($\alpha = 2/9$)	383.470	412.027	413.103	414.137
0.0001($\alpha = 0.0002$)	365.957	385.332	385.056	385.496
Misspecified model in Phase I				
0.8 ($\alpha = 8$)	766.661	841.593	511.873	513.664
0.5 ($\alpha = 2$)	626.859	900.439	921.937	924.302
0.3 ($\alpha = 6/7$)	500.655	559.944	633.908	635.831
0.1 ($\alpha = 2/9$)	383.453	412.309	425.906	425.820
0.0001($\alpha = 0.0002$)	365.935	389.346	390.335	389.686

In the case of known parameters, the ARL is calculated under $\mu = 0$ and $\sigma = 1$ based on 10000 repetitions. Hence, the LCL = -3 and UCL = 3 are fixed. In the estimated parameters case, we use three different methods to estimate the UCL and LCL based on Phase I samples. We simulate 100 Phase I control limits, and for each of them we calculate the ARL based on 100 repetitions. The ARL in these estimated cases is the average of the 100 ARL's. The misspecified models in Phase I uses the *t*-distribution with degree of freedom $\nu = 10$.

References

Bagshow, M. and Johnson, R. A. (1975). The effect of serial correlation on the performance of CUSUM tests II. *Technometrics*, **17**, 73-80.

Box, G. and Narasimhan, S. (2010). Rethinking statistics for quality control. *Quality Engineering*, **22**, 60-72.

Chen, X. and Fan, Y. (2006). Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, **130**, 307-335.

Chen, C. C., Fuh, C. D. and Teng, H. W. (2013). Efficient option pricing with importance sampling. *Journal of the Chinese Statistical Association*, **51**, 253-273.

Darsow, W. F., Nguten, B. and Olsen, E. T. (1992). Copulas and Markov Processes. *Illinois Journal of Mathematics*, **36**, 600-642.

Emura, T. and Konno, Y. (2012). A goodness-of-fit tests for parametric models based on dependently truncated data. *Computational Statistics & Data Analysis*, **56**, 2237-2250.

Emura, T. and Chen, Y.H. (2014). Gene selection for survival data under dependent censoring: a copula-based approach. *Statistical Methods in Medical Research*, doi: 10.1177/0962280214533378.

Emura T. and Lin Y. S. (2013). A comparison of normal approximation rules for attribute control charts. *Quality and Reliability Engineering International*, doi: 10.1002/qre.1601.

Frees, E. W. and Valdez, E. (1998). Understanding the relationships using copulas. *North American Actuarial Journal*, **2**, 1-25.

Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de Institut Henri Poicare-Probabil ites et Statistiques*, **44**, 1096-1127.

Genest, C., Nešlehová, J.G. and Rémillard, B. (2013). On the estimation of Spearman's rho and related tests of independence for possibly discontinuous multivariate data. *Journal of Multivariate Analysis*, **117**, 217-228.

Hung, Y. C. and Tseng, N. F. (2013). Extracting informative variables in the validation of two-group causal relationship. *Computational Statistics*, **28**, 1151-1167.

Hu, Y. H. (2014). Maximum likelihood estimation for double-truncation data under a special exponential family, Master Thesis, Graduate Institute of Statistics, National Central University, Taiwan.

Hryniewicz, O. (2012). *On the robustness of the Shewhart control chart to different types of dependencies in data. Frontiers in Statistical Quality Control 10, Lenz, H.-J. et al. (Eds.)*, Springer-Verlag: Berlin Heidelberg.

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & hall.

Johnson, R. A. and Bagshaw, M. (1974). The effect of serial correlation on the performance of CUSUM tests. *Technometrics*, **16**, 103-112.

Knight, K. (2000). *Mathematical Statistics.* Chapman & Hall.

Knoth, S. and Schmid, W. (2004). *Control charts for time series: a review. Frontiers in Statistical Quality Control 7, Lenz, H.-J. et al. (Eds.)*, Springer-Verlag: Berlin Heidelberg.

Kramer, H. G. and Schmid, W. (2000). The influence of parameter estimation on the ARL of Shewhart type charts for time series. *Statistical Papers*, **41**, 173-196.

Luca L. Quality control charts, R qcc package, version 2.4. 2014.

Montgomery, D. C. (2009a). *Statistical Quality Control, Sixth Edition.* Wiley.

Montgomery, D. C. (2009b). *Introduction to Statistical Quality Control, Sixth Edition.* Wiley.

Nelsen, R. B. (2006). *An Introduction to Copulas, 2nd Edition.* Springer Series in Statistics, Springer-Verlag: New York.

Nešlehová, J. (2007). On rank correlation measures for non-continuous random variables. *Journal of Multivariate Analysis*, **98**, 544-567.

Psarakis, S. and Papaleonida, G. E. A. (2007). SPC procedures for monitoring autocorrelated processes. *Quality Techinology & Quantitative Management*, **4(4)**, 501-540.

Ross, S. M. (2013). *Simulation, Fifth Edition.* Elsevier.

Sari, J. K., Newby, M. J., Brombacher, A. C. and Tang, L. C. (2009). Bivariate constant stress degradation model: led lighting system reliability estimation with two-stage monitoring. *Quality and Reliability Engineering International*, **25**, 1067-1084.

Schmid, W. (1995). On the run length of a Shewhart chart for correlated data. *Statistical Papers*, **36**, 111-130.

Sklar, A. (1959). Fonctions de re'partition a'n dimensions et leurs marges. *Publications de l'Intitut de Statistique de l'Universit de Paris*, **8**, 229-231.

Vasilopoulos, A. V. and Stamboulis, A. P. (1978). Modification of control chart limits in the presence of data correlation. *J. Quality Technology*, **10(1)**, 20-30.

Wardell, D. G., Moskowitz, H. and Plante, R. D. (1994). Run-length distributions of special-cause control charts for correlated process. *Technometrics*, **36**, 3-27.

Wetherill G. B. and Brown D. W. (1991). *Statistical process control, theory and practice*. Chapman and Hall.

Wieringa, J. E. (1999). *Statistical process control for serially correlated data*. Labyrint Publishing.

[Received September 2013; accepted June 2014.]

Journal of the Chinese
Statistical Association
Vol. 52, (2014) 466–496

建立在 Copula 馬可夫鏈模型之下的管制圖的研究

龍庭軒　　江村剛志[1]
國立中央大學統計研究所

摘　要

統計製程管制 (Statistical process control) 對於製成商品以及操作服務品質而言, 是個非常重要且方便的品質管制工具。傳統的休哈特管制圖 (Shewhart control chart) 僅是應用在獨立性的假設之下。然而, 在現實生活中, 存在著許多相關性假設的資料, 因此, 傳統的管制圖在現實生活中常常是不被接受且不實用的。在本文中, 我們提供一個以 Copula 建立在馬可夫鏈之下的模型去衍生相關性資料分析。此外, 我們提供三種方法估計管制圖上的未知參數, 分別爲管制上限 (UCL) 以及管制下限 (LCL)。本文以統計模擬的方式驗證這些過程, 而喬 (Joe) 所提出的最大概似估計方法優於其餘兩者。接著, 我們提出了平均連串長度 (Average run length) 以用來表現出這些管制圖的性質。最後我們利用活塞環數據來進行實例分析。

關鍵詞: 平均連串長度, 克萊頓 (Clayton) 模型, 相關性資料, 肯德爾塔 (Kendall's tau), 馬可夫鏈。
JEL classification: C13, C15, C18. C22.