ECONOMIC DESIGN OF JOINT $\overline{X}$ AND R CONTROL CHARTS: A MARKOV CHAIN METHOD

Su-Fen Yang*

楊 素 芬**

摘 要

本研究以系統化的方式提供一個$\overline{X}$和R管制圖的經濟設計方法，而此兩個管制圖在生產過程中同時被用來管制某一被選定的品質特性。研究中用$\overline{X}$及R圖追踪的一般化製程模式（a generalized process model）首度被建立，且被表示爲再生過程（a renewal process），在再生過程中每一個循環（cycle）又被表示爲馬可夫過程（Markov process）。運用我們所提供的馬可夫鏈方法推導平均循環時間（expected cycle time）和平均循環成本(expected cycle cost），繼而導出成本函數（Asymptotic cost function）會比擴展鄧肯（Duncan）或其他作者的方法簡單容易，尤其是當非隨機因素（assignable causes）有多個時。由於導出的單位時間成本是管制圖設計參數之函數，故用最佳化技巧即可決定設計參數的最佳值。唯非隨機因素越多時，此類問題的計算越複雜，但我們已設計出一般化的福傳（Fortran）程式來簡化此類複雜的計算。文中將給予一個只考慮二個非隨機因素的簡單例子及其資料分析結果。

Abstract

Economic models for the design of control charts based on Duncan's approach have been studied in the recent past. However, the economic design of control charts has not been developed in a systematic manner so far. Consequently, various assumptions and approaches have been made, and most researchers only consider process models involving a single assignable cause, for which a single control chart ($\overline{X}$, P, or S) is used. In practice, these assumptions and approaches are not realistic and flexible, and the application of a single control chart is not sufficient. In this study, the design of control charts for one process variable is treated in a systematic manner. The structure of this study is: (1) To develop a generalized process model (a Markov process) with multiple assignable causes. (2) To apply the joint $\overline{X}$ and R control charts to the generalized process model. (3) To derive a

*The author would like to thank Dr. B. C. Arnold and C.R. Robertson for their helpful comments.

**作者為本校統計研究所客座副教授

cost model depending on the design parameters, (sample size, sampling interval and control limits of $\overline{X}$ and R charts), of joint $\overline{X}$ and R charts using the Markov properties. (4) To obtain optimal design parameters by optimizing the derived cost function

It is believed that the expected cycle time and expected cycle cost are more easily obtained by the proposed Markov chain method than by extending the Duncan's approach and o'hers approaches. The generalized process model, in which we use joint $\overline{X}$ and R charts, shows that it is more reasonable and flexible than a basic process model, in which a single control chart is used. An application of the method is presented using a simple example. A general Fortran program has been written to solved this type of problem. The results of data analyses tell us the critical parameters and show that this design method gives lower quality cost compared to Shewhart's design. The design method can be applied to multiple process variables and to a variety of control charts.

1. INTRODUCTION

A major objective of statistical quality control is the systematic reduction of variability in the quality characteristic of the product of interest (see Montgomery 1985). Control charts present information about two kinds of variability. The first kind is random variability which is usually small and acceptable. This variability is due to a wide variety of chance causes, which cannot be eliminated. The second kind of variability represents a real change in the process. Such a change can be attributed to some so called assignable causes which can be eliminated. (See, for example, Braverman 1981).

Suppose that a quality characteristic of interest has given but different distributions when the process is in control (no assignable causes occur in the process), or out of control (assignable causes occur in the process). For maintaining effective control of the quality characteristic of interest, the appropriate control charts are applied to the process variable and are placed at a necessary observing station on a given production line. A sample of size, n, say X_1, X_2,..., Xn, which are assumed to be independent, is drawn periodically from the process output at the observing station. As long as the sample points plot within the control limits, the process continues, and no action is necessary. If a point plots outside of the control limits, then an investigation is initiated to locate the assignable causes. Once the assignable causes are determined and eliminated, the process returns to an acceptable stable state, as a new system. Suppose that the time between the start of two successive in-control periods is called a cycle. The process can be regarded as a series of cycles. The cycles are independent and identically distributed; it is in fact a renewal process.

The optimal economic design of the $\overline{X}$ control chart was first introduced by Duncan (1956) and has been widely extended by others. Duncan (1971) generalized his single assignable cause model to a situation in which there are several assignable causes. Assumptions are not often realistic in these models, such as; once an assignable cause occurs, the process remains in that out-of-control state until detected and that no further assignable causes occur in a single occurrence model, and a second occurrence of an assignable cause is possible but with the joint effect of the two assignable causes is always to produce a shift of constant magnitude regardless of what two assignable causes have occurred in a double occurrence model (See Montgomery 1980). Saniga (1979) proposed a model for the economic design of joint $\overline{X}$ and R charts when there are two assignable causes. He presented a single occurrence model as Duncan did, but cost model was expressed as expected cost per item produced rather than the expected cost per unit time as in Duncan (1956, 1971). Two different manufacturing process models in which a single $\overline{X}$ chart is used were first introduced by Panagos, Heikes and Montgomery (1985). One of them is a continuous process model which assumes that the process is allowed to continue in operation during the search for an assignable cause. Another is a discontinuous process model which supposes that following an alarm, the process is stopped while a search for the assignable cause is performed. A unified approach for economic design of control charts was developed by Lorenzen and Vance (1986). They considered a general process model with single assignable cause and derived an hourly cost function. A renewal theorem approach for economic designs of control charts was introduced by Banerjee and Rahim (1987). They proposed a renewal equation approach to derive an hourly cost function for a single assignable cause model. Collani and Sheil (1989) first proposed the economic design of an S chart, since they assumed that an assignable cause may only change the process variance. Their objective function was the average profit per item produced in the long run. Unfortunately, papers on the economic design of control charts have not been developed in a systematic manner. We make the following remarks about these earlier papers.

First, different papers have different assumptions and approaches. For example, some authors assume production continues, others assume production stops when looking for assignable causes (See Vance 1986). It is also assumed that assignable causes can only change the process mean or the process variance. Some authors assume the objective function is income, others assume it is cost.

Second, most authors only consider the design of $\overline{X}$ control charts (Saniga is the sole exception). Use of $\overline{X}$ and R charts simultaneously is better than only using one of them for the following reasons:

(1) Duncan (1974) noted that the joint employment of an $\overline{X}$ chart to control process mean and R chart to control process variability will give reasonably good control of the whole process.

(2) Using both $\overline{X}$, R charts gives more information, intuitively.

Use of $\overline{X}$ and R charts is also popular in practice. Saniga and Shirland (1977) in their survey of American industry reported that 71% used $\overline{X}$ chart and 64% used R charts respectively. We are sure the percentage is much higher now, and in practice, they are employed jointly. According to Montgomery (1985) they are the most commonly employed control charts for products with quality measured on a continuous scale.

Third, a generalized process model has not been considered by most researchers. Most authors consider a single assignable cause model, or a multiple assignable causes model with constraint on the number of assignable causes that can occur. They also assume that once the assignable causes have occurred the process needs to be repaired, and cannot repair itself.

In practice, many production processes are affected by several assignable causes (See Montgomery 1980) and some of these assignable causes can correct themselves without any action being taken, so, in such situations, the usual assumptions would seem inappropriate and unrealistic.

However, these problems can be solved by developing the economic design of joint $\overline{X}$ and R control charts in a systematic manner. The design method can be applied to a variety of control charts and the general process model is more reasonable and flexible than others.

2. THE GENERALIZATION OF THE PRODUCTION PROCESS MODEL

Suppose that we are interested in a measurable process variable in a production process. The distribution of the process variable is assumed to be known when the process is in control, the desired state in the process. But the distribution of the process variable will be changed when assignable causes have occurred in the process. Hence, the behavior of the process can be monitored by studying the possible occurrence of assignable causes. For detecting whether the process is in control or not, joint $\overline{X}$ and R control charts should be used to monitor the process. So the behavior of the process in which joint $\overline{X}$ and R charts are used can be monitored by studying both the possible occurrence of assignable causes and test results in the process. A generalized process model in which joint $\overline{X}$ and R charts are used is expressed as a Markov process by giving reasonable and flexible assumptions.

2.1 Assumptions of the Generalized Process Model in Which Joint $\overline{X}$ and R Control Charts Are Used

The generalized process model has the following assumptions:

(1) There is no constraint on the number of assignable causes that can occur in the process.

We assume that the quality of produced items is expressed by the value of a measurable characteristic X, that X has a normal $N(\mu, \delta^2)$ distribution and that the random variable X_i, $i=1,2,3,.,n$ are statistically independent. The normality assumption is made, since $\overline{X}$ and R charts are very robust relative to non-normality (see Burr 1967). Burr (1967) investigated the behavior of $\overline{X}$ and R control charts for 28 non-normal distributions, and found that the charts behaved quite well in many non-normal situations. However, if the distribution is very skew, then the process may appear to be out of control when in fact it is not. The normality assumption also simplifies the calculations of probabilities of type I error and powers because $\overline{X}$ and R are independent for normal samples. We also assume that there are d assignable causes which affect the mean and variance of the process variable. It is possible for more than one assignable cause to occur simultaneously; when one or more causes are in effect, others may occur producing further changes in the distribution of the process variable. It is also possible for one or more assignable causes to correct themselves without any action being taken. So, the distribution of process variable is specified by an arbitrary combination S'' (say) of assignable causes, and the possible one-step transitions are from S'' to S'', say, another combination of assignable causes, with either $S' \subset S''$ or $S'' \subset S'$. Suppose that the d assignable causes can be classified into three types, say A, B and C, where d_1 of the assignable causes belong to type A; d_2 of the assignable causes belong to type B; the others are type C. The occurrence of type A assignable causes will change the process mean; the occurrence of the type B assignable causes increases the process variance and the occurrence of type C assignable causes changes both process mean and variance. When the assignable causes which occur come from type A and type B, then the process mean shifts and the variance increases, and similarly for other combinations of types of assignable causes. The changes in the process mean and process variance depend on which particular assignable causes actually occur. Consequently, the single change of process mean or process variance and both changes of the process mean and process variance could be the effects of the combinations of occurred assignable causes.

Based on the above assumptions, we exhibit the possible distributions of the quality characteristic of interest in the production process.

When there is no assignable cause present in the process the target value is $\mu=\mu_0$ and the inherent process variability is $\delta^2=\delta_0^2$; i.e. $X \sim N\ (\mu_0,\ \delta_0^2)$

When the process is influenced by assignable causes, the distributions of process variable can be classified into seven types:

Let S_1 be a non-empty subset of $WA=\{1,2,3,\ldots.d_1\}$

S_2 be a non-empty subset of $WB=\{d_1+1,\ d_1+2,\ldots,d_1+d_2\}$

S_3 be a non−empty subset of $WC=\{d_1+d_2+1,\ldots,d_1+d_2+d_3\}$

S' be a subset of $WD=\{1,2,3,\ldots,d\}$

where $d=d_1+d_2+d_3$.

Then the most general distribution of the process variable is $N(\mu_{S'},\ \delta^2_{S'})$. But

(i) $N(\mu_{S'},\ \delta^2_{S'}).=N(\mu_{S1},\ \delta^2_0)$, when $S'=S_1$. The number of such cases is $2^{d1}-1$. (The number of non-empty subsets of WA.)

(ii) $N(\mu_{S'},\ \delta^2_{S'})=(\mu_0,\ \delta^2_{S2})$, when $S'=S_2$. The number of such cases is $2^{d2}-1$.

(iii) $N(\mu_{S'},\ \delta^2_{S'})=N(\mu_{S3},\ \delta^2_{S3})$, when $S'=S_3$. The number of such cases is $2^{d3}\ -1$.

(iv) $N(\mu_{S'},\ \delta^2_{S'})=N(\mu_{S1},\ \delta^2_{S2})$, when $S'=S_1US_2$. The number of such cases is $(2^{d1}-1)(2^{d2}-1)$.

(v) $N(\mu_{S'},\ \delta^2_{S'})=N(\mu_{S1US3},\ \delta^2_{S3})$, when $S'=S_1US_3$. The number of such cases is $(2^{d2}-1)(2^{d3}-1)$.

(vi) $N(\mu_{S'}, \delta^2_{S'}) = N(\mu_{S3}, \delta^2_{S2US3})$, when $S' = S_2US_3$. The number of such cases is $(2^{d2}-1)(2^{d3}-1)$.

(vii) $N(\mu_{S'}, \delta^2_{S'}) = N(\mu_{S1US3}, \delta^2_{S2US3})$, when $S' = S_1US_2US_3$. The number of such cases is $(2^{d1}-1)(2^{d2}-1)(2^{d3}-1)$.

So the number of total possible distributions is $2^d(=m)$. Here, the process means $(\mu_{S'})$ and variances $(\delta^2_{S'})$ all are assumed given. The process variances $\delta^2_{S'}$ are usually assumed to be greater than δ^2_0, and the process means $\mu_{S'}$ may be either greater or less than μ_0, where $S' \neq \phi$.

(2) We now define what is meant by being in-control and out-of-control.

As long as no assignable cause has occurred, we say that the process is called in control. During the period after the first assignable cause occurs and before all assignable causes leave the process the process is said to be out of control.

(3) We now define a cycle

It is assumed the production process starts in a state of statistical control. After h units of time from the start of the production process, the engineers take a sample with size, n, then they calculate the sample mean and the range for the quality characteristic of interest. If the two plotted values fall within the control limits of the joint $\overline{X}$ and R charts, the process continues. If at least one of the plotted values falls outside of the control limits of the joint $\overline{X}$ and R charts, they search for assignable causes and then return to an in-control state after repair or renewal. After such a renewal, the process is assumed to start anew. The interval between two successive renewals will be called a renewal cycle. The process is a renewal process. The accumulated cost per cycle is called the cycle cost. The cycle cost will include the costs of making transitions, costs of search and repair, cost of sampling and testing and costs due to the production of conforming and nonconforming items. The cycle costs are independent and identically disttributed. Such a process is known as a renewal reward process (See Ross 1983).

(4) We assume that the inter-occurrence times of assignable causes are exponentially distributed.

Let state variable Y_t $(t>0)$ represent the possible combination of d assignable causes at time t; then the value of Y_t would be from 1 to m $(=2^d)$; each state value j is associated with one subset of WD (S'). The amount of time $(T_{S'})$ it spends in state s' before making a transition into a different state is exponentially distributed with rate $\theta_{s'}$; $T_{S'} \sim \exp(\theta_{S'})$ or $T_j \sim \exp(\theta j)$. The process in a cycle is expressed as a Markov process.

