SHEWHART CONTROL CHARTS TO DETECT MEAN AND STANDARD DEVIATION SHIFTS BASED ON GROUPED DATA

STEFAN H. STEINER
Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

AND

P. LEE GEYER AND GEORGE O. WESOLOWSKY
McMaster University, Hamilton, Ontario L8S 4M4, Canada

SUMMARY

A Shewhart control chart is proposed based on gauging theoretically continuous observations into multiple groups. This chart is designed to monitor the process mean and standard deviation for deviations from stability. By assuming an underlying normal distribution, we derive the optimal grouping criterion that maximizes the expected statistical information available in a sample. Control charts based on grouped observations are superior to standard control charts based on variables, such as $\bar{X}$ and R charts, when the quality characteristic is difficult or expensive to measure precisely, but economical to gauge.

KEY WORDS: likelihood ratio tests; Shewhart control charts; step gauges

1. INTRODUCTION

In industry, the use of step-gauges and similar devices is widespread. A step-gauge classifies continuous observations into groups rather than measure them precisely. A step-gauge usually consists of a number of pins of different diameter; see Figure 1. Gauging articles is often quicker, easier and therefore cheaper than measuring them exactly. Similarly, exact measurements occasionally require costly skilled personnel and sophisticated instruments.[1] In general, if the step-gauge is properly designed, grouped data is an excellent alternative to exact measurement since the small loss in statistical efficiency is often more than offset by savings in the cost of measurement.

At present, most control charts used for grouped data represent each unit by either a group endpoint or group midpoint. These 'variable measurements' are then used in charts designed for true variables measurements, such as an $\bar{X}$ chart. These *ad hoc* solutions have a number of shortcomings. Most importantly, they introduce a bias into the calculations of the sample mean. For example, using a group's right endpoint will consistently overestimate the mean. Using the midpoint of each interval as the 'representative value' results in less bias, but the bias is not eliminated unless the underlying distribution is uniform. The amount of bias depends

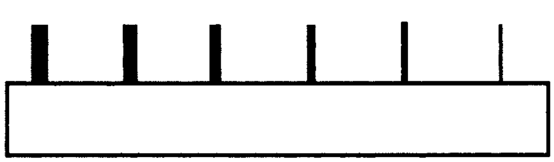

Figure 1. An idealized six step-gauge

on the underlying process distribution and the gauge design, and may be substantial. Consequently, the bias could adversely affect the workings of any control charts that are designed for use with exact measurement. In addition, for either the endpoint or the midpoint approaches we must consider the end-group problem. As the end-group extends to either positive or negative infinity, the group endpoint may not be finite, and the group midpoint will not be finite. As a result, to obtain more reliable results, the fact that the data is actually grouped rather than produced by exact measurement should be taken into account in a more direct manner.

Few control charting techniques have been developed specifically for use with grouped data. The first contributions were made by Tippett,[2] and later, in more detail by Stevens.[3] They proposed two simple control charts for simultaneously monitoring the mean and standard deviation of a normal distribution using a two-step gauge that classifies observations into one of three groups. Stevens[3] found the optimal design of the gauge limits by maximizing the expected Fisher's information in a single observation. However, it is not straightforward to extend Stevens' methodology to more than three groups, and it is difficult to derive an operating characteristic (OC) curve for his charts. Others have pointed out that, when the underlying distribution of the variable of interest is known, savings in inspection costs can be realized by using an attributes plan with compressed (also called artificial) specification limits. Compressed limit sampling plans are discussed by Mace,[4] Ott and Mundel[5] and Duncan.[6]

More recently, advocates of pre-control, some-

CCC 0748-8017/96/050345-09

Received 25 April 1995
Revised 11 March 1996

times called stoplight control, have proposed the use of three classes to monitor the statistical control of a process. See, for example, Traver,[7] Salvia,[8] Shainin and Shainin[9] and Ermer and Roepke.[10] Commonly, the classification is based on specification limits. One class consists of the central half of the tolerance range, another is based on the remaining tolerance range, and the third consists of measurements beyond tolerance limits. Pre-control techniques are appealing since they are very simple to teach and implement. However, the classification and signalling criteria used are arbitrary, and do not reflect process capability and are thus not recommended for process control.

The first to design control charts that are applicable in the general multiple group case were Steiner, Geyer and Wesolowsky.[11] They proposed using the likelihood ratio of multinomial probabilities to design control charts to detect one-sided shifts in a process mean. Essentially these charts are acceptance control charts, and are optimal since by the Neyman–Pearson lemma the optimal partitioning of the accept/reject region is based on the likelihood ratio of the two specific alternatives.[12]

This article extends their work to the two-sided case and thus to 'Shewhart' type control charts. The methodology presented is very generally applicable; however, we illustrate the method by monitoring both the mean and standard deviation of a process whose output follows a normal distribution. In Section 2 two possible approaches to extend the one-sided chart to a two-sided chart are presented and contrasted. Solutions for large sample sizes based on the central limit theorem (CLT) are also developed. As most Shewhart control charts are used with small sample sizes, Section 3 discusses adapting the design when the CLT solution is not adequate. In Section 4 the issue of step-gauge design is addressed and it is shown that the proposed approaches are competitive in terms of efficiency with the traditional variables based approach. Finally, an example is presented from our work in the manufacture of metal fasteners to illustrate the design and implementation of step-gauge Shewhart control charts.

