A Control Chart Based on Copula-Based Markov Time Series Models

Advisor: Takeshi Emura
Presenter: Ting-Hsuan Long
Stat. NCU
05/29/2013

Outline

1. Introduction
2. Methodology
3. Results and Analysis
4. Conclusion and Discussion

Outline

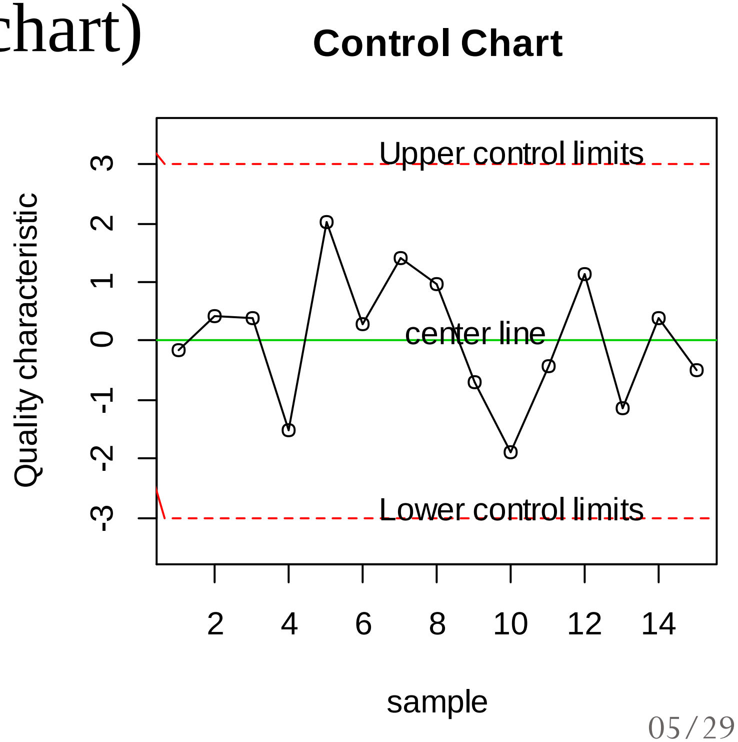
1. Introduction
2. Methodology
3. Results and Analysis
4. Conclusion and Discussion

Introduction-SPC

- Statistical process control (SPC) is an important tool.

➔ Walter Shewhart (1924) provided the most popular tools: (Shewhart control chart)

$\overline{X}$ control chart



Problem for Control Chart

- In tradition Shewhart chart:

  ➔ Only be used in i.i.d case.

- The early work starts with the papers by Johnson and Bagshaw (1974) and Bagshaw and Johnson (1975).

- From 1980s to now:

  ➔ Correlated data are discuss widely.

Introduction for Our Research

- Sklar (1959) proposed a multivariate distribution function with standard uniform marginal called "copula".
- Darsow et al (1992) proposed a copula-based Markov chain model to describe the serially correlated data.

Introduction for Our Research

- The performance of a control chart is evaluated by the average run length (ARL).

  ➔ Method: Monte Carlo. (c.f. Hryniewicz, 2012)

- In addition, we propose to use some variance reduction technique to increase the computational efficiency.

Outline

Methodology

1. Estimation of process parameters
    1. Existing method (Chen and Fan, 2005)
    2. Proposed method (Joe, 1997)
    3. Standard method
2. Control chart based on Copula

What is our purpose?

- We want to estimate the unknown distribution and parameters, especially, (μ, σ, α), which are important parameters in control chart.

1. Existing method

- Chen and Fan (2005) proposed a copula-based Markov chain to describe the dependence structure for time-series data.

Assumption1

- $\{Y_t : t = 1, \ldots, n\}$ is a sample of a stationary first-order Markov process generated from $(G^*(\cdot), C(\cdot,\cdot;\alpha^*))$,

  $G^*(\cdot)$: true invariant distribution .

  $C(\cdot,\cdot;\alpha^*)$: true parametric copula for (Y_{t-1}, Y_t) .

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

Notations

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

$h^*(\cdot \mid y_{t-1})$: the true conditional density function of Y_t given $Y_{t-1} = y_{t-1}$.

$c(\cdot,\cdot;\alpha^*)$: the copula density of $C(\cdot,\cdot;\alpha^*)$.

$g^*(\cdot)$: the density of the true marginal distribution $G^*(\cdot)$ which is unspecified.

Chen & Fan's method

- $\{U_t : U_t \equiv G^*(Y_t)\}$: a stationary parametric Markov process of order 1 in which the joint distribution of U_t and U_{t-1} is given by the copula $C(u_0, u_1; \alpha^*)$.
- $G_n(y) = \frac{1}{n+1}\sum_{t=1}^{n} I\{Y_t \le y\}$: the rescale ecdf.
- $L(\alpha) = \frac{1}{n}\sum_{t=1}^{n}\log g^*(Y_t) + \frac{1}{n}\sum_{t=2}^{n}\log c(G^*(Y_{t-1}), G^*(Y_t); \alpha)$: log-likelihood function.

1. Chen & Fan's method

- G^* Estimated by $G_n(\cdot)$
- $\hat{\alpha} = \arg\max_{\alpha} L_n(\alpha)$

where $L_n(\alpha)$ is $L(\alpha)$ with $G^*(\cdot)$ being replaced by $G_n(\cdot)$.

1. Chen & Fan's method

- We use Stieltjes integral to get the estimates as follows:

1. $\hat{\mu} = \int y dG_n(y) = \sum_{t=1}^{n} y_i [G_n(y_t) - G_n(y_t -)] = \dfrac{n}{n+1}\bar{Y}.$

2. $\hat{\sigma}^2 = \int y^2 dG_n(y) - [\int y dG_n(y)]^2$

$$= \sum_{t=1}^{n} y_t^2 [G_n(y_t) - G_n(y_t -)] - \hat{\mu}^2$$

$$= \frac{1}{n+1}\sum_{t=1}^{n} {y_t}^2 - \left(\frac{n}{n+1}\bar{Y}\right)^2.$$

2. Proposed method (Joe's method)

- We propose to set $G^*(y) = \Phi\left(\dfrac{y-\mu}{\sigma}\right)$,

  where Φ is a CDF of standard normal distribution.

- Hence, the log-likelihood is as follow:

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

2. Proposed method (Joe's method)

- We have to find MLE of (μ, σ, α), which is denoted by $(\hat{\mu}, \hat{\sigma}, \hat{\alpha})$.

- The MLE does not have closed form. This implies that we need to use computational algorithms for finding the MLE.

Newton-Raphson Algorithm

1. Choose the initial value $x_0 = (\mu_0, \sigma_0, \alpha_0)'$.

2. Set $\begin{bmatrix} \mu_{k+1} \\ \sigma_{k+1} \\ \alpha_{k+1} \end{bmatrix} = \begin{bmatrix} \mu_k \\ \sigma_k \\ \alpha_k \end{bmatrix} - \left. \begin{bmatrix} \frac{\partial^2 L}{\partial\mu\partial\mu} & \frac{\partial^2 L}{\partial\mu\partial\sigma} & \frac{\partial^2 L}{\partial\mu\partial\alpha} \\ \frac{\partial^2 L}{\partial\sigma\partial\mu} & \frac{\partial^2 L}{\partial\sigma\partial\sigma} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} \\ \frac{\partial^2 L}{\partial\mu\partial\alpha} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} & \frac{\partial^2 L}{\partial\alpha\partial\alpha} \end{bmatrix}^{-1} \right|_{(\mu_k, \sigma_k, \alpha_k)} \left. \begin{bmatrix} \frac{\partial L}{\partial\mu} \\ \frac{\partial L}{\partial\sigma} \\ \frac{\partial L}{\partial\alpha} \end{bmatrix} \right|_{(\mu_k, \sigma_k, \alpha_k)}, \quad k = 0, 1, \ldots$

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 $(\hat{\mu}, \hat{\sigma}, \hat{\alpha}) = (\mu_{k+1}, \sigma_{k+1}, \alpha_{k+1})$.

