國 立 交 通 大 學

統計學研究所

博士論文

包含區間, 多變量管制圖 及

構面性資料統計品質管制

Coverage Interval, Multivariate Control Chart

and Profile Monitoring

研 究 生 ： 林碩慧

指導教授 ： 陳鄰安 教授

指導教授 ： 洪志真 教授

中華民國一百零一年一月

包含區間，多變量管制圖 及

構面性資料統計品質管制

Coverage Interval, Multivariate Control Chart
and Profile Monitoring

研 究 生 ：林碩慧　　　　　　Student：Shuo-Huei Lin

指 導 教 授：陳鄰安　　　　　Advisor：Dr. Lin-An Chen

指 導 教 授：洪志真　　　　　Advisor：Dr. Jyh-Jen Horng Shiau

國 立 交 通 大 學

統 計 學 研 究 所

博 士 論 文

1896

A Thesis
Submitted to Institute of Statistics
College of Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Ph.D
in
Statistics
Jan 2012
Hsinchu, Taiwan, Republic of China

中 華 民 國 一 百 零 一 年 一 月

包含區間, 多變量管制圖及

構面性資料統計品質管制

研究生：林碩慧　　　指導教授：陳鄰安博士

指導教授：洪志真博士

國立交通大學統計學研究所

摘要

　　包含區間一般以經驗分位數來估計，本論文提出以對稱型分位數來估計，並且在論文中可看出對稱型分位數包含區間在非對稱分佈及離群值資料上有比經驗分位數包含區間有較短的長度及較好的穩健度。在對稱分佈尤其是厚尾型對稱分佈, 對稱型分位數包含區間有較小的變異。本論文亦提出以對稱型分位數建構多分位點管制圖，並且探討其大樣本理論。

本論文亦探討非線性混合性構面型資料統計品質管制。我們以主成份分析來建構統計品質管制 Phase-I 及 Phase-II 的監控統計量。在 Phase-I，我們採用主成份計分建構的 T^2 監控統計量。在 Phase-II，各別主成份計分圖、主成份計分建構的 T^2 管制圖及聯合型主成份計分管制圖被提出及比較，在應用面亦有所建議。
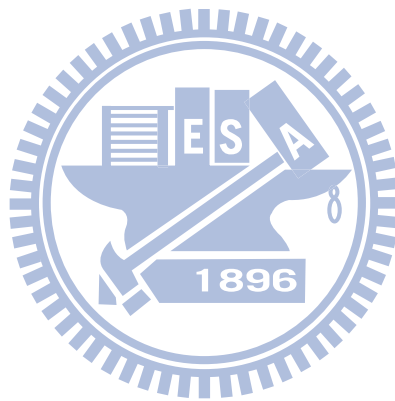
Abstract

Classically the non-parametric coverage interval is estimated by empirical quantiles. We introduce an alternative way for estimating the coverage interval by symmetric quantiles of Chen and Chiang (1996). We further show that this alternative estimator has a better precision in the sense that its asymptotic variances are smaller than the classical one.

In an attempt to develop a scheme for monitoring a vector of distributional quantiles, we propose a symmetric-quantiles-based control chart. Comparative studies in terms of the asymptotic covariance matrix and the average run length show that the proposed control chart is more efficient than the classical empirical-quantiles-based control chart.

The monitoring of process/product profiles is presently a growing and promising area of research in statistical process control. We focus on developing monitoring schemes for nonlinear profiles with random effects in this study. We utilize the technique of principal components analysis to analyze the covariance structure of the profiles and propose monitoring schemes based on principal component (PC) scores. In the Phase I analysis of historical data, due to the dependency of the PC-scores, we adopt the usual Hotelling T^2 chart to check the stability. For Phase II monitoring, we study individual PC-score control charts, a combined chart scheme that combines all the PC-score charts, and a

T^2 chart. Although an individual PC-score chart may be perfect for monitoring a particular mode of variation, a chart that can detect general shifts, such as the T^2 chart and the combined chart scheme, is more feasible in practice. The performances of the schemes under study are evaluated in terms of the average run length.

Contents

2

Chapter 1

Introduction

Statistical control charts have been widely used in many real-life applications and various control chart techniques have been developed to handle different process scenarios and characteristics of data in this literature. In this dissertation, two different types of control charts, a symmetric-quantile based control chart and a nonparametric profile control chart, are proposed and studied. In addition to control charts, other applications of coverage interval are also introduced.

1.1 Coverage Interval

In statistics, we generally have a random sample, $X_1, ..., X_n$, from a distribution with the probability density function (pdf) $f(x, \theta)$, $x \in \Re$, where $\Re$ is the sample space and θ is an unknown parameter (scalar or vector) with the parameter space Θ. The statistical inferences presented in the literature gen-

erally involve estimation and hypothesis about the unknown parameter θ. In many practical problems, we face to deal with statistical problems for inferences about a probabilistic unknown interval or multivariate region, not this single point of parameter θ.

Let X be a random variable or a statistic which is formulated from the random sample $X_1, ..., X_n$. A coverage interval is an interval with two confidence limits which covers values in the population of random quantity W in some probabilistic sense. For example, $C(\theta) = (a(\theta), b(\theta))$ is a γ coverage interval if it satisfies $\gamma = P_\theta(X \in C(\theta))$ for $\theta \in \Theta$. Treating γ coverage interval $C(\theta)$ as the unknown parameter interval of interest for inferences has provided many very important applications in various areas of sciences but still lacking communications in theory and application (see Huang, Chen and Welsh (2010) for a brief review).

This coverage interval has been referred with various terminologies in different areas of applications, for example, the coverage interval $C(\theta)$ is called the reference interval in laboratory chemistry to refer to population-based reference values obtained from a well-defined group of reference individuals. Laboratory test results are commonly compared to a reference interval before care-givers make physiological assessments, medical diagnoses, or management decisions. An individual who is being screened for a disorder based on a measurement is suspected to be abnormal if his/her measurement value lies outside the reference interval.

An experiment for measuring a measurand, a particular quantity subject to measurement, is a method through a process that tries to gain or discover knowledge of the measurand. Measurements always have errors and therefore uncertainties. General rules for evaluating and reporting uncertainty in measurement have been published by the most important and internationally widespread metrological publication – ISO (the International Standards Organization) *Guide to the Expression of Uncertainty in Measurement*. The measurement result should be reported with a specified confidence as an uncertainty interval defining the range of values that could reasonably be attributed to the measurand. Hence, a coverage interval in physical science is called the uncertainty interval. This topic is interesting but not studied in this dissertation.

The control chart is one of the basic quality improvement tools in statistical process control (SPC) for a process to identify special causes of variation and signal the need to take necessary corrective actions. If the variation is due to common causes alone, the process is said to be in statistical control. When special causes are present, the process is said to be out of control. Suppose that W is the statistic to be monitored and $(a(\theta), b(\theta))$ is the γ coverage interval. By letting $LCL = a(\theta)$ and $UCL = b(\theta)$ be the lower control limit and upper control limit, respectively, we may say that special causes may be present when an observation of W falls outside the interval (LCL, UCL).

A control chart is a statistical scheme devised for the purpose of checking and then monitoring the statistical stability of a process. Various control chart

techniques have been developed and widely applied for process control. Widely used Shewhart control chart, EWMA, and CUSUM chart are under the parametric (normal distribution) statistical thinking to perform process control by monitoring the parameters. The design of control chart is based on the concept of the coverage interval of a certain statistic such as sample mean or sample range under the assumption of a specified distribution, mostly the normal distribution.

1.2   Coverage Interval by Symmetric Quantile

The coverage interval can be estimated either parametrically or non-parametrically. The parametric method classically assumes that the underlying distribution of the measurement variable is normal whereas, recently, Chen, Huang and Chen (2007) has proposed a technique for constructing coverage intervals for asymmetric distributions. On the other hand, the non-parametric approach estimates the quantiles (percentiles) directly; the most popular technique for estimating the unknown quantiles is through the empirical quantiles. Most authorities now recommend the nonparametric method because it makes no assumptions concerning the type of the reference variable. It is easy and reliable whether the reference variable follows either a normal or non-normal distribution (see these points in Reed, Henry and Mason (1971) and Solberg (1986)).

It is vitally important to establish a coverage interval so that users can di-

agnose the disease with precision. Some factors that may increase the precision have been considered. Number one hundred and twenty or more of healthy subjects required for the determination of coverage interval has been recommended by International Federation of Clinical Chemistry. The determination of the confidence interval of the quantile, that is, the limits within which true quantile is located with a specified confidence, is strongly recommended. However, Friedberg et al. (2007) has observed that analytic imprecision is a very important factor for the quality of the established coverage interval. Hence, searching for an alternative technique in developing coverage interval to increase the analytic precision of the computed coverage interval is an interesting and important topic.

We consider the non-parametric approach in constructing coverage interval. Is there an alternative technique producing a coverage interval of better precision than that constructed by empirical quantile? For improving the efficiency of the trimmed mean for estimating the location parameter, Kim (1992) and Chen and Chiang (1996) introduced the symmetric quantile to construct an alternative trimmed mean. They observed that this trimmed mean can have asymptotic variances very close to the Cramer Rao lower bounds for several distributions, including heavy tail ones. One aim in this research is to construct an alternative coverage interval by symmetric quantiles and show that it does gain better precision than the classical version constructed by empirical quantiles. This surprising result has been published in Metrologia (Lin, Chan

and Chen (2008)).

1.3　Multivariate Control Chart by Symmetric Quantiles

It is known that the performance of a normal-based control chart is seriously degraded when the underlying distribution departs from normal. However, manufacturing processes with non-normal quality characteristics variable are very common (see, for example, Cheng and Thaga (2006), Shiling and Nelson (1976) and Kanji and Arif (2000)). Nonparametric control charts are then suggested because it makes no assumption concerning the distribution of the monitoring variable. For example, Janacek and Meikle (1997) considered the median chart, Liu and Tang (1996) considered the boostrap control chart, and Grimshaw and Alt (1997) considered using quantile function to construct the control chart.

Among the existing nonparametric control charts, the quantile- based control chart by Grimshaw and Alt (1997) is a nonparametric control chart that can simultaneously monitor distribution parameters. Unlike the confidence interval that estimates the range in which a population parameter falls, the control chart is a coverage interval that estimates the range which covers a certain percentage of the population of a specific statistic. Because coverage intervals are based upon only a sample of the entire population, we cannot be 100% confident that

the interval will contain the specified proportion of the statistic's population.

It is interesting to study if there is an alternative quantile technique that can be used to construct a quantile control chart of better efficiency in some sense than that constructed by empirical quantiles. Again, one of our aims in this research is to construct an alternative control chart by symmetric quantiles and show that it does gain better efficiency than the classical version constructed by empirical quantiles.

1.4   Nonparametric Profile Control Chart

Statistical process control has been widely applied in many areas, especially in industries. Another topic in this research is about profile monitoring. Most statistical process control applications deal with the quality of a process or product can be adequately represented by the distribution of a univariate quality characteristic or a vector of correlated quality characteristics. But in many practical situations, the focus point is the relationship between a response variable and some explanatory variables. Thus, at each sampling stage, a collection of data points that can represent the relationship is a data curve or data profile. And thus a control chart designed for profile data is called a profile control chart.

The monitoring of process/product profiles is presently a growing and promising area of research in statistical process control. One of the aims in this re-

search is to develop monitoring schemes for nonlinear profiles with random effects. We utilize the technique of principal components analysis to analyze the covariance structure of the profiles and propose monitoring schemes based on principal component (PC) scores.

Profile monitoring is a relatively new research area in quality control. Kang and Albin (2000) studied the problem of linear profile monitoring and proposed two control schemes by modelling the profiles with the simple linear regression model, $Y = A_0 + A_1 x + \epsilon$, where Y is the response variable and x is the independent variable; A_0 and A_1 are the parameters to be estimated; the noise variables ϵ's are independent and normally distributed with mean zero and common variance σ^2. By centering the x-values to make the least squares estimators of the Y-intercept (A_0) and slope (A_1) independent of each other, Kim et al. (2003) proposed a combined-chart scheme in which three EWMA charts designed respectively for detecting shifts in intercept, slope, and standard deviation (σ) are used simultaneously. Mahmound and Woodall (2004) presented and compared several control charts for Phase I analysis of linear profiles and applied some of the charts to a calibration application. For more discussions on linear profile monitoring, see the review paper by Woodall et al. (2004).

Shiau and Weng (2004) extended the above linear profile monitoring schemes to a scheme suitable for profiles of more general forms via nonparametric regression. No assumptions are made for the form of the profiles except the smoothness. The nonparametric regression model considered is $Y = g(x) + \epsilon$,

8

where $g(x)$ is a smooth function and ϵ is the random error as before. Spline regression was adopted as the curve fitting/smoothing technique for its simplicity. They proposed an EWMA chart for detecting mean shifts, an R chart for variation changes, and an EWMSD (standard deviation) chart for variation increases.

Note that the models described above all consist of a deterministic line/curve plus random noises. It does not account for some allowable profile-to-profile variations that we often observe in many profile data, e.g., the aspartame example and VDP example, where these profile-to-profile variations should be considered as caused by common causes. A monitoring scheme constructed based on the afore-mentioned "fixed-effect" model may interpret these common-cause variations as caused by some special causes and signal many false alarms. Thus, we need a suitable model that can cope with these common-cause variations and construct a monitoring scheme accordingly.

For this, Shiau, Lin, and Chen (2006) considered a random-effect linear model to develop monitoring schemes for linear profiles. Similarly, Jensen et al. (2006) proposed a linear mixed (effects) model for linear profiles. Williams et al. (2003) fitted nonlinear profiles by nonlinear parametric regression and then monitored profiles with some T^2 statistics of the estimated parameters. Later, Williams et al. (2007) extended this methodology to nonlinear profiles with a nonconstant variance at set points to analyze a set of heteroscedastic dose-response profiles. Adopting a random-effect parametric nonlinear regres-

sion model for profiles, Shiau, Yen, and Feng (2006) proposed a robust nonlinear profile monitoring scheme. Jensen et al. (2009) proposed using nonlinear mixed models to model nonlinear profiles. Note that the parametric approaches mentioned above all need to pre-specify a parametric functional form for profiles, a task often not so easy for practitioners. Qiu, Zou and Wang (2010) proposed a novel control chart, which dealt with mixed effect profile data without assumptions on a parametric functional form. Their control chart is based on local linear kernel smoothing of profile data and on the EWMA weighting scheme at different time points and the heteroscedasticity of observations within each profile.

We extend the nonparametric fixed-effect model of Shiau and Weng (2004) to a random-effect model in order to incorporate some profile-to-profile variability as caused by common causes. With the random-effect model, we focus on the covariance structure and use the principal components analysis (PCA) to analyze it. Ding et al. (2006) also considered modelling profiles nonparametrically for a Phase I analysis, but proposed using ICA (independent components analysis) instead of PCA for monitoring profiles that are in clusters, a situation PCA may fail to preserve the clustering feature of the original data.

