EWMA control charts for monitoring normal variance

Sven Knoth

Fifth Annual Meeting of ENBIS

Newcastle, September 2005

Outline

1. Variance monitoring,
2. EWMA,
3. Numerical challenges,
4. Application,
5. Conclusions.

Statistics & Monitoring

Statistical Process Control (SPC)	industrial statistics
Change Point Detection Schemes	mathematical statistics
Surveillance	biostatistics
Structural Breaks	econometrics

Objective of "Control Charting"

WOODALL/MONTGOMERY (1999)

... in the area of control charting and SPC. As a general definition, we include in this area any statistical method designed to detect changes in a process over time.

Change Point Model (for the mean)

- $\{X_{ij}\}$ – series of random variables with cdf $\{P_i\}$, $j = 1, 2, \ldots, n$, n – subgroup size,

Change Point Model (for the mean)

- $\{X_{ij}\}$ – series of random variables with cdf $\{P_i\}$, $j = 1, 2, \ldots, n$, n – subgroup size,
- assumptions: $P_i(x) = \Phi\big((x - \mu_{(i)})/\sigma_{(i)}\big)$, independence,

Change Point Model (for the mean)

- $\{X_{ij}\}$ – series of random variables with cdf $\{P_i\}$, $j = 1, 2, \ldots, n$, n – subgroup size,
- assumptions: $P_i(x) = \Phi\big((x - \mu_{(i)})/\sigma_{(i)}\big)$, independence,
- model:

$$\mu_{(i)} = \begin{cases} \mu_0 & , \ i < m \\ \mu_1 & , \ i \geq m \end{cases} , \ \sigma_{(i)} = \sigma_0$$

with unkown parameter m (change point),

Change Point Model (for the mean)

- $\{X_{ij}\}$ – series of random variables with cdf $\{P_i\}$, $j = 1, 2, \ldots, n$, n – subgroup size,
- assumptions: $P_i(x) = \Phi\big((x - \mu_{(i)})/\sigma_{(i)}\big)$, independence,
- model:

$$\mu_{(i)} = \begin{cases} \mu_0 & , \ i < m \\ \mu_1 & , \ i \geq m \end{cases} , \ \sigma_{(i)} = \sigma_0$$

  with unkown parameter m (change point),
- more formal objective of SPC:

  detect m as fast and reliable as possible.

Variance monitoring

- variance within samples (e. g., uniformity measures),

Variance monitoring

- variance within samples (e. g., uniformity measures),
- sample statistics such as S^2, S, R or certain robust estimators, e. g.

$$S_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(X_{ij} - \bar{X}_i\right)^2 ,$$

Variance monitoring

- variance within samples (e. g., uniformity measures),
- sample statistics such as S^2, S, R or certain robust estimators, e. g.

$$S_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(X_{ij} - \bar{X}_i\right)^2 ,$$

- change point model:

$$\sigma_{(i)}^2 = \begin{cases} \sigma_0^2 & , \ i < m \\ \sigma_1^2 & , \ i \geq m \end{cases} ,$$

Variance monitoring

- variance within samples (e. g., uniformity measures),
- sample statistics such as S^2, S, R or certain robust estimators, e. g.

$$S_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(X_{ij} - \bar{X}_i\right)^2 ,$$

- change point model:

$$\sigma_{(i)}^2 = \begin{cases} \sigma_0^2 & , \; i < m \\ \sigma_1^2 & , \; i \geq m \end{cases} ,$$

- single variance change or simultaneous mean/variance change.

Upper S^2 EWMA control charts

Exponentially Weighted Moving Average

Roberts (1959), Wortham/Ringer (1971), Sweet (1986)

Upper S^2 EWMA control charts

Exponentially Weighted Moving Average

Roberts (1959), Wortham/Ringer (1971), Sweet (1986)

$$Z_0 = z_0 = E_{\sigma_0}(S_i^2) = \sigma_0^2 \,,$$

Upper S^2 EWMA control charts

Exponentially Weighted Moving Average

Roberts (1959), Wortham/Ringer (1971), Sweet (1986)

$$Z_0 = z_0 = E_{\sigma_0}(S_i^2) = \sigma_0^2 \,,$$

$$Z_i = (1-\lambda)\, Z_{i-1} + \lambda\, S_i^2 \qquad , \; \lambda \in (0,1] \,,$$

