國立交通大學

統計學研究所

碩士論文

驗證實驗室管制圖的有效性

Validating Laboratory Control Charts

研 究 生：吳宗霖

指導教授：陳鄰安　教授

中 華 民 國 九 十 八 年 六 月

驗證實驗室管制圖的有效性

Validating Laboratory Control Charts

研 究 生：吳宗霖　　Student：Zong-Ling Wu

指導教授：陳鄰安　　Advisor：Dr. Lin-An Chen

國 立 交 通 大 學

統計學研究所

碩 士 論 文

A Thesis

Submitted to Institute of Statistics

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

June 2009

Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 八 年 六 月

驗證實驗室管制圖的有效性

學生：吳宗霖　　指導教授：陳鄰安　教授

國立交通大學統計學研究所 碩士班

摘　要

管制圖是統計製程管制中很有效的工具之一，它不只廣泛應用在工業品管上，在實驗室用管制圖的應用也已經越來越多了。在工業上利用管制圖來監控製程包含兩個階段：階段一及階段二，各有其重要的目標。但是在實驗室裡，往往所提供的成本與分析的時間有限，要做類似於工業品管上繁雜的階段一很難執行，基於這種情況下，我們提出了三種統計檢定的方法來執行階段一所要達到的目標，就是檢定所計算出來的管制圖是否能代表實驗室的製程用於未來監控中。最後，用檢定力去比較這三種方法。

Validating Laboratory Control Charts

Student：Zong-Ling Wu Advisor：Dr. Lin-An Chen

Institute of Statistics

National Chiao Tung University

Abstract

Typically control charts should involve phases I and II chartings of different purposes. In consideration of financial cost and analyst's time, the laboratory control charting generally do not follow the process recommended for manufacturing process control. For phase I charting, we propose several tests using for testing if a computed control chart is appropriate for the distribution representing the laboratory process. Power comparisons for the proposed tests are performed and the results are displayed and discussed.

誌　　謝

一轉眼就要畢業了，回想起碩士生涯，讓我充實很多，不僅在學術專業上精進不少，也結識了一些好朋友，更開闊了視野。

首先要非常感謝我的指導教授　陳鄰安老師，他人非常親切好相處，每次我遇到問題，不管他多忙，他總是非常有耐心的跟我解說，每次都不厭其煩的跟我加強很多觀念，不只在課業或論文上，他也常常說一些人生的道理給我分享，讓我受益良多。在此，也感謝口試委員對我論文的指導及建議。

再來要感謝我身旁的同學及朋友，相處及聊天都能消除壓力，使心情放鬆，且會互相討論課業的問題及分享做論文的心得，那種一起奮鬥互相鼓勵的感覺真的很好。

最後，當然要非常感謝我的家人，求學過程中，他們總是讓我無後顧之憂的專心於課業，給我很大的空間，都很支持我所做的決定，在此，將本篇論文獻給我的師長、家人、好朋友以及同學，並致上我最誠摯的謝意。

吳宗霖　謹誌于

國立交通大學統計學研究所

中華民國九十八年六月

Contents

Validating Laboratory Control Charts

Abstract

Typically control charts should involve phases I and II chartings of different purposes. In consideration of financial cost and analyst's time, the laboratory control charting generally do not follow the process recommended for manufacturing process control. For phase I charting, we propose several tests using for testing if a computed control chart is appropriate for the distribution representing the laboratory process. Power comparisons for the proposed tests are performed and the results are displayed and discussed.

Key words: Hypothesis testing; laboratory control chart; phase I charting.

1. Introduction

The control chart is a graphical method that plots results of control samples versus time or sequential run number and evaluates, based on control limits of the chart, whether a measurement procedure is statistical in-control or out-control. Primarily introduced for industrial manufacturing process control by Walter A. Shewhart, the control chart is now also popular as statistical control method in clinical laboratories for clinical quality control. The system of quality control in clinical laboratory is designed to decrease the probability that each result reported by the laboratory analyzer is unvalid and this result may be used with a specified confidence by the physician to make a diagonistic decision. Now, the clinical laboratory routinely uses control charts. For examples, when monitoring analyzer performance in the clinical setting, routinely the laboratories are required, based on this chart, to test concentration of material being monitored. The performance of dual-energy X-ray absorptiometry can be monitored using control charts (see Garland, Lees and Stevenson (1997)). The antigen detection enzyme-linked immunosorbent assays for hog cholera virus, foot and mouth disease virus (see Blacksell et al. (1996)).

The control rules introduced based on industrial control procedures of Shewhart by Levey and Jennings (1950) for comparison of control results

with the control limits have been refined for improvement of analytical precision by a series of papers (see Westgard and Barry (1986)) that is used in most laboratories today.

Standard control chart usage in engineering quality control should involve two phases of different objectives. Basically, the control chart represents the ideal distribution, so, the initial phase, called the phase I, involves a sequence of process including planning, administration, design of the experiment, exploratory work (e.g., graphical) and numerical analysis (e.g., estimation or hypothesis testing) to ensure the control results are drawn from this ideal distribution and process is truely in statistical control. In this stage, it is to see if reliable control limits can be established to monitor future laboratory data. Control charts are used primarily in phase I to assist operating personnel in bringing the process into a state of statistical in-control. In phase II, we use the control chart to quickly detect shifts from the in-control distribution estimated in Phase I by comparing the sample statistic for each successive sample.