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

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

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

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\sigma^2} = \frac{1}{n}\sum_{t=1}^{n}\left\{ -3\frac{(Y_t-\mu)^2}{\sigma^4} + \frac{1}{\sigma^2} \right\}$$

$$+\frac{1}{n}\sum_{t=2}^{n}\left\{ (1+\alpha)\left(\frac{Y_{t-1}-\mu}{\sigma^3}\right)\left(\frac{u_{t-1}}{U_{t-1}}\right)\left(-2 + \frac{(Y_{t-1}-\mu)^2}{\sigma^2} + \left(\frac{Y_{t-1}-\mu}{\sigma}\right)\left(\frac{u_{t-1}}{U_{t-1}}\right) \right) \right\}$$

$$+\frac{1}{n}\sum_{t=2}^{n}\left\{ (1+\alpha)\left(\frac{Y_t-\mu}{\sigma^3}\right)\left(\frac{u_t}{U_t}\right)\left(-2 + \frac{(Y_t-\mu)^2}{\sigma^2} + \left(\frac{Y_t-\mu}{\sigma}\right)\left(\frac{u_t}{U_t}\right) \right) \right\}$$

$$-\frac{1}{n}\sum_{t=2}^{n}\left\{ \frac{(1+2\alpha)}{(U_{t-1}^{\ -\alpha} + U_t^{\ -\alpha} - 1)^2}K_1(U_{t-1}^{\ -\alpha} + U_t^{\ -\alpha} - 1) - \alpha K_2 \right\},$$

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

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

$$K_2 = \frac{Y_{t-1}-\mu}{\sigma^2}U_{t-1}^{-(1+\alpha)}u_{t-1} + \frac{Y_t-\mu}{\sigma^2}U_t^{-(1+\alpha)}u_t,$$

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

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

$$\frac{\partial^2 L(\mu,\sigma,\alpha)}{\partial\mu\partial\sigma}=\frac{-2}{n}\sum_{t=1}^{n}\frac{Y_t-\mu}{\sigma^3}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

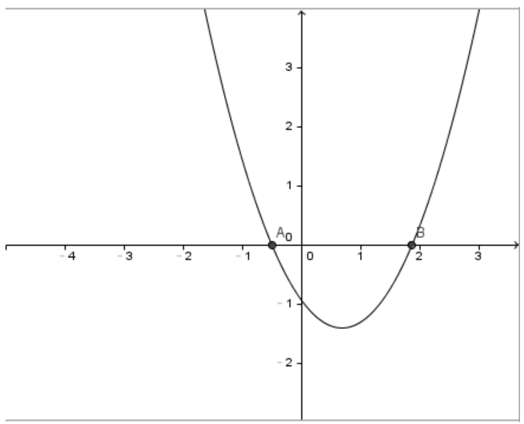
$$-\frac{2\alpha+1}{n}\sum_{t=2}^{n}\left\{ \frac{\left[\left(\frac{Y_{t-1}-\mu}{\sigma^2} \right)U_{t-1}^{-(1+\alpha)}u_{t-1}\right]\left[U_{t-1}^{-\alpha}\log(U_{t-1}) + U_t^{-\alpha}\log(U_t)\right]}{(U_{t-1}^{-\alpha} + U_t^{-\alpha} - 1)^2} \right\}$$

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

We encounter a problem for N-R!!!!!!

1. In our proposal, we also encounter the cases that the algorithm diverges.
2. The function maybe has not only one root.

Ex:



- Choosing a good initial value is very important.

Initial values

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

  where $\tau_0 = \binom{n}{2}^{-1} \sum_{i<j} \left[\operatorname{sgn}(Y_j - Y_i)\operatorname{sgn}(Y_{j+1} - Y_{i+1}) \right]$, which

  is the definition of Kendall's tau.

Modified Newton-Raphson Algorithm

1. Choose the initial value $x_0 = (\mu_0, \sigma_0, \alpha_0)'.$

2. Set
$$\begin{bmatrix} \mu_{k+1} \\ \sigma_{k+1} \\ \alpha_{k+1} \end{bmatrix} = \begin{bmatrix} \mu_k \\ \sigma_k \\ \alpha_k \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 L}{\partial\mu\partial\mu} & \frac{\partial^2 L}{\partial\mu\partial\sigma} & \frac{\partial^2 L}{\partial\mu\partial\alpha} \\ \frac{\partial^2 L}{\partial\sigma\partial\mu} & \frac{\partial^2 L}{\partial\sigma\partial\sigma} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} \\ \frac{\partial^2 L}{\partial\mu\partial\alpha} & \frac{\partial^2 L}{\partial\sigma\partial\alpha} & \frac{\partial^2 L}{\partial\alpha\partial\alpha} \end{bmatrix}^{-1} \Bigg|_{(\mu_k, \sigma_k, \alpha_k)} \begin{bmatrix} \frac{\partial L}{\partial\mu} \\ \frac{\partial L}{\partial\sigma} \\ \frac{\partial L}{\partial\alpha} \end{bmatrix} \Bigg|_{(\mu_k, \sigma_k, \alpha_k)}, \quad k = 0, 1, \ldots$$

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 $(\hat{\mu}, \hat{\sigma}, \hat{\alpha}) = (\mu_{k+1}, \sigma_{k+1}, \alpha_{k+1})$.

Modified Newton-Raphson Algorithm

3. If $|\mu_{k+1}-\mu_k|>10^{20}$ or $|\sigma_{k+1}-\sigma_k|>10^{20}$ or $|\alpha_{k+1}-\alpha_k|>10^{20}$,

   replace $x_0=(\mu_0,\sigma_0,\alpha_0)'$ with $(\mu_0,\sigma_0,\alpha_0+u)'$, where

   $u\sim\text{unif}(-0.1,0.1)$, and repeat step 2.

Simulation Method

- Clayton copula:

  $$C(u_1, u_2; \alpha) = \max\{[u_1^{-\alpha} + u_2^{-\alpha} - 1]^{-\frac{1}{\alpha}}, 0\}, \ \alpha \in (-1, \infty) \setminus \{0\}.$$

- $\alpha \in (-1, 0)$: Y_1 and Y_2 have negative correlation.
- $\alpha \in (0, \infty)$: Y_1 and Y_2 have positive correlation.

- Clayton copula for positive correlation:

  $$C(u_1, u_2; \alpha) = [u_1^{-\alpha} + u_2^{-\alpha} - 1]^{-\frac{1}{\alpha}}, \alpha \in (0, \infty).$$

Simulation Method

- $P(Y_2 \le y_2 \mid Y_1 = y_1) = U_2$, $U_2 \sim \text{unif}(0,1)$.

  Hence, $P(Y_2 \le y_2, Y_1 = y_1) = U_2 P(Y_1 = y_1)$.

  $$U_2 \exp\{-(y_1-\mu)^2/2\sigma\}/\sqrt{2\pi}\sigma = P(Y_2 \le y_2, Y_1 = y_1)$$

  $$= \frac{\partial}{\partial y_1} P(Y_2 \le y_2, Y_1 \le y_1)$$

  $$= \frac{\partial}{\partial y_1}[(\Psi(y_1))^{-\alpha} + (\Psi(y_2))^{-\alpha} - 1]^{-\frac{1}{\alpha}},$$

  Where $\Psi(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$.

