DEPARTMENT OF MATHEMATICS
AND
COMPUTER SCIENCE

TECHNICAL REPORT SERIES

Comparison of Standard and Individual Limits Phase I Shewhart $\overline{X}$, R, and S Charts

by

Charles W. Champ
Department of Mathematics and Computer Science
Georgia Southern University, Statesboro, GA 30460-8093

and

Shou-Peng Chou
15 Lane 22, Han-Hsi Street, Taichung, Taiwan, R.O.C.

Number 2002-006
Submitted: March 20, 2002

Comparison of Standard and Individual Limits Phase I Shewhart $\overline{X}$, R, and S Charts

Charles W. Champ*

Georgia Southern University, Statesboro, GA 30460-8093

Shou-Peng Chou

15 Lane 22, Han-Hsi Street, Taichung, Taiwan, R.O.C.

Abstract

Two Phase I $\overline{X}$ charts are examined. One chart uses all the data to determine the control limits. This chart is referred to as the standard limits Phase I chart. The other, called an individual limits Phase I chart, compares each statistic being plotted with control limits that are estimated from the remaining data. A comparison of these two charting procedures is made and some recommendations are given.

1 Introduction

It is often the case some aspect of the quality of the output of a process can be described in terms of one or more parameters of the distribution of a quality measurement, X. For a continuous quality measurement, the parameters reflecting the quality of the process are usually the mean μ and the standard deviation σ. We will assume this is the case.

The causes of randomness in the distribution of X were classified by Shewhart [1] as either "natural" or "assignable." The natural causes of variability are nonremovable causes of randomness. The assignable causes of variability are removable and when identified and removed, the quality of the process is improved. A process with only natural causes of variability is said to be in statistical control or simply in-control. We

*Correspondence to: Charles W. Champ, Department of Mathematics and Computer Science, Georgia Southern University, Statesboro, GA 30460-8093, USA.

will assume that when the process is in-control that μ and σ take on the values μ_0 and σ_0, respectively. The control chart is a graphical tool that aids in the discovery of assignable causes of variability. That is, when $\mu \neq \mu_0$ or $\sigma \neq \sigma_0$.

The usual discussions in the literature about control charts are concerned with their use in monitoring a production process. These charts are referred to as Phase I charts. Control charts are also used as aids in bringing a process into an in-control state as well as for defining and establishing what is mean by the process being in-control. Charts used in this manor are referred to as Phase I control charts. Phase I as well as Phase II control charts are discussed in Shewhart [1], Duncan [2], and Montgomery [3].

Various authors have examined the performance of Phase I Shewhart $\overline{X}$ charts These include, among others, Hart, Hart, and Philip [4], Chou [5], Chou and Champ [6], Newton [7], Newton and Champ [8], and Quessenberry [9]. Under an autocorrelated data model, the performance of the X chart was studied by Maragah and Woodall [10] and the $\overline{X}$ chart by Hillier [11], Boyles [12]. Chou [5] and Chou and Champ [6] also analyzed the performance of Phase I Shewhart R and S charts. Yang and Hillier [13] examined both Phase I and Phase II Shewhart $\overline{X}$ and R charts and gave some recommendations on designing these charts. Ferrell [14] examined the use of the Phase I Shewhart midrange and $\widetilde{X}$ (sample median) control charts and Clifford [15] proposed methods for constructing Phase I Shewhart $\widetilde{X}$, R, and midrange charts. Phase I Shewhart charts for exponentially distributed times between events data is studied by Champ and Jones [16].

For attribute data models, Woodall [17] (with discussion) reviewed the issues in statistical process control (SPC) for Phase I attribute charts. A performance analysis of Phase I p and np charts is given by Borror and Champ [18].

In this article, we discuss and evaluate two Phase I control charts. One, commonly suggested in the literature, uses all the data to estimate the in-control parameters of the process. The other, only mentioned in passing in the literature, does not use all the data at each sampling stage to estimate the control limits. Before defining these two control charting procedures, we examine how the the parameters of the distribution of the quality measurement, X, are estimated. These two types of Shewhart Phase I $\overline{X}$ chart are discussed and compared in Section 4. In Section 5, Shewhart Phase I R and S charts are examined. Recommendations are given in the conclusion.

2 Estimators of Chart Parameters

Management sometimes is able to specify a (target) values for μ_0 and/or σ_0. If this is not the case then the unspecified parameter(s) must be estimated from a set of preliminary data. We assume this set of preliminary data is in the form of m independent random samples $\{X_{i,1}, ..., X_{i,n}\}$ (each of size n) taken in the order of output. A common assumption about the distribution of X is that it has (at least approximately) a normal distribution. We assume the distribution of X is normal.