2. A k-STEP GAUGE SHEWHART CONTROL CHART

Our objective is to design a control chart to detect significant process shifts from $\theta = \theta_0$, the nominal value, using grouped data. Thus, we consider the hypothesis test:

$$H_0: \theta = \theta_0$$
$$H_1: \theta = \theta_1 \text{ or } \theta = \theta_{-1}$$

where θ_1 and θ_{-1} are significant departures in the upward and downward directions, respectively. Thus, designing Shewhart charts involves testing a composite two-sided hypothesis test. As a result, no single uniformly most powerful test exists.[12]

This design strategy of specifying an alternative hypothesis is consistent with the objective of Shewhart charts to detect any shift from stability. Specifying the alternative hypothesis simply allows us to ensure that our control chart will have sufficient power to detect important parameter shifts. The same design strategy can be used for standard $\bar{X}$ charts. For example, for an $N(\mu,\sigma^2)$ process, choosing the in-control average run length (ARL) equal to 370·4, and the out-of-control ARL equal to 1·56, where the alternative mean values are either $\mu_1 = \mu_0 + 1{\cdot}5\sigma$ or $\mu_{-1} = \mu_0 - 1{\cdot}5\sigma$ results in a traditional $\bar{X}$ chart with three sigma limits and a sample of size five.

The typical Shewhart control chart design problem is to find the appropriate sample size n and control limits so that certain conditions on the average run length (ARL) are satisfied. The ARL is the expected number of samples required before the chart signals. Under H_0 large ARL values are desirable, whereas under H_1 we would like short ARLs so that we are quickly informed that the process has shifted. As a result, our control chart is designed to have at least an ARL of ARL_0^* under H_0 under H_0 and at most ARL_1^* under H_1. Traditional three-sigma-limit Shewhart charts are designed so that $ARL_0^* = 370$. Equivalently these conditions can be expressed in terms of error rates for the hypothesis test. We want false alarms with probability less than α^*, and out-of-control signals with probability at least $1-\beta^*$ whenever the process mean shifts either up to θ_1 or down to θ_{-1}. Since the number of samples until a signal follows a geometric distribution we have $ARL_0^* = 1/\alpha^*$ and $ARL_1^* = 1/(1-\beta^*)$.

We propose two possible approaches. In Section 2.1 we show that the composite test can be considered equivalent to two simple one-sided hypothesis tests for which an optimal testing strategy exists. In Section 2.2, we present an approch based on a single hypothesis test. This second approach, although not optimal, is appealing since it is easier to implement requiring the maintenance of only a single chart. A comparison of these two approaches is made in Section 2.3.

We first provide definitions of some necessary notation. Suppose that the quality characteristic of interest is a random variable X that has a probability function $\phi(x;\mu,\sigma)$, where μ and σ are the location parameter and scale parameter, respectively. In this article we focus on monitoring the mean and standard deviation of the normal distribution, but the aproach is more general and can be applied to monitor any parameter of any distribution. Thus, for example, $\phi(x;\mu,\sigma) = [1/\sigma(2\pi)] \exp[-(x-\mu)^2/2\sigma^2]$ for the normal distribution. A k-step gauge classifies observations into one of $k+1$ distinct intervals. Let the k interval endpoints be denoted by t_j, $j = 1,2,\ldots,k$, and define $t_0 = -\infty$ and $t_{k+1} = \infty$. Then, the probability that an observation is classified as belonging to group j is given by

$$\pi_j(\theta) = \int_{t_{j-1}}^{t_j} \phi(x;\theta)\mathrm{d}x, \quad j = 1, \dots, k+1 \quad (1)$$

where θ is the parameter of interest. Let $\mathbf{Q}$ be a $(k+1)$ column vector whose jth element, Q_j, denotes the number of observations in a sample of size n that are classified into the jth group. Then, the likelihood of any hypothesis about θ, given a sample of size n, is equal to

$$LR(\theta|\mathbf{Q}) = c\prod_{j=1}^{k+1} \pi_j(\theta)^{Q_j} \quad (2)$$

where the constant of proportionality c is arbitrary and $\Sigma_{j=1}^{k} Q_j = n$.

2.1 *Two sets of weights approach*

The optimal solution uses the composite likelihood ratio (CLR) as a test statistic. From (2), and defining $\Omega = \{\theta_1, \theta_{-1}\}$, the log of the composite likelihood ratio can be written as

$$\mathrm{CLR} = \max_{\theta\in\Omega} \ln\left[\frac{L(\theta|\mathbf{Q})}{L(\theta_0|\mathbf{Q})}\right] = \max_{\theta\in\Omega}\left[\sum_{j=1}^{k+1} Q_j \ln\left(\frac{\pi(\theta)}{\pi(\theta_0)}\right)\right] \quad (3)$$

Define z_i^+ and z_i^- as random variables (weights) equal to $\ln(\pi_j(\theta_1)/\pi_j(\theta_0))$ and $\ln(\pi_j(\theta_{-1})/\pi_j(\theta_0))$ respectively when the i^{th} observation belongs to the jth group. Then, since there are only two alternatives in Ω, the test statistic CLR is equivalent (under a rescaling) to max $(\bar{z}^+, \bar{z}^-)$, where

$$\bar{z}^+ = \sum_{i=1}^{n} z_i^+/n \text{ and } \bar{z}^- = \sum_{i=1}^{n} z_i^-/n$$

To attain ARLs of ARL_0^* *and* ARL_1^* under H_0 and H_1 respectively, with the test statistic max $(\bar{z}^+, \bar{z}^-)$, we must find sample size n and control limit λ such that

$$\Pr(\max(\bar{z}^+,\bar{z}^-) > \lambda|\theta = \theta_0) \le 1/ARL^*_0 \quad (4)$$

$$\Pr(\max(\bar{z}^+,\bar{z}^-) > \lambda|\theta = \theta_1) \ge 1/ARL^*_1 \quad (5)$$

$$\Pr(\max(\bar{z}^+,\bar{z}^-) > \lambda|\theta = \theta_{-1}) \ge 1/ARL^*_1 \quad (6)$$

