

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΑΝΑΛΥΣΗ ΦΑΣΗΣ Ι ΓΙΑ ΤΗ ΒΕΛΤΙΩΣΗ ΚΑΙ ΤΗΝ ΠΑΡΑΚΟΛΟΥΘΗΣΗ ΤΩΝ ΠΑΡΑΓΩΓΙΚΩΝ ΔΙΕΡΓΑΣΙΩΝ

Ευαγγελία Κακαβούλη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς

Ιούνιος 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΑΝΑΛΥΣΗ ΦΑΣΗΣ Ι ΓΙΑ ΤΗ ΒΕΛΤΙΩΣΗ ΚΑΙ ΤΗΝ ΠΑΡΑΚΟΛΟΥΘΗΣΗ ΤΩΝ ΠΑΡΑΓΩΓΙΚΩΝ ΔΙΕΡΓΑΣΙΩΝ

Ευαγγελία Κακαβούλη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος
Ειδίκευσης στην *Εφαρμοσμένη Στατιστική*

Πειραιάς

Ιούνιος 2018

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Δ. Αντζουλάκος (Επιβλέπων)
- Χ. Ευαγγελάρας
- Μ. Μπούτσικας

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS

School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**PHASE I ANALYSIS FOR PROCESS
IMPROVEMENT AND MONITORING**

by

Evangelia Kakavouli

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfilment of the
requirements for the degree of Master of Science in Applied
Statistics

Piraeus, Greece

June 2018

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κ. Δημήτριο Αντζουλάκο αφενός μεν για την διδασκαλία του μαθήματος και αφετέρου για την καθοδήγησή του προκειμένου να ολοκληρωθεί η παρούσα εργασία.

Περίληψη

Καταναλωτές και κατασκευαστές επιθυμούν υψηλής ποιότητας προϊόντα τα οποία να ικανοποιούν τις ανάγκες για τις οποίες κατασκευάστηκαν. Μέσω του Στατιστικού Ελέγχου Ποιότητας καταφέρνουμε να επιτύχουμε την έγκαιρη ανακάλυψη μη συμμορφωμένων στις προδιαγραφές προϊόντων και στη συνέχεια μας δίνει τη δυνατότητα διόρθωσης προκειμένου να μην παρατηρηθούν παρόμοια ελαττώματα στο μέλλον. Θεμελιώδες στάδιο της μεθόδου αυτής είναι η Φάση I.

Σκοπός της παρούσας διπλωματικής εργασίας είναι να παρουσιάσει την πρώτη φάση του στατιστικού ελέγχου μιας διεργασίας και να αναλύσει τα τεχνικά μέσα αλλά και τις μεθόδους που οδηγούν στην βελτιστοποίηση των παραγωγικών διεργασιών.

Στην εργασία αυτή περιλαμβάνονται τρία κεφάλαια

- ❖ Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στο Στατιστικό Έλεγχο Ποιότητας.
- ❖ Στο δεύτερο κεφάλαιο παρουσιάζεται το τεχνικό υπόβαθρο της Φάση I του ελέγχου διεργασίας.
- ❖ Τέλος, στο τρίτο κεφάλαιο αναλύονται οι βασικότερες μέθοδοι υπολογισμού των ορίων ελέγχου μιας παραγωγικής διεργασίας στη Φάση I.

Abstract

Both consumers and manufacturers want high quality products to meet the needs for which they were built. Through Statistical Quality Control we manage to detect in time the non-conforming products and then enables us to make the required corrections in order to avoid such defects in the future. A fundamental stage of this method is Phase I.

The purpose of this thesis is to present the first phase of statistical process control and to analyze the technical background and the methods that lead to the optimization of the productive processes.

This chapter includes three chapters

- ❖ In the first chapter there is an introduction to the Statistical Quality Control.
- ❖ The second chapter presents the technical background of the Phase I in process control.
- ❖ Finally, the third chapter analyzes the basic methods of calculating the control limits of a production process in Phase I.

Περιεχόμενα

Κατάλογος Πινάκων	I
Κατάλογος Διαγραμμάτων	II
Κατάλογος Εικόνων	III
Κεφάλαιο 1	1
Εισαγωγή	1
1.1 Η έννοια της ποιότητας και οι διαστάσεις της	1
1.2 Τι είναι ο Στατιστικός Έλεγχος Ποιότητας	2
1.3 Στατιστικός Έλεγχος Διεργασιών	4
1.4 Το πρόβλημα του Στατιστικού Ελέγχου Διεργασιών	11
1.5 Γενικές Αρχές Διαγραμμάτων Ελέγχου	13
1.6 Μέτρα απόδοσης ενός Διαγράμματος Ελέγχου	19
1.7 Άλλα μέτρα απόδοσης	21
1.8 Ταξινόμηση Διαγραμμάτων Ελέγχου	22
1.9 Βασικά Είδη Διαγραμμάτων	23
1.9.1 Διαγράμματα Ελέγχου τύπου Shewhart	24
1.9.2 Διαγράμματα Ελέγχου CUSUM	25
1.9.3 Διαγράμματα Ελέγχου EWMA	27
1.10 Διαγράμματα Ελέγχου Ιδιοτήτων	28
1.11 Φάση I και Φάση II	30

Κεφάλαιο 2	33
2.1 Επισκόπηση	33
2.1.1 Χρήση Ορθολογικής Δειγματοληψίας	34
2.1.2 Μέγεθος δείγματος στη Φάση I	35
2.1.3 Γραφικές Μέθοδοι και διαγράμματα multivari	36
2.1.4 Επιλογή του κατάλληλου εντός ελέγχου μοντέλου	38
2.1.5 Μέτρα απόδοσης	40
2.2 Φάση I – Στατιστικός Έλεγχος Διεργασιών	41
2.2.1 Υπόβαθρο	44
2.2.2 Εφαρμογή δίπλευρων Διαγραμμάτων Ελέγχου τύπου Shewhart – Καθορισμός των ορίων ελέγχου	47
2.3 Μέθοδοι που χρησιμοποιούνται στη Φάση I	49
2.3.1 Έλεγχος ακραίων σημείων στη Φάση I	49
2.3.2 Διαγράμματα Ελέγχου Αυτόματης Εκκίνησης	50
2.3.3 Μέθοδος αλλαγής σημείου	51
2.3.4 Ταξινόμηση και μέθοδοι βασισμένες σε συστάδες	52
2.3.5 Ανθεκτικές εκτιμήσεις για εντός ελέγχου παραμέτρους	53
2.4 Διαγράμματα Ελέγχου για μεταβλητές στη Φάση I	54
2.4.1 Υπόβαθρο	54
2.4.2 Αξιολόγηση της σταθερότητας του μέσου της διεργασίας	55
2.4.3 Αξιολόγηση σταθερότητας για τη διακύμανση της διεργασίας	57
2.5 Διαγράμματα Ελέγχου για Ιδιότητες στη Φάση I	58

2.6	Πολυμεταβλητά διαγράμματα ελέγχου στη Φάση I	59
2.7	Μέθοδοι κατανομής στη Φάση I	61
	Κεφάλαιο 3	63
3.1	Εισαγωγή	63
3.2	Τυπικά όρια Ελέγχου – Μειονεκτήματα της τυπικής μεθόδου	64
3.3	Όρια Ελέγχου το διάγραμμα ελέγχου \bar{X} και R	70
3.3.1	1 ^ο Στάδιο Μεθόδου	70
3.3.2	2 ^ο Στάδιο Μεθόδου	73
3.3.3	Υπολογισμός των σταθερών A_2^* , A_2^{**} , D_3^* , D_3^{**} , D_4^* και D_4^{**}	76
3.4	Εφαρμογή	80
3.5	Διαγράμματα Ελέγχου για το μέσο βασισμένα στις ποσότητες \bar{X} και \bar{u}	82
3.6	Διαγράμματα Ελέγχου βασισμένα στην ποσότητα \bar{u}	86
3.7	Διαγράμματα Ελέγχου για το μέσο βασισμένα στις ποσότητες \bar{X} και s	90
3.8	Διαγράμματα Ελέγχου για τη διασπορά βασισμένα στην ποσότητα s	92
	Βιβλιογραφία	95

Κατάλογος Πινάκων

1.1	Βασικές διαστάσεις της έννοιας της ποιότητας	2
3.1	A_2^{**} : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος (1 ^ο στάδιο)	71
3.2	D_3^{**} : Παράγοντες για το κάτω όριο ελέγχου του R διαγράμματος (1 ^ο στάδιο)	72
3.3	D_4^{**} : Παράγοντες για το άνω όριο ελέγχου του R διαγράμματος (1 ^ο στάδιο)	72
3.4	A_2^* : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος (2 ^ο στάδιο)	74
3.5	D_3^{**} : Παράγοντες για το κάτω όριο ελέγχου του R διαγράμματος (2 ^ο στάδιο)	75
3.6	D_4^{**} : Παράγοντες για το άνω όριο ελέγχου του R διαγράμματος (2 ^ο στάδιο)	75
3.7	A_4^* : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος ($n = 5$)	85
3.8	A_4^{**} : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος ($n = 5$)	85
3.9	$B_{7,\sqrt{B_7^*}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$)	89
3.10	$B_{8,\sqrt{B_8^*}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$)	89
3.11	$B_{7,\sqrt{B_7^{**}}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$)	90
3.12	$B_{8,\sqrt{B_8^{**}}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$)	90
3.13	Τιμές του A_5^* για $n = 5$	92
3.14	Τιμές των B_9^* (ή $\sqrt{B_9^*}$) για $n = 5$	94
3.15	Τιμές των B_{10}^* (ή $\sqrt{B_{10}^*}$) για $n = 5$	94

Κατάλογος Διαγραμμάτων

1.1	Βασική μορφή Διαγράμματος Αιτίας - Αποτελέσματος	5
1.2	Διάγραμμα Pareto για τις αιτίες καθυστέρησης των ταχυδρομικών αποστολών	6
1.3	Scatter Plot/ παράδειγμα συσχέτισης μεταξύ μεταβλητών	7
1.4	Scatter Plot/ παράδειγμα μη συσχετιζόμενων μεταβλητών	7
1.5	Παράδειγμα ενός Defect Concentration Diagram που παρουσιάζει τα ελαττώματα που εμφανίζονται στο τελικό στάδιο παραγωγής ενός πουκαμίσου	11
1.6	Διάγραμμα Ελέγχου για διαδικασία εντός ελέγχου	14
3.1	Παράδειγμα διαγράμματος ελέγχου	63
3.2	Γραφικός προσδιορισμός του C	69

Κατάλογος Εικόνων

1	Παρουσίαση της μηχανής χαρτιού (Shainin (2008))	37
2	Παρουσίαση των σημείων δειγματοληψίας για το Multivari διάγραμμα (Shainin (2008))	38
3	Παρουσίαση του Multivari διαγράμματος για την παραγωγή χαρτιού (Shainin (2008))	38
4	Διάγραμμα Ελέγχου Φάσης I για Case U	43
5	Μέθοδος αλλαγής σημείου	51

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Δεν είναι λίγες οι φορές που με την αγορά ενός προϊόντος εντοπίζουμε ελαττώματα σε αυτό που σχετίζονται είτε με την λειτουργικότητά του, είτε με κάποια διαφοροποίηση στην κατασκευή του. Ως καταναλωτές αναμένουμε το προϊόν που αγοράζουμε να βρίσκεται στη βέλτιστη κατάσταση και να ανταποκρίνεται στις λειτουργίες για τις οποίες κατασκευάστηκε. Οι κατασκευαστές από την πλευρά τους, γνωρίζουν ότι δεν είναι πάντα εφικτό να εξετάζουν κάθε προϊόν ξεχωριστά σε κάθε στάδιο της παραγωγής. Ο τρόπος για να επιτύχουμε τη βέλτιστη (ποιοτικά) κατάσταση ενός προϊόντος είναι με κάποιο τρόπο να το ελέγξουμε κατά την κατασκευή του.

Ο Στατιστικός Έλεγχος Ποιότητας (Statistical Quality Control) αποτελεί την παλαιότερη και γνωστότερη μέθοδο ελέγχου παραγωγικών διεργασιών για τη βελτίωση της ποιότητας των παραγόμενων προϊόντων. Ένας από τους βασικούς στόχους του είναι η έγκαιρη ανακάλυψη μη συμμορφωμένων με τις προδιαγραφές παραγόμενων προϊόντων η οποία σηματοδοτεί τη λήψη διορθωτικών ενεργειών για την απομάκρυνση των αιτιών που είναι υπεύθυνες για τις αποκλίσεις, συμβάλλοντας έτσι στη διατήρηση της ποιότητας των προϊόντων. Θα μπορούσε κανείς να πει πως ο Στατιστικός Έλεγχος Ποιότητας επηρεάζει σημαντικά αποφάσεις σχετικές με τις προδιαγραφές, την παραγωγή και τον έλεγχο των παραγόμενων προϊόντων μιας επιχείρησης.

Το κεφάλαιο που ακολουθεί αποτελεί εισαγωγή στη μέθοδο του Στατιστικού Ελέγχου Ποιότητας (Statistical Quality Control). Μέσω μια σύντομης αναδρομής στις βασικές έννοιες και τεχνικές, ο αναγνώστης θα μπορέσει να κατανοήσει τον τρόπο λειτουργίας αλλά την αποτελεσματικότητα αυτής της μεθόδου.

1.1 Η έννοια της ποιότητας και οι διαστάσεις της

Η έννοια της ποιότητας μπορεί να οριστεί με πολλούς τρόπους. Οι περισσότεροι άνθρωποι έχουν μια διαισθητική άποψη για το τι είναι ποιότητα και τη συνδέουν με τα επιθυμητά χαρακτηριστικά που πρέπει να έχει ένα προϊόν (ή μια υπηρεσία). Τα χαρακτηριστικά αυτά, βάσει των οποίων ο καταναλωτής αξιολογεί τα προϊόντα, καθορίζουν την τελική ποιότητα ενός προϊόντος και σχετίζονται με την ποιότητα του σχεδιασμού και τη διαδικασία

παραγωγής του προϊόντος. Οι βασικές διαστάσεις της ποιότητας παρουσιάζονται αναλυτικά στον ακόλουθο πίνακα

Απόδοση	Το προϊόν κάνει τη δουλειά για την οποία έχει κατασκευασθεί; Κάνει τη δουλειά καλύτερα από άλλα ομοειδή προϊόντα;
Αξιοπιστία	Το προϊόν χρειάζεται συχνή επισκευή;
Διάρκεια	Η διάρκεια ζωής του προϊόντος είναι μεγάλη;
Επισκευή	Πόσο γρήγορη και οικονομική είναι η επισκευή του προϊόντος στην περίπτωση που εμφανιστεί βλάβη;
Αισθητική	Πόσο ικανοποιητικό είναι από άποψη εμφάνισης (χρώμα, σχήμα, περιτύλιγμα κτλ) το προϊόν;
Δυνατότητες	Ποιες είναι οι επιπρόσθετες ικανότητες του προϊόντος;
Φήμη επιχείρησης	Ποια είναι η φήμη της εταιρείας, καλή ή κακή;
Συμμόρφωση με τις προδιαγραφές	Το προϊόν κατασκευάστηκε σύμφωνα με τις προδιαγραφές που έθεσε ο σχεδιαστής του;

Πίνακας 1.1 Βασικές διαστάσεις της έννοιας της ποιότητας

Σύμφωνα με όσα αναφέρθηκαν για τις διαστάσεις της ποιότητας μπορούμε να ορίσουμε ότι η ποιότητα είναι η προσαρμογή των χαρακτηριστικών ενός προϊόντος στις απαιτήσεις του καταναλωτή, δηλαδή ότι ποιότητα σημαίνει καταλληλότητα προς χρήση (ορισμός ποιότητας, Joseph Juran, 1974). Η ποιότητα, λοιπόν, μπορεί να μετρηθεί μέσω της ικανοποίησης που προσφέρει το προϊόν στον καταναλωτή κατά τη χρήση του. Δυστυχώς ο παραπάνω ορισμός έχει συσχετιστεί περισσότερο με την έννοια της συμμόρφωσης των παραγόμενων προϊόντων με προδιαγραφές που πρέπει να πληρούν τα παραγόμενα προϊόντα, και λιγότερο με την έννοια της ποιότητας του σχεδιασμού τους.

1.2 Τι είναι ο Στατιστικός Έλεγχος Ποιότητας

Το να βρίσκουμε τρόπους για την επίλυση προβλημάτων που αφορούν την ποιότητα των παραγόμενων προϊόντων δεν είναι αρκετό. Χρειαζόμαστε εργαλεία τα οποία θα μας δώσουν τη δυνατότητα να πάρουμε τις σωστές αποφάσεις ώστε να μειώσουμε την πιθανότητα παραγωγής ελαττωματικών προϊόντων. Τέτοιου είδους εργαλεία προέρχονται κυρίως από το

χώρο της στατιστικής και μας βοηθούν να ταυτοποιήσουμε προβλήματα που αφορούν την ποιότητα των προϊόντων σε μια παραγωγική διεργασία.

Ο Στατιστικός Έλεγχος Ποιότητας (Statistical Quality Control (SQC)) είναι από τις βασικότερες στατιστικές μεθόδους για τη βελτίωση της ποιότητας των παραγόμενων προϊόντων. Μπορεί να χωριστεί σε τρεις ευρείες κατηγορίες:

1. Περιγραφική Στατιστική (Descriptive Statistics) : Χρησιμοποιείται για να περιγράψει ποσοτικά χαρακτηριστικά και συσχετίσεις (π.χ. μέσος, τυπική απόκλιση, εύρος κτλ.)
2. Στατιστικός Έλεγχος Διεργασιών (Statistical Process Control): Περιλαμβάνει την εξέταση ενός τυχαίου δείγματος από μια διεργασία και την απόφαση για το αν τα παραγόμενα προϊόντα έχουν χαρακτηριστικά τα οποία πέφτουν έξω από κάποιο προκαθορισμένο εύρος. Ο Στατιστικός Έλεγχος Διεργασιών είναι και αυτός που απαντά στο ερώτημα αν η παραγωγή λειτουργεί στο βέλτιστο βαθμό.
3. Δειγματοληψία Αποδοχής (Acceptance Sampling): Περιέχει στατιστικές τεχνικές (δειγματοληπτικές) που είναι απαραίτητες για να αποφασίσουμε αν μια συγκεκριμένη παρτίδα (σωρός) προϊόντων θα γίνει δεκτή ή θα απορριφθεί.

Κάθε μια από τις παραπάνω κατηγορίες μας δίνει διαφορετικές πληροφορίες για την ποιότητα των προϊόντων που αναλύει. Η περιγραφική στατιστική χρησιμοποιείται για να περιγράψει ποσοτικά χαρακτηριστικά όπως για παράδειγμα τη κεντρική τάση και τη μεταβλητότητα του εξεταζόμενου δείγματος. Παρά τη χρησιμότητά τους τα μέτρα αυτά δεν επαρκούν για να εκτιμήσουμε την ύπαρξη ποιοτικών προβλημάτων. Αντίθετα, η δειγματοληψία αποδοχής μπορεί να βοηθήσει σε αυτό. Δίνει τη δυνατότητα να ελέγξουμε αν έχει επιτευχθεί η επιθυμητή ποιότητα για μια συγκεκριμένη παρτίδα ή αν θα πρέπει να απορρίψουμε τα προϊόντα αυτά. Για άλλη μια φορά όμως, παρόλο που η πληροφορία που λαμβάνουμε είναι χρήσιμη για την αποδοχή των προϊόντων μετά την παραγωγή τους, δεν μας βοηθάει στο να έχουμε μια εικόνα της ποιοτικής τους κατάστασης κατά τη διάρκεια της παραγωγικής διεργασίας. Για αυτό το λόγο χρειαζόμαστε τον Στατιστικό Έλεγχο Διεργασιών.

Και οι τρεις προαναφερθείσες κατηγορίες του στατιστικού ελέγχου ποιότητας είναι χρήσιμες για τη μέτρηση και εκτίμηση της ποιότητας των προϊόντων ή υπηρεσιών. Ο Στατιστικός Έλεγχος Διεργασιών όμως χρησιμοποιείται συχνότερα λόγω της δυνατότητάς του να εντοπίζει το πρόβλημα κατά τη διάρκεια της παραγωγικής διεργασίας. Σε επόμενη ενότητα θα δώσουμε μια αναλυτική εικόνα αυτής της μεθόδου.

1.3 Στατιστικός έλεγχος διεργασιών

Η προπαρασκευαστική φάση του Στατιστικού Ελέγχου Διεργασιών περιλαμβάνει μια συλλογή εργαλείων που είναι χρήσιμα για την καλύτερη κατανόηση και βελτίωση της διεργασίας. Τα επτά διαθέσιμα εργαλεία που χρησιμοποιούνται είναι τα εξής:

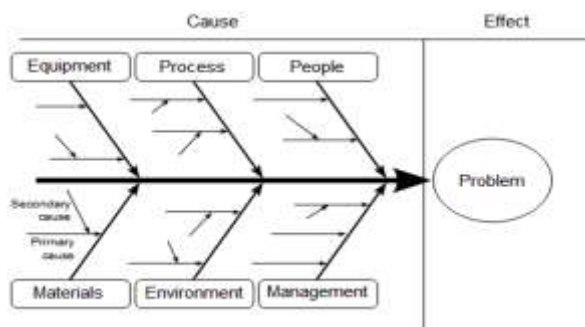
- το Φύλλο Ελέγχου (Check Sheet)
- το Διάγραμμα Αιτίας –Αποτελέσματος (Cause- and- Effect Diagram)
- το Διάγραμμα Pareto (Pareto Chart)
- το Διάγραμμα Διασποράς (Scatter Plot)
- το Ιστόγραμμα, Διάγραμμα Μίσχου-Φύλλων (Histogram, Steam-and-Leaf Plot)
- τα Διαγράμματα Ελέγχου (Control Charts)
- το Διάγραμμα Συγκέντρωσης Ελαττωμάτων (Defect Concentration Diagram)

Το φύλλο ελέγχου (Check Sheet)

Το φύλλο ελέγχου είναι ένα απλό εργαλείο συλλογής και καταγραφής στοιχείων. Σε ένα φύλλο ελέγχου μπορούν να καταχωρηθούν λεπτομέρειες σχετικές με το είδος και το χρόνο των σφαλμάτων που εμφανίζουν τα προϊόντα, να καταγραφούν οι απαραίτητοι έλεγχοι που πρέπει να γίνουν κατά την παραλαβή μιας παρτίδας υλικών ή κατά την επιθεώρηση ενός συστήματος ποιότητας κτλ. Επιπλέον στο φύλλο ελέγχου καταγράφονται διάφορα συμπληρωματικά στοιχεία σχετικά με το είδος των δεδομένων που καταχωρούνται, όπως η μέρα, η παρτίδα, η βάρδια, και γενικά κάθε πληροφορία που αφορά τη διεργασία. Τέλος το φύλλο ελέγχου μπορεί να χρησιμοποιηθεί και ως φύλλο εργασίας για την εισαγωγή των δεδομένων σε υπολογιστή.

Το διάγραμμα Αιτίας-Αποτελέσματος (Cause- and- Effect Diagram)

Το διάγραμμα Αιτίας-Αποτελέσματος χρησιμοποιείται για να ανιχνεύσει πιθανές αιτίες ενός συγκεκριμένου προβλήματος. Αποτελεί μια απεικόνιση των ποικίλων στοιχείων ενός συστήματος (αίτια) τα οποία συνεισφέρουν στη δημιουργία ενός προβλήματος (αποτέλεσμα). Το διάγραμμα αιτίου –αποτελέσματος αναφέρεται και σαν διάγραμμα Ishikawa ή σαν διάγραμμα Fishbone. Το αποτέλεσμα απεικονίζεται σαν το κεφάλι μιας ραχοκοκαλιάς ψαριού και οι πρωταρχικοί παράγοντες (αίτια) απεικονίζονται στα άκρα των πλευρών της. Τα πρωταρχικά αίτια είναι συνήθως οι άνθρωποι (χειριστές μηχανών), τα υλικά (πρώτες ύλες), το περιβάλλον, η μέθοδος και οι μηχανές.



Διάγραμμα 1.1 Βασική Μορφή Διαγράμματος Αιτίας Αποτελέσματος (Fishbone or Ishikawa Diagram)

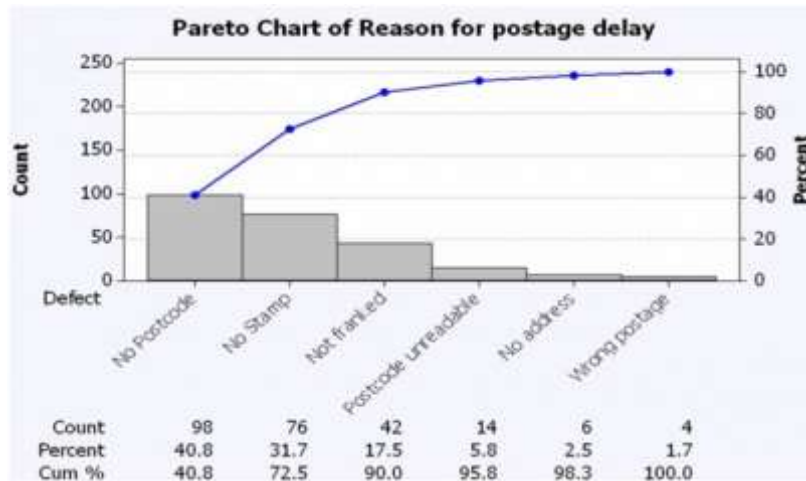
Ανάλογα με τη φύση του προβλήματος μπορούν να προστεθούν ή να αφαιρεθούν και άλλα αίτια. Τα αίτια μπορούν να αναπτυχθούν σε περισσότερα επίπεδα (υπο-αίτια). Τα βασικά βήματα για την κατασκευή και χρήση του διαγράμματος αιτίου –αποτελέσματος είναι τα ακόλουθα:

1. Προσδιορισμός του προβλήματος ή του αποτελέσματος το οποίο πρέπει να αναλυθεί
2. Σύνθεση ομάδας για την αντιμετώπιση του προβλήματος
3. Σχεδίαση της κεντρικής γραμμής και του περιγράμματος του αποτελέσματος
4. Κατασκευή του σκελετού του διαγράμματος μέσω του προσδιορισμού (αναγνώρισης) των πρωταρχικών αιτιών και σύνδεσή τους με την κεντρική γραμμή.
5. Αναγνώριση και καταγραφή των υπο-αιτιών για κάθε κύρια κατηγορία αιτιών
6. Διάταξη των υπο-αιτιών ανάλογα με το ποιες φαίνονται να επηρεάζουν περισσότερο το αποτέλεσμα
7. Ανάλυση πρωτοβουλιών για την αντιμετώπιση του προβλήματος

Το διάγραμμα Pareto (Pareto Diagram)

Το διάγραμμα Pareto οφείλει το όνομά του στον Ιταλό οικονομολόγο Vilfredo Pareto ο οποίος είναι κυρίως γνωστός για την παρατήρηση του ότι η κατανομή του 80%-90% του πλούτου της χώρας του ήταν συγκεντρωμένη στο 20%-10% του πληθυσμού (κανόνας 80/20). Η ίδια συμπεριφορά παρατηρήθηκε και σε άλλους τομείς, όπως στον Στατιστικό Έλεγχο Ποιότητας, όπου φαίνεται να ισχύει ότι το 80% των ελαττωμάτων στα προϊόντα οφείλεται στο 20% των αιτιών. Έτσι η ανάλυση Pareto έχει ως σκοπό να διαχωρίσει τις σημαντικές πλευρές ενός προβλήματος από τις λιγότερο σημαντικές.

Στο ακόλουθο διάγραμμα παρουσιάζεται ένα παράδειγμα διαγράμματος Pareto, το οποίο σχετίζεται με τις αιτίες οι οποίες προκαλούν καθυστέρηση παράδοσης των ταχυδρομικών αποστολών.



Διαγραμμα 1.2 Διαγραμμα Pareto για τις αιτίες καθυστέρησης των ταχυδρομικών αποστολών

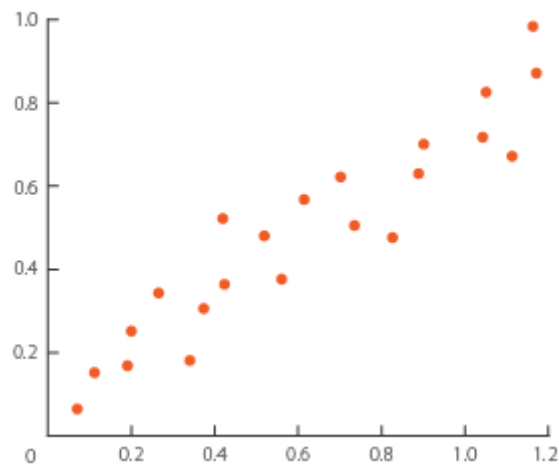
Από το παραπάνω διάγραμμα Pareto προκύπτουν ως σημαντικά αίτια καθυστέρησης μιας ταχυδρομικής αποστολής η έλλειψη ταχυδρομικού κώδικα αλλά και γραμματοσήμου.

Τα βασικά βήματα που ακολουθούνται για την ανάλυση Pareto είναι τα ακόλουθα:

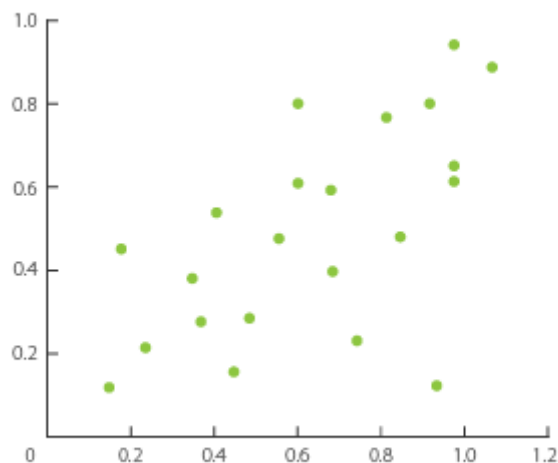
- Καταγραφή όλων των στοιχείων
- Μέτρηση των στοιχείων
- Διάταξη των στοιχείων
- Δημιουργία αθροιστικών κατανομών
- Σχεδίαση διαγράμματος Pareto
- Ερμηνεία του διαγράμματος Pareto

Το Διάγραμμα Διασποράς (Scatter Plot)

Το διάγραμμα διασποράς χρησιμοποιείται για να αποκαλύψει αν υπάρχει σχέση μεταξύ δύο μεταβλητών. Είναι ένα γράφημα που απεικονίζει τις τιμές δύο μεταβλητών με τη μορφή σημείων σε ένα επίπεδο. Αν οι μεταβλητές είναι συσχετισμένες τότε τα σημεία ακολουθούν την πορεία μιας γραμμής ή καμπύλης. Όσο πιο έντονη είναι η συσχέτιση τόσο πιο πολύ τα σημεία «αγκαλιάζουν» τη γραμμή.



Διάγραμμα 1.3 ScatterPlot / παράδειγμα συσχέτισης μεταξύ μεταβλητών



Διάγραμμα 1.4 ScatterPlot/ παράδειγμα μη συσχετισμένων μεταβλητών

Αν διαπιστωθεί ότι υπάρχει συσχέτιση μεταξύ δύο μεταβλητών τότε μπορούμε να επηρεάσουμε την τιμή της μιας μεταβλητής έτσι ώστε να επιτύχουμε μια επιθυμητή τιμή για την άλλη μεταβλητή.

Το Ιστόγραμμα, Διάγραμμα Μίσχου-Φύλλων (Histogram, Steam-and-Leaf Plot)

Το ιστόγραμμα αποτελεί τη γραφική απεικόνιση ενός πίνακα συχνοτήτων και αποκαλύπτει σημαντικές πληροφορίες για το σύνολο των τιμών που αντιπροσωπεύει όπως είναι η κεντρική τάση, η μεταβλητότητα και το σχήμα της κατανομής των τιμών. Για τον τρόπο κατασκευής των πινάκων συχνοτήτων ο αναγνώστης παραπέμπεται σε βιβλία περιγραφικής στατιστικής.

Ωστόσο στη συνέχεια κρίνεται σκόπιμο να υπενθυμίσουμε ορισμένα βασικά στοιχεία κατασκευής ενός πίνακα συχνοτήτων.

Ο καθορισμός του πλήθους των τάξεων σε ένα πίνακα συχνοτήτων (οπότε και του πλήθους των ιστών- ορθογωνίων του αντίστοιχου ιστογράμματος) έγκειται στη κρίση του αναλυτή. Ο αριθμός των τάξεων που χρησιμοποιούνται στην πράξη σε σχέση με τον αριθμό των μετρήσεων που είχαμε στη διάθεσή μας δίνεται στον ακόλουθο πίνακα

Αριθμός μετρήσεων	Αριθμός τάξεων (C)
<50	5-6
50-100	7-9
100-150	10-11
150-200	12-13
>200	14

Αφού επιλεγεί ο αριθμός των τάξεων που θα χρησιμοποιηθούν το επόμενο βήμα είναι ο καθορισμός του πλάτους W κάθε τάξης που υπολογίζεται από τον τύπο

$$W = \frac{R}{C} = \frac{\text{εύρος μετρήσεων}}{\text{αριθμός τάξεων}}$$

(στην περίπτωση που χρησιμοποιούνται ίσα πλάτη τάξεων). Το πλάτος W μπορεί να αυξηθεί έτσι ώστε να προκύψουν εύχρηστα όρια τάξεων. Τέλος η κατασκευή του πίνακα συχνοτήτων μπορεί να ολοκληρωθεί αφού επιλεγθεί κατάλληλα το κατώτερο όριο L_1 της πρώτης τάξης.

Το διάγραμμα μίσχου –φύλλων είναι ένα είδος ιστογράμματος όπου εμφανίζονται και οι τιμές των μετρήσεων.

Επειδή η μορφή του ιστογράμματος μπορεί να μας δώσει ενδιαφέρουσες πληροφορίες για αυτά, στη συνέχεια θα δούμε τις βασικότερες μορφές ιστογραμμάτων που συναντώνται στην πράξη:

Normal Distribution



Σε μια κανονική κατανομή τα σημεία έχουν την ίδια πιθανότητα να πέσουν σε οποιαδήποτε μεριά (αριστερά ή δεξιά) του μέσου με αποτέλεσμα να προκύπτει συμμετρική κατανομή.

Right-skewed Distribution



Λοξές κατανομές (δεξιά ή αριστερά) προκύπτουν συνήθως όταν υπάρχει ένα φυσικό όριο που εμποδίζει να εμφανιστούν τιμές πέρα από αυτό. Έτσι η κορυφή της κατανομής είναι προς την μεριά του φυσικού ορίου και έχει ουρά που απλώνεται μακριά από αυτή.

Binomial (double-peaked) Distribution



Κατανομές με διπλή κορυφή εμφανίζονται όταν μετρήσεις δύο διαδικασιών με διαφορετικές κατανομές συνδυάζονται σε ένα δείγμα.

Plateau Distribution



Κατανομές με πολλές κορυφές εμφανίζονται όταν μετρήσεις διαδικασιών από διαφορετικές «κανονικές» κατανομές συνδυάζονται σε ένα δείγμα.

Comb Distribution



Μοιάζει με κανονική κατανομή αλλά έχει μια υψηλή κορυφή σε μια ουρά. Αυτό οφείλεται συνήθως σε μετρήσεις προς τη μεριά μιας ουράς που έχουν συσσωρευτεί σε μια τάξη με μετρήσεις της μορφής «μεγαλύτερες από...» (ή «μικρότερες από...»).

Truncated or heart-cut Distribution



Η περικομμένη κατανομή μοιάζει με κανονική κατανομή που λείπουν οι δυο ουρές της. Μπορεί να εμφανιστεί σε περιπτώσεις που ο προμηθευτής ενός προϊόντος ξεχωρίζει τα προϊόντα που βρίσκονται εντός των ορίων προδιαγραφών από αυτά που δεν είναι.

Dog food Distribution



Τέτοιου είδους κατανομές προκύπτουν όταν λείπει κάτι που βρίσκεται κοντά στη μέση τιμή και φαίνεται να είναι το κομμάτι που «περισσεύει» από μια περικομμένη κατανομή.

Τα Διαγράμματα Ελέγχου (Control Charts)

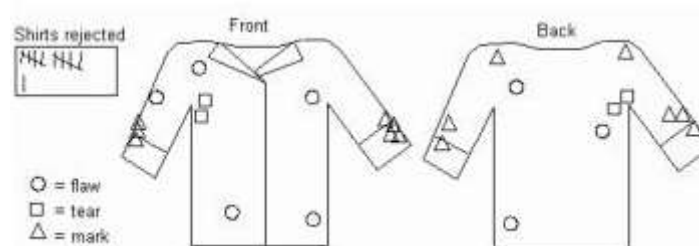
Ένα διάγραμμα ελέγχου είναι μια χρονοσειρά σημαντικών χαρακτηριστικών ενός προϊόντος (ή μιας διαδικασίας) τα οποία λαμβάνονται από τη διαδικασία κατά περιοδικά διαστήματα.

Ένα τέτοιο διάγραμμα είναι:

- Μια γραφική μέθοδος για να καθορίσει αν μια διαδικασία είναι σταθερή.
- Ένα εργαλείο διάκρισης συνήθους μεταβλητότητας από ασυνήθιστες αιτίες.
- Ένας «συναγερμός» που ειδοποιεί για την ανάγκη ανάλυσης των αιτιών των προβλημάτων.
- Μια βοήθεια για την κατανόηση και βελτίωση της διαδικασίας. Η αναγνώριση και η εξάλειψη των πηγών των ασυνήθιστων γεγονότων, βοηθά στη σταθεροποίηση και βελτίωση μιας διαδικασίας. Τα διαγράμματα ελέγχου βοηθούν μια διαδικασία να φτάσει στο καλύτερο δυνατό επίπεδο χωρίς την μεταβολή της βασικής δομής.

Το Διάγραμμα Συγκέντρωσης Ελαττωμάτων (Defect Concentration Diagram)

Το διάγραμμα συγκέντρωσης ελαττωμάτων έχει ως σκοπό να απεικονίσει τις τοποθεσίες που εμφανίζονται τα διάφορα ελαττώματα σε ένα προϊόν. Είναι μια εικόνα όλων των εξωτερικών όψεων του προϊόντος όπου σημειώνονται οι θέσεις (περιοχές) που εμφανίζονται τα ελαττώματα. Έτσι βοηθάει στην ανίχνευση ειδικών αιτιών της παραγωγικής διαδικασίας που είναι υπεύθυνες για τη δημιουργία αυτών των ελαττωμάτων. Η θέση που παρατηρείται το ελάττωμα είτε χρωματίζεται (συνήθως τα διάφορα είδη ελαττωμάτων απεικονίζονται με διαφορετικό χρώμα) είτε σημειώνεται πάνω της ένας κωδικός (διαφορετικοί κωδικοί για κάθε είδος ελαττώματος).



Διάγραμμα 1.5 Παράδειγμα ενός Defect Concentration Diagram που παρουσιάζει τα ελαττώματα που εμφανίζονται στο τελικό στάδιο παραγωγής ενός πουκαμίσου

1.4 Το πρόβλημα του Στατιστικού Ελέγχου Διεργασιών

Σε κάθε παραγωγική διεργασία, θα υπάρχει πάντα μια μορφή φυσικής μεταβλητότητας, ανεξάρτητα από το πόσο καλά είναι σχεδιασμένη και από το πόσο προσεκτικά επιβλέπεται. Η

φυσική αυτή μεταβλητότητα αποτελείται από πολλές μικρές αιτίες οι οποίες αναφέρονται στη βιβλιογραφία ως κοινές ή τυχαίες αιτίες μεταβλητότητας (common or chance causes of variation). Μια διεργασία που λειτουργεί μόνο με την παρουσία φυσικής μεταβλητότητας καλείται εντός ελέγχου διεργασία ή λέμε ότι λειτουργεί σε ευσταθή κατάσταση (stable state).

Αντίθετα, μια παραγωγική διεργασία μπορεί κατά καιρούς να εμφανίσει και άλλες μορφές μεταβλητότητας οι οποίες δεν οφείλονται σε τυχαίες αιτίες όπως περιγράφηκε παραπάνω, αλλά προκαλούν συστηματική αλλαγή στο επίπεδο κάποιων παραγόντων και καθορίζουν την ποιότητα του προϊόντος. Τέτοιου είδους μεταβλητότητα μπορεί να οφείλεται για παράδειγμα σε λανθασμένες ρυθμίσεις, λάθη από τους χειριστές της μηχανής, κακής ποιότητας πρώτη ύλη ή ελαττωματικές πρώτες ύλες κτλ. Τέτοιου είδους μεταβλητότητα είναι σοβαρότερης μορφής καθώς η παρουσία της συνήθως οδηγεί σε μη αποδεκτά επίπεδα λειτουργίας της παραγωγικής διεργασίας. Αυτή η μεταβλητότητα αναφέρεται ως ειδική μεταβλητότητα και οι αιτίες που οδηγούν σε αυτή ονομάζονται ειδικές αιτίες μεταβλητότητας (special or assignable causes of variation). Μια διεργασία που λειτουργεί υπό την παρουσία ειδικής μεταβλητότητας λέμε ότι είναι εκτός στατιστικού ελέγχου διεργασία (out of control) ή ότι λειτουργεί σε ασταθή κατάσταση (unstable state).