Simulation Method

- $y_{t+1} = \Psi^{-1}\{[1+(U_t^{-\frac{\alpha}{\alpha+1}}-1)\Psi(y_t)^{-\alpha}]^{-\frac{1}{\alpha}}\}, t=1,2,\ldots,n$

  $y_1 = \Psi^{-1}(U_1)$, $U_t \sim \text{unif}(0,1)$.

- **Algorithm 1**
  1. Generate $U_1, U_2 \sim \text{unif}(0,1)$, $Y_1 = \Psi^{-1}(U_1)$.
  2. $Y_{t+1} = \Psi^{-1}\{[1+(U_{t+1}^{-\frac{\alpha}{\alpha+1}}-1)\Psi(Y_t)^{-\alpha}]^{-\frac{1}{\alpha}}\}$,

     where $U_{t+1} \sim \text{unif}(0,1), t=1,2,\ldots,n$.

3. Standard method

1. $\hat{\mu} = \overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t.$

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

3. $\hat{\alpha} = -2\tau/(\tau - 1)$, where

$$\tau = \binom{n}{2}^{-1} \sum_{i<j} \left[\operatorname{sgn}(Y_j - Y_i)\operatorname{sgn}(Y_{j+1} - Y_{i+1}) \right]$$

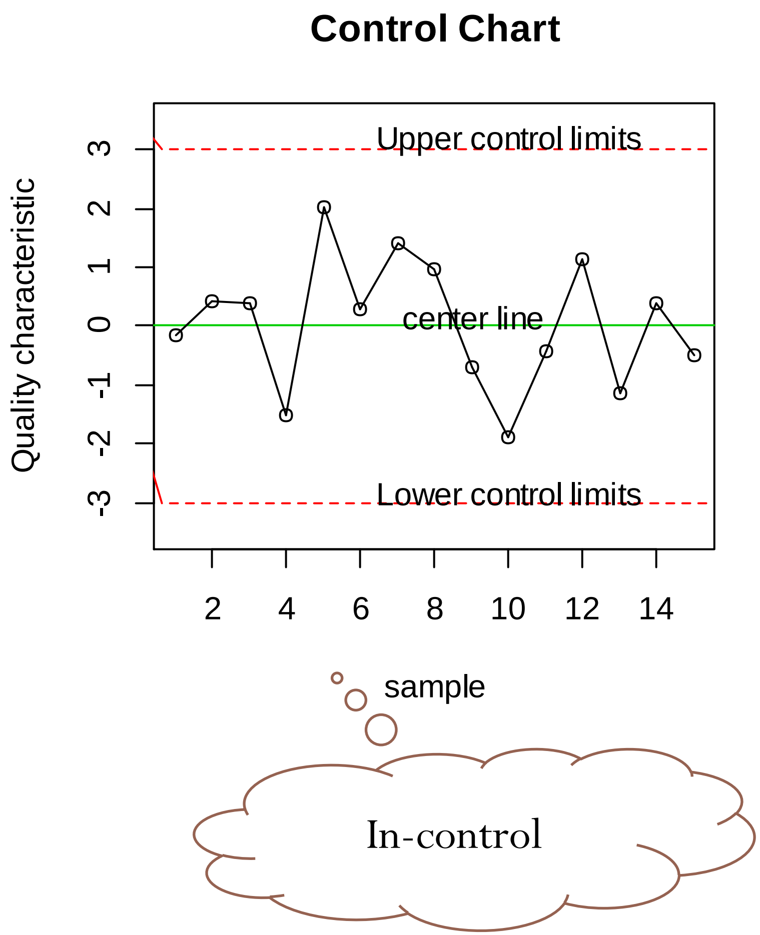
MSE for UCL
($\alpha = 2$, n=1000, 100 repetitions.)

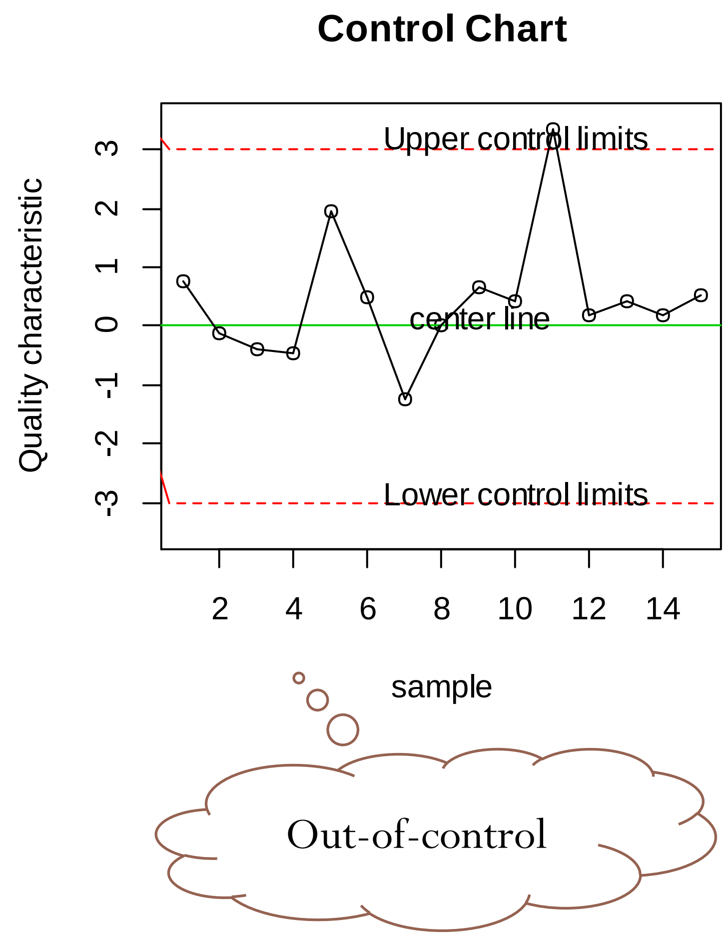
		(μ, σ)					
		UCL=3	UCL=9	UCL=4	UCL=10	UCL=5	UCL=11
Method1 (proposed)	E($\hat{\mu}+3\hat{\sigma}$)	3.0003	9.0010	4.0003	10.0044	5.0003	11.0010
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0083	0.0745	0.0083	0.0907	0.0083	0.0745
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0003	0.0010	0.0003	0.0044	0.0003	0.0010
Method2 (Chen & Fan)	E($\hat{\mu}+3\hat{\sigma}$)	2.9954	8.9863	3.9959	9.9980	4.9994	10.9863
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0190	0.1709	0.0189	0.1673	0.0188	0.1707
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0046	-0.0137	-0.0041	-0.0020	-0.0006	-0.0137
Method3 (standard)	E($\hat{\mu}+3\hat{\sigma}$)	2.9984	8.9952	3.9984	10.0075	4.9984	10.9952
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0190	0.1711	0.0190	0.1678	0.0190	0.1711
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.0016	-0.0046	-0.0016	-0.0075	-0.0016	-0.0048

MSE for UCL
($\alpha = 8$, n=1000, 100 repetitions.)