Since $\bar{z}^+$ and $\bar{z}^-$ are highly negatively correlated $\Pr(\bar{z}+\&\bar{z}^- > \lambda) \cong 0$. Thus, $\Pr(\max(\bar{z}^+,\bar{z}^-)>\lambda)$ can be very closely approximated by $\Pr(\bar{z}^+ > \lambda)$ + $\Pr(\bar{z}^- > \lambda)$. As a result, we may consider the two average weights separately and define two separate control limits λ^+ and λ^-. $\Pr(\bar{z}^+ > \lambda^+|\mu_{-1})$ and mality of the average weights, the equality forms of the equations (5) and (6) can be solved to give the control limits in terms of the unknown sample size. We obtain

$$\lambda^+(n) = -\frac{\sigma_{z^+}(\theta_1)\Phi^{-1}(1/ARL^*_1)}{\sqrt{n}} + \mu_{z^+}(\theta_1) \quad (7)$$

$$\lambda^-(n) = \frac{\sigma_{z^-}(\theta_{-1})\Phi^{-1}(1/ARL^*_1)}{\sqrt{n}} + \mu_{z^-}(\theta_{-1}) \quad (8)$$

where $\mu_{z^+}(\theta), \sigma_{z^+}(\theta)$ and $\mu_{z^-}(\theta), \sigma_{z^-}(\theta)$ are the mean and standard deviation of the weights z_i^+ and z_i^-, respectively, when the true parameter value is θ, and $\Phi^{-1}(\ )$ denotes the inverse of the cumulative distribution function of the standard normal distribution. Substituting equations (7) and (8) into equation (4), the design problem becomes: find the sample size n such that

$$Pr(\bar{z}^+ > \lambda^+(n)|\theta_0) + Pr(\bar{z}^- > \lambda^-(n)|\theta_0) \le 1/ARL^*_0 \quad (9)$$

Again assuming that the central limit theorem is applicable, a solution can be found by starting with $n = 1$ and incrementing n until the inequality (9) is satisfied.

2.2 *One set of weights approach*

Another approach is to derive a control chart for monitoring shifts in a process mean is to compare the likelihood of θ_{-1} against that of θ_1. This technique is quite commonly used in the sequential analysis of multi-hypothesis problems.[13] The advantage of considering only the two alternative parameter values and not the null hypothesis is that it leads to only one set of weights. The likelihood ratio of this hypothesis test for a gauged sample of size n, is given by

$$LR(\mathbf{Q}) = \prod_{j=1}^{k+1}\left(\frac{\pi_j(\theta_1)}{\pi_j(\theta_{-1})}\right)^{Q_j} \quad (10)$$

where

$$\sum_{j=1}^{k+1} Q_j = n$$

Letting λ_U and λ_L denote the critical values (the upper and lower control limits of the plotted statistic, respectively), then the chart signals that the process mean has shifted upwards whenever $LR(\mathbf{Q}) > e^{n\lambda_u}$, and signals a downward process mean shift whenever $LR(\mathbf{Q}) < e^{n\lambda_L}$. Equivalently, the process is deemed in control as long as

$$n\lambda_L \le \sum_{j=1}^{k+1} Q_j \ln\left(\frac{\pi_j(\theta_1)}{\pi_j(\theta_{-1})}\right) \le n\lambda_U \quad (11)$$

Let ARL_0^* and ARL_1^* be the desired ARLs under the null and alternative hypotheses respectively, and let w_i be a random variable such that

$$w_i = \ln\left(\frac{\pi_j(\theta_1)}{\pi_j(\theta_{-1})}\right) \qquad (12)$$

when the ith observation falls into the jth group. Note that the weights w_i are different from the weights z_i^+ and z_i^-. The chart signals if $\bar{w} < \lambda_L$ or $\bar{w} > \lambda_U$ *where* $\bar{w} = \sum_{i=1}^{n} w_i/n$, and we wish to find n, λ_U and λ_L so that

$$1/ARL^*_0 \geq \Pr(\bar{w} > \lambda_U|\theta = \theta_0) + \Pr(\bar{w} < \lambda_L|\theta = \theta_0) \quad (13)$$

$$1/\text{ARL}^*_1 \leq \Pr(\bar{w} > \lambda_U|\theta = \theta_1) + \Pr(\bar{w} < \lambda_L|\theta = \theta_1) \quad (14)$$

and

$$1/ARL^*_1 \leq \Pr(\bar{w} > \lambda_U|\theta = \theta_{-1}) + \Pr(\bar{w} < \lambda_L|\theta = \theta_{-1}) \quad (15)$$

For large sample sizes $\bar{w}$ will have an approximate normal distribution with mean $\mu_w(\theta)$, and variance $\sigma_w^2(\theta)/n$, where

$$\mu_w(\theta) = \sum_{j=1}^{k+1} \pi_j(\theta) \ln\left(\frac{\pi_j(\theta_1)}{\pi_j(\theta_{-1})}\right) \qquad (16)$$

$$\sigma_w^2(\theta) = \sum_{j=1}^{k+1} \pi_j(\theta) \ln\left(\frac{\pi_j(\theta_1)}{\pi_j(\theta_{-1})}\right)^2 - \mu_w^2(\theta) \qquad (17)$$

The solution methodology is similar to that applied in Section 2.1. The second term of the right-hand side of equation (14) and the first term of the right-hand side of equation (15) are insignificant relative to the other terms in the expressions, and thus we can solve for the control limits in terms of the sample size:

$$\lambda_U = -\sigma_w(\theta_1)r(n) + \mu_w(\theta_1)$$
$$\lambda_L = \sigma_w(\theta_{-1})r(n) + \mu_w(\theta_{-1})$$