Η παραγωγή ενός προϊόντος συνδέεται άμεσα με τα όρια προδιαγραφών (specification limits) των ποιοτικών χαρακτηριστικών (quality characteristics) του προϊόντος τα οποία καθορίζονται στη φάση του σχεδιασμού του. Πιο συγκεκριμένα ανάμεσα στο κάτω και το άνω όριο προδιαγραφών (Lower and Upper specification limits, LSL and USL) και εντός αυτών των ορίων πρέπει να βρίσκονται οι τιμές του ποιοτικού χαρακτηριστικού για να θεωρηθεί ποιοτικά αποδεκτό. Στη φάση αυτή καθορίζεται ένα επιπλέον ποιοτικό χαρακτηριστικό που καλείται τιμή στόχος (target value T) που είναι συνήθως το μέσο του διαστήματος [LSL,USL]. Όπως συμπεραίνουμε από τα παραπάνω, υπό συνθήκες φυσικής μεταβλητότητας οι τιμές του ποιοτικού χαρακτηριστικού θα βρίσκονται εντός των ορίων προδιαγραφών, ενώ αντίθετα υπό συνθήκες ειδικής μεταβλητότητας θα βρίσκονται εκτός.

Ο κύριος στόχος του Στατιστικού Ελέγχου Διεργασιών είναι η όσο το δυνατόν πιο έγκαιρη ανίχνευση της εμφάνισης των ειδικών αιτιών μεταβλητότητας έτσι ώστε να προβούμε στις απαραίτητες ενέργειες προκειμένου να αποφύγουμε την κατασκευή επιπλέον μη συμμορφούμενων (non-conforming) προϊόντων. Ένας αποτελεσματικός τρόπος ανίχνευσης των ειδικών αιτιών μεταβλητότητας σε πραγματικό χρόνο είναι τα διαγράμματα ελέγχου (control charts).

Για να κατανοήσουμε καλύτερα τη χρησιμότητα των διαγραμμάτων ελέγχου αρκεί να αναφερθούμε σε δύο έννοιες που είναι άρρηκτα συνδεδεμένες με αυτά. Με τη χρήση του όρου μη συμμορφούμενο ή ελαττωματικό προϊόν, αναφερόμαστε στο προϊόν εκείνο που έχει τουλάχιστον ένα ποιοτικό χαρακτηριστικό με τιμή εκτός των ορίων προδιαγραφών. Ανάλογα με τον αριθμό των ελαττωμάτων (ατελειών) και τη σοβαρότητα τους μπορεί να καθοριστεί αν ένα προϊόν θα τεθεί ή όχι σε πώληση. Υπάρχει επίσης η δυνατότητα κατασκευής διαγραμμάτων ελέγχου για την παρακολούθηση του αριθμού των ελαττωματικών προϊόντων ή ακόμα και για τον αριθμό των ελαττωμάτων ενός προϊόντος. Τέτοια ποιοτικά χαρακτηριστικά δεν μετρώνται σε μια συνεχή κλίμακα, αλλά παίρνουν αριθμήσιμο πλήθος τιμών και περιγράφονται με διακριτές τυχαίες μεταβλητές (attributes).

1.5 Γενικές αρχές διαγραμμάτων ελέγχου

Από στατιστικής πλευράς, τα διαγράμματα ελέγχου (control charts) αποτελούν γραφική παράσταση της πορείας μιας εκτιμήτριας σε συνάρτηση με τον χρόνο ή τον αριθμό δείγματος. Η εκτιμήτρια αφορά σε ορισμένη παράμετρο της κατανομής του ποιοτικού χαρακτηριστικού και υπολογίζεται από τα τυχαία δείγματα που λαμβάνονται κατά διαστήματα από την παραγωγική διαδικασία. Αν το χαρακτηριστικό ποιότητας είναι απαριθμήσιμη εκτιμήτρια αφορά το ποσοστό ελαττωματικών ή τον αριθμό των ελαττωματικών, αν το χαρακτηριστικό μέτρησης αφορά τη θέση ή τη διασπορά της κατανομής. Στην ορολογία του στατιστικού ελέγχου παραγωγικής διαδικασίας η εκτιμήτρια ονομάζεται στατιστική δείγματος (sample statistics).

Η στατιστική δείγματος αφορά:

- σε χαρακτηριστικά μέτρησης όπως η μέση τιμή δείγματος, η τυπική απόκλιση δείγματος ή το εύρος τιμών του δείγματος,
- σε χαρακτηριστικά απαριθμήσιμα όπως το ποσοστό ή ο αριθμός ελαττωματικών στο δείγμα, ο αριθμός ελαττωμάτων στο δείγμα ή ο αριθμός ελαττωμάτων ανά μονάδα δείγματος.

Στο σημείο αυτό, αξίζει να αναφέρουμε πως τα διαγράμματα ελέγχου χωρίζονται σε δυο κατηγορίες:

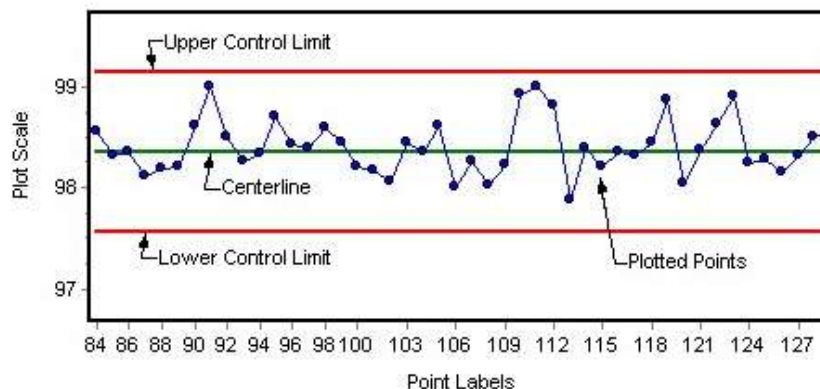
- *Διαγράμματα ελέγχου μεταβλητών* (control charts for variables) τα οποία χρησιμοποιούνται όταν τα χαρακτηριστικά που μας ενδιαφέρουν είναι συνεχή και μετρήσιμα (π.χ. διαστάσεις, θερμοκρασία, πίεση, βάρος, ταχύτητα κτλ.). Κυριότεροι εκπρόσωποι αυτών των

διαγραμμάτων είναι τα διαγράμματα μέσου όρου και εύρους ($\bar{X} - \text{chart}$, $R - \text{chart}$) και το διάγραμμα τυπικής απόκλισης ($s - \text{chart}$).

- *Διαγράμματα ελέγχου χαρακτηριστικών* (control charts for attributes) τα οποία χρησιμοποιούνται όταν τα δεδομένα είναι διακριτά και μη μετρήσιμα (π.χ. αποδεκτό-απορριπτό). Κυριότεροι εκπρόσωποι είναι τα διαγράμματα p , np , c και u .

Σε επόμενη ενότητα θα αναφερθούμε αναλυτικότερα σε αυτά τα διαγράμματα ελέγχου.

Ανεξάρτητα από την κατηγορία στη οποία ανήκουν τα διαγράμματα έχουν κοινή μορφή. Η κεντρική γραμμή (center line, CL) του διαγράμματος αντιστοιχεί στη μέση τιμή της παραμέτρου (π.χ. ποσοστό ελαττωμάτων, μέση τιμή) όταν η διαδικασία βρίσκεται σε στατιστικό έλεγχο. Εκφράζει δηλαδή την κατάσταση ομαλής λειτουργίας, όπου η μεταβλητότητα οφείλεται μόνο σε τυχαίες αιτίες. Οι άλλες δύο οριζόντιες γραμμές είναι τα όρια ελέγχου (control limits) και ειδικότερα το άνω όριο ελέγχου (upper control limit, UCL) και το κάτω όριο ελέγχου (lower control limit, LCL). Από τη στιγμή που τα σημεία του διαγράμματος, δηλαδή οι διαδοχικές τιμές της στατιστικής του δείγματος, βρίσκονται μεταξύ των ορίων ελέγχου και συμπεριφέρονται τυχαία, η διαδικασία θεωρείται ότι βρίσκεται εντός στατιστικού ελέγχου.



Διάγραμμα 1.6 Διάγραμμα ελέγχου για διαδικασία εντός ελέγχου.

Όταν ένα τουλάχιστον σημείο του διαγράμματος είναι εκτός των ορίων ελέγχου, η απόστασή του από την κεντρική γραμμή αποτελεί ένδειξη μεταβολής της κατανομής του ποιοτικού χαρακτηριστικού, που οφείλεται σε ειδική αιτία μεταβλητότητας. Συνεπώς, απαιτείται περαιτέρω διερεύνηση και αποκατάσταση του προβλήματος, έτσι ώστε να επέλθει η διαδικασία εντός στατιστικού ελέγχου.

Ο παραπάνω κανόνας λειτουργίας ενός διαγράμματος ελέγχου είναι απλός αλλά όχι πάντα επαρκής. Συχνά, παρουσιάζονται περιπτώσεις όπου όλα τα σημεία του διαγράμματος βρίσκονται εντός των ορίων ελέγχου, αλλά η διάταξή τους να αποκαλύπτει την παρουσία μιας συστηματικής αιτίας. Στα περισσότερα διαγράμματα τα διαδοχικά σημεία ενώνονται μεταξύ τους με ευθύγραμμα τμήματα ώστε να γίνονται εύκολα αντιληπτές παρόμοιες, μη τυχαίες διατάξεις. Όταν η διαδικασία βρίσκεται πραγματικά εντός στατιστικού ελέγχου, τα σημεία του διαγράμματος εμφανίζονται κατά απολύτως τυχαίο τρόπο, με το μεγαλύτερο ποσοστό τους κοντά στην κεντρική γραμμή αλλά και από τις δύο πλευρές της.

Υπάρχουν διάφορες συστηματικές διατάξεις σημείων που εμφανίζονται συχνά σε διαγράμματα ελέγχου και μπορούν να καταταχθούν σε δυο κατηγορίες ανάλογα με τον τύπο της μεταβολής που τις δημιουργεί:

- χρονοσειρές που οφείλονται σε στιγμιαία μεταβολή της τιμής μιας παραμέτρου, η οποία διατηρείται στη συνέχεια σταθερή στη νέα τιμή ,
- χρονοσειρές που οφείλονται σε συνεχή μεταβολή μιας παραμέτρου.

Οι χρονοσειρές που οφείλονται σε στιγμιαία αλλά παραμένουσα μεταβολή μιας παραμέτρου χαρακτηρίζονται από σχετικά μεγάλο αριθμό διαδοχικών σημείων από τη μία μόνο πλευρά της κεντρικής γραμμής.

Στη βιβλιογραφία αναφέρονται πολλοί κανόνες ροών (run rules), με βάση τους οποίους ένα διάγραμμα ελέγχου σηματοδοτεί την πιθανή παρουσία συστηματικής αιτίας μεταβλητότητας. Οι σημαντικότεροι από αυτούς είναι οι εξής:

- 7 (ή 8) συνεχόμενα σημεία είναι από την ίδια πλευρά της κεντρικής γραμμής (επάνω ή κάτω).
- Τουλάχιστον 10 στα 11 συνεχόμενα σημεία είναι από την ίδια πλευρά της κεντρικής γραμμής.
- Τουλάχιστον 2 στα 3 συνεχόμενα σημεία είναι εκτός ενός προειδοποιητικού ορίου δυο τυπικών αποκλίσεων.
- Τουλάχιστον 4 στα 5 συνεχόμενα σημεία απέχουν περισσότερο από μια τυπική απόκλιση από την κεντρική γραμμή στην ίδια κατεύθυνση (επάνω ή κάτω).

Οι χρονοσειρές που οφείλονται σε συνεχή μεταβολή μιας παραμέτρου ή σε άλλες αιτίες χαρακτηρίζονται γενικά από μη τυχαίες διατάξεις των σημείων στο διάγραμμα ελέγχου. Οι κυριότερες μορφές χρονοσειρών που εμφανίζονται σε διαγράμματα ελέγχου παρουσιάζονται και ερμηνεύονται στη συνέχεια.

✓ Τάσεις

Στα διαγράμματα αυτά παρατηρείται συνεχής μεταβολή μιας παραμέτρου προς την ίδια κατεύθυνση. Τα σημεία του διαγράμματος εμφανίζουν ανοδική τάση (ή καθοδική ανάλογα την περίπτωση). Οι τάσεις οφείλονται σε διάφορους λόγους όπως:

- σταδιακή φθορά κάποιου εργαλείου ή άλλου βασικού συστατικού της παραγωγικής διαδικασίας,
- κόπωση εργαζομένου,
- σταδιακή επιδείνωση εξωτερικών συνθηκών (π.χ. αύξηση υγρασίας ή θερμοκρασίας περιβάλλοντος)

Επειδή η εμφάνιση τάσεων αποτελεί συχνό φαινόμενο και είναι σχετικά εύκολο να ποσοτικοποιηθούν, έχουν αναπτυχθεί και χρησιμοποιούνται δύο κύριοι κανόνες με βάση τους οποίους κρίνεται ότι η διαδικασία υπόκειται στην επίδραση συνεχούς συστηματικής μεταβολής:

- 6 συνεχόμενα σημεία σχηματίζουν μια γνησίως μονότονη ακολουθία (αύξουσα ή φθίνουσα),
- σε σύνολο 10 συνεχόμενων σημείων υπάρχει υποσύνολο 8 σημείων που σχηματίζει γνησίως μονότονη ακολουθία (αύξουσα ή φθίνουσα),

✓ Περιοδικότητα-Επαναλαμβανόμενοι κύκλοι

Εάν τα διαδοχικά σημεία ενός διαγράμματος σχηματίζουν επαναλαμβανόμενους κύκλους, προφανώς δεν είναι τυχαίοι και οφείλονται σε κάποια συστηματική αιτία που μεταβάλλει συνεχώς και με περιοδικότητα τις παραμέτρους της κατανομής του χαρακτηριστικού ποιότητας. Ορισμένες από τις πιθανές αιτίες είναι οι εξής:

- περιοδικές διακυμάνσεις περιβαλλοντικών συνθηκών ή συνθηκών παραγωγής, όπως η θερμοκρασία, η ηλεκτρική τάση κτλ
- περιοδική εναλλαγή εργατών ή μηχανών

✓ Άλματα

Όταν τα διαγράμματα ελέγχου παρουσιάζουν κατά διαστήματα άλματα στο μέσο επίπεδο τιμών της στατιστικής δείγματος, που οφείλονται στους εξής λόγους:

- αλλαγή εργαζομένων, μεθόδου παραγωγής, συσκευής ελέγχου ή μέτρησης,
- διαδοχική επεξεργασία παρτίδων πρώτων υλών διαφορετικής ποιότητας.

Αν η συχνότητα των αλμάτων δεν είναι μεγάλη, η ελεγχόμενη παράμετρος της διαδικασίας παραμένει επί αρκετό διάστημα σε επίπεδο διαφορετικό από την κεντρική γραμμή, οπότε οι μεταβολές αναγνωρίζονται σχετικά εύκολα με τη βοήθεια των κανόνων ροών.

✓ Μεγάλο ποσοστό σημείων κοντά στα όρια ελέγχου

Όταν παρατηρούμε διαγράμματα ελέγχου όπου το ποσοστό των σημείων που βρίσκονται κοντά στα όρια ελέγχου είναι σχετικά μεγάλο και μεγαλύτερο από αυτό που αναμένεται όταν η διαδικασία βρίσκεται σε στατιστικό έλεγχο, συνήθως ευθύνονται τα εξής:

- μεγάλες συστηματικές διαφορές στην ποιότητα των υλικών ή των μεθόδων και συσκευών ελέγχου και μέτρησης,
- υπερβολικά συχνές και άσκοπες ρυθμίσεις των μηχανών,
- ταυτόχρονος έλεγχος δυο διαφορετικών διαδικασιών με το ίδιο διάγραμμα ελέγχου.

Προβλήματα αυτής της μορφής εντοπίζονται ιδιαίτερα αποτελεσματικά με σχεδίαση προειδοποιητικών ορίων στο διάγραμμα και εφαρμογή αντίστοιχων κανόνων ελέγχου.

✓ Έλλειψη μεταβλητότητας

Σε αυτή την περίπτωση όλα τα σημεία βρίσκονται αφύσικα κοντά στην κεντρική γραμμή, δεν παρουσιάζουν δηλαδή τη διασπορά που αναμένεται σύμφωνα με τη φυσική μεταβλητότητα της διαδικασίας. Δυο είναι οι συνηθέστερες αιτίες που προκαλούν αυτό το φαινόμενο:

- εσφαλμένη εκτίμηση των παραμέτρων της κατανομής του χαρακτηριστικού ποιότητας με αποτέλεσμα την εσφαλμένη σχεδίαση των ορίων ελέγχου,
- πρόβλημα της συσκευής ελέγχου που έχει ως αποτέλεσμα όλες οι μετρήσεις να έχουν σχεδόν την ίδια τιμή.

Σε αυτό το σημείο, έχοντας παρουσιάσει τις διάφορες μορφές χρονοσειρών, θα πρέπει να τονίσουμε πως κάθε διαδικασία παραγωγής προϊόντων ή παροχής υπηρεσιών έχει τις ιδιαιτερότητές της και είναι αρκετά δύσκολο να απαλλαγεί η διαδικασία από τη συχνή επίδραση συστηματικών αιτιών, ώστε να λειτουργεί επί μεγάλα χρονικά διαστήματα σε στατιστικό έλεγχο. Απαιτείται αρκετός χρόνος και προσπάθεια για να αναλυθεί μια

διαδικασία, να τεθεί σε κατάσταση στατιστικού ελέγχου και να παρακολουθείται αποτελεσματικά με το κατάλληλο διάγραμμα ελέγχου.

Τέλος πρέπει να επισημάνουμε ότι αν κατά τη σχεδίαση των διαγραμμάτων ελέγχου διαπιστωθεί ότι κάποιο δείγμα βρίσκεται εκτός ορίων ελέγχου για συγκεκριμένους λόγους που εντοπίστηκαν και δεν πρόκειται να επαναληφθούν στο μέλλον, το δείγμα αυτό θα πρέπει να αφαιρεθεί και να μην ληφθεί υπόψη στη στατιστική ανάλυση, έτσι ώστε τα δείγματα που θα απομείνουν να αναφέρονται σε συνθήκες παραγωγής η οποία υπόκειται σε έλεγχο.

Μια σημαντική δυνατότητα των διαγραμμάτων ελέγχου είναι ότι μας παρέχουν ενδείξεις (μέσα σε αρκετά ικανοποιητικά όρια) πότε να αναζητήσουμε αίτια διακυμάνσεων στην παραγωγή και είναι πάντοτε σημαντικό να γνωρίζουμε το πότε. Ωστόσο ανάμεσα στην απόφαση για την αναζήτηση των διαταραχών της παραγωγικής διαδικασίας και στην πραγματική ανακάλυψη και διόρθωση των αιτιών που προξενούν τις διαταραχές υπάρχει μια σημαντική ποσότητα επίπονης τεχνικής εργασίας. Για το λόγο αυτό ο στατιστικός έλεγχος της ποιότητας (SPC) βασίζεται και 80% στην «προσέγγιση μηχανικού» (engineering analysis) και κατά 20% στη στατιστική (statistics).

Συνεχίζοντας την παρουσίαση των βασικών λειτουργιών των διαγραμμάτων ελέγχου, αξίζει να αναφερθούμε στις βασικές παραμέτρους οι οποίες προσδιορίζουν πλήρως τα διαγράμματα αυτά και είναι οι εξής:

- Το χρονικό διάστημα t μεταξύ των λήψεων δυο διαδοχικών δειγμάτων,
- Το μέγεθος δείγματος n .
- Η παράμετρος θέσης των ορίων ελέγχου k .

Τα όρια ελέγχου με παράμετρο θέσης k ονομάζονται όρια k τυπικών αποκλίσεων. Τα περισσότερα διαγράμματα ελέγχου που χρησιμοποιούνται στην πράξη έχουν όρια 3 τυπικών αποκλίσεων. Γενικότερα, η επιλογή της θέσης των ορίων θα πρέπει να στηρίζεται σε οικονομικά κριτήρια. Δεν είναι λογικό να χρησιμοποιούνται διαγράμματα με το ίδιο πλάτος ορίων ελέγχου σε δυο διαδικασίες με διαφορετικά οικονομικά χαρακτηριστικά.

Εκτός από τα όρια ελέγχου, σε ορισμένα διαγράμματα σχεδιάζονται επικουρικά δυο ακόμη οριζόντιες γραμμές μεταξύ των ορίων ελέγχου και της κεντρικής γραμμής, οι οποίες ονομάζονται προειδοποιητικά όρια (warning lines). Τα προειδοποιητικά όρια απέχουν συνήθως ± 2 τυπικές αποκλίσεις της στατιστικής δείγματος από την κεντρική γραμμή.

Αν ένα ή περισσότερα σημεία βρίσκονται μεταξύ ενός προειδοποιητικού ορίου και του αντίστοιχου ορίου ελέγχου, αυτό αποτελεί ένδειξη πιθανής μετατόπισης της μέσης τιμής. Οι συγκεκριμένοι κανόνες λειτουργίας ενός διαγράμματος με προειδοποιητικά όρια (π.χ. έλεγχος για τον εντοπισμό πιθανής βλάβης μετά από δύο ή τρία συνεχόμενα σημεία μεταξύ του προειδοποιητικού ορίου και ορίου ελέγχου) ορίζονται από τον μηχανικό ποιοτικού ελέγχου κατά περίπτωση. Τα προειδοποιητικά όρια αυξάνουν την ευαισθησία του διαγράμματος.

Ενώ οι παράμετροι k και n καθορίζουν τις στατιστικές ιδιότητες του ελέγχου κάθε δείγματος, η παράμετρος t καθορίζει τη συχνότητα της δειγματοληψίας και επηρεάζει το κόστος ελέγχου και την ταχύτητα ανίχνευσης των συστηματικών αιτιών. Το συνολικό κόστος ελέγχου εξαρτάται από τη συχνότητα δειγματοληψίας και το μέγεθος δείγματος. Συχνή δειγματοληψία με μεγάλα δείγματα ανιχνεύει βέβαια ακριβέστερα την επίδραση των συστηματικών αιτιών αλλά έχει υψηλό κόστος. Η μεγάλη πρόοδος στην τεχνολογία των αισθητήρων και των αυτόματων μετρήσεων τα τελευταία χρόνια έχει δημιουργήσει την τάση να λαμβάνονται μικρά αλλά συχνά δείγματα αντί για μεγαλύτερα δείγματα σε αραιότερα διαστήματα. Σε ορισμένες μάλιστα περιπτώσεις είναι ήδη εφικτός ο αυτόματος έλεγχος κάθε παραγόμενης μονάδας αμέσως μετά την κατεργασία.

1.6 Μέτρα απόδοσης ενός διαγράμματος ελέγχου

Μια ακόμα έννοια που σχετίζεται με τα διαγράμματα ελέγχου είναι το μέσο μήκος ροής (ή μέσο μήκος διαδρομής) του διαγράμματος (average run of length, ARL) που ορίζεται από τη σχέση

$$ARL = \frac{1}{p}$$

όπου το p συμβολίζει την πιθανότητα να βρεθεί ένα σημείο του διαγράμματος ελέγχου εκτός των ορίων ελέγχου.

Είναι προφανές ότι το ARL δηλώνει τον αναμενόμενο αριθμό των σημείων (δειγμάτων) που θα πρέπει να σχεδιαστούν σε ένα διάγραμμα ελέγχου για να εμφανιστεί ένα σημείο εκτός των ορίων ελέγχου, αφού το μήκος ροής ή διανομής (run length) ακολουθεί την κατανομή $G(p)$.

Αν η μεταβλητή X ακολουθεί την γεωμετρική κατανομή με πιθανότητα επιτυχίας p , τότε

$$P(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, 3, \dots, n,$$

$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2}.$$

Για μια διεργασία που βρίσκεται εντός ελέγχου και στην περίπτωση που χρησιμοποιούμε 3σ όρια ελέγχου με κανονική κατανομή για την στατιστική συνάρτηση W , έχουμε ότι το εντός ελέγχου μέσο μήκος ροής ARL_{in} (in-control average run length) είναι ίσο με

$$ARL_{in} = \frac{1}{a} = \frac{1}{0,0027} \approx 370,4$$

όπου a είναι η πιθανότητα να βρεθεί ένα σημείο εκτός των ορίων ελέγχου ενώ η διεργασία βρίσκεται εντός ελέγχου και το χαρακτηριστικό X του προϊόντος ακολουθεί κανονική κατανομή με μ , εντός ελέγχου, μέση τιμή ίση με μ και διασπορά σ^2 . Για μια διεργασία που βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου επιπέδου της διεργασίας, το εκτός ελέγχου μέσο μήκος ροής ARL_{out} (out-of-control average run length) είναι ίσο με

$$ARL_{out} = \frac{1}{1-\beta}$$

αφού η πιθανότητα να βρεθεί σε αυτή την περίπτωση ένα σημείο του διαγράμματος ελέγχου εντός των ορίων ελέγχου είναι ίσο με β , οπότε βρίσκεται εκτός των ορίων ελέγχου με πιθανότητα $1-\beta$. Στην πράξη είναι επιθυμητή μια μικρή τιμή του ARL_{out} αφού το ARL_{out} δηλώνει τον αναμενόμενο αριθμό δειγμάτων που πρέπει να ληφθούν για να εντοπισθεί η αλλαγή στο μέσο επίπεδο της διεργασίας από τη στιγμή που η μετατόπιση αυτή συμβεί. Σημαντικό είναι να αναφερθεί πως αυξάνοντας το μέγεθος του δείγματος n η πιθανότητα β μειώνεται και συνεπώς βελτιώνεται η ικανότητα του διαγράμματος ελέγχου να ανιχνεύει εγκαίρως τις μετατοπίσεις του μέσου επιπέδου διεργασίας. Με βάση τη μετατόπιση που θέλουμε να ανιχνεύσουμε, είναι δυνατό να επιλέξουμε το κατάλληλο μέγεθος δείγματος έτσι ώστε να έχουμε συγκεκριμένη τιμή για το ARL_{out} .

Αν αντί των 3σ ορίων είχαμε επιλέξει $k\sigma$ όρια, τότε θα βλέπαμε πως καθώς αυξάνει η τιμή του k μειώνεται η πιθανότητα εσφαλμένης ένδειξης εκτός ελέγχου διεργασίας. Αντίθετα, αν αυξάνεται η πιθανότητα λανθασμένης ένδειξης ότι η διεργασία βρίσκεται εντός ελέγχου ενώ στην πραγματικότητα βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου της διεργασίας.

Η χρήση του ARL ως μέτρου περιγραφής της απόδοσης μια διεργασίας θεωρείται αμφιλεγόμενη, καθώς το ARL που παρατηρείται στην πράξη διαφέρει συνήθως αρκετά από το θεωρητικό ARL λόγω του ότι η κατανομή του μήκους ροής είναι μη-συμμετρική

(παρουσιάζει δεξιά ουρά) και συνεπώς η μέση τιμή δεν μπορεί να χρησιμοποιηθεί ως αντιπροσωπευτικό μέτρο κεντρικής τάσης της κατανομής (ιδιαίτερα για μικρές τιμές του p).

Μια συνήθης πρακτική είναι να συνοδεύεται η τιμή του ARL με την τυπική απόκλιση της κατανομής του μήκους ροής SDRL (standard deviation run length). Στην περίπτωση που η κατανομή του μήκους ροής του διαγράμματος είναι η γεωμετρική η τιμή του SDRL ταυτίζεται με την τιμή του ARL. Επιπλέον, η γνώση της κατανομής του μήκους ροής, μας επιτρέπει τον υπολογισμό και την χρήση ποσοστιαίων σημείων προκειμένου να έχουμε περισσότερη πληροφορία σχετικά με την απόδοση του διαγράμματος.

Θεωρώντας ένα μοντέλο ορίων 3σ που η κατανομή της W είναι κανονική με γνωστές παραμέτρους έχουμε τον ακόλουθο πίνακα για μια εντός ελέγχου διεργασία.

α	ARL	SDRL	Q_1	M	Q_3
0.0027	370.4	370.4	107	257	513

Από τον πίνακα προκύπτει ότι κατά μέσο όρο αναμένονται περίπου 370 δείγματα μέχρι να βρεθεί σημείο εκτός των ορίων ελέγχου όταν η διεργασία είναι εντός ελέγχου. Από την τιμή της διαμέσου M ο διαχειριστής του διαγράμματος γνωρίζει ότι η πιθανότητα να περιμένει 257 δείγματα μέχρι να βρεθεί ένα σημείο εκτός των ορίων ελέγχου είναι τουλάχιστον 50%. Επιπλέον από την τιμή του Q_3 έχουμε ότι η πιθανότητα να μη υπάρξει ένδειξη για σημείο εκτός ελέγχου διεργασίας μετά από 513 δείγματα είναι το πολύ 25%. Έτσι λοιπόν σε πρακτικές εφαρμογές η χρήση των ποσοστιαίων σημείων της κατανομής του μήκους ροής μας παρέχει περισσότερες πληροφορίες σχετικά με την απόδοση του διαγράμματος ελέγχου.

1.7 Άλλα μέτρα απόδοσης

Στην προηγούμενη παράγραφο αναφερθήκαμε στο βασικότερο μέτρο απόδοσης ενός διαγράμματος ελέγχου και είδαμε τα κυριότερα χαρακτηριστικά του. Σε αυτή την παράγραφο θα αναφερθούμε σε εναλλακτικά μέτρα απόδοσης ενός διαγράμματος ελέγχου. Ένα μέτρο που χρησιμοποιείται συχνά είναι ο μέσος χρόνος σήματος ATS (average time to signal) που ορίζεται από τη σχέση

$$ATS = ARL \times \tau$$

όπου το τ συμβολίζει το χρόνο που μεσολαβεί για τη λήψη δυο διαδοχικών δειγμάτων. Έτσι η ποσότητα ATS δηλώνει το μέσο χρόνο που απαιτείται για να δώσει το διάγραμμα ελέγχου ένδειξη εκτός ελέγχου διεργασίας. Αν ο ρυθμός δειγματοληψίας είναι μεταβλητός τότε ορίζεται μια νέα κατηγορία διαγραμμάτων ελέγχου, τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας (adaptive sampling rate control charts).

Σε αυτή την περίπτωση το τ δεν είναι σταθερό αλλά τυχαία μεταβλητή και ο μέσος χρόνος σήματος δίνεται από τη σχέση

$$ATS = ARL \times E(\tau)$$

Είναι σημαντικό να αναφέρουμε ότι τα ARL, ATS που έχουμε αναφέρει έως τώρα υπολογίζουν την απόδοση του διαγράμματος θεωρώντας ότι η μεταβλητή στην εντός ελέγχου τιμή της παραμέτρου συμβαίνει ακριβώς τη χρονική στιγμή που αρχίζει η παρακολούθηση της διεργασίας. Σε αυτή την περίπτωση μιλάμε για μέτρα ARL και ATS αρχικής κατάστασης. Σε πολλές περιπτώσεις όμως η διεργασία ξεκινά ευρισκόμενη εντός ελέγχου και η μεταβολή στην τιμή της υπό παρακολούθηση παραμέτρου συμβαίνει σε κάποιο τυχαίο χρονικό σημείο μεταξύ της λήψης δυο διαδοχικών δειγμάτων.

Σε αυτή την περίπτωση για τον υπολογισμό της απόδοσης ενός διαγράμματος με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο ATS σταθερής κατάστασης το οποίο στη βιβλιογραφία αναφέρεται ως προσαρμοσμένος μέσος χρόνος σήματος (adjusted average time to signal AATS). Αναλόγως στην περίπτωση των διαγραμμάτων με σταθερό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο ARL σταθερής κατάστασης.

Γενικότερα συγκρίσεις μεταξύ των διαφορετικών διαγραμμάτων ελέγχου χωρίς μεταβλητό ρυθμό δειγματοληψίας έχει επικρατήσει να γίνονται χρησιμοποιώντας το ARL αρχικής κατάστασης ενώ για τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το AATS.

1.8 Ταξινόμηση διαγραμμάτων ελέγχου

Όπως αναφέρθηκε και νωρίτερα, μια πρώτη διάκριση των διαγραμμάτων ελέγχου βασίζεται στην κατανομή του ποιοτικού χαρακτηριστικού που εξετάζουμε. Έτσι, αν το ποιοτικό χαρακτηριστικό είναι συνεχές έχουμε διαγράμματα ελέγχου για μεταβλητές (control charts for variables) ενώ αν είναι διακριτό έχουμε διαγράμματα ελέγχου για διακριτά

χαρακτηριστικά (control charts for attributes). Έτσι λοιπόν, μπορούμε να διακρίνουμε τα διαγράμματα ελέγχου ως εξής:

- Αν από την παραγωγική διεργασία λαμβάνονται δείγματα μετρήσεων μεγέθους μεγαλύτερου της μονάδας μιλάμε για διαγράμματα ελέγχου για ομάδες (control charts for rational groups) , ενώ αν λαμβάνονται δείγματα μετρήσεων μεγέθους ίσου με την μονάδα για διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις (control charts for individual observations).
- Αν οι μετρήσεις που λαμβάνονται κάθε χρονική στιγμή t εξαρτώνται από μετρήσεις που ελήφθησαν σε προηγούμενες χρονικές στιγμές τότε αναφερόμαστε σε διαγράμματα ελέγχου για αυτοσυσχετιζόμενες διεργασίες (control charts for autocorrelated processes), ενώ αν οι μετρήσεις που λαμβάνονται σε κάθε χρονική στιγμή είναι ανεξάρτητες από μετρήσεις που ελήφθησαν σε προηγούμενες χρονικές στιγμές τότε αναφερόμαστε σε διαγράμματα ελέγχου για ασυσχέτιστες διεργασίες (control charts for uncorrelated processes).
- Αν οι μετρήσεις που λαμβάνονται αφορούν ένα ποιοτικό χαρακτηριστικό χρησιμοποιούμε μονομεταβλητά διαγράμματα ελέγχου (univariate control charts) ενώ αν οι μετρήσεις αναφέρονται σε περισσότερα ποιοτικά χαρακτηριστικά χρησιμοποιούμε πολυμεταβλητά διαγράμματα ελέγχου (multivariate control charts).
- Αν οι μετρήσεις που λαμβάνονται ακολουθούν μια γνωστή κατανομή τότε μιλάμε για παραμετρικά διαγράμματα ελέγχου ενώ στη αντίθετη περίπτωση μιλάμε για μη παραμετρικά διαγράμματα ελέγχου.
- Αν το διάγραμμα ελέγχου στοχεύει στην παρακολούθηση της μέσης τιμής του ποιοτικού χαρακτηριστικού μιλάμε για διαγράμματα ελέγχου για τη μέση τιμή, ενώ αν στοχεύει στην παρακολούθηση της διασποράς του μιλάμε για διαγράμματα ελέγχου για τη διασπορά.
- Αν οι μετρήσεις (δείγματα) λαμβάνονται σε σταθερούς χρόνους μιλάμε για διαγράμματα ελέγχου σταθερού χρόνου, ενώ στην αντίθετη περίπτωση μιλάμε για διαγράμματα ελέγχου μεταβλητού χρόνου.
- Αν το μέγεθος του δείγματος είναι σταθερό τότε μιλάμε για διαγράμματα ελέγχου σταθερού δείγματος, ενώ στην αντίθετη περίπτωση μιλάμε για διαγράμματα ελέγχου μεταβλητού μεγέθους δείγματος

1.9 Βασικά είδη διαγραμμάτων

Τα βασικότερα είδη διαγραμμάτων ελέγχου που χρησιμοποιούνται στην πράξη είναι τα διαγράμματα ελέγχου τύπου Shewhart, CUSUM και EWMA.

Στο σημείο αυτό αξίζει να τονίσουμε ότι για τον έλεγχο μιας παραγωγικής διαδικασίας με χρήση διαγραμμάτων ελέγχου υπάρχουν δυο φάσεις, η Φάση I και η Φάση II οι οποίες αφορούν στα εξής:

Φάση I: Σε αυτή τη φάση τα διαγράμματα ελέγχου χρησιμοποιούνται αναδρομικά για να ελέγξουμε αν η διεργασία είναι εντός ή εκτός ελέγχου εξετάζοντας δείγματα που συλλέχθηκαν σε παρελθοντικό χρόνο. Σκοπός τους είναι να βοηθήσουνε το διαχειριστή της διεργασίας να φέρει τη διεργασία εντός στατιστικού ελέγχου. Όταν αυτό επιτευχθεί τα διαγράμματα ελέγχου που προκύπτουν (κεντρική γραμμή και όρια ελέγχου) είναι κατάλληλα για μελλοντική παρακολούθηση της συμπεριφοράς της διεργασίας. Αυτή η χρήση των διαγραμμάτων αναφέρεται ως αναδρομική (retrospective). Γενικά, πολύ περισσότερες δράσεις και σκέψεις εξελίσσονται κατά τη διάρκεια αυτής της φάσης πλην της απλής διαγραμματοποίησης κάποιων δεδομένων. Ο διαχειριστής της διεργασίας μελετά σε βάθος τη διεργασία και αυτό γιατί δεν είναι εύκολο κανείς να αποφασίσει αν η διεργασία ήταν εντός ή εκτός στατιστικού ελέγχου κατά τη χρονική περίοδο που τα δεδομένα συλλέχθηκαν.

Φάση II: Σε αυτή τη φάση τα διαγράμματα ελέγχου χρησιμοποιούνται προκειμένου να ελέγξουμε συνεχώς αν η διαδικασία παραμένει εντός ελέγχου. Ο διαχειριστής έχει στα χέρια του ένα πολύτιμο εργαλείο μέσω του οποίου είναι δυνατό να παρακολουθεί συνεχώς την παραγωγική διεργασία και να ανιχνεύει εγκαίρως μια πιθανή αλλαγή στο μέσο επίπεδο των χαρακτηριστικών που καθορίζουν την ποιότητα του παραγόμενου προϊόντος. Δηλαδή σε κάθε χρονική περίοδο που λαμβάνεται ένα δείγμα από τη διεργασία ο διαχειριστής παίρνει μια απάντηση στο ερώτημα εάν η διεργασία παραμένει εντός ελέγχου. Παράλληλα αδιαφορεί για τον τρόπο με τον οποίο προέκυψαν τα όρια ελέγχου και η κεντρική γραμμή του διαγράμματος που χρησιμοποιεί.

1.9.1 Διαγράμματα ελέγχου τύπου Shewhart

Τα διαγράμματα ελέγχου τύπου Shewhart προτάθηκαν από τον Walter A. Shewhart (1931) ο οποίος εργαζόταν ως στατιστικός στα εργαστήρια της εταιρείας Bell και αποτελούν τα πιο γνωστά αλλά και πιο συχνά χρησιμοποιούμενα διαγράμματα ελέγχου.

Αποτελούνται από δύο όρια ελέγχου, το άνω όριο ελέγχου UCL και το κάτω όριο ελέγχου LCL, και την κεντρική γραμμή του διαγράμματος CL η οποία έχει μέση τιμή ίση με το μέσο επίπεδο της διεργασίας χωρίς την παρουσία ειδικής αιτίας μεταβλητότητας (εντός ελέγχου διεργασία). Για την κατασκευή ενός διαγράμματος Shewhart συλλέγονται δείγματα μεγέθους

$n \geq 1$, υπολογίζεται η τιμή κατάλληλης στατιστικής συνάρτησης W η οποία απεικονίζεται σε ένα διάγραμμα όπου τα όρια ελέγχου απέχουν απόσταση k το καθένα (σε μονάδες τυπικής απόκλισης) από την κεντρική γραμμή. Όταν ένα σημείο βρεθεί εκτός των ορίων ελέγχου τότε έχουμε ένδειξη εκτός ελέγχου διεργασίας.

Η πρόταση του Shewhart ήταν για $k = 3$ αν και έκτοτε προτάθηκαν διάφορες επιλογές για την τιμή του k με σκοπό να καλύψουν τις απαιτήσεις στην πράξη. Το διάγραμμα τύπου Shewhart για την παρακολούθηση μέσης τιμής μιας παραγωγικής διεργασίας, το οποίο βασίζεται σε 3σ όρια ελέγχου με μετρήσεις που προέρχονται από κανονική κατανομή με γνωστές παραμέτρους είναι το πιο γνωστό και ευρέως χρησιμοποιούμενο διάγραμμα της κατηγορίας αυτής και αποτελεί μέτρο σύγκρισης με όλα τα νέα προτεινόμενα διαγράμματα παρακολούθησης της μέσης τιμής.

