

# 非正規分布における分散型 EWMA 管理図の ロバストネスに関する研究

## A Study on the Robustness to Non Normality of the EWMA Control Charts for the Dispersion

経営システム工学専攻 松木 崇  
MATSUKI Takashi

### 1 はじめに

品質管理の分野において、品質や製造工程が安定状態で管理されていることを判断するために、管理図が利用される。代表的な管理図として、 $\bar{X}$  管理図や  $\bar{X}$ -R 管理図が挙げられるが、これらの管理図は品質が異常か否かを判断するために用いられる。一方、製造工程が異常か否かを判断する為の管理図として、Cumulative Sum (CUSUM) 管理図 [7] と Exponentially Weighted Moving Average (EWMA) 管理図 [8] が提案された。これらは工程変化の検出力が高い管理図として知られている。本研究では、最新のデータを重視すると共に古いデータを完全に切り捨てないという利点を持つ、EWMA 管理図を扱う。

一般的な EWMA 管理図の研究は、データに正規分布を仮定した研究が多く行われている。しかし、非正規分布を仮定した研究も重要であり、この問題に対して、Borror [1] らは、単一データについて歪みを持つガンマ分布と対称な t 分布のロバスト性を検証した。また、Stoumbos [10] らによって複数データ (サブグループ数が複数) に対応する EWMA 管理図のロバスト性の検証が同様の非正規分布で行われている。一方、データの分散について EWMA 管理図を応用することが可能であり、これも重要な問題である。Maraverakis らの先行研究 [6] では、単一データに対応する分散に関する EWMA 管理図が提案されており、非正規分布のロバスト性を検証している。しかし、t 分布とガンマ分布以外の非正規分布の検証や複数データへの対応など、分散に関する EWMA 管理図の研究は十分に議論されていない。

本研究の目的は、Maraverakis [6] らの先行研究で使用された、単一データに対応する分散に関する EWMA 管理図を用いてガンマ、t 分布以外の非正規分布のロバスト性を検証する。また、先行研究では行われていない「複数データに対応する分散に関する EWMA 管理図」を提案し、非正規分布のロバスト性の検証を行う。

### 2 EWMA 管理図

#### 2.1 EWMA 管理図

工程の安定状態を把握するための管理図として、EWMA 管理図が挙げられる。EWMA 管理図は EWMA 統計量が管理限界 (UCL, LCL) からの逸脱を判断して工程異常を検出する。EWMA 統計量  $Z_i$  は、データを  $x_i$ 、平滑化パラメータ  $\lambda$  を  $0 < \lambda \leq 1$  として以下のように定義する。なお、データ  $x_i$  は  $N(\mu_0, \sigma_0^2)$  の正規分布を仮定する。

$$\begin{aligned} Z_i &= \lambda x_i + (1 - \lambda) Z_{i-1}, \\ Z_0 &= E(x_i) = \mu_0. \end{aligned}$$

EWMA 管理図の管理限界は、 $h$  を管理限界の幅を設定するパラメータとして以下に定義する。

$$\begin{aligned} UCL &= \mu_0 + h\sigma \sqrt{\frac{\lambda}{2 - \lambda}}, \\ LCL &= \mu_0 - h\sigma \sqrt{\frac{\lambda}{2 - \lambda}}. \end{aligned}$$

#### 2.2 上側を考慮した分散に関する EWMA 管理図

EWMA 統計量のデータ  $x_i$  を分散に関する推定量  $V_i$  に代えることで、分散に関する EWMA 管理図として応用できる。分散に関する EWMA 管理図は、工程の分散の傾向を知ることにつながる。即ち、分散に関する EWMA 統計量の減少は工程のバラツキが少ない安定した状態を示し、増加は工程にバラツキが生じた安定性に欠ける状態を示す。

ここで問題となるのが、「分散の増加」の検出である。この問題に対応するために、上側を考慮した分散に関する EWMA 管理図に拡張できる。上側を考慮した分散に関する EWMA 統計量  $UY_i$  は EWMA 統計量  $Z_i$  を拡張して以下に定義する。

$$\begin{aligned} UY_i &= \lambda V_i + (1 - \lambda) \max(UY_{i-1}, UY_0), \\ UY_0 &= E(V_i). \end{aligned}$$

