https://doi.org/10.26219/heal.aueb.6243

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

PROCESS CONTROL TECHNIQUES FOR PROFILE MONITORING

By

Sophia Chrysa I. Mitraka

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
July 2016

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ

Τεχνικές Ελέγχου Διαδικασιών για την Παρακολούθηση Προφίλ

Σοφία Χρύσα Ι. Μητράκα

ΔΙΑΤΡΙΒΗ
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Ιούλιος 2016

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

To my family

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Stelios Psarakis, for the confidence shown in me by assigning a particular topic of interest, as a spiritual culmination of my one-year term in the Master of "Statistics" in Athens University of Economic and Business.

Professor S. Psarakis supported my efforts, guided me, spent much of his time, giving me valuable knowledge and tips to get at this point.

Warm thanks also go to all the professors I had this year of the Master Program. They helped me throughout the duration of my lessons and were operated as standards for me.

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

VITA

I was born in Athens in 1989. I attended “Ionios School” from kindergarten to high school and graduated in 2006. The same year I entered the department of Statistics and Insurance Science in the University of Piraeus. In 2012 and was accepted in the Master Programs in Computer Science in University of Piraeus and in Statistics in Athens University of Economics and Business, which I post-poned until my first Master’s completion in 2014 .

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

ABSTRACT

Sophia Chrysa Mitraka

Process Control Techniques for Profile Monitoring

July 2016

Statistical quality control is a major branch of science of statistics purporting to detect errors occurring in automated production processes. Statistical quality control is under a continuous, critical evaluation of analytical methods and includes a detailed course / process starting from sample input. The most important tool in quality control is the use of control charts (control charts).

This paper attempts a detailed overview of the important elements of Statistical Quality Control. The Statistical Quality Control (Statistical Process Control, SPC) is one of the most reliable tools for the monitoring of the product, ensuring that the processes are under control and also enables business to their long-term survival and especially their profitability. Also, Multivariable Statistical Quality Control (Multivariate Statistical Process Control, MSPC) enables monitoring of two or more variables of a product simultaneously. The need to use multivariate models emerged from the finding that the quality of a product can be associated with more than one quality and measurable characteristics. Of the various techniques resulting MSPC charts, whose analysis and monitoring lead to findings whose presence would otherwise be very difficult or impossible.

The present study deals with the presentation through the literature of multivariate charts as an extension of one-dimensional, which seems to be more prevalent today, and continuing with the description of the profile monitoring. The Profile monitoring is a relatively new technique in statistical quality control is best used when data processing follows a profile (or curve) in each period.

Keywords: Statistical Quality Control, Multivariate Statistical Quality Con trol, Process Control, Profile Monitoring

https://doi.org/10.26219/heal.aueb.6243

ΠΕΡΙΛΗΨΗ

Σοφία Χρύσα Μητράκα

Τεχνικές Ελέγχου Διαδικασιών για την Παρακολούθηση Προφίλ

Ιούλιος 2016

Ο στατιστικός έλεγχος ποιότητας αποτελεί ένα μεγάλο κλάδο της επιστήμης της στατιστικής ο οποίος αποσκοπεί στην ανίχνευση των σφαλμάτων που συμβαίνουν σε αυτοματοποιημένες διαδικασίες παραγωγής. Ο στατιστικός έλεγχος ποιότητας βρίσκεται κάτω από μία συνεχόμενη, κριτική αξιολόγηση των αναλυτικών μεθόδων και περιλαμβάνει την αναλυτική πορεία/διαδικασία ξεκινώντας από την είσοδο του δείγματος. Το πιο σημαντικό εργαλείο στον έλεγχο ποιότητας είναι η χρήση των διαγραμμάτων ελέγχου (control charts).

Στην παρούσα εργασία επιχειρείται μια λεπτομερής ανασκόπηση των σημαντικότερων στοιχείων του Στατιστικού Ελέγχου Ποιότητας. Ο Στατιστικός Έλεγχος Ποιότητας (Statistical Process Control, SPC) είναι ένα από τα πιο αξιόπιστα εργαλεία για την παρακολούθηση του παραγόμενου προϊόντος, διασφαλίζοντας ότι οι διεργασίες βρίσκονται υπό έλεγχο ενώ παράλληλα εξασφαλίζει στις επιχειρήσεις την μακροχρόνια επιβίωσή τους και ειδικότερα την κερδοφορία τους. Επίσης, ο Πολυμεταβλητός Στατιστικός Έλεγχος Ποιότητας (Multivariate Statistical Process Control, MSPC) δίνει την δυνατότητα παρακολούθησης δύο ή περισσοτέρων μεταβλητών ενός προϊόντος ταυτόχρονα. Η ανάγκη χρήσης πολυμεταβλητών μοντέλων προέκυψε από τη διαπίστωση ότι η ποιότητα ενός προϊόντος μπορεί να σχετίζεται με περισσότερα του ενός ποιοτικά και μετρήσιμα χαρακτηριστικά. Από τις διάφορες τεχνικές του MSPC προκύπτουν διαγράμματα, των οποίων η ανάλυση και παρακολούθηση οδηγούν σε διαπιστώσεις των οποίων η παρουσία θα ήταν διαφορετικά πολύ δύσκολη έως και αδύνατη.

Η παρούσα μελέτη ασχολείται με την παρουσίαση μέσα από τη σχετική βιβλιογραφία των πολυμεταβλητών διαγραμμάτων ως προέκταση των μονοδιάστατων, που όπως φαίνεται είναι περισσότερο διαδεδομένα σήμερα, και συνεχίζουμε με την περιγραφή του profile monitoring. Το Profile monitoring είναι μια σχετικά

νέα τεχνική στον στατιστικό έλεγχο ποιότητας που χρησιμοποιείται καλύτερα όταν η επεξεργασία δεδομένων ακολουθεί ένα προφίλ (ή καμπύλη) σε κάθε χρονική περίοδο.

Λέξεις Κλειδιά: Στατιστικός Έλεγχος Ποιότητας, Πολυμεταβλητός Έλεγχος Ποιότητας, Έλεγχος Διαδικασιών, Παρακολούθηση Προφίλ

https://doi.org/10.26219/heal.aueb.6243

TABLE OF CONTENTS

Page

https://doi.org/10.26219/heal.aueb.6243

LIST OF FIGURES

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

CHAPTER 1

Introduction

The Statistical Quality Control is the oldest and best known method of controlling production processes to improve product quality. One of the main objectives is the early detection of non-compliant with the produced product specifications which marks the corrective actions to eliminate the causes responsible for the deviations, thus contributing to maintaining product quality. We could say that the Statistical Quality Control affects important decisions related to the specification, production and testing of products of a company.

For efficient use of Statistical Quality Control required that integration in a framework of operation and administration of the company aims to continuously improve quality at all levels of the company, known as TQM (Total Quality Management or Total Quality Assurance). We could say that the Statistical Quality Control was the harbinger of TQM.

The quality history begins in the early 20th century with the creation of the first laboratory for standards in Great Britain and on the production line at Ford Automotive in Highland Park in the USA (1905). The steps then was rapid and now the quality has become a major concern of both industry and services.

The Statistical Quality Control consists of a set of statistical data analysis. This total can be divided into three key subsets each containing oriented statistical methods in different production phases.

The three subsets are:

- Design of Experiments
- Statistical Process Control
- Acceptance Sampling

The Design and Experimental Analysis contains all those statistical techniques that help us discover the effect of different levels of the factors (variables) that affect the quality parameters of the final product and thus plays an important role in the optimal design of the production process. The Statistical Process Control provides statistical techniques needed to control the production process during

https://doi.org/10.26219/heal.aueb.6243

the production. The Acceptance Sampling provides statistical techniques (sampling) necessary to decide whether a given batch (heap) of products will be accepted or rejected.

Modern quality assurance systems usually focus on Statistical Process Control and Design of Experiments issues despite the Acceptance Sampling since it doesn't feedback the system with critical information that could lead to improved quality. To satisfy the user, a product must be produced in accordance with a "stable repeatable" process. The process should be capable of operating with little variability around some target values set in the qualitative characteristics that must distinguish the final product. The Statistical Process Control is a collection of tools that are useful for monitoring the stability of a process and improve the capability via the reduction of volatility.
The Statistical Process Control can be applied to any process. The seven main tools used are the following:

- The Histogram or Stem-and-Leaf Plot
- The Check Sheet
- The Pareto Chart
- The Cause-Effect Chart
- The Defect Concentration Chart
- The Scatter Plot
- The Control Chart

The most widely used method for monitoring a production process in Statistical Quality Control are the control charts. The X Shewhart control chart is the best known control chart for monitoring the mean of the distribution (often a normal distribution) of a characteristic of the products produced by a production process. With statistical control charts it determines whether the process is under statistical control. When the process is under statistical control, the results are provided and, therefore, within control limits. There are differences from the desired value, but rather by the nature of the process and how it is defined. To provoke change of these results, if desired, the process itself must be changed and the parameters as well. Indication of a problem in the process, may be (if the indicator is not

https://doi.org/10.26219/heal.aueb.6243

incorrect) occured when measuring out of control limits. Then there is a special event that generates most of the expected volatility that is a natural cause of the process. Possible indications of special cause variation are consecutive readings with upward or downward trends and consecutive indications above (or below) the mean (Deming, 1986).

The control chart shows the progress of the quality characteristics as a function of time. Its use is to detect changes in the historical performance of each process and to eliminate the causes that cause. The most common form of control chart consists of a central line (K2), which corresponds to the process smooth operation, the upper control limit (UCL) and the lower control limit (LCL). Any point outside the control limits is an indication that there is a change in the process, so you need to do research on the causes of this change.
Of course, at this point it is worth highlighting that the control charts alone do not result improvement. They must be applied in an environment where there is trust and cooperation, in an environment where the barriers between the various divisions of an organization have been lifted. Thus it becomes obvious how relevant is the Deming philosophy, with 14 points, for any organism (Owen, 2000). In this paper we seek introduction to statistical quality and process control and studying the basic concepts and the main models governing. In the first chapter we present an introduction to the basic concepts of statistical quality control and its importance. In the second chapter we analyze the main tools of statistical control is the control charts. In the third chapter of the work we are expanding to the multivariate control charts. While the two following chapters attempt to import the profile monitoring. Finally in the sixth chapter we present our conclusions.

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

CHAPTER 2

Univariate Control Charts

2.1 Introduction

Control charts can be classified into several categories according to certain characteristics. A basic distinction is made depending on the number of quality characteristics for which measurements are taken. So if the readings taken relating to an attribute, we refer to univariate control charts, and if the readings taken relate to several qualities we refer to multivariate control charts.

This chapter presents the basic univariate control charts and their properties. In subsection 2.2, are presented some basic concepts associated with the construction of a control chart. Subsection 2.3 presents the Shewhart control charts, for samples and for individual observations to monitor the mean and variance of a manufacturing process. Subsection 2.4 presents the CUSUM control charts, for samples and for individual observations to monitor the mean and variance of a process. Finally, subsection 2.5 presents the EWMA control charts for individual observations and samples.

2.2 Basic concepts

Control charts are the main tool of Statistical Quality Control, as their use is possible to test the stability of a process production, and to detect a change in the values of the qualitative characteristic study measurements.

The control chart is the graphical representation of a quality characteristic, which has been measured or calculated from a set of measurements as a function of the serial measurement or time (Montgomery, 2009), and as mentioned above, is used to control the volatility of a process. In this chart depicted as points the values of the quality characteristic measurements, linked together by a zigzag line, and three other usually straight lines. The centreline (CL) usually represents the mean value of the studied qualitative characteristic measurements, which results from the operation of a process that is in control, namely a production process which works with natural, non-systematic variability. The other two lines contained in

https://doi.org/10.26219/heal.aueb.6243

a control chart called upper and lower control limits (UCL and LCL). In the case where all the control plot charts are within these limits, it can be assumed that the process is in control and thus it doesn't needs to be performed corrective action. However, a point outside these limits provide evidence of an except process out of control. Research is therefore essential for the detection of special causes of variation, which led to this behaviour, and their subsequent elimination. It is worth mentioning that it is not enough all the points in a control chart to be within the above limits, but it is further necessary to behave in a random, non-systematic way. For example, if 18 of the 20 points in a control chart located above the centre line but below the upper limit, and only two are located below the centre line, but above the lower limit, the strong suspicion of an out of control operation production process is born(Montgomery, 2009).

For controlling a production process with the use of control charts, there are two phases, Phase I and Phase II. In Phase I of measuring which retrospectively analyzed in order to check whether the process was in control at the time / duration of collection of data. At this phase, the charts play mainly supporting role to the administrator of the process to bring the process in statistical control. When this is achieved, the generated control charts can be used for future monitoring of the process.

In Phase II the control charts are used for the continuous monitoring of the production process and the early detection of a possible shift in the mean level of quality of the product. So, at this phase the focus is on process monitoring, and not to become an out of control process in control. Early detection of a change in the quality of the products obtained by continuous sampling of the production process and comparing statistical function derived from them (usually the mean) with the control chart limits (Montgomery, 2009).

Another concept associated with control charts and to be presented in this subsection, is the Mean Run Length ARL. Except as outlined above, control charts serve two more purposes. One refers to that, when the process is in control, we want the chart to signal for an out of control process as rarely as possible. The second relates to the fact that, when the process is out of control, we want the chart to signal as soon as possible. Referring, then, to evaluate the performance

https://doi.org/10.26219/heal.aueb.6243

of control charts on the achievement of these two objectives, the measure most often used is the mean run length ARL. The amount ARL indicates the expected number of points to be planned in a control chart until a point will signal of an out of control production process. The amount ARL can be calculated from the relation

$ARL = \frac{1}{p}$ wherein p is the probability of finding a point of control chart outside the control limits. ARL_0 is defined as the in control ARL, which indicates the expected number of points to be planned in a control chart until a point that will signal to an out of control production process, when in reality the process is in control. As ARL1 we define the out of control mean run length, which indicates the expected number of points to be planned in a control chart until a point that will signal to an out of control production process, when in reality the process is out of control due displacement of μ in μ*. Namely, the ARL_1 indicates the expected number of measurements / samples to be taken to detect the shift in the mean of the process, from the moment it happened.

Clearly, for a process that is in control we want to have great value for ARL_0, so that the number of false indications for out of control process to be small. Accordingly, we want to have a small value for the ARL_1, so that the number of samples / observations, and therefore the time required to detect the displacement of the process mean to be small.

However, use of the ARL for evaluating the performance of a production process has been reviewed in recent years. This is because the distribution of the Run Length follows the Geometric distribution (G(p)), therefore the standard deviation of the Run Length is quite large and the distribution of the Run Length is highly asymmetric. Therefore, the mean value of the distribution (i.e. ARL) can not be considered as representative of the Run Length (Montgomery, 2009). For this reason, as alternative to the Mean Run Length can be used (MRL), since it is less sensitive to the asymmetry of the distribution aspects of the Run Length.

Even beyond the ARL, as a production process performance evaluation measure is used the Mean Time to Signal ATS, which is given by

$$ATS = ARL \times h$$

wherein h is the time interval between receiving two consecutive measurements/samples and which is constant. Therefore, the ATS indicates the mean time required to provide the control chart display for an out of control process.

2.3 Control Charts Shewhart

In this subsection we study the construction Shewhart-type control charts for variables. These charts are used to monitor the mean value and variance value of a qualitative characteristic, for samples and for individual observations. The following paragraphs are initially presented the Shewhart control charts for samples and then the Shewhart control charts for individual observations.

2.3.1 Shewhart Control Charts for samples

2.3.1.1 Shewhart Control Charts to monitor the mean

In this subsection are presented the Shewhart control charts for monitoring the mean value of a qualitative characteristic X. In each case it is assumed that the variance, and therefore the standard deviation σ of the qualitative characteristic X values remains constant.

Initially it is assumed that the distribution of the qualitative characteristic X follows a normal distribution $N(\mu, \sigma^2)$. If $X_i = (X_{i1}, X_{i2}, \dots, X_{in})$ a random sample of size n from X then the sample mean is calculated as

$$X_i = \frac{X_{i1} + X_{i2} + \cdots + X_{in}}{n},$$

It follows $N(\mu, \sigma^2/n)$ and is an unbiased estimate of the mean value of X. Therefore, for any n sizedsample the sample mean X_i takes values in the interval

$$\left[\mu - Z_{a/2}\frac{\sigma}{\sqrt{n}}, \mu + Z_{a/2}\frac{\sigma}{\sqrt{n}}\right]$$

with probability$1-\alpha$. For α=0.0027, i.e. $Z_{a/2} = 3$, we take the interval

 https://doi.org/10.26219/heal.aueb.6243

$$\left[\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}\right]$$

and the sample mean X_i takes values in this interval with a probability of 99.73%. Therefore, in the case where μ and σ are known, the boundaries of the interval may be used as the upper and lower limits of a control chart. Therefore, in the case where a point on the chart (sample mean) is outside this range, and since the probability of this happening is extremely small (0.0027), then there is an indication that the manufacturing process is out of control due to displacement of the mean from the value μ. These limits are known as three-sigma control limits of Phase II, when the mean's value and the variation of the qualitative characteristic X are known.

2.3.1.1.1 Estimation of the mean

In practice the mean and the variance of the qualitative characteristic X is largely unknown, therefore the focus is on their estimation for the Phase I control charts' construction. For this purpose, it is necessary to collect preliminary m independent random samples $X_{i1}, X_{i2}, \ldots, X_{in}$, n sized each, provided that they were collected when the process was in control.

Suppose, then, that $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_m$ are the sample means of m independent preliminary random samples. Then the best estimate of the process mean is the amount

$$\bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \cdots + \bar{X}_m}{m}$$

which follows the distribution N(μ,σ^2/nm) and is unbiased estimator of the mean of the process. The amount $\bar{\bar{X}}$ is used as the central line in Shewhart control charts. But in order to build them, we need to estimate the process variance. To this end, the following paragraphs present the three most popular methods.

https://doi.org/10.26219/heal.aueb.6243

2.3.1.1.2 Estimation of the standard deviation with Method R

Let $R_1, R_2, \dots, R_m$ the ranges of m preliminary samples, ie

$R_i = X_{i(n)} - X_{i(1)}, \quad 1 \leq i \leq m$.

Then the mean width is calculated as

$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_m}{m}$$

and the estimation of σ is calculated as

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

Where d_2 is a constant that depends on the sample size n.

Thus, using as an estimate of the mean of the process the amount $\bar{\bar{X}}$ and as the estimated standard deviation the amount $\hat{\sigma} = \bar{R}/d_2$, the three sigma control limits of Phase I of a Shewhart chart are as follows:

$$UCL = \bar{\bar{X}} + A_2 \cdot R$$

$$CL = \bar{\bar{X}}$$

$$LCL = \bar{\bar{X}} - A_2 \cdot R$$

Wherein the A_2 constant is equal to $A_2 = \frac{3}{d_2 \cdot \sqrt{n}}$.