PCA is very useful in summarizing and interpreting a set of profile data with the same equally spaced x-values for each profile. We remark that the smoothing step described above can relax this requirement for profile data since the equally-spaced data can be obtained from the smoothed profiles easily.

Some pioneer works on analyzing curves with PCA include Castro et al. (1986), Rice and Silverman (1991), Jones and Rice (1992), and others. For applications, Shiau and Lin (1999) analyzed a set of accelerated LED degradation profiles to estimate the mean lifetime of the product with the techniques of nonparametric regression and PCA.

For Phase I profile monitoring, we propose using the usual Hotelling T^2 chart, a commonly used control chart designed for multivariate process data, by treating the principal component (PC) scores of a profile obtained from PCA as the multivariate data. For Phase II process monitoring, we propose and study three monitoring schemes constructed by utilizing the eigenvalues and eigenvectors obtained from PCA to compute the PC-scores of each incoming profile, including individual PC-score charts, a combined chart that combines all of the PC-score charts and a T^2 chart (different from the T^2 chart of Phase I). The performances of these monitoring schemes are evaluated in terms of the average run length (ARL).

Chapter 2

A Nonparametric Coverage Interval

In this chapter, we propose a coverage interval estimation based on symmetric quantiles. The coverage precision of the proposed symmetric coverage interval is studied and comparisons with empirical quantile coverage interval is also conducted to demonstrate that the symmetric quantile coverage interval is superior to its empirical counterpart in practical usage when the underlying distribution is symmetric.

2.1 Symmetric Coverage Interval

For random variable Y with cumulative distribution function F, the λth quantile is defined as

$$F^{-1}(\lambda) = \inf\{c : F(c) \geq \lambda\}.$$

The classical $1 - \alpha$ coverage interval is defined as

$$C(1 - \alpha) = (F^{-1}(\frac{\alpha}{2}), F^{-1}(1 - \frac{\alpha}{2})).$$

Suppose that we now have a random sample $y_1, ..., y_n$ from distribution F. The corresponding empirical $1 - \alpha$ coverage interval is

$$C_n(1 - \alpha) = (F_n^{-1}(\frac{\alpha}{2}), F_n^{-1}(1 - \frac{\alpha}{2})) \tag{2.1}$$

where we let F_n^{-1} be the empirical quantile, a quantile function with distribution function of the sample type as $F_n(y) = \frac{1}{n} \sum_{i=1}^{n} I(y_i \leq y)$.

Unlike the way in which the empirical quantile is constructed based on the cumulative distribution function, the so-called symmetric quantile of Chen and Chiang (1996) is formulated based on a folded distribution function. Let us consider the folded cumulative function about μ, known or unknown, as

$$F_s(a) = P(|y - \mu| \leq a), a \geq 0.$$

Extending from Chen and Chiang (1996), we define the $1 - \alpha$ symmetric coverage interval as

$$C_s(1 - \alpha) = (\mu - F_s^{-1}(1 - \alpha), \mu + F_s^{-1}(1 - \alpha))$$

where $F_s^{-1}(\lambda) = \inf\{a : F_s(a) \geq \lambda\}$. If F is continuous, the $1 - \alpha$ symmetric coverage interval satisfies $1 - \alpha = P(\mu - F_s^{-1}(1 - \alpha) \leq y \leq \mu + F_s^{-1}(1 - \alpha))$. If we further assume that F is symmetric at μ, it can be seen that

$$C_s(1 - \alpha) = C(1 - \alpha), \tag{2.2}$$

13

the classical one and the symmetric one are identical in the sense of containing the same set of reference individuals.

We interpret the folded cumulative function and the symmetric coverage interval through Figure2.1.
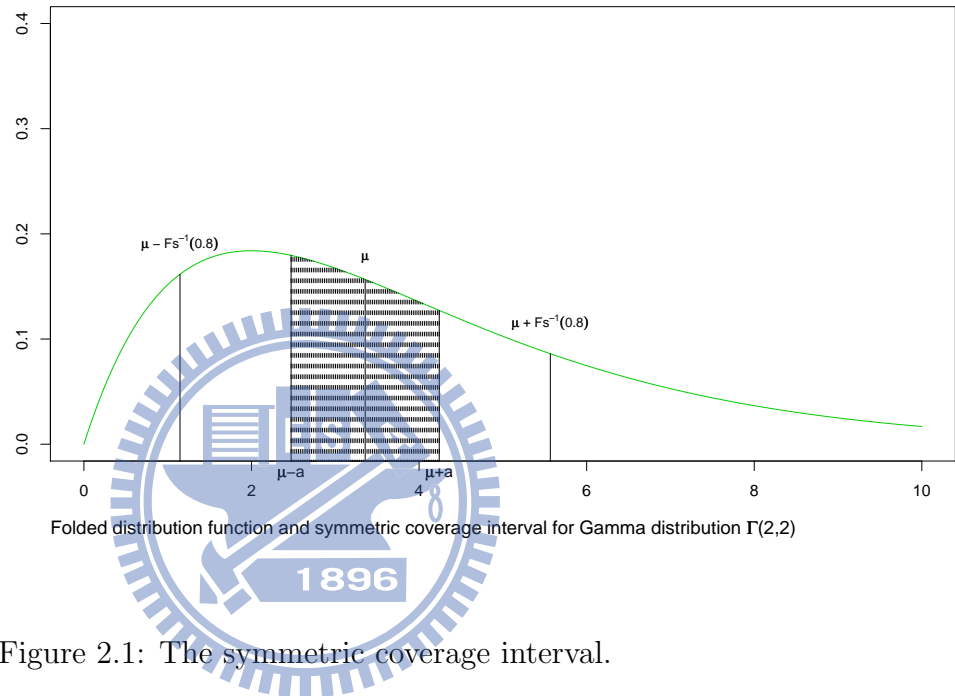
Figure 2.1: The symmetric coverage interval.

Considering the Gamma distribution $\Gamma(2, 2)$ which has probability density function (pdf) as the curve in the figure, we consider the folded distribution about the median. With this distribution, the median μ is 3.36. For a given $a > 0$, the value of this folded distribution at a represents the probability of a region as the part of shadow. Suppose that the interest is 80% coverage interval. For this continuous distribution, we search $F_s^{-1}(0.8) = a^*$ such that $0.8 = P(\mu - a^* \leq y \leq \mu + a^*)$ with $y \sim \Gamma(2, 2)$ which indicates that $a^* = 2.21$.

Hence the 80% symmetric coverage interval is

$$C_s(0.8) = (\mu - F_s^{-1}(0.8), \mu + F_s^{-1}(0.8))$$

$$= (1.15, 5.57)$$

Let $\hat{\mu}$ be an estimate of μ. We may define the sample type $1 - \alpha$ symmetric coverage interval as

$$C_{sn}(1 - \alpha) = (\hat{\mu} - F_{sn}^{-1}(1 - \alpha), \hat{\mu} + F_{sn}^{-1}(1 - \alpha)) \qquad (2.3)$$

where $F_{sn}(a) = \frac{1}{n}\sum_{i=1}^{n} I(|y_i - \hat{\mu}| \leq a)$ is the sample type folded cumulative distribution function and $F_{sn}^{-1}(1 - \alpha) = \inf\{a : F_{sn}(a) \geq 1 - \alpha\}$.

Let's give a simple example to describe the construction of sample symmetric coverage interval. Suppose that we have a set of observations that are ordered as

$$-5, -3, -2, -1, -0.5, 0.5, 1, 3, 50, 100.$$

We want to construct 80% empirical and symmetric coverage intervals. With $F_n^{-1}(0.1) = -5$ and $F_n^{-1}(0.9) = 50$, the 80% empirical coverage interval is

$$C_n(0.8) = (-5, 50). \qquad (2.4)$$

For construction of symmetric coverage interval, we choose sample median as the estimate of μ. That is,

$$\hat{\mu} = F_n^{-1}(0.5) = \inf\{a : \frac{1}{10}\sum_{i=1}^{10} I(y_i \leq a) \geq 0.5\} = -0.5.$$

Let's denote residuals $e_i = y_i - \hat{\mu}, i = 1, ..., 10$. The residuals are

$$-4.5, -2.5, -1.5, -0.5, 0, 1, 1.5, 3.5, 50.5, 100.5.$$

The sample type folded cumulative distribution function is

$$F_{sn}(a) = \frac{1}{10} \sum_{i=1}^{10} I(|e_i| \le a).$$

For examples, $F_{sn}(0) = \frac{1}{10}, F_{sn}(1) = \frac{1}{10}[I(|-0.5| \le 1) + I(|0| \le 1) + I(|1| \le 1)] = \frac{3}{10}$. Then we have

$$F_{sn}^{-1}(0.8) = \inf\{a : \frac{1}{10} \sum_{i=1}^{10} I(|e_i| \le a) \ge 0.8\}$$

$$= 4.5.$$

This indicates that the 80% symmetric coverage interval is

$$C_{sn}(0.8) = (\hat{\mu} - F_{sn}^{-1}(0.8), \hat{\mu} + F_{sn}^{-1}(0.8))$$

$$= (-0.5 - 4.5, -0.5 + 4.5)$$

$$= (-5, 4). \tag{2.5}$$

Comparing the resulted sample empirical and symmetric coverage intervals in (2.4) and (2.5), it is seen the benefit for using the latter one for that it is shorter than the former one. This would happen very often when the observations are drawn from asymmetric distributions.

2.2 Precision Study of Symmetric Coverage Interval

The equality of (2.2) does not hold when the underlying distribution F is not symmetric so that there is no fair criterion to compare their corresponding sample coverage intervals. Hence, we may set the case that F is symmetric to compare the precision of these two coverage intervals through the asymptotic variances of their sample type coverage intervals.

We consider that μ is the median parameter and let $\hat{\mu}$ be the sample median as

$$\hat{\mu} = \mathrm{arginf}_{\mu \in R} \sum_{i=1}^{n} |y_i - \mu|.$$

Suppose that we assume that F is continuous and symmetric at μ. From Ruppert and Carroll (1980), we have a Bahadur representation for this sample median as

$$n^{1/2}(\hat{\mu} - \mu) = n^{-1/2} \frac{1}{f(\mu)} \sum_{i=1}^{n} (0.5 - I(y_i \leq \mu)) + o_p(1). \tag{2.6}$$

On the other hand, a Barhadur representation for $F_{sn}^{-1}(1 - \alpha)$ developed by Chen and Chiang (1996) is

$$n^{1/2}(F_{sn}^{-1}(1 - \alpha) - (F^{-1}(1 - \frac{\alpha}{2}) - \mu)) = \frac{1}{2f(F^{-1}(1 - \frac{\alpha}{2}))} n^{-1/2} \sum_{i=1}^{n} \{1 - \alpha$$

$$- I(F^{-1}(\frac{\alpha}{2}) \leq y_i \leq F^{-1}(1 - \frac{\alpha}{2}))\} + o_p(1). \tag{2.7}$$

The assumption of symmetric distribution indicates that $\hat{\mu} - F_{sn}^{-1}(1 - \alpha)$ and $\hat{\mu} + F_{sn}^{-1}(1 - \alpha)$ have the same asymptotic variance and, from (2.6) and (2.7),

we have a Bahadur representation for $\hat{\mu} - F_{sn}^{-1}(1 - \alpha)$ as

$$n^{1/2}((\hat{\mu} - F_{sn}^{-1}(1 - \alpha)) - F^{-1}(\frac{\alpha}{2})) = n^{-1/2} \sum_{i=1}^{n} \{[-\frac{1}{2f(\mu)} - \frac{1 - \alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))}]$$

$$I(y_i \leq F^{-1}(\frac{\alpha}{2})) + [-\frac{1}{2f(\mu)} + \frac{\alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))}]I(F^{-1}(\frac{\alpha}{2}) \leq y_i \leq \mu)$$

$$+ [\frac{1}{2f(\mu)} + \frac{\alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))}]I(\mu < y_i \leq F^{-1}(1 - \frac{\alpha}{2}))$$

$$+ [\frac{1}{2f(\mu)} - \frac{1 - \alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))}]I(y_i \geq F^{-1}(1 - \frac{\alpha}{2}))\} + o_p(1). \tag{2.8}$$

Since $y_1, ..., y_n$ is a random sample from distribution F, we may see that the

asymptotic variance of $n^{1/2}(\hat{\mu} - F_{sn}^{-1}(1 - \alpha) - F^{-1}(\frac{\alpha}{2}))$ is

$$\sigma_s^2 = \frac{\alpha}{2}[(\frac{1}{2f(\mu)} + \frac{1 - \alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))})^2 + (\frac{1}{2f(\mu)}$$

$$- \frac{1 - \alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))})^2] + (1 - \frac{\alpha}{2})[(-\frac{1}{2f(\mu)} + \frac{\alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))})^2$$

$$+ (\frac{1}{2f(\mu)} + \frac{\alpha}{2f(F^{-1}(1 - \frac{\alpha}{2}))})^2]. \tag{2.9}$$

On the other hand, in this situation that y has a continuous and symmetric

distribution, we may see that $n^{1/2}(F_n^{-1}(\frac{\alpha}{2}) - F^{-1}(\frac{\alpha}{2}))$ and $n^{1/2}(F_n^{-1}(1 - \frac{\alpha}{2}) -$

$F^{-1}(1 - \frac{\alpha}{2}))$ also have the same asymptotic variance (see, for example, Sen and

Singer (1993, p168)) as

$$\sigma_e^2 = \frac{\alpha}{2}(1 - \frac{\alpha}{2})f^{-2}(F^{-1}(1 - \frac{\alpha}{2})). \tag{2.10}$$

Since these two sample coverage intervals estimate the same population cov-

erage interval, it is fair that we evaluate the efficiency of the symmetric type

coverage interval defined as the following

$$Eff = \frac{\sigma_e^2}{\sigma_s^2}. \qquad (2.11)$$

Let's consider several distributions for computation of asymptotic variances of (2.9) and (2.10) to compare their corresponding efficiencies of (2.11) where distributions include standard normal distribution $N(0, 1)$, t-distribution $t(r)$ where r is the degrees of freedom, Cauchy distribution $(Cauchy(s), s > 0)$ with pdf

$$f(y) = \frac{1}{\pi}\frac{s}{y^2 + s^2}, y \in R$$

and the Laplace distribution $(Lap(b))$ with pdf

$$f(y) = \frac{1}{2b}e^{-\frac{|y|}{b}}, y \in R.$$
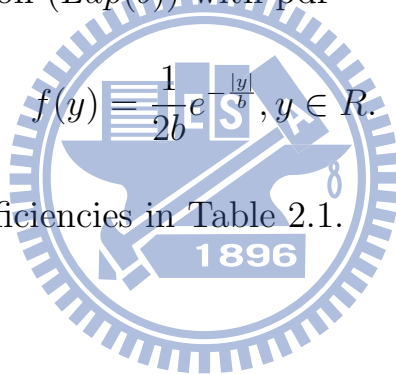
We display the resulted efficiencies in Table 2.1.