Το βασικό μειονέκτημα των διαγραμμάτων ελέγχου τύπου Shewhart είναι ότι δεν είναι ευαίσθητα στην ανίχνευση μικρών ή και μεσαίων μετατοπίσεων στην τιμή της υπό παρακολούθηση παραμέτρου που καθορίζει την ποιότητα των παραγόμενων προϊόντων. Για αυτό το λόγο έχουν προταθεί διαγράμματα τα οποία χρησιμοποιούν σύνθετους κανόνες για την ανακήρυξη μιας διεργασίας εκτός ελέγχου. Πιο συγκεκριμένα, ένας τρόπος να αυξήσουμε την ευαισθησία του διαγράμματος είναι με τη χρήση προειδοποιητικών ορίων ελέγχου και κανόνων διακοπής οι οποίοι βασίζονται στη θεωρία ροών (run rules). Χρησιμοποιώντας, λοιπόν, την επιθυμητή τιμή k μπορούμε να αναπτύξουμε διαγράμματα τύπου Shewhart για την παρακολούθηση της μέσης τιμής και της διασποράς του υπό παρακολούθηση χαρακτηριστικού.

1.9.2 Διαγράμματα Ελέγχου CUSUM

Στα διαγράμματα CUSUM απεικονίζεται η στατιστική συνάρτηση

$$S_t^+ = \max[0, X_t - (\mu_0 + K) + S_{t-1}^+], \quad S_0^+ = 0$$

και έχει ως στόχο την ανίχνευση μετατοπίσεων στη μέση τιμή της παραγωγικής διεργασίας σε υψηλότερο επίπεδο. Στον παραπάνω τύπο μ_0 είναι η εντός ελέγχου μέση τιμή ενώ με X_t συμβολίζεται η τιμή της μεμονωμένης παρατήρησης του λαμβάνεται τη χρονική στιγμή t . Επιπλέον, απεικονίζεται η ποσότητα

$$S_t^- = \max[0, X_t - (\mu_0 - K) + S_{t-1}^-], \quad S_0^- = 0$$

και έχει ως σκοπό την ανίχνευση μετατοπίσεων στη μέση τιμή της παραγωγικής διεργασίας σε χαμηλότερο επίπεδο. Οι τιμές που λαμβάνουν οι ποσότητες S_0^+ και S_0^- ονομάζονται τιμές εκκίνησης (head start values). Η ποσότητα $K \geq 0$ ονομάζεται τιμή αναφοράς (reference value) και η τιμή της προκύπτει από τη σχέση

$$K = \frac{\delta\sigma}{2} = \frac{|\mu_1 - \mu_0|}{2}$$

όπου μ_1 δηλώνει την τιμή μιας εκτός ελέγχου μέσης τιμής. Η ποσότητα S_t^+ συσσωρεύει τις αποκλίσεις των παρατηρήσεων X_t από την ποσότητα $\mu_0 + K$ από τη στιγμή που θα εμφανιστεί θετική απόκλιση και μπορεί να χρησιμοποιηθεί για τον έλεγχο υποθέσεων της μορφής

$$H_0: \mu = \mu_0$$

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

αφού μεγάλες θετικές τιμές της S_t^+ οδηγούν σε αποδοχή της H_1^+ .

Αντίστοιχα, η ποσότητα S_t^- συσσωρεύει τις αποκλίσεις των παρατηρήσεων X_t από την ποσότητα $\mu_0 - K$ από τη στιγμή που θα εμφανιστεί αρνητική απόκλιση και μπορεί να χρησιμοποιηθεί για τον έλεγχο υποθέσεων της μορφής

$$H_0: \mu = \mu_0$$

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

αφού μεγάλες αρνητικές τιμές της S_t^- οδηγούν σε αποδοχή της H_1^- .

Και για τις δυο εναλλακτικές υποθέσεις H_1^+ και H_1^- η απόφασή μας για αποδοχή ή όχι σε κάθε βήμα της διαδικασίας θα εξαρτηθεί από το αν ισχύει η σχέση $S_t^+ > H$ ή $S_t^- > H$, όπου H μια θετική σταθερά ($H > 0$). Η ποσότητα H ονομάζεται διάστημα απόφασης (decision interval).

Συνήθως η τιμή για το διάστημα απόφασης είναι η $H = h\sigma, h = 4, 5$. Είναι προφανές πως έχουμε ένδειξη εκτός ελέγχου διεργασίας (μετατόπιση της μέσης τιμής της παραγωγικής διεργασίας σε υψηλότερο ή χαμηλότερο επίπεδο) αν γίνει αποδεκτή μια εκ των εναλλακτικών υποθέσεων H_1^+ και H_1^- .

Τα διαγράμματα CUSUM χρησιμοποιούνται εναλλακτικά των Shewhart στην ανίχνευση μικρών ή μεσαίων μετατοπίσεων στην τιμή της υπό παρακολούθηση παραμέτρου. Ανήκουν στην κατηγορία διαγραμμάτων ελέγχου με μνήμη που βασικό τους χαρακτηριστικό είναι η ενσωμάτωση πληροφοριών από τις τιμές των προηγούμενων παρατηρήσεων στην απεικονιζόμενη ποσότητα. Τα διαγράμματα CUSUM αποτελούν διαγράμματα με ομοιόμορφη μνήμη αφού δίνουν την ίδια βαρύτητα σε όλες τις προηγούμενες παρατηρήσεις.

1.9.3 Διαγράμματα Ελέγχου EWMA

Στο διάγραμμα EWMA απεικονίζεται η ποσότητα

$$Z_t = \lambda X_t + (1 - \lambda)Z_{t-1}, \quad X_0 = \mu$$

και έχει ως σκοπό την παρακολούθηση της μέσης τιμής μιας παραγωγικής διεργασίας. Στον παραπάνω τύπο το X_t συμβολίζει την τιμή της μεμονωμένης παρατήρησης που λαμβάνεται την χρονική στιγμή t , ενώ η τιμή μ δηλώνει τη μέση τιμή της κατανομής της X_t . Η τιμή της ποσότητας X_0 ονομάζεται τιμή εκκίνησης ενώ η σταθερά λ ($0 \leq \lambda \leq 1$) καθορίζει το βάρος που δίνουμε στην τιμή της πιο πρόσφατης παρατήρησης.

Για να κατασκευάσουμε ένα διάγραμμα EWMA χρειαζόμαστε την μέση τιμή μ_{Z_t} και την τυπική απόκλιση σ_{Z_t} της ποσότητας Z_t που απεικονίζεται στο διάγραμμα. Στην περίπτωση που οι αρχικές παρατηρήσεις X_t είναι ανεξάρτητες τυχαίες μεταβλητές με εντός ελέγχου μέση τιμή μ_0 και διακύμανση σ^2 τότε

$$\sigma_{Z_t}^2 = \sigma^2 \left(\frac{\lambda}{2 - \lambda} \right) [1 - (1 - \lambda)^{2t}].$$

Σε ένα διάγραμμα τύπου EWMA έχουμε τα ακόλουθα όρια ελέγχου και κεντρική γραμμή για την παρακολούθηση της μέσης τιμής μιας παραγωγικής διεργασίας.

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

$$CL = \mu_0$$

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

Για να κατασκευάσουμε ένα διάγραμμα EWMA απαιτείται ο προσδιορισμός των τιμών των παραμέτρων λ και k . Ο προσδιορισμός αυτός συνήθως γίνεται με στατιστικά κριτήρια. Η τιμή της παραμέτρου k καθορίζει την απόσταση των ορίων ελέγχου από την κεντρική γραμμή. Οι συνήθεις τιμές για το λ είναι $\lambda = 0.05, 0.10, 0.20$.

Τα όρια αυτά ονομάζονται χρονομεταβλητά αφού υπάρχει εξάρτηση της τυπικής απόκλισης σ_{Z_t} από το χρόνο. Όμως καθώς αυξάνεται το t , τα όρια ελέγχου προσεγγίζουν μια σταθερή τιμή και έτσι καταλήγουμε στα όρια ελέγχου σταθερής κατάστασης τα οποία είναι τα ακόλουθα

$$LCL = \mu_0 - k\sigma \sqrt{\frac{\lambda}{(2-\lambda)}}$$

$$CL = \mu_0$$

$$UCL = \mu_0 + k\sigma \sqrt{\frac{\lambda}{(2-\lambda)}}$$

Στο σημείο αυτό αξίζει να σημειώσουμε ότι στα διαγράμματα EWMA λόγω του ότι απεικονίζονται σε αυτά σημεία που αποτελούν σταθμισμένο μέσο όρο όλων των προηγούμενων παρατηρήσεων, είναι ανθεκτικά στην παραβίαση της υπόθεσης κανονικότητας των αρχικών παρατηρήσεων. Αποτελούν λοιπόν τα πλέον κατάλληλα διαγράμματα για την παρακολούθηση εργασιών μέσω μεμονωμένων παρατηρήσεων έναντι των αντίστοιχων διαγραμμάτων Shewhart.

1.10 Διαγράμματα ελέγχου ιδιοτήτων

Αν και τα διαγράμματα ελέγχου για το μέσο (\bar{X}) και το εύρος (R) παρουσιάζουν πολλά πλεονεκτήματα, αφού αποτελούν δυναμικά εργαλεία για την διάγνωση των προβλημάτων ποιότητας των βιομηχανικών προϊόντων, η χρήση τους περιορίζεται στις περιπτώσεις που τα χαρακτηριστικά ποιότητας μπορούν να μετρηθούν και να εκφραστούν με ακέραιους αριθμούς.

Στην πράξη τα περισσότερα χαρακτηριστικά ποιότητας παρατηρούνται σαν ιδιότητες, δηλαδή τα προϊόντα που ελέγχονται διακρίνονται σε δυο κατηγορίες: σε εκείνα που συμμορφώνονται με τις τεχνικές προδιαγραφές (conforming) και σε εκείνα που δεν συμμορφώνονται με τις τεχνικές προδιαγραφές (non-conforming).

Ακόμα και για τις περιπτώσεις που τα χαρακτηριστικά ποιότητας μπορούν να μετρηθούν, η αποκλειστική χρήση των διαγραμμάτων \bar{X} και R μπορεί να μην είναι συχνά πρακτική και οικονομική.

Όταν τα δεδομένα της παραγωγικής διαδικασίας δεν αφορούν σε μετρήσιμες μονάδες αλλά σε ποσότητες που καλύπτονται από μια διάκριση ναι/όχι ή με χαρακτηριστικά διαλογής (ελαττωματικά αντικείμενα, ελαττώματα ενός προϊόντος, κτλ) είναι απαραίτητα διαγράμματα ελέγχου ιδιοτήτων (control chart for attributes). Η βασική διαφορά του από διαγράμματα ελέγχου μεταβλητών (control chart for variables) βρίσκεται στον τρόπο δειγματοληψίας και στη μέτρηση των ορίων ελέγχου.

Εάν για μια ορισμένη χρονική περίοδο δεν ελέγχεται ολόκληρη η παραγωγή, τότε το δείγμα που επιλέγεται θα πρέπει να είναι αντιπροσωπευτικό του τρέχοντος ποσοστού ελαττωμάτων. Ο υπολογισμός των ορίων ελέγχου εξαρτάται από το αν στο διάγραμμα παριστάνονται αριθμοί ή αναλογίες (ποσοστά) και από το αν λαμβάνονται υπόψη οι ελαττωματικές μονάδες ή τα ελαττώματα- μια μονάδα μπορεί να έχει ορισμένο αριθμό ελαττωμάτων, προτού ταξινομηθεί ως ελαττωματική.

Προκειμένου να επιλεγεί το κατάλληλο διάγραμμα ελέγχου θα πρέπει πρώτα να απαντήσουμε στις εξής ερωτήσεις:

1. Το μέγεθος του δείγματος είναι σταθερό ή μεταβάλλεται από δείγμα σε δείγμα;
2. Το ποιοτικό χαρακτηριστικό ταυτοποιεί τη μονάδα ως ελαττωματική (μη συμμορφούμενη) ή αναγνωρίζει σε αυτή ελαττώματα;

Ανάλογα με τις απαντήσεις που θα δώσουμε στα παραπάνω ερωτήματα, υπάρχουν τέσσερα είδη διαγραμμάτων ιδιοτήτων:

- ✓ το διάγραμμα p (μεταβλητό μέγεθος δείγματος – ελαττωματικά προϊόντα)
- ✓ το διάγραμμα np (σταθερό μέγεθος δείγματος – ελαττωματικά προϊόντα)
- ✓ το διάγραμμα u (μεταβλητό μέγεθος δείγματος – ελαττώματα)
- ✓ το διάγραμμα c (σταθερό μέγεθος δείγματος – ελαττώματα)

Για να θέσουμε τα ανώτερα και κατώτερα όρια ελέγχου, ακολουθούμε την ίδια λογική με τα διαγράμματα μεταβλητών και προκύπτουν οι ακόλουθες σχέσεις:

Τύπος Διαγράμματος	Όρια Ελέγχου
p	$\bar{p} \pm 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
np	$n\bar{p} \pm 3 \sqrt{n\bar{p}(1-\frac{n\bar{p}}{n})}$
u	$\bar{u} \pm 3 \sqrt{\frac{\bar{u}}{n}}$
c	$\bar{c} \pm 3\sqrt{\bar{c}}$

Όμως καθώς δεν μπορούμε να μιλάμε για αρνητικές αναλογίες ή για αρνητικό αριθμό ελαττωμάτων, όποτε το κάτω όριο προκύπτει αρνητικό, λαμβάνεται ίσο με το μηδέν. Το κόστος για την συλλογή στοιχείων για τα διαγράμματα ιδιοτήτων είναι συνήθως μικρότερο του κόστους συλλογής στοιχείων για τους πίνακες \bar{X} και R , γιατί οι πίνακες ιδιοτήτων χρησιμοποιούν κατά κανόνα στοιχεία που ήδη έχουν συγκεντρωθεί για άλλους σκοπούς. Επίσης το κόστος υπολογισμού μπορεί να είναι μικρότερο, αφού ένας πίνακας ιδιοτήτων μπορεί να εφαρμοστεί για οποιονδήποτε αριθμό χαρακτηριστικών ποιότητας, ενώ χρειάζονται ξεχωριστοί πίνακες \bar{X} και R για κάθε χαρακτηριστικό ποιότητας.

Πέρα από τα πλεονεκτήματα κόστους που παρουσιάζουν οι χάρτες ελέγχου ιδιοτήτων, παρουσιάζουν και το πλεονέκτημα ότι παρέχουν χρήσιμα στοιχεία για την χρονική μεταβολή της ποιότητας των προϊόντων που παράγει.

Ως ελαττωματικό ορίζεται το προϊόν που δεν είναι σύμφωνο με μία ή περισσότερες δοσμένες προδιαγραφές. Ενώ το ελάττωμα αναφέρεται στην αποτυχία κάποιου ειδικού χαρακτηριστικού σ' ένα προϊόν. Σύμφωνα με τα παραπάνω ένα ελαττωματικό προϊόν μπορεί να έχει ένα ή περισσότερα ελαττώματα.

1.11 Φάση I και Φάση II

Έχοντας λάβει μια πρώτη εικόνα για το πώς μπορούμε να ελέγξουμε την ποιότητα των παραγόμενων προϊόντων μέσω του Στατιστικού Ελέγχου Ποιότητας (Statistic Quality Control), είμαστε σε θέση να εμβαθύνουμε σε πιο εξειδικευμένες πτυχές της μεθόδου αυτής.

Όπως αναφέραμε και νωρίτερα, ο έλεγχος μιας παραγωγικής διεργασίας με χρήση διαγραμμάτων ελέγχου γίνεται σε δυο Φάσεις (Φάση I και II) όπου κάθε μια εξυπηρετεί έναν διακριτό σκοπό. Στη Φάση I, έχοντας συλλέξει δεδομένα σε παρελθόντα χρόνο, κατασκευάζουμε δοκιμαστικά όρια ελέγχου προκειμένου να καθορίσουμε αν μια διαδικασία ήταν εντός ελέγχου κατά την συλλογή των δεδομένων και στη συνέχεια να κρίνουμε αν θα μπορούσαμε βάση αυτών να κατασκευάσουμε αξιόπιστα όρια ελέγχου για την παρακολούθηση της μελλοντικής συμπεριφοράς της παραγωγικής διαδικασίας. Αυτό τυπικά είναι και το πρώτο βήμα όταν εφαρμόζουμε όρια ελέγχου σε μια παραγωγική διεργασία. Τα διαγράμματα ελέγχου σε αυτή τη φάση βοηθούν στο να φέρουμε σε κατάσταση εντός ελέγχου την παραγωγική διεργασία. Η Φάση II ξεκινά από την στιγμή που έχουμε συλλέξει ένα σετ δεδομένων που αντιπροσωπεύουν μια εντός ελέγχου διεργασία. Η φάση αυτή, λειτουργεί περισσότερο για να καλύψει την ανάγκη παρακολούθησης της διεργασίας, προκειμένου να σιγουρέψουμε ότι αυτή παραμένει εντός ελέγχου και μελλοντικά.

Έτσι, στη Φάση I συγκρίνουμε έναν σωρό, για παράδειγμα m σημείων με ένα σετ ορίων ελέγχου τα οποία έχουν κατασκευαστεί βάσει αυτών των σημείων. Τυπικά, στη Φάση I θεωρούμε $m=20$ ή 25 υποομάδες. Στη Φάση I θεωρούμε συνήθως ότι η διεργασία ήταν εκτός ελέγχου, και έτσι ο σκοπός του αναλυτή είναι να τη φέρει εντός στατιστικού ελέγχου. Τα όρια ελέγχου υπολογίζονται βάσει των m υποομάδων και τα δεδομένα παρουσιάζονται στα διαγράμματα ελέγχου. Τα εκτός ορίων σημεία χρήζουν περαιτέρω διερεύνησης για πιθανή ειδική αιτία μεταβλητότητας. Οποιαδήποτε ειδική αιτία μεταβλητότητας αναγνωριστεί, πρέπει να γίνει προσπάθεια εξάλειψής της. Σημεία τα οποία βρίσκονται εκτός των ορίων ελέγχου αφαιρούνται και τα όρια υπολογίζονται ξανά. Στη συνέχεια νέα δεδομένα συλλέγονται και συγκρίνονται με τα ανανεωμένα όρια ελέγχου. Η διαδικασία αυτή μπορεί να επαναληφθεί αρκετές φορές. Τελικά, η διεργασία σταθεροποιείται και ένα «καθαρό» σετ δεδομένων που αντιπροσωπεύει τα εντός ελέγχου διαγράμματα συνεχίζει στο επόμενο στάδιο ελέγχου, στη Φάση II.

Για τη Φάση I πιο αποδοτικά θεωρούνται τα διαγράμματα Shewhart, καθώς είναι εύκολα στην κατασκευή και την ερμηνεία. Επιπλέον, είναι αποτελεσματικά όσο αφορά την παρακολούθηση μεγάλων και διαρκών αλλαγών τις παραμέτρους της διεργασίας αλλά και των outliers (μεμονωμένα σημεία που μπορεί να έχουν προκληθεί από ειδικές αιτίες μεταβλητότητας μικρής διάρκειας), σε λάθη μετρήσεων, στην καταγραφή δεδομένων ή /και μεταβίβασης λαθών. Επιπλέον, οι ακολουθίες στα διαγράμματα τύπου Shewhart συχνά είναι εύκολο να ερμηνευθούν και έχουν φυσική ερμηνεία. Τα είδη των ειδικών αιτιών

μεταβλητότητας που συχνά προκύπτουν στη Φάση I, έχουν ως αποτέλεσμα μεγάλες αλλαγές στις οποίες τα διαγράμματα Shewhart αποδίδουν στο μέγιστο. Το μέσο μήκος ροής (Average Run Length, ARL) δεν χρησιμοποιείται συχνά στη Φάση I, καθώς στη φάση αυτή εστιάζουμε περισσότερο στην πιθανότητα μια ειδική αιτία μεταβλητότητα να παρατηρηθεί παρά στην περίπτωση λανθασμένου συναγερμού.

Στη Φάση II, συνήθως θεωρούμε ότι η διεργασία είναι σταθερή. Συχνά, οι ειδικές αιτίες μεταβλητότητας στη φάση αυτή οδηγούν σε μικρές αλλαγές στη διεργασία καθώς οι πιο «επικίνδυνες» αιτίες μεταβλητότητας έχουν επαλειφθεί στη Φάση I. Στη φάση αυτή, σκοπός είναι να εστιάσουμε στην παρακολούθηση της διεργασίας και όχι να φέρουμε τη διεργασία εντός στατιστικού ελέγχου. Το μέσο μήκος ροής (Average Run Length, ARL) είναι ένας έγκυρος τρόπος αξιολόγησης της παρουσίας ενός διαγράμματος ελέγχου στη Φάση II. Τα διαγράμματα Shewhart δεν είναι τόσο αποδοτικά στη Φάση II γιατί δεν είναι ευαίσθητα στις μικρές αλλαγές, για αυτό το λόγο και το ARL είναι σχετικά αδύναμος τρόπος αξιολόγησής τους. Έτσι λοιπόν, στη φάση αυτή έχει επικρατήσει η χρήση των διαγραμμάτων CUSUM και EWMA.

Γίνεται λοιπόν κατανοητή η σημαντικότητα αυτών των μεθόδων για την διεξαγωγή ενός αξιόπιστου Στατιστικού Ελέγχου Ποιότητας (Statistical Quality Control). Αποτελώντας κατά πολλούς το θεμελιώδες στάδιο για έναν έγκυρο και αποτελεσματικό έλεγχο δίνεται η δυνατότητα στη συνέχεια να εμβαθύνουμε σε πιο εξειδικευμένες πτυχές που αφορούν στη Φάση I για τον έλεγχο μιας διεργασίας. Στα επόμενα κεφάλαια θα εστιάσουμε στη φάση αυτή με σκοπό να αναλύσουμε όσο το δυνατό καλύτερα τα τεχνικά μέσα αλλά και τις μεθόδους που σχετίζονται με αυτή.

ΚΕΦΑΛΑΙΟ 2

2.1 Επισκόπηση

Πριν αναπτύξουμε ένα σχέδιο παρακολούθησης της διεργασίας, θα πρέπει πρώτα να καθορίσουμε και να δώσουμε μια λειτουργική περιγραφή των μεταβλητών που είναι σημαντικές για την μέτρηση της ποιότητας της διεργασίας. Μόλις ταυτοποιηθεί (αναγνωριστεί), θα πρέπει να διενεργήσουμε ανάλυση για το σύστημα μέτρησης, προκειμένου να διαβεβαιώσουμε ότι είναι ικανοποιητικό προκειμένου να πάρουμε αξιόπιστες μετρήσεις για τη λειτουργία της διεργασίας. Για περισσότερες πληροφορίες σχετικά με την αξιολόγηση της ικανότητας και αξιοπιστίας των συστημάτων μέτρησης μπορούμε να ανατρέξουμε στον Montgomery (2013, p.p. 379-397) και Steiner and MacKay (2005, p.p. 89-104). Οι Steiner and MacKay (2005) συνέστησαν τη χρήση δεδομένων Φάσης I για την αξιολόγηση του συστήματος μέτρησης, ειδικότερα για την εκτίμηση της διακύμανσης της διεργασίας.

Η πιο συνήθης προσέγγιση για τον καθορισμό της σταθερότητας της διεργασίας είναι η αναδρομική χρήση διαγραμμάτων ελέγχου τύπου Shewhart. Αξίζει να σημειώσουμε ότι τα διαγράμματα ελέγχου τύπου CUSUM και EWMA χρησιμοποιούνται στη Φάση II γιατί μπορούν να εντοπίσουν γρηγορότερα μικρές μετατοπίσεις. Η ταχύτητα εντοπισμού δεν είναι τόσο σημαντική για τα σταθερά ιστορικά σείτ δεδομένων της Φάσης I, γι' αυτό και επικεντρωνόμαστε στην αναδρομική χρήση διαγραμμάτων ελέγχου τύπου Shewhart.

Η στατιστική απόδοση μεθόδων των διαγραμμάτων ελέγχου στη Φάση I εξαρτάται από χαρακτηριστικά σχετικά με την κατανομή της διεργασίας, τη φύση της μεταβλητότητας, τη συχνότητα δειγματοληψίας, το μέγεθος του δείγματος αναφοράς, αν χρησιμοποιούνται ομαδοποιημένα ή μεμονωμένα δεδομένα και τη μέθοδο που χρησιμοποιήθηκε για την εκτίμηση των άγνωστων παραμέτρων της διεργασίας. Στην πραγματικότητα, είναι αδύνατον να εντοπίσουμε την επίδραση του καθενός από αυτά με τη χρήση και ερμηνεία των διαγραμμάτων ελέγχου. Για παράδειγμα, δεν μπορούμε να μελετήσουμε την επίδραση του μεγέθους του δείγματος στη στατιστική απόδοση των διαγραμμάτων ελέγχου χωρίς να λάβουμε υπόψη μας αν τα δεδομένα είναι ομαδοποιημένα, πώς έχουν εκτιμηθεί οι παράμετροι και η κατανομή της διεργασίας.

2.1.1 Χρήση ορθολογικής δειγματοληψίας

Μια στρατηγική δειγματοληψίας, η οποία αναφέρεται συχνά στο γενικό πλαίσιο του στατιστικού ελέγχου διεργασίας είναι η ορθολογική δειγματοληψία (rational subgrouping). Η ορθολογική δειγματοληψία αφορά στο τι μετράμε και στο πώς συλλέγουμε τα δείγματα. Η εφαρμογή της βασίζεται στην γνώση της διεργασίας και στην κοινή λογική. Ο Shewhart (1931, p.p. 298-299) ανέφερε τη σημασία της χρήσης της ορθολογικής δειγματοληψίας, όχι μόνο όσο αφορά τον εντοπισμό εκτός ελέγχου συνθηκών, αλλά και στην εύρεση ειδικών αιτιών μεταβλητότητας για κάθε σημείο εκτός ελέγχου. Η στρατηγική δειγματοληψίας που εφαρμόζεται στη Φάση II μπορεί να καθοριστεί από τη γνώση που αποκτούμε από τη Φάση I. Βέβαια, το σχέδιο δειγματοληψίας στη Φάση I δεν είναι απαραίτητο να αντιστοιχεί με αυτό της Φάσης II. Για παράδειγμα, θα μπορούσαν να χρησιμοποιηθούν ομαδοποιημένες παρατηρήσεις με διαγράμματα ελέγχου τύπου Shewhart στη Φάση I και μεμονωμένες παρατηρήσεις με ένα CUSUM διάγραμμα για γρήγορη ανίχνευση μετατοπίσεων στη Φάση II.

Γενικά, τα δείγματα στη Φάση II, θα μπορούσαν να συλλεχθούν έτσι ώστε, αν υπάρχει ειδική αιτία μεταβλητότητας που οδηγεί σε κατάσταση εκτός ελέγχου, η πιθανότητα να διαφέρουν τα δείγματα μεταξύ τους να είναι μεγαλύτερη από την πιθανότητα να υπάρχουν διαφορές εντός των δειγμάτων. Αυτό ισχύει γιατί είναι ευκολότερο να ανιχνεύσει κανείς και να διαγνώσει μετατοπίσεις του μέσου παρά της διασποράς.

Η ύπαρξη ενός «εντός ελέγχου δείγματος αναφοράς», αποτελεί θεμελιώδη υπόθεση απαραίτητη για την επιλογή και κατασκευή ενός κατάλληλου διαγράμματος ελέγχου Φάσης II. Ουσιαστικά, το να ορίσει κανείς μια εντός ελέγχου κατάσταση και να βρει ένα εντός ελέγχου δείγμα δεν είναι εύκολο. Ο Shewhart (1939, p.76) έγραψε, « Στην πλειοψηφία των πρακτικών παραδειγμάτων, η πιο δύσκολη δουλειά είναι να επιλέξουμε το δείγμα που θα χρησιμοποιηθεί ως βάση για τον ορισμό του εύρους ανοχής (όρια ελέγχου). Αν επιλέξουμε δείγμα χωρίς να λάβουμε υπόψη μας τις ειδικές αιτίες μεταβλητότητας που παρουσιάζονται, είναι πρακτικά αδύνατο να ορίσουμε εύρος ανοχής που να μην υπόκειται σε ένα τεράστιο λάθος».

Ο Montgomery (2013, p.201) έδωσε δύο προσεγγίσεις για τη δειγματοληψία. Στην πρώτη προσέγγιση, έχουμε διαδοχικά δείγματα που ουσιαστικά προκύπτουν την ίδια χρονική στιγμή, ή σε χρονικά κοντινές στιγμές, προκειμένου να σχηματιστούν τα δείγματα. Αυτή η προσέγγιση δίνει στιγμιότυπα της διεργασίας στο χρόνο και είναι χρήσιμη για την παρακολούθηση πιθανών μετατοπίσεων της διεργασίας. Στη δεύτερη προσέγγιση, διαλέγουμε

τυχαία και μετράμε μονάδες οι οποίες παράχθηκαν κατά τη διάρκεια ενός χρονικού διαστήματος, και αυτές πλέον αποτελούν το δείγμα. Η δεύτερη προσέγγιση είναι χρήσιμη, προκειμένου να μπορούμε να βγάλουμε συμπεράσματα σχετικά με ολόκληρο το σετ των μονάδων για μια συγκεκριμένη περίοδο, το οποίο μπορεί να αποτελέσει έναν σημαντικό στόχο για την ανάλυση Φάσης I. Όμως με αυτή την προσέγγιση, χάνεται μέρος της χρονικής ακολουθίας των δεδομένων και έτσι καθίσταται δύσκολη η ανίχνευση συγκεκριμένου τύπου μετατοπίσεων.

Οι ερευνητές θα πρέπει να δώσουν μεγάλη προσοχή στις μεθόδους δειγματοληψίας, ώστε να συλλέγουν δεδομένα που να μπορούν να δώσουν σημαντικές πληροφορίες σχετικά με την διεργασία. Οι Doganaksoy and Hahn (2012) πρότειναν μια διαδικασία με πέντε βήματα προκειμένου να συλλέξουμε τα «σωστά» δεδομένα και συζήτησαν τις πρακτικές δυσκολίες που πρέπει να ληφθούν υπόψη, συμπεριλαμβάνοντας τα επιπλέον κόστη και θέματα ιδιοκτησίας των δεδομένων. Οι Perla et al. (2013) παρείχαν καθοδήγηση για καθορισμό δειγμάτων σε ιατροφαρμακευτικές εφαρμογές. Οι Anderson-Cook and Borror (2013) και Vining (2013) θεώρησαν στρατηγικές συλλογής δεδομένων για ποιοτική παρακολούθηση και ανάλυση.

Είναι σημαντικό να τονίσουμε ότι οι περισσότερες στατιστικές μέθοδοι στη Φάση I μπορούν να εφαρμοστούν μόνο σε ομαδοποιημένες παρατηρήσεις, όμως, ποιοτικά χαρακτηριστικά της διαδικασίας παρατηρούνται συχνά και καταγράφονται και σε μεμονωμένες παρατηρήσεις. Στη βιβλιογραφία δεν γίνεται σαφής αναφορά για τα πλεονεκτήματα και μειονεκτήματα της ανάλυσης Φάσης I σε ομαδοποιημένες ή μεμονωμένες παρατηρήσεις. Σε μεγάλο βαθμό η απόφαση εξαρτάται από την εφαρμογή που μας ενδιαφέρει.

2.1.2 Μέγεθος δείγματος στη Φάση I

Σε κάποιες εφαρμογές μπορεί να μην είναι αναγκαίο να συλλέξουμε καινούργια δεδομένα για την ανάλυση Φάσης I γιατί υπάρχουν διαθέσιμες παρατηρήσεις από ιστορικά δεδομένα. Όμως, είναι συχνά δύσκολο να καθορίσει κανείς εντός ελέγχου δείγματα από ιστορικά δεδομένα μεγάλης έκτασης (βλέπε Zhang et al. (2010)). Είτε συγκεντρώνουμε δεδομένα ειδικά για τη Φάση I είτε χρησιμοποιούμε ένα ιστορικό σετ δεδομένων, ο αριθμός των παρατηρήσεων που είναι διαθέσιμος για την ανάλυση Φάσης I παίζει σημαντικό ρόλο στην εκτίμηση των τιμών των εντός ελέγχου παραμέτρων. Ένας στόχος μείζονος σημασίας στην ανάλυση Φάσης I είναι να εκτιμήσουμε τη μεταβλητότητα της διεργασίας, κάτι που είναι δύσκολο σε δείγματα μικρού μεγέθους και απαιτεί μεγαλύτερα δείγματα από εκείνα για την

εκτίμηση του μέσου. Η ακριβής εκτίμηση των παραμέτρων της διεργασίας είναι ένα πολύ κρίσιμο σημείο προκειμένου να κατασκευάσουμε συγκεκριμένα διαγράμματα Φάσης II. Οι Jensen et al. (2006) έδωσαν μια επισκόπηση της βιβλιογραφίας σχετικά με την επίδραση του σφάλματος εκτίμησης στα διαγράμματα ελέγχου Φάσης II. Το γενικό συμπέρασμα μετά από αρκετές μελέτες πάνω σε αυτό το θέμα, είναι ότι το μέγεθος του δείγματος στη Φάση I, πρέπει συχνά να είναι αρκετά μεγάλο προκειμένου να εξασφαλίσουμε ότι το διάγραμμα ελέγχου που θα κατασκευάσουμε θα προσεγγίζει αρκετά το διάγραμμα ελέγχου που θα προέκυπτε αν οι εντός ελέγχου παράμετροι ήταν γνωστές. Ο όγκος των δεδομένων της Φάσης I μπορεί σε ορισμένες περιπτώσεις να είναι απαγορευτικά μεγάλος (για παράδειγμα δείτε Albers and Kallenberg (2004), Zhang et al. (2013) και Saleh et al. (2014)). Τα προαπαιτούμενα σχετικά με το μέγεθος του δείγματος διαφέρουν για τη Φάση I και τη Φάση II, καθώς εξαρτώνται από τον τύπο του διαγράμματος ελέγχου αλλά και από τη διάσταση των δεδομένων.

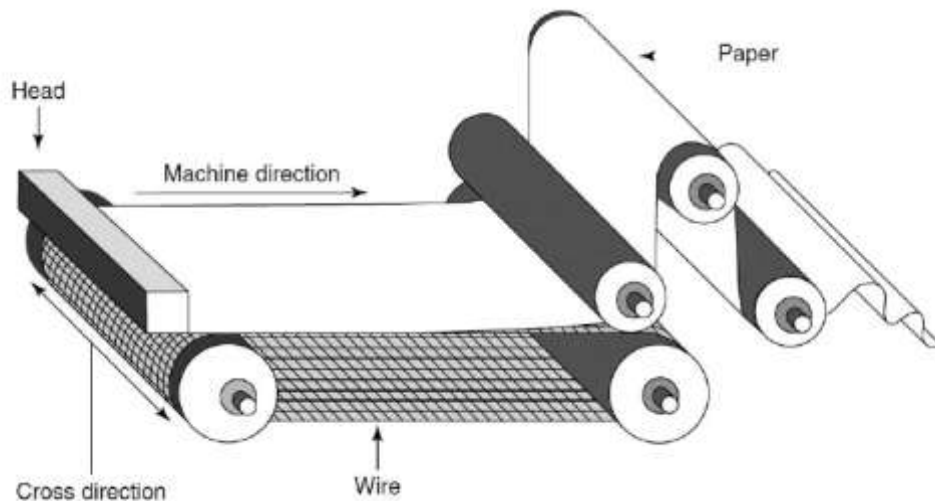
Συμπληρωματικά με το μέγεθος του δείγματος, είναι επίσης σημαντικό να λάβουμε υπόψη μας και το χρονικό εύρος μέσα στο οποίο συγκεντρώσαμε αυτά τα δεδομένα. Οι διεργασίες έχουν μακροχρόνια και βραχυχρόνια χαρακτηριστικά. Τα δείγματα της Φάσης I θα πρέπει να ληφθούν από μια αρκετά μεγάλη χρονική περίοδο για να αξιολογήσουμε και τη βραχυχρόνια αλλά και τη μακροχρόνια μεταβλητότητα του μέσου. Αυτή είναι μια θεμελιώδης αρχή όσο αφορά τη συλλογή δεδομένων για τη μείωση της μεταβλητότητας (δείτε, Steiner and MacKay (2005)).

2.1.3 Γραφικές μέθοδοι και διαγράμματα multivari

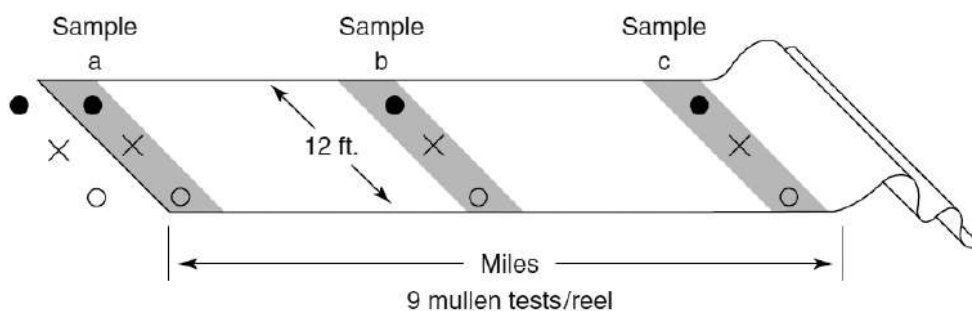
Το πρώτο βήμα σε κάθε ανάλυση δεδομένων είναι η γραφική απεικόνισή τους που περιλαμβάνει ιστογράμματα, διαγράμματα χρονοσειρών και διαγράμματα διασποράς για πολυμεταβλητά δεδομένα. Ένα άλλο πολύ χρήσιμο διάγραμμα είναι το multivari διάγραμμα το οποίο εισήχθη από τον Seder (1950a, 1950b) ως γραφική μέθοδος για την μελέτη των πηγών μεταβλητότητας της διεργασίας. Στην πιο κοινή μορφή τους τα διαγράμματα multivari χρησιμοποιούν εικονογράμματα (pictograma) για τη γραφική παρουσίαση των πολλαπλών πηγών μεταβλητότητας. Ο Shainin (2008) παρουσίασε μια ενδελεχή εικόνα των διαγραμμάτων multivari και αποτύπωσε τη χρήση τους με δύο μελέτες περιπτώσεων (case studies).

Σε ένα παράδειγμα, ο Shainin (2008) κάνει λόγο για τη χρήση ενός multivari διαγράμματος για την μελέτη της μεταβλητότητας της αντοχής του χαρτιού που χρησιμοποιείται για

συσκευασία. Οι πηγές μεταβλητότητας της αντοχής του, έχουν να κάνουν με την «από μηχανή σε μηχανή» μεταβλητότητα, τη μεταβλητότητα ως προς την κατεύθυνση της μηχανής και την από παρτίδα σε παρτίδα μεταβλητότητα ως αποτέλεσμα των πρώτων υλών. Οι Εικόνες 1 και 2 δείχνουν την μηχανή χαρτιού και τα σημεία δειγματοληψίας για τις μετρήσεις της αντοχής του χαρτιού στην έκτασή του αλλά και ως προς την κατεύθυνση της μηχανής.



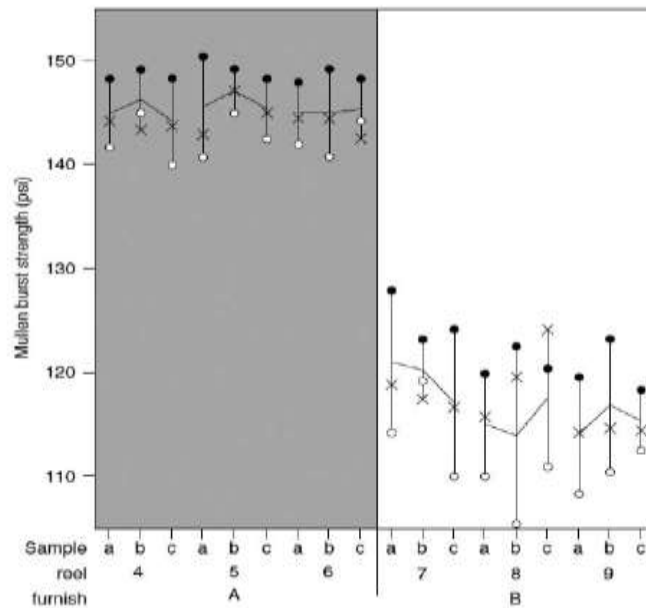
Εικόνα 1. Παρουσίαση της μηχανής χαρτιού (Shainin(2008)).



Εικόνα 2. Παρουσίαση των σημείων δειγματοληψίας για το Multivari διάγραμμα (Shainin(2008)).

Η Εικόνα 3 δίνει το multivari διάγραμμα για την παραγωγή χαρτιού. Όπως ανέφερε ο Shainin (2008) υπήρχαν κάποιες διαφορές κατά μήκος του χαρτιού και της μηχανής, αλλά το μεγαλύτερο μέρος της μεταβλητότητας της αντοχής προκύπτει από τις διαφοροποιήσεις της κάθε παρτίδας. Η μεταβλητότητα που προκύπτει από παρτίδα σε παρτίδα φαίνεται από τις

διαφοροποιήσεις στις μετρήσεις της αντοχής μεταξύ της σκιασμένης περιοχής της Εικόνας 3 που αντιστοιχεί στην Παρτίδα Α και την ασκιαγράφητη περιοχή που αντιστοιχεί στην Παρτίδα Β. Από την Εικόνα 3 φαίνεται ότι οι παρτίδες διαφέρουν, όχι μόνο ως προς τον μέσο αλλά και ως προς τη διασπορά. Αν η πηγή της από παρτίδα σε παρτίδα μεταβλητότητας μπορεί να απαλειφθεί, τότε θα ήταν συνετό να συλλέξουμε περισσότερα δεδομένα προκειμένου να καθορίσουμε ένα κατάλληλο σχέδιο δειγματοληψίας στη Φάση II.



