

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΠΟΛΥΔΙΑΣΤΑΤΑ ΔΙΑΓΡΑΜΜΑΤΑ
ΕΛΕΓΧΟΥ ΔΙΕΡΓΑΣΙΩΝ**

Ελισάβετ Μ. Σοφικίτου

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
Σεπτέμβριος 2012

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΠΟΛΥΔΙΑΣΤΑΤΑ ΔΙΑΓΡΑΜΜΑΤΑ
ΕΛΕΓΧΟΥ ΔΙΕΡΓΑΣΙΩΝ**

Ελισάβετ Μ. Σοφικίτου

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
Σεπτέμβριος 2012

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:

- Καθηγητής Κούτρας Μάρκος (Επιβλέπων)
- Αναπληρωτής Καθηγητής Αντζουλάκος Δημήτριος
- Επίκουρος Καθηγητής Ευαγγελάρας Χαράλαμπος

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



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**MULTIVARIATE
CONTROL CHARTS**

By

Elisavet M. Sofikitou

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
September 2012

РАНЕЕЗНАМО ТЕРПАА

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΔΙΑ

Στην οικογένειά μου

РАНЕЕЗНАМО ПЕРПАА

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Μ. Κούτρα, αρχικά γιατί δέχθηκε να αναλάβει την επίβλεψη της παρούσας εργασίας, κυρίως, όμως, διότι η βοήθειά του υπήρξε καθοριστική για εμένα σε όλα τα στάδια της εκπόνησής της: επιλογή θέματος, συλλογή και επεξεργασία στοιχείων, μελέτη του αντικειμένου και εξαγωγή συμπερασμάτων.

Ευχαριστώ, επίσης, τα μέλη της τριμελούς επιτροπής, τον Αναπληρωτή Καθηγητή κ. Δ. Αντζουλάκο και τον Επίκουρο Καθηγητή κ. Χ. Ευαγγελάρα, για τον πολύτιμο χρόνο που αφιέρωσαν στη μελέτη και διόρθωση της εργασίας αυτής. Θα ήταν παράλειψη να μην ευχαριστήσω τους καθηγητές που είχα σε όλη τη διάρκεια των προπτυχιακών και μεταπτυχιακών μου σπουδών, καθώς η παρούσα εργασία αποτέλεσε το επιστέγασμα όλων των γνώσεων που μου προσέφεραν.

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου και την αδελφή μου, η συμπαράσταση των οποίων υπήρξε πολύτιμη για μένα καθ' όλη τη διάρκεια των σπουδών μου (προπτυχιακών και μεταπτυχιακών) κυρίως, όμως, κατά την εκπόνηση της εργασίας αυτής.

РАНЕЕЗНАМО ТЕРПАА

Περίληψη

Η έννοια του Στατιστικού Ελέγχου Ποιότητας είναι πλέον συνυφασμένη με τα διαγράμματα ελέγχου διεργασιών. Στην πραγματικότητα, τα διαγράμματα αυτά αποτελούν το βασικό στατιστικό εργαλείο, το οποίο χρησιμοποιείται αρχικά για τον εντοπισμό, και στη συνέχεια για τη βελτίωση της ποιότητας ενός προϊόντος και κατ' επέκταση όλης της παραγωγικής διαδικασίας. Σκοπός είναι η έγκαιρη ανίχνευση της «αφύσικης» συμπεριφοράς μιας διεργασίας. Συνήθως, αυτό επιτυγχάνεται με τη χρήση ατομικών-μονοδιάστατων διαγραμμάτων ελέγχου (*univariate control charts*), καθένα από τα οποία αντιστοιχεί σε ένα μόνο από τα προς μελέτη χαρακτηριστικά που εξετάζουμε.

Ωστόσο, η ποιότητα ενός προϊόντος καθορίζεται όχι μόνο από την τιμή ενός ποιοτικού χαρακτηριστικού, αλλά και από περισσότερα χαρακτηριστικά. Η πιθανή εξάρτηση που αναπτύσσεται μεταξύ των χαρακτηριστικών αυτών καθιστά αναγκαία τη χρήση πολυδιάστατων διαγραμμάτων ελέγχου (*multivariate control charts*), τα οποία επιτρέπουν τον ταυτόχρονο έλεγχο όλων των εξαρτημένων χαρακτηριστικών. Στην περίπτωση που ένα τέτοιο διάγραμμα υποδεικνύει ότι η διεργασία είναι εκτός στατιστικού ελέγχου, το ενδιαφέρον έγκειται στον εντοπισμό του χαρακτηριστικού ή των χαρακτηριστικών εκείνων που προκάλεσαν την έξοδο της διεργασίας από την εντός ελέγχου κατάσταση.

Στην παρούσα διπλωματική εργασία γίνεται αρχικά μια εισαγωγή στο Στατιστικό Έλεγχο Ποιότητας, κατόπιν παρουσιάζεται μια ιστορική αναδρομή και περιγράφεται ένα τυπικό διάγραμμα ελέγχου με τις κατηγοριοποιήσεις του. Στη συνέχεια παρουσιάζονται λεπτομερώς τα κυριότερα πολυδιάστατα διαγράμματα ελέγχου, τόσο τα παραμετρικά όσο και τα μη παραμετρικά. Επιπλέον, παρουσιάζονται εν συντομία μόνο εκείνα τα μονοδιάστατα διαγράμματα, των οποίων γενίκευση αποτελούν τα πολυδιάστατα διαγράμματα που αναπτύσσονται στη συνέχεια. Συγχρόνως, αναφέρονται μέθοδοι εντοπισμού του χαρακτηριστικού ή των χαρακτηριστικών που έθεσαν τη διεργασία εκτός ελέγχου. Τέλος, συγκρίνεται η αποτελεσματικότητα μεταξύ διάφορων μονοδιάστατων ή/και πολυδιάστατων διαγραμμάτων ελέγχου που έχουν εμφανιστεί πρόσφατα στη βιβλιογραφία.

РАНЕЕЗНАМО ТЕРПАА

Abstract

Nowadays, the concept of Statistical Quality Control is strongly associated with the process control charts. In fact, these charts are the main statistical tool that is used to identify and improve the quality of a single product and the quality of the entire production process as well. A control chart is used for the early detection of "abnormal" behaviour of a process. Typically, this is achieved by launching separate-univariate control charts, each one monitoring one of the characteristics to be investigated.

However, it is quite common the quality of a product to be determined by the value of more than one quality characteristics. The possible dependence which is usually present among these characteristics, makes it imperative to use multivariate control charts, which allow the simultaneous control of all dependent characteristics of interest. If such a diagram indicates that the process is out of control, one may also be interested to detect which characteristic or characteristics that caused the out of control shift of the process.

The present dissertation initially provides an introduction to Statistical Quality Control and its history, a description of a typical control chart and its categories. Next, the main multivariate control charts, both parametric and non-parametric, are presented analytically. These charts are a straightforward generalization of the univariate ones that are presented in brief as well.

In addition, we illustrate several methods that are capable of identifying the characteristic or characteristics that have caused an out-of-control signal. Finally, we present a detailed comparison of the several univariate and/or multivariate control charts, which have recently appeared in the literature.

РАНЕЕЗНАМО ТЕРПАА

Περιεχόμενα

Περίληψη	σελ. ix
Abstract	σελ. xi
Περιεχόμενα	σελ. xiii
Κατάλογος Πινάκων	σελ. xvii
Κατάλογος Σχημάτων	σελ. xix
Κατάλογος Συντομογραφιών	σελ. xxiii

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Στατιστικός Έλεγχος Ποιότητας	σελ. 1
1.2 Ιστορική Αναδρομή	σελ. 4
1.3 Φιλοσοφία της Ολικής Διαχείρισης Ποιότητας	σελ. 11
1.4 Διαγράμματα Ελέγχου και Μεταβλητότητα	σελ. 13
1.5 Περιγραφή ενός Τυπικού Διαγράμματος Ελέγχου	σελ. 14
1.6 Είδη Διαγραμμάτων Ελέγχου	σελ. 22
1.7 Βασικές Έννοιες σχετικά με τα Διαγράμματα Ελέγχου	σελ. 24
1.7.1 Μοντέλο Ορίων L Σίγμα και Μοντέλο Ορίων Πιθανότητας a	σελ. 24
1.7.2 Σφάλμα Τύπου I και Σφάλμα Τύπου II	σελ. 25
1.7.3 Μέσο Μήκος Ροής και Μέσος Χρόνος Σήματος	σελ. 27
1.7.4 Μέγεθος Δείγματος, Συχνότητα Δειγματοληψίας και Ορθολογικές Υποομάδες	σελ. 28
1.7.5 Φάση I και Φάση II	σελ. 29
1.7.6 Δομή Διπλωματικής Εργασίας	σελ. 30

ΚΕΦΑΛΑΙΟ 2: Μονομεταβλητά Παραμετρικά Διαγράμματα Ελέγχου

2.1 Εισαγωγή	σελ. 33
2.2 Διαγράμματα Ελέγχου Τύπου Shewhart για Μεταβλητές χωρίς μνήμη	σελ. 33
2.2.1 Διαγράμματα Ελέγχου για Δείγματα	σελ. 34
2.2.2 Παράδειγμα Διαγραμμάτων Ελέγχου για Δείγματα	σελ. 41
2.3 Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες χωρίς μνήμη	σελ. 49

2.3.1 Διαγράμματα Ελέγχου για τον Αριθμό των Ελαττωματικών Προϊόντων.....	σελ. 49
2.3.2 Παράδειγμα Διαγραμμάτων Ελέγχου για Ιδιότητες.....	σελ. 50

ΚΕΦΑΛΑΙΟ 3: Πολυμεταβλητά Παραμετρικά Διαγράμματα Ελέγχου

3.1 Εισαγωγή.....	σελ. 53
3.2 Πολυμεταβλητά Διαγράμματα Ελέγχου Τύπου Shewhart για Μεταβλητές χωρίς μνήμη.....	σελ. 55
3.2.1 Εκτίμηση της Μέσης Τιμής μ και του Πίνακα Διακύμανσης-Συνδιακύμανσης Σ	σελ. 56
3.2.2 Πολυμεταβλητά Διαγράμματα Ελέγχου για τη Μέση Τιμή.....	σελ. 57
3.2.3 Πολυμεταβλητά Διαγράμματα Ελέγχου για τη Διασπορά.....	σελ. 67
3.3 Πολυμεταβλητά Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες χωρίς μνήμη.....	σελ. 75
3.4 Πολυμεταβλητά Διαγράμματα Ελέγχου με μνήμη	σελ. 87
3.4.1 Πολυμεταβλητό Διάγραμμα Ελέγχου Συσσωρευμένου Αθροίσματος	σελ. 87
3.4.2 Πολυμεταβλητό Διάγραμμα Ελέγχου για τον Εκθετικά Ζυγισμένο Μέσο	σελ. 103
3.4.3 Πολυμεταβλητό Διάγραμμα Ελέγχου Κινούμενου Μέσου.....	σελ. 113
3.5 Παραδείγματα Πολυμεταβλητών Διαγραμμάτων Ελέγχου	σελ. 117
3.5.1 Διαγράμματα Ελέγχου για Μεταβλητές και Διαγράμματα Ελέγχου με Μνήμη	σελ. 117
3.5.2 Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες	σελ. 127
3.6 Ανίχνευση της Εκτός Ελέγχου Μεταβλητής.....	σελ. 133
3.6.1 Χρήση Μονοδιάστατων Διαγραμμάτων με Τυπικά ή Bonferroni Όρια Ελέγχου	σελ. 133
3.6.2 Χρήση Ελλειπτικής Περιοχής Ελέγχου	σελ. 137
3.6.3 Χρήση της Διάσπασης της Στατιστικής Συνάρτησης T^2	σελ. 144
3.6.4 Διάγραμμα Ελέγχου Minimax, Διάγραμμα Επιλογής-Αιτίας και Προσαρμοσμένες Μεταβλητές Παλινδρόμησης.....	σελ. 156
3.6.5 Ανάλυση Κυρίων Συνιστωσών	σελ. 161
3.6.6 Γραφικές Μέθοδοι.....	σελ. 166

ΚΕΦΑΛΑΙΟ 4: Μη Παραμετρικά Διαγράμματα Ελέγχου

4.1	Εισαγωγή.....	σελ. 177
4.2	Μονομεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου.....	σελ. 180
4.2.1	Διαγράμματα Ελέγχου Τύπου Shewhart.....	σελ. 182
4.3	Πολυμεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου.....	σελ. 186
4.3.1	Διαγράμματα Ελέγχου Τύπου Shewhart.....	σελ. 186
4.3.2	Διαγράμματα Ελέγχου Συσσωρευμένου Αθροίσματος.....	σελ. 193
4.3.3	Διάγραμμα Ελέγχου Εκθετικά Κινούμενου Μέσου.....	σελ. 196
4.3.4	Διάγραμμα Ελέγχου Κινούμενου Μέσου που βασίζεται στο βάθος των δεδομένων.....	σελ. 198
4.3.5	Άλλα Πολυμεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου.....	σελ. 200
4.4	Παραδείγματα Πολυμεταβλητών Μη Παραμετρικών Διαγραμμάτων.....	σελ. 201
	Βιβλιογραφία.....	σελ. 213

РАНЕЕЗНАМО ПЕРПАА

Κατάλογος Πινάκων

2.1	Τιμές διαμέτρου εμβόλου και τιμές στατιστικών συναρτήσεων των μονοδιάστατων διαγραμμάτων R , \bar{X} , S και S^2	σελ. 42
2.2	Ελεγχθέντα μεταλλικά μέρη και αποτελέσματα ελέγχου παραδείγματος διαγραμμάτων ελέγχου για ιδιότητες	σελ. 51
3.1	Τιμές άνω ορίου ελέγχου ενός διαγράμματος $MCUSUM$ με $ARL_0 = 200$	σελ. 94
3.2	Έλεγχος κανονικότητας για κάθε χαρακτηριστικό	σελ. 118
3.3	Τιμές στατιστικών συναρτήσεων των διαγραμμάτων T^2 , $ S^* $ και $MEWMA$	σελ. 123
3.4	Τιμές στατιστικών συναρτήσεων των διαγραμμάτων $MCUSUM$	σελ. 126
3.5	Τιμές στατιστικών συναρτήσεων X και D των διαγραμμάτων MNP και MP	σελ. 131
3.6	Συνεισφορά κάθε χαρακτηριστικού στο εκτός ελέγχου σήμα του T^2	σελ. 156
4.1	FAR και ARL_0 του διαγράμματος SN για διάφορες τιμές του n και t	σελ. 184
4.2	Τιμές άνω ορίου ελέγχου του διαγράμματος SR για καθορισμένο p_0^+ και ARL_0^+ σε δείγματα μεγέθους $n = 4, 5$ και 8	σελ. 185
4.3	Τιμές άνω ορίου ελέγχου του διαγράμματος SR για καθορισμένο p_0^+ και ARL_0^+ σε δείγματα μεγέθους $n = 8$ και 10	σελ. 186
4.4	Τιμές του διαστήματος απόφασης $h_{n,p}$ για το πολυμεταβλητό μη-παραμετρικό διάγραμμα τύπου $CUSUM$ των Dai, Zhou και Wang (2011)	σελ. 196
4.5	Τιμές ορίου ελέγχου L του διαγράμματος $MSEWMA$ για διαφορετικές τιμές των λ , p και ARL_0	σελ. 197
4.6	Τιμές στατιστικών συναρτήσεων των διαγραμμάτων SN^2 και SR^2	σελ. 203
4.7	Τιμές στατιστικών συναρτήσεων των διαγραμμάτων r , Q , S και S^*	σελ. 208
4.8	Τιμές στατιστικής συνάρτησης του διαγράμματος $DDMA$	σελ. 210

РАНЕЕЗНАМО ТЕРПАА

Κατάλογος Σχημάτων

1.1	Διάγραμμα τυπικής διαχρονικής χρήσης των τριών κατηγοριών του Στατιστικού Ελέγχου Ποιότητας (Montgomery (1997)).....	σελ.	3
1.2	Ενεπίγραφή στήλη της Ελευσίνας (Βαρουφάκης (2005))	σελ.	5
1.3	Τυπικό Διάγραμμα Ελέγχου.....	σελ.	15
1.4	Κυκλικό ή περιοδικό πρότυπο.....	σελ.	17
1.5	Μικτό πρότυπο	σελ.	18
1.6	Πρότυπο αλμάτων	σελ.	18
1.7	Πρότυπο τάσης.....	σελ.	19
1.8	Πρότυπο έλλειψης μεταβλητότητας.....	σελ.	19
1.9	Διάγραμμα Ελέγχου με Προειδοποιητικά Όρια Ελέγχου	σελ.	20
2.1	Διάγραμμα R Φάσης I για τη διασπορά του χαρακτηριστικού X	σελ.	43
2.2	Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού (Μέθοδος S).....	σελ.	44
2.3	Διάγραμμα S Φάσης I για τη διασπορά του χαρακτηριστικού X	σελ.	45
2.4	Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού (Μέθοδος S).....	σελ.	46
2.5	Διάγραμμα S^2 Φάσης I για τη διασπορά του χαρακτηριστικού X	σελ.	47
2.6	Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού X (Μέθοδος S^2).....	σελ.	48
2.7	Διάγραμμα np Φάσης I για τον αριθμό των απορριπτόμενων προσόψεων	σελ.	52
3.1	Διάγραμμα <i>MPYRAMID</i> δύο διαστάσεων (Runger and Testik (2004)).....	σελ.	98
3.2	Διάγραμμα <i>MCONE</i> δύο διαστάσεων (Runger and Testik (2004)).....	σελ.	98
3.3	Διάγραμμα <i>MPARABOLOID</i> δύο διαστάσεων (Runger and Testik (2004)).....	σελ.	98
3.4	Διάγραμμα προσαρμοσμένης δισδιάστατης κανονικής κατανομής.....	σελ.	118
3.5	Διάγραμμα προσαρμοσμένων κανονικών κατανομών για κάθε χαρακτηριστικό	σελ.	118
3.6	Διάγραμμα διασποράς μεταξύ της αντοχής στη θραύση και της διαμέτρου μιας υφαντικής ίνας.....	σελ.	119
3.7	Διάγραμμα T^2 του Hotelling Φάσης I για 30 δείγματα.....	σελ.	120
3.8	Διάγραμμα T^2 του Hotelling Φάσης I για 29 δείγματα.....	σελ.	120
3.9	Διάγραμμα Γενικευμένης Διακύμανσης Φάσης I	σελ.	122
3.10	Διάγραμμα <i>MEWMA</i> Φάσης I.....	σελ.	124

3.11	Δεύτερο διάγραμμα <i>MCUSUM</i> του Crosier Φάσης I	σελ. 125
3.12	Πρώτο διάγραμμα <i>MCUSUM</i> των Pignatiello και Runger Φάσης I.....	σελ. 125
3.13	Διάγραμμα <i>MNP</i> Φάσης I για το ζυγισμένο άθροισμα των ελαττωματικών προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 30$)	σελ. 128
3.14	Διάγραμμα <i>MNP</i> Φάσης I για το ζυγισμένο άθροισμα των ελαττωματικών προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 29$)	σελ. 129
3.15	Διάγραμμα <i>MP</i> Φάσης I για το συνολικό αριθμό ελαττωμάτων ή μη συμμορφωμένων προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 30$)	σελ. 130
3.16	Διάγραμμα <i>MP</i> Φάσης I για το συνολικό αριθμό ελαττωμάτων ή μη συμμορφωμένων προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 29$)	σελ. 132
3.17	Διάγραμμα T^2 του Hotelling Φάσης I πριν τη μετατόπιση του μέσου της αντοχής στη θραύση για $a = 0.0027$	σελ. 134
3.18	Διάγραμμα T^2 του Hotelling Φάσης I μετά τη μετατόπιση του μέσου της αντοχής στη θραύση για $a = 0.0027$	σελ. 134
3.19	Διάγραμμα \bar{X} Φάσης I για την αντοχή στη θραύση της υφαντικής ίνας μετά τη μετατόπιση του μέσου της διεργασίας	σελ. 135
3.20	Διάγραμμα \bar{X} Φάσης I για τη διάμετρο της υφαντικής ίνας μετά τη μετατόπιση του μέσου της διεργασίας	σελ. 135
3.21	Διάγραμμα έλλειψης των δύο συσχετισμένων ποιοτικών χαρακτηριστικών.....	σελ. 139
3.22	Ελλειπτική και ορθογώνια περιοχή ελέγχου	σελ. 140
3.23	Ελλειψοειδές διάγραμμα για τρία χαρακτηριστικά (StatPoint, Inc. (2005)).....	σελ. 141
3.24	Μέθοδος υπερ-επιπέδου στον τρισδιάστατο χώρο (Chua and Montgomery (1992).....	σελ. 142
3.25	Διάγραμμα <i>MP</i> για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994)).....	σελ. 167
3.26	Πληροφορίες που περιέχονται σε κάθε θηκόγραμμα τη χρονική στιγμή t (Atienza, Tang and Ang (1998))	σελ. 169
3.27	Διάγραμμα ελέγχου <i>MBTCC</i> (Atienza, Tang and Ang (1998))	σελ. 169
3.28	Πολυδιάγραμμα για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994)).....	σελ. 170
3.29	StarPlot για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994)).....	σελ. 172

3.30	Διάγραμμα ακτίνων για δέκα ποιοτικά χαρακτηριστικά (Nottingham, Cook and Zobel (2001)).....	σελ. 173
3.31	Πολυμεταβλητό διάγραμμα ιστού πέντε χαρακτηριστικών (Adams (1994))	σελ. 173
3.32	Πολυμεταβλητό διάγραμμα ιστού πέντε χαρακτηριστικών για τρεις χρονικές περιόδους (Adams (1994)).....	σελ. 174
4.1	Μη παραμετρικό διάγραμμα SN^2 ($a = 0.05$)	σελ. 202
4.2	Μη παραμετρικό διάγραμμα SR^2 ($a = 0.05$)	σελ. 203
4.3	Παραμετρικό διάγραμμα T^2 του Hotelling ($a = 0.05$)	σελ. 204
4.4	Μη παραμετρικό διάγραμμα r ($a = 0.10$)	σελ. 205
4.5	Μη παραμετρικό διάγραμμα Q ($a = 0.10$)	σελ. 206
4.6	Μη παραμετρικό διάγραμμα S τύπου $CUSUM$ ($a = 0.10$)	σελ. 207
4.7	Μη παραμετρικό διάγραμμα S^* τύπου $CUSUM$ ($a = 0.10$)	σελ. 207
4.8	Μη παραμετρικό διάγραμμα κινούμενου μέσου $DDMA$ ($a = 0.10$)	σελ. 209
4.9	Παραμετρικό διάγραμμα T^2 του Hotelling ($a = 0.10$)	σελ. 211

РАНЕЕЗНАМО ТЕРРА

Κατάλογος Συντομογραφιών

AEM-median	Affine Equivariant Multivariate median – πολυδιάστατη διάμεσος ισο-μεταβολής
AMMA	Arithmetic Multivariate Moving Average – αριθμητικό πολυμεταβλητό διάγραμμα κινούμενου μέσου
ARL	Average Run Length – μέσο μήκος ροής
ARL₀	In Control Average Run Length – εντός ελέγχου μέσο μήκος ροής
ARL₁	Out-of-Control Average Run Length – εκτός ελέγχου μέσο μήκος ροής
ASA	American Statistical Association – Αμερικανική Στατιστική Ένωση
ASQ	American Society for Quality – Αμερικανική Εταιρεία Ποιότητας
ASQC	American Society for Quality Control – Αμερικανική Εταιρεία Ελέγχου Ποιότητας
ATS	Average Time to Signal – μέσος χρόνος σήματος
BSA	Backward Selection Algorithm – προς-τα-πίσω αλγόριθμος επιλογής
cdf	Cumulative Distribution Function – αθροιστική συνάρτηση κατανομής (α.σ.κ.)
CL	Center Line – κεντρική γραμμή
COT	CUSUM of T – συσσωρευμένο άθροισμα του στατιστικού T (του Hotelling)
CUSUM	Cumulative Sum – συσσωρευμένο άθροισμα
DDMA	Nonparametric Multivariate Data Depth Moving Average Chart – μη παραμετρικό πολυδιάστατο διάγραμμα ελέγχου κινούμενου μέσου
EWMA	Exponentially Weighted Moving Average – εκθετικά ζυγισμένος κινούμενος μέσος
FAR	False Alarm Rate – ποσοστό λανθασμένου συναγερμού
FIR	Fast Initial Response – μέθοδος αρχικής αντίδρασης
FIT	Finite Intersections Tests – πεπερασμένοι έλεγχοι τομής
FSA	Forward Selection Algorithm – προς-τα-εμπρός αλγόριθμος επιλογής
ISO	International Standards Organization – Διεθνής Οργανισμός Τυποποίησης
IWL	Inner Warning Limits – εσωτερικά προειδοποιητικά όρια
JUSE	Japanese Union of Scientists and Engineers – Ιαπωνική Ένωση Επιστημόνων και Μηχανικών
LCL	Lower Control Limit – κάτω όριο ελέγχου
MA	Moving Average – κινούμενος μέσος
MA_{CC}	Multivariate Attribute Control Chart – πολυμεταβλητό διάγραμμα ελέγχου για ιδιότητες
MA-GR	Multiattribute Group Runs Chart – πολυμεταβλητό διάγραμμα ομάδων ροών για ιδιότητες
MA-<i>np</i>	Multiattribute <i>np</i> Chart – πολυμεταβλητό διάγραμμα <i>np</i> για ιδιότητες
MA-Syn	Multiattribute Synthetic Chart – πολυμεταβλητό συνθετικό διάγραμμα για ιδιότητες
MBTCC	Multivariate Boxplot T^2 Control Chart – πολυμεταβλητό διάγραμμα ελέγχου θηκόγραμμα- T^2
MCUSUM	Multivariate Cumulative Sum – πολυμεταβλητό συσσωρευμένο άθροισμα
MD_G	Mahalanobis Depth – βάθος δεδομένων που βασίζεται στην απόσταση του Mahalanobis

MD_{Gm}	Sample Mahalanobis Depth – εμπειρικό βάθος δεδομένων που βασίζεται στην απόσταση του Mahalanobis
MEWMA	Multivariate Exponentially Weighted Moving Average – πολυμεταβλητός εκθετικά ζυγισμένος κινούμενος μέσος
MMA	Multivariate Moving Average – πολυμεταβλητός κινούμενος μέσος
MNP	Multivariate Control Chart for Number Nonconforming – πολυμεταβλητό διάγραμμα ελέγχου για τον αριθμό των ελαττωματικών προϊόντων
MP	Multivariate Profile Chart – πολυμεταβλητό διάγραμμα προφίλ
MSEWMA	Multivariate Sign EWMA Control Chart – διάγραμμα ελέγχου εκθετικά κινούμενου μέσου που βασίζεται στον προσημικό έλεγχο
MSPC	Multivariate Statistical Quality Control – πολυμεταβλητός στατιστικός έλεγχος διεργασιών
NSPC	Nonparametric Statistical Process Control – Μη Παραμετρικός Στατιστικός Έλεγχος Ποιότητας
NTL	Natural Tolerance Limits – φυσικά όρια ανοχής
OWL	Outer Warning Limits – εξωτερικά προειδοποιητικά όρια
PCA	Principal Component Analysis – Ανάλυση Κυρίων Συνιστωσών
PLS	Partial Least Squares – Μέθοδος Μερικών Ελαχίστων Τετραγώνων
PP	Projection Pursuit – Μέθοδος Προβολών
RL	Run Length – Μήκος Ροής
RSM	Response Surface Methodology – Μεθοδολογία Επιφανειών Απόκρισης
SD_G	Simplicial Depth – απλό βάθος
SD_{Gm}	Sample Simplicial Depth – απλό δειγματικό βάθος
SN test	Sign Test – προσημικός έλεγχος
SPC	Statistical Process Control – Στατιστικός Έλεγχος Διεργασιών
SQC	Statistical Quality Control – Στατιστικός Έλεγχος Ποιότητας
SR test	Wilcoxon's Signed-Rank Test – προσημικός βαθμολογικός έλεγχος του Wilcoxon
TQM	Total Quality Management – Διαχείριση Ολικής Ποιότητας
UCL	Upper Control Limit – άνω όριο ελέγχου
β.ε.	βαθμοί ελευθερίας
ΚΟΘ	Κεντρικό Οριακό Θεώρημα
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
τ.δ.	τυχαίο διάνυσμα

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Στατιστικός Έλεγχος Ποιότητας

Σε πολλούς από εμάς έχει τύχει να αγοράσουμε ένα προϊόν και να ανακαλύψουμε ότι αυτό είναι ελαττωματικό ή ότι δε λειτουργεί σωστά ή ότι κάποιο εξάρτημά του λείπει. Είναι φυσικό κάθε άνθρωπος, ως καταναλωτής, να περιμένει τα προϊόντα που αγοράζει να λειτουργούν σύμφωνα με τις προδιαγραφές και με τον τρόπο που αυτά είχαν αρχικά σχεδιαστεί. Ωστόσο, δεν είναι πάντα εφικτό να ελέγχονται όλα τα παραγόμενα προϊόντα σε κάθε παραγωγική διαδικασία. Για το σκοπό αυτό κρίθηκε απαραίτητη η ανάπτυξη τεχνικών, οι οποίες μεγιστοποιούν τη δυνατότητα παρακολούθησης της ποιότητας των προϊόντων που παράγονται και ταυτόχρονα μειώνουν τα ελαττώματά τους.

Η έννοια του ελέγχου ποιότητας είναι πλέον συνυφασμένη με την παραγωγικότητα και με κάθε τομέα ελέγχου της παραγωγής. Κάθε βιομηχανία έχει ως στόχο την παραγωγή ενός προϊόντος με την καλύτερη δυνατή σχεδίαση, κατασκευή και εμφάνιση, καθώς επίσης και με το μικρότερο δυνατό κόστος. Συνεπώς, η ανταγωνιστικότητα μεταξύ των βιομηχανιών καθορίζεται από δύο παράγοντες, την ποιότητα και το κόστος, οι οποίοι με τη σειρά τους επηρεάζουν την επιλογή του προϊόντος από τον καταναλωτή. Κάθε επιχείρηση, για να διασφαλίσει την ποιότητα, υποχρεούται να παράγει προϊόντα που συμμορφώνονται με ορισμένες προδιαγραφές, οι οποίες καθορίζονται από το Διεθνή Οργανισμό Τυποποίησης (*International Standards Organization, ISO*). Την τελευταία δεκαετία του 20^{ου} αιώνα, τα πρότυπα της σειράς ISO-9000 (EN-29000 για την Ευρωπαϊκή Ένωση και Q-90 για τις Η.Π.Α.) αποτελούν παγκόσμιο μοντέλο συστήματος ποιότητας (Γεωργακάκος (2002)).

Η ποιότητα, όμως, μπορεί να σημαίνει διαφορετικά πράγματα σε διαφορετικούς ανθρώπους. Μπορεί ακόμη να έχει διαφορετική σημασία στους ίδιους ανθρώπους, σε διαφορετικές χρονικές στιγμές και υπό διαφορετικές συνθήκες (Doty (1996)). Το γεγονός αυτό καθιστά δύσκολο τον ακριβή ορισμό της ποιότητας.

Ο Garvin (1987) προσπάθησε να ορίσει την ποιότητα χρησιμοποιώντας τις επόμενες οχτώ διαφορετικές συνιστώσες-διαστάσεις της έννοιας αυτής:

1. **Επίδοση** (*performance*) του προϊόντος, δηλαδή κατά πόσο το προϊόν αυτό θα ανταπεξέλθει σε συγκεκριμένες λειτουργίες και πόσο καλά θα τις κάνει.
2. **Αξιοπιστία** (*reliability*), δηλαδή πόσο συχνά το προϊόν θα χρειαστεί να επισκευαστεί στο χρόνο ζωής του.
3. **Διάρκεια** (*durability*), δηλαδή η πραγματική διάρκεια ζωής του προϊόντος.
4. **Ικανότητα συντήρησης** (*serviceability*) του προϊόντος και βαθμός ευκολίας της επιδιόρθωσής του.
5. **Αισθητική** (*aesthetics*) του προϊόντος, δηλαδή η εξωτερική του εμφάνιση.
6. **Δυνατότητες** (*features*) του προϊόντος, δηλαδή τί ακριβώς μπορεί κάνει το προϊόν αυτό.
7. **Ποιότητα όπως την αντιλαμβάνεται ο καταναλωτής** ανάλογα με τη φήμη της εταιρίας που παράγει το προϊόν ή ακόμα και τη φήμη του ίδιου του προϊόντος (*perceived quality*).
8. **Συμμόρφωση του προϊόντος σύμφωνα με ορισμένα πρότυπα** (*conformance to standards*). Με άλλα λόγια, κατά πόσο το προϊόν έχει σχεδιαστεί ακριβώς έτσι όπως ο σχεδιαστής του είχε αρχικά προορίσει να γίνει.

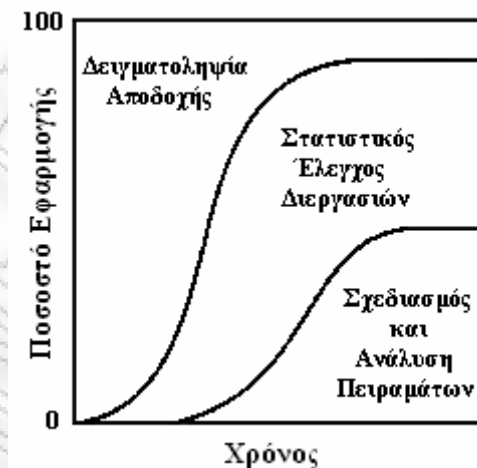
Διαχρονικά, προτάθηκαν αρκετοί διαφορετικοί ορισμοί της ποιότητας. Σύμφωνα με τον παραδοσιακό ορισμό (*fitness for use*), ποιότητα είναι ο βαθμός στον οποίο ένα προϊόν ικανοποιεί τις προσδοκίες του καταναλωτή (Montgomery (1997)). Έχει δοθεί και ένας πιο μοντέρνος ορισμός, σύμφωνα με τον οποίο η ποιότητα είναι αντιστρόφως ανάλογη της ανεπιθύμητης μεταβλητότητας. Επομένως, η βελτίωση της ποιότητας απαιτεί τη μείωση της μεταβλητότητας σε διαδικασίες και προϊόντα. Κατ' επέκταση, ο Στατιστικός Έλεγχος Ποιότητας (*Statistical Quality Control, SQC*) ορίζεται ως η εφαρμογή στατιστικών μεθόδων

σε κατασκευαστικά και βιομηχανικά προβλήματα (Howell (1952)). Με άλλα λόγια, με τον όρο αυτό εννοούμε τη χρήση στατιστικών μεθόδων στον έλεγχο και στη βελτίωση της ποιότητας στη βιομηχανική παραγωγή (Burr (1976)). Συνεπώς, ο όρος Στατιστικός Έλεγχος Ποιότητας περιγράφει ένα σύνολο από στατιστικά εργαλεία που χρησιμοποιούνται για να εντοπίσουν και να διαχειριστούν προβλήματα ποιότητας τόσο στην παραγωγική διαδικασία όσο και στο ίδιο το προϊόν (Reid and Sanders (2010)).

Ο Στατιστικός Έλεγχος Ποιότητας μπορεί να χωριστεί στις επόμενες τρεις μεγάλες κατηγορίες (Montgomery (1997)):

- Σχεδιασμός και Ανάλυση Πειραμάτων (*Design of Experiments*)¹,
- Στατιστικός Έλεγχος Διεργασιών (*Statistical Process Control, SPC*),
- Δειγματοληψία Αποδοχής (*Acceptance Sampling*).

Κάθε μία από τις παραπάνω κατηγορίες αντιστοιχεί σε διαφορετική φάση της Παραγωγικής Διαδικασίας: Σχεδιασμός Προϊόντος, Παραγωγική Διεργασία και Τελικό Προϊόν αντίστοιχα (Αντζουλάκος (2010)). Η τυπική διαχρονική χρήση των τριών παραπάνω κατηγοριών του Στατιστικού Ελέγχου Ποιότητας στις περισσότερες εταιρίες φαίνεται στο ακόλουθο σχήμα (Σχήμα 1.1).



Σχήμα 1.1: Διάγραμμα τυπικής διαχρονικής χρήσης των τριών κατηγοριών του Στατιστικού Ελέγχου Ποιότητας (Montgomery (1997))

¹ Οι Reid και Sanders (2010) αναφέρουν την πρώτη κατηγορία του Στατιστικού Ελέγχου Ποιότητας ως Περιγραφική Στατιστική (*Descriptive Statistics*).

Ο Σχεδιασμός και η Ανάλυση Πειραμάτων χρησιμοποιεί στατιστικές μεθόδους για να περιγράψει τα ποιοτικά χαρακτηριστικά και τις σχέσεις μεταξύ αυτών. Ο Στατιστικός Έλεγχος Διεργασιών περιλαμβάνει την επιθεώρηση ενός τυχαίου δείγματος της παραγωγικής διαδικασίας, στην οποία βασιζόμαστε για να αποφασίσουμε εάν η διαδικασία παράγει προϊόντα με προκαθορισμένα χαρακτηριστικά. Με πιο απλά λόγια, μπορούμε να καταλάβουμε εάν η διαδικασία λειτουργεί σωστά ή όχι. Τέλος, η Δειγματοληψία Αποδοχής είναι η διαδικασία με την οποία επιθεωρούμε τυχαία ένα δείγμα από αγαθά και, με βάση τα αποτελέσματα στα οποία καταλήγουμε, αποφασίζουμε εάν ολόκληρη η παρτίδα των αγαθών θα γίνει αποδεκτή ή θα απορριφθεί (Reid and Sanders (2010)).

Όπως έχουμε ήδη αναφέρει, ο Στατιστικός Έλεγχος Ποιότητας είναι ένα σύνολο στατιστικών εργαλείων και μεθόδων που χρησιμοποιούνται για τη βελτίωση της ποιότητας των παραγόμενων προϊόντων. Τα κυριότερα εργαλεία του Στατιστικού Ελέγχου Ποιότητας είναι 7 και είναι γνωστά ως “magnificent seven” (Montgomery (1997)). Τα εργαλεία αυτά είναι τα ακόλουθα:

- Ιστόγραμμα, Διάγραμμα Μίσχου-Φύλλων (*Histogram, Stem-and-Leaf Plot*)
- Φύλλο Ελέγχου (*Check Sheet*)
- Διάγραμμα Pareto (*Pareto Chart*)
- Διάγραμμα Αιτίας-Αποτελέσματος (*Cause-and-Effect Diagram*)
- Διάγραμμα Συγκέντρωσης Ελαττωμάτων (*Defect Concentration Diagram*)
- Διάγραμμα Διασποράς ή Διασκόρπισης (*Scatter Plot*)
- Διάγραμμα Ελέγχου (*Control Chart*)

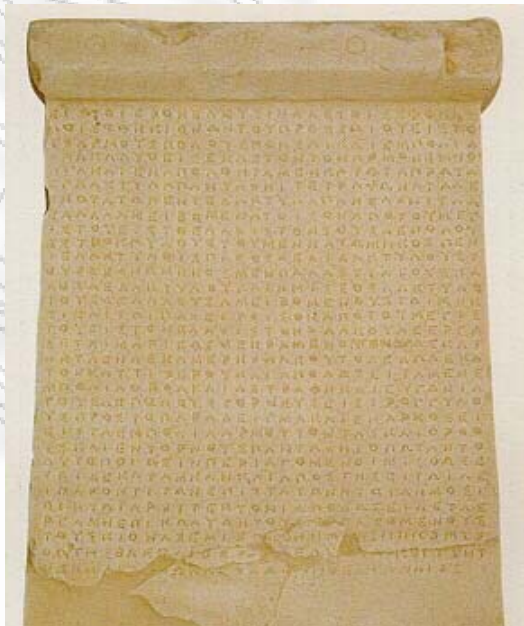
1.2 Ιστορική Αναδρομή

Σύμφωνα με τον ορισμό που δίνει ο Howell (1952), το πιο πρώιμο παράδειγμα Ποιοτικού Ελέγχου εντοπίζεται εκατομμύρια χρόνια πριν, την εποχή του λίθου, όταν ο προϊστορικός άνθρωπος προσπάθησε να κατασκευάσει τον πρώτο πέλεκυ με πρώτη ύλη την πέτρα και το ξύλο. Η ανακάλυψη της επιγραφής της Ελευσίνας¹ (βλ. Σχήμα 1.2), η οποία χρονολογείται τον 4^ο αιώνα π.Χ., αποδεικνύει ότι στην Αρχαία Ελλάδα υπήρχε ένας εμπειρικός τρόπος

¹ Το κείμενό της θεωρείται το αρχαιότερο πρότυπο ελέγχου ποιότητας (Βαρουφάκης (2005))

ελέγχου ποιότητας. Το κείμενο της επιγραφής αυτής αναφέρεται σε μια παραγγελία για την κατασκευή ορειχάλκινων συνδέσμων, που θα έμπαιναν ανάμεσα στους σπονδύλους κίωνων της Φιλώνειας Στοάς. Τόσο οι τεχνικές, όσο και οι χημικές προδιαγραφές για τη σύνθεση των συνδέσμων πόλων και εμπολίων των κίωνων αυτών αναφέρονται με αυστηρότητα στην επιγραφή. Το γεγονός αυτό επιβεβαιώνει ότι οι Αρχαίοι Έλληνες εφάρμοζαν αυστηρές προδιαγραφές στις παραγγελίες τους (Παπαργύρης και Παπαργύρης (2010)).

Την ίδια περίπου περίοδο, ο Αριστοτέλης στο έργο του *Αθηναίων Πολιτεία*, αναφέρεται σε θέματα ποιότητας ελέγχου των προϊόντων της αγοράς. Συγκεκριμένα, αναφέρει ότι υπήρχε ένας ολόκληρος μηχανισμός ελέγχου των προϊόντων που διεξαγόταν με τη βοήθεια δημοσίων λειτουργών, οι οποίοι ορίζονταν με κλήρο. Τα προς πώληση αγαθά έπρεπε να είναι «καθαρά» και «ακίβδηλα», δηλαδή θα έπρεπε να τηρούν συγκεκριμένους όρους υγιεινής και να είναι ανόθευτα. Για την επίτευξη του σκοπού αυτού ορίζονταν δέκα Αγορανόμοι. Υπήρχαν, επίσης, και δέκα Μετρονόμοι (πέντε για την Αθήνα και πέντε για τον Πειραιά) που ήταν υπεύθυνοι ώστε τα μέτρα και τα σταθμά να ανταποκρίνονται σε συγκεκριμένες προδιαγραφές και να χρησιμοποιούνται σωστά από τους εμπόρους. Τέλος, η πολιτεία όριζε και τριανταπέντε Σιτοφύλακες (είκοσι για την Αθήνα και δεκαπέντε για τον Πειραιά), οι οποίοι έπρεπε να ελέγχουν την ποιότητα του σιταριού και του κριθαριού (Camp (2004), Βαρουφάκης (2005)).



Σχήμα 1.2: Ενεπίγραφη στήλη της Ελευσίνας (Βαρουφάκης (2005))

Ο Έλεγχος Ποιότητας, όμως, συνεχίζει να εξελίσσεται δια μέσου των αιώνων. Μια από τις μεγαλύτερες καινοτομίες που έγινε στη βιομηχανία χρονολογείται λίγο πριν το 1800, όταν ο Eli Whitney δημιούργησε κάποια ανταλλάξιμα τμήματα για όπλα. Ήταν σύνηθες φαινόμενο την εποχή εκείνη ένας άνθρωπος να φτιάχνει τα δικά του εργαλεία ή τμήματα αυτών, να τα επιθεωρεί και στη συνέχεια να τα συναρμολογεί (Howell (1952)). Το 1875 ο Frederick W. Taylor εισήγαγε αρχές στην «Επιστημονική Διαχείριση» (*Scientific Management*), σύμφωνα με τις οποίες η εργασία χωρίζεται σε επιμέρους καθήκοντα, ώστε ένα προϊόν να κατασκευάζεται και να συναρμολογείται ευκολότερα. Αυτή η τυποποίηση των μεθόδων εργασίας δημιούργησε την έννοια των εργασιακών προτύπων, όπως ο χρόνος που απαιτούνταν για να επιτευχθεί μία εργασία ή ο προκαθορισμένος αριθμός μονάδων που πρέπει να παραχθούν σε μια συγκεκριμένη χρονική περίοδο. Σε αυτή την ιδέα βασίστηκαν ο Frank Gilbreth και άλλοι για να μελετήσουν την κίνηση και το σχεδιασμό της εργασίας, δίνοντας έμφαση στην παραγωγικότητα (Montgomery (1997)).

Στα επόμενα χρόνια, μεταξύ 1900 και 1930, ο Henry Ford επέκτεινε την καλυτέρευση των εργασιακών μεθόδων με τη βελτίωση της παραγωγικότητας και της ποιότητας. Ανέπτυξε, επίσης, τις έννοιες αυτοελέγχου και επιθεώρησης κατά τη διάρκεια της παραγωγικής διαδικασίας. Το 1901 καθιερώθηκαν τα πρώτα τυποποιημένα εργοστάσια στη Μεγάλη Βρετανία και το 1907 η Αμερικανική Εταιρεία AT&T (*American Telephone & Telegraph*) ξεκινά τη συστηματική επιθεώρηση και το συστηματικό έλεγχο προϊόντων και υλικών. Ένα χρόνο αργότερα, ο W.S. Gosset εισήγαγε την κατανομή t (t -distribution ή *student distribution*) και τα αποτελέσματα που εξήγαγε εφαρμόστηκαν σε θέματα ποιότητας ελέγχου στο εργοστάσιο ζυθοποιίας Guinness. Μεταξύ 1915 και 1919, συστάθηκε ο Σύνδεσμος Τεχνικής Επιθεώρησης (*Technical Inspection Association*) στην Αγγλία, ο οποίος αργότερα μετονομάστηκε σε Ινστιτούτο Διασφάλισης της Ποιότητας (*Institute of Quality Assurance*). Τη δεκαετία του 1920, η εταιρεία AT&T Bell Laboratories δημιούργησε Τμήμα Ποιοτικού Ελέγχου, το οποίο έδινε έμφαση στην ποιότητα, την επιθεώρηση, τον έλεγχο και την αξιοπιστία των προϊόντων. Την ίδια περίοδο, ο B.P. Dudding της εταιρείας General Electric της Αγγλίας χρησιμοποίησε στατιστικές μεθόδους για να ελέγξει την ποιότητα των ηλεκτρικών λαμπτήρων. Μεταξύ 1922 και 1923, ο R.A. Fisher δημοσίευσε μία σειρά από θεμελιώδεις μελέτες στους Πειραματικούς Σχεδιασμούς και στις εφαρμογές τους στις Γεωργικές Επιστήμες (Montgomery (1997)).

Στην πραγματικότητα, η έννοια του Στατιστικού Ελέγχου Ποιότητας καθιερώθηκε επίσημα το 1924, όταν ο Walter A. Shewhart της εταιρείας Bell Telephone Laboratories κυκλοφόρησε ένα υπόμνημα που περιλάμβανε ένα μοντέρνο διάγραμμα ελέγχου. Το 1928 η μεθοδολογία της Δειγματοληψίας Αποδοχής αναπτύχθηκε και τελειοποιήθηκε από τους H.F. Dodge και H.G. Romig (1929). Στη συνέχεια, το 1931 ο Shewhart εξέδωσε το κλασικό του βιβλίο *Economic Control of Quality of Manufactured Product* περιγράφοντας στατιστικές μεθόδους που μπορούν να εφαρμοστούν στην παραγωγή και στα διαγράμματα ελέγχου. Ένα χρόνο μετά, έδωσε διαλέξεις στο Πανεπιστήμιο του Λονδίνου με θέμα τις μεθόδους αυτές. Το 1938, ο W.E. Deming κάλεσε τον Shewhart να παρουσιάσει σεμινάρια στους απόφοιτους του Τμήματος Γεωργίας των Η.Π.Α. με θέμα τα Διαγράμματα Ελέγχου. Μέχρι και το τέλος του 1930, οι μέθοδοι του Στατιστικού Ελέγχου Ποιότητας χρησιμοποιούνταν από την εταιρεία Western Electric, ωστόσο η χρήση τους δεν είχε επεκταθεί στη βιομηχανία γενικότερα (Montgomery (1997), Parr (2008)).

Σύμφωνα με τον Montgomery (1997), η έλευση του Δευτέρου Παγκοσμίου Πολέμου συνετέλεσε στο να γίνει ευρέως γνωστός ο Στατιστικός Έλεγχος Ποιότητας και οι μέθοδοί του να χρησιμοποιηθούν στη βιομηχανία. Κατά τη διάρκεια του πολέμου, οι στατιστικές μέθοδοι για τον έλεγχο και τη βελτίωση της ποιότητας ενός προϊόντος κρίθηκαν απαραίτητες λόγω της έλλειψης εξειδικευμένου προσωπικού στον τομέα της παραγωγής. Ως αποτέλεσμα, το 1946 συστάθηκε η Αμερικανική Εταιρεία Ελέγχου Ποιότητας (*American Society for Quality Control, ASQC*), η οποία προωθούσε τη χρήση μεθόδων βελτίωσης της ποιότητας για όλα τα είδη των προϊόντων και των υπηρεσιών. Επίσης, προσέφερε συνέδρια και δημοσιεύσεις σε τεχνικά θέματα και εκπαιδευτικά προγράμματα στην πιστοποίηση της ποιότητας. Την ίδια χρονιά, το Τμήμα Οικονομικών και Επιστημονικών Υπηρεσιών του Αμερικανικού Τομέα Πολέμου (*Economic and Scientific Services Section of U.S. War Department*) κάλεσε τον Deming στην Ιαπωνία για να βοηθήσει τις δυνάμεις κατοχής να ξαναχτίσουν την ιαπωνική βιομηχανία. Επιπρόσθετα, ιδρύθηκε η Ιαπωνική Ένωση Επιστημόνων και Μηχανικών (*Japanese Union of Scientists and Engineers, JUSE*). Κατά τη διάρκεια 1946 και 1949, ο Deming έδινε σεμινάρια Στατιστικού Ελέγχου Ποιότητας στην ιαπωνική βιομηχανία. Εν τω μεταξύ, το 1948, ο G. Taguchi μελετούσε τους πειραματικούς σχεδιασμούς και τις εφαρμογές τους.

Το 1950, οι μέθοδοι του Στατιστικού Ελέγχου Ποιότητας διδάσκονταν στην Ιαπωνία και συγκεκριμένα ο Deming ήταν αυτός που δίδασκε τους διευθυντές της ιαπωνικής βιομηχανίας.

Την ίδια περίοδο, ο K. Ishikawa εισήγαγε το Διάγραμμα Αιτίας-Αποτελέσματος (*Cause and Effect Diagram*). Τη δεκαετία του 1950, δημοσιεύτηκαν πολλές σημαντικές μελέτες και εργασίες σε θέματα Στατιστικού Ελέγχου Ποιότητας με πρωτεργάτες τους E. Grant και A.J. Duncan. Το 1951, ο A.V. Feigenbaum εξέδωσε την πρώτη έκδοση του βιβλίου του “*Total Quality Control*” και η Ιαπωνική Ένωση Επιστημόνων και Μηχανικών (*JUSE*) καθιέρωσε το βραβείο Deming (“*Deming Prize*”) για επιβράβευση σημαντικών επιτευγμάτων στον Έλεγχο Ποιότητας και τη μεθοδολογία του.

Λίγο αργότερα, άρχισαν να παγιώνονται οι εφαρμογές των Πειραματικών Σχεδιασμών στη χημική βιομηχανία. Αφορμή αποτέλεσε μια σχετική εργασία των G.E.P. Box και K.B. Wilson (1951). Η εργασία αυτή αφορούσε το σχεδιασμό πειραμάτων στη βιομηχανία, εστιάζοντας στη χημική βιομηχανία, και εισήγαγε τη μεθοδολογία των Επιφανειών Απόκρισης (*Response Surface Methodology, RSM*). Το 1954, ο J.M. Juran έδωσε μερικές διαλέξεις σε θέματα διαχείρισης και βελτίωσης της ποιότητας στην Ιαπωνία και ο Βρετανός στατιστικός E.S. Page (1954) εισήγαγε τα Αθροιστικά Διαγράμματα *CUSUM* (*Cumulative Sum*). Το 1957 εκδόθηκε το βιβλίο των J.M. Juran και F.M. Gryna με τίτλο “*Quality Control Handbook*” και δύο χρόνια αργότερα ξεκίνησε η έκδοση του περιοδικού *Technometrics* από τον J.S. Hunter. Το 1959, εισήχθησαν και διαγράμματα *EWMA* (*Exponentially Weighted Moving Average*) από τον S. Roberts (1959).