Typically, in phase I, we assume that the process is initially out of control, so we are comparing a collection of m, typically $m = 20$ or 25, subgroup of samples and the objective is to bring the process into a state of statistical in-control. The control limits obtained early in this phase are viewed as trial limits. Classically in engineering quality control the control limits are revised and refined to ensure that the process is in-control. The phase I analysis is hardly executed in quality control for laboratories. The main reason is the consideration of cost, financial cost of analyses, and more frequently the analyst's time. Cost increasing is raised from the fact that the practice of quality control requires extra analytical effort. As estimated by Analytical Method Committee (1995) the amount of extra work although varies with circumstances but is likely to be at least 15%. With this reason, unlike the online quality control in industry, the frequency with which analysis is undertaken is usually very low so that taking a very long time to collect enough historical record of observations for phase I analysis. Without an appropriate phase I analysis, the resulted control chart is very questionable

to represent the in-control distribution and then the data released from a laboratory are of in-appropriate quality.

In Section 2, we introduce the concept of validating a control chart through the technique of significance test. In Section 3, we introduce two techiques of confidence interval of control limits and, in Section 4, we introduce the higest density significance (HDS) test for control chart validation. In Sections 5 and 6, we display power comparisons for these techniques of control chart validation. Finally, in Section 7, we present examples of data analysis.

2. Validation of Laboratory Control Chart

Let X be the measurement with distribution F from the system representing a characteristic of a subject of interest and $X_1, ..., X_n$ is a random sample drawn from distribution F and we choose a statistic $T = t(X_1, ..., X_n)$ that has mean μ_t and variance σ_t^2. The Shewhart control chart set μ_t as the centre line and placed three standard deviations above and below the centre line as

$$\begin{aligned} UCL &= \mu_t + 3\sigma_t \\ LCL &= \mu_t - 3\sigma_t \end{aligned} \tag{2.1}$$

If statistic T follows a normal or Gaussian distribution, the limits of the chart will cover the values of T in the long run with probability 0.9973. In practice, it is hard to believe that we know μ_t and σ_t. Therefore, we need to estimate them. Control charts are calculated based on a historical record of observations such as m subgroups of n sample and points outside the control limits are excluded and the revised control limits are calculated. Let μ_0 and σ_0 be the corresponding estimates. The estimated control limits are

$$\begin{aligned} UCL &= \mu_{t0} + 3\sigma_{t0} \\ LCL &= \mu_{t0} - 3\sigma_{t0} \end{aligned} \tag{2.2}$$

We say that the laboratory measurement system is stable when the measurement values of T are fell within the limits. On the other hand, a run is rejected and the measurement system is said to be out of control when its measurement T value exceeds the control limits.

Suppose that the measurement variable X follows the normal distribution $N(\mu, \sigma^2)$ and we consider the $\bar{X}$-chart with sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ as the test statistic. Since $\bar{X}$ has normal distribution $N(\mu, \frac{\sigma^2}{n})$. The control limits of the $\bar{X}$-chart in form of (2.1) are:

$$\begin{aligned} UCL &= \mu + 3\frac{\sigma}{\sqrt{n}} \\ LCL &= \mu - 3\frac{\sigma}{\sqrt{n}} \end{aligned} \tag{2.3}$$

By letting μ_0 and σ_0 as, respectively, the sample mean and sample standard deviation based on the historical record of observations, The estimated control limits in form of (2.2) are

$$\begin{aligned} UCL &= \mu_0 + 3\frac{\sigma_0}{\sqrt{n}} \\ LCL &= \mu_0 - 3\frac{\sigma_0}{\sqrt{n}} \end{aligned} \tag{2.4}$$

The difficulty in process control in laboratories is that due the financial cost and analyst's time there is usually no available historical record of enough observations to compute the accurate estimates of μ_0 and σ_0. Statistical inferences may help in checking if a control chart computed from a limited data represents for the distribution of an in-control process.

Suppose that we have a set of observations $x_{1i}, ..., x_{ni}$, $i = 1, ..., m$ and we compute estimates μ_0 and σ_0 to form a control chart with limits of (2.3). The interest now is to test if the measurement system is in statistical control. That is to test the following hypothesis:

$$H_0 : \mu_t - 3\sigma_t = \mu_{t0} - 3\sigma_{t0} \text{ and } \mu_t + 3\sigma_t = \mu_{t0} + 3\sigma_{t0}, \tag{2.5}$$

which is also equivalent to test the hypothesis:

$$H_0 : \mu_t = \mu_{t0}, \sigma_t = \sigma_{t0} \tag{2.6}$$

The validation problem is that we have a random sample $X_1, ..., X_k$ drawn from a distribution F for testing hypothesis (2.5) or (2.6).

3. Confidence Interval Based Tests for Laboratory Control Chart Validation

Suppose that, calculated from the historical data, we have estimates of μ_t and σ_t being μ_{t0} and σ_{t0}. The hypothesis of our concern is;

$$H_0 : (\mu_t - 3\sigma_t, \mu_t + 3\sigma_t) = (\mu_{t0} - 3\sigma_{t0}, \mu_{t0} + 3\sigma_{t0}). \tag{3.1}$$

Let $U_1 = u_1(X_1, ..., X_k)$ and $U_2 = u_1(X_1, ..., X_k)$ be two statistics based on new sample $X_1, ..., X_k$. In the following we define a confidence interval of the true control chart $LCL = \mu_t - 3\sigma_t, UCL = \mu_t + 3\sigma_t$.

Definition 3.1. We say that a random interval (U_1, U_2) is a $100(1-\alpha)\%$ confidence interval of the control chart $(\mu_t - 3\sigma_t, \mu_t + 3\sigma_t)$ if it satisfies

$$1 - \alpha = P_\theta\{U_1 \leq \mu_t - 3\sigma_t < \mu_t + 3\sigma_t \leq U_2\} \text{ for } \theta \in \Theta. \tag{3.2}$$

A rule for testing hypothesis (3.1) is

$$\text{accepting } H_0 \text{ if } u_1 \leq \mu_{t0} - 3\sigma_{t0} < \mu_{t0} + 3\sigma_{t0} \leq u_2. \tag{3.3}$$

This test is with probability, α, of type I error.