(5) We neglect the time of sampling and testing, because they are assumed to be small.

(6) The distribution of search and repair time
The search and repair times (T_{srj}) could be dependent or independent. Their distribution depends on which assignable causes occurred. It is allowed to be any distribution or fixed.

(7) The distribution of search and repair cost
The search and repair costs (C_{srj}) could be dependent or independent. Their distribution depends on which assignable causes occurred. It is allowed to be any distribution or fixed.

(8) The setting of joint $\overline{X}$ and R control charts
Samples of size n, $(x_1, x_2, {}_3, \ldots, x_n)$, are taken at regular time intervals h from current production and the sample mean $\overline{x}=(\sum_{i=1}^{n} x_i)/n$, and the sample range $R=x_{(n)}-x_{(1)}$ are calculated and plotted on an appropriately designed $\overline{X}$ and R charts. The center line of $\overline{X}$ chart is at the target value μ_0 and the control limits are $\mu_0 \pm k_1\delta_0/\sqrt{n}$. The upper control limit of the designed R chart is set at $k_2\delta_0$. In our work we assume that assignable causes may increase process variability not decrease it, and so we do not use a lower control limit. The $\overline{X}$ and R charts are therefore characterized by the parameters, sample size (n), sampling interval (h), and control limits (k_1, k_2), $(n\geq 2,\ h>0,\ k_1>0,\ k_2>0)$ and an alarm is given whenever at least one plotted point exceeds $\mu_0 \pm k_1\delta_0/\sqrt{n}$ or $k_2\mu_0$. Such an alarm constitutes a "false alarm" if it occurs when the process is actually operating in in-control state.

Since $\overline{X}$ and R are independent (See, for example, David 1979), the probability of a false alarm and power are therefore given by (See Saniga 1979):

$$\alpha = \text{Prob}\ (R > k_2\delta_0 \mid \delta=\delta_0) + \text{Prob}(\overline{x} > \mu_0 + k_1\delta_0/\sqrt{n}\ \text{or}\ \overline{x} < \mu_0 - k_1\delta_0/\sqrt{n} \mid \mu=\mu_0)$$
$$- \text{Prob}\ (R > k_1\delta_0 \mid \delta=\delta_0)\text{Prob}(\overline{x} > \mu_0 + k_1\delta_0/\sqrt{n}\ \text{or}\ \overline{x} < \mu_0 - k_1\delta_0/\sqrt{n} \mid \mu=\mu_0)$$

$$= \overline{F}_W(k_2) + 2G(k_1) - 2\overline{F}_W(k_2)G(k_1), \qquad (2.1.2)$$

where G is defined in terms of the standard normal cumulative distribution.

$$G(a) = \int_a^{\infty} \exp\ (-z^2/2)/\sqrt{2}\sqrt{\pi}\,dz$$

F_W is defined as cumulative distribution function of relative range W ($w=R/\delta>0$), $F_W(w)=P(W<w)=1-\overline{F}_w(w)$.

$$\overline{F}_w(w) = \int_a^{\infty} f_w(w)dw \approx 1-(n/\sqrt{\pi})\{\sum_{i=1}^{ns}\alpha_i[F_Z(\sqrt{2}y_i+w)-F_Z(\sqrt{2}y_i)]^{n-1}\}$$
$$(2.1.2)$$

[for a proof of this, see Appendix A]

The cumulative probability distribution of the relative range (2.1.2) can be approximated by the Hermite polynomials method (See Sulzer, Zucker and Capuano 1952), which allows sample size to be any integer unlike the table created by Pearson and Hartley (1942) who only consider sample sizes from 2 to 20. Similarly, the power of the charts to detect a slip to one of the out-of-control states is given by 1-ß. ß depends on the process state S′; and

$$1-\beta = 1-\beta_{s'} = \text{Prob}(\overline{x} > \mu_0 + k_1\delta_0/\sqrt{n}\ \text{or}\ < \mu_0 - k_1\delta_0/\sqrt{n} \mid \mu_{S'},\ \delta_{s'})$$
$$+ \text{Prob}(R > k_2\delta_0 \mid \delta_{S'}) - \text{Prob}(\overline{x} > \mu_0 + k_1S_0/\sqrt{n}$$
$$\text{or} < \mu_0 - k_1\delta_0/\sqrt{n} \mid \mu_{S'},\ \delta_{S'})\ \text{Prob}(R > k_2\delta_0 \mid \delta_{S'})$$
$$= G_{S'} + 1 - G_{S'}' - G_{S'}\overline{F}_W(K_2\delta_0/\delta_{S'}) + G_{S'}'\overline{F}_W(k_2\delta_0/\delta_{S'}), \qquad (2.1.3)$$

where $G_{S'} = G(\sqrt{n}(\mu_0 - \mu_{S'})/\delta_{S'} + k_1\delta_0/\delta_{S'})$,

$G_{S'}' = G(\sqrt{n}(\mu_0 - \mu_{S'})/\delta_{S'} - k_1\delta_0/\delta_{S'})$,

$S' \neq \phi$.

2.2 Description of the Production Process

Given there are 2^d possible distributions in the process, the possible transitions between these distributions in a short time can be expressed as: $N(\mu_{S'}, \delta^2_{S'})$ becomes $N(\mu_{S''}, \delta^2_{S''})$, where s'' is also a subset of WD but s' $\neq$ Wd and $S' \subset S''$ or $S'' \subset S'$. Denote the transition rate from state S' to S'' by $r_{S'S''}$ (>0) if $S' \subset S''$, or $g_{S'S''}$ (>0) if $S'' \subset S'$. Consequently, we can express the possible transition rates $(g_{S'S''})$ as a matrix, the infinitesimal generator matrix.

Let **Q** be a m×m matrix, where $m=2^d$

$$\mathbf{Q} = [q_{S'S''}]$$

$$\begin{aligned} q_{S'S''} &= r_{S'S''} && \text{if } S' \subset S'', \text{ where } S'' \neq \phi \\ &= g_{S'S''} && \text{if } S'' \subset S', \text{ where } S'' \neq Wd \qquad (2.2.1) \\ &= -\Big(\sum_{S''} r_{S'S''} + \sum_{S''} g_{S'S''}\Big) && \text{if } S' = S'' \\ &= 0 && \text{otherwise.} \end{aligned}$$

Here, the sum of the rates out of S' is the exponential parameter $\theta_{S'}$;

$$\theta_{S'} = \sum_{S''} r_{S'S''} + \sum_{S''} g_{S'S''}, \text{ where } S'' \neq S'.$$

The state variable Y_t (t>0) represents the state at time t. A continuous time Markov chain Y_t is a Markov process on the states $1,2,..,2^d$. The transition probability from state i to j in the time h is $\rho_{ij}(h)=P_r(Y_{h+t}=j|Y_t=i)$, i, $j=1,2,\ldots,2^d$. We know that

(i) $\rho_{ij}(h) \geq 0$ (ii) $\sum_j \rho_{ij}(h)=1$

(iii) $\rho_{ik}(t+h)= \sum_j \rho_{ij}(t)\rho_{jk}(h)$

(iv) $\lim_{h\to 0^+} P_{ij}(h)=1$ if $i=j$

$=0$ if $i\neq j$. (See Karlin and Taylor 1976)

Let ρ (h) denote the matrix $\|\rho_{ij}(h)\|_{i,j=1}^{m}$. The transition probability in time interval h can be solved by the standard methods of systems of ordinary differential equations to yield the formula (See Karlin and Taylor 1976 or Ross 1983):

$$\boldsymbol{\rho}(h)=\exp(\mathbf{Q}h)=\mathbf{I}+\sum_{n}(\mathbf{Q}h)^n/n!. \tag{2.2.2}$$

In practical terms, the distinct eigenvalues $\mathbf{a}_1$, $\mathbf{a}_2,\ldots$, $\mathbf{a}_m$ of $\mathbf{Q}$ and a complete system of associated right eignvectors $\mathbf{u}^{(1)},\ldots,\mathbf{u}^{(m)}$ can be determined when possible. The $\boldsymbol{\rho}$ (h) has the representation $\boldsymbol{\rho}(h)=\mathbf{UD(h)U}^{-1}$ (2.2.3)
(See Karlin and Taylor 1976), where $\mathbf{U}$ is the matrix whose column vectors are, respectively, $\mathbf{u}^{(1)},\ldots,\mathbf{u}^{(m)}$ and $\mathbf{D}(h)$ is the diagonal matrix

$$\mathbf{D}(h)=\begin{vmatrix} \exp(\mathbf{a}_1h) & 0 \;\ldots\ldots & 0 \\ 0 & \exp(\mathbf{a}_2h)\;.. & 0 \\ 0 & \ldots\ldots\ldots\ldots & 0 \\ . & \ldots\ldots\ldots\ldots\ldots & 0 \\ 0 & .. & \exp(\mathbf{a}_mh) \end{vmatrix} \tag{2.2.4}$$

The rows of the matrix $\mathbf{U}^{-1}$ can also be identified as a complete system of left eigenvectors normalized to be biorthogonal to the $\{\mathbf{u}^{(\ell)}\}$. (See Karlin and Taylor 1976)

2.3 Description of The Process Which is Monitored by Joint $\overline{X}$ and R Control Charts

For detecting whether the production process is in control or not, the engineers need to take a sample after time h from the start of the production; then they do statistical tests using $\overline{X}$ and R charts. The test results are of two types; signal and no signal. A signal will be a false alarm when in fact the process is in control, or a true alarm indicating that a search and repair is required when the process is in fact out of control. No signal indicates that either the process is still in control, or that it has gone out of control and the test has not detected it. If the test result is no signal then the next sample is taken after a further time h. If the test result is a signal, then a search and repair is performed. After the search and repair, the process returns to its original condition and regular testing continues.

The transition matrix for the process is described in (2.2.1). But the possible transition would be different after we cosider the test results as well as the state of production process. Since a sample and test are taken every h time units, the possible states should be studied at the end of a statistical test. The new state can be defined as: the process variable has distribution $N(\mu_{S'}, \delta^2_{S'})$ with the test giving an alarm indicating (truly or falsely) that we need to search and repair, or no alarm. So, the number of all possible states at the end of a statistical test is b, where b=2m. Define the state τ to be the original state S′ with alarm, or with no alarm, where $\tau=1, 2, \ldots, b$. Based on the definition of a renewal cycle, these states can be grouped into two types: transient states and absorbing states. The states with the alarm showing a need to search and repair are absorbing states, others are transient states. One half of all the states are transient, one half absorbing. Let states 1 to m be transient states and the states from m+1 to b be absorbing states. Now, state variable Y_t (t=0, h, 2h, ...) stands for the states τ at discrete time t, so a discrete time Markov chain Y_t is a Markov process on the states 1, 2, ..., b. The possible transition between two successive ends of statistical tests which with interval time h would be from state τ to state τ', where $\tau=1, 2, \ldots, m$, $\tau'=1, 2, \ldots, b$. The transition probability in time interval h is $P_{\tau\tau'}(h)=Pr(Y_{h+t}=\tau' | Y_t=\tau)$, where t=0, h, 2h,..., and $h \geq 0$, $\tau, \tau'=1, 2, \ldots, b$.

Let **P** denote the matrix $\|P_{\tau\tau'}(h)\|^{b}_{\tau,\tau'=1}$. The Markov property asserts that (i) $\mathbf{P}_{\tau\tau'}(h) \geq 0$ (ii) $\sum_{\tau'} \mathbf{P}_{\tau\tau'}(h)=1$. The transition probability $p_{\tau\tau'}(h)$ can be found by using the transition probabilities $\rho_{S'S''}(h)$ (2.2.2) together with the probability of false alarm (2.1.2) or the powers (2.1.3).

The solution is represented as follows:

$$
\begin{aligned}
p_{\tau\tau'}(h) &= \rho_{ij}(h)\alpha && \text{for } \tau=i,\ \tau'=j+m, \text{ where } i=1, 2\ldots, m,\ j=1, \\
&= \rho_{ij}(h)(1-\alpha) && \text{for } \tau=i,\ \tau'=j=1, \text{ where } i=1, 2,\ldots,m. \qquad (2.3.1)\\
&= \rho_{ij}(h)\beta_j && \text{for } \tau=i,\ \tau'=j, \text{ where } i=1, 2,\ldots,m,\ j\neq 1. \\
&= \rho_{ij}(h)(1-\beta_j) && \text{for } \tau, =i,\ \tau'=j+m, \text{ where } i=1, 2,\ldots,m,\ j\neq 1. \\
&= 1 && \text{for } \tau=i+m=\tau', \text{ where } i=1, 2,\ldots,m. \\
&= 0 && \text{otherwise.}
\end{aligned}
$$

Here, for the process which is monitored by control charts, the state 1 is defined as the process is in control with no alarm, and the state m+1 stands for the process is in control with alarm and a search and repair required. For the unmonitored process, the state 1 is defined as the process is in control.

According to the result (2.3.1), the matrix **P** can be classified into four submatrices (2.3.2):

$$\mathbf{P} = \begin{vmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{O} & \mathbf{I} \end{vmatrix} \tag{2.3.2}$$

where $\mathbf{P}_{11}$ is a mxm matrix, each element of $\mathbf{P}_{11}$ is a transition probability from transient state to transient state in time interval h; $\mathbf{P}_{11} = \|p_{\tau\tau'}(h)\|^{m}_{\tau,\ \tau'=1}$

$\mathbf{P}_{12}$ is a mxm matrix and every element of $\mathbf{P}_{12}$ is a transition probability from transient state to absorbing state in time interval h; $\mathbf{P}_{12} = \|p_{\tau\tau'}(h)\|$, $\tau = 1, 2, \ldots, m$, $\tau' = m+1, \ldots, b$.

O is a mxm zero matrix.

I is a mxm identity Matrix.

2.4 Comparison Between the Generalized Process Model and Others

The advantages of the generalized process model are shown in Table 2.4.1, in which the generalized process model is compared with Duncan (1971), Saniga (1979), Montgomery (1985), Vance (1986) and Banerjee (1987).