2.3.1.1.3 Estimation of the standard deviation with the S method

Let $S_1, S_2, \dots, S_m$ the quantities defined by the relation

$$S_i = \sqrt{\frac{1}{n-1} \cdot \sum_{j=1}^{n} \left(x_{ij} - \bar{x}_j\right)^2}, \quad 1 \leq i \leq m.$$

Posing

$$\bar{S} = \frac{S_1 + S_2 + \cdots + S_m}{m}$$

we take $E(\bar{S}) = \sigma \cdot c_4$, wherein c_4 a constant that depends on the sample size, and it can be shown that the amount $\bar{S}/c_4$ is unbiased estimator of the quantity σ , i.e.

https://doi.org/10.26219/heal.aueb.6243

$$\hat{\sigma} = \frac{\bar{S}}{c_4}$$

(Montgomery, 2009).

Thus, using as an estimate of the process mean the amount $\bar{\bar{X}}$ and as the estimated standard deviation, the amount $\hat{\sigma} = \bar{S}/c_4$, the three sigma control limits of Phase I of a Shewhart chart are as follows:

$$UCL = \bar{\bar{X}} + A_3 \cdot \bar{S}$$

$$CL = \bar{\bar{X}}$$

$$LCL = \bar{\bar{X}} - A_3 \cdot \bar{S}$$

wherein the constant A_3 is equal to

$$A_3 = \frac{3}{c_4 \cdot \sqrt{n}}.$$

2.3.1.1.4 Estimation of the standard deviation to the process S²

Let $S_1^2, S_2^2, \dots, S_m^2$ the quantities defined by the relation

$$S_i^2 = \frac{1}{n-1} \cdot \sum_{j=1}^{n} \left(x_{ij} - \bar{x}_j\right)^2, \ \ 1 \le i \le m$$

for which it holds that

$$E(S_i^2) = \sigma^2 .$$

The amount $\sqrt{\overline{S^2}}$ wherein

$$\overline{S^2} = \frac{S_1^2 + S_2^2 + \cdots + S_m^2}{m}$$

can be used as an estimate (although not impartial)of the quantity σ, i.e. $\hat{\sigma} = \sqrt{\overline{S^2}}$.

Thus, using as an estimate of the process means the amount $\bar{\bar{X}}$ and as the estimated standard deviation the amount $\hat{\sigma} = \sqrt{\overline{S^2}}$, the three sigma control limits Phase I of a Shewhart chart are as follows:

$$UCL = \bar{\bar{X}} + A \cdot \sqrt{\overline{S^2}}$$

$$CL = \bar{\bar{X}}$$

$$LCL = \bar{\bar{X}} - A \cdot \sqrt{\overline{S^2}}$$

https://doi.org/10.26219/heal.aueb.6243

2.3.1.2 Shewhart Control Charts for Monitoring the variance

As mentioned above, control charts described in the previous paragraphs relating to the monitoring of the mean value of a qualitative characteristic X, provided that the variance of the characteristic X was constant throughout the course of production process. However, in practice it is not always certain that the variance of the characteristic X will remain constant throughout the procedure. This sub-division presents Shewhart control charts for monitoring the variances' behaviour of a qualitative characteristic X.
As in previous sections, in each of the cases presented below it is assumed that the X attribute values follow a normal distribution $N(\mu, \sigma^2)$ and that obtained m independent random samples from the X, n each size.

2.3.1.2.1 R control chart for monitoring the variance

Firstly the case is considered in which μ and σ^2 are considered known. Let $X_i = X_{i1}, X_{i2}, \dots, X_{in}$ m random samples of size n from X and even $R_i = R_1, R_2, \dots, R_m$ ranges of samples. It can be calculated that

$\mu_{R_i} = E(R_i) = \sigma \cdot d_2$, $\quad \sigma_{R_i} = \sqrt{Var(R_i)} = \sigma \cdot d_3$

wherein d_2, d_3 amounts depending on the sample size n (Montgomery, 2009). Therefore, a Phase II control chart for monitoring the variance of qualitative characteristic X based on sample ranges R_i are the following:

$$UCL = D_2 \cdot \sigma$$
$$CL = \sigma \cdot d_2$$
$$LCL = D_1 \cdot \sigma$$

wherein $D_1 = d_2 - 3 \cdot d_3, D_2 = d_2 + 3 \cdot d_3$.
If σ is unknown, then using as estimate of, the amount $\hat{\sigma} = \bar{R}/d_2$, a Phase I control chart for monitoring the variance of the qualitative characteristic X values is the following:

https://doi.org/10.26219/heal.aueb.6243

$$UCL = D_4 \cdot \bar{R}$$

$$CL = \bar{R}$$

$$LCL = D_3 \cdot \bar{R}$$

Wherein

$$D_3 = 1 - 3 \cdot \frac{d_3}{d_2}, \qquad D_4 = 1 + 3 \cdot \frac{d_3}{d_2}.$$

2.3.1.2.2 S control chart for monitoring the variance

As before, initially the case is shown in which μ, σ^2 are considered known. Let $X_i = X_{i1}, X_{i2}, \ldots, X_{in}$, m random n-sized samples from X and let the amounts $S_i = S_1, S_2, \ldots, S_m$ defined by the relation

$$S_i = \sqrt{\frac{1}{n-1} \cdot \sum_{j=1}^{n} \left(x_{ij} - \bar{x}_j\right)^2}, \quad 1 \leq i \leq m.$$

It is proved that

$$\mu_{S_i} = \sigma \cdot c_4, \qquad \sigma_{S_i} = \sigma \cdot \sqrt{1 - c_4^2}.$$

Therefore, a Phase II control chart for monitoring the variances' value of qualitative characteristic X based on standard deviations S_i are the following:

$$UCL = B_6 \cdot \sigma$$

$$CL = \sigma \cdot c_4$$

$$LCL = B_5 \cdot \sigma$$

Wherein

$$B_5 = c_4 - 3 \cdot \sqrt{1 - c_4^2}, \qquad B_6 = c_4 + 3 \cdot \sqrt{1 - c_4^2}.$$

If σ is unknown, then using as its estimate, the amount $\hat{\sigma} = \bar{S}/c_4$, a Phase I control chart for monitoring the variances' value of the qualitative characteristic X are the following:

 https://doi.org/10.26219/heal.aueb.6243

$$UCL = B_4 \cdot \bar{S}$$

$$CL = \bar{S}$$

$$LCL = B_3 \cdot \bar{S}$$

Wherein

$$B_3 = 1 - \frac{3}{c_4} \cdot \sqrt{1 - c_4^2}\,, \qquad B_4 = 1 + \frac{3}{c_4} \cdot \sqrt{1 - c_4^2}\,.$$

2.3.1.2.3 S^2 control chart for monitoring the variance

Similarly to the previous cases, initially we show the case where μ and σ^2 are considered known. Let $X_i = X_{i1}, X_{i2}, \dots, X_{in}$ m random samples of size n from X and let $S_1^2, S_2^2, \dots, S_m^2$ the amounts defined by the relation

$$S_i^2 = \frac{1}{n-1} \cdot \sum_{j=1}^{n} \left(x_{ij} - \bar{x}_j\right)^2, \;\; 1 \le i \le m\,.$$

It is proved that

$$E(S_i^2) = \sigma^2 \quad and \quad \frac{(n-1) \cdot S_i^2}{\sigma^2} \sim \chi_{n-1}^2\,.$$

Still, it holds that

$$P\left(\chi_{n-1;1-a/2}^2 \le \frac{(n-1) \cdot S_i^2}{\sigma^2} \le \chi_{n-1;a/2}^2\right) = 1 - a$$

Therefore

$$P\left(\frac{\sigma^2}{n-1} \cdot \chi_{n-1;1-a/2}^2 \le S_i^2 \le \frac{\sigma^2}{n-1} \cdot \chi_{n-1;a/2}^2\right) = 1 - a\,.$$

https://doi.org/10.26219/heal.aueb.6243

Therefore, a Phase II control chart with $a/2$ probability thresholds for the monitoring of variance's values of qualitative characteristic X based on variances S_i^2 is the following:

$$UCL = \frac{\sigma^2}{n-1} \cdot \chi^2_{n-1;a/2}$$

$$CL = \sigma^2$$

$$LCL = \frac{\sigma^2}{n-1} \cdot \chi^2_{n-1;1-a/2}$$

If σ is unknown, then using as an estimate of σ^2, the amount $\widehat{\sigma^2} = \overline{S^2}$, a Phase I control chart with $a/2$ probability thresholds for the monitoring of the variance's value of the qualitative characteristic X is the following:

$$UCL = \frac{\overline{S^2}}{n-1} \cdot \chi^2_{n-1;a/2}$$

$$CL = \overline{S^2}$$

$$LCL = \frac{\overline{S^2}}{n-1} \cdot \chi^2_{n-1;1-a/2}$$

2.3.2 Shewhart Control Charts for individual comments

In the event that no samples collected, but individual observations, control charts presented in the above paragraphs cannot be applied to monitor the behaviour of qualitative characteristic X. Therefore it is needed an appropriate modification. As before, it is assumed that the distribution of the qualitative characteristic X is normal $N(\mu, \sigma^2)$.

2.3.2.1 Control Charts to monitor the mean

Firstly the case is considered in which μ, σ^2, considered known. The Phase II control chart for monitoring the mean of qualitative characteristic X with three sigma control limits presented in the last subsection takes the following form:

https://doi.org/10.26219/heal.aueb.6243

$$UCL = \mu + 3 \cdot \sigma$$

$$CL = \mu$$

$$LCL = \mu - 3 \cdot \sigma$$

and illustrated in this, individual observations X_i.
In the case where μ and σ^2 of the attribute X is unknown, they must be estimated. Suppose, then, a random m sized sample derived from the characteristic X, $X_1, X_2, \dots, X_m$. Then an unbiased estimator of the mean of the sample is given by the following relation:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_m}{m}.$$

To estimate the variance of sample values is not meaningful to calculate the range R. Therefore, in this case the moving range of the individual observations is used, which is calculated as follows:

$$MR_i = |X_i - X_{i-1}| = max(X_{i-1}, X_i) - min(X_{i-1}, X_i), \qquad i \geq 2\,.$$

Posing

$$\overline{MR} = \frac{MR_2 + MR_3 + \cdots + MR_m}{m-1}$$

It turns out $E(\overline{MR}) = \sigma \cdot d_2$, therefore the amount $\overline{MR}/d_2$ is an unbiased estimator of quantity σ. Thus, the Phase I control chart for the monitoring of the mean value of the qualitative characteristic X in three sigma control limits is:

$$UCL = \bar{X} + 3 \cdot \frac{\overline{MR}}{d_2}$$

$$CL = \bar{X}$$

$$LCL = \bar{X} - 3 \cdot \frac{\overline{MR}}{d_2}$$

2.3.2.2 Control charts for monitoring the variance

For the monitoring of the variance of qualitative characteristic X the corresponding R control chart cannot be used, since, as mentioned above there is no sense to calculate the range R. Therefore, in this case the moving range of the individual observations are used. For the moving range MR_i is estimated

$$\mu_{MR_i} = \sigma \cdot d_2, \qquad \sigma_{MR_i} = \sigma \cdot d_3 .$$

Therefore, a Phase II control chart with three sigma limits for monitoring the variance of the qualitative characteristic X values is the following:

$$UCL = D_2 \cdot \sigma$$

$$CL = \sigma \cdot d_2$$

$$LCL = D_1 \cdot \sigma$$

wherein $D_1 = d_2 - 3 \cdot d_3$, $D_2 = d_2 + 3 \cdot d_3$

If σ is unknown, then using as its estimation, the quantity $\hat{\sigma} = \frac{\overline{MR}}{d_2}$, a Phase I control chart with three sigma limits for monitoring the value’s variance of the qualitative characteristic X is as follows:

$$UCL = D_4 \cdot \overline{MR}$$

$$CL = \overline{MR}$$

$$LCL = D_3 \cdot \overline{MR}$$

Wherein

$$D_3 = 1 - 3 \cdot \frac{d_3}{d_2}, \; D_4 = 1 + 3 \cdot \frac{d_3}{d_2}.$$

2.4 CUSUM Control charts

According to the bibliography (Montgomery, 2009), the Shewhart control charts presented in the preceding paragraphs are useful mainly for Phase I of statistical process control, where the production process being out of control is possible.

https://doi.org/10.26219/heal.aueb.6243

However, a major disadvantage of Shewhart control charts is the fact that they use the information received from the last sample each time, ignoring the information provided by previous samples. Thus, these charts are relatively ineffective in detecting small displacements of the mean value of a process, especially when it shifts to a half standard deviation. For this reason, it's not considered suitable for Phase II, if at this stage the production process is largely in control and any mean displacements are small.

By contrast, in CUSUM control charts the design of a point based not only on the most recent sample, but also on information collected from previous samples. For this reason, the CUSUM control chart belongs to a chart class called Memory Control Charts, as well as EWMA charts which will be presented in the next section, and are mainly used in Phase II.

The CUSUM control charts exploit all the information given by earlier samples, showing a chart of the cumulative sums of the deviations of the observations from a target value. If then, collected samples of size n and the amount $\overline{x_j}$ represents the mean value of j-th sample and μ_0 the target value of the production process, then a cumulative control chart is made reflecting the values of the quantity

$$C_i = \sum_{j=1}^{i} (\overline{x_j} - \mu_0)$$

function of the number i of the sample (Montgomery, 2009). According to the way in which the amount C_i is readily calculated found that shifting the process's mean to a value $\mu_1 > \mu_0$ will result a positive or upward drift of the accumulated sum C_i. Similarly, a drift of the process mean to a value $\mu_1 < \mu_0$, it will cause a negative or downward drift of the accumulated sum C_i. Therefore, an upward or downward trend in the chart points indicates displacement of the mean of the production process (Montgomery, 2009) . Figure 2.1 shows a cumulative chart in which the mean of the production process has shifted from position $\mu = 10$ to position $\mu_1 = 11$.

https://doi.org/10.26219/heal.aueb.6243

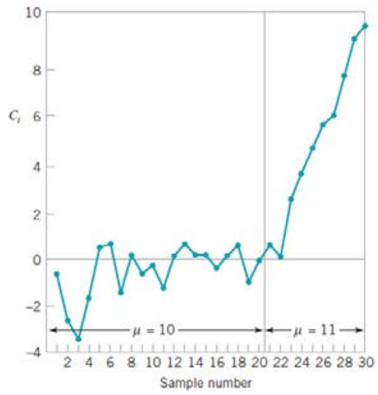

Figure 2.1: Example of a CUSUM chart

Source: Montgomery (2009)

However, as it can be seen, the cumulative chart of Figure 2.1 is not a control chart, as it has no control limits. For the construction of cumulative chart with control limits, there are two methods: algorithmic and method of the V mask. Because the process of V mask is generally not recommended for the construction of cumulative control charts Montgomery (2009), in the present work only the algorithmic method is presented.

2.4.1 Algorithmic Method CUSUM for the mean of a process

The CUSUM control charts are used in either case the measurements from individual observations or events related to the observations samples. However, in practice they are usually used to represent the individual observations. For this reason it will be given more weight in the presentation of CUSUM control chart for individual observations.

Suppose, then, that received individual comments $x_i, i \geq 1$ from an in control process of production of a qualitative characteristic X. For these observations are assumed to come from a normal distribution with mean μ_0 and standard deviation σ (known). The algorithmic CUSUM method for detecting upward displacements of the mean is using the statistic

$$C_i^+ = max[0, x_i - (\mu_0 + K) + C_{i-1}^+],\ \ C_0^+ = 0$$

and for detecting downward displacements of the mean is using the statistic

https://doi.org/10.26219/heal.aueb.6243

$C_i^- = max[0, (\mu_0 - K) - x_i + C_{i-1}^-], \; C_0^- = 0$

where the values of the quantities C_0^+ and C_0^- values called head starting values and K is a positive constant called reference value. The value of the constant K is usually chosen as half the distance between the target value μ_0 and the out of control value of mean, μ_1, which are interested to find. Therefore the value of K usually given by $K = \frac{\delta}{2}\sigma = \frac{|\mu_1 - \mu_0|}{2}$.

The amount C_i^+ accumulates the deviations of the observations x_i from the quantity $\mu_0 + K$ and can be considered suitable for the control of the case

$H_0: \mu = \mu_0$

$H_1^+: \mu = \mu_1 = \mu_0 + \delta\sigma, \qquad \delta > 0$

since large positive values C_i^+ lead to acceptance of H_1^+. Correspondingly, the amount C_i^- accumulates the deviations of the observations x_i from the quantity $\mu_0 - K$ and can be considered suitable for the control of the case

$H_0: \mu = \mu_0$

$H_1^-: \mu = \mu_1 = \mu_0 - \delta\sigma, \qquad \delta > 0$

since large negative values C_i^- lead to acceptance of H_1^-. The decision as to whether the production process is out of control, ie whether to accept any of H_1^+ or H_1^- based on a constant H. More specifically, if $C_i^+ > H$ or $C_i^- < -H$, then the corresponding null hypothesis is rejected and it is concluded that the process is out of control due to displacement of the mean at a higher or lower level respectively. The amount H is called a decision interval and is given by the formula $H = h \cdot \sigma$ where normally $h = 4$ or 5. Figure 2.2 shows a typical cumulative CUSUM chart for one process out of control due to displacement of the mean at a higher level where $H = 5$.

https://doi.org/10.26219/heal.aueb.6243

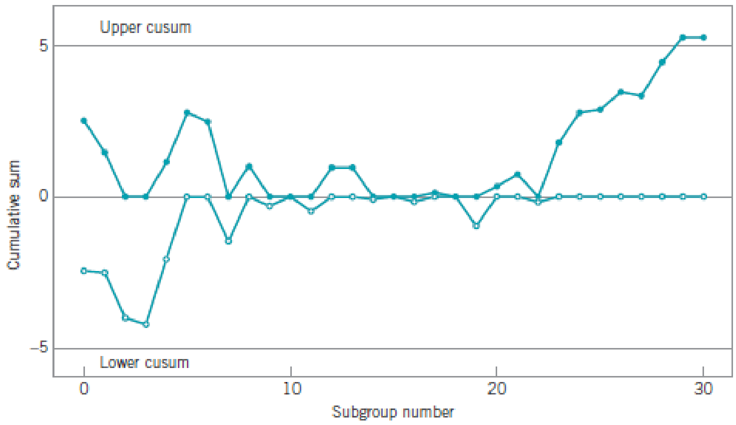

Figure 2.2: Example of a CUSUM chart with decision time H=5

Source: Montgomery (2009)

If the CUSUM control chart detects a displacement of the processing means, then it is useful to calculate the value of the new mean μ_1. This is done using the following relations:

$$\hat{\mu}_1 = \begin{cases} \mu_0 + K + \dfrac{C_i^+}{N^+}, & \alpha\nu\ C_i^+ > H \\ \mu_0 - K - \dfrac{C_i^-}{N^-}, & \alpha\nu\ C_i^- < -H \end{cases}$$

where N^+,N^- the number of consecutive periods (until the indication for out of control process) for which the accumulated totals C_i^+or $C_i^-$is different than zero (Montgomery, 2009).

