Europa-Universität Viadrina Frankfurt (Oder)

Accurate ARL computation for variance monitoring schemes

Sven Knoth

Dep. of Statistics
European University Viadrina

March 2004

Synopsis

1. Some definitions.
2. Drawbacks while computing ARLs for S^2 based schemes.
3. New approach.
4. A few side effects.
5. Resumé.

Some definitions

(i) Sequential change-point model,

(ii) Control charts,

(iii) ARL.

(i) Sequential change-point model

Modeling of a stochastic process with a possible distributional change

Sequence of random variables $X_1, X_2, \ldots$ with pdf $\{F_{(i)}\}$ and a certain (unknown) time point $m =$ **change-point** with

$$F_{(i)} = \begin{cases} F_0 & , \ i < m \\ F_1 & , \ i \geq m \end{cases} .$$

Example: $F_0 = \mathcal{N}(0,1)\,,\ F_1 = \mathcal{N}(1,1)$ + independence

Notation:

$\{X_i\}_{i=1}^{m-1}$ – process **in control**,

$\{X_i\}_{i=m}^{\infty}$ – process **out of control**.

(ii) Control charts

different names, same concepts:

change point detection, continuous inspection, surveillance, monitoring ...

Aim: Detect rapidly and reliably, whether there appeared change-point m!

- Transformation $\left\{X_i\right\}_{i=1,2,\ldots,n} \to Z_n$ and
- Stopping time $L = \min\left\{n \in \mathbb{N} : Z_n \notin \mathcal{O}\right\}$,

$$\mathcal{O} = (-\infty, ucl],\ [lcl, ucl],\ [lcl, \infty) \ldots$$

At time point L observation is stopped & the scheme signals an **alarm**.

Two control chart types under consideration

– the mean monitoring case.

- (1-sided) CUSUM: PAGE (1954)

$$Z_n = \max\left\{0, Z_{n-1} + X_n - k\right\},\ Z_0 = z_0 = 0\,,$$
$$L = \inf\left\{n \in \mathbb{N} : Z_n > h\right\} \qquad \big(k = (\mu_0 + \mu_1)/2\big)\,.$$

- (2-sided) EWMA: ROBERTS (1959)

$$Z_n = (1-\lambda) Z_{n-1} + \lambda X_n\,,\ Z_0 = z_0 = \mu_0\,,$$
$$L = \inf\left\{n \in \mathbb{N} : |Z_n - \mu_0| > c\sqrt{\lambda/(2-\lambda)}\right\}\,.$$

(iii) Average Run Length (ARL)

resembles the most popular performance measure.

Notation: $E_m(.)$ expectation for given change-point m.

Definition:

$$ARL = \begin{cases} E_\infty(L) & \text{, process in control} \\ E_1(L) & \text{, process out of control} \end{cases} .$$

Note that for dealing with the ARL, the sequence $\{X_i\}$ is (strong) stationary with the same probability law for all i. Thus, e. g.,

$$ARL = E_\mu(L) =: \mathcal{L}_\mu .$$

Established methods of ARL computation

(for control charts monitoring normal mean)

Look, e. g., at the CUSUM-ARL integral equation (PAGE, 1954):

$$\mathcal{L}(s) = 1 + F_X(k-s)\,\mathcal{L}(0) + \int_0^h f_X(z+k-s)\,\mathcal{L}(z)\,dz$$

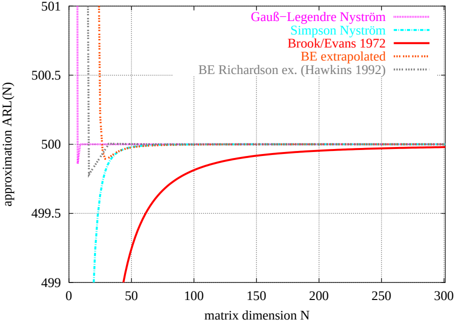
- $\mathcal{N}$, Markov chain approximation, BROOK/EVANS (1972), LUCAS/SACCUCCI (1990)

  (equivalent to Nyström method based on midpoint rule).
- *exponential*, exact, VARDEMAN/RAY (1985), GAN/CHANG (1998).
- $\mathcal{N}$, Gauß-Legendre Nyström method, VANCE (1986), CROWDER (1987).

