EUROPA-UNIVERSITÄT VIADRINA FRANKFURT (ODER)

Simultaneous EWMA charts for controlling mean and variance in the presence of autocorrelation

Sven Knoth

Department of Statistics,
Europe University Viadrina,
Postfach 1786, 15 207 Frankfurt(Oder), Germany

http://w3stat.euv-frankfurt-o.de

1. Simultaneous charts and autocorrelation

2. Modified EWMA charts

3. Residual EWMA charts

4. Comparison and Conclusions

simultaneous charts for mean and variance

Montgomery (1991),

Gan (1995),

Kanagawa and Arizono (1997),

Mittag and Stemann (1997)

control charts and autocorrelated data

Goldsmith and Woodward (1961),

Bagshaw and Johnson (1975),

Nikiforov (1975/79),

Vasilopoulos and Stamboulis (1978),

Alwan (1989),

Amin, Schmid, and Frank (1997),

198x, 199x ...

simultaneous & correlation

Lu and Reynolds (1999),

Knoth, Schmid and Schöne (1998):

$\bar{X}$–S^2 and $\bar{X}$–R Shewhart chart

$\Downarrow$

EWMA ?

Change point model

target process $\{Y_{i,j}\}$, observed $\{X_{i,j}\}$

i – sample number, j – number within the sample (batch size n)

$$X_{i,j} = \begin{cases} Y_{i,j} & \text{for } i < q \\ \mu_0 + \Delta\left(Y_{i,j} - \mu_0\right) + a\sqrt{\gamma_0} & \text{for } i \geq q \end{cases}$$

with change point q and
$\mu_0 = E(Y_{i,j})$, $\gamma_0 = Var(Y_{i,j})$.

$\rightsquigarrow$

$$E(X_{i,j}) = \begin{cases} \mu_0 & \text{for } i < q \\ \mu_0 + a\sqrt{\gamma_0} & \text{for } i \geq q \end{cases} ,$$

$$Var(X_{i,j}) = \begin{cases} \gamma_0 & \text{for } i < q \\ \Delta^2 \gamma_0 & \text{for } i \geq q \end{cases} .$$

EWMA chart (iid)

$$
\begin{aligned}
Z_{\bar{X},i} &\underset{i\geq 1}{=} \left(1-\lambda_1\right) Z_{\bar{X},i-1} + \lambda_1 \bar{X}_i \,, \\
Z_{\bar{X},0} &= E_0(\bar{X}) \;=\; \mu_0 \,, \\
Z_{S^2,i} &\underset{i\geq 1}{=} \left(1-\lambda_2\right) Z_{S^2,i-1} + \lambda_2 S_i^2 \,, \\
Z_{S^2,0} &= E_0(S^2) \;=\; \gamma_0 \,,
\end{aligned}
$$

$$
\begin{aligned}
\tau \;=\; \inf\Bigg\{ & i \in I\!N : \left|Z_{\bar{X},i} - \mu_0\right| > x_{iid} \sqrt{\frac{\lambda_1}{2-\lambda_1}} \sqrt{Var_0(\bar{X})} \\
& \text{or} \quad Z_{S^2,i} - \gamma_0 > s_{iid} \sqrt{\frac{\lambda_2}{2-\lambda_2}} \sqrt{Var_0(S^2)} \Bigg\} .
\end{aligned}
$$

(one–sided for the variance !!)

Average Run Length – ARL

- most popular performance measure of control charts

"expectation of the stopping time τ"

1. in–control
   $\leftrightarrow$ $q = \infty$ or $a = 0$, $\Delta = 1$:

   $E_0(\tau)$

2. out–of–control
   $\leftrightarrow$ $q = 1$ and $a \neq 0$ or $\Delta > 1$

   $E_1(\tau)$

further:

conditional, steady–state delays, quantiles ...

Autocorrelation model

independence between batches

$$\mathbf{Y}_i = (Y_{i,1}, Y_{i,2}, ..., Y_{i,n})' \qquad , \ i = 1, 2, ...$$

dependence within the batch: $AR(1)$

$$Y_{i,j} \;=\; \mu_0 + \alpha\big(Y_{i,j-1} - \mu_0\big) + \varepsilon_{i,j}\,,$$