If the focus is on the detection of the process's mean displacements in one direction, i.e. in a higher (lower) single plane, then the aggregated chart illustrated only the quantities C_i^+ (C_i^-) and the decisions interval H^+ (H^-). These charts are called one-sided CUSUMs.

2.4.2 Standard CUSUM Charts

In addition to the cumulative CUSUM charts developed in the preceding paragraphs, the bibliography indicated the Standardized CUSUMs. So if $Y = \frac{X-\mu_0}{\sigma}$ is the standard value of X, then the standard algorithmic method of cumulative chart uses statistical functions

https://doi.org/10.26219/heal.aueb.6243

$$C_i^+ = max[0, y_i - k + C_{i-1}^+], \qquad C_0^+ = 0$$

$$C_i^- = max[0, -k - y_i + C_{i-1}^-], \qquad C_0^- = 0$$

where if $C_i^+ > h$ ($C_i^- < -h$)then the manufacturing process mean has shifted to a higher (lower) level. One of the advantages of standardized CUSUM chart is that different charts can use the same values for the parameters k and y, and the choice of these parameters does not depend on the variance of the sample values. Another advantage is that a standardized CUSUM chart is suitable for cases in which the variance is not stable (Montgomery, 2009).

2.4.3 CUSUM Control Charts for dispersing a process

According to Hawkins (1981) (as indicated in Montgomery, 2009) it is possible to construct CUSUM control charts to monitor the variance of a production process. If x_i the observations taken from a production process, and $X{\sim}N(\mu_0, \sigma^2)$ then the standardized value is X, as mentioned above, $Y = \frac{X-\mu_0}{\sigma}$ The Hawkins proposed the creation of a new standardized quantity,

$$v_i = \frac{\sqrt{|y_i|} - 0.822}{0.349}$$

which are sensitive to displacements of the production process of variance and which the in control distribution V is the standard normal. Thus, a control chart for the variance of the production process (Scale CUSUM) may be based on the display of the following statistical functions

$$S_i^+ = max[0, v_i - k + S_{i-1}^+], \qquad S_0^+ = 0$$

$$S_i^- = max[0, -k - v_i + S_{i-1}^-], \qquad S_0^- = 0$$

wherein the parameters k, v_i are selected as in the case of manufacturing CUSUM control chart for the monitoring of a process's mean. The interpretation of the Scale CUSUM chart is similar to the interpretation of CUSUM control chart for

https://doi.org/10.26219/heal.aueb.6243

the mean, i.e. an increase $S_i^+$of a value level higher than +h interpreted as increasing the variation of the production process, and a reduction in values $S_i^-$to below $-h$ construe as reducing variance in the production process.

2.4.4 CUSUM Control Charts for samples

The CUSUM control chart presented in the preceding paragraphs may be modified in such a way that they can be used in the case that we do not have individual observations, but sample sizes $n > 1$. In this case the quantity $x_i$is replaced by the quantity $\bar{x}_i$, i.e. the mean of the i sample, and the amount σ is substituted by the amount $\sigma_{\bar{x}} = \sigma/\sqrt{n}$. So in control chart the statistical functions are displayed:

$$C_i^+ = max[0, \bar{x}_i - (\mu_0 + K) + C_{i-1}^+], \qquad C_0^+ = 0$$

$$C_i^- = max[0, (\mu_0 - K) - \bar{x}_i + C_{i-1}^-], \qquad C_0^- = 0$$

wherein $K = k \cdot \frac{\sigma}{\sqrt{n}}$, $H = h \cdot \frac{\sigma}{\sqrt{n}}$

Still, in case we have samples rather than individual observations, it is possible to construct control charts for monitoring the variability of the production process, based on the sampling variance. Suppose, then, that the observations of the samples are normally distributed and that the in and out of control variance is $\sigma_0^2$and σ_1^2respectively. If $S_i^2$the variance of the sample, then the CUSUM control chart statistics functions are

$$C_i^+ = max[0, C_{i-1}^+ + S_i^2 + k], \qquad C_0^+ = 0$$

$$C_i^- = max[0, C_{i-1}^- + S_i^2 - k], \qquad C_0^- = 0$$

wherein $k = \frac{2\ln(\sigma_0/\sigma_1)\sigma_0^2\sigma_1^2}{\sigma_0^2 - \sigma_1^2}$ (Montgomery, 2009).

 https://doi.org/10.26219/heal.aueb.6243

2.5 Control Charts EWMA

2.5.1 EWMA control charts for the mean of a process

The EWMA control charts were introduced by Roberts (1959) and are an alternative to the Shewhart control charts when we care to detect small shifts of a production process mean. In EWMA control charts show the statistic

$$Z_i = \lambda \cdot X_i + (1-\lambda) \cdot Z_{i-1}$$

wherein λ is a constant $0 < \lambda \leq 1$. As a starting value Z_0shall be the value target of the production process μ_0 while can be used the value of the mean $\bar{X}$ calculated by preliminary data. Using sequentially the above formula we get

$$Z_i = \lambda \cdot \sum_{j=0}^{i-1}(1-\lambda)^j X_{i-j} + (1-\lambda)^i Z_0$$

As can be seen, the amount Z_i is a weighted mean of the observations $Z_0, X_1, X_2, \ldots, X_i$ with corresponding weights$(1-\lambda)^i, \lambda(1-\lambda)^{i-1}, \ldots, \lambda$. Weights $\lambda(1-\lambda)^j$ decline geometrically as we move from observation X_i to observation X_1 and their sum is equal to one.

For the variance of the statistical function Z_i we have

$$\sigma_{Z_i}^2 = Var(Z_i) = \sigma^2\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]$$

Therefore, an EWMA control chart shows the value of the statistic Z_i and the control limits are given by the following relations:

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$

$$CL = \mu_0$$

$$LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$

https://doi.org/10.26219/heal.aueb.6243

In case $\lambda = 1$ then the above EWMA control chart, is reduced to a Shewhart control chart. Also, as can be appreciated, in the case where $\lambda \neq 1$ the above control chart limits are variable. However, the quantity $(1-\lambda)^{2i}$ tends to zero as i grows, so after a relatively small number of periods, the control limits are fixed and given by the following relations:

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}}$$

$$CL = \mu_0$$

$$LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}}$$

It is worth noting that the EWMA control charts are not significantly affected by the breach of the case of normal distributed observations therefore an EWMA control chart is a good non parametric solution for the detection of small displacements of a production process mean.

2.5.2 EWMA control charts for samples

The EWMA control charts although initially designed for individual observations can be modified to be used in case of sample sizes $n > 1$. In this case the quantity $X_i$is replaced by the amount $\overline{X_\iota}$, i.e. the i sample mean and the amount σ with the quantity $\sigma/\sqrt{n}$. Therefore, in a EWMA control chart for samples illustrated quantity

$$Z_i = (1-\lambda)\cdot Z_{i-1} + \lambda \cdot \overline{X_\iota}$$

with control limits and central line given by the following relations:

https://doi.org/10.26219/heal.aueb.6243

$$UCL = \mu_0 + L\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$

$$CL = \mu_0$$

$$LCL = \mu_0 - L\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$

Amin & Searcy (1991) proposed a nonparametric formula of an EWMA control chart, based on the rating of the observations in the sample using the grouped signed ranked statistic or otherwise GRS-EWMA. The statistics function depicted given the following relation $Zi = \lambda\ SRi + (1 - \lambda)\ Zi\text{-}1$, $0 < \lambda < 1$,
wherein the initial value Z_0 is the target value that has been determined.

The process is considered to be out of control whenever any Z_i either exceeds the upper control limit (UCL) or is below the lower control limit (LCL). The control limits are given by the relation $\mu_0 \pm L$. Additional, as reference the EWMA non-parametric control chart, this has been assessed and compared on the basis of the ARL. Uniform, normal, double exponential, gamma, and Cauchy distributions are selected. Control limits for standardized GRS-EWMA and the GRS - EWMA are chosen so that the frequency of the points that are out of the control limits are the same for procedures when the process is in control.
The addition of the warning control limits improve chart's performance. It is proposed control scheme that the control chart for the variance can be used with the GRS - EWMA process.
As reference the ARL, is influenced by the values of lambda to the proposed non-parametric GRS - EWMA control chart. So we could say that the method of GRS - EWMA is a good alternative approach non parametric control chart.

https://doi.org/10.26219/heal.aueb.6243

CHAPTER 3

Multivariable Statistical Quality Control

3.1 Introduction

Monitoring correlated variables is the main object of study processed by Multivariable Statistical Process Control (MSPC) (Montgomery (2005)). That enables simultaneous monitoring of two or more variables of a product. The need to use the Multivariable Statistical Quality Control arose from the finding that the quality of a product can be associated with more than one qualitative and measurable characteristics.

The Control Charts monitor the quality of a process and how it changes over time. In practice, however, many if not most of the production processes or services are multivariate, are dependent on more than one variable. Application of Control Charts in each variable separately to monitor the quality of such a process, it is insufficient and can lead to erroneous conclusions. So multivariate methods are needed to examine variables together. However, the multivariate control charts are less popular than univariate because besides the relative difficulty they have in their calculation, they have some additional difficulties in their implementation. For example, unlike univariate control charts, the range of values that appear in multivariate charts are not related to the scales of each variable monitored. Thus, when a signal is given out of control by the multivariate chart, it is difficult to determine which variable caused this signal.

In many cases, however, monitoring a single attribute is deemed unreliable because it takes no account of the relation between two or more variables. The monitoring of each control chart individually can not provide information about whether the variables we manage are related. Similarly, we can say that the consumer does not perceive each variable of a product separate, but considers them all together (as one) when selecting a product. In no case the variables are uncorrelated. This led MacGregor and Kourti (1995) to doubt the ability of univariate control charts in determining the quality of a manufactured product. So there are

https://doi.org/10.26219/heal.aueb.6243

many cases in which the simultaneous monitoring of two or more correlated qualitative characteristics are necessary. But as the number of measurable qualitative variables increases, the difficulty to monitor also arises. Suppose we have p statistically independent qualitative variables to produce a separate product and Control Graph is P{Type I error} = a, then the actual probability of the fault I incorporating the whole production process is a'= 1- (1-α)p and the probability that all the p variables will be drawn simultaneously into the control limits when the process is in control is P {all p means are in control} = (1 - a) p. If the variables are not independent, then it is very difficult to calculate α' of the production process.

The first dealt with the Multivariable Statistical Quality Control (Multivariate Statistical Process Control, MSPC) was Harold Hotelling (1947), who applied the procedure to data collected from the Second World War on bombed areas. Three of the most important Multivariable Control Charts are Multivariable Shewhart Control Charts (for example, Hotelling T^2, which monitors the mean vector of a process), the Multivariate Cumulative Sum (MCUSUM) and the Multivariate Exponentially Weighted Moving Mean (MEWMA).

Several works reported in the bibliography relating multivariate test procedures (examples include: (Hicks, 1955), (Alt, 1985), (Jackson, 1985), (Healy, 1987), (Crossier, 1988), (Lowry et al., 1992), (Tracy, Young and Mason, 1992), (Pignatiello and Runger, 1990), (Montgomery, 1996), (Sullivan and Woodall, 1996)).

In the previous chapter we looked at the parameters and the use of one-dimensional Shewhart control charts, EWMA and CUSUM. This chapter will analyze the multidimensional form in the context of multivariate quality control analysis.

https://doi.org/10.26219/heal.aueb.6243

3.2 Multivariable charts Shewhart

3.2.1 The p-dimensional normal distribution

To p-dimensional random vector (Montgomery (2005))

$$\mathbf{X}=\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

following p-dimensional normal distribution with parameters μ and Σ (notation X ~ Np (μ,Σ)), wherein μ′ = [μ1, μ2, …, μp] is a vector of real numbers and

$$\mathbf{\Sigma}=\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix}$$

is a symmetric positive definite matrix, if the density function is given by

$$f(\mathbf{x})=\frac{1}{(2\pi)^{p/2}\,|\,\mathbf{\Sigma}\,|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\,\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right),\quad \mathbf{x}\in R^p$$

(with | Σ | denotes the determinant of the matrix Σ and Σ-1 the inverse of the matrix Σ). It can be shown that

$$E(\mathbf{X})=\boldsymbol{\mu},\quad E[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})']=\mathbf{\Sigma},\quad (\mathbf{X}-\boldsymbol{\mu})'\,\mathbf{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})\sim\chi_p^2.$$

In the special case p = 2 shows the two-dimensional normal distribution whose density function is given by

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}Q(x_1,x_2)\right),\quad x_1,x_2\in R,$$

wherein

$$Q(x_1,x_2)=\left\{\frac{1}{1-\rho^2}\cdot\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

and $\rho\sigma_1\sigma_2 = \sigma_{12}$. In the special case p = 1 it shows the one-dimensional normal distribution.

Suppose now that we have a random n sized sample from p-dimensional normal distribution with parameters μ and Σ. The sample mean X and the sampling variance - covariance matrix S are given by formulas

$$\overline{\mathbf{X}} = \begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_p \end{bmatrix} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i ,$$

$$\mathbf{S} = \begin{bmatrix} S_1^2 & S_{12} & \cdots & S_{1p} \\ S_{21} & S_2^2 & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_p^2 \end{bmatrix} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \overline{\mathbf{X}})(\mathbf{X}_i - \overline{\mathbf{X}})'$$

For details of the S matrix we have that

$$S_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij} - \overline{X}_j)^2, \quad 1 \le j$$

$$S_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij} - \overline{X}_j)(X_{ik} - \overline{X}_k), \quad 1 \le j,k \le p .$$

It's known that

$$E(\overline{\mathbf{X}}) = \boldsymbol{\mu}, \quad E(\mathbf{S}) = \boldsymbol{\Sigma}, \quad \overline{\mathbf{X}} \sim N_p(\boldsymbol{\mu},(1/n)\boldsymbol{\Sigma}), \quad \sqrt{n}(\overline{\mathbf{X}} - \boldsymbol{\mu}) \sim N_p(\mathbf{0},\boldsymbol{\Sigma})$$

$$n(\overline{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2, \quad (n-1)\mathbf{S} \sim W_p(n-1,\boldsymbol{\Sigma}) .$$

https://doi.org/10.26219/heal.aueb.6243

3.2.2 Control Charts for the vector of mean values

Sample size n> 1

Suppose that the qualitative characteristic X '= [X_1, X_2, ... X_p] follows the p-dimensional normal distribution with known parameters μ and Σ, and suppose that we have samples from the qualitative characteristic n>1 sized. The general form of the sample k (k ≥1) are the following

$$k\ sample: \left\{ \begin{bmatrix} X_{11k} \\ X_{12k} \\ \vdots \\ X_{1pk} \end{bmatrix}, \begin{bmatrix} X_{21k} \\ X_{22k} \\ \vdots \\ X_{2pk} \end{bmatrix}, \ldots, \begin{bmatrix} X_{n1k} \\ X_{n2k} \\ \vdots \\ X_{npk} \end{bmatrix} \right\}$$

For each sample, we define the quantity (sample mean)

$$\overline{\mathbf{X}}'_k = [\overline{X}_{1k}, \overline{X}_{2k}, \ldots, \overline{X}_{pk}], \quad k \geq 1$$

wherein $\overline{X}_{jk} = \frac{1}{n}\sum_{i=1}^{n} X_{ijk}, \quad 1 \leq j \leq p$.

We define $D_k^2 = n(\overline{\mathbf{X}}_k - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{X}}_k - \boldsymbol{\mu}) \sim \chi_p^2$.

The Statistical D^2_k function represents the weighted distance (Mahalanobis distance) point $\bar{X}_k$ from the value μ. Large values of statistical D^2_k function indicate that the distance of points $\bar{X}_k$ and μ is large thereby we adopt the view that the process is out of statistical control due to displacement of process's mean. In contrast, small or zero values statistical D^2_k function indicate that the process is in statistical control.

Therefore, a control chart for the monitoring the mean μ of the process can be based on a graph wherein the amount shown is the statistical D^2_k. The control limits are given below:

X^2 control chart **Phase II control limits**
$UCL = X^2_{p;a}$ $LCL = 0$

https://doi.org/10.26219/heal.aueb.6243

The upper control limit UCL of X^2 control chart is calculated to ensure that the type I error is equal to a, i.e. P (D^2_k> UCL) = a

Thus, the chart's in control ARL is equal to

ARLin = 1 / a

Assume now that the vector of the means of process values is shifted from the in control value μ to the out of control value $\mu^* = \mu + \delta$, $\delta \neq 0$ (i.e., assume that E (X) = μ + δ), while variance - covariance matrix Σ, has remained the same. In this case the distribution of the statistical function D^2_k is the non - central X^2 distribution with p degrees of freedom and non – centrality parameter

$\lambda = \lambda(\mu^*) = n(\mu^* - \mu)'\Sigma^{-1}(\mu^* - \mu) = n\delta'\Sigma^{-1}\delta$.

Thus, out of control ARL is calculated by the formula $ARL_{out} = 1 / (1-\beta)$ wherein b denotes the type II error, i.e.