Substituting the results into (13) we obtain the inequality

$$1/ARL^*_0 \geq \Pr(\bar{w} > -\sigma_w(\theta_1)r(n) + \mu_w(\theta_1)|\theta_0) \quad (18)$$
$$+ \Pr(\bar{w} < \sigma_w(\theta_{-1})r(n) + \mu_w(\theta_{-1})|\theta_0$$

where

$$r(n) = \frac{\Phi^{-1}(1/ARL^*_1)}{\sqrt{n}}$$

As in the two sets of weights case, the appropriate sample size can be found assuming that $\bar{w}$ is normally distributed with mean and standard deviation as defined by (16) and (17), and incrementing n from unity until inequality (18) holds.

2.3 *Comparison of the one and two sets of weights approaches*

The two approaches are ideally compared by determining the best ARL_1^* value that can be achieved with the approach for a given fixed ARL_0^* level. However, owing to the discreteness inherent in the problem, making this comparison is difficult. For any given **t** and n values there are a limited number of possible ARL_0^* values that can be achieved. Unfortunately, this is a disadvantage of any method that relies on discrete data. As a result, a more satisfactory comparison can be made on the basis of the theoretical sample size that each requires (based on the given CLT solutions) to achieve the same ARLs. In this manner, we avoid the discreteness in the sample size. For an underlying N(0,1) process, Figure 2 plots the theoretical sample sizes required to detect various mean shifts. Figure 2 shows that the one set of weights approach typically requires only marginally larger sample sizes than the two sets of weights approach.

The percentage increase in sample size required for the suboptimal one set of weights approach grows as μ_1 increases. However, gauging units is typically very easy and inexpensive; thus, we recommend the one set of weights approach since the decrease in efficiency is small and only one chart needs to be maintained.

3. SOLUTION FOR SMALL SAMPLE SIZES

Since sample sizes must be integers, and the random variables z_i^+, z_i^- and w_i are discrete, the ARLs will not be exactly as desired. For large sample sizes, the solution methodology, presented in the previous section, based on the CLT, gives good results. For small sample sizes, however, the effect of the problem discreteness may be significant, and we may

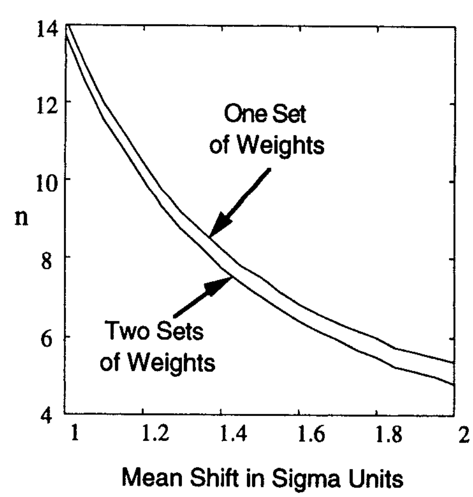

Figure 2. Sample size required versus mean shift for the two proposed approaches: $\mathbf{t} = [-1,0,1]$, $ARL_0^* = 1000$, $ARL_1^* = 2$, $\mu_0 = 0$, $\mu_1 = -\mu_{-1}$

need to consider the distribution of $\bar{z}^+$ and $\bar{z}^+$ or $\bar{w}$. This analysis is important since samples of size five are typically for Shewhart charts. For the purposes of illustration, in this section, we will restrict attention to $\bar{w}$, although a similar analysis is possible for $\bar{z}^+$ and $\bar{z}^-$.

The question of what sample sizes are large enough to give good solutions via the CLT solution is a difficult one. However, for any problem with a moderate sample size and number of groups, we can find the true ARL levels through enumeration. To determine the actual ARLs for the one set of weights approach given a sample of size n, a critical value λ and gauge steps vector **t**, we must determine all partitions of the n observations into $(k+1)$ groups, without regard to order, where either $LR(\mu|\mathbf{Q}) > e^{n\lambda_U}$, or $LR(\mu|\mathbf{Q}) < e^{n\lambda_L}$. Let $\{S\}$ represent the set of all such partitions that cause the chart to signal. The probability that the chart signals when the true value of the parameter is θ, is given by

$$\Pr(LR(\theta|\mathbf{Q}) > e^{n\lambda_U} \text{ or } LR(\theta|\mathbf{Q}) < e^{n\lambda_L}) \quad (19)$$

$$= \sum_{\{S\}} \frac{n!}{Q_1!Q_2!\ldots Q_{k+1}!} \prod_{j=1}^{k+1} \pi_j(\theta)^{Q_j}$$

Then, the ARL at the null is $ARL_0 = 1/\Pr[LR(\mu_0|\mathbf{Q}) > e^{n\lambda_U}$ or $LR(\mu_0|\mathbf{Q}) < e^{n\lambda_L}]$, and the ARLs when θ equals θ_1 and θ_{-1} are $ARL_{1\text{up}} = 1/\Pr[LR(\mu_1|\mathbf{Q}) > e^{n\lambda_U}]$ and $ARL_{1\text{down}} = 1/\Pr[LR(\mu_{-1}|\mathbf{Q}) < e^{n\lambda_L}]$, respectively (this assumes that the probability of signalling a downward mean shift when a significant upward mean shift has taken place, and vice versa, is negligible). The number of partitions in $\{S\}$ grows exponentially as the number of groups increases and polynomially as the sample size increases. For moderatly large n and k, we can enumerate all possible partitions of a sample to find the set $\{S\}$. Given $\{S\}$, we can determine the true ARL_0 and ARL_1 (maximum of $ARL_{1\text{up}}$ and $ARL_{1\text{down}}$) levels, or determine the control limits that result in the best possible ARL_1 level given the constraint $ARL_0^* > ARL_0$.