$$j = 1, 2, ..., n\,,$$

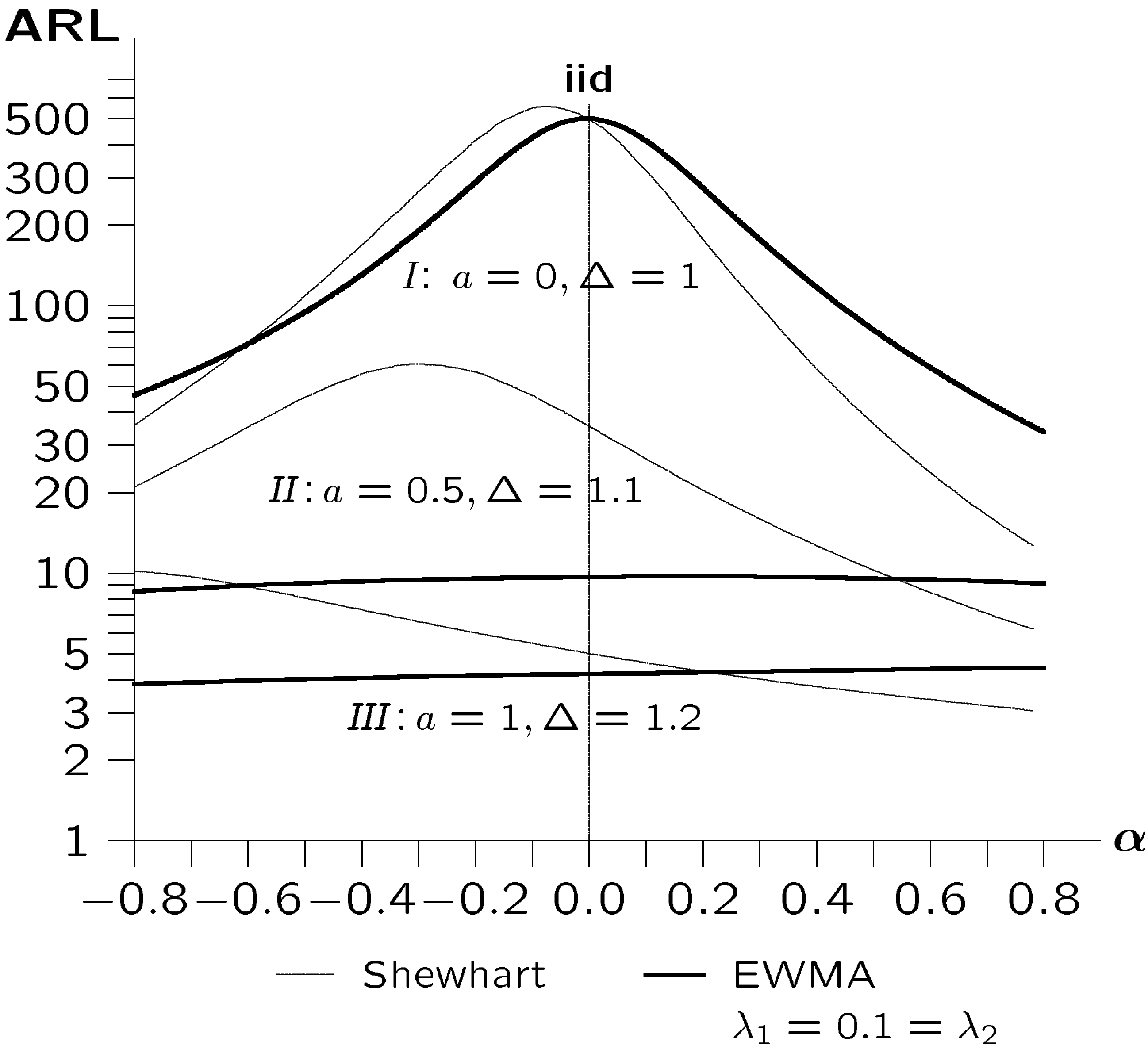
$$|\alpha| < 1\,,$$

$$\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_0^2) \text{ (iid)}\,.$$

$\leadsto$

$$\begin{aligned} E(Y_{i,j}) &= \mu_0\,, \\ Var(Y_{i,j}) &= \frac{\sigma_0^2}{1-\alpha^2} = \gamma_0\,. \end{aligned}$$

Influence of the autocorrelation coefficient α on the ARL



Main concepts of control charts for correlated data

start with fit of appropriate time series model – here AR(1)

modified charts	**residual charts**
adapt variance and critical value	model residuals are iid and
$\rightarrow$ correct in–control ARL	the empirical residuals approximately as well (starting problems)
$\Downarrow$	$\Downarrow$
transformed threshold for original data	classical charts on transformed data

special case:

cusum charts of Nikiforov (1975/79), Schmid (1997)

