**Economic Statistical Design of Multivariate Control Chart with Variable Sample Sizes**

N. Najmi Sarooghi, M. Bamenimoghadam[[1]](#footnote-1)

**Received:** May 2, 2011 ; **Accepted:** August 15, 2011

**Abstract** Today, quality improvement and cost reduction are key factors for achieving business success, growth and position. One of the primary tools for quality improvement and cost reduction in online activities of statistical process control is control charts. As the need for monitoring several correlated quality characteristics is extensively growing, the use of multivariate control charts becomes more popular. Among these, the control chart is the most popular due to many advantages. In this paper, the variable sampling scheme is studied to increase the power of the fixed sampling rate () control chart to show small shifts. In this study, one of the parameters of control chart, e.g., the sample size, vary between three values, depending on the most recent value. The design of control chart with three variable sample sizes () from the statistical and economic perspective has been investigated, when the shift in the process mean does not occur at the beginning but at some random time in the future. Therefore, the Markov Chain approach used to develop the cost function. Then, the genetic algorithm () is applied to search for the optimal values of the parameters of the control chart. Finally, numerical comparisons among and schemes are made and discussed.

**Keywords** Multivariate *T2* Control Chart with Three Variable Sample Sizes (*3VSS*), Small Shifts, Markov Chain, Genetic Algorithm.

**1 Introduction**

Statistical process control (SPC) is an effective method to improve a firm’s quality and productivity. The primary tool of SPC is the control chart. Recent advances in industrial technology such as modern data-acquisition and on-line computer use during the production have made it common to monitor several correlated quality characteristics simultaneously. This has formed the basis of extensive work performed in the field of multivariate quality control (MQC). Shewhart, famous for the development of the statistical control chart (Shewhart charts) was the first one who recognized the need to consider quality control problems as multivariate in character.

A great deal of work on multivariate statistical control procedures was performed in the 1930’s and in the 1940’s by Hotelling [1]. The traditional practice in applying a control chart to monitor a process is to obtain samples of fixed size, . Which is known as Fixed Ratio Sampling () scheme. However, variable Sample Size procedure () is a scheme that varies the sampling size as a function of prior sample results.

The design of univariate Shewhart charts with adaptive sample sizes was studied by Burr [2], Daudin [3], Prabhu et al. [4], and Costa [5]. In these studies, two sample sizes are used and this procedure shows better power in detecting shifts in the mean. Zimmer et al. [6] presented a three-state adaptive sample size control chart and compared it with both the standard Shewhart control chart and the developed two-state adaptive sample size control chart. They concluded that the three-state procedure is only slightly better than the two-state scheme, and so the two-state scheme is likely adequate in most applications.

In the Multivariate SPC, The -*FRS* control chart shows a good performance to detect large shifts in the process mean. However, in many practical situations it is necessary to be able to detect even moderate shifts in the processmean. In such cases, the statistical efficiency of the chart (in terms of the speed with which process mean shifts are detected) is poor. Aparisi [7] studied the control chart. He considered an adaptive strategy for the subgroup size based on the data trends. He divided the area between the and the origin, into two zones for the use of two different sample sizes. If the current sample value falls in a particular zone, then the corresponding sample size is to be applied for the successive sampling. He showed that the Hotelling’s control chart with scheme, significantly improves the efficiency of the standard Hotelling’s control chart in detecting small changes in the process mean.

When a control chart is used to monitor a process, some design parameters should be determined, e.g., the sample size, the sampling interval between successive samples, and the control limits or critical region of the chart. Since the development of control charts, they were designed from a statistical perspective. In statistical design of control charts, the statistical properties such as type I error, type II error, and average run length (ARL) are checked. Designing a control chart has several economic consequences. This is because sampling, testing and inspection costs, costs of warnings for out of control and eliminating reasonable deviations and costs related to the defective product received by customer on all the select control chart parameters. Thus, it seems logical that control charts are designed from an economic perspective. In economic design of control charts, Typical three class of common costs are considered:

* Sampling and testing costs
* Cost of a warning for out of control and correction process
* The cost of producing defective product

Duncan [8] proposed the first economic model for determining the design parameters for the control charts that minimizes the average cost when a single out-of-control state (assignable cause) exists. Duncan’s cost model includes the cost of sampling and inspection, the cost of defective products, the cost of false alarm, the cost of searching assignable cause, and the cost of process correction. Lorenzen and Vance [9] presented a unified approach for economic design of control charts.