Accuracy of different methods for ARL computation (mean)

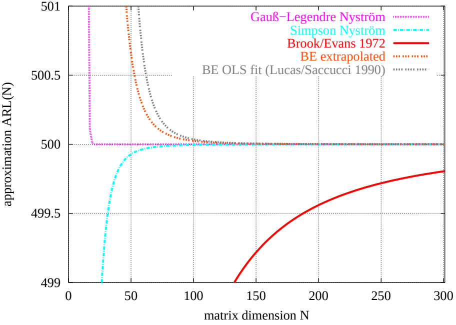
CUSUM with $k = 0.5\,,\ h = 4.389\,13$

(MC: 499.987 $\pm$ .016, 10^9 rep., 1450')



EWMA with $\lambda = 0.1\,,\ c = 2.814\,31$

(MC: 500.007 $\pm$ .016, 10^9 rep., 1400')



Variance monitoring and ARL computation

To begin with,

consider batches of size $N \geq 1$ and take

$$S_n^2 = \frac{1}{N-1} \sum_{j=1}^{N} \left(X_{nj} - \bar{X}_n\right)^2 \quad , \; \bar{X}_n = \frac{1}{N} \sum_{j=1}^{N} X_{nj}$$

$$\text{or} \quad \tilde{S}_n^2 = \frac{1}{N} \sum_{j=1}^{N} \left(X_{nj} - \mu_0\right)^2$$

as variance estimator V_n with df degrees of freedom.

Two control chart types under consideration – the variance monitoring case.

- (1-sided upper) CUSUM:

$$Z_n = \max\left\{0, Z_{n-1} + V_n - k\right\},\ Z_0 = z_0 = 0\,,$$
$$L = \inf\left\{n \in \mathbb{N} : Z_n > h\right\} \qquad \left(k = \frac{\sigma_1^2(\ln\sigma_1^2 - \ln\sigma_0^2)}{\sigma_1^2/\sigma_0^2 - 1}\right)\,.$$

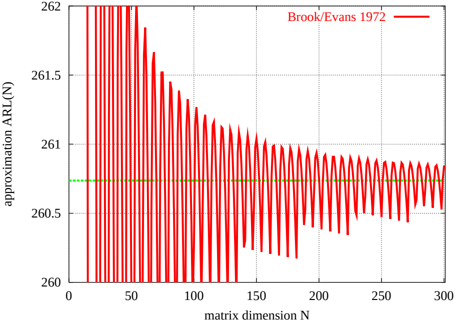
- (1-sided upper) EWMA:

$$Z_n = (1-\lambda)Z_{n-1} + \lambda V_n\,,\ Z_0 = z_0 = \sigma_0^2\,,$$
$$L = \inf\left\{n \in \mathbb{N} : Z_n > \sigma_0^2 + c\sqrt{\lambda/(2-\lambda)}\sqrt{2/df}\,\sigma_0^2\right\}\,.$$

Accuracy of different methods for ARL computation (variance)

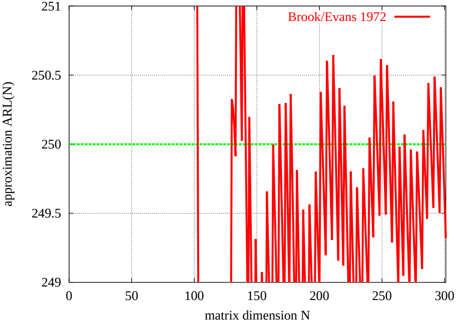
$$df = 1$$

CUSUM-S^2 with $k = 1.46\,,\ h = 10$



Ramírez/Juan 1989, BE, $N < 500$

EWMA-S^2 with $\lambda = 0.025\,,\ c = 1.661\,865$

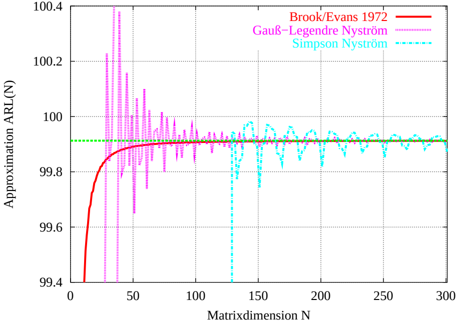


Domangue/Patch 1991, BE, $N \geq 200$

Accuracy of different methods for ARL computation (variance)