The most commonly used unbiased estimator for μ_0 is the average of the sample mean given by

$$\overline{\overline{X}} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} X_{i,j} = \frac{1}{m}\sum_{i=1}^{m}\overline{X}_i$$

where $\overline{X}_i$ is the ith sample mean. Under the assumptiom the process is in-control, the estimator $\overline{\overline{X}}$ has a normal distribution with mean μ_0 and variance $\sigma_0^2/(mn)$.

The commonly recommended estimators for σ_0 found in the literature are $\overline{R}$ and $\overline{S}$, the average of the sample ranges and average of the standard deviations, respectively. We define these as follows

$$\overline{R} = \frac{1}{m}\sum_{i=1}^{m} R_i \text{ and } \overline{S} = \frac{1}{m}\sum_{i=1}^{m} S_i,$$

where R_i and S_i are respectively the ith sample range and standard deviation. Under the normality assumption and the sample is from and in-control process, unbiased estimators of σ_0 are obtained from the sample range and standard deviation by dividing R_i by d_2 and S_i by c_4. Harter [19] gives tables which include the values for d_2 and d_3^2, where the mean and variance of the normal sample range are given respectively by $d_2\sigma_0$ and $d_3^2\sigma_0^2$. The constant c_4 is determined by

$$c_4 = \frac{\sqrt{2}\Gamma\left(\frac{n}{2}\right)}{\sqrt{n-1}\Gamma\left(\frac{n-1}{2}\right)}.$$

These constants are also tabled in Duncan [2] and Montgomery [3]. Patnaik [20] demonstrates a method for approximating the distribution of $\nu\overline{R}^2/\left(c\sigma_0^2\right)$ as that of a chi square variate with ν degrees of freedom where c is a constant depending on ν. He states that "... there is at present no direct means of judging the accuracy of the ... approximation."

Another biased estimator for σ_0 found in the literature is

$$\overline{V}^{1/2} = \sqrt{\frac{1}{m}\sum\nolimits_{i=1}^{m} S_i^2},$$

where S_i^2 is the ith sample variance. Dividing this estimator by the constant $c_{4,m}$ gives an unbiased estimator of σ_0, where

$$c_{4,m} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)+1}{2}\right)}{\sqrt{m\left(n-1\right)}\Gamma\left(\frac{m(n-1)}{2}\right)}.$$

Note that $c_{4,m} = c_4\left(m\left(n-1\right)+1\right)$. It is easy to show that $m\left(n-1\right)\overline{V}/\sigma_0^2$ has a chi square distribution with $m\left(n-1\right)$ degrees of freedom. Further, of the three estimators for σ_0, it can be demonstrated that

$$Var\left[\overline{V}^{1/2}/c_{4,m}\right] \leq Var\left[\overline{S}/c_4\right] \leq Var\left[\overline{R}/d_2\right].$$

3 Phase I Shewhart $\overline{X}$ Control Charts

Similar to charts for monitoring a process mean, a Phase I Shewhart control chart is defined by a lower control limit (LCL), a center line (CL), and an upper control limit (UCL). These values are defined for the Phase I Shewhart $\overline{X}$ chart by

$$LCL_{\overline{X}} = \widehat{\mu}_0 - k_{\overline{X}}\frac{\widehat{\sigma}_0}{\sqrt{n}};\ CL_{\overline{X}} = \widehat{\mu}_0;\ UCL_{\overline{X}} = \widehat{\mu}_0 + k_{\overline{X}}\frac{\widehat{\sigma}_0}{\sqrt{n}}, \tag{1}$$

where $\widehat{\mu}_0$ and $\widehat{\sigma}_0$ are estimators of μ_0 and σ_0, respectively. Typically, it is recommended that the positive contant $k_{\overline{X}}$ be set to 3. In this article, we will select $k_{\overline{X}}$ such that if $\mu = \mu_0$ and $\sigma = \sigma_0$ the probability at least one sample mean is outside the control limits is at most α with $0 < \alpha < 1$. The estimators $\widehat{\mu}_0$ and $\widehat{\sigma}_0$ are functions of the m preliminary samples and the statistic $\overline{X}_i$ (ith sample mean) is plotted verses the sample number i. If any point plots below the $LCL_{\overline{X}}$ or above the $UCL_{\overline{X}}$, this is taken as evidence that the associated sample is from an out-of-control process. When all the preliminary data is used to determine the center line and control limits, we will refer to such a chart as a "standard limits" Phase I chart.

