The Use of Sequential Sampling in Process Monitoring

Marion R. Reynolds, Jr.
Departments of Statistics and Forestry
Virginia Tech, USA

Quality and Productivity Research Conference

June, 2009

1. The monitoring problem

- Consider the problem of detecting changes in the mean μ and/or the variance σ^2 of a normal variable X.

- Let μ_0 and σ_0^2 be the in-control values (assumed, for simplicity, to be known).

- Assume that the objective is to detect small and large changes in μ from μ_0, or small or large increases in σ^2 above σ_0^2.

- Samples of n observations are taken from the process at regular intervals.

- Let $X_{k1}, X_{k2}, \ldots, X_{kn}$ be the n observations at sampling point k, and let $\bar{X}_k$ the sample mean and let S_k^2 be the sample variance.

- Traditional Shewhart control charts for monitoring μ and σ^2 can be based on plotting $\bar{X}_k$ and S_k^2.

2. The EWMA chart for μ

- Shewhart charts are not effective for detecting small shifts in μ or σ^2 unless n is large. More effective detection of small shifts can be obtained by using EWMA charts or CUSUM charts.

- Here we focus on EWMA charts.

- At sample k the EWMA statistic for monitoring μ is

$$E_k^{\bar{X}} = (1-\lambda)E_{k-1}^{\bar{X}} \quad + \quad \lambda \bar{X}_k$$

- λ is a weighting or tuning parameter satisfying $0 < \lambda \le 1$. The value of λ determines the weight give to the current data relative to past data.

- This chart signals if $E_{kj}^{\bar{X}} < \text{LCL}$ or $E_{kj}^{\bar{X}} > U\text{CL}$.

3. An EWMA control chart for σ^2

- An EWMA control chart for monitoring σ^2 can be based on the sample variances S_1^2, S_2^2, $S_3^2 \ldots$.

- At sample k consider the EWMA statistic

$$E_k^{S^2} = (1-\lambda)\max\{E_{k-1}^{S^2}, \sigma_0^2\} \quad + \quad \lambda S_k^2$$

- This EWMA statistic is designed to detect increases in σ^2, and has a reset back to σ_0^2 whenever the statistic drops below σ_0^2 (which is the in-control expected value of S_k^2).

- This EWMA control chart signals if $E_k^{S^2} > \mathrm{UCL}$.

4. Another EWMA chart for σ^2.

- Instead of using sample variances to monitor σ^2, better performance can be achieved by using

$$D_k^2 = \frac{1}{n} \sum_{j=1}^{n} (X_{kj} - \mu_0)^2 ,$$

the sample **squared deviations from target**.

- At sample k consider the EWMA statistic

$$E_k^{D^2} = (1-\lambda)\max\{E_{k-1}^{D^2}, \sigma_0^2\} \quad + \quad \lambda D_k^2$$

- This EWMA statistic also has a reset back to σ_0^2 whenever the statistic drops below σ_0^2, and signals if $E_k^{D^2} > \text{UCL}$.

- This EWMA chart can be used in combination with the EWMA chart for μ effective detection of changes in μ and increases in σ^2.

- For chart combinations using squared deviations see Stoumbos and Reynolds (2000, 2004, 2005), Reynolds and Stoumbos (2001a, 2001b, 2004a, 2004b, 2005, 2006), and Stoumbos, Reynolds, and Woodall (2003).

5. Comparisons of chart combinations

- The average run length of (ARL) of EWMA charts, counted from the time of the shift, depends on the value of the EWMA statistic at the time that the shift occurs. Thus, we use the **steady state ARL** as a measure of the time required to detect a parameter shift.

- Measure the shift size in terms of

$$\delta = (\mu - \mu_0) / \sigma_0$$

$$\psi = \sigma / \sigma_0 .$$

- For purposes of comparison with other control charts, the control limits of these two charts will be adjusted so that they have the same in-control ARL, and the in-control ARL of the charts used together in combination is 200.