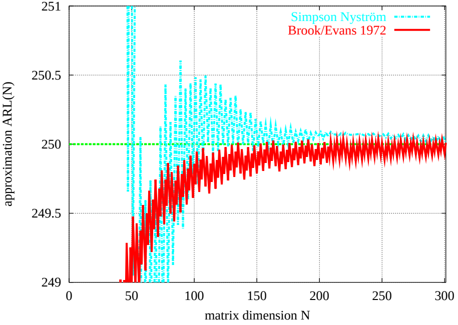
$$df = 4$$

CUSUM-S^2 with $k = 1.285\,,\ h = 2.921$



Chang/Gan 1995, MC, Acosta-Mejía et al. 1999, GL

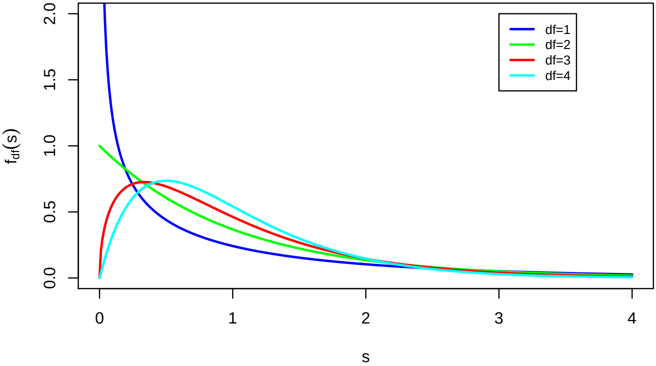
EWMA-S^2 with $\lambda = 0.18\,,\ c = 2.909\,223$



Mittag et al. 1998, BE, $N = 126$

Cause of the accuracy problem

is the behavior of the density of S^2 *at 0.*



New approach

What could we do?

- Product Nyström method: better than established ones, but worse than
- collocation.

Up to now, collocation for ARL computation was used by:

- Fellner (1990),
- Gianino/Champ/Rigdon (1990),
- Calzada/Scariano (2003).

(ARL of mean monitoring control charts only!)

Collocation for ARL computation of variance control charts

- Approximation of $\mathcal{L}(s)$ by linear combinations of
- (piecewise) Chebyshev polynomials (with additional exponential term for CUSUM),
- numerical quadrature of all integrals (less than 30 nodes are needed!),
- and solution of a linear equation system.

Sketch for the one-sided upper EWMA-S^2 control chart

$$c_u^* = \sigma_0^2 + c\sqrt{\lambda/(2-\lambda)}\sqrt{2/df}\,\sigma_0^2\,,$$

$$T_j(z) = \cos\left(j\arccos(z)\right)\,,\quad j = 0,1,\ldots,N-1\,,\ z\in[-1,1]\,,$$

$$T_j^*(z) = T_{j-1}\big((2z-c_u^*)/c_u^*\big)\,,\quad j = 1,2,\ldots,N\,,\ z\in[0,c_u^*]\,,$$

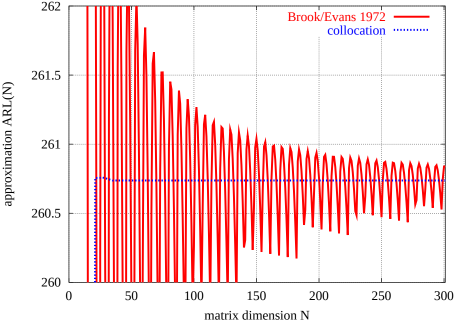
$$z_i = \frac{c_u^*}{2}\left[1+\cos\left((2i-1)/(2N)\,\pi\right)\right]\,,\quad i = 1,2,\ldots,N$$

$$\sum_{j=1}^{N} c_j\, T_j^*(z_i) = 1 + \sum_{j=1}^{N} c_j \int_{(1-\lambda)\,z_i}^{c_u^*} T_j^*(x)\,\frac{1}{\lambda}\,f_{df}\left(\frac{x-(1-\lambda)\,z_i}{\lambda}\right)\,dx\,.$$