また、管理限界は上側のみを利用し、以下に定義する。

$$UCL = E(V_i) + h \sqrt{\frac{\lambda}{2 - \lambda} \text{Var}(V_i)}.$$

### 3 扱う EWMA 管理図

本節では、本研究で扱った上側を考慮した分散に関する EWMA 管理図の統計量と UCL を説明する。ここで、データ数 (サブグループ数  $j$ ) が「単一」と「複数」のものに分けて標記した。なお、データ  $x_i$  の平均は  $\mu_0$ 、標準偏差は  $\sigma_0$  である。

#### 3.1 単一データに対応する分散に関する EWMA 管理図

Maraverakis [6] らの先行研究では、5 種類の「単一データ」に対応する上側を考慮した分散に関する EWMA 管理図を提案している。本研究でも、先行研究で使用した既存の管理図を利用した。

##### 3.1.1 WR 統計量

工程の推移を検出するために偏差平方を利用した EWMA 管理図が提案された [12]。この管理図の統計量を WR 統計量とし、その統計量と管理限界を以下に示す。

$$\begin{aligned} WR_i &= \lambda(x_i - \mu_0)^2 + (1 - \lambda) \max(WR_{i-1}, \sigma_0^2), \\ WR_0 &= \sigma_0^2, \\ UCL &= \sigma_0^2 + h_{wr} \sigma_0^2 \sqrt{\frac{2\lambda}{2-\lambda}}. \end{aligned}$$

##### 3.1.2 SR 統計量

分散に関する推定量に絶対偏差を利用した EWMA 管理図が提案された [10]。この管理図の統計量を SR 統計量とし、その統計量と管理限界を以下に示す。

$$\begin{aligned} SR_i &= \lambda |x_i - \mu_0| + (1 - \lambda) \max\left(SR_{i-1}, \sigma_0 \sqrt{\frac{2}{\pi}}\right), \\ SR_0 &= \sigma_0 \sqrt{\frac{2}{\pi}}, \\ UCL &= \sigma_0 \sqrt{\frac{2}{\pi}} + h_{sr} \sigma_0 \sqrt{1 - \frac{2}{\pi}} \sqrt{\frac{\lambda}{2-\lambda}}. \end{aligned}$$

##### 3.1.3 HO 統計量

Hawkins and Olwell [5] は scale の変化に対応する CUSUM 型の統計量を提案し、この統計量を Maraverakis [6] らは EWMA 管理図タイプに応用した。これを HO 統計量とし、その統計量と管理限界を以下に示す。

$$\begin{aligned} HO_i &= \lambda \sqrt{|x_i - \mu_0|} \\ &\quad + (1 - \lambda) \max\left(HO_{i-1}, 2^{(\frac{3}{4})} \Gamma\left(\frac{3}{4}\right) \sqrt{\frac{\sigma_0}{2\pi}}\right), \\ HO_0 &= 2^{(\frac{3}{4})} \Gamma\left(\frac{3}{4}\right) \sqrt{\frac{\sigma_0}{2\pi}}, \\ UCL &= 2^{(\frac{3}{4})} \Gamma\left(\frac{3}{4}\right) \sqrt{\frac{\sigma_0}{2\pi}} \\ &\quad + h_{ho} \sqrt{\sigma_0 \left(\sigma_0 \sqrt{\frac{2}{\pi}} - \frac{\sqrt{2} \Gamma^2(\frac{3}{4})}{\pi}\right) \frac{\lambda}{2-\lambda}}. \end{aligned}$$

#### 3.1.4 DP 統計量

分散に関する推定量を  $D_i = \frac{x_i - \mu_0}{\sigma_0}$  の絶対値の  $\alpha$  乗となる統計量が提案された [3]。この統計量を DP 統計量し、以下に示す。

$$\begin{aligned} DP_i^\alpha &= \lambda |D_i|^\alpha + (1 - \lambda) \max(DP_{i-1}, DP_0), \\ DP_0^\alpha &= E(DP_i^\alpha). \end{aligned}$$

ここで、 $\alpha=0.5$  の場合を DP1、 $\alpha=2$  の場合を DP2 する。

DP1 統計量の期待値、分散は、

$$\begin{aligned} E[DP_i^{0.5}] &= \left(\frac{\sqrt{2}}{\pi}\right)^{\frac{1}{2}} \Gamma\left(\frac{3}{4}\right), \\ \text{Var}[DP_i^{0.5}] &= \frac{\sqrt{2}\lambda}{(2-\lambda)\pi} \left[\sqrt{\pi} - \Gamma^2\left(\frac{3}{4}\right)\right] \end{aligned}$$