6. Table 1. Shewhart vs. EWMA, n = 4, $\lambda = .4$

δ and ψ	Shew $\bar{X}$ & S^2	EWMA $\bar{X}$ & S^2	EWMA $\bar{X}$ & D^2
in control	200.0	200.0	200.0
δ = 0.2	139.7	75.6	72.6
0.4	63.9	20.6	19.9
0.6	27.3	8.5	8.2
0.8	12.5	4.8	4.7
1.0	6.4	3.3	3.2
1.4	2.4	2.1	2.0
2.0	1.2	1.4	1.3
3.0	1.0	1.0	1.0
ψ = 1.2	32.7	27.6	25.0
1.4	10.9	9.1	8.0
1.6	5.4	4.8	4.2
1.8	3.4	3.2	2.9
2.0	2.5	2.4	2.2
3.0	1.3	1.3	1.3

7. Sequential sampling

- In the traditional control charts discussed above, the sampling rate is fixed. The charts in the table above are based on always taking a sample of n = 4 observations at each sampling point.

- Improved performance can be achieved by using **sequential sampling** at each sampling point. With sequential sampling observations are taken sequentially at each sampling point

- At each sampling point sample until we have either good evidence that the process is in control, or enough evidence to signal that there has been a change in the process.

- Thus, the actual sample size at each sampling point may be smaller or larger than 4, depending on the data obtained.

- Sequential sampling would be particularly appropriate for situations in which items in the sample must be tested one by one, the results of one test are available before the next item is tested, and testing is expensive and/or destructive.

8. Applying an SPRT

- One of the simplest cases in which sequential sampling can be used is when we have a specific shift in μ and σ^2 that should be detected.

- Suppose we want to detect a shift in (μ, σ^2) from (μ_0, σ_0^2) to (μ_1, σ_1^2).

- At each sampling point apply a sequential probability ratio test (SPRT) for testing $H_0 : \mu = \mu_0, \sigma^2 = \sigma_0^2$ against $H_1 : \mu = \mu_1, \sigma^2 = \sigma_1^2$.

- If the SPRT accepts H_0 then conclude that there is no evidence that the process has changed, and wait until the next sampling point to sample again and apply another SPRT.

- If the SPRT rejects H_0 then signal that the process has changed.

- The SPRT could be designed so that the expected sample size is 4 observations when the process is in control. This would give the same average sampling rate as the standard sampling procedure that always takes a sample of n = 4 observations.

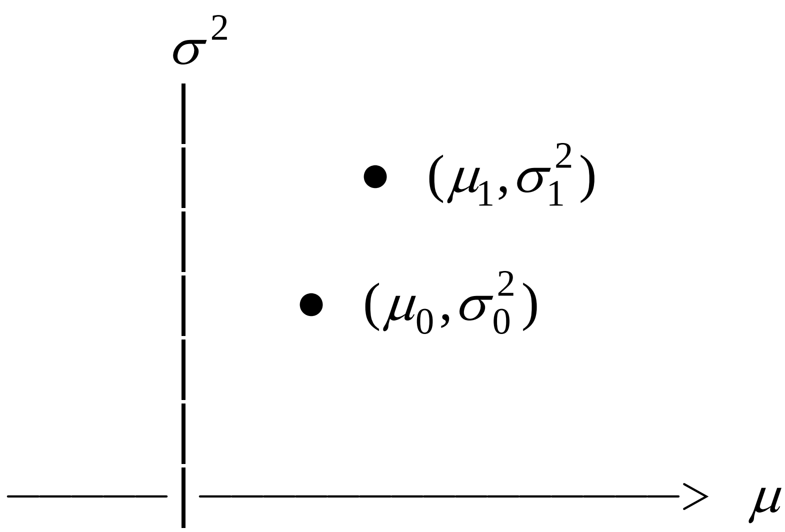
9. Groups of observations

- An SPRT would normally be based on taking observations individually, but groups of $m > 1$ observations could be used if this would be more convenient.

- If the expected sample size must be 4 at each sampling point, then we must have $m < 4$.

10. Example of applying an SPRT

- Suppose that we want to detect a change in the process that increases μ from μ_0 to $\mu_0 + 0.5\sigma_0$, and at the same time increases σ^2 from σ_0^2 to $1.5^2\sigma_0^2$.



- An SPRT can be designed so that, when the process is in control, the expected sample size is 4.0 and the probability of rejecting H_0 is .005.

- When the in-control probability of rejecting is .005, the in-control ARL will be 1/.005 = 200, the same as used for the standard control charts.