Το 1960, οι G.E.P. Box και J.S. Hunter (1961) έγραψαν μία θεμελιώδη εργασία για τους 2^{k-p} παραγοντικούς σχεδιασμούς και ο K. Ishikawa εισήγαγε την έννοια του Κύκλου της Ποιότητας Ελέγχου (*Quality Control Circle*) στην Ιαπωνία. Ένα χρόνο αργότερα, σχηματίστηκε στη Μεγάλη Βρετανία το Εθνικό Συμβούλιο Ποιότητας και Παραγωγικότητας (*National Council for Quality and Productivity*) ως μέρος του Βρετανικού Συμβουλίου Παραγωγικότητας (*British Productivity Council*). Το 1969, διακόπηκε η έκδοση του περιοδικού “*Industrial Quality Control*” και αντικαταστάθηκε από το “*Quality Process*” και το “*Journal of Quality Technology*”. Γενικότερα, τη δεκαετία του 1960, εισήχθησαν μαθήματα Στατιστικού Ελέγχου Ποιότητας σε Πανεπιστημιακά Προγράμματα Βιομηχανικής Μηχανικής (*Industrial Engineering Academic Programs*) και στις βιομηχανίες των Η.Π.Α. εισήχθησαν προγράμματα, γνωστά ως «Μηδέν Ελαττώματα» (*Zero Defects, ZD*).

Τη δεκαετία του 1970, ιδρύθηκε ο Βρετανικός Σύνδεσμος Ποιότητας (*British Quality Association*). Μεταξύ 1975 και 1978, άρχισαν να εμφανίζονται βιβλία με θέμα τους

Πειραματικούς Σχεδιασμούς που απευθύνονταν σε μηχανικούς και άλλους επιστήμονες. Επιπλέον, ο Κύκλος της Ποιότητας φάνηκε να αποκτά μεγάλο ενδιαφέρον στη Βόρεια Αμερική. Αυτό είχε ως αποτέλεσμα την ανάπτυξη της Διαχείρισης της Ολικής Ποιότητας (*Total Quality Management, TQM*).

Τη δεκαετία του 1980, εισήχθησαν μέθοδοι Πειραματικών Σχεδιασμών και υιοθετήθηκαν από ένα ευρύ σύνολο οργανισμών που περιλάμβανε βιομηχανίες στον κλάδο των ηλεκτρονικών, του αεροδιαστήματος, των ημιαγωγών και των αυτοκινήτων. Επίσης, για πρώτη φορά εμφανίστηκαν στις Η.Π.Α. εργασίες και μελέτες του G. Taguchi με θέμα τους Πειραματικούς Σχεδιασμούς. Το 1984, η Αμερικανική Στατιστική Ένωση (*American Statistical Association, ASA*) καθιέρωσε μία Ad Hoc επιτροπή για την Ποιότητα και την Παραγωγικότητα (*Ad Hoc Committee on Quality and Productivity*), η οποία αργότερα εξελίχθηκε σε ολόκληρο τομέα της ASA. Δύο χρόνια αργότερα, ο G.E.P. Box και άλλοι επιστήμονες επισκέφθηκαν την Ιαπωνία με σκοπό να τονίσουν τη χρησιμότητα των Πειραματικών Σχεδιασμών και άλλων στατιστικών μεθόδων, ενώ οι Geladi and Kowalski (1986) ανέπτυξαν τη μέθοδο Μερικών Ελαχίστων Τετραγώνων (*Partial Least Squares, PLS*). Το 1988, το Κογκρέσο των Ηνωμένων Πολιτειών καθιέρωσε το Εθνικό Βραβείο Ποιότητας “*Malcolm Baldrige*”. Το 1987 άρχισε να εκδίδεται το περιοδικό “*Quality & Reliability Engineering International*” και δύο χρόνια αργότερα το “*Quality Engineering*”.

Στις αρχές της δεκαετίας του 1990, η Motorola εφάρμοσε το προγράμμα ποιότητας «Έξι Σίγμα» (6σ), το οποίο άρχισε σιγά σιγά να διαδίδεται και σε άλλες βιομηχανίες. Επίσης, καθιερώθηκαν διάφορα Πρότυπα Ποιότητας, όπως για παράδειγμα το ISO 9000 της βιομηχανίας των Η.Π.Α.. Παράλληλα, παρατηρήθηκε σταθερή αύξηση των αιτούντων για το βραβείο Baldrige και πολλές πολιτείες των Η.Π.Α. βασίστηκαν στα κριτήρια του βραβείου αυτού για να υποστηρίξουν Βραβεία Ποιότητας. Την ίδια περίοδο, οι στατιστικές τεχνικές, με έμφαση στα Διαγράμματα Ελέγχου και τους Πειραματικούς Σχεδιασμούς, άρχισαν να γίνονται μέρος της βασικής εκπαίδευσης πολλών προπτυχιακών προγραμμάτων Μηχανικής.

Συγκεκριμένα, το 1990, ο Snee περιγράφει πώς η έννοια της στατιστικής σκέψης μπορεί να χρησιμοποιηθεί στα διάφορα επίπεδα οργάνωσης. Τον ίδιο χρόνο, ο J.F. MacGregor έθεσε τους στόχους που αφορούσαν στην ανάλυση δεδομένων από παρτίδες. Αυτό είχε ως αποτέλεσμα να αναπτυχθούν μέθοδοι ανάλυσης διεργασιών με μεγάλο αριθμό δεδομένων και διαδικασίες για τη συνεχή παρακολούθηση διεργασιών. Επιπλέον, δημιουργήθηκαν απλά

εργαλεία που εφαρμόστηκαν στη βιομηχανία και αναπτύχθηκαν διαγνωστικά εργαλεία για την ταχεία ταυτοποίηση της φύσης οποιουδήποτε σφάλματος. Το 1991, τέθηκε η βασική μεθοδολογία εφαρμογής των μεθόδων Ανάλυσης Κυρίων Συνιστωσών (*Principal Component Analysis, PCA*) και Μερικών Ελαχίστων Τετραγώνων (*PLS*) για την παρακολούθηση συνεχών διεργασιών από τους Kresta, J.F. MacGregor (Kresta, MacGregor and Marlin (1991)) και B.M. Wise (Wise (1991), Wise, Veltkamp, Ricker, Kowalski, Barnes and Arakali (1991), Wise and Ricker (1991)).

Η χρήση της γραφικής αναπαράστασης με σκοπό τη διάγνωση των ειδικών αιτιών μεταβλητότητας εισήχθη από τους Miller και MacGregor το 1993 και το 1994 αντίστοιχα (Amin and Miller (1993), Miller, Swanson and Hecler (1993), MacGregor, Jaeckle, Kiparissides and Koutoudi (1994), MacGregor, Nomikos and Kourti (1994), Nomikos and MacGregor (1994)). Οι τεχνικές PCA και PLS επεκτάθηκαν, ώστε να διαχειριστούν μεγάλο πλήθος δεδομένων που αφορούν διεργασίες και το 1995 οι Nomikos και MacGregor δημιούργησαν τις μεθόδους Multiblock PCA και PLS (Nomikos and MacGregor (1995a), Nomikos and MacGregor (1995b)). Αυτές οι μέθοδοι χρησιμοποιήθηκαν και σε διεργασίες παρτίδων. Το ίδιο έτος, οι Rao και Arthanari (1995) εισήγαγαν τις τρεις ακόλουθες κατευθυντήριες γραμμές στις οποίες πρέπει να βασίζεται η στατιστική σκέψη: οι αποφάσεις πρέπει να βασίζονται σε σχετικές και αξιόπιστες πληροφορίες, οι πληροφορίες αυτές πρέπει να βασίζονται σε δεδομένα επαρκούς στατιστικής ποιότητας και να εξάγονται με κατάλληλα στατιστικά εργαλεία. Δε θα πρέπει να παραλείψουμε ότι το 1997 η Αμερικανική Εταιρεία Ελέγχου Ποιότητας (*American Society for Quality Control, ASQC*) μετονομάστηκε σε Αμερικανική Εταιρεία Ποιότητας (*American Society for Quality, ASQ*). Το 1999, οι Baardseth, Dalen και Tandberg μετέφεραν τη γνώση και την τεχνολογία που είχε αναπτυχθεί σε μικρές και μεγάλες επιχειρήσεις. Δέκα χρόνια μετά την καθιέρωση Προτύπων ποιότητας, δηλαδή το 2000, γίνεται η δεύτερη αναθεώρηση της σειράς Προτύπων ISO 9000.

Από τα παραπάνω, γίνεται φανερό ότι ο Στατικός Έλεγχος Ποιότητας αναπτύχθηκε ραγδαία τις τελευταίες δεκαετίες και η χρήση των μεθόδων του επεκτάθηκε σχεδόν σε όλο το φάσμα της βιομηχανίας. Σε αυτό, βέβαια, βοήθησε και η ανάπτυξη της τεχνολογίας. Όπως αναφέρει και ο Ishikawa (1989), η καλή ποιότητα δεν είναι δυνατό να επιτευχθεί χωρίς την κατάλληλη τεχνολογία.

1.3 Φιλοσοφία της Ολικής Διαχείρισης Ποιότητας

Παρόλο που οι στατιστικές τεχνικές αποτελούν στατιστικά εργαλεία κρίσιμης σημασίας, για να χρησιμοποιούνται πιο αποτελεσματικά θα πρέπει να εφαρμοστούν σε ένα πλαίσιο διαχείρισης και να είναι μέρος του πλαισίου αυτού. Η φιλοσοφία της βελτίωσης της ποιότητας και η εφαρμογή της σε όλους τους τομείς των επιχειρήσεων συνιστούν την Ολική Διαχείριση Ποιότητας (*Total Quality Management, TQM*).

Πολλοί ήταν οι άνθρωποι που συνέβαλαν στην ανάπτυξη στατιστικών μεθόδων για τη βελτίωση της ποιότητας, όμως τρεις είναι αυτοί που θεωρούνται «πατέρες» της φιλοσοφίας και εφαρμογής της ποιότητας: ο W.E. Deming, ο J.M. Duran και ο A.V. Feigenbaum.

Ο W.E. Deming θεωρεί ότι η επίτευξη της ποιότητας εξαρτάται σε μεγάλο βαθμό από τις ενέργειες της διαχείρισής της και όχι τόσο από το εργατικό δυναμικό. Η φιλοσοφία του αυτή βασίζεται σε δεκατέσσερα σημεία, τα οποία αναφέρονται παρακάτω:

- Συνεχής βελτίωση στην εμφάνιση και το σχεδιασμό του προϊόντος, η οποία επιτυγχάνεται μέσω της επένδυσης στην έρευνα, στην ανάπτυξη και στην καινοτομία. Με τον τρόπο αυτό, η εταιρία ή η επιχείρηση θα έχει μακροπρόθεσμα οφέλη.
- Υιοθέτηση νέας φιλοσοφίας που θα απορρίπτει τα ελαττωματικά προϊόντα, η παραγωγή των οποίων μπορεί να κοστίζει περισσότερο στην εταιρία από ότι η παραγωγή μη ελαττωματικών προϊόντων.
- Ο έλεγχος της ποιότητας των προϊόντων δε θα πρέπει να βασίζεται στη μαζική επιθεώρηση, διότι όταν διαπιστωθεί ότι υπάρχει κάποιο ελαττωματικό προϊόν, η επιχείρηση έχει ήδη ξοδέψει κάποιο ποσό για την παραγωγή του.
- Συνεχής βελτίωση της παραγωγής, η οποία επιτυγχάνεται με τη χρήση στατιστικών μεθόδων και στατιστικών εργαλείων.
- Πρακτική εφαρμογή νέων μεθόδων εκπαίδευσης των εργαζομένων.
- Πρακτική εφαρμογή μοντέρνων μεθόδων επίβλεψης.
- Εξάλειψη του φόβου που αισθάνονται οι εργαζόμενοι, ώστε να νιώθουν άνετα να θέτουν ερωτήσεις και να αναφέρουν τυχόν προβλήματα που προκύπτουν και

σχετίζονται με την ποιότητα και την παραγωγή. Με τον τρόπο αυτό, η επιχείρηση δε θα ζημιωθεί οικονομικά.

- Για να μπορέσει να υπάρξει αποτελεσματική βελτίωση της ποιότητας και της παραγωγικότητας, θα πρέπει να υπάρχει ομαδική συνεργασία μεταξύ διαφορετικών επιχειρησιακών μονάδων.
- Οι επιχειρήσεις θα πρέπει να αποφεύγουν να θέτουν στόχους και να έχουν συνθήματα (*slogans*), διότι αυτά είναι συνήθως αντιπαραγωγικά.
- Εξάλειψη αριθμητικών ποσοστών και τυποποιημένων εργασιακών προτύπων, τα οποία έχουν καθοριστεί ιστορικά και δεν έχουν λάβει υπ' όψιν τους την ποιότητα.
- Θα πρέπει να ξεπεραστούν τα εμπόδια που αποθαρρύνουν τους εργαζομένους από το να κάνουν τη δουλειά τους και ταυτόχρονα θα πρέπει οι εργοδότες να ακούνε τις προτάσεις, τα σχόλια και τα παράπονα των εργαζομένων τους.
- Εφαρμογή εκπαιδευτικών προγραμμάτων, ώστε όλοι οι εργαζόμενοι να μάθουν να χρησιμοποιούν στατιστικά εργαλεία και κυρίως τα διαγράμματα ελέγχου.
- Δημιουργία δομής σε θέματα διαχείρισης, η οποία θα υποστηρίζει σθεναρά τα πρώτα δεκατρία σημεία.

Από τα παραπάνω σημεία της φιλοσοφίας του Deming παρατηρούμε ότι ο ίδιος δίνει ιδιαίτερη έμφαση στην αλλαγή. Επιπλέον, γίνεται εμφανές ότι είναι κυρίαρχης σημασίας ο ρόλος της διαχείρισης που καθοδηγεί αυτή τη διαδικασία αλλαγής. Επίσης, θεωρεί ότι η χρήση στατιστικών μεθόδων, όπως οι πειραματικοί σχεδιασμοί και τα διαγράμματα ελέγχου, θα συνεισφέρουν στη βελτίωση της παραγωγικής διαδικασίας.

Ο J.M. Duran εστιάζει λιγότερο στις στατιστικές μεθόδους σε σχέση με τον Deming. Θεωρεί ότι η αλλαγή και η βελτίωση μπορεί να επέλθει μέσω του διευθυντικού επιτεύγματος (*managerial breakthrough*), όπως το αποκαλεί ο ίδιος. Δηλαδή συμμερίζεται την ίδια άποψη με τον Deming, ότι η ποιότητα μπορεί να βελτιωθεί σε αρκετά μεγάλο βαθμό (σε ποσοστό 80%) με την κατάλληλη διαχείριση.

Από την άλλη πλευρά, ο A.V. Feigenbaum πιστεύει ότι, για να βελτιωθεί η ποιότητα, πρέπει να εστιάσουμε περισσότερο σε οργανωτικές δομές ή συστήματα και λιγότερο σε

στατιστικές μεθόδους. Η φιλοσοφία του είναι αντίθετη με την πιο μοντέρνα άποψη, σύμφωνα με την οποία πρέπει να διαδοθεί ευρέως η γνώση και η χρήση στατιστικών εργαλείων.

Οι φιλοσοφίες των Deming, Duran και Feigenbaum παρουσιάζουν κάποιες διαφορές, αλλά έχουν και αρκετές ομοιότητες μεταξύ τους. Συνοψίζοντας, τόσο οι στατιστικές μέθοδοι και τεχνικές όσο και η διαχείριση μπορούν να μετατρέψουν την ποιότητα σε ένα ανταγωνιστικό εργαλείο μεταξύ των επιχειρήσεων (Montgomery (1997)).

1.4 Διαγράμματα Ελέγχου και Μεταβλητότητα

Σε κάθε παραγωγική διαδικασία, ανεξάρτητα από το πόσο καλά σχεδιασμένη είναι και από το πόσο προσεκτικά συντηρείται, παρατηρούνται συσσωρευτικές επιδράσεις που προκαλούνται από πολλές μικρές και ουσιαστικά αναπόφευκτες αιτίες, γνωστές ως **κοινές ή τυχαίες αιτίες μεταβλητότητας** (*common or chance causes of variation*)¹. Οι επιδράσεις αυτές συνήθως αναφέρονται με τον όρο **φυσική μεταβλητότητα** (*natural variability*) και θεωρούνται αναπόσπαστο μέρος της διαδικασίας. Μια διεργασία λέμε ότι είναι **εντός στατιστικού ελέγχου** (*in statistical control*) ή ότι λειτουργεί σε **ευσταθή κατάσταση** (*stable state*), όταν σε αυτή είναι παρούσα μόνο η φυσική μεταβλητότητα.

Όμως, σε μια παραγωγική διαδικασία, εκτός της φυσικής μεταβλητότητας, μπορεί να εμφανιστούν και άλλες μορφές μεταβλητότητας, οι οποίες επηρεάζουν την ποιότητα του χαρακτηριστικού. Οι αιτίες που προκαλούν αυτές τις μορφές μεταβλητότητας μπορεί να είναι οι ακατάλληλα προσαρμοσμένες-ρυθμισμένες μηχανές, τα σφάλματα των χειριστών μηχανημάτων, η ελαττωματική πρώτη ύλη κ.τ.λ.. Αυτού του είδους η μεταβλητότητα είναι κατά πολύ μεγαλύτερη σε σχέση με τη φυσική μεταβλητότητα και συνήθως αντιπροσωπεύει ένα μη αποδεκτό επίπεδο της απόδοσης της διεργασίας. Η μεταβλητότητα αυτή αναφέρεται ως **ειδική μεταβλητότητα** (*special variability*) και οι αιτίες που την προκαλούν είναι γνωστές ως **ειδικές ή συστηματικές αιτίες μεταβλητότητας** (*special or assignable causes*

¹ Ο όρος **τυχαίες αιτίες μεταβλητότητας** (*chance causes of variation*) προτάθηκε από τον Dr. Walter A. Shewhart. Σήμερα, στη βιβλιογραφία, αντί του όρου αυτού μπορεί να συναντήσουμε τον όρο **κοινές αιτίες μεταβλητότητας** (*common causes of variation*).

of variation)¹. Μια διεργασία λέμε ότι είναι **εκτός στατιστικού ελέγχου** (*out of statistical control*) ή ότι λειτουργεί σε **ασταθή κατάσταση** (*unstable state*), όταν σε αυτή υπάρχουν και συστηματικές αιτίες μεταβλητότητας.

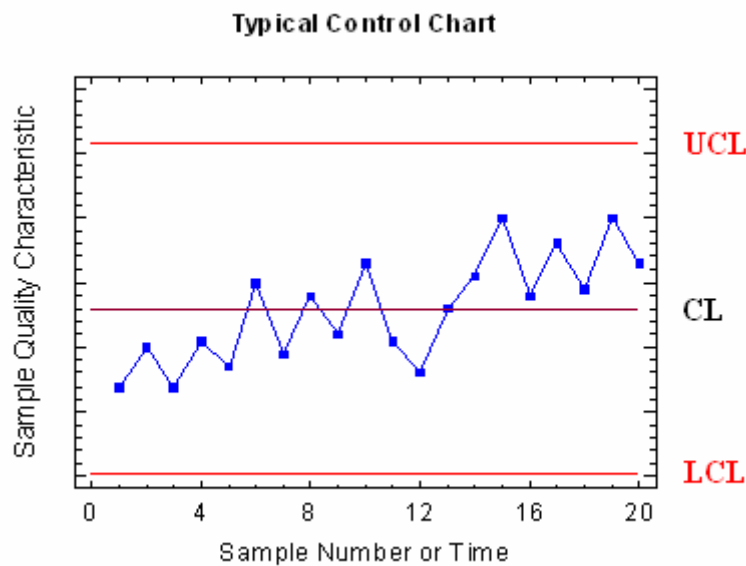
Ο κύριος στόχος του Στατιστικού Ελέγχου Διεργασιών είναι η άμεση ανίχνευση συστηματικών αιτιών μεταβλητότητας σε μια διεργασία έτσι, ώστε να πραγματοποιηθούν απαραίτητες διορθωτικές ενέργειες πριν κατασκευαστεί μεγάλος αριθμός προϊόντων που δεν είναι συμμορφωμένα (*nonconforming*) με τις προδιαγραφές. Για την επίτευξη του στόχου αυτού, η πιο διαδεδομένη τεχνική που χρησιμοποιείται είναι τα Διαγράμματα Ελέγχου Διεργασιών, τα οποία μπορούν να ανιχνεύσουν σε πραγματικό χρόνο τις ειδικές αιτίες μεταβλητότητας (*on-line process-monitoring*). Επίσης, μπορούν να εκτιμήσουν τις παραμέτρους μιας παραγωγικής διαδικασίας, να προσδιορίσουν την ικανότητα μιας διεργασίας και να παρέχουν χρήσιμες πληροφορίες για τη βελτίωσή της. Συνεπώς, τα Διαγράμματα Ελέγχου αποτελούν, εάν όχι το σημαντικότερο εργαλείο, ένα από τα πιο χρήσιμα εργαλεία του Στατιστικού Ελέγχου Διεργασιών, διότι μειώνουν τη μεταβλητότητα όσο το δυνατόν περισσότερο, δεδομένου ότι η αυτή δεν μπορεί να εξαλειφθεί τελείως.

1.5 Περιγραφή ενός Τυπικού Διαγράμματος Ελέγχου

Αυτό που μας ενδιαφέρει σε μια παραγωγική διαδικασία είναι να παρακολουθήσουμε τη συμπεριφορά μιας κρίσιμης ποσότητας, η οποία χαρακτηρίζει την ποιότητα ενός προϊόντος. Η ποσότητα αυτή αναφέρεται συνήθως σε ένα (μετρήσιμο) χαρακτηριστικό X των παραγόμενων προϊόντων, όπου X είναι τυχαία μεταβλητή. Για να μπορέσουμε να παρακολουθήσουμε τη διαδικασία, βασιζόμαστε σε μετρήσεις του χαρακτηριστικού X , οι οποίες προκύπτουν επιλέγοντας τυχαία δείγματα προϊόντων από την παραγωγή σε διαφορετικές χρονικές στιγμές. Έστω $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ τυχαία δείγματα τιμών του χαρακτηριστικού X , τα οποία αντιστοιχούν στα δείγματα προϊόντων που επιλέξαμε τυχαία. Με βάση τα τυχαία δείγματα $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ υπολογίζουμε την τιμή $W_i = g(\mathbf{X}_i)$,

¹ Ο όρος **συστηματικές αιτίες μεταβλητότητας** (*assignable causes of variation*) προτάθηκε κι αυτός από τον Dr. Walter A. Shewhart. Σήμερα, στη βιβλιογραφία, αντί του όρου αυτού μπορεί να συναντήσουμε τον όρο **ειδικές αιτίες μεταβλητότητας** (*special causes of variation*).

$i=1, 2, \dots, n$ μιας κατάλληλης στατιστικής συνάρτησης (τυχαίας μεταβλητής). Αυτή η στατιστική συνάρτηση είναι συνήθως αμερόληπτη εκτιμήτρια της κρίσιμης ποσότητας για την οποία ενδιαφερόμαστε. Συνεπώς, με την παρακολούθηση των τιμών της στατιστικής συνάρτησης W στα διάφορα δείγματα, επιτυγχάνουμε τη (διαχρονική) παρακολούθηση της συμπεριφοράς της κρίσιμης ποσότητας.



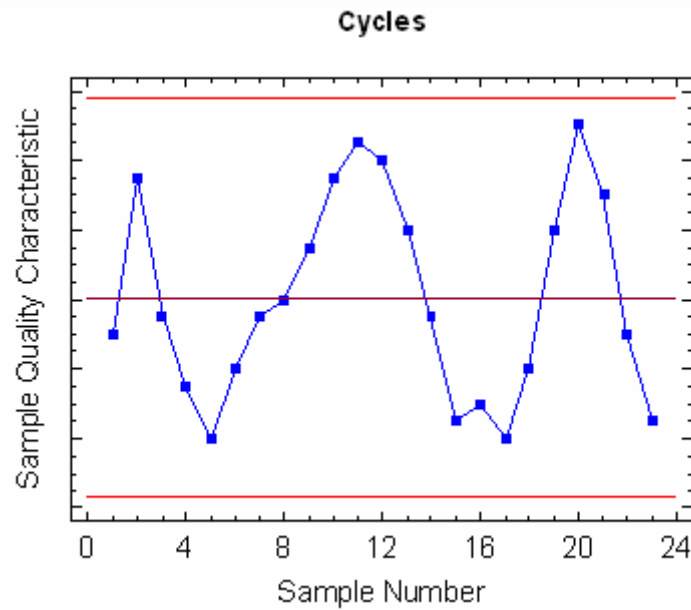
Σχήμα 1.3: Τυπικό Διάγραμμα Ελέγχου

Ένα τυπικό διάγραμμα ελέγχου είναι μία δισδιάστατη γραφική παράσταση της εξέλιξης μιας διεργασίας στην πάροδο του χρόνου. Στο Σχήμα 1.3, μπορούμε να δούμε τη μορφή που έχει ένα διάγραμμα ελέγχου. Ο οριζόντιος άξονας του γραφήματος αυτού αντιπροσωπεύει είτε το χρόνο είτε τα δείγματα προϊόντων που επιλέχθηκαν τυχαία από την παραγωγή (σε διαφορετικές χρονικές στιγμές). Τα δείγματα αυτά τοποθετούνται σε χρονολογική σειρά έτσι, ώστε αριστερά να εμφανίζεται το πρώτο δείγμα που λαμβάνεται και κάθε νέο αποκτηθέν δείγμα να εμφανίζεται προς τα δεξιά. Ο κάθετος άξονας αντιπροσωπεύει την τιμή της στατιστικής συνάρτησης W . Στο γράφημα παρατηρούμε ότι οι τιμές της W_i , $i=1, 2, \dots, n$ απεικονίζονται με σημεία, τα οποία είναι ενωμένα με μια τεθλασμένη γραμμή. Ένα τυπικό διάγραμμα ελέγχου περιλαμβάνει, επίσης, τρεις παράλληλες οριζόντιες γραμμές: μία κεντρική γραμμή και δύο όρια ελέγχου. Η **κεντρική γραμμή** (*Center Line*, *CL*)

αντιπροσωπεύει το επίπεδο της διεργασίας υπό σταθερές συνθήκες, δηλαδή όταν στη διεργασία υπάρχει μόνο φυσική μεταβλητότητα. Συνήθως, το επίπεδο της διεργασίας παριστάνει το μέσο όρο της στατιστικής συνάρτησης. Οι δύο γραμμές πάνω και κάτω από την κεντρική γραμμή ονομάζονται **άνω όριο ελέγχου** (*Upper Control Limit, UCL*) και **κάτω όριο ελέγχου** (*Lower Control Limit, LCL*) αντίστοιχα, και δηλώνουν το φυσικό εύρος της διακύμανσης της στατιστικής συνάρτησης. Τα όρια ελέγχου βρίσκονται συνήθως σε απόσταση τριών τυπικών αποκλίσεων από την κεντρική γραμμή και στις δύο πλευρές.

Εάν τα σημεία, δηλαδή οι τιμές της στατιστικής συνάρτησης, συμπεριφέρονται με «τυχαίο» τρόπο και βρίσκονται εντός των ορίων ελέγχου, τότε λέμε ότι η διεργασία είναι εντός στατιστικού ελέγχου. Συνεπώς, δε χρειάζεται να προβούμε σε διορθωτικές κινήσεις. Ωστόσο, εάν κάποιο σημείο βρεθεί εκτός των ορίων, τότε αυτό αποτελεί ένδειξη ότι η διεργασία είναι εκτός στατιστικού ελέγχου και αντιμετωπίζουμε **κατάσταση συναγερμού** (*alarm*). Στην περίπτωση αυτή πρέπει να ανακαλύψουμε τις ειδικές αιτίες μεταβλητότητας, οι οποίες δημιούργησαν την εκτός ελέγχου κατάσταση, και στη συνέχεια να προβούμε σε διορθωτικές κινήσεις, εφόσον αυτές κριθούν απαραίτητες. Στο σημείο αυτό, θα πρέπει να σημειώσουμε ότι μια διεργασία μπορεί να είναι εκτός στατιστικού ελέγχου ακόμη κι αν όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου, αλλά συμπεριφέρονται με **συστηματικό ή μη τυχαίο τρόπο**. Οι βασικότερες μορφές προτύπων ή μοτίβων δίνουν ενδείξεις για μη τυχαία συμπεριφορά είναι τα κυκλικά ή περιοδικά πρότυπα, τα μικτά πρότυπα, τα πρότυπα αλμάτων, τα πρότυπα τάσης και τα πρότυπα έλλειψης μεταβλητότητας, τα οποία περιγράφονται παρακάτω.

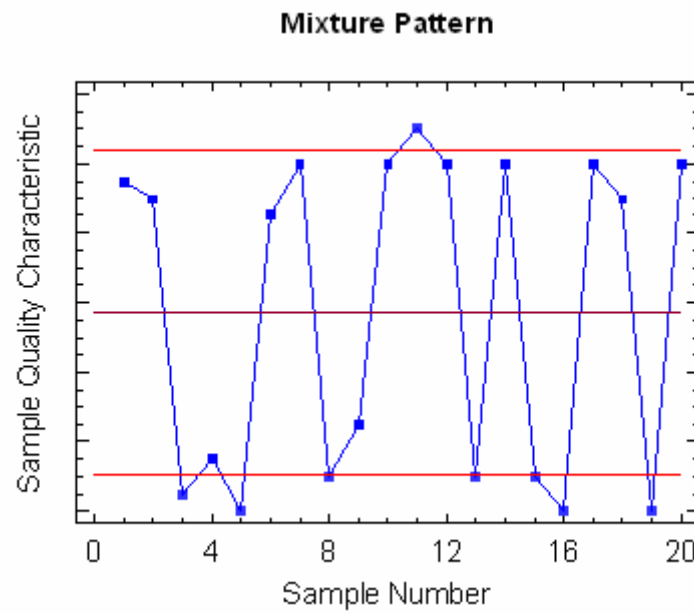
Σε ένα **κυκλικό πρότυπο** (*cycles*), η διάταξη των σημείων είναι περιοδική και αυτό έχει ως αποτέλεσμα τον επαναλαμβανόμενο σχηματισμό κύκλων. Ένα τέτοιο πρότυπο βλέπουμε στο Σχήμα 1.4. Οι περιοδικές-συστηματικές αλλαγές των συνθηκών περιβάλλοντος ή παραγωγής, όπως η θερμοκρασία, η διακύμανση της πίεσης ή της τάσης, η κόπωση των χειριστών και η εναλλαγή των μηχανημάτων, θεωρούνται μερικές από τις αιτίες εμφάνισης κυκλικών προτύπων σε ένα διάγραμμα ελέγχου.



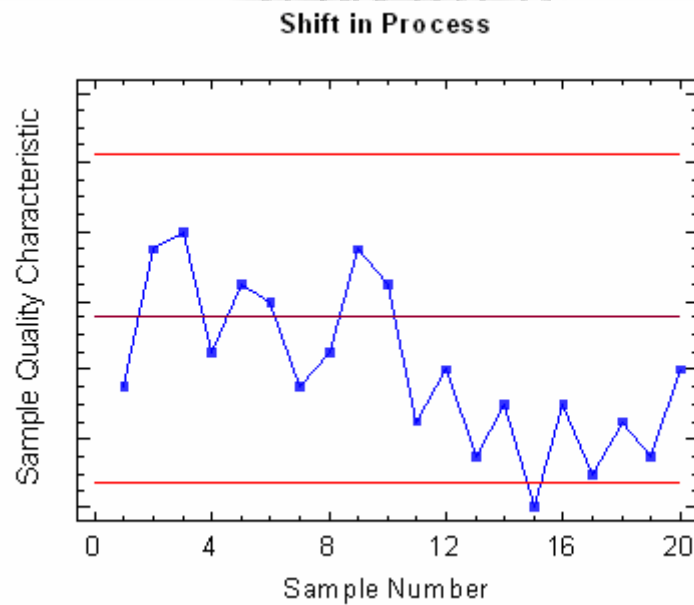
Σχήμα 1.4: Κυκλικό ή περιοδικό πρότυπο

Σε ένα **μικτό πρότυπο** (*mixture pattern*), τα σημεία έχουν την τάση να είναι είτε κοντά στα όρια ελέγχου είτε οριακά έξω από αυτά, ενώ σχετικά λίγα βρίσκονται κοντά στην κεντρική γραμμή. Όπως μπορούμε να παρατηρήσουμε και στο Σχήμα 1.5, τα σημεία ενός τέτοιου προτύπου φαίνεται να προέρχονται από δύο (ή καμιά φορά και περισσότερες από δύο) διαφορετικές κατανομές του χαρακτηριστικού X , μεταξύ των οποίων παλινδρομεί η διεργασία. Μια πιθανή αιτία μπορεί να είναι ο υπερβολικός έλεγχος (*overcontrol*). Αρκετά συχνά οι χειριστές προβαίνουν σε προσαρμογές της διαδικασίας που ανταποκρίνονται σε τυχαία διακύμανση και όχι σε συστηματικές αιτίες. Άλλη μια πιθανή αιτία μπορεί να θεωρηθεί το γεγονός ότι για την παρακολούθηση μιας διεργασίας συλλέγονται δείγματα προϊόντων, τα οποία προέρχονται από διαφορετικές πηγές (παράλληλες μηχανές).

Σε ένα **πρότυπο αλμάτων** (*shifts in process level*), παρατηρείται μετατόπιση του μέσου του χαρακτηριστικού X από μια χρονική στιγμή και μετά. Ένα πρότυπο αλμάτων απεικονίζεται στο Σχήμα 1.6 και μπορεί να προκύψει ύστερα από την εισαγωγή νέων εργαζομένων, νέων μεθόδων παραγωγής, νέων μηχανημάτων ή ακόμη και μετά από κάποια αλλαγή που οφείλεται στις ικανότητες, την προσοχή ή το κίνητρο των εργαζομένων.



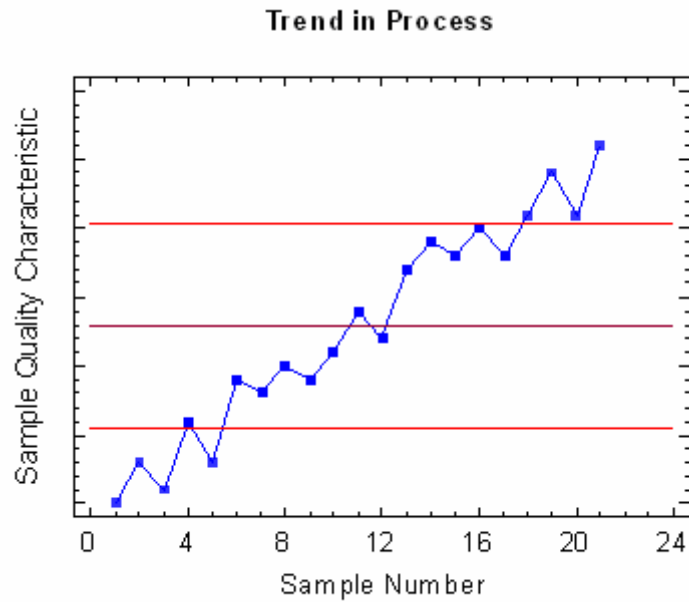
Σχήμα 1.5: Μικτό πρότυπο



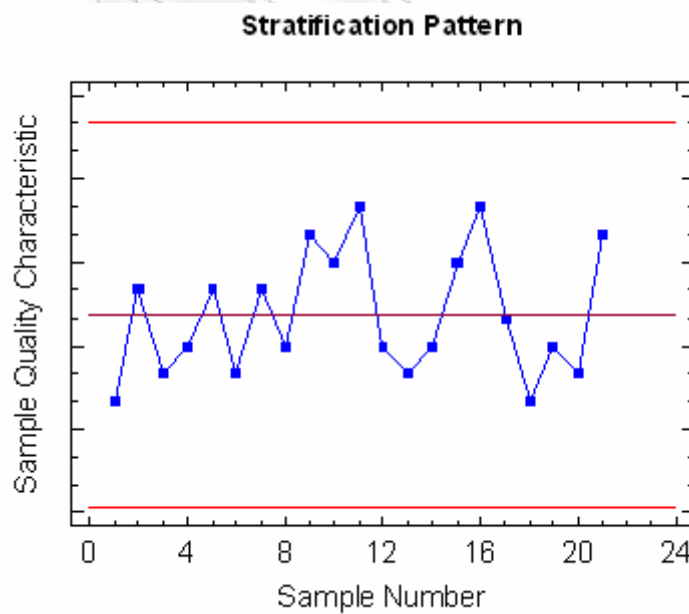
Σχήμα 1.6: Πρότυπο αλμάτων

Σε ένα **πρότυπο τάσης** (*trend in process level*), τα σημεία φαίνεται να παρουσιάζουν μια συνεχή κίνηση προς μία κατεύθυνση. Αυτό γίνεται φανερό στο Σχήμα 1.7, όπου τα σημεία

κινούνται προς τα πάνω. Πιθανές αιτίες μπορούν να θεωρηθούν η σταδιακή φθορά εργαλείων, η κόπωση των εργαζομένων ή ακόμη και η παρουσία επιθεωρητή.



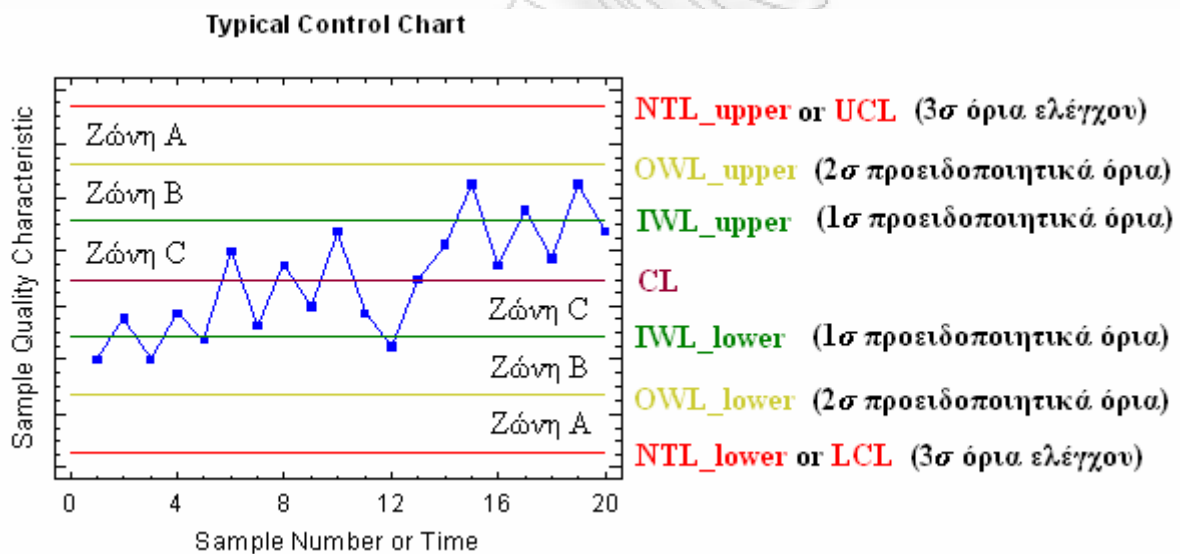
Σχήμα 1.7: Πρότυπο τάσης



Σχήμα 1.8: Πρότυπο έλλειψης μεταβλητότητας

Τέλος, σε ένα **πρότυπο έλλειψης μεταβλητότητας** (*Stratification Pattern*), τα σημεία φαίνεται να κινούνται με έναν τεχνητό τρόπο γύρω από την κεντρική γραμμή, λόγω έλλειψης φυσικής μεταβλητότητας της διεργασίας. Ένα πρότυπο σαν κι αυτό μπορούμε να δούμε στο Σχήμα 1.8. Η πιο πιθανή αιτία σχηματισμού προτύπου έλλειψης μεταβλητότητας είναι η λανθασμένη σχεδίαση των ορίων ελέγχου.

Όπως είδαμε παραπάνω, υπάρχουν αρκετές περιπτώσεις στις οποίες τα σημεία ενός διαγράμματος ελέγχου εμφανίζουν **ειδικές ακολουθίες** (*patterns*). Σκοπός μας είναι να δημιουργήσουμε ένα πιο ευαίσθητο διάγραμμα ελέγχου, το οποίο να έχει την ικανότητα να ανιχνεύει γρηγορότερα τις εκτός ελέγχου διεργασίες. Αυτό μπορεί να γίνει σχεδιάζοντας εκτός από τα όρια ελέγχου και τα λεγόμενα προειδοποιητικά όρια (*warning limits*), τα οποία βρίσκονται εσωτερικά των ορίων ελέγχου. Παρατηρώντας το Σχήμα 1.9 που ακολουθεί καταλαβαίνουμε πως ακριβώς σχεδιάζεται ένα ευαίσθητο διάγραμμα ελέγχου.



Σχήμα 1.9: Διάγραμμα Ελέγχου με Προειδοποιητικά Όρια Ελέγχου

Τα όρια ελέγχου που βρίσκονται σε απόσταση 1σ από την κεντρική γραμμή του διαγράμματος ελέγχου ονομάζονται **εσωτερικά προειδοποιητικά όρια** (*Inner Warning Limits, IWL*), ενώ τα όρια που βρίσκονται σε απόσταση 2σ ονομάζονται **εξωτερικά προειδοποιητικά όρια** (*Outer Warning Limits, OWL*). Επίσης, τα όρια που βρίσκονται σε

απόσταση 3σ από την κεντρική γραμμή ονομάζονται **φυσικά όρια ανοχής** της διεργασίας (*Natural Tolerance Limits, NTL*).

Τα προειδοποιητικά όρια είναι χρήσιμα στη δημιουργία κανόνων, οι οποίοι εντοπίζουν τις ειδικές ακολουθίες που μπορεί να εμφανίσουν τα σημεία ενός διαγράμματος ελέγχου. Συνεπώς, δεν είναι απαραίτητο να υπάρχει κάποιο σημείο εκτός των ορίων ελέγχου (*UCL* και *LCL*) για να θεωρηθεί μια διεργασία εκτός στατιστικού ελέγχου.

Οι πιο σημαντικοί κανόνες ευαισθητοποίησης ενός διαγράμματος ελέγχου, όπως αυτοί αναφέρονται από τον Montgomery (1997), είναι οι εξής:

- Ένα ή περισσότερα σημεία εκτός των ορίων ελέγχου.
- Δύο από τρία συνεχόμενα σημεία στην Ζώνη A (σε μια από τις δύο ζώνες A).
- Τέσσερα από πέντε συνεχόμενα σημεία πέραν της Ζώνης C (σε μια από τις δύο περιοχές).
- Οκτώ συνεχόμενα σημεία στην ίδια μεριά (επάνω ή κάτω) της κεντρικής γραμμής.
- Έξι συνεχόμενα σημεία σε αύξουσα ή φθίνουσα διάταξη.
- Δεκαπέντε συνεχόμενα σημεία στην ολική Ζώνη C.
- Δεκατέσσερα συνεχόμενα σημεία σε εναλλασσόμενη μορφή «πάνω-κάτω».
- Οκτώ συνεχόμενα σημεία εκτός της ολικής Ζώνης C.
- Οποιαδήποτε ασυνήθιστη ή μη τυχαία ακολουθία σημείων.
- Ένα ή περισσότερα σημεία κοντά στα προειδοποιητικά όρια ή τα όρια ελέγχου.

Οι τέσσερις πρώτοι κανόνες είναι γνωστοί ως Western Electric Rules (Western Electric (1956)). Όπως έχουμε ήδη αναφέρει, οι παραπάνω κανόνες μπορούν να φανούν χρήσιμοι ώστε να ανιχνεύσουμε γρηγορότερα τις εκτός ελέγχου διεργασίες. Ωστόσο, η ταυτόχρονη χρήση πολλών από αυτούς τους κανόνες μπορεί να προκαλέσει λανθασμένους συναγερμούς. Ως αποτέλεσμα, όσους περισσότερους λανθασμένους συναγερμούς έχουμε, τόσες περισσότερες λανθασμένες διακοπές της παραγωγικής διαδικασίας θα έχουμε, οι οποίες θα γίνονται με σκοπό να βρεθούν οι ειδικές αιτίες που προκάλεσαν την εκτός ελέγχου κατάσταση. Τελικά, κάτι τέτοιο θα επιφέρει αύξηση του κόστους της παραγωγής.

1.6 Είδη Διαγραμμάτων Ελέγχου

Υπάρχουν πολλές κατηγορίες διαγραμμάτων ελέγχου. Η διάκρισή τους βασίζεται σε ορισμένα χαρακτηριστικά, όπως το είδος (συνεχές ή διακριτό) και το πλήθος των ποιοτικών χαρακτηριστικών, η κατανομή των μετρήσεων του χαρακτηριστικού, η εξάρτηση των μετρήσεων μεταξύ διαδοχικών χρονικών στιγμών, το μέγεθος δείγματος, καθώς και η κρίσιμη ποσότητα που παρακολουθούμε. Δύο επιπλέον κατηγορίες διαγραμμάτων ελέγχου μπορούν να προκύψουν λαμβάνοντας υπ' όψιν μας τις πληροφορίες που δίνει το δείγμα στο οποίο βασίζεται ο σχεδιασμός ενός σημείου του διαγράμματος ελέγχου.

Εάν το ποιοτικό χαρακτηριστικό X είναι συνεχές, δηλαδή περιγράφεται από μία συνεχή τυχαία μεταβλητή, τότε αναφερόμαστε σε **διαγράμματα ελέγχου για μεταβλητές** (*control charts for variables*). Ενώ, εάν το χαρακτηριστικό X περιγράφεται από μια διακριτή τυχαία μεταβλητή, τότε αναφερόμαστε σε **διαγράμματα ελέγχου για ιδιότητες** (*control charts for attributes*).

Εάν οι μετρήσεις που λαμβάνονται αφορούν ένα ποιοτικό χαρακτηριστικό ($p=1$), τότε αναφερόμαστε σε **μονομεταβλητά διαγράμματα ελέγχου** (*univariate control charts*). Αν οι μετρήσεις αφορούν περισσότερα χαρακτηριστικά ($p>1$), τότε αναφερόμαστε σε **πολυμεταβλητά διαγράμματα ελέγχου** (*multivariate control charts*).

Όταν οι μετρήσεις που λαμβάνονται ακολουθούν κάποια γνωστή κατανομή, τότε αναφερόμαστε σε **παραμετρικά διαγράμματα ελέγχου** (*parametric control charts*). Αντίθετα, εάν δε γνωρίζουμε ποια κατανομή ακολουθούν οι μετρήσεις που λαμβάνονται, τότε αναφερόμαστε σε **μη παραμετρικά διαγράμματα ελέγχου** (*nonparametric control charts*).

Στην περίπτωση που οι μετρήσεις που λαμβάνονται σε κάθε χρονική στιγμή t είναι εξαρτημένες από τις μετρήσεις που λαμβάνονται στο χρόνο $t-1$, τότε αναφερόμαστε σε **διαγράμματα ελέγχου για αυτοσυσχετιζόμενες διεργασίες** (*control charts for autocorrelated processes*). Αν οι μετρήσεις είναι ανεξάρτητες, τότε αναφερόμαστε σε **διαγράμματα ελέγχου για ασυσχετίστες διεργασίες** (*control charts for uncorrelated processes*).

Τέλος, αν από την παραγωγική διαδικασία λαμβάνουμε μία μέτρηση (παρατήρηση) σε κάθε χρονική στιγμή, τότε αναφερόμαστε σε **διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις** (*control charts for individual observations*). Ενώ, εάν τα δείγματα μετρήσεων

που λαμβάνονται από την παραγωγική διαδικασία είναι μεγαλύτερα της μονάδας ($n > 1$), τότε αναφερόμαστε σε **διαγράμματα ελέγχου για δείγματα ή ομάδες** (*control charts for rational subgroups*).

Τα διαγράμματα ελέγχου για μεταβλητές περιλαμβάνουν **διαγράμματα ελέγχου για τη μέση τιμή** (*control charts for process mean*) και **διαγράμματα ελέγχου για τη διασπορά** (*control charts for process dispersion or variability*). Τα διαγράμματα ελέγχου για ιδιότητες μπορούν να είναι **διαγράμματα ελέγχου για το ποσοστό των ελαττωματικών προϊόντων** (*control charts for the proportion of nonconforming units or control charts for fraction nonconforming, p-chart*), **διαγράμματα ελέγχου για τον αριθμό των ελαττωματικών προϊόντων** (*control charts for the number of nonconforming units, np-chart*), **διαγράμματα ελέγχου για τον αριθμό των ελαττωμάτων** (*counts control charts for the total number of nonconformities per unit or control charts for nonconformities, c-chart*) και **διαγράμματα ελέγχου για το μέσο αριθμό ελαττωμάτων** (*control charts for the average number of nonconformities per unit or control charts for nonconformities per unit, u-chart*).

Εάν τα σημεία που σχεδιάζονται σε ένα διάγραμμα ελέγχου βασίζονται μόνο σε πληροφορίες που δίνει μόνο το πιο πρόσφατο δείγμα, αγνοώντας τις πληροφορίες που περιέχονται σε προηγούμενα δείγματα, τότε αναφερόμαστε σε **διαγράμματα ελέγχου χωρίς μνήμη** (*control charts without memory*). Αντίθετα, εάν τα σημεία που σχεδιάζονται σε ένα διάγραμμα ελέγχου βασίζονται τόσο σε πληροφορίες που δίνει το πιο πρόσφατο δείγμα, όσο και σε πληροφορίες που προέρχονται από προγενέστερα δείγματα, αναφερόμαστε σε **διαγράμματα ελέγχου με μνήμη** (*control charts with memory*).

Στην πρώτη κατηγορία ταξινομούνται τα **διαγράμματα ελέγχου τύπου Shewhart** (*Shewhart type control charts*), τα οποία χρησιμοποιούνται κυρίως όταν οι αλλαγές του μέσου επιπέδου που ανιχνεύουμε είναι μεγάλες, όταν το μέγεθος δείγματος είναι τουλάχιστον δύο και όταν είναι γνωστή η κατανομή πιθανότητας των αρχικών παρατηρήσεων ή του δειγματικού χαρακτηριστικού που απεικονίζεται το διάγραμμα. Τα διαγράμματα Shewhart δεν είναι κατάλληλα για μικρές μετατοπίσεις του μέσου, είναι πιο αποτελεσματικά για μετατοπίσεις του μέσου της τάξης μεταξύ 1.5σ και 2σ ή μεγαλύτερες. Τα διαγράμματα αυτά παρουσιάστηκαν για πρώτη φορά από τον Shewhart το 1931.

Στη δεύτερη κατηγορία ανήκουν τα **αθροιστικά διαγράμματα ελέγχου ή διαγράμματα ελέγχου τύπου CUSUM** (*Cumulative Sum, CUSUM type control charts*) και τα **διαγράμματα**

ελέγχου για τον εκθετικά ζυγισμένο κινούμενο μέσο ή διαγράμματα ελέγχου τύπου *EWMA* (*Exponentially Weighted Moving Average, EWMA type control charts*), οποία χρησιμοποιούνται κυρίως όταν οι αλλαγές του μέσου επιπέδου που ανιχνεύουμε είναι μικρές. Τα διαγράμματα *CUSUM* είναι χρήσιμα όταν είναι γνωστή η κατανομή πιθανότητας των αρχικών παρατηρήσεων ή του δειγματικού χαρακτηριστικού που απεικονίζεται το διάγραμμα. Όταν είναι άγνωστη η κατανομή πιθανότητας των αρχικών παρατηρήσεων ή του δειγματικού χαρακτηριστικού που απεικονίζεται το διάγραμμα, τότε τα πλέον κατάλληλα διαγράμματα είναι τα *EWMA*. Τα διαγράμματα ελέγχου με μνήμη αναπτύχθηκαν πριν από πενήντα περίπου χρόνια και χρησιμοποιούνται κυρίως στη Φάση II.

