Open Access Library Journal

Effect of Sample Size on the Control Limits of Exponentially Weighted Moving Average Distance Square Scheme

Athambawa Mohamed Razmy

Department of Mathematics and Statistics, Sultan Qaboos University, Muscat, Oman
Email: mrazmy@squ.edu.om, amrazmy@seu.ac.lk

Received 22 May 2016; accepted 17 June 2016; published 21 June 2016

Abstract

In the series of quality monitoring schemes with exponentially weighted moving average, the exponentially weighted moving average distance square scheme was introduced for joint monitoring of process mean and variance. This scheme claims that it has a special feature that the control limits of the scheme are independent of sample size and therefore it gives more freedom to the users. However, this claim was not studied in detail. In this study, the control limits were found for this scheme through simulations, for different sample sizes with different combination of other scheme parameters. This study concludes that the control limits for designing this scheme are independent of sample size.

Keywords

Average Run Length, Control Limit, Exponential Weighted Moving Average, Joint Monitoring

Subject Areas: **Mathematical Statistics**

1. Introduction

The exponentially weighted moving average (EWMA) chart was introduced for monitoring the sample mean of a quality parameter by Robrts in 1959 [1]. This chart is more sensitive in detecting any small shift in mean compared to the conventional Shewhart control chart which was introduced by Shewhart in 1939 [2]. In EWMA chart, EWMA Variable A_t is plotted against the sample number $t(t = 1, 2, \cdots)$, where

$$A_t = (1-\lambda_m) A_{t-1} + \lambda_m \bar{X}_t. \quad (1)$$

Normally A_0, is considered as the target mean μ_0, λ_m is a constant such that $0 < \lambda_m < 1$ and it is selected based on the shift in the mean to be detected quickly for any process for a given in-control average run length (ARL).

How to cite this paper: Razmy, A.M. (2016) Effect of Sample Size on the Control Limits of Exponentially Weighted Moving Average Distance Square Scheme. *Open Access Library Journal*, **3**: e2663. http://dx.doi.org/10.4236/oalib.1102663

$\overline{X}_t$ is the t^{th} sample mean of the quality parameter, to be monitored. This chart issues an out-of-control signal if A_t is greater than the upper control limit (UCL) or lower than the lower control limit (LCL). For designing this chart, λ_m values and control limits for detecting different shifts in mean under different sample sizes and in-control ARLs, can be found in Crowder (1989) [3].

Chang and Gan introduced an EWMA chart for monitoring sample variance of a quality parameter in 1993 [4]. In their chart, the EWMA variable B_t is plotted against the sample number t ($t = 1, 2, \cdots$), where

$$B_t = (1-\lambda_v) B_{t-1} + \lambda_v \log\left(S_t^2\right). \tag{2}$$

Normally, the value of B_t is taken as $E\left[\log\left(S_t^2\right)\right]$ and S_t^2 is the sample variance of the quality parameter interested. λ_v is a positive constant which has the possible values of $0 < \lambda_v < 1$ and it is based on the shift in the variance to be detected quickly for a given in-control ARL. Like in the EWMA chart for monitoring sample mean, this chart also issues an out-of-control signal if B_t is greater than UCL or the lower than the LCL.

The above discussed two EWMA charts are used for monitoring the process mean and variance independently. In 1997, it was understood that monitoring the sample mean and variance was a bivariate problem and these two had to be monitored jointly [5]. A joint monitoring scheme for monitoring mean and variance simultaneously using EWMA technique was introduced by Gan [6]. Another two joint monitoring schemes, called max EWMA scheme and EWMA semicircle scheme for joint monitoring of process mean and variance were proposed by Chen *et al.* [7] [8]. The scheme parameters to design these joint monitoring schemes differ based on the sample sizes. This issue restricts the users in selecting a convenient sample size.

Subsequently, a new joint monitoring scheme named EWMA distance square scheme (EWMAD2) was introduced with a claim that its control limits were independent of sample size [9]. This scheme uses the standardized sample mean U_t and variance V_t such that

$$U_t = \frac{\overline{X}_t - \mu_0}{\sigma_0 / \sqrt{n}} \tag{3}$$

and

$$V_t = \Phi^{-1}\left[H\left(\frac{(n-1)S_t^2}{\sigma_0^2} : n-1\right)\right]. \tag{4}$$

$$H\left[\frac{(n-1)S_t^2}{\sigma_0^2} : n-1\right] = H(w;v) = P(W \le w) \text{ for } W \sim \chi_v^2 \tag{5}$$

the chi-square distribution with v degrees of freedom and $\phi(.)$ is the cumulative distribution function of a standard normal random variable [10]. S_t is the sample standard deviation, σ_0 is the population standard deviation and n is the sample size.