Upper S^2 EWMA control charts

Exponentially Weighted Moving Average

Roberts (1959), Wortham/Ringer (1971), Sweet (1986)

$$Z_0 = z_0 = E_{\sigma_0}(S_i^2) = \sigma_0^2\,,$$

$$Z_i = (1-\lambda)\, Z_{i-1} + \lambda\, S_i^2 \qquad ,\ \lambda \in (0,1]\,,$$

$$E_{\sigma_0}(Z_i) = \sigma_0^2\,,$$

$$Var_{\sigma_0}(Z_i) = \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right) \times Var_{\sigma_0}(S_i^2)$$

$$\to \frac{\lambda}{2-\lambda} \times \frac{2}{n-1}\,\sigma_0^4 =: \sigma_Z^2\,,$$

Upper S^2 EWMA control charts

Exponentially Weighted Moving Average

Roberts (1959), Wortham/Ringer (1971), Sweet (1986)

$$Z_0 = z_0 = E_{\sigma_0}(S_i^2) = \sigma_0^2\,,$$

$$Z_i = (1-\lambda)\,Z_{i-1} + \lambda\,S_i^2 \qquad ,\ \lambda \in (0,1]\,,$$

$$E_{\sigma_0}(Z_i) = \sigma_0^2\,,$$

$$Var_{\sigma_0}(Z_i) = \frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2i}\right) \times Var_{\sigma_0}(S_i^2)$$

$$\to \frac{\lambda}{2-\lambda} \times \frac{2}{n-1}\,\sigma_0^4 =: \sigma_Z^2\,,$$

$$L = \inf\left\{i \in \mathbb{N} : Z_i > \sigma_0^2 + c\,\sigma_Z\right\}.$$

Modifications

$$Z_i^* = \max\left\{(1-\lambda)\, Z_{i-1}^* + \lambda\, S_i^2, \sigma_0^2\right\},$$

Modifications

$$Z_i^* = \max\left\{(1-\lambda)\, Z_{i-1}^* + \lambda\, S_i^2, \sigma_0^2\right\},$$

$$L_{\text{two-sided}} = \inf\left\{i \in \mathbb{N} : Z_i - \sigma_0^2 \notin [c_{\text{lower}}, c_{\text{upper}}] \times \sigma_Z\right\},$$

Modifications

$$Z_i^* = \max\left\{(1-\lambda)\, Z_{i-1}^* + \lambda\, S_i^2, \sigma_0^2\right\},$$

$$L_{\text{two-sided}} = \inf\left\{i \in \mathbb{N} : Z_i - \sigma_0^2 \notin [c_{\text{lower}}, c_{\text{upper}}] \times \sigma_Z\right\},$$

use S_i or $\ln S_i^2$ or R_i or $\ldots$

Performance measures for control charting

Denote $E_m(\cdot)$ the expectation for a change point at m.

1. **Average Run Length (ARL):** $E_\infty(L)$ & $E_1(L) \rightsquigarrow E_\sigma(L)$

Performance measures for control charting

Denote $E_m(\cdot)$ the expectation for a change point at m.

1. **Average Run Length (ARL):** $E_\infty(L)$ & $E_1(L) \rightsquigarrow E_\sigma(L)$

2. Lorden's worst case: $E_\infty(L)$ & $\sup_m \operatorname{ess\,sup} E_m\big((L-m+1)^+ | \mathcal{F}_{m-1}\big)$

Performance measures for control charting

Denote $E_m(\cdot)$ the expectation for a change point at m.

1. **Average Run Length (ARL):** $E_\infty(L)$ & $E_1(L) \rightsquigarrow E_\sigma(L)$
2. Lorden's worst case: $E_\infty(L)$ & $\sup_m \operatorname{ess\,sup} E_m\big((L-m+1)^+|\mathcal{F}_{m-1}\big)$
3. Roberts' steady state: $E_\infty(L)$ & $\lim_{m\to\infty} E_m\big(L-m+1|L \geq m-1\big)$
4. ...