19

Table 2.1. The efficiencies, Eff, of symmetric coverage interval

dist/$1-\alpha$	0.6	0.8	0.9	0.95	0.98
$N(0,1)$	0.87	0.87	1.02	1.21	1.48
$t(r)$					
$r=1$	0.98	1.78	2.13	2.1	2.04
$r=5$	0.89	1.01	1.31	1.61	1.85
$r=10$	0.88	0.94	1.16	1.42	1.7
$Cauchy(s)$ $s=1,5,10$	0.98	1.78	2.13	2.1	2.04
$Lap(b)$ $b=1,5,10$	1.2	1.6	1.8	1.9	1.96

It is relatively efficient to use the empirical quantile to construct coverage interval when the quantile percentage is close to 0.5 in either direction. This means that when we want a $1-\alpha$ coverage interval with coverage probability $1-\alpha$ as the value of 0.6 or even smaller. The one estimated by empirical quantiles is the right choice. On the other hand, we see that it gains more precision to use symmetric quantile to construct coverage interval when $1-\alpha$ is with the value of 0.8 or larger. This alternative coverage interval is then attractive since it is very common that we apply coverage interval only for large $1-\alpha$, for example, the reference interval in medical diagnosis chooses the value of 0.95. In fact, the case that when the underlying distribution is the Laplace one the coverage interval constructed by symmetric quantiles totally dominate the one

by empirical quantiles.

This interesting result is not surprising. The surprising fact is that, unlike estimation of location and scale parameters that have been received very much attention in statistical literature for proposing techniques and developing theories in gaining better precisions, the attention for developing alternative ways in constructing coverage intervals for gaining better precision than the classical one has not been paid in statistical and metrological literatures.

2.3   Concluding Remarks
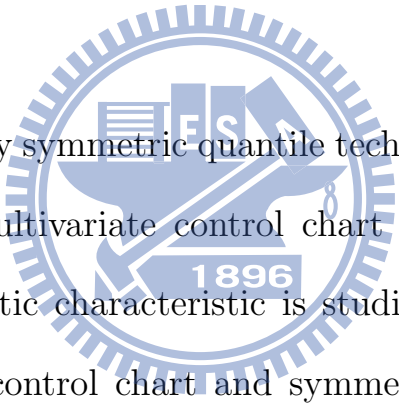
As for constructing a coverage interval, a symmetric quantile coverage interval performs better than the commonly used empirical quantile coverage interval both empirically and theoretically. The robust estimation on the distribution mass and the symmetric folding feature of the symmetric quantile coverage interval prevent the estimation of coverage interval heavily impacted by outliers. When an underlying distribution is not symmetric, with center on the distribution median, using symmetric quantiles to construct a coverage interval can cover more percentage of the higher density part compared with the other side with lower distribution density, and thus a symmetric quantile coverage interval gives a shorter coverage interval than the empirical one. Even when the distribution is symmetric, we still can see the symmetric quantile coverage interval is with smaller asymptotical variance than the empirical one

for most coverage intervals with large $1 - \alpha$.
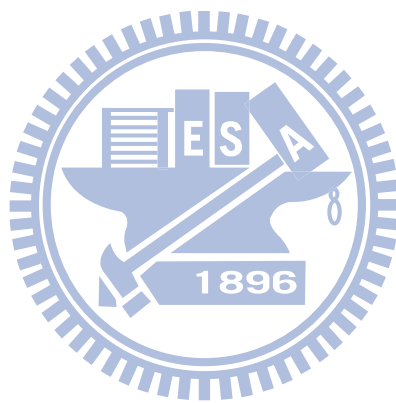
Chapter 3

Multivariate Control Chart by Symmetric Quantiles

In this chapter, we apply symmetric quantile techniques in constructing control chart schemes. A multivariate control chart by symmetric quantiles is proposed and its asymptotic characteristic is studied. Comparisons between empirical quantile based control chart and symmetric quantile control chart are also studied through efficiency in terms of asymptotic variances and ARL. Symmetric quantile based multivariate control chart is more efficient than the empirical quantile based multivariate control chart in detecting small distributional shifts even when underlying distribution is not symmetric.

3.1 Empirical Quantiles Based Control Chart for Monitoring Quantile Vector

With the interest of control charts comparison, let us first introduce, in detail, the empirical quantiles based control chart of Grimshaw and Alt (1997). For percentages $\alpha_1, ..., \alpha_k$ with $\alpha_1 < \alpha_2 < ... < \alpha_k$, let us consider the population quantile vector

$$Q(\alpha_1, ..., \alpha_k) = \begin{pmatrix} F^{-1}(\alpha_1) \\ F^{-1}(\alpha_2) \\ \vdots \\ F^{-1}(\alpha_k) \end{pmatrix} \tag{3.1}$$

for monitoring that can be estimated by the corresponding empirical quantile vector

$$Q_e(\alpha_1, ..., \alpha_k) = \begin{pmatrix} F_n^{-1}(\alpha_1) \\ F_n^{-1}(\alpha_2) \\ \vdots \\ F_n^{-1}(\alpha_k) \end{pmatrix}.$$

Grimshaw and Alt (1997) proposed to apply $Q_e(\alpha_1, ..., \alpha_k)$ to monitor the population quantile vector $Q(\alpha_1, ..., \alpha_k)$. The asymptotic property of $Q_e(\alpha_1, ..., \alpha_k)$ relies on the empirical quantiles $F_n^{-1}(\alpha_j)$'s. A Bahhadur representation for $F_n^{-1}(\alpha)$ (see, for example, Ruppert and Carroll (1980)) is

$$\sqrt{n}(F_n^{-1}(\alpha) - F^{-1}(\alpha)) = \frac{1}{f(F^{-1}(\alpha))} n^{-1/2} \sum_{i=1}^{n} (\alpha - I(y_i \leq F^{-1}(\alpha))) + o_p(1)$$

where f is the probability density function of variable y and I is the indicator function. This leads to that $n^{1/2}(Q_e(\alpha_1,...,\alpha_k) - Q(\alpha_1,...,\alpha_k))$ converges to a normal distribution $N_k(0_k, \Sigma)$ with $\Sigma = (\sigma^e_{jk})$ and where $\sigma^e_{ij} = \frac{\alpha_i(1-\alpha_j)}{f(F^{-1}(\alpha_i))f(F^{-1}(\alpha_j))}$ for $i \leq j$. This further implies that the following

$$n(Q_e(\alpha_1,...,\alpha_k) - Q(\alpha_1,...,\alpha_k))'\Sigma^{-1}(Q_e(\alpha_1,...,\alpha_k) - Q(\alpha_1,...,\alpha_k)) \to \chi^2(k)$$

holds asymptotically in distribution.

Suppose that we have a training sample $y_{ij}, i = 1,...,n, j = 1,...,m$ that represents an in-control data set of m samples of size n from distribution F so that estimate of $Q(\alpha_1,...,\alpha_k)$ and Σ are available. Generally we let $Q_{ej}(\alpha_1,...,\alpha_k)$ and $\hat{\Sigma}_j$ be estimates, respectively, based on sample $y_{ij}, i = 1,...,n$ and define $Q_0(\alpha_1,...,\alpha_k) = \frac{1}{m}\sum_{j=1}^{m} Q_{ej}(\alpha_1,...,\alpha_k)$ and $\Sigma_0 = \frac{1}{m}\sum_{j=1}^{m} \hat{\Sigma}_j$. Treated estimates Q_0 and Σ_0 as true values of $Q(\alpha_1,...,\alpha_k)$ and Σ, the control statistic and upper control limit proposed by Grimshaw and Alt (1997) are

Control statistic $T_e = n(Q_e(\alpha_1,...,\alpha_k) - Q_0(\alpha_1,...,\alpha_k))'\Sigma_0^{-1}(Q_e(\alpha_1,...,\alpha_k)$

$- Q_0(\alpha_1,...,\alpha_k))$

$UCL_e = \chi^2_\alpha$ \hfill (3.2)

\hfill (3.3)

where χ^2_α satisfies $1 - \alpha = P(\chi^2(k) \leq \chi^2_\alpha)$. With this proposal, if a sample point $y_1,...,y_n$ has a value of T_e lying below the upper limit UCL_e and a set of $T_e's$ does not exhibit any systematic pattern, we say that the process is in statistical control at the level $1 - \alpha$.

With $Q(\alpha_1, ..., \alpha_k)$ of (3.1) as the target to be monitored, the idea behind this study is why should we estimate it by the empirical quantiles.

3.2 Symmetric Quantile Control Chart

From the study in chapter 2, we anticipate that using the symmetric quantiles to estimate the distributional characteristic vector $Q(\alpha_1, ..., \alpha_k)$ is more efficient than that of empirical quantiles. Considering a number ℓ decreasing percentages $\gamma_1 > \gamma_2 > ... > \gamma_\ell$, we define its corresponding 2ℓ symmetric quantile vector and population symmetric quantiles, respectively, as

$$
Q_{sn}(\gamma_1, ..., \gamma_\ell) = \begin{pmatrix} F_{sn}^{(-)}(\gamma_1) \\ F_{sn}^{(-)}(\gamma_2) \\ \vdots \\ F_{sn}^{(-)}(\gamma_\ell) \\ F_{sn}^{(+)}(\gamma_\ell) \\ F_{sn}^{(+)}(\gamma_{\ell-1}) \\ \vdots \\ F_{sn}^{(+)}(\gamma_1) \end{pmatrix} \text{ and } Q_s(\gamma_1, ..., \gamma_\ell) = \begin{pmatrix} F_s^{(-)}(\gamma_1) \\ F_s^{(-)}(\gamma_2) \\ \vdots \\ F_s^{(-)}(\gamma_\ell) \\ F_s^{(+)}(\gamma_\ell) \\ F_s^{(+)}(\gamma_{\ell-1}) \\ \vdots \\ F_s^{(+)}(\gamma_1) \end{pmatrix}.
$$

where, the γ symmetric quantile pair is defined as

$$
\{F_s^{(-)}(\gamma), F_s^{(+)}(\gamma)\} = \{\mu - F_s^{-1}(\gamma), \mu + F_s^{-1}(\gamma)\} \tag{3.4}
$$

. For the random sample $y_1, ..., y_n$ from distribution F, let $\hat{\mu}$ be an estimate of

μ. We may define the sample type γ symmetric quantile pair as

$$\{F_{sn}^{(-)}(\gamma), F_{sn}^{(+)}(\gamma)\} = \{\hat{\mu} - F_{sn}^{-1}(\gamma), \hat{\mu} + F_{sn}^{-1}(\gamma)\} \qquad (3.5)$$

with sample folded distribution function $F_{sn}(a) = \frac{1}{n}\sum_{i=1}^{n} I(|y_i - \hat{\mu}| \leq a)$

and

$$F_{sn}^{-1}(\gamma) = \inf\{a : F_{sn}(a) \geq \gamma\}. \qquad (3.6)$$

From Chen and Chiang (1996), we may see that $n^{1/2}(Q_{sn}(\gamma_1, ..., \gamma_\ell) - Q_s(\gamma_1, ..., \gamma_\ell))$ is asymptotically normal $N_{2\ell}(0_{2\ell}, \Sigma_s)$ for some matrix Σ_s that will be given explicitly latter. This further implies that the following

$$n(Q_{sn}(\gamma_1, ..., \gamma_\ell) - Q_s(\gamma_1, ..., \gamma_\ell))'\Sigma_s^{-1}(Q_{sn}(\gamma_1, ..., \gamma_\ell)$$
$$- Q_s(\gamma_1, ..., \gamma_\ell)) \to \chi^2(2\ell)$$

holds asymptotically in distribution.

Again, from the training sample $y_{ij}, i = 1, ..., n, j = 1, ..., m$ that represents an in-control data set of m samples of size n, we let $Q_{s0}(\gamma_1, ..., \gamma_\ell) = \frac{1}{m}\sum_{j=1}^{m} Q_{sn,j}(\gamma_1, ..., \gamma_\ell)$ and $\Sigma_{s0} = \frac{1}{m}\sum_{j=1}^{m} \hat{\Sigma}_{s,j}$ where $Q_{sn,j}(\gamma_1, ..., \gamma_\ell)$ and $\hat{\Sigma}_{s,j}$ are estimates of, respectively, $Q_s(\gamma_1, ..., \gamma_\ell)$ and Σ_s. Let us denote these two estimates by Q_{s0} and Σ_{s0}. Based on these estimates, we proposed control statistic

27

and upper control limit as

$$\text{Control statistic } T_s = n(Q_{sn}(\gamma_1, ..., \gamma_\ell) - Q_{s0}(\gamma_1, ..., \gamma_\ell))'\Sigma_{s0}^{-1}$$

$$(Q_{sn}(\gamma_1, ..., \gamma_\ell) - Q_{s0}(\gamma_1, ..., \gamma_\ell))$$

$$UCL = \chi_\alpha^2(2\ell) \tag{3.7}$$

The asymptotic covariance matrix Σ_s varies with the predetermined estimator of μ. From a comparison in Chen and Chiang (1996), although various estimators of μ lead to different asymptotic distributions for their corresponding symmetric quantiles, however, their performances in constructing statistical procedures such as trimmed means are very competitive. Hence from here after, for simplicity, we consider that μ is the median parameter and let $\hat{\mu}$ be the ℓ_1-norm estimate as

$$\hat{\mu} = \text{arginf}_{\mu \in R} \sum_{i=1}^n |y_i - \mu|.$$

A description of the asymptotic covariance matrix Σ_s under this setting of predetermined estimator will be developed in next subsection.