1.7 Βασικές Έννοιες σχετικά με τα Διαγράμματα Ελέγχου

1.7.1 Μοντέλο Ορίων L Σίγμα και Μοντέλο Ορίων Πιθανότητας a

Το γενικό μοντέλο που χρησιμοποιείται για την κατασκευή ενός διαγράμματος ελέγχου καλείται **μοντέλο ορίων σίγμα** (*sigma limits model*), γιατί χρησιμοποιεί ως όρια ελέγχου που προκύπτουν απομακρυνόμενοι από το μέσο κατά ένα πολλαπλάσιο της τυπικής απόκλισης (L σίγμα). Πιο συγκεκριμένα το LCL και το UCL του μοντέλου καθορίζονται ως εξής

$$\text{Μοντέλο Ορίων } L \text{ Σίγμα: } \begin{cases} LCL & = \mu_w - L \cdot \sigma_w \\ CL & = \mu_w \\ UCL & = \mu_w + L \cdot \sigma_w \end{cases},$$

όπου τα σύμβολα μ_w και σ_w παριστάνουν τη μέση τιμή και την τυπική απόκλιση της στατιστικής συνάρτησης W , όταν η διεργασία είναι εντός στατιστικού ελέγχου. Όπως έχουμε ήδη αναφέρει, η στατιστική συνάρτηση W είναι μια (αμερόληπτη συνήθως) εκτιμήτρια της μέσης τιμής της ποσότητας (χαρακτηριστικού) X , η οποία απεικονίζεται στο διάγραμμα ελέγχου. Με την ποσότητα L συμβολίζουμε την απόσταση, σε μονάδες τυπικής απόκλισης πάντα, των ορίων ελέγχου από την κεντρική γραμμή.

Εκτός από το γενικό μοντέλο ορίων σίγμα, μπορούμε να χρησιμοποιήσουμε και το **μοντέλο ορίων πιθανότητας a** (*probability limits model*) για την κατασκευή ενός

διαγράμματος. Εάν η κατανομή της στατιστικής συνάρτησης W είναι κανονική το μοντέλο ορίων πιθανότητας έχει τη μορφή

$$\text{Μοντέλο Ορίων Πιθανότητας } a: \begin{cases} LCL = w_{a/2} = \mu_W - z_{a/2} \cdot \sigma_W \\ CL = w_{0,5} = \mu_W \\ UCL = w_{1-(a/2)} = \mu_W + z_{a/2} \cdot \sigma_W \end{cases},$$

όπου w_a το άνω ποσοστιαίο σημείο της στατιστικής συνάρτησης W .

Στην περίπτωση που η κατανομή της στατιστικής συνάρτησης W δεν είναι κανονική, η κανονικότητα μπορεί να εξασφαλιστεί χρησιμοποιώντας μεγάλο μέγεθος δείγματος. Με τον τρόπο αυτό και μέσω του Κεντρικού Οριακού Θεωρήματος (ΚΟΘ), η ποσότητα W ακολουθεί προσεγγιστικά την τυπική κανονική κατανομή.

1.7.2 Σφάλμα Τύπου I και Σφάλμα Τύπου II

Σε αναλογία με την έννοια του σφάλματος στον έλεγχο υποθέσεων, με τον όρο **Σφάλμα Τύπου I** (*Type I Error*) εννοούμε τον κίνδυνο ή την πιθανότητα ένα σημείο του διαγράμματος να βρεθεί εκτός των ορίων ελέγχου ενώ η διεργασία είναι εντός στατιστικού ελέγχου. Ως **Σφάλμα Τύπου II** (*Type II Error*) ορίζουμε την πιθανότητα ένα σημείο του διαγράμματος να βρεθεί εντός των ορίων ελέγχου, ενώ στην πραγματικότητα η κατανομή της X και κατ' επέκταση η κατανομή της W έχει αλλάξει. Το Σφάλμα Τύπου II εκφράζει την πιθανότητα το διάγραμμα να δίνει εσφαλμένα ένδειξη ότι η διαδικασία είναι εντός στατιστικού ελέγχου, ενώ είναι εκτός ελέγχου, λόγω μετατόπισης του μέσου επιπέδου της διεργασίας από μ_W σε μ^* .

Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι καθώς μεγαλώνει η τιμή του L , δηλαδή η απόσταση των ορίων ελέγχου από την κεντρική γραμμή, μειώνεται η πιθανότητα του Σφάλματος Τύπου I και ταυτόχρονα αυξάνεται η πιθανότητα του Σφάλματος Τύπου II. Αντίστοιχα, για μικρές τιμές του L , αυξάνεται η πιθανότητα του Σφάλματος Τύπου I και ταυτόχρονα μειώνεται η πιθανότητα του Σφάλματος Τύπου II.

Το Σφάλμα Τύπου I εμφανίζεται συχνά όταν τα όρια ενός διαγράμματος ελέγχου είναι πολύ στενά, οπότε μια τυχαία αιτία μεταβλητότητας θεωρείται συστηματικό αίτιο. Αντίθετα,

το Σφάλμα Τύπου II εμφανίζεται συχνά όταν τα όρια ενός διαγράμματος ελέγχου έχουν αρκετά μεγάλο εύρος, οπότε ένα συστηματικό αίτιο θεωρείται τυχαίο.

Το Σφάλμα Τύπου I όταν χρησιμοποιούμε όρια ελέγχου $L\sigma$, η κατανομή της στατιστικής συνάρτησης W είναι κανονική και η διεργασία είναι εντός στατιστικού ελέγχου υπολογίζεται από τον τύπο

$$\begin{aligned}
 a &= P(W \notin [LCL, UCL] | W \sim N(\mu_w, \sigma_w^2)) = P(W < LCL \text{ ή } W > UCL | W \sim N(\mu_w, \sigma_w^2)) = \\
 &= 1 - P(LCL \leq W \leq UCL | W \sim N(\mu_w, \sigma_w^2)) = \\
 &= 1 - P\left(\frac{LCL - \mu_w}{\sigma_w} \leq \frac{W - \mu_w}{\sigma_w} \leq \frac{UCL - \mu_w}{\sigma_w} \middle| \frac{W - \mu_w}{\sigma_w} \sim N(0,1)\right) = \\
 &= 1 - P\left(\frac{\mu_w - L \cdot \sigma_w - \mu_w}{\sigma_w} \leq Z \leq \frac{\mu_w + L \cdot \sigma_w - \mu_w}{\sigma_w} \middle| Z \sim N(0,1)\right) = \\
 &= 1 - P(-L \leq Z \leq L | Z \sim N(0,1)) = 1 - [\Phi(L) - \Phi(-L)] = 1 - \Phi(L) + \Phi(-L) = \\
 &= \Phi(-L) + \Phi(-L) = 2 \cdot \Phi(-L)
 \end{aligned}$$

Στην περίπτωση $L=3$ (όρια ελέγχου 3σ) παίρνουμε $a = 2 \cdot \Phi(-3) = 0.0027$, το οποίο πρακτικά σημαίνει ότι ένας λανθασμένος συναγερμός συμβαίνει 27 φορές ανά 10,000 σημεία.

Αντίστοιχα, όταν χρησιμοποιούμε όρια ελέγχου $L\sigma$, η κατανομή της στατιστικής συνάρτησης W είναι κανονική και η διεργασία είναι εκτός στατιστικού ελέγχου (λόγω μετατόπισης του μέσου από τη θέση μ_w στη θέση μ^* , δεδομένου ότι η ποσότητα σ_w^2 παραμένει σταθερή), το Σφάλμα Τύπου II υπολογίζεται από τον τύπο

$$\begin{aligned}
 \beta &= P(W \in [LCL, UCL] | W \sim N(\mu^*, \sigma_w^2)) = P(LCL \leq W \leq UCL | W \sim N(\mu^*, \sigma_w^2)) = \\
 &= P\left(\frac{LCL - \mu^*}{\sigma_w} \leq \frac{W - \mu^*}{\sigma_w} \leq \frac{UCL - \mu^*}{\sigma_w} \middle| \frac{W - \mu^*}{\sigma_w} \sim N(0,1)\right) = \\
 &= P\left(\frac{\mu_w - L \cdot \sigma_w - \mu^*}{\sigma_w} \leq Z \leq \frac{\mu_w + L \cdot \sigma_w - \mu^*}{\sigma_w} \middle| Z \sim N(0,1)\right) = \\
 &= P\left(\frac{(\mu_w - \mu^*) - L \cdot \sigma_w}{\sigma_w} \leq Z \leq \frac{(\mu_w - \mu^*) + L \cdot \sigma_w}{\sigma_w} \middle| Z \sim N(0,1)\right) = \\
 &= P\left(\frac{\mu_w - \mu^*}{\sigma_w} - \frac{L \cdot \sigma_w}{\sigma_w} \leq Z \leq \frac{\mu_w - \mu^*}{\sigma_w} + \frac{L \cdot \sigma_w}{\sigma_w} \middle| Z \sim N(0,1)\right) =
 \end{aligned}$$

$$= P\left(\frac{\mu_w - \mu^*}{\sigma_w} - L \leq Z \leq \frac{\mu_w - \mu^*}{\sigma_w} + L \mid Z \sim N(0,1)\right) = \Phi\left(\frac{\mu_w - \mu^*}{\sigma_w} + L\right) - \Phi\left(\frac{\mu_w - \mu^*}{\sigma_w} - L\right)$$

1.7.3 Μέσο Μήκος Ροής και Μέσος Χρόνος Σήματος

Το **μέσο μήκος ροής** (*Average Run Length, ARL*) είναι μία πολύ χρήσιμη έννοια που σχετίζεται με τα διαγράμματα ελέγχου και ορίζεται ως εξής

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

όπου p η πιθανότητα ένα σημείο να βρεθεί εκτός των ορίων ελέγχου. Το μέσο μήκος ροής δηλώνει τον αναμενόμενο αριθμό των σημείων (δειγμάτων) που πρέπει να τοποθετηθούν σε ένα διάγραμμα μέχρις ότου να εμφανιστεί ένα σημείο (το πρώτο σημείο) εκτός των ορίων ελέγχου. Είναι προφανές ότι το **μήκος ροής** (*Run Length*) ακολουθεί Γεωμετρική κατανομή με πιθανότητα επιτυχίας p (συμβ. $RL \sim G(p)$).

Όταν μία διεργασία είναι εντός στατιστικού ελέγχου και χρησιμοποιούμε μοντέλο ορίων $L\sigma$, το **εντός ελέγχου μέσο μήκος ροής** (*in-control Average Run Length, ARL₀*) υπολογίζεται από τον τύπο

$$ARL_0 = \frac{1}{a},$$

όπου το a υπολογίζεται από τον τύπο του Σφάλματος Τύπου I. Το εντός ελέγχου μέσο μήκος ροής όταν χρησιμοποιούμε όρια ελέγχου 3σ , η κατανομή της στατιστικής συνάρτησης W είναι κανονική και η διεργασία είναι εντός στατιστικού ελέγχου είναι

$$ARL_0 = \frac{1}{a} = \frac{1}{0.0027} = 370.$$

Όταν μία διεργασία βρίσκεται εκτός των ορίων ελέγχου, διότι το μέσο επίπεδο της διεργασίας μετατοπίστηκε από μ_w σε μ^* και χρησιμοποιούμε μοντέλο ορίων $L\sigma$, το **εκτός ελέγχου μέσο μήκος ροής** (*out-of-control Average Run Length, ARL₁*) δίνεται από τον τύπο

$$ARL_1 = \frac{1}{1 - \beta},$$

όπου το β υπολογίζεται από τον τύπο του Σφάλματος Τύπου II. Το ARL_1 δηλώνει τον αναμενόμενο αριθμό δειγμάτων που πρέπει να ληφθούν για να εντοπιστεί η αλλαγή στο μέσο επίπεδο της διεργασίας από τη στιγμή που αυτή συνέβη. Αυτό είναι λογικό, αφού όπως αναφέραμε προηγουμένως β είναι η πιθανότητα ένα σημείο του διαγράμματος να βρεθεί εντός των ορίων ελέγχου, ενώ στην πραγματικότητα η κατανομή της X , και κατ' επέκταση η κατανομή της W , έχει αλλάξει. Συνεπώς, η πιθανότητα το σημείο αυτό να βρίσκεται εκτός των ορίων ελέγχου είναι $1 - \beta$.

Στην πράξη θέλουμε να έχουμε μεγάλο εντός ελέγχου μέσο μήκος ροής (ARL_0) και μικρό εκτός ελέγχου μέσο μήκος ροής (ARL_1). Αυτό πρακτικά σημαίνει ότι πρέπει να έχουμε όσα περισσότερα σημεία γίνεται εντός των ορίων ελέγχου και ταυτόχρονα όσο το δυνατόν λιγότερα μέχρι να εντοπιστεί η αλλαγή στο μέσο επίπεδο της διεργασίας από τη στιγμή που αυτή συνέβη.

Συχνά εκτός από το μέσο μήκος ροής χρησιμοποιούμε και μία άλλη έννοια σχετική με τα διαγράμματα ελέγχου, το **μέσο χρόνο σήματος** (*Average Time to Signal, ATS*). Η ποσότητα ATS ορίζεται από τον τύπο

$$ATS = ARL \times h,$$

όπου h είναι ο χρόνος που μεσολαβεί για τη λήψη δύο διαδοχικών δειγμάτων και θεωρείται σταθερός. Συνεπώς, ο μέσος χρόνος σήματος (ATS) δηλώνει το μέσο (πραγματικό) χρόνο που απαιτείται για να δώσει το διάγραμμα ένδειξη ότι η διεργασία είναι εκτός ελέγχου.

1.7.4 Μέγεθος Δείγματος, Συχνότητα Δειγματοληψίας και Ορθολογικές Υποομάδες

Πριν ξεκινήσουμε να σχεδιάζουμε ένα διάγραμμα ελέγχου πρέπει πρώτα να καθορίσουμε το μέγεθος δείγματος και τη συχνότητα δειγματοληψίας. Το ιδανικό θα ήταν να έχουμε όσο το δυνατόν πιο μεγάλο μέγεθος δείγματος γίνεται, έτσι ώστε η ανίχνευση μικρών μετατοπίσεων του μέσου επιπέδου της διεργασίας να γίνεται ευκολότερα. Όσο το μέγεθος δείγματος αυξάνεται, η πιθανότητα β μικραίνει και αυτό έχει ως αποτέλεσμα την αύξηση της ικανότητας της διεργασίας να ανιχνεύει ταχύτερα τις μετατοπίσεις του μέσου επιπέδου της διεργασίας. Συνεπώς, το είδος της μετατόπισης, είτε αυτή είναι μεγάλη είτε μικρή, του μέσου

της διεργασίας που θέλουμε να ανιχνεύσουμε είναι αυτό που καθορίζει το μέγεθος δείγματος (Αντζουλάκος (2010)).

Όσον αφορά τη συχνότητα δειγματοληψίας, το ιδανικό θα ήταν να γίνεται αρκετά συχνά, και μάλιστα να ελέγχονται όλα τα δείγματα. Κάτι τέτοιο, όμως, συνεπάγεται και μεγάλο οικονομικό κόστος. Για το λόγο αυτό, είναι προτιμότερο να ελέγχουμε μικρά δείγματα αρκετά συχνά (Αντζουλάκος (2010)).

Ο W.A. Shewhart ανέπτυξε μια έννοια, αυτή των **ορθολογικά επιλεγμένων υποομάδων** (*rational subgroups*), σχετική με τη συλλογή δεδομένων που θα χρησιμοποιηθούν για να κατασκευαστεί ένα διάγραμμα ελέγχου. Σύμφωνα με την έννοια αυτή, τα δεδομένα θα πρέπει να συλλέγονται με τέτοιο τρόπο ώστε, εάν εμφανίζονται ειδικές αιτίες μεταβλητότητας σε μια διεργασία, οι διαφορές μεταξύ (*between*) των δειγμάτων να μεγιστοποιούνται και ταυτόχρονα οι διαφορές μέσα (*within*) στα δείγματα να ελαχιστοποιούνται.

Η επιλογή των ορθολογικών υποομάδων γίνεται με δύο τρόπους. Οι τιμές κάθε δείγματος προκύπτουν από προϊόντα που είτε έχουν παραχθεί την ίδια χρονική στιγμή, στο βαθμό που αυτό γίνεται εφικτό, είτε έχουν επιλεγεί τυχαία από τη συνολική παραγωγή των προϊόντων που παράχθηκαν από τη στιγμή που επιλέχθηκε το προηγούμενο δείγμα. Η πρώτη μέθοδος χρησιμοποιείται όταν θέλουμε να ανιχνεύσουμε αλλαγές στη διεργασία, ενώ η δεύτερη όταν θέλουμε να πάρουμε αποφάσεις σχετικά με την αποδοχή όλων των προϊόντων που παράχθηκαν από τη στιγμή που επιλέχθηκε το προηγούμενο δείγμα (Αντζουλάκος (2010)).

1.7.5 Φάση I και Φάση II

Στη βιβλιογραφία, έχει προταθεί η χρήση δύο διαφορετικών φάσεων σχετικά με τη σχεδίαση ενός διαγράμματος ελέγχου διεργασιών, η Φάση I και η Φάση II (Woodall (2000)).

Στη **Φάση I** (*Phase I or Off-Line Control Phase or Initial Study Phase or Retrospective Analysis*) χρησιμοποιούνται διαγράμματα ελέγχου για να ελέγξουμε αναδρομικά εάν η διεργασία ήταν εντός ή εκτός στατιστικού ελέγχου όταν συλλέχθηκαν τα πρώτα δείγματα. Εφόσον η διεργασία διαπιστωθεί ότι όντως είναι εντός στατιστικού ελέγχου, τα όρια ελέγχου του διαγράμματος που προέκυψαν θεωρούνται κατάλληλα για την μελλοντική παρακολούθηση της συμπεριφοράς της διεργασίας αυτής. Η Φάση I συχνά αναφέρεται και ως

αναδρομική (*retrospective*). Σε αυτή τη φάση ο διαχειριστής της διαδικασίας αυτής μελετά τη διαδικασία εντατικά προσπαθώντας να απαντήσει στο ερώτημα: «Τα δείγματα αυτά που συλλέχθηκαν προέρχονται από μία εντός ελέγχου διεργασία;». Στο σημείο αυτό, θα πρέπει να σημειώσουμε ότι οι παράμετροι της διεργασίας, μ_0 και Σ_0 , μπορεί να είναι γνωστές ή μπορούν να εκτιμηθούν.

Στη **Φάση II** (*Phase II or On-Line Control Phase or Control to Standard Phase or Process Monitoring*) χρησιμοποιούνται διαγράμματα ελέγχου για να ελέγξουμε εάν η διεργασία εξακολουθεί να παραμένει εντός στατιστικού ελέγχου όταν συλλέγονται μελλοντικά δείγματα. Στη φάση αυτή ο διαχειριστής της διαδικασίας παρακολουθεί τη διεργασία για οποιαδήποτε πιθανή αλλαγή που θα επηρεάσει την εντός ελέγχου κατάσταση, καθώς τα νέα δείγματα προστίθενται στη διεργασία. Συνεπώς, ο διαχειριστής σε κάθε στάδιο δειγματοληψίας θα πρέπει να αναρωτιέται εάν η κατάσταση της διεργασίας άλλαξε. Επιπλέον, στη φάση αυτή τα δεδομένα δε συλλέγονται από μία εντός ελέγχου διεργασία, εκτός εάν υπάρχει ξεκάθαρη ένδειξη ότι δεν υπάρχει καμία αλλαγή στη διεργασία. Αξίζει να επισημάνουμε ότι τα όρια ελέγχου που χρησιμοποιούνται στα διαγράμματα ελέγχου της Φάσης II, είναι αυτά που προέκυψαν από τη Φάση I (Bersimis, Psarakis & Panaretos (2007), Αντζουλάκος (2010)).

1.8 Δομή Διπλωματικής Εργασίας

Η παρούσα διπλωματική εργασία πραγματεύεται τα διαγράμματα ελέγχου διεργασιών, τα οποία αποτελούν το βασικό, εάν όχι το βασικότερο, εργαλείο του Στατιστικού Ελέγχου Διεργασιών. Συγκεκριμένα, στο παρόν κεφάλαιο γίνεται μια σύντομη εισαγωγή στο Στατιστικό Έλεγχο Ποιότητας, καθώς και στην ιστορική του διαδρομή. Επιπλέον, γίνεται αναφορά στους διαφορετικούς ορισμούς της έννοιας της ποιότητας, οι οποίοι έχουν προταθεί κατά καιρούς, και στη φιλοσοφία της ολικής διαχείρισής της. Επίσης, περιγράφεται η μορφή ενός τυπικού διαγράμματος ελέγχου, καθώς και οι κατηγοριοποιήσεις του.

Στο δεύτερο κεφάλαιο παρουσιάζονται εν συντομία τα μονοδιάστατα παραμετρικά διαγράμματα ελέγχου, άμεση γενίκευση των οποίων αποτελούν τα κυριότερα πολυδιάστατα παραμετρικά διαγράμματα, τα οποία αναπτύσσονται αναλυτικά στο τρίτο κεφάλαιο. Συγχρόνως, αναφέρονται οι μέθοδοι εντοπισμού του χαρακτηριστικού ή των

χαρακτηριστικών εκείνων που έθεσαν τη διεργασία εκτός στατιστικού ελέγχου. Στο τελευταίο κεφάλαιο παρουσιάζονται μερικά πολυδιάστατα, μη παραμετρικά διαγράμματα ελέγχου. Τα εν λόγω διαγράμματα θεωρούνται άμεση γενίκευση κάποιων μονοδιάστατων διαγραμμάτων της ίδιας κατηγορίας, τα οποία περιγράφονται σύντομα στο τέταρτο κεφάλαιο. Συμπληρωματικά γίνεται σύγκριση μεταξύ παραμετρικών και μη παραμετρικών διαγραμμάτων ελέγχου, αναλύοντας τόσο τα πλεονεκτήματα, όσο και τα μειονεκτήματά τους.

Επιπρόσθετα, τα βασικότερα διαγράμματα κάθε κατηγορίας απεικονίζονται γραφικά στο αντίστοιχο κεφάλαιο, μέσω παραδειγμάτων. Τέλος, συγκρίνεται η αποτελεσματικότητα μεταξύ διαφόρων μονοδιάστατων ή/και πολυδιάστατων διαγραμμάτων ελέγχου, τα οποία έχουν εμφανιστεί πρόσφατα στη βιβλιογραφία.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΔΑ

ΚΕΦΑΛΑΙΟ 2

Μονομεταβλητά Παραμετρικά Διαγράμματα Ελέγχου

2.1 Εισαγωγή

Στο κεφάλαιο αυτό θα αναφερθούμε στα **μονομεταβλητά** (ή **μονοδιάστατα**) **παραμετρικά διαγράμματα ελέγχου** (*univariate parametric control charts*). Οι μετρήσεις που λαμβάνονται στα διαγράμματα αυτά αφορούν ένα ποιοτικό χαρακτηριστικό ($p = 1$), το οποίο γνωρίζουμε ότι προέρχεται από κάποια γνωστή κατανομή. Εάν και υπάρχουν αρκετά μονομεταβλητά διαγράμματα ελέγχου, παρακάτω θα παρουσιάσουμε μόνο τα διαγράμματα εκείνα, τα οποία αποτελούν άμεση γενίκευση των πολυμεταβλητών διαγραμμάτων που θα αναπτυχθούν στο επόμενο κεφάλαιο.

Τα μονομεταβλητά παραμετρικά διαγράμματα ελέγχου χωρίζονται σε δύο μεγάλες κατηγορίες ανάλογα με το εάν ο σχεδιασμός ενός σημείου (του διαγράμματος) βασίζεται στις πληροφορίες που δίνει το πιο πρόσφατο δείγμα ή προγενέστερα δείγματα. Η πρώτη κατηγορία αναφέρεται στα διαγράμματα ελέγχου **χωρίς μνήμη** (*control charts without memory*) και η δεύτερη στα διαγράμματα ελέγχου **με μνήμη** (*control charts with memory*).

2.2 Διαγράμματα Ελέγχου Τύπου Shewhart για Μεταβλητές χωρίς μνήμη

Τα διαγράμματα ελέγχου **τύπου Shewhart για μεταβλητές** (*control charts for variables*) χρησιμοποιούνται όταν θέλουμε να παρακολουθήσουμε τη μέση τιμή και τη διασπορά του

συνεχούς ποιοτικού χαρακτηριστικού X και φυσικά όταν θέλουμε να ανιχνεύσουμε οποιαδήποτε πιθανή μεταβολή της διεργασίας. Για την παρουσίαση των διαγραμμάτων της ενότητας αυτής βασιζόμαστε στους παρακάτω συγγραφείς: Montgomery (1997), Mitra (2008) Γεωργακάκος (2002) και Αντζουλάκος (2010).

2.2.1 Διαγράμματα Ελέγχου για Δείγματα

Όπως αναφέραμε στο πρώτο κεφάλαιο, τα **διαγράμματα ελέγχου για δείγματα** ή **ομάδες** (*control charts for individual observations*) χρησιμοποιούνται όταν το μέγεθος δείγματος είναι μεγαλύτερο από τη μονάδα ($n > 1$). Τα διαγράμματα αυτά ανάλογα με την κρίσιμη ποσότητα χωρίζονται σε διαγράμματα ελέγχου για τη μέση τιμή και διαγράμματα ελέγχου για τη διασπορά.

❖ Διαγράμματα Ελέγχου για τη Μέση Τιμή

Τα **διαγράμματα ελέγχου για τη μέση τιμή** (*control charts for the process mean*) χρησιμοποιούνται με σκοπό την παρακολούθηση της συμπεριφοράς της μέσης τιμής του συνεχούς ποιοτικού χαρακτηριστικού X , με την προϋπόθεση ότι η διακύμανση παραμένει σταθερή σε όλη τη διαδικασία.

Θεωρούμε ότι η κατανομή του ποιοτικού χαρακτηριστικού X είναι κανονική με γνωστές παραμέτρους, συμβολικά $X \sim N(\mu, \sigma^2)$. Εάν $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $i \geq 1$ είναι ένα τυχαίο δείγμα του χαρακτηριστικού X με μέγεθος n , τότε ο δειγματικός μέσος

$$\bar{X}_i = \frac{X_{i1} + X_{i2} + \dots + X_{in}}{n}$$

ακολουθεί την κανονική κατανομή με παραμέτρους μ και σ^2/n , συμβολικά $\bar{X}_i \sim N(\mu, \sigma^2/n)$ και επιπλέον είναι αμερόληπτος εκτιμητής της μέσης τιμής μ (του χαρακτηριστικού X). Συνεπώς, ο δειγματικός μέσος $W_i = \bar{X}_i$ θα παίρνει τιμές στο διάστημα

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

με πιθανότητα $1 - a$, δηλαδή

$$P\left(\bar{X}_i \in \left[\mu - z_{a/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu + z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} \right] \right) = 1 - a.$$

Για $a = 0.0027$, και άρα $z_{a/2} = 3$, το διάστημα παίρνει την ακόλουθη μορφή

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

Επομένως, η πιθανότητα ο δειγματικός μέσος $W_i = \bar{X}_i$ να βρεθεί στο διάστημα αυτό είναι 99.73%.

Τα 3σ όρια ελέγχου ενός διαγράμματος, στο οποίο θα απεικονίζονται οι τιμές του δειγματικού μέσου $W_i = \bar{X}_i$ είναι τα εξής

$$\begin{cases} LCL = \mu_{W_i} - 3 \cdot \sigma_{W_i} \\ CL = \mu_{W_i} \\ UCL = \mu_{W_i} + 3 \cdot \sigma_{W_i} \end{cases} \Rightarrow \begin{cases} LCL = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} \\ CL = \mu \\ UCL = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} \end{cases}.$$

Εάν ένα σημείο του διαγράμματος αυτού βρεθεί εκτός των ορίων ελέγχου, τότε λέμε ότι υπάρχει ένδειξη ότι η διεργασία είναι εκτός στατιστικού ελέγχου, με την προϋπόθεση πάντα ότι η διασπορά του συνεχούς χαρακτηριστικού X παρέμεινε σταθερή σε όλη τη διάρκεια της διεργασίας. Συνεπώς, τα 3σ όρια ελέγχου ενός διαγράμματος για τη μέση τιμή του χαρακτηριστικού X είναι τα ακόλουθα

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης II} \end{array} \begin{cases} LCL = \mu - A \cdot \sigma \\ CL = \mu \\ UCL = \mu + A \cdot \sigma \end{cases}, \quad (2.2.1)$$

όπου $A = 3/\sqrt{n}$. Τα όρια ελέγχου του διαγράμματος αυτού είναι Φάσης II, διότι η μέση τιμή μ και η διακύμανση σ^2 του ποιοτικού χαρακτηριστικού X είναι γνωστά.

Τα $L\sigma$ όρια ελέγχου του διαγράμματος για τη μέση τιμή του χαρακτηριστικού X είναι τα ακόλουθα

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με } L\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης II} \end{array} \left\{ \begin{array}{l} LCL = \mu - L \cdot \frac{\sigma}{\sqrt{n}} \\ CL = \mu \\ UCL = \mu + L \cdot \frac{\sigma}{\sqrt{n}} \end{array} \right.$$

Αντίστοιχα, τα όρια ελέγχου πιθανότητας a του διαγράμματος για τη μέση τιμή του χαρακτηριστικού X είναι τα εξής

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με Όρια Πιθανότητας } a \\ \text{Φάσης II} \end{array} \left\{ \begin{array}{l} LCL = \mu - z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} \\ CL = \mu \\ UCL = \mu + z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} \end{array} \right. \quad (2.2.2)$$

Στην πραγματικότητα, όμως, η μέση τιμή μ και η διακύμανση σ^2 του ποιοτικού χαρακτηριστικού X δεν είναι γνωστά και πρέπει να εκτιμηθούν.

Χρησιμοποιώντας τη Μέθοδο R και παίρνοντας ως εκτίμηση της μέσης τιμής την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \bar{R}/d_2$, τα όρια έλεγχου του διαγράμματος για τη μέση τιμή του χαρακτηριστικού X έχουν τη μορφή

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \\ \text{(Μεθοδος } R) \end{array} \left\{ \begin{array}{l} LCL = \bar{\bar{X}} - A_2 \cdot \bar{R} \\ CL = \bar{\bar{X}} \\ UCL = \bar{\bar{X}} + A_2 \cdot \bar{R} \end{array} \right. \quad (2.2.3)$$

όπου η σταθερά A_2 υπολογίζεται από τον τύπο

$$A_2 = \frac{3}{d_2 \cdot \sqrt{n}}.$$

Χρησιμοποιώντας τη Μέθοδο S και παίρνοντας ως εκτίμηση της μέσης τιμής την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \bar{S}/c_4$, τα όρια έλεγχου του διαγράμματος για τη μέση τιμή του χαρακτηριστικού X έχουν την ακόλουθη μορφή

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \\ \text{(Μεθοδος } S) \end{array} \left\{ \begin{array}{l} LCL = \bar{\bar{X}} - A_3 \cdot \bar{S} \\ CL = \bar{\bar{X}} \\ UCL = \bar{\bar{X}} + A_3 \cdot \bar{S} \end{array} \right., \quad (2.2.4)$$

όπου η σταθερά A_3 υπολογίζεται από τον τύπο

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

Η ποσότητα c_4 εξαρτάται από το μέγεθος δείγματος και υπολογίζεται από τον τύπο

$$c_4 = \left[\frac{2}{(n-1)} \right]^{1/2} \cdot \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}.$$

Τέλος, χρησιμοποιώντας τη Μέθοδο S^2 και παίρνοντας ως εκτίμηση της μέσης τιμής την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \sqrt{\bar{S}^2}$, τα όρια έλεγχου του διαγράμματος για τη μέση τιμή του χαρακτηριστικού X είναι τα εξής

$$\begin{array}{l} \text{Διάγραμμα } \bar{X} \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \\ \text{(Μεθοδος } S^2) \end{array} \left\{ \begin{array}{l} LCL = \bar{\bar{X}} - A \cdot \sqrt{\bar{S}^2} \\ CL = \bar{\bar{X}} \\ UCL = \bar{\bar{X}} + A \cdot \sqrt{\bar{S}^2} \end{array} \right. \quad (2.2.5)$$

❖ Διαγράμματα Ελέγχου για τη Διασπορά

Τα διαγράμματα ελέγχου για την διασπορά (*control charts for the process variability-dispersion*) χρησιμοποιούνται με σκοπό την παρακολούθηση της συμπεριφοράς της διασποράς του ποιοτικού χαρακτηριστικού X , το οποίο είναι συνεχές.

Θεωρούμε ότι η κατανομή του χαρακτηριστικού X είναι κανονική με γνωστή την τυπική απόκλιση σ (συμβ. $X \sim N(\mu, \sigma^2)$). Εάν $X_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $i \geq 1$ είναι ένα τυχαίο δείγμα του χαρακτηριστικού X με μέγεθος n , τότε για τη στατιστική συνάρτηση

$$W_i = R_i = X_{i(n)} - X_{i1}$$

ισχύει ότι $\mu_{R_i} = E(R_i) = \sigma \cdot d_2$ και $\sigma_{R_i} = \sqrt{V(R_i)} = \sigma \cdot d_3$, όπου d_2 και d_3 είναι ποσότητες που εξαρτώνται από το μέγεθος δείγματος n .

Τα 3σ όρια ελέγχου ενός διαγράμματος, στο οποίο θα απεικονίζονται οι τιμές του εύρους R_i των δειγμάτων, το οποίο θεωρείται μέτρο διασποράς, είναι τα εξής

$$\begin{cases} LCL = \mu_{R_i} - 3 \cdot \sigma_{R_i} \\ CL = \mu_{R_i} \\ UCL = \mu_{R_i} + 3 \cdot \sigma_{R_i} \end{cases} \Rightarrow \begin{cases} LCL = (d_2 - 3 \cdot d_3) \cdot \sigma \\ CL = d_2 \cdot \sigma \\ UCL = (d_2 + 3 \cdot d_3) \cdot \sigma. \end{cases}$$

Θέτοντας $D_1 = d_2 - 3 \cdot d_3$ και $D_2 = d_2 + 3 \cdot d_3$, τα 3σ όρια ελέγχου ενός διαγράμματος για τη διασπορά του χαρακτηριστικού X είναι τα ακόλουθα

$$\begin{array}{l} \text{Διάγραμμα } R \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης II} \end{array} \begin{cases} LCL = D_1 \cdot \sigma \\ CL = d_2 \cdot \sigma \\ UCL = D_2 \cdot \sigma. \end{cases} \quad (2.2.6)$$

Όταν $n \leq 6$, τότε το $D_1 < 0$. Στην περίπτωση αυτή και επειδή δεν μπορούμε να έχουμε αρνητικά όρια ελέγχου, θέτουμε $D_1 = 0$.

Για να μπορέσουμε να αναπτύξουμε ένα διάγραμμα πιθανότητας a , χρησιμοποιούμε τον τύπο

$$P(\sigma \cdot w_{1-a/2} \leq R_i \leq \sigma \cdot w_{a/2}) = 1 - a,$$

όπου w_a είναι το άνω a -ποσοστιαίο σημείο της στατιστικής συνάρτησης $W = R / \sigma$.

Για να μπορέσουμε να αναπτύξουμε ένα διάγραμμα παρακολούθησης της διασποράς του χαρακτηριστικού X , χρησιμοποιούμε τη στατιστική συνάρτηση

$$W_i = S_i = \sqrt{\frac{1}{n-1} \cdot \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2},$$

για την οποία ισχύει ότι

$$\mu_{S_i} = E(S_i) = \sigma \cdot c_4 \text{ και } \sigma_{S_i} = \sqrt{V(S_i)} = \sigma \cdot \sqrt{1 - c_4^2}.$$

Συνεπώς, τα 3σ όρια ελέγχου ενός διαγράμματος, στο οποίο θα απεικονίζονται οι δειγματικές τυπικές αποκλίσεις S_i , το οποίο αποτελεί μέτρο διασποράς, είναι τα εξής

$$\begin{cases} LCL = \mu_{S_i} - 3 \cdot \sigma_{S_i} \\ CL = \mu_{S_i} \\ UCL = \mu_{S_i} + 3 \cdot \sigma_{S_i} \end{cases} \Rightarrow \begin{cases} LCL = (c_4 - 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma \\ CL = c_4 \cdot \sigma \\ UCL = (c_4 + 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma. \end{cases}$$

Θέτοντας $B_5 = c_4 - 3 \cdot \sqrt{1 - c_4^2}$ και $B_6 = c_4 + 3 \cdot \sqrt{1 - c_4^2}$, τα 3σ όρια ελέγχου ενός διαγράμματος για τη διασπορά του χαρακτηριστικού X είναι τα ακόλουθα

$$\begin{array}{l} \text{Διάγραμμα } S \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης II} \end{array} \begin{cases} LCL = B_5 \cdot \sigma \\ CL = c_4 \cdot \sigma \\ UCL = B_6 \cdot \sigma. \end{cases} \quad (2.2.7)$$

Όταν $n \leq 5$, τότε το $B_5 < 0$, οπότε θέτουμε $B_5 = 0$.

Για να μπορέσουμε να αναπτύξουμε ένα διάγραμμα ελέγχου πιθανότητας a , χρησιμοποιούμε τον τύπο

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

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

και για την παρακολούθηση της διασποράς του χαρακτηριστικού X , μπορούμε να χρησιμοποιήσουμε και τη στατιστική συνάρτηση

$$W_i = S_i^2 = \frac{1}{n-1} \cdot \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2,$$

για την οποία ισχύει ότι

$$\mu_{S_i} = E(S_i^2) = \sigma^2 \text{ και } \frac{(n-1) \cdot S_i^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Οπότε έχουμε

$$P\left(\chi^2_{n-1,1-a/2} \leq \frac{(n-1) \cdot S_i^2}{\sigma^2} \leq \chi^2_{n-1,a/2}\right) = 1-a \quad \Leftrightarrow$$

$$P\left(\frac{\sigma^2}{n-1} \cdot \chi^2_{n-1,1-a/2} \leq S_i^2 \leq \frac{\sigma^2}{n-1} \cdot \chi^2_{n-1,a/2}\right) = 1-a.$$

Συνεπώς, τα 3σ όρια ελέγχου ενός διαγράμματος, στο οποίο θα απεικονίζονται οι τιμές της δειγματικής διακύμανσης S_i^2 , το οποίο αποτελεί μέτρο διασποράς, είναι τα εξής

$$\begin{array}{l} \text{Διάγραμμα } S^2 \\ \text{με Όρια Πιθανότητας } \alpha \\ \text{Φάσης II} \end{array} \left\{ \begin{array}{l} LCL = \frac{\sigma^2}{n-1} \cdot \chi^2_{n-1,1-a/2} \\ CL = \sigma^2 \\ UCL = \frac{\sigma^2}{n-1} \cdot \chi^2_{n-1,a/2} \end{array} \right. \quad (2.2.8)$$

Στην πραγματικότητα, όμως, η τυπική απόκλιση σ του ποιοτικού χαρακτηριστικού X δεν είναι γνωστή και πρέπει να εκτιμηθεί. Για να μπορέσουμε να εκτιμήσουμε την ποσότητα σ , επιλέγουμε m προκαταρκτικά δείγματα μεγέθους n το καθένα $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $1 \leq i \leq m$, τα οποία είναι ανεξάρτητα μεταξύ τους. Κατά τη διάρκεια επιλογής των δειγμάτων, θεωρούμε ότι η διεργασία ήταν εντός στατιστικού ελέγχου. Συνήθως, τα προκαταρκτικά δείγματα είναι μεταξύ 20 και 25 με μέγεθος από 4 έως 6. Επιπλέον, θεωρούμε ότι το χαρακτηριστικό X ακολουθεί κανονική κατανομή, συμβολικά $X \sim N(\mu, \sigma^2)$.

Χρησιμοποιώντας ως εκτίμηση της τυπικής απόκλισης σ την ποσότητα $\hat{\sigma} = \bar{R}/d_2$, ο τύπος (2.2.6) γίνεται

$$\begin{array}{l} \text{Διάγραμμα } R \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \end{array} \left\{ \begin{array}{l} LCL = D_3 \cdot \bar{R} \\ CL = \bar{R} \\ UCL = D_4 \cdot \bar{R} \end{array} \right. \quad (2.2.9)$$

όπου

$$D_3 = 1 - 3 \cdot \frac{d_3}{d_2} \quad \text{και} \quad D_4 = 1 + 3 \cdot \frac{d_3}{d_2}.$$

Όταν $n \leq 6$, τότε το $D_3 < 0$, οπότε θέτουμε $D_3 = 0$.

Χρησιμοποιώντας ως εκτίμηση της τυπικής απόκλισης σ την ποσότητα $\hat{\sigma} = \bar{S}/c_4$, ο τύπος (2.2.7) γίνεται

$$\begin{array}{l} \text{Διάγραμμα } S \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \end{array} \left\{ \begin{array}{l} LCL = B_3 \cdot \bar{S} \\ CL = \bar{S} \\ UCL = B_4 \cdot \bar{S} \end{array} \right. \quad (2.2.10)$$

όπου

$$B_3 = 1 - \frac{3}{c_4} \cdot \sqrt{1 - c_4^2} \quad \text{και} \quad B_4 = 1 + \frac{3}{c_4} \cdot \sqrt{1 - c_4^2}.$$

Όταν $n \leq 5$, τότε το $B_3 < 0$, συνεπώς θέτουμε $B_3 = 0$.

Χρησιμοποιώντας ως εκτίμηση της διακύμανσης σ^2 την ποσότητα $\hat{\sigma}^2 = \bar{S}^2$, ο τύπος (2.2.8) γίνεται

$$\begin{array}{l} \text{Διάγραμμα } S^2 \\ \text{με Όρια Πιθανότητας } \alpha \\ \text{Φάσης I} \end{array} \left\{ \begin{array}{l} LCL = \frac{\bar{S}^2}{n-1} \cdot \chi_{n-1; 1-\alpha/2}^2 \\ CL = \bar{S}^2 \\ UCL = \frac{\bar{S}^2}{n-1} \cdot \chi_{n-1; \alpha/2}^2 \end{array} \right. \quad (2.2.11)$$

2.2.2 Παράδειγμα Διαγραμμάτων Ελέγχου για Δείγματα

Θεωρούμε ως χαρακτηριστικό ποιότητας τη διάμετρο εμβόλου που χρησιμοποιεί μια αυτοκινητοβιομηχανία. Για την παρακολούθηση της διεργασίας λαμβάνονται ανά 2 ώρες δείγματα μεγέθους 10 εμβόλων και ελέγχεται η διάμετρός τους. Κωδικοποιούμε τις παρατηρήσεις έτσι ώστε η επιθυμητή διάμετρος του εμβόλου να αντιστοιχεί στην τιμή 0.0. Από δεκαπέντε δείγματα μεγέθους 10 που συγκεντρώθηκαν καταγράφηκαν οι παρακάτω τιμές (σε mm).

Δείγμα	Παρατηρήσεις										\bar{X}_i	R_i	S_i	S_i^2
1	1.0	-1.0	-1.0	-1.0	2.0	1.5	0.0	1.0	0.0	0.0	0.25	3.0	1.0865	1.1806
2	2.0	-1.5	-0.5	1.5	2.0	2.0	0.0	-1.0	2.5	-0.5	0.65	4.0	1.4916	2.2250
3	2.0	0.5	2.0	-1.0	1.0	-1.0	0.5	1.5	0.5	-1.5	0.45	3.5	1.2572	1.5806
4	0.0	0.0	0.5	1.0	1.5	1.0	-1.0	1.0	1.5	-1.0	0.45	2.5	0.9265	0.8583
5	1.5	1.0	1.0	-1.0	0.0	-1.5	-1.0	-1.0	1.0	-1.0	-0.10	3.0	1.1255	1.2667
6	0.0	0.5	-2.0	0.0	-1.0	1.5	-1.2	0.0	-2.0	-1.5	-0.57	3.5	1.1519	1.3268
7	0.0	0.0	-2.0	-0.5	0.5	1.0	-2.5	-0.5	0.0	0.0	-0.40	3.5	1.0750	1.1556
8	1.0	-0.5	0.0	0.0	0.0	0.5	-1.0	1.0	-2.0	1.0	0.00	3.0	0.9718	0.9444
9	-2.0	-1.5	1.5	1.5	0.0	0.0	2.0	1.0	0.0	1.0	0.35	4.0	1.3134	1.7250
10	1.0	0.0	1.5	1.5	1.0	-1.0	2.0	2.5	-2.0	-1.5	0.50	4.5	1.5456	2.3889
11	-0.5	3.5	0.0	1.0	-1.5	1.5	-1.0	-1.0	1.0	0.5	0.35	5.0	1.4916	2.2250
12	0.0	1.5	0.0	0.0	2.0	-1.5	0.5	-0.5	2.0	-1.0	0.30	3.5	1.2065	1.4556
13	0.0	-2.0	-0.5	0.0	-0.5	2.0	1.5	0.0	0.5	-1.0	0.00	4.0	1.1547	1.3333
14	-1.0	-0.5	-0.5	-1.0	0.0	0.5	0.5	-1.5	-1.0	-1.0	-0.55	2.0	0.6852	0.4694
15	0.5	1.0	-1.0	-0.5	-2.0	-1.0	-1.5	0.0	1.5	1.5	-0.15	3.5	1.2483	1.5583

Πίνακας 2.1: Τιμές διαμέτρου εμβόλου και τιμές στατιστικών συναρτήσεων των μονοδιάστατων διαγραμμάτων R , \bar{X} , S και S^2 .

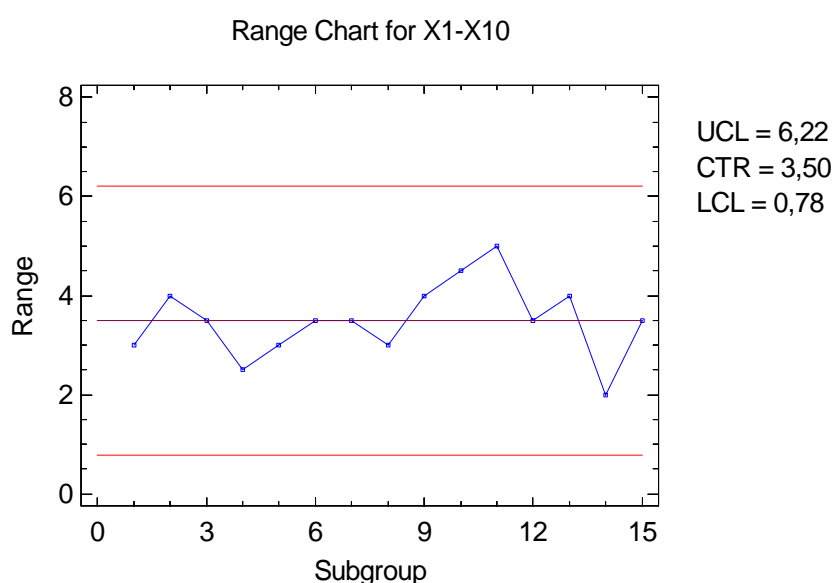
Σύμφωνα με το παραπάνω παράδειγμα, έχουμε στη διάθεση μας $m=15$ προκαταρκτικά δείγματα μεγέθους $n=10$ το καθένα. Για κάθε προκαταρκτικό δείγμα έχει υπολογιστεί ο δειγματικός μέσος όρος \bar{X}_i , το εύρος R_i , οι δειγματικές τυπικές αποκλίσεις S_i και η δειγματική διακύμανση S_i^2 . Η κατανομή της διαμέτρου X του εμβόλου, καθώς η μέση τιμή και η τυπική απόκλιση της κατανομής, είναι άγνωστα. Για να μπορέσουμε να κατασκευάσουμε τα διαγράμματα ελέγχου (Φάσης I), υποθέτουμε ότι η κατανομή της διαμέτρου X είναι κανονική.

Για την κατασκευή ενός διαγράμματος ελέγχου για τη διασπορά του χαρακτηριστικού X , στο οποίο η απεικονιζόμενη ποσότητα θα είναι το εύρος R_i , απαιτείται ο υπολογισμός της

ποσότητας $\bar{R} = (R_1 + R_2 + \dots + R_{10})/10 = 3.5$. Συνεπώς, τα όρια ελέγχου του διαγράμματος R Φάσης I, που προκύπτουν από τον τύπο (2.2.9), είναι τα εξής

$$\begin{cases} LCL = D_3 \cdot \bar{R} = 0.223 \times 3.5 = 0.781 \\ CL = \bar{R} = 3.5 = 3.5 \\ UCL = D_4 \cdot \bar{R} = 1.777 \times 3.5 = 6.2195. \end{cases}$$

Από το Σχήμα 2.1 που ακολουθεί είναι φανερό ότι η διασπορά του χαρακτηριστικού X είναι εντός στατιστικού ελέγχου κατά τη λήψη των προκαταρκτικών δειγμάτων.



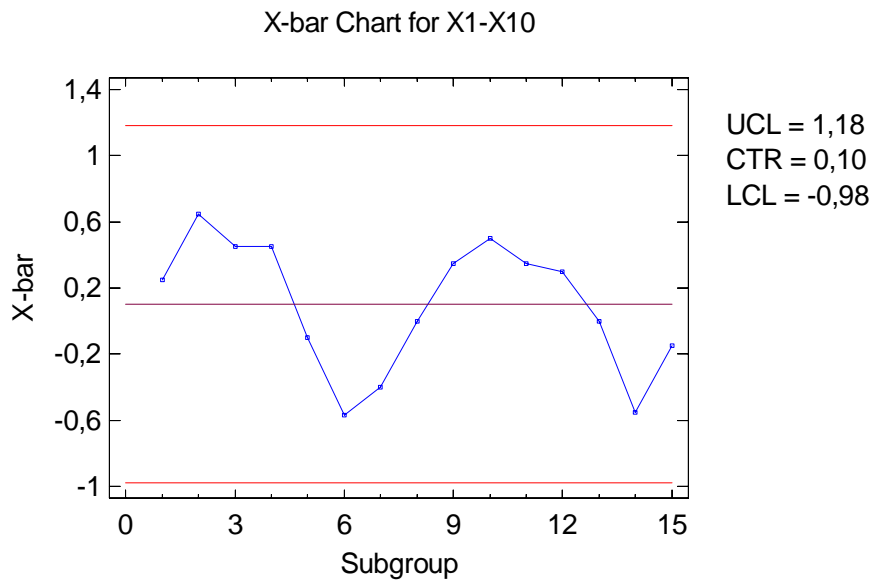
Σχήμα 2.1: Διάγραμμα R Φάσης I για τη διασπορά του χαρακτηριστικού X

Στο σημείο αυτό θα πρέπει να αναφέρουμε ότι κατασκευάσαμε πρώτα το διάγραμμα R , διότι τα όρια ελέγχου του διαγράμματος \bar{X} εξαρτώνται από τη διασπορά του της X . Για το λόγο αυτό, εάν η διασπορά της X δεν είναι εντός στατιστικού ελέγχου, τα όρια που θα προκύψουν για το διάγραμμα \bar{X} δε θα είναι αξιόπιστα.