Let's consider the normal $\bar{x}$ control chart. Suppose that the $\bar{X}$-chart developed from a historical record is $LCL = \mu_0 - 3\frac{\sigma_0}{\sqrt{n}}, UCL = \mu_0 + 3\frac{\sigma_0}{\sqrt{n}}$. One way to construct confidence interval of the true control chart $LCL = \mu - 3\frac{\sigma}{\sqrt{n}}, UCL = \mu + 3\frac{\sigma}{\sqrt{n}}$ is through the sample mean $\bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i$ and sample standard deviation S with $S^2 = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - \bar{X})^2$. Hence, the hypothesis of interest is

$$H_0 : (\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}) = (\mu_0 - 3\frac{\sigma_0}{\sqrt{n}}, \mu_0 + 3\frac{\sigma_0}{\sqrt{n}}). \tag{3.4}$$

We know that $\sqrt{k}\frac{\bar{X} - \mu + 3\frac{\sigma}{\sqrt{n}}}{S} \sim t(k-1, 3\sqrt{\frac{k}{n}})$ and $\sqrt{k}\frac{\bar{X} - \mu - 3\frac{\sigma}{\sqrt{n}}}{S} \sim t(k-1, -3\sqrt{\frac{k}{n}})$, where $t(\nu, \eta)$ is the noncentral t-distribution with ν degrees of freedom and noncentrality parameter η. We also denote $t_\delta(\nu, \eta)$ as the δth quantile of noncentral t distribution $t(\nu, \eta)$.

Theorem 3.2.

$$(\bar{X} - t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}, \bar{X} + t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}})$$

is a $100(1-\alpha)\%$ confidence interval of $\bar{X}$ control chart $(\mu-3\frac{\sigma}{\sqrt{n}},\mu+3\frac{\sigma}{\sqrt{n}})$.

Proof.

$$
\begin{aligned}
1-\alpha &= P_{\mu,\sigma}\{\sqrt{k}\frac{\bar{X}-\mu+3\frac{\sigma}{\sqrt{n}}}{S} \le t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\} \\
&- P_{\mu,\sigma}\{\sqrt{k}\frac{\bar{X}-\mu-3\frac{\sigma}{\sqrt{n}}}{S} \le t_{\frac{\alpha}{2}}(k-1,-3\sqrt{\frac{k}{n}})\} \\
&= P_{\mu,\sigma}\{\bar{X}-\mu+3\frac{\sigma}{\sqrt{n}} \le t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}\} \\
&- P_{\mu,\sigma}\{\bar{X}-\mu-3\frac{\sigma}{\sqrt{n}} \le t_{\frac{\alpha}{2}}(k-1,-3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}\} \\
&= P_{\mu,\sigma}\{t_{\frac{\alpha}{2}}(k-1,-3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}+3\frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}-3\frac{\sigma}{\sqrt{n}}\} \\
&= P_{\mu,\sigma}\{-t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}} \le \bar{X}-\mu-3\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu+3\frac{\sigma}{\sqrt{n}} \le t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}\} \\
&= P_{\mu,\sigma}\{\bar{X}-t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}} \le \mu-3\frac{\sigma}{\sqrt{n}} < \mu+3\frac{\sigma}{\sqrt{n}} \le \bar{X}+t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}\} \quad \square
\end{aligned}
$$

The rule for testing H_0 through the confidence interval technique is:

accepting H_0 if $\bar{x}-t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{s}{\sqrt{k}} \le \mu_0-3\frac{\sigma_0}{\sqrt{n}} \le \mu_0+3\frac{\sigma_0}{\sqrt{n}} \le \bar{x}+t_{1-\frac{\alpha}{2}}(k-1,3\sqrt{\frac{k}{n}})\frac{s}{\sqrt{k}}$ (3.5)

The probability of type I error for this test is α.

Note that testing hypothesis $H_0:(\mu_t-3\sigma_t,\mu_t+3\sigma_t)=(\mu_{t0}-3\sigma_{t0},\mu_{t0}+3\sigma_{t0})$ if and only if to test the hypothesis $H_0:\mu_t=\mu_{t0},\sigma_t=\sigma_{t0}$. Then, in the normal case, testing $H_0:(\mu-3\frac{\sigma}{\sqrt{n}},\mu+3\frac{\sigma}{\sqrt{n}})=(\mu_0-3\frac{\sigma_0}{\sqrt{n}},\mu_0+3\frac{\sigma_0}{\sqrt{n}})$ is equivalent to test the hypothesis $H_0:\mu=\mu_0,\ \sigma=\sigma_0$. The classical technique of exact level α test for testing hypothesis $H_0:\mu=\mu_0,\ \sigma=\sigma_0$ uses the product of tests, respectively, for single parameter hypothesis $H_0^{\mu}:\mu=\mu_0$ and another single parameter hypothesis $H_0^{\sigma}:\sigma=\sigma_0$. When H_0^{μ} is true,

$$(\mu_0-z_{\frac{1-\sqrt{1-\alpha}}{2}}\frac{\sigma_0}{\sqrt{k}},\mu_0+z_{\frac{1-\sqrt{1-\alpha}}{2}}\frac{\sigma_0}{\sqrt{k}})$$

is a $\sqrt{1-\alpha}$ prediction interval for $\bar{X}$ and, when H_0^{σ} is true,

$$(\sigma_0\sqrt{\frac{1}{k-1}\chi^2_{\frac{1+\sqrt{1-\alpha}}{2}}},\sigma_0\sqrt{\frac{1}{k-1}\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}}\ )$$

is a $\sqrt{1-\alpha}$ prediction interval for S with $S^2 = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - \bar{X})^2$. From the fact that $\bar{X}$ and S are independent and with (3.3) and (3.4), an exact level α test for hypothesis H_0 is

$$\text{rejecting } H_0 \text{ if } |\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{k}}| > z_{\frac{1-\sqrt{1-\alpha}}{2}} \text{ or }, \frac{(k-1)s^2}{\sigma_0^2} < \chi^2_{\frac{1+\sqrt{1-\alpha}}{2}} \text{ or } > \chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}. \tag{3.6}$$

This test is generally called the combination test. We would not specify its acceptance region since our interest is its power performance.