		(μ, σ)					
		UCL=3	UCL=9	UCL=4	UCL=10	UCL=5	UCL=11
Method1 (proposed)	E($\hat{\mu}+3\hat{\sigma}$)	3.0175	9.0096	3.9972	9.9935	4.9871	10.9783
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0150	0.1202	0.0142	0.0996	0.0120	0.1599
	Bias($\hat{\mu}+3\hat{\sigma}$)	0.0175	0.0096	-0.0028	-0.0065	-0.0129	-0.0217
Method2 (Chen & Fan)	E($\hat{\mu}+3\hat{\sigma}$)	2.8502	8.3811	3.7998	9.5026	4.7973	10.5276
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0800	1.0722	0.0990	0.7594	0.0976	0.8773
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.1498	-0.6189	-0.2002	-0.4974	-0.2027	-0.4724
Method3 (standard)	E($\hat{\mu}+3\hat{\sigma}$)	2.8530	8.3891	3.8017	9.5111	4.7949	10.5354
	MSE($\hat{\mu}+3\hat{\sigma}$)	0.0793	1.0650	0.0983	0.7520	0.0986	0.8711
	Bias($\hat{\mu}+3\hat{\sigma}$)	-0.1470	-0.6109	-0.1983	-0.4889	-0.2051	-0.4646

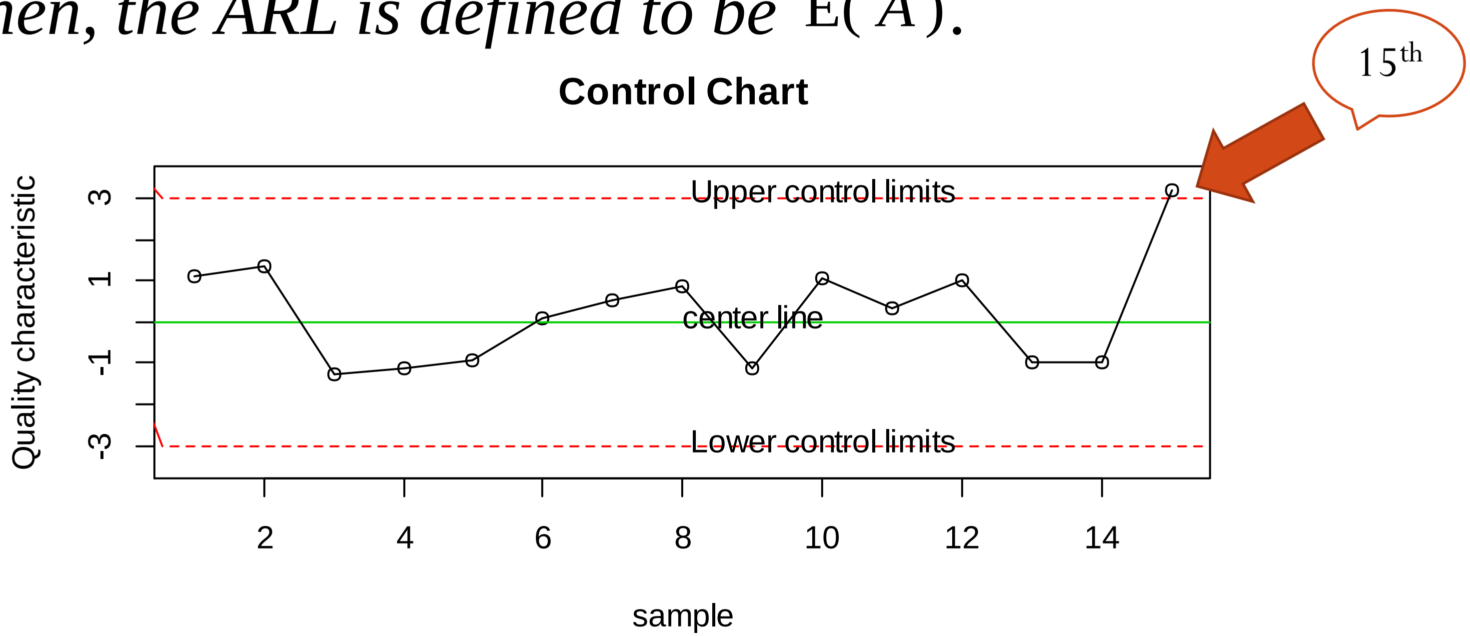
3-Sigma Control Chart





Average Run Length (ARL)

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



Average Run Length (ARL)

- i.i.d case:

  $$\text{ARL} = \frac{1}{\alpha}, \text{ where } \alpha \text{ is type I error.}$$

  - Example: $\alpha = 0.0027$

    $z_{0.0027} = 3$

  $$\text{ARL} = \frac{1}{0.0027} = 370.3704$$

- correlated case:

  It is hard to solve!!! (See Schmid, 1995)

Algorithm 2 (Average run length)

1. Set $t = 1$ and $Y_1 \sim N(\mu, \sigma)$.

2. $U_t \sim \text{unif}(0,1)$, and then

   $$Y_{t+1} = \Psi^{-1}\{[1+(U_t^{-\frac{\alpha}{\alpha+1}} - 1)\Psi(Y_t)^{-\alpha}]^{-\frac{1}{\alpha}}\}, \quad t = 1, 2, \ldots\ldots,$$

   where $\Psi(y) = \Phi\left(\dfrac{y-\mu}{\sigma}\right)$.

3. Set $A = \min\{t : Y_t < \mu - 3\sigma \text{ or } Y_t > \mu + 3\sigma\}$.

4. Repeat Step 1 ~ Step 3 m times. The average of the m run length is the ARL.

Known vs. Unknown

Table 1 ARLs-Shewhart control chart for Known and estimated parameters.

Kendall's tau	**known**	**Estimated by standard method**	**Estimated by our proposed method**	**Estimated by Chan & Fan's method**
0.8	753.821	485.236	747.418	479.568
0.5	623.152	605.173	630.533	606.277
0.3	495.731	516.062	510.375	511.695
0.1	382.190	385.635	382.594	383.363
0	369.989	368.442	369.214	367.545

Variance Reduction

- Why do we want to reduce variance?

  ➔ Get more accurate results.

- Method:

  Antithetic variables. (Ross, 2006)

Variance Reduction

- $A = h\{U_t; t = 1, 2, \ldots\}$, $B = h\{1 - U_t; t = 1, 2, \ldots\}$.

➔ $\dfrac{(A+B)}{2}$ is an unbiased estimator of E(A).

$$\operatorname{var}\left(\frac{A+B}{2}\right) = \frac{\operatorname{var}(A) + \operatorname{var}(B) + 2\operatorname{cov}(A, B)}{4} < \frac{\operatorname{var}(A)}{2}.$$

Algorithm 3

1. Set $t = 1$ and $Y_1 \sim N(\mu, \sigma)$.
2. $U_1 \sim \text{unif}(0, 1)$, $Y_{1,1} = \Psi^{-1}(U_1)$ and $Y_{2,1} = \Psi^{-1}(1 - U_1)$,

   where $\Psi(y) = \Phi\left(\dfrac{y-\mu}{\sigma}\right)$.
3. $U_t \sim \text{unif}(0, 1)$, and then

   $$Y_{1,t+1} = \Psi^{-1}\{[1 + (U_{t+1}^{-\frac{\alpha}{\alpha+1}} - 1)\Psi(Y_{1,t})^{-\alpha}]^{-\frac{1}{\alpha}}\}, t = 1, 2, \ldots\ldots$$
4. Set $A = \min\{t : Y_{1,t} < \mu - 3\sigma \text{ or } Y_{1,t} > \mu + 3\sigma\}$.