Για την κατασκευή ενός διαγράμματος ελέγχου για τη μέση τιμή του χαρακτηριστικού X , στο οποίο η απεικονιζόμενη ποσότητα θα είναι ο δειγματικός μέσος \bar{X}_i , απαιτείται ο υπολογισμός της ποσότητας $\bar{\bar{X}} = (\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_{10})/10 = 0.102$. Συνεπώς, τα όρια ελέγχου του διαγράμματος \bar{X} Φάσης I, που προκύπτουν από τον τύπο (2.2.3) με τη μέθοδο R , είναι

$$\begin{cases} LCL = \bar{\bar{X}} - A_2 \cdot \bar{R} = 0.102 - 0.3083 \times 3.5 = -0.97705 \\ CL = \bar{\bar{X}} = 0.102 = 0.102 \\ UCL = \bar{\bar{X}} + A_2 \cdot \bar{R} = 0.102 + 0.3083 \times 3.5 = 1.18105. \end{cases}$$

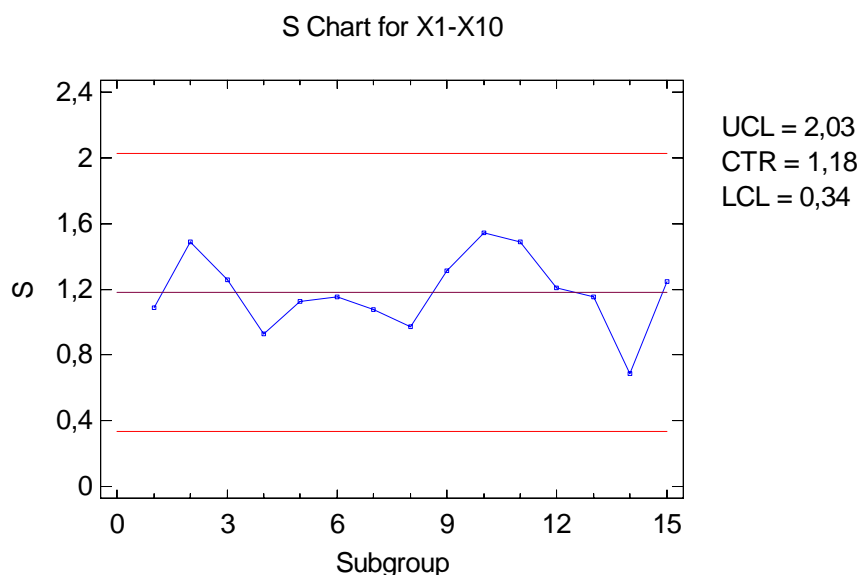
Υπενθυμίζουμε ότι με τη μέθοδο R χρησιμοποιούμε ως εκτίμηση της μέσης τιμής την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \bar{R}/d_2$. Από το Σχήμα 2.2 είναι φανερό ότι όλα τα σημεία του διαγράμματος \bar{X} είναι εντός των ορίων ελέγχου. Ωστόσο, διαφαίνεται η ύπαρξη κυκλικού μοτίβου. Συνεπώς, δεν μπορούμε να πούμε ότι η διεργασία είναι εντός στατιστικού ελέγχου.



Σχήμα 2.2: Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού X (Μέθοδος R)

Για την κατασκευή ενός διαγράμματος ελέγχου για τη διασπορά του χαρακτηριστικού X , στο οποίο η απεικονιζόμενη ποσότητα θα είναι οι δειγματικές τυπικές αποκλίσεις S_i , απαιτείται ο υπολογισμός της ποσότητας $\bar{S} = (S_1 + S_2 + \dots + S_{10})/10 = 1.18208$. Συνεπώς, τα όρια ελέγχου του διαγράμματος S Φάσης I, που προκύπτουν από τον τύπο (2.2.10), είναι

$$\begin{cases} LCL = B_3 \cdot \bar{S} = 0.2837 \times 1.18208 = 0.33536 \\ CL = \bar{S} = 1.18208 = 1.18208 \\ UCL = B_4 \cdot \bar{S} = 1.7163 \times 1.18208 = 2.02880. \end{cases}$$

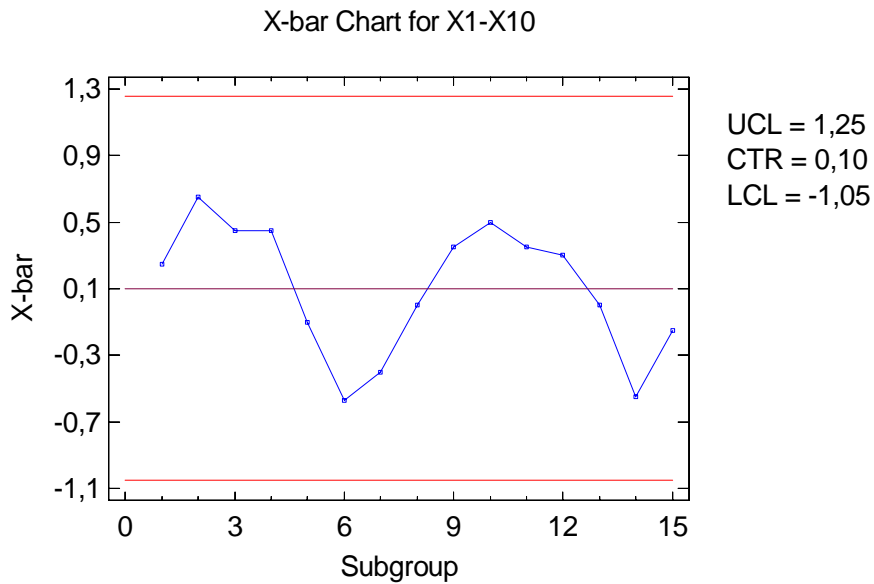


Σχήμα 2.3: Διάγραμμα S Φάσης I για τη διασπορά του χαρακτηριστικού X

Από το Σχήμα 2.3 είναι φανερό ότι η διασπορά του χαρακτηριστικού X είναι εντός στατιστικού ελέγχου κατά τη λήψη των προκαταρκτικών δειγμάτων.

Κατασκευάζουμε ένα ακόμη διάγραμμα ελέγχου για τη μέση τιμή της X , χρησιμοποιώντας τη μέθοδο S . Οπότε ως εκτίμηση της μέσης τιμής παίρνουμε την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \bar{S}/c_4$. Στο διάγραμμα αυτό η ποσότητα που θα απεικονίζεται είναι ο δειγματικός μέσος \bar{X}_i . Συνεπώς, τα όρια ελέγχου του διαγράμματος \bar{X} Φάσης I, που προκύπτουν από τον τύπο (2.2.4) με τη μέθοδο S , είναι

$$\begin{cases} LCL = \bar{\bar{X}} - A_3 \cdot \bar{S} = 0.102 - 0.9754 \times 1.18208 = -1.0510 \\ CL = \bar{\bar{X}} = 0.102 = 0.102 \\ UCL = \bar{\bar{X}} + A_3 \cdot \bar{S} = 0.102 + 0.9754 \times 1.18208 = 1.2550. \end{cases}$$



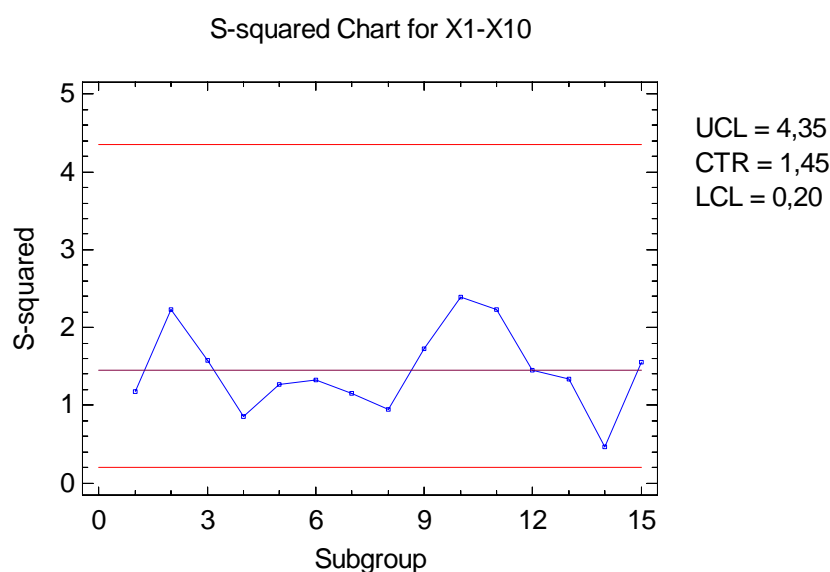
Σχήμα 2.4: Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού X (Μέθοδος S)

Και από το Σχήμα 2.4 διαφαίνεται η ύπαρξη κυκλικού μοτίβου. Συνεπώς, δεν μπορούμε να πούμε ότι η διεργασία είναι εντός στατιστικού ελέγχου.

Για την κατασκευή ενός διαγράμματος ελέγχου για τη διασπορά του χαρακτηριστικού X , στο οποίο η απεικονιζόμενη ποσότητα θα είναι οι δειγματική διακύμανση S_i^2 , απαιτείται ο υπολογισμός της ποσότητας $\bar{S}^2 = (S_1^2 + S_2^2 + \dots + S_{10}^2) / 10 = 1.44623$. Άρα, τα όρια ελέγχου πιθανότητας $\alpha = 0.002$ του διαγράμματος S Φάσης I, που προκύπτουν από τον τύπο (2.2.11), είναι τα εξής

$$\begin{cases} LCL = \frac{\bar{S}^2}{n-1} \cdot \chi_{n-1; 1-\alpha/2}^2 = \frac{1.44623}{9} \times \chi_{9, 0.999}^2 = 0.16069 \times 1.151950 = 0.199458 \\ CL = \bar{S}^2 = 1.44623 = 1.44623 = 1.44623 \\ UCL = \frac{\bar{S}^2}{n-1} \cdot \chi_{n-1; \alpha/2}^2 = \frac{1.44623}{9} \times \chi_{9, 0.001}^2 = 0.16069 \times 27.87716 = 4.435367. \end{cases}$$

Από το Σχήμα 2.5 που ακολουθεί, είναι φανερό ότι η διασπορά του χαρακτηριστικού X είναι εντός στατιστικού ελέγχου κατά τη λήψη των προκαταρκτικών δειγμάτων.

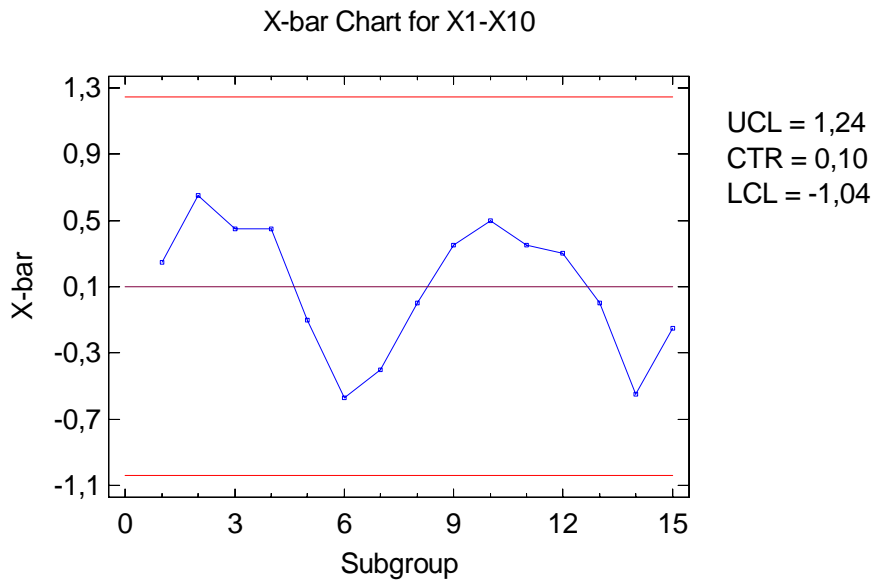


Σχήμα 2.5: Διάγραμμα S^2 Φάσης I για τη διασπορά του χαρακτηριστικού X

Κατασκευάζουμε ένα διάγραμμα ελέγχου για τη μέση τιμή της X , χρησιμοποιώντας τη μέθοδο S^2 . Οπότε ως εκτίμηση της μέσης τιμής παίρνουμε την ποσότητα $\hat{\mu} = \bar{\bar{X}}$ και ως εκτίμηση της τυπικής απόκλισης την ποσότητα $\hat{\sigma} = \sqrt{\bar{S}^2}$. Στο διάγραμμα αυτό η ποσότητα που θα απεικονίζεται είναι ο δειγματικός μέσος \bar{X}_i . Συνεπώς, τα όρια ελέγχου του διαγράμματος \bar{X} Φάσης I, που προκύπτουν από τον τύπο (2.2.5) με τη μέθοδο S^2 , είναι

$$\begin{cases} LCL = \bar{\bar{X}} - A \cdot \sqrt{\bar{S}^2} = 0.102 - \frac{3}{\sqrt{10}} \times \sqrt{1.44623} = -1.038 \\ CL = \bar{\bar{X}} = 0.102 \\ UCL = \bar{\bar{X}} + A \cdot \sqrt{\bar{S}^2} = 0.102 + \frac{3}{\sqrt{10}} \times \sqrt{1.44623} = 1.2429. \end{cases}$$

Στο Σχήμα 2.6, που απεικονίζεται παρακάτω, διαφαίνεται η ύπαρξη κυκλικού μοτίβου. Συνεπώς, δεν μπορούμε να πούμε ότι η διεργασία είναι εντός στατιστικού ελέγχου.



Σχήμα 2.6: Διάγραμμα \bar{X} Φάσης I για τη μέση τιμή του χαρακτηριστικού X (Μέθοδος S^2)

Από τα διαγράμματα ελέγχου Φάσης I μπορούμε να εκτιμήσουμε τη μέση τιμή μ και την τυπική απόκλιση σ της κατανομής του χαρακτηριστικού X , υπό την προϋπόθεση φυσικά ότι η διεργασία είναι εντός στατιστικού ελέγχου. Η εκτίμηση του μ θα είναι $\hat{\mu} = \bar{\bar{X}} = 0.102$ και η εκτίμηση του σ θα είναι $\hat{\sigma} = \bar{R}/d_2 = 3.5/3.0775 = 1.13729$, σύμφωνα με τη μέθοδο R , $\hat{\sigma} = \bar{S}/c_4 = 1.18208/0.9727 = 1.21526$, με τη μέθοδο S και $\hat{\sigma} = \sqrt{\bar{S}^2} = \sqrt{.44623} = 1.20259$, σύμφωνα με τη μέθοδο S^2 . Συνεπώς, θεωρώντας ότι $X \sim N(0.102, 1.13729^2)$ μπορούμε να κατασκευάσουμε τα διαγράμματα \bar{X} και R Φάσης II, χρησιμοποιώντας τα όρια ελέγχου των τύπων (2.2.1) και (2.2.6), αντίστοιχα. Θεωρώντας ότι $X \sim N(0.102, 1.21526^2)$ μπορούμε να κατασκευάσουμε τα διαγράμματα \bar{X} και S Φάσης II, χρησιμοποιώντας τα όρια ελέγχου των τύπων (2.2.1) και (2.2.7), αντίστοιχα. Τέλος, θεωρώντας ότι $X \sim N(0.102, 1.20259^2)$ κατασκευάζουμε τα διαγράμματα \bar{X} και S^2 Φάσης II, χρησιμοποιώντας τα όρια ελέγχου των τύπων (2.2.1) και (2.2.8), αντίστοιχα. Εάν είχαμε στη διάθεση μας επιπρόσθετα δείγματα, με τα διαγράμματα Φάσης II θα μπορούσαμε να ελέγξουμε εάν η διαδικασία παραμένει εντός στατιστικού ελέγχου. Προφανώς, τα όρια ελέγχου Φάσης II που θα προκύψουν θα είναι ίδια με αυτά της Φάσης I.

2.3 Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες χωρίς μνήμη

Τα διαγράμματα ελέγχου για ιδιότητες (*control charts for attributes*) χρησιμοποιούνται όταν το ποιοτικό χαρακτηριστικό περιγράφεται με διακριτές τυχαίες μεταβλητές. Ανάλογα με το χαρακτηριστικό που περιγράφουν, τα διαγράμματα αυτά χωρίζονται σε τέσσερις κατηγορίες: τα **διαγράμματα ελέγχου για το ποσοστό ελαττωματικών προϊόντων** (*control charts for fraction nonconforming, p-chart*), τα **διαγράμματα ελέγχου για τον αριθμό των ελαττωματικών προϊόντων** (*control charts for number nonconforming, np-chart*), τα **διαγράμματα ελέγχου για τον αριθμό των ελαττωμάτων** (*control charts for defects-nonconformities, c-chart*) και τα **διαγράμματα ελέγχου για το μέσο αριθμό των ελαττωμάτων** (*control charts for fraction nonconforming, u-chart*). Παρακάτω θα αναπτύξουμε μόνο τη δεύτερη κατηγορία, όπως αυτή παρουσιάζεται από τους Montgomery (1997), Jones-Farmer (2008), Γεωργακάκο (2002) και Αντζουλάκο (2010).

2.3.1 Διαγράμματα Ελέγχου για τον Αριθμό των Ελαττωματικών Προϊόντων

Θεωρούμε την τυχαία μεταβλητή $X_i = X_{i1} + X_{i2} + \dots + X_{in}$, η οποία εκφράζει τον **αριθμό των ελαττωματικών προϊόντων** (*number nonconforming*) στο i δείγμα και ακολουθεί τη Διωνυμική κατανομή με παραμέτρους n και p , συμβολικά $X_i \sim B(n, p)$.

Συνεπώς, τα 3σ όρια ελέγχου ενός διαγράμματος για την παρακολούθηση του αριθμού των ελαττωματικών προϊόντων, στο οποίο θα απεικονίζονται οι τιμές της τυχαίας μεταβλητής $W_i = X_i$ στα διάφορα δείγματα που επιλέγονται από την παραγωγή, είναι τα εξής

$$\begin{array}{l} \text{Διάγραμμα } np \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης II} \end{array} \left\{ \begin{array}{l} LCL = np - 3 \cdot \sqrt{np \cdot (1-p)} \\ CL = np \\ UCL = np + 3 \cdot \sqrt{np \cdot (1-p)}. \end{array} \right. \quad (2.3.1)$$

Όταν ο αριθμός των ελαττωματικών προϊόντων σε μια διεργασία δεν είναι γνωστός, θα πρέπει να τον εκτιμήσουμε. Χρησιμοποιώντας, όπως και προηγουμένως, την ποσότητα \bar{P} ως εκτίμηση του ποσοστού p , δηλαδή $\hat{p} = \bar{P}$, τότε ο τύπος (2.3.1) γίνεται

$$\begin{array}{l} \text{Διάγραμμα } np \\ \text{με } 3\sigma \text{ Όρια Ελέγχου} \\ \text{Φάσης I} \end{array} \left\{ \begin{array}{l} LCL = n\bar{P} - 3 \cdot \sqrt{n\bar{P} \cdot (1-\bar{P})} \\ CL = n\bar{P} \\ UCL = n\bar{P} + 3 \cdot \sqrt{n\bar{P} \cdot (1-\bar{P})} \end{array} \right. \quad (2.3.2)$$

2.3.2 Παράδειγμα Διαγραμμάτων Ελέγχου για Ιδιότητες

Μια μονάδα κατασκευής ασυρμάτων αμυντικών προδιαγραφών, χρησιμοποιεί τρία μεταλλικά μέρη (πρόσοψη με επιφάνεια $0,12 \text{ m}^2$, πλευρικό κάλυμμα με επιφάνεια $0,34 \text{ m}^2$ και οπίσθιο κάλυμμα με επιφάνεια $0,11 \text{ m}^2$) κράματος αλουμινίου, συγκεκριμένων προδιαγραφών, ως πρώτη ύλη για τη συναρμολόγηση των ασυρμάτων.

Τα μεταλλικά μέρη, αφού παραληφθούν από τη μονάδα, υφίστανται μια διαδικασία εμβάπτισης σε ειδικά λουτρά και στη συνέχεια στεγνώματος σε ειδικούς φούρνους. Με τη διαδικασία αυτή, τα μεταλλικά μέρη αποκτούν μια προστατευτική επίστρωση από ένα υλικό με βασικό συστατικό το χρώμιο, έτσι ώστε να εξασφαλίζεται αντοχή σε διαβρωτικές συνθήκες.

Το Τμήμα Ελέγχου Ποιότητας της μονάδας, ελέγχει δειγματοληπτικά τα μεταλλικά μέρη μετά τη φάση του στεγνώματος για να εντοπίσει τυχόν ελαττώματα της προστατευτικής επίστρωσης (π.χ. ακάλυπτα σημεία με εμβαδόν μεγαλύτερο του 0.5 cm^2). Οι ελεγκτές επιλέγουν κάθε ημέρα, ορισμένα μεταλλικά μέρη και από τις τρεις κατηγορίες τα ελέγχουν οπτικά και καταγράφουν τα τυχόν ελαττώματα επίστρωσης σε κατάλληλο έντυπο. *Ειδικά για την μεταλλική πρόσοψη του ασυρμάτου, ο πελάτης απαιτεί την καταγραφή των απορριπτόμενων τεμαχίων και για το πλευρικό κάλυμμα απαιτεί την καταγραφή του αριθμού των ελαττωμάτων επίστρωσης.*

Στον παρακάτω πίνακα παρατίθεται το έντυπο που συμπληρώνεται, μετά από παραγωγή 15 ημερών. Σε αυτό καταγράφεται η ημέρα παραγωγής (και δειγματοληψίας), ο αριθμός των ελεγχθέντων μεταλλικών μερών για κάθε μία από τις τρεις κατηγορίες, και το συνολικό πλήθος ελαττωμάτων επίστρωσης που εντοπίστηκαν τη συγκεκριμένη μέρα σε όλα τα ελεγχθέντα μεταλλικά μέρη. Επίσης, στην προτελευταία στήλη του εντύπου, καταγράφονται τα ελαττώματα επίστρωσης του πλευρικού καλύμματος και στην τελευταία στήλη καταγράφονται τα απορριπτόμενα τεμάχια της πρόσοψης που προέκυψαν κατά τον παραπάνω έλεγχο.

Ημέρα παραγωγής	Ελεγχθέντα μεταλλικά μέρη			Αποτελέσματα ελέγχου		
	Πρόσοψη	Πλευρικό	Οπίσθιο	Συνολικά Ελαττώματα επίστρωσης	Ελαττώματα επίστρωσης πλευρικού καλύμματος	Απορριπτόμενες προσόψεις
1	10	4	5	8	2	2
2	10	6	6	8	1	3
3	10	5	7	9	1	4
4	10	3	4	10	0	2
5	10	2	7	11	1	4
6	10	4	8	6	0	3
7	10	7	3	12	2	3
8	10	5	4	7	1	4
9	10	4	5	14	2	2
10	10	2	2	8	1	4
11	10	6	8	6	4	3
12	10	5	4	9	4	2
13	10	5	6	11	4	9
14	10	5	2	16	5	3
15	10	2	4	10	2	4

Πίνακας 2.2: Ελεγχθέντα μεταλλικά μέρη και αποτελέσματα ελέγχου παραδείγματος διαγραμμάτων ελέγχου για ιδιότητες.

Τα 3σ όρια ελέγχου ενός διαγράμματος ελέγχου Φάσης I για τον αριθμό των απορριπτόμενων προσόψεων (np -chart) της παραπάνω διεργασίας (με κοινό μέγεθος δείγματος), τα οποία προκύπτουν από τον τύπο (2.3.2), είναι τα εξής

$$LCL = n\bar{P} - 3 \cdot \sqrt{n\bar{P} \cdot (1 - \bar{P})} = 10 \times 0.3466 - 3 \times \sqrt{10 \times 0.3466 \times (1 - 0.3466)} = -1.0486$$

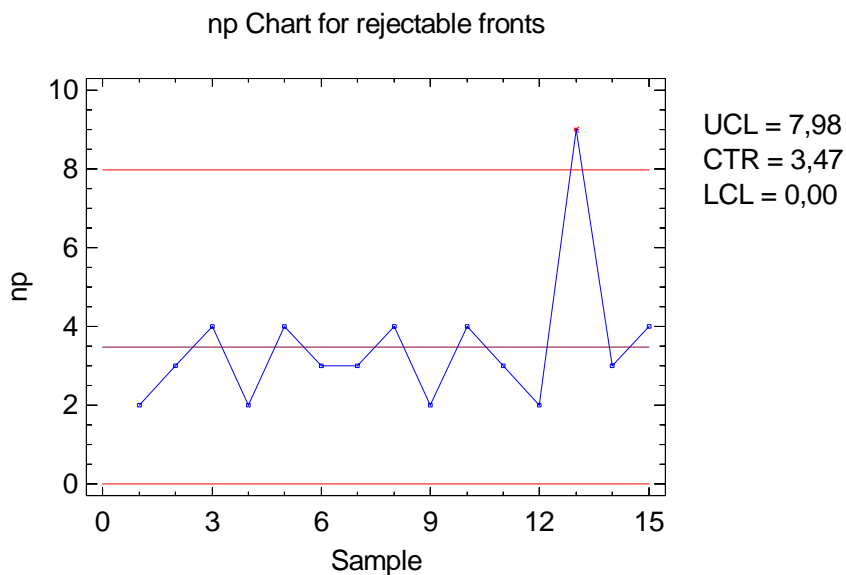
$$CL = \bar{P} = 10 \times 0.3466 = 3.466$$

$$UCL = n\bar{P} + 3 \cdot \sqrt{n\bar{P} \cdot (1 - \bar{P})} = 10 \times 0.3466 + 3 \times \sqrt{10 \times 0.3466 \times (1 - 0.3466)} = 7.9806.$$

Επειδή το κάτω όριο είναι αρνητικό, το θέτουμε ίσο με το μηδέν. Τα τελικά όρια ελέγχου που προκύπτουν είναι

$$\begin{cases} LCL = 0 \\ CL = 0.47 \\ UCL = 0.98. \end{cases}$$

Από το Σχήμα 2.7 που ακολουθεί, συμπεραίνουμε ότι το σημείο που αντιστοιχεί στη 13^η ημέρα είναι εκτός των ορίων ελέγχου. Συνεπώς, η διεργασία είναι εκτός στατικού ελέγχου και θα πρέπει να διερευνηθούν οι συνθήκες που επικρατούσαν τη συγκεκριμένη ημέρα (οι ειδικές αιτίες μεταβλητότητας).



Σχήμα 2.7: Διάγραμμα *np* Φάσης I για τον αριθμό των απορριπτόμενων προσόψεων

ΚΕΦΑΛΑΙΟ 3

Πολυμεταβλητά Παραμετρικά Διαγράμματα Ελέγχου

3.1 Εισαγωγή

Τα τελευταία χρόνια, η ραγδαία ανάπτυξη της τεχνολογίας και ιδιαίτερα η ολοένα αυξανόμενη ικανότητα των υπολογιστών να λαμβάνουν και να αποθηκεύουν μεγάλο όγκο δεδομένων αύξησε το ενδιαφέρον για την ταυτόχρονη επίβλεψη (έλεγχος) πολλών συσχετισμένων ποιοτικών χαρακτηριστικών ή πολλών μεταβλητών μιας διεργασίας (Lowty and Montgomery (1995), Runger and Montgomery (1997)). Αυτό γίνεται εφικτό με την εφαρμογή κάποιων τεχνικών που είναι γνωστές ως **πολυμεταβλητές στατιστικές τεχνικές ελέγχου διεργασιών** (*multivariate statistical process control procedures*). Τα βασικά εργαλεία του **πολυμεταβλητού στατιστικού ελέγχου διεργασιών** (*Multivariate Statistical Quality Control, MSPC*) είναι τα **πολυμεταβλητά (ή πολυδιάστατα) διαγράμματα ελέγχου** (*multivariate control charts*), τα οποία θα παρουσιάσουμε στο κεφάλαιο αυτό.

Η χρήση πολυμεταβλητών διαγραμμάτων ελέγχου καθίσταται αναγκαία, διότι αρκετά συχνά στον έλεγχο ποιότητας η ποιότητα ενός προϊόντος καθορίζεται όχι μόνο από την τιμή ενός ποιοτικού χαρακτηριστικού, αλλά από περισσότερα, τα οποία είναι συνήθως συσχετισμένα μεταξύ τους (Toralidou and Psarakis (2009)). Για παράδειγμα, οι φυσικές διαστάσεις, το πάχος, η στιλπνότητα, η ομοιομορφία και το φινίρισμα μιας επιφάνειας μπορούν να θεωρηθούν πολλαπλά χαρακτηριστικά που καθορίζουν την ποιότητα ενός προϊόντος. Γενικά, στον πολυμεταβλητό έλεγχο ποιότητας, μπορεί να υπάρχουν διαφορετικές μονάδες μέτρησης στα υπό εξέταση χαρακτηριστικά (π.χ. η θερμοκρασία, η πίεση, η

συγκέντρωση, οι υλικές ιδιότητες, ο ρυθμός ροής, η τάση κ.α.) μπορεί, όμως, και να προκύψουν μετρήσεις στην ίδια μονάδα, όπως μετρήσεις του πάχους σε διαφορετικές πλευρές μιας επιφάνειας ή ένα φάσμα απορρόφησης για διαφορετικά μήκη κύματος του φωτός.

Η πρώτη μορφή των σύγχρονων διαγραμμάτων ελέγχου προτάθηκε το 1931 από τον Shewhart, ο οποίος τα χρησιμοποίησε στη βιομηχανία (Shewhart (1931)). Ωστόσο, οι πολυμεταβλητές τεχνικές ελέγχου διεργασιών ουσιαστικά καθιερώθηκαν αρκετά χρόνια αργότερα. Η πρώτη πολυμεταβλητή προσέγγιση του ελέγχου διεργασιών έγινε στις δημοσιεύσεις του Hotelling το 1947 και το 1951 (Hotelling (1947), (1951)). Ο Hotelling εφάρμοσε την τεχνική του στις δοκιμές και τον έλεγχο των συσκευών (*bombsights*) που χρησιμοποιούνταν από βομβαρδιστικά αεροσκάφη για τη ρίψη βομβών σε επίγειους στόχους.

Στη συνέχεια, παρατηρήθηκε σημαντική αύξηση του ενδιαφέροντος των ερευνητών για την εφαρμογή πολυμεταβλητών μεθόδων στη βιομηχανική παραγωγή. Συγκεκριμένα, οι Knoth and Schmid (2002) υποστήριξαν ότι η επιτήρηση πολλών παραμέτρων μιας κατανομής, όπως η μέση τιμή και η διακύμανση, εντάσσεται στον τομέα του πολυμεταβλητού ελέγχου διεργασιών. Η Wärmefjord (2004) περιέγραψε την εφαρμογή πολυμεταβλητών διαγραμμάτων στη διαδικασία συναρμολόγησης τμημάτων της αυτοκινητοβιομηχανίας Saab. Ένα χρόνο αργότερα, οι Sahni, Aastveit and Naes (2005) πρότειναν ότι οι πρώτες ύλες και οι διάφορες μεταβλητές μιας διεργασίας της βιομηχανίας τροφίμων πρέπει να αναλυθούν με σκοπό τη διασφάλιση της ποιότητας του τελικού προϊόντος. Οι Tsung, Li and Jin (2008) υπογράμμισαν την ανάγκη εφαρμογής των πολυμεταβλητών διαγραμμάτων ελέγχου στις διαδικασίες παραγωγής και υπηρεσιών.

Τα τελευταία χρόνια, η εφαρμογή πολυμεταβλητών στατιστικών τεχνικών έχει επεκταθεί και σε άλλους τομείς πέρα από τη βιομηχανική παραγωγή. Για παράδειγμα, στην Αμερική, η μεθοδολογία επιτήρησης έχει αποκτήσει ιδιαίτερο ενδιαφέρον μετά την τρομοκρατική επίθεση της 11ης Σεπτεμβρίου του 2001. Επίσης, η χωρική επιτήρηση μπορεί να θεωρηθεί πολυμεταβλητή, από τη στιγμή που εμπλέκονται σε μία διεργασία διάφορες τοποθεσίες. Τέλος, προσπάθειες εφαρμογής πολυμεταβλητών τεχνικών επιτήρησης σε στρατηγικές λήψης αποφάσεων που αφορούν οικονομικά ζητήματα έχουν γίνει από τους Golosnoy, Schmid and Okhrin (2008) και τους Okhrin and Schmid (2008). Κατά τη Frisé (2011), αυτό οφείλεται στην ανάγκη για συνεχή επιτήρηση χρονικών σειρών με σκοπό την άμεση ανίχνευση μιας

σημαντικής αλλαγής της υποκείμενης διεργασίας από τη στιγμή που αυτή η αλλαγή συνέβη και μετά, στο βαθμό βέβαια που είναι εφικτό κάτι τέτοιο.

Εν κατακλείδι, θα μπορούσε κανείς να συμπεράνει ότι τα πολυμεταβλητά διαγράμματα ελέγχου έχουν κεντρίσει το ενδιαφέρον πολλών ερευνητών, οι οποίοι δημοσίευσαν κάποιες σημαντικές ανασκοπήσεις (*reviews*). Ενδεικτικά, αναφέρουμε τα άρθρα ανασκοπήσεις των Anderson (1984), Alt (1985), Jackson (1985), Alt and Smith (1988), Wierda (1994), Lowry and Montgomery (1995), MacGregor and Kourti (1995), Montgomery (1997), Yeh, Lin and McGrath (2006) και Bersimis, Psarakis and Panaretos (2007), καθώς και τη Διπλωματική Εργασία του Σταθάκη (2006).

3.2 Πολυμεταβλητά Διαγράμματα Ελέγχου Τύπου Shewhart για Μεταβλητές χωρίς μνήμη

Θεωρούμε ότι έχουμε στη διάθεσή μας m ανεξάρτητα τυχαία δείγματα από μία διεργασία, έστω $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \dots, \mathbf{X}_m$. Οι μετρήσεις που λαμβάνονται κάθε χρονική στιγμή αφορούν περισσότερα από ένα ποιοτικά χαρακτηριστικά ($p > 1$), τα οποία είναι συσχετισμένα μεταξύ τους. Επιπλέον θεωρούμε ότι η ποιότητα μιας παραγωγικής διαδικασίας μπορεί να μετρηθεί από το από κοινού επίπεδο των p συσχετισμένων ποιοτικών χαρακτηριστικών. Κάθε ένα από τα δείγματα αυτά εκτός από p ποιοτικά χαρακτηριστικά έχει και κοινό, συνήθως, μέγεθος n . Συνεπώς, κάθε δείγμα αποτελείται από $(p \times 1)$ διανύσματα, πλήθους n , έστω τα $\mathbf{X}_{1k}, \mathbf{X}_{2k}, \dots, \mathbf{X}_{nk}$. Οπότε, κάθε δείγμα θα έχει την ακόλουθη μορφή

$$\mathbf{X}_k = \{\mathbf{X}_{1k}, \mathbf{X}_{2k}, \dots, \mathbf{X}_{nk}\} = \left\{ \begin{pmatrix} X_{11k} \\ X_{12k} \\ \vdots \\ X_{1pk} \end{pmatrix}_{p \times 1}, \begin{pmatrix} X_{21k} \\ X_{22k} \\ \vdots \\ X_{2pk} \end{pmatrix}_{p \times 1}, \dots, \begin{pmatrix} X_{n1k} \\ X_{n2k} \\ \vdots \\ X_{npk} \end{pmatrix}_{p \times 1} \right\}, \quad k = 1, 2, \dots, m.$$

Η βασική υπόθεση που κάνουμε είναι ότι η κατανομή των p ποιοτικών χαρακτηριστικών είναι μία p -διάστατη κανονική κατανομή με μέση τιμή το διάνυσμα $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$ και πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$, δηλαδή $\mathbf{X}_{ik} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), i = 1, 2, \dots, n$. Θα

συμβολίζουμε τη μέση τιμή μ με μ_0 , όταν η διεργασία είναι εντός ελέγχου και με μ_1 , όταν η διεργασία είναι εκτός ελέγχου. Ο Σ είναι ένας $(p \times p)$ συμμετρικός, θετικά ορισμένος, πίνακας και ορίζεται ως εξής

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

Με $|\Sigma|$ και με Σ^{-1} συμβολίζονται η ορίζουσα και ο αντίστροφος του πίνακα Σ , αντίστοιχα.

Η συνάρτηση πυκνότητας του διανύσματος \mathbf{X}_{ik} δίνεται από τον τύπο

$$f_{\mathbf{X}}(\mathbf{X}_{ik}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \cdot e^{-\frac{1}{2}(\mathbf{X}_{ik}-\mu)' \Sigma^{-1}(\mathbf{X}_{ik}-\mu)}, \quad \mathbf{X}_{ik} \in \mathbb{R}^p.$$

Από τις ιδιότητες της πολυμεταβλητής κανονικής κατανομής αποδεικνύεται ότι $E(\mathbf{X}_{ik}) = \mu$,

$$\text{Cov}(\mathbf{X}_{ik}) = E\left\{(\mathbf{X}_{ik} - \mu)(\mathbf{X}_{ik} - \mu)'\right\} = \Sigma \quad \text{και} \quad (\mathbf{X}_{ik} - \mu)' \Sigma^{-1}(\mathbf{X}_{ik} - \mu) \sim \chi_p^2.$$

3.2.1 Εκτίμηση της Μέσης Τιμής μ και του Πίνακα Διακύμανσης-Συνδιακύμανσης Σ

Σύμφωνα με τον Montgomery (1997), στην πράξη οι ποσότητες μ και Σ δεν είναι γνωστές και πρέπει να εκτιμηθούν από τα προκαταρκτικά δείγματα μεγέθους n , τα οποία συλλέγονται από μία εντός ελέγχου διαδικασία. Θεωρούμε ότι έχουμε στη διάθεσή μας m τέτοια δείγματα. Οι δειγματικοί μέσοι και οι δειγματικές διακυμάνσεις υπολογίζονται από κάθε δείγμα ως εξής

$$\bar{X}_{jk} = \frac{1}{n} \sum_{i=1}^n X_{ijk}, \quad \begin{cases} j = 1, 2, \dots, p \\ k = 1, 2, \dots, m \end{cases}$$

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

όπου X_{ijk} είναι η i παρατήρηση του j χαρακτηριστικού στο k δείγμα. Η συνδιακύμανση ανάμεσα στο j και στο h ποιοτικό χαρακτηριστικό του k δείγματος είναι οι εξής

$$S_{jhk} = \frac{1}{n-1} \sum_{i=1}^n (X_{ijk} - \bar{X}_{jk})(X_{ihk} - \bar{X}_{hk}), \quad \begin{cases} k=1, 2, \dots, m \\ j \neq h. \end{cases}$$

Οι μέσοι όροι των παραπάνω στατιστικών συναρτήσεων για όλα τα δείγματα είναι

$$\bar{\bar{X}}_j = \frac{1}{m} \cdot \sum_{k=1}^m \bar{X}_{jk}, \quad j=1, 2, \dots, p,$$

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

και

$$\bar{S}_{jh} = \frac{1}{m} \cdot \sum_{k=1}^m S_{jhk}, \quad j \neq h.$$

Οι ποσότητες $\{\bar{\bar{X}}_j\}$ είναι τα στοιχεία του διανύσματος $\bar{\bar{\mathbf{X}}}$. Ο ολικός δειγματικός πίνακας διακύμανσης-συνδιακύμανσης $\mathbf{S}_{(p \times p)}$ σχηματίζεται ως εξής

$$\mathbf{S} = \begin{bmatrix} \bar{S}_1^2 & \bar{S}_{12} & \bar{S}_{13} & \cdots & \bar{S}_{1p} \\ \bar{S}_{21} & \bar{S}_2^2 & \bar{S}_{23} & \cdots & \bar{S}_{2p} \\ \bar{S}_{31} & \bar{S}_{32} & \bar{S}_3^2 & \cdots & \bar{S}_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{S}_{p1} & \bar{S}_{p2} & \bar{S}_{p3} & \cdots & \bar{S}_p^2 \end{bmatrix},$$

όπου $\bar{S}_{ij} = \bar{S}_{ji}$ για κάθε $i \neq j$ με $i, j=1, 2, \dots, p$. Όταν η διεργασία είναι εντός στατιστικού ελέγχου, ο πίνακας \mathbf{S} είναι αμερόληπτος εκτιμητής του πίνακα Σ .

3.2.2 Πολυμεταβλητά Διαγράμματα Ελέγχου για τη Μέση Τιμή

Ο σκοπός χρήσης των διαγραμμάτων αυτών είναι η παρακολούθηση της μέσης τιμής της διεργασίας. Στην περίπτωση των πολυμεταβλητών διαγραμμάτων η μέση τιμή είναι ένα $(p \times 1)$ διάνυσμα. Δε θα πρέπει να παραλείψουμε να αναφέρουμε ότι για την ανάπτυξη των διαγραμμάτων ελέγχου για τη μέση τιμή, θεωρούμε ότι η διασπορά παραμένει σταθερή σε όλη τη διάρκεια της διεργασίας.

❖ Διαγράμματα Ελέγχου Μέσης Τιμής για Δείγματα

Θεωρούμε ότι $\mathbf{X}_{ik} = (X_{i1k}, X_{i2k}, \dots, X_{ipk})'$ είναι ένα $(p \times 1)$ τυχαίο διάνυσμα (τ.δ.), το οποίο αναπαριστά p συσχετισμένα ποιοτικά χαρακτηριστικά που προέρχονται από μία πολυδιάστατη διεργασία. Σκοπός μας είναι να ελέγξουμε την ποιότητα της διεργασίας. Εάν η διεργασία είναι εντός ελέγχου, το \mathbf{X}_{ik} ακολουθεί την p -διάστατη κανονική κατανομή με μέσο το διάνυσμα $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$ και πίνακα συνδιακύμανσης $\boldsymbol{\Sigma}_{(p \times p)}$, συμβολικά $\mathbf{X}_{ik} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Έστω

$$\mathbf{X}_k = (\mathbf{X}_{1k}, \mathbf{X}_{2k}, \dots, \mathbf{X}_{nk}) = \begin{pmatrix} X_{11k} & X_{21k} & \dots & X_{n1k} \\ X_{12k} & X_{22k} & \dots & X_{n2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1pk} & X_{2pk} & \dots & X_{npk} \end{pmatrix}, k = 1, 2, \dots, m$$

ένα τυχαίο δείγμα μεγέθους n από διανύσματα \mathbf{X}_{ik} . Επιπλέον, θεωρούμε ότι τα $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ είναι m ανεξάρτητα δείγματα (μεγέθους n) διαθέσιμα από τη διεργασία.

Η κατασκευή ενός πολυμεταβλητού διαγράμματος ελέγχου για την παρακολούθηση της μέσης τιμής μπορεί να βασιστεί στην εξής στατιστική συνάρτηση

$$\chi_k^2 = n(\bar{\mathbf{X}}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}}_k - \boldsymbol{\mu}), k = 1, 2, \dots, m, \quad (3.2.1)$$

όπου $\bar{\mathbf{X}}_k$ είναι το διάνυσμα των δειγματικών μέσων της k ορθολογικής υποομάδας, $\boldsymbol{\mu}$ το διάνυσμα στόχος των μέσων και $\boldsymbol{\Sigma}$ ο πίνακας διακύμανσης-συνδιακύμανσης. Το διάνυσμα $\boldsymbol{\mu}$ και ο πίνακας $\boldsymbol{\Sigma}$ είναι γνωστά. Η ποσότητα χ_k^2 είναι ανάλογη προς την απόσταση του Mahalanobis κάθε σημείου από το διάνυσμα στόχο $\boldsymbol{\mu}$. Το διάγραμμα ελέγχου δίνει σήμα ότι παρουσιάστηκε μια στατιστικά σημαντική μετατόπιση του μέσου, όταν ισχύει

$$n(\bar{\mathbf{X}}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}}_k - \boldsymbol{\mu}) > UCL,$$

όπου $UCL > 0$ είναι ένα καθορισμένο άνω όριο ελέγχου. Συνεπώς, η διεργασία τίθεται εκτός ελέγχου. Η στατιστική συνάρτηση χ_k^2 ακολουθεί μία κατανομή χ^2 με p βαθμούς ελευθερίας (β.ε.), συμβολικά χ_p^2 .

Για το λόγο αυτό, ένα πολυμεταβλητό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση του μέσου της διεργασίας (*multivariate Shewhart control chart for the process mean*), με γνωστό διάνυσμα μέσων μ και γνωστό πίνακα διακύμανσης-συνδιακύμανσης Σ , έχει τα ακόλουθα όρια ελέγχου

$$\begin{aligned} \text{Διάγραμμα } \chi^2 \\ \text{Φάσης II} \end{aligned} \begin{cases} LCL = 0 \\ UCL = \chi_{p,\alpha}^2. \end{cases} \quad (3.2.2)$$

Το διάγραμμα αυτό είναι γνωστό ως **διάγραμμα ελέγχου χ^2** (χ^2 -control chart) ή ως **διάγραμμα ελέγχου χ^2 Φάσης II** (*Phase II χ^2 -chart*). Συνήθως, χρησιμοποιείται μόνο το άνω όριο ελέγχου του παραπάνω διαγράμματος, επειδή ακραίες τιμές της στατιστικής συνάρτησης χ_k^2 αντιστοιχούν σε σημεία μακριά από τις τιμές στόχου του διανύσματος μ , ενώ μικρές ή μηδενικές τιμές της συνάρτησης χ_k^2 αντιστοιχούν σε σημεία κοντά στο μ (Bersimis, Psarakis and Panaretos (2007)).

Το διάγραμμα ελέγχου χ^2 είναι ουσιαστικά μια οπτική απεικόνιση του ελέγχου λόγου πιθανοφανειών ($H_0: \mu = \mu_0$ κατά $H_1: \mu \neq \mu_0$ ή $H_1: \mu = \mu_1$) για διαδοχικά σημεία (Alt (1985)). Θα πρέπει να αναφέρουμε ότι το διάγραμμα ελέγχου χ^2 Φάσης II αποτελεί γενίκευση του διαγράμματος \bar{X} με όρια πιθανότητας α Φάσης II του τύπου (2.2.2).

Στην πράξη, το διάνυσμα μέσων μ και ο πίνακας διακύμανσης-συνδιακύμανσης Σ δεν είναι γνωστά και πρέπει να εκτιμηθούν. Οι αμερόληπτες εκτιμήσεις που χρησιμοποιούμε είναι $\bar{\bar{X}}$ και \bar{S} , αντίστοιχα. Τα διαγράμματα ελέγχου που χρησιμοποιούνται στην περίπτωση αυτή είναι διαγράμματα ελέγχου Φάσης I. Η Φάση I μπορεί να χωριστεί σε δύο στάδια. Στο **Στάδιο I** (*Stage I*) ελέγχουμε εάν η διεργασία ήταν εντός στατιστικού ελέγχου όταν επιλέχθηκαν οι πρώτες υποομάδες, ενώ στο **Στάδιο II** (*Stage II*) ελέγχουμε εάν η διεργασία παραμένει εντός στατιστικού ελέγχου όταν επιλέγονται μελλοντικές υποομάδες.

Θα παρουσιάσουμε πρώτα τα διαγράμματα ελέγχου Φάσης I για το Στάδιο I. Σύμφωνα με τον Ryan (1989), αντικαθιστώντας το διάνυσμα μ με το ολικό δειγματικό διάνυσμα $\bar{\bar{X}}$ και τον πίνακα Σ με το συνολικό (*pooled*) δειγματικό πίνακα διακύμανσης-συνδιακύμανσης \bar{S} , ο τύπος (3.2.1) παίρνει την ακόλουθη μορφή

$$T_{k,1}^2 = n(\bar{X}_k - \bar{\bar{X}})' S^{-1} (\bar{X}_k - \bar{\bar{X}}), \quad k = 1, 2, \dots, m, \quad (3.2.3)$$

όπου $\bar{\mathbf{X}}_k$ είναι το διάνυσμα των δειγματικών μέσων της k ορθολογικής υποομάδας. Αποδεικνύεται ότι η ποσότητα $T_{k,1}^2 / c_1(m, n, p)$ ακολουθεί την κατανομή F με p και $mn - m - p + 1$ β.ε., συμβολικά γράφουμε $T_{k,1}^2 \sim c_1(m, n, p) \times F_{p, mn-m-p+1}$. Η ποσότητα $c_1(m, n, p)$ είναι μία σταθερά που εξαρτάται από τον αριθμό των ποιοτικών χαρακτηριστικών p , το μέγεθος δείγματος n και τον αριθμό των δειγμάτων m που λαμβάνονται στο χρόνο και υπολογίζεται από τον τύπο

$$c_1(m, n, p) = \frac{p(m-1)(n-1)}{mn - m - p + 1}.$$

Ένα πολυμεταβλητό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση του μέσου της διεργασίας, με άγνωστο διάνυσμα μέσων $\boldsymbol{\mu}$ και άγνωστο πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$, έχει τα όρια ελέγχου που ακολουθούν

$$\begin{aligned} \text{Διάγραμμα } T^2 \\ \text{Φάσης I (Στάδιο I)} \end{aligned} \begin{cases} LCL = 0 \\ UCL = c_1(m, n, p) \times F_{p, mn-m-p+1; \alpha} \end{cases} \quad (3.2.4)$$

Το παραπάνω διάγραμμα είναι γνωστό ως **διάγραμμα ελέγχου T^2 του Hotelling Φάσης I-Σταδίου I** (*Phase I-Stage I Hotelling's T^2 -chart*).

Ο Alt (1985) δεν πολλαπλασιάζει τη στατιστική συνάρτηση $T_{k,1}^2$ με n , διότι διαιρεί τη σταθερά $c_1(m, n, p)$ με n . Συνεπώς, ο τύπος (3.2.1) παίρνει τη μορφή

$$T_{k,1}^{*2} = (\bar{\mathbf{X}}_k - \bar{\bar{\mathbf{X}}})' \mathbf{S}^{-1} (\bar{\mathbf{X}}_k - \bar{\bar{\mathbf{X}}}), \quad k = 1, 2, \dots, m,$$

όπου $\bar{\mathbf{X}}_k$ είναι το διάνυσμα των δειγματικών μέσων της k ορθολογικής υποομάδας. Οι Alt, Goode and Wadsworth (1976) έδειξαν ότι η ποσότητα $T_{k,1}^{*2} / c_1^*(m, n, p)$ ακολουθεί την κατανομή F με p και $mn - m - p + 1$ β.ε., συμβολικά $T_{k,1}^{*2} \sim c_1^*(m, n, p) \times F_{p, mn-m-p+1}$, όπου η σταθερά $c_1^*(m, n, p)$ υπολογίζεται από τον τύπο

$$c_1^*(m, n, p) = \frac{p(m-1)(n-1)}{n(mn - m - p + 1)}.$$

Επομένως, τα όρια του **διαγράμματος T^2 του Hotelling Φάσης I (Στάδιο I)** έχουν τη μορφή

$$\begin{aligned} \text{Διάγραμμα } T^2 \\ \text{Φάσης I (Στάδιο I)} \end{aligned} \begin{cases} LCL = & 0 \\ UCL = & c_1^*(m, n, p) \times F_{p, mn-m-p+1; a} \end{cases} \quad (3.2.5)$$

Στο σημείο αυτό, θα πρέπει να αναφέρουμε ότι οι Hillier (1969) και Yang and Hillier (1970) πρότειναν μία μονοδιάστατη μέθοδο και για τα δύο στάδια της Φάσης I, χρησιμοποιώντας όρια πιθανότητας. Στη μέθοδο αυτή βασίστηκαν οι Alt, Goode and Wadsworth (1976) και την γενίκευσαν για την πολυμεταβλητή περίπτωση. Παρατηρούμε ότι για $p=1$, το άνω όριο ελέγχου και η στατιστική συνάρτηση του διαγράμματος ελέγχου T^2 του Hotelling Φάσης I (Στάδιο I) είναι αντίστοιχα

$$UCL = \frac{m-1}{m} \times F_{1, m(n-1); a} \quad \text{και} \quad T_{k,1}^2 = \frac{(\bar{X}_k - \bar{\bar{X}})^2}{\bar{S}^2}, \quad k=1, 2, \dots, m,$$

όπου \bar{S}^2 είναι ο μέσος των δειγματικών διακυμάνσεων που προκύπτει για κάθε δείγμα.

Επειδή $F_{1, m(n-1); a} = t_{m(n-1); a/2}^2$, για $A_4 = \sqrt{(m-1)/m} \cdot t_{m(n-1); a/2}$ θα ισχύει

$$\begin{aligned} 1-a &= P\left[(\bar{X}_k - \bar{\bar{X}})^2 / \bar{S}^2 \leq \frac{m-1}{m} \cdot F_{1, m(n-1); a} \right] = P\left[|\bar{X}_k - \bar{\bar{X}}| \leq \sqrt{\bar{S}^2 \cdot \frac{m-1}{m}} \cdot t_{m(n-1); a/2} \right] = \\ &= P\left[\bar{\bar{X}} - A_4 \cdot \sqrt{\bar{S}^2} \leq \bar{X}_k \leq \bar{\bar{X}} + A_4 \cdot \sqrt{\bar{S}^2} \right] = P\left[\bar{\bar{X}} - A_4 \cdot \bar{S} \leq \bar{X}_k \leq \bar{\bar{X}} + A_4 \cdot \bar{S} \right]. \end{aligned}$$

Επιπλέον, τα διαστήματα εμπιστοσύνης Bonferroni για κάθε επιμέρους χαρακτηριστικό προκύπτουν χρησιμοποιώντας $A_4 = \sqrt{(m-1)/m} \cdot t_{m(n-1); a/2p}$. Για $p=2$, τα όρια ελέγχου για κάθε μεταβλητή υπολογίζονται ως εξής

$$\begin{cases} LCL = & \bar{\bar{X}} - A_4 \cdot \bar{S} \\ UCL = & \bar{\bar{X}} + A_4 \cdot \bar{S}. \end{cases}$$

Στη συνέχεια, θα παρουσιάσουμε τα διαγράμματα ελέγχου Φάσης I για το Στάδιο II. Σύμφωνα με τον Ryan (1989), όταν θέλουμε να απεικονίσουμε και να αναλύσουμε **μελλοντικές υποομάδες** (*future subgroups*), η στατιστική συνάρτηση που χρησιμοποιείται έχει την ακόλουθη μορφή

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

όπου $\bar{\mathbf{X}}_f$ είναι ένα $(p \times 1)$ διάνυσμα των δειγματικών μέσων μιας μελλοντικής υποομάδας.