2.4.1 The Assumptions Common to This Work and Other Work

The common assumptions of this research and other authors except Saniga are:

(1) The process variable is normally distributed.
(2) The effects of assignable causes are given.
(3) The cycle cost is the sum of (i) accumulated cost during the in-control period (ii) accumulated cost during the out-of-control period (iii) cost of sampling

TABLE 2.4.1 COMPARISON TABLE

	General process model	Dunch papers	Montgomery paper
1. no. of assignable causes	≥ 2	1('56) ≥ 2('71)	1
2. allowable no. in process	1, ≥ 2	1,2	1
3. used control charts	$\overline{X}$,R	$\overline{X}$	$\overline{X}$
4. faults may correct themselves & occur simultaneously?	yes	no	no
5. Renewal process?	yes	yes	yes
6. design parameters	n,h, k_1, k_2	n, h, k_1	n, h, k_1
7. production process	may or may not break for search	may not break for search	may or may not break for search
8. charge repair cost?	yes	no	yes
9. charge transition cost?	yes	no	no
10. sampling time	negligible	constant	constant
11. Lower Control limit of R chart	no	—	—
12. Markov chain model?	yes	no	no
13. time, cost of search and repair	any distribution	fixed	fixed
14. sample size limitation?	no	no	no

TABLE 2.4.1 COMPARISON TABLE (CONTINUOUS)

	Saniga paper	Vance paper	Banerjee paper
1. no. of assignable causes	2	1	1
2. allowable no. in process	1	1	1
3. used control charts	$\overline{X}$, R	$\overline{X}$	$\overline{X}$
4. faults may correct themselves & occur simultaneously?	yes	no	no
5. renewal process?	no	yes	yes
6. design parameters	n, h k_1, k_2, $k_3=0$	n, h, k_1	n, h, k_1
7. production process	may or may not break for search	may or may not break for search	may or may not break for search
8. charge repair cost?	yes	yes	yes
9. charge transition cost?	no	no	no
10. sampling time	not available	constant	negligible
11. lower control limit of R chart	$k_3=0$	—	—
12. Markov chain model?	no	no	renewal equation
13. time, cost of search and repair	fixed	fixed	fixed
14. sample size limitation?	2—20	no	no

and testing (iv) search and repair cost.

(4) The cycle time is the sum of (i) in-control time (ii) out-of-control time (iii) search and repair time (iv) sampling time.

2.4.2 Advantages of the General Process Model

The proposed generalized process model for the economic design of joint $\overline{X}$ and R control charts has the following advantages compared to other models.

(1) Our model considers multiple assignable causes and there is no constraint on how many assignable causes can be in effect at the same time. It would be more appropriate and realistic than single assignable cause model and single occurrence model because many production processes are affected by several assignable causes.

(2) Our model allow assignable causes to influence either the process mean or variance or both. This makes the application of the general process model more realistic.

(3) the presented assignable causes can leave the process without repairing and it is possible for more than one to occur at the same time. These assumptions make the process be more realistic and flexible.

(4) We consider the transition cost in the model. Suppose that when the process makes a transition from state i to another state j it receives a cost of d_{ij} dollars. The consideration of transition cost makes the model more realistic and flexible.

(5) Our model uses both $\overline{X}$ and R control charts, which is better than only using one of these.

(6) The model allows the search and repair times or costs to have arbitrary distributions, not necessary independent. This makes the model more flexible.

3. DEVELOPMENT OF COST MODEL

From an economic design viewpoint, the values of the design parameters, (n, h, k_1, k_2), should be chosen to minimize (or maximize) an objective function. The objective function could be income, profit or cost. We will take the asymptotic expected cost per unit time as our objective function.

From section 2, we know the generalized process model can be regarded as a series of cycles; it is in fact a renewal process, and a Markov process within cycles. So the expected cycle time and the expected cycle cost can be derived applying the Markov property, and the asymptotic expected cost per unit time is derived using the properties of renewal reward processes; the asymptotic expected cost per unit time is the ratio of the expected cycle cost and the expected cycle time (see Ross 1983). The expected cycle cost and expected cycle time are all functions of design parameters, so the asymptotic cost function also depends on the design parameters.

3.1 The Derivation of Expected Cycle Time

The cycle time is the time from the process starting in control until an alarm is detected and repaired, or equivalently it is the time from transient state 1 to reach an absorbing state (state 1 represents that the process is in control with no alarm). The state variable Y_t ($t=0$, h, 2h,...) (as in section 2.3) is a Markov chain on the states 1, 2,..., 2m and so the Markov property can be effectively used to find the expected cycle time.

Let random variable T_k be the time up to absorption (in any of the m absorbing states) from transient state k, ($k=1,2,\ldots,m$). Then using the Markov property and conditioning on the first step, we find:

$$T_k \stackrel{d}{=} h + T_\tau \quad \text{w.p. } p_{k\tau}(h),\ \tau=1,\ldots,m \tag{3.1.1}$$

$$\stackrel{d}{=} h + T_{s\gamma\tau} \quad \text{w.p. } p_{k\tau}(h),\ \tau=m+1,\ldots,b.$$

($T_{s\gamma\tau}$ is the time of search and repair for state τ. Where $\stackrel{d}{=}$ means has the same distribution as.)

To determine the expected time up to absorption from transient state k, ($k=1,2,\ldots,m$), first note that from (3.1.1) we have:

$$E(T_k) = h + \sum_{\tau=m+1}^{2m} p_{k\tau}(h)ET_{s\gamma\tau} + \sum_{\tau=1}^{m} p_{k\tau}(h)ET_\tau, \tag{3.1.2}$$

where $k=1,2,\ldots,m$.

We express (3.1.2) in matrix form:

$$\mathbf{M}_{m\times 1}=h\mathbf{1}_{m\times 1}+\mathbf{P}_{12\ m\times m}(\mathbf{M}_{sr})_{m\times 1}+\mathbf{P}_{11\ m\times m}\mathbf{M}_{m\times 1}$$

So

$$\mathbf{M}=h(\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{1}+(\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{P}_{12}\mathbf{M}_{sr} \qquad (3.1.3)$$

where

$\mathbf{M}$: a $(m\times 1)$ vector, with kth element ET_k

$\mathbf{1}$: a $(m\times 1)$ vector, with elements 1

$\mathbf{M}_{s\gamma}$: a $(m\times 1)$ vector, with τth element $ET_{s\gamma\tau}$, $\tau=m+1,\ldots,b$

$\mathbf{P}_{11}$, $\mathbf{P}_{12}$ are as defined in (2.3.3).

The expected cycle length is the first element of vector $\mathbf{M}$, i.e. $\mathbf{M}_1$ or $\mathbf{ET}_1$. After we obtain the expected cycle length, we need to calculate the expected cost of samples and tests which occur in the cycle, then the expected cycle cost can be obtained.

The expected cost of samples and tests in a cycle is the product of the expected number of samples and tests (EN) and the cost of sampling and testing (a+bn). The number (N) of samples and tests in a cycle is a random variable because it depends on cycle time. The cost of sampling and testing is assumed to be a linear function of sample size, with fixed cost, a (>0), and variable cost per unit sampled and tested, b(>0).

The expected number of samples and tests (EN) depends on whether the manufacturing process is continuous or discontinuous. In the continuous process, the process continues when engineers look for the assignable causes, and stops when they repair assignable causes; during the repair period, there is no sample and test taken. Hence, we find that the expected cycle time divided by h is the expected number of samples and tests;

$$\begin{aligned} EN &= \frac{(h(\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{1}+(\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{P}_{12}\mathbf{M}_s)_1}{h} \\ &= ((\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{1})_1+\frac{((\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{P}_{12}\mathbf{M}_s)_1}{h} \end{aligned} \qquad (3.1.4)$$

$((\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{1})_1$ means the first element of that matrix. Where $\mathbf{M}_s$ is a $m\times 1$ vector, whose τth element is the expected search time for state τ, $\tau=m+1,\ldots,b$. (This is similar to 3.1.3, except that $\mathbf{M}_s$ appears instead of $\mathbf{M}s\gamma$.)

In the discontinuous process, the process stops when engineers search and repair assignable causes; there is no sample and test performed during the search and repair period. So the expected number of samples and tests will not be influenced by the expected search and repair time;

$$EN = ((\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{1})_1. \tag{3.1.5}$$

3.2 The Derivation of Expected Cycle Cost

The derivation of expected cycle cost uses the Markov property in a similar way to that used for expected cycle time. Let $\gamma_{ij}(h)$ be the expected cost that would be associated with transition from state i to j in time interval h if the process were not monitored by joint $\overline{X}$ and R charts, i, j=1,2,...,m. Let $c_{\tau\tau'}(h)$ be the expected cost associated with transition from state τ to state τ' in time interval h for the process which is monitored by control charts, $\tau,\tau'=1,2,\ldots,b(=2m)$. Let EC_{srj} be the expected cost of search and repair for state j, j=1,2,...,m. For explaining the relationship between $\gamma_{ij}(h)$ and $c_{\tau\tau'}(h)$, we introduce matrices $\Gamma=\|\gamma_{ij}(h)\|_{i,\,j=1}^{m}$ and $\mathbf{C}=\|c_{\tau\tau'}(h)\|_{\tau,\tau'=1}^{b}$. Denote the jth column of the mxm matrix, **CS**, by cs_j and the m elements of column cs_j be EC_{srj}, j=1,2,...,m. The relationship between the three matrices is expressed as:

$$\mathbf{C} = \begin{vmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{vmatrix} = \begin{vmatrix} \Gamma & \Gamma+\mathbf{CS} \\ \mathbf{O} & \mathbf{O} \end{vmatrix} \tag{3.2.1}$$

Where matrix **C** is expressed by four submatrices: $\mathbf{C}_{11}$, $\mathbf{C}_{12}$, $\mathbf{C}_{21}$, and $\mathbf{C}_{22}$.

$\mathbf{C}_{11}$ is a mxm matrix, it is in fact the matrix Γ.

$\mathbf{C}_{12}$ is a mxm matrix, it is matrix $\Gamma+\mathbf{CS}$.

$\mathbf{C}_{21}$ is a mxm zero matrix.

$\mathbf{C}_{22}$ is a mxm zero matrix.

Their relationship can also be expressed as in (3.2.2).

$$c_{\tau\tau'}(h) = \begin{cases} \gamma_{ij}(h)+EC_{srj} & \text{for } \tau'=m+j \\ & \quad \tau=i \\ & \quad i,j=1,2,\ldots,m \\ \gamma_{ij}(h) & \text{for } \tau=i \\ & \quad \tau'=j \\ & \quad i,j=1,2,\ldots,m \\ 0 & \text{otherwise} \end{cases} \tag{3.2.2}$$

For calculating cumulative cost in a fixed time, we should consider the expected cost $\check{c}_k$ per unit time when the process is in state k, the time that the process stays in state k, and the transition cost, d_{ij}, incurred when the process moves from a state i to another state j. So the expected cost, $\gamma_{ij}(h)$, associated with transition from state i to j in time interval h if the process were not monitored would be the sum of two terms. The first is the expected cumulative cost associated with time spent in various states passed through starting from state i and ending in state j in the time interval h. Denote this by $E(Z_1(h)I(Y_h=j)|Y_0=i)$. The second is the expected cumulative transition cost associated with all transitions in the interval h starting in state i and ending in state j. Denote this by $E(Z_2(h)I(Y_h=j)|Y_0=i)$. Thus $Z_1(h)$ is a random variable which represents cumulative cost associated with time spent in various states in time interval h, $Z_2(h)$ is a random variable representing cumulative transition cost in time h.

$$\begin{aligned} \text{Define } \gamma_{ij}(h) &= E(Z_1(h)I(Y_h=j)|Y_0=i)+E(Z_2(h)I(Y_h=j)|Y_0=i) \\ &= \gamma_{1ij}(h)+\gamma_{2ij}(h), \\ \text{Where } \gamma_{1ij}(h) &= E(Z_1(h)I(Y_h=j)|Y_0=i), \\ \gamma_{2ij}(h) &= E(Z_2(h)I(Y_h=j)|Y_0=i). \end{aligned}$$

Theorem:

$$\begin{aligned} \gamma_{1ij}(h) &= E(Z_1(h)I(Y_h=j)|(Y_0=i)) = \sum_k \check{c}_k \int_0^h p_{ik}(t)P_{kj}(h-t)dt \\ &= (\mathbf{UB}_1\mathbf{U}^{-1})_{ij}, \end{aligned} \tag{3.2.3}$$

where $\rho_{ik}(t)$, $\mathbf{U}$ and $\mathbf{D}(t)$ are defined in section 2.2, $\mathbf{B}_1=\|b_{ij}\|^m_{i,j=1}$, $b_{ii}=ha_{ii}\exp(\mathbf{a}_i h)$, $b_{ij}=[a_{ij}/(\mathbf{a}_i-\mathbf{a}_j)]\ \bullet\ [\exp(\mathbf{a}_i h)-\exp(\mathbf{a}_j h)]$, where a_{ij} is the element of matrix $\mathbf{A}_1=\|a_{ij}\|^m_{i,j=1}=\mathbf{U}^{-1}\mathbf{C}_d\mathbf{U}$, $\mathbf{C}_d=\mathrm{diag}(\check{c}_1, \check{c}_2,\ldots,\check{c}_m)$.