A statistics D_t^2 is defined by

$$D_t^2 = U_t^2 + V_t^2 \tag{6}$$

The EWMAD2 scheme is obtained by plotting the EWMAD2 statistic C_t against the sample number t where

$$C_t = (1-\lambda_d) C_{t-1} + \lambda_d D_t^2. \tag{7}$$

$C_0 = E\left(D_t^2\right) = 0$ and λ_d is a constant selected based on the shift in D_t^2, to be detected quickly. The optimum λ_d values for detecting various shifts in D_t^2, for the selected in-control ARLs can be read in Razmy (2005) [9]. This research paper examines the effect of sample size on the control limits of the EWMAD2 scheme for different ARLs.

2. Methodology

For this study, commonly used in-control ARLS100, 250, 300, 370, 500 and 1000 were selected. The sample sizes (n) studied were 5, 10, 50, 100 and 150. Samples were simulated in SAS using proc RANNO with sample sizes n. For each sample, the statistics D_t^2 was calculated from the variables u_t and v_t. The statics C_t was obtained for different values of λ_d ($\lambda_d = 0.05, 0.1, 0.2, \cdots, 0.9, 1.0$).

Initially arbitrary control limits (*CL*) were assumed for each combination of λ_d and in-control ARLs. Then, if $C_1 < CL$, then the second sample was simulated. This procedure was continued till $C_t > CL$ where C_t is an out-of-control point. Each time, the number of samples generated till to find an out-of-control point was recorded and it is the run length. In the same way 100,000 runs were performed and the ARLs were found. Subsequently, the assumed *CL* values were adjusted till to obtain the required in-control ARLS of 100, 250, 300, 370, 500 and 1000 for different λ_d values. By this procedure, the *CL*s for different combinations of in-control ARLs, sample sizes and λ_d values were found. In all cases, simulations were run until the standard error of the ARL was less than 1%of the pre-specified ARL.A sample program with sample size 5, is given in the Appendix.

3. Results and Discussion

The obtained *CL*s were plotted against the lambda value for the selected in-control ARLS and sample sizes. From the **Figures 1-5**, it could be observed that the control limits was increased with λ_d. It is obvious that the *CL* increases with in-control ARLs for a given λ_d value. The main observation is, for given λ_d and in-control ARL, the *CL* limit is constant with any sample size. This proves the claim that in the EWMAD2 scheme, *CLs* are independent of sample size and this property eases the procedure of designing this scheme with unrestricted sample sizes.

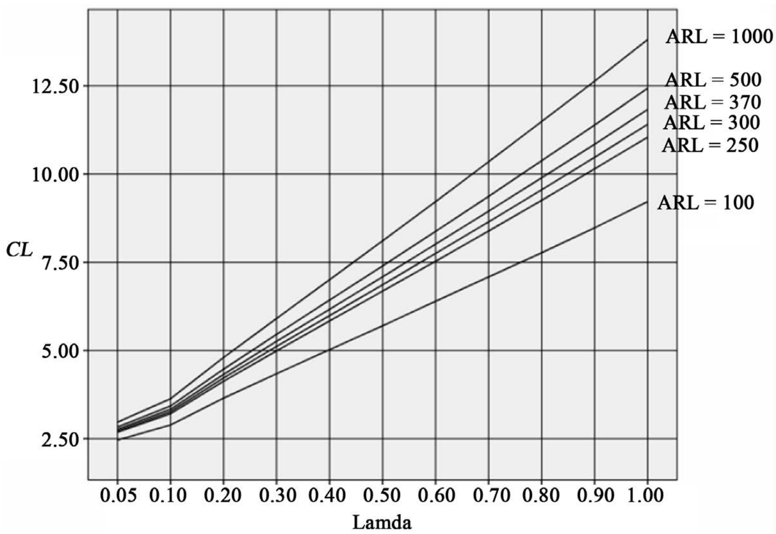

Figure 1. Control limits for selected in-control ARLS, $n = 5$.