4. Highest Density Test for Laboratory Control Chart Validation

For developing a HDS test for hypothesis (2.6), we consider it in a general distributional situation. Suppose that we have a random sample $X_1, ..., X_k$ drawn from a distribution having a probability density function (pdf) $f(x, \theta_1, ..., \theta_m)$ where parameters $\theta_1, ..., \theta_m$, $m \geq 1$, including location and scale ones, are unknown. It has been an important question in applications to develop tests for hypothesis simultaneously dealing with all parameters such as

$$H_0 : \theta_1 = \theta_{10}, ..., \theta_m = \theta_{m0}, \tag{4.1}$$

where $\theta_{10}, ..., \theta_{m0}$ are specified constants.

Definition 4.1. Consider the null hypothesis $H_0 : \theta_1 = \theta_{10}, ..., \theta_m = \theta_{m0}$. Suppose that there exists a constant a_α such that

$$1 - \alpha = \int_{\{(x_1,...,x_k): L(x_1,...,x_k,\theta_{10},...,\theta_{m0}) > a_\alpha\}} L(x_1, ..., x_k, \theta_{10}, ..., \theta_{m0})dx.$$

Then we call the test with acceptance region

$$A_{hds} = \{(x_1, ..., x_k) \in \Lambda : L(x_1, ..., x_k, \theta_{10}, ..., \theta_{m0}) > a_\alpha\}$$

a level α highest density significance (HDS) test. The acceptance region A_{hds} is called a level α HDS acceptance region and its corresponding rejection region $C_{hds} = \Lambda - A_{hds}$ is called the HDS rejection region.

The method of highest density for significance test is appealing for that it uses probability ratio to determine acceptance region, for example, if

$$\frac{L(x_a, \theta_{10}, ..., \theta_{m0})}{L(x_b, \theta_{10}, ..., \theta_{m0})} > 1$$

and x_b is in acceptance region, then x_a must also be in acceptance region. This appealing also indicates that the test statistic for hypothesis H_0 is derived through the joint probability (pdf).

If the joint pdf $L(x_1, ..., x_k, \theta_{10}, ..., \theta_{m0})$ of the random sample may be reformulated as an increasing function of statistic $T = t(X_1, ..., X_k)$, then a level α HDS test has acceptance region $A_{hds} = \{(x_1, .., x_k) : t(x_1, .., x_k) \geq t_\alpha\}$ with $1 - \alpha = P_{H_0}(t(X_1, ..., X_k) \geq t_\alpha)$.

Let $X_1, ..., X_k$ be a random sample drawn from a normal distribution $N(\mu, \sigma^2)$. Consider the null hypothesis $H_0 : \mu = \mu_0, \ \sigma = \sigma_0$. With the fact that $L(x_a, \mu_0, \sigma_0) \geq L(x_b, \mu_0, \sigma_0)$ if and only if $\sum_{i=1}^k (x_{ia} - \mu_0)^2 \leq \sum_{i=1}^k (x_{ib} - \mu_0)^2$ for $x'_a = (x_{1a}, ..., x_{ka})$ and $x'_b = (x_{1b}, ..., x_{kb})$, the level α HDS test searches t_α such that

$$\alpha = P_{\mu_0, \sigma_0}(\sum_{i=1}^k (X_i - \mu_0)^2 \geq t_\alpha) = P(\chi^2(k) \geq \frac{t_\alpha}{\sigma_0^2})$$

where $\chi^2(k)$ is the random variable with chi-square distribution of degrees of freedom k. Hence, the HDS level α test has acceptance region

$$A_{hds} = \{(x_1, ..., x_k) : \sum_{i=1}^k (x_i - \mu_0)^2 \leq \sigma_0^2 \chi_\alpha^2(k)\} \tag{4.2}$$

where χ_α^2 satisfies $\alpha = P(\chi^2(k) \geq \chi_\alpha^2(k))$.

5. Power Performance Comparisons for Laboratory Control Chart Validation

With tests developed in Sections 3 and 4, it is then interesting to compare these tests in terms of power when there is distributional shift. We consider a process with normal distribution $N(\mu, \sigma)$ where the in-control distribution parameters are $\mu = \mu_0$ and $\sigma = \sigma_0$. The null hypothesis for the $\bar{X}$ chart is of (3.4). We now set the $\bar{X}$ chart under the alternative situation is

$$H_1 : (LCL, UCL) = (\mu_0 + a - 3\frac{b\sigma_0}{\sqrt{n}}, \mu_0 + a + 3\frac{b\sigma_0}{\sqrt{n}}) \tag{5.1}$$

In the following, we develop the power functions for the three corresponding tests when H_1 is true.