A second Phase I chart is discussed in Yang and Hillier [13]. The center line and control limits for the statistic plotted at time i are not functions of the ith sample. One or more of the other $m-1$ samples are used to estimate μ_0 and σ_0. When only the i sample is removed, these estimators for μ_0 and σ_0 will be denoted respectively by $\widehat{\mu}_{0,[i]}$ and $\widehat{\sigma}_{0,[i]}$. This type of Phase I Shewhart $\overline{X}$ is defined for each i by

$$LCL_{\overline{X},[i]} = \widehat{\mu}_{0,[i]} - k_{\overline{X},[i]}\frac{\widehat{\sigma}_{0,[i]}}{\sqrt{n}};\ CL_{\overline{X},[i]} = \widehat{\mu}_0;\ UCL_{\overline{X},[i]} = \widehat{\mu}_{0,[i]} + k_{\overline{X},[i]}\frac{\widehat{\sigma}_{0,[i]}}{\sqrt{n}},$$

where $\widehat{\mu}_{0,[i]}$ and $\widehat{\sigma}_{0,[i]}$ are unbiased estimator of μ_0 and σ_0 respectively based on the preliminary set of data with the ith sample removed. Chou [5] and Chou and Champ [6] refer to these charts as "individual limits" Phase I Shewhart $\overline{X}$ charts.

A Phase I control chart is designed assuming the data is from an in-control process. To evaluate the effectiveness of a Phase I $\overline{X}$ chart in identifying an out-of-control process, we consider a simple out-of-control senerio This senerio was chosen partly because various in-control and out-of-control probabilities can be calculated. These probabilities are compared to give some indication of relative effectiveness of each chart.

The out-of-control situation considered is that one of the samples is out-of-control and the other $m-1$ samples are in-control. Without loss of generality, we can take the first sample to be the one out-of-control. We suppose $\overline{X}_1 \backsim N\left(\mu_1, \sigma_0^2/\sqrt{n}\right)$ and $\overline{X}_i \backsim N\left(\mu_0, \sigma_0^2/n\right)$, $i = 2, ..., m$, where $\mu_1 \neq \mu_0$. In other words, the first sample is out-of-control and this is reflected only in the mean and the remaining $m-1$ samples are from an in-control process. It will be convenient to define the value δ_1 to be the number of standard deviations, $\sigma_0/\sqrt{n}$, of the sample mean that the value μ_1 differs from μ_0. That is, $\delta_1 = (\mu_1 - \mu_0)/(\sigma_0/\sqrt{n})$. Under these assumptions

$$\overline{X}_1 - \overline{\overline{X}} \backsim N\left(\frac{m-1}{m}(\mu_1 - \mu_0), \frac{m-1}{m}\frac{\sigma_0^2}{n}\right); \ \frac{m(n-1)\overline{V}}{\sigma_0^2} \backsim \chi^2_{m(n-1)};$$
$$\overline{X}_i - \overline{\overline{X}} \backsim N\left(-\frac{1}{m}(\mu_1 - \mu_0), \frac{m-1}{m}\frac{\sigma_0^2}{n}\right), \ i = 2, 3, ..., m.$$

Further, it should be noted that $\overline{X}_i - \overline{\overline{X}}$ and $m(n-1)\overline{V}/\sigma_0^2$ are independent for $i = 1, 2, ..., m$.

We now use these results to obtain some out-of-control probabilities for the standards limits Phase I Shewhart $\overline{X}$ chart defined in (1) with $\widehat{\mu}_0 = \overline{\overline{X}}$, $\widehat{\sigma}_0 = \sqrt{\overline{V}}/c_{4,m}$, and $k_{\overline{X}} = \sqrt{(m-1)/m}c_{4,m}t_{m(n-1),0,\alpha/(2m)}$. The value $t_{\nu,0,1-\gamma}$ is the 100γth percentage point of a central t-distribution with ν degrees of freedom. These limits are chosen using Boole's well-known inequality (see Chou [5] and Chou and Champ [6]). We express these limits and center line as

$$LCL_{\overline{X}} = \overline{\overline{X}} - A_4^{***}\sqrt{\overline{V}}; \ CL_{\overline{X}} = \overline{\overline{X}}; \ UCL_{\overline{X}} = \overline{\overline{X}} + A_4^{***}\sqrt{\overline{V}}$$

where $A_4^{***} = \sqrt{(m-1)/(mn)}t_{m(n-1),0,\alpha/(2m)}$. This simple method of choosing the control limits gives a probability of at least one false alarm at no more than α when the process is in-control. Various values of A_4^{***} are given in Table 1.

