DE GRUYTER OPEN

Scientific Annals
of the "Alexandru Ioan Cuza" University of Iaşi
Economic Sciences
61 (1), 2014, 55-66
DOI 10.2478/aicue-2014-0007

ON DETERMINING THE OPTIMAL SAMPLING FREQUENCY FOR FEEDBACK QUALITY CONTROL SYSTEMS

C. P. KARTHA
University of Michigan-Flint
Michigan, United States
cpk@umflint.edu

Abstract

This paper considers the choice of the most profitable sampling frequency for a feedback quality control system. It is assumed that the disturbance can be adequately modeled by a first-order Integrated Moving Average (IMA (0, 1, 1)) process. The cost model includes two terms, one for the cost of being off-target and one for the cost of sampling and adjustment. An analytical solution is obtained and the sensitivity of parameters examined.

Keywords: Integrated Moving Average Processes, Feedback Control Systems.
JEL classification: C44, C83, L15

1. INTRODUCTION

The quest for quality is probably more widespread and intense globally today than at anytime in history. Organizations worldwide have realized that the key to increased productivity and profitability is improving quality. They have been actively involved in developing and implementing strategies for improving quality of their goods and services by employing techniques that reduce product and process variability.

Statistical Process Control (SPC) and Automatic Process Control (APC) (which is also referred to as Engineering Process Control (EPC)) are two quality improvement strategies that focus on reduction of variability as their objective. SPC is used to continuously monitor a process to detect signals representing possible assignable causes that increase variability. Control charts are used to detect assignable causes of variation. APC, on the other hand, is used to reverse the effect of process disturbances by making regular adjustments to process variables that can be manipulated. The objective of the adjustments is to keep the output quality characteristic close to its target. Box and Jenkins (1976) developed optimum control strategies using Minimum Mean Square Error (MMSE) control by modeling the process dynamics and the disturbances at the output. Box and Kramer (1992) provided an excellent comparison of SPC and APC and discussed the differences between the two strategies. SPC performs a monitoring function that signals when control is needed by identifying and removing assignable causes of variation while APC, rather than removing the assignable causes, uses continuous adjustments to keep the process variable on target. MacGregor (1987) showed that integrating these two strategies can improve the overall quality performance. This was further emphasized by MacGregor and Harris (1990) Montgomery et al. (1994) and Vander Wiel (1996). A related problem that is relevant to both SPC and APC is the selection of the best sampling interval for monitoring the process behavior. This has been examined under various scenarios by Bessegato et al. (2011),Duncan (1956), Dasgupta and Mandal (2008), Hunter and Kartha (1984) and MacGregor (1976).

This study investigates the problem of selecting the optimum sampling frequency for a feedback quality control system in which the process disturbance can be represented by an IMA(0,1,1) model. An analytical solution for optimum sampling interval is obtained by minimizing the overall cost that includes both the cost of sampling and adjustment as well as the cost of being off the target. The method discussed here provides a cost-effective alternative to the trial-and-error procedure that is often used in practice.

2. PROBLEM

Consider a manufacturing process in which the product is manufactured in the form of discrete units (for a continuously produced product such as a liquid, one can imagine that the product stream is divided into appropriately sized parcels, and the theory below applies). Suppose these discrete units are sampled at regular intervals to check whether a particular specification, such as minimum weight, is being satisfied. By "sampling" we mean the act of making a measurement of a single quality characteristic on a sample taken from the process. Suppose there is a target value for this specification. At each time t the process is sampled and the deviation ε_t from this target value is observed. An input variable X_t (for example, the setting of a valve) is adjusted, the amount of adjustment at the time t being a function of ε_t . The purpose of the adjustment X_t is to try to maintain the process as close to its target value as possible. Such a system, used for process control, is often referred to as a closed loop feedback control system.

In general, sampling the process more frequently would permit tighter control, that is, the ε_t`s in some sense would be smaller. The most frequently one could sample a process would be to sample each item produced. But the higher the rate of sampling and adjustment, the higher the cost. Therefore, the most economic sampling frequency must be a compromise between the cost of being off-target and the cost of sampling and adjustment. The solution must depend on the cost parameters and the stochastic nature of the production process.