Modified $\bar{X}$–S^2 EWMA chart I

adapted moments

with autocorrelation function
$\gamma_h = Cov_0(X_{i,j}, X_{i,j+h})$

$$Var_0(\bar{X}) \;=\; \frac{1}{n}\sum_{|j|<n}\left(1 - \frac{|j|}{n}\right)\gamma_j\,,$$

$$E_0(S^2) \;=\; \gamma_0 - \frac{2}{n-1}\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\gamma_j\,,$$

$$Var_0(S^2) \;=$$

$$\frac{2}{(n-1)^2}\left[\sum_{v,j=1}^{n}\gamma_{v-j}^2 - \frac{2}{n}\sum_{v=1}^{n}\left(\sum_{j=1}^{n}\gamma_{v-j}\right)^2 + \frac{1}{n^2}\left(\sum_{v,j=1}^{n}\gamma_{v-j}\right)^2\right].$$

AR(1): $\gamma_h = \dfrac{\alpha^{|h|}\,\sigma_0^2}{1-\alpha^2}$

Modified $\bar{X}$–S^2 EWMA chart *II*

adapted critical values

prespecify in–control ARL $\rightarrow$ determine related critical values (x_m, s_m)

1. approximate ARL by Markov chain approach (Brook/Evans 1972):
   - 2dimensional chain,
   - transition probabilities:
     joint distribution of $(\bar{X}, S^2)$
     $\rightsquigarrow$ Schöne/Schmid (1997/99)

$$P_{\bar{X},S^2}(|\bar{X}| \leq x, S^2 \leq s) = \sum_{i,j=0}^{\infty} \frac{f_{i+1,j}}{\sqrt{2\pi}} \frac{4\beta}{2j+1} g_i^{(n+1)}(s)\, x^{2j+1} + \\ + \chi^2_{n-1}\left(\frac{s}{\beta}\right) \left[\Phi\left(\frac{x}{\sqrt{b_0}}\right) - \Phi\left(-\frac{x}{\sqrt{b_0}}\right)\right] .$$

2. additional condition:
   univariate charts behave symmetrically

3. two nonlinear equations,
   two unknown parameters
   $\rightarrow$ solution by 2dimensional secant rule

4. in study:
   maximal dimension $\approx 1600 = 40 \cdot 40$.

Residual $\bar{X}$–S^2 EWMA chart I

standardized residuals – AR(1)

(X – observed, Y – target)

$$\widehat{\varepsilon}_{i,j} = \frac{X_{i,j} - \widehat{X}_{i,j}}{\sqrt{Var_0(X_{i,j} - \widehat{X}_{i,j})}},$$

$$\widehat{X}_{i,j} = \begin{cases} \mu_0 & , j = 1 \\ \mu_0 + \alpha \left(X_{i,j-1} - \mu_0 \right) & , j > 1 \end{cases},$$

$$Var_0(X_{i,j} - \widehat{X}_{i,j}) = \begin{cases} \gamma_0 & , j = 1 \\ \sigma_0^2 & , j > 1 \end{cases},$$

$$X_{i,j} - \widehat{X}_{i,j} \underset{i \geq q}{=} \begin{cases} \Delta \left(Y_{i,1} - \mu_0 \right) + a \sqrt{\gamma_0} & , j = 1 \\ \Delta \, \varepsilon_{i,j} + a \sqrt{\gamma_0} \, (1 - \alpha) & , j > 1 \end{cases}.$$

$$\bar{\widehat{\varepsilon}}_i = \frac{1}{n} \sum_{j=1}^{n} \widehat{\varepsilon}_{i,j} \quad , \; \widehat{S}_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(\widehat{\varepsilon}_{i,j} - \bar{\widehat{\varepsilon}}_i \right)^2,$$

$$E_0(\bar{\widehat{\varepsilon}}_i) = 0 \quad , \; Var_0(\bar{\widehat{\varepsilon}}_i) = 1/n \, ,$$

$$E_0(\widehat{S}_i^2) = 1 \quad , \; Var_0(\widehat{S}_i^2) = 2/(n-1) \, .$$

Residual $\bar{X}$–S^2 EWMA chart *II*

computation of the ARL

1. Gan (1995): $\sim$ Waldmann (1986)

2. here:

$$
\begin{aligned}
E_0(\tau_{biv}) &= \sum_{i=0}^{\infty} P_0(\tau_{biv} > i) \\
&= \sum_{i=0}^{\infty} P_0(\tau_1 > i) \cdot P_0(\tau_2 > i) \\
&\approx \sum_{i=0}^{\infty} \mathbf{p}_1' \, \mathbf{P}_1^i \, \mathbf{1} \times \mathbf{p}_2' \, \mathbf{P}_2^i \, \mathbf{1} \\
&= \mathbf{p}_1' \, \mathbf{Z} \, \mathbf{p}_2 \quad , \; \mathbf{Z} \;=\; \sum_{i=0}^{\infty} \mathbf{P}_1^i \, \mathbf{1} \, \mathbf{1}' \, (\mathbf{P}_2')^i
\end{aligned}
$$