ARL computation

1. Markov chain approximation (Brook/Evans, 1972, CUSUM) Lucas/Saccucci (1990),

ARL computation

1. Markov chain approximation (Brook/Evans, 1972, CUSUM)
   Lucas/Saccucci (1990),
2. Nyström method (Vance, 1986, CUSUM)
   Crowder (1987),

ARL computation

1. Markov chain approximation (Brook/Evans, 1972, CUSUM)
   Lucas/Saccucci (1990),
2. Nyström method (Vance, 1986, CUSUM)
   Crowder (1987),
3. Champ/Rigdon (1991): comparison of the previous ones,

ARL computation

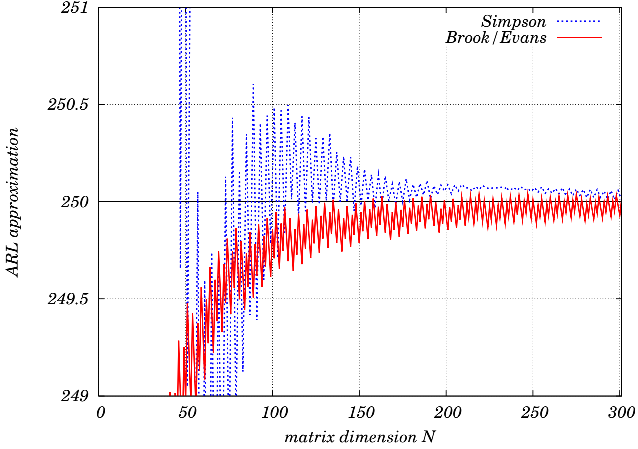
1. Markov chain approximation (Brook/Evans, 1972, CUSUM)
   Lucas/Saccucci (1990),
2. Nyström method (Vance, 1986, CUSUM)
   Crowder (1987),
3. Champ/Rigdon (1991): comparison of the previous ones,
4. Monte Carlo,
5. ...,

ARL computation

1. Markov chain approximation (Brook/Evans, 1972, CUSUM)
   Lucas/Saccucci (1990),
2. Nyström method (Vance, 1986, CUSUM)
   Crowder (1987),
3. Champ/Rigdon (1991): comparison of the previous ones,
4. Monte Carlo,
5. ...,
6. Collocation (Fellner, 1990, CUSUM)
   Gianino/Champ/Rigdon (1990).

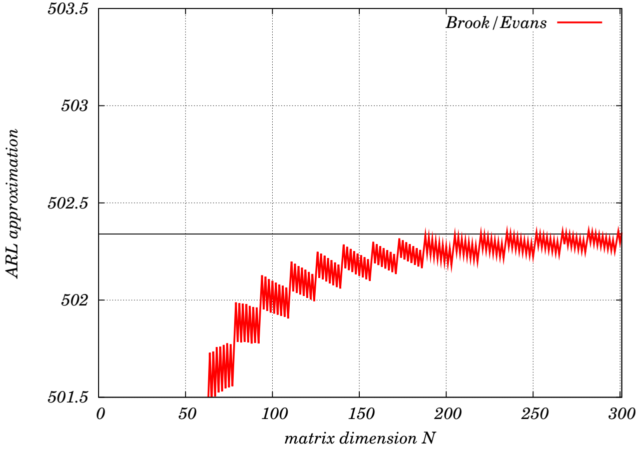
Application to upper S^2 EWMA ARLs I

$n = 5,\ \lambda = 0.18,\ c = 2.90922$ (final $E_\infty(L)$ value is 250)



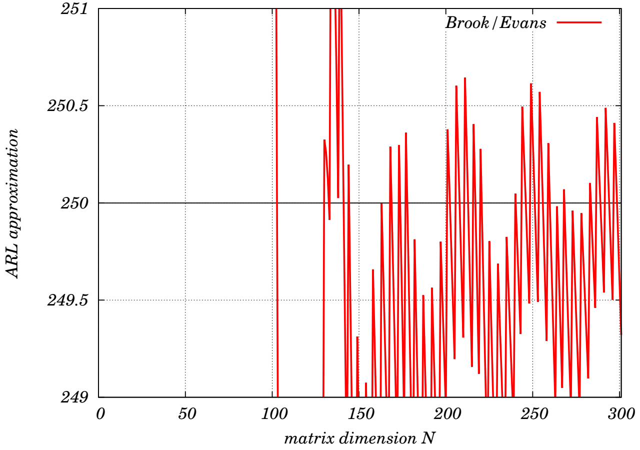
Application to upper S^2 EWMA ARLs II

$n = 3$, $\lambda = 0.18$, $c = 3.61384$ (final $E_\infty(L)$ value is 502.34)



Application to upper S^2 EWMA ARLs III

$n = 2,\ \lambda = 0.025,\ c = 1.66186$ (final $E_\infty(L)$ value is 250)