$1-\beta = 1 - P(D^2_k < X^2 p,a \mid \lambda(\mu^*)) = P(D^2_k > X^2 p,a \mid \lambda(\mu^*))$

A calculation formula for the type II error is the following:

$$\beta = e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{(\lambda/2)^j}{j!2^{(p/2)+j}\Gamma((p/2)+j)}\int_0^{\chi^2_{p;a}} y^{(p/2)+j-1}e^{-y/2}dy .$$

In the development Phase I of control charts, the parameters μ and Σ of qualitative characteristic X '= [X_1, X_2, ... X_p] are unknown and must be estimated. Suppose we have m samples of the qualitative characteristic, n>1 sized. We define the quantity (total sample mean) $\overline{\overline{\mathbf{X}}}' = [\overline{\overline{X}}_1, \overline{\overline{X}}_2, \ldots, \overline{\overline{X}}_p]$

Wherein

$$\overline{\overline{X}}_j = \frac{1}{m}\sum_{k=1}^{m}\overline{X}_{jk} = \frac{1}{mn}\sum_{k=1}^{m}\sum_{i=1}^{n}X_{ijk}$$

and the matrix (cumulative (pooled) variance - covariance matrix) (Montgomery (2005))

$$\overline{\mathbf{S}} = \begin{bmatrix} \overline{S}_1^2 & \overline{S}_{12} & \cdots & \overline{S}_{1p} \\ \overline{S}_{21} & \overline{S}_2^2 & \cdots & \overline{S}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{S}_{p1} & \overline{S}_{p2} & \cdots & \overline{S}_p^2 \end{bmatrix}$$

 https://doi.org/10.26219/heal.aueb.6243

wherein

$$\bar{S}_j^2 = \frac{1}{m}\sum_{k=1}^{m} S_{jk}^2, \quad j = 1,2,\ldots,p$$

$$S_{jk}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ijk} - \bar{X}_{jk})^2, \quad j = 1,2,\ldots,p, \quad k = 1,2,\ldots,m$$

$$\bar{S}_{jh} = \frac{1}{m}\sum_{k=1}^{m} S_{jhk}, \quad 1 \le j \ne h \le p$$

$$S_{jhk} = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ijk} - \bar{X}_{jk})(X_{ihk} - \bar{X}_{hk}), \quad 1 \le j \ne h \le p, \quad k = 1,2,\ldots,m\,.$$

The quantities X and S are estimates of the parameters μ and Σ the qualitative characteristic. Substituting the unknown μ and Σ with $\bar{\bar{X}}$ and $\bar{S}$ corresponding to the statistic D^2_k shows the following statistic

$$T_k^2 = n(\bar{\mathbf{X}}_k - \bar{\bar{\mathbf{X}}})' \bar{\mathbf{S}}^{-1} (\bar{\mathbf{X}}_k - \bar{\bar{\mathbf{X}}}), \quad k = 1,2,\ldots,m$$

whereas

$$\frac{mn-m-p+1}{p(m-1)(n-1)} T_k^2 \sim F_{p,mn-m-p+1}$$

the control chart for the monitoring of the mean of the process described in Phase I as follows

Hotelling T^2 control chart **Phase I control limits**
$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{p,mn-m-p+1;a}$ $LCL = 0$

For Phase II control charts wherein the parameters μ and Σ qualitative characteristic $X' = [X_1, X_2, \ldots X_p]$ have been evaluated in Phase I by the quantities $\bar{\bar{X}}$ and $\bar{S}$the statistical function is used

$$T_f^2 = n(\bar{\mathbf{X}}_f - \bar{\bar{\mathbf{X}}})' \bar{\mathbf{S}}^{-1} (\bar{\mathbf{X}}_f - \bar{\bar{\mathbf{X}}})$$

Wherein $\bar{X}_f$ is the sample mean of a future sample.

Whereas

https://doi.org/10.26219/heal.aueb.6243

$$\frac{mn-m-p+1}{p(m+1)(n-1)}T_f^2 \sim F_{p,mn-m-p+1}$$

the control chart for the monitoring mean of the process described in Phase II as follows (Montgomery (2005))

Hotelling T^2 control chart **Phase II control limits**
$UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{p,mn-m-p+1;a}$ $LCL = 0$

3.2.3 Individual Observations

Suppose that the qualitative characteristic X '= [X_1, X_2, ... X_p] follows the p-dimensional normal distribution with known parameters μ and Σ, and suppose that we have samples from the qualitative characteristic, size n = 1 each. The general form of the k-th sample (k ≥1) is the following (Montgomery, (2005))

$$\mathbf{X}_k = \begin{bmatrix} X_{k1} \\ X_{k2} \\ \vdots \\ X_{kp} \end{bmatrix}, \quad k \geq 1.$$

The statistic

$$D_k^2 = (\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu}) \sim \chi_p^2.$$

Thus, a control chart for the monitoring the mean of the process is that which the amount shown is the statistical D^2_k and chart control limits are given below. In the development Phase I control charts parameters μ and Σ of the qualitative characteristic X '= [X_1, X_2, ... X_p] is unknown and must be estimated. Suppose we have m samples of the qualitative characteristic, sized n = 1. We define the quantity (sample mean)

 https://doi.org/10.26219/heal.aueb.6243

$$\overline{\mathbf{X}} = \begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_p \end{bmatrix} = \frac{1}{m}\sum_{k=1}^{m} \mathbf{X}_k$$

and matrix (Sample variance- covariance matrix)

$$\mathbf{S} = \begin{bmatrix} S_1^2 & S_{12} & \cdots & S_{1p} \\ S_{21} & S_2^2 & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_p^2 \end{bmatrix} = \frac{1}{m-1}\sum_{k=1}^{m} (\mathbf{X}_k - \overline{\mathbf{X}})(\mathbf{X}_k - \overline{\mathbf{X}})'.$$

Quantities$\bar{X}$ and S are estimates of the parameters μ and Σ of the qualitative characteristic. Substituting the unknown μ and Σ with $\bar{X}$ and S corresponding, the statistic D^2_k shows the following statistic

$$T_k^2 = (\mathbf{X}_k - \overline{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X}_k - \overline{\mathbf{X}}), \quad k = 1,2,\ldots,m$$

Whereas

$$\frac{m}{(m-1)^2} T_k^2 \sim B_{p/2,(m-p-1)/2}$$

the control chart for the monitoring of the mean of the process in Phase II is described below

Hotelling T^2 control chart **Phase I control limits**
$UCL = \frac{(m-1)^2}{m} B_{p/2,(m-p-1)/2;a}$ $LCL = 0$

For Phase II control charts where the parameters μ and Σ qualitative characteristic X '= [X_1, X_2, ... X_p] have been evaluated in Phase I of the quantities $\bar{X}$ and S,using the statistic

$$T_f^2 = (\mathbf{X}_f - \overline{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X}_f - \overline{\mathbf{X}})$$

where X_f is a future sample. Given that:

$$\frac{m(m-p)}{p(m+1)(n-1)}T_f^2 \sim F_{p,m-p}$$

the control chart for the monitoring of the mean of the process in Phase II is described below

Hotelling T^2 control chart **Phase II control limits**
$UCL = \frac{p(m+1)(n-1)}{m(m-p)}F_{p,m-p;a}$ $LCL = 0$

3.2.4 Control Charts for the Variance

Suppose that the qualitative characteristic X'= [X_1, X_2, ... X_p] follows the p- dimensional normal distribution with known parameters μ and Σ, and suppose that we have samples from the qualitative characteristic, n> 1 sized.

To test the hypothesis

$H_0 : \Sigma = \Sigma_0$ Vs $H_1 : \Sigma \neq \Sigma_0$

Alt (1985) proposed the statistic

$$W_k = -pn + pn\ln(n) - n\ln\left(\frac{|\mathbf{A}_k|}{\mathbf{\Sigma}_0^{-1}}\right) + tr(\mathbf{\Sigma}_0^{-1}\mathbf{A}_k) \sim \chi^2_{p(p+1)/2}$$

where A_k = (n-1) S_k, S_k is the sampling variance - covariance matrix, for the k-th sample, and tr(X) denotes the trace of a matrix. In the control chart for monitoring the variance of the process (the mean μ of the process remains constant) the amount shown is the statistic W_k.

Here are the chart control limits:

W control chart **Phase II control limits**
$UCL = X^2_{p(p+1)/2;a}$ $LCL = 0$

 https://doi.org/10.26219/heal.aueb.6243

Another approach for monitoring the variance of a process is based on sample generalized variance defined as |S|, where S a pxp Sample covariance matrix. Specifically, if the qualitative characteristic is of the form X'= [X_1, X_2] then, in accordance with Alt (1985), the statistic

$$\frac{[2(n-1)\,|\mathbf{S}|^{1/2}]}{|\boldsymbol{\Sigma}_0|^{1/2}} \sim \chi^2_{2n-4}$$

So

$$P\left(\frac{|\boldsymbol{\Sigma}_0|^{1/2}\,\chi^2_{2n-4;1-(a/2)}}{2(n-1)} \leq |\mathbf{S}|^{1/2} \leq \frac{|\boldsymbol{\Sigma}_0|^{1/2}\,\chi^2_{2n-4;a/2}}{2(n-1)}]\right) = 1-a.$$

Thus, a control chart for monitoring the process of variance is that the amount shown is the statistic | S_k | and the chart control limits are given below

|S| control chart **Phase II control limits**
$UCL = \frac{\lvert\Sigma_0\rvert (X^2_{2n-4;a/2})^2}{4(n-1)^2}$ $LCL = \frac{\lvert\Sigma_0\rvert (X^2_{2n-4;1-a/2})^2}{4(n-1)^2}$

A different S control chart for monitoring the process of variance is based on the observation that the interval $E[|S|] \pm 3\sqrt{V[|S|]}$ is distributed across the probability of the statistical function |S|

wherein

$$E[|\mathbf{S}|] = b_1 |\boldsymbol{\Sigma}_0|$$

$$V[|S|] = b_2 |\boldsymbol{\Sigma}_0|^2$$

$$b_1 = (n-1)^{-p} \prod_{i=1}^{p} (n-i)$$

$$b_2 = (n-1)^{-2p} \prod_{i=1}^{p} (n-i) \times \left[\prod_{j=1}^{p} (n-j+2) - \prod_{j=1}^{p} (n-j)\right]$$

https://doi.org/10.26219/heal.aueb.6243

The control chart limits are listed below:

|S| control chart Phase II control limits
$UCL = \|\Sigma_0\| b_1 + \|\Sigma_0\| 3 b_2^{1/2}$ $LCL = \|\Sigma_0\| b_1 - \|\Sigma_0\| 3 b_2^{1/2}$

If the lower threshold is negative then we set LCL = 0.

Another approach is based on the square root of the matrix of sample variance - covariance $|S|^{1/2}$ with two versions.

According to the first embodiment, two variables by Alt and Smith (1988) gave the following control limits:

$\|S\|^{1/2}$ control chart Phase II control limits
$UCL = \frac{\|\Sigma_0\|^{1/2} X^2_{2n-4;1-a/2}}{2(n-1)}$ $LCL = \frac{\|\Sigma_0\|^{1/2} (X^2_{2n-4;a/2})}{2(n-1)}$

For more features (variables) are other options (Anderson (1985)).

The second version again by Alt and Smith (1988) is based on relations

$$E[|S|^{1/2}] \pm 3\sqrt{V\left[|S|^{1/2}\right]}$$

$$E[|S|^{1/2}] = |\Sigma_0|^{1/2} \left(2/(n-1)\right)^{p/2} \Gamma(n/2)/\Gamma\left((n-p)/2\right) = |\Sigma_0|^{1/2} b_3$$

$$V\left[|S|^{1/2}\right] = |\Sigma_0| \left(b_1 - b_3^2\right)$$

Here are the control limits

$\|S\|^{1/2}$ **control chart** **Phase II control limits**
$UCL = \|\Sigma_0\|^{1/2} b_3 + \|\Sigma_0\|^{1/2} 3\sqrt{b_1 - b_3^2}$ $LCL = \|\Sigma_0\|^{1/2} b_3 - \|\Sigma_0\|^{1/2} 3\sqrt{b_1 - b_3^2}$

If the lower limit is negative we set LCL = 0.

For Phase I control chart, wherein the matrix | S_0 | is unknown, Alt (1985) used as unbiased estimator the quantity | S | + B_1 where

$$|\bar{S}| = \frac{1}{m}\sum |S_i|$$

and S_i is the sampling variance - covariance matrix of the sample i where i = 1,2,.., m. The new resulting limits follows

$\|\bar{S}\|$ **control chart** **Phase I control limits**
$UCL = (\|\bar{S}\| + b_1)b_1 + (\|\bar{S}\| + b_1)3b_2^{1/2}$ $LCL = (\|\bar{S}\| + b_1)b_1 - (\|\bar{S}\| + b_1)3b_2^{1/2}$

https://doi.org/10.26219/heal.aueb.6243

3.3 Multivariable CUSUM Charts

3.3.1 CUSUM control with sequential likelihood ratio

Healy (1987) examined the CUSUM charts using sequence of sequential probability tests ratio(SPRT) and applied them to multivariate normal distribution in order to obtain the MCUSUM charts.

Suppose we watch a sequence of random variables X_1, X_2, ... following p-dimensional normal distribution. Suppose the first m-1 observations X_1, X_2, ... X_m-1 state in control distribution Np (μ_0, Σ) and subsequent X_m, X_m+ 1 are out of control distribution Np (μ_1, Σ). The problem is to identify the point m when the distribution of observations changes position timely. According to the methodology of sequential likelihood ratio tests have

$$f_1(\mathbf{X}_t) / f_0(\mathbf{X}_t) = \frac{\exp\left(-0.5(\mathbf{X}_t - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_1)\right)}{\exp\left(-0.5(\mathbf{X}_t - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)\right)}$$

where f is the probability density of the multivariate normal distributions. Using the log function we have

$$log(f_1(\mathbf{X}_t) : f_0(\mathbf{X}_t)) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} \mathbf{X}_t - 0.5(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) \quad .$$

Thus, according to Healy (1987) CUSUM chart is the following wherein the control chart points from the statistic

$$S_t = \max[(S_{t-1} + \mathbf{a}' \mathbf{X}_t - K), 0]$$

Wherein

$$\mathbf{a}' = \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1}}{\left[(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right]^{1/2}}$$

and

$$K = \frac{1}{2} \frac{(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}{\left[(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right]^{1/2}}$$

The variable $a'X_t$ follows dimensional normal distribution with variance one. The statistic St can be written in the form

$$S_t = \max\left[\left(S_{t-1} + \mathbf{a}'(\mathbf{X}_t - \boldsymbol{\mu}_0) - 0.5\lambda(\boldsymbol{\mu}_1)\right), 0\right]$$

https://doi.org/10.26219/heal.aueb.6243

where $\lambda^2(\mu_1) = (\mu_1 - \mu_0)\ \Sigma^{-1}{}_0(\mu_1 - \mu_0)$

The MCUSUM will give out of control signal when St ≥ H, wherein H is calculated in Pignattielo's and Rugner's simulation. If the out of control vector M1 cannot be determined then it is recommended Crosier's methodology (1988). This approach examines the vector of means of a multivariate normal distribution. Below we describe the event that monitoring variance - covariance matrix of a multivariate normal distribution is necessary.

Assume that Xt is a point in time t, from the p-dimensional normal distribution with in control variance - covariance matrix Σ_0, and out of control matrix $\Sigma_1 = C\ \Sigma_0$ and common (in and out of control) vector of means μ_0. Then

$$f_1(\mathbf{X}_t) / f_0(\mathbf{X}_t) = \frac{(2\pi)^{-p/2} |C\Sigma_0|^{-1/2} \exp\left(-0.5(\mathbf{X}_t - \boldsymbol{\mu}_0)' (C\Sigma_0)^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)\right)}{(2\pi)^{-p/2} |\Sigma_0|^{-1/2} \exp\left(-0.5(\mathbf{X}_t - \boldsymbol{\mu}_0)' \Sigma_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)\right)}$$

and we take using the log function

$$log(f_1(\mathbf{X}_t) / f_0(\mathbf{X}_t)) = -\frac{p}{2}\log C + 0.5(\mathbf{X}_t - \boldsymbol{\mu}_0)' \Sigma_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)\left(1 - \frac{1}{C}\right).$$

Eventually we have the statistic

$$S_t = \max\left[\left(S_{t-1} + (\mathbf{X}_t - \boldsymbol{\mu}_0)' \Sigma_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0) - K\right), 0\right]$$

wherein

$$K = p \log C \frac{C}{C-1}.$$

If

$$T_t^2 = (\mathbf{X}_t - \boldsymbol{\mu}_0)' \Sigma_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)$$

then the statistic shows

$$S_t = \max\left[\left(S_{t-1} + T_t^2 - K\right), 0\right].$$

The MCUSUM control chart will give signal for an out of control process when $S_t \geq H$. q

 https://doi.org/10.26219/heal.aueb.6243

3.3.2 The Crosier approach

Crosier (1988) proposed two new multivariate MCUSUM approaches. The first one is based on the square root of the statistical function of Hotelling-T^2 while the second one is obtained by suitable modification of dimensional CUSUM charts.

Assume that an observation X_t is at time t from the p-dimensional multivariate normal distribution with known variance - covariance matrix Σ0 and μ0.

The first Crosier approach uses statistic

$$S_t = \max\left[\left(S_{t-1} + \sqrt{(\mathbf{X}_t - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{X}_t - \boldsymbol{\mu}_0)} - K\right), 0\right]$$

If

$$T_t = \sqrt{(\mathbf{X}_t - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{X}_t - \boldsymbol{\mu}_0)}$$

Then $S_t = \max[S_{t-1} + T_t - K), 0]$

where $S_0 \geq 0$ and $K \geq 0$.

This process will enable us to mark out of control process when St ≥ H.

The second approach proposed by Crosier has better ARL. It uses the statistic

$$Y_t = \sqrt{\mathbf{S}_t'\boldsymbol{\Sigma}_0^{-1}\mathbf{S}_t}$$

Where $S_t = 0, if\ C_t \leq K$

$$\mathbf{S}_t = (S_{t-1} + \mathbf{X}_t - \boldsymbol{\mu}_0)\left(1 - \frac{K}{C_t}\right),\ \alpha\nu\ C_t > K$$

And $C_t = \sqrt{(S_{t-1} + X_t - \mu_0)'\Sigma_0^{-1}(S_{t-1} + X_t - \mu_0)}$

out of control signal we get when Yt> H. The value of H is selected to have predetermined in control ARL with simulation's help. Crosier (1988) showed that the distribution of Yt and ARL depend on the vector of means and variance - covariance matrix only via the parameter of the non-centrality. Also reported that $K = \lambda(\mu_1)/2$ where $\lambda(\mu_1)$ is the square root of the non - centrality parameter that minimizes out of control ARL.

Comparing the MCUSUM charts with multivariate Shewhart charts we conclude that they excel in giving a signal for an out of control process for small to mean shifts faster.

https://doi.org/10.26219/heal.aueb.6243

3.3.3 The J. Pignatiello and G. Runger approach

Pignatiello and Runger (1990) introduced two new multivariate CUSUM approaches, MCUSUM # 1 (MC#1) and MCUSUM # 2 (MC#2).

The MC#1 is the plan that has the best ARL and is as follows:

Suppose, Xt an observation at time t, from the p-dimensional normal distribution with known variance - covariance matrix Σ0 and a known mean μ0.

Let $S_t = \sum_{i=t-n_t+1}^{t}(X_t - \mu_0)$

Wherein

$$n_t = \begin{Bmatrix} n_{t-1}+1, & if\ MC\#1_{t-1} > 0 \\ 1, & else \end{Bmatrix}$$

A multivariate control chart can be constructed by setting $C_t = \sqrt{\mathbf{S}_t'\boldsymbol{\Sigma}_0^{-1}\mathbf{S}_t}$ and $MC\#1_t = \max\{C_t - kn_t, 0\}$

wherein $MC\#1_0 = 0$ and k are selected to be half the distance between μ1 and μ0, 0.5λ2 (μ_1), wherein μ1 is an out of control mean.