Algorithm 3

5. $Y_{2,r+1} = \Psi^{-1}\{[1+((1-U_{r+1})^{-\frac{\alpha}{\alpha+1}}-1)\Psi(Y_{2,r})^{-\alpha}]^{-\frac{1}{\alpha}}\}, r=1, 2, ...$
6. Set $B = \min\{t : Y_{2,t} < \mu - 3\sigma \text{ or } Y_{2,t} > \mu + 3\sigma\}$.
7. Repeat step 1~step 6 m times. We get two sequences $(A_1, ..., A_m)$ and $(B_1, ..., B_m)$. The ARL is $\sum_{i=1}^{m} \frac{(A_i + B_i)}{2m}$.

ARL & S.D

- ARL:
  - Monte Carlo: $\dfrac{\sum_{i=1}^{2m} A_i}{2m}$
  - Antithetic variables: $\dfrac{\sum_{i=1}^{m} (A_i + B_i)}{2m}$
- S.D:
  - Monte Carlo: $\sqrt{\sum_{i=1}^{2m} A_i^{\,2} / 2m - \left(\sum_{i=1}^{2m} A_i / 2m\right)^2}$
  - Antithetic variables: $\sqrt{\sum_{i=1}^{m} (A_i^{\,2} + B_i^{\,2}) / 2m - \left(\sum_{i=1}^{m} (A_i + B_i) / 2m\right)^2}$

Simulation Result

Table 6 Standard deviation (S.D) comparison of ARL (simulate 10000 times); $\alpha = 2$

(μ, σ)	Monte Carlo Method		Antithetic variate			
	ARL	S.D	ARL	S.D	cor(A, B)	Ratio
(0, 1)	620.9303	632.5059	616.3838	627.1665	0.0714	1.0085
(0, 3)	621.3078	616.1965	615.4543	613.8501	0.0508	1.0038
(1, 1)	616.6959	615.5491	605.9013	606.0645	0.0532	1.0156
(1, 3)	613.5204	622.5679	615.3911	615.2539	0.0445	1.0118
(2, 1)	623.7823	638.1642	629.4922	633.7178	0.0577	1.0070
(2, 3)	616.6412	613.5663	620.6297	616.9739	0.0394	0.9944

NOTE: Ratio > 1 corresponds to better performance of the Antithetic variables.

Simulation Result

Table 6 Standard deviation (S.D) comparison of ARL (simulate 10000 times); $\alpha = 2$

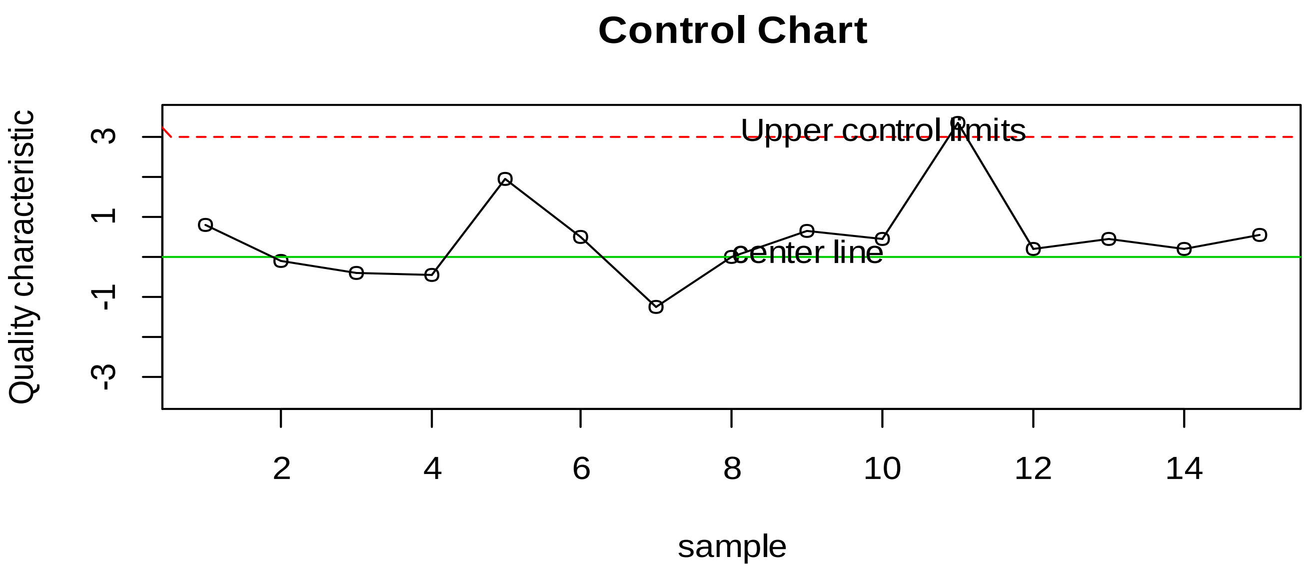
(μ, σ)	Monte Carlo Method		Antithetic variate			
	ARL	S.D	ARL	S.D	cor(A, B)	Ratio
(0, 1)	763.1522	772.7252	753.8036	755.4667	-0.0515	1.0228
(0, 3)	764.4263	771.4628	754.8563	753.3652	-0.0536	1.0240
(1, 1)	759.8151	748.3154	760.9389	757.5643	-0.0515	0.9877
(1, 3)	762.0303	761.0947	753.4320	751.9185	-0.0786	1.0122
(2, 1)	762.6541	767.7091	764.2615	770.0238	-0.0423	0.9970
(2, 3)	766.4103	758.8996	781.1320	772.5143	-0.0430	0.9824

NOTE: Ratio > 1 corresponds to better performance of the Antithetic variables.

We also show one-sided case!!!

- We only set upper control limit, the definition of ARL is as follow:

$$A = \min\{\, t : Y_t > \mu + 3\sigma \,\}$$



Simulation Result

Table 7 Standard deviation (S.D) comparison of ARL for one-sided case. (simulate 10000 times); $\alpha = 2$

(μ, σ)	Monte Carlo Method		Antithetic variate			
	ARL	S.D	ARL	S.D	$cor(A, B)$	Ratio
(0, 1)	748.4771	750.8413	744.3434	742.4753	-0.0219	1.0112
(0, 3)	743.8884	732.3003	738.7775	736.0809	-0.0201	0.9948
(1, 1)	746.2283	743.1858	745.7316	746.3437	-0.0020	0.9957
(1, 3)	752.2325	759.6817	751.4091	770.7314	-0.0229	0.9856
(2, 1)	742.6974	735.7492	749.0812	744.1737	-0.0129	0.9886
(2, 3)	744.6714	743.7506	740.9985	736.7277	-0.0477	1.0095

Simulation Result

Table 7 Standard deviation (S.D) comparison of ARL for one-sided case. (simulate 10000 times); $\alpha = 8$

(μ, σ)	Monte Carlo Method		Antithetic variate			
	ARL	S.D	ARL	S.D	cor(A, B)	Ratio
$(0, 1)$	786.5697	793.4528	796.9079	815.5159	-0.0867	0.9729
$(0, 3)$	783.6691	773.3309	776.2317	777.2251	-0.0995	0.9949
$(1, 1)$	786.7654	783.5356	787.9072	784.7020	-0.1082	0.9985
$(1, 3)$	784.1199	787.1412	791.7128	786.2797	-0.0731	1.0010
$(2, 1)$	784.0967	790.8294	785.9121	779.8624	-0.0856	1.0140
$(2, 3)$	777.7631	793.051	777.5777	793.7792	-0.0771	0.9990

Outline

Data 1

- "Statistical Quality Control: A Modern Introduction, 6th edition" written by D. C. Montgomery (2009). The processing costs for the most recent 40 weeks.