Table 1. Values of $A_4^{***}, \alpha = 0.05$

n	5	10	15	20	25
2	2.55014	2.40248	2.37979	2.38149	2.39035
3	1.63660	1.72719	1.77863	1.81572	1.84493
4	1.31814	1.43716	1.49745	1.53814	1.56889
5	1.13814	1.26056	1.32089	1.36085	1.39066
6	1.01783	1.13748	1.19588	1.23424	1.24610
7	0.92967	1.04507	1.10113	1.12318	1.15211
8	0.86134	0.97226	1.01225	1.04937	1.07666
9	0.80631	0.91294	0.95320	0.98846	1.01435
10	0.76073	0.86336	0.90344	0.93708	0.96176

The probability $\overline{X}_1$ falls outside the control limits is given by $p^{***}_{\overline{X},1}$. It is shown in Chou [5] that

$$p^{***}_{\overline{X},1} = 1 - F_{m(n-1),\theta_1}\left(\sqrt{\frac{mn}{m-1}}A_4^{***}\right) + F_{m(n-1),\theta_1}\left(-\sqrt{\frac{mn}{m-1}}A_4^{***}\right),$$

where $F_{m(n-1),\theta_1}$ is the cumulative distribution function (CDF) of a noncentral t-distribution with $m\,(n-1)$ degrees of freedom and noncentrality parameter $\theta_1 = \sqrt{(m-1)\,/m}\delta_1$. Table 2 gives these probabilities for $\delta_1 = 0.0\,(0.1)\,2.0$, $m = 5\,(5)\,25$, $n = 5$, and $\alpha = 0.05$.

Table 2. Values of $p^{***}_{\overline{X},1}, n = 5, \alpha = 0.05$

	m				
δ_1	5	10	15	20	25
0.0	0.01000	0.0050	0.00330	0.00250	0.00200
0.1	0.01025	0.00518	0.00347	0.00261	0.00210
0.2	0.01102	0.00572	0.00389	0.00296	0.00239
0.3	0.01232	0.00665	0.00461	0.00355	0.00290
0.4	0.01418	0.00800	0.00567	0.00443	0.00365
0.5	0.01664	0.00981	0.00711	0.00563	0.00468
0.6	0.01976	0.01216	0.00899	0.00721	0.00606
0.7	0.02359	0.01512	0.01139	0.00925	0.00783
0.8	0.02822	0.01877	0.01440	0.01182	0.01010
0.9	0.03372	0.02321	0.01811	0.01503	0.01294
1.0	0.04017	0.02856	0.02264	0.01898	0.01647
1.1	0.04766	0.03494	0.02811	0.02380	0.02080
1.2	0.05629	0.04246	0.03465	0.02962	0.02607
1.3	0.06614	0.05125	0.04240	0.03658	0.03241
1.4	0.07730	0.06144	0.05151	0.04483	0.03998
1.5	0.08986	0.07316	0.06211	0.05452	0.04894
1.6	0.10389	0.08652	0.07435	0.06579	0.05943
1.7	0.11944	0.10163	0.08835	0.07880	0.07160
1.8	0.13658	0.11858	0.10422	0.09366	0.08560
1.9	0.15532	0.13743	0.12206	0.11051	0.10156
2.0	0.17568	0.15822	0.14194	0.12941	0.11959

Let us now consider the probability any one of the in-control sample means plots outside the control limits in (1). Since this probability will be the same for each sample mean, we will calculate the probability $\overline{X}_2$ falls outside the control limits. This probability, denoted by $p^{***}_{\overline{X},2}$, is given by

$$p^{***}_{\overline{X},2} = 1 - F_{m(n-1),\theta_2}\left(\sqrt{\frac{mn}{m-1}}A^{***}_4\right) + F_{m(n-1),\theta_2}\left(-\sqrt{\frac{mn}{m-1}}A^{***}_4\right),$$

where $F_{m(n-1),\theta_2}$ is the cumulative distribution function (CDF) of a noncentral t-distribution with $m(n-1)$ degrees of freedom and noncentrality parameter $\theta_2 = \sqrt{1/[m(m-1)]}\delta_1$. Table 3 gives these probabilities, $p^{***}_{\overline{X},2}$, for $\delta_1 = 0.0\,(0.1)\,2.0$, $m = 5\,(5)\,25$, $n = 5$, and $\alpha = 0.05$.

Table 3. Values of $p^{***}_{\overline{X},2}, n = 5, \alpha = 0.05$

	m				
δ_1	5	10	15	20	25
0.0	0.01000	0.00500	0.00333	0.00250	0.00200
0.1	0.01002	0.00500	0.00333	0.00250	0.00200
0.2	0.01006	0.00501	0.00334	0.00250	0.00200
0.3	0.01014	0.00502	0.00334	0.00250	0.00200
0.4	0.01025	0.00504	0.00334	0.00250	0.00200
0.5	0.01040	0.00505	0.00335	0.00251	0.00200
0.6	0.01057	0.00508	0.00336	0.00251	0.00201
0.7	0.01078	0.00511	0.00337	0.00252	0.00201
0.8	0.01102	0.00514	0.00338	0.00252	0.00201
0.9	0.01129	0.00518	0.00339	0.00253	0.00201
1.0	0.01160	0.00522	0.00340	0.00253	0.00202
1.1	0.01194	0.00527	0.00342	0.00254	0.00202
1.2	0.01232	0.00532	0.00343	0.00254	0.00202
1.3	0.01273	0.00537	0.00345	0.00255	0.00203
1.4	0.01317	0.00543	0.00347	0.00256	0.00203
1.5	0.01366	0.00550	0.00349	0.00257	0.00204
1.6	0.01418	0.00557	0.00351	0.00258	0.00204
1.7	0.01473	0.00564	0.00354	0.00259	0.00205
1.8	0.01533	0.00572	0.00356	0.00260	0.00205
1.9	0.01596	0.00580	0.00359	0.00261	0.00206
2.0	0.01664	0.00589	0.00362	0.00262	0.00207