Both statistical and economic designs have their unique strengths and weaknesses. Statistical design builds charts that have a high power and low type I error rate to identify a specific change in process but these designs have cost more than economic design. On the other hand, economic design focuses only on costs and ignores statistical features. Economic statistical design of an economic plan minimizes the average cost per unit time due to some limitations on statistical power and type I error rate in the control chart. These limitations are determined according to the needs of the designer. Sanyga [10] is the founder of the economic statistical design. Faraz et al. [11] studied the economic statistical design of control chart with two variable sample sizes ().

The present study presents the economic statistical design of chart in which the design parameters are determined such that the average cost associated with control chart operation is minimized and statistical properties are met. In the next section, a brief description of the use of control chart with variable sample sizes to maintain current control of a process mean is given. The cost model is then established by modifying the cost model in Lorenzen and Vance [9]. The genetic algorithm (GA) is employed to obtain the optimal values of the test parameters for control chart, and an example is provided to illustrate the solution procedure and compare the results with .

**2 The control chart**

Consider correlated quality character that is supposed to be monitored using the control chart. It is assumed that the distribution of these characteristics is multivariate normal with in-control mean vecter = () and covariance matrix . Thus we have;

When s represents the independent sample vectors, the Hotelling’s multivariate statistic with assuming that and ∑ are known, is calculated as

,

When process parameters, i.e, and ∑ are known, , the action limit is . Here, if , the control chart warns.

If and ∑ are unknown, then they are estimated by the averaged sample mean vector and sample covariance matrix from m initial () random vectors prior to on-line process monitoring, and a statistic defined by

is the approximate statistic for the Hotelling’s multivariate control chart. The action limit for Hotelling’s control chart to monitor future random vectors is given by Alt [12] as

where , , and is the upper α percentage point of distribution with and degrees of freedom if the sample size Moreover, and if the sample size .

If the process is out of control, the chart statistic is distributed as a noncentral distribution with and degrees of freedom with non-centrality parameter , where *d* is the Mahalanobis distance that is used as a measure of process shift in multivariate statistical quality control.

We denote in-control state with , hence we have and so the Average Run Length (ARL) is given by When the process is out of control with shift the ARL is , where *β* is the probability of the type II errors, i.e.

Throughout this article, it is assumed that the process starts in a state of statistical control with mean vector and covariance matrix ∑. The occurrence of assignable cause results in a shift in the process vector mean, which is measured by *d*. The time before the assignable cause occurs has an exponential distribution with parameter . Thus, the mean time that the process remains in state of statistical control is . We use three sample sizes *<*  *<* , and two warning limits . The position of each sample point on the chart establishes the size of the next sample. The Procedure is as follows:

1. If then the next sample is taken with size .
2. If then the next sample is taken with size .
3. If then the next sample is taken with size .

When the process has just started, or after a false alarm, the first sample size and sampling interval is chosen at random.

Control Limit:

If then the chart will show that the process is out of control. However, if then the signal is a false alarm.

**2.1 Performance measure**

Following Costa [5], the speed with which a control chart detects process mean shifts measures its statistical efficiency. When the intervals between samples are fixed, the speed can be measured by the average run length (ARL). However, when the process starts in a state of statistical control and the time before the assignable cause occurs has an exponential distribution with parameter *λ*, it must be measured by modification of ARL, namely the adjusted average time to signal (AATS), which is sometimes called the steady state ATS. The ARL is the expected number of samples before the chart produces a signal, and the AATS is the average time from the process means shift until the chart produces a signal. The average time of the cycle (ATC) is the average time from the start of the production until the first signal after the process shift. If the assignable cause occurs according to an exponential distribution with parameter *λ* then

The memoryless property of the exponential distribution allows the computation of the ATC using the Markov chain approach.

**2.2 Markov chain approach**

In sampling scheme, at each sampling stage, one of the following eight transient states is reached according to the status of the process (in or out of control) and size of the sample:

State 1: and the process is in control ();

State 2: and the process is in control ();

State 3: and the process is in control ();

State 4: and the process is in control ();

State 5: and the process is out of control ();

State 6: and the process is out of control ();

State 7: and the process is out of control ();

State 8: and the process is out of control ().

