Integrating the improvement and the control phase of Six Sigma for categorical responses through application of Mahalanobis-Taguchi System (MTS)

Tirthankar Dasgupta

Department of Statistics,
Harvard University,
Cambridge, MA 02138, USA
Fax: 617 495 1712
E-mail: dasgupta@stat.harvard.edu

Abstract: A unified framework for achieving improvement and control of processes with categorical responses during implementation of Six Sigma is proposed. The proposed framework exploits the versatility of MTS as a tool for classification, variable selection and monitoring, and is explained and demonstrated using a simulated example.

Keywords: Six Sigma; DOE; design of experiments; process control; MD; Mahalanobis distance; simulation.

Reference to this paper should be made as follows: Dasgupta, T. (xxxx) ‘Integrating the improvement and the control phase of Six Sigma for categorical responses through application of Mahalanobis-Taguchi System (MTS)’, *Int. J. Industrial and Systems Engineering*, Vol. x, No. x, pp.xxx–xxx.

Biographical notes: Tirthankar Dasgupta is an Assistant Professor of Statistics at Harvard University. He completed his PhD in Industrial Engineering from Georgia Institute of Technology, Atlanta. His research interests include design of experiments (DOE), statistical and engineering process control, applications of statistics in nanotechnology, quality engineering and quality management. He also has eight years of consulting experience in Indian industries and has dual Master’s Degrees in Applied Statistics and Quality, Reliability and Operations Research from the Indian Statistical Institute.

1 Introduction

Six-sigma methodology is being successfully used by companies across the globe to achieve breakthrough product and process improvements (Antony et al., 2005; Raisinghani et al., 2005; Hensley and Dobie, 2005; Lloyd and Holsenback, 2006). Companies implementing six-sigma methodology use a five-phase problem solving process viz., Define, Measure, Analyse, Improve, and Control (DMAIC).

The improvement phase is usually initiated by selecting those product performance characteristics $Y$ that must be improved to achieve the goal. Once this is done, the characteristics are diagnosed to reveal the major sources of variation. Next, the important process variables $X_1, X_2, \ldots, X_p$ are identified through brainstorming and/or analysis of historical data. However, all of these $p$ variables and their interactions
normally do not affect the response significantly. Often, a statistically designed experiment is conducted to identify a subset \( \{X_1, \ldots, X_k\} \) of \( \{X_1, \ldots, X_p\} \), \( k \leq p \) that consists of the ‘active’ factors that have significant effects on the response. Such experiments are often called ‘screening’ experiments, as they screen out active factors from inert ones. The next stage of experimentation is modelling and optimisation. Optimum values and performance specifications are established for each of \( X_1, \ldots, X_k \).

The objective of the control phase is to ensure that the input variables are controlled as per the performance specifications identified in the improvement phase. Statistical Process Control or SPC (Montgomery, 2000) is often recommended at this stage. There are various types of control charts that can be used to keep the input variables in control. When the number of input variables to be monitored is large, a multivariate control chart may be appropriate.

The tasks involved in the improvement and control phase of Six Sigma implementation can be summarised as in Figure 1. The statistical tools that are often used to execute these tasks are also shown in the figure. As is seen from the figure, these tasks involve three key elements – screening (variable selection), optimisation and monitoring.

Therefore, successful implementation of statistical Design of Experiments (DOE) followed by an appropriate monitoring strategy is crucial to the eventual success in Six Sigma implementation.

**Figure 1** Improvement and control phases in Six Sigma implementation

 Conducting successful screening and parameter design experiments is, however, not a trivial task when the product characteristic \( Y \) is categorical (for example, occurrence or non-occurrence of a particular type of defect). Although it is well known that continuous data contain more information than categorical data, time and cost constraints often force engineers to collect categorical data. In situations where the failure probability is very small (or very large) and the pool of variables is large, obtaining effect estimates, screening and optimisation can be very difficult, and often impossible. For such optimisation experiments, an information maximisation approach like the Failure Amplification Method (FAM) proposed by Joseph and Wu (2004) may have to be used. Since the complexity of such specialised optimisation experiments increases almost exponentially with the number of factors, a pragmatic approach would be to...
screen important factors from historical data and conduct an optimisation experiment with a small group of factors.