For small sample sizes, the distribution of $\bar{w}$ is typically skewed toward the null value. As a result, the CLT approximation will usually overestimate tail probabilities, and yield control limits that are larger in absolute value than necessary. This problem is most pronounced when trying to detect large mean shifts with few groups and small samples. When the sample is of moderate size we can correct the CLT solution by deriving the set $\{S\}$ to determine precisely the required n and control limits. The following algorithm will determine an appropriate control chart design. In Step 3 we use the fact that the distribution of $\bar{w}$ is discrete, and thus there are only a finite number of values for the control limits that will change the average run length properties.

1. Determine the smallest integer n that satisfies
2. For the current choice of n, determine the set $\{S\}$.
3. Use $\{S\}$ to compute the best control limits λ_U and λ_L that ensure that $ARL_0 > ARL_0^*$.
4. Given n and λ_U and λ_L, use $\{S\}$ to determine $ARL_{1\text{up}}$ and $ARL_{1\text{down}}$.
5. Incrementally increase n and repeat procedure starting at Step 2 until $ARL_{1\text{up}}$ and $ARL_{1\text{down}}$ are both less than ARL_1^*.

The above algorithm has been coded in MATLAB® and is available from the authors. On a Macintosh SE/30, determining the ARL_0 and ARL_1 values in Table I for each iteration took less than 75 seconds. To illustrate, we present an example with three cases where we wish to detect mean shifts in a normal process. Assume that we desired a control chart to detect a shift of 2 sigma units in a standard normal mean with $ARL_0^* = 1000$, using a three-step gauge defined by $\mathbf{t} = (-1,0,1)$. From (1) and (9) we have group weights $\mathbf{w} = (-6{\cdot}43, -1{\cdot}85, 1{\cdot}85, 6{\cdot}43)$ since, for example,

$$w_1 = \ln(0{\cdot}00135/0{\cdot}84135) \text{ and}$$
$$0{\cdot}00135 = \int_{-\infty}^{-3} \phi(x)\mathrm{d}x.$$

In this example, the gauge limits are symmetric about $\mu_0 = 0$ and $\mu_1 = \mu_{-1}$, thus we have $\lambda_U = -\lambda_L$ and we will only consider the upper control limit. Table I presents the results for $ARL_1^* = 2$, 1·05 and 1·005. For each proposed design we determine the actual ARL_0 and ARL_1 using enumeration. We first solve equation (18) to determine the CLT solution for the required sample size and control limit. If the resulting actual error rates are not sufficiently small (i.e. if $ARL_0 < ARL_0^*$ or $ARL_1 > ARL_1^*$) we use the enumeration algorithm presented above to find the optimal location for the control limit, and increment n until the desired error rates are obtained. Note that in all cases the ARL_0 level was sufficiently large. However, when $ARL_1^* = 1{\cdot}005$, although the CLT solution suggested $n = 11$, due to skewness in the distribution of $\bar{w}$ we actually required $n = 12$ to obtain $ARL_0 > 1000$ and $ARL_1 < 1{\cdot}005$.

4. OPTIMAL STEP-GAUGE DESIGN FOR SHEWHART CHARTS

There are two decisions to be made in specifying the grouping criteria: we must decide how many groups

Table I. ARLs for different designs where $ARL_0^* = 1000$

ARL_1^*	Solution type	n	λ_U	ARL_0	ARL_1
2	CLT	6	5·61	2252	1·43
1·05	CLT	9	4·53	5076	1·1
1·05	enumeration	9	4·25	2273	1·03
1·005	CLT	11	4·08	2584	1·015
1·005	enumeration	11	3·9	2179	1·007
1·005	enumeration	12	3·6	1124	1·003

to use, and how these groups are distinguished. In general, a k-step gauge classifies units into $(k+1)$ groups. Clearly, as more groups are used, more information becomes available about the parameters of the underlying distribution. The limiting case occurs when the variable is measured to infinite precision. However, given that a k-step gauge is to be used, not all gauge limits will provide the same amount of information about the parameters of the underlying distribution. It is not intuitively clear how best to place the k-steps of the gauge. We consider these questions in more detail for the mean and standard deviation of a normal process.

Shewhart control charts are designed to detect whenever the process is no longer stable; thus we are interested in detecting small upward or downward parameter shifts. As a result, the problem of determining the best gauge limits for control may be thought of as equivalent to designing the step gauge limits to obtain the best estimate of θ in the neighbourhood of the standard (in-control) value.[3] This is also justified because the process should operate in an in-control state for the majority of the time. Designing the step-gauge in such a manner is accomplished by maximizing the expected Fisher information of the gauged sample at the standard value. Fisher information is defined as the square of the derivative of the log-likelihood. However, the log-likelihood for a sample of size n will be formed by the sum of n log-likelihoods, each of identical expectation.[14] As a result, it is equivalent for our purposes to consider the expected information in a single observation. The expected information in a sample of size one at θ, $E(I(\theta))$, may be obtained by conditioning on the group into which the observation is classified. In particular, if $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{k+1}$ denote the unit vectors of length $(k+1)$, then

$$E(I(\theta)) = \sum_{j=1}^{k+1} I(\theta|\mathbf{e}_j)\pi_j(\theta) = \sum_{j=1}^{k+1} \frac{1}{\pi_j(\theta)} \left(\frac{\mathrm{d}\pi_j(\theta)}{\mathrm{d}\theta}\right)^2 \quad (20)$$