- The next table gives the ARL when there is a shift in (μ, σ^2) in the direction (μ_1, σ_1^2).

11. Table 2. Shifts in the (μ_1, σ_1^2) direction

δ and ψ	Shew	EWMA	SPRT chart	
	$\bar{X}$ & S^2	$\bar{X}$ & D^2	$m = 1$	$m = 2$
in control	200.0	200.0	200.0	200.0
(0.05,1.05)	113.6	100.5	68.1	68.7
(0.10,1.10)	67.8	52.5	27.9	28.1
(0.15,1.15)	42.7	30.1	13.5	13.6
(0.20,1.20)	28.3	18.9	7.6	7.7
(0.25,1.25)	19.7	12.9	5.0	5.0
(0.30,1.30)	14.3	9.4	3.6	3.7
(0.40,1.40)	8.4	5.8	2.4	2.4
(0.50,1.50)	5.6	4.1	1.9	1.9
(0.60,1.60)	4.0	3.2	1.6	1.6
(0.70,1.70)	3.1	2.6	1.5	1.5
(0.80,1.80)	2.6	2.2	1.4	1.4
(1.00,2.00)	1.9	1.8	1.2	1.3
(2.00,3.00)	1.1	1.1	1.1	1.1

12. References to sequential sampling

- Sequential sampling has been widely used in quality control in the area of acceptance sampling, but not in the area of process monitoring.

- Most previous work on sequential sampling is based on using double sampling (i.e., take either one or two group of observations at each sampling point). See Croasdale (1974), Daudin (1992), Irianto and Shinozaki (1998), Carot, Jabaloyes, and Carot (2002), He, Grigoryan, and Sigh (2002), and He (2002).

- Stoumbos (1993),Stoumbos and Reynolds (1996, 1997a, 1997b, 2001) and Reynolds and Stoumbos (1998, 2001, 2007) investigated the use of the more efficient approach of applying an SPRT at each sampling point.

- Reynolds and Kim (2005, 2007) considered the use of sequential sampling in the multivariate setting.

13. Extensions of sequential sampling

- As shown in the previous table, the SPRT chart works extremely well when the shift is in (μ, σ^2) is in the direction for which the SPRT was designed.

- The approach of applying an SPRT at each sampling point also works extremely well when the objective is detecting one-sided shifts in a single parameter (for example detecting increases in the proportion of nonconforming items).

- However, when monitoring (μ, σ^2) it will usually not be possible to specify that exact direction of a parameter shift that might occur.

- Thus when monitoring (μ, σ^2) we need a sequential sampling method that will work for shifts in any direction.

- A straightforward approach is to use the EWMA statistics to determine the amount of sampling to be done at each sampling point.

14. Sequential sampling using EWMA statistics

- To apply sequential sampling using EWMA statistics, continue to sample as long as an EWMA statistic is close to a control limit.

- Stop sampling at the current sampling point and go to the next sampling point when the EWMA statistics are not close to a control limit.

- As in the case of the SPRT chart, assume that at each sampling point observations are taken in groups of $m \geq 1$ observations.

- When using the combination of two EWMA charts, continue sampling if either chart indicates that sampling should continue.

15. EWMA statistics for sequential sampling

- For the j^{th} group at sampling point k let $\bar{X}_{kj}$ be the sample mean.

- Let $N(k)$ be the total number of groups sampled at sampling point k.

- After the first group at sampling point k the EWMA statistic for monitoring μ is

$$E_{k,1}^{\bar{X}} = (1-\lambda)E_{k-1,N(k-1)}^{\bar{X}} + \lambda\bar{X}_{k1}$$