3.3 Derivation of Asymptotic Covariance Matrix of Symmetric Quantile Vector

We assume that F is continuous and symmetric at median μ. From Ruppert and Carroll (1980), we have a Bahadur representation for this ℓ_1-norm estimate

as

$$n^{1/2}(\hat{\mu} - \mu) = n^{-1/2}\frac{1}{f(\mu)}\sum_{i=1}^{n}(0.5 - I(y_i \leq \mu)) + o_p(1). \qquad (3.8)$$

Furthermore, a Barhadur representation for $F_{sn}^{-1}(\gamma)$ of (3.6) that has been developed by Chen and Chiang (1996) is

$$n^{1/2}(F_{sn}^{-1}(\gamma) - F^{-1}(\gamma)) = \frac{1}{2f(F^{-1}(\frac{1+\gamma}{2}))}n^{-1/2}\sum_{i=1}^{n}(\gamma$$

$$- I(-F^{-1}(\frac{1+\gamma}{2}) \leq y_i \leq F^{-1}(\frac{1+\gamma}{2})) + o_p(1). \qquad (3.9)$$

From (3.8) and (3.9), Bahadur representations, respectively, for $F_{sn}^{(-)}(\gamma)$ and $F_{sn}^{(+)}(\gamma)$ are

$$n^{1/2}(F_{sn}^{(-)}(\gamma) - F^{-1}(\frac{1-\gamma}{2})) = n^{-1/2}\sum_{i=1}^{n}\{[-\frac{1}{2f(\mu)} - \frac{\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]$$

$$I(y_i \leq F^{-1}(\frac{1-\gamma}{2})) + [-\frac{1}{2f(\mu)} + \frac{1-\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(F^{-1}(\frac{1-\gamma}{2}) \leq y_i \leq \mu)$$

$$+ [\frac{1}{2f(\mu)} + \frac{1-\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(\mu < y_i \leq F^{-1}(\frac{1+\gamma}{2}))$$

$$+ [\frac{1}{2f(\mu)} - \frac{\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(y_i \geq F^{-1}(\frac{1+\gamma}{2}))\} + o_p(1). \qquad (3.10)$$

and

$$n^{1/2}(F_{sn}^{(+)}(\gamma) - F^{-1}(\frac{1+\gamma}{2})) = n^{-1/2}\sum_{i=1}^{n}\{[-\frac{1}{2f(\mu)} + \frac{\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]$$

$$I(y_i \leq F^{-1}(\frac{1-\gamma}{2})) + [-\frac{1}{2f(\mu)} - \frac{1-\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(F^{-1}(\frac{1-\gamma}{2}) \leq y_i \leq \mu)$$

$$+ [\frac{1}{2f(\mu)} - \frac{1-\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(\mu < y_i \leq F^{-1}(\frac{1+\gamma}{2}))$$

$$+ [\frac{1}{2f(\mu)} + \frac{\gamma}{2f(F^{-1}(\frac{1+\gamma}{2}))}]I(y_i \geq F^{-1}(\frac{1+\gamma}{2}))\} + o_p(1). \qquad (3.11)$$

By letting $\ell = \frac{k}{2}$ and $\alpha_1, ..., \alpha_k$ satisfying $1-\alpha_1 = \alpha_k, 1-\alpha_2 = \alpha_{k-1}, ..., 1-\alpha_{k/2} = \alpha_{\frac{k}{2}+1}$, then $Q_s(1-2\alpha_1, 1-2\alpha_2, ..., 1-2\alpha_{\frac{k}{2}}) = Q(\alpha_1, \alpha_2, ..., \alpha_k)$. This allows us to compare estimators of $Q(\alpha_1, ..., \alpha_k)$ that are constructed by empirical quantiles and symmetric quantiles. For $\alpha_i < \alpha_j$, we need to develop the asymptotic variances and covariances for $F_{sn}^{(-)}(1 - 2\alpha_i)$ and $F_{sn}^{(+)}(1 - 2\alpha_j)$ for $i \neq j$. Let's denote $\sigma_{\alpha_1}^2, \sigma_{1-\alpha_1}^2, \sigma_{\alpha_1 1-\alpha_2}$ as the asymptotic variance of $n^{1/2}(F_{sn}^{(-)}(1 - 2\alpha_1) - F^{-1}(\alpha_1))$ and $n^{1/2}(F_{sn}^{(+)}(1 - 2\alpha_1) - F^{-1}(1 - \alpha_1))$ and asymptotic covariance of $n^{1/2}(F_{sn}^{(-)}(1 - 2\alpha_1) - F^{-1}(\alpha_1))$ and $n^{1/2}(F_{sn}^{(+)}(1 - 2\alpha_2) - F^{-1}(1 - \alpha_2))$. Since $y_1, ..., y_n$ is a random sample from distribution F, we may derive the followings from (3.10) and (3.11),

$$\sigma_{\alpha_1}^2 = \sigma_{1-\alpha_1}^2, \sigma_{\alpha_1 1-\alpha_2} = \sigma_{1-\alpha_2 \alpha_1}, \sigma_{\alpha_1 \alpha_2} = \sigma_{\alpha_2 \alpha_1}, \sigma_{1-\alpha_1 1-\alpha_2} = \sigma_{1-\alpha_2 1-\alpha_1},$$

$$\sigma_{\alpha_1}^2 = (\frac{1}{2f(\mu)})^2 + 2\alpha_1(1 - 2\alpha_1)(\frac{1}{2f(F^{-1}(1 - \alpha_1))})^2,$$

$$\sigma_{\alpha_1 \alpha_2} = (\frac{1}{2f(\mu)})^2 + 2\alpha_1(1 - 2\alpha_2)(\frac{1}{4f(F^{-1}(1 - \alpha_1))f(F^{-1}(1 - \alpha_2))}),$$

$$(3.12)$$

$$\sigma_{\alpha_1 1-\alpha_2} = \begin{cases} (\frac{1}{2f(\mu)})^2 - 2\alpha_1(1 - 2\alpha_2)(\frac{1}{4f(F^{-1}(1-\alpha_1))f(F^{-1}(1-\alpha_2))}) & \text{if } \alpha_1 < \alpha_2 \\ (\frac{1}{2f(\mu)})^2 - 2\alpha_2(1 - 2\alpha_1)(\frac{1}{4f(F^{-1}(1-\alpha_1))f(F^{-1}(1-\alpha_2))}) & \text{if } \alpha_1 > \alpha_2 \end{cases},$$

$$\sigma_{1-\alpha_1 1-\alpha_2} = (\frac{1}{2f(\mu)})^2 + 2\alpha_1(1 - 2\alpha_2)(\frac{1}{4f(F^{-1}(1 - \alpha_1))f(F^{-1}(1 - \alpha_2))}).$$

With careful arrangement and derivations from (3.8) - (3.12) and denoting $\alpha_{ij} = \alpha_i(1 - \alpha_j)$ and $f_{ij} = f(F^{-1}(1 - \frac{\alpha_i}{2}))f(F^{-1}(1 - \frac{\alpha_j}{2}))$, the asymptotic covariance

30

matrix for the symmetric quantile vector $Q_{sn}(1 - 2\alpha_1, 1 - 2\alpha_2, ..., 1 - 2\alpha_\ell)$ is stated in the following theorem.

Theorem 3.3.1 When the distribution F is symmetric about its median μ, then the asymptotic covariance matrix for the symmetric quantile vector is

$$\Sigma_s = \frac{1}{4f(\mu)^2} 11' + \frac{1}{4} \left(\begin{pmatrix} A_1 \\ A_1' \\ A_3' \\ A_2' \end{pmatrix} \begin{pmatrix} A_2 \\ A_3 \\ A_4 \\ A_4' \end{pmatrix} \right)$$

where

$$A_1 = \begin{pmatrix} \frac{\alpha_{11}}{f_{11}} & \frac{\alpha_{12}}{f_{12}} & \cdots & \frac{\alpha_{1(\ell-1)}}{f_{1(\ell-1)}} & \frac{\alpha_{1\ell}}{f_{1\ell}} \\ & \frac{\alpha_{22}}{f_{22}} & \cdots & \frac{\alpha_{2(\ell-1)}}{f_{2(\ell-1)}} & \frac{\alpha_{2\ell}}{f_{2\ell}} \\ & & \ddots & \vdots & \vdots \\ & & & \frac{\alpha_{(\ell-1)(\ell-1)}}{f_{(\ell-1)(\ell-1)}} & \frac{\alpha_{(\ell-1)\ell}}{f_{(\ell-1)\ell}} \\ & & & & \frac{\alpha_{\ell\ell}}{f_{\ell\ell}} \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -\frac{\alpha_{1\ell}}{f_{1\ell}} & -\frac{\alpha_{1(\ell-1)}}{f_{1(\ell-1)}} & \cdots & -\frac{\alpha_{12}}{f_{12}} & -\frac{\alpha_{11}}{f_{11}} \\ & -\frac{\alpha_{2(\ell-1)}}{f_{2(\ell-1)}} & \cdots & -\frac{\alpha_{22}}{f_{22}} & -\frac{\alpha_{21}}{f_{21}} \\ & & \vdots & \vdots & \vdots \\ & & & -\frac{\alpha_{(\ell-1)2}}{f_{(\ell-1)2}} & -\frac{\alpha_{(\ell-1)1}}{f_{(\ell-1)1}} \\ & & & & -\frac{\alpha_{\ell1}}{f_{\ell1}} \end{pmatrix}$$

$$
A_3 = \begin{pmatrix}
-\dfrac{\alpha_{\ell 2}}{f_{\ell 2}} & & & & \\[2ex]
-\dfrac{\alpha_{\ell 3}}{f_{\ell 3}} & -\dfrac{\alpha_{(\ell-1)3}}{f_{(\ell-1)3}} & & & \\[2ex]
\vdots & \vdots & & & \\[2ex]
-\dfrac{\alpha_{\ell(\ell-1)}}{f_{\ell(\ell-1)}} & -\dfrac{\alpha_{(\ell-1)(\ell-1)}}{f_{(\ell-1)(\ell-1)}} & \cdots & -\dfrac{\alpha_{3(\ell-1)}}{f_{3(\ell-1)}} & \\[2ex]
-\dfrac{\alpha_{\ell\ell}}{f_{\ell\ell}} & -\dfrac{\alpha_{(\ell-1)\ell}}{f_{(\ell-1)\ell}} & \cdots & -\dfrac{\alpha_{3\ell}}{f_{3\ell}} & -\dfrac{\alpha_{2\ell}}{f_{2\ell}}
\end{pmatrix}
$$

$$
A_4 = \begin{pmatrix}
\dfrac{\alpha_{\ell\ell}}{f_{\ell\ell}} & \dfrac{\alpha_{(\ell-1)\ell}}{f_{(\ell-1)\ell}} & \cdots & \dfrac{\alpha_{2\ell}}{f_{2\ell}} & \dfrac{\alpha_{1\ell}}{f_{1\ell}} \\[2ex]
 & \dfrac{\alpha_{(\ell-1)(\ell-1)}}{f_{(\ell-1)(\ell-1)}} & \cdots & \dfrac{\alpha_{2(\ell-1)}}{f_{2(\ell-1)}} & \dfrac{\alpha_{1(\ell-1)}}{f_{1(\ell-1)}} \\[2ex]
 & & \vdots & \vdots & \vdots \\[2ex]
 & & & \dfrac{\alpha_{22}}{f_{22}} & \dfrac{\alpha_{12}}{f_{12}} \\[2ex]
 & & & & \dfrac{\alpha_{11}}{f_{11}}
\end{pmatrix}, \text{ and } \ell = \frac{k}{2}.
$$

3.4 A Comparison between Empirical and Symmetric Quantile Control Charts

When the interest is of constructing a symmetric quantiles based control chart, we wish to establish evidences for supporting the use of this new control chart. Since the symmetric quantile and the empirical quantile are both asymptotically normal and both are consistent for a same population quantile vector $Q(\alpha_1, ..., \alpha_k)$, one interesting criterion in comparing these two control charts is to study the efficiencies in terms of the ratio of sizes of their asymptotic covariance matrices.

Grimshaw and Alt (1997) pointed out that for effective use of a quantile chart in detection of distributional shift we should select quantile percentages α_i's so that their corresponding quantile differences $F_O^{-1}(\alpha_i) - F_I^{-1}(\alpha_i)$, with F_O and F_I respectively representing the distribution function of in-control and likely out-of-control processes, are large. For verification of this concern, Chiang et al. (2006) showed that $\alpha_i's$ should be lie away of 0.5. We consider comparing the estimator of quantile pair

$$\begin{pmatrix} F^{-1}(\alpha) \\ F^{-1}(1-\alpha) \end{pmatrix}. \tag{3.13}$$

The asymptotic covariance matrices of the estimators of empirical quantile and symmetric quantile provide us to define the efficiency as

$$Eff = \frac{\text{Det}\left(Cov\begin{pmatrix} F_n^{-1}(\alpha) \\ F_n^{-1}(1-\alpha) \end{pmatrix}\right)}{\text{Det}\left(Cov\begin{pmatrix} F_{sn}^{(-)}(1-2\alpha) \\ F_{sn}^{(+)}(1-2\alpha) \end{pmatrix}\right)} = \frac{\alpha(1-\alpha)/(f(F^{-1}(\alpha)f(F^{-1}(1-\alpha)))}{\sigma_\alpha^4 - \sigma_{\alpha 1-\alpha}^2} \tag{3.14}$$

where the covariance matrix of empirical quantiles is decribed in subsection 3.3 and that of the symmetric quantiles and notations of σ_α^4 and $\sigma_{\alpha 1-\alpha}^2$ are listed in (3.12). When $Eff > 1$, the estimation of quantile vector (3.13) is more efficient by the symmetric quantiles. On the other hand, if $Eff < 1$, it prefers to estimate quantile vector by empirical quantiles. For computation of efficiencies of (3.14), we consider the distributions including normal distribution, Cauchy

distribution, Laplace distribution ($Lap(b)$) with pdf

$$f(y) = \frac{1}{2b}e^{-\frac{|y|}{b}}, y \in R$$

and the Cauchy distribution with pdf

$$f(x, \delta, \gamma) = \frac{1}{\pi}[\frac{\gamma}{(x-\delta)^2 + \gamma^2}], x \in R.$$

From (3.14), the computation of efficiency requires only the density function and its corresponding distribution function. We display the resulted efficiencies in Table 3.1.

Table 3.1. The efficiencies, Eff, for symmetric quantiles based estimator

Dist/$1 - 2\alpha$	0.6	0.7	0.8	0.9	0.95	0.98
Normal	0.85	0.79	0.81	1.03	1.50	2.72
Laplace	1.25	1.67	2.5	5	10	25
Cauchy	0.98	1.4	3.35	21.93	166.98	2573
$t(r)$						
$r = 1$	0.98	1.4	3.35	21.93	166.98	2573
$r = 2$	0.9	1.01	1.53	4.29	14.58	83.01
$r = 5$	0.87	0.86	1.02	1.75	3.54	10.24

The estimation of quantile vector (3.13) by empirical quantiles is more efficient when the quantile percentage is 0.6. However, it is impressed that it gains more precision to use symmetric quantile to construct the quantile vector estimator when percentage is equal or more than 0.8. In fact, the case that when

34

the underlying distribution is the Laplace one the estimator of quantile vector constructed by symmetric quantiles totally dominate the one by empirical quantiles.

3.5 Comparison of Average Run Length for Two Quantile Charts

The average run length (ARL), representing the average number of samples taken before an action signal is given, is the most popular technique in evaluating a control chart or comparison of alternative control charts. In this section, we make further comparison of $ARL's$ for control charts constructed by empirical quantile and symmetric quantile.