であり、DP2 統計量の期待値、分散は以下を満たす。

$$\begin{aligned} E[DP_i^2] &= 1, \\ \text{Var}[DP_i^2] &= \frac{2\lambda}{2-\lambda}. \end{aligned}$$

また、管理限界は上記の値を用いて、以下に与える。

$$UCL = E[DP_i^\alpha] + h_{dp} \sqrt{\text{Var}[DP_i^\alpha]}.$$

#### 3.2 複数データに対応する分散に関する EWMA 管理図

管理図は「複数データ (1 つの群に複数のデータ: サブグループ数が複数のデータ)」に対応する必要があるが、分散に関する EWMA 管理図は十分な対応がなされていない。

そこで本研究では、複数データに対応する上側を考慮した分散に関する EWMA 管理図を提案した。本節では提案した管理図の統計量と管理限界を説明する。ここで、データ  $x_{i,j}$  は  $i$  行  $j$  列のデータ ( $j$  個のサブグループを持つデータ) を仮定している。また、mean と Var は平均と分散を示す。

##### 3.2.1 SD 統計量

複数データに対応する分散に関する EWMA 管理図の基準となる管理図として、分散に関する推定量に、データの各行の標準偏差 ( $j$  個のデータの標準偏差) を取り入れた管理図を提案する。この管理図の統計量を SD 統計量とし、管理限界と共に以下に示す。

$$\begin{aligned} SD_i &= \lambda \times \text{sd}(x_{i \cdot}) + (1 - \lambda) \max(SD_0, SD_{i-1}), \\ SD_0 &= \text{mean}(\text{sd}(x_{i \cdot})), \\ UCL &= \text{mean}(\text{sd}(x_{i \cdot})) + h_{sd} \sqrt{\frac{\lambda}{2-\lambda} \times \text{Var}(\text{sd}(x_{i \cdot}))}. \end{aligned}$$

### 3.2.2 MAD 推定量を利用した統計量

Hample [4] により MAD (Median Absolute Deviation) 推定量が提案された. MAD 推定量  $M_i$  を以下に与える.

$$M_i = 1.4826 \times \text{median}_i | x_{i,j} - \text{median}_j(x_{i,\cdot}) |.$$

ここで, MAD 推定量を適応した EWMA 管理図を提案する. この管理図の統計量を MAD 統計量とし, 管理限界と共に以下に示す.

$$MAD_i = \lambda \times M_i + (1 - \lambda) \max(MAD_0, MAD_{i-1}),$$

$$MAD_0 = \text{mean}(M_i),$$

$$UCL = \text{mean}(M_i) + h_{mad} \sqrt{\frac{\lambda}{2 - \lambda} \times \text{Var}(M_i)}.$$

### 3.2.3 S 推定量を利用した統計量

Rousseeuw and Croux [9] により S 推定量が提案された. S 推定量  $S_i$  を以下に与える.

$$S_i = 1.1926 \times \text{median}_i(\text{median}_j | x_i - x_j |).$$

ここで, S 推定量を適応した EWMA 管理図を提案する. この管理図の統計量を S 統計量とし, 管理限界と共に以下に示す.

$$Sn_i = \lambda \times S_i + (1 - \lambda) \max(Sn_0, Sn_{i-1}),$$

$$Sn_0 = \text{mean}(S_i),$$

$$UCL = \text{mean}(S_i) + h_s \sqrt{\frac{\lambda}{2 - \lambda} \times \text{Var}(S_i)}.$$

### 3.2.4 FQAD 推定量を利用した統計量

MAD 推定量の median の代わりに, 第 1 四分位点 (25% 点) を利用した推定量を FQAD (First Quartile Absolute Deviation) 推定量として提案する. FQAD 推定量  $F_i$  を, 以下に定義する.

$$F_i = 3.1383 \times \text{fquartile} | x_{i,j} - \text{median}(x_{i,\cdot}) |.$$

fquartile は第 1 四分位点 (25% 点) を意味している. ここで, FQAD 推定量を適応した EWMA 管理図を提案する. この管理図の統計量を FQAD 統計量とし, 管理限界と共に以下に示す.

$$FQ_i = \lambda \times F_i + (1 - \lambda) \max(FQ_0, FQ_{i-1}),$$

$$FQ_0 = \text{mean}(F_i),$$

$$UCL = \text{mean}(F_i) + h_f \sqrt{\frac{\lambda}{2 - \lambda} \text{Var}(F_i)}.$$