Numerics

- $f(z_{i-1} \to z_i) = 0$ for $z_i < (1-\lambda)z_{i-1} \rightsquigarrow$ Nyström fails,

Numerics

- $f(z_{i-1} \to z_i) = 0$ for $z_i < (1-\lambda) z_{i-1} \rightsquigarrow$ Nyström fails,
- product Nyström,

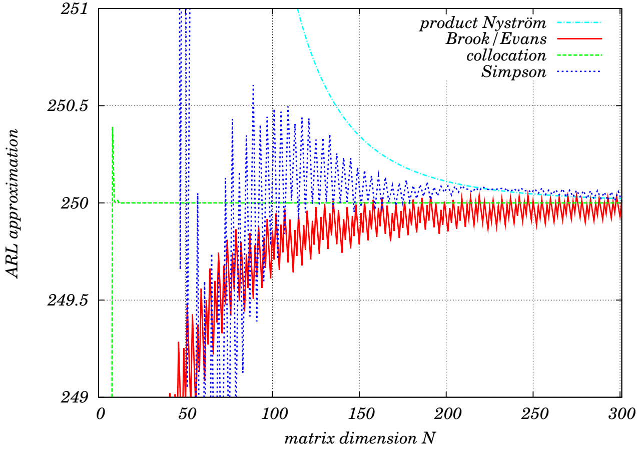
Numerics

- $f(z_{i-1} \to z_i) = 0$ for $z_i < (1-\lambda) z_{i-1} \rightsquigarrow$ Nyström fails,
- product Nyström,
- collocation: $ARL(z) = \sum_{i=1}^{N} c_i T_i(z)$,

  $T_i(z)$ – monomials, Lagrange or Chebyshev polynomials

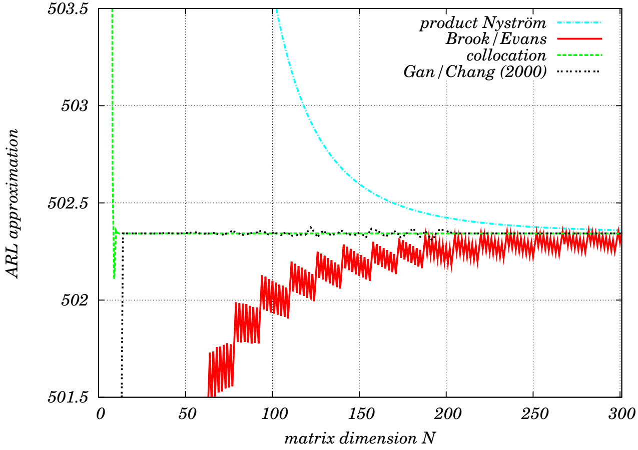
Application to upper S^2 EWMA ARLs I

$n = 5$, $\lambda = 0.18$, $c = 2.90922$ (final $E_\infty(L)$ value is 250)



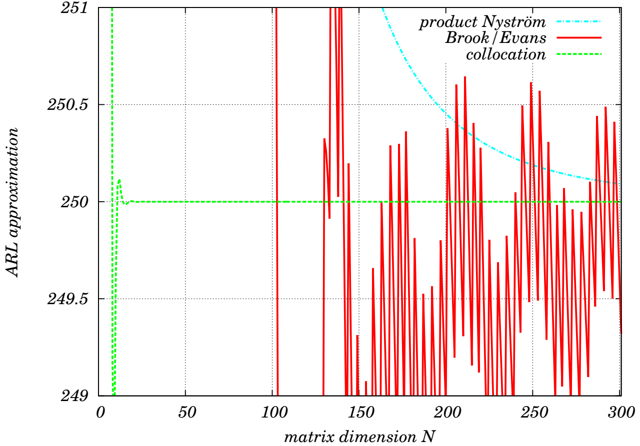
Application to upper S^2 EWMA ARLs II

$n = 3$, $\lambda = 0.18$, $c = 3.61384$ (final $E_\infty(L)$ value is 502.34)



Application to upper S^2 EWMA ARLs III

$n = 2,\ \lambda = 0.025,\ c = 1.66186$ (final $E_\infty(L)$ value is 250)



How fast is the collocation approach?