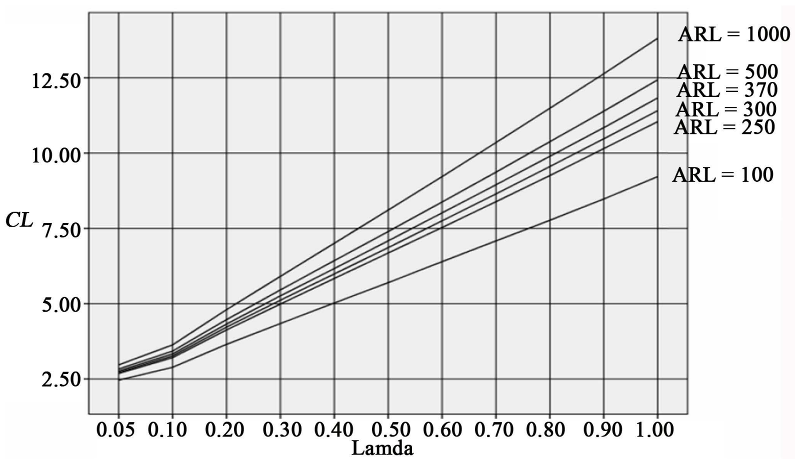

Figure 2. Control limits for selected in-control ARLS, $n = 10$.

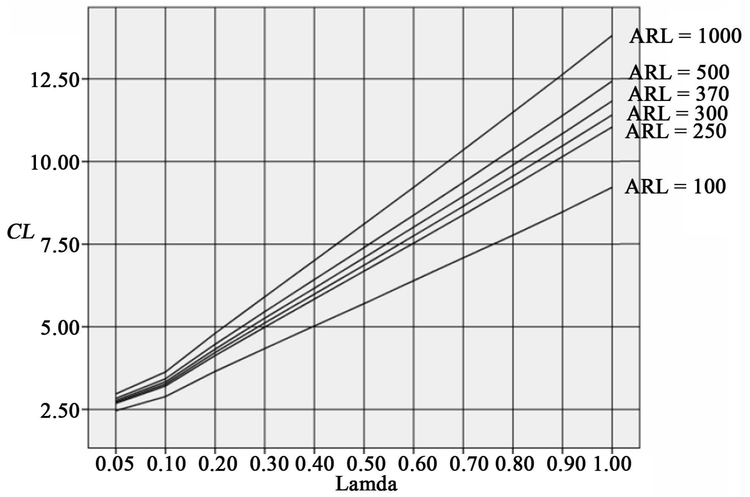

Figure 3. Control limits for selected in-control ARLS, *n* = 50.

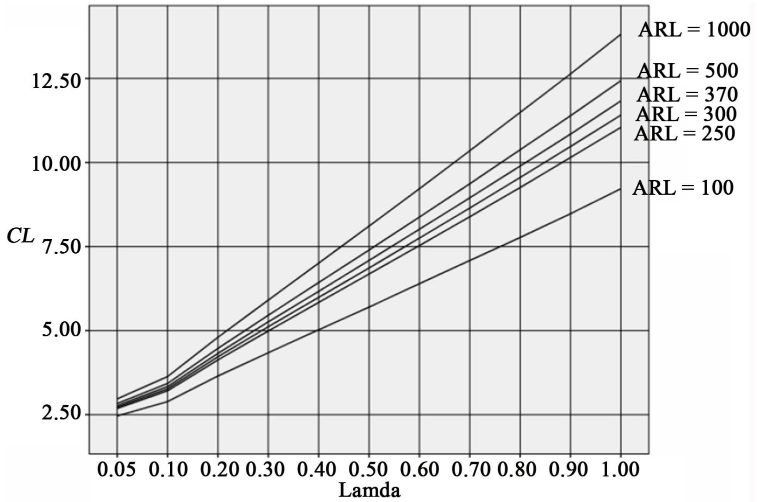

Figure 4. Control limits for selected in-control ARLS, *n* = 100.

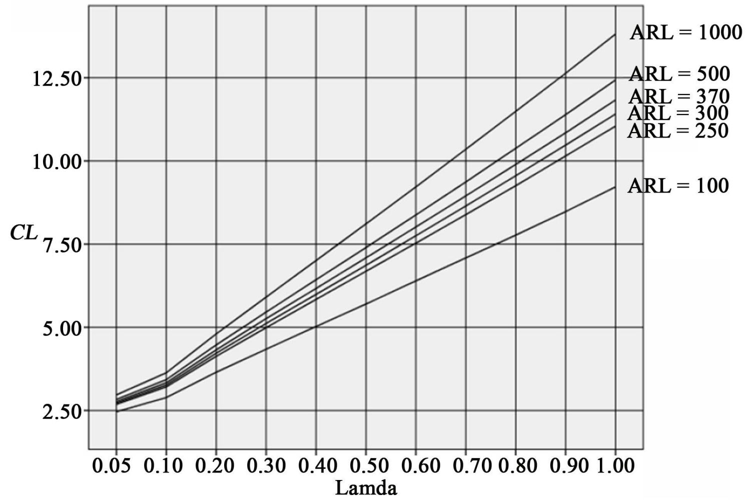

Figure 5. Control limits for selected in-control ARLS, *n* = 150.