Consider a closed loop feedback control system in which the ε_t`s are observed at equal intervals of time. At time t, an unobserved disturbance N_t enters the system at the output. The disturbances can be counteracted by making adjustments X_t of a manipulated input variable. These adjustments are determined by the control equation f (ε_t, ε_{t-1} . . .). It is assumed that X_t and N_t are deviations from reference values which are such that if the conditions $X_t = 0$ and $N_t = 0$ are maintained, the output would remain on the target value, that is, $\varepsilon_t = 0$. A variety of situations that occur in practice can be approximated as follows:

$$\varepsilon_t = \{g\ (1-\delta) / (1-\delta B)\}\ X_{t-1} + N_t \tag{1}$$

where g = steady state gain of the dynamic system, δ = dynamic constant, B is the backshift operator such that $BX_t = X_{t-1}$, and

$$N_t = \{(1-\theta_1 B) / (1-B)\}\ a_t.\ 0<|\theta_1|< 1, \tag{2}$$

where θ_1 = the moving average parameter and $\{a_1\}$ = white noise series with mean zero and variance σ_1^2. Equations (1) and (2) represent a feedback control system with first order dynamics and no delay in which the disturbance follows an IMA (0,1,1) integrated moving average model (Box and Jenkins, 1976). For a system described by Equations (1) and (2), the adjustment to be made for minimum mean square error control (MMSE) is

$$X_t = \{ (1-\delta B) / g(1-\delta) \}\ (1-\theta_1)\ \varepsilon_t \tag{3}$$

With such adjustments, the observed deviations from target $\{\varepsilon_t\}$ will be a white noise sequence $\{a_1\}$ with mean zero and variance σ_1^2, a_t being the one-step ahead forecast error. The variance of adjustments is given by

$$\sigma^2_x =\{ (1-\theta_1^2) / g^2\}\ \{ (1+ \delta_1^2)\ \sigma_1^2 / (1- \delta_1)^2]\} \tag{4}$$

Often, in practice, the process is sampled to determine whether or not a particular quality characteristic is on target. Suppose that only every s^{th} value of ε_t is observed where s is a positive integer. Denote this sampled process by M_t. In this paper, we consider the selection of the optimal sampling frequency that minimizes the cost per unit time.

3. COST FUNCTION

If the N_t process defined by (2), which is IMA (0,1,1), is observed at intervals of s time units (where s, the sampling interval, is a positive integer), the resulting sampled process M_t is also IMA (0,1,1), that is,

$$M_t = \{ (1-\theta_s B) / (1-B) \}\ e_t \tag{5}$$

where $\{e_t\}$ is a white noise series with mean zero and variance σ_s^2 . The parameters of the two processes are related according to the following two equations:

$$(1-\theta_s)^2 / s\theta_s = (1-\theta_1)^2 /\theta_1 \tag{6}$$

$$\sigma_s^2\theta_s =\sigma_1^2\theta_1 \tag{7}$$

We assume the sampled process is invertible, which implies that $|\theta_1|< 1$. Note that the adjustment that ensures MMSE control for M_t process is

$$X_t (s) = -\{ (1-\theta_s) / g(1-\delta^{s)} \}\ (1-\delta^s B)\ \varepsilon_t (s). \tag{8}$$

Under MMSE control, the observed deviations from target { ε_t (s) } are uncorrelated random variables with mean zero and variance σ_s^2 .

In general, sampling the process more frequently would permit tighter control, that is, the ε_t`s in some sense would be smaller. The most frequently one could sample a process would be to sample each item produced. But the higher the rate of sampling and adjustment, the higher the cost. The overall costs C(s) consists of two parts, (a) that associated with production of items which are different from the target value and (b) that of sampling and adjustment. A smaller sampling interval decreases cost (a) while increasing cost (b). On the other hand, a larger sampling interval increases cost (a) and decreases cost (b). The control scheme ensures that, on the average, the process is on target, that is, the expected value of ε_t is zero. Therefore, variance of ε_t , the deviations from the target, may be used as a measure of cost (a). The cost per unit time, C(s), is defined as

$$C(s) = R\sigma_s^2 + 1/s \tag{9}$$

where $R = C_1 / C_2$, C_1 = cost of being off-target per unit time and C_2 = unit cost of sampling and adjustment.

The optimal sampling interval $s = s^*$ is such that $C(s^{*)}$ = min { c(s) } for $s \geq 1$.