These days many industries have well-organised databases that contain information on all relevant input variables corresponding to information on each product. In this paper, we propose a unified framework for achieving improvement and control of categorical responses based on such historical data that would ensure execution of all the tasks shown in Figure 1. Such a framework will help identify the most critical input variables, optimise them and design a monitoring scheme for them using a three-step approach. The proposed framework (shown in Figure 2) would use Mahalanobis Taguchi System (MTS) (Taguchi and Jugulum, 2000, 2002; Taguchi et al., 2001; Rai et al., 2008) as a common tool for variable screening and monitoring. The intermediate task of optimisation will be performed by conducting a small experiment with an information maximisation approach.

Figure 2  An overview of the proposed framework

The paper is organised as follows. Section 2 describes the proposed framework and the motivation for using MTS. In Section 3, the proposed approach is explained, demonstrated and compared with other methods using a simulated example. Section 4 consists of summary and conclusions.

2  A unified framework for the improvement and control phases using MTS

As mentioned in the introductory section, the framework will be based on historical data collected on the categorical response \( Y \) and the corresponding input variables \( X_1, X_2, \ldots, X_p \). Suppose the database contains information on \( N \) products in the format shown in Table 1.

Here we consider the simplest situation where the response \( Y \) is either 0 (non-defective) or 1 (defective). It is possible to extend the framework to situations where the number of classes (or categories) is more, e.g., different types of defects. A ‘normal’ group would indicate observations corresponding to non-defective products \((Y = 0)\). Assume that the number of observations in the normal group is \( m \ (<N)\).
The following notations are introduced with respect to Table 1.

- $\bar{X}_i$: Arithmetic mean of all observations in the column corresponding to $X_i$, $i = 1, \ldots, p$
- $s_i$: Standard deviation of all observations in the column corresponding to $X_i$, $i = 1, \ldots, p$
- $Z_{ij} = \frac{X_{ij} - \bar{X}_i}{s_i}$, $j = 1, \ldots, N_i$, $i = 1, \ldots, p$

### Table 1  Historical data on process variables and response

<table>
<thead>
<tr>
<th>Subgroup no.</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$\ldots$</th>
<th>$X_p$</th>
<th>Response ($Y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>$\ldots$</td>
<td>$X_{1p}$</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>$\ldots$</td>
<td>$X_{2p}$</td>
<td>1</td>
</tr>
<tr>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$N$</td>
<td>$X_{N1}$</td>
<td>$X_{N2}$</td>
<td>$\ldots$</td>
<td>$X_{Np}$</td>
<td>0</td>
</tr>
</tbody>
</table>

### 2.1 Improvement phase

As seen from Figure 2, this phase will consist of two key tasks:

- **i** identify a subset $X_1, X_2, \ldots, X_l$ of input variables that are really significant with respect to $Y$ from the data in Table 1
- **ii** determine their optimum combination that would yield the desired $Y$.

Task (i) is basically a variable search problem in a classification framework. Here, we explore the ability of MTS to execute this task. Although Woodall et al. (2003) discussed some limitations of MTS as a tool for classification and variable selection, in comparison to standard statistical classification and variable selection procedures, the reasons behind using it in our framework are as follows.

- MTS is, by definition, a variable selection and monitoring procedure in a classification framework. More traditional statistical variable search methods with binary response (e.g., stepwise logistic regression) would involve examining possible interactions, which may be very large in number. For more number of classes and with multinomial models, this will be even more complicated. Being a completely model-free procedure, MTS may therefore have an advantage.
- The most important motivation lies in the natural extension of MTS to the monitoring stage, where it becomes very much identical to standard SPC procedures. This will be discussed in Section 2.2.

A MTS analysis usually involves four stages (Taguchi and Jugulum, 2002), out of which, the first three will be relevant to the improvement phase of Six sigma. In stage-1 a measurement scale is constructed from a standardised (by subtracting the mean and
Integrating the improvement and the control phase of six sigma

dividing by the standard deviation) ‘normal’ group of \( p \) features (predictor variables) using Mahalanobis Distances (MDs) given by,

\[
MD_j = D_j^2 = \frac{1}{p} Z_j^\top C^{-1} Z_j
\]

where

\( j \): Observation number in the normal group (1 to \( m \))

\( Z_j = (z_{j1}, z_{j2}, \ldots, z_{jm})^\top \): Standardised vector obtained by subtracting the mean and dividing by the standard deviation

\( C^{-1} \): Inverse of the correlation matrix of \( X_1, X_2, \ldots, X_p \) corresponding to the \( m \) data points in the normal group.