The MC#1 chart works designing the MC # 1_t points with upper control limit H. When a point MC # 1t exceed the then our process is out of control.

Pignatiello and Runger (1990) demonstrated that ARL is dependent on the square root of the non-centrality parameter and the value H is calculated according to the methodology of Markov chains. They also concluded that all multivariate CUSUM charts behave better than those of multivariate Shewhart except where we have large values for the parameter of non - centrality.

The MC # 2 is the second approach of Pignatiello and Runger (1990). At least,

$$X_t^2 = (\mathbf{X}_t - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0)$$

A onesided test resulting from the relation

$$MC\#2_t = \max\{0, MC\#2_{t-1} + X_t^2 - k\}$$

wherein MC # 20 = 0 and k being selected to be half the distance between μ_1 and μ_0 + p i.e.

$$k = p + 0.5\,\lambda^2(\mu_1)$$

 https://doi.org/10.26219/heal.aueb.6243

The MC#2 chart works by drawing the points t MC # 2 with an upper control limit H. When a point t MC # 2 exceeds H then our process is out of control. The value of H is calculated by the methodology of Marcov chains simulation.

3.4 Multivariable Charts EWMA

3.4.1 Control Charts MEWMA

The multivariate EWMA charts presented in this section have been proposed by Lowry et al. (1992). Let X_t follows the p- dimensional normal distribution with known variance - covariance matrix Σ0 and a known μ0.

The multivariate EWMA control chart uses statistical

$$\mathbf{z}_t = \mathbf{R}(\mathbf{X}_t - \boldsymbol{\mu}_0) + (\mathbf{I} - \mathbf{R}) z_{t-1}$$

or equivalent

$$\mathbf{z}_t = \sum_{j-1}^{t} \mathbf{R}(\mathbf{I} - \mathbf{R})^{t-j} (\mathbf{X}_j - \boldsymbol{\mu}_0)$$

wherein t = 1,2,3, ... and R are pxp matrix designated

$$\mathbf{R} = \begin{bmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & r_p \end{bmatrix}$$

where $0 \le r_k \le 1$ for $k = 1,2,3,..p.$

If there is no reason to weigh different weights, the p qualitative characteristics we monitor, then we set r1 = r2 = ... = rp = r. The initial value z0 is usually equal to the in control mean. Obviously, if R = 1 then the multivariate EWMA chart is equivalent to the T^2 control chart (or x^2 control chart where we know the vector of the means and the variance - covariance matrix).

The multivariate EWMA chart gives an off control signal if

$$\mathbf{z}_t' \sum_{z_t}^{-1} \mathbf{z}_t > h$$

where Σz_t is the variance - covariance matrix of z_t.

The h value is determined by simulation so as to achieve a predetermined in control ARL. Variance - covariance matrix Σzt of zt is calculated by the following formula $\boldsymbol{\Sigma}_{z_t} = \sum_{j=1}^{t} \left[\mathbf{R}(\mathbf{I} - \mathbf{R})^{t-j} \boldsymbol{\Sigma}_0 (\mathbf{I} - \mathbf{R})^{t-j} \mathbf{R} \right]$ if r1 = r2 = ... = rp = r

https://doi.org/10.26219/heal.aueb.6243

then $\Sigma_{z_t} = \frac{r\left[1-(1-r)^{2t}\right]}{2-r}\Sigma_0$

and t → + ∞ we have that $\Sigma_{z_t} = \frac{r}{2-r}\Sigma_0$

3.4.2 ARL for MEWMA

The ARL in multivariate analysis of EWMA charts depends solely from μ0, Σ0 and the value of the non-centrality parameter. For calculating the ARL have been proposed two methodologies by Rigdon as follows:

- Integral Equation for the in Control ARL

In this section, we give the complete equation Rigdon (1995a), provided that the mean of the in-process control is 0, and the variance - covariance matrix is unitary I.

Suppose, Xi, i = 1,2, .., remarks following distribution p N (μ, Σ). If the multivariate EWMA control chart with parameters r and h is applied to Xi, then in control ARL is the same as in the case of observation with mean 0 and variance-covariance matrix I. The in-control ARL of a multivariate EWMA chart that is applied to a process with mean vector **0** and variance - covariance matrix I, depends on the initial value z_0 only through the quantity δ = z'0z0.

Suppose we have a multivariate EWMA chart with r parameter, applied to a p-dimensional process with mean 0 and variance - covariance matrix I. Let, L (δ / h) the ARL, corresponding to the quantities δ and h. Then the function L satisfies the integral equation

$$L(\delta / h) = 1 + \int_0^{hr/(2-r)} L(y / h) \times f\left(y \mid z_0' z_0 = \delta\right) dy$$

wherein f ($y|z'_0z_0$=δ) is the probability density function of the non-central x2 distribution with p degrees of freedom and non-centrality parameter {$[(1- r)/r]^2\delta$}.

https://doi.org/10.26219/heal.aueb.6243

- Dual Integral Equation for out of control ARL

In this section we give the double integral equation of Rigdon (1995b), for calculating the out of control ARL under the condition that in the control mean is 0 and that in control variance - covariance matrix is I.
We assume that the production process following the N_P (0, I) distribution. If the mean shifts from 0 to μ, then the ARL depends on z_0 through only the quantities $a = (z'_0 z_0)^2$ and b = μ'z0.
Suppose that X_i, = 1,2, .., i are observations from the N_p (m, T) distribution, where μ = 0 when the process is in control. Suppose $\mu'\mu = \delta \neq 0$.
Then for a and b the ARL, L (a, b), the multivariate EWMA chart with parameters r and h satisfy the integral equation

$$L(\alpha,\beta) = 1 + \iint_R L(u,v)\left(\frac{1}{\sqrt{2\pi\delta r}}\right)\exp\left[-\frac{1}{2\delta r^2}(v - r\delta - (1-r)\beta)\right]^2\left(\frac{1}{r^2}\right)h(./v,r)dudv$$

wherein h (./ v, r)

$$h(./v,r) = h\left(\frac{(u - v^2/\delta)}{r^2}/p-1, \left(\frac{1-r}{r}\right)^2(\alpha - \beta^2/\delta)\right)$$

and h (./ v, l) is the probability density function of the non-central x^2 distribution with v degrees of freedom and non-centrality parameter λ. Furthermore, R is the region where $v^2/\delta < u < rh/(2 - r)$.

https://doi.org/10.26219/heal.aueb.6243

CHAPTER 4

Profile Monitoring (Part A ’)

4.1 Introduction

In many practical situations, the quality of a process or a product can be classified and summarized in a functional relation between a dependent variable and one or more explanatory (independent) variables. A large study has recently been in the area of the process of statistical monitoring. Moreover, the range of applications has been greatly expanded. The Profile Monitoring is used to understand and control the stability of this relation over time. The Profile Monitoring is a relatively new technique in statistical quality control and is best used when data processing follows a profile (or curve) in each period. The methods we examine have the following three characteristics: the collection of time data, rapid detection of specific process changes due to respective causes, usually represented by the changes to the parameter(s) of a probability distribution representing the variation, common cause of the process and the determination of performance control, measured by the false alarm rate or other metric

The Profile Monitoring includes two phases. In Phase I, the statistician collects a sample of the time series data from the process of interest. These Phase I data are used to gain an understanding of the process. You must check for unusual or extreme results. It must also assess the stability of the process to select a suitable in-control model to assess the parameter(s) of this model, and determine the design parameters of the monitoring methodology to be used in Phase II. The monitoring process is then implemented with data collected sequentially over time in Phase II, in order to detect changes in the process by the assumed-control model. In many applications, it is possible to regulate processes using forward data or feedback control. To do this, there must be an adjustment variable with some information available on impact. For information and perspectives on this issue, we refer the reader to del Castillo (2002), Box et al. (2009) and Box and Narasimhan (2010). It is important to raise the role of monitoring the process in perspective. We believe that the monitoring process is important to understand the

https://doi.org/10.26219/heal.aueb.6243

variation in a process and to assess the current situation. The stability should be evaluated prior to any study of the process-ability. In many cases, monitoring of the process before and after a change of the process required to evaluate the effect of the change process. Monitoring process alone, however, is usually not sufficient to significantly improve the process. We strongly support the use of designed experiments in efforts to achieve innovations in process and product performance.

We also strongly support the use of Six Sigma define, measure, analyze, improve and control the process (DMAIC). For a description of this approach, see Montgomery and Woodall (2008). Under the DMAIC process, the profile monitoring plays a big role. It is useful in measuring the phase estimate current performance and monitoring of performance measurement systems. It is also useful in the control phase monitoring input variables, so that the performance of the phase to be maintained over time. We also support variations in reduction approaches of Steiner and MacKay (2005), which include the collection of process data over time.

It is extremely important during Phase I of the control chart to determine which of the data points are similar and which are somewhat remote. This ensures that the implementation of Phase II will be sufficient for real-time monitoring. Due to technological progress, it becomes increasingly more common to acquire profiles (a series of data points forming a curve) in each period, representing the state of process's quality. Therefore, recent research has focused on how to determine which profile is remotely i.e. outside the boundaries during Phase I.

At each sampling phase one observes a collection of data points that can be represented by a curve (or profile). In some calibration applications, the profile can be adequately represented by a simple linear model, while other applications require more complex models.

The sampling schemes using information from a random sample of size n units selected from a batch of size N units ($n<N$) for a decision on acceptance or rejection of the lot called single sampling plans. The rule by which we decide whether the lot is accepted based on the number D_n of defective units (products) found in the random sample.

https://doi.org/10.26219/heal.aueb.6243

If the number D_n does not exceed a number c ($D_n \leq c$), which is called the acceptance number (acceptance number), then the lot is accepted. Otherwise the batch is rejected. A single sampling plan with the given values for the quantities of N, n, c will be called (simple) sampling plan (N;n;c) The quantities N, n, c satisfy the relation $0 \leq c \leq n \leq N$.

The probability of accepting a batch ($P(D_n \leq c)$) depends on the (unknown) number M ($0 \leq M \leq N$) of defective units contained in the specific lot or equivalent by the percentage (or ratio, or fraction) p=M/N of defective units containing the batch. A large (small) value of p, which indicates the low (high) batch quality, should logically be associated with low (high) probability of lot acceptance. Therefore, if we denote by L (p), the probability of acceptance of a batch, ie

$$L(p) = P(D_n \leq c \mid p) = \sum_{d=0}^{c} P(D_n = d \mid p), \quad p = 0, \frac{1}{N}, \frac{2}{N}, \ldots, 1$$

we expect that the function L (p) should be a decreasing function with respect to p and the exact form depending on the distribution of the random variable D_n. The graph of the function L (p) is called operating-characteristic curve and contains all the information we need to analyze of a single sampling plan.

The ideal characteristic curve is the graph below

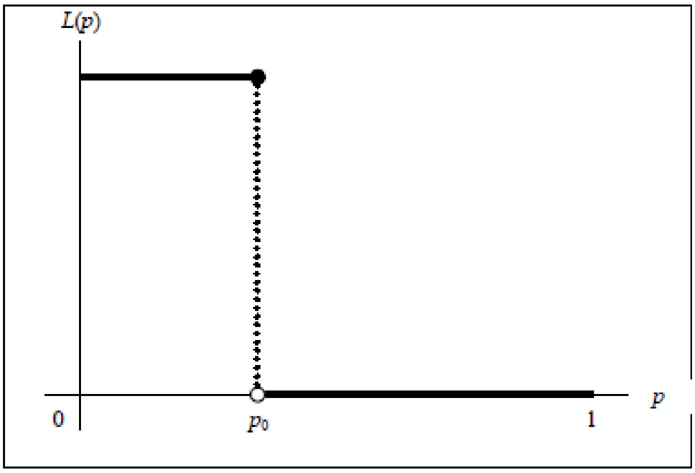

Figure 4.1: Ideal characteristic curve

According to the above characteristic, lots with quality less than or equal to p0 (high quality batches) have acceptance probability 1 while lots with quality higher than p0 (low quality batches) have acceptance probability of 0. Of course,

 https://doi.org/10.26219/heal.aueb.6243

the ideal characteristic curve does not meet in practice. Instead we find a smooth descending curve with the following form (we consider that p can take any value in the interval [0,1])

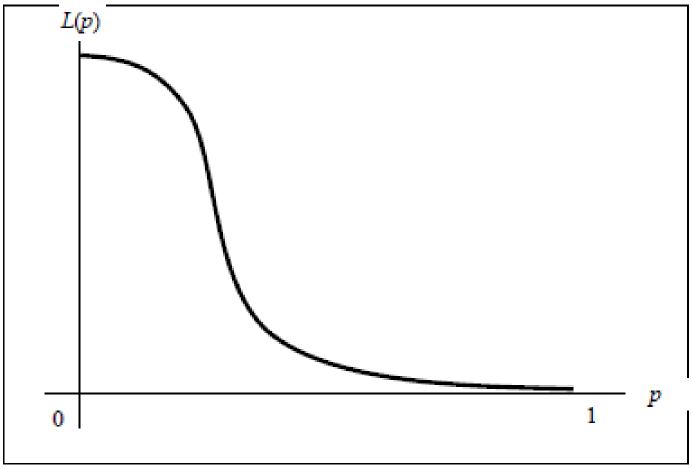

Figure 4.2: Caracteristic curve we meet in practice

For this curve we want to have a sharp slope (steepness) so that it is as close as possible to a greater extent the ideal characteristic curve. The steeper the slope the more wins a separating force of the characteristic curve, i.e., increases its ability to separate batches by high low grade batches.

For the design of the curve (acceptance probability) L (p) the knowledge of the distribution of the random variable D_n is required.

Sometimes the relation between a dependent variable and a number of explanatory variables, called profile will be monitored over time instead of the quality characteristic itself. Depending on the type of this relation, the profiles are classified into several categories such as simple linear profiles, multiple linear profile, polynomial profile, etc. In recent years, many investigations have been made in the profile monitoring. Mestek et al., Stover and Brill, and Amiri et al. Mestek show some embodiments of the profiles. Kang and Albin, Kim et al. and Mahmoud et al. have suggested some methods for monitoring simple linear profile (simple linear profiles). Zhou et al. and Mahmoud suggest some methods for monitoring multiple linear profile (multiple linear profiles). Monitoring multivariate profile (multivariate profiles) examined by authors such as Noorossana et

https://doi.org/10.26219/heal.aueb.6243

al. and Eyvazian et al. Authors including Vaghefi et al., and Williams et al. propose methods for monitoring non-linear profile.

In an increasing number of industrial applications, the quality of a process or product is best described by a function, called "profiles". In these embodiments, a response variable associated with one or more explanatory variables. In these cases, the profile changes depending on time are of interest. Various types of models are used to represent profiles, including simple linear regression, nonlinear regression, multiple regression, non-parametric regression, mixed models, and wavelet models.

An example of a profile is shown in Figure below. This figure shows the force against the position of the tool to a broaching process in a production application. Broaching is a machining process using a toothed tool, called pin, for material removal. Each embodiment of the broaching process yields a profile and changes the shape of the profile over time may indicate wear or other problems with the tool.

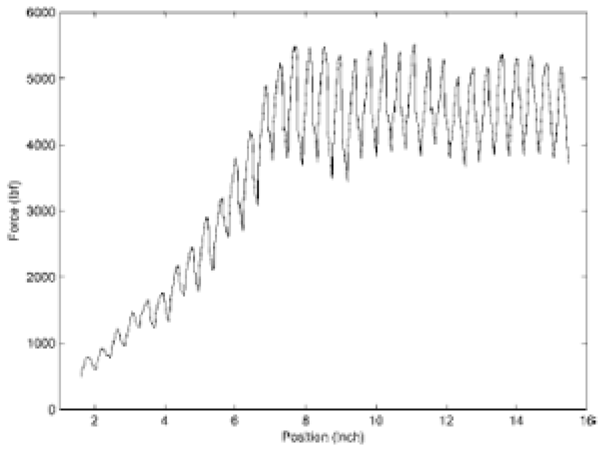

Figure 4.3: Example of Profile Data from a Broaching Process

Source: Jaime A. Camelio, Broaching Process

If the profile is represented by a parametric model, then usually a model parameters are monitored with separate control charts, if the parameter estimators are independent, or the parameters are monitored together with a multivariate control chart. In very few cases, the profile shape cannot be adequately represented by a parametric model, so a non-parametric approach is needed. See Qiu et al. (2010)

https://doi.org/10.26219/heal.aueb.6243

and Zou et al. (2008). Sometimes it is useful and easier to monitor certain characteristics of the profile, such as the maximum value. The profile monitoring now includes monitoring of two-dimensional shapes and three-dimensional surfaces. Effective monitoring in practical applications requires expertise, subject and an emphasis on root of the cause, analysis. Even simpler manufacturing processes consist of several production steps. In many industrial applications, one must bear in mind the multiple steps in the process and model the impact of a process step to the next. If one monitors only the result of a secondary process without taking into consideration the input, then it may result in erroneous conclusions. The use of multiple data levels allows one to understand the transmission variation through a process referred to as an analysis of stream-of-variation.

The profile monitoring is the use of control charts for cases in which the quality of a process or product can be characterized by a functional relation between a response variable and one or more explanatory variables. The image monitoring could be viewed as a natural extension of profile monitoring methods in cases where the explanatory variables indicate the position of the measuring volume within the image. In general, it is worth noting that the letters written to the image monitoring can be much more difficult to understand from the documents, relating to the monitoring profile involving only one explanatory variable. One tends to have much more data to video applications, more sophisticated methods of analysis, and more broad potential for monitoring purposes.

4.2 Simple linear profiles

The majority of work in profile monitoring has focused on situations where the profile is linear. For example, works by Kang & Albin (2000), Kim et al. (2003), Mahmoud & Woodall (2004), Wang & Tsung (2005), and Jensen et al. (2006b). These methods often fit linear regression models separately and monitor the regression model coefficients for determining the remote profile. Thus, the profiles are reduced to a smaller set of values that simplifies the monitoring system. However, more often the profiles are best described by a nonlinear function rather than a linear function.

https://doi.org/10.26219/heal.aueb.6243

Kang and Albin (2000) proposed two strategies to control chart for monitoring the Phase II and a process or a product characterized by a linear profile of Phase II. One approach involves a multivariate T^2 chart. Other statistical uses based on successive samples deviate from the control line in conjunction with a exponentially weighted moving mean (EWMA) chart to monitor the mean deviation and a series of charts (R-) to monitor the variation of deviations. Both approaches are described in the next section.