Εικόνα 3. Παρουσίαση του Multivari διαγράμματος για την παραγωγή χαρτιού (Shainin (2008)).

Γραφικές μέθοδοι όπως τα multivari διαγράμματα μπορούν να χρησιμοποιηθούν στα αρχικά στάδια ελέγχου της διεργασίας προκειμένου να μάθουμε περισσότερα στοιχεία σχετικά με αυτή. Βοηθούν τον αναλυτή να βρει λύσεις και να βελτιώσει τη διεργασία, έτσι ώστε να μειώσει τη μεταβλητότητα και να φέρει τη διεργασία εντός ελέγχου.

2.1.4 Επιλογή του κατάλληλου εντός ελέγχου μοντέλου

Μια από τις μεγαλύτερες προκλήσεις στη Φάση I είναι η ανάγκη να αποτιμήσουμε τη σταθερότητα της διεργασίας χωρίς να έχουμε υποθέσει κάποιο μοντέλο για αυτή. Δεν είναι ξεκάθαρο τι από τα δύο προηγείται, αφού από την μία πρέπει να αποτιμήσουμε τη σταθερότητα της διεργασίας αλλά από την άλλη για να το κάνουμε αυτό απαιτείται να υποθέσουμε κάποιο ικανοποιητικό μοντέλο για αυτή. Το να καθορίσουμε ένα κατάλληλο μοντέλο είναι πρόωρο αν η διεργασία δεν είναι σταθερή. Συνεπώς, ο αναλυτής θα πρέπει να καθορίσει το μοντέλο και ταυτόχρονα να εκτιμήσει και τη σταθερότητα της διεργασίας. Δεν

υπάρχει κάποιο ξεκάθαρο σχέδιο δράσης για αυτό το δύσκολο πρόβλημα, αλλά έχουν επισημανθεί κάποια πρακτικά σημεία που μπορούμε να λάβουμε υπόψη μας. Ο Wheeler (2011) πρότεινε τη χρήση μεμονωμένων διαγραμμάτων κινητού μέσου χωρίς την επιλογή κάποιου μοντέλου. Αυτή η προσέγγιση μπορεί να χρησιμοποιηθεί αλλά δεν είναι πάντα η καταλληλότερη επιλογή.

Κάτι που θα πρέπει να θυμόμαστε είναι ότι η σταθερότητα της διεργασίας σε μεγάλο βαθμό εξαρτάται από την «εικόνα» της. Μια διεργασία θα μπορούσε να είναι σταθερή δεδομένων κάποιων ποιοτικών χαρακτηριστικών και με συγκεκριμένο σχέδιο δειγματοληψίας, αλλά ασταθής δεδομένων κάποιων άλλων επιλογών. Κάτι τέτοιο μπορεί να προκαλέσει σύγχυση στην πρακτική σε βιομηχανικό επίπεδο.

Πολλές παραμετρικές στατιστικές μέθοδοι στη Φάση I έχουν στατιστική συμπεριφορά η οποία είναι αρκετά ευαίσθητη σε αποκλίσεις από τις υποθέσεις του μοντέλου (βλέπε για παράδειγμα Jones-Farmer et al. (2009)). Για συνεχείς τυχαίες μεταβλητές υποθέτουμε συνήθως την κανονική κατανομή. Αν πρόκειται να χρησιμοποιήσουμε ένα \bar{X} διάγραμμα στη μονοδιάστατη περίπτωση ή ένα Hotelling's T^2 διάγραμμα στην πολυδιάστατη περίπτωση, αναδρομικά, είναι σημαντικό να εκτιμήσουμε την κατανομή των δεδομένων. Για συνεχείς πολυδιάστατες μεταβλητές, τα διαγράμματα ελέγχου Φάσης I δεν είναι ανθεκτικά σε αποκλίσεις από την κανονικότητα (βλέπε για παράδειγμα Bell et al. (2014)). Τα διαγράμματα Q-Q ή άλλες μέθοδοι μπορούν να χρησιμοποιηθούν για να σιγουρευτούμε ότι η υπόθεση της κανονικότητας ισχύει, όμως, οι εκτός ελέγχου τιμές μπορεί να μας δώσουν την εντύπωση ότι υπάρχουν αποκλίσεις από την κανονικότητα ενώ στη πραγματικότητα δεν υπάρχουν. Αν η υπόθεση κανονικότητας δεν ισχύει, θα μπορούσαμε να προβούμε σε μετασχηματισμό των δεδομένων ή σε χρήση μη παραμετρικών μεθόδων.

Συμπληρωματικά, θα πρέπει να ελέγξουμε τα δεδομένα για αυτοσυσχέτιση και υπερσκέδαση. Στο παραδοσιακό μοντέλο στατιστικού ελέγχου, η τιμή οποιασδήποτε παραμέτρου υποθέτουμε ότι είναι σταθερή στο χρόνο και οι παρατηρήσεις υποθέτουμε ότι είναι ασυσχέτιστες. Σε πολλές εφαρμογές τα δεδομένα συλλέγονται με τόσο γρήγορο ρυθμό που η υπόθεση της ανεξαρτησίας στο χρόνο δεν ισχύει και οι αυτοσυσχετίσεις θα πρέπει να μας απασχολήσουν. Για αυτοσυσχετιζόμενες διεργασίες οι αναλυτές θα πρέπει να παρατηρήσουν το επίπεδο της διεργασίας και τα κατάλοιπα από ένα κατάλληλο μοντέλο χρονοσειρών. Το πρόβλημα της υπερσκέδασης προκύπτει όταν υψηλή μεταβλητότητας εμφανίζεται μεταξύ των δειγμάτων. Αυτό μπορεί να συνεπάγεται ότι το σχέδιο δειγματοληψίας που επιλέξαμε δεν

είναι κατάλληλο ή ότι οι παράμετροι για κάποιο λόγο μεταβάλλονται στο χρόνο. Επίσης θα μπορούσε να χρησιμοποιηθεί μια πολυσταδιακή προσέγγιση στην οποία το αποτέλεσμα της διεργασίας θα παρακολουθείται μόνο αφού προσαρμοστεί στις μεταβλητές που εισέρχονται στη διεργασία.

2.1.5 Μέτρα απόδοσης

Τα μέτρα απόδοσης ενός διαγράμματος ελέγχου στη Φάση II, όπως το μέσο μήκος ροής (ARL) και ο μέσος χρόνος σήματος (ATS) δεν δίνουν πληροφορίες σχετικές με την εκτίμηση, σύγκριση και σχεδίαση μεθόδων στη Φάσης I. Συχνά, οι μέθοδοι της Φάσης I συγκρίνονται με βάση τη συνολική πιθανότητα εντοπισμού κάποιων συγκεκριμένων εκτός ελέγχου συνθηκών όπως μια απομακρυσμένη τιμή ή μιας συνεχής αλλαγή μιας εξεταζόμενης μεταβλητής. Η εντός ελέγχου λειτουργία χαρακτηρίζεται από μια ολική πιθανότητα λανθασμένου συναγερμού (FAP, False Alarm Probability), η οποία ορίζεται ως το σήμα συναγερμού που προκύπτει από τη μέθοδο που χρησιμοποιείται για την ανάλυση του σετ των ιστορικών δεδομένων Φάσης I όταν η διεργασία είναι στην πραγματικότητα σταθερή.

Όταν οι ερευνητές αποτιμούν τις μεθόδους της Φάσης I χρησιμοποιούν πιθανότητες συναγερμού ή την πιθανότητα παρατήρησης τουλάχιστον ενός συναγερμού όταν η διεργασία βρίσκεται εκτός ελέγχου. Αυτή η μέθοδος έχει χρησιμοποιηθεί από αρκετούς ερευνητές, για παράδειγμα από τους Sullivan and Woodall (1996), Vargas (2003), Woodall et al. (2004), Jensen et al. (2007), Alfaro and Ortega (2008), Jobe and Pokojovy (2009), Jones – Farmer et al. (2009) και Graham et al. (2010). Σε καταστάσεις εκτός ελέγχου είναι πιο χρήσιμο να αποτιμήσουμε τις μεθόδους Φάσης I βασιζόμενοι στην ικανότητά τους να αναγνωρίζουν σωστά τις ακραίες τιμές. Η δεύτερη προσέγγιση προτάθηκε από τους Shiau and Sun (2009) και Chen et al. (2014). Επιπλέον, οι Bell et al. (2014) χρησιμοποίησαν την πιθανότητα σωστού εντοπισμού για να εκτιμήσουν την απόδοση ενός ειδικού πολυμεταβλητού διαγράμματος για το μέσο βαθμό (mean rank) που εισήγαγαν.

Το μεγαλύτερο πρόβλημα που προκύπτει όταν χρησιμοποιούμε αναδρομικά διαγράμματα ελέγχου είναι ότι αυξάνονται σε ακραίο βαθμό οι τιμές του FAP. Οι Chakraborti et al. (2009) ανέπτυξαν αυτό το θέμα λεπτομερώς. Όταν υπάρχουν αρκετοί λανθασμένοι συναγερμοί, είναι δύσκολο να εμπιστευτείς την σημαντικότητα των σημάτων. Όταν οι εντός ελέγχου παράμετροι υποτεθούν ότι είναι γνωστές και οι παρατηρήσεις είναι ανεξάρτητες, οι διαδοχικές συγκρίσεις της στατιστικής συνάρτησης που απεικονίζεται στο διάγραμμα Shewhart με τα όρια ελέγχου είναι ανεξάρτητες. Σε αυτή την περίπτωση όταν υπάρχει μια

σταθερή πιθανότητα λανθασμένου συναγερμού για κάθε δεδομένη στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα, έστω p , έχουμε ότι

$$FAP = 1 - (1 - p)^m$$

όπου m είναι ο αριθμός των σημείων του διαγράμματος. Για παράδειγμα, αν μια διεργασία κατανέμεται κανονικά με άγνωστες εντός ελέγχου παραμέτρους και χρησιμοποιηθούν 3σ όρια ελέγχου, η πιθανότητα ενός τουλάχιστον λανθασμένου συναγερμού σε ένα \bar{X} διάγραμμα με $m = 30$ σημεία είναι $FAP = 0.0780$. Με $m = 50$ έχουμε $FAP = 0.126$. Γενικά, όσο αυξάνεται το m , αυξάνεται και το FAP . Παρόμοια, όταν οι παράμετροι είναι άγνωστες, το FAP των αναδρομικών διαγραμμάτων μπορεί να παραμείνει αυξημένο πάνω από το επιθυμητό επίπεδο όσο ο αριθμός των δειγμάτων αυξάνεται.

2.2 Φάση I-Στατιστικός Έλεγχος Διεργασιών

Όπως έχουμε αναφέρει το μεγαλύτερο μέρος της προκαταρκτικής (στατιστικής) ανάλυσης γίνεται στη Φάση I, η οποία περιλαμβάνει το σχεδιασμό, τη διαχείριση, τη σχεδίαση της μελέτης, τη συλλογή δεδομένων, τη διαχείρισή τους και τη διερευνητική δουλειά (συμπεριλαμβανομένης της γραφικής και αριθμητικής ανάλυσης, ανάλυση καλής προσαρμογής, κ.λπ.) προκειμένου να σιγουρευτούμε ότι η διεργασία είναι στη πραγματικότητα σε κατάσταση στατιστικού ελέγχου. Μπορούμε να ανατρέξουμε στους Woodall (2000) και Montgomery (2005) για περαιτέρω συζήτηση σχετικά με αυτά τα θέματα. Στόχος μας είναι να σιγουρευτούμε ότι η διεργασία λειτουργεί γύρω από μια αποδεκτή τιμή στόχο κάτω από κάποιες συνήθειες (κοινές) αιτίες μεταβλητότητας και δεν παρουσιάζονται ειδικές αιτίες ή ανησυχίες για αυτή. Γνωρίζουμε ότι η ανάλυση Φάσης I είναι μια επαναληπτική διαδικασία στην οποία τα διαγράμματα ελέγχου παίζουν πολύ σημαντικό ρόλο. Τα όρια ελέγχου που παίρνουμε στο αρχικό στάδιο της φάσης αυτής θεωρούνται δοκιμαστικά όρια και συχνά αναθεωρούνται και βελτιώνονται με σκοπό να έρθει η διεργασία εντός ελέγχου.

Αν οι τιμές των παραμέτρων που ενδιαφερόμαστε να εξετάσουμε είναι γνωστές ή δίνονται (συχνά η κατάσταση αυτή αναφέρεται ως Περίπτωση με Γνωστό Πρότυπο (standards known case) ή **Περίπτωση K** (CaseK)), χρειάζεται να βεβαιωθούμε ότι η διεργασία βρίσκεται ή κινείται γύρω από τη δοθείσα τιμή στόχο υπό την παρουσία φυσικών αιτιών μεταβλητότητας. Αντίθετα, αν οι παράμετροι είναι άγνωστες, η εφαρμογή ελέγχου συνεπάγεται εκτίμηση των παραμέτρων καθώς επίσης και των ορίων ελέγχου. Μια τέτοια κατάσταση αναφέρεται ως

Περίπτωση με Άγνωστο Πρότυπο (standards unknown case) ή **Περίπτωση U** (Case U). Και οι δυο αυτές περιπτώσεις (Case K, Case U) μπορούν να εφαρμοστούν στην πράξη, αλλά η Case U παρουσιάζεται πιο συχνά, αφού συνήθως δεν έχουμε στη διάθεσή μας τις απαραίτητες ιστορικές πληροφορίες.

Το πρόβλημα απόφασης υπό το σενάριο του διαγράμματος ελέγχου Φάσης I, είναι θεωρητικά παρόμοιο με το πρόβλημα ομογένειας ενός τεστ πολλαπλών δειγμάτων όπου εξετάζουμε αν τα δεδομένα που προέρχονται από διαφορετικές ομάδες προέρχονται από την ίδια κατανομή (εντός ελέγχου διεργασίας). Κάτι τέτοιο έχει επισημανθεί και από τους Champ and Jones (2004). Παρόλα αυτά, το πρόβλημα της ταυτόχρονης σύγκρισης ενός αριθμού στατιστικών συναρτήσεων που απεικονίζονται σε ένα διάγραμμα με το ίδιο σύνολο ορίων ελέγχου είναι παρόμοιο με αυτό που συναντάμε στην βιβλιογραφία ως «πολλαπλός έλεγχος» (multiple testing). Έτσι, η πιθανότητα λανθασμένου συναγερμού (FAP) η οποία είναι η (συνολική) πιθανότητα τουλάχιστον ενός λανθασμένου συναγερμού, χρησιμοποιείται για την κατασκευή (σχεδίαση) και αποτίμηση των διαγραμμάτων ελέγχου Φάσης I. Συνεπώς, ένα διάγραμμα ελέγχου Φάσης I σχεδιάζεται προσδιορίζοντας μια ονομαστική τιμή FAP, η οποία συμβολίζεται ως FAP_0 , που τυπικά παίρνει τις τιμές 0.01, 0.05 ή 0.10. Θα μπορούσε να σημειωθεί ότι αυτή η προσέγγιση διαφέρει από τη δημιουργία διαγραμμάτων ελέγχου Φάσης II, που βασίζονται σε εντός ελέγχου δεδομένα, όπου τυπικά προσδιορίζονται κάποια γνωρίσματα της εντός ελέγχου κατανομής του μήκους ροής, όπως το μέσο μήκος ροής, προκειμένου να καθοριστούν τα όρια ελέγχου.

Όπως αναφέρει ο Montgomery (2005), τα διαγράμματα τύπου Shewhart ταιριάζουν σε εφαρμογές της Φάσης I καθώς είναι διαγράμματα ελέγχου γενικού σκοπού που μπορούν να εντοπίσουν επιτυχώς μεγάλες μετατοπίσεις. Με βάση τη βιβλιογραφία, για τη χρήση διαγραμμάτων ελέγχου τύπου Shewhart υποθέτουμε αρχικά κανονικότητα.

Υπάρχουν κατά κύριο λόγο δυο μέθοδοι για την κατασκευή διαγραμμάτων Φάσης I στην Case U.

1η μέθοδος: Η Hillier (1969) και οι Yang and Hillier (1970) πρότειναν μια μεθοδολογία στην οποία το FAR, το οποίο ορίζεται ως η πιθανότητα λανθασμένου συναγερμού σε κάθε στάδιο δειγματοληψίας (με κάθε δείγμα), ρυθμίζεται σε κάποιο επιθυμητό επίπεδο (για παράδειγμα 0.0027) προκειμένου να καθορίσουμε τα όρια ελέγχου. Παρότι έλαβαν υπόψη τις επιδράσεις της εκτίμησης παραμέτρων, δεν μελέτησαν λεπτομερώς

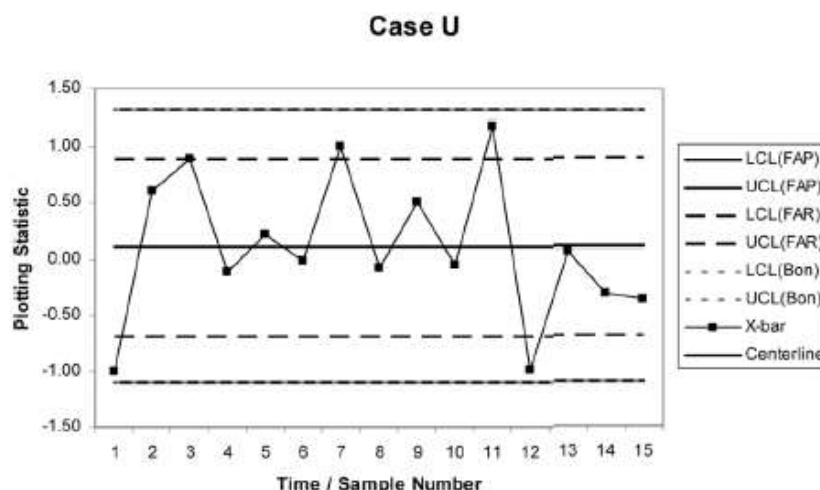
(i) το γεγονός ότι τα δείγματα συγκρίνονται όλα με τα ίδια όρια ελέγχου και έτσι τα σήματα ή οι περιπτώσεις συναγερμού είναι στατιστικά εξαρτημένες, και

(ii) ότι πολλά δείγματα συγκρίνονται με τα όρια ελέγχου την ίδια χρονική στιγμή (έτσι προκύπτει το θέμα των ταυτόχρονων συγκρίσεων που πρέπει να λάβουμε υπόψη μας).

Όπως έχει τονιστεί από πολλούς ερευνητές, αυτά περιορίζουν τη δυνατότητα εφαρμογής των μεθόδων αυτών στη Φάση I, καθώς θα υπάρξουν πολλοί λανθασμένοι συναγερμοί στα διαγράμματα που κατασκευάστηκαν με σταθερό FAR.

2η μέθοδος: Υπό αυτή τη μέθοδο που μπορεί να αποδοθεί στον King (1954), τα όρια ελέγχου υπολογίζονται για συγκεκριμένο FAR. Η βασική ιδέα είναι να λάβουμε υπόψη μας το γεγονός ότι τα ενδεχόμενα που οδηγούν σε συναγερμό είναι εξαρτημένα και επίσης να διαχειριστούμε το θέμα του «ταυτόχρονου». Αυτή η μέθοδος προτιμάται πρακτικά και κυρίως επικεντρωνόμαστε σε αυτή και σε παρόμοιες διαδικασίες στη συνέχεια.

Οι δυο μέθοδοι παρουσιάζονται στην Εικόνα 4 για την περίπτωση Case U για ένα σύνολο τεχνητών δεδομένων. Φαίνεται ότι μπορεί να υπάρχουν περισσότεροι λανθασμένοι συναγερμοί αν τα όρια ελέγχου βασίζονται στο κριτήριο FAR, δηλαδή στην 1η μέθοδο. Αυτό συμβαίνει επειδή τα όρια ελέγχου που βασίζονται στο FAR είναι πιο στενά σε σχέση με αυτά που βασίζονται στο FAP. Συνεπώς, στη Φάση I εστιάζουμε σε μεθόδους που χρησιμοποιούν το FAP.



Εικόνα 4. Διαγράμματα ελέγχου Φάσης I για CaseU (Chakraborti et al. (2008))

Άλλες μέθοδοι διαγραμματικού ελέγχου για την CaseU περιλαμβάνουν για παράδειγμα

- (i) Q-charts (Quesenberry, 1991),
- (ii) διαγράμματα ελέγχου που χρησιμοποιούν ακολουθιακά σχέδια δειγματοληψίας (βλέπε Zhangetal. 2006), και
- (iii) τα model-based διαγράμματα ελέγχου (Koning, 2006).

2.2.1 Υπόβαθρο

Υποθέτουμε ότι έχουμε στη διάθεσή μας m (>1) ανεξάρτητα τυχαία δείγματα, το καθένα με μέγεθος n (>1). Τα δεδομένα παρουσιάζονται στη μορφή X_{ij} , η οποία υποδηλώνει τη j παρατήρηση από το i δείγμα για $i = 1, 2, \dots, m$ και $j = 1, 2, \dots, n$. Υποθέτουμε επίσης η τ.μ. X_{ij} είναι συνεχής (π.χ. κανονική) και η κατανομή της είναι συνάρτηση των άγνωστων παραμέτρων $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Έτσι για παράδειγμα, εάν η διεργασία κατανέμεται κανονικά με άγνωστο μέσο μ και άγνωστη διακύμανση σ^2 , έχουμε $\theta = (\theta_1 = \mu, \theta_2 = \sigma^2)$.

Ένα διάγραμμα ελέγχου αποτελείται από μια στατιστική συνάρτηση που απεικονίζεται σε αυτό και κάποια όρια ελέγχου. Διαισθητικά, η στατιστική συνάρτηση C_i για το i δείγμα επιλέγεται με τέτοιο τρόπο έτσι ώστε να αποτελεί αποτελεσματικό εκτιμητή της παραμέτρου που μας ενδιαφέρει. Έτσι για παράδειγμα, εάν ο μέσος της διεργασίας είναι η παράμετρος που μας ενδιαφέρει, τότε ο δειγματικός μέσος \bar{X}_i , είναι ο καλύτερος εκτιμητής, και είναι μια λογική υποψήφια στατιστική συνάρτηση για να απεικονιστεί διάγραμμα ελέγχου. Επικεντρωνόμαστε στα δίπλευρα διαγράμματα – έτσι, είτε μια πολύ χαμηλή είτε μια πολύ υψηλή τιμή του στατιστικής συνάρτησης μπορεί να προκαλέσει ανησυχία – αλλά και τα μονόπλευρα διαγράμματα μπορούν να χρησιμοποιηθούν με ανάλογο τρόπο. Η εκτίμηση των παραμέτρων είναι ένα σημαντικό βήμα για τη δημιουργία ενός διαγράμματος ελέγχου. Γενικά προτιμώνται οι αμερόληπτοι εκτιμητές και αν περισσότεροι από ένας τέτοιου είδους εκτιμητές είναι διαθέσιμοι, θα επιλέξουμε αυτόν με την μικρότερη διακύμανση (minimum variance unbiased (MVU) estimator). Έστω $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$ οι εκτιμητές των άγνωστων παραμέτρων. Τα εκτιμώμενα κάτω όριο ελέγχου ($L\hat{C}L$), κεντρική γραμμή ($\hat{C}L$) και άνω όριο ελέγχου ($U\hat{C}L$) είναι προφανώς συναρτήσεις του $\hat{\theta}$. Αυτά τα όρια ελέγχου βασίζονται στην δειγματοληπτική κατανομή της στατιστικής συνάρτησης που απεικονίζεται στο διάγραμμα και κάποιο κριτήριο απόδοσης του διαγράμματος. Ένα διάγραμμα ελέγχου με όλες τις στατιστικές συναρτήσεις (από τα m δείγματα) μαζί με τα (εκτιμημένα) όρια ελέγχου αποτελούν το διάγραμμα ελέγχου Φάσης I.

Εάν όλες οι στατιστικές συναρτήσεις βρίσκονται μεταξύ των ορίων ελέγχου, δηλαδή αν $L\hat{C}L < C_i < U\hat{C}L$, για $i = 1, 2, \dots, m$ και δεν παρουσιάζεται κανένα συστηματικό μοτίβο, η διεργασία θεωρείται εντός ελέγχου (in control (IC)). Αντίθετα, αν κάποιο C_i πέσει πάνω ή έξω από τα εκτιμημένα όρια ελέγχου, δηλαδή αν $C_i \leq L\hat{C}L$ ή $C_i \geq U\hat{C}L$, για τουλάχιστον ένα i , η διεργασία θεωρείται ότι βρίσκεται εκτός ελέγχου (out-of-control (OOC)) και θα πρέπει να παρέμβουμε. Όπως αναφέραμε νωρίτερα, τα διαγράμματα ελέγχου Φάσης I αποτελούν μια επαναληπτική διαδικασία όπου τα δείγματα που προκαλούν πρόβλημα (out-of-control, OOC) πρώτα ερευνώνται και στη συνέχεια είναι πιθανό να εξαιρεθούν. Στη συνέχεια, με τα εναπομείναντα δείγματα οι παράμετροι εκτιμώνται ξανά, τα όρια ελέγχου υπολογίζονται εκ νέου και η διαγραμματική διαδικασία επαναλαμβάνεται. Αυτή η μέθοδος δοκιμής και σφάλματος (trial and error) συνεχίζεται μέχρι το σημείο όπου όλες οι εναπομείναντες στατιστικές συναρτήσεις να βρίσκονται ανάμεσα στα πιο πρόσφατα όρια ελέγχου, και καταλήγοντας στο συμπέρασμα ότι η διεργασία είναι εντός ελέγχου. Τα δεδομένα στα οποία καταλήξαμε συχνά αναφέρονται ως εντός ελέγχου δεδομένα ή δεδομένα αναφοράς, από τα οποία κάθε απαραίτητη παράμετρος εκτιμάται και χρησιμοποιείται για να βρεθούν τα κατάλληλα όρια ελέγχου τα οποία χρησιμοποιούνται για την μελλοντική παρακολούθηση στη Φάση II. Υποθέτουμε ότι το m δηλώνει τον τελικό αριθμό των δειγμάτων αναφοράς μετά από μια επιτυχημένη ανάλυση Φάσης I. Έτσι, τα δεδομένα αναφοράς θεωρείται ότι έχουν $N = mn$ το πλήθος παρατηρήσεις.

Όταν η παρατηρούμενη τιμή μιας στατιστικής συνάρτησης βρεθεί είτε πάνω από το άνω όριο ελέγχου, είτε χαμηλότερα από το κάτω όριο ελέγχου έχουμε μία εκτός ελέγχου κατάσταση. Αυτό το γεγονός καλείται «σήμα» (signal) ή «γεγονός σήματος» (signaling event). Για το σκοπό μας είναι πιο βολικό να θεωρήσουμε το συμπληρωματικό γεγονός, δηλαδή όταν το δείγμα δεν δίνει σήμα «γεγονός μη-σήματος» (non-signaling event). Έτσι, για το i δείγμα, το ενδεχόμενο $E_i = \{L\hat{C}L < C_i < U\hat{C}L\}$ δηλώνει ένα γεγονός μη-σήματος. Όπως συζητήσαμε νωρίτερα, το FAP είναι το ενδεδειγμένο κριτήριο σχεδίασης ενός διαγράμματος ελέγχου Φάσης I. Η πιθανότητα FAP μπορεί να εκφραστεί ως ακολούθως :

FAP = (τουλάχιστο ένας λανθασμένος συναγεργμός στα m δείγματα)

= $1 - P(\text{κανένας λανθασμένος συναγεργμός στα } m \text{ δείγματα})$

$$= 1 - P\left(\bigcap_{i=1}^m E_i \mid IC\right)$$

$$\begin{aligned}
&= 1 - P\left(\bigcap_{i=1}^m \{L\hat{C}L < C_i < U\hat{C}L\} \middle| IC\right) \\
&= 1 - \int_a^b \int_a^b \dots \int_a^b f_{C_1, C_2, \dots, C_m}(c_1, c_2, \dots, c_m) dc_1 dc_2 \dots dc_m \quad (1)
\end{aligned}$$

όπου $f_{C_1, C_2, \dots, C_m}(c_1, c_2, \dots, c_m)$ δηλώνει την από κοινού συνάρτηση πυκνότητας πιθανότητας των στατιστικών συναρτήσεων C_1, C_2, \dots, C_m όταν η διεργασία είναι εντός ελέγχου και $a = L\hat{C}L$ και $b = U\hat{C}L$.

Έτσι, το FAP σχετίζεται με m ταυτόχρονα γεγονότα μη-σήματος και με όρια ελέγχου που έχουν εκτιμηθεί (Περίπτωση U). Να σημειώσουμε ότι αυτά τα γεγονότα είναι εξαρτημένα καθώς οι στατιστικές συναρτήσεις συγκρίνονται με το ίδιο ζευγάρι εκτιμημένων ορίων ελέγχου. Συνεπώς, ο υπολογισμός του FAP απαιτεί γνώση της από κοινού κατανομής των στατιστικών συναρτήσεων όταν η διεργασία βρίσκεται εντός ελέγχου. Η εξαγωγή αυτής της κατανομής και ακολούθως ο καθορισμός των ορίων ελέγχου, θέτουν κάποιες πρακτικές προκλήσεις, ειδικότερα όταν ο αριθμός των συγκρίσεων (m) μπορεί συχνά να είναι χαμηλός, έως 25, ή υψηλός, 300 ή παραπάνω.

Στο σημείο αυτό αναφέρουμε ότι ένα κοινό μέτρο που χρησιμοποιούμε στα διαγράμματα ελέγχου είναι ο ρυθμός λανθασμένου συναγερμού (false alarm rate (FAR)) ο οποίος ορίζεται ως ακολούθως :

$$FAR = 1 - P(E_i | IC) \quad (2)$$

Έτσι, το FAR είναι η πιθανότητα λανθασμένου συναγερμού σε κάθε στάδιο δειγματοληψίας (σε κάθε δείγμα) και ο υπολογισμός της πιθανότητας αυτής απαιτεί μόνο την περιθώρια κατανομή της i στατιστικής συνάρτησης C_i όταν η διεργασία είναι εντός ελέγχου. Αυτή η εντός ελέγχου περιθώρια κατανομή είναι τυπικά η ίδια για όλα τα $i = 1, 2, \dots, m$, με αποτέλεσμα το FAR να είναι το ίδιο για όλα τα δείγματα. Για παράδειγμα, για ένα 3σ διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση της μέσης τιμής όταν οι παράμετροι είναι γνωστές, το FAR είναι 0.0027 (κανονική στατιστική συνάρτηση). Ωστόσο, πρέπει να τονίσουμε ότι το FAP είναι το πιο χρήσιμο μέτρο στη Φάση I επειδή με αυτό τον τρόπο τυπικά συγκρίνουμε έναν μεγάλο αριθμό στατιστικών συναρτήσεων ταυτόχρονα έναντι των ίδιων ορίων ελέγχου και αυτό όπως θα ήταν λογικό διογκώνει σε πολύ μεγάλο βαθμό το FAR. Για παράδειγμα, σχετικά με το 3σ διάγραμμα ελέγχου που αναφέρθηκε παραπάνω, αν συγκρίνουμε 25 δείγματα ταυτόχρονα στη Φάση I, όπως γίνεται συχνά στην

πράξη, η πιθανότητα τουλάχιστον ενός λανθασμένου συναγερμού, το FAP δηλαδή, ισούται με $1 - (1 - 0.0027)^{25} = 0.0654$, το οποίο είναι υπερβολικά υψηλό. Όπως είδαμε στην Εικόνα 4, τα 3σ όρια ελέγχου βασισμένα στο FAR είναι πιο στενά και παρατηρούνται περισσότεροι λανθασμένοι συναγερμοί. Στο παράδειγμά μας, προκειμένου το FAP να είναι ίσο με 0.0027, θα πρέπει να χρησιμοποιήσουμε 3.8715σ όρια ελέγχου (πιο πλατιά) και τότε το FAR που προκύπτει να είναι ίσο με $FAR = 1 - (1 - 0.0027)^{1/25} = 0.0001081402$ (προκύπτει από τη λύση της εξίσωσης $1 - (1 - FAR)^{25} = 0.0027$).

Συνοψίζοντας, η εκτίμηση των παραμέτρων (και των ορίων ελέγχου) επηρεάζει και το FAP και το FAR, ωστόσο, είναι πιο λογικό κριτήριο για να ελέγξει κανείς τη Φάση I είναι το FAP, επειδή λαμβάνει υπόψη του θέματα εκτίμησης και πολλαπλότητας (πολλές ταυτόχρονες συγκρίσεις και εξάρτηση).

2.2.2 Εφαρμογή δίπλευρων διαγραμμάτων ελέγχου τύπου Shewhart - Καθορισμός των ορίων ελέγχου

Η εφαρμογή ενός διαγράμματος ελέγχου Φάσης I απαιτεί τον προσδιορισμό μιας επιθυμητής τιμής για το FAP (FAP_0) και στη συνέχεια των αντίστοιχων ορίων ελέγχου. Αυτό σημαίνει ότι χρειάζεται να λύσουμε ως προς δυο σταθερές a και b την εξίσωση

$$FAP_0 = 1 - \int_a^b \int_a^b \dots \int_a^b f_{C_1, C_2, \dots, C_m}(c_1, c_2, \dots, c_m) dc_1 dc_2 \dots dc_m \quad (3)$$

για δοθείσα τιμή FAP_0 (πρακτικά 0.01, 0.05 ή 0.1) του FAP. Ωστόσο σε μερικές περιπτώσεις οι στατιστικές συναρτήσεις κατανέμονται συμμετρικά γύρω από το μηδέν (χωρίς να επηρεάζει αυτό τη γενικότητα της προσέγγισής μας) και αυτό λογικά συνεπάγεται ότι μπορούμε να χρησιμοποιήσουμε συμμετρικά όρια ελέγχου, δηλαδή, θέτουμε $L\hat{C}L = -U\hat{C}L = -d$. Έτσι το d προκύπτει από τη λύση της εξίσωσης

$$FAP_0 = 1 - \int_{-d}^d \int_{-d}^d \dots \int_{-d}^d f_{C_1, C_2, \dots, C_m}(c_1, c_2, \dots, c_m) dc_1 dc_2 \dots dc_m. \quad (4)$$

Εάν οι στατιστικές συναρτήσεις C_i δεν κατανέμονται συμμετρικά και επιθυμούμε δίπλευρα διαγράμματα ελέγχου, ένας τρόπος είναι να χρησιμοποιήσουμε την προσέγγιση των ίσων ουρών στην οποία το μισό FAP_0 αντιστοιχεί σε κάθε ουρά (μισό πάνω από το $U\hat{C}L$ και μισό κάτω από το $L\hat{C}L$, αντίστοιχα). Αυτή η προσέγγιση μπορεί να εξηγηθεί πιο εύκολα με τον επαναπροσδιορισμό του FAP μέσω της σχέσης

$$FAP = P(\min(C_1, C_2, \dots, C_m) \leq a|IC) + P(\max(C_1, C_2, \dots, C_m) \geq b|IC) \quad (5)$$

Η παραπάνω σχέση δικαιολογείται από το γεγονός ότι η εμφάνιση τουλάχιστον ενός λανθασμένου συναγερμού μπορεί ισοδύναμα να εκφραστεί ως η ένωση δύο ξένων ενδεχομένων. Το ένα είναι το $\{\max(C_1, C_2, \dots, C_m) \geq b|IC\}$ και το άλλο είναι το $\{\min(C_1, C_2, \dots, C_m) \leq a|IC\}$. Έτσι τα a, b υπολογίζονται από τις σχέσεις

$$P(\min(C_1, C_2, \dots, C_m) \leq a|IC) = FAP_0/2,$$

$$P(\max(C_1, C_2, \dots, C_m) \geq b|IC) = FAP_0/2.$$

Μια πιο απλή εναλλακτική προσέγγιση, η οποία προσφέρει μια προσεγγιστική λύση και εφαρμόζεται είτε οι στατιστικές συναρτήσεις C_i κατανέμονται συμμετρικά είτε όχι, είναι η χρήση προσέγγισης τύπου Bonferroni, έτσι ώστε το FAP να είναι το πολύ όσο προσδιορίστηκε. Σε αυτή την περίπτωση, τα όρια ελέγχου δίνονται από το $100(FAP_0/2m)$ και το $100[1 - (\frac{FAP_0}{2m})]$ ποσοστιαία σημεία της περιθώριας (μονοδιάστατης) εντός ελέγχου κατανομής μιας στατιστικής συνάρτησης C_i . Να σημειώσουμε ότι αυτή η προσέγγιση δεν χρησιμοποιεί την από κοινού κατανομή των στατιστικών συναρτήσεων C_i . Τα όρια ελέγχου Bonferroni υπολογίζονται και παρουσιάζονται στην Εικόνα 4 όπου φαίνεται να συμπίπτουν με τα όρια ελέγχου που βασίζονται στο FAP που προκύπτουν από τη σχέση (4).

Κλείνοντας την παρούσα παράγραφο σημειώνουμε ότι κάποιος θα μπορούσε να προσεγγίσει τα όρια ελέγχου αγνοώντας την εξάρτηση μεταξύ των γεγονότων σήματος. Στην πραγματικότητα, βλέπουμε ότι σε πολλές περιπτώσεις όταν ο αριθμός των δειγμάτων m , είναι μεγάλος, η συσχέτιση μεταξύ των στατιστικών συναρτήσεων εξαλείφεται. Όταν οι στατιστικές συναρτήσεις είναι ανεξάρτητες, από την (1) μπορούμε εύκολα να δούμε ότι

$$FAP = 1 - \prod_{i=1}^m P(E_i|IC) = 1 - [P(E_i|IC)]^m = 1 - (1 - FAR)^m$$

δηλαδή

$$FAR = 1 - (1 - FAP)^{\frac{1}{m}}$$

και έτσι τα προσεγγιστικά όρια ελέγχου δίνονται από το $100 \left[\frac{FAR_0}{2} \right] = 100 \left[\{1 - (1 - FAP_0)^{\frac{1}{m}}\} / 2 \right]$ και το $100 \left[1 - \frac{FAR_0}{2} \right] = 100 \left[1 - \{1 - (1 - FAP_0)^{\frac{1}{m}}\} / 2 \right]$ ποσοστιαία

σημεία της περιθώριας (μονοδιάστατης) εντός ελέγχου κατανομής μιας στατιστικής συνάρτησης C_i . Για άλλη μια φορά, αυτή η προσέγγιση είναι πιο εύκολη καθώς δεν χρειάζεται να δουλέψουμε με την από κοινού κατανομή των στατιστικών συναρτήσεων C_i , ωστόσο η ακριβής μέθοδος προτιμάται για μικρό m .

2.3 Μέθοδοι που χρησιμοποιούνται στη Φάση I

2.3.1 Έλεγχος ακραίων σημείων στη Φάση I

Όταν χρησιμοποιούμε αναδρομικά διαγράμματα ελέγχου τύπου Shewhart, υπολογίζουμε δοκιμαστικά όρια ελέγχου και διερευνούμε κάθε τιμή που βρίσκεται εκτός των ορίων ελέγχου. Εάν η αιτία για την οποία τα σημεία βρίσκονται εκτός των ορίων ελέγχου προσδιοριστεί και αξιολογηθούν οι ειδικές αιτίες μεταβλητότητας, τότε τα σημεία αυτά αφαιρούνται από τα δεδομένα και τα όρια ελέγχου υπολογίζονται ξανά. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να βεβαιωθούμε ότι όλα τα δεδομένα προέρχονται από μια εντός ελέγχου διεργασία.

Με αυτή την προσέγγιση αναπτύχθηκαν μερικές εύλογες σκέψεις. Αρχικά, οι αναλυτές πολλές φορές αφαιρούν τα εκτός ελέγχου σημεία αυτόματα, χωρίς να προβούν σε κάποια αξιολόγηση ή έρευνα. Αυτή είναι μια ανόητη πρακτική για την ανάλυση Φάσης I, όταν ξέρουμε λίγα πράγματα σχετικά με την διεργασία. Το μοντέλο στο οποίο βασίστηκαν τα διαγράμματα ελέγχου θα μπορούσε να είναι εσφαλμένο και επομένως τα εκτός ορίων ελέγχου σημεία θα μπορούσαν να είναι αντανάκλαση μιας φυσιολογικής μεταβλητότητας των δεδομένων. Ο Montgomery (2013, p.p. 206-207, 238-239) και οι Shiau and Sun (2009) ασχολήθηκαν επίσης με αυτό το θέμα. Ανεξάρτητα με τη μέθοδο ανάλυσης, συστήνεται προσεκτική μελέτη πριν την εξάλειψη δεδομένων στην ανάλυση Φάσης I.