Let $\alpha = (\alpha_1, ..., \alpha_k)'$ be the percentage vector considered for constructing a quantile control chart. Let $Q_n(\alpha)$ be the quantile control chart for monitoring the distributional shift of random variable X. We assume that when the process is in control, the quantile control chart satisfies the following asymptotic property

$$n(Q_n(\alpha) - Q_0(\alpha))'\Sigma_0^{-1}(Q_n(\alpha) - Q_0(\alpha)) \to \chi^2(k) \text{ in distribution.}$$

Hence we actually consider the following hypothesis

$$H_0 : Q(\alpha) = Q_0(\alpha) \text{ vs } Q(\alpha) \neq Q_0(\alpha), \tag{3.15}$$

where $Q(\alpha)$ is the true population quantile vector estimated by statistic $Q_n(\alpha)$. Suppose that the significance level for the control chart is α. Then, the quantile control chart indicates rejecting H_0 when

$$n(Q_n(\alpha) - Q_0(\alpha))'\Sigma_0^{-1}(Q_n(\alpha) - Q_0(\alpha) \geq \chi_\alpha^2(k).$$

We note here that control chart by symmetric quantiles and by empirical quantiles have the same vector $Q_0(\alpha)$ but have different asymptotic covariance matrices Σ_0.

To study the performance of the control charts in terms of ARL, we consider linear function $aX + b$ for representation of distributional shift. Note that the population quantile vector and covariance matrix for $aX + b$ are, respectively, $Q_{aX+b}(\alpha) = aQ_0(\alpha) + b$ and $\Sigma_{aX+b} = a^2\Sigma_0$. Then the probability that we claim for an out of control under the linear function $aX + b$ is

$$P_{Q_{aX+b}(\alpha), \Sigma_{aX+b}}(n(Q_n(\alpha) - Q_0(\alpha))'\Sigma_0^{-1}(Q_n(\alpha) - Q_0(\alpha)) \geq \chi_\alpha^2(k))$$

$$= P_{Q_{aX+b}(\alpha), \Sigma_{aX+b}}(n(Q_n(\alpha) - Q_{aX+b}(\alpha) + (-(1-a)Q_0(\alpha) + b))'\Sigma_{aX+b}^{-1}$$

$$(Q_n(\alpha) - Q_{aX+b}(\alpha) + (-(1-a)Q_0(\alpha) + b)) \geq \frac{1}{a^2}\chi_\alpha^2(k))$$

$$= P(\chi^2(k, \frac{n}{a^2}(-(1-a)Q_0(\alpha) + b)'\Sigma_0^{-1}(-(1-a)Q_0(\alpha) + b)) \geq \frac{1}{a^2}\chi_\alpha^2(k))$$

$$(3.16)$$

$$(3.17)$$

since $\sqrt{n}\Sigma_{aX+b}^{-1/2'}(Q_n(\alpha) - Q_{aX+b}(\alpha) + (-(1-a)Q_0(\alpha) + b)) \stackrel{\text{app}}{\sim} N_k(\sqrt{n}\Sigma_{aX+b}^{-1/2'}(-(1-a)Q_0(\alpha) + b), I_k) = N_k(\frac{\sqrt{n}}{a}\Sigma_0^{-1/2'}(-(1-a)Q_0(\alpha) + b), I_k)$ where $\chi^2(k, \lambda)$ has a

noncentral chi-square distribution with non-centrality parameter λ. The ARL for quantile vector $Q_n(\alpha)$ is

$$ARL = \frac{1}{P(\chi_k^2(\lambda = \frac{n}{a^2}(-(1-a)Q_0(\alpha)+b)'\Sigma_0^{-1}(-(1-a)Q_0(\alpha)+b)) \geq \frac{1}{a^2}\chi_\alpha^2(k))}$$

where a level α control chart is expected to have $ARL = \frac{1}{\alpha}$ when the process is in control.

Since the asymptotic covariance matrices Σ_{aX+b} for control charts constructed by symmetric quantiles and by empirical quantiles are different that leads to different non-centrality parameters resulting varied performances in $ARL's$. We then compute the $ARL's$ for the symmetric quantile control chart and the empirical quantile control chart for comparison.

Let us fix significance level at $\alpha = 0.005$ and $n = 1000$, for a non-parametric study, we consider the Laplace distribution $Lap(2)$ as the in control distribution with percentage vectors given as

$k = 4 : (\alpha_1, \alpha_2, ..., \alpha_4) = (0.02, 0.05, 0.95, 0.98),$

$k = 10 : (\alpha_1, \alpha_2, ..., \alpha_{10}) = (0.02, 0.05, 0.13, 0.25, 0.37, 0.63, 0.75, 0.87, 0.95, 0.98)$

For easiness of expression, we denote the $ARL's$ for empirical quantile charts and symmetric quantile charts, respectively, by ARL_e and ARL_s. The computed $ARL's$ are displayed in Table 3.2 and Table 3.3.

Table 3.2. Comparison of symmetric and empirical quantile charts by ARL (k=4)

(a,b)	ARL_s	ARL_e	(a,b)	ARL_s	ARL_e
$(1,0)$	200	200	$(1.2,0.5)$	17.00	21.73
$(1,0.2)$	172.14	196.81	$(1.2,1)$	8.32	18.54
$(1,0.5)$	90.89	181.30	$(1.2,2)$	2.20	11.91
$(1,1)$	22.74	138.83	$(1.5,1)$	3.02	3.91
$(1,2)$	2.68	61.77	$(2,1)$	1.44	1.54
$(1,5)$	1.00	5.32	$(2,2)$	1.25	1.47

Table 3.3. Comparison of symmetric and empirical quantile charts by ARL (k=10)

(a,b)	ARL_s	ARL_e	(a,b)	ARL_s	ARL_e
$(1,0)$	200	200	$(1.2,0.5)$	2.27	4.15
$(1,0.2)$	102.4	140.7	$(1.2,1)$	1.08	1.64
$(1,0.5)$	13.05	37.01	$(1.2,2)$	1	1
$(1,1)$	1.42	3.97	$(1.5,0.5)$	1.09	1.29
$(1,2)$	1.00	1.01	$(2,1)$	1	1
$(1,5)$	1	1	$(2,2)$	1	1

The case $(a,b) = (1,0)$ represents the process being in-control and both $ARL's$ are the expected number 200 for setting $\alpha = 0.05$. Surprisingly $ARL'_s s$ are all smaller than the corresponding $ARL'_e s$ unless they are number $1's$. This

indicates that the symmetric quantiles based control chart can detect the distributional shift with smaller number of samples.

We see that in this setting of coverage interval the symmetric quantiles based control chart is still more efficient than the empirical quantiles based control chart in detection of distributional shift.

3.6 Concluding Remarks

In contrast with empirical quantile based control chart of Grimshaw and Alt (1997), a symmetric quantile based control chart is proposed in this dissertation to monitor the popoulation-quantile vector aiming for monitoring more detailed features of a population distribution. The asymptotic theorem is also derived.

The symmetric quantile based control chart totally dominates the empirical one across all α_i of a population quantile vector $Q(\alpha_1, ..., \alpha_k)$, when the underlying distribution is Laplace distribution, which is widely used in modelling spectral vector of speech signals in speech recognition.

Chapter 4

Monitoring Nonlinear Profiles with Random Effects by Nonparametric Regression

In this chapter, we study nonlinear profile monitoring schemes. Principal component analysis is conducted, and a T^2 chart and a combined chart based on principal component scores are studied as well as individual Principal Component charts.

4.1 Proposed Monitoring Schemes

4.1.1 A Motivated Example

This study was motivated by the aspartame example given in Kang and Albin (2000). Since no data are available, a profile of the form $Y = I + Me^{N(x-1)^2} + \epsilon$ is used to mimic an aspartame profile. Then the idea is to perturb the parameters I, M, N randomly to create allowable profile-to-profile variations for an in-control process.

Thus the following random-effect model *was* considered to generate aspartame profiles:

$$Y_j = I + Me^{N(x_j-1)^2} + \epsilon_j, \quad j = 1, \cdots, p, \tag{4.1}$$

where $I \sim N(\mu_I, \sigma_I^2)$, $M \sim N(\mu_M, \sigma_M^2)$, $N \sim N(\mu_N, \sigma_N^2)$, $\epsilon \sim N(0, \sigma_\epsilon^2)$, and all the random components are independent of each other. Unfortunately, the response profile $\boldsymbol{Y} = (Y_1, \cdots, Y_p)'$ of model (4.1) has a complicated distribution with mean $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_p)'$ and covariance matrix $\boldsymbol{\Sigma}$ as follows. For $i, j = 1, \cdots, p$,

$$\mu_j = E(Y_j) = \mu_I + \mu_M e^{\mu_N(x_j-1)^2 + \frac{\sigma_N^2(x_j-1)^4}{2}},$$

$$Cov(Y_i, Y_j) = \sigma_I^2 + (\mu_M^2 + \sigma_M^2)\left[e^{\mu_N[(x_i-1)^2+(x_j-1)^2]+\frac{\sigma_N^2[(x_i-1)^2+(x_j-1)^2]^2}{2}}\right]$$
$$-\mu_M^2 e^{\mu_N(x_i-1)^2+\frac{\sigma_N^2(x_i-1)^4}{2}+\mu_N(x_j-1)^2+\frac{\sigma_N^2(x_j-1)^4}{2}} + \sigma_\varepsilon^2 \delta_{ij}, \tag{4.2}$$

where $\delta_{ij} = 1$ if $i = j$; and 0 otherwise. Note that, by (4.2), the covariance matrix $\boldsymbol{\Sigma}$ will be changed if the mean of M or N shifts, a situation too complicated

to analyze the performance of the control charts under study.

So, instead, we model the aspartame profiles as realizations of a Gaussian stochastic process with the mean function

$$\mu(x) = \mu_I + \mu_M e^{\mu_N (x-1)^2} \tag{4.3}$$

and a covariance function $G(s,t)$, where s, t are in the domain of x. To retain a similar profile-to-profile variation as it would be in the random-effect model (4.1), we let the in-control profiles follow $MVN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}_0 = (\mu_{01}, \cdots, \mu_{0p})'$ with

$$\mu_{0j} = \mu_I + \mu_M e^{\mu_N (x_j - 1)^2}, j = 1, \cdots, p, \tag{4.4}$$

and $\boldsymbol{\Sigma}$ is the covariance matrix given by equation (4.2).

When the mean function (4.3) is shifted, say, μ_I to $\mu_I + \alpha \sigma_I$, μ_M to $\mu_M + \beta \sigma_M$, and μ_N to $\mu_N + \gamma \sigma_N$, μ_{0j} is shifted from $\mu_I + \mu_M e^{\mu_N (x_j - 1)^2}$ to

$$\widetilde{\mu}_j \equiv (\mu_I + \alpha \sigma_I) + (\mu_M + \beta \sigma_M) e^{(\mu_N + \gamma \sigma_N)(x_j - 1)^2}, \quad j = 1, \cdots, p.$$

Let $\widetilde{\boldsymbol{\mu}} = (\widetilde{\mu}_1, \cdots, \widetilde{\mu}_p)'$. Then the shift on the mean of $\boldsymbol{Y}$ is $\boldsymbol{\delta} \equiv \widetilde{\boldsymbol{\mu}} - \boldsymbol{\mu}_0$.

4.1.2 Data Smoothing

In order to extend nonlinear profiles of a fixed parametric form to smooth profiles of a flexible nonparametric form, a smoothing technique is needed for denoising sample profiles. The idea of smoothing is to fit a smooth function whose final form is determined by the data and the chosen level of smoothness for the

curve. One popular approach is to fit noisy data by splines. Frequently, cubic splines (i.e., piecewise cubic polynomials with continuous second derivatives) are used for such approximations. Two commonly used spline smoothing techniques are smoothing splines and B-spline regression, both are available in popular statistical packages like R, Splus, and others. Other smoothing techniques such as local polynomial smoothing and wavelets can be used as well. We remark based on our experiences that, by filtering out noises, the actual signals can be better extracted from the data and PCA can explore the variation among profiles a lot better. In particular, smoothing tends to be more advantageous as the noise level (σ_ϵ^2) gets larger.

4.1.3 Phase I Monitoring

Assume that a set of n historical profiles is available for Phase I analysis. We first apply a smoothing technique to each of the n profiles to filter out the noise, and then apply PCA to the smoothed profiles as follows. Denote the $(p \times 1)$ data vector of the i-th profile by $\boldsymbol{y}_i$ and the usual sample covariance matrix of $\{\boldsymbol{y}_i, i = 1, \cdots, n\}$ by $\boldsymbol{S}$. Apply the eigenanalysis to $\boldsymbol{S}$. The eigenvector $\boldsymbol{v}_r$ corresponding to the r-th largest eigenvalue λ_r is the r-th principal component and $S_{ir} \equiv \boldsymbol{v}_r' \boldsymbol{y}_i$ is called the score of the r-th principal component of the i-th profile, $r = 1, \cdots, p, i = 1, \cdots, n$.

We select the number of the "effective" principal components by considering the total variation explained by the chosen principal components along with

the principle of parsimoniousness that we often use in the variable selection problem. Denote this number by K and the $(K \times 1)$ score vector $(S_{i1}, \cdots, S_{iK})'$ by $\boldsymbol{s}_i$.

For Phase I monitoring, due to the dependency of the K PC-scores, we adopt the usual Hotelling T^2 statistic described below. For the i-th profile, $i = 1, \ldots, n$, the T^2 statistic is defined as

$$T_i^2 = (\boldsymbol{s}_i - \bar{\boldsymbol{s}})' \boldsymbol{B}^{-1} (\boldsymbol{s}_i - \bar{\boldsymbol{s}}), \qquad (4.5)$$

where $\bar{\boldsymbol{s}} = \sum_{i=1}^{n} \boldsymbol{s}_i / n$ and $\boldsymbol{B} = \sum_{i=1}^{n} (\boldsymbol{s}_i - \bar{\boldsymbol{s}})(\boldsymbol{s}_i - \bar{\boldsymbol{s}})'/(n-1)$, the usual sample mean and sample covariance matrix of the score vectors.

Since score vectors are distributed as multivariate normal asymptotically (Anderson, 2003), according to Tracy et al. (1992), also Sullivan and Woodall (1996), we have

$$\frac{n}{(n-1)^2} T_i^2 \sim Beta\left(\frac{K}{2}, \frac{n-K-1}{2}\right) \text{ approximately.}$$

Thus, an approximate α-level upper control limit can be set at the $100(1-\alpha)$ percentile of the beta distribution with $K/2$ and $(n-K-1)/2$ as parameters.

For Phase I analysis, perform control-charting with the T^2 statistic of the score vectors in (4.5) to detect the out-of-control profiles in the historical data set. If there are any, remove them and redo PCA and control-charting with the remaining profiles. Repeat this procedure until all the remaining profiles are within the control limit. These remaining profiles are considered as "in-control" profiles and can be used to characterize the in-control process. The resulting

principal components and eigenvalues can then be used to set up the control limit for Phase II on-line monitoring.