Suppose jth random sample collected over time have the observations (xi, yij), i = 1, 2, ..., n. It is assumed that when the process is in statistical control, the underlying model is

$$Y_{ij} = A_0 + A_1 X_i + \varepsilon_{ij} \text{ , } i = 1, 2, \ldots, n, \quad \textbf{(1)}$$

where ε 's are independent random variables (iid) that follow a normal distribution with mean zero and variance s^2. For simplicity, we must first consider the case for adopting the values X and get the same set of values for each sample. In this section we consider the case of Phase II in which the values of the controlled parameters A_0, A_1, and s_2 in the equation (1) is known.

Kang and Albin (2000) suggested two strategies to follow a procedure when all the regression parameters are known. The first strategy is to use a bivariate T^2 diagram for monitoring the regression coefficients. This diagram is based on the fact that the least squares estimators A0 and A1 follow bivariate normal distribution. The least squares estimators A0 and A1 for sample j is given by the following formulas:

$$a_{0j} = \bar{y}_j - a_{1j}\bar{x} \quad \text{and} \quad a_{1j} = \frac{S_{xy(j)}}{S_{xx}},$$

wherein

$$\bar{y}_j = \frac{1}{n}\sum_{i=1}^{n} y_{ij} \text{ , } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ , } S_{xy(j)} = \sum_{i=1}^{n}(x_i - \bar{x})y_{ij} \text{ , and } S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 \text{ .}$$

The least squares estimators A_0 and A_1 follows bivariate normal distribution with mean $\mu=(A_0,A_1)^T$ and variance - covariance table

https://doi.org/10.26219/heal.aueb.6243

$$S = \begin{pmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{pmatrix}, \quad S = \begin{pmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{pmatrix}, \quad S = \begin{pmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{pmatrix},$$

wherein

$$\sigma_0^2 = \sigma^2\left(\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}\right), \ \sigma_1^2 = \sigma^2 \frac{1}{S_{xx}} \quad \text{and} \quad \sigma_{01}^2 = -\sigma^2 \frac{\overline{x}}{S_{xx}}$$

the variances of α_{0j}, α_{1j} and covariance of α_{0j} and α_{1j} respectively.

The first tracking strategy compute the vector of estimators $Z_j = (a_{0j}, a_{1j})^T$ for sample j, wherein j, a_0 and a_1 are the constants of the sample and the gradient defined in equation. Then T^2 is given by

$$T_j^2 = (Z_j - \mu)^T S^{-1} (Z_j - \mu),$$

When the process is in control, T^2_j follows a central chi-square distribution with two degrees of freedom. Therefore, the proposed upper limit control diagram is UCL = $\chi^2_{2,\alpha}$, where in $\chi^2_{2,\alpha}$, is 100 (1-a) percentile of the chi-square distribution with two degrees of freedom. When there are shifts in the values of parameters to control Kang and Albin (2000) stressed that T^2_j follows a non-central chi-square distribution with non-centrality parameter, $\tau = n(\lambda + \beta\ \overline{x})^2 + \beta^2 S_{xx}$, where λ is the displacement of the constant A0 and b the shift of the slope A1. Both λ, b are measured in standard deviation units when the actual displacement is λσ and βσ of A0 and A1 respectively.

Also Kang and Albin (2000) studied the ARL in the performance of control charts T^2 using a simulation study. However, since T^2_j follows a non-central distribution Chisquare, we can evaluate the exact ARL of T^2 graph using the following formula:

$$\text{ARL} = \frac{1}{\Pr(T_j^2 > \chi^2_{2\alpha})}.$$

The second strategy is to implement common control charts to regression of j sample residuals using the following formula:

$$e_{ij} = y_{ij} - A_0 - A_1 x_i \ , i = 1, 2, \ldots, n.$$

https://doi.org/10.26219/heal.aueb.6243

They used an EWMA chart to monitor the mean of these deviations. They proposed an R-chart to be used in conjunction with the EWMA chart.

The mean residual for sample j calculated using

$$\bar{e}_j = \frac{\sum_{i=1}^{n} e_{ij}}{n}.$$

where z_j, j = 1,2,…., $z_j = \theta\ \bar{e}_j + (1-\theta) z_{j-1}$,

with θ ($0< \theta \leq 1$) and $z_0 = 0$. Signal out of range is given when the z_j is below the lower limit (LCL) or above the upper limit (UCL), wherein

$$LCL = -L\sigma\sqrt{\frac{\theta}{(2-\theta)n}}\ ,\quad UCL = L\sigma\sqrt{\frac{\theta}{(2-\theta)n}}$$

where L (> 0) is a constant chosen to give the signal within the ARL limits.

Sometimes you need to apply more complex models from simple linear regression model to describe a profile. Kazemzadeh et al. developed some methods for polynomial profile monitoring in Phases I and II. Mahmoud developed a technical parameter reduction and was able to expand some of the simple methods of linear regression profile for the analysis of multiple linear regression profile in Phase I. In addition, in certain applications the quality of the process or the product may be characterized by a multivariate regression model. In this case, there are some dependent qualities as response variables, to be modelled as functions of one or more explanatory variables.

Subsequent studies are those of Chang and Gan (2006) who proposed Shewhart control charts to monitor the slope of the relation between two or more measuring processes in order to ensure its accuracy but also the study of Zhu and Lin the later time also dealt the slope of the regression line in Phases I and II. Earlier, in 2004 Noorossana et al. They have proposed a new method for monitoring of simple linear regression based on multivariate Healy control charts method (1987). Also studied the effect of non-normality of error terms, using t-distributions as alternatives on the performance of EWMA / R of Kang and Albin method (2000). Mahmoud et al. (2007) and Zou et al. (2006) used methods by changing points to detect changes in the parameters of the simple linear regression model while Zhou

https://doi.org/10.26219/heal.aueb.6243

et al. (2007), proposed a method which avoids as much as possible distinction between Phase I and Phase II.

Kim et al. (2003) proposed a method for monitoring linear profiles in Phase II and recommended the application of this method and in Phase I This method involves the use of three separate Shewhart-type control charts for monitoring constant, slope, and residual variance. The performance of the method was first proposed by Kim et al. (2003) analyzed in detail by Mahmoud and Woodall (2004), who showed how to be much more effective than others. Thus Mahmoud and Woodall (2004) recommended the use of the method of Kim et al. (2003), not only because it outperforms against competing approaches, but also because it is simple and much more interpretable than other methods.

Saghaei et al. in 2009, introduced a new method (CUSUM) based on aggregated statistics amounts to enhance the monitoring of the linear profiles in phase II. The performance of the proposed method is evaluated on the basis of the ARL. Also conducted a comprehensive comparison between the performance of the proposed method and other methods for simple linear profile monitoring. The results show that the proposed method performs satisfactorily. In addition, the benchmark results are investigated, the sample size, and the sum of the squares of the explanatory variables on the performance of the proposed method. Used together three CUSUM control charts for monitoring the constant, the slope, and the variance of the error simultaneously. Their study showed that the process has a remarkable performance in the detection of a broad spectrum of different species of changes with different sizes. Furthermore, simulation studies have shown that the performance of the proposed method for detecting shifts with different sizes can be improved by choosing an appropriate reference value, K.

The proposed method performs better in detecting all kinds of shifts when the sample size increases. In the case of the shift in the slope, increase S_{xx} led to the reduction of ARL beyond permissible limits.

In the competitive life of today, scientists and professionals have come to realize that efficiency and product or process quality are more important than ever. As a result, the statistical process control techniques (SPC), an approach long prac-

 https://doi.org/10.26219/heal.aueb.6243

ticed in different sectors, is to be improved to ensure efficient and effective compliance with a wide range of requirements. In some NFC applications in which the variable quality depends functionally independent or explanatory variable (s), the quality of a process or product cannot be adequately represented by the standard use of the distribution of the same qualitative characteristic or a general body of many qualitative variables. In these cases, more often in practical applications, profile monitor, the relation between the response variable and explanatory variable(s), is highly recommended in the bibliography.

To monitor the Phase II of linear profiles that can adequately model the characteristic in many applications, several approaches have been proposed. In phase II, the process parameters are assumed to be known, and the objective is to detect displacements as possible. Kang and Albin introduced T^2 and exponentially weighted moving mean (EWMA) -R methods for monitoring the phase II of simple linear profile. For the same purpose, Kim et al. applied three independent EWMA control charts using coded values x and introduced EWMA-3 method. Noorossana and Amiri used a multivariate sum control chart (CUSUM) with a x^2 control chart to improve the performance of existing ones. Niaki et al. used generalized linear statistical model and an R control chart for this purpose. Gupta et al. compared a method developed by Croarkin and Varner performance of Kim et al. [10]. Zhou et al. proposed a multivariate exponentially weighted moving mean (MEWMA) control chart. Saghaei et al. proposed CUSUM-3 method which applies three CUSUM control charts for monitoring the intersection, the slope, and the standard deviation of a simple linear profile. Zhang et al. proposed a method based on statistical likelihood ratio to follow simple linear profile in Phase II. Noorossana et al. and Soleimani et al. suggested some methods based on time series approach in case there is autocorrelation between and within simple linear profile, respectively. Noorossana et al. also studied the performance of control charts for monitoring the simple linear profile with the assumption of irregularity in the distribution of the regression error conditions.

Although both rise and fall shifts have negative consequences for the process, all the methods proposed, have been studying only the increasing shifts. EWMA-R, MEWMA and likelihood ratio test (LRT) are effective methods to monitor growth

https://doi.org/10.26219/heal.aueb.6243

and reduce shifts. However, the methods T^2, EWMA-3, and CUSUM- 3 should be modified to be able to detect increasing and reducing shifts. On the other hand, all of the profile monitoring techniques have focused only on the detection of the stepped changes. However, drift, which is a time-varying change, often in industrial applications, deterioration graduate equipment, littering, aging of the catalyst, or human causes, such as fatigue or close supervision function are some causes of displacement (Reynolds and Stubb). Many authors, including Davis and Woodall, Divoky and Taylor, Fahmy and Elsayed, Gan, and Zhou et al. studied the performance of different control charts where there is a shift in process mean. Bissell presented the results of the evaluation length runs under linear trend charts CUSUM. Davis and Woodall have shown that the rules, concerning an ordinary long series of successive reductions and increases in operational performance, it is not effective in detecting linear displacement in the middle of the monitored process. Aerne et al. studied the performance of the Shewhart chart with supplementary runs rules, CUSUM and EWMA charts when the process mean changes as a linear trend.

Gan presented a numerical procedure to compute the mean run length (ARL) of an EWMA chart below a linear displacement of the process mean. Divoky and Taylor showed that the simulation of the process means to evaluate the sensitivity of the trend of rules when applied together with the combinations of the standard test for additional runs. Extended previously presented rules tend to a larger set of rules, and Reynolds and Stubb attended the mean and standard deviation of the process according to drift together. They showed that at detecting slow and moderate drift rate, the combination of two EWMA control charts for monitoring both the mean and variance performs better than the I / MR control chart. Fahmy and Elsayed used a statistic based on the deviation between the target means and the expected mean of the proceedings in their proposed approach.

Also, compared with the performance of their approach with CUSUM, EWMA, Shewhart, and generalized likelihood ratio (GLR) charts. Zhou et al. compared five controls charts, including EWMA, CUSUM, generalized EWMA, GLR-S, and GLR-L to monitor the movement through the presentation and the asymptotic estimation and numerical simulation of the ARL.

https://doi.org/10.26219/heal.aueb.6243

As previously mentioned, the slip detection bibliography, the focus is mainly in the middle of the process. However, displacement may occur when an operation between two or more variables that characterize process performance. The particular focus of this study is on Phase II methods including EWMA-R, MEWMA, LRT, T^2, EWMA-3, and CUSUM- 3. T^2, EWMA-3, and CUSUM-3 methods modified to become effective for monitoring the increase and decrease of the shifts in the parameters of straight linear profile.

4.3 General linear profiles

The profile monitoring is a new area of statistical quality control. In this embodiment, the quality of the process is best characterized by a linear or non-linear relation between a response variable and one or more explanatory variables. This relation is commonly known as a profile or function. A good introduction to the concept of profile monitoring, and application examples can be found in Woodall et al. (2004) and Woodall (2007). In linear profiles monitoring, the control charts that are usually applied for monitoring of the estimated parameters of the profile is represented by a regression line, such as the slope and intercept. Non-parametric models are often applied to monitor non-linear profile. Nonparametric methods do not require a specific functional form for the profile. In nonparametric regression methods, one obtains a normalized curve can be expressed as the weighted mean of the observed responses. One can monitor the metrics that measure the departures of the observed profile of a baseline reference profile established based on historical data.

In some embodiments, the quality of a process characterized by the functional relation between a response variable and one or more explanatory variables. The profile monitoring is a technique for controlling the stability of this relation over time. The general linear tracking profile is especially useful in practice because of its simplicity and its flexibility. However, existing monitoring methods suffer from the disadvantage that they all assume that the distribution of the error is normal. Where distribution followed is not appropriate, consistency and effectiveness of the more usual least squares estimator (LSE) is likely to be ineffective

https://doi.org/10.26219/heal.aueb.6243

and therefore the detection capability of the procedures based on the LSE will be significantly reduced.

In this section we describe the modelling of a general linear profile and revise existing profile monitoring systems in the bibliography.

Suppose the j-th random sample collected over time, have the observations (X_j, Y_j), wherein Y_j is nj -variate vector and X_j is a NJ × p (NJ> σ) matrix. It is assumed that when the process is in statistical control, the underlying model is

$$\mathbf{Y}_j = \mathbf{X}_j\beta + \varepsilon_j,$$

wherein

$$\beta = (\beta^{(1)}, \beta^{(2)}, \cdots, \beta^{(p)})$$

is the p-dimensional vector of coefficients and ε_j's are iid as nj multivariate normal random vector with mean zero and covariance matrix σ^2I. For the sake of generality we assume that Xj is of the form (1, X * j) wherein X * j is orthogonal to 1 and 1 is p-variate vector of all 1.

The simplest model, the straight line regression model was analyzed by Kang and Ablin (2000), Kim et al. (2003), Mahmoud and Woodall (2004) and Mahmud et al . (2005). Denote by {(x_i, y_{ij}), i = 1, 2, · · ·, n} the j-th random sample collected over time. When the process is controlled, the relation between the response variable and explanatory supposedly

$$y_{ij} = A_0 + A_1 x_i + \varepsilon_{ij}, \ i = 1, 2, \cdots, n,$$

where ε_{ij}/σ random variable that follows a standard normal distribution (IID)

$$y_{ij} = B_0 + B_1 x_i^* + \varepsilon_{ij}, \ i = 1, 2, \cdots, n,$$

Where $B_0 = A_0 + A_1\bar{x}$, $B_1 = A_1$, $x_i^* = (x_i - \bar{x})$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

While on the j-th to random appraisers least squares are B^0, B^1 and σ^2 respectively.

$$b_{0j} = \bar{y}_j, \quad b_{1j} = \frac{S_{xy(j)}}{S_{xx}}, \quad MSE_j = \frac{1}{n-2}\sum_{i=1}^{n}(y_{ij} - b_{1j}x_i^* - b_{0j})^2,$$

Where

https://doi.org/10.26219/heal.aueb.6243

$\bar{y}_j = \frac{1}{n}\sum_{i=1}^{n} y_{ij},\ S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$

and $S_{xy(j)} = \sum_{i=1}^{n}(x_i - \bar{x})y_{ij}$.

Just because these three estimators are independent, we are using three charts EWMA ($EWMA_I$, $EWMA_S$, $EWMA_E$) to detect whether the constant-Y (B_0), the slope (B_1) and the standard deviation (σ) were changed respectively. It is

$$\text{EWMA}_I(j) = \theta b_{0j} + (1-\theta)\text{EWMA}_I(j-1)$$

$$\text{EWMA}_S(j) = \theta b_{1j} + (1-\theta)\text{EWMA}_S(j-1)$$

$$\text{EWMA}_E(j) = \max\left\{\theta ln(MSE_j) + (1-\theta)\text{EWMA}_E(j-1), \ln(\sigma^2)\right\},$$

where $EWMA_I(0) = B_0$, $EWMA_S(0) = B_1$, $EWMA_E(0) = \ln(\sigma^2)$ and θ are constantly balancing.

The three EWMA charts are used together, and the change of profile is detected as one of the chart signals. The ARL comparisons show that the three EWMA charts are more effective than the methods of Kang and Albin (2000) in the detection of prolonged displacement in the Y-fixed and slope, and increase the error variance. The three charts EWMA is particularly effective in the detection of changes in the slope of the line, i.e., changes in the parameter B1 of the equation. Also, the authors argue that their method seemed more interpretable. Thus, the approach of combined multiple charts EWMA (denoted KMW chart) is used as a benchmark for comparisons

Zhou et al. (2007) investigated the performance monitoring of the proposed NEWMA system through ARL comparisons. Although the proposed control chart, may be used to monitor a general profile model, the bibliography does not appear to contain any other effective comparable process for such a model. As we can see, the KMW approach requires three charts EWMA handle one next to another, each have a statistic to be updated and designed for each sample. Such a system may be manageable for a simple linear profile case, but can become quite complicated and impractical for a general linear profile as the arrangement of the diagrams of the general profile requires more control charts to be combined

https://doi.org/10.26219/heal.aueb.6243

simultaneously to monitor such additional parameters. Therefore, the design, implementation and evaluation of the performance of such a system is rather complex and impractical.

 https://doi.org/10.26219/heal.aueb.6243

CHAPTER 5

Profile Monitoring (Part B')

5.1 Introduction

The monitoring of each process, fault detection and diagnosis at an early stage is very important for quality assurance in all organizations. Particular attention is required in most industrial processes where there is often a large number of variables, not necessarily independent of each other. In this case the traditional methods are inadequate and may lead to incorrect results. This was the impetus for the development of multivariate methods (MSPC- Multivariate statistical process control), which are applied when the quality of the process or the product is determined by more than one capacity, which should be controlled by joint and uniform way.

Some of the advantages provided by such methods is the ability to handle large numbers of variables with strong correlation, the ability to troubleshoot errors that may occur during the measurements and the possibility of applying them in cases where data are lacking. Also can reduce the size of the monitoring area, using one, two and / or three-dimensional control charts, which is more simple in presentation and interpretation. In some embodiments, the profiles can be described quite well by linear, non-linear and polynomial regression models. However, in some other embodiments, more flexible models are essential for the correct description of the profile. Much of the work that has been done for the profile monitoring is summarized in Woodall's article (2007). This chapter will introduce a large amount of research that has been made on the use of control charts to monitor the quality of a profile process of Woodall's research (2007) to date, with emphasis on multiple linear and polynomial profile, as well as some other approaches of profile monitoring.