Δεύτερον, όταν αρκετά σημεία βρίσκονται εκτός των ορίων ελέγχου υπάρχει ανησυχία σχετικά με την ευστάθεια της διεργασίας. Όμως, αυτό θα μπορούσε να είναι αντανάκλαση ενός λανθασμένου μοντέλου για την εξεταζόμενη μεταβλητή ή ότι η μέθοδος δειγματοληψίας είναι ανεπαρκής. Για παράδειγμα, όταν χρησιμοποιούμε \bar{X} διάγραμμα ελέγχου στη Φάση I, δεν θεωρούμε απαραίτητα ότι ο μέσος είναι σταθερός και εντός ελέγχου. Μπορούν να υπάρχουν επιπλέον στοιχεία μεταβλητότητας που επηρεάζουν τον μέσο της διεργασίας όπως αναφέρουν οι Woodall and Thomas (1995) και άλλοι. Μια ακόμα λογική εξήγηση για τον μεγάλο αριθμό των εκτός ελέγχου σημείων στα \bar{X} ή X διαγράμματα ελέγχου, μπορεί να είναι η παρουσία θετικής αυτοσυσχέτισης όπως επιδείχθηκε από τους Maragah and Woodall

(1992). Θα πρέπει να αφαιρέσουμε την πηγή αυτής της αυτοσχέτισης αν αυτό είναι δυνατό. Αν η αφαίρεσή της δεν είναι δυνατή, τότε μπορεί κάποιος να εφαρμόσει την διαδικασία ελέγχου των Boxand Luceño (1997) δεδομένου ότι υπάρχει μεταβλητή ελέγχου. Εάν η αυτοσχέτιση δεν μπορεί να εξαλειφθεί από τα δεδομένα, τότε κάθε παρακολούθηση της διεργασίας στη Φάση II θα πρέπει να λάβει υπόψη της αυτό το χαρακτηριστικό των δεδομένων.

2.3.2 Διαγράμματα ελέγχου αυτόματης εκκίνησης

Τα διαγράμματα ελέγχου αυτόματης εκκίνησης (self-starting) είναι διαγράμματα ελέγχου στα οποία χρησιμοποιούνται διαδοχικές παρατηρήσεις για να ενημερωθούν οι εκτιμώμενες παραμέτρους και ταυτόχρονα να ελέγξουν για εκτός ελέγχου σημεία. Ο σκοπός αυτών των διαγραμμάτων είναι να ξεκινήσει η παρακολούθηση της Φάσης II όσο το δυνατό γρηγορότερα με τον ελάχιστο αριθμό δεδομένων. Αυτές οι μέθοδοι παρουσιάστηκαν από τον Hawkins (1987) προκειμένου να μειώσει την ανάγκη για πιθανά δαπανηρό μέγεθος δείγματος στη Φάση I. Τα διαγράμματα ελέγχου αυτόματης εκκίνησης είναι χρήσιμα, όμως, μόνο όταν η συλλογή δεδομένων είναι αργή και τα ιστορικά δεδομένα της διεργασίας είναι ανεπαρκή για να εκτιμηθούν οι εντός ελέγχου παράμετροι της διεργασίας. Μια τέτοια κατάσταση θα μπορούσε να προκύψει, για παράδειγμα, σε μια βιομηχανία μικρού όγκου παραγόμενων προϊόντων.

Με τα μονοδιάστατα διαγράμματα ελέγχου αυτόματης εκκίνησης έχουν ασχοληθεί οι Hawkins (1987), Quesenberry (1991), del Castilloetal. (1996), Zou et al. (2007), Li et al. (2010),Zhang et al.(2012) και άλλοι.

Τα πολυμεταβλητά διαγράμματα ελέγχου αυτόματης εκκίνησης έχουν μελετηθεί από τους Sullivan and Jones (2002), Capizzi and Masarotto (2010), Hawkins and Maboudou – Tchao (2007), Maboudou – Tchao and Hawkins (2011) και άλλους.

Τα περισσότερα διαγράμματα ελέγχου αυτόματης εκκίνησης ελέγχονται ως προς την εκτίμηση σφάλματος που παρουσιάζεται σε διεργασίες που είναι είτε εκτός ελέγχου από την αρχή της παρακολούθησής τους είτε σε αλλαγές που παρατηρήθηκαν νωρίς στη διεργασία. Η εσφαλμένη εκτίμηση των παραμέτρων μπορεί να είναι μεγάλη στα αρχικά στάδια παρακολούθησης όταν δεν έχουμε στη διάθεσή μας αρκετές πληροφορίες.

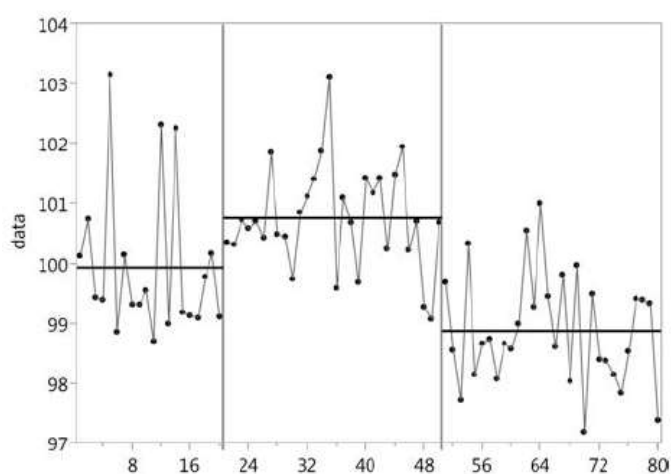
Οι Sullivan and Jones (2002) σημείωσαν το πρόβλημα της πρώιμης εσφαλμένης εκτίμησης των παραμέτρων και πρότειναν οι μέθοδοι αυτόματης εκκίνησης να συμπληρωθούν με

ενδεδεγμένη αναδρομική ανάλυση μόνις συγκεντρωθούν επαρκή δεδομένα. Οι Hawkins and Maboudou – Tchao (2007) και Maboudou –Tchao and Hawkins (2011) εισήγαγαν το πρόβλημα της εσφαλμένης εκτίμησης των παραμέτρων, προτείνοντας επιπροσθέτως με την μελλοντική παρακολούθηση ενός διαγράμματος ελέγχου αυτόματης εκκίνησης να εφαρμοστούν οι μέθοδοι αυτόματης εκκίνησης ξεκινώντας με την πιο πρόσφατη παρατήρηση από τη διεργασία και χρησιμοποιώντας τις προηγούμενες παρατηρήσεις πηγαίνοντας προς τις αρχικές παρατηρήσεις.

Οι μέθοδοι αυτόματης εκκίνησης είναι χρήσιμα εργαλεία σε διεργασίες με δεδομένα που αυξάνονται αργά και έχουν μικρή ιστορία αλλά δεν έχουν μειώσει την ανάγκη για ενδεδεγμένη ανάλυση Φάσης I. Η ανάλυση Φάσης I δίνει πολύτιμες πληροφορίες σχετικά με τη διεργασία και έτσι θα μπορούσαμε να την παραλείψουμε εκτός και αν είναι απολύτως απαραίτητο.

2.3.3 Μέθοδος αλλαγής σημείου

Η ανάλυση αλλαγής σημείου (change-point) είναι μια προτεινόμενη στατιστική προσέγγιση για την αξιολόγηση της σταθερότητας μιας διεργασίας στη Φάση I. Οι μέθοδοι ανάλυσης αλλαγής σημείου χρησιμοποιούνται για να ελέγξουμε αν έχει προκύψει κάποια μετατόπιση στις παραμέτρους της διεργασίας σε κάθε βήμα της αλλά και να εκτιμήσουμε τον χρόνο των μετατοπίσεων αυτών. Ένα παράδειγμα δίνεται στην Εικόνα 5 για δεδομένα που παράγονται από μια κανονική κατανομή με σταθερή διακύμανση ίση με 1. Ο μέσος είναι 100 για τις πρώτες 20 παρατηρήσεις μετατοπίζεται στο 101 για τις επόμενες 30 τιμές και στη συνέχεια πέφτει στο 99.



Εικόνα 5. Μέθοδος αλλαγής σημείου (Jones-Farmer et al. (2014))

Στην ανάλυση Φάσης I, αν εντοπισθούν σημεία αλλαγής στο δείγμα, θα ήταν λάθος να συνδυάσουμε όλα τα δεδομένα της Φάσης I μαζί προκειμένου να εκτιμήσουμε τις εντός ελέγχου παραμέτρους. Αν υπάρχουν μετατοπίσεις στην κατανομή μέσα στο χρόνο, τότε θα πρέπει να κάνουμε κάποια προσπάθεια για να αντιμετωπίσουμε την βασική αιτία αυτού του φαινομένου έτσι ώστε η διακύμανση της διεργασίας να μειωθεί.

Ο Hinkley (1970) έδωσε ένα τεστ πιθανοφάνειας για να ελέγξει υποθέσεις γύρω από σημεία αλλαγής. Επιπλέον, ο Hinkley (1971) έδειξε τη σχέση μεταξύ του διαγράμματος CUSUM και αυτού του τεστ πιθανοφάνειας. Οι Chenand Gupta (2011) έδωσαν μια καλή παρουσίαση των μεθόδων αλλαγής σημείου και ανέφεραν αρκετές περιοχές εφαρμογής της μεθόδου. Αρκετοί ερευνητές εφάρμοσαν τη μέθοδο αλλαγής σημείου στο Στατιστικό Έλεγχο Διεργασιών. Ενδεικτικά αναφέρουμε τα άρθρα των Sullivan and Woodall (2000), Mahmoud and Woodall (2004), Mahmoud et al. (2007) , Paynabaretal. (2012) και Panand Rigdom (2012).

Οι μέθοδοι αλλαγής σημείου είναι ιδιαίτερα χρήσιμες στην ανάλυση Φάσης I όταν αναμένονται μία ή περισσότερες διαρκείς αλλαγές στη διεργασία (παρά η τυχαία εμφάνιση ακραίων τιμών). Υπάρχουν πολλές ευκαιρίες για μελλοντική έρευνα σχετικά με τις μεθόδους αλλαγής σημείου στη Φάση I. Οι Zhang et al. (2013) και οι Jones-Farmeretal. (2014) συζήτησαν την έλλειψη μεθόδων αλλαγής σημείου για την διεξαγωγή ανάλυσης Φάσης I για ιδιότητες (attributes). Επιπλέον, έχουν γίνει κάποιες μελέτες στην ανάπτυξη μη παραμετρικών μεθόδων αλλαγής σημείου (Zhouetal. (2009), Hawkins and Deag (2010) και Zou et al. (2013)) ωστόσο στο συγκεκριμένο πεδίο χρειάζεται περισσότερη έρευνα. Για μεθόδους αλλαγής σημείου που χρησιμοποιούνται από ερευνητές στο Στατιστικό Έλεγχο Ποιότητας, και λόγω της πολυπλοκότητας των μεθόδων, κρίνεται απαραίτητη η ανάπτυξη λογισμικού.

2.3.4 Ταξινόμηση και μέθοδοι βασισμένες σε συστάδες

Μια προσέγγιση είναι να σετάρουμε την ανάλυση Φάσης I ως ένα πρόβλημα ταξινόμησης (classification), προσπαθώντας να ταξινομήσουμε τις παρατηρήσεις σε δυο ή περισσότερες ομάδες (π.χ. εντός και εκτός ελέγχου). Σημειώνοντας ότι τα διαγράμματα ελέγχου Φάσης I συχνά αποτυγχάνουν όταν παρατηρούνται πολλαπλές μετατοπίσεις στο δείγμα αναφοράς, ο Sullivan (2002) εισήγαγε την προσέγγιση κατά συστάδες για να εντοπίζει πολλαπλές μετατοπίσεις στο μέσο της διεργασίας. Οι Zhang et al. (2010) ανέπτυξαν μια μέθοδο για να καθορίσουν τη μακροπρόθεσμη σταθερότητα στις παρατηρήσεις της διεργασίας σε ένα σετ ιστορικών δεδομένων. Πρότειναν μια ανθεκτική μέθοδο που ταυτοποιεί το εντός ελέγχου

δείγμα αναφοράς χρησιμοποιώντας συνδυάζοντας την εμπειρική κατανομή πιθανότητας και μεθόδους κατά συστάδες. Οι Jobe and Pokojony (2009) πρότειναν μια εντατική υπολογιστική και κατά συστάδες αλγοριθμική προσέγγιση για μεμονωμένες πολυμεταβλητές παρατηρήσεις στη Φάση I. Έδειξαν ότι οι μέθοδοί τους ήταν καλύτερες στο να εντοπίζουν τυχαίες ακραίες τιμές και κάποιες μετατοπίσεις του μέσου παρά τη χρήση των αναδρομικών διαγραμμάτων Hotelling's T^2 .

Υπάρχουν πολλοί τρόποι με τους οποίους ένα δείγμα Φάσης I μπορεί να βρεθεί με εκτός ελέγχου σημεία, συμπεριλαμβάνοντας τα ακραία σημεία, διαρκείς αλλαγές στη διεργασία, σποραδικές αλλαγές στη διεργασία, έλλειψη σταθερότητας κτλ. Προτείνονται απλές ιεραρχικές μέθοδοι κατά συστάδες στις οποίες οι παρατηρήσεις τοποθετούνται σε ομάδες παρέχοντας πιο σαφή αποτελέσματα σε περίπτωση που προκύψουν τυχαία ακραίες τιμές. Όταν οι συστάδες έχουν σχηματιστεί, μπορούμε να ψάξουμε για μοτίβα στο χρόνο. Έχουν αναπτυχθεί μερικές μέθοδοι ταξινόμησης που διατηρούν την χρονική ακολουθία των παρατηρήσεων (βλέπε π.χ. Liao (2005)). Μια σημαντική ευκαιρία για μελλοντική έρευνα είναι να επεκτείνουμε τη μελέτη για ταξινόμηση και για μεθόδους βασισμένες σε συστάδες για τη Φάση I.

2.3.5 Ανθεκτικές εκτιμήσεις για εντός ελέγχου παραμέτρους

Επειδή είναι σύνηθες να υπάρχουν ακραίες τιμές στα δεδομένα της Φάσης I, είναι σημαντική η χρήση ανθεκτικών εκτιμητών για τις εντός ελέγχου παραμέτρους μαζί με μια επιτυχημένη σχεδίαση μεθόδων Φάσης I. Όπως εξηγήθηκε από τους Mahmoud et al. (2010), δεν συστήνεται η ιστορικά δημοφιλής και συνεχής πρακτική της χρήσης του εύρους των δειγμάτων για να εκτιμήσουμε την τυπική απόκλιση στα \bar{X} διαγράμματα ελέγχου. Είναι καλύτερο να χρησιμοποιούμε τη δειγματική τυπική απόκλιση γιατί επηρεάζεται λιγότερο από τις ακραίες τιμές. Επίσης μπορούμε να ανατρέξουμε στους Schoonhoven and Does (2012, 2013) και Schoonhoven et al. (2011) για περαιτέρω συζήτηση σχετικά με τους ανθεκτικούς εκτιμητές σε μονοδιάστατα δεδομένα Φάσης I.

Στο μοντέλο της πολυμεταβλητής κανονικής κατανομής, οι αρθρογράφοι χρησιμοποίησαν διαφορετικές στρατηγικές για την εκτίμηση του πίνακα διακυμάνσεων-συνδιακυμάνσεων στη Φάση I για την εφαρμογή διαγραμμάτων Hotelling's T^2 . Ο Vargas (2003) πρότεινε τη χρησιμοποίηση μια ελάχιστης όγκου ελλειψοειδούς (minimum –volume ellipsoid (MVE)) εκτίμηση του πίνακα διακυμάνσεων-συνδιακυμάνσεων για τον εντοπισμό πολλαπλών

ακραίων τιμών στη Φάση I. Αργότερα, οι Jensen et al. (2007) σύγκριναν τους MVE εκτιμητές με τους εκτιμητές ορίζουσας ελάχιστης συνδιακύμανσης (minimum covariance determinant (MCD)), δείχνοντας ότι οι MVE εκτιμητές είναι καλύτεροι όταν τα ποσοστά των ακραίων τιμών είναι μικρό και οι MCD εκτιμητές προτιμώνται όταν υπάρχει μεγάλο ποσοστό ακραίων τιμών στο δείγμα Φάσης I. Οι Oyeyemi and Ipinoyomi (2010) χρησιμοποίησαν ένα εναλλακτικό εκτιμητή για τον πίνακα διακυμάνσεων-συνδιακυμάνσεων για μεμονωμένα T^2 διαγράμματα στη Φάση I που έχουν καλύτερη απόδοση από τις MVE και MCD μεθόδους σε ένα περιορισμένο αριθμό περιπτώσεων. Οι Yanez et al. (2010) πρότειναν τη χρήση ενός διπλού βάρους εκτιμητή για το S για τη θέση (location) και διασπορά σε ένα T^2 διάγραμμα για μεμονωμένα δεδομένα με προσομοιωμένα όρια που βασίζονται σε MVE εκτιμητές για μικρά δείγματα.

2.4 Διαγράμματα ελέγχου για μεταβλητές στη Φάση I

Επειδή συχνά προτείνεται τα διαγράμματα ελέγχου τύπου Shewhart να εφαρμόζονται αναδρομικά σε δεδομένα Φάσης I, κάποια από τα τεχνικά θέματα που εμπλέκονται με αυτή την προσέγγιση θα συζητηθούν σε αυτή και σε επόμενες υποενότητες. Γενικά, υποθέτουμε ότι οι παρατηρήσεις στο χρόνο είναι ανεξάρτητες και ότι οι εντός ελέγχου τιμές των παραμέτρων είναι άγνωστες. Ο στόχος είναι να εντοπίσουμε κάθε παρέκκλιση από την σταθερότητα της διεργασίας με συγκεκριμένο FAP. Όλες οι προτεινόμενες μέθοδοι βασίζονται στην υπόθεση ότι η σταθερότητα αντιστοιχεί σε σταθερές εντός ελέγχου τιμές των παραμέτρων. Τα δυο πρωταρχικά θέματα με αυτές τις μεθόδους είναι ότι (α) τα όρια ελέγχου πρέπει να «φαρδύνουν» προκειμένου να αποφύγουμε τον υπέρμετρο αριθμό λανθασμένων συναγερμών και (β) η απόδοση των μεθόδων δεν φαίνεται να είναι ανθεκτική στις αποκλίσεις από τις υποθέσεις που αφορούν την κατανομή της μεταβλητής που συνήθως είναι η κανονική.

2.4.1 Υπόβαθρο

Οι Chou and Champ (2005) και Champ and Chou (2003) μελέτησαν δυο προσεγγίσεις για την αντιμετώπιση της συσχέτισης μεταξύ των αναδρομικών συγκρίσεων των στατιστικών συναρτήσεων που απεικονίζονται στο διάγραμμα και τα όρια ελέγχου για τα \bar{X} , R και S διαγράμματα. Στην πρώτη προσέγγιση, η οποία αναφέρεται ως «τυπικά όρια» (standard limits), τα όρια υπολογίζονται με τον συνήθη τρόπο, αλλά η πιθανότητα σήματος σε κάθε σύγκριση είναι προσαρμοσμένη με χρήση διόρθωσης Bonferroni. Αυτή η συντηρητική προσέγγιση έχει ως αποτέλεσμα ένα πραγματικό FAP που είναι παραπάνω από το επιθυμητό,

δεδομένου ότι ισχύουν οι υποθέσεις για τις κατανομές. Άλλη μια προσέγγιση, η οποία αναφέρεται ως «ατομικά όρια» (individual limits), θεωρεί διαφορετικούς εκτιμητές παραμέτρων για κάθε διαδοχική σύγκριση των στατιστικών συναρτήσεων που απεικονίζονται στο διάγραμμα ελέγχου με τα όρια ελέγχου. Για μία δοθείσα σύγκριση της στατιστικής συνάρτησης με τα όρια ελέγχου, οι παρατηρήσεις της διεργασίας για αυτή τη στατιστική συνάρτηση εξαλείφονται από αυτές που χρησιμοποιούνται για τις εκτιμήσεις παραμέτρων. Αυτή η προσέγγιση απομακρύνει την εξάρτηση μεταξύ των διαδοχικών συγκρίσεων. Η μελέτη από τους Champ and Chou (2003) προτείνει ότι τα τυπικά όρια βασισμένα στη διόρθωση Bonferroni είναι καλύτερα σε σχέση με αυτά των ατομικών ορίων. Έτσι, οι μέθοδοι που θα αναπτύξουμε στη συνέχεια βασίζονται στη μέθοδο των τυπικών ορίων προκειμένου να υπολογιστούν τα όρια ελέγχου Φάσης I.

2.4.2 Αξιολόγηση της σταθερότητας του μέσου της διεργασίας

Οι Champ and Jones (2004) πρότειναν όρια ελέγχου για αναδρομικά \bar{X} διαγράμματα ελέγχου διεργασιών που κατανέμονται κανονικά βασισμένα στην πολυδιάστατη κατανομή t η οποία λαμβάνει υπόψη την εκτίμηση των παραμέτρων καθώς επίσης και την εξάρτηση των διαδοχικών συγκρίσεων των στατιστικών συναρτήσεων που απεικονίζονται στο διάγραμμα με τα όρια ελέγχου. Επίσης πρότειναν προσεγγιστικά όρια για αναδρομικά διαγράμματα ελέγχου βασισμένα στην μονοδιάστατη t κατανομή για την εκτίμηση των παραμέτρων της διεργασίας αγνοώντας την εξάρτηση μεταξύ των στατιστικών συναρτήσεων. Με χρήση προσομοίωσης, απέδειξαν ότι και οι δυο προσεγγίσεις καταλήγουν σε εμπειρικές τιμές για το FAP, οι οποίες βρίσκονται πολύ κοντά στο επιθυμητό FAP.

Οι Newton and Champ (1997) θεώρησαν τη χρήση της ανάλυσης μέσων (ANOM) ως μέθοδο για την κατασκευή ορίων ελέγχου για αναδρομικά \bar{X} διαγράμματα ελέγχου. Οι Nedumaran and Pignatiello (2005) μελέτησαν την απόδοση αναδρομικών \bar{X} διαγραμμάτων ελέγχου βασισμένα στην ANOM μέθοδο για διεργασίες που κατανέμονται κανονικά με άγνωστες παραμέτρους. Με χρήση προσομοίωσης, απέδειξαν ότι τα όρια ελέγχου βασισμένα σε ANOM διατήρησαν ένα FAP το οποίο βρισκόταν πιο κοντά στο επιθυμητό επίπεδο με απόδοση ελαφρώς καλύτερη από εκείνη που τα όρια βασίζονται σε μια τυπική κανονική κατανομή με μια διόρθωση Bonferroni βασισμένη στον αριθμό των στατιστικών συναρτήσεων. Όμως, οι Champ and Jones (2004), έδειξαν ότι όταν χρησιμοποιήθηκε η προσέγγιση Bonferroni για τον καθορισμό των \bar{X} ορίων ελέγχου για ένα πλήθος m δειγμάτων μεγέθους n το καθένα, το εμπειρικό FAP ήταν εξαρτημένο από το n . Χρησιμοποιώντας

προσομοίωση, κατέληξαν σε εμπειρικές διαφορές από το επιθυμητό FAP που κυμαίνονταν από 50% πιο κάτω από το επιθυμητό FAP για δείγματα μεγέθους $n = 10$ έως 36% πάνω από το επιθυμητό FAP για δείγματα μεγέθους $n = 3$. Στην πρώτη περίπτωση, τα όρια θα ήταν πολύ φαρδιά οδηγώντας σε χαμηλότερες πιθανότητες εντοπισμού των εκτός ελέγχου καταστάσεων. Στη δεύτερη περίπτωση, τα όρια θα ήταν πολύ στενά οδηγώντας σε πολλούς εσφαλμένους συναγερμούς.

Οι Champ and Jones (2004) και Nedumaran and Pignatiello (2005) μελέτησαν \bar{X} διαγράμματα ελέγχου για διεργασίες κανονικά κατανομημένες. Συμπληρωματικά, σχετικά με την επίδραση των εκτιμημένων παραμέτρων στην απόδοση των αναδρομικών διαγραμμάτων ελέγχου, η κατάσταση περιπλέκεται περισσότερο από την κατανομή της διεργασίας. Οι Jones – Farmeretal. (2009) μελέτησαν την εντός ελέγχου απόδοση της χρήσης των \bar{X} ορίων ελέγχου που υπολογίστηκαν με τις μεθόδους που αναπτύχθηκαν από τους Champ and Jones (2004) κάτω από συγκεκριμένες συνθήκες μη κανονικότητας των δεδομένων. Στην περίπτωση ενός διαγράμματος που σχεδιάστηκε με FAP 0.10, έδειξαν ότι, όταν η κατανομή της διεργασίας είχε βαριά ουρά (κατανομή t με 3 βαθμούς ελευθερίας) , το εμπειρικό FAP αυξήθηκε δραματικά, φτάνοντας κοντά στο 50% για κάποιους συνδυασμούς των m και n . Παρόμοια, αλλά λιγότερο δραματικά αποτελέσματα, δόθηκαν για το ίδιο διάγραμμα με FAP 0.10 αλλά και για λοξές κατανομές, που έδωσαν εμπειρικό FAP γύρω στο 30% μεγαλύτερο για $m = 50$ και $n = 3$.

Οι Jones- Farmeretal.(2009) πρότειναν τη χρήση τυποποιημένων μέσων βαθμών (mean ranks) παρόμοιους με αυτούς της διαδικασίας Kruskal –Wallis (Kruskal and Wallis (1952)) με αναδρομικά όρια ελέγχου. Με χρήση προσομοίωσης, οι Jones- Farmer et al. (2009) έδειξαν ότι το δικό τους διάγραμμα μέσων βαθμών διατήρησε το επιθυμητό FAP κάτω από διαφορετικές κατανομές της διεργασίας και πέτυχε υψηλότερες πιθανότητες σήματος σε σχέση με το αναδρομικό διάγραμμα \bar{X} σε ορισμένες περιπτώσεις όταν η κατανομή της διεργασίας δεν ήταν κανονική.

Οι Graham et al. (2010) θεώρησαν ένα μη παραμετρικό διάγραμμα ελέγχου για τη Φάση I για ομαδοποιημένα δεδομένα βασισμένο στη διάμεσο των ομαδοποιημένων (pooled) δεδομένων. Με χρήση προσομοίωσης, απέδειξαν ότι οι τιμές του εμπειρικού FAP ήταν πιο κοντά στα επιθυμητά επίπεδα σε σύγκριση με το \bar{X} διάγραμμα όταν η κατανομή της διεργασίας ήταν μη κανονική, όμως, αυτή η μέθοδος που βασίζεται στη διάμεσο απαιτεί μέγεθος δείγματος n μεγαλύτερο από τον αριθμό των δειγμάτων m προκειμένου οι τιμές του FAP να βρίσκονται

κοντά στα επιθυμητά επίπεδα. Ωστόσο, στην πράξη, τις περισσότερες φορές έχουμε να κάνουμε με μεμονωμένες παρατηρήσεις ή με δείγματα μικρού μεγέθους.

Πρόσφατα, οι Capizzi and Masaratto (2013) εισήγαγαν ένα μη παραμετρικό διάγραμμα ελέγχου Φάσης I που μπορεί να εφαρμοστεί σε μεμονωμένες παρατηρήσεις. Οι μέθοδοί τους χρησιμοποιούν δεδομένα με βάση το χρόνο συλλογής των μεμονωμένων παρατηρήσεων και είναι παρόμοιες με τις μεθόδους των Sullivan (2002) και Zhang et al. (2010). Οι Capizzi and Masaratto (2013) χρησιμοποίησαν μια μέθοδο μετάθεσης (permutation) για τον καθορισμό των ορίων ελέγχου και οι μέθοδοι τους φαίνεται να εντοπίζουν διαφορετικού τύπου μετατοπίσεις των παραμέτρων της διεργασίας.

Κάποιες προσαρμογές στα όρια των διαγραμμάτων ελέγχου θα πρέπει να εφαρμόζονται σε αναδρομικά διαγράμματα ελέγχου Φάσης I. Εάν η υπόθεση κανονικότητας μας ενδιαφέρει, προτείνεται η χρήση μιας μη παραμετρικής μεθόδου. Εάν η σταθερότητα είναι αναγκαία η μορφή της υπό μελέτη εντός ελέγχου κατανομής θα πρέπει να αξιολογηθεί στο σχεδιασμό της Φάσης II. Προτείνονται γραφικές μέθοδοι για τον καθορισμό των κατάλληλων μοντέλων για την εντός ελέγχου κατανομή.

2.4.3 Αξιολόγηση σταθερότητας για τη διακύμανση της διεργασίας.

Το \bar{X} διάγραμμα ελέγχου για τη θέση μιας διεργασίας, γενικά συμπληρώνεται με τα R , S και S^2 διαγράμματα ελέγχου για την παρακολούθηση της μεταβλητότητας. Με παρόμοιο τρόπο, ένα αναδρομικό διάγραμμα ελέγχου θέσης μπορεί να συμπληρωθεί με ένα διάγραμμα για την παρακολούθηση της μεταβλητότητας της διεργασίας. Δεν χρειάζεται ιδιαίτερη δουλειά για την αναδρομική χρήση των διαγραμμάτων ελέγχου για την παρακολούθηση της μεταβλητότητας της διεργασίας. Ο Hillier (1969) και οι Yang and Hillier (1979) θεώρησαν διαγράμματα ελέγχου για τη διακύμανση διεργασιών που κατανέμονται κανονικά με εκτιμημένες παραμέτρους, αλλά δεν έλαβαν υπόψη την εξάρτηση των διαδοχικών συγκρίσεων των στατιστικών συναρτήσεων με τα όρια ελέγχου. Οι Champ and Chou (2003) εξέτασαν τα R και S διαγράμματα ελέγχου Φάσης I στην περίπτωση του κανονικού μοντέλου χωρίς εξάρτηση. Οι Human et al. (2010) χρησιμοποίησαν προσομοίωση για να βρουν εμπειρικά όρια ελέγχου που λαμβάνουν υπόψη την εκτίμηση των παραμέτρων και την εξάρτηση των συγκρίσεων σε μια αναδρομική ανάλυση. Παρείχαν εκτενείς πίνακες ορίων ελέγχου για διαφορετικό αριθμό δειγμάτων (m) και μεγέθη δειγμάτων (n) για τα R , S και S^2 διαγράμματα βασισμένα σε κανονικές διεργασίες.

Οι Jones – Farmer and Champ (2010) χρησιμοποίησαν προσομοίωση για να δείξουν ότι οι εμπειρικές τιμές του FAP για τα αναδρομικά διαγράμματα R και S ήταν αρκετά υψηλές πάνω από το επιθυμητό FAP στην περίπτωση των κανονικών, με βαριά ουρά και λοξών κατανομών. Για παράδειγμα, στην περίπτωση $m = 30$ και $n = 5$ το διάγραμμα R κατασκευασμένο για $FAP = 0.1$ κατέληξε σε ένα εμπειρικό FAP ίσο με 0.215 στην περίπτωση των κανονικών παρατηρήσεων και 0.918 στην περίπτωση των παρατηρήσεων με βαριά ουρά (κατανομή t με 3 βαθμούς ελευθερίας). Οι Jones – Farmer and Champ (2010) σύγκριναν την απόδοση διαφόρων στατιστικών συναρτήσεων βασισμένων στο μέσο βαθμό δείγματος, $|X_{ij} - M|$, όπου X_{ij} είναι η j παρατήρηση από το i δείγμα και M είναι η διάμεσος από όλες τις παρατηρήσεις. Πρότειναν την χρήση του τετραγώνου των ομαδοποιημένων βαθμών (pooled ranks) παρόμοια με το διάγραμμα των μέσων βαθμών που προτάθηκε από τους Jones – Farmer et al. (2009). Αυτή η μέθοδος διατήρησε την τιμή του FAP στα επιθυμητά επίπεδα, ανεξάρτητα από την κατανομή της υπό μελέτη διεργασίας και εντόπισε αυξήσεις και μειώσεις στη διακύμανση με υψηλότερη πιθανότητα από τα S^2 διαγράμματα στην περίπτωση της μη κανονικά κατανεμημένης διεργασίας. Οι Jones – Farmer and Champ (2010) πρότειναν τη χρήση των scale-rank διαγραμμάτων σε συνδυασμό με το διάγραμμα μέσων βαθμών (mean-rank) των Jones – Farmer et al. (2009).

Ένας περιορισμός στα αναδρομικά διαγράμματα ελέγχου τύπου Shewhart για την ανάλυση Φάσης I είναι ότι όλες οι μέθοδοι εφαρμόζονται καλύτερα για μεγάλα μεγέθη δειγμάτων. Επιπλέον, οι περισσότερες από τις μη παραμετρικές μεθόδους φαίνεται να μη δουλεύουν καλά όταν τα μεγέθη των δειγμάτων είναι μικρά. Οι Jones – Farmer et al. (2009) και Jones and Champ (2010) δεν συνιστούν τις μεθόδους τους όταν το μέγεθος του δείγματος είναι μικρότερο από 5 όπως και η μέθοδος του διαγράμματος διαμέσου των Graham et al. (2010). Σε αρκετές διεργασίες τα δεδομένα συλλέγονται μεμονωμένα. Η δυνατότητα εφαρμογής των αναδρομικών διαγραμμάτων ελέγχου τύπου Shewhart για ποιοτικές μεταβλητές είναι κάτι που πρέπει να ερευνηθεί. Η επίδραση της κατανομής της διεργασίας έχει σοβαρό αντίκτυπο στα διαγράμματα Φάσης I για μεμονωμένα δεδομένα. Επειδή δεν υπάρχουν αρκετές πληροφορίες για την κατανομή της διεργασίας στη Φάση I, συνίσταται οι ερευνητές να αναπτύξουν διαγράμματα Φάσης I απαλλαγμένα από κατανομές για μεμονωμένα δεδομένα.

2.5 Διαγράμματα ελέγχου για ιδιότητες στη Φάση I

Δεν έχουν βρεθεί πολλές αναφορές σχετικά με τη χρήση αναδρομικών διαγραμμάτων ελέγχου για δεδομένα ιδιοτήτων (attribute data). Οι Borror and Champ (2001) μελέτησαν την

αναδρομική χρήση των p και np διαγραμμάτων ελέγχου χρησιμοποιώντας προσομοίωση για να δείξουν ότι το FAP είναι αρκετά υψηλό σε αρκετές περιπτώσεις, ειδικά για μεγάλο αριθμό δειγμάτων ($m > 50$). Πρότειναν να χρησιμοποιούνται τέτοια διαγράμματα στην ανάλυση Φάσης I με ιδιαίτερη προσοχή.

Οι Jones and Champ (2012) και οι Donoedo and Chakraborti (2012) πρότειναν διαγράμματα ελέγχου Φάσης I για την παρακολούθηση του χρόνου μεταξύ σπάνιων ενδεχομένων, με μια μέθοδο που βασίζεται στην εκθετική κατανομή. Οι Jones and Champ (2012) ασχολήθηκαν με περιπτώσεις όπου οι εντός ελέγχου παράμετροι της διεργασίας είναι γνωστές ή άγνωστες, με προσεγγιστικά όρια ελέγχου στην περίπτωση που οι παράμετροι είναι άγνωστες. Γενικά, τα διαγράμματα για δεδομένα που κατανέμονται εκθετικά έχουν χαμηλή ισχύ στην ανίχνευση μετατοπίσεων της διεργασίας κατά τη διάρκεια της Φάσης I.

Υπάρχουν αρκετά ανοιχτά θέματα όσο αφορά την ανάλυση Φάσης I για δεδομένα ιδιοτήτων. Χρειάζονται μεγάλα δείγματα για τη Φάση I για να εκτιμήσουμε τις παραμέτρους με αρκετή ακρίβεια προκειμένου τα διαγράμματα για τη Φάση II να λειτουργούν παρόμοια με τις περιπτώσεις που οι παράμετροι είναι γνωστές. Επειδή δεν έχει γίνει αρκετή έρευνα σε αυτό το θέμα, είναι σημαντικό οι ερευνητές να μελετήσουν μεθόδους βασισμένες σε μεγάλα δείγματα αναφοράς. Οι Szarka and Woodall (2011) μελέτησαν κάποιες μεθόδους που είχαν προταθεί για την ανίχνευση σημείων αλλαγών (change points) σε ακολουθίες δεδομένων Bernoulli.

2.6 Πολυμεταβλητά διαγράμματα ελέγχου στη Φάση I

Τα περισσότερα από τα πολυμεταβλητά διαγράμματα ελέγχου που έχουν αναπτυχθεί για τη Φάση I είναι παραλλαγές των Hotelling T^2 διαγραμμάτων ελέγχου και βασίζονται στην υπόθεση μιας πολυμεταβλητής κανονικά κατανομημένης διεργασίας. Οι Tracy et al. (1992) περιέγραψαν μια μέθοδο κατασκευής ορίων ελέγχου Φάσης I για τα αναδρομικά T^2 διαγράμματα ελέγχου. Οι Nedumaran and Pignatiello (2000) έδωσαν οδηγίες για την κατασκευή των T^2 ορίων ελέγχου για αναδρομική ανάλυση ομαδοποιημένων δεδομένων που κατανέμονται κανονικά όταν οι εντός ελέγχου παράμετροι είναι γνωστές.

Είναι μικρός ο αριθμός των δημοσιευμένων ερευνών σχετικά με την ανθεκτικότητα, με μη παραμετρικά ή με παραμετρικά πολυμεταβλητά διαγράμματα ελέγχου για χρήση στη Φάση I. Οι Bell et al. (2014) αξιολόγησαν το αναδρομικό Hotelling T^2 διαγραμμάτων κάτω από ορισμένες αποκλίσεις από την κανονικότητα και έδειξαν ότι δεν λειτουργούν ικανοποιητικά,

με αποτέλεσμα το FAP να ανεβαίνει κατά 90% σε ορισμένες περιπτώσεις. Οι Bell et al. (2014) πρότειναν ένα αναδρομικό διάγραμμα ελέγχου μέσω βαθμών (mean-rank) για συμμετρικά ελλειπτικά πολυμεταβλητά δεδομένα, παρόμοιο με το μονοδιάστατο διάγραμμα ελέγχου μέσω βαθμών των Jones – Farmer et al. (2009). Στα πολυμεταβλητά διαγράμματα μέσω βαθμών των Bell et al. (2014), η οι βαθμοί βασίζονται στην ιδέα του βάθους των δεδομένων (Tukey (1975)), που μετρά το βάθος ενός σημείου εντός του πολυμεταβλητού δείγματος. Ο Liu (1995) εισήγαγε την ιδέα του βάθους των δεδομένων στα διαγράμματα ελέγχου αναπτύσσοντας αρκετά διαγράμματα ελέγχου Φάσης II βασισμένα στην έννοια του απλού βάθους (simplicial depth). Οι Stoumbos and Jones (2000) εκτίμησαν διαγράμματα ελέγχου βασισμένα στο απλό βάθος, σημειώνοντας ότι το απλό βάθος έχει περιορισμούς στο να διακρίνει εκτός ελέγχου σημεία στην ανάλυση Φάσης I. Ομοίως, οι Bell et al. (2014) έδειξαν ότι τα διαγράμματα Φάσης I που βασίζονται στο απλό βάθος δεν εντοπίζουν μετατοπίσεις στη διεργασία όπως επίσης που βασίζονται σε άλλες μεθόδους απλού βάθους.

Υπάρχουν πολλά σημαντικά ερευνητικά θέματα σχετικά με τα αναδρομικά διαγράμματα ελέγχου για πολυμεταβλητές διεργασίες στη Φάση I. Παρόλο που έχουν γίνει κάποιες μελέτες για μεμονωμένες παρατηρήσεις, αυτές βασίζονται στην υπόθεση της πολυδιάστατης κανονικής κατανομής. Η εμπειρία έχει δείξει ότι τα πολυμεταβλητά κανονικά μοντέλα είναι σπάνια επαρκή στην πράξη, έτσι, μια σημαντική περιοχή έρευνας είναι η ανάπτυξη μη παραμετρικών διαγραμμάτων ελέγχου Φάσης I, για παρατηρήσεις από συνεχείς πολυμεταβλητές διαδικασίες. Ο Coleman (1997) εξέφρασε αυτή την άποψη πιο έντονα λέγοντας εν κατακλείδι ότι «ο μη παραμετρικός πολυμεταβλητός στατιστικός έλεγχος διεργασιών είναι αυτό που χρειαζόμαστε».

Ειδικότερα, πρέπει να δοθεί περισσότερη προσοχή στην καταλληλότητα των μεθόδων που βασίζονται σε βαθμούς (ranks) όπως το βάθος των δεδομένων για πολυμεταβλητά διαγράμματα Φάσης I. Παρότι κάποιες μετρήσεις για το βάθος των δεδομένων μπορεί να είναι δύσκολες ή και χρονοβόρες για να υπολογιστούν σε μεγαλύτερες διαστάσεις, άλλες είναι πιο εύκολο να υπολογιστούν. Οι Bell et al. (2014) θεώρησαν κάποιες μετρήσεις για το βάθος των δεδομένων, συμπεριλαμβανομένης και του βάθους Mahalanobis, το ανθεκτικό βάθος Mahalanobis και απλά βάθη. Οι Bell et al. (2014) έδειξαν ότι τα δικά τους διαγράμματα ελέγχου Φάσης I με χρήση ανθεκτικού βάθους Mahalanobis εντοπίζουν αλλαγές με υψηλότερη πιθανότητα σε σχέση με αυτά που χρησιμοποιούν απλό βάθος για να αποδώσουν βαθμούς στα δεδομένα. Το βάθος Mahalanobis απαιτεί περισσότερους υπολογισμούς από τη μέτρηση της απόστασης Mahalanobis, ενώ είναι αρκετά δύσκολο να

υπολογίσουμε το απλό βάθος σε δύο διαστάσεις και δεν υπάρχει αλγόριθμος για τον υπολογισμό απλού βάθους πέρα από τις τρεις διαστάσεις. Από τα αποτελέσματα των Bell et al. (2014) προκύπτει ότι η απόδοση των πολυμεταβλητών διαγραμμάτων ελέγχου βαθμών εξαρτάται από το μέτρο που χρησιμοποιείται για τη μέτρηση του βάθους των δεδομένων, έτσι, είναι αναγκαία να γίνει περαιτέρω έρευνα για να καθοριστεί ποιο μέτρο βάθους των δεδομένων δουλεύει καλύτερα με την κάθε περίπτωση.