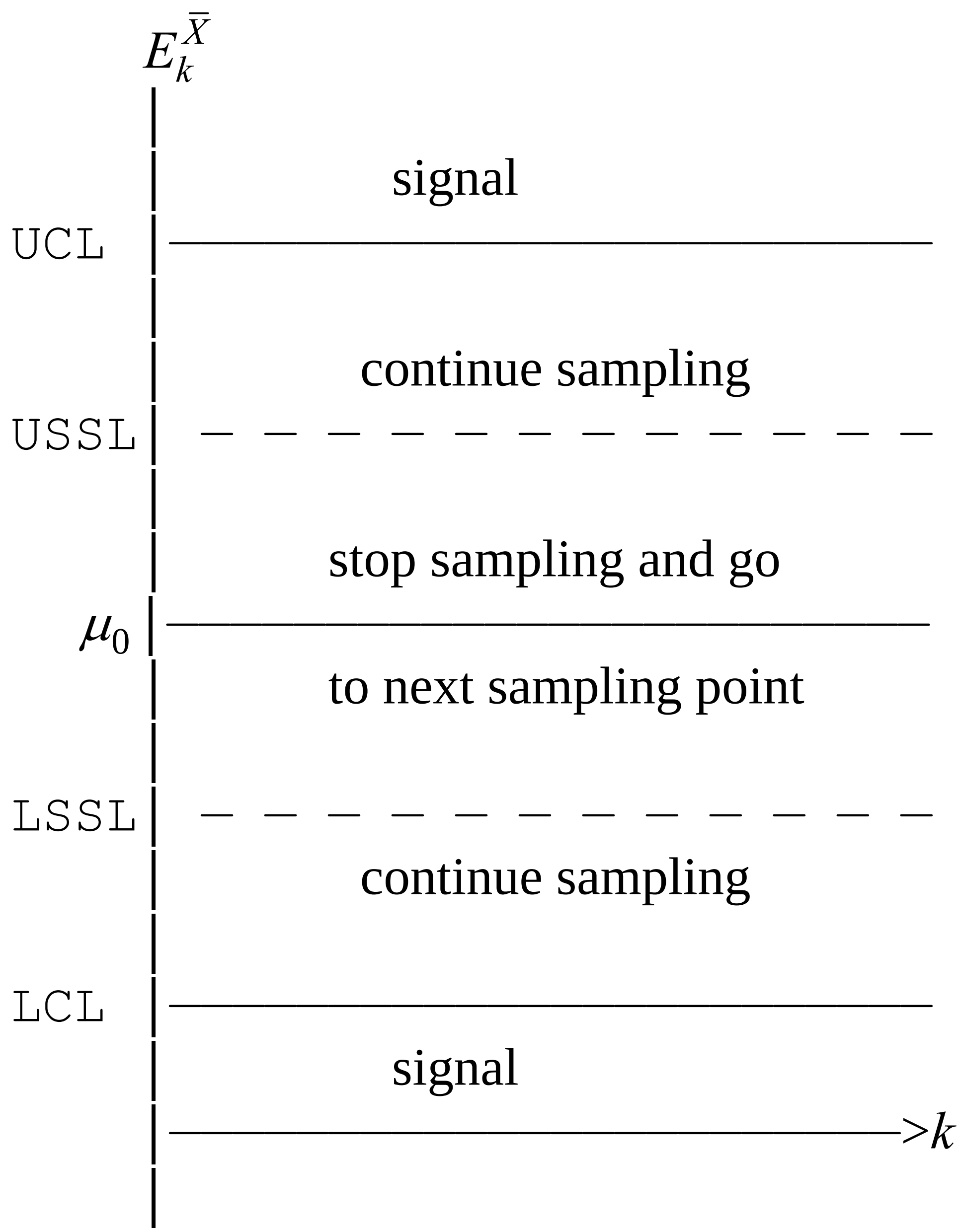
  and after the j^{th} group, $j = 2,3,\ldots$, the EWMA statistic is

$$E_{k,j}^{\bar{X}} = (1-\lambda)E_{k,j-1}^{\bar{X}} + \lambda\bar{X}_{kj}.$$

- We have a similar definition for $E_{kj}^{D^2}$, the EWMA statistic for monitoring σ^2.

16. Define sequential sampling limits for μ

- For the EWMA chart based on $E_{kj}^{\bar{X}}$ define a **lower sequential sampling limit** LSSL and an **upper sequential sampling limit** USSL. At each k we have $E_{kj}^{\bar{X}}$ plotted for $j = 1, 2, \ldots$