In stage-2, larger values of MDs obtained from an abnormal group (defective products) are used for validating the measurement scale developed in stage-1. In stage-3, Orthogonal Arrays (OA) and signal-to-noise ratio values are used for identifying useful input variables from those under study. The simulation study in Section 3 will help in understanding these stages better. Thus, using MTS, we expect to be able to identify the key input variables that have a significant impact on \( Y \).

Having performed task (i), that is, screening of process variables, we move on to task (ii) of the improvement phase, which is optimisation. Since MTS is not designed to perform optimisation, we propose using an optimisation experiment using an information maximisation approach similar to FAM (Joseph and Wu, 2004) to obtain the optimum settings of \( X_1, X_2, \ldots, X_k \). Such an approach requires a factor (of no direct interest to the experimenter) with known effect on the failures, based on physical knowledge of the process. The average values of \( X_1, X_2, \ldots, X_k \) normal and abnormal groups may provide useful information in selecting factor levels in this optimisation experiment. The optimum combination ** * can be obtained by maximising a suitable performance measure \( PM(X_1, X_2, \ldots, X_k) \) with respect to \( X_1, X_2, \ldots, X_k \).

2.2 Control phase

Having identified the right input variables from a pool of potential candidates, the task now is to set up a control scheme with these variables. The traditional procedure is to use a multivariate control chart for controlling \( X_1, X_2, \ldots, X_k \). Multivariate control charts based on Hotelling’s \( T^2 \) statistic (Hotelling, 1947) is one of the most extensively used tools in multivariate process control (Montgomery, 2000). Tracy et al. (1992) developed multivariate control charts for individual observations based on Hotelling’s \( T^2 \) statistic.

Let \( X = (X_{ij}) \) denote the \( N \times k \) matrix of observations on the \( k \) input variables. Then, the Hotelling’s \( T^2 \) statistic for the \( j \)th observation is given by

\[
T_j^2 = (X_j - \bar{X})^\top S^{-1}(X_j - \bar{X}),
\]

where \( \bar{X} \) is the sample mean vector and \( S \) is the covariance matrix. Depending on the phase of control charting, \( \gamma T_j^2 \) will have either a Beta distribution or an \( F \) distribution where \( \gamma \) is a suitable constant (see Tracy et al., 1992). Note that equations (1) and (2) are almost identical except for the fact that in equation (1), the expression for MD
uses vectors scaled by the standard deviation of the normal group, and uses the correlation matrix instead of covariance matrix. This similarity between the Hotelling’s $T^2$ statistic and the MD, discussed in details by Tracy et al. (1997), is one of the compelling reasons that prompt us to use MTS for our framework.

The MTS approach to diagnose future observations is as follows. For each future observation vector $\mathbf{X}_f = (X'_{f_1}, X'_{f_2}, ..., X'_{f_k})^T$, compute

$$\text{MD}_f = \frac{1}{k}(\mathbf{Z}_f^T \mathbf{C}^{-1}\mathbf{Z}_f)$$

where

$$\mathbf{Z}_f = (\mathbf{X}_f - \overline{\mathbf{X}})^T \mathbf{C}^{-1}(\mathbf{X}_f - \overline{\mathbf{X}}),$$

and $\overline{\mathbf{X}}$ and $\mathbf{C}^{-1}$ are the mean vector and inverse of the correlation matrix respectively, obtained from the normal group earlier and used in equation (1). The $j$th future observation is declared abnormal if

$$\text{MD}_f > \tau.$$  

The threshold $\tau$ in equation (5) is given by Taguchi et al. (2001)

$$\tau = \sqrt{\frac{A}{A_0}D},$$

where $A$ is the cost of obtaining a complete examination, $A_0$ is the loss caused by not obtaining a complete examination; and $D$ is the mid-value of the MD of the abnormal group obtained in stage 2 of MTS.