Αποδεικνύεται ότι η ποσότητα $T_{k,2}^2 / c_2(m, n, p)$ ακολουθεί την κατανομή F με p και

$mn - m - p + 1$ β.ε., συμβολικά γράφουμε $T_{k,2}^2 \sim c_2(m, n, p) \times F_{p, mn-m-p+1}$, όπου η $c_2(m, n, p)$ είναι μία σταθερά που εξαρτάται από τον αριθμό των ποιοτικών χαρακτηριστικών p , το μέγεθος δείγματος n και τον αριθμό των δειγμάτων m που λαμβάνονται στο χρόνο και υπολογίζεται από τον τύπο

$$c_2(m, n, p) = \frac{p(m+1)(n-1)}{mn-m-p+1}.$$

Ένα πολυμεταβλητό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση του μέσου της διεργασίας, με άγνωστο διάνυσμα μέσων μ και άγνωστο πίνακα διακύμανσης-συνδιακύμανσης Σ , έχει όρια ελέγχου

$$\begin{aligned} \text{Διάγραμμα } T^2 \\ \text{Φάσης I (Στάδιο II)} \end{aligned} \begin{cases} LCL = 0 \\ UCL = c_2(m, n, p) \times F_{p, mn-m-p+1; \alpha} \end{cases} \quad (3.2.6)$$

Το παραπάνω διάγραμμα είναι γνωστό ως **διάγραμμα ελέγχου T^2 του Hotelling Φάσης I-Σταδίου II** (*Phase I-Stage II Hotelling's T^2 -chart*).

Ο Alt (1985) αντί να πολλαπλασιάσει τη στατιστική συνάρτηση $T_{k,1}^2$ με n , διαιρεί τη σταθερά $c_2(m, n, p)$ με n , όπως έκανε αντίστοιχα και στο Στάδιο I. Συνεπώς, η στατιστική συνάρτηση μπορεί να πάρει τη μορφή

$$T_{k,2}^{*2} = (\bar{\mathbf{X}}_f - \bar{\bar{\mathbf{X}}})' \mathbf{S}^{-1} (\bar{\mathbf{X}}_f - \bar{\bar{\mathbf{X}}}),$$

όπου $\bar{\mathbf{X}}_f$ είναι ένα $(p \times 1)$ διάνυσμα των δειγματικών μέσων μιας μελλοντικής υποομάδας.

Ο Alt (1985) απέδειξε ότι η ποσότητα $T_{k,2}^{*2} / c_2^*(m, n, p)$ ακολουθεί την κατανομή F με p και $mn - m - p + 1$ β.ε., συμβολικά γράφουμε $T_{k,2}^{*2} \sim c_2^*(m, n, p) \times F_{p, mn-m-p+1}$, όπου η σταθερά $c_2^*(m, n, p)$ υπολογίζεται από τον τύπο

$$c_2^*(m, n, p) = \frac{p(m+1)(n-1)}{n(mn-m-p+1)}.$$

Επομένως, τα όρια ελέγχου του **διαγράμματος T^2 του Hotelling Φάσης I (Στάδιο II)** είναι

$$\begin{aligned} \text{Διάγραμμα } T^2 \\ \text{Φάσης I (Στάδιο II)} \end{aligned} \begin{cases} LCL = 0 \\ UCL = c_2^*(m, n, p) \times F_{p, mn-m-p+1; \alpha} \end{cases} \quad (3.2.7)$$

Οι Yang and Hillier (1970) πρότειναν ότι τα $\bar{\bar{X}}$, \bar{S} και το άνω όριο ελέγχου του τύπου (3.2.7) θα πρέπει να ενημερώνονται αρκετά συχνά στην αρχή της διεργασίας και λιγότερο συχνά από τη στιγμή που η διεργασία σταθεροποιείται και μετά. Το διάγραμμα ελέγχου T^2 του Hotelling Φάσης I (Στάδιο II) μπορεί να κατασκευαστεί από επιμέρους διαγράμματα για κάθε ποιοτικό χαρακτηριστικό. Για $p = 2$, τα όρια ελέγχου για κάθε μεταβλητή είναι τα εξής

$$\begin{cases} LCL &= \bar{\bar{X}} - A_4^* \cdot \bar{S} \\ UCL &= \bar{\bar{X}} + A_4^* \cdot \bar{S}, \end{cases}$$

όπου $A_4^* = \sqrt{(m+1)/m} \cdot t_{m(n-1), a/2p}$.

Υπογραμμίζουμε το γεγονός ότι πολλοί ερευνητές, όπως για παράδειγμα ο Montgomery (1997), δε διαχωρίζουν τη Φάση I σε επιμέρους στάδια, για αποφυγή οποιασδήποτε σύγχυσης. Με αποτέλεσμα, να χαρακτηρίζουν όλα τα διαγράμματα του Σταδίου II της Φάσης I ως διαγράμματα Φάσης II. Επιπρόσθετα, αναφέρουμε ότι όταν το διάλυσμα μ και ο πίνακας Σ εκτιμώνται από μεγάλο αριθμό προκαταρκτικών δειγμάτων, συνηθίζεται να χρησιμοποιείται το $UCL = \chi_{p,a}^2$ ως άνω όριο ελέγχου, τόσο στη Φάση I, όσο και στη Φάση II. Στη μονοδιάστατη περίπτωση, και συγκεκριμένα στο διάγραμμα \bar{X} , γνωρίζουμε ότι όταν το m παίρνει τιμές μεγαλύτερες από 20 ή 25, η διάκριση μεταξύ της Φάσης I και II δεν είναι συνήθως απαραίτητη, αφού οι τιμές των ορίων ελέγχου των δύο Φάσεων είναι πολύ κοντά. Ωστόσο, με τα πολυμεταβλητά όρια ελέγχου θα πρέπει να είμαστε αρκετά προσεκτικοί.

Οι Lowry and Montgomery (1995) έδειξαν ότι σε πολλές περιπτώσεις χρειάζεται ένας αρκετά μεγάλος αριθμός προκαταρκτικών δειγμάτων, ώστε τα ακριβή όρια ελέγχου Φάσης II να προσεγγίζονται με μεγάλη ακρίβεια από τα χ^2 -όρια ελέγχου. Οι συγγραφείς αυτοί παρουσίασαν, επίσης, πίνακες στους οποίους προτείνουν την ελάχιστη τιμή που πρέπει να πάρει το m για μέγεθος δείγματος $n = 2, 3, 10$ και για $p = 2, 3, 4, 5, 10, 20$ ποιοτικά χαρακτηριστικά. Οι προτεινόμενες τιμές του m είναι πάντα μεγαλύτερες από 20 και πολλές φορές άνω των 50 προκαταρκτικών δειγμάτων.

Σύμφωνα με τους Lowry and Montgomery (1995) και Montgomery (1997), θα πρέπει να είμαστε ιδιαίτερα προσεκτικοί με τα στατιστικά προγράμματα που χρησιμοποιούμε, διότι αρκετά συχνά ο τρόπος υπολογισμού των ορίων ελέγχου δεν είναι σωστός. Για παράδειγμα, το άνω όριο ελέγχου του διαγράμματος T^2 του Hotelling υπολογίζεται μέσω του τύπου

$$UCL = \frac{p(m-1)}{m-p} \times F_{p, m-p; a}$$

και όχι μέσω των τύπων (3.2.4) και (3.2.5) για τη Φάση I ή μέσω των τύπων (3.2.6) και (3.2.7) για τη Φάση II. Προφανώς, το παραπάνω όριο δεν είναι σωστό. Μπορεί να δημιουργεί τη σωστή κρίσιμη περιοχή για έναν πολυμεταβλητό στατιστικό έλεγχο υπόθεσης που αφορά το μέσο διάνυσμα μ , όταν ένα δείγμα μεγέθους m επιλέγεται τυχαία από μία p -διάστατη κανονική κατανομή, αλλά δεν μπορεί να εφαρμοστεί άμεσα στο διάγραμμα ελέγχου.

❖ Διαγράμματα Ελέγχου Μέσης Τιμής για Μεμονωμένες Παρατηρήσεις

Οι βασικές υποθέσεις που κάνουμε τώρα είναι ότι έχουμε στη διάθεσή μας m δείγματα μεγέθους $n=1$ το καθένα και ότι p είναι ο αριθμός των ποιοτικών χαρακτηριστικών που επιθεωρούνται σε κάθε δείγμα. Η κατασκευή διαγραμμάτων ελέγχου για την παρακολούθηση της μέσης τιμής στην περίπτωση των μεμονωμένων παρατηρήσεων μπορεί να βασιστεί στην ακόλουθη στατιστική συνάρτηση

$$\chi_k^2 = (\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu}), \quad k=1, 2, \dots, m,$$

όπου \mathbf{X}_k είναι η k παρατήρηση που ακολουθεί p -διάστατη κανονική κατανομή $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Το $\boldsymbol{\mu}$ είναι το διάνυσμα των μέσων και ο $\boldsymbol{\Sigma}$ είναι ο πίνακας διακύμανσης-συνδιακύμανσης, τα οποία είναι γνωστά. Υποθέτουμε ακόμη ότι οι παρατηρήσεις \mathbf{X}_k είναι ανεξάρτητες μεταξύ τους. Σύμφωνα με τον Seber (1984), η συνάρτηση χ_k^2 ακολουθεί τη χ^2 κατανομή με p β.ε., συμβολικά χ_p^2 .

Ένα πολυμεταβλητό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση του μέσου της διεργασίας στην περίπτωση μεμονωμένων παρατηρήσεων, με γνωστό διάνυσμα μέσων $\boldsymbol{\mu}$ και γνωστό πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$, έχει τα όρια ελέγχου Φάσης II που ακολουθούν

$$\begin{array}{l} \text{Διάγραμμα } \chi^2 \\ \text{Φάσης II} \end{array} \begin{cases} LCL = 0 \\ UCL = \chi_{p; a}^2. \end{cases} \quad (3.2.8)$$

Το διάγραμμα αυτό είναι γνωστό ως **διάγραμμα ελέγχου χ^2 Φάσης II** (*Phase II χ^2 -chart*).

Αντικαθιστώντας το διάνυσμα μ με τον ολικό δειγματικό μέσο $\bar{\mathbf{X}}$ και τον πίνακα Σ με τον ολικό δειγματικό πίνακα διακύμανσης-συνδιακύμανσης \mathbf{S} , η στατιστική συνάρτηση παίρνει την ακόλουθη μορφή

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

Η ποσότητα $T_k^2 / d_0(m)$, όπου η σταθερά $d_0(m)$ υπολογίζεται από τον τύπο

$$d_0(m) = (m-1)^2 / m,$$

ακολουθεί την κατανομή Βήτα με $p/2$ και $m-p-1$ β.ε., συμβολικά $T_k^2 \sim d_0(m) \times B_{p/2, (m-p-1)/2}$. Οπότε, ένα πολυμεταβλητό διάγραμμα τύπου Shewhart για την παρακολούθηση του μέσου της διεργασίας στην περίπτωση των μεμονωμένων παρατηρήσεων, με άγνωστο διάνυσμα μέσων μ και άγνωστο πίνακα διακύμανσης-συνδιακύμανσης Σ , έχει όρια ελέγχου Φάσης I

$$\begin{aligned} \text{Διάγραμμα } T^2 \\ \text{Φάσης I} \end{aligned} \begin{cases} LCL = d_0(m) \times B_{p/2, (m-p-1)/2; 1-a/2} \\ UCL = d_0(m) \times B_{p/2, (m-p-1)/2; a/2} \end{cases} \quad (3.2.9)$$

Το παραπάνω διάγραμμα είναι γνωστό ως **διάγραμμα ελέγχου T^2 (του Hotelling) Φάσης I** (*Phase I Hotelling's T^2 -chart*). Τα όρια ελέγχου του τύπου (3.2.9) προτάθηκαν από τους Tracy, Young and Mason (1992). Ωστόσο, το κάτω όριο ελέγχου τίθεται συνήθως ίσο με μηδέν και το άνω όριο υπολογίζεται για a και όχι για $a/2$.

Όταν θέλουμε να απεικονίσουμε και να αναλύσουμε **μελλοντικές υποομάδες** (*future subgroups*), η στατιστική συνάρτηση που χρησιμοποιείται έχει τη μορφή

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

όπου \mathbf{X}_f είναι μία μελλοντική μεμονωμένη παρατήρηση, η οποία είναι ανεξάρτητη από τις εκτιμήσεις $\bar{\mathbf{X}}$ και \mathbf{S} . Αποδεικνύεται ότι η ποσότητα $T_k^2 / d_1(m, p)$, όπου $d_1(m, p)$ είναι

$$d_1(m, p) = p(m+1)(m-1) / m(m-p),$$

ακολουθεί την κατανομή F με p και $(m-p)$ β.ε., συμβολικά $T_k^2 \sim d_1(m, p) \times F_{p, m-p; a}$. Ένα πολυμεταβλητό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση του μέσου της

διεργασίας στην περίπτωση των μεμονωμένων παρατηρήσεων, με άγνωστες παραμέτρους, έχει τα όρια ελέγχου που ακολουθούν

$$\begin{array}{l} \text{Διάγραμμα } T^2 \\ \text{Φάσης II} \end{array} \begin{cases} LCL = d_1(m, p) \times F_{p, m-p; 1-a/2} \\ UCL = d_1(m, p) \times F_{p, m-p; a/2} \end{cases}$$

Το παραπάνω άνω όριο προτάθηκε από τον Ryan (1989), ενώ το αντίστοιχο κάτω όριο ελέγχου από τους Tracy, Young and Mason (1992). Συνήθως, όμως, το κάτω όριο ελέγχου τίθεται ίσο με μηδέν και το άνω όριο δεν υπολογίζεται για $a/2$, αλλά για a .

Στην περίπτωση που ο αριθμός των προκαταρκτικών δειγμάτων m είναι μεγάλος ($m > 100$), ο Jackson (1985) προτείνει ως όρια τα εξής

$$\begin{array}{l} \text{Διάγραμμα } T^2 \\ \text{Φάσης II} \end{array} \begin{cases} LCL = 0 \\ UCL = \frac{p(m-1)}{m-p} \times F_{p, m-p; a} \end{cases} \quad (3.2.10)$$

Άλλοι ερευνητές προτείνουν εναλλακτικά τα όρια ελέγχου του τύπου (3.2.8). Τα όρια του τύπου (3.2.10) είναι μία λογική προσέγγιση, σε αντίθεση με τα όρια του τύπου (3.2.8) που θα πρέπει να χρησιμοποιούνται με προσοχή. Σύμφωνα με τους Lowry and Montgomery (1995) και Montgomery (1997), εάν το p είναι μεγάλο ($p > 10$), τότε θα πρέπει να έχουμε στη διάθεση μας τουλάχιστον 250 δείγματα ($m \geq 250$), ώστε το χ -τετράγωνο άνω όριο ελέγχου να είναι μία λογική προσέγγιση της σωστής τιμής.

Τα διαγράμματα χ^2 και T^2 , ως διαγράμματα ελέγχου τύπου Shewhart, χρησιμοποιούν μόνο τις πληροφορίες που δίνει το πιο πρόσφατο δείγμα. Κατά συνέπεια, δεν είναι πολύ αποτελεσματικά στην ανίχνευση μεσαίων ή μικρών μετατοπίσεων του μέσου διανύσματος της διεργασίας. Συχνά, για την αύξηση της ευαισθησίας των διαγραμμάτων αυτών, χρησιμοποιείται η θεωρία των κανόνων ροών (*runs rules*). Για παράδειγμα, οι Koutras, Bersimis and Antzoulakos (2006), εφαρμόζοντας τους κανόνες ροής σε ένα πολυμεταβλητό διάγραμμα χ^2 τύπου Shewhart, παρουσιάζουν δύο τροποποιημένα διαγράμματα χ^2 , τα οποία βελτιώνουν την επίδοση του κλασσικού διαγράμματος χ^2 . Το πρώτο διάγραμμα ελέγχου χρησιμοποιεί έναν κανόνα ροής k -από- k ($k|k$) και έχει καλύτερο *ARL* από το διάγραμμα $1|1$, δηλαδή από το κλασσικό διάγραμμα χ^2 . Το δεύτερο διάγραμμα που πρότειναν είναι το συνδυασμένο διάγραμμα τύπου $r|r-k|k$, το οποίο στην ειδική περίπτωση $r=1$ με $k \geq 2$

είναι πιο αποτελεσματικό και από τα δύο διαγράμματα $1|1$ και $k|k$. Επιπλέον, οι Rakitzis and Antzoulakos (2011) προτείνουν το διάγραμμα $r|m$, το οποίο βασίζεται σε έναν κανόνα ροής r -από- m και αποτελεί καλύτερη εναλλακτική όχι μόνο του κλασσικού διαγράμματος χ^2 , αλλά και άλλων διαγραμμάτων που βασίζονται σε κανόνες ροής. Η αποτελεσματικότητα του διαγράμματος αυτού βελτιώνεται καθώς αυξάνεται ο αριθμός των ποιοτικών χαρακτηριστικών.

3.2.3 Πολυμεταβλητά Διαγράμματα Ελέγχου για τη Διασπορά

Ο σκοπός χρήσης των διαγραμμάτων αυτών είναι η παρακολούθηση της διασποράς της διεργασίας. Στην περίπτωση των πολυμεταβλητών διαγραμμάτων ελέγχου η διασπορά είναι ένας $(p \times p)$ πίνακας συνδιακύμανσης Σ . Τα στοιχεία της κύριας διαγωνίου του πίνακα είναι οι διακυμάνσεις των μεμονωμένων μεταβλητών της διεργασίας, ενώ τα μη διαγώνια στοιχεία είναι οι συνδιακυμάνσεις. Ο Alt (1985) παρουσιάζει δύο πολύ χρήσιμες μεθόδους.

Η πρώτη μέθοδος είναι η επέκταση του μονοδιάστατου διαγράμματος έλεγχου S^2 του τύπου (2.2.8). Η μέθοδος αυτή είναι ισοδύναμη με την εφαρμογή επαναλαμβανόμενων ελέγχων σημαντικότητας για την υπόθεση ότι ο πίνακας συνδιακύμανσης της διεργασίας είναι ίσος με ένα συγκεκριμένο πίνακα Σ_0 , του οποίου τα στοιχεία είναι σταθερές ($H_0: \Sigma = \Sigma_0$ κατά $H_1: \Sigma \neq \Sigma_0$). Χρησιμοποιώντας την εν λόγω μέθοδο, η απεικονιζόμενη στατιστική συνάρτηση ενός διαγράμματος για το k δείγμα, υπό τη μηδενική υπόθεση, είναι

$$W_k = -pn + pn \cdot \ln(n) - n \cdot \ln\left(\frac{|\mathbf{A}_k|}{|\Sigma|}\right) + tr(\Sigma^{-1}\mathbf{A}_k), \quad k=1, 2, \dots, m, \quad (3.2.11)$$

όπου $\mathbf{A}_k = (n-1)\mathbf{S}_k$, \mathbf{S}_k είναι ο $(p \times p)$ δειγματικός πίνακας διακύμανσης για το k δείγμα και $tr(\Sigma^{-1}\mathbf{A}_k)$ είναι το ίχνος¹ του πίνακα $\Sigma^{-1}\mathbf{A}_k$. Στο σημείο αυτό αναφέρουμε ότι οι Alt and Smith (1988) ορίζουν τη στατιστική συνάρτηση ελέγχου ως εξής

$$W_k^* = -p \cdot (n-1) - (n-1) \cdot \ln(|\mathbf{S}|) + (n-1) \cdot \ln(|\Sigma|) + (n-1) \cdot tr(\Sigma^{-1}\mathbf{S}), \quad k=1, 2, \dots, m, \quad (3.2.12)$$

¹ Το ίχνος ενός πίνακα είναι το άθροισμα των στοιχείων της κύριας διαγωνίου του.

Η στατιστική συνάρτηση W_k ακολουθεί τη χ^2 -κατανομή με $p \times (p+1)/2$ β.ε., συμβολικά $W_k \sim \chi_{p(p+1)/2}^2$. Το ίδιο ισχύει και για τη στατιστική συνάρτηση W_k^* , δηλαδή $W_k^* \sim \chi_{p(p+1)/2}^2$. Ένα πολυμεταβλητό διάγραμμα τύπου Shewhart για τη διασπορά της διεργασίας, με γνωστό διάνυσμα μέσων μ και γνωστό πίνακα διακύμανσης-συνδιακύμανσης Σ , έχει όρια

$$\begin{array}{l} \text{Διάγραμμα } W \text{ ή } W^* \\ \text{Φάσης II} \end{array} \begin{cases} LCL = 0 \\ UCL = \chi_{p(p+1)/2; a}^2 \end{cases} \quad (3.2.13)$$

Η διεργασία είναι εκτός στατιστικού ελέγχου, εάν η τιμή της στατιστικής W_k είναι μεγαλύτερη από το άνω όριο του τύπου (3.2.13). Το παραπάνω διάγραμμα ονομάζεται **διάγραμμα W ή W^* Φάσης II** (*Phase II W or W^* -chart*) και είναι ανάλογο του μονοδιάστατου διαγράμματος S^2 Φάσης II με όρια πιθανότητας a του τύπου (2.2.8).

Η δεύτερη μέθοδος βασίζεται στη γενικευμένη δειγματική διακύμανση $|\mathbf{S}|$, όπου \mathbf{S} είναι ο δειγματικός πίνακας διακύμανσης-συνδιακύμανσης. Ένας τρόπος για την κατασκευή του διαγράμματος $|\mathbf{S}|$ είναι χρησιμοποιώντας τις ιδιότητες της κατανομής που ακολουθεί το $|\mathbf{S}|$. Ο Hoel (1937) απέδειξε ότι στην περίπτωση που έχουμε δύο ποιοτικά χαρακτηριστικά ($p=2$), η ποσότητα $2(n-1)|\mathbf{S}|^{1/2} |\Sigma|^{-1/2}$ ακολουθεί τη χ^2 -κατανομή με $2n-4$ β.ε., δηλαδή

$$\frac{2(n-1)|\mathbf{S}|^{1/2}}{|\Sigma|^{1/2}} \sim \chi_{2n-4}^2. \quad (3.2.14)$$

Σύμφωνα με τους Alt and Smith (1988), αντιστρέφοντας τον τύπο (3.2.14) προκύπτει ότι ένα διάγραμμα τύπου Shewhart για τη διασπορά της διεργασίας, στην περίπτωση που έχουμε μόνο δύο ποιοτικά χαρακτηριστικά και οι παράμετροι είναι γνωστές, έχει τα εξής όρια πιθανότητας

$$\begin{array}{l} \text{Διάγραμμα } |\mathbf{S}|^{1/2} \\ \text{Φάσης II} \\ \text{(Όρια Πιθανότητας } a) \end{array} \begin{cases} LCL = \frac{|\Sigma|^{1/2}}{2(n-1)} \cdot \chi_{2n-4; 1-(a/2)}^2 \\ UCL = \frac{|\Sigma|^{1/2}}{2(n-1)} \cdot \chi_{2n-4; a/2}^2 \end{cases} \quad (3.2.15)$$

Το παραπάνω διάγραμμα είναι γνωστό ως **διάγραμμα ελέγχου $|S|^{1/2}$ Φάσης II** (*Phase II $|S|^{1/2}$ -chart with probability limits*) και είναι ανάλογο του μονοδιάστατου διαγράμματος ελέγχου S Φάσης II με 3σ όρια ελέγχου του τύπου (2.2.7). Θα πρέπει να αναφέρουμε ότι ο Alt (1985) παρουσιάζει το **διάγραμμα ελέγχου $|S|$ Φάσης II** (*Phase II $|S|$ -chart with probability limits*), το οποίο έχει τα όρια ελέγχου του διαγράμματος $|S|^{1/2}$ υψωμένα στο τετράγωνο. Τα όρια αυτά δίνονται από τον παρακάτω τύπο

$$\begin{array}{l} \text{Διάγραμμα } |S| \\ \text{Φάσης II} \\ \text{(Όρια Πιθανότητας } \alpha) \end{array} \left\{ \begin{array}{l} LCL = \frac{|\Sigma|}{4(n-1)^2} \cdot \left(\chi_{2n-4; 1-(\alpha/2)}^2 \right)^2 \\ UCL = \frac{|\Sigma|}{4(n-1)^2} \cdot \left(\chi_{2n-4; \alpha/2}^2 \right)^2 \end{array} \right. \quad (3.2.16)$$

Τα διαγράμματα $|S|^{1/2}$ και $|S|$ είναι ισοδύναμα.

Η ακριβής κατανομή για περισσότερα από δύο χαρακτηριστικά δεν είναι κάποια γνωστή κατανομή, εάν και έχουν προταθεί ορισμένες προσεγγίσεις από τους Anderson (1984) και Gnanadesikan and Gupta (1970). Σύμφωνα με την πρώτη προσέγγιση του Anderson, όταν έχουμε p χαρακτηριστικά μεγέθους n και η διεργασία είναι εντός ελέγχου, η κατανομή του $|S|$ δίνεται από τον τύπο

$$\frac{|S| \cdot (n-1)^p}{|\Sigma|} \equiv \prod_{i=1}^p \chi_{n-1}^2,$$

όπου $|\Sigma|$ είναι η ορίζουσα του εντός ελέγχου πίνακα διακύμανσης-συνδιακύμανσης και οι χι-τετράγωνο μεταβλητές είναι ανεξάρτητες. Θα πρέπει να αναφέρουμε ότι δεν υπάρχουν πίνακες για την κατανομή αυτή, αφού δεν είναι γνωστή. Οι Aparisi, Jabaloyes and Carrión (1999) παρουσίασαν δύο διαγράμματα $|S|$ για p χαρακτηριστικά, ένα που έχει μόνο το άνω όριο ελέγχου και ένα που έχει και τα δύο όρια ελέγχου. Αρχικά, όρισαν μία νέα μεταβλητή, η οποία υπολογίζεται από τον παρακάτω τύπο

$$J_{n,p} = \frac{|S| \cdot (n-1)^p}{|\Sigma|}.$$

Στη συνέχεια, για να καθορίσουν τα όρια ελέγχου των διαγραμμάτων $|\mathbf{S}|$, έπρεπε να ορίσουν τα ποσοστιαία σημεία της νέας μεταβλητής $J_{n,p}$. Για το σκοπό αυτό, προχώρησαν στον προσδιορισμό της συνάρτησης πυκνότητας πιθανότητας (σ.π.π.) της $J_{n,p}$.

Θεώρησαν ότι $X_j \sim \chi_{n-1}^2, j=1, 2, \dots, p$, το X_j είναι ανεξάρτητο από το X_l για $j \neq l$ και έθεσαν $y_1 = \prod_{j=1}^p X_j = J_{n,p}$. Για το λόγο αυτό, η σ.π.π. του y_1 είναι ίδια με τη σ.π.π. της $J_{n,p}$. Επίσης, θεώρησαν ότι $y_2 = x_2, \dots, y_p = x_p$ και απέδειξαν ότι η από κοινού σ.π.π. των y_1, y_2, \dots, y_p δίνεται από τον τύπο που ακολουθεί

$$f(y_1, y_2, \dots, y_p) = \frac{\prod_{j=2}^p y_j^{(n-j+1)/2} \cdot \prod_{j=2}^p \frac{1}{y_j} \cdot \exp\left[-\frac{1}{2}\left(\sum_{j=2}^p y_i + \frac{y_1}{2\sum_{j=2}^p y_i}\right)\right] \cdot \left(\frac{y_j}{\sum_{j=2}^p y_i}\right)^{\frac{n-3}{2}}}{2^{\sum_{j=2}^p \frac{n-j}{2}} \cdot \prod_{j=1}^p \Gamma\left(\frac{n-j}{2}\right)}.$$

Η κατανομή του y_1 είναι η περιθώρια (*marginal*) κατανομή του $f(y_1, y_2, \dots, y_p)$, δηλαδή

$$f(y_1) = \int_{y_2=0}^{\infty} \int_{y_3=0}^{\infty} \dots \int_{y_p=0}^{\infty} f(y_1, y_2, \dots, y_p) dy_2 dy_3 \dots dy_n.$$

Για το διάγραμμα $|\mathbf{S}|$ που έχει μόνο το άνω όριο ελέγχου, το σφάλμα τύπου Ι είναι

$$\begin{aligned} a &= P(|\mathbf{S}| > UCL) = P\left(\frac{|\mathbf{S}| \cdot (n-1)^p}{|\mathbf{\Sigma}|} > \frac{UCL \cdot (n-1)^p}{|\mathbf{\Sigma}|}\right) = P\left(J_{n,p} > \frac{UCL \cdot (n-1)^p}{|\mathbf{\Sigma}|}\right) \\ &= P(J_{n,p} > J_{n,p;a}), \end{aligned}$$

όπου $J_{n,p;a}$ είναι το άνω a -ποσοστιαίο σημείο της εντός ελέγχου κατανομής της τυχαίας μεταβλητής $J_{n,p}$. Συνεπώς, το άνω όριο ελέγχου του **διαγράμματος** $|\mathbf{S}|$, στην περίπτωση που έχουμε p χαρακτηριστικά, υπολογίζεται ως εξής

$$\left. \begin{array}{l} \text{Διάγραμμα } |\mathbf{S}| \\ \text{Φάσης II} \\ (\mathbf{p} > 2) \end{array} \right\} UCL = \frac{|\mathbf{\Sigma}|}{(n-1)^p} \cdot J_{n,p;a}.$$

Για το διάγραμμα $|\mathbf{S}|$ με δύο όρια ελέγχου, το σφάλμα τύπου I υπολογίζεται ως εξής

$$\begin{aligned} 1-a &= P(LCL < |\mathbf{S}| < UCL) = P\left(\frac{LCL \cdot (n-1)^p}{|\Sigma|} < \frac{|\mathbf{S}| \cdot (n-1)^p}{|\Sigma|} < \frac{UCL \cdot (n-1)^p}{|\Sigma|}\right) = \\ &= P\left(\frac{LCL \cdot (n-1)^p}{|\Sigma|} < J_{n,p} < \frac{UCL \cdot (n-1)^p}{|\Sigma|}\right) = P\left(J_{n,p; 1-(a/2)} < J_{n,p} < J_{n,p; a/2}\right), \end{aligned}$$

όπου $J_{n,p; a/2}$ είναι το άνω $(a/2)$ -ποσοστιαίο σημείο της εντός ελέγχου κατανομής της τυχαίας μεταβλητής $J_{n,p}$. Συνεπώς, το άνω και κάτω όριο ελέγχου του **διαγράμματος** $|\mathbf{S}|$, στην περίπτωση που έχουμε p χαρακτηριστικά, είναι

$$\text{Διάγραμμα } |\mathbf{S}| \begin{cases} LCL = \frac{|\Sigma|}{(n-1)^p} \cdot J_{n,p; 1-(a/2)} \\ UCL = \frac{|\Sigma|}{(n-1)^p} \cdot J_{n,p; a/2} \end{cases} \begin{matrix} \text{Φάσης II} \\ (p > 2) \end{matrix}$$

Ο Alt (1985), βασιζόμενος και πάλι στη γενικευμένη δειγματική διακύμανση $|\mathbf{S}|$, χρησιμοποίησε ένα δεύτερο τρόπο για να κατασκευάσει ένα ακόμη διάγραμμα διασποράς, χωρίς να χρησιμοποιήσει τις ιδιότητες της κατανομής του $|\mathbf{S}|$. Ακολουθώντας τον τρόπο αυτό, βασίζεται μόνο στις ροπές πρώτης και δεύτερης τάξης του $|\mathbf{S}|$, δηλαδή στη μέση τιμή $E(|\mathbf{S}|)$ και στη διακύμανση $V(|\mathbf{S}|)$ του $|\mathbf{S}|$, αλλά και στο γεγονός ότι το μεγαλύτερο μέρος της κατανομής πιθανότητας του $|\mathbf{S}|$ περιέχεται στο διάστημα $E(|\mathbf{S}|) \pm 3\sqrt{V(|\mathbf{S}|)}$. Μπορεί να αποδειχθεί ότι $E(|\mathbf{S}|) = b_1 \cdot |\Sigma|$ και $V(|\mathbf{S}|) = b_2 \cdot |\Sigma|^2$, όπου

$$b_1 = \frac{1}{(n-1)^p} \cdot \prod_{i=1}^p (n-i). \quad (3.2.17)$$

και

$$b_2 = \frac{1}{(n-1)^{2p}} \cdot \prod_{i=1}^p (n-i) \cdot \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right]. \quad (3.2.18)$$

Συνεπώς, ένα διάγραμμα ελέγχου τύπου Shewhart για τη διασπορά της διεργασίας, στην περίπτωση που έχουμε περισσότερα από δύο ποιοτικά χαρακτηριστικά και οι παράμετροι είναι γνωστές, έχει τα ακόλουθα 3σ όρια ελέγχου

$$\begin{array}{l} \text{Διάγραμμα } |S| \\ \text{Φάσης II} \\ \text{(3σ Όρια Ελέγχου)} \end{array} \left\{ \begin{array}{l} LCL = |\Sigma| \cdot (b_1 - 3\sqrt{b_2}) \\ CL = |\Sigma| \cdot b_1 \\ UCL = |\Sigma| \cdot (b_1 + 3\sqrt{b_2}). \end{array} \right. \quad (3.2.19)$$

Το παραπάνω διάγραμμα είναι γνωστό ως **διάγραμμα ελέγχου $|S|$ Φάσης II** (*Phase II $|S|$ -chart with 3σ limits*).

Θα πρέπει να αναφέρουμε ότι οι Alt and Smith (1988) βασίστηκαν στις ροπές πρώτης και δεύτερης τάξης του $|S|^{1/2}$ και στο γεγονός ότι το μεγαλύτερο μέρος της κατανομής πιθανότητας του $|S|^{1/2}$ περιέχεται στο διάστημα $E(|S|^{1/2}) \pm 3\sqrt{V(|S|^{1/2})}$. Αποδεικνύεται ότι

$$E(|S|^{1/2}) = b_3 \cdot |\Sigma|^{1/2} \text{ και } V(|S|^{1/2}) = (b_1 - b_3^2) \cdot |\Sigma|,$$

όπου το b_1 υπολογίζεται από τον τύπο (3.2.17) και το b_3 υπολογίζεται ως εξής

$$b_3 = \left(\frac{2}{n-1} \right)^{p/2} \frac{\Gamma(n/2)}{\Gamma((n-p)/2)}. \quad (3.2.20)$$

Με τον τρόπο αυτό, προκύπτει το **διάγραμμα ελέγχου $|S|^{1/2}$ Φάσης II με 3σ όρια ελέγχου** (*Phase II $|S|^{1/2}$ -chart with 3σ limits*), το οποίο δίνεται από τον παρακάτω τύπο

$$\begin{array}{l} \text{Διάγραμμα } |S|^{1/2} \\ \text{Φάσης II} \\ \text{(3σ Όρια Ελέγχου)} \end{array} \left\{ \begin{array}{l} LCL = |\Sigma|^{1/2} \cdot (b_3 - 3\sqrt{b_1 - b_3^2}) \\ CL = |\Sigma|^{1/2} \cdot b_3 \\ UCL = |\Sigma|^{1/2} \cdot (b_3 + 3\sqrt{b_1 - b_3^2}). \end{array} \right. \quad (3.2.21)$$

Το παραπάνω διάγραμμα είναι κατάλληλο για οποιοδήποτε αριθμό ποιοτικών χαρακτηριστικών και είναι ανάλογο του μονοδιάστατου διαγράμματος ελέγχου S Φάσης II

με 3σ όρια ελέγχου, αυτά του τύπου (2.2.7). Προφανώς, τα διαγράμματα $|\mathbf{S}|^{1/2}$ και $|\mathbf{S}|$ δεν είναι ισοδύναμα, αφού $(E(|\mathbf{S}|))^{1/2} \neq E(|\mathbf{S}|^{1/2})$ και $(V(|\mathbf{S}|))^{1/2} \neq V(|\mathbf{S}|^{1/2})$.

Στη συνέχεια, θα αναπτύξουμε διαγράμματα ελέγχου για τη διασπορά Φάσης II. Αφού η ακριβής τιμή $|\Sigma|$ δεν είναι γνωστή, θα πρέπει στη θέση του να χρησιμοποιήσουμε έναν αμερόληπτο εκτιμητή. Ο Alt (1985) όρισε τη μέση γενικευμένη δειγματική διακύμανση ως

$$|\mathbf{S}^*| = \frac{1}{m} \cdot \sum_{k=1}^m |\mathbf{S}_k|, \quad k = 1, 2, \dots, m,$$

όπου $|\mathbf{S}_k|$ είναι ο δειγματικός πίνακας διακύμανσης-συνδιακύμανσης του k δείγματος. Αφού $E(|\mathbf{S}_k|) = b_1 \cdot |\Sigma|$, ο $|\mathbf{S}^*|/b_1$ είναι αμερόληπτος εκτιμητής του $|\Sigma|$, όπου το b_1 υπολογίζεται μέσω του τύπου (3.2.17). Οι Alt and Smith (1988) όρισαν ανάλογα το μέσο όρο της τετραγωνικής ρίζας της γενικευμένης δειγματικής διακύμανσης ως εξής

$$|\mathbf{S}^*|^{1/2} = \frac{1}{m} \cdot \sum_{k=1}^m |\mathbf{S}_k|^{1/2}, \quad k = 1, 2, \dots, m.$$

Επομένως, ο $|\mathbf{S}^*|^{1/2}/b_3$ είναι ένας αμερόληπτος εκτιμητής του $|\Sigma|^{1/2}$, όπου το b_3 υπολογίζεται από τον τύπο (3.2.20). Ο Kshirsagar (1972) με τη σειρά του έδειξε ότι ένας αμερόληπτος εκτιμητής του $|\Sigma|^{-1}$ είναι η ποσότητα $(n-p-2) \cdot \mathbf{S}_k^{-1}/(n-1)$, όπου \mathbf{S}_k^{-1} είναι ο αντίστροφος δειγματικός πίνακας διακύμανσης-συνδιακύμανσης του k δείγματος για $k = 1, 2, \dots, m$. Κατά συνέπεια, εάν

$$\mathbf{S}_*^{-1} = \frac{1}{m} \cdot \sum_{k=1}^m \mathbf{S}_k^{-1}, \quad k = 1, 2, \dots, m,$$

η ποσότητα $(n-p-2) \cdot \mathbf{S}_*^{-1}/(n-1)$ θα είναι αμερόληπτος εκτιμητής του $|\Sigma|^{-1}$, που λαμβάνεται από m ορθολογικά δείγματα.

Αντικαθιστώντας το $|\Sigma|$ με $|\mathbf{S}^*|/b_1$ και το $|\Sigma|^{-1}$ με $(n-p-2) \cdot \mathbf{S}_*^{-1}/(n-1)$ στους τύπους (3.2.11) και (3.2.12), οι νέες τιμές των στατιστικών συναρτήσεων W_k και W_k^* θα απεικονίζονται σε ένα **διάγραμμα W ή W^*** , αντίστοιχα, **Φάσης I**, με τα εξής όρια ελέγχου

$$\begin{array}{l} \text{Διάγραμμα } W \text{ ή } W^* \\ \text{Φάσης I} \end{array} \left\{ \begin{array}{l} LCL = 0 \\ UCL = \chi_{p(p+1)/2; a}^2 \end{array} \right.$$

Τα παραπάνω όρια ελέγχου μπορούν να χρησιμοποιηθούν τόσο για ένα, όσο και για περισσότερα από ένα ποιοτικά χαρακτηριστικά.

Εάν στον τύπο (3.2.15) αντικαταστήσουμε το $|\Sigma|^{1/2}$ με $|\mathbf{S}^*|^{1/2} / b_3$, προκύπτει το **διάγραμμα ελέγχου $|\mathbf{S}^*|^{1/2}$ Φάσης I με όρια πιθανότητας a** (*Phase I $|\mathbf{S}^*|^{1/2}$ -chart with probability limits*), τα οποία υπολογίζονται ως εξής

$$\begin{array}{l} \text{Διάγραμμα } |\mathbf{S}^*|^{1/2} \\ \text{Φάσης I} \\ \text{(Όρια Πιθανότητας } a) \end{array} \left\{ \begin{array}{l} LCL = \frac{|\mathbf{S}^*|^{1/2}}{2b_3(n-1)} \cdot \chi_{2n-4; 1-(a/2)}^2 \\ UCL = \frac{|\mathbf{S}^*|^{1/2}}{2b_3(n-1)} \cdot \chi_{2n-4; a/2}^2 \end{array} \right.$$

Αντίστοιχα, εάν στον τύπο (3.2.16) αντικαταστήσουμε το $|\Sigma|$ με $|\mathbf{S}^*| / b_1$, προκύπτει το **διάγραμμα ελέγχου $|\mathbf{S}^*|$ Φάσης I με όρια πιθανότητας a** (*Phase I $|\mathbf{S}^*|$ -chart with probability limits*), τα οποία υπολογίζονται ως εξής

$$\begin{array}{l} \text{Διάγραμμα } |\mathbf{S}^*| \\ \text{Φάσης I} \\ \text{(Όρια Πιθανότητας } a) \end{array} \left\{ \begin{array}{l} LCL = \frac{|\mathbf{S}^*|}{4b_1(n-1)^2} \cdot (\chi_{2n-4; 1-(a/2)}^2)^2 \\ UCL = \frac{|\mathbf{S}^*|}{4b_1(n-1)^2} \cdot (\chi_{2n-4; a/2}^2)^2 \end{array} \right.$$

Εάν στον τύπο (3.2.19) κάνουμε την ίδια αντικατάσταση που κάναμε προηγουμένως, δηλαδή όπου $|\Sigma|$ βάλουμε το $|\mathbf{S}^*| / b_1$, προκύπτει το **διάγραμμα ελέγχου $|\mathbf{S}^*|$ Φάσης I με 3σ όρια ελέγχου** (*Phase I $|\mathbf{S}^*|$ -chart with 3σ limits*) που υπολογίζονται μέσω του τύπου

$$\begin{array}{l} \text{Διάγραμμα } |\mathbf{S}^*| \\ \text{Φάσης I} \\ \text{(} 3\sigma \text{ Όρια Ελέγχου)} \end{array} \left\{ \begin{array}{l} LCL = |\mathbf{S}^*| \cdot (1 - 3\sqrt{b_2} / b_1) \\ CL = |\mathbf{S}^*| \\ UCL = |\mathbf{S}^*| \cdot (1 + 3\sqrt{b_2} / b_1) \end{array} \right. \quad (3.2.22)$$

Τέλος, εάν στον τύπο (3.2.21) αντικαταστήσουμε το $|\Sigma|^{1/2}$ με $|\mathbf{S}^*|^{1/2}/b_3$, προκύπτει το **διάγραμμα ελέγχου $|\mathbf{S}^*|^{1/2}$ Φάσης I με 3σ όρια ελέγχου** (*Phase I $|\mathbf{S}^*|^{1/2}$ -chart with 3σ limits*), του οποίου τα όρια υπολογίζονται από τον παρακάτω τύπο

$$\begin{array}{l} \text{Διάγραμμα } |\mathbf{S}^*|^{1/2} \\ \text{Φάσης I} \\ \text{(3σ Όρια Ελέγχου)} \end{array} \left\{ \begin{array}{l} LCL = |\mathbf{S}^*|^{1/2} \cdot \left(1 - 3\sqrt{\frac{b_1}{b_3^2} - 1}\right) \\ CL = |\mathbf{S}^*|^{1/2} \\ UCL = |\mathbf{S}^*|^{1/2} \cdot \left(1 + 3\sqrt{\frac{b_1}{b_3^2} - 1}\right) \end{array} \right.$$

Το παραπάνω διάγραμμα είναι ανάλογο του διαγράμματος S Φάσης I με 3σ όρια ελέγχου του τύπου (2.2.10). Τα διαγράμματα $|\mathbf{S}^*|$ και $|\mathbf{S}^*|^{1/2}$, είτε έχουν 3σ όρια ελέγχου είτε όρια πιθανότητας a , δεν είναι ισοδύναμα.

3.3 Πολυμεταβλητά Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες χωρίς μνήμη

Τα πολυμεταβλητά διαγράμματα ελέγχου για ιδιότητες (*multiattribute control charts*) δεν έχουν αναπτυχθεί αρκετά, παρόλο που ο έλεγχος πολυμεταβλητών διεργασιών για ιδιότητες κρίνεται ιδιαίτερα σημαντικός στις πρακτικές παραγωγικές διαδικασίες. Ο Patel (1973) ήταν ο πρώτος που ασχολήθηκε με μεθόδους ποιοτικού ελέγχου, όταν τα p -διάστατα ποιοτικά χαρακτηριστικά προέρχονται από μια πολυδιάστατη Διωνυμική ή Poisson κατανομή.

Ο Patel (1973) πρότεινε ένα διάγραμμα χ^2 τύπου Hotelling για την παρακολούθηση παρατηρήσεων, οι οποίες προέρχονται από πολυδιάστατη Διωνυμική ή πολυδιάστατη Poisson κατανομή στην περίπτωση που έχουμε είτε χρονικά ανεξάρτητα (*time independent samples*) είτε χρονικά εξαρτημένα δείγματα (*time dependent or auto-correlated samples*). Υπέθεσε ότι εάν επιλέξουμε ένα κατάλληλο μέγεθος δείγματος, τα διανύσματα θα ακολουθούν προσεγγιστικά την πολυδιάστατη κανονική κατανομή. Συνεπώς, είναι σωστό να χρησιμοποιήσουμε την πολυδιάστατη κανονική τεχνική ελέγχου για την κατασκευή του

διαγράμματος. Στη συνέχεια, θα παρουσιάσουμε το διάγραμμα χ^2 του Patel, μόνο για την περίπτωση που τα δείγματα είναι χρονικά ανεξάρτητα.

Θεωρούμε ότι έχουμε ένα δείγμα μεγέθους m , έστω $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$. Τα ανεξάρτητα διανύσματα $\mathbf{X}_k = (X_{k1}, X_{k2}, \dots, X_{kp})$ για $k=1, \dots, m$, διάστασης p , προέρχονται από πολυδιάστατη διωνυμική κατανομή. Η στατιστική συνάρτηση που χρησιμοποιείται είναι

$$G_k = (\mathbf{X}_k - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X}_k - \bar{\mathbf{X}}), \quad k=1, \dots, m,$$

όπου $\bar{\mathbf{X}}$ είναι η εκτίμηση του δειγματικού μέσου $\boldsymbol{\mu}$ και \mathbf{S} είναι η εκτίμηση του πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$ της διεργασίας, ο οποίος θεωρείται ότι παραμένει σταθερός σε όλη τη διάρκεια της διεργασίας. Για μεγάλο m , η στατιστική συνάρτηση ελέγχου G_k προσεγγίζει τη χ^2 -κατανομή με p β.ε., όπου p είναι ο αριθμός των ιδιοτήτων της διεργασίας. Συνεπώς, τα όρια ελέγχου του διαγράμματος χ^2 του Patel είναι τα εξής

$$\text{Διαγράμμα } \chi^2 \begin{cases} LCL = 0 \\ UCL = \chi_{p, \alpha}^2 \end{cases}$$

Σύμφωνα με τους Topalidou and Psarakis (2009), η μέθοδος του Patel παρουσιάζει μειονεκτήματα, όπως η ασυμμετρία των κατανομών για τις ιδιότητες και η δυσκολία εντοπισμού της μετατοπισμένης μεταβλητής. Ένα άλλο μειονέκτημα είναι ότι το διάγραμμα που πρότεινε ο Patel έχει μόνο άνω όριο ελέγχου, γεγονός που κάνει αδύνατη τη χρήση του για την ανίχνευση της υποβάθμισης ή της βελτίωσης της διεργασίας. Τέλος, εξαιτίας της πολυπλοκότητάς της και της υπόθεσης της κανονικότητας και της ισότητας της διακύμανσης της διεργασίας, η μέθοδος του Patel δε χρησιμοποιείται εκτεταμένα στην πράξη.

Οι Lu, Xie, Goh and Lai (1998) πρότειναν ένα **πολυμεταβλητό διάγραμμα ελέγχου για τον αριθμό των ελαττωματικών προϊόντων** (*multivariate control chart for number nonconforming or MNP-chart*), το οποίο αποτελεί γενίκευση του μονοδιάστατου διαγράμματος np . Αρχικά, θεωρούν ότι υπάρχουν p διαθέσιμα ποιοτικά χαρακτηριστικά και ορίζουν με p_j την πιθανότητα ένα στοιχείο να είναι ελαττωματικό σε σχέση με το j χαρακτηριστικό. Τα ποιοτικά χαρακτηριστικά μπορεί να μην είναι ανεξάρτητα, οπότε η συσχέτιση μεταξύ των χαρακτηριστικών i, j συμβολίζεται με δ_{ij} . Επίσης, το διάνυσμα του

ποσοστού των ελαττωμάτων είναι $\mathbf{P} = (p_1, p_2, \dots, p_p)$ και $\mathbf{\Sigma} = [\delta_{ij}]_{p \times p}$ είναι ο πίνακας του συντελεστή συσχέτισης. Επιπλέον, $\mathbf{C} = (C_1, C_2, \dots, C_p)$ είναι το διάνυσμα των μετρήσεων των ελαττωματικών μονάδων, όπου C_j είναι η μέτρηση των ελαττωματικών μονάδων του j χαρακτηριστικού στο δείγμα. Η στατιστική συνάρτηση ελέγχου

$$X = \sum_{j=1}^p \frac{C_j}{\sqrt{p_j}}$$

αποτελεί το σταθμισμένο άθροισμα των ελαττωματικών μονάδων όλων των ποιοτικών χαρακτηριστικών στο δείγμα. Όσο μικρότερο είναι το ποσοστό ελαττωματικών προϊόντων p_j , τόσο περισσότερο συνεισφέρει η μέτρηση C_j στη στατιστική συνάρτηση ελέγχου X της πολυμεταβλητής διεργασίας για ιδιότητες. Για το λόγο αυτό, επιλέγουμε την ποσότητα $1/\sqrt{p_j}$ ως βάρος για τη συνεισφορά του j χαρακτηριστικού.