17. Table 3. Sequential sampling with the EWMA

	Shew	EWMA	EWMA – Seq Samp	
δ and ψ	$n = 4$	$n = 4$	$m = 1$	$m = 2$
in control	200.0	200.0	200.0	200.0
$\delta = 0.2$	139.7	72.6	58.3	60.7
0.4	63.9	19.9	10.6	11.5
0.6	27.3	8.2	4.0	3.9
0.8	12.5	4.7	2.6	2.3
1.0	6.4	3.2	2.0	1.8
1.4	2.4	2.0	1.5	1.3
2.0	1.2	1.3	1.2	1.1
3.0	1.0	1.0	1.0	1.0
$\psi = 1.2$	32.7	25.0	14.0	15.7
1.4	10.9	8.0	4.0	4.1
1.6	5.4	4.2	2.5	2.3
1.8	3.4	2.9	2.0	1.8
2.0	2.5	2.2	1.8	1.6
3.0	1.3	1.3	1.4	1.2

18. Multivariate monitoring

- Consider the situation in which p variables are measured, so that an observation consists of a vector

$$\mathbf{X} = (X_1, X_2, \ldots, X_p)'$$

- Assume that $\mathbf{X}$ has a **multivariate normal** distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)'$ and variance-covariance matrix $\boldsymbol{\Sigma}$. Let $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)'$ be the vector of variances of the variables (the diagonal elements of $\boldsymbol{\Sigma}$).

- Let $\boldsymbol{\mu}_0$, $\boldsymbol{\sigma}_0^2$, and $\boldsymbol{\Sigma}_0$ be the in-control values of $\boldsymbol{\mu}$, $\boldsymbol{\sigma}^2$, and $\boldsymbol{\Sigma}$, respectively.

- Let

$$\bar{\mathbf{X}}_k = (\bar{X}_{k1}, \bar{X}_{k2}, \ldots, \bar{X}_{kp})$$

  be the vector of sample means for the p variables at sampling point k.

- The first multivariate control for monitoring $\boldsymbol{\mu}$ (developed by Hotelling (1947)) is a Shewhart-type chart based on the statistic

$$(\bar{\mathbf{X}}_k - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\bar{\mathbf{X}}_k - \boldsymbol{\mu}_0)$$

19. A multivariate EWMA chart

- Another multivariate control chart, called the **multivariate** EWMA (MEWMA) chart, is similar to Hotelling's chart except that the vector $\bar{\mathbf{X}}_k$ of sample means is replaced in the quadratic form with a vector
$$\mathbf{E}_k^{\bar{X}} = (E_{k1}^{\bar{X}}, E_{k2}^{\bar{X}}, \ldots, E_{kp}^{\bar{X}})$$
of EWMA statistics for the p variables.

- This gives the MEWMA statistic
$$M_k^{\bar{X}} = (\mathbf{E}_k^{\bar{X}} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{E}_k^{\bar{X}} - \boldsymbol{\mu}_0)$$

- This chart signals if $M_k^{\bar{X}} > \text{UCL}$.

- This chart was developed by Lowry, Woodall, Champ, and Rigdon (1992).

- The MEWMA chart has better performance than Hotelling's chart except for large shifts in $\boldsymbol{\mu}$

20. Multivariate EWMA charts for $\boldsymbol{\sigma}^2$

- An MEWMA-type chart for monitoring $\boldsymbol{\sigma}^2$ can be developed by using EWMA statistics based on S^2 or D^2 in the quadratic form in place of the EWMA statistics based on $\bar{X}$

- Let

$$\mathbf{E}_k^{D^2} = (E_{k1}^{D^2}, E_{k2}^{D^2}, \ldots, E_{kp}^{D^2})$$

  be the vector of EWMA statistics of squared deviations from target for the p variables.

- An MEWMA-type statistic could be based on

$$(\mathbf{E}_k^{D^2} - \boldsymbol{\sigma}_0^2)'(\boldsymbol{\Sigma}_0^{(2)})^{-1}(\mathbf{E}_k^{D^2} - \boldsymbol{\sigma}_0^2)$$

  where $\Sigma_0^{(2)}$ represents the matrix whose elements are the squares of the elements of Σ_0.

- Instead, we found that better performance is usually obtained by using

$$M_k^{D^2} = (\mathbf{E}_k^{D^2})'(\boldsymbol{\Sigma}_0^{(2)})^{-1}(\mathbf{E}_k^{D^2})$$

- This chart signals if $M_k^{D^2} > \text{UCL}$.