Proof:

$$\gamma_{1ij}(h) = E(Z_1(h)I(Y_h=j)|Y_0=i),\ \text{where}\ Z_1(h)=\int_0^h\Sigma\check{c}_k I(Y_t=k)dt.$$

Let $CG_{ij}(h)=E(Z_1(h)|Y_0=i,\ Y_h=j)$

The relationship between $\gamma_{1ij}(h)$ and $CG_{ij}(h)$ is:

$$\gamma_{1ij}(h) = \int Z_1 f_{z1(h),\ yo,\ yh}(z_1),\ i,\ j,\ dz_1/f_{yo}(i)$$
$$CG_{ij}(h) = \int Z_1 f_{z1(h),\ yo,\ yh}(Z_1,\ i,\ j),\ dz_1/f_{yo,\ yh}(i,\ j),$$

So, $\gamma_{1ij}(h) = CG_{ij}(h)\rho_{ij}(h)$,

$$CG_{ij}(h) = \sum_k \check{c}_k\int_0^h \Pr(Y_t=k|Y_0=i,\ Y_h=j)dt,$$

where

$$\begin{aligned}\Pr(Y_t=k|Y_0=i,\ Y_h=j) &= \Pr(Y_t=k|Y_0=i)\Pr(Y_h=j|Y_t=k)/\Pr(Y_h=j|Y_0=i)\\ &= \rho_{ik}(t)\rho_{kj}(h-t)/\rho_{ij}(h).\end{aligned}$$

$$\begin{aligned}\text{so}\ CG_{ij}(h) &= E(Z_1(h)|Y_0=i,\ Y_h=j)\\ &= \int_0^h (\mathbf{U}\mathbf{D}(t)\mathbf{U}^{-1}\mathbf{C}_d\mathbf{U}\mathbf{D}(h-t)\mathbf{U}^{-1})_{ij}\ /\ \rho_{ij}(h)\\ &= (\mathbf{U}\mathbf{B}_1\mathbf{U}^{-1})_{ij}\ /\ \rho_{ij}(h), \qquad (3.2.4)\end{aligned}$$

Consequently,

$$\begin{aligned}\gamma_{1ij}(h) &= \sum_k\check{c}_k\int_0^h\rho_{ik}(t)\rho_{kj}(h-t)dt\\ &= \int_0^h(\mathbf{U}\mathbf{D}(t)\mathbf{U}^{-1}\mathbf{C}_d\mathbf{U}\mathbf{D}(h-t)\mathbf{U}^{-1})_{ij}dt\\ &= (\mathbf{U}\int_0^h\mathbf{D}(t)\mathbf{A}_1\mathbf{D}(h-t)dt\mathbf{U}^{-1})_{ij}\\ &= (\mathbf{U}\mathbf{B}_1\mathbf{U}^{-1})_{ij}.\end{aligned}$$

We also need to determine $\gamma_{2ij}(h)=E(Z_2(h)I(Y_h=j)|Y_0=i)$.

Theorem:

$$\gamma_{2ij}(h) = \sum_{Kl}\sum \int_0^h \rho_{ik}(t)\rho_{lj}(h-t)q_{kl}d_{kl}\ dt \quad (3.2.5)$$
$$= (\mathbf{U}\mathbf{B}_2\mathbf{U}^{-1})_{ij},$$

where d_{kl} is the expected transition cost from state k to l, q_{kl} is the transition rate from state k to l, k,l=1,...,m, but $l\neq k$, $\mathbf{B}_2=||b_{2ij}||^m_{i,j=1}$, where $b_{2ii}=ha_{2ii}\exp(\mathbf{a}_i h)$, $b_{2ij}=[a_{2ij}/(\mathbf{a}_i-\mathbf{a}_j)]\ [\exp(\mathbf{a}_i h-\exp(\mathbf{a}_j h)]$. The a_{2ij} is the element of $\mathbf{A}_2$ matrix, where $\mathbf{A}_2=||a_{2ij}||^m_{i,\ j=1}=\mathbf{U}^{-1}\mathbf{C}_q\mathbf{U}$. The matrix $\mathbf{C}_q=||q_{ij}d_{ij}||_{i,\ j=1}^{\ m}$. The ith diagonal element of the $\mathbf{C}_q$ is zero, because there is no transition cost occurred when transition does not happen; i.e., $d_{ii}=0$, i=1,2,...,m. The nondiagonal element of the $\mathbf{C}_q$ is a transition rate times a transition cost; i.e. $q_{ij}d_{ij}$, i,j=1,2,...,m, $i\neq j$.

Proof:

Let $\gamma_{2ij}(t) = E(Z_2(t)I(Y_h=j)|Y_0=i),\ t \in (0,h).$

$$d[\gamma_{2ij}(t)]/dt = \lim_{\triangle t\to 0}\ (1/\triangle t)[\gamma_{2ij}(t+\triangle t)-\gamma_{2ij}(t)]$$
$$= \lim_{\triangle t\to 0}\ (1/\triangle t)[E((Z_2(t+\triangle t)-Z_2(t))I(Y_h=j)|Y_0=i)]$$
$$= \lim_{\triangle t\to 0}\ (1/\triangle t)\sum_k \rho_{ik}(t)E((Z_2(t+\triangle t)-Z_2(t))I(Y_h=j)|Y_0=i)$$
$$= \lim_{\triangle t\to 0}\ (1/\triangle t)\sum_k \rho_{ik}(t)\sum_l \triangle t q_{kl}d_{kl}\rho_{lj}(h-t-\triangle t)$$
$$= \lim_{\triangle t\to 0}\ \sum_k\ \rho_{ik}(t)\sum_l q_{kl}d_{kl}\rho_{lj}(h-t)$$

So,

$$\gamma_{2ij}(h) = \int_0^h\sum_k \rho_{ik}(t)\sum_l q_{kl}d_{kl}\rho_{lj}(h-t)\ dt = (\int_o^h\ \mathbf{U}\mathbf{D}(t)\mathbf{A}_2\mathbf{D}(h-t)\mathbf{U}^{-1}\ dt)_{ij}$$
$$= (\mathbf{U}\mathbf{B}_2\mathbf{U}^{-1})_{ij}$$

Combining (3.2.3) and (3.2.5) we find:

$$\gamma_{ij}(h) = (\mathbf{U}\mathbf{B}_1\mathbf{U}^{-1})_{ij}+(\mathbf{U}\mathbf{B}_2\mathbf{U}^{-1})_{ij}$$

Where may be expressed in matrix form as follows:

$$\Gamma = \|(\mathbf{UB}_1\mathbf{U}^{-1})_{ij}+(\mathbf{UB}_2\mathbf{U}^{-1})_{ij}\| = \|\gamma_{ij}(h)\|^m_{i,\,j=1} = \mathbf{UB}_1\mathbf{U}^{-1}+\mathbf{UB}_2\mathbf{U}^{-1} = \mathbf{UBU}^{-1}, \quad (3.2.6)$$

Where $\mathbf{B}=\mathbf{B}_1+\mathbf{B}_2$.

Alternatively we may write

$$\begin{aligned}\Gamma &= \int_o^h \mathbf{UD}(t)\mathbf{U}^{-1}\mathbf{C}_d\mathbf{UD}(h-t)\mathbf{U}^{-1}\,dt+\int_o^h \mathbf{UD}(t)\mathbf{U}^{-1}\mathbf{C}_q\mathbf{UD}(h-t)\mathbf{U}^{-1}\,dt \\ &= \int_o^h \mathbf{UD}(t)\mathbf{U}^{-1}(\mathbf{C}_d+\mathbf{C}_q)\mathbf{UD}(h-t)\mathbf{U}^{-1}\,dt \\ &= \int_o^h \mathbf{UD}(t)\mathbf{U}^{-1}\mathbf{C}_{dq}\mathbf{UD}(h-t)\mathbf{U}^{-1}\,dt = \int_o^h \mathbf{UD}(t)\mathbf{AD}(h-t)\mathbf{U}^{-1}dt \\ &= \mathbf{UBU}^{-1}, \end{aligned} \quad (3.2.7)$$

where $\Gamma = \|\gamma_{ij}(h)\|^m_{i,j=1}$, $\quad \mathbf{C}_{dq} = \mathbf{C}_d+\mathbf{C}_q$,

$\mathbf{A} = \mathbf{U}^{-1}\mathbf{C}_{dq}\mathbf{U}$, $\quad \mathbf{B} = \int_o^h \mathbf{D}(t)\mathbf{AD}(h-t)\,dt$.

Note that the diagonal element is the expected cost per unit time in a state and the nondiagonal element is the product of an expected transition cost and a transition rate for matrix $\mathbf{C}_{dq}$.

After we obtain the value of $\gamma_{ij}(h)$, the value of $c_{\tau\tau'}(h)$ can be calculated as in (3.2.1).

For finding expected cycle cost excluding cost of samples and tests, we let random variable ϵ_k be the cost incurred up to absorption starting from transient state k, k=1,2,…,m. Then using Markov property and conditioning on the first step, we find that:

$$\begin{aligned}\epsilon_k &\overset{d}{=} c_{k\tau}(h) &&\text{w.p.} \quad p_{k\tau}(h),\ \tau=m+1,\ldots,b \\ &\overset{d}{=} c_{k\tau}(h)+\epsilon_\tau &&\text{w.p.} \quad p_{k\tau}(h),\ \tau=1,2,\ldots,m.\end{aligned} \quad (3.2.8)$$

From (3.2.8), the expected cost up to absorption from transient state k ($E\epsilon_k$) is derived:

$$E\epsilon_k = \sum_{\tau=1}^{b} P_{k\tau}(h)c_{k\tau}(h)+\sum_{\tau=1}^{m} p_{k\tau}(h)E\epsilon_\tau. \quad (3.2.9)$$

We express (3.2.9) in matrix form:

$$\check{\mathbf{U}}_{m\times 1} = [\mathbf{P}_{11\ m\times m}*\Gamma_{m\times m} : \mathbf{P}_{12\ m\times m}*(\Gamma+\mathbf{CS})_{m\times m}]\ \mathbf{1}_{b\times 1}+\mathbf{P}_{11\ m\times m}\check{\mathbf{U}}_{m\times 1}$$

$$= \mathbf{W}_{m\times b}\mathbf{1}_{b\times 1}+\mathbf{P}_{11\ m\times m}\check{\mathbf{U}}_{m\times 1},$$

where * denotes the Hadamard product of the two matrices, obtained by multiplication element by element.

So $\check{\mathbf{U}} = (\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{W}\mathbf{1}$,

where

$$\mathbf{P}*\mathbf{C} = \begin{vmatrix} \mathbf{P}_{11}*\mathbf{C}_{11} & \mathbf{P}_{12}*\mathbf{C}_{12} \\ \mathbf{O} & \mathbf{O} \end{vmatrix} = \frac{\mathbf{W}}{\mathbf{O}}, \tag{3.2.10}$$

$\check{\mathbf{U}}$: a $m\times 1$ vector, the kth element is $E\epsilon_k$,

$\mathbf{W}$: The combination of submatrices $\mathbf{P}_{11}*\mathbf{C}_{11}$ and $\mathbf{P}_{12}*\mathbf{C}_{12}$,

$\mathbf{P}_{11}$, $\mathbf{P}_{12}$ are defined as in (2.3.3),

$\mathbf{C}_{11}$, $\mathbf{C}_{12}$, Γ, $\mathbf{CS}$ are defined as in (3.2.1).

The expected cost up to absorption from state 1 is the expected cycle cost excluding cost of samples and tests; i.e. $\check{U}_1=E\epsilon_1$ (where $\check{U}_1$ means the first elements of vector $\check{\mathbf{U}}$). To obtain the expected cycle cost, we need to add the expected cost of samples and tests. The expected cost of samples and tests in a cycle is the product of the expected value of N, the number of samples and tests and the cost of sampling and testing. Hence, the expected cycle cost ($E\epsilon$) is:

$$E\epsilon = EN\cdot(a+bn) + \check{U}_1, \tag{3.2.11}$$

3.3 The Derivation of Cost Function

Based on the theory of renewal reward processes the asymptotic expected cost per unit time denoted by EV_∞ can be derived by obtaining the ratio of expected cycle cost and expected cycle time. As a result, the derived asymptotic cost function, denoted by EV_∞, is:

$$EV_\infty = \frac{((\mathbf{I}-\mathbf{P}_{11})^{-1}\mathbf{W1})_1 + EN(a+bn)}{\mathbf{M}_1} = \frac{\mathbf{E}\epsilon}{\mathbf{ET}_1}. \qquad (3.3.1)$$

Note that EV_∞ depends on the design parameters n, h, k_1 and k_2, since the expected cycle time and expected cycle cost are functions of design parameters. Hence, the optimal economic design parameters n, h, k_1, k_2, can be obtained by applying some optimization techniques to minimize the cost function.

3.4 Conclusion

We have derived the asymptotic cost function for the generalized process model in which both $\overline{X}$ and R control charts are used using the Markov property. Compared to the traditional method, the Markov chain method is much more effective, flexible and simple. In practice, the calculation of the expected cycle cost would be complex when there are many assignable causes occurring in the process, but the availability of computer programs and associated optimization schemes simplifies the complicated computations. This gives us a practical method for dealing with multiple assignable causes.

4. AN EXAMPLE

In this section, an example is given to illustrate the method which was described in sections 2 and 3.

4.1 Computation of Expected Cycle Time

Clearly the process might have multiple assignable causes. Our model is capable of dealing with this. However, for ease of exposition we consider the case in which there are several assumptions; (1) There are only two assignable causes, One of them causes a shift of process mean, the original $(X \sim N(\mu_0, \delta_0^2))$ changes to

$(X \sim N(\mu_1, \delta_0^2))$; the other causes an increases in variance, the original $(X \sim N(\mu_0, \delta_0^2))$ changes to $(X \sim N(\mu_0, \delta_2^2))$. and a double occurrence changes both process mean and variance the original $(X \sim N(\mu_0, \delta_0^2))$ changes to $(X \sim N(\mu_1, \delta_2^2))$ (2) once assignable causes occur, they cannot be removed from the process except by repair, and more than one assignable cause can not happen in a short time. (3) The transition cost will be taken to be zero in this example.

4.1.1 Definition of the State Space and of the Possible Transitions

Before calculating the expected cycle time, we need to define the possible states at the end of testing and to calculate transition probabilities in time h. In our example, there are eight possible states that could happen at the end of testing. These states are defined as follows:

state 1: no occurrence of any assignable cause and no alarm.

state 2: occurrence of assignable cause 1, but no alarm

state 3: occurrence of assignable cause 2, but no alarm

state 4: occurrence of assignable causes 1, 2, but no alarm

state 5: no occurrence of any assignable cause and an alarm, a search and repair required.

state 6: occurrence of assignable cause 1 and an alarm, a search and repair required.

state 7: occurrence of assignable cause 2 and an alarm, a search and repair required.

state 8: occurrence of assignable causes 1, 2 and an alarm, a search and repair required.

In this list, states 1-4 are transient states and states 5-8 are absorbing states. Before calculating the transition probabilities, $p_{\tau\tau'}(h)$, in time interval h for the process which is monitored, we compute the probabilities of type I and type II errors which will be needed to determine the transition probabilities, $p_{\tau\tau'}(h)$, $(\tau, \tau' = 1,2,\ldots,8)$ for a monitored process, from the transition probabilities, $\rho_{ij}(h)$, in time h, (i, j=1,2,3,4), for a process which is not monitored by $\overline{X}$ and R charts.