Accuracy of different methods for ARL computation (variance)

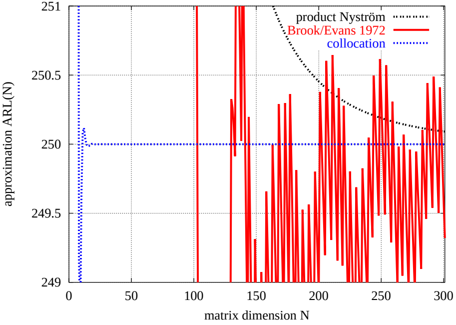
$$df = 1$$

CUSUM-S^2 with $k = 1.46\,,\ h = 10$



Ramírez/Juan 1989, BE, $N < 500$

EWMA-S^2 with $\lambda = 0.025\,,\ c = 1.661\,865$

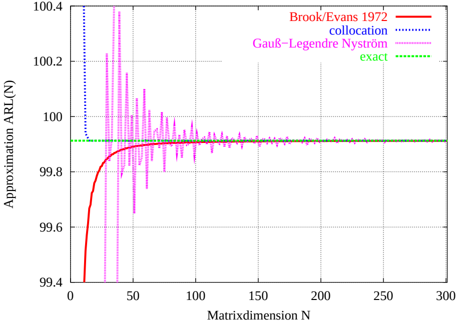


Domangue/Patch 1991, BE, $N \geq 200$

Accuracy of different methods for ARL computation (variance)

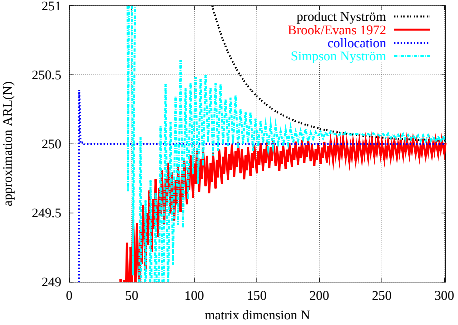
$df = 4$

CUSUM-S^2 with $k = 1.285\,,\ h = 2.921$



Chang/Gan 1995, MC

EWMA-S^2 with $\lambda = 0.18\,,\ c = 2.909\,223$



Mittag et al. 1998, BE, $N = 126$

Postscript for CUSUM-S^2 schemes

While χ^2_{df} random variables

with even degrees of freedom *df*

are Erlang distributed,

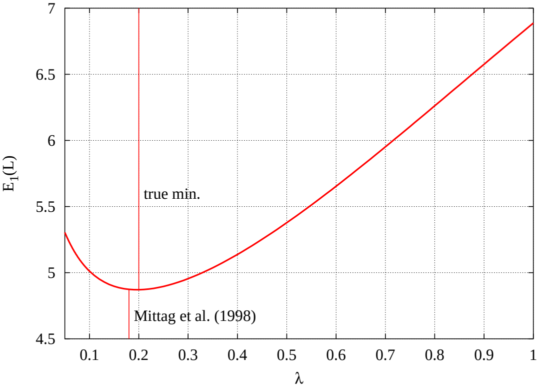
we could apply the results of Knoth (1998).

One-sided EWMA-S^2 chart for monitoring variance

$$Z_n = (1-\lambda) Z_{n-1} + \lambda S_n^2\,,\ n \ge 1\,,\ Z_0 = z_0 = \sigma_0^2 = 1\,,\ S_n^2 = 1/4 \sum_{j=1}^{5} (X_{nj} - \bar{X}_n)^2\,,$$

$$L = \inf\left\{ n \in \mathbb{N} : Z_n > \sigma_0^2 + c\sqrt{\lambda/(2-\lambda)}\sqrt{2/4}\,\sigma_0^2 \right\}\,,$$

$E_\infty(L) = 250\,,\ \sigma_1^2 = 1.5^2$ (as in Mittag et al., 1998).