Ωστόσο, ο Montgomery (1997) πρότεινε εναλλακτικά να συμβολίσουμε με d_j τον αριθμό των ελαττωμάτων, υποδηλώνοντας έτσι την αυστηρότητα της μη συμμόρφωσης του j χαρακτηριστικού. Συνεπώς, η στατιστική συνάρτηση ελέγχου X μπορεί να έχει την εξής πιο γενική μορφή

$$X_D = \sum_{j=1}^p \frac{d_j C_j}{\sqrt{p_j}}$$

Τα στοιχεία $d_j, j = 1, \dots, p$, συνιστούν το διάνυσμα μη συμμόρφωσης $\mathbf{D} = (d_1, d_2, \dots, d_p)$. Η μέση τιμή και η διακύμανση της στατιστικής συνάρτησης X_D υπολογίζονται από τους τύπους που ακολουθούν

$$E(X_D) = E\left(\sum_{j=1}^p \frac{d_j C_j}{\sqrt{p_j}}\right) = \sum_{j=1}^p \frac{d_j E(C_j)}{\sqrt{p_j}} = \sum_{j=1}^p \frac{d_j n p_j}{\sqrt{p_j}} = n \sum_{j=1}^p d_j \sqrt{p_j}$$

και

$$V(X_D) = V\left(\sum_{j=1}^p \frac{d_j C_j}{\sqrt{p_j}}\right) = \sum_{j=1}^p \frac{d_j^2 V(C_j)}{p_j} + 2 \sum_{l < j} \frac{d_l d_j \delta_{lj} \sqrt{V(C_l) V(C_j)}}{\sqrt{p_l p_j}} =$$

$$\begin{aligned}
&= \sum_{j=1}^p \frac{d_j^2 n p_j (1-p_j)}{(\sqrt{p_j})^2} + 2 \sum_{l < j} \frac{d_l d_j \delta_{lj} \sqrt{n p_l (1-p_l) n p_j (1-p_j)}}{\sqrt{p_l p_j}} = \\
&= n \left\{ \sum_{j=1}^p d_j^2 (1-p_j) + 2 \sum_{l < j} d_l d_j \delta_{lj} \sqrt{(1-p_l)(1-p_j)} \right\}.
\end{aligned}$$

Γενικά για το συντελεστή συσχέτισης ισχύει ότι $\delta_{ij} = \delta_{ji}$, $|\delta_{ij}| \leq 1$ και $\delta_{ij} = 1$ για $i = j$. Το δ_{ij} είναι είτε γνωστό είτε εκτιμάται από τα προκαταρκτικά δείγματα μέσω του τύπου

$$\delta_{ij} = \frac{\text{Cov}(C_i, C_j)}{\sqrt{V(C_i)V(C_j)}} = \frac{E[(C_i - E(C_i))(C_j - E(C_j))]}{\sqrt{V(C_i)V(C_j)}}.$$

Η γενική μορφή των ορίων ελέγχου ενός διαγράμματος τύπου Shewhart για τη συνάρτηση X_D είναι $E(X_D) \pm L\sqrt{V(X_D)}$. Συνεπώς, εάν το διάνυσμα $\mathbf{P} = (p_1, p_2, \dots, p_p)$ και ο πίνακας Σ είναι γνωστά, τα **3σ όρια ελέγχου ενός διαγράμματος MNP τύπου Shewhart** είναι

$$\begin{aligned}
\text{Διάγραμμα} & \left\{ \begin{aligned} LCL &= n \sum_{j=1}^p d_j \sqrt{p_j} - 3 \sqrt{n \left\{ \sum_{j=1}^p d_j^2 (1-p_j) + 2 \sum_{l < k} d_l d_k \delta_{lk} \sqrt{(1-p_l)(1-p_k)} \right\}} \\ MNP & \left\{ \begin{aligned} CL &= n \sum_{j=1}^p d_j \sqrt{p_j} \\ \text{Φάσης II} & \left\{ \begin{aligned} UCL &= n \sum_{j=1}^p d_j \sqrt{p_j} + 3 \sqrt{n \left\{ \sum_{j=1}^p d_j^2 (1-p_j) + 2 \sum_{l < k} d_l d_k \delta_{lk} \sqrt{(1-p_l)(1-p_k)} \right\}}. \end{aligned} \right. \end{aligned} \right. \end{aligned} \right. \quad (3.3.1)
\end{aligned}$$

Όταν το διάνυσμα $\mathbf{P} = (p_1, p_2, \dots, p_p)$ και ο πίνακας Σ δεν είναι γνωστά, θα πρέπει να εκτιμηθούν από τα δεδομένα. Αρχικά, παίρνουμε m προκαταρκτικά δείγματα μεγέθους n . Οι Lu, Xie, Goh and Lai (1998) προτείνουν το m να είναι περίπου 25, ενώ για το n θα πρέπει να ισχύει

$$\frac{9 \text{tr} \left(\sqrt{(\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P})' \Sigma'} \right)}{(\sqrt{\mathbf{P}\mathbf{I}})^2},$$

όπου $\mathbf{I} = [1, 1, \dots, 1]$ είναι ένα $(p \times 1)$ διάνυσμα. Επίσης, ορίζουμε με C_{jk} , για $j = 1, 2, \dots, p$ και $k = 1, 2, \dots, m$, τις μη συμμορφωμένες μονάδες του j χαρακτηριστικού του k

δείγματος. Το διάνυσμα του ποσοστού των ελαττωμάτων του k δείγματος (\mathbf{P}_k) εκτιμάται από τον τύπο που ακολουθεί

$$\bar{\mathbf{P}}_k = \left(\frac{C_{1k}}{n}, \frac{C_{2k}}{n}, \dots, \frac{C_{pk}}{n} \right) = (\bar{p}_{1k}, \bar{p}_{2k}, \dots, \bar{p}_{pk}), \quad k = 1, 2, \dots, m$$

και το ολικό διάνυσμα του ποσοστού ελαττωματικών προϊόντων \mathbf{P} εκτιμάται ως εξής

$$\bar{\mathbf{P}} = \frac{\sum_{k=1}^m \bar{\mathbf{P}}_k}{m} = \left(\frac{\sum_{k=1}^m C_{1k}}{n \times m}, \frac{\sum_{k=1}^m C_{2k}}{n \times m}, \dots, \frac{\sum_{k=1}^m C_{mk}}{n \times m} \right) = (\bar{p}_1, \bar{p}_2, \dots, \bar{p}_m).$$

Ο πίνακας συσχέτισης Σ εκτιμάται από τα προκαταρκτικά δείγματα μέσω του τύπου

$$\bar{\Sigma} = [\bar{\delta}_{ij}]_{p \times p},$$

όπου ο συντελεστής συσχέτισης δ_{ij} εκτιμάται από τον τύπο

$$\bar{\delta}_{ij} = \frac{\text{Cov}(C_i, C_j)}{\sqrt{V(C_i) \times V(C_j)}} = \frac{4 \sum_{k=1}^m \left(C_{kj} - \frac{\sum_{l=1}^m C_{lk}}{m} \right) \left(C_{jh} - \frac{\sum_{l=1}^m C_{jl}}{m} \right)}{\sqrt{\left[\sum_{k=1}^m \left(C_{lk} - \frac{\sum_{l=1}^m C_{lk}}{m} \right)^2 \right] \times \left[\sum_{k=1}^m \left(C_{jk} - \frac{\sum_{l=1}^m C_{jl}}{m} \right)^2 \right]}}$$

Συνεπώς, όταν $\bar{\mathbf{P}}$ είναι η εκτίμηση του διανύσματος \mathbf{P} και $\bar{\Sigma}$ η εκτίμηση του πίνακα Σ , τα 3σ όρια ελέγχου ενός διαγράμματος MNP τύπου Shewhart παίρνουν τη μορφή

$$\left. \begin{array}{l} \text{Διάγραμμα} \\ \text{MNP} \\ \text{Φάσης I} \end{array} \right\} \begin{cases} LCL = n \sum_{j=1}^p \sqrt{\bar{p}_j} - 3 \sqrt{n \left\{ \sum_{j=1}^p (1 - \bar{p}_j) + 2 \sum_{l < j} \bar{\delta}_{lj} \sqrt{(1 - \bar{p}_l)(1 - \bar{p}_j)} \right\}} \\ CL = n \sum_{j=1}^p \sqrt{\bar{p}_j} \\ UCL = n \sum_{j=1}^p \sqrt{\bar{p}_j} + 3 \sqrt{n \left\{ \sum_{j=1}^p (1 - \bar{p}_j) + 2 \sum_{l < j} \bar{\delta}_{lj} \sqrt{(1 - \bar{p}_l)(1 - \bar{p}_j)} \right\}} \end{cases} \quad (3.3.2)$$

Εάν όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου, τότε η διεργασία είναι εντός στατιστικού ελέγχου και τα δοκιμαστικά όρια είναι κατάλληλα για τον έλεγχο της τρέχουσας ή μελλοντικής διεργασίας. Όταν η διεργασία είναι εκτός ελέγχου, τα δοκιμαστικά όρια ελέγχου πρέπει να αναθεωρηθούν. Αυτό μπορεί να επιτευχθεί εξετάζοντας κάθε σημείο που βρέθηκε εκτός των ορίων και αναζητώντας ειδικές αιτίες μεταβλητότητας. Εφόσον βρεθεί κάποια ειδική αιτία, το αντίστοιχο σημείο δε λαμβάνεται υπ' όψιν και τα δοκιμαστικά όρια ελέγχου υπολογίζονται ξανά, χρησιμοποιώντας μόνο τα υπόλοιπα σημεία. Στη συνέχεια, εξετάζουμε ξανά εάν τα σημεία που μένουν βρίσκονται εντός των νέων ορίων. Η διαδικασία αυτή συνεχίζεται μέχρις ότου όλα τα σημεία να είναι εντός των ορίων. Στο στάδιο αυτό τα όρια είναι έτοιμα για κανονική χρήση. Θα πρέπει να αναφέρουμε ότι τα διαγράμματα *MNP* Φάσης I και II (των τύπων 3.3.1 και 3.3.2) είναι προφανώς ανάλογα των διαγραμμάτων *np* Φάσης I και II (των τύπων 2.3.1 και 2.3.2), αντίστοιχα.

Οι Lu, Xie, Goh and Lai (1998) παρουσίασαν, επίσης, μια μεθοδολογία για την ανίχνευση της μεταβλητής που είναι υπαίτια για την εκτός ελέγχου κατάσταση της διεργασίας. Πρότειναν μία νέα συνάρτηση X που βασίζεται στο ζυγισμένο άθροισμα των μετρήσεων των μη συμμορφωμένων μονάδων όλων των ποιοτικών χαρακτηριστικών. Το βάρος του j χαρακτηριστικού είναι μία φθίνουσα συνάρτηση του ποσοστού ελαττωματικών προϊόντων του j χαρακτηριστικού. Σύμφωνα με τη μέθοδο αυτή, όσο μικρότερο είναι το ποσοστό p_j , τόσο περισσότερο συνεισφέρει η αλλαγή του j χαρακτηριστικού στη μετατόπιση της συνάρτησης X . Συνεπώς, η νέα στατιστική συνάρτηση, που είναι γνωστή ως **συνάρτηση των σκορ** (*score statistic*), ορίζεται ως εξής

$$Z_j = \frac{C_j - np_j}{\sqrt{p_j}} \quad \text{ή} \quad Z_{D_j} = \frac{d_j (C_j - np_j)}{\sqrt{p_j}}.$$

Η συνάρτηση Z_{D_j} χρησιμοποιείται όταν η μη συμμόρφωση δεν είναι ίδια σε κάθε δείγμα.

Εάν ένα εκτός ελέγχου σήμα εντοπιστεί, υπολογίζουμε πρώτα το σκορ για κάθε ποιοτικό χαρακτηριστικό ξεχωριστά. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι Z_l και Z_j είναι το μικρότερο και το μεγαλύτερο σκορ, αντίστοιχα. Όταν η στατιστική συνάρτηση X είναι πάνω από το άνω όριο, το Z_j δίνει το μεγαλύτερο θετικό σκορ και το j ποιοτικό χαρακτηριστικό θεωρείται υπεύθυνο για την ανοδική μετατόπιση (ή υποβάθμιση) της

διεργασίας. Αντίθετα, όταν η στατιστική συνάρτηση X είναι κάτω από το κάτω όριο, το Z_l δίνει το μικρότερο αρνητικό σκορ και το l ποιοτικό χαρακτηριστικό θεωρείται υπεύθυνο για την καθοδική μετατόπιση (ή βελτίωση) της διεργασίας.

Ο Jolayemi (1999) πρότεινε ένα **πολυμεταβλητό διάγραμμα ελέγχου για ιδιότητες** (*Multivariate Attribute Control Chart, MACC*), το οποίο βασίζεται σε μία προσέγγιση συνέλιξης¹ (*convolution*) ανεξάρτητων διωνυμικών μεταβλητών και σε μία επέκταση του μονοδιάστατου διαγράμματος np . Σύμφωνα με τη μεθοδολογία αυτή, θεωρούμε ότι έχουμε m ανεξάρτητες ιδιότητες, κάθε μία από τις οποίες ακολουθεί Διωνυμική κατανομή. Πιο συγκεκριμένα ισχύει $X_1 \sim B(n, p_1)$, $X_2 \sim B(n, p_2)$, ..., $X_m \sim B(n, p_m)$. Χρησιμοποιώντας την προσέγγιση J (Jolayemi (1991), (1994)), παρατηρούμε ότι το άθροισμα ή η συνέλιξη του αριθμού των ελαττωματικών προϊόντων μιας διεργασίας, που αποτελείται από m ιδιότητες μεγέθους n , προσεγγίζεται ικανοποιητικά από τη Διωνυμική κατανομή $B(nm, \bar{p})$, όπου \bar{p} είναι ο μέσος των p_1, p_2, \dots, p_m . Συνεπώς, αντί να χρησιμοποιήσουμε m διαφορετικά διαγράμματα np , η οποία είναι μια μη αποτελεσματική-αποδοτική μεθοδολογία που εφαρμόζεται συχνά στην παραγωγική βιομηχανία, χρησιμοποιούμε την προσέγγιση του Jolayemi (1994) για να αναπτύξουμε ένα πολυμεταβλητό διάγραμμα ελέγχου για ιδιότητες. Τα όρια ελέγχου του διαγράμματος αυτού υπολογίζονται από τον τύπο

$$\begin{array}{l} \text{Διάγραμμα} \\ \text{MACC} \end{array} \begin{cases} LCL = nm\bar{p}_0 - L \cdot \sqrt{nm\bar{p}_0(1-\bar{p}_0)} \\ UCL = nm\bar{p}_0 + L \cdot \sqrt{nm\bar{p}_0(1-\bar{p}_0)}, \end{cases}$$

όπου L είναι το πλάτος των ορίων ελέγχου, \bar{p}_0 είναι ο μέσος των \bar{p}_{0k} , $k=1, 2, \dots, m$ και \bar{p}_{0k} είναι το αναμενόμενο ποσοστό ελαττωματικών προϊόντων που αντιστοιχεί στην ιδιότητα X_k , $k=1, 2, \dots, m$, όταν η διεργασία είναι εντός ελέγχου.

Κατά τον Jolayemi (1999) η πιθανότητα ανίχνευσης της μετατόπισης της διεργασίας σε ένα μόνο δείγμα, όταν η διεργασία είναι εντός ελέγχου, δίνεται από τον τύπο

¹ Η συνέλιξη είναι μία πράξη που εφαρμόζεται σε δύο συναρτήσεις, έστω ϕ , γ και συμβολίζεται με $\phi * \gamma$. Εάν οι συναρτήσεις είναι διακριτές η συνέλιξη ορίζεται από τον τύπο $(\phi * \gamma)(\chi) = \sum_{\kappa=-\infty}^{+\infty} \phi(\kappa) \cdot \gamma(\chi - \kappa)$, ενώ εάν οι συναρτήσεις είναι συνεχείς από τον τύπο $(\phi * \gamma)(\chi) = \int_{-\infty}^{+\infty} \phi(\kappa) \cdot \gamma(\chi - \kappa) d\kappa$.

$$P = 1 - \sum_{k=0}^{c_m} \binom{nm}{k} \bar{p}_1^k (1 - \bar{p}_1)^{nm-k} = 1 - \beta, \quad (3.3.3)$$

όπου \bar{p}_1 είναι ο μέσος των \bar{p}_{1k} , $k=1, 2, \dots, m$, \bar{p}_{1k} είναι το αναμενόμενο ποσοστό ελαττωματικών προϊόντων που αντιστοιχεί στην ιδιότητα X_k , $k=1, 2, \dots, m$, όταν η διεργασία είναι εκτός ελέγχου, P είναι η ισχύς του διαγράμματος και β είναι το ρίσκο του καταναλωτή. Επιπλέον, ισχύει

$$\sum_{k=0}^m \binom{nm}{k} \bar{p}_0^k (1 - \bar{p}_0)^{nm-k} = 1 - a, \quad (3.3.4)$$

όπου a είναι η πιθανότητα σήμανσης λανθασμένου συναγερμού, όταν η διεργασία είναι εντός ελέγχου, δηλαδή το ρίσκο του παραγωγού. Επειδή, όμως, είναι δύσκολη η επίλυση των παραπάνω δύο σχέσεων ως προς c_m και $n' = mn$, ο Jolayemi πρότεινε τους εξής τύπους

$$n' = (a^2 z_\beta^2 \bar{p}_1 \bar{q}_1) / (b - a\bar{p}_1)^2 \quad (3.3.5)$$

και

$$c = (n'b) / a, \quad (3.3.6)$$

όπου $a = z_\beta \sqrt{\bar{p}_1 \bar{q}_1 / \bar{p}_0 \bar{q}_0} - z_{1-a}$, $b = \bar{p}_0 z_\beta \sqrt{\bar{p}_1 \bar{q}_1 / \bar{p}_0 \bar{q}_0} - z_{1-a} \bar{p}_1$, $\bar{q}_0 = 1 - \bar{p}_0$, $\bar{q}_1 = 1 - \bar{p}_1$ και $z_\beta < 0$. Αφού υπολογίσουμε τις ποσότητες c_m και n' , τις αντικαθιστούμε στους τύπους (3.3.3) και (3.3.4) για να διαπιστώσουμε εάν τους ικανοποιούν.

Επίσης, ο Jolayemi (1999) πρότεινε μία μέθοδο για την ανίχνευση της μεταβλητής που προκαλεί το συναγερμό, η οποία βασίζεται σε κάποιες ιδιότητες. Όταν η ποιοτική ιδιότητα (*quality attribute*) προσδιορίζεται, ο τύπος του ελαττώματος (που αντιστοιχεί στην ιδιότητα αυτή) γίνεται εύκολα αναγνωρίσιμος. Το ποιοτικό χαρακτηριστικό μιας ιδιότητας μπορεί να μην εμπλέκει ακριβείς μετρήσεις. Ως αποτέλεσμα, όταν ένα προϊόν, του οποίου το ποιοτικό χαρακτηριστικό κρίνεται από μία ή περισσότερες ιδιότητες, είναι ελαττωματικό, τότε ο αριθμός των ελαττωμάτων που εμπλέκεται με κάθε ιδιότητα μπορεί εύκολα να καταμετρηθεί στο προϊόν. Συνεπώς, ο αριθμός αποδοχής c_k της X_k ιδιότητας είναι ο μεγαλύτερος ακέραιος μικρότερος ή ίσος της ποσότητας $np_{0k} + z_{1-a} \sqrt{np_{0k} q_{0k}}$, όπου τα n , \bar{q}_0 , \bar{p}_0 ορίζονται από τους

τύπους (3.3.5) και (3.3.6). Εάν στο διάγραμμα *MACC* σημάνει συναγερός, ο αριθμός των ελαττωμάτων d_k της ιδιότητας X_k , $k=1, 2, \dots, m$, όλων των ελαττωματικών προϊόντων d ($d > c$) του δείγματος μεγέθους n , πρέπει να καταμετρηθεί. Ενώ, εάν κάποιο από τα d_k , $k=1, 2, \dots, m$ είναι μεγαλύτερο από το αντίστοιχο c_k , θα πρέπει να ελεγχθούν οι ειδικές αιτίες μεταβλητότητας της X_k ιδιότητας.

Οι Chiu and Kuo (2007) κατασκεύασαν ένα διάγραμμα για τον έλεγχο πολυμεταβλητών αριθμήσιμων δεδομένων που προέρχονται από την κατανομή Poisson (*multivariate Poisson count data*) και το ονόμασαν **διάγραμμα MP** (*MP-chart*). Έστω ότι X_j είναι ο αριθμός ελαττωμάτων ή των μη συμμορφώσεων του j χαρακτηριστικού για $j=1, 2, \dots, p$. Υποθέτουμε ότι το διάνυσμα $\mathbf{X}=(X_1, X_2, \dots, X_n)$ ακολουθεί μία p -διάστατη κατανομή Poisson. Κάθε X_j ακολουθεί κατανομή Poisson με μέσο λ_j και η συνδιακύμανση μεταξύ δύο μεταβλητών (X_j, X_l) είναι θ_0 για $j \neq l$. Επιπρόσθετα, ορίζουμε τη στατιστική συνάρτηση D ως άθροισμα όλων των X_j

$$D = \sum_{j=1}^p X_j, \quad j=1, 2, \dots, p. \quad (3.3.7)$$

Οι Kemp and Kemp (1965), χρησιμοποιώντας το διμεταβλητό μοντέλο του Holgate (1964), απέδειξαν ότι για $p=2$ η συνάρτηση πυκνότητας πιθανότητας (σ.π.π.) της D έχει τη μορφή

$$P(D=d) = \exp[-(\lambda_1 + \lambda_2 - \theta_0)] \sum_{i=0}^{d/2} \frac{(\lambda_1 + \lambda_2 - 2\theta_0)^{d-2i} \theta_0^i}{(d-2i)! i!}, \quad d=0, 1, 2, \dots$$

Οι Chiu and Kuo (2007) χρησιμοποιώντας τη μέθοδο μετασχηματισμού των μεταβλητών, απέδειξαν ότι για $p=3$ η σ.π.π. της D έχει τη μορφή

$$P(D=d) = \exp[-(\lambda_1 + \lambda_2 + \lambda_3 - 2\theta_0)] \sum_{i=0}^{d/3} \frac{(\lambda_1 + \lambda_2 + \lambda_3 - 3\theta_0)^{d-3i} \theta_0^i}{(d-3i)! i!}, \quad d=0, 1, 2, \dots,$$

ενώ χρησιμοποιώντας μαθηματική επαγωγή, έδειξαν ότι για $p > 3$ η σ.π.π. της D είναι

$$P(D = d) = \exp\left\{-\left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0\right]\right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0\right)^{d-pi} \theta_0^i}{(d-pi)! i!}, \quad d = 0, 1, 2, \dots$$

Συνεπώς, τα όρια ελέγχου του διαγράμματος μπορούν να προσδιοριστούν, εάν πάρουμε τα άνω και κάτω ($a/2$)-ποσοστιαία σημεία της ακριβούς κατανομής του τύπου (3.3.7). Τα όρια ελέγχου πρέπει να ικανοποιούν τις ακόλουθες σχέσεις

$$P(D < LCL) = \sum_{d=0}^{LCL} \exp\left\{-\left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0\right]\right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0\right)^{d-pi} \theta_0^i}{(d-pi)! i!} \leq \frac{a}{2}$$

και

$$\begin{aligned} P(D > UCL) &= \sum_{d=UCL}^{\infty} \exp\left\{-\left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0\right]\right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0\right)^{d-pi} \theta_0^i}{(d-pi)! i!} = \\ &= 1 - \sum_{d=0}^{UCL-1} \exp\left\{-\left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0\right]\right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0\right)^{d-pi} \theta_0^i}{(d-pi)! i!} \leq \frac{a}{2}. \end{aligned}$$

Γενικά, η πολυμεταβλητή διαδικασία Poisson για μικρές τιμές στο μέσο κάθε μεταβλητής X_j , έχει σαν αποτέλεσμα η στατιστική συνάρτηση D να παίρνει μικρές ή μηδενικές τιμές. Αυτό σημαίνει ότι τις περισσότερες φορές η πιθανότητα $P(D = 0)$ θα είναι μεγαλύτερη από $a/2$. Εάν, η πιθανότητα $P(D = 0)$ πάρει τιμή μεγαλύτερη από $a/2$, τότε το κάτω όριο ελέγχου τίθεται ίσο με το μηδέν ($LCL = 0$) και χρησιμοποιείται μόνο το άνω όριο. Στην περίπτωση που a είναι ο λανθασμένος συναγερμός, το άνω όριο ελέγχου μπορεί να ληφθεί από τον τύπο

$$P(D > UCL) = \sum_{d=UCL}^{\infty} \exp\left\{-\left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0\right]\right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0\right)^{d-pi} \theta_0^i}{(d-pi)! i!} =$$

$$= 1 - \sum_{d=0}^{UCL-1} \exp \left\{ - \left[\sum_{j=1}^p \lambda_j - (p-1)\theta_0 \right] \right\} \sum_{i=0}^{d/p} \frac{\left(\sum_{j=1}^p \lambda_j - p\theta_0 \right)^{d-pi} \theta_0^i}{(d-pi)! i!} < a.$$

Εάν τα λ_j και θ_0 είναι άγνωστα, μπορούν να εκτιμηθούν από τα δεδομένα. Όταν καθοριστεί η τιμή του άνω ορίου, τότε μπορούμε να απεικονίσουμε τη στατιστική συνάρτηση D στο διάγραμμα MP για να ελέγξουμε τη διεργασία.

Οι Xie, Goh and Kuralmani (2002) έδειξαν ότι μπορεί να χρησιμοποιηθεί η κανονική κατανομή για να προσεγγίσει την κατανομή Poisson, όταν η μέση τιμή είναι μεγαλύτερη από πέντε. Κατά συνέπεια, για $D > 5$ οι Chiu and Kuo (2007) χρησιμοποίησαν την προσέγγιση κανονικής κατανομής. Με τον τρόπο αυτό, από τη μέση τιμή και τη διακύμανση της συνάρτησης D , προσδιόρισαν 3σ όρια ελέγχου τύπου Shewhart της μορφής $E(D) \pm 3 \cdot \sqrt{V(D)}$. Η μέση τιμή της στατιστικής συνάρτησης D υπολογίζεται από τον τύπο

$$E(D) = E \left(\sum_{j=1}^p X_j \right) = \sum_{j=1}^p E(X_j) = \sum_{j=1}^p \lambda_j,$$

ενώ η διακύμανση της D προκύπτει ως εξής

$$\begin{aligned} V(D) &= V \left(\sum_{j=1}^p X_j \right) = \sum_{j=1}^p V(X_j) + 2 \sum_{j < l} Cov(X_j, X_l) = \sum_{j=1}^p \lambda_j + 2 \sum_{j < l} \rho_{jl} \sqrt{V(X_j)V(X_l)} = \\ &= \sum_{j=1}^p \lambda_j + 2 \sum_{j < l} \rho_{jl} \sqrt{\lambda_j \lambda_l}, \end{aligned}$$

όπου λ_j είναι ο μέσος του j τύπου ελαττωμάτων ή μη συμμορφώσεων και ρ_{jl} είναι ο συντελεστής συσχέτισης μεταξύ των X_j και X_l . Εάν οι τιμές των λ_j και ρ_{jl} είναι άγνωστες, μπορούν να εκτιμηθούν από τα δεδομένα. Συνεπώς, τα όρια ελέγχου τύπου Shewhart που προκύπτουν έχουν την ακόλουθη μορφή

$$\begin{array}{l}
 \text{Διάγραμμα} \\
 \text{MP} \\
 \text{Τύπου Shewhart}
 \end{array}
 \left\{ \begin{array}{l}
 LCL = \sum_{j=1}^p \lambda_j - 3 \sqrt{\sum_{j=1}^p \lambda_j + 2 \sum_{j<l} \rho_{jl} \sqrt{\lambda_j \lambda_l}} \\
 CL = \sum_{j=1}^p \lambda_j \\
 UCL = \sum_{j=1}^p \lambda_j + 3 \sqrt{\sum_{j=1}^p \lambda_j + 2 \sum_{j<l} \rho_{jl} \sqrt{\lambda_j \lambda_l}}
 \end{array} \right. \quad (3.3.8)$$

Σύμφωνα με τους Chiu and Kuo (2007), όταν η διεργασία είναι εντός ελέγχου, το εντός ελέγχου μέσο μήκος ροής (ARL_0) είναι ο μέσος αριθμός των δειγμάτων πριν δοθεί ένα σήμα. Ενώ, το εκτός ελέγχου μέσο μήκος ροής (ARL_1) είναι ο μέσος αριθμός των δειγμάτων που απαιτούνται για να ανιχνεύσουν μία μετατόπιση του μέσου, όταν η διεργασία είναι εκτός ελέγχου. Επιπλέον, οι συγγραφείς υποστηρίζουν το διάγραμμα *MP* είναι πιο κατάλληλο από το διάγραμμα ελέγχου τύπου Shewhart, όταν η συσχέτιση μεταξύ των μεταβλητών υπάρχει και ότι είναι πιο ευαίσθητο σε μεγάλες τιμές του συντελεστή συσχέτισης για την εκτός ελέγχου κατάσταση.

Οι Gadre and Rattihalli (2005) πρότειναν τρία πολυμεταβλητά διαγράμματα ελέγχου για ιδιότητες, τα οποία βασίζονται σε ομάδες επιθεώρησης για την ανίχνευση οποιασδήποτε υποβάθμισης της διεργασίας. Τα διαγράμματα αυτά είναι το *np* (*Multiattribute np-chart or MA-np chart*), το *συνθετικό* (*Multiattribute synthetic or MA-Syn chart*) και το *διάγραμμα ομάδων ροών* (*Multiattribute Group Runs or MA-GR chart*), τα οποία αναπτύσσονται βάσει του *MP*-ελέγχου. Οι ίδιοι έδειξαν ότι το διάγραμμα *MA-GR* συμπεριφέρεται καλύτερα τόσο από το *MA-Syn* όσο και από το *MA-np*. Μεταξύ των *MA-Syn* και *MA-np*, το δεύτερο διάγραμμα έχει τη χειρότερη απόδοση. Επιπρόσθετα, παρουσίασαν δύο τεχνικές για τον εντοπισμό των ιδιοτήτων εκείνων που είναι υπεύθυνες για την εκτός ελέγχου κατάσταση της διεργασίας. Οι τεχνικές αυτές βασίζονται στο εάν οι μη συμμορφωμένες ομάδες υπήρχαν πριν ή μετά τη σήμανση του συναγερμού.

Τέλος, οι Niaki and Abbasi (2007b) παρουσίασαν μία τεχνική για τη μείωση της ασυμμετρίας των πολυμεταβλητών διαγραμμάτων για ιδιότητες, μιας και η ασυμμετρία είναι το κυριότερο μειονέκτημα της Διωνυμικής και της Poisson κατανομής. Τον επόμενο χρόνο, οι ίδιοι συγγραφείς (Niaki and Abbasi (2008)) παρουσίασαν μία τεχνική μετασχηματισμού για την κατασκευή πολυμεταβλητών διαγραμμάτων *C* για ιδιότητες. Θα πρέπει να

αναφέρουμε ότι και οι δύο τεχνικές βασίζονται στον ίδιο αλγόριθμο. Επιπλέον, οι Niaki and Abbasi (2007a) πρότειναν για την κατασκευή πολυμεταβλητών διαγραμμάτων για ιδιότητες τη χρήση της μεθόδου Bootstrap.

3.4 Πολυμεταβλητά Διαγράμματα Ελέγχου με μνήμη

Τα πολυμεταβλητά διαγράμματα ελέγχου με μνήμη (*multivariate control charts with memory*) χρησιμοποιούν πληροφορίες όχι μόνο από το πιο πρόσφατο δείγμα, αλλά και από προγενέστερα δείγματα. Κατά συνέπεια, είναι ευαίσθητα σε μικρές και μέτριες μετατοπίσεις του διανύσματος των μέσων, σε αντίθεση με τα διαγράμματα ελέγχου τύπου Shewhart.

3.4.1 Πολυμεταβλητό Διάγραμμα Ελέγχου Συσσωρευμένου Αθροίσματος

Στη βιομηχανία, συχνά, χρησιμοποιούνται μονοδιάστατα διαγράμματα *CUSUM* για την παρακολούθηση διαφορετικών επιπέδων μεταβλητών, τα οποία μετρούν από κοινού την ποιότητα της διεργασίας. Τέτοιου είδους τεχνικές προτάθηκαν από τους Jackson and Mudholkan (1979) και Elder, Provost and Ecker (1981). Οι Woodall and Ncube (1985) πρότειναν την ταυτόχρονη χρήση μονοδιάστατων διαγραμμάτων ελέγχου *CUSUM*, την οποία αναφέρουν ως **ενιαία πολυμεταβλητή διαδικασία συσσωρευμένου αθροίσματος** (*single Multivariate CUSUM procedure, MCUSUM*).

Κατά τους Woodall and Ncube (1985), θεωρούμε ότι $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ είναι m ανεξάρτητα δείγματα (μεγέθους n το καθένα) διαθέσιμα από μία διεργασία, όπου $\mathbf{X}_k = (\mathbf{X}_{1k}, \dots, \mathbf{X}_{nk})$, $k = 1, 2, \dots, m$. Οι μεταβλητές $\mathbf{X}_{ik} = (X_{i1k}, \dots, X_{ipk})'$, για $i = 1, 2, \dots, n$ και $k = 1, 2, \dots, m$, προέρχονται από μία p -διάστατη κανονική κατανομή, η οποία έχει μέσο το διάνυσμα $E(\mathbf{X}_{ik}) = \boldsymbol{\mu}_{ik} = (\mu_{i1k}, \mu_{i2k}, \dots, \mu_{ipk})'$ και πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}_{(p \times p)}$, τα οποία είναι γνωστά. Χωρίς βλάβη της γενικότητας θεωρούμε ότι $\boldsymbol{\mu}_{ik} = (0, 0, \dots, 0)' = \mathbf{0}$ με $i = 1, \dots, n$ και $k = 1, \dots, m$. Για την εφαρμογή μιας πολυδιάστατης διαδικασίας *MCUSUM*,

για κάθε ακολουθία τυχαίων μεταβλητών $\{\mathbf{X}_{ik}\}$, $i=1, \dots, n$ και $k=1, \dots, m$, εφαρμόζεται είτε ένα μονόπλευρο είτε ένα δίπλευρο διάγραμμα *CUSUM*. Το **μήκος ροής ενός μονόπλευρου διαγράμματος *CUSUM*** για την k μεταβλητή είναι $N(k) = \min\{n: S_{i,k} \geq h_k\}$, όπου $k=1, \dots, m$ και $h_k > 0$. Η επιλογή του h_k γίνεται ώστε να έχουμε το επιθυμητό ARL_0 . Η στατιστική συνάρτηση $S_{i,k}$ υπολογίζεται μέσω του τύπου

$$S_{i,k} = \max\left[0, S_{i-1,k} + X_{ik} - K_k\right], \begin{cases} i=1, 2, \dots, n \\ k=1, 2, \dots, m \end{cases}$$

όπου $0 \leq S_{0,k} \leq h_k$ και $K_k \geq 0$. Για την ανίχνευση μετατοπίσεων σε οποιαδήποτε κατεύθυνση της ακολουθίας $\{\mu_{ik}\}$, ένα **δίπλευρο διάγραμμα *CUSUM*** δίνει σήμα εκτός ελέγχου στο βήμα $N(k) = \min\{n: S_{i,k} \geq h_k \text{ ή } T_{i,k} \leq -h_k\}$, όπου η στατιστική συνάρτηση $T_{i,k}$ είναι η εξής

$$T_{i,k} = \min\left[0, T_{i-1,k} + X_{ik} + K_k\right], \begin{cases} i=1, 2, \dots, n \\ k=1, 2, \dots, m \end{cases}$$

και $-h_k \leq T_{0,k} \leq 0$. Η τιμή αναφοράς K_k , που επιλέγεται για την k μεταβλητή, καθορίζεται από τη μετατόπιση του μέσου της μεταβλητής που είναι εκτός ελέγχου και χρήζει άμεσης ανίχνευσης. Τελικά, το **μήκος ροής της διαδικασίας *MCUSUM*** είναι

$$N = \min\{N(1), N(2), \dots, N(m)\}.$$

Σύμφωνα με την μέθοδο των Woodall and Ncube (1985), η διεργασία του πολυμεταβλητού συσσωρευμένου αθροίσματος τίθεται εκτός στατιστικού ελέγχου, εάν σε κάποιο από τα μονοδιάστατα διαγράμματα *CUSUM* σημάνει συναγερμός. Οι Woodall and Ncube έδειξαν, επίσης, ότι το μέσο μήκος ροής *ARL* εξαρτάται από την κατεύθυνση της μετατόπισης του μέσου της διεργασίας. Ωστόσο, οι Pignatiello and Runger (1990) έδειξαν ότι το παραπάνω διάγραμμα ελέγχου δεν έχει καλό *ARL*, όταν ο μέσος μετατοπίζεται ταυτόχρονα σε πολλά χαρακτηριστικά.

Ο Healy (1987) ανέπτυξε δύο πολυμεταβλητά διαγράμματα *CUSUM*, βασιζόμενος στο γεγονός ότι τα μονοδιάστατα διαγράμματα *CUSUM* είναι μια σειρά διαδοχικών ελέγχων πιθανοφάνειας. Υποθέτουμε ότι οι μεταβλητές $\mathbf{X}_{ik} = (X_{i1k}, X_{i2k}, \dots, X_{ipk})'$ για $i=1, 2, \dots, n$ και $k=1, 2, \dots, m$ προέρχονται από μία p -διάστατη κανονική κατανομή, με μέση τιμή το

εντός ελέγχου $(p \times 1)$ διάνυσμα μ_0 και γνωστό $(p \times p)$ πίνακα διακύμανσης-συνδιακύμανσης Σ . Επιπλέον, υποθέτουμε ότι μ_1 είναι το εκτός ελέγχου $(p \times 1)$ μέσο διάνυσμα. Ο μετατοπισμένος μέσος μ_1 είναι γνωστός υπό την έννοια ότι το μέγεθος και η κατεύθυνση της μετατόπισης είναι γνωστά. Ένα **πολυμεταβλητό διάγραμμα ελέγχου CUSUM για την ανίχνευση της μετατόπισης του μέσου από το μ_0 στο μ_1** , βασίζεται στην ακόλουθη στατιστική συνάρτηση

$$MC_k = \max \left[MC_{k-1} + a' \cdot (X_{ik} - \mu_0) - \frac{1}{2} \cdot \lambda(\mu_1), 0 \right], \quad \begin{cases} i = 1, 2, \dots, n \\ k = 1, 2, \dots, m \end{cases}$$

όπου

$$a' = \frac{(\mu_1 - \mu_0)' \Sigma^{-1}}{\left[(\mu_1 - \mu_0)' \Sigma^{-1} (\mu_1 - \mu_0) \right]^{1/2}} = \frac{(\mu_1 - \mu_0)' \Sigma^{-1}}{\lambda(\mu_1)}$$

και η παράμετρος μη κεντρικότητας (*no centrality parameter*) υπολογίζεται ως εξής

$$\lambda(\mu_1) = \sqrt{(\mu_1 - \mu_0)' \Sigma^{-1} (\mu_1 - \mu_0)}. \quad (3.4.1)$$

Στην πραγματικότητα, το παραπάνω διάγραμμα ελέγχει την υπόθεση $H_0 : \mu = \mu_0$ έναντι της εναλλακτικής $H_1 : \mu_0 = \mu_1$ με $\mu_0 \neq \mu_1$. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, όταν η παραπάνω στατιστική συνάρτηση ελέγχου είναι ίση ή ξεπεράσει τη σταθερά απόφασης H , δηλαδή $MC_k \geq H$.

Ο Healy (1987) πρότεινε και ένα πολυμεταβλητό διάγραμμα ελέγχου CUSUM για την ανίχνευση αλλαγής στον πίνακα διακύμανσης-συνδιακύμανσης Σ . Θεώρησε ότι ο πίνακας διακύμανσης-συνδιακύμανσης μετατοπίστηκε από τη θέση $\Sigma = \Sigma_0$ στη θέση $\Sigma_1 = C\Sigma_0 = C\Sigma$, όπου C είναι μία σταθερά. Συνεπώς, ένα **διάγραμμα MCUSUM για την παρακολούθηση του πίνακα διακύμανσης-συνδιακύμανσης**, δεδομένου ότι το μέσο διάνυσμα μ παραμένει σταθερό σε όλη τη διεργασία, βασίζεται στη στατιστική συνάρτηση που ακολουθεί

$$MC_k = \max [MC_{k-1} + Y_k - K, 0], \quad k = 1, 2, \dots, m,$$

όπου $Y_k = (\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu})$ και $K = p \cdot \log(C) \cdot [C / (C - 1)]$. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, όταν $MC_k \geq H$. Η προαναφερθείσα μέθοδος υποθέτει ότι έχουμε n παρατηρήσεις σε κάθε χρονική στιγμή και ότι \mathbf{X}_{ik} είναι η i παρατήρηση της k μεταβλητής. Η προηγούμενη στατιστική συνάρτηση μπορεί εναλλακτικά να γραφεί ως εξής

$$MC_k = \max \left[MC_{k-1} + \sum_{i=1}^n (\mathbf{X}_{ik} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_{ik} - \boldsymbol{\mu}) - K', 0 \right], \quad k = 1, 2, \dots, m,$$

όπου $K' = pm \cdot \log(C) \cdot [C / (C - 1)]$.

Ο Crosier (1988) πρότεινε ένα πολυμεταβλητό διάγραμμα ελέγχου *CUSUM* για την ανίχνευση της μετατόπισης του μέσου το οποίο χρησιμοποιεί την τετραγωνική ρίζα του Y_k , δηλαδή την ποσότητα $\sqrt{Y_k} = \left[(\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu}) \right]^{1/2}$. Στην περίπτωση αυτή, η στατιστική συνάρτηση δίνεται από τον τύπο

$$MC_k = \max \left[MC_{k-1} + \sqrt{Y_k} - K', 0 \right], \quad k = 1, 2, \dots, m,$$

όπου $C_0 \geq 0$ και $K' > 0$. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, όταν $MC_k \geq H$. Το παραπάνω διάγραμμα ονομάζεται *COT* (*CUSUM of T*), επειδή χρησιμοποιεί το T του Hotelling $\left(T_k^2 = (\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu}) = Y_k \Rightarrow T_k = \left[(\mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_k - \boldsymbol{\mu}) \right]^{1/2} = \sqrt{Y_k} \right)$. Στην προσπάθειά του να βρει τη βέλτιστη τιμή του K' , ο Crosier δημιούργησε μια ακολουθία τιμών πολύ κοντά στην τετραγωνική ρίζα του αριθμού των χαρακτηριστικών $(K' \cong \sqrt{p})$.

Ο Crosier (1988) πρότεινε και ένα δεύτερο πολυμεταβλητό διάγραμμα ελέγχου *CUSUM* για την ανίχνευση της μετατόπισης του μέσου, που προκύπτει αντικαθιστώντας τις αριθμητικές ποσότητες του μονοδιάστατου *CUSUM* με διανύσματα. Η στατιστική συνάρτηση του διαγράμματος αυτού δίνεται από τον τύπο

$$MC_k = \left[\mathbf{S}'_k \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{S}_k \right]^{1/2}, \quad (3.4.2)$$

όπου

$$C_k = \left[(\mathbf{S}_{k-1} + \mathbf{X}_k - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{S}_{k-1} + \mathbf{X}_k - \boldsymbol{\mu}) \right]^{1/2}$$

και

$$\mathbf{S}_k = \begin{cases} 0 & , \text{ για } C_k \leq K \\ (\mathbf{S}_{k-1} + \mathbf{X}_k - \boldsymbol{\mu})(1 - K/C_k) & , \text{ για } C_k > K. \end{cases}$$

Επιπλέον, $\mathbf{S}_0 = \mathbf{0}$ και $K > 0$. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, όταν $MC_k \geq H$, όπου η σταθερά H είναι τέτοια ώστε το εντός ελέγχου μήκος ροής ARL_0 να πάρει την προκαθορισμένη τιμή που προέκυψε μέσω προσομοίωσης. Κατά τον Crosier, το μέσο μήκος ροής ARL μειώνεται για $K = \lambda(\boldsymbol{\mu}_1)/2$, όπου $\lambda(\boldsymbol{\mu}_1)$ είναι η παράμετρος μη κεντρικότητας του τύπου (3.4.1).

Και τα δύο διαγράμματα που πρότεινε ο Crosier ανιχνεύουν γρηγορότερα μικρές μετατοπίσεις του μέσου διανύσματος σε σχέση με τα πολυμεταβλητά διαγράμματα Shewhart. Ωστόσο, είναι προτιμότερο να χρησιμοποιούμε το δεύτερο διάγραμμα έναντι του διαγράμματος COT , διότι δίνει μία ένδειξη της κατεύθυνσης της μετατόπισης του μέσου. Στο συμπέρασμα αυτό κατέληξε ο Crosier μετά από υπολογισμούς του μέσου μήκους ροής ARL για κάθε ένα από τα διαγράμματα αυτά, το οποίο εξαρτάται από το μέσο διάνυσμα και στη συσχετισμένη δομή των δεδομένων μόνο μέσω της παραμέτρου μη κεντρικότητας.

Επιπρόσθετα, ο Crosier υποστήριξε ότι η ανάπτυξη ενός **συνδυασμένου πολυμεταβλητού διαγράμματος *Shewhart-CUSUM*** (*combined multivariate Shewhart-CUSUM chart*) μπορεί να γίνει θέτοντας σε λειτουργία ένα $MCUSUM$ και ένα πολυμεταβλητό διάγραμμα Shewhart ταυτόχρονα. Στην περίπτωση αυτή, υπάρχει ένδειξη μετατόπισης του μέσου της διεργασίας, όταν σε ένα από δύο διαγράμματα σημάνει συναγερμός. Επίσης, τόνισε ότι ένα πολυμεταβλητό διάγραμμα $CUSUM$ μπορεί να είναι ισχυρό-ανθεκτικό (*robust*) εφαρμόζοντας τον κανόνα «δύο στη σειρά» (*two-in-a row rule*) στο T του Hotelling για τις παρατηρήσεις. Σε ορισμένες εφαρμογές, αντί της στατιστικής συνάρτησης T , μπορούν να εφαρμοστούν ειδικά κριτήρια για ακραίες τιμές (*special outlier criteria*), ώστε να προσαρμόσουν την ανθεκτικότητα μιας διεργασίας. Συμπληρωματικά, ο Crosier αναφέρει ότι η μέθοδος αρχικής αντίδρασης (FIR) εφαρμόζεται σε ένα πολυμεταβλητό διάγραμμα ελέγχου $CUSUM$, αλλάζοντας την αρχική τιμή του H . Συνήθως, τίθεται $H_0 = H/2$ και

$H_k = \min \left[H, H_{k-1} + \max(0, K^* - T_k) \right]$, $k=1, 2, \dots, m$, όπου T_k η τιμή της T στατιστικής συνάρτησης του Hotelling του k δείγματος και K^* η τιμή της παραμέτρου K ενός COT διαγράμματος σχεδιασμένο, ώστε να ανιχνεύει την ίδια απόκλιση που ανιχνεύει και ένα $MCUSUM$. Η τελευταία μέθοδος μειώνει το εντός ελέγχου ARL περισσότερο για μεγάλες τιμές του p παρά για μικρότερες, χωρίς να επιφέρει τις δυσμενείς συνέπειες του διαγράμματος COT , διότι και το εκτός ελέγχου ($\lambda(\boldsymbol{\mu}_1) = 1$) ARL μειώνεται περισσότερο για μεγάλες τιμές του p παρά για μικρότερες.

Οι Pignatiello and Runger (1990) πρότειναν δύο **πολυμεταβλητά διαγράμματα $CUSUM$, τα οποία χρησιμοποιούν τον πίνακα διακύμανσης-συνδιακύμανσης**. Και τα δύο πολυμεταβλητά συσσωρευμένα αθροίσματα βασίζονται σε τετραγωνικές μορφές του διανύσματος του μέσου. Ωστόσο, στο πρώτο διάγραμμα ελέγχου συσσωρεύονται πρώτα τα διανύσματα \mathbf{X}_k και στη συνέχεια δημιουργούνται οι τετραγωνικές μορφές, ενώ στο δεύτερο διάγραμμα πρώτα υπολογίζεται η τετραγωνική μορφή για κάθε \mathbf{X}_k και μετά συσσωρεύονται αυτές οι τετραγωνικές μορφές. Το δεύτερο διάγραμμα των Pignatiello and Runger είναι παρόμοιο με το πρώτο διάγραμμα που προτάθηκε από τον Crosier.

Το πρώτο διάγραμμα των Pignatiello and Runger έχει την εξής στατιστική συνάρτηση

$$MC_k = \max \left[\|\mathbf{C}_k\| - K \cdot n_k, 0 \right], \quad (3.4.3)$$

όπου το πολυμεταβλητό άθροισμα και η νόρμα του υπολογίζονται αντίστοιχα ως εξής

$$\mathbf{C}_k = \sum_{l=k-n_k+1}^k (\mathbf{X}_l - \boldsymbol{\mu}), \quad \|\mathbf{C}_k\| = \sqrt{\mathbf{C}'_k \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{C}_k}.$$

Το n_k είναι ο αριθμός των υποομάδων από την πιο πρόσφατη ανανέωση του $CUSUM$ (π.χ. την τιμή μηδέν) και ορίζεται ως εξής

$$n_k = \begin{cases} n_{k-1} + 1, & \text{για } MC_{k-1} > 0 \\ 1, & \text{για } MC_{k-1} \leq 0 \end{cases} \quad \text{ή} \quad \frac{1}{n_k} \mathbf{C}_k = \left(\frac{1}{n_k} \sum_{l=k-n_k+1}^k \mathbf{X}_l \right) - \boldsymbol{\mu}.$$

Οι Pignatiello and Runger προτείνουν το K να είναι το μισό της απόστασης του εκτός ελέγχου μέσου διανύσματος $\boldsymbol{\mu}_1$ από το εντός ελέγχου μέσο διάνυσμα $\boldsymbol{\mu}_0$, δηλαδή

$K = \lambda(\boldsymbol{\mu}_1)/2 = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}/2$. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, όταν $MC_k \geq H$, όπου το H προκύπτει μέσω προσομοίωσης.

Στο δεύτερο διάγραμμα των Pignatiello and Runger απεικονίζεται η συνάρτηση

$$MC_k = \max[MC_{k-1} + D_k^2 - K, 0]. \quad (3.4.4)$$

Η παραπάνω στατιστική συνάρτηση βασίζεται στο τετράγωνο της απόστασης του k δειγματικού μέσου από το μέσο $\boldsymbol{\mu}_0$ (υπό τη μηδενική υπόθεση $\boldsymbol{\mu} = \boldsymbol{\mu}_0$), όπου $MC_0 = 0$ και $D_k^2 = (\mathbf{X}_k - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1}(\mathbf{X}_k - \boldsymbol{\mu}_0)$, $k = 1, 2, \dots, m$. Η τιμή που συστήνουν οι Pignatiello και Runger για το K είναι

$$K = \frac{1}{2} \cdot \lambda^2(\boldsymbol{\mu}_1) + p = \frac{1}{2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) + p.$$

Οι Pignatiello and Runger απέδειξαν ότι το ARL του πρώτου διαγράμματος (τύπος (3.4.3)) που πρότειναν εξαρτάται μόνο από την τετραγωνική ρίζα της παραμέτρου μη κεντρικότητας και είναι καλύτερο σε σχέση με αυτό του δεύτερου διαγράμματός τους (τύπος (3.4.4)). Επιπλέον, υποστήριξαν ότι το ARL του πρώτου διαγράμματος (τύπος (3.4.3)) που πρότειναν είναι περίπου ίδιο με αυτό του ARL του δεύτερου διαγράμματος του Crosier (τύπος (3.4.4)), υποθέτοντας ότι οι παράμετροι είναι γνωστές. Ωστόσο, συγκρίνοντας τα διαγράμματα αυτά, οι Mahmoud and Maravelakis (2011) κατέληξαν στο συμπέρασμα ότι η επίδοση του εκτός ελέγχου μέσου μήκους ροής (ARL_1) του διαγράμματος των Pignatiello and Runger είναι καλύτερη από αυτή του διαγράμματος του Crosier για $\lambda \geq 2K$, όπου λ είναι το μέγεθος της μετατόπισης της παραμέτρου μη κεντρικότητας. Για $\lambda < 2K$ συμβαίνει το αντίθετο, υποθέτοντας πάντα ότι το μέσο διάνυσμα και ο πίνακας διακύμανσης-συνδιακύμανσης της διεργασίας είναι γνωστά ή μπορούν να εκτιμηθούν.

Επιπρόσθετα, οι Mahmoud and Maravelakis (2011) πρότειναν διαφορετικές τιμές για το άνω όριο ελέγχου (UCL) του πρώτου διαγράμματος των Pignatiello and Runger (MCI) και του δεύτερου διαγράμματος του Crosier ($MCUSUM$), οι οποίες παράγουν εντός ελέγχου μέσο μήκος ροής ίσο με 200 ($ARL_0 = 200$), όταν οι παράμετροι είναι γνωστές, για διαφορετικές τιμές p των και K . Οι τιμές αυτές δίνονται στον Πίνακα 3.1 που ακολουθεί.