Now consider the individual limits Shewhart $\overline{X}$ chart where the limits are based on the statistics $\overline{\overline{X}}_{[i]}$

and $\sqrt{\overline{V}_{[i]}}$, where

$$\overline{\overline{X}}_{[i]} = \frac{1}{m} \sum_{j=1, j\neq i}^{m} \overline{X}_i \text{ and } \overline{V}_{[i]} = \frac{1}{m-1} \sum_{j=1, j\neq i}^{m} S_i^2.$$

This chart is defined by the control limits and centerline given by

$$LCL_{\overline{X},[i]} = \overline{\overline{X}}_{[i]} - A_{4,[i]}^{***}\sqrt{\overline{V}_{[i]}};\ CL_{\overline{X},[i]} = \overline{\overline{X}}_{[i]};\ UCL_{\overline{X},[i]} = \overline{\overline{X}}_{[i]} + A_{4,[i]}^{***}\sqrt{\overline{V}_{[i]}}$$

where $A_{4,[i]}^{***} = \sqrt{m/\left(\left(m-1\right)n\right)}t_{(m-1)(n-1),\alpha/(2m)}$ for $i = 1, ..., m$. Values for $A_{4,[i]}^{***}$ are given in Table 4.

Table 4. Values of $A_{4,[i]}^{***}, \alpha = 0.05$

n	5	10	15	20	25
2	3.63986	2.75009	2.58340	2.52547	2.50193
3	2.16589	1.94537	1.91740	1.91803	1.92622
4	1.70754	1.61092	1.61084	1.62282	1.63673
5	1.46039	1.40973	1.41944	1.43493	1.45024
6	1.29871	1.27037	1.28433	1.30097	1.29846
7	1.18192	1.16613	1.18209	1.18282	1.20045
8	1.09228	1.08422	1.10113	1.10502	1.12179
9	1.02057	1.01759	1.02191	1.04083	1.05684
10	0.96148	0.96198	0.96848	0.98669	1.00202

We will represent the probabilities the means of the first and second samples fall outside their individual control limits, respectively, by $p_{\overline{X},[1]}^{***}$ and $p_{\overline{X},[2]}^{***}$. These probabilities can be calculated as follows

$$p_{\overline{X},[1]}^{***} = 1 - F_{(m-1)(n-1),\theta_{[1]}}\left(\sqrt{\frac{m}{\left(m-1\right)n}}A_{[4]}^{***}\right) + F_{(m-1)(n-1),\theta_{[1]}}\left(-\sqrt{\frac{m}{\left(m-1\right)n}}A_{[4]}^{***}\right)$$

and

$$p_{\overline{X},[2]}^{***} = 1 - F_{(m-1)(n-1),\theta_{[2]}}\left(\sqrt{\frac{m}{\left(m-1\right)n}}A_{[4]}^{***}\right) + F_{(m-1)(n-1),\theta_{[2]}}\left(-\sqrt{\frac{m}{\left(m-1\right)n}}A_{[4]}^{***}\right)$$

where $\theta_{[1]} = \sqrt{\left(m-1\right)/m}\delta_1$ and $\theta_{[2]} = -\sqrt{1/\left[m\left(m-1\right)\right]}\delta_1$. For $\delta_1 = 0.0\left(0.1\right)2.0$, $m = 5\left(5\right)25$, $n = 5$, and $\alpha = 0.05$, Tables 5 and 6 give, respectively, the values for $p_{\overline{X},[1]}^{***}$ and $p_{\overline{X},[2]}^{***}$.