In the current framework, we propose the following modification to the above approach. Recall that at the end of the improvement stage, the optimum setting of the screened process variables was found as $\mathbf{X}^* = (X^*_{1}, X^*_{2}, ..., X^*_{k})^T$. The monitoring methodology will remain exactly the same as described in the previous paragraph, except for the fact that instead of $\overline{\mathbf{X}}$, the target $\mathbf{X}^*$ will be used to standardise the observations. Thus, equation (4) will change to

$$\mathbf{Z}_f = (\mathbf{X}_f - \mathbf{X}^*)^T \mathbf{C}^{-1}(\mathbf{X}_f - \mathbf{X}^*).$$

Note that, this idea is conceptually very similar to control chart implementation, where although the chart is set up around the average of all observations, for future control, it is more prudent to use the target value of the variable under consideration.

### 3 A simulation study

Effectiveness of the proposed framework is examined using a simulated example. The simulation model should consist of a complex input-output relationship where the output is a categorical (binary in the simplest case) variable and the input variables are continuous. Here a complex relationship means, the probability that a binary response $Y$ takes value 0 or 1 should not be a simple or commonly used function (e.g., a logit or a probit model) of the input variables. The number of input variables should be large;
Integrating the improvement and the control phase of six sigma however, most of them should be inert, and a few of them should actually affect the response. Also, as is the case in real-life process control, all input variables will be controlled tightly within specific tolerances around their set points; however, they would go out of control with certain probabilities. Such out-of-control situations may result in defectives or non-conforming items or abnormal situations.

The simulation model consists of one latent output variable denoted by \( W \) and ten process variables denoted by \( X_1, X_2, \ldots, X_{10} \), among which only the first four are \( X_1, X_2, X_3, \) and \( X_4 \) are active, and the rest are inert. We assume the following functional relationship between \( W \) and \( \mathbf{X} = (X_1, X_2, \ldots, X_{10}) \):

\[
W(X_1, X_2, \ldots, X_{10}) = \sum_{j=1}^{10} \left( \sum_{i=1}^{4} (X_j - A_{ij})^2 + c_j \right), \quad 1 \leq X_j \leq 10
\]  

(8)

where \( A_{ij} \) and \( c_j \) are the \((i, j)\)th and \(j\)th elements of the following matrix and vector, respectively.

\[
\mathbf{A} = \begin{bmatrix}
4 & 1 & 8 & 6 & 3 & 2 & 5 \\
4 & 1 & 8 & 6 & 7 & 9 & 5 \\
4 & 1 & 8 & 6 & 3 & 2 & 3 \\
4 & 1 & 8 & 6 & 7 & 9 & 3
\end{bmatrix} \quad \text{and} \quad \mathbf{c} = \frac{1}{10} (1, 2, 2, 4, 4, 6, 3)^T.
\]

The function given by equation (8) is a standard global optimisation test function known as Shekel function (Shekel, 1971), which has several local optima and a single global maximum at \((X_1, X_2, X_3, X_4) = (4, 4, 4, 4)\). The maximum value is 10.4029. The response variable for our study will be a binary random variable \( Y \) defined as

\[
Y = \begin{cases} 
0 & \text{if } W \geq 7.7 \\
1 & \text{otherwise}
\end{cases}
\]  

(9)

To relate the simulation to a real-life problem, let us assume that we have an unobservable Larger-The-Better (LTB) latent variable \( W \), which determines the state of each product. If \( W \) falls below a certain threshold, a defective product is generated. Brainstorming sessions conducted in the improvement phase of six-sigma implementation have identified \( X_1, X_2, X_3, \ldots, X_{10} \) as potentially important, and based on past experience, engineers try to control each of these variables around four. We assume that ‘a controlled process’ means the following.

\[ X_i \sim \text{TN}(4, \sigma^2), \quad 3.9 \leq X_i \leq 4.1, \quad i = 1, \ldots, 10. \]

Here \( \text{TN}(\mu, \sigma^2) \) denotes a truncated normal distribution with mean \( \mu \) and variance \( \sigma^2 \), truncated within limits \([a, b]\). Simulation of 200 observations shows that the in-control distribution of \( W \) is more or less symmetric (see Figure 3). An in-control state would ensure a negligible probability of a defective item, and the process can be called a ‘Six Sigma’ process. It is natural that the mean of \( W \) will be less that 10.4029 due to fluctuations in \( \mathbf{X} \).
Let us now define ‘out of control’ situations. Assume that an out-of-control situation with respect to any $X_i$ means the following distribution:

$$X_i \sim \text{TN}(3.9, \sigma^2), \quad 3.7 \leq X_i \leq 4.1, \quad i = 1, \ldots, 10.$$ 