5.1. Power Function of Confidence Interval of Control Chart

We know that $\sqrt{k}\frac{\bar{X}-\mu_0+3\frac{\sigma_0}{\sqrt{n}}}{S} \sim t(k-1, \frac{\sqrt{k}(a+3\frac{\sigma_0}{\sqrt{n}})}{b\sigma_0})$ and $\sqrt{k}\frac{\bar{X}-\mu_0-3\frac{\sigma_0}{\sqrt{n}}}{S} \sim t(k-1, \frac{\sqrt{k}(a-3\frac{\sigma_0}{\sqrt{n}})}{b\sigma_0})$ when H_1 is true.

$$
\begin{aligned}
&1-\pi(\mu_0+a, b\sigma_0) = P_{\mu_0+a,b\sigma_0}\{\bar{X} - t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}} \le \mu_0 - 3\frac{\sigma_0}{\sqrt{n}} < \\
&\mu_0 + 3\frac{\sigma_0}{\sqrt{n}} \le \bar{X} + t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{S}{\sqrt{k}}\} \\
&= P_{\mu_0+a,b\sigma_0}\{\sqrt{k}\frac{\bar{X}-\mu_0+3\frac{\sigma_0}{\sqrt{n}}}{S} \le t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\} \\
&- P_{\mu_0+a,b\sigma_0}\{\sqrt{k}\frac{\bar{X}-\mu_0-3\frac{\sigma_0}{\sqrt{n}}}{S} \le t_{\frac{\alpha}{2}}(k-1, -3\sqrt{\frac{k}{n}})\} \\
&= P\{t(k-1, \frac{\sqrt{k}(a+3\frac{\sigma_0}{\sqrt{n}})}{b\sigma_0}) \le t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\} \\
&- P\{t(k-1, \frac{\sqrt{k}(a-3\frac{\sigma_0}{\sqrt{n}})}{b\sigma_0}) \le t_{\frac{\alpha}{2}}(k-1, -3\sqrt{\frac{k}{n}})\}
\end{aligned}
$$

5.2. Power Function of Classical Test

The power function of this classical test is

$$
\begin{aligned}
&\pi_{class}(\mu_0+a, b\sigma_0) \qquad (5.2) \\
&= P_{\mu_0+a,b\sigma_0}(\{|\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{k}}| > z_{\frac{1-\sqrt{1-\alpha}}{2}}\} \cup \{\frac{(k-1)S^2}{\sigma_0^2} < \chi^2_{\frac{1+\sqrt{1-\alpha}}{2}} \text{ or } > \chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}\}) \\
&= P_{\mu_0+a,b\sigma_0}(|\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{k}}| > z_{\frac{1-\sqrt{1-\alpha}}{2}}) + P_{\mu_0+a,b\sigma_0}(\frac{(k-1)S^2}{\sigma_0^2} < \chi^2_{\frac{1+\sqrt{1-\alpha}}{2}} \text{ or } > \chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}) \\
&- P(|\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{k}}| > z_{\frac{1-\sqrt{1-\alpha}}{2}})\ P(\frac{(k-1)S^2}{\sigma_0^2} < \chi^2_{\frac{1+\sqrt{1-\alpha}}{2}} \text{ or } > \chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}) \\
&= P(|N(a,1)| > \frac{1}{b}z_{\frac{1-\sqrt{1-\alpha}}{2}}) + P(\chi^2(k-1) < \frac{1}{b^2}\chi^2_{\frac{1+\sqrt{1-\alpha}}{2}}) + P(\chi^2(k-1) > \frac{1}{b^2}\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}) \\
&- P(|N(a,1)| > \frac{1}{b}z_{\frac{1-\sqrt{1-\alpha}}{2}})(P(\chi^2(k-1) < \frac{1}{b^2}\chi^2_{\frac{1+\sqrt{1-\alpha}}{2}}) + P(\chi^2(k-1) > \frac{1}{b^2}\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}})).
\end{aligned}
$$

5.3. Power Function of HDS Test

We consider a power comparison for this example of distribution where we let the sample be drawn from normal distribution with mean $\mu = \mu_0 + a$

and standard deviation $\sigma = b\sigma_0, b > 0$. The power function of the HDS test of (3.1) may be seen as

$$\pi_{hds}(\mu_0 + a, b\sigma_0) = P(\chi^2(k, \frac{ka^2}{b^2\sigma_0^2}) \geq b^{-2}\chi^2_\alpha(k)) \tag{5.3}$$

where $\chi^2(k, c)$ is a random variable with noncentral chi-square distribution with degrees of freedom k and noncentrality parameter c.

With sample size $k = 30$ and significance level $\alpha = 0.05$, we list the results of powers computed from (5.2) and (5.3) for the two tests and display them in Table 1.

Table 1. Power comparison for HDS test, classical combination test and confidence interval based test

(a, b)	π_{hds}	π_{class}	π_{ci} $(n = 2)$	π_{ci} $(n = 3)$	π_{ci} $(n = 5)$
$(0, 1)$	0.05	0.05	0.05	0.05	0.05
$(0, 1.5)$	0.9299	0.8472	1.39×10^{-4}	2.35×10^{-4}	4.87×10^{-4}
$(0, 2)$	0.9994	0.9978	1.58×10^{-6}	4.58×10^{-6}	1.87×10^{-5}
$(0, 5)$	1	1	0	0	0
$(1, 1)$	0.8950	0.1307	0.668	0.785	0.892
$(1, 1.5)$	0.9970	0.8871	0.019	0.044	0.108
$(1, 2)$	0.9999	0.9986	3.22×10^{-4}	0.001	0.005
$(1, 5)$	1	1	0	0	0
$(2, 1)$	1	0.4216	0.996	0.999	1
$(2, 1.5)$	1	0.9499	0.331	0.555	0.813
$(2, 2)$	1	0.9995	0.017	0.057	0.190
$(2, 5)$	1	1	1.98×10^{-8}	3.37×10^{-7}	9.02×10^{-6}
$(5, 1)$	1	0.9972	1	1	1
$(5, 1.5)$	1	1	1	1	1
$(5, 2)$	1	1	0.923	0.992	0.999
$(5, 5)$	1	1	7.80×10^{-5}	9.87×10^{-4}	0.013

The powers when $(a, b) = (0, 1)$ represent, respectively, the significance levels of these two test and they are, as designed, equal to 0.05. However, when value a moves away from zero and $b > 1$, the alternative distribution indicates wilder than the null one. Surprisingly the powers of the HDS test is uniformly better than or equal to the classical combination test. This fully supports the use of HDS test for hypothesis test of multiple parameters.