4. PARAMETERS OF THE SAMPLED PROCESS

For the sampled process M_t defined by (5) explicit expressions for the parameters θ_s and σ_s^2 are given below (Hunter and Kartha, 1984). The M_t process will be invertible of θ_s is chosen such that

$$\theta_s = \begin{cases} \theta^+_s & \text{if } 0<\theta_1<1 \\ \theta^-_s & \text{if } -1<\theta_1<0 \end{cases} \tag{10}$$

where $s \geq 1$ and θ_s are defined by

$$\theta^+_s = [2 - \alpha s + u(s)]/2 \tag{11}$$

$$\theta^-_s = [2 + \alpha s + u(s)]/2 \tag{12}$$

where

$$u(s) = (\alpha^2 s^2 + 4\alpha s)^{1/2} \tag{13}$$

and

$$\alpha = (1 - \theta_1)^2/\theta_1 \tag{14}$$

Note that θ_s^+ is the correct root when θ_1 is positive, and θ_s^- is the correct root when θ_1 is negative, which explains the use of the superscripts + and - for θ_s. Once θ_s is determined, σ_s^2 can be obtained from (7).

The invertibility condition for the M_t process (Box and Jenkins, 1976) is that $0 \leq |\theta_s| < 1$. For processes that are invertible, the current observation does not depend overwhelmingly on observations in the remote past. The condition also leads to a unique autocorrelation structure.

5. DERIVATION OF THE OPTIMAL SAMPLING INTERVAL

The parameters θ_s and σ_s^2 of M_t the sampled process are related to those of N_t the original process as follows.

$$\theta_s = \{ 2+ \alpha s\text{-sgn}\,(\theta_1)\, u(s) \} / 2 \tag{15}$$

$$\sigma_s^2 = \theta_1\sigma_1^2 / \theta_s \tag{16}$$

where $\alpha = (1 - \theta_{1)}^2 / \theta_1$, $u\,(s) = (\alpha^2 s^2+4\alpha s)^{1/2}$ and sign (θ_1) is +1 if $0<\theta_1<1$ and -1 if $-1<\theta_1<0$. Substituting (16) into (9) and differentiating with respect to s, we get

$$C'\,(s) = \text{sgn}\,(\theta_1)\,(R\theta_1\sigma_1^2/\theta_s)\,\{\alpha / \alpha^2 s^2 + 4\alpha s)^{1/2} - 1/s^2 \tag{17}$$

and differentiating again, we get

$$C''\,(s) = \text{sgn}\,(\theta_1)\,(2R\theta_1\sigma_1^2\,\alpha^2) / (\alpha^2 s^2 + 4\alpha s)^{3/2} + 2/s^3 \tag{18}$$

It can be shown that, for $0 <|\theta_1|< 1$, the equation $C'\,(s) = 0$ leads to the cubic equation

$$F\,(s) = \alpha\beta\,(\alpha + \beta)\, s^3 + 4\alpha\beta s^2 - \alpha s - 4 = 0 \tag{19}$$

where $\beta = R\theta_1\sigma_1^2$. Cardan's solution for the general cubic equation in s,

$$s^3 + ps^3 + qs + r = 0 \tag{20}$$

Gives the following three roots:

$$s_1 = A + B - p/3 \tag{21}$$

$$s_2 = -(A+B)/2 + \{A-B)\sqrt{-3/2} - p/3 \tag{22}$$

$$s_3 = -(A+B)/2 - \{A-B)\sqrt{-3/2} - p/3 \tag{23}$$

where $A = [-b/2) + b^2/4) + (a^3/27)\}^{1/2}]^{1/3}$ $B = [-b/2) - \{b^2/4) + (a^3/27)\}^{1/2}]^{1/3}$, $a = (3q-p^2)/3$, and $b = (2p^3 - 9pq + 27r)/27$. Note that s_1 is always real. In general, recalling that $\alpha = (1-\gamma_1)^2 / \theta_1$ and

$\beta = R\theta_1\sigma_1^2$, one can directly obtain the three roots of (19) by setting $p=4/(\alpha+\beta)$, $q= -1/\beta(\alpha+\beta)$, and

$r = -4/\alpha\beta(\alpha+\beta)$. Which of the three roots s_1, s_2 and s_3 is the desired solution s^* depends on the properties of C (s) and therefore we now study this function. Note that C(s) is a continuous function of s and that the derivatives C' (s) and C" (s) exist for all permissible values of s.