Assume that all the input variables are not equally susceptible to control problems. Let $\pi_i$ denote the probability that the $i$th input variable would go out of control. This means,

$$X_i \sim \begin{cases} 
\text{TN}(4, \sigma^2), & 3.9 \leq X_i \leq 4.1, \quad \text{with probability } 1 - \pi_i \\
\text{TN}(3.9, \sigma^2), & 3.7 \leq X_i \leq 4.1, \quad \text{with probability } \pi_i 
\end{cases}$$

It is assumed that $\pi = (0.10, 0.15, 0, 0.05, 0.05, 0.07, 0, 0.10, 0, 0.20)$. Thus, $X_1, X_2, X_8,$ and $X_{10}$ are most likely to go out of control, whereas $X_3, X_7,$ and $X_9$ would never go out of control. This means $X_1, X_2,$ and $X_4$ are the candidate input variables that should necessarily be included in the eventual monitoring scheme.

Before we present the simulation results, let us clearly distinguish ‘out of control’ and ‘abnormal situations’. Out of control situations represent deviations of the input variables from their in-control distributions. Abnormal conditions pertain to the response variable $Y$ and represent a defective item, i.e., $Y = 1$. As discussed earlier, and seen from Figure 3, an in-control process will result in a negligible percentage of defectives. Loosely speaking, if the process is in control, there will be no defectives, and hence no abnormal condition. The converse is, however, not true. An out-of-control situation would not necessarily give rise to an abnormal condition; although the probability that $Y = 1$ would increase as the active input variables deviate more and more from their stable state.

### 3.1 Using MTS for variable selection in the improvement phase: a single simulation

Suppose the historical data consists of 200 data points, i.e., a $200 \times 10$ matrix of observations on the ten input variables, and a $200 \times 1$ vector of observations on the latent variable $W$. Figure 4 shows the histogram and run chart obtained from a single simulation where the dispersion parameter $\sigma$ is taken to be 0.1. In the simulated data set, 13 values...
of \( W \) are found to fall below 7.7. These correspond to \( Y = 1 \) and are labelled as abnormal observations. Thus, the sizes of the normal and abnormal group are 187 and 13, respectively.

**Figure 4** Histogram and run chart of 200 simulated observations on \( W \)

For constructing and validating the ‘measurement scale’, data on \( X_1, X_2, \ldots, X_{10} \) corresponding to 187 in-control situations representing ‘normal’ group are used. After normalising each process variable, MD values are obtained using equation (1). The measurement scale thus obtained is validated with the help of MDs calculated by normalising 13 cases representing ‘abnormal’ group using mean and standard deviations from normal group. The results are shown in Figure 5.

**Figure 5** Scaled Mahalanobis Distances based on data on all ten process variables

From Figure 5, it can be observed that although some amount of separation between normal and abnormal situations is achieved, there is also a substantial overlap between the two groups. Table 2 provides a summary of the validation of the measurement scale in terms of average, maximum, and minimum MDs.

**Table 2** Summary of validation of the measurement scale

<table>
<thead>
<tr>
<th></th>
<th>Average MD</th>
<th>Minimum MD</th>
<th>Maximum MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>With process variables based on normal data</td>
<td>0.9947</td>
<td>0.2309</td>
<td>3.0591</td>
</tr>
<tr>
<td>With process variables based on abnormal data</td>
<td>2.4654</td>
<td>1.3361</td>
<td>3.5667</td>
</tr>
</tbody>
</table>
Identification of useful process variables.

The discrimination between the two groups may be further improved by identifying useful variables out of the ten process variables. The useful variables are identified using OA and Signal-to-Noise (S/N) ratios. The process variables $X_1$, $X_2$, …, $X_{10}$ are studied at two-levels each by assigning them to the first ten columns of $L_{12}(2^{11})$ orthogonal array (Taguchi, 1987). Note that a full-factorial design would have required $2^{10} = 1024$ experimental combinations. The $L_{12}(2^{11})$ orthogonal array is given in Table 3.