4.1.4 Phase II Monitoring

As in most of Phase II studies, we assume the in-control process distribution of the profiles after de-noising has been characterized as $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, either from prior experiences or estimated from the Phase I analysis.

Our Phase II monitoring schemes are also based on PCA. Apply PCA to $\boldsymbol{\Sigma}_0$ to obtain eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$, and the corresponding eigenvectors, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p$. Similar to that in Phase I analysis, choose the number of effective principal components K based on the parsimoniousness and the total variation that the first K PCs account for. More specifically, since the r-th PC accounts for $\lambda_r / \sum_{r=1}^{p} \lambda_r$ of the total variation, we can simply choose the first K such that $\sum_{r=1}^{K} \lambda_r / \sum_{r=1}^{p} \lambda_r$ reaches a desired level.

Now for each of the incoming profiles in Phase II monitoring, first smooth and then project it onto the first K PCs to obtain K PC-scores. Denote these scores by $S_1, \cdots, S_K$. Since these scores are independent and S_r follows a normal distribution with mean $\boldsymbol{v}_r' \boldsymbol{\mu}_0$ and variance λ_r when the process is in control, it is easy to construct a control chart for each of the K PC-scores accordingly. Denote the desired in-control false-alarm rate by α. Then the control limits for the r-th PC-score chart, which monitors the statistic S_r, is

$$\boldsymbol{v}_r' \boldsymbol{\mu}_0 \pm Z_{\alpha/2} \sqrt{\lambda_r}, r = 1, \cdots, K.$$

If a particular mode of process change can be caught by one of the first K principal components, then we can use that particular PC-score chart to monitor it. However, very often a process shift is reflected in more than one principal component. When this happens, we can consider a combined chart scheme by combining all K PC-score charts. A combined chart scheme signals out-of-control when any of the K individual charts signals. Thus, the proposed combined chart is equivalent to monitoring the statistic

$$\max_{1 \leq r \leq K} |\frac{S_r - \boldsymbol{v}'_r \boldsymbol{\mu}_0}{\sqrt{\lambda_r}}|.$$

This chart signals out-of-control when $\max_{1 \leq r \leq K} |(S_r - \boldsymbol{v}'_r \boldsymbol{\mu}_0)/\sqrt{\lambda_r}| > Z_{\alpha'/2}$, where the individual false-alarm rate α' should be chosen at the level of $1 - (1-\alpha)^{1/K}$ so that the overall false-alarm rate is at the desired level α.

We can also consider a T^2 chart by monitoring the statistic

$$T^2 = \sum_{r=1}^{K} \frac{(S_r - \boldsymbol{v}'_r \boldsymbol{\mu}_0)^2}{\lambda_r}, \tag{4.6}$$

which follows the chi-square distribution with K degrees of freedom (denoted by χ^2_K) when the process is in control. Thus, the upper control limit is the $100(1-\alpha)$ percentile of χ^2_K.

4.1.5 ARL of the Proposed Schemes

We evaluate the performances of the proposed Phase II monitoring schemes described above in terms of ARL, the average run length. The ARL values of the individual PC-score chart can be computed as follows. Assume that the

mean of the profile has been shifted from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_0 + \boldsymbol{\delta}$. The probability of detecting the shift by the r-th PC-score chart is

$$
\begin{aligned}
p &= 1 - P(|\frac{S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0}{\sqrt{\lambda_r}}| \le Z_{\alpha/2}) = 1 - P(-\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} - Z_{\alpha/2} \le Z \le -\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} + Z_{\alpha/2}) \\
&= 1 - \Phi(-\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} + Z_{\alpha/2}) + \Phi(-\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} - Z_{\alpha/2}),
\end{aligned}
$$

where Φ is the cumulative distribution function of the standard normal variate Z and $Z_{\alpha/2}$ is the $100(1 - \alpha/2)$ percentile of Z. Then the value $1/p$ is the ARL of the r-th PC-score chart.

Since the PC-scores $S_1, \cdots, S_K$ are independent, the ARL of the combined chart also can be computed easily by the reciprocal of

$$
\begin{aligned}
p &= 1 - P(\max_{1 \le r \le K} |\frac{S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0}{\sqrt{\lambda_r}}| \le Z_{\alpha'/2}) = 1 - \prod_{r=1}^{K} P(|\frac{S_r - \boldsymbol{v}_r'\boldsymbol{\mu}_0}{\sqrt{\lambda_r}}| \le Z_{\alpha'/2}) \\
&= 1 - \prod_{r=1}^{K} [\Phi(-\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} + Z_{\alpha'/2}) - \Phi(-\frac{\boldsymbol{v}_r'\boldsymbol{\delta}}{\sqrt{\lambda_r}} - Z_{\alpha'/2})],
\end{aligned}
$$

where $\alpha' = 1 - (1 - \alpha)^{1/K}$ is the individual false-alarm rate.

Since T^2 statistic in (4.6) follows a noncentral chi-square distribution with K degrees of freedom and non-centrality $\xi = \sum_{r=1}^{K}(\boldsymbol{v}_r'\boldsymbol{\delta})^2/\lambda_r$ (denoted by $\chi_K^2(\xi)$). Then the detecting power of the T^2 chart can be easily calculated by

$$
p = P(T^2 > \chi_{K,\alpha}^2) = P(\chi_K^2(\xi) > \chi_{K,\alpha}^2),
$$

where $\chi_{K,\alpha}^2$ denotes the $100(1 - \alpha)$ percentile of the central chi-square distribution χ_K^2.

4.2 Simulation and Comparative Studies

4.2.1 Settings for Simulation

In our simulation study, we generate profiles from $MVN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}_0$ is given in (4.4) and $\boldsymbol{\Sigma}$ is given in (4.2) with $\mu_I = 1, \sigma_I = 0.2, \mu_M = 15, \sigma_M = 1$, $\mu_N = -1.5, \sigma_N = 0.3$, and $x = 0.64, 0.8, \ldots, 3.52$. Both of x and y values are scaled variables, not the actual temperature levels and the amount of aspartame dissolved in the dissolving process. Denote the in-control ARL by ARL_0. All charts are designed to have the same $ARL_0 = 370.3704$, which corresponds to the false-alarm rate of $\alpha = 0.0027$.

4.2.2 A Study on the Number of Principal Components

To study how the choice of the number of effective principal components affects the detecting power of the monitoring scheme, we conduct a simulation study. In this study, the detecting power is measured by the ability of the monitoring scheme in detecting the real out-of-control profiles. For example, in a data set of fifty simulated profiles with three out-of-control profiles, if the scheme catches two of the three, then the detecting power measured is 2/3. The false-alarm rate can be measured in a similar way. Let the number of the principal components used be k. Then for each data set, compute the detecting power and the percentage of the total variation explained by k principal components for various values of k.

We choose some settings of shifts. For each setting, we generate fifty profiles within which some profiles are generated from the shifted population. Then repeat each setting 20,000 times to get the averaged detecting power and the averaged percentage of the total variation explained.

As one would expect, the result (not shown) of the study indicates that having more principal components does explain more variation, but not necessarily has more detecting power. In fact, the power of the T^2 statistic starts to drop when k gets to a certain level, which usually is a fairly small number. So it is necessary to choose an appropriate number of principal components.

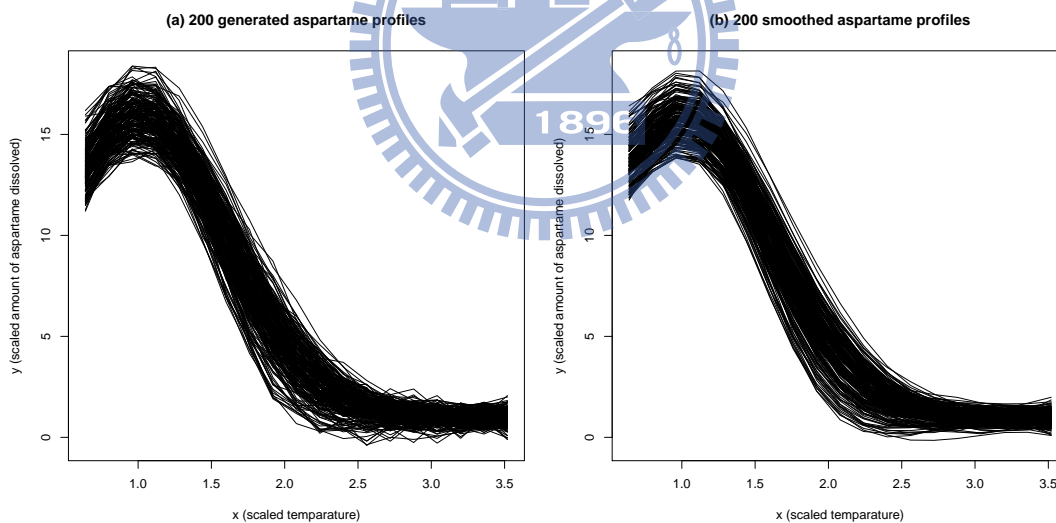### 4.2.3 A Simulated Aspartame Example–Phase I Monitoring

Figure 4.1: (a) 200 generated and (b) smoothed in-control aspartame profiles.

We now demonstrate Phase I analysis with the aspartame example described before. In Phase I, we generate 200 in-control historical profiles, each with 19

set points. First de-noise these profiles by smoothing splines using statistical package R. Figures 4.1(a) and 4.1(b) display the 200 simulated profiles and their smoothed profiles respectively. Then we apply PCA to the sample covariance matrix of these smoothed profiles and get the 19 eigenvalues, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{19} \geq 0$ and their associated eigenvectors. The first four principal components account for 75.16%, 19.41%, 2.60%, and 0.92% of the total variation in the profiles, respectively. For profile monitoring, we decide to choose $K = 3$ for parsimoniousness since with three PCs, it has already accounted for 97.17% of the total variation. In practice, it is also fine to choose $K = 4$. Now project each (smoothed) profile onto the first three eigenvectors to get the scores and then compute T^2 by (4.5). The resulting T^2 control chart (not shown) indicates that the process is in control.

4.2.4 An ARL Comparison Study–Phase II Monitoring

To compare the performances of the proposed schemes for Phase II monitoring described in Subsection 4.1.4, we compute the ARL values of each scheme as derived in Subsection 4.1.5.

Let the (i, j)-th entry of the in-control covariance matrix $\mathbf{\Sigma}_0$ be (4.2) without the $\sigma_\epsilon^2 \delta_{ij}$ term. Apply PCA to this "population" covariance matrix. It is found that the first four principal components respectively account for 74.82%, 22.58%, 2.30%, and 0.29%, which totals 99.99%, of the total variation; and other components practically explain nothing. This is mainly because we have

only three degrees of freedom in varying profiles (without measurement error), namely, the values of I, M, and N in model (4.1).

Figure 4.2(a) depicts the first three eigenvectors of $\mathbf{\Sigma}_0$. To see the effect of a particular principal component, say $\boldsymbol{v}_r$, Ramsay and Silverman (2005) presented a visualizing tool that plots $\boldsymbol{\mu}_0 \pm L\boldsymbol{v}_r$, where L is a suitable multiple. Figures 4.2(b)-4.2(d) illustrate respectively the corresponding features captured by the first three principal components with $L = 3$. From the figures, we can see that the PC1 captures the variation of the vertical shifting of the profiles except for the right tail; PC2 reflects the variation in the height of the peak and the declining rate of the curve; PC3 captures mainly the vertical shifting in the tail and slightly the variation at the peak.

For ARL comparison, consider the I-shift from μ_I to $\mu_I + \delta\sigma_I$, M-shift from μ_M to $\mu_M + \delta\sigma_M$, and N-shift from μ_N to $\mu_N + \delta\sigma_N$, for $\delta = 0, 0.25, \ldots, 3$. Figures 4.3(a)-4.3(c) display the ARL curves for the shifts in I, M, N, respectively. We observe the followings:

- Both PC1 and PC3 have some power in catching the shift in I because both represent the mode of variation in vertical shifting (but in different areas) of the profile as shown in Figures 4.2(a) and 4.2(c). We are a little bit surprised to see that PC1 is less powerful than PC3. This may be explained by: (i) PC3 explains almost all the variation in vertical shifting in the tail area; (ii) although PC1 can explain the vertical shifting for $x < 2.5$, the other two PCs also pick up some, especially around the peak
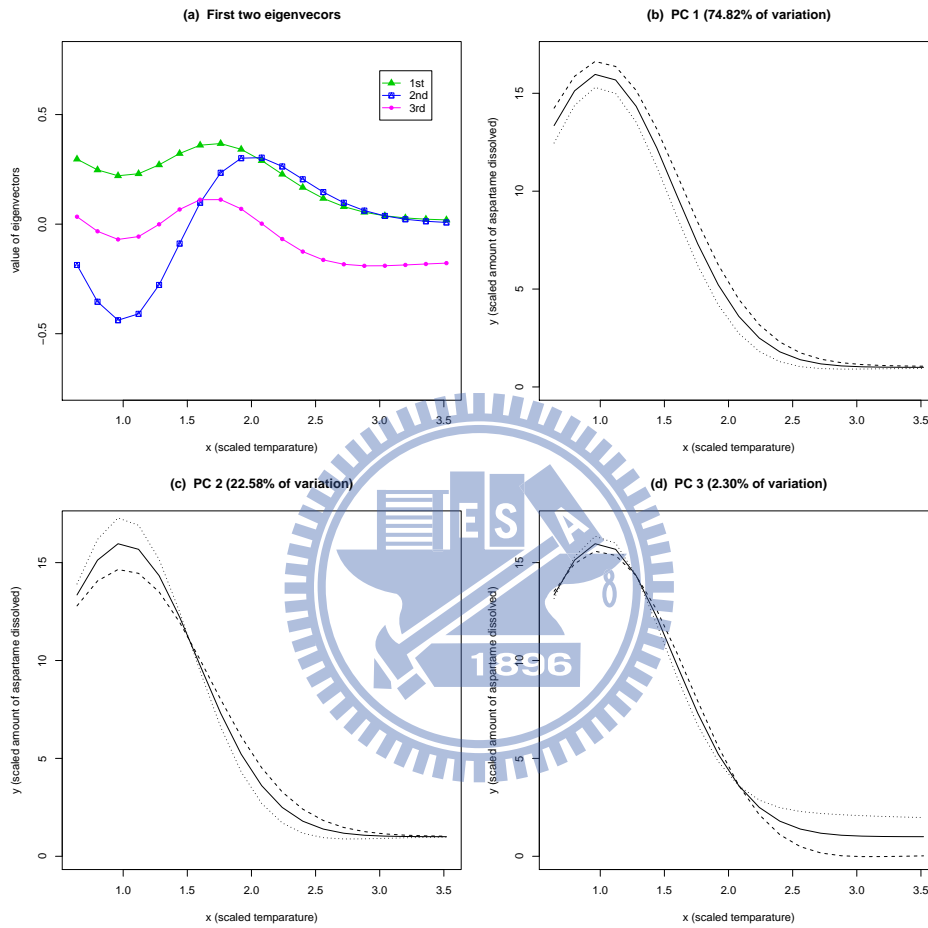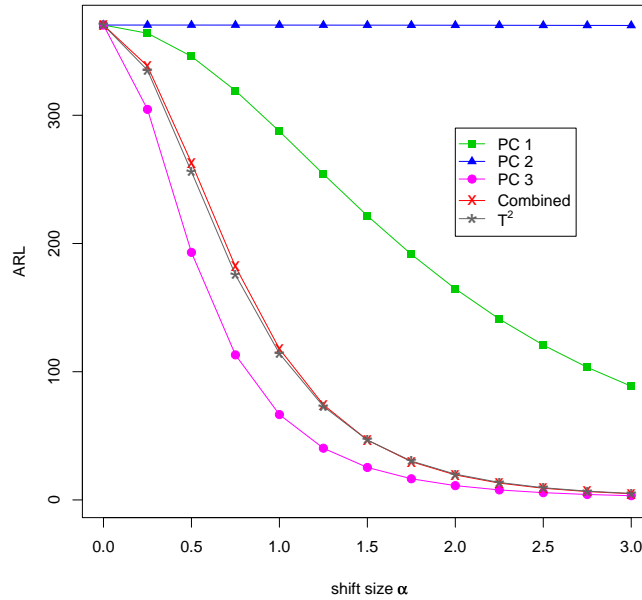
Figure 4.2: Aspartame example. (a) Plot of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$; (b)-(d) $\boldsymbol{\mu}_0 \pm 3\boldsymbol{v}_r, r = 1, 2, 3$.