There are four states in the process if we do not consider the monitoring process (which may either given an alarm or not):

state 1: the process is in control;

state 2: assignable cause 1 occurs in the process;

state 3: assignable cause 2 occurs in the process;

states 4: assignable causes 1 and 2 occur in the process.

Based on the assumptions that assignable causes cannot occur simultaneously and they cannot correct themselves, the transitions from state 1 to state 4, state 2 to state 1, state 2 to state 3, state 3 to state 1, state 3 to state 2 and state 4 to all other states are impossible. So their transition rates are zero. The nonzero transition rates q_{ij} are expressed as follows:

$$
\begin{aligned}
q_{12} &= r_{12} \\
q_{13} &= r_{13} \\
q_{11} &= -(r_{12}+r_{13}) \\
q_{22} &= -r_{24} \\
q_{24} &= r_{24} \\
q_{33} &= -r_{34} \\
q_{34} &= r_{34}
\end{aligned}
$$

To simplify the algebra, and sensitivity analyses, we make the number of parameters as small as possible. So we consider the special case, $r_{12}=r_{34}=r_1$ and $r_{13}=r_{24}=r_2$. The four eigenvalues of the matrix $\mathbf{Q}$ are $\mathbf{a}_1=0$, $\mathbf{a}_2=-r_1$, $\mathbf{a}_3=-r_2$and $\mathbf{a}_4=-(r_1+r_2)$ and the eigenvectors are $\mathbf{u}'^{(1)}=(1,1,1,1)$, $\mathbf{u}'^{(2)}=(1,0,1,0)$, $\mathbf{u}'^{(3)}=(1,1,0,0)$, $\mathbf{u}'^{(4)}=(1,0,0,0)$ (in a general case, we would use an eigenvalues and eigenvectors routine). From (2.2.3), the solution for $\rho_{ij}(h)$, (i, j=1, 2,3,4) is:

$$
\begin{aligned}
\rho_{11}(h) &= \exp((-r_1-r_2)h) \\
\rho_{12}(h) &= \exp(-r_2h)(1-\exp(-r_1h)) \\
\rho_{13}(h) &= \exp(-r_1h)(1-\exp(-r_2h)) \\
\rho_{14}(h) &= (1-\exp(-r_1h))(1-\exp(-r_2h)) \\
\rho_{22}(h) &= \exp(-r_2h) \\
\rho_{24}(h) &= 1-\exp(-r_2h)
\end{aligned}
$$

$$\rho_{33}(h) = \exp(-r_1 h)$$
$$\rho_{34}(h) = 1-\exp(-r_1 h)$$
$$\rho_{44}(h) = 1.$$

(Of course the other $\rho_{ij}(h)$'s are zero.)

From (2.1.1), the probability of type I error is:

$$\alpha = 2G(k_1)+\overline{F}_W(K_2)-2G(k_1)\overline{F}_W(k_2). \tag{4.1.1.1}$$

From (2.1.3), the probabilities of type II errors are:

$$\beta_1 = G((\mu_0-\mu_1)\sqrt{n}/\delta_0+k_1)\ (\overline{F}_W(k_2)-1) + G((\mu_0-\mu_1)\sqrt{n}/\delta_0-k_1)\ (1-\overline{F}_W(k_2)) \tag{4.1.1.2}$$

$$\beta_2 = 1+2G(k_1\delta_0/\delta_1)\overline{F}_W(k_2\delta_0/\delta_1)-2G(k_1\delta_0/\delta_1)-\overline{F}_W(k_2\delta_0/\delta_1) \tag{4.1.1.3}$$

$$\beta_{12} = G((\mu_0-\mu_1)\sqrt{n}/\delta_1+k_1\delta_0/\delta_1)\ (\overline{F}_w(k_2\delta_0/\delta_1)-1) + G((\mu_0-\mu_1)\sqrt{n}/\delta_1-k_1\delta_0/\delta_1)\ (1-\overline{F}_W(k_2\delta_0/\delta_1)). \tag{4.1.1.4}$$

Since the value of $p_{\tau\tau'}(h)$ is zero if a transition from state τ to state τ' in time interval h is impossible, (τ, τ'=1,2,...,8). The structure of the transition probability matrix is as follows:

$$[p_{\tau\tau'}]_{8\times 8} = \begin{vmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} & p_{17} & p_{18} \\ 0 & p_{22} & 0 & p_{24} & 0 & p_{26} & 0 & p_{28} \\ 0 & 0 & p_{33} & p_{34} & 0 & 0 & p_{37} & p_{38} \\ 0 & 0 & 0 & p_{44} & 0 & 0 & 0 & p_{48} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

where

$$p_{11}(h) = \rho_{11}(h)(1-\alpha) \qquad p_{12}(h) = \rho_{12}(h)\beta_2$$

$$p_{13}(h) = \rho_{13}(h)\beta_3 \qquad p_{14}(h) = \rho_{14}(h)\beta_4$$

$$p_{15}(h) = \rho_{11}(h)\alpha \qquad p_{16}(h) = \rho_{12}(h)(1-\beta_2)$$

$$p_{17}(h) = \rho_{13}(h)(1-\beta_3) \qquad p_{18}(h) = \rho_{14}(h)(1-\beta_4)$$

$$p_{22}(h) = \rho_{22}(h)\beta_2 \qquad p_{24}(h) = \rho_{24}(h)\beta_4$$

$$p_{26}(h) = \rho_{22}(h)(1-\beta_2) \qquad p_{28}(h) = \rho_{24}(h)(1-\beta_4)$$

$$p_{33}(h) = \rho_{33}(h)\beta_3 \qquad p_{34}(h) = \rho_{34}(h)\beta_4$$

$$p_{37}(h) = \rho_{33}(h)(1-\beta_3) \qquad p_{38}(h) = \rho_{34}(h)(1-\beta_4)$$

$$p_{44}(h) = \beta_4 \qquad p_{48}(h) = 1-\beta_4$$

$$p_{55}(h) = p_{66}(h) = p_{77}(h) = p_{88}(h) = 1$$

4.1.2 Expected Cycle Time

Having calculated the transition probability from state τ to τ' in time interval h for the process as monitored by $\overline{X}$ and R charts, we may determine the expected cycle time as follows.

Based on the model (3.1.1), the cycle time is T_1, which satisfies the following equations:

$$\begin{aligned} T_1 &\overset{d}{=} h+T_\tau \quad \text{w.p.} \quad p_{1\tau}(h), \quad \tau=1,2,3,4 \\ &\overset{d}{=} h+T_{sr\tau} \quad \text{w.p} \quad p_{1\tau}(h), \quad \tau=5,6,7,8 \end{aligned} \tag{4.1.2.1}$$

($T_{sr\tau}$ is the search and repair time in state τ).

Consequently, as shown in appendix B, the expected cycle time is:

$$\begin{aligned} ET_1 &= (1/1-p_{11}(h))(h+\sum_{5}^{8} p_{1\tau}(h)ET_{sr\tau}+\sum_{2}^{4} p_{1\tau}(h)ET_\tau) \qquad (4.1.2.2) \\ &= [1/1-p_{11}(h)]\ \{h[1+p_{12}(h)/(1-p_{22}(h))+p_{12}(h)p_{24}(h)/(1-p_{22}(h))p_{48} \\ &\quad +p_{13}(h)/(1-p_{33}(h))+p_{13}(h)p_{34}(h)/(1-p_{33}(h))p_{48}(h) \\ &\quad +p_{14}(h)/p_{48}(h)]+T_{sr6}(p_{12}(h)p_{26}(h)/(1-p_{22}(h))+p_{16}(h)) \\ &\quad +T_{sr7}(p_{13}(h)p_{37}(h)/(1-p_{33}(h))+p_{17}(h)) \end{aligned}$$

$+T_{sr8}[p_{12}(h)(p_{24}(h)+p_{28}(h))/(1-p_{22}(h))+$

$P_{13}((h)(P_{34}(h)+p_{38}(h))/(1-p_{33}(h))+p_{14}(h)+p_{18}(h)]+p_{15}(h)T_{sr5}\}.$

Next, we need to calculate the expected cycle cost. In order to do this, we first focus on the expected cost of samples and tests.

4.1.3. Expected Number of Samples And Tests

The expected cost of samples and tests is the product of expected number of samples and tests (EN) and cost of sampling and testing (a+bn) (Recall that each test is based on a sample of size n). The expected number of samples and tests depends on whether the production process is continuous or discontinuous.

For continuous process, the expected number of samples and tests is:

$$EN = \{(1/(1-p_{11}(h))\ (h+\sum_{5}^{8}p_{1\tau}(h)ET_{s\tau}+\sum_{2}^{4}p_{1\tau}(h)ET_{\tau})\}/h. \qquad (4.1.3.1)$$

In the discontinuous process, the expected number of samples and tests is:

$$EN = [(1/(1-p_{11}(h))\ (h+\sum_{2}^{4}p_{1\tau}(h)ET_{\tau})]/h. \qquad (4.1.3.2)$$

Note, the difference between (4.1.3.1) and (4.1.3.2) is that there is no $ET_{s\tau}$ in (4.1.3.2). The difference between (4.1.2.2) and (4.1.3.1) is that $ET_{s\tau}$ appears instead of $ET_{sr\tau}$.

4.2 Computation of Expected Cycle Cost

The expected cumulative cost up to absorption from state 1 is the expected cycle cost excluding cost of samples and tests. To find this, the expected cost associated with transition from state τ to τ', $c_{\tau\tau'}(h)$, in time h, (τ, $\tau'=1,2,\ldots,8$), for the monitored process should be computed first. The procedure to compute the $c_{\tau\tau'}(h)$ is complex. It will be described in the next section.

4.2.1 The Structure of the Cost Matrix

Let the cost matrix **C** be $\|c_{\tau\tau'}(h)\|_{\tau,\tau'=1}^{8}$. The value of $c_{\tau\tau'}(h)$ will be zero if the corresponding transition probability in time h, $p_{\tau\tau'}(h)$, is zero, or if state τ is absorbing, $\tau,\tau'=1,2,\ldots,8$. So the cost matrix has following structure.

$$[c_{\tau\tau'}] = \begin{vmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & c_{17} & c_{18} \\ 0 & c_{22} & 0 & c_{24} & 0 & c_{26} & 0 & c_{28} \\ 0 & 0 & c_{33} & c_{34} & 0 & 0 & c_{37} & c_{38} \\ 0 & 0 & 0 & c_{44} & 0 & 0 & 0 & c_{48} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix} \qquad (4.2.1.1)$$

(note that all the $c_{\tau\tau'}$'s are actually functions of h)

In order to obtain the expected cost, $c_{\tau\tau'}$, associated with transition from state τ to τ' in time interval h, we need to know the expected cost, $\gamma_{ij}(h)$, associated with transition from state i to state j in time h for an unmonitored process,and the expected cost, $EC_{sr\tau'}$, of search and repair, where i, j=1,2,3,4, and τ, $\tau'=1,2,\ldots,8$.

For calculating $\gamma_{ij}(h)$, we need to know

(1) eigenvalues and eigenvectors for infinitesimal generator matrix.

(2) expected cost per hour for state k, k=1,2,3,4.

(3) transition probability, $\rho_{ij}(h)$, from state i to state j in time h for the unmonitored process, i, j=1,2,3,4.

Then the solution of $\gamma_{ij}(h)$ can be derived using the model (3.2.1).
The $\gamma_{ij}(h)$ are calculated as follows (from model 3.2.1):

$$\gamma_{11}(h) = h\check{c}_1\exp((-r_1-r_2)h)$$

$$\gamma_{12}(h) = h\check{c}_2\exp(-r_2h)-h\check{c}_1\exp((-r_1-r_2)h)$$
$$-(\check{c}_1-\check{c}_2)(\exp((-r_1-r_2)h)-\exp(-r_2)h/r_1$$

$$\gamma_{13}(h) = h\check{c}_3\exp(-r_1h)-h\check{c}_1\exp((-r_1-r_2)h)$$
$$-(\check{c}_1-\check{c}_3)(\exp((-r_1-r_2)h)-\exp(-r_1h))/r_2$$

$$\gamma_{14}(h) = h\check{c}_1 - h\check{c}_3\exp(-r_1h) - h\check{c}_2\exp(-r_2h) + h\check{c}_1\exp((-r_1-r_2)h)$$
$$+(\check{c}_1-\check{c}_3)(\exp((-r_1-r_2)h)-\exp(-r_1h))$$
$$+(\check{c}_1-\check{c}_2)(\exp((-r_1-r_2)h)-\exp(-r_2h))/r_1$$
$$\gamma_{22}(h) = h\check{c}_2\exp(-r_2h)$$
$$\gamma_{24}(h) = h\check{c}_4 - h\check{c}_2\exp(-r_2h) - (\check{c}_2-\check{c}_4)(\exp(-r_2h)-1)/r_2$$
$$\gamma_{33}(h) = h\check{c}_3\exp(-r_1h)$$
$$\gamma_{34}(h) = h\check{c}_4 - h\check{c}_3\exp(-r_1h) - (\check{c}_3-\check{c}_4)(\exp(-r_1h)-1)/r_1$$
$$\gamma_{44}(h) = h\check{c}_4.$$

($\check{c}_i$ is expected cost per hour, where $i=1,2,3,4$.)