<table>
<thead>
<tr>
<th>SL</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
<th>$X_6$</th>
<th>$X_7$</th>
<th>$X_8$</th>
<th>$X_9$</th>
<th>$X_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

In Table 3 level-1 indicates the presence and level-2 indicates the absence of the process variable. For each experimental combination, 13 MD values corresponding to the 13 abnormal points are obtained using equation (1). Since larger MD values help to obtain a better separation between process variables corresponding to normal and abnormal situations, a LTB S/N ratio is obtained from the MDs. Let $MD_j$ denote the MD value obtained from the $j$th abnormal observation corresponding to the $i$th experimental run. The S/N ratio for the $i$th ($i = 1, \ldots, 12$) experimental run with $t$ ($t = 13$) abnormal data points can be obtained as (Taguchi and Jugulum, 2002, 2000),

$$
\eta_i = -10\log \left[ \frac{1}{t} \sum_{j=1}^{t} \frac{1}{MD_{ij}} \right].
$$

(10)

The S/N ratios obtained for the 12 experimental combinations are used for calculating average at level 1 and level 2 for each of the ten process variables. Subsequently, gain in S/N ratio values is obtained by taking difference of the two average values as,

$$
\text{Gain} = (\text{Average S/N Ratio})_{\text{level-1}} - (\text{Average S/N Ratio})_{\text{level-2}}.
$$

(11)

A positive gain for a process variable indicates its usefulness and vice-versa. The gain in average S/N ratio values obtained for all the ten process variables under study is shown in Figure 5.
Integrating the improvement and the control phase of six sigma

From Figure 6 it can be observed that process variables $X_1$, $X_2$, $X_4$, and $X_{10}$ provide positive gain values in the S/N ratios. The gain corresponding to $X_{10}$ is, however, very marginal. These four process variables are therefore identified as useful in discriminating the abnormal situations in terms of the process variables. The effectiveness of these four process variables identified as useful in providing a better discrimination as compared to all ten process variables is depicted in Figure 7 and summarised in Table 4.

**Figure 6** Average gain in Signal-to-Noise ratio for all ten process variables (see online version for colours)

**Figure 7** Scaled Mahalanobis Distances based on $X_1$, $X_2$, $X_4$, and $X_{10}$

<table>
<thead>
<tr>
<th>Variables included</th>
<th>$S/N$ ratio</th>
<th>Average MD of abnormal group</th>
</tr>
</thead>
<tbody>
<tr>
<td>All ten variables</td>
<td>3.440</td>
<td>2.465</td>
</tr>
<tr>
<td>Only $X_1$, $X_2$, $X_4$, and $X_{10}$</td>
<td>6.339</td>
<td>4.824</td>
</tr>
</tbody>
</table>

Table 4 Comparison after confirmatory trial

It can be observed from Figure 7 that the overlap between MDs corresponding to normal and abnormal situations is lower than that achieved with all ten variables. Table 4 also points to $S/N$ ratio values and average $MD$ values that are almost two times based only on four process variables identified as useful compared to all ten variables.
What if experimental design was used for variable screening?

To compare the performance of MTS with respect to variable screening with a usual experimental design approach, a set of simulation experiments were conducted with the 10 factors $X_1, X_2, \ldots, X_{10}$, each at three levels $4 - \Delta, 4, 4 + \Delta$. A 27-run orthogonal array that can accommodate 13 three-level factors was used (Wu and Hamada, 2000, Chapter 7), and the ten factors were assigned to the first ten columns of the array. For each experimental combination, assuming an internal noise of $\sigma$ associated with each input variable, the latent response variable $W$ and the corresponding observed binary response $Y$ was computed. Three replications of the basic experiment were conducted. Logistic regression was used to identify the significant factors. The results of nine such screening experiments with three different values of $\Delta$: 0.10, 0.15 and 0.20 and three different values of $\sigma$: 0 (no noise), 0.05 and 0.1 are shown in Table 5. None of the nine experiments are able to correctly identify the significant variables.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>$\sigma$</th>
<th>Significant variables identified (5% level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0</td>
<td>$X_6, X_9$</td>
</tr>
<tr>
<td>0.10</td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>0.10</td>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>0.15</td>
<td>0.05</td>
<td>$X_4, X_5, X_7, X_8$</td>
</tr>
<tr>
<td>0.15</td>
<td>0.05</td>
<td>None</td>
</tr>
<tr>
<td>0.15</td>
<td>0.05</td>
<td>None</td>
</tr>
<tr>
<td>0.20</td>
<td>0.10</td>
<td>None</td>
</tr>
<tr>
<td>0.20</td>
<td>0.10</td>
<td>None</td>
</tr>
<tr>
<td>0.20</td>
<td>0.10</td>
<td>None</td>
</tr>
</tbody>
</table>