Table 5. Values of $p^{***}_{\overline{X},[1]}, n = 5, \alpha = 0.05$

δ_1	m = 5	10	15	20	25
0.0	0.01000	0.00500	0.00333	0.00250	0.00200
0.1	0.01024	0.00518	0.00347	0.00261	0.00210
0.2	0.01098	0.00571	0.00389	0.00296	0.00239
0.3	0.01222	0.00663	0.00460	0.00355	0.00289
0.4	0.01400	0.00796	0.00566	0.00442	0.00364
0.5	0.01636	0.00976	0.00709	0.00562	0.00468
0.6	0.01934	0.01207	0.00896	0.00719	0.00605
0.7	0.02300	0.01499	0.01134	0.00922	0.00782
0.8	0.02741	0.01859	0.01433	0.01178	0.01008
0.9	0.03263	0.02297	0.01801	0.01498	0.01291
1.0	0.03876	0.02825	0.02251	0.01891	0.01643
1.1	0.04586	0.03452	0.02794	0.02371	0.02075
1.2	0.05403	0.04192	0.03443	0.02950	0.02600
1.3	0.06335	0.05058	0.04212	0.03643	0.03232
1.4	0.07390	0.06060	0.05116	0.04464	0.03987
1.5	0.08576	0.07213	0.06167	0.05428	0.04879
1.6	0.09900	0.08527	0.07381	0.06550	0.05925
1.7	0.11369	0.10013	0.08769	0.07844	0.07138
1.8	0.12986	0.11680	0.10343	0.09323	0.08533
1.9	0.14755	0.13534	0.12113	0.10999	0.10124
2.0	0.16678	0.15580	0.14086	0.12881	0.11920

Table6. Values of $p^{***}_{\overline{X},[2]}, n = 5, \alpha = 0.05$

			m		
δ_1	5	10	15	20	25
0.0	0.01000	0.00500	0.00333	0.00250	0.00200
0.1	0.01002	0.00500	0.00333	0.00250	0.00200
0.2	0.01006	0.00501	0.00334	0.00250	0.00200
0.3	0.01014	0.00502	0.00334	0.00250	0.00200
0.4	0.01024	0.00503	0.00334	0.00250	0.00200
0.5	0.01038	0.00505	0.00335	0.00251	0.00200
0.6	0.01055	0.00508	0.00336	0.00251	0.00201
0.7	0.01075	0.00511	0.00337	0.00252	0.00201
0.8	0.01098	0.00514	0.00338	0.00252	0.00201
0.9	0.01124	0.00518	0.00339	0.00253	0.00201
1.0	0.01154	0.00522	0.00340	0.00253	0.00202
1.1	0.01186	0.00526	0.00342	0.00254	0.00202
1.2	0.01222	0.00531	0.00343	0.00254	0.00202
1.3	0.01262	0.00537	0.00345	0.00255	0.00203
1.4	0.01304	0.00543	0.00347	0.00256	0.00203
1.5	0.01351	0.00549	0.00349	0.00257	0.00204
1.6	0.01400	0.00556	0.00351	0.00258	0.00204
1.7	0.01454	0.00563	0.00354	0.00259	0.00205
1.8	0.01511	0.00571	0.0356	0.00260	0.00205
1.9	0.01571	0.00580	0.00359	0.00261	0.00206
2.0	0.01636	0.00588	0.00361	0.00262	0.00207

Now consider how well these charts are able to detect a certain out-of-control condition. For example from Table 2 we can see that for $m = 10$ there is a 2.856% chance that, if the first sample comes from a distribution with mean one standard deviation above the in-control mean and the other four samples are in-control, its mean will plot outside the control limits. The chart being considered is a standard Phase I

Shewhart chart. On the other hand, the corresponding individual limits Phase I Shewhart chart has a 2.825% chance the mean of the sample coming from this out-of-control situation will plot outside the individual control limits. In general, $p^{***}_{\overline{X},1} > p^{***}_{\overline{X},[1]}$ for all size shifts in the mean, number of samples m, and sample size n. Looking at Tables 3 and 5 we see that $p^{***}_{\overline{X},2} \approx p^{***}_{\overline{X},[2]}$, but in general, $p^{***}_{\overline{X},2}$ is slightly greater than $p^{***}_{\overline{X},[2]}$. These results demonstrate that the standard limits Phase I Shewhart $\overline{X}$ performs better (but only slightly better) than the corresponding individual limits chart. Also, we note that more computations are required in computing the control limits of an individual limits Phase I chart than for the corresponding standards limits chart.

4 Phase I Shewhart *R* and *S* Control Charts

The standard limits Phase I R and S charts are defined, respectively, by

$$
\begin{aligned}
LCL_{R,i} &= d_2\widehat{\sigma}_0 - k^*_{R,L,i}d_3\widehat{\sigma}_0 = \left(d_2 - k^*_{R,L,i}d_3\right)\widehat{\sigma}_0 = k_{R,L,i}\widehat{\sigma}_0; \\
CL_{R,i} &= d_{2,n_i}\widehat{\sigma}_0; \\
UCL_{R,i} &= d_2\widehat{\sigma}_0 + k^*_{R,U,i}d_3\widehat{\sigma}_0 = \left(d_2 + k^*_{R,U,i}d_3\right)\widehat{\sigma}_0 = k_{R,U,i}\widehat{\sigma}_0;
\end{aligned}
$$