Assuming a normal distribution, and using (1), it can be shown that the expected information for a single observation about the mean μ and standard deviation σ are respectively

$$E(I(\mu)) = \frac{\phi(t_1;\mu)^2}{\pi_1(\mu)} + \sum_{j=2}^{k} \frac{(\phi(t_{j-1};\mu) - \phi(t_j;\mu))^2}{\pi_j(\mu)} + \frac{\phi(t_k;\mu)^2}{\pi_{k+1}(\mu)} \quad (21)$$

$$E(I(\sigma)) = \frac{t_1^2\phi(t_1;\sigma)^2}{\sigma\pi_1(\sigma)} + \sum_{j=2}^{k} \frac{(t_{j-1}\phi(t_{j-1};\sigma) - t_j\phi(t_j;\sigma))^2}{\sigma\pi_j(\sigma)} + \frac{t_k^2\phi(t_k;\sigma)^2}{\sigma\pi_{k+1}(\sigma)} \quad (22)$$

Without loss of generality, we assume that 'in-control' $\mu = 0$ and $\sigma = 1$. The function $E(I(\theta))$ is in general not concave. If we fix the first $(k-1)$ gauge limits and allow the kth to become arbitrarily large, the expected information asymptotically approaches a minimum. Extensive experimentation has indicated, however, that the expected information function is unimodal, at least in the cases represented by (21) and (22). As a result, the gauge limits that maximize the expected information can be found efficiently using a quasi-Newton method such as the BFGS algorithm or the Fletcher Reeves algorithm if the gradient is determined.[15]

Table II shows the optimal gauge limits for detecting mean shifts away from $\mu = 0$ and the efficiency of that group gauge design relative to exact measurement. The efficiency of the gauged sample is computed by taking the ratio of information available in the gauged sample to the information in an exactly measured sample. Table III gives similar results for detecting standard deviation shifts away from $\sigma = 1$.

Note that in the σ shift case for an even number of groups, the middle step gauge placement has arbitrary sign, and is not zero as in the mean shift case. This is because a group limit placed at $t = 0$ (the mean) will provide no additional information about σ.

Examining Tables II and III we notice that the optimal group limits to detect σ shifts are much more spread out than those that are best for monitoring shifts in μ. This suggests that a compromise gauge design be considered in cases where we wish to monitor for mean and standard deviation shifts simultaneously. There are many possible ways to derive a reasonable compromise. We proposed maximizing the weighted sum of efficiency ratings for mean and standard deviation estimation. In other words, maximize

$$\mathrm{Eff}(\mu,\sigma;d) = d\,\mathrm{Eff}(\mu) + (1-d)\,\mathrm{Eff}(\sigma) \quad (23)$$

where d is the weight, $\mathrm{Eff}(\mu)$ is the efficiency of mean estimation, and $\mathrm{Eff}(\sigma)$ is the efficiency of standard deviation estimation. Often in practice the detection of mean shifts is given priority. Table IV gives the optimal compromise group designs for various number of group limits when the mean estimation is given a greater weight ($d = 0{\cdot}7$).

As shown in Tables II, III and IV the use of more than two or three groups significantly increases the efficiency of an observation. For example, near the target value, more information about μ is available in ten five-group optimally gauged observations, than is available in nine exact measurements. Clearly, if exact measurement is uneconomical a properly designed gauge may be an excellent alternative.

Tables II, III and IV present optimal designs for a standard normal in-control process, but they may also be used to calculate the optimal gauge limits for any normal process. If the vector t represents

Table II. Optimal group limits to detect mean shifts assuming that when the process in 'in control' $\mu = 0 (\sigma = 1)$

Number of groups	Efficiency	t_1	t_2	t_3	t_4	t_5	t_6
2	0·6366	0·0					
3	0·8098	−0·6120	0·6120				
4	0·8825	−0·9817	0·0	0·9817			
5	0·9201	−1·244	−0·3824	0·3824	1·244		
6	0·9420	−1·4468	−0·6589	0·0	0·6589	1·4468	
7	0·9560	−1·6108	−0·8744	−0·2803	0·2803	0·8744	1·6108

Table III. Optimal group limits to detect sigma shifts assuming that when the process is 'in control' $\sigma = 1$ $(\mu = 0)$

Number of groups	Efficiency	t_1	t_2	t_3	t_4	t_5	t_6
2	0·3042	±1·5758					
3	0·6522	−1·4825	1·4825				
4	0·7074	−1·4520	1·1855	2·0249			
4	0·7074	−2·0249	−1·1855	1·4520			
5	0·8244	−1·9956	−1·1401	1·1401	1·9956		
6	0·8588	−1·9827	−1·1193	0·9837	1·6189	2·3267	
6	0·8588	−2·3269	−1·6190	−0·9837	1·1190	1·9821	
7	0·8943	−2·3130	−1·6002	−0·9558	0·9558	1·6002	2·3130

Table IV. Optimal group limits to detect mean and sigma shifts $d = 0·7$, assuming that if the process is 'in control' $\mu = 0$ and $\sigma = 1$

Number of groups	Eff(μ)	Eff(σ)	t_1	t_2	t_3	t_4	t_5	t_6
2	0·6366	0·0	0·0					
3	0·7822	0·4664	−0·8487	0·8487				
4	0·8685	0·6262	−1·2529	0·0	1·2529			
5	0·9082	0·7384	−1·5500	−0·5295	0·5295	1·5500		
6	0·9333	0·8039	−1·7703	−0·8768	0·0	0·8768	1·7703	
7	0·9489	0·8486	−1·9481	−1·1366	−0·3889	0·3889	1·1366	1·9481

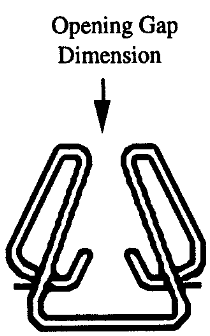

Figure 3. Robotics clamp

the optimal gauge limits for an N(0,1) process then $t' = t\sigma + \mu$ will be the corresponding optimal limits for an N(μ,σ) process.