The control chart produces a signal when . If the current state is 1, 2, 3 or 4, the signal is a false alarm; if the current state is 5, 6, 7 or 8, the signal is a true alarm. The absorbing state, state 8, is reached when the true alarm occurs. The transition probability matrix is given by

where denotes the transition probability that *i* is the prior state and *j* is the current state. In what follows, , , , if sample size and if sample size . Also, will denote the cumulative probability distribution function of a non-central *F* distribution with *p* and degrees of freedom and noncentrality parameter . In calculating ’s consider that statistic in subgroup and statistic in subgroup are independent. Then, s which are conditional probabilities on the prior states are (see, e.g., Cinlar [13])

 ; ; ;

 ; ; ;

 ;

**3 Model assumption**

To simplify the mathematical manipulation and analysis, the following assumptions are made:

1. The process characteristic monitored by the 3VSS chart follows a multivariate normal distribution with mean vector and covariance matrix *∑.*
2. In the start of the process, the process is assumed to be in the safe state; that is, . This state is defined as State 1.
3. Just one assignable cause in process mean occurs and this assignable cause will lead to change the process mean (), of to .
4. The process covariance matrix *∑* remains unchanged.
5. Before the change in the process mean, assuming the process is in control.
6. The occurrences of the assignable cause are exponential with λ rate. In other words, if the process starts from in control state, time of the process remains under control which is a random variable with exponential distribution with mean .
7. The process is not self-correcting.

**4 The cost model**

The cost model developed by Lorenzen and Vance [9] is modified in the present paper to form the objective function for the economic statistical design of the chart. The expected cost per hour derived by Lorenzen and Vance [9] includes the quality cost during production, the cost of false alarm, the cost of searching assignable cause, the cost of process correction, and the cost of sampling and inspection. Lorenzen and Vance [9] in their model, presented the production cycle which is shown in the fig 1.



**Fig. 1** Production cycle in Lorenzen and Vance model

In this connection, the periods of control and out of control with regard to Lorenzen and Vance economic model assumptions, the average cycle time and its expected cost can be calculated as fallows

where

 The average sample size after a warning out of control which is equal to

 The average number of false warnings considering the characteristics of the Markov chain, is calculated as

where

 The average number of inspection items,

where

 The average number of samples taken,

where

 Expected search time when the signal is a false alarm,

 Expected time to discover the assignable cause,

 = Expected time to correct the process,

 Fixed and variable cost of sampling,

 The cost of finding an assignable cause,

 Cost of study false alarm,

 Cost of non-conforming products as long as the process is in control,

 Cost of non-conforming products as long as the process is out of control,

 The expected time for the interpretation sample where is Expected time to interpret a piece.

Economic design of chart is to determine the optimal values of the 7 design parameters such that the average total cost in Equation (3) may be minimized.

**5 Optimization problem using genetic algorithms**

The purpose of the design of the economic statistical control chart is to find the value of control chart parameters, i.e., (), so that the equation 1-3 will be minimized where process parameters, i.e, (λ) and cost parameters, i.e, () are assumed to be known. Therefore, the optimization problem can be written as

.

The solution procedure is carried out using the genetic algorithm to obtain the optimal values of and that minimize the average total cost in Equation (3). With adding constraints and in solving the optimal model, an economic statistical design can be achieved.

The GA, based on the concept of natural genetics, is directed toward a random optimization search technique. The GA solves problems using the approach inspired by the process of Darwinian evolution. The current GA in science and engineering refers to the models introduced and investigated by Holland [14]. In GA, a set of solution of a problem is called a “chromosome”. A chromosome is composed of genes (i.e, features or characters). GA is considered as an appropriate technique for solving the problems of combinatorial optimization and has been successfully applied in many areas to solve optimization problems. The solution procedure for optimization problems using the GA is described as follows:

**Step 1. Initialization.** Thirty initial solutions that satisfy the constraint condition of each test parameter are randomly produced.

**Step 2. Evaluation.** The fitness of each solution is evaluated by calculating the value of fitness function. The fitness function for our example is the cost model in Equation (3).

**Step 3. Selection.** The survivors (i.e., 30 solutions) are selected for the next generation according to the better fitness of chromosomes. In the first generation, the chromosome with the highest cost is replaced by the chromosome with the lowest cost.

**Step 4. Crossover.** A pair of survivors (from the 30 solutions) are selected randomly as the parents used for crossover operations to produce new chromosomes (or children) for the next generation. In this paper, we apply the arithmetical crossover method with crossover rate 0.8. If 30 parents are randomly selected, then there are 60 children. Thus, the population size increases to 90, i.e., 30 parents + 60 children.

**Step 5. Mutation.** Suppose that the mutation rate is 0.1. In this paper, we use a non-uniform method to carry out the mutation operation. Since we have 90 solutions, we can randomly select 9 chromosomes (i.e., 90×0.1=9) to mutate some parameters (or genes).

**Step 6.** Repeat Step 2 to 5 until the stopping criteria is found. In this paper, we use “50 generations” as our stopping criteria.