$n = 2$, $\lambda = 0.025$, $c = 1.66186$ (final $E_\infty(L)$ value is 250)

Method	matrix dimension N				
	25	51	101	201	301
Markov chain	103.7077	307.4809	254.6729	250.3782	249.3206
	< 1[1]	1	3	13	39
factor Nyström	< 0	376.0594	256.6795	250.4456	250.0908
	< 1	1	2	9	27
Collocation	249.9999	249.9997	249.9997	249.9997	249.9997
	1	3	11	40	101

[1] CPU time in hundredth seconds on an Athlon XP 1.4 GHz

Modifications etc.

- two-sided or one-sided with reflection barrier: piecewise collocation,

Modifications etc.

- two-sided or one-sided with reflection barrier: piecewise collocation,
- similar ideas could be employed for $\bar{X}$-S^2 EWMA charts – enhance Gan (1995),

Modifications etc.

- two-sided or one-sided with reflection barrier: piecewise collocation,
- similar ideas could be employed for $\bar{X}$-S^2 EWMA charts – enhance Gan (1995),
- for $\ln S^2$ EWMA charts the Gauss-Legendre Nyström method is the most powerful – Crowder/Hamilton (1992).

Comparison S^2 and $\ln S^2$ EWMA charts

Setup: derived from Crowder/Hamilton (1992)

Comparison S^2 and $\ln S^2$ EWMA charts

Setup: derived from Crowder/Hamilton (1992)

1. $\ln S^2$ EWMA: $z_{\text{reflect}} = \ln \sigma_0^2$ or $E_{\sigma_0}(\ln S_i^2)$

$$Z_0 = z_{\text{reflect}} \,,\; Z_i = \max\{(1-\lambda) Z_{i-1} + \lambda \ln S_i^2, z_{\text{reflect}}\},$$

Comparison S^2 and $\ln S^2$ EWMA charts

Setup: derived from Crowder/Hamilton (1992)

1. $\ln S^2$ EWMA: $z_{\text{reflect}} = \ln \sigma_0^2$ or $E_{\sigma_0}(\ln S_i^2)$

$$Z_0 = z_{\text{reflect}}\ ,\ Z_i = \max\{(1-\lambda)Z_{i-1} + \lambda \ln S_i^2, z_{\text{reflect}}\},$$

2. S^2 EWMA: $z_{\text{reflect}} = \sigma_0^2 = E_{\sigma_0}(S_i^2)$

$$Z_0 = z_{\text{reflect}}\ ,\ Z_i = \max\{(1-\lambda)Z_{i-1} + \lambda S_i^2, z_{\text{reflect}}\},$$

Comparison S^2 and $\ln S^2$ EWMA charts

Setup: derived from Crowder/Hamilton (1992)

1. $\ln S^2$ EWMA: $z_{\text{reflect}} = \ln \sigma_0^2$ or $E_{\sigma_0}(\ln S_i^2)$

$$Z_0 = z_{\text{reflect}}\ ,\ Z_i = \max\{(1-\lambda)Z_{i-1} + \lambda \ln S_i^2, z_{\text{reflect}}\},$$

2. S^2 EWMA: $z_{\text{reflect}} = \sigma_0^2 = E_{\sigma_0}(S_i^2)$

$$Z_0 = z_{\text{reflect}}\ ,\ Z_i = \max\{(1-\lambda)Z_{i-1} + \lambda S_i^2, z_{\text{reflect}}\},$$

3. both: $\lambda \in \{0.05, 0.16, 0.32\}$, $E_\infty(L) = 200$, $\sigma_0^2 = 1$.

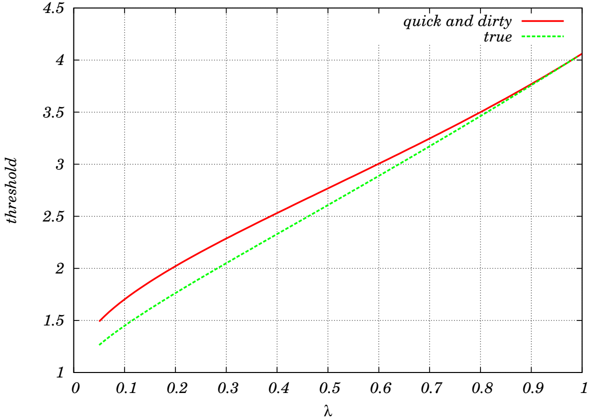
Comparison S^2 and $\ln S^2$ EWMA charts