We now investigate the conditions under which C (s) is a convex function of s and the properties of C (s) when these conditions are not satisfied.

Lemma 1 A necessary and sufficient condition that C (s) is a convex function of s for $s \geq 1$, $R > 0$, σ_1^2 (>0), is that

$$R\sigma^2_s \leq \begin{cases} \{(1-\theta_1)/\theta_1\}^2 & \text{if } 0<\theta_1<1 \\ \{(1-\theta_1)^3/\theta^2_1(1-\theta_1)\} & \text{if } -1<\theta_1<0 \end{cases} \tag{24}$$

Proof A necessary and sufficient condition that C (s) is convex is that C" (s) $\geq$ 0. Using (18) it can be shown that a necessary and sufficient condition for C (s) to be convex is that

$$R\theta_1\sigma_1^2\alpha^{1/2} \leq \{\alpha + (4/s)\}^{3/2} \qquad \text{if } 0<\theta_1<1 \tag{25}$$

and

$$R\beta_1\sigma_1^2\gamma^{1/2} \leq \{r - (4/s)\}^{3/2} \qquad \text{if } -1<\theta_1<0 \tag{26}$$

where

$$\alpha = \{(1-\theta_1)^2/\theta_1\} \tag{27}$$

$$\gamma = \{(1+\beta_1)^2/\beta_1\} \tag{28}$$

and $\beta_1 = -\theta_1$. It can be shown that these inequalities imply (24) and vice versa.

Note that if (24) is satisfied C(s) is convex and has one distinct real minimum. But in practice C(s) is not always convex so it is necessary to treat the case when (24) is not satisfied.

Lemma 2 C (s) is an increasing function of s for $s \geq 1$, $R > 0$, σ_1^2 (>0) and $0<|\theta_1|<1$ if

$$R\sigma^2_1 > \begin{cases} \{(1+\theta_1)^3/\theta_1^2(1-\theta_1)\} & \text{if } 0<\theta_1<1 \\ \{(1-\theta_1)/\theta_1\}^2 & \text{if } -1<\theta_1<0 \end{cases} \tag{29}$$

Proof The proof consists of showing that C' (s) > 0 under (29) for $s \geq 1$. Under conditions of this lemma C(s) is minimized when s = 1.

Lemma 3 For $R > 0$, σ_1^2 (>0), $0<\theta_1< 1$

$$\text{and } \{ (1-\theta_1) / \theta_1 \}^2 < R\sigma_1^2 < \{ (1 + \theta_1)^3 / \theta_1^2 (1- \theta_1) \} \tag{30}$$

there exists an $s \in (1, \infty)$, say s_α, such that C (s) is convex in the interval $(1, s_\alpha)$ and is an increasing function of s in the interval (s_α, ∞). s_α is given by $4 / \{ (R^2\theta_1^2\sigma_1^4\alpha)^{1/3} - \alpha \}$ where $\alpha = (1-\theta^2_1) / \theta_1$.

Proof The inequality $\{ (1-\theta_1) / \theta_1\}^2< R\sigma_1^2< \{ (1 + \theta_1)^3 / \theta_1^2 (1- \theta_1) \}$ can be shown to be equivalent to $\alpha^3< R^2\theta_1^2\sigma_1^4\alpha< (\alpha + 4)^3$. Since $(\alpha + 4/s)^3$ is decreasing in s for all $s \geq 1$ and $R^2\theta_1^2\sigma_1^4\alpha< (\alpha + 4/s_\alpha)^3$, there exists an s, s_α such that $R^2\theta_1^2\sigma_1^4\alpha = (\alpha + 4/s_\alpha)^3>\alpha^3$. Now

$R^2\theta_1^2\sigma_1^4\alpha< (\alpha +4/s)^3$ for $1 < s < s_\alpha$ and by (25) is convex in $(1, s_\alpha)$ while $R^2\theta_1^2\sigma_1^4\alpha> (\alpha +4/s)^3$

for $s_\alpha< s<\infty$ which implies that C (s) is increasing in this interval.