Eventually, the expected cost, $c_{\tau\tau'}(h)$, associated with transition from state τ to τ' in time interval h, for the monitored process is computed from model 3.2.5 as:

$$c_{11}(h) = \gamma_{11}(h) \qquad c_{15}(h) = \gamma_{11}(h) + EC_{sr5}$$
$$c_{12}(h) = \gamma_{12}(h) \qquad c_{16}(h) = \gamma_{12}(h) + EC_{sr6}$$
$$c_{13}(h) = \gamma_{13}(h) \qquad c_{17}(h) = \gamma_{13}(h) + EC_{sr7}$$
$$c_{14}(h) = \gamma_{14}(h) \qquad c_{18}(h) = \gamma_{14}(h) + EC_{sr8}$$

$$c_{22}(h) = \gamma_{22}(h) \qquad c_{33}(h) = \gamma_{33}(h)$$
$$c_{24}(h) = \gamma_{24}(h) \qquad c_{34}(h) = \gamma_{34}(h)$$
$$c_{26}(h) = \gamma_{22}(h) + EC_{sr6} \qquad c_{37}(h) = \gamma_{33}(h) + EC_{sr7}$$
$$c_{28}(h) = \gamma_{24}(h) + EC_{sr8} \qquad c_{38}(h) = \gamma_{34}(h) + EC_{sr8}$$

$$c_{44}(h) = \gamma_{44}(h)$$
$$c_{48}(h) = \gamma_{48}(h) + EC_{sr8}.$$

4.2.2 Expected Cycle Cost Excluding Cost of Samples and Tests

After we have the cost matrix (4.2.1.1), let the random variable ϵ_1 be the cost up to absorption from state 1, then the expectation ($E\epsilon_1$) can be obtained by applying the formula in (3.2.8).

$$\epsilon_1 \overset{d}{=} c_{1\tau}(h) \quad \text{w.p.} \quad p_{1\tau}(h), \quad \tau=5,\ 6,\ 7,\ 8$$
$$\overset{d}{=} c_{1\tau}(h)+\epsilon_\tau \quad \text{w.p.} \quad p_{1\tau}(h), \quad \tau=1,\ 2,\ 3,\ 4$$

The expected cycle cost, as shown in appendix C, excluding cost of samples and tests is:

$$E\epsilon_1 = \sum_{1}^{8} p_{1\tau}(h)C_{1\tau}(h)+\sum_{1}^{4} p_{1\tau}(h)E\epsilon_\tau.$$

or

$$E\epsilon_1 = (1/p_{11}(h))(\sum_{1}^{8} p_{1\tau}(h)c_{1\tau}(h)+\sum_{2}^{4} p_{1\tau}(h)E\epsilon_\tau). \qquad (4.2.2.1)$$

$$= [1/p_{11}(h)]\{\Sigma p_{1\tau}(h)c_{1\tau}(h)+p_{12}(h)/(1-p_{22}(h))$$
$$[p_{22}(h)c_{22}(h)+p_{24}(h)c_{24}(h)+p_{26}(h)c_{26}(h)+p_{28}(h)c_{28}(h)+$$
$$p_{24}(h)p_{44}(h)c_{44}(h)/p_{48}(h)+p_{24}(h)c_{48}(h)]+$$
$$(p_{13}(h)/1-p_{33}(h)))[p_{33}(h)c_{33}(h)+p_{34}(h)c_{34}(h)+$$
$$p_{37}(h)c_{37}(h)+p_{38}(h)c_{38}(h)+p_{34}(h)p_{44}(h)c_{44}(h)/p_{48}(h)+$$
$$p_{34}(h)c_{48}(h)]+p_{14}(h)p_{44}(h)c_{44}(h)/p_{44}(h)+p_{14}(h)c_{48}(h)$$

$E\epsilon_1$ does not include the cost of samples and tests; hence, to obtain the expected cycle cost we need to add the expected cost of samples and tests.

4.2.3 Expected Cycle Cost

The expected cycle ($E\epsilon$) is the sum of the expected cost up to absorption from state 1 ($E\epsilon_1$) and the expected cost of samples and tests.
So the expected cycle cost is:

$$E\epsilon = (1/p_{11}(h))(\sum_{1}^{8} p_{1\tau}(h)C_{1\tau}(h)+\sum_{2}^{4} p_{1\tau}(h)E\epsilon_\tau)+EN(a+bn). \qquad (4.2.3.1)$$

4.3 Cost Model

The asymptotic expected hourly cost (EV_∞) is obtained by dividing the expected cycle cost ($E\epsilon$) by the expected cycle time (ET_l) applying the property of renewal reward processes (Ross 1983).

$$EV_\infty = E\epsilon/ET_l = f_v(n,\ h,\ k_1,\ k_2) \tag{4.3.1}$$

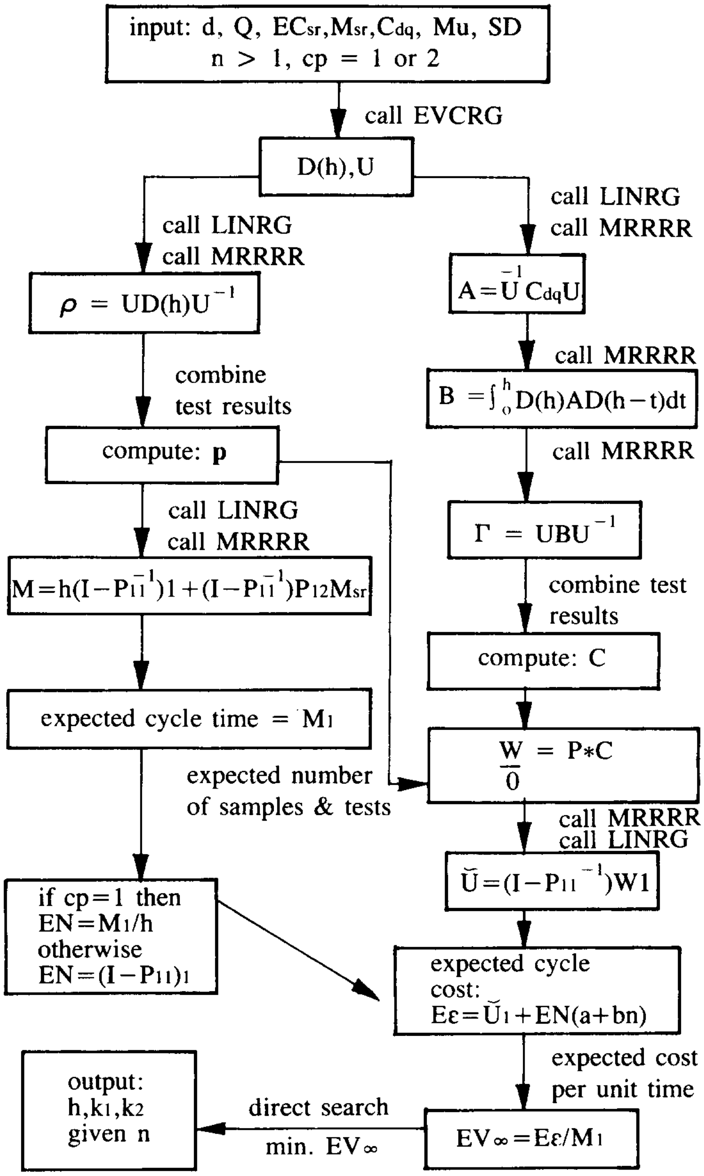
The asymptotic hourly cost function depends on the design parameters, since the expected cycle cost and the expected cycle time are both functions of design parameters. The asymptotic expected hourly cost function may be minimized with respect to design parameters, (n, h, k_1, k_2), by optimization techniques.

5. DATA ANALYSES

A general Fortran program has been written to solve this type of problem. To use this Fortran program (Figure 5.1.1),we need to specify the number of assignable causes, the expected costs and the expected times of search and repair at various states, the transition rates and the transition costs, the costs per unit time at various states, changes in process mean and variance, a specified sample size and whether the process is continuous or discontinuous. The eigenvalues and eigenvectors of the infinitesimal generator matrix can be calculated by routine EVCRG. The transition probabilities and expected cost associated with transition from a state to another state in time h for the process which has not been monitored can be obtained using the eigenvectors and eigenvalues in routines LINRG and MRRRR (for inverting and multiplying matrices). Expected cycle time and expected cycle cost can also be calculated using routine LINRG and MRRRR. The optimal design parameters for a specified sample size can be obtained using direct search algorithm which minimizes the asymptotic hourly cost function. These subroutines of the program are obtained from the IMS Library (1989). Using this program, the data analyses can be performed. In the data analyses, we perform the sensitivity analyses for the example which is described in Chapter 4, but with the sample size restricted to be less than 21 and with the continuous process model. Montgomery (1985) suggested that the continuous process model is most likely to occur in practice.

In this chapter, the global minima in the parameter region considered are presented. The optimization scheme, subject to constraints used, follows the direct search method which is described in the figure 5.1.2. The advantages of our model for economic design of $\overline{X}$ and R charts are shown by comparing them with the cost of Shewhart control charts. We also discuss how the optimal design

FIGURE 5.1.1
FLOW CHART FOR FORTRAN PROGRAM



Note on Figure 5.1.1:

d=number of assignable causes.

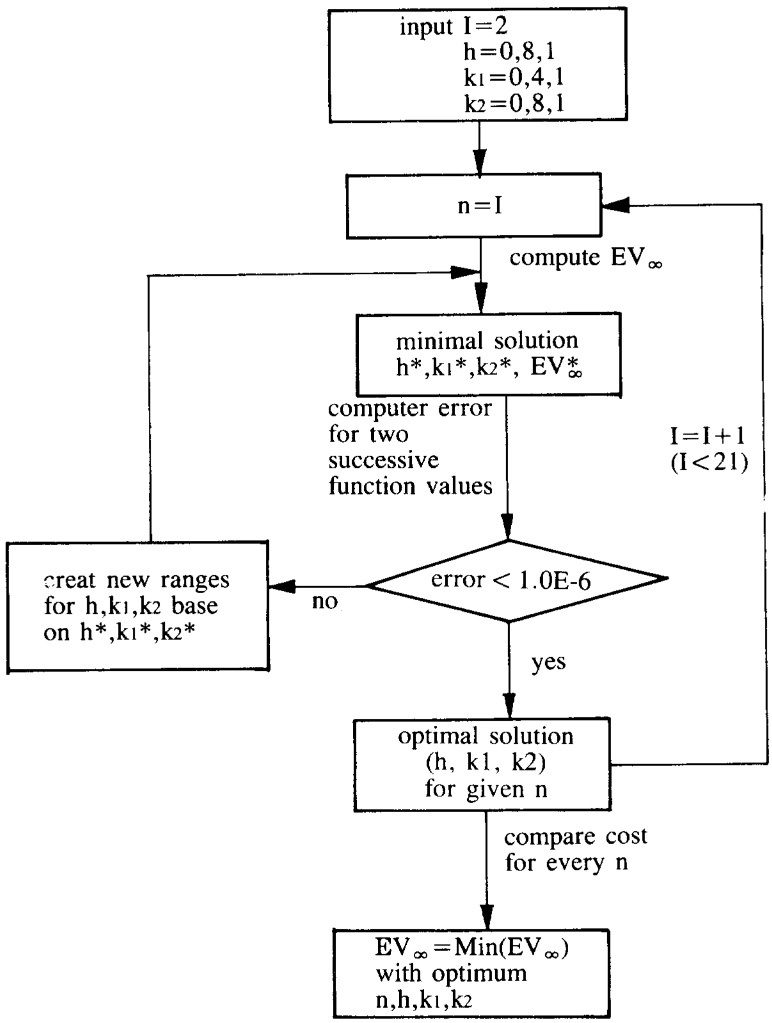
cp=1 means the manufacturing process model is continuous, otherwise it is discontinuous.

EVCRG: A routine to obtain eigenvalues and eigenvectors for a matrix.

LINRG: A routine for finding the inverse matrix.

MRRRR: A routine for multiplying two real matrices.

FIGURE 5.1.2
THE PROCEDURE TO FIND THE OPTIMAL PARAMETERS



parameters of the $\overline{X}$ and R control charts are affected by various parameters.

5.1 Optimization Technique: Direct Search Method

After deriving the cost function which is a function of the design parameters n, h, k_1, k_2, the optimal design parameters, h, k_1, k_2, are obtained using the direct search method for a specified integer n, and specified values of parameters in the estimated range. The model requires that eighteen cost and process parameters be specified by the user. In practice, the values of the parameters will need to be supplied by experts in the production process, such as production engineers. Under the estimated ranges of the parameters, the results of data analyses can offer reliable information to understand the effect of each parameter and help decision-makers make more precise decisions.

Table 5.1.1 shows the ranges of the parameters used in the Numerical examples. The ranges of most parameters are based on Duncan (1971) and Saniga (1979). Others are set up to satisfy reasonable requirements. For example, the cost per unit time when the process is in control should be smaller than that when the process is out of control, the cost of search and repair when the process is influenced by multiple assignable causes should be larger than that when the process is influenced by single assignable cause. Similar considerations apply to the times of search and repair. For minimizing the cost function which is subject to constraints, the method starts by searching a coarse grid to find an approximate solution; the process is repeated with finer grids until sufficient accuracy is obtained. The iterative procedure is terminated when the difference of successive cost values is less than 1.0E-6 for every given sample size n $(1<n)$, so that the values of h, k_1, k_2, for the minimum point are computed. The value of h is limited to be at most eight, becaue eight hours is the usual length of an industrial shift, and in most circumstances the quality control engineer does not like a shift to go by without some information on the process (See Duncan 1971). The values of k_1 and k_2 are less than 4 and 8 respectively because larger values give the cumulative Normal and Relative range probabilities approximately 1. The objective function and constraints expressed mathematically are:

$$\begin{aligned} \text{Min.} \quad & EV_\infty = f_v\,(n,\ h,\ k_1,\ k_2) \\ \text{st.} \quad & 0 < h \le 8 \\ & 0 < k_1 \le 4 \\ & 0 < k_2 \le 8 \\ & 1 < n,\ n \text{ is integer.} \end{aligned}$$

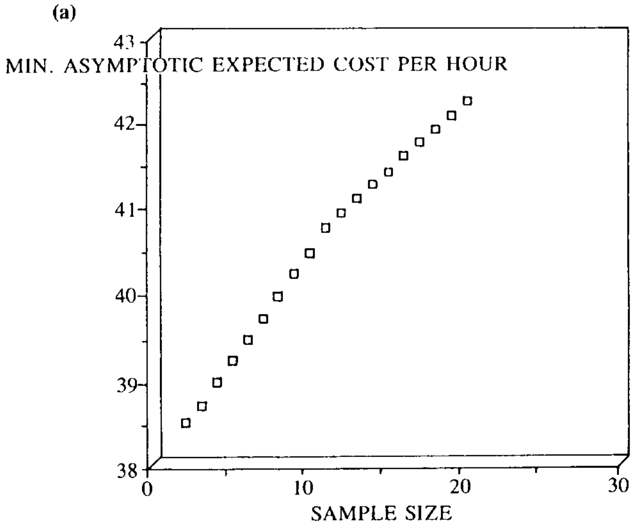
TABLE 5.1.1.
NUMERICAL EXAMPLE: THE RANGES OF PARAMETERS

Parameters	Lower bound	Upper bound
r_1	0.1	1.0
r_2	0.1	0.5
$Mu=\mu_0-\mu_1$	0.5	4.0
$SD=\delta_1/\delta_0$	1.5	4.0
EC_{sr5}	50.0	500.0
a	5.0	100.0
b	0.5	10.0
ET_{sr5}	0.3	0.6
ET_{sr6}	1.0	2.0
ET_{sr7}	1.5	2.0
ET_{sr8}	2.0	3.5
EC_{sr6}	80.0	500.0
EC_{sr7}	100.0	300.0
EC_{sr8}	150.0	400.0
$\check{c}_0$	15.0	25.0
$\check{c}_1$	20.0	30.0
$\check{c}_2$	40.0	60.0
$\check{c}_3$	50.0	70.0