6. METAL FASTENERS EXAMPLE

This example is suggested by our study of the manufacture of metal fasteners in a progressive die environment, and is an extension of the example presented by Steiner *et al.*[11] The opening gap dimension of a metal clamp, called a robotics clamp, was considered critical; see Figure 3.

However, obtaining exact measurements of the ... difficult and expensive. The metal used in the clamp is fairly pliable, and as a result, using calipers distorts the opening gap dimension. Another alternative, an optical measuring device, is expensive and not practical for on-line quality monitoring. As a result, the only economical alternative on the shop floor is to use step gauges, where clamps are classified into different groups based on the smallest pin that the clamp's opening gap does not fall through.

The process mean is currently stable, producing clamps with an open gap dimension of 54·2 thousandths of an inch and standard deviation of 1·3. We wish to monitor for stability in both the mean and standard deviation of the width of opening gap. It was desired to creat a control chart that has $ARL_0^* = 200$ and $ARL_1^* = 2$ or better when detecting mean shifts of one standard deviation unit in an upward or downward direction. We also wish to simultaneously test the hypothesis $H_0 : \sigma = \sigma_0 = 1·3$ versus $H_1 : \sigma = 2\sigma_0$ or $\sigma = 0·5\sigma_0$. For the standard deviation chart we will use the sample size suggested by our design of the mean shift chart, and set the control limits so that $ARL_0^* \geq 200$. In this case, the step gauge employed has gauge limits (or pins) with diameters of 53, 54, and 55 thousandths of an inch. As a result, using this standard gauge, units are

Table V. Step-gauge design for metal fasteners example

Interval	Group	μ-weight	σ-weight
$(-\infty,53]$	I	$-2{\cdot}97$	$2{\cdot}30$
$(53,54]$	II	$-1{\cdot}03$	$-0{\cdot}86$
$(54,55]$	III	$0{\cdot}44$	$-1{\cdot}22$
$(55,\infty)$	IV	$2{\cdot}50$	$1{\cdot}24$

classified into four intervals with corresponding weights calculated from equation (9); see Table V.

Using this step-gauge the design procedure presented in Section 2.2 suggests a sample size of $n = 12$ and upper and lower control limits $\lambda_U = 1{\cdot}55$, and $\lambda_L = -1{\cdot}66$ for the mean chart. Enumeration shows that this design results in $ARL_0 = 384{\cdot}6$, $ARL_{1\mathrm{up}} = 1{\cdot}92$, and $ARL_{1\mathrm{down}} = 1{\cdot}96$ for the mean shift which satisfies our requirement. For the standard deviation chart, choosing $\lambda_U = 1{\cdot}30$, and $\lambda_L = -0{\cdot}87$ results in $ARL_0 = 200$, $ARL_{1\mathrm{up}} = 6{\cdot}96$, and $ARL_{1\mathrm{down}} = 3{\cdot}45$. The resulting control charting procedure for each sample can be summarized as follows:

1. Take a sample of 12 units from the process.
2. Assign each of the 12 units a μ-weight and a σ-weight as given in Table V.
3. Calculate and plot the average μ and σ weight for the sample.
4. Search for an assignable cause if the average μ-weight plots above $1{\cdot}55$ or below $-1{\cdot}66$ or the average σ-weight lies outside the range $(-0{\cdot}87,\ 1{\cdot}3)$.

Figure 4 shows the results of this procedure. The first ten samples were taken from the process at 30 minute interval, and suggest an 'in control' process. The next ten samples are simulated average weights and show the result of an 'out of control' situation on the control chart. The simulated 'out of control' values are based on a $N(52{\cdot}9,\ 1{\cdot}3)$ process, i.e. the mean has shifted to $-\mu_1$. At this 'out-of-control' setting our mean shift chart has a 51 per cent $(1 - 1/ARL_{1\mathrm{down}})$ change of detecting this magnitude of shift in one sample. Notice that the standard deviation chart also begins to signal when the mean shifts down, since many observations in the lowest (or highest) group may also be indicative of increased variance. In this example, with the mean at $-\mu_1$, there is an 18 per cent chance for each sample that the standard deviation chart will signal. However, a quick check of the actual sample would quickly confirm that the actual shift must have been in the mean. The efficiency relative to the classical variables based approach of these grouped data control charts is $\mathrm{Eff}(\mu) = 0{\cdot}8689$ and $\mathrm{Eff}(\sigma) = 0{\cdot}4203$.

Note that considering only the mean shift problem using the optimal gauge limits for this problem (from Table II) allows a reduction in sample size. With gauge limits at 52·9, 54·2, and 55·5 thousands of inch, a sample of size 11 with control limits at 1·66 and −1·66 we obtain ARLs of $ARL_0 = 294{\cdot}1$ and $ARL_{1\mathrm{up}} = ARL_{1\mathrm{down}} = 1{\cdot}89$. A variables-based control chart ($\bar{X}$ chart) designed for the same propose would require a sample size of 9. Our proposed three-step gauge control chart requires a larger sample size to match the power of an $\bar{X}$ chart, but due to the savings in measurement costs is an excellent alternative in this situation.