Διάγραμμα	p	K		
		0.25	0.50	1.00
<i>MCUSUM</i>	2	8.64	5.49	3.00
	3	10.85	6.90	3.77
	5	14.81	9.40	5.21
<i>MCI</i>	2	7.49	4.78	2.71
	3	8.75	5.52	3.14
	5	10.80	6.81	3.92

Πίνακας 3.1: Τιμές άνω ορίου ελέγχου ενός διαγράμματος *MCUSUM* με $ARL_0 = 200$.

Το 1991 ο Hawkins πρότεινε δύο διαγράμματα *CUSUM* για προσαρμοσμένες μεταβλητές παλινδρόμησης. Στην πραγματικότητα βασίστηκε στην ιδέα ότι στην πράξη οι μετατοπίσεις μιας διεργασίας έχουν γνωστή δομή. Συγκεκριμένα, θεώρησε ότι η μετατόπιση του μέσου σε μία μόνο μεταβλητή έχει μέγεθος δ . Επιπλέον, θεώρησε την πολλαπλή παλινδρόμηση της X_i , η οποία είναι η i μεταβλητή του \mathbf{X} , σε όλες τις άλλες μεταβλητές του \mathbf{X} . Η παλινδρόμηση μπορεί να μοντελοποιηθεί ως εξής

$$X_i - \mu_i = \sum_{j \neq i} \beta_{ij} \cdot (X_j - \mu_j) + \varepsilon_i.$$

Σύμφωνα με τον Hawkins, ο εκτός ελέγχου μέσος μίας μόνο μεταβλητής μπορεί να μετατοπίσει την κατανομή της διεργασίας στην $N(\mu_0 + \delta \cdot \mathbf{e}_i, \Sigma_0)$, όπου \mathbf{e}_i είναι ένα διάνυσμα γραμμή με τιμή 1 στην i γραμμή και τιμή ίση με 0 διαφορετικά. Ο αριθμός δ είναι το μέγεθος της μετατόπισης. Συνεπώς, η στατιστική συνάρτηση του ελέγχου πιθανοφάνειας της (μηδενικής) υπόθεσης $\delta = 0$ δίνεται από τον τύπο

$$Z_i = \left[(X_i - \mu_i) - \sum_{i \neq j} \beta_{ij} \cdot (X_i - \mu_i) \right] / \tau_{ii}^{1/2}.$$

Ο αριθμός Z_i είναι το υπόλοιπο όταν η μεταβλητή X_i παλινδρομείται σε όλες τις άλλες μεταβλητές του \mathbf{X} . Η ποσότητα Z_i ακολουθεί την τυπική κανονική κατανομή. Με τ_{ii} συμβολίζονται τα κατάλοιπα της διακύμανσης των ε_i , τα οποία σύμφωνα με τον Weisberg (1985) υπολογίζονται από τον τύπο

$$\tau_{ii} = \sigma_{ii} - \sum_{i \neq j} \beta_{ij} \cdot \sigma_{ij}.$$

Ισχύει ότι $\sigma^{ii} = \tau_{ii}^{-1}$ και $\sigma^{ij} = -\beta_{ij} / \tau_{ii}$ για $i \neq j$ (Raveh (1985)). Θα πρέπει να αναφέρουμε ότι σ_{ij} και σ^{ij} είναι το ij στοιχείο των πινάκων Σ_0 και Σ_0^{-1} , αντίστοιχα. Το διάνυσμα των καταλοίπων αυτών είναι

$$\mathbf{Y} = \Sigma^{-1}(\mathbf{X} - \mu_0).$$

Το i στοιχείο του \mathbf{Y} είναι το κατάλοιπο της παλινδρόμησης, όταν η i μεταβλητή παλινδρομείται σε όλες τις άλλες μεταβλητές, χρησιμοποιώντας ως κλίμακα τον παράγοντα τ_{ii}^{-1} . Όταν η διεργασία είναι εντός ελέγχου, το διάνυσμα $\mathbf{Y} \sim N(0, \Sigma_0^{-1})$. Αλλάζοντας την κλίμακα του \mathbf{Y} προκύπτει το διάνυσμα \mathbf{Z} , το οποίο υπολογίζεται ως εξής

$$\mathbf{Z} = \mathbf{A}(\mathbf{X} - \mu_0) = [\text{diag}(\Sigma^{-1})]^{-1/2} \Sigma^{-1}(\mathbf{X} - \mu_0),$$

όπου $\mathbf{A} = [\text{diag}(\Sigma^{-1})]^{-1/2} \Sigma^{-1}$. Εάν $\mathbf{B} = [\text{diag}(\Sigma^{-1})]^{-1/2} \Sigma^{-1} [\text{diag}(\Sigma^{-1})]^{-1/2}$, για το διάνυσμα \mathbf{Z} ισχύει ότι $\mathbf{Z} \sim N(0, \mathbf{B})$.

Το διάνυσμα \mathbf{Z} παρέχει τη δυνατότητα ξεχωριστού ελέγχου των μεμονωμένων μεταβλητών του \mathbf{X} . Με τον τρόπο αυτό, ελέγχουμε ξεχωριστά p διαγράμματα, ένα για κάθε μεταβλητή του \mathbf{Z} (Hawkins (1991),(1993)). Ορίζουμε τις στατιστικές συναρτήσεις

$$L_{i,n}^+ = \max[L_{i,n-1}^+ + Z_{n,i} - K, 0] \quad (3.4.5)$$

$$L_{i,n}^- = \min[L_{i,n-1}^- + Z_{n,i} - K, 0] \quad (3.4.6)$$

$$S_{i,n}^+ = \max[S_{i,n-1}^+ + W_{n,i} - K, 0] \quad (3.4.7)$$

$$S_{i,n}^- = \min[S_{i,n-1}^- + W_{n,i} - K, 0], \quad (3.4.8)$$

όπου $L_{i,0}^+ = L_{i,0}^- = S_{i,0}^+ = S_{i,0}^- = 0$ και $W_{n,i} = (|Z_{n,i}|^{1/2} - 0.822) / 0.349$. Το διάστημα απόφασης των συσσωρευμένων αθροισμάτων L^+ και L^- ελέγχουν για μετατοπίσεις της θέσης ανοδικής και καθοδικής κατεύθυνσης, αντίστοιχα. Τα συσσωρευμένα αθροίσματα S^+ και S^- ελέγχουν για μετατοπίσεις της μεταβλητότητας ανοδικής και καθοδικής κατεύθυνσης, αντίστοιχα. Οι Hawkins and Wixley (1986) έδειξαν ότι εάν $Z \sim N(0, 1)$, τότε η $|Z|^{1/2}$ προσεγγίζει την

κατανομή $N(0.822, 0.349^2)$ και ότι αλλαγές στην κλίμακα του Z επηρεάζουν τη θέση του $|Z|^{1/2}$. Εάν απεικονίσουμε σε ένα κοινό διάγραμμα και τα τέσσερα παραπάνω συσσωρευμένα αθροίσματα, δημιουργείται ένα **ισχυρό συσσωρευμένο άθροισμα για τον έλεγχο της θέσης και της κλίμακας**. Η διεργασία τίθεται εκτός ελέγχου εάν κάποιο από τα τέσσερα αθροίσματα πάρει τιμή μεγαλύτερη από το διάστημα απόφασης H . Η επιλογή της τιμής του H μπορεί να γίνει για συγκεκριμένη τιμή αναφοράς K , ώστε να έχουμε το επιθυμητό ARL . Ο Lucas (1976) έδωσε πίνακες του ARL για διαγράμματα ελέγχου $CUSUM$ για διαφορετικές τιμές των H και K .

Ο Hawkins (1991), για να μπορέσουμε να ελέγξουμε ταυτόχρονα τα p μεμονωμένα διαγράμματα, πρότεινε τη χρήση των δύο ακόλουθων στατιστικών συναρτήσεων

$$MCZ = \max [L_{i,n}^+, -L_{i,n}^-]$$

και

$$ZNO = \sum_{i=1}^p (L_{i,n}^+ + L_{i,n}^-)^2,$$

όπου τα συσσωρευμένα αθροίσματα $L_{i,n}^+$ και $L_{i,n}^-$ υπολογίζονται από τους τύπους (3.4.5) και (3.4.6), αντίστοιχα. Η στατιστική συνάρτηση MCZ προέκυψε εφαρμόζοντας την $MCUSUM$ διαδικασία των Woodall and Ncube (1985) για το \mathbf{Z} , ενώ η συνάρτηση ZNO είναι το τετράγωνο της Ευκλείδειας νόρμας των διανυσμάτων που προκύπτουν από το συσσωρευμένο άθροισμα για ανοδικές και καθοδικές μετατοπίσεις του μέσου. Αντί για τα συσσωρευμένα αθροίσματα $L_{i,n}^+$ και $L_{i,n}^-$ θα μπορούσαμε να χρησιμοποιήσουμε τα αθροίσματα $S_{i,n}^+$ και $S_{i,n}^-$ (των τύπων (3.4.7) και (3.4.8)), με τα οποία ελέγχουμε μετατοπίσεις στην κλίμακα.

Σύμφωνα με τον Wierda (1994), η συνάρτηση MCZ υπερτερεί έναντι της ZNO . Την άποψή του αυτή βασίζει στο γεγονός ότι με τη χρήση της MCZ τα σήματα ερμηνεύονται άμεσα, ενώ δεν υπάρχει κάποιο ξεκάθαρο θεωρητικό υπόβαθρο-θεμέλιο για τη συνάρτηση ZNO . Η διεργασία τίθεται εκτός ελέγχου εάν το MCZ και το ZNO ξεπεράσουν κάποια σταθερά απόφασης H για προκαθορισμένη τιμή του ARL . Η τιμή αυτή του ARL για δοσμένο H δεν μπορεί να προκύψει εύκολα, αφού τα Z_i είναι συσχετισμένα. Το πρόβλημα αυτό αντιμετωπίζεται χρησιμοποιώντας προσομοίωση.

Οι Runger and Testik (2004) πρότειναν τέσσερις επεκτάσεις των διαγραμμάτων *MCUSUM*. Αρχικά, εφάρμοσαν το διάγραμμα των Woodall and Ncube (1985) για το τυποποιημένο διάνυσμα $\mathbf{Z}_t = \Sigma^{-1/2}(\mathbf{X}_t - \boldsymbol{\mu}_0)$ των αρχικών χαρακτηριστικών και $K = \lambda(\boldsymbol{\mu}_1)/2$, όπου $\lambda(\boldsymbol{\mu}_1) = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}$. Το πλεονέκτημα χρήσης του \mathbf{Z} είναι ότι αποτελεί ένα μετασχηματισμό από κύριες συνιστώσες που λαμβάνει υπόψη τη συσχέτιση μεταξύ των μεταβλητών. Η διεργασία τίθεται εκτός στατιστικού ελέγχου, εάν για $j = 1, 2, \dots, p$ συμβεί κάποιο από τα ακόλουθα

$$MC_t^+(j) = \max \left[MC_{t-1}(j) + \mathbf{Z}_t - \frac{\lambda(\boldsymbol{\mu}_1)}{2}, 0 \right] > H$$

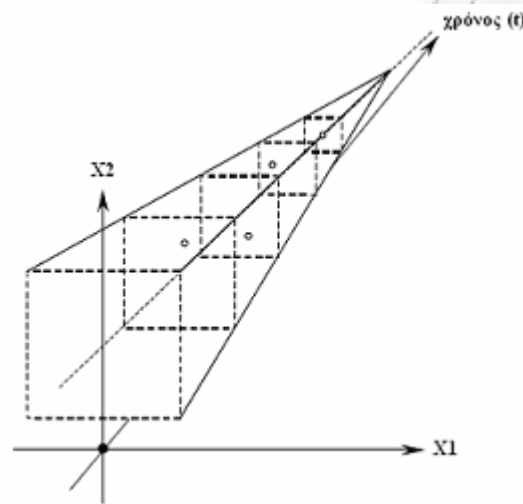
$$MC_t^-(j) = \max \left[MC_{t-1}(j) - \mathbf{Z}_t - \frac{\lambda(\boldsymbol{\mu}_1)}{2}, 0 \right] > H,$$

όπου $MC_0^+(j) = MC_0^-(j) = 0$. Σε κάθε ένα από τα παραπάνω δίπλευρα συσσωρευμένα αθροίσματα μπορεί να εφαρμοστεί η μέθοδος της *V*-μάσκας. Γεωμετρικά, αυτό ορίζει σχήμα πυραμίδας σε $(p+1)$ διαστάσεις, περιλαμβάνοντας τον άξονα του χρόνου. Οι Runger and Testik (2004) ονόμασαν αυτό το πολυμεταβλητό διάγραμμα *MPYRAMID* (*multivariate pyramid-shell chart*), λόγω της χαρακτηριστικής γεωμετρικής του ιδιότητας που απεικονίζεται στο Σχήμα 3.1. Οι Testik and Runger (2008) αναφέρουν ότι η επίδοση του *ARL* του διαγράμματος *MPYRAMID* εξαρτάται από την κατεύθυνση του διανύσματος της μετατόπισης και ως αποτέλεσμα μπορεί να μην είναι τόσο αποτελεσματικό στην ανίχνευση μετατοπίσεων σε διαφορετικές κατευθύνσεις από αυτές των αξόνων των κύριων συνιστωσών, όπως για παράδειγμα οι κατευθύνσεις των αξόνων των αρχικών μεταβλητών.

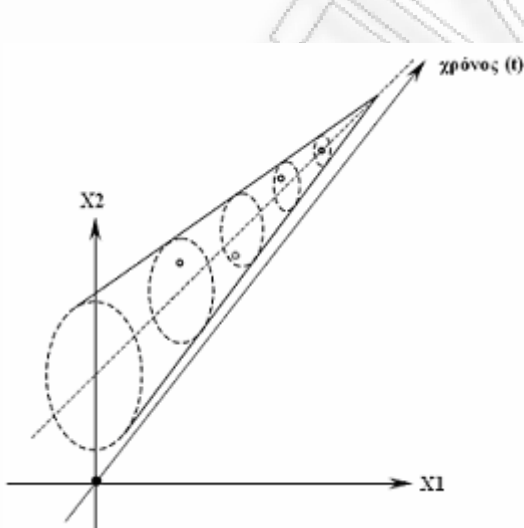
Στη συνέχεια, οι Runger and Testik (2004) υποστήριξαν ότι εάν το μέγεθος της παραμέτρου μη κεντρικότητας είναι γνωστό, αλλά όχι η κατεύθυνσή της, μπορεί να εφαρμοστεί η γενικευμένη μέθοδος πιθανοφάνειας. Στην περίπτωση αυτή η διεργασία είναι εκτός στατιστικού ελέγχου, εάν

$$\max_{1 \leq \tau \leq t} \left[\lambda(\boldsymbol{\mu}_1) \left\| \sum_{i=\tau}^t \mathbf{Z}_i \right\| - \frac{\lambda(\boldsymbol{\mu}_1)}{2}(t - \tau + 1) \right] > c,$$

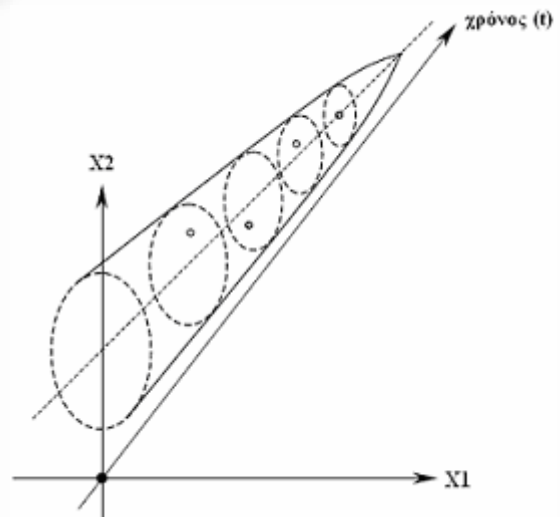
όπου $\mathbf{Z}_i = \Sigma^{-1/2}(\mathbf{X}_i - \boldsymbol{\mu}_0)$, τ είναι μία άγνωστη χρονική μετατόπιση ($1 \leq \tau \leq t$) και c είναι μια σταθερά. Ανάλογα της μεθόδου της V -μάσκας, ορίζεται σχήμα κώνου σε $(p+1)$ διαστάσεις, περιλαμβάνοντας τον άξονα του χρόνου. Το παραπάνω πολυμεταβλητό διάγραμμα ονομάζεται **MCONE** (*multivariate cone-shell chart*), λόγω της χαρακτηριστικής γεωμετρικής του ιδιότητας που απεικονίζεται στο Σχήμα.3.2



Σχήμα 3.1: Διάγραμμα *MPYRAMID* δύο διαστάσεων (Runger and Testik (2004))



Σχήμα 3.2: Διάγραμμα *MCONE* δύο διαστάσεων (Runger and Testik (2004))



Σχήμα 3.3: Διάγραμμα *MPARABOLOID* δύο διαστάσεων (Runger and Testik (2004))

Επιπλέον, οι Runger and Testik (2004) υποστήριξαν ότι και στην περίπτωση που ούτε η παράμετρος μη κεντρικότητας είναι γνωστή ούτε η κατεύθυνσή της, μπορεί να εφαρμοστεί η γενικευμένη μέθοδος πιθανοφάνειας. Η διεργασία θεωρείται ότι είναι εκτός στατιστικού ελέγχου, εάν

$$\max_{1 \leq \tau \leq t} \frac{1}{2(t-\tau+1)} \left(\sum_{i=\tau}^t (\mathbf{X}_i - \boldsymbol{\mu}_0)' \right) \boldsymbol{\Sigma}^{-1} \left(\sum_{i=\tau}^t (\mathbf{X}_i - \boldsymbol{\mu}_0) \right) = \max_{1 \leq \tau \leq t} \frac{1}{2(t-\tau+1)} \left\| \sum_{i=\tau}^t \boldsymbol{\Sigma}^{-1/2} (\mathbf{X}_i - \boldsymbol{\mu}_0)' \right\|^2 > c$$

ή ισοδύναμα εάν υπάρχει τ τέτοιο, ώστε να ισχύει η σχέση

$$\left\| \sum_{i=\tau}^t \boldsymbol{\Sigma}^{-1/2} (\mathbf{X}_i - \boldsymbol{\mu}_0)' \right\| > \sqrt{2(t-\tau+1)c}.$$

Η παραπάνω διαδικασία προκύπτει εύκολα από τη διαδικασία κατασκευής του διαγράμματος *MCONE*. Το νέο διάγραμμα ορίζεται από ένα παραβολοειδές σχήμα σε $(p+1)$ διαστάσεις, περιλαμβάνοντας τον άξονα του χρόνου. Το διάγραμμα αυτό ονομάζεται *MPARABOLOID* (*multivariate paraboloid-shell chart*), λόγω της χαρακτηριστικής γεωμετρικής του ιδιότητας που απεικονίζεται στο Σχήμα.3.3.

Τέλος, σύμφωνα με τους Runger and Testik (2004), εάν το μέγεθος της μετατόπισης του εκτός ελέγχου μέσου είναι γνωστή και μόνο η κατεύθυνση είναι άγνωστη, τότε το πρόβλημα είναι αμετάβλητο υπό ορθογώνιους μετασχηματισμούς. Η στατιστική συμπερασματολογία θα πρέπει να βασιστεί σε μία μέγιστη αναλλοίωτη στατιστική συνάρτηση, η οποία είναι $\|\bar{z}_i\|$ για συγκεκριμένη χρονική μετατόπιση τ . Εάν υποθέσουμε ότι η μετατόπιση του εκτός ελέγχου μέσου είναι γνωστή, έστω λ , τότε υπό την H_0 (H_1) η πυκνότητα της $\|\bar{z}_i\|$ ακολουθεί μία κεντρική (μη κεντρική) χ^2 -κατανομή με p β.ε. (και παράμετρο μη κεντρικότητας λ^2). Η συνάρτηση πιθανοφάνειας έχει την ακόλουθη μορφή

$$\ln \left\{ \left[\int \exp \left(-\frac{1}{2} \sum_{i=1}^{\tau-1} (\Gamma z_i)' (\Gamma z_i) \right) \exp \left(-\frac{1}{2} \sum_{i=\tau}^t (\Gamma z_i - \boldsymbol{\mu})' (\Gamma z_i - \boldsymbol{\mu}) \right) d\Gamma \right] \times \left[\int \exp \left(-\frac{1}{2} \sum_{i=\tau}^t (\Gamma z_i)' (\Gamma z_i) \right) d\Gamma \right]^{-1} \right\}$$

ή εναλλακτικά τη μορφή

$$\ln \left\{ \exp \left[-\frac{1}{2} (t-\tau+1) \lambda^2 \right] \int \exp \left[\lambda \left\| \sum_{i=1}^t z_i \right\| \varepsilon_1' \Gamma \varepsilon_1 \right] d\Gamma \right\},$$

όπου ε_i είναι το μοναδιαίο διάνυσμα κατά μήκος του πρώτου άξονα συντεταγμένων. Εάν θέσουμε $q = p/2 - 1$, τότε η παραπάνω έκφραση γράφεται ως εξής

$$\ln \left\{ (2\pi)^{p/2} \exp \left[-\frac{1}{2}(t-\tau+1)\lambda^2 \right] \left(I_q \left[\lambda \left\| \sum_{i=\tau}^t z_i \right\| \right] \right) \left(\left[\lambda \left\| \sum_{i=\tau}^t z_i \right\|^q \right]^{-1} \right) \right\}, \quad (3.4.9)$$

όπου $I_q(\cdot)$ υποδηλώνει την τροποποιημένη συνάρτηση Bessel πρώτου είδους. Συμβολίζοντας με $g^*(x)$ το φυσικό λογάριθμο του ολοκληρώματος του τύπου (3.4.9) ως συνάρτηση της ποσότητας $\lambda \left\| \sum_{i=\tau}^t z_i \right\|$, η διεργασία δίνει εκτός ελέγχου σήμα εάν υπάρχει τ τέτοιο ώστε

$$g^* \left(\lambda \left\| \sum_{i=\tau}^t \Sigma^{-1/2} (y_i - \mu_0) \right\| \right) - \frac{\lambda^2}{2}(t-\tau+1) > c.$$

Το παραπάνω διάγραμμα ονομάζεται **MINVARIANT** (*multivariate invariant-shell chart*). Θα πρέπει να αναφέρουμε ότι το ίδιο διάγραμμα προτάθηκε από τους Basseville and Nikiforov (1993). Επειδή $g^*(x) \approx \|x\|$, η προηγούμενη διαδικασία θεωρείται σχεδόν ίδια με τη διαδικασία που ακολουθείται για την κατασκευή του διαγράμματος **MCONE**.

Οι Ngai and Zhang (2001) πρότειναν δύο πολυμεταβλητά διαγράμματα **CUSUM** για την ανίχνευση μετατοπίσεων του μέσου της διεργασίας. Θεωρώντας ότι μ_0 είναι ο μέσος και Σ_0 ο πίνακας διακύμανσης-συνδιακύμανσης, όταν η διεργασία είναι εντός στατιστικού ελέγχου, όρισαν την τυποποιημένη ποσότητα

$$\bar{Y}_k = \sqrt{n} \Sigma_0^{-1} (\bar{X}_k - \mu_0),$$

όπου \bar{X}_k είναι ο δειγματικός μέσος του k δείγματος μεγέθους n . Τα \bar{Y}_k είναι ανεξάρτητα μεταξύ τους και για κάθε ένα από αυτά ισχύει $\bar{Y}_k \sim N(0, \mathbf{I})$. Όταν η διεργασία είναι εκτός ελέγχου, ο πίνακας $\mathbf{I}_{p \times p}$ παραμένει σταθερός και η ποσότητα \bar{Y}_k γίνεται $\sqrt{n} \Sigma_0^{-1} (\mu_1 - \mu_0)$, όπου μ_1 είναι ο εκτός ελέγχου μέσος της διεργασίας. Συνεπώς, η συσσωρευμένη αθροιστική στατιστική συνάρτηση μπορεί να έχει τη μορφή

$$MC_k^\alpha = \max \left[MC_{k-1}^\alpha + \alpha' \cdot \bar{Y}_k - K, 0 \right], \quad k = 1, 2, \dots, m,$$

όπου $MC_0^\alpha = 0$ και α είναι το διάνυσμα της κατεύθυνσης με Ευκλείδεια νόρμα $\|\alpha\| = 1$. Θα πρέπει να αναφέρουμε ότι η απόκλιση μεταξύ $\alpha' \Sigma_0^{-1/2} \mu_1$ και $\alpha' \Sigma_0^{-1/2} \mu_0$ λαμβάνει ως μέγιστη τιμή το α_0 , για το οποίο ισχύει $\Sigma_0^{-1/2} (\mu_1 - \mu_0) = \alpha_0 \|\Sigma_0^{-1/2} (\mu_1 - \mu_0)\|$. Οι Ngai and Zhang (2001) έδειξαν ότι το MC_k^α είναι σχεδόν ανάλογο της ποσότητας $|\alpha' \mu_1 - \alpha' \mu_0|$, όταν $|\alpha' \mu_1 - \alpha' \mu_0| > K$ και το k είναι αρκετά μεγάλο. Αυτό έχει ως αποτέλεσμα η εκτίμηση του α_0 να προκύπτει από την κατεύθυνση που δίνει η μέγιστη τιμή του MC_k^α . Η μέθοδος που χρησιμοποιήθηκε για την εκτίμηση του α_0 είναι η μέθοδος προβολών *Projection Pursuit* (*PP*). Άρα, το συσσωρευμένο άθροισμα $MC_k^{\hat{\alpha}_0}$ μπορεί να εκτιμηθεί από το άθροισμα

$$MC_k^{\hat{\alpha}_0} = \max_{\|\alpha\|=1} MC_k^\alpha. \quad (3.4.10)$$

Ένα πολυμεταβλητό διάγραμμα *PPCUSUM* (*Projection Pursuit CUSUM*) βασίζεται στην παραπάνω στατιστική συνάρτηση. Η διεργασία τίθεται εκτός ελέγχου εάν $MC_k^{\hat{\alpha}_0} > H$, όπου H είναι το όριο ελέγχου.

Όταν η διεργασία είναι εντός στατιστικού ελέγχου, η στατιστική συνάρτηση $MC_k^{\hat{\alpha}_0}$ που απεικονίζεται σε ένα διάγραμμα *PPCUSUM* παίρνει τιμές στο διάστημα $[0, H]$ και μακριά από το άνω όριο H όταν $1 \leq k \leq m$ και το m είναι μεγάλο. Οι Ngai and Zhang (2001) έδειξαν μέσω προσομοίωσης ότι η ποσότητα $H(1-\lambda)^{k-t_k+1}$ είναι πολύ μικρή για μία κατάλληλα επιλεγμένη σταθερά $0 < \lambda < 1$ και $t_k = \min [j : 1 \leq j \leq k, MC_{kj}^{\hat{\alpha}_0} = MC_j^{\hat{\alpha}_0}]$. Συνεπώς, εάν προσθέσουμε την ποσότητα $H(1-\lambda)^{k-t_k+1}$ σε κάθε $MC_k^{\hat{\alpha}_0}$ για $1 \leq k \leq m$, και η ακολουθία $\{MC_k^{\hat{\alpha}_0}\}$ που θα προκύψει θα έχει τιμές εντός του διαστήματος $[0, H]$. Με τον τρόπο αυτό θα έχουμε ένδειξη ότι κάποιο σημείο είναι πάνω από το όριο H πιο γρήγορα. Το νέο διάγραμμα ονομάζεται *FPCUSUM* και βασίζεται στη στατιστική συνάρτηση

$$MC_k^{\hat{\alpha}_0} + H(1-\lambda)^{k-t_k+1}.$$

Οι Chan and Zhang (2001) πρότειναν ένα πολυμεταβλητό διάγραμμα *CUSUM* για την παρακολούθηση του πίνακα διακύμανσης-συνδιακύμανσης. Βασιζόμενοι στη μέθοδο *Projection Pursuit* έκαναν δύο σημαντικές παρατηρήσεις. Η πρώτη ήταν ότι ο πίνακας

διακύμανσης-συνδιακύμανσης παραμένει εντός ελέγχου ($\Sigma = \Sigma_0$), εάν και μόνο εάν οι ποσότητες $\alpha'_{\max} \Sigma_0^{-1/2} \mathbf{X}$ και $\alpha'_{\min} \Sigma_0^{-1/2} \mathbf{X}$ έχουν κοινή διακύμανση, όπου α_{\max} και α_{\min} είναι τα ιδιοδιανύσματα που αντιστοιχούν στη μέγιστη και την ελάχιστη ιδιοτιμή του πίνακα $\Sigma_0^{-1/2} \Sigma \Sigma_0^{-1/2}$, αντίστοιχα, και α' ο ανάστροφος του α . Η δεύτερη παρατήρηση ήταν ότι τα α_{\max} και α_{\min} δίνουν την μέγιστη και ελάχιστη (προσημική) διαφορά (*signed difference*) μεταξύ της διακύμανσης του $\alpha' \Sigma_0^{-1/2} \mathbf{X}$ και του 1, αντίστοιχα. Για το λόγο αυτό, ένα πολυμεταβλητό διάγραμμα *CUSUM* που βασίζεται στη μέθοδο *Projection Pursuit* μπορεί να αναπτυχθεί επιλέγοντας ένα μονοδιάστατο διάγραμμα *CUSUM* για την παρακολούθηση της διακύμανσης και εφαρμόζοντας το στις ποσότητες $\hat{\alpha}'_{\max} \Sigma_0^{-1/2} \mathbf{X}$ και $\hat{\alpha}'_{\min} \Sigma_0^{-1/2} \mathbf{X}$, όπου $\hat{\alpha}'_{\max}$ και $\hat{\alpha}'_{\min}$ είναι εκτιμήσεις του α'_{\max} και του α'_{\min} , αντίστοιχα. Το μονοδιάστατο διάγραμμα που πρότειναν οι Chan και Zhang (2001), για την παρακολούθηση των ποσοτήτων $\hat{\alpha}'_{\max} \Sigma_0^{-1/2} \mathbf{X}$ και $\hat{\alpha}'_{\min} \Sigma_0^{-1/2} \mathbf{X}$, είναι αυτό που ανέπτυξαν οι Johnson and Leone (1962a, b, c).

Στην περίπτωση των μεμονωμένων παρατηρήσεων, υπέθεσαν ότι $\mu_0 = \mathbf{0}$ και $\Sigma_0 = \mathbf{I}_p$. Επίσης, όρισαν με λ_{kl}^{\max} και λ_{kl}^{\min} τις μέγιστες και ελάχιστες ιδιοτιμές του δειγματικού πίνακα $\mathbf{X}_l \mathbf{X}'_l + \mathbf{X}_{l+1} \mathbf{X}'_{l+1} + \dots + \mathbf{X}_k \mathbf{X}'_k$ για $1 \leq l \leq k$, αντίστοιχα. Επιπρόσθετα, θεώρησαν τις ποσότητες $MCU_{kl} = \lambda_{kl}^{\max} - (k-l+1)r_{upper}$ και $MCL_{kl} = \lambda_{kl}^{\min} - (k-l+1)r_{lower}$, όπου r_{upper} και r_{lower} είναι δύο τιμές αναφοράς (*reference values*). Συνεπώς, οι στατιστικές συναρτήσεις ενός πολυμεταβλητού διαγράμματος *CUSUM* με τη μέθοδο *Projection Pursuit* είναι οι εξής

$$\begin{cases} MCU_k = \max_{\|a\|=1} MCU_k^a = \max [0, MCU_{k1}, MCU_{k2}, \dots, MCU_{km}] \\ MCL_k = \min_{\|a\|=1} MCL_k^a = \min [0, MCL_{k1}, MCL_{k2}, \dots, MCL_{km}] \end{cases}, \quad k = 1, 2, \dots, m, \quad (3.4.11)$$

όπου $MCU_0 = MCL_0 = 0$. Η διεργασία τίθεται εκτός ελέγχου, όταν $MCU_k > H_{upper}$ ή όταν $MCL_k > H_{lower}$, όπου H_{upper} και H_{lower} είναι τιμές απόφασης (*decision values*).

Δε θα πρέπει να παραλείψουμε να αναφέρουμε τις βασικές υποθέσεις της παραπάνω μεθοδολογίας, οι οποίες είναι ότι ο πίνακας Σ_0 είτε είναι γνωστός είτε μπορεί να εκτιμηθεί από το $\bar{\mathbf{S}}$ και ότι ο μέσος της διεργασίας παραμένει σταθερός σε όλη τη διαδικασία.

3.4.2 Πολυμεταβλητό Διάγραμμα Ελέγχου για τον Εκθετικά Ζυγισμένο Μέσο

Οι Lowry, Woodall, Champ and Rigdon (1992) πρότειναν ένα πολυμεταβλητό διάγραμμα ελέγχου για τον εκθετικά ζυγισμένο μέσο (*Multivariate Exponentially Weighted Moving Average, MEWMA*) ως φυσική επέκταση του μονοδιάστατου *EWMA*. Θεωρούμε ότι έχουμε στη διάθεσή μας m ανεξάρτητα δείγματα μεγέθους n , έστω $\mathbf{X}_t, t=1, 2, \dots, m$. Το \mathbf{X}_t είναι ένα $(p \times 1)$ διάνυσμα με εντός ελέγχου μέσο $\boldsymbol{\mu} = \boldsymbol{\mu}_0 = \mathbf{0}$ και γνωστό πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$ που εξαρτάται μόνο από την παράμετρο μη κεντρικότητας

$$\lambda = (\boldsymbol{\mu}'_1 \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1)^{1/2},$$

όπου $\boldsymbol{\mu}_1$ ο μετατοπισμένος μέσος. Ένα διάγραμμα *MEWMA* βασίζεται στα διανύσματα

$$\mathbf{Z}_t = \mathbf{R} \cdot \mathbf{X}_t + (\mathbf{I} - \mathbf{R}) \cdot \mathbf{Z}_{t-1}, \quad t=1, 2, \dots, m,$$

όπου $\mathbf{Z}_0 = \boldsymbol{\mu}_0 = \mathbf{0}$, \mathbf{I} είναι ο ταυτοτικός πίνακας και $\mathbf{R} = \text{diag}(r_1, r_2, \dots, r_p)$ με $0 < r_j \leq 1$ και $j=1, 2, \dots, p$. Συνεπώς, ο παραπάνω τύπος μπορεί εναλλακτικά να γραφεί ως εξής

$$\mathbf{Z}_t = \begin{bmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_p \end{bmatrix} \mathbf{X}_t + \begin{bmatrix} 1-r_1 & 0 & \dots & 0 \\ 0 & 1-r_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1-r_p \end{bmatrix} \mathbf{Z}_{t-1}, \quad t=1, 2, \dots, m.$$

Εάν δεν υπάρχει λόγος κάθε παρελθοντική παρατήρηση να έχει διαφορετικό βάρος, τότε όλα τα διαγώνια στοιχεία του πίνακα \mathbf{R} θα είναι ίσα, δηλαδή θα ισχύει $r_1 = r_2 = \dots = r_p = r$. Στην περίπτωση αυτή, ο τύπος (3.4.11) παίρνει τη μορφή

$$\mathbf{Z}_t = r \cdot \mathbf{X}_t + (1-r) \cdot \mathbf{Z}_{t-1}, \quad t=1, 2, \dots, m.$$

Στην ειδική περίπτωση που $r=1$, παρατηρούμε ότι $\mathbf{Z}_t = \mathbf{X}_t$ για $t=1, 2, \dots, m$. Θα ήταν παράλειψη να μην αναφέρουμε ότι, όταν έχουμε δείγματα και όχι μεμονωμένες παρατηρήσεις, αντικαθιστούμε το \mathbf{X}_t των προηγούμενων σχέσεων με $\bar{\mathbf{X}}_t$ και επιπλέον θα ισχύει $\mathbf{Z}_0 = \bar{\mathbf{X}}_t$. Σε κάθε περίπτωση, η ποσότητα που απεικονίζεται στο διάγραμμα είναι

$$\mathbf{T}_t^2 = \mathbf{Z}'_t \boldsymbol{\Sigma}_t^{-1} \mathbf{Z}_t, \quad t=1, 2, \dots, m.$$

Η διεργασία είναι εκτός ελέγχου εάν $T_t^2 > h$, όπου η επιλογή του $h > 0$ γίνεται έτσι ώστε να έχουμε το επιθυμητό ARL_0 . Οι Lowry, Woodall, Champ and Rigdon (1992) δίνουν πίνακες, στους οποίους ορίζουν το h για διαφορετικές τιμές των p , r και λ για $ARL_0 \cong 200$. Με Σ_{Z_t} συμβολίζεται ο πίνακας συνδιακύμανσης των Z_t , ο οποίος υπολογίζεται από τον τύπο

$$\Sigma_{Z_t} = \frac{r}{2-r} [1 - (1-r)^{2t}] \Sigma.$$

Η ασυμπτωτική μορφή του πίνακα συνδιακύμανσης ($t \rightarrow \infty$) δίνεται από τον τύπο

$$\Sigma_{Z_t} = \frac{r}{2-r} \Sigma.$$

Οι Lowry, Woodall, Champ and Rigdon (1992) προτείνουν τιμές για τα r , p , h και λ , ώστε το μέσο μήκος ροής να είναι περίπου διακόσια ($ARL_0 \cong 200$). Επίσης, υποστηρίζουν ότι η τιμή του r εξαρτάται από το μέγεθος της μετατόπισης του μέσου και προτείνουν ότι μια λογική τιμή είναι $r = 0.10$. Επιπλέον, μέσω προσομοίωσης έδειξαν ότι το *MEWMA* διάγραμμα έχει μικρότερο ARL_0 σε σχέση με το διάγραμμα T^2 του Hotelling.

Ο Yumin (1996) υπογραμμίζει το γεγονός ότι η συσχέτιση μεταξύ των X_t μπορεί να περιοριστεί χρησιμοποιώντας ορθογώνιο μετασχηματισμό. Πιο συγκεκριμένα, υπάρχει ορθογώνιος πίνακας Q τέτοιος ώστε

$$Q' Cov(\mathbf{X}', \mathbf{X}) Q = \mathbf{R},$$

όπου ο $\mathbf{R} = diag(r_1, r_2, \dots, r_p)$ και $Cov(\mathbf{X}', \mathbf{X})$ ο πίνακας συσχέτισης του \mathbf{X} . Αν θέσουμε $\mathbf{Y} = \mathbf{Q}\mathbf{X}$ θα ισχύει $\mathbf{R} = Cov(\mathbf{Y}', \mathbf{Y})$ και όλες οι μεταβλητές Y_t , $t = 1, 2, \dots, m$, είναι ασυσχέτιστες μεταξύ τους. Αυτές συνήθως καλούνται κύριες συνιστώσες των αρχικών ποιοτικών χαρακτηριστικών. Το διάγραμμα του Yumin βασίζεται στα διανύσματα

$$\mathbf{Z}_t = \mathbf{R}^* \cdot \mathbf{Y}_t + (\mathbf{I} - \mathbf{R}^*) \cdot \mathbf{Z}_{t-1}, \quad t = 1, 2, \dots, m,$$

όπου $\mathbf{Z}_0 = \mathbf{Q} \cdot E(\mathbf{X})$, \mathbf{I} είναι ο ταυτοτικός πίνακας και $\mathbf{R}^* = diag(r_1^*, r_2^*, \dots, r_p^*)$ για $0 < r_j^* \leq 1$ και $j = 1, 2, \dots, p$. Η βέλτιστη επιλογή των r_j^* γίνεται είτε σύμφωνα του μονοδιάστατου διαγράμματος *EWMA* των Lucas and Saccucci (1990) είτε βάσει του κανόνα

του Woodall (Lowry, Woodall, Champ and Rigdon (1992)). Εναλλακτικά, ο παραπάνω τύπος γράφεται ως εξής

$$\mathbf{Z}_t = \begin{bmatrix} r_1^* & 0 & \cdots & 0 \\ 0 & r_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_p^* \end{bmatrix} \mathbf{Y}_t + \begin{bmatrix} 1-r_1^* & 0 & \cdots & 0 \\ 0 & 1-r_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-r_p^* \end{bmatrix} \mathbf{Z}_{t-1}, \quad t=1, 2, \dots, m.$$

Προφανώς, ο πίνακας των παραμέτρων εξομάλυνσης που αντιστοιχεί στο διάνυσμα \mathbf{X} είναι

$$\mathbf{B} = \mathbf{R}^* \mathbf{Q} = \begin{bmatrix} r_1^* & 0 & \cdots & 0 \\ 0 & r_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_p^* \end{bmatrix} \mathbf{Q}.$$

Ο Yimin έδειξε ότι η μέση τιμή του \mathbf{Z}_t και ο πίνακας συνδιακύμανσης $Cov(\mathbf{Z}'_t, \mathbf{Z}_t)$ υπολογίζονται αντίστοιχα μέσω των τύπων

$$\begin{aligned} E(\mathbf{Z}_t) &= \left\{ \left[\mathbf{I} + (\mathbf{I} - \mathbf{R}^*) + \dots + (\mathbf{I} - \mathbf{R}^*)^{t-1} \right] \mathbf{R}^* + (\mathbf{I} - \mathbf{R}^*)^t \right\} \mathbf{Q} E(\mathbf{X}) = \\ &= \left\{ \left[\mathbf{I} - (\mathbf{I} - \mathbf{R}^*)^t \right] \left[\mathbf{I} - (\mathbf{I} - \mathbf{R}^*) \right]^{-1} \mathbf{R}^* + (\mathbf{I} - \mathbf{R}^*)^t \right\} \mathbf{Q} E(\mathbf{X}) = \\ &= \mathbf{Q} E(\mathbf{X}) \end{aligned}$$

και

$$\begin{aligned} Cov(\mathbf{Z}'_t, \mathbf{Z}_t) &= (\mathbf{R}^* \mathbf{Q})(\mathbf{R}^* \mathbf{Q})' Cov(\mathbf{X}', \mathbf{X}) + \left((\mathbf{I} - \mathbf{R}^*) \mathbf{R}^* \mathbf{Q} \right) \left((\mathbf{I} - \mathbf{R}^*) \mathbf{R}^* \mathbf{Q} \right)' Cov(\mathbf{X}', \mathbf{X}) + \\ &\quad + \dots + \left((\mathbf{I} - \mathbf{R}^*)^{t-1} \mathbf{R}^* \mathbf{Q} \right) \left((\mathbf{I} - \mathbf{R}^*)^{t-1} \mathbf{R}^* \mathbf{Q} \right)' Cov(\mathbf{X}', \mathbf{X}) = \\ &= \left[\mathbf{I} + (\mathbf{I} - \mathbf{R}^*)^2 + (\mathbf{I} - \mathbf{R}^*)^4 + \dots + (\mathbf{I} - \mathbf{R}^*)^{2(t-1)} \right] \mathbf{R}^* \mathbf{R}^* Cov(\mathbf{X}', \mathbf{X}) = \\ &\cong (\mathbf{2I} - \mathbf{R}^*)^{-1} \mathbf{R}^* Cov(\mathbf{X}', \mathbf{X}). \end{aligned}$$

Αφού το \mathbf{X} ακολουθεί κανονική κατανομή με μέση τιμή $E(\mathbf{X})$ και πίνακα διακύμανσης-συνδιακύμανσης $Cov(\mathbf{X}', \mathbf{X})$, το \mathbf{Z}_t , το οποίο είναι γραμμικός συνδυασμός του \mathbf{X} , θα ακολουθεί κανονική κατανομή με μέση τιμή $E(\mathbf{X})$ και πίνακα διακύμανσης-συνδιακύμανσης $(\mathbf{2I} - \mathbf{R}^*)^{-1} \mathbf{R}^* Cov(\mathbf{X}', \mathbf{X})$, δηλαδή

$$\mathbf{Z}_t \sim N\left(E(\mathbf{X}), (2\mathbf{I} - \mathbf{R}^*)^{-1} \mathbf{R}^* \text{Cov}(\mathbf{X}', \mathbf{X})\right).$$

Τελικά, η στατιστική συνάρτηση που χρησιμοποιείται για την κατασκευή του διαγράμματος είναι

$$\chi_t^2 = (\mathbf{Z}_t - E(\mathbf{X}))' \left[(2\mathbf{I} - \mathbf{R}^*)^{-1} \mathbf{R}^* \text{Cov}(\mathbf{X}', \mathbf{X}) \right]^{-1} (\mathbf{Z}_t - E(\mathbf{X})),$$

η οποία ακολουθεί την κατανομή χ_p^2 . Το διάγραμμα που πρότεινε ο Yumin ονομάζεται **γενικευμένο πολυμεταβλητό διάγραμμα για τον εκθετικά ζυγισμένο μέσο** (*Generalized Multivariate Exponentially Weighted Moving Average or Generalized MEWMA*) και έχει τα ακόλουθα όρια ελέγχου

$$\begin{array}{l} \text{Γενικευμένο Διάγραμμα} \\ \text{MEWMA} \end{array} \begin{cases} CL = 0 \\ UCL = \chi_{p, \alpha}^2. \end{cases}$$

Εάν υπάρχει \mathbf{Z}_t για το οποίο να ισχύει $\chi_t^2 > \chi_{\alpha, p}^2$, η διεργασία τίθεται εκτός ελέγχου.

Οι Yeh, Lin, Zhou and Venkataramani (2003) πρότειναν ένα πολυμεταβλητό διάγραμμα ελέγχου *EWMA* για την ανίχνευση μικρών αλλαγών της μεταβλητότητας, το οποίο ονόμασαν **διάγραμμα *EWMA V*** και ένα πολυμεταβλητό διάγραμμα *EWMA* για την παρακολούθηση του μέσου της διεργασίας, το οποίο ονόμασαν **διάγραμμα *EWMA M***.

Σύμφωνα με τους Yeh, Lin, Zhou and Venkataramani, θεωρούμε ότι έχουμε στη διάθεσή μας m ανεξάρτητα δείγματα (μεγέθους n το καθένα) από μία διεργασία, έστω $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$, όπου $\mathbf{X}_k = (\mathbf{X}_{1k}, \mathbf{X}_{2k}, \dots, \mathbf{X}_{nk})$, $k = 1, 2, \dots, m$. Όταν η διεργασία είναι εντός στατιστικού ελέγχου, οι μεταβλητές $\mathbf{X}_{ik} = (X_{i1k}, X_{i2k}, \dots, X_{ipk})'$, για $i = 1, 2, \dots, n$ και $k = 1, 2, \dots, m$, προέρχονται από μία p -διάστατη κανονική κατανομή με μέσο διάνυσμα $\boldsymbol{\mu}_0$ και πίνακα διακύμανσης συνδιακύμανσης $\boldsymbol{\Sigma}_0$, συμβολικά $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. Επειδή τα $\boldsymbol{\mu}_0$ και $\boldsymbol{\Sigma}_0$ είναι άγνωστα, εκτιμώνται από τα m δείγματα. Επίσης, θεωρούμε τις ποσότητες

$$\bar{\bar{\mathbf{X}}} = \sum_{k=1}^m \frac{\bar{\mathbf{X}}_k}{m} \quad \text{και} \quad \bar{\mathbf{S}} = \sum_{k=1}^m \frac{\mathbf{S}_k}{m},$$

όπου

$$\bar{\mathbf{X}}_k = \sum_{i=1}^n \frac{X_{ik}}{n} \quad \text{και} \quad \mathbf{S}_k = \sum_{i=1}^n \frac{(X_{ik} - \bar{\mathbf{X}}_k)(X_{ik} - \bar{\mathbf{X}}_k)'}{n}$$

είναι το δειγματικό μέσο διάνυσμα και ο δειγματικός πίνακας διακύμανσης-συνδιακύμανσης του k δείγματος, για $k = 1, 2, \dots, m$, αντίστοιχα.

Για $t \geq 1$, όρισαν την επόμενη πιθανότητα v_t

$$v_t = P \left(\prod_{j=1}^p F_{n-j, N-i+1-j} \leq \left(\prod_{j=1}^p \frac{N-m+1-j}{n-j} \right) \times \frac{|n\mathbf{S}_t|}{|N\bar{\mathbf{S}}|} \right),$$

όπου $|n\mathbf{S}_t|$ και $|N\bar{\mathbf{S}}|$ είναι οι ορίζουσες των πινάκων $n\mathbf{S}_t$ και $N\bar{\mathbf{S}}$, αντίστοιχα, $N = n \cdot m$ και $F_{n-j, N-i+1-j}$ είναι η κατανομή F με $n-j$ και $N-i+1-j$ βαθμούς ελευθερίας. Οι Yeh, Lin, Zhou and Venkataramani (2003) απέδειξαν ότι όταν η διεργασία είναι εντός ελέγχου, το γινόμενο των τυχαίων μεταβλητών που υπάρχει στην πιθανότητα v_t κατανέμεται ομοιόμορφα στο διάστημα $(0, 1)$. Επομένως, ένα πολυμεταβλητό διάγραμμα ελέγχου $EWMA$ μπορεί να βασιστεί στην εξής συνάρτηση

$$S_v(t) = \omega \times (v_t - 0.5) + (1 - \omega) \times S_v(t-1), \quad t \geq 1,$$

όπου $0 < \omega < 1$ και $S_v(0) = 0$. Για $t \geq 1$ και δοσμένο ω ισχύει

$$E(S_v(t)) = 0 \quad \text{και} \quad V(S_v(t)) = \frac{1}{12} \left(\frac{\omega}{2 - \omega} \right) \left(1 - (1 - \omega)^{2t} \right).$$

Καθώς το $S_v(t)$ είναι συμμετρικό γύρω από το μηδέν, τα όρια ελέγχου που προκύπτουν για το πολυμεταβλητό διάγραμμα ελέγχου $EWMA V$ (Multivariate $EWMA V$ -chart) για την ανίχνευση μικρών αλλαγών της μεταβλητότητας δίνονται παρακάτω

$$\text{Πολυμεταβλητό} \begin{cases} LCL = -L \times \sqrt{\frac{1}{12} \left(\frac{\omega}{2 - \omega} \right) \left(1 - (1 - \omega)^{2t} \right)} \\ \text{Διάγραμμα} \\ \text{EWMA V} \quad \begin{cases} CL = 0 \\ UCL = +L \times \sqrt{\frac{1}{12} \left(\frac{\omega}{2 - \omega} \right) \left(1 - (1 - \omega)^{2t} \right)}, \end{cases} \end{cases}$$

όπου $t = 1, 2, \dots$ και L επιλέγεται ώστε να ελέγχεται το εντός ελέγχου μέσο μήκος ροής. Για το παραπάνω διάγραμμα, οι συγγραφείς προτείνουν $m \geq 50$, $n \geq 10$ και $3 \leq p \leq 8$.

Οι Yeh, Lin, Zhou and Venkataramani (2003) βασιζόμενοι στο γεγονός ότι το $S_v(t)$ μπορεί να χρησιμοποιηθεί για την ανίχνευση αλλαγών της μεταβλητότητας της διεργασίας, επέκτειναν την ιδέα αυτή στην περίπτωση που **θέλουμε να ανιχνεύσουμε μικρές μετατοπίσεις του μέσου της διεργασίας**. Για $t \geq 1$, όρισαν την πιθανότητα u_t , η οποία υπολογίζεται από τον τύπο που ακολουθεί

$$u_t = P \left(\prod_{j=1}^p F_{p, N-m-p+1} \leq \frac{N-m-p+1}{p(m+1)} (\bar{\mathbf{X}}_t - \bar{\mathbf{X}})' \bar{\mathbf{S}}^{-1} (\bar{\mathbf{X}}_t - \bar{\mathbf{X}}) \right).$$

Εάν η διεργασία είναι εντός ελέγχου, η πιθανότητα u_t κατανέμεται ομοιόμορφα στο διάστημα $(0, 1)$. Συνεπώς, ένα **πολυμεταβλητό διάγραμμα ελέγχου EWMA M** (*Multivariate EWMA M-chart*) για την ανίχνευση μικρών μετατοπίσεων του μέσου της διεργασίας μπορεί να βασιστεί στην εξής συνάρτηση

$$S_m(t) = \omega \times (u_t - 0.5) + (1 - \omega) \times S_m(t-1), \quad t \geq 1,$$

όπου $0 < \omega < 1$ και $S_m(0) = 0$. Για το παραπάνω διάγραμμα, οι συγγραφείς προτείνουν το $N = n \cdot m$ να είναι μεταξύ 200 και 250 για $n \leq 10$. Επίσης, για την αύξηση της ευαισθησίας του διαγράμματος αυτού στην ανίχνευση τυχόν αύξησης (ή μείωσης) της ορίζουσας του πίνακα διακύμανσης-συνδιακύμανσης συνιστούν η ποσότητα $S_m(0)$ να πάρει μία μικρή αρχική θετική (ή αρνητική) τιμή, την οποία δεν καθορίζουν.