Results: ARL

σ^2	$\lambda = 0.05$			$\lambda = 0.16$			$\lambda = 0.32$		
	$\ln S^2$		S^2	$\ln S^2$		S^2	$\ln S^2$		S^2
	0	-.267		0	-.267		0	-.267	
1	200.00	200.00	200.00	200.00	200.00	200.00	200.00	200.00	200.00
1.1^2	**43.04**	41.55	37.78	45.59	43.19	43.44	48.93	46.90	50.50
1.2^2	**18.10**	19.92	16.71	18.54	18.46	17.42	19.63	19.12	20.05
1.3^2	10.75	13.11	10.32	**10.52**	11.11	9.85	10.73	10.79	10.74
1.4^2	7.63	9.93	7.39	7.20	7.96	6.68	**7.08**	7.34	6.93
1.5^2	5.97	8.11	5.74	5.49	6.27	5.03	5.24	5.57	5.03
2^2	3.17	4.67	2.74	2.77	3.37	2.33	2.44	2.78	2.18

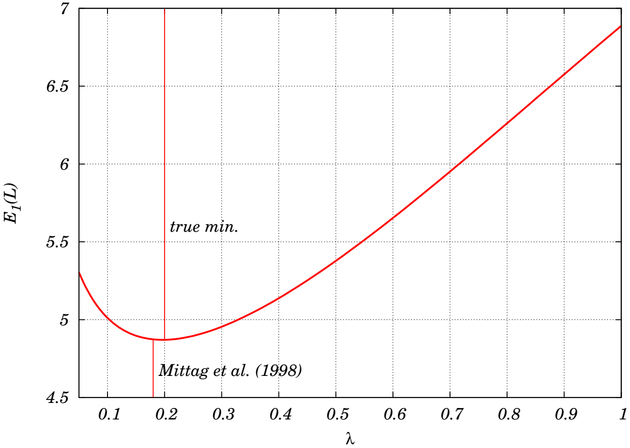
Calculating limits: quick and dirty vs. costly and correct

$$n = 5,\ E_\infty(L) = 370\,,\ \textit{threshold} = \sigma_0^2 + \sigma_0^2\sqrt{\frac{\lambda}{2-\lambda}} \times \begin{cases} c_{\text{coll}}\sqrt{\frac{2}{n-1}} & \text{, correct} \\ \frac{\chi^2_{n-1;1-1/370}}{n-1} - 1 & \text{, quick and dirty} \end{cases}$$



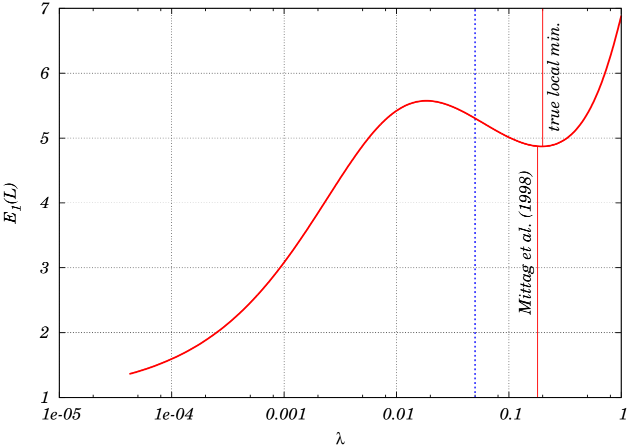
Running for the "optimal" λ

$n = 5$, $E_\infty(L) = 250$, $\sigma_1 = 1.5$, Mittag et al. (1998)



Very small λ

$n = 5,\ E_\infty(L) = 250,\ \sigma_1 = 1.5$, Mittag et al. (1998)



Very small λ

"winning scheme"

$$\lambda = 0.000\,042\,,$$

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

Very small λ

"winning scheme"

$$\lambda = 0.000\,042\,,$$

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

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

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

$$10^9 \text{ rep.}\,,$$

Very small λ

"winning scheme"

$$\lambda = 0.000\,042\,,$$

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

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

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

$$10^9 \text{ rep.}\,,$$

$$P_\infty(L=1) \approx 0.4\,.$$

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).

Conclusions

- EWMA control charts for the variance were often applied, but the accuracy in computing their ARL values was poor.
- Collocation allows to cure these accuracy problems.
- The algorithm will be implemented in R library spc.
- The quick and dirty approach of deriving S^2 EWMA control limits provides suitable results.
- Now, we are able to compute the ARL of S^2 EWMA control charts for very small λ.
- S^2 EWMA seems to be better than EWMA $\ln S^2$ (supported by earlier, more extensive comparison studies).