2.7 Μέθοδοι κατατομής στη Φάση I

Η παρακολούθηση της κατατομής (profile) με διαγράμματα ελέγχου είναι μια προσέγγιση του Στατιστικού Ελέγχου Διεργασιών που χρησιμοποιείται όταν η ποιότητα ενός προϊόντος ή μιας διεργασία μπορεί να χαρακτηριστεί καλύτερα από μια συναρτησιακή σχέση μεταξύ μιας μεταβλητής απόκρισης και μιας ή περισσότερων επεξηγηματικών μεταβλητών. Οι Woodall et al. (2004) και οι Noorossana et al. (2011) έδωσαν μια επισκόπηση και βιβλιογραφικές αναφορές σχετικά με την παρακολούθηση της κατατομής. Η περισσότερη δουλειά σε δεδομένα κατατομής γίνεται στη Φάση I επειδή πρέπει να χρησιμοποιηθούν ιστορικά δεδομένα για να καθορίσουν την μορφή της συνάρτησης κατατομής. Επιπλέον, πρέπει να αποφασίσουμε αν είναι αναγκαίο να υποθέσουμε ότι η εξεταζόμενη συνάρτηση κατατομής παραμένει σταθερή στο χρόνο όταν η διεργασία είναι εντός ελέγχου ή αν υπάρχει κάποια κοινή αιτία μεταβλητότητας που αναμένεται να εμφανιστεί στο χρόνο.

Στην περίπτωση των γραμμικών κατατομών, η σχέση μεταξύ της μεταβλητής απόκρισης και της επεξηγηματικής μεταβλητής (ή μεταβλητών) συχνά μοντελοποιείται με χρήση ανάλυσης παλινδρόμησης. Στην Φάση II, τις παραμέτρους της παλινδρόμησης μπορούμε να τις παρακολουθήσουμε με χρήση διάφορων μονοδιάστατων διαγραμμάτων ή με ένα πολυμεταβλητό διάγραμμα ελέγχου, π.χ. το διάγραμμα Hotelling T^2 , με σκοπό να εντοπίσουμε αλλαγές στις παραμέτρους της παλινδρόμησης όσο πιο γρήγορα γίνεται.

Προκειμένου να παρακολουθήσουμε τις αλλαγές στις παραμέτρους κατατομής του μοντέλου, είναι σημαντικό να ορίσουμε αναφορικές (baseline) εντός ελέγχου τιμές για τις παραμέτρους κατατομής. Οι Mahmoud and Woodall (2004) έκαναν λόγο για την σημασία της ανάπτυξης μεθόδων Φάσης I για το απλό γραμμικό μοντέλο κατατομής και εισήγαγαν μια μέθοδο που βασίζεται σε ένα έλεγχο F για μοντέλα απλής γραμμικής παλινδρόμησης. Οι Mahmoud et al. (2007) πρότειναν μια μέθοδο Φάσης I για γραμμικές κατατομές βασισμένη σε μεθόδους αλλαγής σημείου και έδειξαν ότι η μέθοδος αλλαγής σημείου προσφέρει βελτιωμένο

εντοπισμό έναντι διατηρήσιμων αλλαγών στις παραμέτρους κατατομής όταν συγκριθούν με τις παραδοσιακές μεθόδους.

Οι Ding et al. (2006) και οι Williams et al. (2007) ανέπτυξαν διαδικασίες για την ανάλυση Φάσης I για μη γραμμικά δεδομένα κατατομής. Η μέθοδος των Ding et al. (2006) αποτελείται από δυο στοιχεία: μια τεχνική ελάττωσης των δεδομένων για να διαχειριστούμε την υψηλή διάσταση των μη γραμμικών δεδομένων κατατομής και μια μέθοδο διαχωρισμού των δεδομένων για να ξεχωρίσουμε τις εντός και εκτός ελέγχου παρατηρήσεις. Οι Chen et al. (2014) πρότειναν μια μέθοδο βασισμένη σε συστάδες που μπορεί να εφαρμοστεί για την ανάλυση γραμμικών και μη γραμμικών κατατομών στη Φάση I. Πρότειναν πρώτα την προσαρμογή μοντέλων στις κατατομές σε ένα σετ ιστορικών δεδομένων και μετά να χρησιμοποιήσουν τις συστάδες για την εκτίμηση των παραμέτρων του μοντέλου για να ξεχωρίσουμε τις εντός και εκτός ελέγχου κατατομές. Οι εκτιμήσεις των εντός ελέγχου παραμέτρων του κατατομών βρέθηκαν με τη χρήση της προσέγγισης των μικτών μοντέλων.

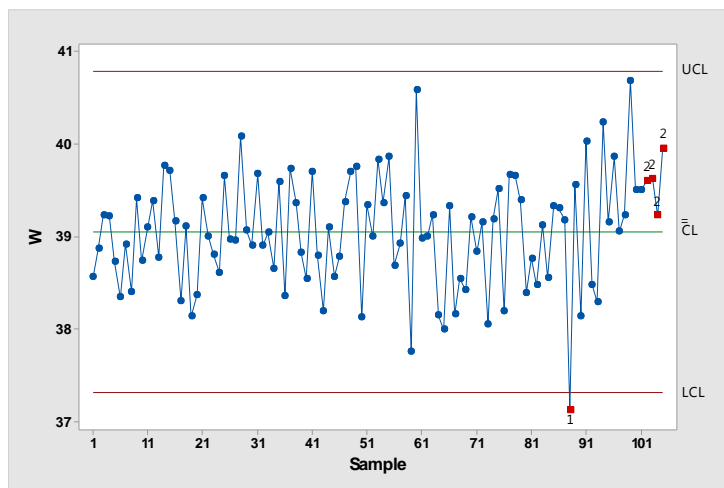
Κάποιες μέθοδοι κατατομών Φάσης I έχουν αναπτυχθεί για μη παραμετρικά και ημιπαραμετρικά μοντέλα. Γενικά, υπάρχει μεγάλο ενδιαφέρον για αυτά τα μοντέλα.

ΚΕΦΑΛΑΙΟ 3

3.1 Εισαγωγή

Κατά την μελέτη των παραγωγικών διεργασιών στόχος μας είναι να παρακολουθήσουμε την συμπεριφορά μιας κρίσιμης ποσότητας ενός χαρακτηριστικού X (τυχαία μεταβλητή) των παραγόμενων προϊόντων. Λαμβάνοντας τυχαία δείγματα X_1, X_2, \dots από την εξεταζόμενη μεταβλητή υπολογίζουμε την τιμή $W_i = g(X_i)$, $i = 1, 2, \dots$ μιας κατάλληλης στατιστικής συνάρτησης που εκτιμά (συνήθως αμερόληπτα) την κρίσιμη ποσότητα. Έτσι, η διαχρονική μελέτη της συμπεριφοράς της κρίσιμης ποσότητας επιτυγχάνεται με την παρακολούθηση των τιμών που λαμβάνει η στατιστική συνάρτηση W στο κάθε δείγμα.

Στοδιάγραμμα ελέγχου απεικονίζουμε τις διαδοχικές παρατηρούμενες τιμές της W_i , οι οποίες ενώνονται με μια τεθλασμένη γραμμή. Στο σχήμα που ακολουθεί παρουσιάζεται ένα διάγραμμα το οποίο χωρίζεται σε επιμέρους ζώνες από τρεις παράλληλες ευθείες με τον οριζόντιο άξονα. Αυτές είναι η κεντρική γραμμή (centerline, CL) ή μέσο επίπεδο και τα άνω και κάτω όρια ελέγχου (upper and lower control limits, UCL και LCL).



Διάγραμμα 3.1 Παράδειγμα διαγράμματος ελέγχου.

Πιο συγκεκριμένα, η κεντρική γραμμή του διαγράμματος ελέγχου παριστάνει συνήθως τη μέση τιμή (mean value) της στατιστικής συνάρτησης W όπως προκύπτει από τη λειτουργία μιας εντός ελέγχου διεργασίας. Όσο οι τιμές της W βρίσκονται εντός των ορίων ελέγχου με «τυχαία» συμπεριφορά, μπορούμε να υποθέσουμε ότι η διεργασία παραμένει εντός στατιστικού ελέγχου και δεν χρειάζεται να επέμβουμε με κάποια διορθωτική κίνηση. Σε

περίπτωση που υπάρξει ένδειξη για εκτός ορίων ελέγχου σημεία, τότε αντιμετωπίζουμε κατάσταση συναγερμού (alarm) και πρέπει να προβούμε σε αναζήτηση των ειδικών αιτιών μεταβλητότητας που δικαιολογεί αυτή τη συμπεριφορά και αν κριθεί αναγκαίο να προβούμε σε διορθωτικές κινήσεις. Αξίζει να σημειώσουμε πως ακόμα και στην περίπτωση όπου όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου αλλά συμπεριφέρονται με ένα συστηματικό (μη τυχαίο) τρόπο, τότε και αυτό αποτελεί ένδειξη ότι η διεργασία είναι εκτός ελέγχου.

Στο κεφάλαιο αυτό, θα αναπτύξουμε διάφορες μεθόδους υπολογισμού των ορίων ελέγχου μιας παραγωγικής διεργασίας στη Φάση I. Αρχικά, θα αναλύσουμε την εφαρμογή των τυπικών ορίων ελέγχου και στη συνέχεια θα εμβαθύνουμε σε πιο εξειδικευμένες μεθόδους όπως τα όρια ελέγχου για μικρό μέγεθος δείγματος.

3.2 Τυπικά όρια ελέγχου - Μειονεκτήματα της τυπικής μεθόδου

Έστω ότι επιθυμούμε να μελετήσουμε κάποιο ποιοτικό χαρακτηριστικό ενός προϊόντος (π.χ. διάμετρος, ύψος, βάρος, κτλ). Το πρώτο βήμα είναι να συλλέξουμε το απαιτούμενο δείγμα από την παραγωγική διαδικασία. Οι παρατηρήσεις αυτές συλλέγονται περιοδικά και σε μικρά δείγματα συνήθως μεγέθους πέντε (5), έτσι ώστε καθένα από τα δείγματα αυτά να είναι όσο το δυνατόν πιο ομογενές. Έχοντας συλλέξει τα δεδομένα, συνεχίζουμε με τον υπολογισμό δυο βασικών στατιστικών μέτρων για κάθε δείγμα, τον μέσο του δείγματος \bar{X} (δηλαδή τον μέσο των μετρήσεων εντός του δείγματος) και το εύρος R (δηλαδή τη διαφορά μεταξύ της μικρότερης από τη μεγαλύτερη μέτρηση εντός του δείγματος).

Αρχικά, θα πρέπει να ορίσουμε τα όρια ελέγχου της διεργασίας. Έχοντας συλλέξει τα απαραίτητα δεδομένα και έχοντας υπολογίσει τον ολικό μέσο $\bar{\bar{X}}$ και το μέσο εύρος \bar{R} της υπό μελέτη διεργασίας, δεν μένει παρά να ορίσουμε τα (δοκιμαστικά) όρια ελέγχου ως εξής:

- $\bar{\bar{X}} \pm A_2 \bar{R}$ για το \bar{X} διάγραμμα
- $D_3 \bar{R}$ για το κάτω όριο του R διαγράμματος
- $D_4 \bar{R}$ για το άνω όριο του R διαγράμματος

όπου A_2 , D_3 , D_4 είναι κατάλληλες σταθερές που εξαρτώνται από το μέγεθος του δείγματος. Εάν στο i – οστό δείγμα παρατηρηθεί ότι το \bar{X}_i ή το R_i βρίσκονται εκτός των ορίων ελέγχου στα αντίστοιχα διαγράμματα, συμπεραίνουμε ότι η διεργασία πιθανώς ήταν «εκτός στατιστικού ελέγχου» όταν συλλέχθηκε το συγκεκριμένο δείγμα. Έτσι, το i – οστό δείγμα και τα αντίστοιχα σημεία το \bar{X}_i και R_i θα πρέπει να αφαιρεθούν. Η διαδικασία αυτή θα

επαναληφθεί για όλα τα δείγματα έως ότου μείνουν τα δείγματα που φαίνονται να προέρχονται από μια εντός ελέγχου διεργασία. Μετά θα υπολογισθούν ξανά τα \bar{X}, \bar{R} από τα εναπομείναντα δείγματα και τα όρια ελέγχου για μελλοντική χρήση θα πρέπει αντιστοίχως να προσαρμοσθούν. Η διαδικασία αυτή θα επαναληφθεί έτσι ώστε τα δείγματα που θα απομείνουν να προέρχονται από σταθερό σύστημα.

Έχοντας υπολογίσει πλέον τα τελικά όρια ελέγχου για κάθε νέο δείγμα υπολογίζουμε τις ποσότητες \bar{X}, R και αν ο μέσος βρίσκεται εκτός των ορίων του \bar{X} διαγράμματος, υπάρχει ένδειξη ότι η διεργασία έχει μετατοπισθεί. Σε αντίστοιχο συμπέρασμα καταλήγουμε και για το R διάγραμμα. Και στις δυο περιπτώσεις η διεργασία πρέπει να ελεγχθεί και να ληφθούν οι απαραίτητες διορθωτικές κινήσεις. Αν οι αλλαγές που προέκυψαν είναι μόνιμες, τα όρια ελέγχου θα πρέπει να ορισθούν ξανά με τη διαδικασία που περιγράψαμε στην προηγούμενη παράγραφο.

Συμβολίζοντας με S τις ποσότητες \bar{X} και R , θα αναπτύξουμε στη συνέχεια το στατιστικό υπόβαθρο των διαγραμμάτων ελέγχου. Υποθέτουμε ότι η διεργασία βρίσκεται εντός στατιστικού ελέγχου, δηλαδή οι μεμονωμένες παρατηρήσεις έχουν συλλεχθεί από την ίδια κατανομή. Έτσι, η ποσότητα S (αντιπροσωπεύει το \bar{X} στο \bar{X} διάγραμμα και το R στο διάγραμμα R) για τα αντίστοιχα δείγματα ακολουθεί κάποια κοινή κατανομή πιθανότητας με μέση τιμή \bar{S}' και τυπική απόκλιση σ'_S . Υπό αυτές τις συνθήκες η πιθανότητα το S να είναι εντός του διαστήματος με όρια από $(\bar{S}' - 3\sigma'_S)$ έως $(\bar{S}' + 3\sigma'_S)$ είναι κοντά στη μονάδα, έτσι ώστε μόνο μια από μερικές εκατοντάδες παρατηρούμενες τιμές του S να βρίσκεται εκτός του διαστήματος αυτού (η πιθανότητα να βρεθεί εκτός είναι 0.0027 όταν $S = \bar{X}$ και 0.0046 όταν $S = R$, οι μεμονωμένες παρατηρήσεις κατανέμονται κανονικά και το μέγεθος των δειγμάτων είναι πέντε (5)). Συνεπώς, όταν οι τιμές του S είναι εκτός, προφανώς η υπόθεση ότι η διεργασία είναι εντός στατιστικού ελέγχου δεν πρέπει να ευσταθεί. Εάν τα \bar{S}' και σ'_S μπορούν να καθορισθούν, τότε τα διαστήματα $\bar{S}' \pm 3\sigma'_S$ θα είναι κατάλληλα όρια ελέγχου καθώς μια τιμή του S εκτός αυτών των ορίων θα είναι ένας αξιόπιστος δείκτης ότι η διεργασία είναι εκτός στατιστικού ελέγχου. Επιπλέον, τα όρια ελέγχου πιθανώς να μην πρέπει να είναι φαρδύτερα καθώς κάτι τέτοιο θα μείωνε την ευαισθησία στις αλλαγές της διεργασίας.

Έτσι, με αντικατάσταση τα κατάλληλα όρια ελέγχου θα είναι $\bar{X}' \pm 3\sigma'_{\bar{X}}$ για το διάγραμμα \bar{X} ($S = \bar{X}$) και $\bar{R}' \pm 3\sigma'_{\bar{R}}$ για το R διάγραμμα ($S = R$).

Συνήθως, η μέση τιμή και η τυπική απόκλιση του \bar{X} και του R δεν είναι γνωστά. Για το λόγο αυτό συνηθίζουμε να χρησιμοποιούμε τις εκτιμήσεις τους από τα αρχικά δείγματα. Έτσι, τα $\bar{\bar{X}} - A_2\bar{R}$, $\bar{\bar{X}} + A_2\bar{R}$, $D_3\bar{R}$, και $D_4\bar{R}$ εκτιμούν τα «ιδανικά» όρια ελέγχου $\bar{X}' - 3\sigma'_{\bar{X}}$, $\bar{X}' + 3\sigma'_{\bar{X}}$, $\bar{R}' - 3\sigma'_{\bar{R}}$, (ή μηδέν (0) αν είναι αρνητικό) και $\bar{R}' + 3\sigma'_{\bar{R}}$, αντίστοιχα.

Όταν τα δείγματα που εξετάσαμε είναι λίγα, τα όρια ελέγχου διαφέρουν από αυτά που θα έπρεπε να είναι. Ένα πιθανό πρόβλημα, είναι να υπάρξει μεγάλο ποσοστό μελλοντικών τιμών των \bar{X} ή R , που μπορεί να βρεθούν έξω από τα όρια ελέγχου όταν η διεργασία είναι εντός ελέγχου. Έτσι λοιπόν, για μικρό αριθμό δειγμάτων τα όρια ελέγχου που προτείνουμε παραπάνω δεν είναι αξιόπιστα για να κρίνουμε το αν τελικά η διεργασία είναι εντός ή εκτός στατιστικού ελέγχου.

Στον πίνακα που ακολουθεί παρουσιάζονται οι πιθανότητες σφάλματος τύπου I (δηλαδή η πιθανότητα ένα νέο δείγμα να θεωρηθεί εκτός στατιστικού ελέγχου ενώ στην πραγματικότητα δεν είναι) για τα διαγράμματα \bar{X} και R για διάφορες τιμές του m (αριθμός δειγμάτων που χρησιμοποιήθηκε για τον υπολογισμό των ορίων ελέγχου) και για μέγεθος δείγματος $n = 5$.

m	\bar{X} διάγραμμα	R διάγραμμα
5	0.012	0.018
10	0.0067	0.0102
15	0.0051	0.0081
20	0.0044	0.0072
25	0.0040	0.0066
50	0.0033	0.0056
100	0.0030	0.0050
∞	0.0027	0.0046

Οι πιθανότητες που δίνονται στον πίνακα βρίσκουν εφαρμογή μόνο σε μελλοντικά δείγματα. Πιο συγκεκριμένα, το πρώτο βήμα είναι να ελέγξουμε αναδρομικά εάν η διεργασία ήταν εντός στατιστικού ελέγχου όταν συλλέχθηκε το αρχικό δείγμα και καθένα από τα \bar{X} και R περιλαμβάνονται στα $\bar{\bar{X}}$ και \bar{R} . Έτσι, μια ασυνήθιστα μεγάλη ή μικρή τιμή των \bar{X} ή R θα επηρεάσει προς την ίδια κατεύθυνση και τα $\bar{\bar{X}}$ και \bar{R} και συνεπώς η πιθανότητα σφάλματος τύπου I θα είναι ελαφρά μικρότερη από αυτή που παρουσιάζεται στον παραπάνω πίνακα.

Για την αξιολόγηση των τιμών του πίνακα θα πρέπει να γνωρίζουμε ότι η πιθανότητα που υποθέσαμε είναι κάτι λιγότερο από 0.005 (συγκεκριμένα 0.0027 για το \bar{X} διάγραμμα και 0.0046 για το R διάγραμμα). Αυτό σημαίνει ότι τα τυπικά όρια ελέγχου έχουν δημιουργηθεί κάτω από την υπόθεση ότι τα δείγματα σπάνια θα παρουσιάσουν λανθασμένο πρόβλημα. Σε αντίθεση, ο πίνακας δείχνει, για παράδειγμα, ότι κατά μέσο όρο ένα δείγμα ανά 50 δείγματα θα δίνει λανθασμένο συναγερμό σε ένα από τα δύο διαγράμματα εάν τα όρια ελέγχου υπολογισθούν χρησιμοποιώντας μόνο 10 αρχικάδείγματα (μεγέθους 5).

Δεν μπορούμε να προβούμε σε γενικά συμπεράσματα σχετικά με τον αριθμό των δειγμάτων που θα χρησιμοποιήσουμε προτού υπολογίσουμε τα τυπικά όρια ελέγχου. Ωστόσο ο παραπάνω πίνακας δίνει συγκεκριμένες πληροφορίες για κάθε μεμονωμένη περίπτωση για γρήγορο έλεγχο και το κόστος για τον έλεγχο του προβλήματος ενώ δεν υπάρχει.

Για να μπορέσουμε να ελέγξουμε αποτελεσματικά τους δυο τύπους σφαλμάτων θα πρέπει τα όρια ελέγχου στο \bar{X} διάγραμμα να παρουσιάζουν τα εξής δύο αντικρουόμενα χαρακτηριστικά:

- να είναι τόσο φαρδιά ώστε σπάνια να προκύπτουν σημεία εκτός αυτών λόγω μετατόπισης της μέσης τιμής, και
- να είναι τόσο στενά ώστε οι ξαφνικές και σημαντικές αλλαγές να έχουν ως αποτέλεσμα ένα ή περισσότερα σημεία να βρεθούν εκτός των ορίων ελέγχου.

Όταν το \bar{X} διάγραμμα κατασκευάζεται βάσει της τυπικής μεθόδου που περιγράφηκε παραπάνω, η ύπαρξη των χαρακτηριστικών που αναφέραμε νωρίτερα εξαρτάται από τον αριθμό και το μέγεθος των δειγμάτων που χρησιμοποιήσαμε. Πρακτικά, έχει παρατηρηθεί ότι ένας αποτελεσματικός συνδυασμός είναι η εφαρμογή διαγραμμάτων ελέγχου τύπου Shewhart με χρήση 25 δειγμάτων με μέγεθος 5 το καθένα. Υποθέτοντας την ύπαρξη κανονικότητας, η πιθανότητα σφάλματος τύπου I υπολογίζεται ίση με 0.05 (εναλλακτικά, η πιθανότητα να βρεθούν και οι 25 δειγματικοί μέσοι εντός των ορίων ελέγχου είναι ίση με 0.95). Πιο συγκεκριμένα, όταν η διεργασία είναι εντός στατιστικού ελέγχου σε περίπου ένα διάγραμμα επί συνόλου 20 διαγραμμάτων, ένας τουλάχιστον από τους 25 δειγματικούς μέσους θα βρεθεί εκτός των ορίων $\bar{\bar{X}} \pm A_2\bar{R}$. Έτσι, τα όρια σε αυτή την περίπτωση είναι αρκετά φαρδιά για να παρέχουν καλή προστασία έναντι των σφαλμάτων τύπου I. Επίσης, με βάση κάποια πρακτικά αλλά και θεωρητικά αποτελέσματα επιβεβαιώνεται η δυνατότητα

παρακολούθησης ξαφνικών αλλαγών στα επίπεδα της διεργασίας και επομένως υπάρχει προστασία έναντι των σφαλμάτων τύπου II.

Όμως, το «σκηνικό» αλλάζει όταν το \bar{X} διάγραμμα βασιστεί σε μικρότερο αριθμό δεδομένων. Καθώς το πλήθος των δειγμάτων m μειώνεται (για δεδομένο n) η πιθανότητα σφάλματος τύπου I μειώνεται. Για παράδειγμα, όταν το διάγραμμα βασίζεται σε μόνο δύο δείγματα ($m = 2$) δεν παρουσιάζεται σημαντική αλλαγή στο σφάλμα τύπου I. Αυτό όμως έχει ως αποτέλεσμα στην αύξηση του σφάλματος τύπου II, δηλαδή μείωση της ισχύος. Τα όρια ελέγχου είναι τόσο φαρδιά που περιλαμβάνουν τη φυσιολογική (κανονική) μεταβλητότητα των δειγματικών μέσων και αφήνουν κάποια περιθώρια για εντοπισμό αλλαγών της διεργασίας. Εξ αιτίας του μεγάλου πλάτους των ορίων ελέγχου μόνο βίαιες αλλαγές στη διεργασία προκαλούν την εμφάνιση σημείου εκτός των ορίων ελέγχου.

Τα μειονεκτήματα της χρήσης μικρού δείγματος μπορούν να περιοριστούν χρησιμοποιώντας διαφορετικό πολλαπλάσιο του \bar{R} στη σχέση $\bar{X} \pm A_2 \bar{R}$. Ο πολλαπλασιαστής αυτός θα πρέπει να επιλεγεί με τρόπο ώστε η πιθανότητα σφάλματος τύπου I να είναι προσεγγιστικά 0.05 ανεξάρτητα από το πλήθος των δειγμάτων. Πιο συγκεκριμένα, ο παράγοντας A_2 , που επηρεάζεται μόνο από το μέγεθος n του δείγματος, μπορεί να αντικατασταθεί από τον παράγοντα C που εξαρτάται και από το πλήθος τους (m). Παρόλο που η αλλαγή από A_2 σε C προσφέρει πιο στενά όρια όταν ο αριθμός των δειγμάτων είναι μικρός σημειώνουμε ότι το C προσεγγίζει το A_2 ραγδαία όσο το m αυξάνει. Συγκεκριμένα, στην περίπτωση όπου υπάρχουν 25 δείγματα, τα A_2 και C είναι παρόμοια και τα όρια ελέγχου είναι ουσιαστικά $\bar{X} \pm A_2 \bar{R}$.

Παράδειγμα

Υποθέτουμε τέσσερα (4) δείγματα, το καθένα μεγέθους δυο(2), τα οποία συλλέχθηκαν σε διάστημα μιας ώρας και κάνοντας τους ανάλογους υπολογισμούς προέκυψαν τα ακόλουθα αποτελέσματα

\bar{X}	R
1.17	0.05
0.14	1.64
1.13	1.40
0.54	0.76

Συνεπώς

$$\bar{\bar{X}} = 0.745 \text{ και } \bar{R} = 0.9625.$$

Τα τυπικά όρια ελέγχου θα υπολογισθούν ως εξής ($A_2 = 1.88$):

$$\bar{\bar{X}} - A_2 \bar{R} = -1.0645$$

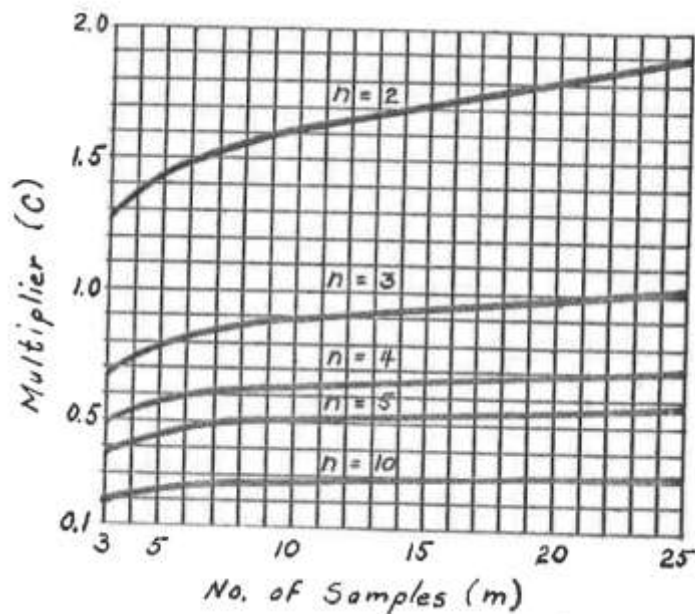
$$\bar{\bar{X}} + A_2 \bar{R} = 2.5545$$

Κάνοντας χρήση του κάτωθι διαγράμματος στην καμπύλη για $n = 1$ και για $m = 4$, προκύπτει ότι $c = 1.36$.

Υπολογίζοντας τα όρια ελέγχου με χρήση της σταθεράς c , παίρνουμε

$$\bar{\bar{X}} - C \bar{R} = -0.564$$

$$\bar{\bar{X}} + C \bar{R} = 2.054$$



Διάγραμμα 3.2 Γραφικός προσδιορισμός του C (E.P. King (1954))

Για τον υπολογισμό του C έχουμε τα ακόλουθα. Ο παράγοντας A_2 καθορίζεται από τη σχέση

$$A_2 = 3/(\sqrt{nd_2})$$

όπου d_2 είναι η αναμενόμενη τιμή του εύρους ενός τυχαίου δείγματος μεγέθους n από κανονική κατανομή με διακύμανση ίση με τη μονάδα. Ο αριθμός 3 στην παραπάνω θα

αντικατασταθεί από τον παράγοντα k_m , ο οποίος εξαρτάται από το πλήθος των δειγμάτων. Ο παράγοντας αυτός έχει επιλεγεί με τρόπο τέτοιο, ώστε όταν η κατανομή της διεργασίας είναι κανονική με μέση τιμή μ και διακύμανση σ^2 , να ισχύει ότι

$$P\{|\bar{X}_1 - \bar{X}| \leq k_m \sigma / \sqrt{n}, \dots, |\bar{X}_m - \bar{X}| \leq k_m \sigma / \sqrt{n}\} \cong 0.95 \quad (1)$$

όπου $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ είναι οι μέσοι των m διαδοχικών δειγμάτων μεγέθους n και $\bar{X} = (\bar{X}_1 + \dots + \bar{X}_m) / m$. Οι τιμές των k_3 και k_4 δίνονται από τον King (1952). Για $m > 4$ οι τιμές του k_m προκύπτουν από την κατανομή της στατιστικής

$$\frac{\max |\bar{X}_i - \bar{X}|}{\sigma / \sqrt{n}}$$

για κάθε m . Το 95% ποσοστιαίο σημείο αυτής της δειγματοληπτικής κατανομής, χρησιμοποιείται ως εκτιμήτρια του k_m , αφού η παραπάνω έκφραση (1) ισοδυναμεί με

$$P\{|\bar{X}_1 - \bar{X}| \leq k_m \sigma / \sqrt{n}, \dots, |\bar{X}_m - \bar{X}| \leq k_m \sigma / \sqrt{n}\} = P\left\{\frac{\max |\bar{X}_i - \bar{X}|}{\sigma / \sqrt{n}} \leq k_m\right\} \cong 0.95$$

Οι τιμές του C προκύπτουν από τη σχέση $C = k_m / (\sqrt{n} d_2)$.

3.3 Όρια ελέγχου για το διάγραμμα ελέγχου \bar{X} και R

Στην ενότητα αυτή θα παρουσιάσουμε μια μέθοδο για την δημιουργία αξιόπιστων ορίων ελέγχου, τα οποία μπορούν να χρησιμοποιηθούν ανεξάρτητα από τον αριθμό των δειγμάτων που εξετάζουμε. Η μέθοδος αυτή χωρίζεται σε δύο στάδια. Το πρώτο στάδιο αφορά στη δημιουργία ορίων ελέγχου για αναδρομικό έλεγχο, προκειμένου να ελέγξουμε αν η διεργασία ήταν εντός ελέγχου με βάση τα αρχικά δεδομένα. Στο δεύτερο στάδιο, αφού έχουν ταυτοποιηθεί τα εντός ελέγχου δείγματα, ορίζουμε νέα όρια ελέγχου προκειμένου να διαπιστώσουμε αν η διεργασία παραμένει εντός ελέγχου για μελλοντικές παρατηρήσεις. Τα δυο αυτά στάδια αναπτύσσονται αναλυτικά στις υποενότητες που ακολουθούν.

3.3.1 1^ο στάδιο της μεθόδου

Αρχικά, θα πρέπει να εξετάσουμε έναν μικρό αριθμό αρχικών δειγμάτων (m) και με βάση αυτά να υπολογίσουμε τις ποσότητες \bar{X} και \bar{R} . Στο σημείο αυτό απαιτείται η λήψη τριών αποφάσεων σχετικά με τις κατάλληλες πιθανότητες σφάλματος τύπου I για κάθε αρχικό δείγμα. Πιο συγκεκριμένα, θα πρέπει να επιλέξουμε τις κατάλληλες τιμές για τις ποσότητες

- a_2 : πιθανότητα σφάλματος τύπου I για το \bar{X} διάγραμμα
- a_3 : πιθανότητα σφάλματος τύπου I για το κάτω όριο ελέγχου του R διαγράμματος
- a_4 : πιθανότητα σφάλματος τύπου I για το άνω όριο ελέγχου του R διαγράμματος

Δηλαδή, για κάθε αρχικό δείγμα, a_2 είναι η πιθανότητα το \bar{X} να πέσει έξω από τα όρια ελέγχου του διαγράμματος \bar{X} , a_3 είναι η πιθανότητα το R να πέσει χαμηλότερα από το κάτω όριο του διαγράμματος R και a_4 είναι η πιθανότητα το R να πέσει ψηλότερα από το άνω όριο του διαγράμματος R , δεδομένου ότι η διεργασία βρίσκεται σταθερά εντός στατιστικού ελέγχου. Για να επιτύχουμε τις παραπάνω πιθανότητες τα όρια ελέγχου $\bar{X} \pm A_2\bar{R}$ για το \bar{X} διάγραμμα, $D_3\bar{R}$ για το κάτω όριο του R διαγράμματος και $D_4\bar{R}$ για το άνω όριο του R διαγράμματος υπολογιστούν με χρήση κατάλληλων σταθερών οι οποίες συμβολίζονται ως A_2^{**}, D_3^{**} και D_4^{**} στη θέση των A_2, D_3 και D_4 . Δοθέντων των τιμών m, a_2, a_3, a_4 (θεωρούμε ότι $n = 5$) και ανατρέχοντας στους Πίνακες 3.1, 3.2, 3.3 για την εύρεση των τιμών $A_2^{**}, D_3^{**}, D_4^{**}$ που ικανοποιούν τα δεδομένα μας, υπολογίζουμε τα όρια ελέγχου για το \bar{X} και το R διάγραμμα ως εξής:

- Κάτω όριο ελέγχου για το \bar{X} διάγραμμα (1^ο στάδιο): $\bar{X} - A_2^{**}\bar{R}$
- Άνω όριο ελέγχου για το \bar{X} διάγραμμα (1^ο στάδιο): $\bar{X} + A_2^{**}\bar{R}$
- Κάτω όριο ελέγχου για το R διάγραμμα (1^ο στάδιο): $D_3^{**}\bar{R}$
- Άνω όριο ελέγχου για το R διάγραμμα (1^ο στάδιο): $D_4^{**}\bar{R}$

Πίνακας 3.1. A_2^{**} : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος (1^ο στάδιο), (Hillier (1969))

n	α_2				
	0.001	0.0027	0.01	0.025	0.05
2	0.684	0.575	0.449	0.366	0.306
3	0.679	0.590	0.476	0.397	0.337
4	0.669	0.589	0.483	0.408	0.349
5	0.662	0.588	0.487	0.414	0.356
6	0.658	0.587	0.489	0.418	0.360
7	0.655	0.586	0.490	0.420	0.362
8	0.652	0.585	0.491	0.421	0.363
9	0.650	0.584	0.492	0.423	0.366
10	0.648	0.584	0.493	0.425	0.368
15	0.643	0.581	0.493	0.426	0.370
20	0.639	0.579	0.493	0.427	0.372
25	0.637	0.578	0.493	0.428	0.372
50	0.636	0.578	0.495	0.430	0.375
100	0.634	0.577	0.495	0.430	0.375
∞	0.633	0.577	0.495	0.431	0.377

Πίνακας 3.2. D_3^{**} : Παράγοντες για το κάτω όριο ελέγχου του R διαγράμματος (1° στάδιο),(Hillier (1969))

m	α_1				
	0.001	0.005	0.01	0.025	0.05
2	0.236	0.339	0.394	0.487	0.560
3	0.202	0.296	0.349	0.434	0.513
4	0.189	0.279	0.330	0.414	0.493
5	0.182	0.270	0.320	0.403	0.482
6	0.177	0.264	0.314	0.396	0.475
7	0.174	0.259	0.309	0.391	0.470
8	0.172	0.257	0.307	0.388	0.466
9	0.170	0.255	0.304	0.385	0.463
10	0.169	0.253	0.302	0.383	0.461
15	0.165	0.248	0.296	0.377	0.455
20	0.163	0.246	0.294	0.374	0.452
25	0.162	0.244	0.292	0.372	0.450
50	0.160	0.241	0.289	0.369	0.446
100	0.159	0.240	0.287	0.367	0.445
∞	0.158	0.239	0.286	0.365	0.443

Πίνακας 3.3. D_4^{**} : Παράγοντες για το άνω όριο ελέγχου του R διαγράμματος (1° στάδιο),(Hillier (1969))

m	α_1				
	0.001	0.005	0.01	0.025	0.05
2	1.78	1.67	1.62	1.53	1.44
3	2.02	1.84	1.76	1.63	1.52
4	2.12	1.92	1.82	1.68	1.56
5	2.18	1.96	1.85	1.71	1.58
6	2.21	1.98	1.88	1.72	1.60
7	2.23	2.00	1.89	1.73	1.60
8	2.25	2.01	1.90	1.74	1.61
9	2.26	2.02	1.91	1.75	1.62
10	2.27	2.03	1.92	1.76	1.62
15	2.30	2.06	1.94	1.77	1.63
20	2.32	2.07	1.95	1.78	1.64
25	2.32	2.07	1.96	1.79	1.64
50	2.34	2.09	1.97	1.80	1.65
100	2.35	2.09	1.97	1.80	1.65
∞	2.36	2.10	1.98	1.80	1.66

Για κάθε δείγμα θα πρέπει να σιγουρευτούμε ότι ο μέσος \bar{X} και το εύρος R βρίσκονται εντός των ορίων ελέγχου στο πρώτο στάδιο της μεθόδου για τα διαγράμματα \bar{X} και R αντίστοιχα. Αν κάτι τέτοιο δεν επαληθευτεί, τότε θα πρέπει να απομακρύνουμε τα προβληματικά δείγματα από τα δεδομένα μας και να υπολογίσουμε ξανά τα όρια ελέγχου. Επιπλέον, η διεργασία θα πρέπει να εξεταστεί, προκειμένου να σιγουρευτούμε ότι η αλλαγή/μετατόπιση που έγινε ήταν προσωρινή, αλλιώς θα πρέπει να επαναλάβουμε ολόκληρη τη διαδικασία με νέα δείγματα όταν θεωρήσουμε ότι η διεργασία επανήλθε εντός στατιστικού ελέγχου. Όταν όλα τα

δείγματα βρεθούν εντός των τρεχόντων ορίων ελέγχου και αφού σιγουρευτούμε ότι τα δείγματα που χρησιμοποιήσαμε αντιπροσωπεύουν την εξεταζόμενη διεργασία, μπορούμε να προχωρήσουμε στον υπολογισμό των ορίων ελέγχου στο δεύτερο στάδιο της μεθόδου, με σκοπό να ελέγξουμε αν η διεργασία θα παραμείνει εντός στατιστικού ελέγχου και στο μέλλον.

Η επιλογή του αριθμού των δειγμάτων m , θα πρέπει να γίνει πριν τον υπολογισμό των ορίων ελέγχου και θα πρέπει να λάβουμε υπόψη μας δυο συνθήκες. Αρχικά, τα όρια ελέγχου θα πρέπει να είναι αξιόπιστα, ανεξάρτητα από το πόσο μικρό είναι το m . Μπορεί το m να πάρει ακόμα και τις τιμές $m = 1$ ή $m = 2$ (παρόλο που όταν το $m = 1$ ο αναδρομικός έλεγχος δεν έχει νόημα και πηγαίνουμε κατευθείαν στο δεύτερο στάδιο της μεθόδου). Επιπλέον, ένα ακόμα σημαντικό στοιχείο, είναι η ισχύς στην ανίχνευση διεργασιών εκτός ελέγχου. Όσο πιο μεγάλο είναι το m , τόσο μεγαλύτερη ισχύς παρουσιάζεται (τόσο μικρότερη είναι η πιθανότητα σφάλματος τύπου II) όπως αποδεικνύεται και από το γεγονός ότι τα A_2^{**} και D_4^{**} μειώνονται και το D_3^{**} αυξάνεται καθώς το m αυξάνει. Έτσι λοιπόν, ο κίνδυνος που αντιμετωπίζουμε όταν επιλέγουμε μικρό m είναι ότι στο αρχικό δείγμα μπορεί τα στατιστικά μέτρα να βρεθούν εντός των ορίων ελέγχου και στη συνέχεια με επαναπροσδιορισμό των ορίων στα διαγράμματα ελέγχου να βρεθούν εκτός ελέγχου. Για το λόγο αυτό, προτιμούμε σε πολλές περιπτώσεις να περιμένουμε για μερικά δείγματα, έστω $m = 5$.