(a) ARL comparison for I–shift

area. PC2 hardly has any power in detecting I-shift. The power of the T^2 and combined chart are between that of PC1 and PC3, and as the shift gets larger, the difference in ARL between PC3 and the T^2/combined chart gets smaller. Also, the T^2 and combined chart are very close with T^2 slightly better for the shift size $\alpha \leq 1.5$ and the combined chart better for $\alpha \geq 1.75$.

• As to the M-shifts, except for PC3 that does not have much power, the other four charts are not too far from each other. The order of the performance is about PC2 > T^2 > combined > PC1, with the exception that T^2 finally beats PC2 for larger shifts. Here ">" means "performs better than".

• For detecting N-shift, PC3 has a strange ARL curve (see Figure 4.3(c)),

(b) ARL comparison for M–shift
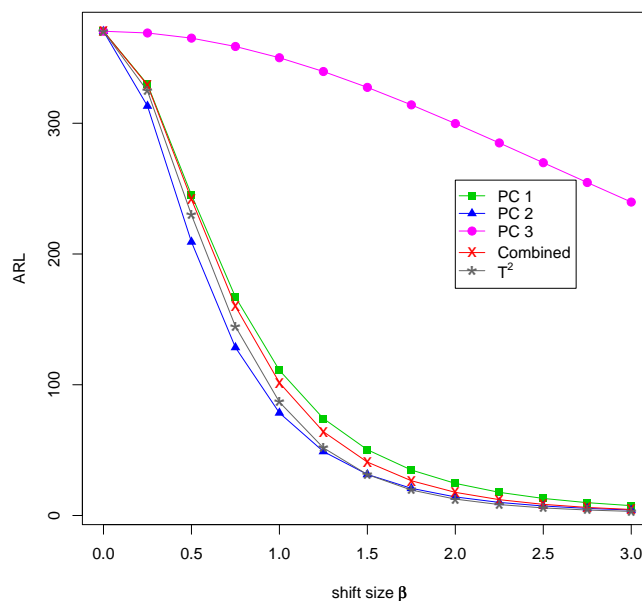
which may be caused by the fact that the shift in the mean vector when projected onto PC3, $\boldsymbol{v}_3'\boldsymbol{\delta}$, is not monotone in the shift multiple δ. PC1 performs the best for small shifts but gets worse as the shift gets larger and eventually becomes the worst one for large shifts. T^2 is the second best for small shifts and then quickly becomes the best. PC2 and the combined chart are fairly close for small shifts; PC2 wins when the shift size is small and soon loses it to the combined chart for moderate to large shifts.

We learn from this study that a mode of variation often cannot be captured by a single PC-score chart. Even for the variation as simple as the vertical shifting like the I-shift, it takes more than one principal component to represent this effect. Also the T^2 and combined chart are fairly close to each other and comparable with the best PC-score chart.
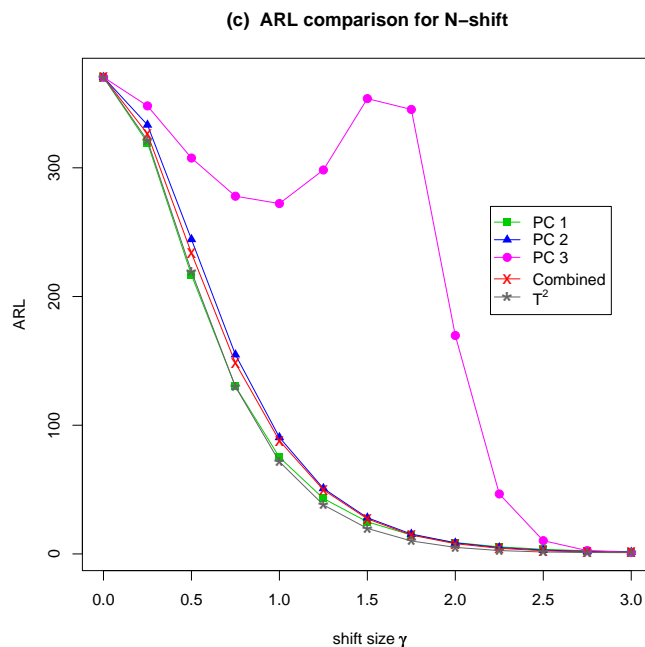
(c) ARL comparison for N–shift

Figure 4.3: ARL comparisons of aspartame example. (a) I, (b) M, (c) N shifts.

4.3   A Case Study–VDP Example

The VDP data set described in Section 1 contains $n = 24$ profiles, each was measured at $p = 314$ set points. Figure 4.4(a) is the plot of the VDP data. First de-noise these profiles by smoothing splines using statistical package R. See Figure 4.4(b) for the plot of the smoothed profiles. Apply PCA to the sample covariance matrix of the smoothed profiles. And the first four principal components account for 85.26%, 10.83%, 1.90%, 0.84% of the total variation in the profiles, respectively. We select $K = 3$ principal components for Phase I process monitoring because the total variation explained by the first three PCs is already as high as 97.99%. Figure 4.5(a) is the plot of the first three eigenvectors. Figures 4.5(b)-4.5(d) show the modes of variation they capture by
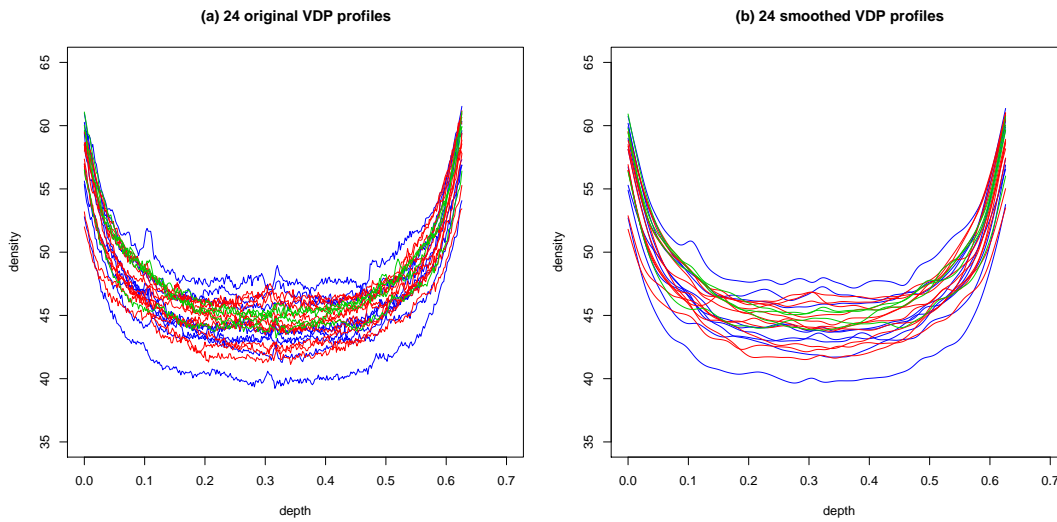
55

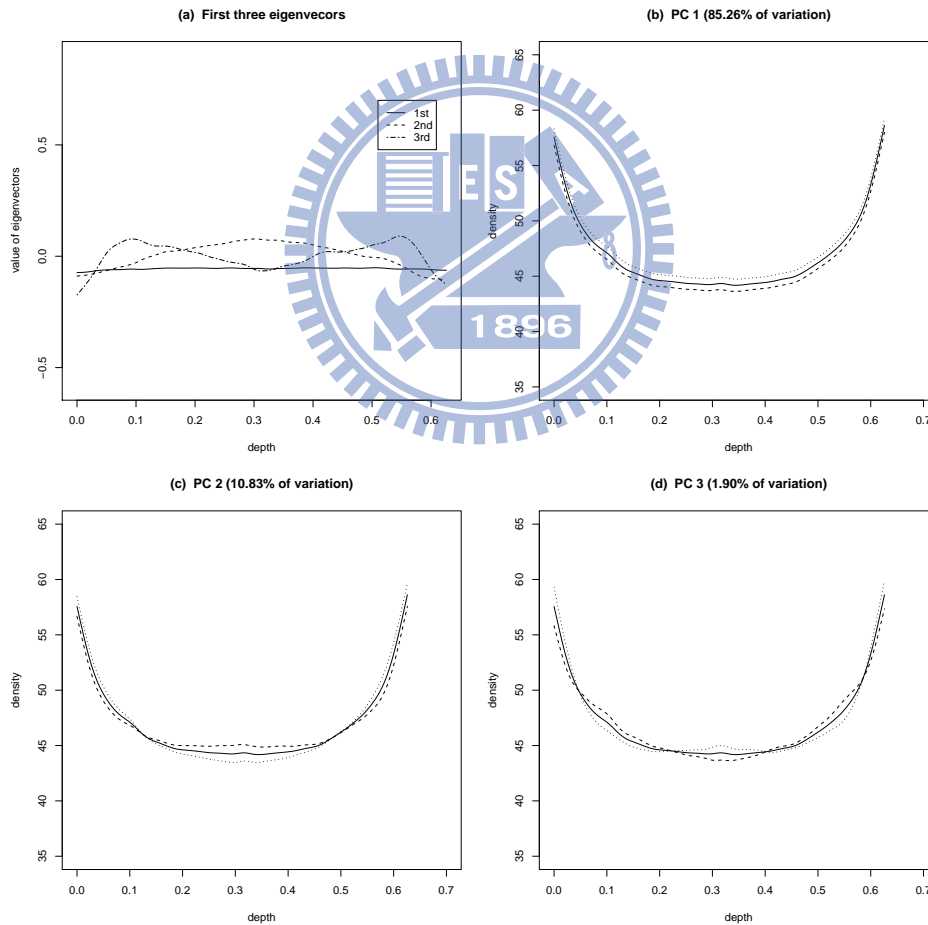Figure 4.4: (a) 24 VDP profiles (b) 24 smoothed VDP profiles.

Figure 4.5: VDP example. (a) Plot of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$; (b)-(d) $\boldsymbol{\mu} \pm 10\boldsymbol{v}_r, r = 1, 2, 3$.

plotting the mean vector $\boldsymbol{\mu}$ and $\boldsymbol{\mu} \pm 10 \, \boldsymbol{v}_r, r = 1, 2, 3$. The figures show that: (i) PC1 represents the variation in the ground level of the VDP profiles, especially the bottom part of the "bathtub"; (ii) PC2 can catch some variation in the roundness of the bottom part of the bathtub; and (iii) PC3 may be able to describe the variation in the roundness on the two ends and the central part of the bathtub bottom. All these interpretations can also be seen from the three eigenvectors shown in Figure 4.5(a). As to Phase I monitoring, the T^2 control chart (not shown) indicates no out-of-control profiles in the VDP data.

In Phase II monitoring, we treat the average profile vector $\boldsymbol{\mu}$ of the 24 smoothed VDP profiles and the sample covariance matrix Σ as the in-control process parameters to perform our simulation. Here, $\boldsymbol{\mu} = (\mu(t_1), \cdots, \mu(t_{314}))'$ is a 314×1 vector and the covariance matrix Σ is a 314×314 matrix. We can generate the in-control profiles by $\mathbf{Y} \sim MVN(\boldsymbol{\mu}, \Sigma)$. For out-of-control conditions, we shift the profiles in their first two principal components. That is, generate new profile data by

$$\tilde{\mathbf{Y}} \sim k\sqrt{\lambda_i} \times \boldsymbol{v}_i + MVN(\boldsymbol{\mu}, \Sigma) \, ,$$

where k=0, 0.25, ..., 2 and $\boldsymbol{v}_i$ is the i-th eigenvector of Σ, i=1, 2. We simulated 200,000 profiles to compute an ARL estimate for each out-of-control conditions considered. Then we repeat the procedure 1000 times to get our final ARL estimate along with its standard error.

ARL results for shifts in principal component 1 and 2 are listed separately

in Table(4.1) and Table(4.2). As we can see that shifts in principal component 1 are solely captured by the first principal component score. Likewise, shifts in principal component 2 are captured by the second principal component score. The other three principal component scores make no contributions to the power of detection.

4.4 Concluding Remarks

In this study, we propose and discuss monitoring schemes for nonlinear profiles. We use the principal components analysis to analyze the covariance matrix of the profiles and then utilizing the principal component scores that capture the main features of the profile data for process monitoring.

In addition to the individual PC score charts, we study two charts that utilize the overall information contained in the K effective principal components, namely, the combined chart and T^2 chart. The T^2 chart performs somewhat better than the combined chart in terms of the average run length, but not too far off. However, by providing charts for all of the effective components, the combined chart gives more clues for finding assignable causes than the T^2 chart.