3.2 The optimisation stage

We do not include the optimisation step in the simulation, because including an amplification factor requires explicit assumptions about the physics of the process. In this situation, where we have a single latent variable $W$ that determines the failure, and there is only one type of failure, a loss function for higher the better characteristics and single latent variable modeling approach (Joseph and Wu, 2004) may be used. The experimental design employed may be a fixed design or a sequential design. Let us assume that the optimisation experiment further screens out $X_1, X_2, \text{ and } X_4$ as the important factors and declares $X_{10}$ as insignificant and identifies the optimal value of each variable as $X_i^* = 4$.

3.3 Using MTS for monitoring in the control phase

Based on the results of the optimisation experiment, the task is to develop a monitoring scheme that would ensure that input variables $X_1, X_2, \text{ and } X_4$ are in control around the target value of four. Dropping variable $X_{10}$ results in even better separation. The average MD values for the normal and abnormal group are obtained as 0.991 and 5.926 respectively. We now generate 200 future observations from the process, and
Integrating the improvement and the control phase of six sigma

examine each of them using the monitoring scheme described in Section 2.2, using $X^2 = 4$, but the same correlation matrix $C$ computed and used in Section 3.1. An abnormal situation is detected if the computed value of MD exceeds the threshold $\tau = \sqrt{A/A_0} 5.926$, given by equation (6). Among the new 200 observations, 14 are found to belong to the abnormal group (that is $W < 7.7$). The performance of the monitoring scheme versus the ratio $A : A_0$ is shown in Table 6.

**Table 6**  Performance of the monitoring scheme

<table>
<thead>
<tr>
<th>$A : A_0$</th>
<th>Abnormal conditions detected</th>
<th>Number of abnormal situations undetected (false negatives)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correctly</td>
<td>Incorrectly (false positives)</td>
</tr>
<tr>
<td>1 : 1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>1 : 2</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>1 : 3</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>1 : 4</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>1 : 5</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>1 : 6</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>

As expected, the false positives (declaring a normal condition as abnormal) increase as the $A : A_0$ ratio decreases, that is, the loss due to non-detection of failure relative to inspection cost increases. Decrease in $A : A_0$ ratio means lowering of the threshold line separating abnormals from normals (see Figure 8), which results in a higher instance of false positives. In Figure 8, the solid dots represent the abnormal outcomes. Note that, there are four such outcomes that are indistinguishable from the normal outcomes which are represented by small white dots.

**Figure 8**  Separating future observations with MTS

3.4 **Effect of noise and comparison with ordinary and stepwise logistic regression**

In the simulation done so far, it was assumed that the variables are stochastic, and the variation around their set values represents noise in the model. This type of noise
(variation around set value) is often referred to as ‘internal noise’. The level of internal noise in our simulation model is determined by the parameter $\sigma$ associated with the distribution of each input variable $X$. Now we consider the following updated model that includes an external noise component $\varepsilon$.

$$W(X_1, X_2, ..., X_{10}) = \sum_{j=1}^{10} \left( \sum_{i=1}^{4} (X_j - A_i)^2 + c_j \right) + \varepsilon, \quad \text{where } \varepsilon \sim N(0, \sigma^2). \quad (12)$$

The external noise can be physically interpreted as the combined impact of many uncontrollable variables apart from the ten input variables. Larger the external noise, the more difficult it will be to correctly identify the significant inputs. We now examine the ability of MTS to correctly identify the influential variables at various levels of noise. We also employ two other popular statistical methods – ordinary logistic regression and stepwise logistic regression and compare their performances with MTS with respect to variable selection ability. For this comparison a $4 \times 3$ full factorial experiment was conducted with four values of $\sigma_\varepsilon (0, 0.5, 1.0, 1.5)$ and three values of $\sigma (0.05, 0.10, 0.15)$. For each of the 12 combinations, one set of 200 observations was generated, and influential variables were selected using the three methods. For ordinary logistic regression, the variables that were significant at 5% level of significance were chosen. For stepwise logistic regression, the Akaike Information Criterion (AIC) was used (Akaike, 1974). The results are summarised in Table 7, from which we observe the following:

- In most of the cases, MTS selected supersets of the set of variables that are actually significant, whereas ordinary and stepwise logistic regression selected subsets of the correct set. Of course, for the latter methods, this depends on the choice of the level of significance.
- When the internal noise level is small ($\sigma = 0.05$), logistic regression performs poorly due to insufficient variation in values of the predictor variables. In comparison, the performance of MTS seems to be more robust to internal noise.
- Performances of all the three methods, expectedly, worsen, as external noise level $\sigma_\varepsilon$ increases.