We optimize EV_∞ with respect to h, k_1, k_2 for integers n starting at $n=2$ and increasing, and then choose the optimum value of n. Considered as a function of n, the graph of minimal asymptotic expected cost per hour for a specific case is illustrated in Figure 5.1.3. Typical shapes of graphs are presented in Figure 5.1.3. (a) to (b). In picture (a) the global minimum is at the beginning; $n=2$. The asymptotic expected cost per hour increases when sample size increases. The objective function in picture (b) has an interior minimum; the asymptotic expected cost per hour falls to a minimum and then rises. The shapes of the graphs are sensitive to changes in the values of (i) the expected costs of search and repair, (EC_{sr5}, EC_{sr6}, EC_{sr7}, EC_{sr8}), (ii) the fixed cost (a) and unit cost (b) of sampling and testing, (iii) the size of the shift of process mean, (Mu), (iv) the ratio of the in and out of control process standard deviations, (SD), and (v) the transition rates, (r_1, r_2). For example, if the expected costs of search and repair are high the graph

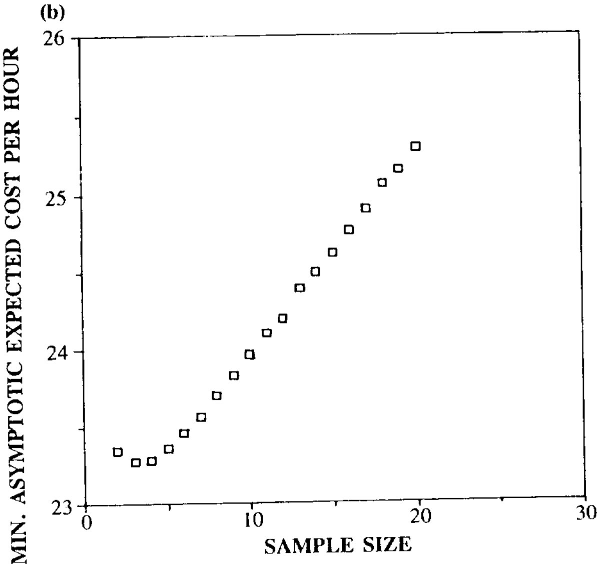
FIGURE 5.1.3
GRAPH OF MINIMUM ASYMPTOTIC EXPECTED COST PER HOUR AND SAMPLE SIZE

(a)



$A=5$ $b=.9$ $Mu=.5$ $SD=1.5$ $r_1=.33$ $r_2=1.0$
$\check{c}_0=15$ $\check{c}_1=20$ $\check{c}_2=40$ $\check{c}_3=50$ $EC_{sr5}=50$
$EC_{sr6}=80$ $EC_{sr7}=100$ $EC_{sr8}=150$ $ET_{sr5}=.5$
$ET_{sr6}=1$ $ET_{sr7}=1.5$ $ET_{sr8}=2.0$

(b)



$a=5$ $b=0.9$ $EC_{sr5}=50$ $EC_{sr6}=80$ $EC_{sr7}=100$ $EC_{sr8}=150$
$r_1=0.5$ $r_2=0.1$ $\check{c}_0=15$ $\check{c}_1=20$ $\check{c}_2=40$ $\check{c}_3=50$ $ET_{sr5}=.5$
$ET_{sr6}=2$ $ET_{sr7}=2.5$ $ET_{sr8}=3.0$ $Mu=0.5$ $SD=1.5$

will have an interior minimum; if these costs are low, then the minimum is achieved for n=2. However, a high value of the shift in the process mean will cause the graph to have a global minimum at the beginning; if the shift is small, then the graph will have an interior minimum. We find similar behavior for the parameters SD, a, b and r_1; high values of them will cause the graphs to have global minima at the beginning, but low values of them will be associated with interior minima. For low values of EC_{srj}, (j=5, 6, 7), changing the values of penalty costs, ($\check{c}_0$, $\check{c}_1$, $\check{c}_2$, $\check{c}_3$), cannot change the shapes of the graphs; the graphs all have global minimum at the beginning; n=2.

5.2 A Cost Comparison of Shewhart's Design and Our Model for Economic Design

The control limits of Shewhart's control charts are fixed at $k_1=3.0$ and $k_2=5.4$ with the probability of type I error being .0027. (Choosing k_1 to take the quite natural value $k_1=3$, we find a type I error probability of .0027; this leads to a value 5.4 for k_2. These values are the ones customarily used. Also, see Saniga 1979) To compare the asymptotic expected cost of the Shewhart's design (EVS) and our model for economic design (EV_∞), the optimal cost of Shewhart's design is obtained by fixing $k_1=3$, $k_2=5.4$, n=5 and optimal h (Figure 5.2.1). Then the cost difference is computed. In Table 5.2.1 we present results showing the gains due to using our economic design method. These are expressed in terms of ic ($=EVS-EV_\infty$) and p ($=100 \times ic/EV_\infty$).

Table 5.2.1 shows that the improvements are quite large, and correspond to percentage increases of between 4—103%. Most of numerical examples have the values of p higher than 20%. We find from Table 5.2.1 that the percentage increase (p) is not sensitive to change in the costs of sampling and testing (a and b) and the costs of search and repair (EC_{sr5}, EC_{sr6}, EC_{sr7}, and EC_{sr8}), but it is sensitive to penalty costs ($\check{c}_0$, $\check{c}_1$, $\check{c}_2$, and $\check{c}_3$), to change in the transition rates (r_1 and r_2), to the process mean (Mu) and the process standard deviation (SD). Changing a and b causes a small change in p. Increasing EC_{sr5}, EC_{sr6}, EC_{sr7} EC_{sr8} all decrease the value of p. Increasing $\check{c}_0$, $\check{c}_1$ and $\check{c}_2$ all decrease the value of p. Small changes in r_1 cause larger increases in the p value. Small changes in r_2 cause larger changes in the value of p. Increases in the shift of process mean (Mu) and changes of process standard deviation (SD) increase the p values.

TABLE 5.2.1

The OPTIMUM h, INCREASE IN COST (ic), AND PERCENTAGE (p)

(1) For various costs of sampling and testing, and costs of false alarm:

a	b	EC_{sr5}	h	ic	p
5	0.5	50	1.58	11.33	21.11
		250	1.58	7.21	18.40
		500	1.58	6.25	15.56
5	5.0	50	8.00	13.70	36.10
		250	8.00	9.73	23.83
		500	8.00	9.4	22.20
50	0.5	50	8.00	7.17	15.16
		250	8.00	6.98	14.70
		300	8.00	6.73	14.10
50	5.0	50	8.00	8.88	18.34
		250	8.00	8.67	17.83
		500	8.00	8.07	16.39
100	0.5	50	8.00	7.17	13.39
		250	8.00	6.96	12.95
		500	8.00	6.73	12.47
100	5.0	50	8.00	8.26	15.01
		250	8.00	8.63	15.79
		500	8.00	8.43	15.36

$EC_{sr6}=80$ $EC_{sr8}=150$ $ET_{sr6}=1$ $EC_{sr7}=2$ $Mu=.5$ $r_2=.25$ $EC_{sr7}=100$
$ET_{sr5}=.5$ $ET_{sr7}=1.5$ $SD=1.5$ $r_1=.33$ $\check{c}_0=15$ $\check{c}_1=20$ $\check{c}_2=40$ $\check{c}_3=50$

(2) For various costs of search and repair:

EC_{sr6}	EC_{sr7}	EC_{sr8}	h	ic	p
80	100	150	1.58	11.32	20.70
250	100	150	1.58	3.65	8.50
500			3.02	3.20	7.25
80	200	150	3.02	6.47	15.82
	300		3.02	3.08	6.92
80	100	250	8.00	9.22	22.11
	200	250	8.00	2.76	5.73
	100	400	8.00	3.65	7.26

$a=5$ $EC_{sr5}=50$ $ET_{sr7}=1.5$ $r_1=.33$ $Mu=.5$ $\check{c}_1=20$ $b=.5$ $ET_{sr6}=1$
$ET_{sr8}=2.0$ $r_2=.25$ $SD=1.5$ $\check{c}_2=40$ $ET_{sr5}=.5$ $\check{c}c_3=50$ $\check{c}_0=15$

(3) For various penalty costs:

$\check{c}_0$	$\check{c}_1$	$\check{c}_2$	$\check{c}_3$	h	ic	p
15	20	40	50	3.02	12.83	36.64
20				3.02	7.12	17.27
25				3.02	6.51	14.30
15	25	40	50	3.02	8.61	21.9
	30			5.46	8.57	21.25
15	20	50	50	3.02	8.60	21.57
		60		8.00	8.59	21.32
15	20	40	55	1.58	12.18	31.92
15	20	40	70	1.28	15.94	43.26

$a=5$ $EC_{sr5}=50$ $EC_{sr7}=100$ $ET_{sr6}=1$ $ET_{sr8}=2$ $r_1=.25$ $b=.9$ $EC_{sr6}=80$
$EC_{sr8}=150$ $ET_{sr7}=1.5$ $ET_{sr5}=0.5$ $r_2=.33$ $SD=1.5$ $Mu=.5$

(4) For various transition rates:

r_1	r_2	h	ic	0
0.1	0.1	1.49	12.15	27.30
	0.25	2.32	12.73	29.40
	0.50	2.79	13.36	37.34
0.33	0.1	1.55	15.82	65.10
	0.25	3.01	12.54	34.28
	0.50	8.00	9.12	21.46
0.5	0.1	1.99	16.31	66.14
	0.25	3.27	11.61	30.96
	0.5	3.27	5.14	11.13
1.0	0.1	2.14	20.92	103.64
	0.25	4.73	12.30	33.62
	0.50	8.00	3.05	6.36

$a=8$, $b=0.5$, $\check{c}_0=15$, $\check{c}_1=20$, $\check{c}_2=40$, $\check{c}_3=50$, $EC_{sr5}=50$, the values of other parameters are same as Table in the last page.

(5) For various process mean (Mu) and process variance (SD)

Mu	SD	h	ic	p
0.5	1.5	8.00	2.45	4.71
1.5		8.00	2.32	4.73
2.0		8.00	2.41	4.92
3.0		8.00	5.29	11.46
0.5	2.0	8.00	3.96	8.38
	3.0	5.46	4.92	11.11
	4.0	5.46	5.48	12.25

$a=5$ $\check{c}_0=15$ $\check{c}_2=40$ $EC_{sr6}=100$ $EC_{sr8}=250$ $ET_{sr6}=1$ $ET_{sr8}=2$ $b=.9$ $\check{c}_1=20$ $\check{c}_3=50$ $EC_{sr7}=180$ $EC_{sr5}=100$ $ET_{sr7}=1.6$ $r_1=.33$ $ET_{sr5}=.5$ $r_2=.25$

In practice, when Shewhart's design is used the value of h is decided by the manager. Under this condition, the chosen value of h may not be the optimum. Table 5.2.2 shows the effects of eight values on h on the percentage increase in cost for a specified case resulting from use of Shewhart's design instead of the economic design (Figure 5.2.2). We found the larger the difference between the chosen h value and optimum h value, the larger the percentage increase in cost will be. Besides, our model is designed to minimize cost (which is not a specific objective of Shewhart's); it monitors process behavior, and considers the effects of assignable causes. The advantage of the economic design is apparent (Table 5.2.3).

FIGURE 5.2.1
FLOW CHART FOR CALCULATING THE DIFFERENTIAL COST

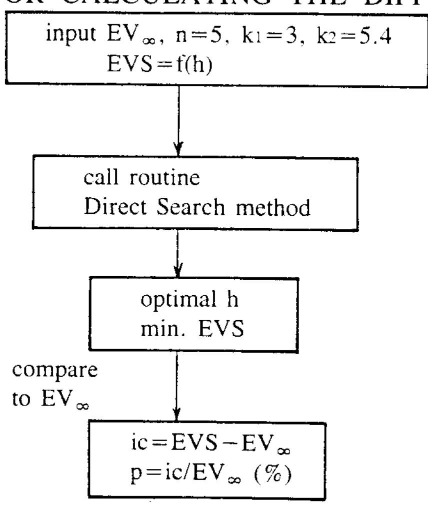


TABLE 5.2.2
PERCENTAGE INCREASE IN COST (p_1, p_2) FOR SHEWHART'S DESIGN AND ECONOMIC DESIGN WITH 8 VALUES OF h FOR A SPECIFICED CASE

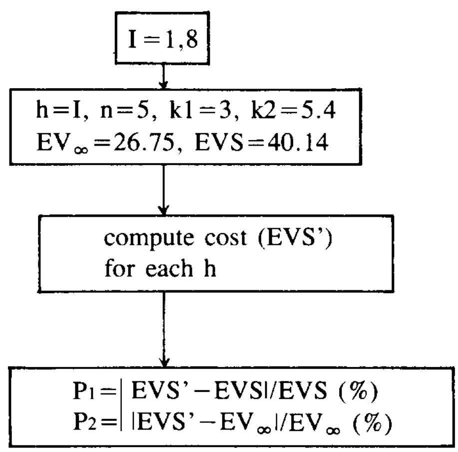
h	1	2	3	4	5	6	7	8
p_1	2.2	0.1	3.1	6.1	8.15	10.15	11.68	13.08
p_2	40.31	39.70	40.72	41.84	42.66	43.45	44.05	44.61

$a=8$, $b=0.5$, $EC_{sr5}=50$, $Mu=.5$, $SD=1.5$, $\check{c}_0=15$, $\check{c}_1=20$, $\check{c}_2=40$, $\check{c}_3=50$, $ET_{sr5}=.5$, $ET_{sr6}=1$, $ET_{sr7}=1.5$, $ET_{sr8}=2$, $EC_{sr6}=80$, $EC_{sr7}=100$, $EC_{sr8}=150$, $r_1=.1$, $r_2=.1$. The optimum h is 1.49 for Shewhart's design. The optimum h is 5.85 for economic design

TABLE 5.2.3
COMPARISON OF ECONOMIC AND SHEWHART'S CONTROL CHARTS

	Economic Design	Shewhart's Design
1. motivation	min. cost	$\alpha = .0027$
2. k_1	always < 4	$k_1 = 3$
3. k_2	always < 5.4	$k_2 = 5.4$
4. n	always < 10	$n = 4$ or 5
5. h	depends on cost	depends on manager
6. monitor process behavior?	yes	no
7. consider effects of assignable causes?	yes	no
8. cost	low	high

FIGURE 5.2.2
FLOW CHART FOR CALCULATING THE PERCENTAGE INCREASE IN COST FOR VARYING VALUES OF h



6. SUMMARY

An expected cost model for a generalized process model in which we use both $\overline{X}$ and R charts to test whether the process is out of control, is formulated. Numerical problems have been solved by using a general Fortran program and the sensitivity of the optimal design to changes in the model inputs can be down. Compared with other models, our model is more reasonable and flexible, since it considers multiple assignable causes, allows multiple occurrence of assignable causes simultaneously, allows the departure of several assignable causes at the same time, and it considers the transition costs for all possible transitions. The production process is expressed as a renewal process and Markov process within cycles. This makes the derivation of expected cycle time and expected cycle cost easier than if we try to extend the Duncan's approach or other approaches to cover the case of multiple assignable causes.