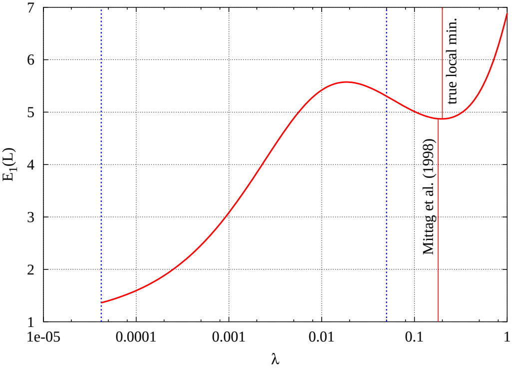
$\lambda_{\min} = 0.197$

One-sided EWMA-S^2 chart for monitoring variance II

$$Z_n = (1-\lambda) Z_{n-1} + \lambda S_n^2 \,,\; n \geq 1 \,,\; Z_0 = z_0 = \sigma_0^2 = 1 \,,\; S_n^2 = 1/4 \sum_{j=1}^{5} (X_{nj} - \bar{X}_n)^2 \,,$$

$$L = \inf \left\{ n \in \mathbb{N} : Z_n > \sigma_0^2 + c\sqrt{\lambda/(2-\lambda)}\sqrt{2/4}\,\sigma_0^2 \right\} \,,$$

$E_\infty(L) = 250 \,,\; \sigma_1^2 = 1.5^2$ (as in Mittag et al., 1998).



$\lambda = 0.000\,042$,

$c = 0.000\,064\,375\,308$,

$\widehat{E_\infty(L)} = 250.103 \pm 0.091$,

$\widehat{E_1(L)} = 1.3628 \pm 0.0000$,

10^9 rep.

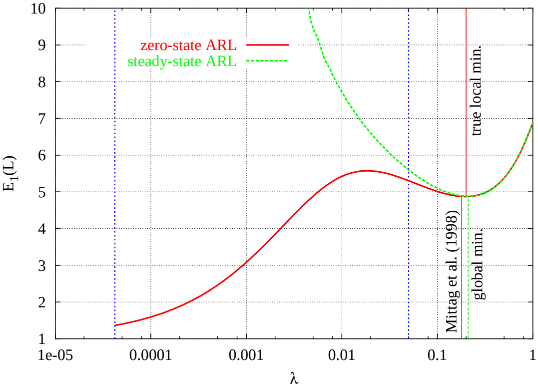
$P_\infty(L = 1) \approx 0.4!$

One-sided EWMA-S^2 chart for monitoring variance III

$$Z_n = (1-\lambda) Z_{n-1} + \lambda S_n^2 \,,\; n \ge 1 \,,\; Z_0 = z_0 = \sigma_0^2 = 1 \,,\; S_n^2 = 1/4 \sum_{j=1}^{5} (X_{nj} - \bar{X}_n)^2 \,,$$

$$L = \inf \left\{ n \in \mathbb{N} : Z_n > \sigma_0^2 + c\sqrt{\lambda/(2-\lambda)}\sqrt{2/4}\,\sigma_0^2 \right\} \,,$$

$E_\infty(L) = 250 \,,\; \sigma_1^2 = 1.5^2$ (as in Mittag et al., 1998).



$\lambda = 0.000\,042$,

$c = 0.000\,064\,375\,308$,

$\widehat{E_\infty(L)} = 250.103 \pm 0.091$,

$\widehat{E_1(L)} = 1.3628 \pm 0.0000$,

10^9 rep.

$P_\infty(L = 1) \approx 0.4$!

EWMA: S^2 based vs. $\ln S^2$ based control charts

CROWDER/HAMILTON (1992)

Note:
Here, all schemes possess a reflecting barrier at σ_0^2 and $\ln \sigma_0^2$, resp.