### Table 7  Influential variables identified at different noise levels

<table>
<thead>
<tr>
<th>External noise ($\sigma_\varepsilon$)</th>
<th>Internal noise ($\sigma$)</th>
<th>Number of defectives (out of 200)</th>
<th>Ordinary logistic regression</th>
<th>Stepwise logistic regression</th>
<th>MTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>5</td>
<td>2, 4</td>
<td>2, 4, 8, 10</td>
<td>1, 2, 4, 8, 9</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>13</td>
<td>1, 2</td>
<td>1, 2, 4</td>
<td>1, 2, 4, 10</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>22</td>
<td>1, 2, 4, 5</td>
<td>1, 2, 4, 5</td>
<td>1, 2, 3, 4, 6</td>
</tr>
<tr>
<td>0.5</td>
<td>0.05</td>
<td>7</td>
<td>1, 2, 7</td>
<td>1, 2, 3, 4, 6, 7, 10</td>
<td>1, 2, 4, 6</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>14</td>
<td>1, 2, 4</td>
<td>1, 2, 4, 6</td>
<td>1, 2, 3, 4, 7</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>25</td>
<td>1, 2, 4</td>
<td>1, 2, 4</td>
<td>1, 2, 3, 4, 9</td>
</tr>
<tr>
<td>1.0</td>
<td>0.05</td>
<td>15</td>
<td>None</td>
<td>5, 7</td>
<td>1, 2, 3, 4, 10</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>27</td>
<td>2, 4</td>
<td>2, 4, 5, 9</td>
<td>1, 2, 3, 4, 7, 8</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>34</td>
<td>2</td>
<td>2, 4, 8</td>
<td>1, 2, 3, 4, 7, 9</td>
</tr>
</tbody>
</table>
Integrating the improvement and the control phase of six sigma

Table 7 Influential variables identified at different noise levels (continued)

<table>
<thead>
<tr>
<th>External noise ($\sigma_e$)</th>
<th>Internal noise ($\sigma$)</th>
<th>Number of defectives (out of 200)</th>
<th>Variables identified by using Ordinary logistic regression</th>
<th>Stepwise logistic regression</th>
<th>MTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.05</td>
<td>30</td>
<td>None</td>
<td>6, 7</td>
<td>1, 2, 3, 4, 8, 10</td>
</tr>
<tr>
<td>0.10</td>
<td>44</td>
<td>2, 4</td>
<td>2, 4</td>
<td>1, 2, 3, 4, 7, 9, 10</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td>52</td>
<td>None</td>
<td>1, 2, 4, 8</td>
<td>2, 3, 4, 6, 7, 9, 10</td>
<td></td>
</tr>
</tbody>
</table>

4 Summary and conclusions

In this paper, we propose the use of MTS in the integration of improvement and control phases of Six Sigma projects that involve categorical responses. Screening, optimisation and monitoring are three major tasks associated with the improvement and control phases of six sigma. When the output characteristic is categorical with a very small or very large failure probability, conducting screening and optimisation experiments is not easy. Two key reasons for developing an integrated framework based on MTS are:

- MTS deals with variable screening in a classification framework and process monitoring
- Monitoring a process with MTS is similar in principle to standard statistical methods for multivariate control.

The proposed framework uses MTS for screening and monitoring, and advanced experimentation techniques for optimisation. The simulation study shows that MTS does a reasonable job in situations where the underlying statistical model is fairly complex and non-standard, and the size of the abnormal group is small compared to the size of the normal group. It has to be remembered, though, that MTS is not a statistical approach, as it does not allow users to make statistical inferences. Future research may consist of:

- Studying the statistical behaviour of MTS in a probabilistic framework to allow statistical inference
- Seeking better variable search methods that boost the screening ability of MTS, based on the recommendations of Woodall et al. (2003).

Acknowledgement

I would like to thank the referee for his helpful comments that were instrumental in improving the quality of this paper.

References