The cost of the optimal design has been compared to the cost of Shewhart's design. We examine in detail the case where two assignable causes lead to changes in the mean, variance or both, and where the k_1, k_2 for Shewhart's design are known. Here, Shewhart's design results in high cost. Our numerical results show that the economic design is preferred. Although the computed procedure is complex, the availability of computer programs and associated optimization schemes make it easier.

The economic design method can be used for a variety of control charts and various process models. It can also be applied to nonnormally distributed process variables.

We could also apply our method in the case of multiple process variables, but there might be some difficult distributional problems.

APPENDICES

A. Calculation of the distribution of the sample range

For sample size n, $X_i \sim N(\mu, \delta^2)$, $i=1,2,\ldots,n$. The range is $R=X_{(n)}-X_{(1)}$.

The p.d.f. of R is

$$f_R(r)=n(n-1)\int_{-\infty}^{\infty} f_x(x)[F_x(x+r)-F_x(x)]^{n-1}f_x(x+r)dx,$$

and the c.d.f. of R is

$$F_R(r)=n\int_{-\infty}^{\infty} f_x(x)[F_x(x+r)-F_x(x)]^{n-1}dx.$$

(see, for example, David 1970)

Since $X_{(n)}-X_{(1)}=(X_{(n)}-\mu)-(X_{(1)}-\mu)$ the distribution of R does not depend on μ. In fact, the distribution of R/δ depends only on n, and not on (μ, δ). Therefore it is sufficient to consider the standardized case where $\mu=0$, $\delta=1$. Writing $W=R/\delta$, the relative range, we work in term of W.

Let random variable $Z \sim N(0,1)$.

In this case, we find the c.d.f. of W is

$$F_w(w) = n\int_{-\infty}^{\infty} f_z(z)[F_z(z+w)-F_z(z)]^{n-1}dz$$

$$= (n/\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp(-z^2/2)[F_z(z+w)-F_z(z)]^{n-1}\, dz, \quad w>0$$

$$\text{Let } Y = Z/\sqrt{2},$$

then $F_w(w)=(n/\sqrt{\pi})\int_{-\infty}^{\infty}\exp(-y^2)[F_z(\sqrt{2}y+w)-F_z(\sqrt{2}y)]^{n-1}dy$ and we approximate this using Hermite polynomials as follows:

$$F_w(w) = (n/\sqrt{\pi}) \sum_{i=1}^{ns} \alpha_i[F_z(\sqrt{2}y+w)-F_z(\sqrt{2}y_i)]^{n-1}.$$

The value of ns, α_i and y_i can be obtained from the table of the zero and weight factors of the first twenty Hermite polynomials in the paper by Sulzer, Zucker and Capuano (1952).

B. Calculation of expected times to absorption, starting from transient states

Conditioning on the first step and using the Markov property, we see that the absorption times T_k satisfy the following equations.

$$
\begin{aligned}
T_1 &\overset{d}{=} h+T_{sr6} && \text{w.p. } p_{16} \\
&\overset{d}{=} h+T_{sr7} && \text{w.p. } p_{17} \\
&\overset{d}{=} h+T_{sr8} && \text{w.p. } p_{18} \\
&\overset{d}{=} h+T_1 && \text{w.p. } p_{11} \\
&\overset{d}{=} h+T_2 && \text{w.p. } p_{12} \\
&\overset{d}{=} h+T_3 && \text{w.p. } p_{13} \\
&\overset{d}{=} h+T_4 && \text{w.p. } p_{14} \\
&\overset{d}{=} h+T_{sr5} && \text{w.p. } p_{15}
\end{aligned}
$$

$$
\begin{aligned}
T_3 &\overset{d}{=} h+T_{sr7} && \text{w.p. } p_{37} \\
&\overset{d}{=} h+T_{sr8} && \text{w.p. } p_{38} \\
&\overset{d}{=} h+T_3 && \text{w.p. } p_{33} \\
&\overset{d}{=} h+T_4 && \text{w.p. } p_{34}
\end{aligned}
$$

$$
\begin{aligned}
T_4 &\overset{d}{=} h+T_{sr8} && \text{w.p. } p_{48} \\
&\overset{d}{=} h+T_4 && \text{w.p. } p_{44}.
\end{aligned}
$$

$$
\begin{aligned}
T_2 &\overset{d}{=} h+T_{sr6} && \text{w.p. } p_{36} \\
&\overset{d}{=} h+T_{sr8} && \text{w.p. } p_{38} \\
&\overset{d}{=} h+T_2 && \text{w.p. } p_{22} \\
&\overset{d}{=} h+T_4 && \text{w.p. } p_{24}
\end{aligned}
$$

(Note that $P\tau\tau'$ are actually functions of h, where τ, $\tau'=1,2,\ldots,8$.)

The expected time for T_k, $k=1,\ldots,4$ can be expressed as follows:

$$
\begin{aligned}
ET_1 &= h+p_{11}ET_1+p_{12}ET_2+p_{13}ET_3+p_{14}ET_4+p_{15}T_{sr5}+p_{16}T_{sr6} \\
&\quad +p_{17}T_{sr7}+p_{18}T_{sr8} \\
ET_2 &= h+p_{22}ET_2+p_{24}ET_4+p_{26}T_{sr6}+p_{28}T_{sr8} \\
ET_3 &= h+p_{33}ET_3+p_{34}ET_4+p_{37}T_{sr7}+p_{38}T_{sr8} \\
ET_4 &= h+p_{44}ET_4+p_{48}T_{sr8}.
\end{aligned}
$$

From this we may verify that ET_1 is as given in (4.1.2.2).

C. Calculation of expected costs to absorption, starting from transient states

Condistioning on the first step and using the Markov property. we see that the absorption cost ϵ_k satisfy the following equations.

$$\epsilon_1 \overset{d}{=} c_{1\tau}(h) \quad \text{w.p. } p_{1\tau},\ \tau=5, 6, 7, 8 \qquad \epsilon_2 \overset{d}{=} c_{2\tau}(h) \quad \text{w.p. } p_{2\tau},\ \tau=6, 8$$
$$\overset{d}{=} c_{11}(h)+\epsilon_1 \quad \text{w.p. } p_{11} \qquad \overset{d}{=} c_{22}(h)+\epsilon_2 \text{ w.p. } p_{22}$$
$$\overset{d}{=} c_{12}(h)+\epsilon_2 \quad \text{w.p. } p_{12} \qquad \overset{d}{=} c_{24}(h)+\epsilon_4 \text{ w.p. } p_{24}$$
$$\overset{d}{=} c_{13}(h)+\epsilon_3 \quad \text{w.p. } p_{13}$$
$$\overset{d}{=} c_{14}(h)+\epsilon_4 \quad \text{w.p. } p_{14}$$

$$\epsilon_3 \overset{d}{=} c_{3\tau}(h) \quad \text{w.p. } p_{3\tau}\ \tau=7,\ 8 \qquad \epsilon_4 \overset{d}{=} c_{48}(h) \quad \text{w.p. } p_{48}$$
$$\overset{d}{=} c_{33}(h)\epsilon_3 \quad \text{w.p. } p_{33} \qquad \overset{d}{=} c_{44}(h)+\epsilon_4 \quad \text{w.p. } p_{44}$$
$$\overset{d}{=} c_{34}(h)\epsilon_4 \quad \text{w.p. } p_{34}$$

(Note that $p_{\tau\tau}$' are actually functions of h, where $\tau,\tau=1,\ 2,\ldots,8$.)

The expected cost for ϵ_k can be obtained

$$E\epsilon_1 = \sum_{\tau=1}^{8} p_{1\tau}c_{1\tau}(h) + \sum_{\tau=1}^{4} p_{1\tau}E\epsilon_\tau$$
$$E\epsilon_2 = p_{22}c_{22}(h)+p_{24}c_{24}(h)+p_{26}c_{26}(h)+p_{28}c_{28}(h)+p_{22}E\epsilon_2+p_{24}E\epsilon_4$$
$$E\epsilon_3 = p_{33}c_{33}(h)+p_{37}c_{37}(h)+p_{38}c_{38}(h)+p_{33}E\epsilon_3+p_{34}E\epsilon_4+p_{34}c_{34}(h)$$
$$E\epsilon_4 = p_{44}c_{44}(h)+p_{48}c_{48}(h)+p_{44}E\epsilon_4.$$

From the above results the $E\epsilon_1$ is as given in (4.2.2.1)

REFERENCES

1. Banerjee, P.K. & Rahim, M.A. (1987), "The Economic Design of Control Charts: A Renewal Theory Approach", Engineering Optimization, vol. 12, pp 63-73.
2. Braverman, J.D. (1981), _Fundamentals of Statistical Quality Control,_ Reston Publishing Co., Inc.
3. Brook, D. and Evans, D.A. (1972), "An approach to the probability distribution of cusum run length", Biomerika, vol. 59, No. 3, pp 539-549.
4. Burr, I.W. (1967), "The Effect of Non-normality on Constants for $\overline{X}$ and R Charts", Industrial Quality Control, Vol. 23, No. 11, May, pp 563-569.
5. Collani, V.E. and Sheil, J. (1989), "An Approach to Controlling Process Variability", Journal of Quality Technology, Vol 21, No. 2, April, pp 87-96.
6. David. H.A. (1970), _Order Statistics_, Wiley, New York.

7. Duncan, A.J. (1956), "The Economic Design of $\overline{X}$-Chart Used to Maintain Current Control of a Process', Journal of the American Statistical Association, Vol. 51, pp 228-242.
8. Duncan, A.J. (1971), "The Economic Design of $\overline{X}$ Charts When There is A Multiplicity of Assignable Causes", Journal of the American Statistical Association, Vol. 66, No. 333, pp 107-121.
9. Gibra, I.N. (1975), "Recent Developments in Control Chart Techniques", Journal of Quality Technology, Vol. 7, No. 4, Oct., pp 183-192.
10. Hill, I.D. (1973), "The Normal Integral", Journal of Applied Statistics, V22, No. 3, pp 422-427.
11. Howard, R.A. (1966), *Dynamic Programming and Markov Processes*, Cambridge, Massachusetts Institute of Technology.
12. Karlin, S. and Taylor, H.M. (1976), *A First Course in Stochastic Processes*, Academic Press Inc, New York.
13. Kemeny, J.G. and Snell, J.L. (1976), *Finite Markov Chains*, Springer Verlag, New York.
14. IMS Library (1989), User's Manual Math./Library Fortran Subroutines, IMSL, Inc.
15. Lorenzen, T.J. and Vance, L.C. (1986), "The Economic Design of Control Charts: A Unified Approach", Technometrics, Vol 28, No 1, pp 3-10.
16. Masuda, Y. and Sumita, U. (1990), "A Multivariate Reward Process Defined On A Semi-Markov Process and Its First Passage Time Distribution", Technical report, Graduate School of Management, University of California, Riverside.
17. Montgomery, D.C. (1980), "The Economic Design of Control Charts: A Review and Literature Survey", Journal of Quality Technology, Vol. 12 , No. 2, April, pp 75-86.
18. Montgomery, D.C. (1985), *Introduction to Statistical Quality Control*, John Willey & Sons, Inc.
19. Montgomery, D.C. and Scorer R.H. (1986), "Economic Models and Process Quality Control", Quality and Reliability Engineering International, Vol. 2., pp 221-228.
20. Panagos, M.R., Russell, H.G. and Montgomery, D.C. (1985), "Economic Design of $\overline{X}$ Control Charts for Two Manufacturing Process Models", Naval Research Logistics Quarterly, Vol. 32, pp 631-646.
21. Pearson, E.S. (1942), "The Probability Integral of the Range in Samples of n Observations from a Normal Population, I Forward Tables", Biometrika, April, pp 301-310.
22. Pitt, H. (1985), "A Modern Strategy for Process Improvement", Quality Progress, May, pp 22-28.
23. Ross, S. M. (1983), *Stochastic Processes*, Academic Press, Inc.
24. Saniga, E.M. (1979), "Joint Economically Optimal Design of $\overline{X}$ and R Control Charts", Journal of Management Science, Vol. 24, No. 4, pp 420-431.
25. Saniga, E.M. & Shirland, L.E. (1977), "Quality Control in Practice: A Survey", Quality Progress, May, pp 30-33.
26. Sulzer, H.E., Zucker, R. and Capuano, R. (1952), "Table of the Zero and Weight Factors of the First Twenty Hermite Polynomials", Journal of Research of the National Bureau of Standards, Vol 48, No. 2, Feb., pp 111-116.
27. Vance, L.C. (1983), "A Bibliography of Statistical Quality Control Chart Techniques, 1970-1980", Journal of Quality Technology, Vol. 15. No. 2, April, pp 59-62