Όταν διαλέγουμε τα a_2, a_3 και a_4 συνήθως επιλέγουμε μικρές τιμές αυτών. Όμως αυτό έχει ως συνέπεια τα όρια ελέγχου που θα προκύψουν να είναι φαρδιά. Ιδανικά, επιθυμούμε τα όρια ελέγχου να είναι σχετικά στενά, έτσι ώστε τα διαγράμματα ελέγχου να είναι ευαίσθητα σε αλλαγές της διεργασίας (μικρή πιθανότητα σφάλματος τύπου II). Έτσι, τα a_2, a_3 και a_4 θα πρέπει να πάρουν τέτοιες τιμές ώστε να δίνουν έναν επαρκή βαθμό αξιοπιστίας. Όταν διαλέγουμε τις τιμές αυτές θα πρέπει να υπάρχει ισορροπία μεταξύ της μικρότερης πιθανότητας σφάλματος τύπου I και της ευαισθησίας των διαγραμμάτων ελέγχου, λαμβάνοντας υπόψη και τα κόστη που υπεισέρχονται. Εάν θέλουμε να επιλέξουμε συμβατικές τιμές για μεγάλο m , θα έχουμε $a_2 = 0.0027$, $a_3 = 0$ (δηλαδή $D_3^{**}=0$) και $a_4 = 0.005$.

3.3.2 2^ο στάδιο της μεθόδου

Έχοντας ολοκληρώσει τον αναδρομικό έλεγχο και έχοντας καταλήξει πως τα αρχικά δείγματα που τελικά προέκυψαν βρίσκονταν εντός ελέγχου, είμαστε έτοιμοι να υπολογίσουμε τα όρια

ελέγχου στο δεύτερο στάδιο της μεθόδου, τα οποία αφορούν σε μελλοντικές παρατηρήσεις. Η διαδικασία αυτή απαιτεί για ακόμα μια φορά την επιλογή τιμών για τις ποσότητες a_2, a_3 και a_4 , όμως στο στάδιο αυτό η πιθανότητα σφάλματος τύπου I θα αφορά στα μελλοντικά δείγματα. Επιπλέον, είναι προφανές πως τα μελλοντικά μεγέθη \bar{X} και R δεν περιλαμβάνονται στα ήδη υπολογισμένα στη Φάση $I\bar{X}$ και \bar{R} , έτσι οι παράγοντες που χρησιμοποιήσαμε για τα όρια ελέγχου στο πρώτο στάδιο ($A_2^{**}, D_3^{**}, D_4^{**}$) δεν ισχύουν πλέον. Θα πρέπει να αντικατασταθούν από τους κατάλληλους παράγοντες, έστω τα A_2^*, D_3^* και D_4^* που δίνονται από τους Πίνακες 3.4, 3.5 και 3.6, για μέγεθος δείγματος ίσο με πέντε (η διαδικασία που απαιτείται για την εξαγωγή των παραγόντων αυτών περιγράφεται στην επόμενη ενότητα). Συνεπώς, τα όρια ελέγχου για στο δεύτερο στάδιο της μεθόδου υπολογίζονται ως ακολούθως:

- Κάτω όριο ελέγχου για το \bar{X} διάγραμμα (2^ο στάδιο): $\bar{X} - A_2^* \bar{R}$
- Άνω όριο ελέγχου για το \bar{X} διάγραμμα (2^ο στάδιο): $\bar{X} + A_2^* \bar{R}$
- Κάτω όριο ελέγχου για το R διάγραμμα (2^ο στάδιο): $D_3^* \bar{R}$
- Άνω όριο ελέγχου για το R διάγραμμα (2^ο στάδιο): $D_4^* \bar{R}$

Πίνακας 3.4. A_2^* : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος (2^ο στάδιο), (Hillier (1969))

n	α_2				
	0.001	0.0027	0.01	0.025	0.05
1	2.27	1.74	1.21	0.911	0.720
2	1.19	1.00	0.781	0.637	0.532
3	0.960	0.834	0.673	0.562	0.477
4	0.864	0.760	0.624	0.527	0.451
5	0.811	0.720	0.596	0.507	0.436
6	0.779	0.695	0.579	0.495	0.426
7	0.756	0.677	0.566	0.485	0.418
8	0.739	0.662	0.557	0.477	0.412
9	0.726	0.653	0.550	0.473	0.409
10	0.717	0.645	0.545	0.470	0.407
15	0.687	0.621	0.527	0.455	0.396
20	0.672	0.609	0.518	0.449	0.391
25	0.663	0.602	0.513	0.445	0.387
50	0.649	0.590	0.505	0.439	0.383
100	0.640	0.583	0.500	0.434	0.379
∞	0.633	0.577	0.495	0.431	0.377

Πίνακας 3.5. D_3^{**} : Παράγοντες για το κάτω όριο ελέγχου του R διαγράμματος (2^ο στάδιο),(Hillier (1969))

m	α_3				
	0.001	0.005	0.01	0.025	0.05
1	0.134	0.204	0.245	0.317	0.389
2	0.144	0.219	0.263	0.338	0.413
3	0.148	0.225	0.270	0.346	0.422
4	0.151	0.228	0.274	0.351	0.427
5	0.152	0.230	0.276	0.353	0.430
6	0.153	0.231	0.278	0.355	0.432
7	0.154	0.232	0.279	0.357	0.433
8	0.154	0.233	0.280	0.358	0.434
9	0.155	0.234	0.280	0.359	0.435
10	0.155	0.234	0.281	0.359	0.436
15	0.156	0.236	0.282	0.361	0.438
20	0.156	0.236	0.283	0.362	0.439
25	0.157	0.237	0.284	0.363	0.440
50	0.157	0.238	0.285	0.364	0.441
100	0.158	0.238	0.285	0.365	0.442
∞	0.158	0.239	0.286	0.365	0.443

Πίνακας 3.6. D_4^{**} : Παράγοντες για το άνω όριο ελέγχου του R διαγράμματος (2^ο στάδιο),(Hillier (1969))

m	α_4				
	0.001	0.005	0.01	0.025	0.05
1	7.94	5.10	4.20	3.22	2.60
2	4.14	3.19	2.83	2.39	2.07
3	3.39	2.76	2.50	2.17	1.92
4	3.08	2.57	2.36	2.07	1.85
5	2.91	2.47	2.27	2.01	1.81
6	2.81	2.40	2.22	1.98	1.78
7	2.74	2.35	2.18	1.95	1.77
8	2.69	2.32	2.16	1.93	1.75
9	2.65	2.30	2.14	1.92	1.74
10	2.62	2.27	2.12	1.91	1.73
15	2.53	2.21	2.07	1.87	1.71
20	2.48	2.18	2.05	1.85	1.70
25	2.46	2.17	2.03	1.84	1.69
50	2.41	2.13	2.01	1.82	1.67
100	2.38	2.12	1.99	1.81	1.67
∞	2.36	2.10	1.98	1.80	1.66

Αφού υπολογισθούν τα όρια αυτά, θα πρέπει στο μέλλον να επιθεωρούνται/αξιολογούνται περιοδικά. Εάν για οποιοδήποτε δείγμα το \bar{X} ή το R βρεθεί εκτός των ορίων ελέγχου, η διαδικασία θα πρέπει να εξετασθεί άμεσα προκειμένου να βρεθεί η αιτία της μετατόπισης και να ληφθεί η ανάλογη διορθωτική απόφαση. Εάν η διεργασία δεν είναι εφικτό να

αποκατασταθεί, τότε η διαδικασία θα πρέπει να επαναληφθεί από την αρχή προκειμένου να βρεθούν νέα όρια ελέγχου για το πρώτο και δεύτερο στάδιο της μεθόδου.

Στους Πίνακες 3.4, 3.5 και 3.6 παρατηρούμε, πως όταν το m αυξάνεται, τα A_2^* και D_4^* μειώνονται, ενώ το D_3^* αυξάνεται. Επιπλέον, τα όρια ελέγχου τείνουν να γίνουν πιο στενά και πιο ευαίσθητα στις μετατοπίσεις όταν το m (το οποίο χρησιμοποιήθηκε για τον υπολογισμό των \bar{X} και \bar{R}) αυξάνεται. Κάτι τέτοιο θα μπορούσε να σημαίνει τον συχνό υπολογισμό των ορίων ελέγχου, πιθανόν μετά από κάθε δείγμα. Κάτι τέτοιο όμως πρακτικά δεν θα ήταν επιθυμητό να συμβεί. Για το λόγο αυτό θα μπορούσαμε για παράδειγμα να υπολογίζουμε ξανά τα όρια ελέγχου για $m = 5, 10, 25$ και 100 , και πιθανόν οποτεδήποτε ένα σημείο πέσει κοντά στα όρια ελέγχου που δεν έχουν διορθωθεί πρόσφατα.

Για την επιλογή των ποσοτήτων a_2, a_3 και a_4 η διαδικασία είναι η ίδια που περιγράφηκε και στο προηγούμενο στάδιο. Όμως, δεν χρειάζεται να επιλέξουμε τις ίδιες τιμές με πριν. Το ιδανικό θα ήταν να μειώσουμε τις τιμές τους, ειδικά όταν επαναυπολογίζονται τα όρια ελέγχου με μεγαλύτερο αριθμό δειγμάτων m , προκειμένου να υπάρξει ισορροπία μεταξύ της πιθανότητας σφάλματος τύπου I και τύπου II.

Στην περίπτωση που κάποιο καινούργιο \bar{X} και R βρεθεί εκτός των ορίων ελέγχου στο δεύτερο στάδιο της μεθόδου η ερμηνεία που δίνουμε είναι παρόμοια με αυτά που ήδη αναφέραμε παραπάνω. Μια μικρή διαφοροποίηση παρουσιάζεται όταν κάποιο από τα δείγματα που αρχικά ήταν εντός ελέγχου, βρεθεί εκτός αυτών μετά από την τελευταία τροποποίηση των ορίων ελέγχου. Όταν συμβεί κάτι τέτοιο έχουμε ενδείξεις ότι η διεργασία πιθανώς ήταν και είναι ακόμα εκτός στατιστικού ελέγχου, παρόλο που δεν υπήρχαν επαρκή στοιχεία κατά την διάρκεια της μελέτης του δείγματος αυτού.

Επιπλέον, όταν σε κάποιο από τα δείγματα το \bar{X} ή το R βρεθούν εκτός των ορίων ελέγχου, θα πρέπει να υπολογισθούν ξανά τα όρια ελέγχου χωρίς όμως να ληφθούν υπόψη αυτά τα \bar{X} και R στον υπολογισμό των \bar{X} και \bar{R} . Συνεπώς, θα πρέπει να χρησιμοποιήσουμε τον αριθμό των εντός ελέγχου δειγμάτων που περιλαμβάνονται στον υπολογισμό των \bar{X} και \bar{R} για τα τρέχοντα όρια ελέγχου, και όχι τον συνολικό αριθμό των δειγμάτων που μελετάμε.

3.3.3 Υπολογισμός των σταθερών $A_2^*, A_2^{**}, D_3^*, D_3^{**}, D_4^*$ και D_4^{**}

Θεωρούμε έναν κανονικό πληθυσμό με μέσο \bar{X}' και τυπική απόκλιση σ' . Υποθέτουμε nm το πλήθος παρατηρήσεις έχουν επιλεγεί τυχαία από αυτόν τον πληθυσμό που διαιρούνται σε m

δείγματα (ομάδες) μεγέθους n , με ολικό μέσο $\bar{\bar{X}}$ και μέσο εύρος \bar{R} . Επιπλέον, οι ποσότητες \bar{X}_i και R_i συμβολίζουν αντίστοιχα το μέσο και το εύρος του i δείγματος, $i = 1, 2, \dots, m$. Τώρα υποθέτουμε ότι επιλέγονται *νεπιπλέον* τυχαίες παρατηρήσεις από τον ίδιο πληθυσμό, δηλαδή επιλέγεται ένα επιπλέον δείγμα. Συμβολίζουμε το μέσο και το εύρος αυτού του επιπλέον δείγματος με \bar{X} και R αντίστοιχα.

Για τα όρια ελέγχου στο δεύτερο στάδιο της μεθόδου για το \bar{X} διάγραμμα, η πιθανότητα a_2 ικανοποιεί τη σχέση

$$P\{\bar{\bar{X}} - A_2^* \bar{R} \leq \bar{X} \leq \bar{\bar{X}} + A_2^* \bar{R}\} = 1 - a_2$$

Παρατηρούμε ότι

$$1 - a_2 = P\{\bar{\bar{X}} - A_2^* \bar{R} \leq \bar{X} \leq \bar{\bar{X}} + A_2^* \bar{R}\} = P\left\{-A_2^* \leq \frac{\bar{X} - \bar{\bar{X}}}{\bar{R}} \leq A_2^*\right\}.$$

Για την κατανομή της στατιστικής συνάρτησης $(\bar{X} - \bar{\bar{X}})/\bar{R}$ έχουμε ότι ο αριθμητής $(\bar{X} - \bar{\bar{X}})$ έχει κανονική κατανομή με μέση τιμή μηδέν και τυπική απόκλιση $\sigma' \sqrt{(m+1)/nm}$. Ο Patnaik (1950) έδειξε ότι η στατιστική συνάρτηση $\bar{R}^2/c^2\sigma'^2$ ακολουθεί προσεγγιστικά την κατανομή χ^2 με ν βαθμούς ελευθερίας, όπου τα ν και c είναι συναρτήσεις των m και n . Ο Resnikoff (1954) απέδειξε ότι για $n = 5$ αυτή είναι μια άριστη προσέγγιση ακόμα και για μικρές τιμές του m . Σε αυτό το σημείο αξίζει να σημειώσουμε ότι η κατανομή τείνει ο λόγος μιας τυποποιημένης κανονικής τυχαίας μεταβλητής διά την τετραγωνική ρίζα ενός κλάσματος που έχει ως αριθμητή μια ανεξάρτητη τυχαία μεταβλητή με κατανομή χ τετράγωνο και ως παρονομαστή τους βαθμούς ελευθερίας αυτής της τυχαίας μεταβλητής. Συνεπώς, δεδομένου ότι τα $(\bar{X} - \bar{\bar{X}})$ και \bar{R} είναι ανεξάρτητα, η στατιστική συνάρτηση

$$\frac{(\bar{X} - \bar{\bar{X}})/\sqrt{\frac{m+1}{nm}} \sigma'}{\sqrt{\frac{\nu \bar{R}^2/c^2 \sigma'^2}{\nu}}} = c \sqrt{\frac{nm}{m+1}} \left(\frac{(\bar{X} - \bar{\bar{X}})}{\bar{R}} \right)$$

ακολουθεί προσεγγιστικά την t κατανομή με ν βαθμούς ελευθερίας.

Έστω ότι η ποσότητα $t_{\beta, \nu}$ συμβολίζει εκείνο το ποσοστιαίο σημείο της τυχαίας μεταβλητής τ με κατανομή $(t_{\nu}, \tau \sim t_{\nu})$ που ικανοποιεί τη σχέση

$$P\{\tau > t_{\beta, \nu}\} = \beta$$

όπου $\tau \sim t_\nu$. Από τη σχέση a_2

$$P(-t_{a_2/2, \nu} \leq \tau \leq t_{a_2/2, \nu}) = 1 - a_2$$

παίρνουμε ότι

$$P(-t_{a_2/2, \nu} \leq c \sqrt{\frac{nm}{m+1}} \left(\frac{\bar{X} - \bar{\bar{X}}}{\bar{R}} \right) \leq t_{a_2/2, \nu}) = 1 - a_2$$

και μπορούμε πλέον να συμπεράνουμε ότι

$$t_{a_2/2, \nu} = A_2^* c \sqrt{\frac{nm}{m+1}}$$

ή ισοδύναμα

$$A_2^* = \frac{\sqrt{\frac{m+1}{nm}}}{c} t_{a_2/2, \nu}.$$

Εάν τώρα, θέλουμε να προσδιορίσουμε το A_2^{**} για τον υπολογισμό των ορίων ελέγχου στο πρώτο στάδιο της μεθόδου για το \bar{X} διάγραμμα, το μόνο που αλλάζει είναι ότι η τυχαία μεταβλητή που παρατηρείται είναι ο μέσος από καθένα από τα αρχικά δείγματα \bar{X}_i και όχι ο μελλοντικός μέσος \bar{X} .

Έτσι, η πιθανότητα που μας ενδιαφέρει να εξετάσουμε είναι της μορφής

$$P \left\{ -A_2^{**} \leq \frac{\bar{X}_i - \bar{\bar{X}}}{\bar{R}} \leq A_2^{**} \right\} = 1 - a_2.$$

Παρατηρούμε ότι

$$\bar{X}_i - \bar{\bar{X}} = \bar{X}_i - \frac{\sum_{j=1}^m \bar{X}_j}{m} = \frac{m-1}{m} \bar{X}_i - \frac{1}{m} \sum_{\substack{i=1 \\ j \neq i}}^m \bar{X}_j.$$

Έτσι

$$E(\bar{X}_i - \bar{\bar{X}}) = 0,$$

$$\text{Var}(\bar{X}_i - \bar{\bar{X}}) = \frac{(m-1)^2 + (m-1)}{m^2} \cdot \frac{\sigma'^2}{n} = \frac{m-1}{nm} \sigma'^2.$$

Ακολουθώντας την ίδια διαδικασία με πριν παίρνουμε ότι

$$A_2^{**} = \frac{\sqrt{\frac{m-1}{nm}}}{c} t_{a_2/2, \nu} = \sqrt{\frac{m-1}{m+1}} A_2^*.$$

Στη συνέχεια, θα θεωρήσουμε τα όρια ελέγχου για το δεύτερο στάδιο της μεθόδου για το R διάγραμμα. Η πιθανότητα που θα μελετήσουμε είναι της μορφής

$$P\{R \leq K_a \bar{R}\} = P\left\{\frac{R}{\bar{R}} \leq K_a\right\} = a$$

όπου K_a συμβολίζει το a -ποσοστιαίο σημείο της κατανομής του R/\bar{R} . Έτσι λοιπόν, το πρόβλημα επικεντρώνεται στο να βρούμε την κατανομή του R/\bar{R} . Με εφαρμογή και σε αυτή την περίπτωση της προσέγγισης του Patnaik που περιγράψαμε νωρίτερα, καταλήγουμε πως η ποσότητα cR/\bar{R} έχει την κατανομή του studentized εύρους του η οποία παρουσιάζεται με εκτενείς πίνακες από τους Harter, Clemm, και Guthrie (1959) και Harter (1960). Εάν η τυχαία μεταβλητή Q ακολουθεί την κατανομή του studentized εύρους, τότε

$$P\{Q \leq q_{\alpha, \nu}\} = a.$$

Συνεπώς,

$$a = P\left\{\frac{cR}{\bar{R}} \leq q_{\alpha, \nu}\right\} = P\left\{\frac{R}{\bar{R}} \leq \frac{q_{\alpha, \nu}}{c}\right\}.$$

Έτσι,

$$D_3^* = \frac{q_{a_3, \nu}}{c}, \quad D_4^* = \frac{q_{1-a_4, \nu}}{c}$$

Δοθέντων των τιμών m, a_3 και a_4 , για τον υπολογισμό των ορίων ελέγχου στο πρώτο στάδιο του R διαγράμματος, είναι ξεκάθαρο ότι οι απαιτούμενοι παράγοντες D_3^{**} και D_4^{**} θα δοθούν συναρτήσεις κάποιων κατάλληλων τιμών D_3^* και D_4^* . Πιο συγκεκριμένα, έστω ότι οι ποσότητες $D_{3, m-1}^*$ και $D_{4, m-1}^*$ συμβολίζουν τις τιμές των D_3^* και D_4^* για τα όρια ελέγχου στο δεύτερο στάδιο, βασισμένα σε $m - 1$ δείγματα, αλλά με τις ίδιες τιμές στα a_3 και a_4 .

Για κάθε μελλοντικό δείγμα με εύρος R έχουμε

$$\alpha_3 = P(R \leq \bar{R}_{m-1} D_{3, m-1}^*).$$

Έτσι για το m -οστό δείγμα με εύρος R_m και θέτοντας $(\bar{R} = (R_1 + R_2 + \dots + R_m)/m)$ έχουμε

$$\begin{aligned} \alpha_3 &= P \left\{ R_m \leq D_{3;m-1}^* \left(\frac{m\bar{R} - R_m}{m-1} \right) \right\} = P \{ R_m (m-1 + D_{3;m-1}^*) \leq D_{3;m-1}^* m\bar{R} \} \\ &= P \left\{ R_m \leq \frac{mD_{3;m-1}^*}{m-1 + D_{3;m-1}^*} \bar{R} \right\} \end{aligned}$$

Έτσι,

$$D_3^{**} = \frac{mD_{3;m-1}^*}{m-1 + D_{3;m-1}^*}.$$

Ακολουθώντας την ίδια διαδικασία παίρνουμε

$$D_4^{**} = \frac{mD_{4;m-1}^*}{m-1 + D_{4;m-1}^*}.$$

3.4 Εφαρμογή

Υποθέτουμε πως μια βιομηχανία επιθυμεί να ελέγξει ένα συγκεκριμένο ποιοτικό χαρακτηριστικό ενός από τα προϊόντα της, για το οποίο φαίνεται να παρουσιάζεται συχνά πρόβλημα κατά την διαδικασία παραγωγής. Η προσέγγιση αυτού του ελέγχου θα γίνει βάσει της δισταδιακής μεθόδου που περιγράφηκε στην προηγούμενη ενότητα.

Αποφασίστηκε να χρησιμοποιηθεί δείγμα μεγέθους πέντε, συνεπώς για τον υπολογισμό των ορίων μπορεί να γίνει χρήση των πινάκων που δόθηκαν παραπάνω. Προκειμένου να υπάρξει ισορροπία μεταξύ αξιοπιστίας και ευαισθησίας στα διαγράμματα, επιλέγουμε για το πρώτο και το δεύτερο στάδιο της μεθόδου $\alpha_2 = 0.0027$, $\alpha_3 = 0.001$ και $\alpha_4 = 0.005$. Επιπλέον, θα ενεργήσουμε ξεκινώντας από το πρώτο στάδιο της μεθόδου και στη συνέχεια, αφού εξετάσουμε τα πρώτα πέντε δείγματα θα προχωρήσουμε στο δεύτερο στάδιο της μεθόδου, με μια ενημέρωση των ορίων ελέγχου στο δεύτερο στάδιο μετά από 10,25 και 100 εντός ελέγχου ορίων.

Τα αποτελέσματα αυτού του σχεδίου παρουσιάζονται στον κάτωθι πίνακα.

Subgroup Number	\bar{X}	R	$\bar{\bar{X}}$	\bar{R}	m	A_2^{**}	A_2^*	D_3^{**}	D_3^*	D_4^{**}	D_4^*	\bar{X} chart		R chart		
												LCL	UCL	LCL	UCL	
1	505.6	17														
2	501.2	9														
3	508.0	14														
4	507.6	37														
5	504.4	12	505.36	17.8	5	0.588		0.182		1.96		494.9	515.8	3.2	34.9	
			504.80	13.0	4	0.580		0.189		1.92		497.1	512.5	2.5	25.0	
					4		0.760		0.151		2.57	494.9	514.7	2.0	32.6	
6	504.8	19										494.9	514.7	2.0	32.6	
7	503.6	11										494.9	514.7	2.0	32.6	
8	493.2	29										494.9	514.7	2.0	32.6	
9	509.0	15										494.9	514.7	2.0	32.6	
10	500.2	8										494.9	514.7	2.0	32.6	
11	504.6	21										494.9	514.7	2.0	32.6	
12	501.8	16	504.32	14.2	10		0.645		0.155		2.27	495.2	513.5	2.2	32.2	

Όπως προκύπτει από τα αποτελέσματα, όταν στήθηκαν τα αρχικά όρια ελέγχου με βάση το πρώτο στάδιο της μεθόδου, βρέθηκε ότι το εύρος του τέταρτου δείγματος έπεφτε πάνω από το άνω όριο ελέγχου του R διαγράμματος. Αυτό είχε ως αποτέλεσμα, η βιομηχανία να διερευνήσει τις αιτίες οι οποίες μπορεί να προκάλεσαν αυτή την προσωρινή απόκλιση και στη συνέχεια να μεριμνήσει ώστε να μην προκληθεί ξανά κάτι παρόμοιο. Τα διαγράμματα \bar{X} και R για το συγκεκριμένο δείγμα απορρίφθηκαν και τα όρια υπολογίστηκαν ξανά με βάση τα άλλα τέσσερα δείγματα. Από τη στιγμή που τα \bar{X} και R για τα εναπομείναντα δείγματα έπεφταν εντός των ορίων, και δεν παρατηρήθηκε κάποιο άλλο στοιχείο το οποίο να προκαλέσει αμφιβολία για τη σταθερότητα της διεργασίας, τα όρια ελέγχου στο δεύτερο στάδιο υπολογίστηκαν για τον έλεγχο μεταγενέστερων δειγμάτων. Η διαδικασία συνεχίστηκε χωρίς κάποια επιπλέον ένδειξη, μέχρι το τρίτο στη σειρά από τα καινούργια δείγματα, του οποίου ο μέσος έπεφτε πιο κάτω από το κάτω όριο του διαγράμματος \bar{X} . Μετά από την απαραίτητη διερεύνηση, ανακαλύφθηκε πως επρόκειτο για ειδική αιτία μεταβλητότητας, η οποία στη συνέχεια αποκαταστήθηκε. Λαμβάνοντας επιπλέον δείγματα, προέκυψε πως η διεργασία παραμένει εντός στατιστικού ελέγχου. Τα όρια για το δεύτερο στάδιο υπολογίστηκαν ξανά στο δωδέκατο δείγμα με βάση τα δέκα εντός ελέγχου δείγματα, όπως προκύπτει και από τον παραπάνω πίνακα.

Λαμβάνοντας υπόψη αυτή την εφαρμογή, γίνεται αντιληπτό πως η έγκαιρη διάγνωση μέσω της διαδικασίας που περιγράφηκε στις παραπάνω ενότητες, οδήγησε σε μια σταθερή παραγωγική διεργασία πιο γρήγορα σε σχέση με την εφαρμογή των τυπικών ορίων ελέγχου.

3.5 Διαγράμματα ελέγχου για το μέσο βασισμένα στις ποσότητες \bar{X} και \bar{u}

Όπως έχουμε ήδη δει και σε προηγούμενες ενότητες, τα όρια ελέγχου για το διάγραμμα ελέγχου του μέσου συνήθως βασίζονται στα \bar{X} και \bar{R} . Όμως, σύμφωνα με την προσέγγιση του Patnaik (1950) καθώς και βάσει του πίνακα D3 του Duncan (1965, p. 910) αποδείχθηκε ότι υπάρχει προσεγγιστικά μια διαφορά της τάξης του 10% στις διακυμάνσεις των εκτιμητών του σ^2 , $\bar{u} = \bar{s}^2 = \frac{\sum_{i=1}^m s_i^2}{m}$ και ενός πολλαπλασίου του \bar{R}^2 στην περίπτωση που $n = 5$. Σε κάθε περίπτωση ένα πολλαπλάσιο του \bar{R}^2 είναι ελαφρώς λιγότερο αποτελεσματικός εκτιμητής του σ^2 σε σχέση με το \bar{u} .

Στη συνέχεια θεωρούμε ότι η διεργασία που μελετάμε βρίσκεται ήδη εντός στατιστικού ελέγχου και θέλουμε να βρούμε όρια πιθανοτήτων για να ελέγξουμε τις μελλοντικές παρατηρήσεις. Δοθέντων των m , n και a , έστω ότι τα όρια ελέγχου είναι της μορφής

$$LCL = \bar{X} - A_4^* \sqrt{\bar{u}}$$

$$UCL = \bar{X} + A_4^* \sqrt{\bar{u}}.$$

Έτσι, το πρόβλημα επικεντρώνεται στην εύρεση του A_4^* , έτσι ώστε

$$P\{\bar{X} - A_4^* \sqrt{\bar{u}} \leq \bar{X}_f \leq \bar{X} + A_4^* \sqrt{\bar{u}}\} = 1 - a$$

όπου το (μελλοντικό) δείγμα που έδωσε το \bar{X}_f ήταν εντός ελέγχου δείγμα. Αρχικά, παρατηρούμε ότι

$$1 - a = P\{\bar{X} - A_4^* \sqrt{\bar{u}} \leq \bar{X}_f \leq \bar{X} + A_4^* \sqrt{\bar{u}}\} = P\left\{-A_4^* \leq \frac{\bar{X}_f - \bar{X}}{\sqrt{\bar{u}}} \leq A_4^*\right\}.$$

Έτσι είναι απλά αναγκαίο να βασιστούμε στα ποσοστιαία σημεία της κατανομής του $(\bar{X}_f - \bar{X})/\sqrt{\bar{u}}$. Δεδομένου ότι υποθέτουμε πως οι μεμονωμένες παρατηρήσεις είναι ανεξάρτητες και προέρχονται από την ίδια κανονική κατανομή με διακύμανση σ^2 , η στατιστική συνάρτηση $(\bar{X}_f - \bar{X})$ ακολουθεί την κανονική κατανομή με μέσο μηδέν και διακύμανση $((1/n) + (1/N))\sigma^2$, όπου $N = nm$. Επιπλέον, η ποσότητα

$(N - m)\bar{u}/\sigma^2$ ακολουθεί την χ^2 κατανομή με $N - m$ βαθμούς ελευθερίας αφού μπορεί να γραφεί ως άθροισμα m ανεξάρτητων κατανομών χ^2 με $n - 1$ βαθμούς ελευθερίας καθεμιά. Συνεπώς, η στατιστική συνάρτηση

$$\frac{(\bar{X}_f - \bar{X})/\sigma \left(\frac{1}{n} + \frac{1}{N}\right)^{1/2}}{\left[\frac{(N-m)\bar{u}}{\sigma^2}\right]^{1/2}/(N-m)^{1/2}} = \left(\frac{N}{m+1}\right)^{1/2} \left(\frac{\bar{X}_f - \bar{X}}{\sqrt{\bar{u}}}\right)$$

ακολουθεί την κατανομή t με $N - m$ βαθμούς ελευθερίας, δηλαδή την κατανομή t_{N-m} .

Έστω ότι η ποσότητα $t_{\beta,\nu}$ συμβολίζει εκείνο το ποσοστιαίο σημείο της τυχαίας μεταβλητής τ με κατανομή t_ν ($\tau \sim t_\nu$) που ικανοποιεί τη σχέση

$$P\{\tau > t_{\beta,\nu}\} = \beta$$

όπου $\tau \sim t_\nu$. Έτσι $t_{1-\beta,\nu} = -t_{\beta,\nu}$. Επομένως, από τις σχέσεις

$$P\left\{-A_4^* \leq \frac{\bar{X}_f - \bar{X}}{\sqrt{\bar{u}}} \leq A_4^*\right\} = 1 - \alpha$$

και

$$P\left\{-t_{\alpha/2, N-m} \leq \left(\frac{N}{m+1}\right)^{1/2} \left(\frac{\bar{X}_f - \bar{X}}{\sqrt{\bar{u}}}\right) \leq t_{\alpha/2, N-m}\right\} = 1 - \alpha$$

παίρνουμε άμεσα ότι

$$t_{\alpha/2, N-m} = \left(\frac{N}{m+1}\right)^{1/2} A_4^*$$

δηλαδή

$$A_4^* = \left(\frac{m+1}{N}\right)^{1/2} t_{\alpha/2; N-m}.$$

Στην περίπτωση που επιθυμούμε να ελέγξουμε αναδρομικά αν η διεργασία ήταν εντός στατιστικού ελέγχου όταν συλλέχθηκαν τα δείγματα, θα χρησιμοποιήσουμε μια εναλλακτική μέθοδο. Η προσέγγιση αυτή, κάνει χρήση ενός μόνο ζεύγους ορίων, έστω της μορφής $\bar{X} \pm A_4^{**} \sqrt{\bar{u}}$ που θα αποδώσουν την επιθυμητή πιθανότητα σφάλματος τύπου I για κάθε δείγμα i όταν το \bar{X}_i συμπεριλαμβάνεται στον υπολογισμό του \bar{X} .

Αφού

$$\bar{X}_i - \bar{\bar{X}} = \bar{X}_i - \frac{\sum_{j=1}^m \bar{X}_j}{m} = \frac{m-1}{m} \bar{X}_i - \frac{1}{m} \sum_{\substack{j=1 \\ j \neq i}}^m \bar{X}_j$$

τότε

$$\text{Var}(\bar{X}_i - \bar{\bar{X}}) = \frac{(m-1)^2 + (m-1)}{m^2} \cdot \frac{\sigma^2}{n} = \frac{m-1}{N} \sigma^2, \quad i = 1, 2, \dots, m.$$

Συνεπώς, εργαζόμενοι όπως πριν, καταλήγουμε στη σχέση

$$A_4^{**} = \left[\frac{m-1}{N} \right]^{1/2} t_{\alpha/2; N-m} = \left[\frac{m-1}{m+1} \right]^{1/2} A_4^*.$$

Η μέθοδος που αναπτύξαμε παραπάνω μπορεί να προκαλέσει κάποιες εύλογες ενστάσεις.

- Το γεγονός ότι ο δειγματικός μέσος \bar{X}_i συμπεριλαμβάνεται στην ποσότητα $\bar{\bar{X}}$, οδηγεί στην μείωση της ισχύος του ελέγχου, καθώς ένας εκτός ελέγχου δειγματικός μέσος \bar{X}_i μπορεί να «τραβήξει» μαζί του και το $\bar{\bar{X}}$, ειδικά για μικρές τιμές του m .
- Ο αριθμητικός μέσος $\bar{\bar{X}}$ μπορεί να περιλαμβάνει και εκτός ελέγχου δειγματικούς μέσους \bar{X}_i και έτσι αυξάνεται ο κίνδυνος να έχουμε κάποια εντός ελέγχου \bar{X}_i τα οποία να βρεθούν τελικά εκτός των ορίων ελέγχου.
- Όταν έχουν βρεθεί κάποια εκτός ελέγχου \bar{X}_i και έχουν αφαιρεθεί από το δείγμα, πρέπει να υπολογίσουμε ξανά την ποσότητα $\sqrt{\bar{u}}$ ενδεχομένως και παραπάνω από μια φορές εάν υπάρχουν ακόμα εκτός ελέγχου παρατηρήσεις.
- Επειδή ο βασικός στόχος της προσέγγισης αυτής είναι να ελέγξουμε εάν η διεργασία ήταν εντός ελέγχου όταν συλλέξαμε το δείγμα και όχι να ταυτοποιήσουμε τα μεμονωμένα εκτός ελέγχου δείγματα, θα ήταν χρήσιμο να ορίσουμε ένα κατάλληλο επίπεδο σημαντικότητας για τον συνολικό έλεγχο και όχι για κάθε μεμονωμένο δείγμα.

Ωστόσο, υπάρχουν και αρκετά στοιχεία της προσέγγισης αυτής για τον έλεγχο των αρχικών δειγμάτων που αιτιολογούν την εφαρμογή της.

- Το γεγονός ότι όλα τα δείγματα μπορούν να ελεγχθούν ταυτόχρονα καθιστά την μέθοδο αυτή ιδιαίτερα πρακτική.
- Λόγω της χρήσης της ποσότητας \bar{u} κρίνεται ως αποτελεσματική.

- Χρησιμοποιεί ένα ολικό επίπεδο σημαντικότητας όταν το m λαμβάνει μικρές τιμές.
- Όταν προκύπτει πως η διεργασία που μελετάμε βρισκόταν εκτός ελέγχου, η μέθοδος αυτή μας βοηθάει να καθορίσουμε και να διορθώσουμε τις αιτίες που το προκάλεσαν.
- Αν οι μετατοπίσεις είναι προσωρινές, μπορούμε να χρησιμοποιήσουμε τα υπόλοιπα δείγματα για τον υπολογισμό των ορίων ελέγχου.

Οι τιμές των A_4^* και A_4^{**} παρουσιάζονται στους δύο ακόλουθους πίνακες για μέγεθος δείγματος ίσο με 5.

Πίνακας 3.7. A_4^* : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

m	α			
	0.001	0.002	0.01	0.05
1	5.44	4.54	2.91	1.76
2	2.76	2.47	1.84	1.26
3	2.23	2.03	1.58	1.13
4	2.01	1.84	1.46	1.06
5	1.89	1.74	1.39	1.02
6	1.80	1.67	1.35	1.00
7	1.76	1.63	1.32	0.981
8	1.72	1.60	1.30	0.966
9	1.69	1.57	1.28	0.956
10	1.67	1.55	1.27	0.948
15	1.60	1.49	1.23	0.924
20	1.56	1.46	1.21	0.912
25	1.55	1.45	1.20	0.905
50	1.51	1.41	1.17	0.891
100	1.49	1.40	1.16	0.884
∞	1.47	1.38	1.15	0.877

Πίνακας 3.8. A_4^{**} : Παράγοντες για τα όρια ελέγχου του \bar{X} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

m	α			
	0.001	0.002	0.01	0.05
2	1.59	1.42	1.06	0.729
3	1.58	1.43	1.12	0.795
4	1.55	1.43	1.13	0.821
5	1.54	1.42	1.14	0.834
6	1.52	1.41	1.14	0.840
7	1.53	1.42	1.15	0.850
8	1.52	1.41	1.15	0.852
9	1.51	1.40	1.15	0.855
10	1.51	1.40	1.15	0.857
15	1.49	1.40	1.15	0.864
20	1.49	1.39	1.15	0.867
25	1.48	1.39	1.15	0.869
50	1.48	1.39	1.15	0.873
100	1.47	1.38	1.15	0.875
∞	1.47	1.38	1.15	0.877

3.6 Διαγράμματα ελέγχου για τη διασπορά βασισμένα στην ποσότητα \bar{u}

Πέρα από τον έλεγχο του μέσου μιας διεργασίας μέσω του αντίστοιχου διαγράμματος, είναι εξίσου σημαντικός και ο έλεγχος της διασποράς της διεργασίας. Αυτό συνήθως γίνεται μέσω του διαγράμματος R . Όμως, προκειμένου να διεξάγουμε έναν πιο ισχυρό έλεγχο, αξίζει να απεικονίσουμε τη δειγματική διακύμανση (ή τη δειγματική τυπική απόκλιση) στο διάγραμμα και να βασίσουμε τα όρια ελέγχου στην ποσότητα \bar{u} . Τέτοιου είδους διαγράμματα περιγράφονται στη συνέχεια.

Λαμβάνοντας υπόψη τον τρόπο που βρίσκουμε όρια πιθανοτήτων για μικρά δείγματα για τον έλεγχο μελλοντικών δειγμάτων, έχουμε

$$LCL = B_7^* \bar{u}$$

$$UCL = B_8^* \bar{u}$$

Αφού δοθούν τα m, n, a_L (πιθανότητα να βρεθεί ένα σημείο κάτω από το κάτω όριο ελέγχου) και a_U (πιθανότητα να βρεθεί ένα σημείο πάνω από το πάνω όριο ελέγχου), το πρόβλημα είναι ο καθορισμός των B_7^* και B_8^* με τέτοιο τρόπο ώστε

$$P\{s_f^2 < B_7^* \bar{u}\} = a_L$$

$$P\{s_f^2 > B_8^* \bar{u}\} = a_U$$

όταν τα \bar{u} και s_f^2 επιλέχθηκαν όταν η διαδικασία βρισκόταν εντός ελέγχου. Για να βρούμε τις ποσότητες B_7^* και B_8^* , παρατηρούμε ότι

$$P\{B_7^* \bar{u} \leq s_f^2 \leq B_8^* \bar{u}\} = P\left\{B_7^* \leq \frac{s_f^2}{\bar{u}} \leq B_8^*\right\}$$

και επιπλέον η στατιστική s_f^2/\bar{u} ακολουθεί την κατανομή F με $n-1$ και $m(n-1) = N-m$ βαθμούς ελευθερίας, αφού

$$\frac{s_f^2}{\bar{u}} = \frac{\frac{(n-1)s_f^2}{\sigma^2}/(n-1)}{\frac{(n-1)\sum_{i=1}^m s_i^2/\sigma^2/m(n-1)}$$

Έστω ότι η ποσότητα F_{β, ν_1, ν_2} , συμβολίζει εκείνο το ποσοστιαίο σημείο της τυχαίας μεταβλητής Y με κατανομή F_{ν_1, ν_2} ($Y \sim F_{\nu_1, \nu_2}$) που ικανοποιεί τη σχέση

$$P\{Y > F_{\beta, \nu_1, \nu_2}\} = \beta.$$

Αυτό συνεπάγεται ότι

$$B_7^* = F_{1-\alpha_L, n-1, N-m}$$

$$B_8^* = F_{1-\alpha_U, n-1, N-m}$$

Το διάγραμμα διασποράς που αναφέρθηκε παραπάνω, μπορεί να μετατραπεί και στο αντίστοιχο διάγραμμα τυπικής απόκλισης, χρησιμοποιώντας τα αντίστοιχα όρια αυτή τη φορά όμως βάση της ποσότητας $\sqrt{\bar{u}}$.