Lemma 4 For $R > 0$, σ_1^2 (> 0), $-1<\theta_1< 0$ and

$$\{ (1 + \theta_1)^3 / \theta_1^2 (1-\theta_1) \} < R\sigma_1^2< \{ (1-\theta_1) / \theta_1 \}^2 \tag{31}$$

there exists an $s \in (1,\infty)$, say s_γ, such that C (s) is concave in the interval $(1, s_\gamma)$ and convex in the interval (s_γ, ∞). s_γ is given by $4 / \{ \gamma- (R^2\beta_1^2\sigma_1^4\gamma)^{1/3} \}$ where $\beta_1 = -\theta_1$ and $\gamma = (1+ \beta_1)^2 / \beta_1$.

Proof The inequality $\{ (1+ \theta_1)^3 / \theta_1^2 (1-\theta_1) \}, R\sigma_1^2 < \{ (1-\theta_1) / \theta_1 \}^2$ can be shown to be equivalent to $(\gamma-4)^3< R^2\beta_1^2\sigma_1^4\gamma<\gamma^3$. Since $(\gamma - 4 / s^3)$ is increasing in s for all $s \geq 1$ and $R^2\beta_1^2\sigma_1^4\gamma<\gamma^3$ there exists an s , s_γ , such that $R^2\beta_1^2\sigma_1^4\gamma =(\gamma-4/s_\gamma)^3<\gamma^3$. Then $R^2\beta_1^2\sigma_1^4\gamma> (\gamma - 4 / s^3)$ for

$1< s< s_\gamma$ and C(s) is concave there while $R^2\beta_1^2\sigma_1^4\gamma< (\gamma - 4 / s^3)$ for $s_\gamma< s <\infty$ and C(s) is convex at this interval. The expression for s_γ is obtained from the equality above. Under conditions of this lemma, any root of $C'(s) =0$ which is in the interval $(1, s_\gamma)$ will be a relative maximum while a root of $C'(s)=0$ in the interval (s_γ, ∞) will minimize C(s). If no minimizing root exists, $s^* = 1$.

Which of the three roots s_1, s_2, s_3 is the desired solution s^* depends on the properties of C(s) discussed in the lemmas above, the consideration of which leads to the following theorem.

Theorem For an M_t process (where $R> 0$, $\sigma_1^2> 0$, $s \geq 1$) the optimal sampling interval s^* is given by

$$s^* = \begin{cases} s_0 & \text{when condition (1) is satisfied} \\ s_2' & \text{when condition (2) is satisfied} \\ s_1 & \text{otherwise} \end{cases} \tag{32}$$

where condition (1) implies

$R\sigma^2_1>\{(1+ \theta_1)^3 / \theta^2_1(1 - \theta_1)\}$ for $0 <\theta_1< 1$

and

$$R\sigma^2_1>\{(1- \theta_1) / \theta_1\}^2 \quad \text{for } -1 <\theta_1< 0 \tag{33}$$

and condition (2) implies

$\{(1+ \theta_1)^3 / \theta^2_1(1 - \theta_1)\}< R\sigma^2_1<\{(1- \theta_1) / \theta_1\}^2$ for $-1 <\theta_1< 0$

and

$$C` (s_2'> s_\gamma) = 0 \tag{34}$$

where $s_0 = 1$ is the basic unit sampling interval, $s\gamma = 4 / \{\gamma - (R^2\beta_1^2\sigma_1^4\gamma)^{1/3}\}$ and s_1 is given by (21).