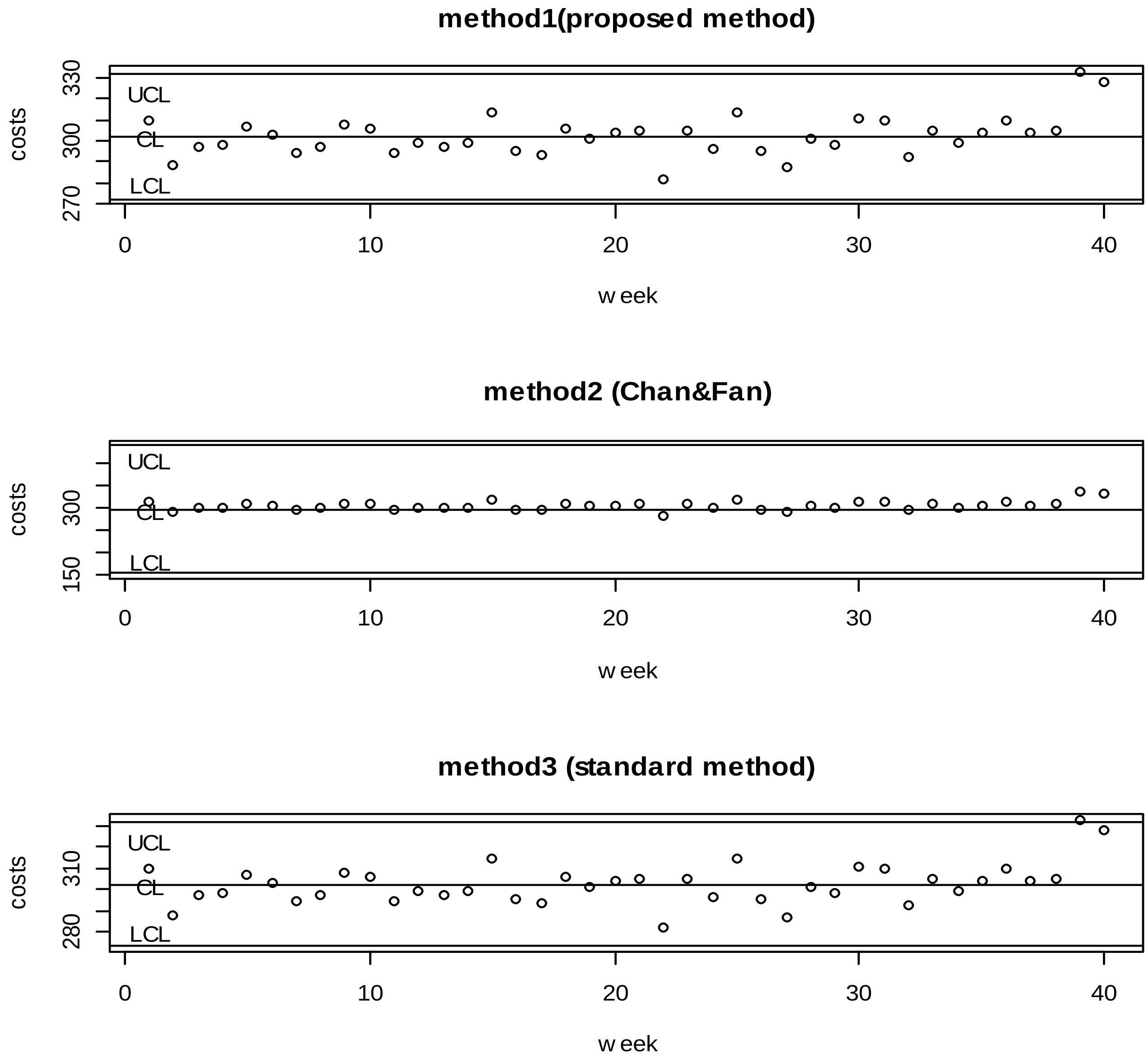
- The mortgage loan processing unit of a bank monitors the costs of processing loan applications. The quantity tracked is the average weekly processing costs, obtained by dividing total weekly costs by the number of loans processed during the week.

Numerical result

Table 9 Estimates of process parameters based on processing costs data.

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	UCL
Method1 (proposed)	302.1790	10.0173	-0.2757	332.2308
Method2 (Chen & Fan)	294.9756	47.6064	-0.2113	437.7947
Method3 (standard)	302.3500	9.7863	-0.0879	331.7089

A little strange


method1(proposed method)
costs
270
300
330
UCL
CL
LCL
0
10
20
30
40
week
method2 (Chan&Fan)
costs
150
300
UCL
CL
LCL
0
10
20
30
40
week
method3 (standard method)
costs
280
310
UCL
CL
LCL
0
10
20
30
40
week

What's happened??

- The result in method 2 is very strange!

- Reason: Sample size is too small.