7. CONCLUSION

A multi-step gauge control chart has been presented to monitor the stability of the parameter of interest when observations are classified into one of several groups. These proposed charts are very easy to use and interpret on the production floor. We show how a control chart can be designed to satisfy any specified ARL criteria. The implementation of this approach in a progressive die operation shows that the k-step gauge control chart is a viable alternative to other control charts, approaching the variables-based control charts in efficiency. The optimal gauge design has been derived for monitoring the mean

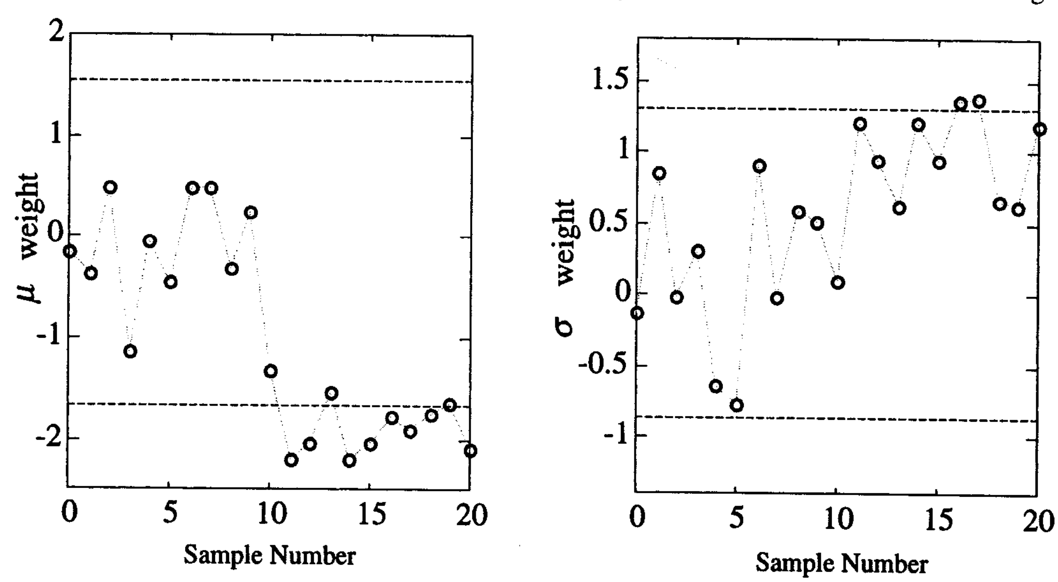

Figure 4. Sample control charts

and standard deviation of a normal process by maximizing the expected Fisher information at the target value. These charts are applicable in situations where variables measurements are expensive or impossible, and yet classifying units in groups is economical.

ACKNOWLEDGEMENTS

This research was supported, in part, by the Natural Sciences and Engineering Research Council of Canada. Our sincere thanks also go to two anonymous referees of this journal whose comments were very helpful in improving this paper. MATLAB® is a registered trademark of The MathWorks Inc.

REFERENCES

1. S. P. Ladany and Z. Sinuary-Stern, 'Curtailment of artificial attribute sampling plans', *International Journal of Production Research*, **23**, (5), 847–856 (1985).
2. L. H. C. Tippett, 'The efficient use of gauges in quality control', *Engineer*, **177**, 481–483 (1944).
3. W. L. Stevens, 'Control by gauging' with discussion, *Journal of the Royal Statistical Society: Series B*, **10**, (1), 54–108 (1948).
4. A. E. Mace, 'The use of limit gages in process control, *Industrial Quality Control*, January 1952, pp. 24–31.
5. E. R. Ott and A. B. Mundel, 'Narrow-limit gaging', *Industrial Quality Control*, March 1954, pp. 21–28.
6. A. J. Duncan, *Quality Control and Industrial Statistics*, Fifth edn, Richard D. Irwin, Homewood, IL, 1986.
7. R. W. Traver, 'Pre-control: a good alternative to $\bar{X}$-R charts', *Quality Progress*, September 1985, pp. 11–14.
8. A. A. Salvia, 'Stoplight control', *Quality Progress*, September 1988, pp. 39–42.
9. D. Shainin and P. Shainin, 'Pre-control versus $\bar{X}$ & R charting: continuous or immediate quality improvement?', *Quality Engineering*, **1**, (4), 419–429 1989.
10. D. S. Ermer and J. R. Roepke, 'An analytical analysis of pre-control', *ASQC Quality Congress Transactions*, Milwaukee, 1991, pp. 522–527.
11. S. H. Steiner, P. L. Geyer and G. O. Wesolowksy, 'Control charts based on grouped data', *International Journal of Production Research*, **32**, (1), 75–91 (1994).
12. M. Kendall and D. Stuart, *The Advanced Theory of Statistics, Volume 2: Inference and Relationship*, Fourth edn, Charles Griffin and Company Ltd., London, 1979.
13. B. K. Ghosh and P. K. Sen (eds), *Handbook of Sequential Analysis*, Marcel Dekker Inc., New York, 1991.
14. A. W. F. Edwards, *Likelihood*, Cambridge University Press, Cambridge, 1972.
15. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1988.

Author's biographies:

Stefan H. Steiner is an assistant professor in the Department of Statistics and Actuarial Science at the University of Waterloo. He obtained his Ph.D. in Management Science from McMaster University and M.Sc. and BMATH degrees from the University of British Columbia and the University of Waterloo respectively. His research interests included industrial statistics and applications of statistics in operations research.

P. Lee Geyer is a Ph.D. student at McMaster University in the Management Science/Systems Area of the Faculty of Business. He also consults extensively in the area of health care funding and utilization analysis. His research interests include process control and statistical modeling.

George O. Wesolowsky is a professor of Management Science in the Faculty of Business at McMaster University in Hamilton, Ontario, Canada. His research interests include quality control and location models.