Acknowledgements

The outputs for numerous programs were recorded and arranged by Ms. Iresha Dilhani.

References

[1] Roberts, S.W. (1959) Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, **1**, 239-250. http://dx.doi.org/10.1080/00401706.1959.10489860

[2] Shewhart, W. A. (1939) Statistical Methods from the Viewpoint of Quality Control. Graduate School, Department of Agriculture, Washington DC, 75.

[3] Crowder, S.V. (1989) Design of Exponentially Weighted Moving Average Schemes. *Journal of Quality Technology*, **21**, 155-162.

[4] Chang, T.C. and Gan, F.F. (1993) Optimal Designs of One-Sided EWMA Charts for Monitoring a Process Variance. *Journal of Statistical Computing* & *Simulations*, **49**, 33-48. http://dx.doi.org/10.1080/00949659408811559

[5] Gan, F.F. (1997) Joint Monitoring of Process Mean and Variance. *Nonlinear Analysis*, *Theory*, *Methods and Applications*, **30**, 4017-4024. http://dx.doi.org/10.1016/S0362-546X(97)00224-1

[6] Gan, F.F. (1995) Joint Monitoring of Process Mean and Variance Using Exponentially Weighted Moving Average Control Charts. *Technometrics*, **37**, 446-453. http://dx.doi.org/10.1080/00401706.1995.10484377

[7] Chen, G., Cheng, S.W. and Xie, H.W. (2001) Monitoring Process Mean and Variability with One EWMA Chart. *Journal of Quality Technology*, **33**, 223-233.

[8] Chen, G., Cheng, S.W. and Xie, H.W. (2004) A New EWMA Control Chart for Monitoring Both Location and Dispersion. *Quality Technology* &*Quantitative Management*, **1**, 217-231. http://dx.doi.org/10.1080/16843703.2004.11673074

[9] Razmy, A.M. (2005) Joint Monitoring of Process Mean and Variance. MSc Thesis, Department of Statistics and Applied Probability, National University of Singapore, Singapore.

[10] Quesenberry, C.P. (1995) On Properties of Q Charts Variables. *Journal of Quality Technology*, **21**, 242-250.

Appendix

```
data raw;
mu0 = 0;
sigma = 1.0;
sigmasq = sigma*sigma;
nnn = 5;
nnnsqrt = sqrt(nnn);
nnn1 = nnn -1;
totalrun = 100,000;
maxsamp = 1,000,000;
seed = 86,924,569;
lam = 0.7;
onelm= 1.0 − lam;
bigh = 7.0856;
dobigd = 0;
dosmalld = 1;
actuals = smalld * sigma;
actualm = mu0 + bigd * sigma/nnnsqrt;
message = 'OK';
dorunnum = 1tototalrun;
qqq = 2;
runlen = 0;
do sample = 1tomaxsamp;
x1 = actualm + actuals * rannor (seed);
x2 = actualm + actuals * rannor (seed);
x3 = actualm + actuals * rannor (seed);
x4 = actualm + actuals * rannor (seed);
x5 = actualm + actuals * rannor (seed);
xsum = x1 + x2 + x3 + x4 + x5;
x2sum = x1*x1 + x2*x2 + x3*x3 + x4*x4 + x5*x5;
xbar = xsum/nnn;
s = x2sum − xsum*xsum/nnn;
s = s/nnn1;
uu = (xbar − mu0)/sigma*nnnsqrt;
uuu = uu*uu;
h = nnn1*s/sigmasq;
hhh= CDF('CHISQUARE', h, nnn1);
ifhhh > 0.0000000000001 and hhh < 0.999999999999999then
vv = probit(hhh); else
vv = 8.21;
vvv = vv*vv;
rrr = uuu + vvv;
runlen = runlen + 1;
qqq = onelm * qqq + lam * rrr;
ifqqq > bighthendo;
output;
goto DONE1;
end;
keep lam runlenbigdsmalldbighxbarqqqrrrvvvuuu message ;
end;
Message = 'Maximum # of sample Exceeded';
output;
```

```
DONE1:
end;
end;
end;
procmeans data = raw;
by bigdsmalld;
run;
```