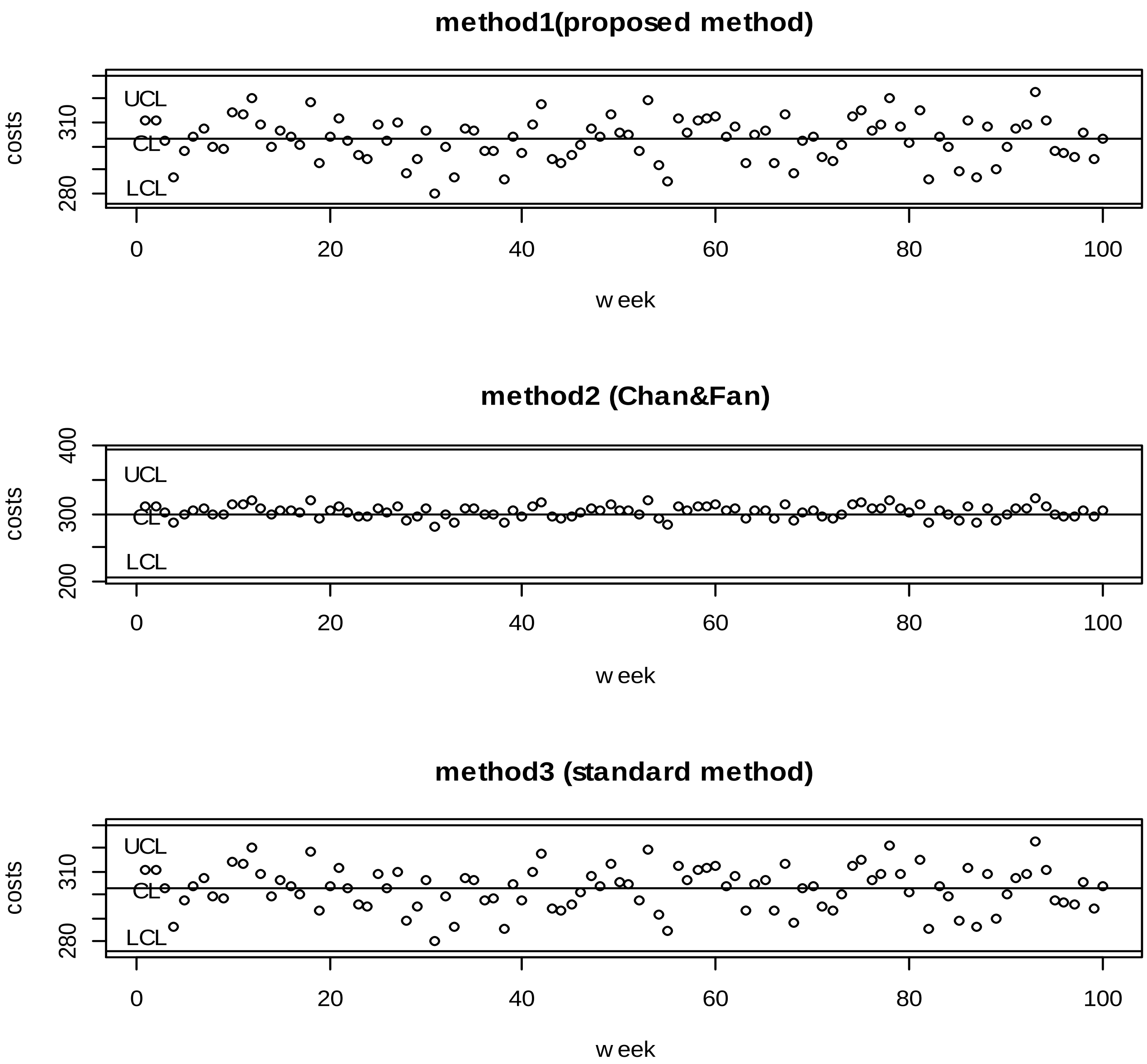
$$\hat{\mu} = \frac{n}{n+1}\bar{Y}$$

$$\hat{\sigma}^2 = \frac{n}{n+1}\bar{Y}^2 - \left(\frac{n}{n+1}\bar{Y}\right)^2$$

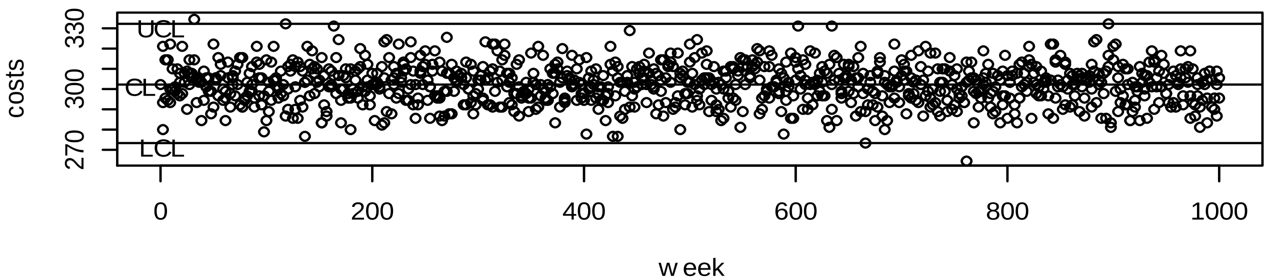
Numerical result

Table10 Simulation results with $\hat{\mu}$=302.35 and $\hat{\sigma}$ =9.79 for three method.

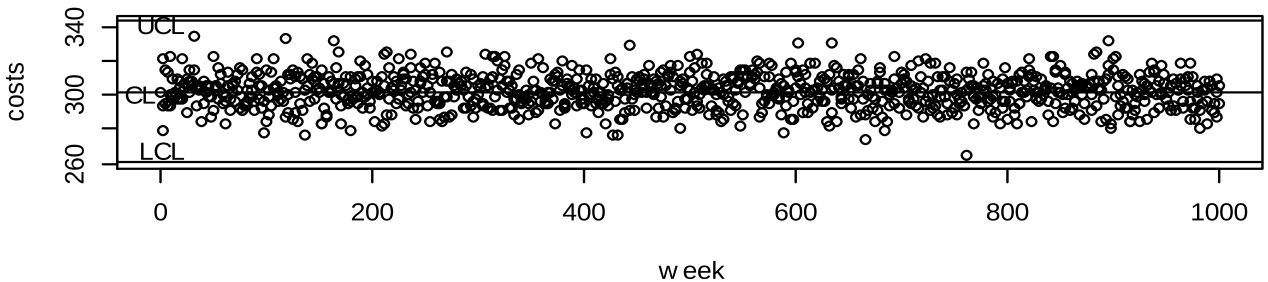
		$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	UCL
Method1 (proposed)	n = 100	301.8767	8.8276	-0.0560	328.3595
	n = 1000	302.2952	9.3321	-0.0747	330.2915
Method2 (Chen & Fan)	n = 100	298.8750	31.1527	-0.0441	392.3331
	n = 1000	301.9915	13.3481	-0.0699	342.0358
Method3 (standard)	n = 100	301.8638	8.8763	-0.0272	328.4927
	n = 1000	302.2935	9.3352	-0.0937	330.2991


method1(proposed method)
UCL
CL
LCL
costs
310
280
0
20
40
60
80
100
week
method2 (Chan&Fan)
UCL
CL
LCL
costs
400
300
200
0
20
40
60
80
100
week
method3 (standard method)
UCL
CL
LCL
costs
310
280
0
20
40
60
80
100
week

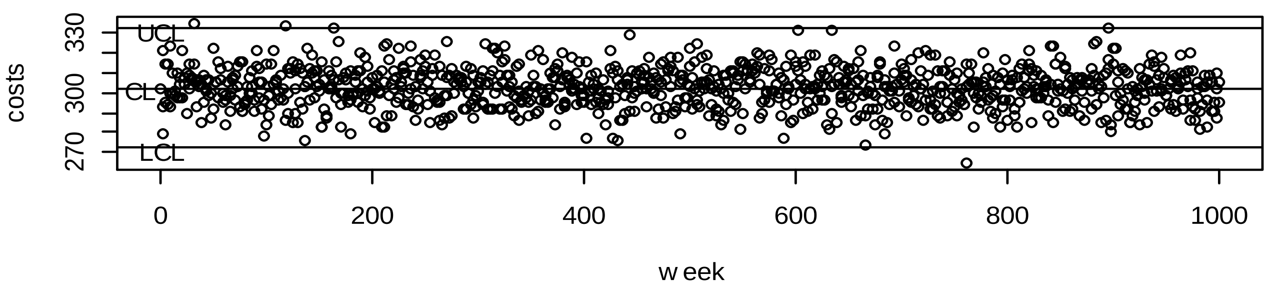
method1(proposed method)



method2 (Chan&Fan)



method3 (standard method)



Normality Test

- Why do we need to check normality?
  - our proposed method sets the unknown distribution G^* as $\Phi\left(\frac{y-\mu}{\sigma}\right)$.
- Chernoff and Lehmann (1954) proposed chi-square goodness-of-fit test.
- Kendall's tau is -0.0460 ➔ close to i.i.d dataset case.

Chi-Square Goodness-of-Fit Test

- $H_0 : Y \sim N\left(\mu, \sigma^2\right)$ v.s $H_1 : not\ H_0$

  test statistic: $Q = \sum_{i=1}^{n} \frac{\left(e_i - o_i\right)^2}{e_i}$

  where o_i is the observed frequency for bin *i* and e_i is the expected frequency for bin *i*.