Studies focused on simple linear profile of particular prosperity and acceptance, for example, Kang and Albin (2000), Kim et al. (2003), Mahmoud and Woodall (2004), Zhu et al. (2006, 2007b), among many others. Studies of multiple and polynomial models of regression profiles are studied by Zhou et al. (2007a),

https://doi.org/10.26219/heal.aueb.6243

Kazemzadeh et al. (2008), Mahmoud (2008), Jensen et al. (2008) and Jensen and Birch (2009). The nonlinear profile models have been studied by Williams et al. (2007). Recently, also the profile for the general model profile has also attracted much attention. The reader is referred to Zhu et al. (2008, 2009) and Qiu et al. (2010) for Phase II methods based on nonparametric regression and to Ding et al. (2006), Colosimo et al. (2008), Chicken et al. (2009) and Zhang and Albin (2009) on procedures using various techniques, various dimension - reduction techniques, wavelet transformations and independent component analysis. A recent review of the bibliography is given by Woodall (2007).

5.2 Multiple linear profiles

In some statistical processing applications, the quality of a process or a product it is represented by a functional relation between a response variable and one or more explanatory variables. Multiple linear regression is one of these functional relations. In some studies in the bibliography, several methods are proposed for monitoring the multiple linear profile's Phase II. However, there are no surveys when the sample size is smaller than the number of explanatory variables in the model. We introduce two methods to solve such a problem. The process I depends on the control of multiple linear model from a simple linear model. Two charts of exponentially weighted moving averages are used to monitor the intercept and slope of the simple linear model separately, and an EWMA chart based on the mean of squared deviations for monitoring the variation of the process. Method II uses two EWMA charts: one based on the mean of standardized residuals and the other based on the estimated variance of the standardized residuals. The simulations indicate that both methods perform very well under these conditions. Moreover, they have higher performance compared to other methods used in the bibliography to monitor multiple linear profile when the sample size is greater than the number of explanatory variables. The proposed approaches may also be used to monitor multiple linear profile when only two variables are used to determine a profile. The diagnostic tools used to detect the variability's source when a signal occurs. It will depend on the process's point of change to detect if the signal is due to a shift of the variance, the intersection, or one of the slopes.

https://doi.org/10.26219/heal.aueb.6243

Kazemzadeh et al. (2007a) and Mahmoud (2007) proposed methods for the Phase I profile represented by polynomials and multiple regression models, respectively. These authors have extended the methods of simple linear regression. Zhou et al. (2007) studied multiple regression models in the general linear model. They dealt with multivariate EWMA charts and analyzed diagnostics after an out of control's limits signal, including a sign change approach. They used an example of the semiconductor manufacturing while an example that requires multiple regression calibration methods was the motivation of the work of Parker and Finley (2007).

In some cases, multiple profiles are needed to effectively modulate the quality of a product or process. Generally the multivariate linear profiles are especially useful in practice because of its simplicity and flexibility. However, in such cases, the existing parametric methods for monitoring profiles suffer from the disadvantage that, when the parameter of the dimension is large, the detection capability of the procedures commonly used i.e. T^2 statistical charts are likely to decline significantly. Because multivariate linear profile belongs in the category of general linear as seen above, we will not expand in this category.

Moreover, Amiri et al in 2012 focused specifically on monitoring Phase II of multiple linear regression profile and suggested a new dimension reduction method to overcome the problem of the dimension in some of the existing methods. Assumed that the jth random sample collected in some time, we have n observations as follows: $x_{i1},x_{i2},\ldots,x_{ip},y_{ij}$, where i= 1,2,…,n and p is the number of the explanatory variables. The relation between the response variable and the explanatory variables is characterized by a multiple linear regression model as follows:

$\mathbf{y_j = X\beta_j + \varepsilon_j}$, where y_j is a nx1 vector of response variables for the jth random sample, X is an array nx (p + 1) of the explanatory variables, βj is a vector 1 x (p+1) of the regression parameters and ε_j is a nx1 vector of error terms taken as IID with mean μ and standard deviation σ^2. The X values are constant for each sample. The Ordinary least squares (OLS) of βj vector is as given

$$\hat{\boldsymbol{\beta}}_j = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}_j,$$

https://doi.org/10.26219/heal.aueb.6243

As we saw in the previous section Mahmoud and Woodall (2004) suggested the simplest linear model and reduced the number of variables. Therefore, the dimension of the regression parameters is reduced by (p+1) x 1 to 2 x 1. In the proposed method, Amiri et al propose the following model of the data set profile in time j

$$\begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{nj} \end{bmatrix} = \begin{bmatrix} 1 & u_1 \\ 1 & u_2 \\ 1 & \vdots \\ 1 & u_n \end{bmatrix} \begin{bmatrix} A_{0j} \\ A_{1j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix}$$

where, ui = E (yi) and is equal to $\beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}$. The A_{0j} and A_{1j} estimated for each profile using OLS. It should be noted that the OLS estimators A_{0j} and A_{1j} are linear functions of the corresponding dependent values. Since the response variables correlate, the estimators of the parameters also correlated and therefore can not be controlled separately. After fitting the model to the above equation in each sample, the estimated parameters can be written as follows:

$$\widehat{\mathbf{a}}_j = (\widehat{A}_{0j}, \widehat{A}_{1j}, z_j(\sigma))^T,$$

when the process is in statistical control, the value and variance - covariance matrix given as $\mathrm{E}(\widehat{\mathbf{a}}_j) = (0, 1, 0)^T$

and

$$\sum_{\widehat{\mathbf{a}}_j} = \begin{pmatrix} \sum_{\widehat{A}_{0j}, \widehat{A}_{1j}} & 0 \\ 0 & 1 \end{pmatrix}$$

It can be shown $\sum_{\widehat{A}_{0j}, \widehat{A}_{1j}}$ is a 2x2 matrix and equals $[U^T U]^{-1} \sigma^2$, where U is nx2 table.

Finally, we use the following statistical MEWMA for monitoring the vector of parameters and illustrated below:

$$\mathbf{w}_j = \theta \widehat{\mathbf{a}}_j + (1 - \theta) \mathbf{w}_{j-1},$$ where $w_0 = (0, 1, 0)^T$

And the upper control limit is calculated from:

$$U_j = \mathbf{w}_j^T \sum\nolimits_{\widehat{\mathbf{a}}_j}^{-1} \mathbf{w}_j > L \frac{\theta}{2 - \theta},$$

where L> (0) is selected to achieve ARL within limits.

Zhou et al. (2011) proposed a new method for monitoring the multivariate multiple linear regression profile's Phase II to overcome the problem of perspective in monitoring multivariate profile. This method, which is the extension of the work of Zhou and Qiu uses MEWMA control chart based on the LASSO (LEWMA) for monitoring the regression coefficients and the profile's variation. This control process can determine the direction of change and is capable of detecting various shifts in parameters.

5.3 Polynomial linear profiles

Because some applications in the real world, the quality of the process or product can be characterized by a polynomial relation, Kazemzadeh et al. (2008) suggested three methods for polynomial profile monitoring in Phase I.
The general polynomial profile based on the development of these three methods are:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k + \varepsilon$$

where ε follows a standard normal distribution.
Furthermore, it should be noted that for this model, the mean value χ is subtracted from each of χ and that χ values are equally spaced. After that he presented three methods for monitoring polynomial profile in Phase I. The approach of the change point, the F statistic approach and T^2
Kazemzadeh et al. (2009) proposed a new method for monitoring the polynomial profile in Phase II, and then compared the performance of this new method with existing methods, i.e. T^2 proposed by Kang and Albin (2000), with MEWMA of Zhou et al. (2007) and MCUSUM / χ^2 proposed by Noorossana and Amiri (2007). The general idea of this new method is that polynomials profiles are transformed into orthogonal polynomials profile, and then the regression parameters of the transformed model are independent and can be monitored separately. Moreover, this method is recommended when the order of the polynomial regression is not large, so there are not much control charts.
In conclusion, they showed that the proposed method performs better than the three other existing methods. Also, since there is no one - to - one relation between the parameters of the main and the transformed model, the explanation of

https://doi.org/10.26219/heal.aueb.6243

the results may be difficult in some cases. For this reason, the proposed method is not recommended for diagnostic purposes. Although, this diagnostic method could be the subject of future research of polynomial regression or other profile. Kazemzadeh et al. (2007a) compared the method of the F-statistic of Mahmoud and Woodall (2004), by the method of changing the sign of Mahmoud et al. (2007), and a method based on statistical information T^2 to detect a displacement of the regression parameters. Mahmoud (2007), on the other hand, compared the F-statistic of Mahmoud and Woodall method (2004), by a T^2 based method. Kazemzadeh et al. (2007b) examined the Phase II monitoring of polynomial regression profile.

A polynomial profile provides flexibility and customization through the choice of the boundary conditions for a desired movement. The general form of a polynomial profile is given by

$$s(u) = C_0 + C_1 u + C_2 u^2 + \cdots + C_n u^n \quad (1)$$

The coefficients of the polynomial, Cn, are selected to satisfy the boundary conditions and limit the dynamic response. The function s (u) is the displacement where u is the normalized profile time such as u = 0 at the beginning of motion and u = 1 at the end of the movement. Two common polynomial profiles are 3-4-5 and 4-5-6-7, named in the order of the terms of the polynomial. The 3-4-5 polynomial provides the continuity of the initial and final terms of displacement, velocity v (u), and the acceleration α(u):

$$s(0) = 0,\ v(0) = 0,\ a(0) = 0$$
$$s(1) = h,\ v(1) = 0,\ a(1) = 0 \quad (2)$$

where h is the size of the profile's increase. Solving the coefficients of (1) using the boundary conditions in (2) gives the 3-4-5 polynomial profile:

$$s(u) = h(10u^3 - 15u^4 + 6u^5) \quad (3)$$

The 4-5-6-7 polynomial profile can be achieved by the inclusion of additional constraints on the continued jolt j (u), the third time, derivative of the position:

$$j(0) = 0$$
$$j(1) = 0 \quad (4)$$

https://doi.org/10.26219/heal.aueb.6243

The 4-5-6-7 polynomial profile:

$$s(u) = h(35u^4 - 84u^5 + 70u^6 - 20u^7) \quad (5)$$

The rule of thumb can be used to select polynomial profile, and additional restrictions may be placed on transient characteristics to achieve a desired path. However, the polynomial profile does not guarantee sufficient reduction of vibration without the full dynamic system analysis. A good example of using polynomial profile is provided by Amiri et al. (2009) from the automobile industry. Consider the basic relation on a car engine using a second degree polynomial however there was autocorrelation between profiles. After data reduction in a number of parameter estimators, they performed a step by step Phase I analysis of the polynomial tracking using a T^2 based procedure in order to check the stability of the process.

5.4 Nonlinear profiles

A non-linear profile typically is defined by a nonlinear regression model. Suppose that there are m data profiles, each of which has n measurements. The relation between the response variable and the independent variable can be modeled as

$y_{ij} = f(x_{ij}, \beta) + \varepsilon_{ij}$, i = 1, 2,…,m; j = 1, 2,…, n

where f (;) is a nonlinear function, xij is a single regression variable, β is a p × 1 vector of parameters for each profile, and εij's are random independent variables, that follow a normal distribution with mean μ and variance σ^2 . For example, a non-linear profile was investigated by Williams et al. is given by

$$f(x_{ij},\beta) = \begin{cases} a_1(x_{ij}-c)^{b_1}+d, & x_{ij} > c \\ a_2(-x_{ij}+c)^{b_2}+d, & x_{ij} \le c \end{cases} \quad i = 1,2,\ldots,m; j = 1,2,\ldots,n$$

where β = (a1, a2, b1, b2, c, d).

Ding et al. (2006) used various types of data reduction methods for nonlinear profile data, especially the main component analysis and independent component analysis. These authors also studied methods for the analysis of variance for the profile data. Williams et al. (2007) used the approach of non-linear regression of

 https://doi.org/10.26219/heal.aueb.6243

Williams et al. (2007) using indicative data supplied by DuPont. A four-parameter regression model was used to represent the profile. These articles were focused on the monitoring of the parameters of the nonlinear regression model. Only a few types of nonlinear regression are studied to date in this manner.

Because nonlinear profiles are mainly applied in the real world's applications, Vaghefi et al. (2009) developed two approaches for monitoring the process of nonlinear profiles' Phase II. The general non-linear model is assumed:

$$y_{ij} = f(x_{ij}, \beta_i) + \varepsilon_{ij}$$

where,

- xij is k × 1 vector of independent variables for the jth observation in the ith profile
- εij are random errors, where εij ~ $N(0,\sigma^2)$, IID
- β_i is a p × 1 vector of parameters for each i profile with p> 1
- f is nonlinear in the parameters.

They even assumed for the model, n observations in the ith sample collected and presented as (xij, yij). In the first approach, they developed various control charts for non-linear profiles monitoring using parametric estimates of the regression model, and the second approach they used the measurement method for measuring the deviation from the reference curve. Below are two approaches.

- Multivariable T^2 control chart

This T^2 control chart is based on the T^2 control chart of Kang and Albin (2000), proposed for monitoring linear profiles. The ith statistical chart by Kang and Albin (2000) is as follows:

$$T_i^2 = (\hat{\beta}_i - \beta)\Sigma_\beta^{-1}(\hat{\beta}_i - \beta)$$

where applicable:

- the min squared estimators of the vector regression coefficient β = [βoβ1] follow bivariate normal distribution.
- In Phase II it is assumed that the covariance matrix of estimators Σ_β is known.

 https://doi.org/10.26219/heal.aueb.6243

- When the process is in control, the T^2 is a chi square random variable with two degrees of freedom.

Other methods suggested, is a combination of a S and EWMA control chart and a combination of an MCUSUM and x^2 control chart.

- Metrics Method

In the second approach they suggest metrics method. This approach uses non-parametric methods to overcome some problems arising from the complexity of the estimation of the nonlinear profile coefficient. Especially Williams et al. (2007) proposed several distance measurements that can be expanded in Phase II applications for nonlinear models.

Several studies are conducted by researchers, leading Amiri and Kazemzadeh, on using polynomial models for profile monitoring. Kazemzadeh et al. (2009) proposed the use of orthogonal polynomial regression to polynomials profile monitoring in phase II. In this method, one must turn the polynomial regression in orthogonal polynomial regression model. Then, the regression parameters are independent and can be tracked separately.

This method is proposed when the number of the polynomial regression (k) is not so great. If the command was great, this method can be a little difficult, because we will have to deal with many control charts. However, Montgomery et al. claim that, as a general rule, using a higher-order polynomial (k> 2) should be avoided unless it can be justified for outside the data reasons. Further, they believe that a low-order model in a transformed variable is almost always preferable to a high-order model to the initial measurement, and a sense of parsimony must maintained. Namely, we should use the simplest possible model that is consistent with the data and knowledge of the environmental problem. Therefore, in practice, one should try to characterize the quality of a process or product of lower order polynomial regression model. Therefore, the proposed method can be used easily in most applications. They proved with their research that the proposed method performs better than the other methods of point detection. As there is no one-to-one relation between the parameters of the main and the transformed model, the explanatory of the results could be difficult in some cases.

https://doi.org/10.26219/heal.aueb.6243

Therefore, the proposed method can be used in most applications easily. They proved with their research that the proposed method performs better than the other methods of detection point. As there is no one-to-one relation between the parameters of the main and the transformed model, the explanatory of the results could be difficult in some cases. Therefore, the proposed method is not recommended for diagnostic purposes. The development of a diagnostic method can be considered as a future investigation especially in polynomial or multiple linear regression cases. For example, one can use an artificial neural network as a tool for this purpose.

For nonlinear models, after their research Jensen and Birch suggested a NLM model to eliminate random effects and correlation errors. They concluded that an approach that uses separate NL regression models for determining the effects of random functions well in creating the analysis models with NLM profile. The proposed method uses an easy way to calculate the control limit and therefore does not require extensive simulation to calculate the correct control limit as the approach of Williams et al. (2006a). They end up that they can ignore the correlation of errors and focus their efforts on random effects modelling. The random effects modelling allows to use a Phase I control limit that will not be required to be obtained by simulation as it would be required if we only received the estimates from each NL regression models.

Moguerza et al. (2007) used appropriate vector machines to monitor themselves applied curves instead watch the model parameters being adapted to the curves. Their approach contrasts with that of Williams et al. (2007). Chen and Nembhard (2007) proposed a large-scale approach control chart in the profile based on the statistic Neyman test of rates.

The traditional statistical process control (SPC) charts have been widely used to monitor the processes of changes in means, standard deviations, and Poisson values. The multivariate control charts are also available to watch the changes in mean vectors and covariance matrices. These control charts have been used successfully to monitor processes for many years. In many modern industrial processes, vast amounts of data are often readily available and due to the complex data structures cannot be easily addressed by traditional control charts.

 https://doi.org/10.26219/heal.aueb.6243

One problem with the monitoring of the parameters of the linear profile is that a process could be changed so that the resulting profile would be no longer linear or even that a straight line could yield the same parameter estimates. Thus, the lack of linearity could pass unnoticed. While it appears that the EWMA charts work well in the monitoring of changes in linear profile parameters, there are many cases where the form of the profile is non-linear or otherwise too complicated to be expressed as a parametric function. It proposed a method for detecting changes in a series of profiles that do not depend on strong assumptions about the shape of the profile. Such a method necessarily involves the use of non-parametric tools.

Chicken et al. suggested an approach that allows much more general differences between profiles, requiring only that the differences are functions in the class of finite L^2. That is, the differences between the profiles in the f0 control and out of control f_0 profile called only to be such that

$$\|f^1 - f^0\|^2_{L^2} = \int_a^b (f^1(x) - f^0(x))^2 dx < \infty,$$

where [a, b] is the common support profile. This represents a very large class of profiles, including non-parametric normalizations as those of Williams et al. (2007) and Gardner et al. (1997) who did not manage well.

Wang and Guo showed in a recent study of 2013 a new method for evaluating the performance of the process for non-linear profile and a simple formula to obtain its confidence interval. They conducted an evaluation of the performance of the proposed method and a real example is used to illustrate the application.

Moreover, Amiri et al. with their study in 2009 provide an example of a polynomial model application in industry. In their article they explain that one of the most important qualities of a car engine is the relation between the torque produced by the engine and the engine speed in revolutions per minute (RPM). For the type of engine (TU3, assembled on a Peugeot car), the engine operates at different RPM values and obtained respective values of torque. In other words, the torque generated by each motor is regarded as the response variable and the speed values correspondent considered as explanatory variables. If the manufacturing process is in control, the profiles explaining the relation between RPM and

https://doi.org/10.26219/heal.aueb.6243

torque must be similar. An engine with mechanical defects or other issues will yield a profile that is different from the good engines. It is desired to be applied a process for monitoring the quality level of engine. Because there are many RPM values obtained for each engine, it is natural to try to apply a multivariate quality control procedure for the detection of engines that are not acceptable. However, if the number of RPM values is large, it is better to use a parametric model to describe the relation and monitor the estimated parameters instead of the actual values of torque.