We also observe that the validation technique of confidence interval od true control chart performs poorly when there is scale shift.

6. Power Simulation Study for Laboratory Control Chart Validation

It is also interesting to see the power performance of three tests through a Monte Carlo study. We first consider the normal $\bar{x}$ control chart. Suppose that the $\bar{X}$-chart developed from a historical record is $LCL = \mu_0 - 3\frac{\sigma_0}{\sqrt{n}}, UCL = \mu_0 + 3\frac{\sigma_0}{\sqrt{n}}$. Hence, the hypothesis of interest is

$$H_0 : (\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}) = (\mu_0 - 3\frac{\sigma_0}{\sqrt{n}}, \mu_0 + 3\frac{\sigma_0}{\sqrt{n}}). \tag{6.1}$$

With replication $m = 10{,}000$, we select a random sample $x_{j1}, ..., x_{jk}$ of size $k = 30$ from a distribution G and we conduct the tests stated above of higest density test, classical test and the test based on confidence interval. The first case, we consider G is the normal distribution $N(\mu_0 + a, b^2\sigma_0^2)$ where we choose $\mu_0 = 0$ and $\sigma_0 = 1$ in this simulation. Let $\bar{x}_j = \frac{1}{k}\sum_{i=1}^{k} x_{ji}$ and $s_j^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_{ji} - \bar{x}_j)^2$ be the sample mean and sample variance for the sample of jth replication. With this simulation, the simulated powers of tests defined in (3.5), (3.6) and (4.2) are

$$\pi_{hds} = \frac{1}{10,000}\sum_{i=1}^{10,000} I[\sum_{i=1}^{k}(x_{ji} - \mu_0)^2 > \sigma_0^2\chi_\alpha^2(k)]$$

$$\pi_{class} = \frac{1}{10,000}\sum_{i=1}^{10,000} I[|\frac{\bar{x}_j - \mu_0}{\sigma_0/\sqrt{k}}| > z_{\frac{1-\sqrt{1-\alpha}}{2}} \text{ or } , \frac{(k-1)s_j^2}{\sigma_0^2} < \chi^2_{\frac{1+\sqrt{1-\alpha}}{2}} \text{ or } > \chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}],$$

and

$$\pi_{ci} = 1 - \frac{1}{10,000}\sum_{i=1}^{10,000} I[\bar{x}_j - t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{s_j}{\sqrt{k}} \le \mu_0 - 3\frac{\sigma_0}{\sqrt{n}} < \mu_0 + 3\frac{\sigma_0}{\sqrt{n}}$$

$$\le \bar{x}_j + t_{1-\frac{\alpha}{2}}(k-1, 3\sqrt{\frac{k}{n}})\frac{s_j}{\sqrt{k}}]$$

The simulated results are displayed in Table 2.

Table 2. Simulated powers for three tests when there are location and scale shifts

(a,b)	π_{hds}	π_{class}	π_{ci} $(n=2)$	π_{ci} $(n=3)$	π_{ci} $(n=5)$
$(0,1)$	0.0501	0.0512	0.0337	0.0342	0.0370
$(0,1.5)$	0.9355	0.8837	0	2×10^{-4}	3×10^{-4}
$(0,2)$	0.9994	0.9990	0	0	1×10^{-4}
$(0,5)$	1	1	0	0	0
$(1,1)$	0.9015	0.9999	0.6323	0.7447	0.8702
$(1,1.5)$	0.9966	0.9973	0.0142	0.0381	0.0926
$(1,2)$	1	0.9999	3×10^{-4}	6×10^{-4}	0.0041
$(1,5)$	1	1	0	0	0
$(2,1)$	1	1	0.994	0.9994	1
$(2,1.5)$	1	1	0.2964	0.5129	0.7742
$(2,2)$	1	1	0.0136	0.0486	0.167
$(2,5)$	1	1	0	0	0
$(5,1)$	1	1	1	1	1
$(5,1.5)$	1	1	0.9999	1	1
$(5,2)$	1	1	0.9033	0.9891	0.9998
$(5,5)$	1	1	0	7×10^{-4}	0.0107

We have several comments drawn from the results displaying in Table 2:

1. The powers for all tests for case $(a,b)=(0,1)$ are expected to be 0.05 since it indicates H_0 is true. It shows that the HDS test is the most accurate in sense of preserving the significance level. The confidence interval based tests are all too conservative in this sense.

2. For cases other than $(0,1)$, the HDS and classical tests are all very powerful and the confidence interval based tests are almost very poor unless distribution shifting occured only in location.

Next we consider to perform a simulation study assuming that the observations are drawn from non-normal G as $G=t(a)+b$ where t is t distribution and we let $a=1,3,10$ and $b=0,3,10$. We apply the same tests stated above and compute the simulated powers.