$\mathbf{Z}$ solves matrix equation

$$\mathbf{Z} = \mathbf{1}\,\mathbf{1}' + \mathbf{P}_1 \, \mathbf{Z} \, \mathbf{P}_2'$$

or equivalent Sylvester matrix equation

$$
\begin{aligned}
(\mathbf{I} - \mathbf{P}_1)^{-1} \, \mathbf{Z} + \mathbf{Z} \, \mathbf{P}_2' \, (\mathbf{I} - \mathbf{P}_2')^{-1} \quad &= \\
&(\mathbf{I} - \mathbf{P}_1)^{-1} \, \mathbf{1} \, \mathbf{1}' \, (\mathbf{I} - \mathbf{P}_2')^{-1} \,.
\end{aligned}
$$

Comparison Study

$$\begin{aligned}
\mu_0 &= 0 \\
\Rightarrow X_{i,j} &= \Delta\, Y_{i,j} + a\sqrt{\gamma_0}\,, \\
Y_{i,j} &= \alpha\, Y_{i,j-1} + \varepsilon_{i,j}\,.
\end{aligned}$$

$$\begin{aligned}
q &= 1\,, \\
a &\in \left\{0, .25, ..., 1.5, 2\right\}, \\
\Delta &\in \left\{1, 1.1, ..., 1.5, 1.75, 2\right\}.
\end{aligned}$$

$$E_0(\tau) = 500\,.$$

$$(n, \alpha) = (5, .3) \text{ or } (10, .5)\,.$$

Monte Carlo with size 10^6.

Optimal (λ_1, λ_2) and related ARL

Modified chart

$\triangle$	a 0.00	0.25	0.50	...	2.00
1.00	**500**	41.59 *(.05,.50)*	14.21 *(.10,.05)*	... ...	1.68 *(1.0,.05)*
1.10	64.00 *(1.0,.05)*	31.81 *(.05,.05)*	13.42 *(.10,.10)*	... ...	1.70 *(1.0,.05)*
1.20	23.59 *(1.0,.05)*	19.63 *(.10,.05)*	11.66 *(.10,.10)*	... ...	1.71 *(.50,1.0)*
⋮					
2.00	2.48 *(1.0,.25)*	2.44 *(1.0,.25)*	2.37 *(1.0,.25)*	... ...	1.47 *(1.0,.50)*

Residual chart

$\triangle$	a 0.00	0.25	0.50	...	2.00
1.00	**500**	41.31 *(.05,1.0)*	14.17 *(.10,1.0)*	... ...	1.69 *(1.0,1.0)*
1.10	59.40 *(1.0,.05)*	31.08 *(.05,.05)*	13.32 *(.10,.25)*	... ...	1.70 *(.50,1.0)*
1.20	21.94 *(1.0,.05)*	18.70 *(.10,.05)*	11.37 *(.10,.10)*	... ...	1.71 *(.50,1.0)*
⋮					
2.00	2.30 *(1.0,.50)*	2.27 *(1.0,.50)*	2.21 *(1.0,.50)*	... ...	1.42 *(1.0,.50)*

Conclusions

- pure shifts: both schemes act similarly,
  $a \uparrow \Rightarrow$ optimal $\lambda_1 \uparrow$
  ($\Delta = 2 \rightarrow$ overall optimal $\lambda_1 = .25/.5$)

- $\Delta > 1$: residual chart performs better

- modified chart:

  1. more attractive for practitioner,

  2. extremely time–consuming determination of critical values and worse performance.

- residual chart:

  1. more artificial appearance,

  2. quick determination of critical values and better performance.

- as usual, EWMA is better for small changes than Shewhart chart.

References

F. F. Gan (1995) Joint monitoring of process mean and variance using exponentially weighted moving average control charts.
Technometrics **37**, *446-453.*

S. Knoth, W. Schmid & A. Schöne (1998) Simultaneous Shewhart–type charts for the mean and the variance of a time series.
To appear in: H.–J. Lenz & P.–Th. Wilrich (Eds.) Frontiers of Statistical Quality Control **6**, *Physica Verlag, Heidelberg, Germany.*

S. Knoth & W. Schmid (1999) Monitoring the mean and the variance of a stationary process.
Arbeitsbericht 130, Europa–Universität Viadrina, Frankfurt(Oder), Germany.

C.–W. Lu & M. R. Reynolds, Jr. (1999) Control charts for monitoring the mean and variance of autocorrelated processes.
J. Qual. Tech. **31**, *no. 3, 259-274.*

A. Schöne & W. Schmid (1999) On the joint distribution of a quadratic and a linear form in normal variables.
To appear in J. Multiv. Anal.