Έτσι τα όρια πιθανοτήτων για το διάγραμμα της τυπικής απόκλισης θα είναι

$$LCL = \sqrt{B_7^* \bar{u}}$$

$$UCL = \sqrt{B_8^* \bar{u}}.$$

Αυτός ο τρόπος ελέγχου για το αν μια διαδικασία είναι εντός στατιστικού ελέγχου είναι προφανώς ισοδύναμος με αυτόν της διασποράς. Το πλεονέκτημα όμως του διαγράμματος τυπικής απόκλισης είναι ότι η κλίμακα υπολογισμού είναι πιο εύκολο να εξηγηθεί, κάτι που βοηθάει την αξιολόγηση της σημαντικότητας των σημείων κοντά στα όρια ελέγχου. Από την άλλη, χρειάζεται επιπλέον υπολογισμό (τετραγωνική ρίζα) για κάθε δείγμα προκειμένου να μετατραπεί η δειγματική διακύμανση σε δειγματική τυπική απόκλιση.

Προκειμένου να χρησιμοποιήσουμε την ποσότητα \bar{u} για να ελέγξουμε αναδρομικά τα αρχικά δείγματα, παρατηρούμε ότι

$$\frac{s_i^2 / \sigma^2}{[m(n-1)\bar{u} - (n-1)s_i^2] / (m-1)(n-1)\sigma_i^2} = \frac{(m-1)s_i^2}{m\bar{u} - s_i^2}$$

ακολουθεί κατανομή F με βαθμούς ελευθερίας ίσους με $n-1$ και $(m-1)(n-1)$, δηλαδή

$$\frac{(m-1)s_i^2}{m\bar{u} - s_i^2} \sim F_{n-1, (m-1)(n-1)}.$$

Συνεπώς,

$$s_i^2 \sim \frac{mF_{n-1, (m-1)(n-1)}}{m-1 + F_{n-1, (m-1)(n-1)}} \bar{u}.$$

Παρατηρούμε ότι συνάρτηση $mx/(m-1+x)$ είναι μια μονότονη γνησίως αύξουσα συνάρτηση του x για $m > 1$. Έτσι, κάνοντας χρήση της προσέγγισης που εφαρμόσαμε για το διάγραμμα μέσου βασισμένοι στις ποσότητες $\bar{\bar{X}}$ και \bar{u} , προκύπτει ότι τα προσαρμοσμένα όρια ελέγχου για τον έλεγχο των πρωταρχικών δειγμάτων με τη χρήση διαγραμμάτων διασποράς είναι

$$LCL = B_7^{**} \bar{u}$$

$$UCL = B_8^{**} \bar{u}$$

όπου

$$B_7^{**} = \frac{mF_{1-\alpha_L; n-1, (m-1)(n-1)}}{m-1 + F_{1-\alpha_L; n-1, (m-1)(n-1)}},$$

$$B_8^{**} = \frac{mF_{1-\alpha_U; n-1, (m-1)(n-1)}}{m-1 + F_{1-\alpha_U; n-1, (m-1)(n-1)}}.$$

Αντίστοιχα, τα προσαρμοσμένα όρια ελέγχου για την τυπική απόκλιση διαμορφώνονται ως εξής

$$LCL = \sqrt{B_7^{**} \bar{u}},$$

$$UCL = \sqrt{B_8^{**} \bar{u}}.$$

Σε άλλες περιπτώσεις, η προσέγγιση αυτή για τον έλεγχο των δειγμάτων έχει ουσιαστικά τα ίδια πλεονεκτήματα και μειονεκτήματα με αυτά του διαγράμματος μέσου με χρήση της ποσότητας \bar{u} .

Οι τιμές των $B_7^*, \sqrt{B_7^*}, B_8^*, \sqrt{B_8^*}, B_7^{**}, \sqrt{B_7^{**}}$ και $B_8^{**}, \sqrt{B_8^{**}}$ παρουσιάζονται στους ακόλουθους πίνακες για μέγεθος δείγματος ίσο με 5.

Πίνακας 3.9 $B_7^*, \sqrt{B_7^*}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

α_L	0.001		0.005		0.025	
	B_7^*	$\sqrt{B_7^*}$	B_7^*	$\sqrt{B_7^*}$	B_7^*	$\sqrt{B_7^*}$
1	0.0187	0.137	0.0431	0.208	0.104	0.322
2	0.0204	0.143	0.0467	0.216	0.111	0.333
3	0.0221	0.145	0.0483	0.220	0.114	0.338
4	0.0215	0.147	0.0490	0.221	0.116	0.341
5	0.0217	0.147	0.0495	0.222	0.117	0.342
6	0.0218	0.148	0.0500	0.224	0.118	0.344
7	0.0220	0.148	0.0503	0.224	0.118	0.344
8	0.0221	0.149	0.0503	0.224	0.118	0.344
9	0.0221	0.149	0.0505	0.225	0.118	0.344
10	0.0222	0.149	0.0505	0.225	0.119	0.345
15	0.0223	0.149	0.0510	0.226	0.120	0.346
20	0.0224	0.150	0.0513	0.226	0.120	0.346
25	0.0225	0.150	0.0513	0.226	0.120	0.346
50	0.0226	0.150	0.0515	0.227	0.121	0.348
100	0.0227	0.151	0.0515	0.227	0.121	0.348
∞	0.0227	0.151	0.0518	0.228	0.121	0.348

Πίνακας 3.10 $B_8^*, \sqrt{B_8^*}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

α_U	0.001		0.005		0.025	
	B_8^*	$\sqrt{B_8^*}$	B_8^*	$\sqrt{B_8^*}$	B_8^*	$\sqrt{B_8^*}$
1	53.4	7.31	23.2	4.82	9.60	3.10
2	14.4	3.79	8.81	2.97	5.05	2.25
3	9.63	3.10	6.52	2.55	4.12	2.03
4	7.94	2.82	5.64	2.37	3.73	1.93
5	7.10	2.66	5.17	2.27	3.51	1.87
6	6.59	2.57	4.89	2.21	3.38	1.84
7	6.25	2.50	4.70	2.17	3.29	1.81
8	6.02	2.45	4.56	2.14	3.22	1.79
9	5.84	2.42	4.46	2.11	3.17	1.78
10	5.70	2.39	4.37	2.09	3.13	1.77
15	5.31	2.30	4.14	2.03	3.01	1.73
20	5.13	2.26	4.03	2.01	2.95	1.72
25	5.01	2.24	3.96	1.99	2.92	1.71
50	4.81	2.19	3.84	1.96	2.85	1.69
100	4.71	2.71	3.78	1.94	2.82	1.68
∞	4.62	2.15	3.72	1.93	2.79	1.67

Πίνακας 3.11 $B_7^{**}, \sqrt{B_7^{**}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

α_L	0.001		0.005		0.025	
	B_7^{**}	$\sqrt{B_7^{**}}$	B_7^{**}	$\sqrt{B_7^{**}}$	B_7^{**}	$\sqrt{B_7^{**}}$
2	0.0368	0.192	0.0826	0.287	0.189	0.435
3	0.0303	0.174	0.0685	0.262	0.158	0.397
4	0.0279	0.167	0.0634	0.252	0.147	0.383
5	0.0267	0.163	0.0605	0.246	0.141	0.375
6	0.0259	0.161	0.0588	0.242	0.137	0.370
7	0.0254	0.159	0.0579	0.241	0.134	0.366
8	0.0250	0.158	0.0570	0.239	0.133	0.365
9	0.0248	0.157	0.0562	0.237	0.131	0.362
10	0.0245	0.157	0.0558	0.236	0.130	0.361
15	0.0239	0.155	0.0545	0.233	0.127	0.356
20	0.0236	0.154	0.0538	0.232	0.126	0.355
25	0.0234	0.153	0.0533	0.231	0.125	0.354
50	0.0230	0.152	0.0525	0.229	0.123	0.351
100	0.0229	0.151	0.0520	0.228	0.122	0.349
∞	0.0227	0.151	0.0518	0.228	0.121	0.348

Πίνακας 3.12 $B_8^{**}, \sqrt{B_8^{**}}$: Παράγοντες για τα όρια ελέγχου του \bar{u} διαγράμματος ($n = 5$), (Yang, Hillier (1970))

α_U	0.001		0.005		0.025	
	B_8^{**}	$\sqrt{B_8^{**}}$	B_8^{**}	$\sqrt{B_8^{**}}$	B_8^{**}	$\sqrt{B_8^{**}}$
2	1.96	1.40	1.92	1.39	1.81	1.35
3	2.63	1.62	2.44	1.56	2.15	1.47
4	3.05	1.75	2.74	1.66	2.31	1.52
5	3.32	1.82	2.93	1.71	2.41	1.55
6	3.52	1.88	3.05	1.75	2.47	1.57
7	3.66	1.91	3.14	1.77	2.52	1.59
8	3.77	1.94	3.21	1.79	2.56	1.60
9	3.86	1.96	3.27	1.81	2.58	1.61
10	3.94	1.98	3.31	1.82	2.60	1.61
15	4.15	2.04	3.44	1.85	2.67	1.63
20	4.27	2.07	3.51	1.87	2.70	1.64
25	4.33	2.08	3.55	1.88	2.71	1.65
50	4.47	2.11	3.63	1.91	2.75	1.66
100	4.54	2.13	3.68	1.92	2.77	1.66
∞	4.62	2.15	3.72	1.93	2.79	1.67

3.7 Διαγράμματα ελέγχου για το μέσο βασισμένα στις ποσότητες \bar{X} και s

Στην προηγούμενη ενότητα ασχοληθήκαμε με τα διαγράμματα μέσου που βασίζονται στις ποσότητες \bar{X} και \bar{u} . Η χρήση τους ήταν αποτέλεσμα της διαπίστωσης ότι ένα πολλαπλάσιο του \bar{R}^2 είναι λιγότερο αποτελεσματικός εκτιμητής του σ^2 , σε σχέση με το \bar{u} . Όμως, γνωρίζουμε ότι η απλή διακύμανση s^2 , όπου

$$s^2 = \frac{\sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \bar{\bar{X}})^2}{N - 1},$$

είναι αμερόληπτος εκτιμητής ελάχιστης διακύμανσης του σ^2 και είναι ακόμα πιο αποτελεσματικός του \bar{u} . Πιο συγκεκριμένα, εάν το m δεν είναι αρκετά μικρό, το s^2 έχει σχεδόν 25% μικρότερη διακύμανση από το \bar{u} για $n = 5$. Στην συνέχεια θα μελετήσουμε διαγράμματα ελέγχου για το μέσο βασισμένα στις ποσότητες \bar{X} και s .

Δοθέντων των m , n και a , και κάνοντας χρήση του A_5^* ως το κατάλληλο πολλαπλάσιο για το s , προκύπτουν τα ακόλουθα όρια ελέγχου για το διάγραμμα ελέγχου του μέσου

$$LCL = \bar{\bar{X}} - A_5^* s,$$

$$UCL = \bar{\bar{X}} + A_5^* s.$$

Θα πρέπει λοιπόν να βρούμε την τιμή του A_5^* έτσι ώστε

$$P\{\bar{\bar{X}} - A_5^* s \leq \bar{X}_f \leq \bar{\bar{X}} + A_5^* s\} = P\left\{-A_5^* \leq \frac{\bar{X}_f - \bar{\bar{X}}}{s} \leq A_5^*\right\} = 1 - a$$

όταν η επιλογή του \bar{X}_f έγινε ενώ η διεργασία βρισκόταν εντός ελέγχου.

Η διαδικασία είναι η ίδια με αυτή που περιγράφηκε στην προηγούμενη παράγραφο. Παρατηρούμε ότι η ποσότητα

$$\frac{(\bar{X}_f - \bar{\bar{X}})/\sigma \left(\frac{1}{n} + \frac{1}{N}\right)^{1/2}}{[(N-1)s^2/\sigma^2]^{1/2}/(N-1)^{1/2}} = \left[\frac{N}{m+1}\right]^{1/2} \left[\frac{\bar{X}_f - \bar{\bar{X}}}{s}\right]^{1/2}$$

ακολουθεί την κατανομή t με $(N-1)$ βαθμούς ελευθερίας. Έτσι καταλήγουμε λοιπόν στη σχέση

$$A_5^* = \left[\frac{m+1}{N}\right]^{1/2} t_{a/2, N-1}.$$

Η χρήση του s έχει κάποια μειονεκτήματα. Αρχικά, ένα πολύ σημαντικό πρόβλημα είναι αυτό της αποτελεσματικότητας του s^2 , η οποία είναι ιδιαίτερα ευαίσθητη στην υπόθεση ότι η διεργασία ήταν εντός στατιστικού ελέγχου κατά τη συλλογή των δεδομένων. Αν ο μέσος της διεργασίας αλλάζει σε κάποια δείγματα και αυτό δεν εντοπιστεί, τότε το s^2 θα υπερεκτιμήσει το σ^2 ενώ με το \bar{u} θα συνέχιζε να αποτελεί αμερόληπτο εκτιμητή. Επιπλέον, με τη χρήση του

συμπάχει αυξημένος υπολογιστικός κόπος στην περίπτωση που θα χρειαστεί να υπολογίσουμε ξανά και ξανά τα όρια ελέγχου.

Υπάρχουν αρκετοί λόγοι που δείχνουν ότι η χρήση του s είναι ιδιαίτερα αποτελεσματική. Ας υποθέσουμε ότι ένας εύλογος αριθμός ελέγχων (αναδρομικών ελέγχων ή και μελλοντικών ελέγχων) υποδεικνύει ότι η διεργασία βρίσκεται μόνιμα εντός στατιστικού ελέγχου και δεν υπάρχει λόγος να υποπτευθούμε ότι το αποτέλεσμα αυτό είναι λανθασμένο. Υπό αυτή τη συνθήκη, τα μειονεκτήματα που παρουσιάσαμε παραπάνω δεν υπάρχουν πλέον και είναι λογικό να χρησιμοποιήσουμε το s και για μελλοντικούς ελέγχους. Τέλος, πολύ σημαντικό είναι και το γεγονός ότι με τη χρήση του s πετυχαίνουμε πιο ισχυρούς ελέγχους (μικρότερη πιθανότητα σφάλματος τύπου II).

Πίνακας 3.13. Τιμές του A_5^* για $n = 5$, (Yang, Hillier (1970))

m	α			
	0.001	0.002	0.01	0.05
1	5.44	4.54	2.91	1.76
2	2.62	2.35	1.78	1.24
3	2.14	1.96	1.54	1.11
4	1.94	1.79	1.43	1.05
5	1.83	1.70	1.37	1.01
6	1.77	1.64	1.33	0.988
7	1.72	1.61	1.31	0.973
8	1.69	1.57	1.28	0.960
9	1.66	1.55	1.27	0.950
10	1.64	1.53	1.26	0.943
15	1.58	1.48	1.22	0.920
20	1.55	1.45	1.20	0.909
25	1.54	1.44	1.19	0.902
50	1.50	1.41	1.17	0.890
100	1.49	1.40	1.16	0.883
∞	1.47	1.38	1.15	0.877

3.8 Διαγράμματα ελέγχου για τη διασπορά βασισμένα στην ποσότητα s

Σε συνέχεια των παραπάνω θα εστιάσουμε στα διαγράμματα διασποράς. Αφού δοθούν τα m, n, a_L (πιθανότητα να βρεθεί ένα σημείο κάτω από το κάτω όριο ελέγχου) a_U (πιθανότητα να βρεθεί ένα σημείο πάνω από το πάνω όριο ελέγχου) και κάνοντας χρήση των κατάλληλων παραγόντων B_9^* και B_{10}^* , έχουμε

$$LCL = B_9^* s^2$$

$$UCL = B_{10}^* s^2$$

τα οποία αποτελούν τα επιθυμητά όρια για τον έλεγχο μελλοντικών δειγμάτων. Έτσι,

$$P\{s_f^2 < B_9^* s^2\} = a_L$$

$$P\{s_f^2 > B_{10}^* s^2\} = a_U$$

όταν τα s^2 και s_f^2 επιλέχθηκαν όταν η διεργασία ήταν εντός στατιστικού ελέγχου.

Ακολουθώντας παρόμοια διαδικασία με αυτή που δείξαμε νωρίτερα, παρατηρούμε ότι η στατιστική s_f^2/s^2 ακολουθεί την κατανομή F με $n-1$ και $N-1$ βαθμούς ελευθερίας, Προκύπτει λοιπόν, ότι

$$B_9^* = F_{1-a_L, n-1, N-1},$$

$$B_{10}^* = F_{a_U, n-1, N-1}.$$

Εάν επιθυμούμε την δημιουργία διαγραμμάτων για την τυπική απόκλιση s_f για τα μελλοντικά δείγματα, τα όρια πιθανοτήτων για την τυπική απόκλιση όπως είναι προφανές, προέρχονται από τη σχέση

$$LCL = \sqrt{B_9^*} s$$

$$UCL = \sqrt{B_{10}^*} s.$$

Οι τιμές των B_9^* (ή $\sqrt{B_9^*}$) και B_{10}^* (ή $\sqrt{B_{10}^*}$) παρουσιάζονται στους ακόλουθους πίνακες για μέγεθος δείγματος ίσο με 5

Πίνακας 3.14. Τιμές των B_9^* (ή $\sqrt{B_9^*}$) για $n = 5$, (Yang, Hillier (1970))

α_L	0.001		0.005		0.025	
	B_9^*	$\sqrt{B_9^*}$	B_9^*	$\sqrt{B_9^*}$	B_9^*	$\sqrt{B_9^*}$
1	0.0187	0.137	0.0431	0.208	0.104	0.322
2	0.0206	0.144	0.0474	0.218	0.112	0.335
3	0.0213	0.146	0.0488	0.221	0.115	0.339
4	0.0216	0.147	0.0495	0.222	0.117	0.342
5	0.0218	0.148	0.0500	0.224	0.118	0.344
6	0.0220	0.148	0.0503	0.224	0.118	0.344
7	0.0221	0.149	0.0505	0.225	0.118	0.344
8	0.0222	0.149	0.0505	0.225	0.119	0.345
9	0.0222	0.149	0.0508	0.225	0.119	0.345
10	0.0223	0.149	0.0508	0.225	0.119	0.345
15	0.0224	0.150	0.0513	0.226	0.120	0.346
20	0.0225	0.150	0.0513	0.226	0.120	0.346
25	0.0225	0.150	0.0515	0.227	0.120	0.346
50	0.0226	0.150	0.0515	0.227	0.121	0.348
100	0.0227	0.151	0.0515	0.227	0.121	0.348
∞	0.0227	0.151	0.0518	0.228	0.121	0.348

Πίνακας 3.15. Τιμές των B_{10}^* (ή $\sqrt{B_{10}^*}$) για $n = 5$, (Yang, Hillier (1970))

α_U	0.001		0.005		0.025	
	B_{10}^*	$\sqrt{B_{10}^*}$	B_{10}^*	$\sqrt{B_{10}^*}$	B_{10}^*	$\sqrt{B_{10}^*}$
1	53.4	7.31	23.2	4.82	9.60	3.10
2	12.6	3.55	7.96	2.82	4.72	2.17
3	8.62	2.94	6.00	2.45	3.89	1.97
4	7.26	2.69	5.27	2.30	3.56	1.89
5	6.59	2.57	4.89	2.21	3.38	1.84
6	6.18	2.49	4.66	2.16	3.27	1.81
7	5.92	2.43	4.50	2.12	3.19	1.79
8	5.73	2.39	4.39	2.10	3.14	1.77
9	5.59	2.36	4.31	2.08	3.09	1.76
10	5.47	2.34	4.24	2.06	3.06	1.75
15	5.17	2.27	4.06	2.01	2.97	1.72
20	5.01	2.24	3.96	1.99	2.92	1.71
25	4.93	2.22	3.91	1.98	2.89	1.70
50	4.77	2.18	3.82	1.95	2.82	1.68
100	4.69	2.17	3.76	1.94	2.81	1.68
∞	4.62	2.15	3.72	1.93	2.79	1.67

ΒΙΒΛΙΟΓΡΑΦΙΑ

Abdel-Salam, A. G.; Birch, J. B.; and Jensen, W. A. (2013). "A Semiparametric Mixed Model Approach to Phase I Profile Monitoring". *Quality and Reliability Engineering International* 29, pp. 555–569.

Albers, W. and Kallenberg, W. C. (2004). "Are Estimated Control Charts in Control?" *Statistics* 38(1), pp. 67–79.

Alfaro, J. L. and Ortega, J. F. (2008). "A Robust Alternative to Hotelling's T² Control Chart Using Trimmed Estimators". *Quality and Reliability Engineering International* 24, pp. 601–611.

Anderson-Cook, C. M., and Borror, C. M. (2013). "Paving the Way: Seven Data Collection Strategies to Enhance your Quality Analysis". *Quality Progress* April, pp. 18–19.

Bayes, T., "An Essay Toward Solving a Problem in the Doctrine of Chances", *Philosophical Transactions of the Royal Society*, Vol. 53, 1763, pp. 370-418.

Bell, R. C.; Jones-Farmer, L. A.; and Billor, N. (2014). "A Distribution-Free Multivariate Phase I Location Control Chart for Subgrouped Data from Elliptical Distribution". *Technometrics*, to appear.

Beyer, W. H. (Ed.), (1966), *Handbook of Tables for Probability and Statistics*, The Chemical Rubber Publishing Co.

Borror, C. M., Champ, C. W. (2001). *Phase I control charts for independent Bernoulli data*. *Quality and Reliability Engineering International*, 17:391–396.

Box, G. and Luceño, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. New York, NY: John Wiley & Sons, Inc.

Boyles, R. A. (2000). *Phase I analysis for autocorrelated processes*. *Journal of Quality Technology*, 32:395–409.

- Capizzi, G. and Masarotto, G. (2010). "Self-Starting CUSCORE Charts for Individual Multivariate Observations". *Journal of Quality Technology* 42(2), pp. 136–151.
- Capizzi, G. and Masarotto, G. (2013). "Phase I Distribution- Free Analysis of Univariate Data". *Journal of Quality Technology* 45, pp. 273–284.
- Chakraborti, S.; Human, S. W.; and Graham, M. A. (2009). "Phase I Statistical Process Control Charts: An Overview and Some Results". *Quality Engineering* 21, pp. 52–62.
- Champ, C. W., Chou, S. P. (2003). *Comparison of standard and individual limits Phase I Shewhart X, R and S charts*. *Quality and Reliability Engineering International*, 19:161–170.
- Champ, C. W., Jones, L. A. (2004). *Designing phase I X charts with small sample sizes*. *Quality and Reliability Engineering International*, 20:497–510.
- Chen, J. and Gupta, A. K. (2011). *Parametric Statistical Change Point Analysis: With Applications in Genetics, Medicine, and Finance, 2nd edition*. New York, NY: Birkh user.
- Chen, Y.; Birch, J. B.; and Woodall, W. H. (2014), "Cluster-Based Profile Monitoring in Phase I Analysis". *Journal of Quality Technology*, to appear.
- Chou, S. P., Champ, C. W. (1995). *A comparison of two phase I control charts*. *Proceedings of the Quality and Productivity Section of the American Statistical Association, Orlando, FL*, 31–35
- Coleman, D. E. (1997). "Individual Contributions to 'A Discussion on Statistically-Based Process Monitoring and Control'". *Journal of Quality Technology* 29, pp. 148–156.
- David, H. A.(1962), "Special Problems in Testing Hypotheses," Ch. 7 in Sarhan and Greenberg (Eds.), *Contributions to Order Statistics*.
- Del Castillo, E.; Grayson, J. M.; Montgomery, D. C.; and Runger, G. C. (1996). "A Review of Statistical Process Control Techniques for Short Run Manufacturing Systems". *Communications in Statistics—Theory and Methods* 25(11), pp. 2723–2737.

Ding, Y.; Zeng, L.; and Zhou, S. (2006). "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing". *Journal of Quality Technology* 38(3), pp. 199–216.

Doganaksoy, N. and Hahn, G. J. (2012). "Getting the Right Data Up Front: A Key Challenge." *Quality Engineering* 24, pp. 446–459.

Dovoedo, Y. H. and Chakraborti, S. (2012). "Boxplot- Based Phase I Control Charts for Time Between Events". *Quality and Reliability Engineering International* 28, pp. 123–130.

Duncan, Acheson J. (1958), "Design and Operation of a Double-Limit Variables Sampling Plan," *Journal of the American Statistical Association*, Vol. 53, No. 282, pp. 543-550.

Duncan, Acheson J.(1965), *Quality Control and Industrial Statistics*, 3,d ed., Irwin, Homewood.

Graham, M. A.; Human, S. W.; and Chakraborti, S. (2010). "A Phase I Nonparametric Shewhart-Type Control Chart Based on the Median". *Journal of Applied Statistics* 37(11), pp. 1795–1813.

Grant, Eugene L.(1965), *Statistical Quality Control*, 3,d ed., McGraw-Hill, New York.

Grubbs, Frank E., and Weaver, Chalmers L. (1947), "The Best Unbiased Estimate of Population Standard Deviation Based on Group Ranges," *Journal of the American Statistical Association*, Vol. 42, No. 238, pp.224-241.

Hard, A.(1952), *Statistical Tables and Formulas*, Wiley, New York.

Harter, H. Leon, and Clemm, Donald S. (1959), "The Probability Integrals of the Range and of the Studentized Range-Probability Integral, Percentage Points, and Moments of the Range," *Wright Air Development Center Technical Report 58-484*. Vol. I (ASTIA Document No. AD 215024).

Harter, H. Leon, Clemm, Donald S., and Guthrie, Eugene H. (1959), "The Probability Integrals of the Range and the Studentized Range-Probability Integral and Percentage Points of the Studentized Range; Critical Values for Duncan's New Multiple Range Test," *Wright Air Development Center Technical Report 58- 484*, Vol. II.

Harter, H. LEON, (1960) "Tables of Range and Studentized Range," *Annals of Mathematical Statistics*, Vol. 31,pp. 1122-1147.

Hawkins, D. M. (1977). *Testing a sequence of observations for a shift in location*. *Journal of the American Statistical Association*, 72:180–186.

- Hawkins, D. M. (1987). "Self-Starting CUSUM Charts for Location and Scale". *The Statistician* 36, pp. 299–316.
- Hawkins, D. M., Qiu, P., Kang, C. W. (2003). *The change point model for statistical process control*. *Journal of Quality Technology*, 35:355–366.
- Hawkins, D. M., Zamba, K. D. (2005). *Statistical process control for shift in mean or variance using a change-point formulation*. *Technometrics*, 47:164–173.
- Hawkins, D. M. and Maboudou-Tchao, E. M. (2007). "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting". *Technometrics* 49(2), pp. 199–209.
- Hawkins, D. M. and Deng, Q. (2010). "A Nonparametric Change-Point Control Chart". *Journal of Quality Technology* 42(2), pp. 165–173.
- Hillier, Frederick S. (1967), "Small Sample Probability Limits for the Range Chart," *Journal of the American Statistical Association*, Vol. 62, No. 320, pp. 1488-1493.
- Hillier, Frederick S., and Lieberman, Gerald J. (1967), *Introduction to Operations Research*, Holden-Day, San Francisco
- Hillier, F. S. (1969). *X and R chart control limits based on a small number of subgroups*. *Journal of Quality Technology*, 1:17–26.
- Hinkley, D. V. (1970). "Inference About the Change-Point in a Sequence of Random Variables". *Biometrika* 57(1), pp. 1–17.
- Hinkley, D. V. (1971). "Inference About the Change-Point from Cumulative Sum Tests". *Biometrika* 58(3), pp. 509– 523
- Human, S. W., Chakraborti, S., Smit, C. F. (2008). *S², S and R control charts for phase I applications*. *Submitted*.

Human, S. W.; Chakraborti, S.; and Smit, C. F. (2010). "Shewhart-Type Control Charts for Variation in Phase I Data Analysis". *Computational Statistics and Data Analysis* 54(4), pp. 863–874.

Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., Woodall, W. H. (2006). *Effects of parameter estimation on control chart properties: A literature review. Journal of Quality Technology*, 38:349–364.

Jensen, W. A.; Birch, J. B.; and Woodall, W. H. (2007). "High Breakdown Estimation for Phase I Multivariate Control Charts". *Quality and Reliability Engineering International* 23, pp. 615–629.

Jobe, J. M. and Pokojovy, M. A. (2009). "Multistep, Cluster-Based Multivariate Chart for Retrospective Monitoring of Individuals". *Journal of Quality Technology* 41(4), pp. 323–339.

Jones, L. A., Champ, C. W. (2002). *Phase I control charts for times between events. Quality and Reliability Engineering International*, 18:479–488.

Jones-Farmer, L. A.; Jordan, V.; and Champ, C. W. (2009). "Distribution-Free Phase I Control Charts for Subgroup Location". *Journal of Quality Technology* 41, pp. 304–316.

Jones-Farmer, L. A. and Champ, C. W. (2010). "A Distribution-free Phase I Control Charts for Subgroup Scale". *Journal of Quality Technology* 42, pp. 373–387.

Jones-Farmer, L. A.; Ezell, J. D.; and Hazen, B. T. (2014). "Applying Control Chart Methods to Enhance Data Quality". *Technometrics* 56(1), pp. 29–41.

Jones-Farmer, L.A ;Woodall W.H.; Steiner S.H. and Champ C.W. (2014) "An Overview of Phase One for Process Improvement and Monitoring"

King, E. P. (1952), "The Operating Characteristic of the Control Chart for Sample Means," *Annals of Mathematical Statistics*, Vol. 23 pp. 384-395.

- King, E. P. (1954). *Probability limits for the average chart when process standards are unspecified*. *Industrial Quality Control*, 10:62–64.
- Koning, A. J. (2006). *Model-based control charts in phase I statistical process control*. *Statistica Neerlandica*, 60:327–338.
- Kruskal, W. H. and Wallis, W. A. (1952). “*Use of Ranks in One-Criterion Variance Analysis*”. *Journal of the American Statistical Association* 47, pp. 583–621.
- Li, Z.; Zhang, J.; and Wang, Z. (2010). “*Self-Starting Control Chart for Simultaneously Monitoring Process Mean and Variance*”. *International Journal of Production Research* 48(15), pp. 4537–4553.
- Liao, T. W. (2005). “*Clustering of Time Series Data—A Survey*”. *Pattern Recognition* 38(11), pp. 1857–1874.
- Liu, R. Y. (1995). “*Control Charts for Multivariate Processes*”, *Journal of the American Statistical Association* 90(432), pp. 1380–1387.
- Maboudou-Tchao, E. M. and Hawkins, D. M. (2011). “*Self-Starting Multivariate Control Charts for Location and Scale*”. *Journal of Quality Technology* 43(2), pp. 113–126.
- Mahmoud, M. A. and Woodall, W. H. (2004). “*Phase I Analysis of Linear Profiles with Calibration Applications*”. *Technometrics* 46(4), pp. 377–391.
- Mahmoud, M. A.; Parker, P. A.; Woodall, W. H.; and Hawkins, D. M. (2007). “*A Change Point Method for Linear Profile Data*”. *Quality and Reliability Engineering International* 23(2), pp. 247–268.
- Mahmoud, M. A.; Henderson, G. R.; Epprecht, E. K.; and Woodall, W. H. (2010). “*Estimating the Standard Deviation in Quality Control Applications*”. *Journal of Quality Technology* 42(4), pp. 348–357.

Maragah, H. D., Woodall, W. H. (1992). *The effect of autocorrelation on the retrospective X chart*, *Journal of Statistical Computation and Simulation*, 40:29–42.

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

Nedumaran, G. and Pignatiello, Jr., J. J. (2000). “On Constructing T2 Control Charts for Retrospective Examination”. *Communications in Statistics—Simulation and Computation* 29(2), pp. 621–632.

Nedumaran, G., Pignatiello, J. J. (2005). *On constructing retrospective X control chart limits*. *Quality and Reliability Engineering International*, 21:81–89.

Newton, P. B. and Champ, C. W. (1997). “Probability Limits for Phase I X-Bar Charts”. In *Proceedings of the Southeast Decision Sciences Institute Annual Conference*, pp. 234-236. Atlanta, GA.

Noorossana, R.; Saghaei, A.; and Amiri, A., (eds.) (2011). *Statistical Analysis of Profile Monitoring*. Hoboken, NJ: John Wiley & Sons, Inc.

Nelson, P. R. (1982). *Multivariate normal and t distributions with qjk O ajak*: *Communications in Statistics: Simulation and Computation*, 11:239–248.

Nelson, L. S. (1988). “Control Charts: Rational Subgroups and Effective Applications”. *Journal of Quality Technology* 20, pp. 73–75.

Quesenberry, C. P. (1991). *SPC Q charts for start-up processes and short or long runs*. *Journal of Quality Technology*, 23(3):213–246.

Owen, D. B. (1962), *Handbook of Statistical Tables*, AddisonWesley, Reading, Mass.

Oyeyemi, G. M. and Ipinyomi, R. A. (2010). “A Robust Method of Estimating Covariance Matrix in Multivariate Data Analysis”. *African Journal of Mathematics and Computer Science Research* 3(1), pp. 1–18.

- Palm, A. (1992). "Some Aspects of Sampling for Control Charts". *Statistics Division Newsletter Summer*. Available at <http://rube.asq.org/statistics/1992/06/asq-statistics-division-newsletter-v12-i04-full-issue.pdf>.
- Pan, R. and Rigdon, S. E. (2012). "A Bayesian Approach to Change Point Estimation in Multivariate SPC". *Journal of Quality Technology* 44(3), pp. 231–248.
- Patnaik, P. B. (1950), "The Use of Mean Range as an Estimator of Variance in Statistical Tests," *Biometrika*, Vol. 37, pp. 78-87.
- Paynabar, K.; Jin, J. H.; and Yeh, A. B. (2012). "Phase I Risk-Adjusted Control Charts for Monitoring Surgical Performance by Considering Categorical Covariates". *Journal of Quality Technology* 44(1), pp. 39–53.
- Perla, R. J.; Provost, L. P.; and Murray, S. K. (2013). "Sampling Considerations for Health Care Improvement". *Quality Management in Healthcare* 22(1), pp. 36–47.
- Proschan, F., and Savage, I. R. (1960), "Starting a Control Chart," *Industrial Quality Control*, Vol. 17, No.3, Sept.)pp. 12-13
- Quesenberry, C. P. (1991). "SPC \bar{X} Charts for Start-Up Processes and Short or Long Runs". *Journal of Quality Technology* 23, pp. 213–224.
- Resnikoff, George J.(1954), "The Distribution of the Average-Range for Subgroups of Five," *Technical Report No. 15, Contract N60nr-25126, Applied Mathematics and Statistics Laboratories, Stanford University*
- Ryan, T. P. (1989). *Statistical Methods for Quality Improvement*. New York: John Wiley.
- Saleh, N. A.; Mahmoud, M. A.; Keefe, M. J.; and Woodall, W. H. (2014). "The Difficulty in Designing \bar{X} -Bar and \bar{X} -Control Charts with Estimated Parameters". *Journal of Quality Technology*, to appear.
- Sceffe, Henry (1949), "Operating Characteristics of Average and Range Charts," *Industrial Quality Control*, Vol. 5, No.6, pp. 13-18.
- Schoonhoven, M. and Does, R. J. M. M. (2012). "A Robust Standard Deviation Control Chart". *Technometrics* 54(1), pp. 73–82.

Schoonhoven, M.; Nazir, H. Z.; Riaz, M.; and Does, R. J. M. M. (2011). “Robust Location Estimators for the X-bar Control Chart”. *Journal of Quality Technology* 43(4)

Schoonhoven, M. and Does, R. J. M. M. (2013). “A Robust X-Bar Control Chart”. *Quality and Reliability Engineering International*, to appear.

Seder, L. (1950a). “Diagnosis with Diagrams. Part I”. *Industrial Quality Control* 7(1), pp. 11–19.

Seder, L. (1950b). “Diagnosis with Diagrams. Part II”. *Industrial Quality Control* 7(2), pp. 7–10.

Shainin, R. D. (2008). “Multi-Vari Charts”. In *Encyclopedia of Statistics in Quality and Reliability*, Ruggeri, F.; Kenett, R. S.; and Faltin F. W., eds. Hoboken, NJ: John Wiley & Sons, Inc. (<http://dx.doi.org/10.1002/9780470061572.eqr405>)

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. New York, NY: D. Van Nostrand Company, Inc.

Shewhart, W. A. (1939). *Statistical Method from the Viewpoint of Quality Control*. New York, NY: Dover Publications.

Shiau, J.-J. H. and Sun J.-H. (2009). “A New Strategy for Phase I Analysis in SPC”. *Quality and Reliability Engineering International* 26, pp. 475–486.

Steiner, S. H. and MacKay, R. J. (2005). *Statistical Engineering: An Algorithm for Reducing Variation in Manufacturing Processes*. Milwaukee, WI: ASQ Quality Press.

Stoumbos, Z. G., and Jones, L. A. (2000). “On the Properties and Design of Individual Control Charts Based on Simplicial Depth”. *Nonlinear Studies* 7(2), pp. 147–178.

Sullivan, J. H., Woodall, W. H. (1996). A control chart for preliminary analysis of individual observations. *Journal of Quality Technology*, 28(3):265–278.

Sullivan, J. H. and Woodall, W. H. (2000). "Change-Point Detection of Mean Vector or Covariance Matrix Shifts Using Multivariate Individual Observations". *IIE Transactions—Quality and Reliability Engineering* 32(6), pp. 537–549.

Sullivan, J. H. (2002). "Estimating the Locations of Multiple Change Points in the Mean". *Computational Statistics* 17, pp. 289–296.

Sullivan, J. H. and Jones, L. A. (2002). "A Self-Starting Control Chart for Multivariate Individual Observations". *Technometrics* 44(1), pp. 24–33.

Szarka, III, J. L. and Woodall, W. H. (2011). "A Review and Perspective on Surveillance of Bernoulli Processes". *Quality and Reliability Engineering International* 27, pp. 735–752.

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

Tukey, J. W. (1975). "Mathematics and Picturing Data". In *Proceedings of the 1974 International Congress of Mathematicians*, James, R., ed., vol. 2, pp. 523–531. Vancouver, BC.

Vargas, J. A. (2003). "Robust Estimation in Multivariate Control Charts for Individual Observations". *Journal of Quality Technology* 35(4), pp. 367–376.

Vining, G. (2013). "Technical Advice: Scientific Method and Approaches for Collecting Data". *Quality Engineering* 25, pp. 194–201.

Wheeler, D. J. (2011). "What About p-Charts?". *Quality Digest*. Available on-line at <http://www.qualitydigest.com/inside/quality-insider-article/what-about-p-charts.html>, accessed on 5/18/2013.

Wheeler, D. J. and Chambers, D. S. (1992). *Understanding Statistical Process Control*, 2nd edition. Knoxville, TN: SPC Press, Inc.

Williams, J. D.; Woodall, W. H.; and Birch, J. B. (2007). "Statistical Monitoring of Nonlinear Product and Process Quality Profiles". *Quality and Reliability Engineering International* 23(7), pp. 925–941.

Woodall, W. H. (2000). *Controversies and contradictions in statistical process control*. *Journal of Quality Technology*, 32:341–350.

Woodall, W. H. and Montgomery, D. C. (2014). "Some Current Directions in the Theory and Application of Statistical Process Monitoring". *Journal of Quality Technology* 46(1), pp. 78–94.

Woodall, W. H.; Spitzner, D. J.; Montgomery, D. C.; and Gupta, S. (2004). "Using Control Charts to Monitor Process and Product Quality Profiles". *Journal of Quality Technology* 36(3), pp. 309–320.

Woodall, W. H. and Thomas, E. V. (1995). "Statistical Process Control with Several Components of Common Cause Variability". *IIE Transactions* 27(6), pp. 757–764.

Yanez, S.; Gonzalez, N.; and Vargas, J. A. (2010). "Hotelling's T₂ Control Charts Based on Robust Estimators". *Dyna* 163, pp. 239–247.

Yang, C. H., Hillier, F. S. (1970). *Mean and variance control chart limits based on a small number of subgroups*. *Journal of Quality Technology*, 2:9–16.

Zhang, C. W., Xie, M., Goh, T. N. (2006). *Design of exponential control charts using a sequential sampling scheme*. *IIE Transactions*, 38:1105–1116.

Zhang, H.; Albin, S. L.; Wagner, S. R.; Nolet, D. A.; and Gupta, S. (2010). "Determining Statistical Process Control Baseline Periods in Long Historical Data Streams". *Journal of Quality Technology* 42(1), pp. 21–35.

Zhang, M.; Peng, Y.; Schuh, A.; Megahed, F.; and Woodall, W. H. (2013). "Geometric Charts with Estimated Control Limits". *Quality and Reliability Engineering International* 29, pp. 209–223.

Zhou, C.; Zou, C.; Zhang, Y.; and Wang, Z. (2009). “Nonparametric Control Chart Based on Change-Point Model”. *Statistical Papers* 50(1), pp. 13–28.

Zou, C.; Zhou, C.; Wang, Z.; and Tsung, F. (2007). “A Self- Starting Control Chart for Linear Profiles”. *Journal of Quality Technology* 39(4), pp. 364–375.

Zou, C.; Yin, G.; Feng, L.; and Wang, Z. (2014). “Nonparametric Maximum Likelihood Approach to Multiple Change- Point Problems”. *The Annals of Statistics*, to appear.