Table 3. Power Simulation when distributional shifted to t distribution

(a,b)	π_{hds}	π_{class}	π_{ci} $(n=2)$	π_{ci} $(n=3)$	π_{ci} $(n=5)$
$(1,0)$	0.9991	0.9986	0	0	0
$(1,3)$	1	1	0.0155	0.0284	0.0552
$(1,10)$	1	1	0.2895	0.3620	0.4496
$(3,0)$	0.8078	0.7546	0.0010	0.0012	0.0022
$(3,3)$	1	1	0.6928	0.8008	0.8933
$(3,10)$	1	1	0.9923	0.9960	0.9970
$(10,0)$	0.2547	0.1975	0.0149	0.0166	0.0169
$(10,3)$	1	1	0.9955	0.9995	0.9997
$(10,10)$	1	1	1	1	1

The HDS and classical tests are still very efficient and the confidence interval based tests are relatively poor but shifting to more wild case such as this t distribution is better than shifting to other normals.

7. Real Data Analyses

Let us consider two real data analses. First, a data set of control materials with size 100 (20 observed monthly) in five months is available in Westgard, Barry and Hunt (1981). They performed in constructing the in-control chart of one control material (single observation control chart) and discussed the rules applying the control chart in clinical chemistry. A data set of size 100 to perform a Phase I analysis, as recommended in statistical quality control, is not enough. The validation technique provides a scientific method for constructing an in-control chart for use in laboratory quality control. We now choose 60 observations observed from the first three months to construct the 3-sigma control chart that is

$$UCL = 99.67 + 3 \times 4.77822 = 114.0$$
$$LCL = 99.67 - 3 \times 4.77822 = 85.33$$

and is sketched in Figure 1 where the 60 observations are also displayed.

Figure 1 is here

Since there is no observation lying outside the control limits, we then conscern if this control chart is appropriate for use for quality control in clinical

chemistry. We then perform the three available tests stated above and their corresponding results are listed below:

1. HDS test: $\frac{\sum_{i=1}^{40}(x_i-99.67)^2}{4.77822^2} = 24.30847 < \chi^2_{0.05}(40) = 55.75848$, we do not reject the hypothesis of in-control process.

2. Classical test: $\left|\frac{\bar{x}-99.67}{4.77822/\sqrt{40}}\right| = 0.56916 < z_{(1-\sqrt{0.95})/2} = 2.236477$. On the other hand, $\frac{(k-1)s^2}{4.77822^2} = 24.6$, and $\chi^2_{\frac{1+\sqrt{0.95}}{2}}(39) = 21.959$ and $\chi^2_{\frac{1-\sqrt{0.95}}{2}}(39) = 61.353$ indicating $21.959 < \frac{(k-1)s^2}{4.77822^2} < 61.353$. We do not reject the hypothesis of in-control process.

3. 95% confidence interval: $(\bar{x} - \frac{s}{\sqrt{40}}t_{0.975}(39, 3\times\sqrt{40}), \bar{x} + \frac{s}{\sqrt{40}}t_{0.975}(39, 3\times\sqrt{40}))$, we do not reject the hypothesis of in-control process since $(LCL, UCL) = (85.33, 114.0)) \subset (85.2059, 114.9941)$.

In the second example, we consider a laboratory measurement data set displayed in Mullins (1999). The data set is composed with $m = 29$ runs and, for each run, three observations are measured by one analyte. So, totally there are number 87 observations. The interest in Mullins (1999) is the analytical precision and the phase I range chart, considering the difference between the larget and the smallest values in one run since they are measured with the same analyte, was developed. The quality control of central tendency is is also important and then we consider using this data set to perform the phase I $\bar{X}$ chart validation.

Following Mullins (1999), we let $n = 3$ for consideration the quality of measurements observed by the same analyte. We may observe that observations on run 28 is $\{198.92, 479.95, 492.15\}$ where observation 198.92 is an extreme outlier that should be an typing error like observation. Hence we drop this run of data and we use the rest of data set of runs 28 for analysis. We use the first 20 runs to construct 3σ $\bar{X}$ chart. The control limits of the chart and observed 20 sample means $\bar{x}$ are plotted showing in Figure 2.

Figure 2 is here

From the figure, we see that the observed $\bar{x}$'s of numbers $5, 11, 15$ and 20 lied outsider the control limits and then we removed these observations and

re-compute the $\bar{X}$ chart from the data in the rest of 16 runs of data. The resulted control limits and the $\bar{x}$'s are displaed in Figure 3.

Figure 3 is here

There are no observed $\bar{x}$ that lies outside the control limits and we consider if the resulted control limits as

$$UCL = 494.1035 + 3 \times \frac{7.79323}{\sqrt{3}} = 507.6$$
$$LCL = 494.1035 - 3 \times \frac{7.79323}{\sqrt{3}} = 480.61$$

can be used for phase II control chart. Then we consider to use observations of size 24 in 8 runs to test if the above control limits are valid for phase II control chart. We then perform the three available tests stated above and their corresponding results are listed below:

1. HDS test: $\frac{\sum_{i=1}^{24}(x_i - 494.1035)^2}{7.79323^2} = 26.7293 < \chi^2_{0.05}(24) = 36.415$, we do not reject the hypothesis of in-control process.

2. Classical test: $\left|\frac{\bar{x} - 494.1035}{7.79323/\sqrt{24}}\right| = 1.5536 < z_{(1-\sqrt{0.95})/2} = 2.236477$. On the other hand, $\frac{(k-1)s^2}{7.79323^2} = 25.373$, and $\chi^2_{\frac{1+\sqrt{0.95}}{2}}(23) = 10.549$ and $\chi^2_{\frac{1-\sqrt{0.95}}{2}}(23) = 40.746$ indicating $10.549 < \frac{(k-1)s^2}{7.793233^2} < 40.746$. we do not reject the hypothesis of in-control process.