21. A combination of MEWMA charts

- The MEWMA chart for $\boldsymbol{\mu}$ can be used in combination with the MEWMA-type chart for $\boldsymbol{\sigma}^2$ to detect process changes in $\boldsymbol{\mu}$ or increases in $\boldsymbol{\sigma}^2$.

- When the variables are not independent we used regression adjustment of the variables (Hawkins (1991, 1993)) because it improves the average performance of the MEWMA chart for $\boldsymbol{\sigma}^2$.

- Combinations of MEWMA charts for monitoring $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$ have been investigated by Reynolds and Cho (2006), Reynolds and Kim (2007), and Reynolds and Stoumbos (2008, 2009).

- Sequential sampling can be applied using the combination of two MEWMA charts, by taking observations sequentially and adding a USSL to each of the two MEWMA charts.

- When using the combination of two MEWMA charts, continue sampling if either chart indicates that sampling should continue.

22. Measuring the multivariate parameter shift.

- The SSARL of the control charts being considered **depends on the direction of the shift** in $\boldsymbol{\mu}$ or $\boldsymbol{\sigma}^2$ in addition to the size of the shift.

- We assume that there **is no particular shift direction of interest**, and thus consider a random direction for the shift in $\boldsymbol{\mu}$, and a random direction corresponding to an increase for $\boldsymbol{\sigma}^2$.

- Measure the size of the shift in $\boldsymbol{\mu}$ terms of the non-centrality parameter

$$\delta = \sqrt{(\boldsymbol{\mu} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)}$$

so that the in-control case corresponds to $\delta = 0$.

- For variable i let $\gamma_i = \sigma_i / \sigma_{0i}$ be the standardized shift in the standard deviation, and let

$$\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)'.$$

- Measure the size of the increase in $\boldsymbol{\sigma}^2$ in terms of

$$\psi = 1 + \sqrt{(\boldsymbol{\gamma} - \mathbf{1})'(\boldsymbol{\gamma} - \mathbf{1})}$$

so that in-control case corresponds to $\psi = 1$.

23. Table 4. $p = 4$, $n = 4$, $\rho = 0.0$, $\lambda = .4$

	MEWMA	MEWMA - Seq Samp	
δ and ψ	$n = 4$	$m = 1$	$m = 2$
in control	200.0	200.0	200.0
$\delta = 0.2$	116.4	99.9	104.6
0.4	38.6	22.8	25.9
0.6	14.4	6.7	7.3
0.8	7.2	3.6	3.5
1.0	4.5	2.7	2.4
1.4	2.6	2.0	1.6
2.0	1.6	1.5	1.2
3.0	1.1	1.1	1.0
$\psi = 1.2$	35.7	19.6	22.9
1.4	10.3	4.5	4.9
1.6	4.9	2.5	2.4
1.8	3.1	1.9	1.7
2.0	2.3	1.7	1.5
3.0	1.2	1.2	1.1

24. Table 5. $p = 4$, $n = 4$, $\rho = 0.9$, $\lambda = .4$, Reg. Adj.

	MEWMA	MEWMA - Seq Samp	
δ and ψ	$n = 4$	$m = 1$	$m = 2$
in control	200.0	200.0	200.0
$\delta = 0.2$	115.8	99.8	104.4
0.4	38.4	22.8	25.8
0.6	14.3	6.7	7.3
0.8	7.2	3.6	3.5
1.0	4.5	2.7	2.4
1.4	2.6	1.9	1.6
2.0	1.6	1.4	1.2
3.0	1.1	1.1	1.0
$\psi = 1.2$	31.5	16.8	19.7
1.4	7.9	3.5	3.7
1.6	3.6	2.1	1.9
1.8	2.3	1.7	1.5
2.0	1.7	1.5	1.3
3.0	1.1	1.2	1.1

25. Conclusions

- Sequential sampling allows the sample size at each sampling point to depend on the data obtained form the process.

- Sequential sampling can be used when it is feasible to take observations sequentially at each sampling point.

- The SPRT chart can be used when a specific shift direction is of interest.

- When there is no specific shift direction of interest, a combination of EWMA charts in the univariate setting or MEWMA charts in the multivariate setting gives very good overall performance.

- The EWMA and MEWMA charts for monitoring variability considered here are based on squared deviations from target.

- EWMA or MEWMA charts based on sequential sampling provide significant improvements in performance compared to standard charts based on fixed sample sizes.