	$\lambda = 0.05$		$\lambda = 0.16$		$\lambda = 0.32$	
σ^2	$\ln S^2$	S^2	$\ln S^2$	S^2	$\ln S^2$	S^2
1	200.00	200.00	200.00	200.00	200.00	200.00
1.1^2	**43.04**	37.78	45.59	43.44	48.93	50.50
1.2^2	**18.10**	16.71	18.54	17.42	19.63	20.05
1.3^2	10.75	10.32	**10.52**	9.85	10.73	10.74
1.4^2	7.63	7.39	7.20	6.68	**7.08**	6.93
1.5^2	5.97	5.74	5.49	5.03	5.24	5.03
2^2	3.17	2.74	2.77	2.33	2.44	2.18

Resumé

- Variance control charts were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- Now, we are able to compute the ARL of EWMA-S^2 control charts for very small λ.
- It is not restricted to variance schemes. All distributions of chart statistics with restricted supports could be treated in this way.
- Collocation could be used for the steady-state ARL as well.
- Similar ideas could be used for (simultaneous) $\bar{X}$-S^2 schemes.

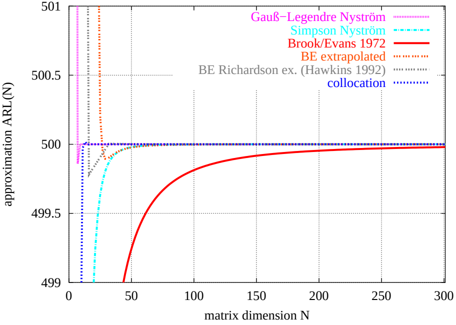
Accuracy of different methods for ARL computation (mean)

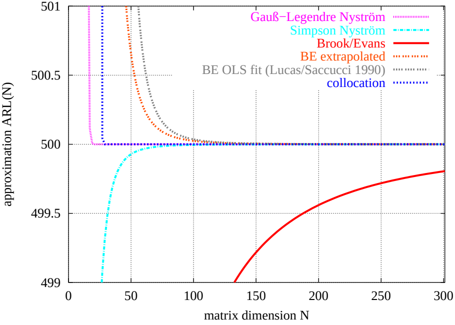
CUSUM with $k = 0.5\,,\ h = 4.389\,13$

(MC: 499.987 $\pm$ .016, 10^9 rep., 1450')

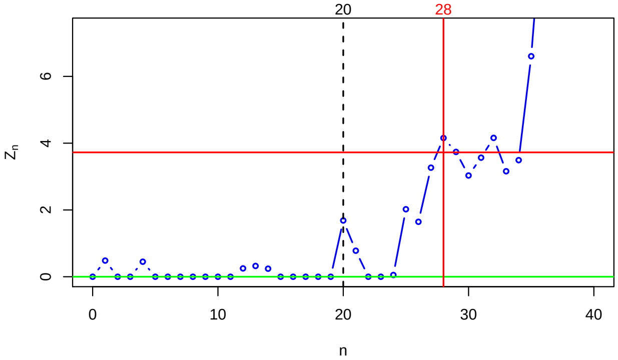
EWMA with $\lambda = 0.1\,,\ c = 2.814\,31$

(MC: 500.007 $\pm$ .016, 10^9 rep., 1400')





One-sided CUSUM-S^2 scheme



$$\sigma_0 = 1\,,\ \sigma_1 = 1.5\,,\ \text{batch size } N = 5\,,$$

$$Z_0 = 0\,,\ Z_n = \max\{0, Z_{n-1} + S_n^2 - k\}\,,$$

$$L = \min\{n \in \mathbb{N} : Z_n > h\}\,,$$

$$k = 1.46\,,\ h = 3.725\,,\ E_\infty(L) = 500\,,\ E_1(L) = 5.86\,.$$

Measuring control chart performance

- In order to evaluate "true" online procedures, first of all one has to take into account
  - elapsed time or
  - number of proceeded statistical objects.
- The control chart raises a signal with Prob. 1, whether there was a change or not.

$\rightsquigarrow$ Thus, assess properties of stopping time L!