In Phase I of a system of control chart, we have a set of historical data at our disposal from which we try to determine whether the profiles form a consistent set of similar profile without remote profiles. Once we are satisfied that the data set does not contain remote profiles, we can estimate the parameters of the process that were fixed for the Phase II control system chart. In Phase II, we follow the procedure and care to detect displacements as quickly as possible to continue to collect profile in real time. So our goal in Phase II is to know when the process is shifted from the historical, steady state, so that we can react quickly to these changes and prevent problems.

The study showed that this relation may be modelled by a second order polynomial profile and there is autocorrelation structure in each profile. They proposed an IMM approach to estimate the polynomial regression parameters. Next they used a control process T^2 for conducting Phase I studies. The results showed that the process produced by the engine is constant. Furthermore, using the standard procedure, the estimated mean vector and variance - covariance matrix parameter estimators, which are necessary to monitor the process in the future.

https://doi.org/10.26219/heal.aueb.6243

5.5 Mixed profiles

Most linear models applied to profile data assume that the errors are independent and identically distributed random variables. In some cases, however, errors may exhibit autocorrelation. This generalization studied by Jensen et al. (2007) and Jensen and Birch (2007) for a straightforward and non-linear regression, respectively, using mixed models. Jensen and Birch (2007) showed that the use of mixed models could have significant advantages if nonlinear regression models are used.

Mosesova et al. (2007) also suggested a mixed model approach with an example from the automotive industry. Among other issues, the authors discussed the issue of registration curve i.e. aligning the curves before analysis. This step is necessary in certain applications. Similarly Shiau et al. (2009) proposed a method for monitoring non-linear profile with random results using nonparametric regression methods. They used the technique of principal components analysis to analyse the covariance structure of the profile.

Abdel-Salam et al. 2012 proposed a semi parametric process combining P (parametric) and NP (nonparametric) profile. They refer to the semi parametric process as mixed robust model monitoring profile (MRRPM). These three methods (P, NP and MRRPM) may account for the autocorrelation between profiles and manage the collection of profiles as a random sample from the population. Also they offered a version of the Hotelling T^2 for use and analysis in Phase I to determine unusual profile based on the estimated random results and take the corresponding control limits. The simulation results showed that the method MRRPM performs well in decisions concerning the unusual profile compared to methods based on a poorly defined model P or NP based on a regression. Additionally, however, the MMRPM method is a model capable of good specificity because it also performs well compared to a well-defined model of P. The proposed chart is able to detect changes in the Phase I data and the control limits is easily calculated.

Peihua Qiu et al. (2010), studied the case of non-parametric monitoring profile when the data in the profile are associated. Specifically, in this method, within -

https://doi.org/10.26219/heal.aueb.6243

profile relation is described by a nonparametric mixed model. Finally, based on the estimated structure of the variation from one IC (within control limits) data set, which proposes a new Phase II control chart, incorporating the local linear smoothing of core profile data within an EWMA control chart. Initially, the estimated relative IC process by a set IC data using nonparametric mixed model. They presented the following model:

$$y_{ij} = g(x_{ij}) + f_i(x_{ij}) + \varepsilon_{ij} \quad j = 1, \dots . n_i \, , i = 1, \dots . m$$

They concluded that the proposed for Phase II control chart is efficient in detecting shifts in various cases. Also, the proposed method is very useful and effective in industrial applications.

Moreover, they suggested various topics for future research, such as the extension of the proposed method for the analysis of Phase I, in which the detection of outliers is also interesting. Another subject of future research proposed is the detection of possible changes in the variance - covariance matrix of the profile. Finally, as a future research topic is proposed the generalization of the proposed control chart for multivariate cases.

Statistical process control is important and challenging to track profiles with categorical data and random predictor variables. They used the GLM to model the relation between the binary response variable and random predictor variables. They proposed a new control system, EWMA-GLM, to track profile with binary data and random explanatory variables. The EWMA-GLM system integrates EWMA system and logistic regression likelihood ratio test. As shown by the simulation results of the EWMA-GLM system it runs almost always better than the Shewhart-GLM system and NEWMA-GLM system, developed as a benchmark for comparison of performance based on existing research.

https://doi.org/10.26219/heal.aueb.6243

5.6 Wavelets profiles

Wavelets are still a popular way to represent the data profile when simple models cannot adequately depict the shape of the profile. Reis and Saraiva (2006) used the wavelets to represent the surface of the paper. Zhou et al. (2007), Jeong et al. (2006), and Chicken and Pignatiello (2007) have also proposed methods based on wavelets. Chicken et al. (2009) suggested a likelihood ratio test based on wavelets to monitor nonlinear and nonparametric profiles and to estimate the actual time change. Paynabar and Jin (2011) extended the LRT model with the inclusion of wavelets coefficients with random results to explain the variability between profiles. They used this method to group successive profile and assess the actual change time from a single or multiple points. Wavelets are usually recommended when the shape of the profile is too complex for linear and non-linear models. As noted by Woodall et al. (2004), however, monitoring only a subset of important wavelets coefficients based on the data control can be dangerous in the sense that some changes in the operation outside the control limits will not be detectable. Finally in 2014, Chang et al, gave a real example monitoring the condensation-water temperature profiles to demonstrate the proposed framework. In normal profile monitoring change detection occurs when a complete profile is created. In their own study, the detection of a possible change profile occurs before all the information about the interest profile become fully available. The main objective of the research is to make a proper decision-making process as soon as possible.

They studied two possible solutions, but favoured the first approach, due to the high accuracy and fast scanning time. The proposed method comprises the following steps: determining the desired fold decision (t's), modelling the exponential decay function to the set of profiles in time t, the calculation of the profile areas' rates over time, calculate the AR(1) filter indoors and standardization of statistics for each profile using the mean and standard deviation of statistical data obtained from the Phase I profile. Finally, the statistics obtained from these procedures for each of the profiles are fed to IX and MR control charts for monitoring process. Simulations provide various controlled scenarios for out of control values that are mixed with normal profiles. The simulation results showed that

https://doi.org/10.26219/heal.aueb.6243

the proposed methods are able to detect these abnormal profiles although it also created some false alarms. Both methods provide a high accuracy rate.

The proposed method can be extended to other applications in which the characteristic of interest demonstrate an oscillating pattern. For example, Hammond et al. (2013) found that a flow pattern of a syringe pump of an automated tool for veterinary diagnostics. When you go extreme ranges in the standard wave, a mode of failure occurs and leads to performance loss.

 https://doi.org/10.26219/heal.aueb.6243

CHAPTER 6

Conclusions

In this study we sought the introduction to statistical quality and process control and studied the basic concepts and the main governing models.

Statistical quality control is a method of monitoring, management, maintenance and improvement of the performance of a process (either production or service) through the use of statistical methods. The statistical quality control tool traditionally used in the industry, but in recent years become more and more obvious contribution to the quality assurance and the service sector. Failed use of statistical quality control can lead to increased product recalls, more defective, an increased number of customer complaints, guarantees and reduced costs margins, reduce productivity and market share etc. (Little, 2001).

By contrast, the successful application of statistical quality control can bring significant benefits, both in outcome level directly (fewer defects, fewer cost losses, efficiency increase, reducing quality costs) and long term (increase market share, reduce complaints, consistent increase in product and procedures, establish a reputation of high quality products) (Mason and Antony, 2000). It also creates a common language between employees of different levels on the performance of the process, finds when there is natural and when specific volatility cause, it helps to better understand the process by the involved and helps to better understand the quality concept in the usage of the client because of the possibility provided for compliance to the designated requirements (Antony et al, 2000). More generally, it creates a framework within which decisions based on facts, not assumptions. (Antony et al., 2000) and it happens dynamically, enabling operation in semi-finished products and perform corrective actions (Dimancescu and Dwenger, 1996).

The difficulties in the application of statistical quality control can be grouped into two categories. The first category concerns the difficulties arising during the introduction of statistical control in a process generally. The second category relates to the obstacles that appear when there is wider application of effort and expansion of control.

https://doi.org/10.26219/heal.aueb.6243

In the first chapter we presented an introduction to the basic concepts of statistical quality control and its importance. In the second chapter we studied the main tools of statistical control, the control charts. The control charts are the graphical representation of a qualitative characteristic, which has been measured or calculated from a set of measurements as a function of the serial measurement or time (Montgomery, 2009), and as mentioned above, are used to control the volatility of a process. Specifically we saw the Shewhart, EWMA and CUSUM control charts.

In the third chapter of the work we expanded in multivariate control charts. Monitoring correlated variables is the main object of study processed by Multivariable Statistical Process Control (MSPC) (Montgomery (2005)). That enables simultaneous monitoring of two or more variables of a product. The need to use the Multivariable Statistical Process Control arose from the finding that the quality of a product can be associated with more than one quality and measurable characteristics. Specifically we saw the multivariate Shewhart control charts, EWMA and CUSUM.

In the fourth and fifth chapter of our work we attempted insertion into the profile monitoring. However, due to lack of relative bibliography we stayed in theoretical context. Moreover, the complexity of the subject exceeds our work purposes. In many applications, the quality of a process or a product is characterized by the functional relation between a dependent variable and one or more explanatory variables. Specifically, at each sampling stage one can observe a set of data points from these variables, which can be represented by a profile (curve). The monitoring profile is mainly used for controlling the stability of the comparison over time based on the data observed in the profile. In some embodiments, the profile can be described quite well by linear, non-linear and polynomial regression models. However, in some other embodiments, more flexible models are essential for the correct description of the profile. The profile monitoring is a very useful tool in a growing number of practical applications. Much of the work in recent years has focused on the use of more efficient mapping methods, the study of the shapes of the profiles, and the study of the effects of cases of violations. There are many issues to be solved and the promising research should continue as the wide range

https://doi.org/10.26219/heal.aueb.6243

of profile shapes and possible patterns. Some profiles are considering whether and in theory are mainly Simple linear profiles, Multiple linear profiles, General linear profiles, Polynomial profiles and Nonlinear profiles as well as some very recently in the bibliography.

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243

Bibliography

Abdel-Salam, Jeffrey B. Birch and W. Jensen "A Semiparametric Mixed Model Approach to Phase I Profile Monitoring" Published online 10 May 2012 in Wiley Online Library

Alt, F.B. (1985). Multivariate Quality Control. *The Encyclopedia of Statistical Sciences*, Kotz S., Johnson, N.L. and Read, C.R. (eds), New York: John Wiley, 110-122.

Alt, F.B. and Smith, N.D. (1988). Multivariate Process Control. *Handbook of Statistics*, P.R. Krishnaiah and C.R. Rao (eds), North-Holland: Elsevier Science Publishers B.V., 7, 333-351.

Amiri A, Jensen W, Kazemzadeh R. (2009) "A Case Study on Monitoring Polynomial Profiles in the Automotive Industry" Published online 21 October 2009 in Wiley InterScience

Amiri, A., Eyvazian, M., Zou, C., Noorosana, R.(2011) A parameters reduction method for monitoring multiple linear regression profiles, Published online 10 June 2011 in Springer –Verlag, 621-629

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*, 2nd ed, New York: John Wiley.

Chang, T. C.; Gan, F. F. (2006) Monitoring Linearity of Measurement Gauges. Journal of Statistical Computation and Simulation, v. 76, p. 889-911.

Chang, Tavakkol, Chou and Tsai (2014) "Real-time detection of wave profile changes" Computers & Industrial Engineering 75, 187–199

Crosier, R.B. (1988). Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. *Technometrics*, 30, 3, 291-303.

Hawkins, D.M. (1991). Multivariate Quality Control Based on Regression-Adjusted Variables. *Technometrics*, 33, 1, 61-75.

Hawkins, D.M. (1993). Regression Adjustment for Variables in Multivariate Quality Control. *Journal of Quality Technology*, 25, 3, 170-182.

Healy, J.D. (1987). A Note on Multivariate CUSUM Procedures. *Technometrics*, 29,4, 409-412.

Hicks, C.R. (1955). *Some Applicationsof Hotelling's* T^2. *Industrial Quality Control, Vol 11(9), 23-29.*

Hotelling, H. (1947). Multivariate Quality Control - Illustrated by the Air Testing of Sample Bombsights. *In Techniques of Statistical Analysis*, Eisenhart, C., Hastay, M.W., Wallis, W.A. (eds), New York: MacGraw-Hill, 111-184.

Kang, L. and Albin, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile", *Journal of Quality Technology* 32, pp. 418-426.

Kazemzadeh, R.B., Noorosana, R., Amiri, A.,(2009), Monitoring Polynomial Profiles in Quality Control Applications, The International Journal of Advanced Manufacturing Technology, 42, 703-712

Kourti, T. and MacGregor, J.F. (1996). Multivariate SPC Methods for Process and Product Monitoring. *Journal of Quality Technology*, 28, 4, 409-428.

Kourti, T., Nomikos, P. and MacGregor, J.F. (1996). Analysis, Monitoring and Fault Diagnosis of Batch Process Using Multiblock and Multiway PLS. *Journal of Production Control*, 5, 4, 277-284.

Lowry, C.A. and Montgomery, D.C (1995). A Review of Multivariate Control Charts. *IIE Transactions*, 27, 800-810.

Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992). A Multivariate EWMA Control Chart. *Technometrics*, 34, 1, 46-53.

Lu, X.S., Xie, M., Goh, T.N. and Lai, C.D. (1998). Control Charts for Multivariate Attribute Processes. *International Journal of Production Research*, 36, 12, 3477-3489.

MacGregor, J.F. and Kourti, T. (1995). Statistical Process Control of Multivariate Processes. *Control Engineering Practice*, 3, 3, 403-414.

Mahmoud, M.A., (2008), "Phase I Analysis of Multiple Linear Regression Profiles," Communications in Statistics, Simulation and Computation, 37(10), 2106-2130

Mahmoud, M. A. and Woodall, W. H. (2004). "Phase I Analysis of Linear Profiles with Calibration Applications", *Technometrics*, November 2004, Vol.46, No.4, 380-391.

Mahmoud, M.A., Parker, P.A., and Woodall, W.H., (2007), "A Change Point Method for Linear Profile Data," Quality and Reliability Engineering International, 23(2), 247-268.

https://doi.org/10.26219/heal.aueb.6243

Mason, R.L., Tracy, N.D. and Young, J.C. (1995). Decomposition of T^2 for Multivariate Control Chart Interpretation. *Journal of Quality Technology*, 27, 2, 99-108.

Mason, R.L., Tracy, N.D. and Young, J.C. (1997). A Practical Approach for Interpreting Multivariate T^2 Control Chart Signals. *Journal of Quality Technology*, 29, 4, 396-406.

Mason, R.L. and Young, J.C. (1999). Improving the Sensitivity of the T^2 Statistic in Multivariate Process Control. *Journal of Quality Technology*, 31, 2, 155-165.

Montgomery, D.C. (1991). *Introduction to Statistical Quality Control*. 2nd ed, New York: John Wiley.

Montgomery, D.C. (2005). *Introduction to Statistical Quality Control*. 5nd ed, New York: John Wiley & Sons, Inc.

Montgomery, D.C., and H.M. Wadsworth, Jr. (1972). Some Techniques for Multivariate Quality Control Applications, ASQC Technical Conference Transactions, Washington, DC, May, pp. 427-435.

Noorosana, R., Eyvazian, M., Vaghefi, A. (2010). Phase II monitoring of multivariate simple linear profiles, Journal of Computers & Industrial Engineering,58, 563-570

Pignatiello, J.J. and Runger, G.C. (1990). Comparisons of Multivariate CUSUM Charts. *Journal of Quality Technology*, 22, 1, 173-186.

Rigdon, S.E. (1995). ,A double-integral Equation for the Mean run length of a mewma Control Chart. *Statistcs and Probability Letters*, 24, 365-373.

Rigdon, S.E. (1995). An Integral Equation for the In-Control Mean Length of a MEWMA Control Chart. *Journal of Statistical Computations and Simulation*, 52, 351-365.

Runger, G.C., and Alt, F.B. (1996). Choosing Principal Components for Multivariate SPC. *Communications in Statistics - Theory and Methods*, 25, 5, 909-922.

Runger, G.C., Alt, F.B. and Montgomery, D.C. (1996a). Contributors to a Multivariate SPC Chart signal. *Communications in Statistics - Theory and Methods*, 25, 10, 2203-2213.

Runger, G.C., Alt, F.B. and Montgomery, D.C. (1996b). Controlling Multiple Stream Processes with Principal Components. *International Journal of Production Research*, 34, 11, 2991-2999.

Runger, G.C. and Prabhu, S.S. (1996). A Markov Chain Model for the Multivariate EWMA Control Chart. *Journal of the American Statistical Association*, 91, 436, 1701-1706.

Saghei, A.,Mehrjoo, M., Amiri, A. (2009). A CUSUM-based method for monitoring simple linear profiles. Published online 13 May 2009 in Springer- Verlag

Tracy, N.D., Young, J.C. and Mason, R.L. (1992). Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 24, 2, 88-95.

Vaghefi, A., Tajbakhsh, S.D., and Noorossana, R., (2009). "Phase II Monitoring of Nonlinear Profiles," Communications in Statistics-Theory and Methods, 38(11), 1834-1851.

Wade, M. R., and Woodall, W. H. (1993). "A Review and Analysis of Cause-Selecting Control Charts". *Journal of Quality Technology* 25, pp. 161-169

Woodall, W.H. and Montgomery, D.C. (1999). Research Issues and Ideas in Statistical Process Control. *Journal of Quality Technology*, 31, 4, 376-386.

Woodall, W.H. and Ncube, M.M. (1985). "Multivariate CUSUM Quality Control Procedures". *Technometrics*, 27, 3, 285-292.

Zhang, G. X. (1992). "Cause-Selecting Control Chart and Diagnosis, Theory and Practice". Aarhus School of Business, Department of Total Quality Management, Aarhus, Denmark.

ZHU, J.; LIN, D. K. J. (2007). Monitoring the Slopes of Linear Profiles, paper presented at the Joint Statistical Meetings, Salt Lake City, Utah and under revision for resubmission to the Journal of Quality Technology.

ZOU, C.; ZHANG, Y.; WANG, Z. (2006). A Control Chart Based on a Change-Point Model for Monitoring Profiles. IIE Transactions v. 38, p. 1093-1103.

ZOU, C.; TSUNG, F.; WANG, Z. (2007). Monitoring General Linear Profiles Using Multivariate Exponentially Weighted Moving Mean Schemes. Technometrics, v. 49, n. 4, p. 395-408.

https://doi.org/10.26219/heal.aueb.6243

https://doi.org/10.26219/heal.aueb.6243