### 3.2.5 TQAD 統計量

MAD 推定量の median 部分を第 3 四分位点 (75% 点) に変えた推定量を, TQAD (Third Quartile Absolute Deviation) 推定量として提案する. TQAD 推定

量  $T_i$  を以下に定義する.

$$T_i = 0.8693 \times \text{tquartile} | x_{i,j} - \text{median}(x_{i,\cdot}) |.$$

tquartile は第 3 四分位点 (75% 点) を意味している. ここで, TQAD 推定量を適応した EWMA 管理図を提案する. この管理図の統計量を TQAD 統計量とし, 管理限界と共に以下に示す.

$$TQ_i = \lambda \times T_i + (1 - \lambda) \max(TQ_0, TQ_{i-1}),$$

$$TQ_0 = \text{mean}(T_i),$$

$$UCL = \text{mean}(T_i) + h_t \sqrt{\frac{\lambda}{2 - \lambda} \text{Var}(T_i)}.$$

## 4 評価・検証方法

### 4.1 ARL

非正規分布を仮定してロバスト性を検証するために, ARL (Average Run Length) を用いる. ARL は管理図のパフォーマンスを計る指標としてよく知られており, 管理限界を超えるまでの連の長さの平均を求めることで与えられる. つまり, ARL は異常値の発生間隔を意味する.

また, 正規分布を仮定すると, ARL=370.4 となることは良く知られている [6]. 本研究では非正規分布を仮定した ARL 値が 370.4 にどの程度近くなるかで, ロバスト性を検証する.

### 4.2 検証

本研究ではシミュレーションから ARL 値を求めロバスト性を検証した. なお, 平滑化パラメータ  $\lambda$  は 0.05, 0.1, 0.2, 0.3 の 4 種類とし, 非正規分布はガンマ分布 (G(4,1), G(2,1), G(0.5,1)), t 分布 (t(4), t(10), t(30)), ワイブル分布 (W(0.5,1), W(2,1), W(2.6,1), W(4,1)), 対数正規分布 (LN(0,0.5), LN(0,1)), ベータ分布 (B(1,1), B(10,10)) の計 14 種類を扱った.

## 5 結果

3 章で説明した各統計量は, 平滑化パラメータ  $\lambda$  や非正規分布の違いによって結果に特徴が現れた. ここでは,  $\lambda = 0.2$  の場合の結果を示す.

数値結果の「NA」は観測不能を表す. また, 図は横軸を各非正規分布 (左から G(4,1), t(10), W(2.6,1), LN(0,1), B(10,10)), 縦軸を ARL 値, 各線を統計量とし, ARL=370.4 は横の実線で, 「NA」は ARL=0 で表現した.

表 1 単一データに対応する EWMA 管理図

	WR	SR	HO	DP1	DP2
G(4,1)	113.5	173.3	280.4	320.3	121.7
G(2,1)	83.7	131.2	240.8	295.7	90.7
G(0.5,1)	52.2	81.1	179.8	291.0	59.0
t(4)	86.9	130.2	239.7	383.7	96.1
t(10)	152.1	212.2	301.1	353.0	161.8
t(30)	256.5	303.4	344.8	361.6	263.8
W(0.5,1)	57.5	81.1	12.7	463.1	31.4
W(2,1)	208.7	286.1	349.2	348.9	16.0
W(2.6,1)	433.6	437.0	NA	376.5	14.1
W(4,1)	823.7	502.5	NA	381.3	13.4
LN(0,0.5)	63.5	95.7	192.9	315.7	70.7
LN(0,1)	63.2	92.5	186.9	419.6	72.3
B(1,1)	NA	NA	1331.5	689.0	NA
B(10,10)	801.8	504.7	404.4	381.8	731.2

表 2 複数データに対応する EWMA 管理図

	SD	MAD	S	FQAD	TQAD
G(4,1)	217.9	314.0	300.8	324.7	265.8
G(2,1)	184.5	267.9	250.1	288.9	215.3
G(0.5,1)	139.3	160.1	148.5	188.9	148.4
t(4)	151.0	252.3	231.6	295.1	218.6
t(10)	241.5	315.2	301.2	340.8	307.5
t(30)	325.1	351.0	344.9	358.9	352.3
W(0.5,1)	106.1	126.3	117.7	153.7	120.4
W(2,1)	316.7	373.1	366.9	364.9	324.2
W(2.6,1)	397.3	403.2	408.3	385.5	376.9
W(4,1)	442.2	401.5	409.1	380.9	391.0
LN(0,0.5)	124.3	199.6	185.0	230.1	158.9
LN(0,1)	116.6	158.2	146.7	188.9	135.4
B(1,1)	1339.0	714.9	775.8	495.7	392.0
B(10,10)	438.8	403.1	409.4	384.5	391.9