Proof The proof of this theorem follows from the lemmas stated above. Note that under the conditions stated in lemma 1, when (24) is true C(s) is a convex function of s for all $s \geq 1$, $R > 0$, $\sigma_1^2 > 0$ and for θ_1 such that $0 < |\theta_1| < 1$. This implies that there is one and only one real distinct minimum of C(s) for all permissible values of R, θ_1, σ_1^2 and for all $s \geq 1$ under (24). Now s* is the root of the cubic equation F(s) given by (19) such that $C(s^*)$ is a minimum for all permissible values of R, θ_1, σ_1^2 and $s \geq 1$. Since F(s) = 0 has three roots of which only s_1 given by (21) is real for all permissible values of R, θ_1, σ_1^2 and $s \geq 1$, it follows that $s^* = s_1$ is the root which gives the minimum for C(s) whenever (24) is true. Under the conditions of Lemma 3, C(s) is convex in $(1, s_\alpha)$ and is increasing in (s_α, ∞). Since C(s) is convex in $(1, s_\alpha)$, there is one and only one real distinct minimum of C(s) for all permissible values of R, θ_1, σ_1^2 and for all $s \geq 1$ and $s^* = s_1$ where $s_1 \varepsilon (1, s_\alpha)$ and is given by (21). Since C(s) is an increasing function in the interval (s_α, ∞), $s^*=s_0$in this interval. When θ_1 is negative, C(s) is convex in (s_γ, ∞) under the conditions of Lemma 4 and any root of C'(s) = 0 which is in the interval $(1, s_\gamma)$ will be a relative maximum. Since C(s) is convex in (s_γ, ∞), C(s) has only one distinct real minimum for all permissible values of R, θ_1, σ_1^2 and for all $s \geq 1$ when $s\varepsilon (s_\gamma, \infty)$ and is given by $s_2' > s_\gamma$ such that $C'(s_2') = 0$ and is given by (21). If no minimizing root exists, s*=1. Again, under conditions of Lemma 2 C(s) is minimized when s* = 1.

Note that if unity is found to be the optimal sampling interval, then the current sampling interval is better than any values greater than unity. That is, by using any sampling interval greater than unity one increases the cost. Suppose, for example, measurements are made on a trial basis every five minutes to collect the necessary data from which the optimal sampling frequency is to be determined. In that case the basic unit sampling interval is five minutes and in terms of the basic unit s = 1. If it appears from the data that C(s) might be reduced if s < 1, it may be desirable to collect data with a smaller basic unit sampling interval to determine the best value of s. These data will also allow the assessment of the model, which may be inadequate at the higher frequency of sampling. When the parameters θ and σ_1^2 are unknown, estimates for them may be obtained (Box and Jenkins, 1976).

6. SENSITIVITY ANALYSIS

In this section we study the behavior of s* as a function of θ_1 and R. Table 1 below summarizes cost as a function of s, R and θ for values of s from 1 to 15 for four values of θ_1 (0.9, 0.5, -0.5, and -0.9) and for three values of R (0.1, 1.0 and 3.0). It can be observed that the cost in an increasing function of R.

Table 2 provides values of s* as a function of θ_1 and R which is provided graphically in Figure 1. For any s* < 1, the optimal value is s* = 1. This means that by using a sampling interval greater than the current one, the cost will be increased. As an example, suppose measurements were taken every 30 minutes on a trial basis of a manufacturing process in order to collect data from which s* is to be determined. In this case, the basic unit sampling interval is 30 minutes and if s* = 1.5, that means that the optimum interval is 45 minutes.

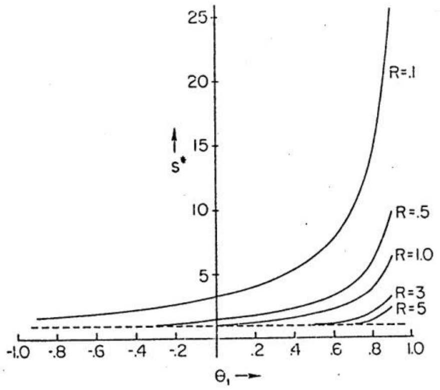

Figure no. 1 s* as a function of R and θ_1

Table no. 1 Cost as a function of s, R and σ_1

		θ_1=-.9			θ_1=-.5	
S	R=.1	R=1	R=3	R=.1	R=1	R=3
1	1.10	2.00	4.00	1.10	2.00	4.00
2	10.3	5.77	16.00	.84	3.93	10.78
3	1.23	9.27	27.15	.90	6.04	17.45
4	1.51	12.83	37.98	1.05	8.22	24.16
5	1.82	16.40	48.80	1.22	10.43	30.88
6	2.15	19.99	59.62	1.41	12.65	37.61
7	2.49	23.58	70.45	1.62	14.88	44.34
8	2.83	27.18	81.28	1.82	17.11	51.08
9	3.18	30.77	92.10	2.03	19.35	57.82
1	3.53	34.38	102.93	2.25	21.59	64.56
1	3.88	37.98	113.76	2.46	23.83	71.31
1	4.23	41.58	124.58	2.68	26.07	78.05
1	4.59	45.19	135.41	2.90	28.32	84.80
1	4.94	48.79	146.24	3.12	30.56	91.55
1	5.30	52.40	157.07	3.34	32.81	98.29