and

$$
\begin{aligned}
LCL_{S,i} &= c_4\widehat{\sigma}_0 - k^*_{S,L,i}\sqrt{1-c_4^2}\widehat{\sigma}_0 = \left(c_4 - k^*_{S,L,i}\sqrt{1-c_4^2}\right)\widehat{\sigma}_0 = k_{S,L,i}\widehat{\sigma}_0; \\
CL_{S,i} &= c_{4,}\widehat{\sigma}_0; \\
UCL_{S,i} &= c_4\widehat{\sigma}_0 + k^*_{S,U,i}\sqrt{1-c_4^2}\widehat{\sigma}_0 = \left(c_4 + k^*_{S,U,i}\sqrt{1-c_4^2}\right)\widehat{\sigma}_0 = k_{S,U,i}\widehat{\sigma}_0.
\end{aligned}
$$

The individual limits Phase I Shewhart R and S charts are defined by

$$
\begin{aligned}
&LCL_{R,[i]} = k_{R,L,[i]}\widehat{\sigma}_{0,[i]};\ CL_{R,[i]} = d_2\widehat{\sigma}_{0,[i]};\ UCL_{R,[i]} = k_{R,U,[i]}\widehat{\sigma}_{0,[i]};\ \text{and} \\
&LCL_{S,[i]} = k_{S,L,[i]}\widehat{\sigma}_{0,[i]};\ CL_{S,[i]} = c_4\widehat{\sigma}_{0,[i]};\ UCL_{S,[i]} = k_{S,U,[i]}\widehat{\sigma}_{0,[i]}.
\end{aligned}
$$

where $\widehat{\mu}_{0,[i]}$ and $\widehat{\sigma}_{0,[i]}$ are unbiased estimator of μ_0 and σ_0 respectively based on the preliminary set of data with the ith sample removed.

In general, a standard limits Phase I Shewhart chart and its corresponding individual limits chart are not equivalent. However, the standard and individual limits Phase I R charts are equivalent when the unbiased estimators

$$\widehat{\sigma}_0 = \frac{\overline{R}}{d_2} \text{ and } \widehat{\sigma}_{0,[i]} = \frac{\overline{R}_{[i]}}{d_2} = \frac{1}{(m-1)\,d_2} \sum_{j=1,j\neq i}^{m} R_j$$

are used for σ_0. Similarly, the standard and individual limits Phase I Shewhart S charts are equivalent when

$$\widehat{\sigma}_0 = \frac{\overline{S}}{c_4} \text{ and } \widehat{\sigma}_{0,[i]} = \frac{\overline{S}_{[i]}}{c_4} = \frac{1}{(m-1)\,c_4} \sum_{j=1,j\neq i}^{m} S_j;$$

or

$$\widehat{\sigma}_0 = \frac{\overline{V}^{1/2}}{c_{4,m}} \text{ and } \widehat{\sigma}_{0,[i]} = \frac{\overline{V}_{[i]}}{c_{4,m-1}} = \frac{1}{c_{4,m-1}} \sqrt{\frac{1}{m-1} \sum_{j=1,j\neq i}^{m} S_j^2}$$

To see this for the R chart, first we note that

$$\frac{R_i}{\overline{R}} = \frac{m}{1 + (m-1)\left(R_i/\overline{R}_{[i]}\right)^{-1}}.$$

Now consider the event that the ith sample range, R_i, of the standard limits Phase I Shewhart R chart is within the lower and upper control limits. It is easy to see that

$$\begin{aligned}&\left\{k_{R,L,i}\overline{R} < R_i < k_{R,U,i}\overline{R}\right\}\\ &= \left\{\frac{m-1}{mk_{R,L,i}^{-1} - 1} < \frac{R_i}{\overline{R}_{[i]}} < \frac{m-1}{mk_{R,U,i}^{-1} - 1}\right\}\end{aligned}$$

If $k_{R,L,[i]}$ and $k_{R,U,[i]}$ are defined in terms of $k_{R,L,i}$ and $k_{R,U,i}$ by

$$k_{R,L,[i]} = \frac{m-1}{mk_{R,L,i}^{-1} - 1} \text{ and } k_{R,U,[i]} = \frac{m-1}{mk_{R,U,i}^{-1} - 1},$$

then standard and individuals limits Phase I Shewhart R charts are equivalent. Similar results hold for the standard and individuals Phase I S charts when

$$k_{S,L,[i]} = \frac{m-1}{mk_{S,L,i}^{-1} - 1}; \text{ and } k_{S,U,[i]} = \frac{m-1}{mk_{S,U,i}^{-1} - 1}.$$

5 Conclusion

We have discussed two Phase I control charts found in the literature. These charts have been referred to as standard and individual limits Phase I Shewhart charts. For the individual limits Phase I Shewhart chart, a more precise description is given. We use a simple procedure for determining the control limits proposed in Chou [5] and Chou and Champ [6] such that the probability of a false alarm is at most a given value α. This procedure was developed using Boole's inequality. Based on our analysis of these two Phase I Shewhart charts, we recommend the standard limits Phase I Shewhart $\overline{X}$ charts. It was shown that individual and standard limits Phase I Shewhart R charts can be designed to be equivalent. Similarly, individual and standard limits Phase I Shewhart S charts can be designed to be equivalent.