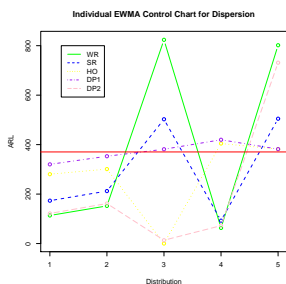


図 1 単一データに対応する EWMA 管理図

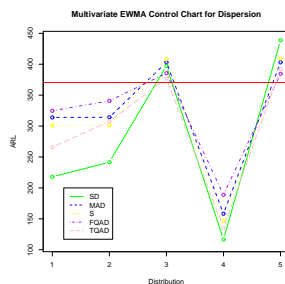


図 2 複数データに対応する EWMA 管理図

表 1 から、単一データに対応する EWMA 管理図は DP1 統計量以外の統計量で、「NA」が観測された。その中で、DP1 統計量は他の統計量よりも全体的に ARL= 370.4 に近い値を得ている。一方、先行研究で推奨された HO 統計量は「NA」を観測するなど、DP1 統計量と比較した際に良いとは言えない結果となった。また、図 1 から上記の特徴が視覚的に見て取れる。

表 2 から、複数データに対応する EWMA 管理図は、提案した全ての統計量は SD 統計量よりも良い結果となった。その中で、FQAD 統計量が最も良い結果となった分布が多い。また、単一データに対応する EWMA 管理図よりも全体を通じて ARL=370.4 に近づいたことも特徴的である。図 2 から、FQAD 統計

量は他の統計量とかなり近い値が存在するが、最も良い結果を収めていることが分かる。

## 6 まとめ

本研究で得られた成果として、単一データに対応する分散に関する EWMA 管理図は、 $\lambda = 0.05$  の場合は SR 統計量、 $\lambda = 0.2, 0.3$  の場合は DP1 統計量の利用を推奨する。一方、複数データに対応する分散に関する EWMA 管理図は、如何なる  $\lambda$  の違いに関わらず FQAD 統計量の利用を推奨する。また、 $\lambda$  の利用に関しては ARL 値の精度の観点から  $\lambda = 0.2$  の利用を推奨する。

## 参考文献

- [1] Borror, C. M., Montgomery, D. C., Runger, G. C. (1999). Robustness of the EWMA control chart to non normality. *J. Qual. Technol.* 31:309-316.
- [2] Burr, J. T. (2005). *Elementary Statistical Quality Control 2nd Edition*. New York: Marcel Dekker.
- [3] Domangue, R., Patch, S. C. (1991). Some omnibus exponentially weighted moving average statistical process monitoring schemes. *Technometrics* 33:299-313.
- [4] Hampel, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Stat. Association*. 69:383-393.
- [5] Hawkins, D. M., Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*. New York: Springer-Verlag.
- [6] Maraverakis, P. E., Panaretos, J., Psrakis, S. (2005). An Examination Robustness to Non Normality of the EWMA Control Charts for the Dispersion. *Commun. Statist. Simul. Computat.* 34:1069-1079.
- [7] Page, E. S. (1954). Continuous inspection schemes. *Biometrika*. 41:100-115.
- [8] Roberts, S. W. (1959). Control chart tests based on geometric moving averages. *Technometrics*. 1:239-250.
- [9] Rousseeuw, P. J., Croux, C. (1993). Alternatives to the median absolute deviation. *J. Amer. Stat. Association*. 88:1273-1283.
- [10] Stoumbos, Z. G., Reynolds, M. R. Jr. (2000). Robustness to non normality and autocorrelation of individual control charts. *J. Stat. Compu. Simu.* 66:145-187.
- [11] Stoumbos, Z. G., Sullivan, J. H. (2002). Robustness to non normality of the multivariate EWMA control chart. *J. Qual. Technol.* 34:260-276.
- [12] Wortham, A. W., Ringer, L. J. (1971). Control via exponential smoothing. *Transportation and Logistic Rev.* 7:33-39.