		θ_1=.5			θ_1=.9	
S	R=.1	R=1	R=3	R=.1	R=1	R=3
1	1.10	2.00	4.00	1.00	2.00	4.00
2	.63	1.81	4.43	.60	1.54	3.63
3	.49	1.93	5.11	.44	1.41	3.57
4	.44	2.17	5.85	.35	1.36	3.58
5	.41	2.33	6.60	.31	1.34	3.67
6	.41	2.56	7.35	.28	1.33	3.66
7	.41.	2.80	8.11	.26	1.33	3.71
8	.42	3.04	8.87	.25	1.34	3.76
9	.43	3.28	9.62	.23	1.34	3.81
1	.44	3.53	10.38	.23	1.35	3.86
1	.46	3.77	11.14	.22	1.37	3.91
1	.48	4.02	11.89	.21	1.38	3.97

1	.50	4.27	12.65	.21	1.39	4.02
1	.52	4.52	13.40	.20	1.40	4.07
1	.54	4.76	14.16	.20	1.47	4.12

Table no. 2 s as a function of R and θ_1*

θ_1	R=.1	R=.5	R=1	R=3	R=5
.9	25.51	9.82	6.43	3.24	2.35
.8	14.19	5.71	3.81	1.97	1.44
.7	9.91	4.11	2.77	1.46	1.08
.6	7.62	3.23	2.20	1.18	
.5	6.19	2.67	1.84	1.00	
.4	5.21	2.27	1.58		
.3	4.49	1.98	1.38		
.2	3.94	1.75	1.23		
.1	3.51	1.57	1.11		
-.1	2.87	1.28			
-.2	2.63	1.16			
-.3	2.42	1.05			
-.4	2.24				
-.5	2.09				
-.6	1.95				
-.7	1.83				
-.8	1.72				
-.9	1.62				

If s* is less than the basic unit, it will be necessary to collect data with a smaller basic unit to determine the value of s*. In table 2, values of s* $\geq$ 1 are given. Figure 1 plots the same information for values of R equal to 0.1, 0.5, 1.0, 3.0 and 5.0.

As can be observed from table 1, for large values of θ_1, the cost function is relatively flat. For smaller values of θ_1, the cost function is steeper. For the sake of comparison, let us consider the case for R = 1 and θ_1 = 0.9. The optimum value is 6 with minimum cost equal to 1.33 for this case. The cost is less than 10% higher than the minimum of the value of s is between 3 and 12. For the same comparison with the minimum cost, for θ_1 = 0.5, the value of s is between 1 and 3, a much shorter range. The optimum value in this case is 2. Therefore, the choice of a sampling interval only slightly different from the optimal value will result in considerably increased cost for small θ, while for large θ, the situation is not so sensitive.

Since R is the ratio between C_1, the cost of being off the target and C_2, the cost of sampling, and adjustment, for larger R, s* is smaller. That is, in case cost of sampling is relatively small, and hence a larger R, the best policy is to sample more frequently. On the other hand, if sampling is costly, the optimum policy is to sample less frequently, that is, with a larger s*.

Given the values of the parameters θ_1, σ_1^2 and R a corresponding optimal sampling interval can be obtained by using techniques for estimating a dynamic stochastic model from input and the output data collected over periods of time (Box and Jenkins, 1976). In either case, information regarding the effect of these parameter values on s* could be of considerable help in the final choice of the sampling interval for a given situation.

7. SPECIAL CASES

In the previous sections we considered the case where the range of the parameter θ_1 in the model for the disturbance N_t defined by (2) was restricted to $0 < |\theta_1| < 1$. We shall now consider two special cases when N_t is a random walk or a white noise series. First consider the case when $\theta_1 = 0$. The noise model is then $(1\text{-}B)N_t = a_t$, where $\{a_t\}$ is a white noise series with mean zero and variance σ_1^2.