When the shift corresponds to a mode of variation that a particular principal component represents, then it would be ideal to use the individual PC-score chart for process monitoring because this particular individual chart will have

the best power among the charts under study. Unfortunately, this ideal situation is rare in practice. Moreover, by just monitoring one individual PC chart, one is running a risk of not being able to detect other types of process changes. One may argue that we should monitor the process with all K individual charts simultaneously in order to catch all potential changes. But with the same α for each of the individual charts, the overall false-alarm rate is greatly increased to $1 - (1 - \alpha)^K$, which is about K times of the original false-alarm rate. Thus, for being more practical and conservative, we recommend using the T^2 chart or the combined chart scheme, because they still have comparably good power to monitor these particular types of shifts and have a lot better power than the individual chart for general types of shifts.

It is noted that the degree of smoothness used in the data smoothing step has a great impact on the result of the subsequent PCA step. The high total explanation power of the first few principal components demonstrated in the paper is in fact caused by the high degree of smoothness in the smoothed curves. This argument is supported by our finding that if B-spline regression is used for smoothing, the number of B-spline bases used is exactly the number of the principal components with nonzero eigenvalues. So if the underlying profiles (i.e., with no noises) are fairly smooth as what we have in the hypothetical aspartame example (in which a data profile is a 3-parameter exponential function plus noises), then the data dimension can be well reduced by applying PCA to the smoothed curves. The situation in the VDP data is similar. On

the other hand, if the underlying profiles are not very smooth and data profiles are not too much over-smoothed, then it might take quite a few principal components to explain good enough proportion of variation. We remark that, regardless of which K, the number of effective principal components, is chosen, the false-alarm rate for each of our schemes stays at α. However, the diagnosis of out-of-control conditions would likely become more complicated when K is large.

The degree of smoothness may affect the effectiveness of the Phase II process monitoring as well. If the noise levels are about the same for both Phase I and Phase II profile data, we suggest applying the same degree of smoothness to them. In this way, the results of Phase II monitoring are somehow not that sensitive to the extent of smoothing. However, when profiles are over-smoothed to the extent that some local features are lost, then the process changes associated with these vanished local features cannot be detected. On the other hand, when profiles are under-smoothed to an extent, some spurious features may appear in the fitted curves. Unfortunately, these spurious features may not appear at the same place and may not have the same form across profiles. Then the estimated in-control model obtained from Phase I data may not suitable for effective Phase II monitoring. We remark that even for the case that the in-control process is known or appropriately characterized, the spurious features in the "smoothed" Phase II profiles caused by under-smoothing will cause more false alarms to signal.

In practice, when one employs the T^2 chart or the combined chart scheme and detects significant shifts, it is desirable to find the sources responsible for the shifts. For this, we suggest to rank the standardized PC-scores and investigate the corresponding principal components in order, starting from the largest PC score. With the help of the plots like Figures 4.2(a)-4.2(d) and 4.5(a)-4.5(d) and engineers' expertise, the characteristics of the principal components sometimes can reveal potential root causes of the shifts.

The monitoring of process or product profiles has become a popular and promising area of research in statistical process control in recent years. At the same time, functional data analysis (FDA) is also gaining lots of attentions and applications. We believe many techniques developed for FDA may be extended to developing new profile monitoring techniques in SPC.

			k		
chart	0	0.25	0.5	0.75	1
PC1	201.141	157.7244	91.6269	50.1074	28.1935
	0.6537	0.4263	0.1869	0.0773	0.0339
PC2	202.5093	201.1011	202.7793	201.3586	202.3857
	0.6897	0.6687	0.6295	0.6704	0.6432
PC3	202.2185	202.1651	201.4955	203.475	202.0911
	0.6937	0.6648	0.6495	0.6207	0.6461
PC4	202.4143	200.7331	201.7661	202.0241	202.1318
	0.6594	0.6526	0.6303	0.6551	0.6248
chart	1.25	1.5	1.75	2	
PC1	16.7264	10.4605	6.8874	4.7689	
	0.0153	0.0071	0.0037	0.0020	
PC2	202.5093	201.1011	202.7793	201.3586	
	0.6897	0.6687	0.6295	0.6704	
PC3	202.2185	202.1651	201.4955	203.475	
	0.6937	0.6648	0.6495	0.6207	
PC4	202.4143	200.7331	201.7661	202.0241	
	0.6594	0.6526	0.6303	0.6551	

Table 4.1: ARL comparison for shifts in principal component 1.

			k		
chart	0	0.25	0.5	0.75	1
PC1	203.1434	202.2938	202.2304	201.9739	202.0307
	0.6599	0.6496	0.6643	0.6826	0.6423
PC2	202.3857	157.2349	91.368	50.138	28.2533
	0.6432	0.4475	0.1938	0.0783	0.0318
PC3	202.0911	201.3132	202.4239	202.4192	201.6323
	0.6461	0.6695	0.6455	0.6580	0.6622
PC4	202.1318	201.4323	201.5589	203.1134	201.6549
	0.6248	0.6797	0.6407	0.6532	0.6349
chart	1.25	1.5	1.75	2	
PC1	200.6637	202.6417	202.2779	201.5532	
	0.6327	0.6492	0.6344	0.6624	
PC2	16.75	10.455	6.8888	4.771	
	0.0141	0.0073	0.0038	0.0020	
PC3	201.7072	202.043	202.6156	202.3369	
	0.6355	0.6418	0.6803	0.6310	
PC4	203.0749	202.5458	202.4285	201.5443	
	0.6780	0.6515	0.6505	0.6427	

Table 4.2: ARL comparison for shifts in principal component 2.

Chapter 5

Conclusion

5.1 Symmetric Quantile Coverage Interval

Symmetric quantile coverage interval performs better than empirical quantile coverage interval in terms of the followings:

a. symmetric quantile coverage interval can cover more of the high density part of distribution function, when the underlying distribution is asymmetric.

b. symmetric quantile coverage interval is more robust against outliers than empirical quantile coverage interval.

c. Even when the underlying distribution is symmetric, symmetric quantile coverage interval is with smaller variance when the coverage percentage is with large $1 - \alpha$, which is the common practices of most applications of coverage interval.

5.2 Multivariate Control Chart by Symmetric Quantile

An application of coverage interval is a control chart, with the upper and lower control limits constructed by the two ends of a coverage interval. As for constructing a control chart with interests on the multiple characteristics of one distribution, multivariate control chart by symmetric quantiles is proposed and its asymptotic theorem is derived.

When the underlying distribution is with heavy tails, the proposed symmetric quantile control chart performs better than the empirical quantile control chart, proposed by Grimshaw and Alt(1997).

5.3 Monitoring Nonlinear Profiles by Nonparametric Regression

PCA score based T^2 monitoring schemes is proposed to deal with nonlinear profile data in phase-1 monitoring. For phase-II monitoring, three PCA score based monitoring schemes(Individual score charts, T^2 chart, the combined chart) are investigated and compared. T^2 chart is recommended in practical uses with the plotting of

$$\boldsymbol{\mu} \pm L\, \boldsymbol{v}_r$$

.

Data smoothing is recommended to be deone before PCA analysis, for it has great impacts on the effectiveness of the construction of phase-I control schemes and the effectiveness of phase-II monitoring.

5.4   Future Research

PAT(Part Average Testing) is the post-process defined in AEC-Q001, where Automotive Eletronics Council(AEC) was established for the purpose of establishing common part-qualification and quality system standards. PAT is primarily for detecting outlier parts which tend to be higher contributors to long term quality and reliability problems. Current static PAT limits= (Robust Mean) $\pm$ 6*(Robust Sigma) where the Robust Mean is the median, and the Robust Sigma qual to $(Q3 - Q1)/1.35$.

PAT has recently been adopted by a number of semiconductor companies, with the typical applications on wafer sort and final test data. Most semiconductor wafer sort and final test parameters are not normal and with heavy tails or asymmetric. Application of symmetric qunatile coverage interval as outlier detection tool can be investigated and compared with the one proposed by AEC standard.

Tolerance analysis is adopted by semiconductor companies to assist designing specs. The output performance transfer function and component tolerance intervals are the inputs into tolerance analysis and output performance toler-

ance interval is the desired result for important references in spec definition. However, when major output performance is a I-V curve under specific conditions, the solution to the tolerance region of the I-V curve is a topic which needs to be explored. Curve/Profile tolerance region will contribute an important I-V curve performance measure in semiconductor IC performance evaluation.

Bibliography

[1] ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd edition. Wiley, New York.

[2] CASTRO, P. E., LAWTON, W. H., and SYLVESTRE, E. A. (1986). Principal modes of variation for processes with continuous sample curves. *Technometrics*, **28**, 329-337.

[3] CHAO, M. T. and CHENG, S. W. (1996). Semicircle control chart for variables data. *Quality Engineering*, **8**, 441-446.

[4] CHEN, L.-A. and CHIANG, Y. C.(1996). Symmetric type quantile and trimmed means for location and linear regression model. *Journal of Nonparametric Statistics*, **7**, 171-185.

[5] CHEN, L.-A., HUANG, J.-Y. and CHEN, H.-C. (2007). Parametric coverage interval. *Metrologia*, **44**, L7-L9.

[6] CHENG, S. W. and THAGA, K. (2005). Multivariate Max-CUSUM chart. *Quality Technology and Quantitative Management International*, **2**, 191-206.

[7] CHIANG, Y. C., CHEN, L.-A. and YANG, H.-C. (2006). Symmetric Quantiles and their Applications. *Journal of Applied Statistics*, **33**, 807-817.

[8] DING, P. L., ZENG, L., and ZHOU, S. Y. (2006). Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Quality Engineering*, **5**, 619-625.

[9] FRIEDBERG, R. C., SOUERS, R., WAGAR, E. A., STANKOVIC, A. K. and VALENSTEIN, P. N. (2007). The origin of reference intervals. *Arch Pathol Lab Med*, **131**, 348-357.

[10] GRIMSHAW, S. D. and ALT, F. B. (1997). Control charts for quantile function values. *Journal of Quality Technology*, **29**, 1-7.

[11] HUANG, J.-Y., CHEN, L.-A. and WELSH, A. H. (2010). A note on reference limits. *IMS Collections, Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jureckova*, **7**, 84-94.

[12] JANACEK, G. J. and MEIKLE, S. E. (1997). Control charts based on medians. *The Statistician*, **46**, 19-31.

[13] JENSEN, W. A., BIRCH, J. B., and WOODALL W. H. (2006). Profile monitoring via linear mixed models. Technical Report No. 06-2, Department of Statistics, Virginia Polytechnic Institute and State University, Virginia.

[14] JENSEN, WILLIS A. and BIRCH, JEFFREY B. (2009). Profile monitoring via nonlinear mixed models. *Journal of Quality Technology*, **41**, 18-35.

[15] JONES, M. C. and RICE, J. A. (1992). Displaying the important features of large collections of similar curves. *The American Statistician*, **46**, 140-145.

[16] KANG, L. and ALBIN, S. L. (2000). On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, **32**, 418-426.

[17] KANJI, G. K. and ARIF, O. H. (2000). Median ranki control chart by quantile approach. *Journal of Applied Statistics*, **27**, 757-770.

[18] KIM, K., MAHMOUD, M. A., and WOODALL, W. H. (2003). On the monitoring of linear profiles. *Journal of Quality Technology*, **35**,317-328.

[19] KIM, S. J. (1992). The metrically trimmed means as a robust estimator of location. *Annals of Statistics*, **20**, 1534-1547.

[20] LIN, S.-H., CHAN, W. and CHEN, L.-A. (2008). A nonparametric coverage interval. *Metrologia*, **45**, 1-4.

[21] LIU, R. Y. and TANG, J. (1996). Control charts for dependent and independent measurements based on bootstrap methods. *Journal of the American Statistical Association*, **91**, 1694-1700.

[22] MAHMOUD, M. A. and WOODALL, W. H. (2004). Phase I analysis of linear profiles with calibration applications. *Technometrics*, **46**, 380-391.

[23] QIU, PEIHUA, ZOU, CHANGLIANG and WANG, ZHAUJUN (2010). Nonparametric Profile Monitoring by Mixed Effects Modeling. *Technometrics*, **52**, 265-293.

[24] RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Applied Functional Data Analysis*, 2nd edition. Springer, New York.

[25] REED, A. H., HENRY, R. J. and MASON, W. B. (1971). Influence of statistical method used on the resulting estimate of normal range. *Clinical Chemistry*, **17**, 275.

[26] REPCO, J. (1986). Process capability plot. *The Proceedings of the 330th EQQC Conference*, 373-381.

[27] RICE, J. A. and SILVERMAN, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society, Series B*, **53**, 233-243.

[28] RUPPERT, D. and CARROLL, R. J. (1980). Trimmed least squares estimation in the linear model. *Journal of the American Statistical Association*, **75**, 828-838.

[29] SEN, P. K. and SINGER, J. M. *Large Sample Methods In Statistics*. Capman & Hall: New York.

[30] SHIAU, J.-J. H. and LIN, H.-H. (1999). Analyzing accelerated degradation data by nonparametric regression. *IEEE Transactions on Reliability*, **48**, 149-158.

[31] SHIAU, J.-J. H., LIN, S.-H., and CHEN, Y.-C. (2006). Monitoring linear profiles based on a random-effect model. Technical Report. Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

[32] SHIAU, J.-J. H. and WENG, Z.-P. (2004). Profile monitoring by nonparametric regression. Technical Report. Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

[33] Shiau, J.-J. H., Yen, C.-L., and Feng, Y.-W. (2006). A new robust method for Phase I monitoring of nonlinear profiles. Technical Report. Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

[34] Shiling, E. G. and Nelson, P. R. (1976). The effect of non-normality on the control limits of $\bar{X}$ control chart. *Journal of Quality Technology*, **8**, 183-187.

[35] Solberg, H. E. Establishment and use of reference values. In Tietz, N. W., ed. *Textbook of Clinical Chemistry* (1986). Philadephia.

[36] Spiring, F. A. and Cheng, S. W. (1998). An alternative variables control chart: The univariate and multivariate case. *Statistica Sinica*, **8**, 273-287.

[37] Sullivan, J. H. and Woodall, W. H. (1996). A comparision of multivariate control charts for individual observations. *Journal of Quality Technology*, **28**, 398-408.

[38] Tracy, N. D., Young, J. C., and Mason, R. I. (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, **24**, 88-95.

[39] Van Nuland, Y. (1992). ISO 9002 and the circle technique. *Quality Engineering*, **5**, 269-291.

[40] Walker, E. and Wright, S. P. (2002). Comparing curves using additive models. *Journal of Quality Technology*, **34**, 118-129.

[41] Williams, J. D., Woodall, W. H., and Birch, J. B. (2003). Phase I analysis of nonlinear product and process quality profiles. Technical Report No. 03-5, Department of Statistics, Virginia Polytechnic Institute and State University, Virginia.

[42] Williams, J. D., Birch, J. B., Woodall, W. H., and Ferry, N. M. (2007). Statistical monitoring of heteroscedastic dose-response profiles from high-throughput screening. *Journal of Agricultural, Biological, and Environmental Statistics*, **2**, 216-235.

[43] Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004). Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, **36**, 309-320.