  Reject H_0 if $Q > \chi^2_{n-1-2,1-\alpha}$

Chi Square Goodness-of-Fit Test

Table 11 The results for chi square goodness of fit test.
p.s. p_i is the probability for bin *i*.

	p_i	o_i	$e_i = np_i$	Result for hypothesis
$(-\infty, \mu-2\sigma)$	0.025	1	1.0	$\sum_{i=1}^{8} \frac{(e_i - o_i)^2}{e_i} = 5.054 < \chi^2_{8-1-2,\,1-0.05} = 11.07$
$(\mu-2\sigma, \mu-\sigma)$	0.135	3	5.4	
$(\mu-\sigma, 0)$	0.340	16	13.6	
$(0, \mu+\sigma)$	0.340	16	13.6	
$(\mu+\sigma, \mu+2\sigma)$	0.135	2	5.4	
$(\mu+2\sigma, \infty)$	0.025	2	1.0	

Data 2

- "Time Series Analysis and Its Applications With R Example" written by Shumway & Stoffer. (2010)

- Its is shown that global mean land-ocean temperature deviations, measured in degrees centigrade, for the years 1880-2009.

We encounter a problem!!

- Kendall's tau for data 2 is 0.6451
  - Strong correlation!
  - Can not use chi-square goodness-of-fit test.
- we use the Cramér-von Mises and Kolmogorov-Smirnov test statistic with the null distribution approximated by the parametric bootstrap (see Genest and Remillard, 2008) to check normality.

Cramér-von Mises and Kolmogorov-Smirnov test

- $H_0 : Y \sim N\left(\mu, \sigma^2\right)$ v.s $\quad H_1 : not\ H_0$

We use bootstrap method to check normality:

$$\hat{\alpha}_1^{(b)} - \hat{\alpha}_2^{(b)}, b = 1, \ldots, 1000.$$

where $\hat{\alpha}_i$ is the estimator of α in method i, i=1, 2.

if $\dfrac{\sum_{b=1}^{1000} \left(\mathrm{I}\left\{ \hat{\alpha}_1^{(b)} - \hat{\alpha}_2^{(b)} > 0 \right\} \right)}{1000} < 0.05$, reject H_0.

Result

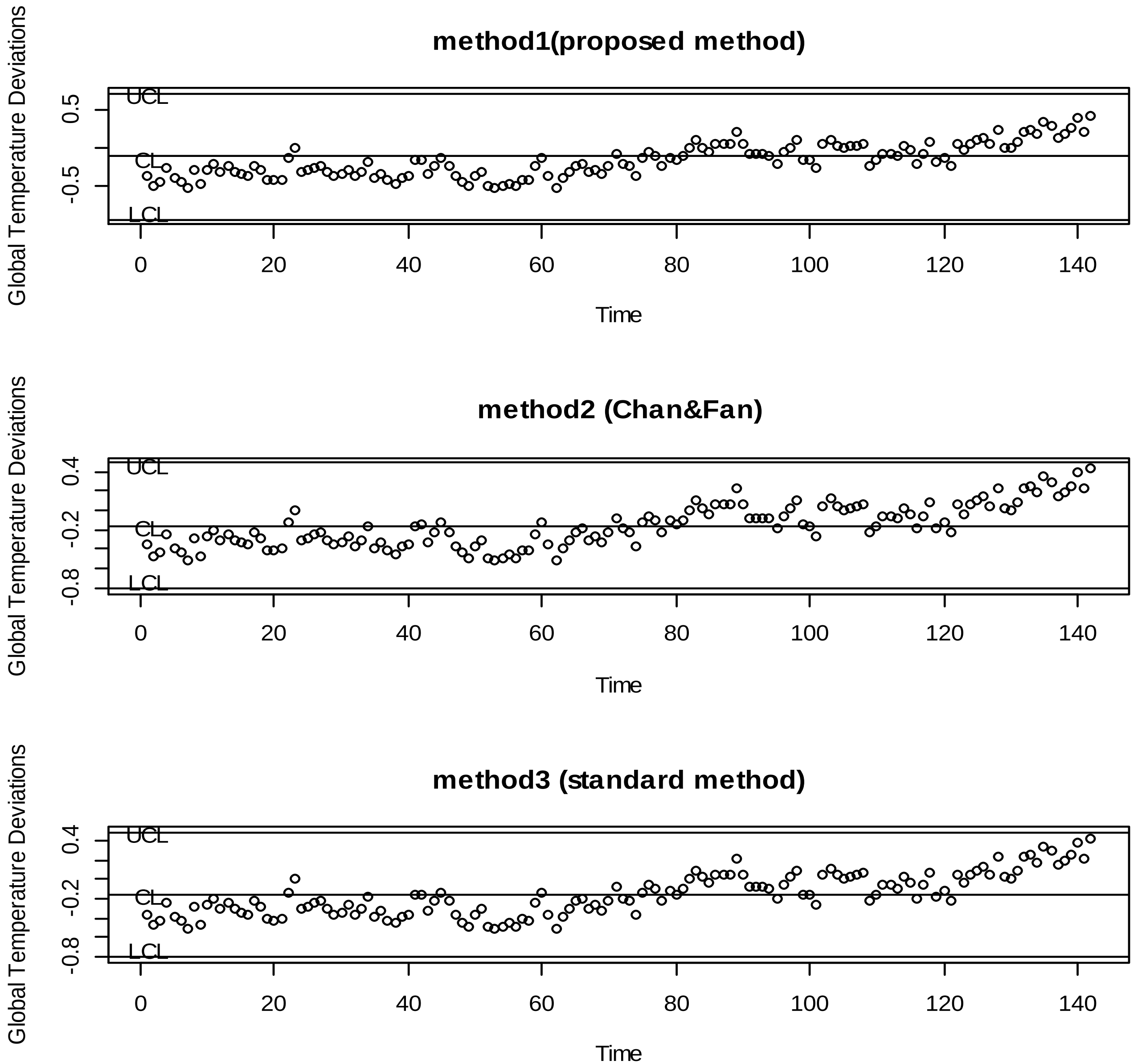
- $$\frac{\sum_{b=1}^{1000}\left(\mathrm{I}\left\{ \hat{\alpha}_1^{(b)} - \hat{\alpha}_2^{(b)} > 0 \right\} \right)}{1000} = 0.4589 > 0.05$$

There is no enough evidence to reject the null hypothesis. Hence, the data are normal distribution with $\mu = -0.1581$ and $\sigma^2 = 0.0475$.

Numerical result

Table 12 Estimates of process parameters based on data 2.

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	UCL
Method1 (proposed)	-0.1051	0.2734	3.0569	0.7151
Method2 (Chen & Fan)	-0.1569	0.2167	1.7228	0.4933
Method3 (standard)	-0.1581	0.2179	3.6360	0.4955


method1(proposed method)
Global Temperature Deviations
0.5
-0.5
UCL
CL
LCL
0
20
40
60
80
100
120
140
Time
method2 (Chan&Fan)
Global Temperature Deviations
0.4
-0.2
-0.8
UCL
CL
LCL
0
20
40
60
80
100
120
140
Time
method3 (standard method)
Global Temperature Deviations
0.4
-0.2
-0.8
UCL
CL
LCL
0
20
40
60
80
100
120
140
Time

Outline

1. Introduction
2. Methodology
3. Results and Analysis
4. Conclusion and Discussion

Conclusion

1. Setting appropriate estimators of UCL and LCL is very important for engineers.
2. Our proposed method give most accurate estimates of UCL and LCL.
3. It is appropriate only when sample size is very large in Chen & Fan's method.
4. In our proposed method, it is important to check whether the real data are normality or not.

Future Work

1. In variance reduction, we have no useful method to reduce them. We leave this part to the future work. (Fuh et. al., 2013)
2. We can try to do our research in two groups of multivariate time series data case. (Hung and Tseng, 2012)

Thank you for your attention!