Equivalently, N_t can be said to follow a random walk. If we sample s units apart from this process, the sampled process M_t will be $(1\text{-}B)M_t = e_t$, where $\{e_t\}$ is a white noise series with mean zero and variance $\sigma_s^2 = s\sigma_1^2$. The cost function becomes, therefore, $c(s) = R\, s\sigma_1^2 + 1/s$, and the optimal sampling interval s* in this case is given by

$$s^* = 1/\sigma_1 R^{1/2} \tag{35}$$

Let us examine the practical implications of (35). Recall that $R = C_1/C_2$ where C_1 is the cost of being off-target per unit time and C_2 is the cost of sampling adjustment. From (35) if R is large (the cost of sampling and adjustment is relatively small), s* will be small. The optimal policy therefore will be to sample with a relatively high frequency. On the other hand, from (35) if R is small (the cost of sampling is relatively large), s* will be large. In this case the suggested optimal policy is to sample with a relatively low frequency. These results agree with common sense.

Another case of special interest is when the disturbance N_t is white noise a_t, where a_t is defined as above. The sampled process M_t will also be white noise in this case. Then the cost function (9) is

$$C(s) = =R\sigma_s^2 + (1/s) = R\sigma_1^2 + (1/s) \tag{36}$$

Since $\sigma_s^2 = \sigma_1^2$. C(s) is minimized with $s \to \infty$, which means the best policy is never to sample at all! At first glance this result may seem unreasonable, because there is a danger that real processes will go out of control if not checked periodically (this is, after all, the rationale for quality control programs). But the reason for this result is clear: the assumption of a stationary time series comprised of uncorrelated errors is unrealistic for most manufacturing processes. If the relevant time series is stationary, it will always tend to return to its fixed mean value, and there is little danger that it will "wander off" causing an out-of-control condition. (In fact, generally for all stationary processes the long-run expression for σ_s^2 is a constant and $s^* = \infty$.) For this artificial stationary case, then, the mathematics does give the right answer: sampling is a waste of money. Curiously enough, a majority of the literature in quality control assumes, unrealistically, that industrial time series are stationary. This perhaps explains why algorithms for determining optimal sampling frequencies for practical quality control problems are few in the literature.

Acknowledgements

I express my gratitude to the late Professor William G. Hunter, with whom this work was initially undertaken and developed, for his contributions. I am also thankful to S. C. Wu and W.F. Lamboy for their contributions related to this work.

References

Bessegato, L., Quinino, R. Ho L.L. Duczmal, L., 2011. “Variable Interval Sampling in Economical Designs for Online Process Control of Attributes with Misclassification Errors”, Journal of Operational Research Society, 62, 1365-1375.

Box, G.E.P. and Jenkins, G.M., 1976. “Time Series Analysis, Forecasting and Control,” Holden Day, San Francisco.

Box, G. E. P. and Kramer, T., 1992. “Statistical Process Monitoring and Feedback Adjustment - A Decision”, Technometrics, 34, 251-267.

Dasgupta, T. and Mandal A., 2008. “Estimation of Process Parameters to Determine the Optimum Diagnosis Interval for Control of Defective Items”, Technometrics, 50, 2, 167-181.

Duncan, A. J., 1956. “The Economic Design of x- Charts Used to Maintain Current Control of a Process”, Journal of American Statistical Association, 51, 228-242.

Hunter, W. G. and Kartha, C.P., 1977. “Determining the Most Profitable Target Value for Production Process.” Journal of Quality Technology, 9, No. 4, 176-181.

Hunter, W. G. and Kartha, C. P., 1984. “On Sampling From an IMA (0,1,1) Process”, American Journal of Mathematical and Management Sciences, 3, No. 1, 35-46.

MacGregor, J.F., 1976. “Optimal choice of the Sampling Interval for Discrete Process Control,” Technometrics, 18, 4, 151-160.

MacGregor, J. F., 1987. “Interfaces Between Process Control and Online Statistical Process Control”, American Institute of Chemical Engineers CAST Newsletter, 9-19.

MacGregor, J. F. and T.J. Harris, 1990. “A Different View of the Funnel Experiment”, Journal of Quality Technology, 22, 255-259.

Montgomery, D.C., Keats, J.B., Runger, G. C., and Messina, W. S., 1994. “Integrating Statistical Process Control and Engineering Process Control”, Journal of Quality Technology, 26, No. 2, 79- 87.

Vander Wiel, S. A., 1996. “Monitoring Processes That Wander Using Integrated Moving Average Models”, Technometrics, 38, 2, 139-151.