In this study, we code GA program under the MATLAB (version R2010a) environment, in which real coding is applied.

**6 Comparisons**

In this section we compare and in terms of average cost per hour.

**6.1 Fair comparisons**

For a fair comparison between and designs, two designs have been planned to be under control and have the same average sample size values.

In design, when and ∑ are known, is calculated as

where is identity matrix with 7 ranking and Q is

where

then

In design, when process is in control, is calculated as

By equating equations (\*) and (\*'), is defined as follows:

**6.2 Numerical comparisons**

Table 1 provides optimal control scheme for -3*VSS* along with a comparison with the and -2*VSS* control charts in term of performance. According to Table 1, the -3*VSS* chart has lower values for detecting shifts in comparison to the other procedure and hence, one can expect great savings due to the reduction in the production of non-conforming parts in an out-of-control condition. For example, consider the case where
 *p* = 4*,*  = 2and . In this case, three-state adaptive sample size procedure reduces the average cost from 180.63 $ to 108.99 $.

**Table 1** Numerical camparisons

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.5 | 15.24 | 1.4 | 3.56 | 1 | 5 | 41 | 0.84 | 2.55 | 108.99\* | 109.03 |  | 180.63 |
|  | 1 | 14.39 | 0.34 | 2.25 | 1 | 2 | 15 | 1.22 | 1.82 | 105.13\* | 106.84 |  | 127.14 |
| 2 | 1.5 | 14.36 | 0.81 | 2.24 | 1 | 2 | 9 | 1.24 | 0.82 | 103.95\* | 104.53 |  | 112.64 |
|  | 2 | 15.08 | 0.69 | 1.24 | 1 | 2 | 6 | 1.16 | 0.69 | 103.43\* | 103.96 |  | 105.50 |
|  | 2.5 | 15.37 | 0.49 | 0.5 | 1 | 2 | 4 | 1.06 | 0.62 | 103.13\* | 103.39 |  | 103.72 |
|  | 3 | 16.28 | 5.8 | 11.11 | 2 | 3 | 4 | 1 | 0.74 | 103.41\* | 103.15 |  | 103.03 |

**References**

1. Hotelling, H., (1947). Multivariate quality control. In: Eisenhart, C., Hastay, M., Wallis, W.A. (eds.) Techniques of Statistical Analysis, 111–184. McGraw-Hill, New York.
2. Burr, I. W., (1969). Control charts for measurements with varying sample sizes. J. Qual. Technol, 1, 163–167.
3. Daudin, J. J., (1992). Double sampling charts. J. Qual. Technol, 24, 78–87.
4. Prabhu, S. S., Runger, G. C., Keats, J. B., (1993). chart with adaptive sample sizes. Int. J. Prod. Res, 31, 2895–2909.
5. Costa, A. F. B., (1994). charts with variable sampling size. J. Qual. Technol, 26, 155–163.
6. Zimmer, L. S., Montgomery, D. C., Runger, G. C., (1998). Evaluation of a Three-state adaptive sample size control chart. Int. J. Prod. Res, 36, 733–743.
7. Aparisi, F., (1996). Hotelling’s control chart with adaptive sample sizes. Int. J. Prod. Res, 34(10), 2853–2862.
8. Duncan, A. J., (1956). The economic design of charts used to maintain current control of a process. Journal of American Statistical Association, 51, 228–242.
9. Lorenzen, T. J., Vance, L. C., (1986). The economic design of control charts: a unified approach. Technometrics, 28, 3–10.
10. Saniga, (1991). Economic Statistical Control Charts Design with an Application to and R Charts. Journal of Quality Technology, 21(11), 1271-1282.
11. Faraz, A., Saniga, E., (2009). Economic statistical design of the control chart with two adaptive sample sizes. Accepted to appear in Journal of Statistical Computation and Simulation.
12. Alt, F. B., (1984). Multivariate quality control. In: Kotz, S., Johnson, N.L., Read, C.R. (Eds.), The Encyclopedia of Statistical Sciences. John Wiley, New York, 110–122.
13. Cinlar, E., (1975). Introduction to stochastic Process. Prentice Hall, Englewood Cliff NJ.
14. Holland, J., (1975). Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor.
1. **Corresponding Author. (**🖂**)**

**E-mail:** Bamenimoghadam@gmail.com (M. Bamenimoghadam)

**N. Najmi Sarooghi, M. Bamenimoghadam**

Research Scholar, Department of Statistics, Allameh Tabataba’i University. [↑](#footnote-ref-1)