6 References

1. Shewhart WA (1931). *Economic Control of Quality of Manufactured Product*, D. Van Nostrand: New York.

2. Duncan AJ. *Quality Control and Industrial Statistics*, 4th ed., McGraw-Hill: New York 1989.

3. Montgomery DM. *Introduction to Statistical Quality Control*, 4th ed., John Wiley & Sons, Inc.: New York 2001.

4. Hart MK, Hart RF, Philip GC. Control charts: False indications of lack of control. *American Statistical Association Winter Conference*, 4-6 January 1990.

5. Chou S-P. Retrospective control charts. *Masters Project*, Department of Mathematics and Computer Science, Georgia Southern University, 1994.

6. Chou S-P Champ CW. A comparison of two phase I control charts. *Proceedings of the Quality and Productivity Section of the American Statistical Association*, Orlando, FL 1995; 31-35,.

7. Newton PB. Probability limits for Shewhart phase I $\overline{X}$ chart. *Masters Project*, Department of Mathematics and Computer Science, Georgia Southern University, 1996.

8. Newton PB, Champ CW. Probability limits for Shewhart phase I $\overline{X}$ chart. *Proceedings of the Southeast Decision Sciences Institute Annual Conference*. Decision Sciences Institute: Atlanta, GA, 1996; 26-28.

9. Quesenberry CP. SPC Q-charts for start-up processes and short and long runs. *Journal of Quality Technology* 1991: **23**:213-224.

10. Maragah HD, Woodall WH. The effect of autocorrelation on the retrospective X chart. *Journal of Quality Technology* 1992: **40**:29-42.

11. Hillier FS. Small sample probability limits for the range chart. *Journal of the American Statistical Association* 1967: **62**:1488-1493.

12. Boyles RA. Phase I analysis for autocorrelated processes. *Journal of Quality Technology* 2000: **32**:395-409.

13. Yang C-H, Hillier FS. Mean and variance control chart limits based on a small number of subgroups. *Journal of Quality Technology* 1970; **2**:9-16.

14. Ferrell EB. Control charts using midranges and medians. *Industrial Quality Control* 1953: **9**:30-34.

15. Clifford PC. Control charts without calculations: some modifications and some extensions. *Industrial Quality Control* 1959: **15**:40-44.

16. Jones LA, Champ CW. The design and performance of phase I control charts for times between events. *Quality and Reliability Engineering International* 2002: **18**. In press.

17. Woodall WH. Control charts based on attribute data: bibliography and review. *Journal of Quality Technology* 1997: **29**:172-183.

18. Borror CM, Champ CW. Phase I control charts for independent bernoulli data. *Quality and Reliability Engineering International* 2001: **17**:391-396.

19. Harter HL. Tables of range and studentized range. *Annals of Mathematical Statistics* 1960: **31**:1122-1147.

20. Patnaik PB. The use of mean range as an estimator of variance in statistical tests. *Biometrika* 1950; **37**:78-87.

Key Words: Retrospective charts, Shewhart charts.

Dr Charles W. Champ received his PhD in Statistics in 1986 from the University of Louisiana at Lafayette. His research area is statistical quality control. He has made several contributions through the introduction and performance analyses of control charting procedures. Other areas of interest include design of experiments, software quality assurance and computer arithmetic.

Mr Shou-Peng Chou received his MS in Mathematics with an emphasis in Statistics in 1994 from the Department of Mathematics and Computer Science, Georgia Southern University.

SUMMARY

Phase I control charts are used as aids in bringing a process into and defining the meaning of a process being in a state of statistical in-control. This done by looking at the data in retrospect to answer the question "were these data taken from an in-control process?" Typically the data is collected as independent random samples taken periodically from the output of the process. Commonly it is recommended that each sample be compared via a statistic(s) with all the data. We referred to the Phase I Shewhart chart that uses all the data to determine the control limits as "standard limits" charts. An alternative method is to compare each sample with the remaining or some subcollection of the remaining samples. These charts are referred to as "individual" limits charts. The individual limits charts appear to be attractive in detecting a sample from an out-of-control process when the other samples were taken when the process was in-control. We demonstrate using a simple probability analysis that standard limits Phase I Shewhart $\overline{X}$ chart performs better than the individual limits Phase I Shewhart $\overline{X}$ chart. Further, it is shown that the individual limits Phase I Shewhart R charts can be designed to be equivalent. This also holds for the individual and standard limits Phase I Shewhart S charts.