3. 95% confidence interval: $(\bar{x} - \frac{s}{\sqrt{24}}t_{23}(0.975, 3\times\sqrt{\frac{24}{3}}), \bar{x} + \frac{s}{\sqrt{24}}t_{23}(0.975, 3\times\sqrt{\frac{24}{3}}))$, we do not reject the hypothesis of in-control process since $(LCL, UCL) = (480.61, 507.6)) \subset (470.507, 512.757)$.

References

Analytical Methods Committee (1995). Internal quality control of analytical data. *Analyst*, 120, 29-34.

Blacksell, S. D., Cameron, A. R. and Chamnanpood, P. etal. (1996). Implementation of internal laboratory quality control procedures for the monitoring of ELISA performance at regional veterinary laboratory. *Veterinary Microbiology*, 51, 1-9.

Garland, S. W., Lees, B. and Stevenson, J. C. (1997). Dxa longitudinal quality control: a comparison of inbuilt quality assurance, visual inspection, multirule Shewhart charts and cusum analysis. *Osteoporosis International*, 7, 231-237.

Huang, J.-Y., Chen, L.-A. and Welsh, A. H. (2008). Reference Limits from the Mode Interval. Submitted to *JSPI* for publication (In revision).

Levey, S. and Jennings, E. R. (1950). The use of control charts in the clinical laboratory. *American Journal of Clinical Pathology*. 20, 1059-1066.

Mullins, E. (1994). Introduction to control charts in the analytical laboratory: tutorial review. *Analyst*, 119, 369-375.

Mullins, E. (1999). Getting more from your laboratory control charts. *Analyst*, 124, 433-442.

Westgard, J. O. and Barry, P. L. (1986). *Cost-Effective Quality Control: Managing the Quality and Productivity of the Analytical Process.* AACC Press: Washington DC.

Westgard, J. O., Barry, P., L., Hunt, M. R. and Groth, T. (1981). A multirule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry*, 27, 493-501.

Figure 1: Phase I individual control chart

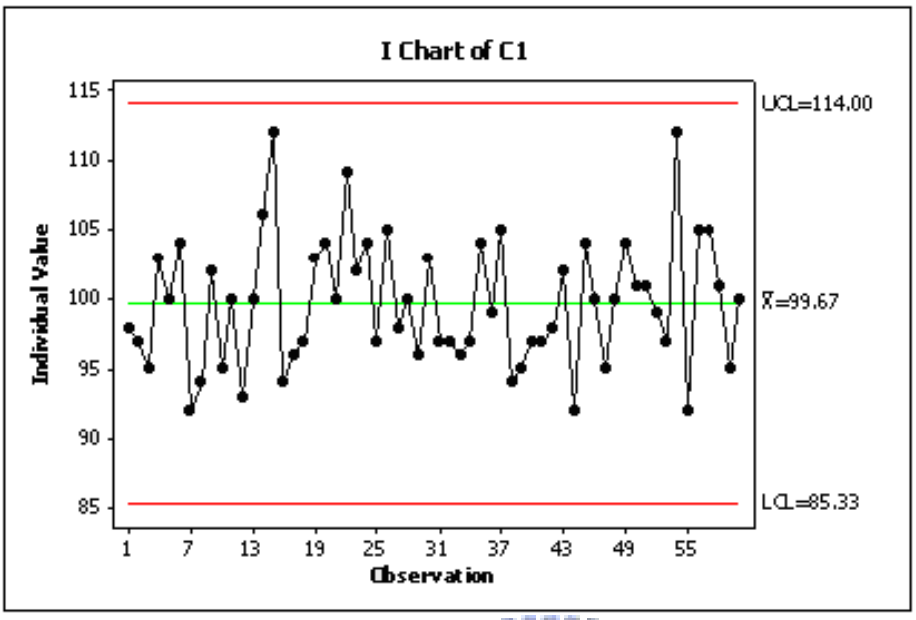

Figure 2 : Phase I $\overline{X}$ control chart for quality data

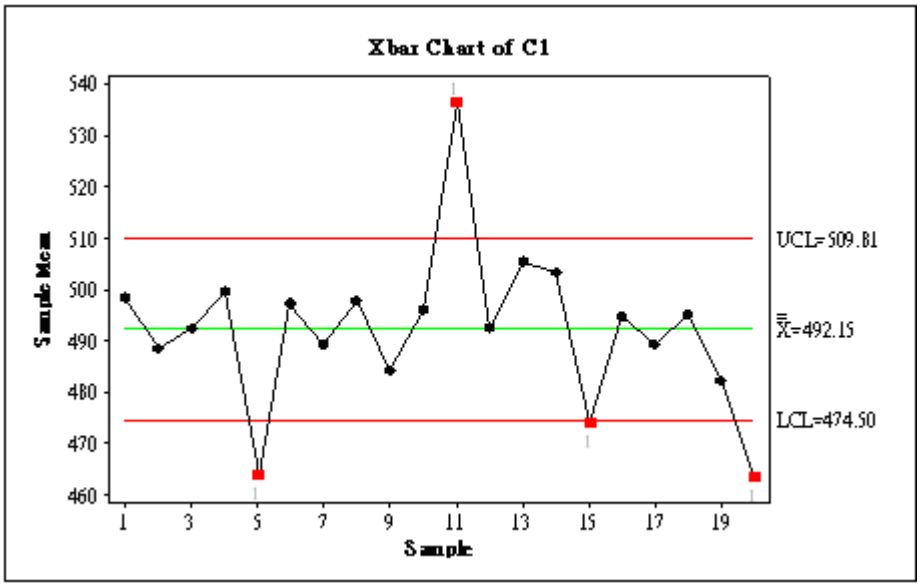

Figure 3 : Revised $\bar{X}$ control chart for quality data