Παρατηρούμε ότι όταν η διεργασία είναι εντός στατιστικού ελέγχου οι συναρτήσεις $S_v(t)$ και $S_m(t)$ έχουν την ίδια κατανομή και ως αποτέλεσμα τα ίδια όρια ελέγχου. Επομένως, σε ένα κοινό γράφημα μπορούν να απεικονιστούν οι συναρτήσεις $S_v(t)$ και $S_m(t)$ με όρια ελέγχου που υπολογίζονται από τον τύπο (3.4.25). Προφανώς, τα πολυμεταβλητά διαγράμματα EWMA V και M έχουν κοινό εντός και εκτός ελέγχου μέσο μήκος ροής. Σύμφωνα με τους Yeh, Lin, Zhou and Venkataramani (2003), το διάγραμμα EWMA V υπερέχει σε σχέση με το διάγραμμα $|\mathbf{S}|$ σχεδόν πάντα. Στην περίπτωση, όμως, που το $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_0|$ είναι μεγάλο, τα δύο διαγράμματα έχουν σχεδόν ίδια ARL. Επιπλέον, το διάγραμμα EWMA

M υπερέρχει σε σχέση με το $MEWMA$ των Lowry, Woodall, Champ and Rigdon (1992), εκτός από την περίπτωση που παρατηρείται μικρή αύξηση στην τυπική απόκλιση της πρώτης μεταβλητής και στην περίπτωση που η συσχέτιση είναι μέτρια και είτε οι τυπικές αποκλίσεις αυξάνονται κατά ένα μικρό ποσοστό είτε αυξάνεται μόνο η τυπική απόκλιση της πρώτης μεταβλητής.

Οι Chen, Cheng and Xie (2005) προτείνουν ένα διάγραμμα ελέγχου $MEWMA$ για την ταυτόχρονη παρακολούθηση της μέσης τιμής και του πίνακα διακύμανσης-συνδιακύμανσης της διεργασίας. Θεωρούν ότι η διεργασία αποτελείται από m ανεξάρτητα δείγματα μεγέθους n το καθένα, έστω $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$, όπου $\mathbf{X}_k = (\mathbf{X}_{1k}, \mathbf{X}_{2k}, \dots, \mathbf{X}_{nk})$, $k=1, 2, \dots, m$ και $\mathbf{X}_k \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Επίσης, θεωρούμε ότι $\boldsymbol{\mu}_0$ είναι το επιθυμητό μέσο διάνυσμα, $\boldsymbol{\Sigma}_0$ ο επιθυμητός πίνακας διακύμανση-συνδιακύμανσης και τα τυχαία διανύσματα \mathbf{X}_{ik} είναι ανεξάρτητα μεταξύ τους.

Αρχικά, για την παρακολούθηση του μέσου διανύσματος της διεργασίας θεωρούν το δειγματικό μέσο που υπολογίζεται από τον τύπο

$$\bar{\mathbf{X}}_k = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_{ik}.$$

Ορίζουν την ποσότητα

$$\mathbf{Z}_k = (1-\lambda) \cdot \mathbf{Z}_{k-1} + \lambda \cdot (\bar{\mathbf{X}}_k - \boldsymbol{\mu}_0),$$

όπου \mathbf{Z}_0 είναι το σημείο εκκίνησης και λ η παράμετρος εξομάλυνσης για την οποία ισχύει $0 < \lambda \leq 1$. Όταν $\mathbf{Z}_0 = \boldsymbol{\mu} - \boldsymbol{\mu}_0$ ισχύουν τα εξής

$$E(\mathbf{Z}_k) = \boldsymbol{\mu} - \boldsymbol{\mu}_0 \quad \text{και} \quad \text{Cov}(\mathbf{Z}_k) = \frac{\lambda [1 - (1-\lambda)^{2k}]}{n(2-\lambda)} \boldsymbol{\Sigma}.$$

Για το λόγο αυτό έχουμε ότι

$$\mathbf{Z}_k \sim N_p(\boldsymbol{\mu} - \boldsymbol{\mu}_0, \text{Cov}(\mathbf{Z}_k)) \quad \text{και} \quad T_k = \frac{n(2-\lambda)}{\lambda [1 - (1-\lambda)^{2k}]} \mathbf{Z}_k' \boldsymbol{\Sigma}^{-1} \mathbf{Z}_k \sim \chi_{p, \delta^2}^2,$$

όπου $\text{Con}(\cdot)$ είναι η συνάρτηση συνδιακύμανσης της πολυμεταβλητής κατανομής, Σ^{-1} είναι ο αντίστροφος του πίνακα Σ και p είναι οι β.ε. της χ^2 -κατανομής με παράμετρο μη κεντρικότητας

$$\delta^2 = \left\{ n(2-\lambda) / \lambda \left[1 - (1-\lambda)^{2k} \right] \right\} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)' \Sigma^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0).$$

Συνεπώς, η στατιστική συνάρτηση για την παρακολούθηση του μέσου της διεργασίας είναι

$$\mathbf{U}_k = \Phi^{-1} \left[H_p \left\{ \frac{n(2-\lambda)}{\lambda \left[1 - (1-\lambda)^{2k} \right]} \mathbf{Z}'_k \Sigma_0^{-1} \mathbf{Z}_k \right\} \right],$$

όπου $H_p(\cdot)$ είναι η συνάρτηση κατανομής της χ^2_p , $\Phi(\cdot)$ είναι η συνάρτηση της τυπικής κανονικής κατανομής και $\Phi^{-1}(\cdot)$ είναι η αντίστροφη της $\Phi(\cdot)$.

Στη συνέχεια, για την παρακολούθηση της μεταβλητότητας της διεργασίας ορίζουν

$$\mathbf{W}_k = \sum_{i=1}^n (\mathbf{X}_{ik} - \bar{\mathbf{X}}_k)' \Sigma_0^{-1} (\mathbf{X}_{ik} - \bar{\mathbf{X}}_k),$$

έτσι ώστε $\mathbf{W}_k \sim \chi^2_{p(n-1)}$, όταν $\Sigma = \Sigma_0$. Επιπλέον, ορίζουν την ποσότητα

$$\mathbf{Y}_k = (1-\lambda) \cdot \mathbf{Y}_{k-1} + \lambda \cdot \Phi^{-1} \left\{ H_{p(n-1)}(\mathbf{W}_k) \right\},$$

όπου \mathbf{Y}_0 είναι το σημείο εκκίνησης και $0 < \lambda \leq 1$. Όταν η μεταβλητότητα της διεργασίας είναι εντός ελέγχου ($\Sigma = \Sigma_0$) και $\mathbf{Y}_0 = 0$ ισχύουν τα εξής

$$E(\mathbf{Y}_k) = 0 \quad \text{και} \quad V(\mathbf{Y}_k) = \frac{\lambda \left[1 - (1-\lambda)^{2k} \right]}{2-\lambda}.$$

Η στατιστική συνάρτηση για την παρακολούθηση της μεταβλητότητας της διεργασίας είναι

$$\mathbf{V}_k = \sqrt{\frac{2-\lambda}{\lambda \left[1 - (1-\lambda)^{2k} \right]}} \cdot \mathbf{Y}_k.$$

Τα \mathbf{U}_k και \mathbf{V}_k είναι ανεξάρτητα. Όταν $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\Sigma = \Sigma_0$, $\mathbf{Z}_0 = 0$ και $\mathbf{Y}_0 = 0$, οι συναρτήσεις \mathbf{U}_k και \mathbf{V}_k ακολουθούν την τυπική κανονική κατανομή. Συνεπώς, από το

συνδυασμό των συναρτήσεων \mathbf{U}_k και \mathbf{V}_k προκύπτει η στατιστική συνάρτηση του νέου διαγράμματος, η οποία υπολογίζεται ως εξής

$$\mathbf{M}_k = \max[|\mathbf{U}_k|, |\mathbf{V}_k|].$$

Επειδή η \mathbf{M}_k είναι η μέγιστη τιμή των $|\mathbf{U}_k|$ και $|\mathbf{V}_k|$, το διάγραμμα (στο οποίο απεικονίζεται η \mathbf{M}_k) ονομάζεται *Max-MEWMA*. Εάν η συνάρτηση \mathbf{M}_k πάρει μεγάλες τιμές, τότε ή το μέσο διάνυσμα της διεργασίας έχει μετατοπιστεί μακριά από το $\boldsymbol{\mu}_0$ ή η μεταβλητότητα της διεργασίας έχει αλλάξει ή έχουν συμβεί και τα δύο ενδεχόμενα. Αντίθετα, εάν η συνάρτηση \mathbf{M}_k πάρει μικρές τιμές, τότε τόσο το μέσο διάνυσμα όσο και η μεταβλητότητα της διεργασίας έχουν παραμείνει κοντά στις αρχικές προκαθορισμένες τιμές τους.

Η εντός ελέγχου συνάρτηση πυκνότητας πιθανότητας της \mathbf{M}_k είναι

$$f(y; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = 4\Phi(y)[2\Phi(y)-1]^2,$$

όπου $\Phi(\cdot)$ είναι η συνάρτηση της τυπικής κανονικής κατανομής. Επειδή η συνάρτηση \mathbf{M}_k δεν μπορεί να πάρει αρνητικές τιμές, το διάγραμμα *Max-MEWMA* θα έχει μόνο άνω όριο ελέγχου, το οποίο υπολογίζεται από τον τύπο που ακολουθεί

$$\text{Διάγραμμα Max - MEWMA} \quad \left\{ UCL = E(\mathbf{M}_k) + L\sqrt{V(\mathbf{M}_k)} \right\}.$$

Το L αποτελεί έναν πολλαπλασιαστή που σε συνδυασμό με την παράμετρο λ , ελέγχουν την επίδοση του διαγράμματος. Επίσης, η μέση τιμή $E(\mathbf{M}_k)$ και η διακύμανση $V(\mathbf{M}_k)$ λαμβάνονται αριθμητικά. Οι Chen, Cheng and Xie (2005) δίνουν πίνακες στους οποίους προτείνουν τιμές για το ζεύγος (λ, L) . Τέλος, υποστηρίζουν ότι το διάγραμμα *Max-MEWMA* έχει καλύτερη απόδοση στην ανίχνευση μέτριων αλλαγών της διεργασίας σε σχέση με ένα συνδυασμό του διαγράμματος χ^2 (για την παρακολούθηση του μέσου της διεργασίας) και του διαγράμματος $|\mathbf{S}|$ (για την παρακολούθηση της μεταβλητότητας της διεργασίας).

Στη βιβλιογραφία έχουν προταθεί εναλλακτικές χρήσεις των διαγραμμάτων *MEWMA* από διαφορετικούς ερευνητές. Για παράδειγμα, η Jones (1966) παρουσίασε μία μέθοδο για την εκτίμηση της βέλτιστης μήτρας για εκθετική εξομάλυνση και πρόβλεψη πολυδιάστατων χρονικών σειρών, ενώ οι Kramer and Schmid (1997) πρότειναν μία γενίκευση του

διαγράμματος *MEWMA* των Lowry, Woodall, Champ and Rigdon (1992) για πολυμεταβλητές χρονικές σειρές. Επίσης, οι Enns, Machak, Spivey and Wroblewski (1982) πρότειναν ένα πολλαπλά μοντέλα εκθετικής εξομάλυνσης χρήσιμα σε βιομηχανικά συστήματα πρόβλεψης και ο Harvey (1986) παρουσίασε μία πολυμεταβλητή γενίκευση ανάλογη του πολυμεταβλητού εκθετικού μοντέλου εξομάλυνσης.

Οι Sullivan and Woodall (1998) πρότειναν τη χρήση ενός *MEWMA* για την προκαταρκτική ανάλυση πολυμεταβλητών παρατηρήσεων. Επιπλέον, οι Runger, Keats, Montgomery and Scranton (1999) έδειξαν ότι η ικανότητα ανίχνευσης μιας μετατόπισης ενός διαγράμματος *MEWMA* βελτιώνεται σημαντικά τροποποιώντας τις αρχικές μεταβλητές της διεργασίας σε ένα υπόχωρο χαμηλότερης διάστασης μέσω της χρήσης του μετασχηματισμού U , ο οποίος είναι όμοιος με το μετασχηματισμό των κύριων συνιστωσών. Συμπληρωματικά, ο Fassò (1999) ανέπτυξε ένα μονόπλευρο διάγραμμα *MEWMA*, το οποίο βασίζεται σε έναν περιορισμένο εκτιμητή μέγιστης πιθανοφάνειας.

Το 2002 οι Tseng, Chou and Lee πρότειναν έναν ελεγκτή πολυμεταβλητού εκθετικά ζυγισμένου κινούμενου μέσου (*MEWMA controller*) υπό ένα γραμμικό μοντέλο πολλαπλής εισόδου-πολλαπλής εξόδου (*linear multiple-input-multiple-output*), ενώ την ίδια χρονιά οι Del Castigio and Rajagopal πρότειναν μία επέκταση πολλαπλής εισόδου-πολλαπλής εξόδου του μονοδιάστατου διπλού *EWMA* (*univariate double EWMA, dEWMA*), το οποίο είχε αρχικά προταθεί από τους Butler and Stefani (1994). Επιπρόσθετα, το 2004 οι Yeh, Huwang and Wu πρότειναν ένα διάγραμμα για τον εκθετικά ζυγισμένο κινούμενο μέσο του λόγου πιθανοφάνειας (*likelihood-ratio-based EWMA*), το οποίο είναι αρκετά αποτελεσματικό στην ανίχνευση μικρών αλλαγών της μεταβλητότητας πολυδιάστατων κανονικών διεργασιών, ενώ οι Testik and Borror (2004) έκαναν μία ανασκόπηση και σύγκριση των τριών μεθόδων σχεδιασμού ενός *MEWMA*. Οι τεχνικές αυτές είναι ο στατιστικός σχεδιασμός (*statistical design*), ο οικονομικός-στατιστικός σχεδιασμός (*economic-statistical design*) και ο ισχυρός σχεδιασμός (*robust design*).

Θα ήταν παράλειψη να μην αναφέρουμε ότι έχουν προταθεί τρεις διαφορετικές μέθοδοι υπολογισμού του μέσου μήκους ροής ενός διαγράμματος *MEWMA*. Το *ARL* μπορεί να προσεγγιστεί μέσω προσομοίωσης ή μέσω ολοκληρωτικής εξίσωσης ή μέσω Μαρκοβιανής Αλυσίδας. Αρχικά, οι Lowry, Woodall, Champ and Rigdon (1992) χρησιμοποίησαν προσομοίωση για τον προσδιορισμό του *ARL* και στη συνέχεια οι Linderman and Love

(2000a, 2000b) υπολόγισαν το εντός και εκτός ελέγχου ARL ενός $MEWMA$ μέσω προσομοίωσης ως μέρος οικονομικού μοντέλου. Επιπρόσθετα, ο Rigdon πρότεινε μία εξίσωση διπλού ολοκληρώματος (1995a) και μία μονού ολοκληρώματος (1995b) για τον υπολογισμό του ARL , ενώ οι Runger and Prabhu (1996) προσδιόρισαν το μέσο μήκος ροής ενός $MEWMA$ μέσω μιας προσέγγισης Μαρκοβιανής Αλυσίδας δύο διαστάσεων. Στη συνέχεια, οι Molnau, Runger, Montgomery, Skinner, Loreda and Prabhu (2001) παρουσίασαν ένα πρόγραμμα, το οποίο επιτρέπει τον υπολογισμό του ARL ενός $MEWMA$, όταν είναι γνωστές οι μετατοπίσεις στο μέσο διάνυσμα, τα όρια ελέγχου και οι παράμετροι εξομάλυνσης.

3.4.3 Πολυμεταβλητό Διάγραμμα Ελέγχου Κινούμενου Μέσου

Ο Margavio (1990) πρότεινε μία γενική πολυμεταβλητή στατιστική συνάρτηση για τον κινούμενο μέσο, η οποία υπολογίζεται ως εξής

$$\mathbf{S}_t = \sum_{i=1}^t \alpha_i (\bar{\mathbf{X}}_i - \boldsymbol{\mu}_0), \quad (3.4.12)$$

όπου α_i είναι το βάρος (σταθερά) που αντιστοιχεί στο i δειγματικό μέσο διάνυσμα. Τα βάρη για κάθε μεταβλητή δεν πρέπει να είναι ίσα. Οι Lowry, Woodall, Champ and Rigdon (1992) προτείνουν τη χρήση της μεθόδου του Hawkins (1991).

Οι Margavio and Conerly (1995) θεωρούν ότι $\bar{\mathbf{X}}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$, όπου $i=1, 2, 3, \dots$ είναι οι χρονικές περίοδοι και n είναι ο αριθμός των υποομάδων ή το δειγματικό μέγεθος. Μία τετραγωνική μορφή που έχει προταθεί για την ανίχνευση μετατοπίσεων στο μέσο διάνυσμα είναι η

$$D_t = \mathbf{S}_t' [\text{Var}(\mathbf{S}_t)]^{-1} \mathbf{S}_t,$$

όπου το \mathbf{S}_t υπολογίζεται από τον τύπο (3.4.12). Εάν η διεργασία είναι εντός ελέγχου και τα \mathbf{S}_t είναι ανεξάρτητα, τότε $D_t \sim \chi_p^2$. Ωστόσο, η ακριβής κατανομή του D_t είναι πιο σύνθετη, αφού τα \mathbf{S}_t είναι εξαρτημένα και η D_t μπορεί να παρατηρηθεί μόνο εάν η D_{t-1} βρίσκεται εντός των ορίων ελέγχου.

Οι Margavio and Conerly (1995) πρότειναν δύο διαγράμματα κινούμενου μέσου, σε καθένα από τα οποία η επιλογή της συνάρτησης βάρους S_t είναι διαφορετική. Το πρώτο διάγραμμα που πρότειναν είναι ένα απλό **αριθμητικό πολυμεταβλητό διάγραμμα κινούμενου μέσου** (*Arithmetic Multivariate Moving Average, AMMA*). Το διάγραμμα αυτό βασίζεται στον αριθμητικό μέσο των k τελευταίων παρατηρήσεων. Η στατιστική συνάρτηση έχει τη μορφή

$$S_t = \frac{1}{\tau} \sum_{i=t-\tau+1}^t (\bar{X}_i - \mu_0),$$

όπου το τ υπολογίζεται ως εξής

$$\tau = \begin{cases} k, & t \geq k \\ t, & t < k \end{cases}.$$

Το δεύτερο διάγραμμα που πρότειναν ονομάζεται **CMEWMA** και θεωρείται μια περικεκομμένη εκδοχή (*truncated version*) του διαγράμματος *MEWMA* των Lowry, Woodall, Champ and Rigdon (1992). Όπως γνωρίζουμε, το *MEWMA* λαμβάνει πληροφορίες και από παρελθοντικά δείγματα και από τα πιο πρόσφατα. Εντούτοις, οι πληροφορίες που συλλέγονται από τα πρώτα δείγματα μπορεί να μην είναι σχετικές ή να μην έχουν νόημα. Περικόπτοντας το *MEWMA* μετά από ένα αρκετά σύντομο χρονικό διάστημα, έστω μετά από k χρονικές περιόδους, η στατιστική συνάρτηση *CEWMA* δίνει περισσότερο βάρος στους μέσους των πιο πρόσφατων υποομάδων. Με τον τρόπο αυτό, όταν ο μέσος της διεργασίας μετατοπίζεται, αυτή η στατιστική συνάρτηση αντιδρά πιο γρήγορα. Η προαναφερθείσα στατιστική συνάρτηση υπολογίζεται από τον τύπο που ακολουθεί

$$S_t = \sum_{i=t-k+1}^t \frac{r(1-r)^{t-i}}{1-(1-r)^k} (\bar{X}_i - \mu_0), \quad t > k.$$

Εάν $t \leq k$, τότε η παραπάνω στατιστική συνάρτηση συμπίπτει με αυτή ενός διαγράμματος *MEWMA*, δηλαδή θα έχει τη μορφή

$$S_t = r(\bar{X}_t - \mu_0) + (1-r)S_{t-1} = \sum_{i=1}^t r(1-r)^{t-i} (\bar{X}_i - \mu_0), \quad t = 1, 2, 3, \dots,$$

όπου $S_0 = \mathbf{0}_{(p \times 1)}$. Θα πρέπει να αναφέρουμε, επίσης, ότι η επιλογή τόσο του k όσο και του r επηρεάζει την απόδοση του *ARL* του παραπάνω διαγράμματος.

Οι Margavio and Conerly (1995) υποστηρίζουν ότι όταν το r είναι μεγάλο, οι τιμές του ARL των διαγραμμάτων $MEWMA$ και $CMEWMA$ είναι σχεδόν ίδιες. Επιπλέον, όταν τοποθετούνται μικρά βάρη στις πιο πρόσφατες παρατηρήσεις, το $CMEWMA$ ανιχνεύει πολύ μικρές μετατοπίσεις του μέσου γρηγορότερα από ένα διάγραμμα $MEWMA$ ή ένα διάγραμμα $AMMA$. Ενώ, αντίθετα όταν τοποθετούνται μεγάλες τιμές στις πιο πρόσφατες παρατηρήσεις, το $CMEWMA$ και το $MEWMA$ είναι στην ουσία ισοδύναμα και ταυτόχρονα είναι και τα δύο πιο γρήγορα στην ανίχνευση μικρών και μεσαίων μετατοπίσεων του μέσου της διεργασίας. Επίσης, οι συγγραφείς αναφέρουν ότι το διάγραμμα $AMMA$ ανιχνεύει γρηγορότερα μεγάλες μετατοπίσεις του μέσου.

Τέλος, τονίζουν ότι η επιλογή του διαγράμματος που θα χρησιμοποιήσουμε εξαρτάται κάθε φορά από το σενάριο που περιγράφει καλύτερα την κατάσταση της διεργασίας. Για παράδειγμα, όταν η διεργασία είναι σταθερή, τα διαγράμματα $CMEWMA$ και $MEWMA$ ανιχνεύουν όλες τις μετατοπίσεις πιο γρήγορα από ένα διάγραμμα $AMMA$. Συνεπώς, το $AMMA$ δεν είναι ποτέ καλύτερο από τα άλλα δύο διαγράμματα. Παρ' όλα αυτά, είναι προτιμότερο να χρησιμοποιούμε τα διαγράμματα $CMEWMA$ και $MEWMA$ για την ανίχνευση μικρών μετατοπίσεων του μέσου.

Ο Pongchavalit (2009) εφάρμοσε ένα **πολυδιάστατο διάγραμμα κινούμενου μέσου** (*Multivariate Moving Average control chart or MMA control chart*) σε ηλεκτρικές μετρήσεις φωτοβολταϊκών κυττάρων και τονίζει τη βελτιωμένη ευαισθησία, που μπορεί να προσφέρει αυτή η μέθοδος ελέγχου, στη μεταβλητότητα της διεργασίας. Το διάγραμμα αυτό ονομάζεται MMV και η κατασκευή του γίνεται σε τέσσερα βήματα.

Στο πρώτο βήμα, κατασκευάζεται η στατιστική συνάρτηση T_k^2 . Προηγουμένως, όμως, θα πρέπει να ορίσουμε το διάνυσμα των δειγματικών μέσων των κινούμενων μέσων ως εξής

$$\bar{\mathbf{M}}_k = \begin{bmatrix} \bar{\mathbf{M}}_{1k} \\ \bar{\mathbf{M}}_{2k} \\ \vdots \\ \bar{\mathbf{M}}_{pk} \end{bmatrix}, \quad k = 1, 2, \dots, m.$$

Το $\bar{\mathbf{M}}_{jk}$ είναι ο δειγματικός μέσος του κινούμενου μέσου του j χαρακτηριστικού του k δείγματος και υπολογίζεται από τον τύπο

$$\bar{\mathbf{M}}_{jk} = \frac{\sum_{i=1}^n M_{ijk}}{n}, \quad \begin{cases} j = 1, 2, \dots, p \\ k = 1, 2, \dots, m \end{cases}$$

όπου M_{ijk} είναι η τιμή της i παρατήρησης του j χαρακτηριστικού του k δείγματος. Η στατιστική συνάρτηση ελέγχου θα έχει την ακόλουθη μορφή

$$T_k^2 = \mathbf{M}'_k \Sigma_{\mathbf{M}_k}^{-1} \mathbf{M}_k.$$

Στο δεύτερο βήμα, υπολογίζουμε τις δειγματικές διακυμάνσεις για το j χαρακτηριστικό του k δείγματος κινούμενου μέσω από τον τύπο

$$S_{jk}^2 = \frac{1}{n-1} \sum_{i=1}^n (M_{ijk} - \bar{\mathbf{M}}_{jk})^2, \quad \begin{cases} j = 1, 2, \dots, p \\ k = 1, 2, \dots, m \end{cases}$$

και τη συνδιακύμανση μεταξύ του j και του h χαρακτηριστικού του k δείγματος από τον τύπο που ακολουθεί

$$S_{jhk} = \frac{1}{n-1} \sum_{i=1}^n (M_{ijk} - \bar{\mathbf{M}}_{jk})(M_{ihk} - \bar{\mathbf{M}}_{hk}), \quad \begin{cases} k = 1, 2, \dots, m \\ j \neq h. \end{cases}$$

Το διάνυσμα $\bar{\bar{\mathbf{M}}}$ των τιμών στόχων των μέσων κάθε χαρακτηριστικού για m δείγματα μπορεί να εκτιμηθεί μέσω του τύπου

$$\bar{\bar{\mathbf{M}}}_j = \frac{\sum_{k=1}^m \bar{\mathbf{M}}_{jk}}{m}, \quad j = 1, 2, \dots, p.$$

Στο τρίτο βήμα, υπολογίζουμε τις (ζυγισμένες) εκτιμήσεις των στοιχείων του πίνακα διακύμανσης-συνδιακύμανσης \mathbf{S} από τους δύο τύπους που ακολουθούν

$$S_j^2 = \frac{\sum_{k=1}^m S_{jk}^2}{m}, \quad j = 1, 2, \dots, p \quad \text{και} \quad S_{jh} = \frac{\sum_{k=1}^m S_{jhk}}{m}, \quad j \neq h.$$

Στο τελευταίο βήμα, το διάνυσμα $\bar{\bar{\mathbf{M}}}$ εκτιμάται χρησιμοποιώντας τα στοιχεία $\{\bar{\bar{\mathbf{M}}}_k\}$ και ο τετραγωνικός πίνακας \mathbf{S} εκτιμάται ως εξής

$$\mathbf{S} = \begin{bmatrix} S_1^2 & S_{12} & \cdots & S_{1p} \\ & S_2^2 & \cdots & S_{2p} \\ & & \ddots & \vdots \\ & & & S_p^2 \end{bmatrix} \quad \text{ή} \quad \mathbf{S} = \begin{bmatrix} \frac{S_1^2}{nw_1} & S_{12} & \cdots & S_{1p} \\ & \frac{S_2^2}{nw_2} & \cdots & S_{2p} \\ & & \ddots & \vdots \\ & & & \frac{S_p^2}{nw_p} \end{bmatrix}$$

Η δεύτερη μορφή του πίνακα \mathbf{S} προκύπτει, επειδή ισχύει ο τύπος

$$V(M_t) = \frac{1}{w^2} \sum_{i=t-w+1}^t V(\bar{X}_i) = \frac{1}{w^2} \sum_{i=t-w+1}^t \frac{\sigma^2}{n} = \frac{\sigma^2}{nw}$$

Συνεπώς, στο **διάγραμμα MMV** απεικονίζονται οι τιμές της στατιστικής συνάρτησης T_k^2 και το άνω όριο ελέγχου από $UCL = [(mnp - mp - np + p)/(mn - m - p + 1)] \times F_{\alpha; p, mn-m-p+1}$ γίνεται

$$\left. \begin{array}{l} \text{Διάγραμμα} \\ \text{MMV} \end{array} \right\} UCL = \frac{mwp - mp - wp + p}{mw - m - p + 1} \times F_{\alpha; p, mw-m-p+1}$$

όπου p είναι ο αριθμός των μεταβλητών και το εντός ελέγχου μέσο μήκος ροής.

3.5 Παραδείγματα Πολυμεταβλητών Διαγραμμάτων Ελέγχου

3.5.1 Διαγράμματα Ελέγχου για Μεταβλητές και Διαγράμματα Ελέγχου με μνήμη

Για τη δημιουργία των πολυμεταβλητών δεδομένων, βάσει των οποίων κατασκευάστηκαν τα διαγράμματα της παρούσας παραγράφου, χρησιμοποιήθηκε ως πρότυπο το Παράδειγμα 8-1 του Montgomery (1997). Σύμφωνα με το παράδειγμα αυτό, ενδιαφερόμαστε για την ταυτόχρονη παρακολούθηση δύο ποιοτικών χαρακτηριστικών, της αντοχής στη θραύση (*tensile strength*) και της διαμέτρου μιας υφαντικής ίνας (*diameter of a textile fiber*). Το πρώτο χαρακτηριστικό μετριέται σε λίμπρες ανά τετραγωνική ίντσα (*pound per square inch, p.s.i.*)¹ και το δεύτερο σε χιλιοστά (*mm*). Για κάθε χαρακτηριστικό ($p = 2$), δημιουργήσαμε

¹ Μία λίμπρα ανά τετραγωνική ίντσα ισούται με 0.068 (φυσικές) ατμόσφαιρες ($1 \text{ p.s.i.} = 0.068 \text{ atm}$) και αποτελεί μονάδα μέτρησης της πίεσης.

τριάντα πρακαταρκτικά δείγματα ($m = 30$), καθένα από τα οποία περιλαμβάνει δεκαπέντε διαφορετικούς τύπους ινών ($n = 15$).

Τα δεδομένα αποτελούν ψευδοτυχαίους αριθμούς που προέρχονται από τη διδιάστατη κανονική κατανομή με μέση τιμή το διάνυσμα $\mu = (110 \ 10)'$ και πίνακα διακύμανσης $\Sigma = \begin{pmatrix} 4 & 1.6 \\ 1.6 & 1 \end{pmatrix}$, συμβολικά $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \sim N_{p=2}(\mu, \Sigma)$. Το Σχήμα 3.4 απεικονίζει τη διδιάστατη κανονική κατανομή. Επιπλέον, κάθε χαρακτηριστικό ακολουθεί τη μονοδιάστατη κανονική κατανομή. Στον πίνακα *Tests for Normality* γίνεται ο έλεγχος της υπόθεσης:

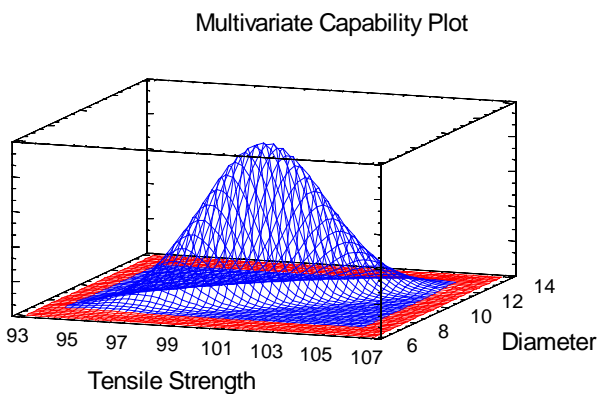
H_0 : τα δεδομένα κάθε χαρ/κού είναι ανεξάρτητα δείγματα από κανονική κατανομή

H_1 : τα δεδομένα κάθε χαρ/κού δεν είναι ανεξάρτητα δείγματα από κανονική κατανομή.

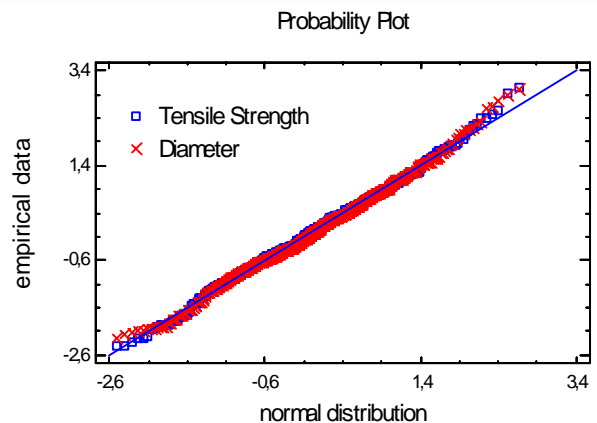
Tests for Normality

P-Values	Shapiro-Wilk
Tensile Strength	0.515583
Diameter	0.194273

Πίνακας 3.2: Έλεγχος κανονικότητας για κάθε χαρακτηριστικό



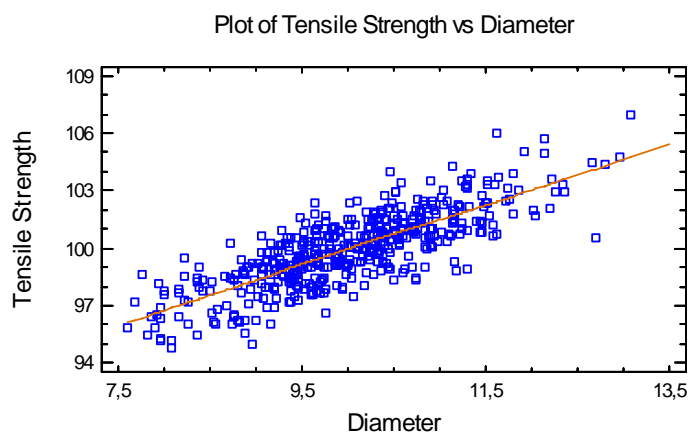
Σχήμα 3.4: Διάγραμμα προσαρμοσμένης δισδιάστατης κανονικής κατανομής



Σχήμα 3.5: Διάγραμμα προσαρμοσμένων κανονικών κατανομών για κάθε χαρακτηριστικό

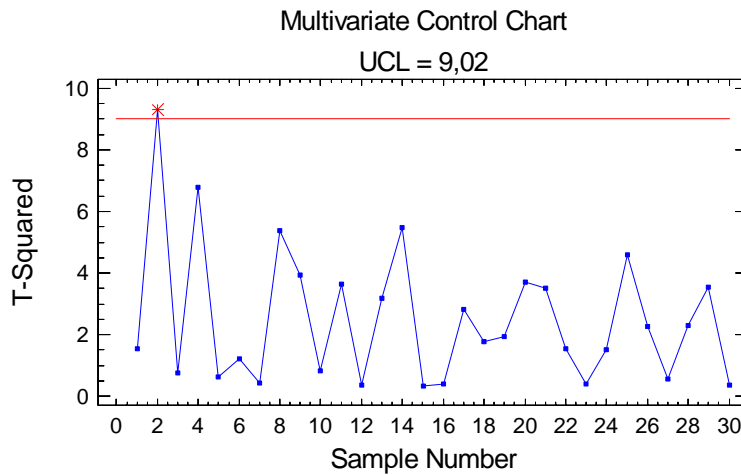
Σύμφωνα με τις τιμές των p_values του Πίνακα 3.2 δεν υπάρχει στατιστικά σημαντική ένδειξη για την απόρριψη της μηδενικής υπόθεσης. Συνεπώς, οδηγούμαστε στο συμπέρασμα ότι η κανονική κατανομή αποτελεί λογικό μοντέλο για κάθε χαρακτηριστικό. Αυτό φαίνεται και από το PP -Plot του Σχήματος 3.5, αφού τα σημεία, τόσο της αντοχής στη θραύση όσο και της διαμέτρου, βρίσκονται κατά προσέγγιση κοντά στην ευθεία και εκατέρωθεν αυτής.

Αρχικά, θα εξετάσουμε εάν υπάρχει κάποιου είδους σχέση μεταξύ των δύο ποιοτικών χαρακτηριστικών, κατασκευάζοντας το διάγραμμα διασποράς. Από το Σχήμα 3.6 είναι φανερό ότι η σχέση μεταξύ της αντοχής στη θραύση και της διαμέτρου μιας υφαντικής ίνας είναι σχεδόν γραμμική. Συνεπώς, αφού υπάρχει σχέση μεταξύ των δύο χαρακτηριστικών, θα προχωρήσουμε στην κατασκευή πολυμεταβλητών διαγραμμάτων ελέγχου.

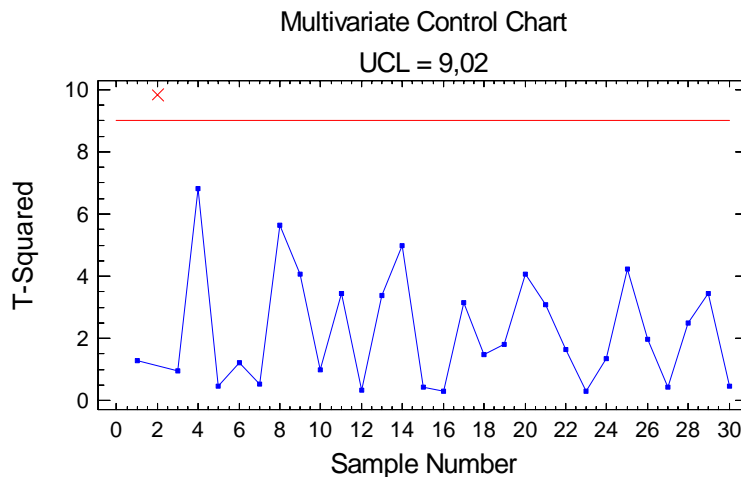


Σχήμα 3.6: Διάγραμμα διασποράς μεταξύ της αντοχής στη θραύση και της διαμέτρου μιας υφαντικής ίνας

Θα κατασκευάσουμε δύο διαγράμματα ελέγχου, το διάγραμμα T^2 του Hotelling για την παρακολούθηση του μέσου της διεργασίας, και το διάγραμμα $|S^*|$, για τον έλεγχο της μεταβλητότητας της διεργασίας. Στο διάγραμμα T^2 απεικονίζεται η στατιστική συνάρτηση του τύπου (3.2.3), ενώ στο διάγραμμα $|S^*|$ η δειγματική γενικευμένη διακύμανση. Για την κατασκευή των διαγραμμάτων Φάσης I, θα πρέπει να εκτιμήσουμε τα μ , Σ ολόκληρης της διεργασίας. Οι εκτιμήσεις αυτές θα χρησιμοποιηθούν για τον έλεγχο της μελλοντικής παραγωγής, εφ' όσον η διεργασία βρίσκεται εντός ελέγχου.



Σχήμα 3.7: Διάγραμμα T^2 του Hotelling Φάσης I για 30 δείγματα



Σχήμα 3.8: Διάγραμμα T^2 του Hotelling Φάσης I για 29 δείγματα

Στο Σχήμα 3.7 απεικονίζεται το διάγραμμα T^2 του Hotelling, το οποίο αφού κατασκευάστηκε από προκαταρκτικά δείγματα, αποτελεί διάγραμμα ελέγχου Φάσης I (Στάδιο I). Είναι φανερό ότι ένα σημείο, που αντιστοιχεί στο δεύτερο δείγμα ($m = 2$), βρίσκεται εκτός των ορίων ελέγχου. Συνεπώς η διεργασία είναι εκτός στατιστικού ελέγχου. Αυτό σημαίνει ότι πρέπει να ελέγξουμε την ύπαρξη ειδικών αιτιών μεταβλητότητας. Υποθέτοντας ότι στο συγκεκριμένο δείγμα ανακαλύφθηκαν ειδικές αιτίες μεταβλητότητας, θα πρέπει να κατασκευάσουμε ξανά το διάγραμμα T^2 , αποκλείοντας το σημείο αυτό (Σχήμα 3.8). Με τον τρόπο αυτό, η διεργασία είναι εντός ελέγχου και οι εκτιμήσεις $\bar{\bar{X}}_1 = 99.963 \text{ p.s.i.}$,

$\bar{\bar{X}}_2 = 10.021 \text{ mm}$, $\bar{S}_1^2 = 3.852$, $\bar{S}_2^2 = 1.0019$ και $\bar{S}_{12} = 1.57$ που προκύπτουν, θα χρησιμοποιηθούν για την κατασκευή του διαγράμματος Φάσης II.

Για $a = 0.01$, τα όρια ελέγχου Φάσης I υπολογίζονται από τον τύπο (3.2.4) ως εξής

$$\begin{cases} UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} \times F_{p, mn-m-p+1; a} = \frac{2(30-1)(15-1)}{30 \cdot 15 - 30 - 2 + 1} \times F_{2, 419; 0.01} = \frac{812}{419} \times 4.656 = 9.02 \\ LCL = 0. \end{cases}$$

Τα όρια ελέγχου του Σταδίου II της Φάσης I ή απλούστερα Φάσης II, υπολογίζονται από τον τύπο (3.2.6). Για $a = 0.01$, έχουμε

$$\begin{cases} UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} \times F_{p, mn-m-p+1; a} = \frac{2(30+1)(15-1)}{30 \cdot 15 - 30 - 2 + 1} \times F_{2, 419; 0.01} = \frac{868}{419} \times 4.656 = 9.645 \\ LCL = 0. \end{cases}$$

Εάν είχαμε χρησιμοποιήσει τα όρια ελέγχου του διαγράμματος χ^2 Φάσης II (τύπος (3.2.2)), θα παρατηρούσαμε ότι οι τιμές που προκύπτουν είναι λογικά κοντά στις τιμές των σωστών ορίων ελέγχου Φάσης I και Φάσης II, παρόλο που το δείγμα μας δεν είναι αρκετά μεγάλο.

$$\begin{cases} UCL = \chi_{p, a}^2 = \chi_{2; 0.01}^2 = 9.21 \\ LCL = 0. \end{cases}$$

Για την παρακολούθηση της μεταβλητότητας της διεργασίας (*process variability*), θα κατασκευάσουμε το διάγραμμα της γενικευμένης διακύμανσης. Ο πίνακας διακυμάνσεων-συνδιακυμάνσεων, που εκτιμήθηκε από τα $m = 30$ προκαταρκτικά δεδομένα, είναι

$$\mathbf{S}^* = \begin{bmatrix} 3.85238 & 1.57 \\ 1.57 & 1.0019 \end{bmatrix},$$

με ορίζουσα $|\mathbf{S}^*| = 1.3948$. Οι σταθερές b_1 και b_2 υπολογίζονται από τους τύπους (3.2.17) και (3.2.18), αντίστοιχα. Για $n = 15$, οι τιμές των σταθερών είναι

$$b_1 = \frac{1}{(n-1)^p} \cdot \prod_{i=1}^p (n-i) = \frac{1}{(15-1)^2} \cdot (15-1) \cdot (15-2) = \frac{1}{14^2} \cdot 14 \cdot 13 = 0.9286$$

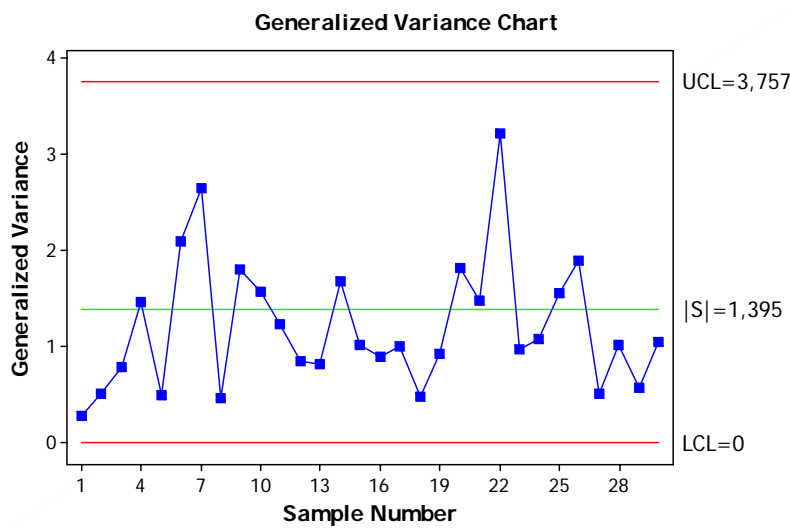
και

$$\begin{aligned} b_2 &= \frac{1}{(n-1)^{2p}} \cdot \prod_{i=1}^p (n-i) \cdot \left[\prod_{j=1}^p (n-j+2) - \prod_{j=1}^p (n-j) \right] = \\ &= \frac{1}{(15-1)^4} \cdot (15-1) \cdot (15-2) \cdot [(15-1+2) \cdot (15-2+2) - (15-1) \cdot (15-2)] = 0.2748. \end{aligned}$$

Συνεπώς, τα 3σ όρια ελέγχου Φάσης I του διαγράμματος $|S^*|$, που υπολογίζονται από τον τύπο (3.2.22), είναι τα εξής

$$\begin{cases} UCL = |S^*| \times (1 + 3\sqrt{b_2}/b_1) = 1.3948 \times (1 + 3\sqrt{0.2748}/0.9286) = 3.757 \\ CL = |S^*| = 1.3948 \\ LCL = |S^*| \times (1 - 3\sqrt{b_2}/b_1) = 1.3948 \times (1 - 3\sqrt{0.2748}/0.9286) = -0.967 \end{cases}$$

και αφού $LCL < 0$, θέτουμε $LCL = 0$.



Σχήμα 3.9: Διάγραμμα Γενικευμένης Διακύμανσης Φάσης I

Από το παραπάνω σχήμα παρατηρούμε ότι κανένα σημείο δε βρίσκεται εκτός των ορίων ελέγχου, γεγονός που υποδηλώνει ότι η δισδιάστατη μεταβλητότητα (μέσα σε κάθε δείγμα) δεν είναι ασυνήθιστα μεγάλη. Αφού η διεργασία ήταν εντός ελέγχου, τα όρια ελέγχου Φάσης II υπολογίζονται μέσω του τύπου (3.2.10) για $|\Sigma| = |S^*| = 1.3948$. Συνεπώς, τα όρια Φάσης II του διαγράμματος γενικευμένης διακύμανσης θα έχουν τις ακόλουθες τιμές

$$\begin{cases} UCL = |\Sigma| \times (b_1 + 3\sqrt{b_2}) = 1.3948 \times (0.9286 + 3\sqrt{0.2748}) = 3.4887 \\ CL = |\Sigma| \times b_1 = 1.3948 \times 0.9286 = 1.2952 \\ LCL = |\Sigma| \times (b_1 - 3\sqrt{b_2}) = 1.3948 \times (0.9286 - 3\sqrt{0.2748}) = -0.898 \end{cases}$$

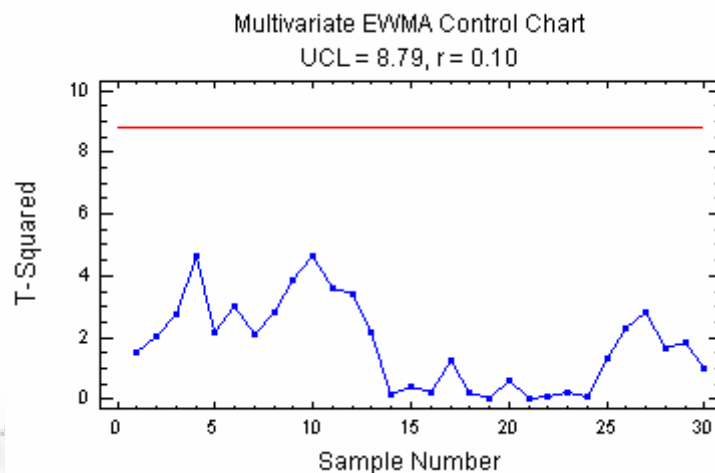
και αφού $LCL < 0$, θέτουμε $LCL = 0$.

Αριθμός Δείγματος (k ή t)	Δειγματικοί Μέσοι		Διακυμάνσεις και Συνδιακυμάνσεις			Στατιστικές Συναρτήσεις		
	Αντοχή στη θραύση (\bar{X}_{1k})	Διάμετρος (\bar{X}_{2k})	S_{1k}^2	S_{2k}^2	S_{12k}	T_k^2	$ S_k $	T_t^2
1	100.529	10.337	1.814	0.778	1.060	1.547	0.287	1.547
2	98.566	9.249	4.117	0.508	1.256	9.301	0.514	2.059
3	99.548	9.805	1.676	0.782	0.722	0.762	0.790	2.759
4	99.283	10.090	4.220	1.995	2.637	6.771	1.464	4.644
5	100.354	10.201	2.214	1.157	1.435	0.613	0.502	2.146
6	99.738	10.085	4.251	0.946	1.389	1.202	2.093	3.024
7	99.837	9.876	6.984	1.410	2.685	0.424	2.639	2.085
8	99.421	9.481	3.045	0.839	1.443	5.370	0.471	2.834
9	99.910	9.692	3.600	1.036	1.389	3.942	1.802	3.887
10	99.690	9.796	5.914	0.891	1.922	0.830	1.574	4.667
11	100.782	10.199	3.035	0.754	1.025	3.627	1.239	3.597
12	100.138	10.017	2.531	0.627	0.854	0.352	0.857	3.440
13	99.276	9.921	5.150	0.691	1.654	3.188	0.823	2.163
14	100.487	10.560	3.676	1.346	1.810	5.471	1.673	0.179
15	99.684	9.890	2.740	1.155	1.464	0.315	1.020	0.394
16	100.132	10.172	1.847	0.807	0.775	0.390	0.892	0.194
17	99.113	9.662	4.656	1.015	1.928	2.821	1.007	1.243
18	100.616	10.336	2.575	0.912	1.366	1.760	0.481	0.240
19	100.548	10.139	4.786	1.280	2.280	1.931	0.928	0.031
20	99.061	9.541	4.471	0.766	1.269	3.692	1.814	0.636
21	100.579	10.493	3.274	1.188	1.553	3.502	1.478	0.060
22	99.962	9.827	5.520	1.331	2.033	1.554	3.214	0.116
23	100.259	10.110	2.192	0.775	0.850	0.383	0.975	0.192
24	100.020	10.234	4.975	0.651	1.471	1.506	1.074	0.086
25	101.049	10.453	5.281	0.908	1.798	4.593	1.561	1.322
26	100.346	10.380	6.136	1.272	2.432	2.271	1.891	2.320
27	100.340	10.182	0.874	0.870	0.495	0.557	0.516	2.802
28	99.339	9.904	5.971	1.444	2.757	2.294	1.023	1.654
29	100.610	10.071	4.307	1.335	2.275	3.533	0.575	1.826
30	99.671	9.927	3.735	0.589	1.073	0.358	1.049	0.981
Μέσοι	$(\bar{\bar{X}}_1)$	$(\bar{\bar{X}}_2)$	$(\bar{\bar{S}}_1^2)$	$(\bar{\bar{S}}_2^2)$	$(\bar{\bar{S}}_{12})$			
	99.963	10.021	3.852	1.0019	1.57			

Πίνακας 3.3: Τιμές στατιστικών συναρτήσεων των διαγραμμάτων T^2 , $|S^*|$ και $MEWMA$.

Παραπάνω, στον Πίνακα 3.3, παραθέτουμε τις τιμές των στατιστικών συναρτήσεων των διαγραμμάτων T^2 , $|\mathbf{S}^*|$ και $MEWMA$, καθώς και των ποσοτήτων εκείνων που ήταν απαραίτητες για τον υπολογισμό τους. Οι τιμές των συναρτήσεων προέκυψαν από τα 30 προκαταρκτικά δείγματα

Αντίστοιχο του διαγράμματος ελέγχου T^2 είναι το πολυμεταβλητό $EWMA$. Ωστόσο, υπάρχει μια σημαντική διαφορά. Στο $MEWMA$ τα σημεία που απεικονίζονται αποτελούν ζυγισμένους μέσους παροντικών και παρελθοντικών παρατηρήσεων. Σύμφωνα με το παραπάνω παράδειγμα, έχουμε στη διάθεσή μας δείγματα. Υποθέτοντας ότι $r_1 = r_2 = r = 0.10$, τα διανύσματα, στα οποία θα βασίζεται το διάγραμμα $MEWMA$, θα έχουν τη μορφή $\mathbf{Z}_t = r \cdot \bar{\mathbf{X}}_t + (1-r) \cdot \mathbf{Z}_{t-1}$ και η ποσότητα που θα απεικονίζεται στο διάγραμμα θα είναι $\mathbf{T}_t^2 = \mathbf{Z}_t' \Sigma_{\mathbf{Z}_t}^{-1} \mathbf{Z}_t$ για $t=1, 2, \dots, m$, όπου $\Sigma_{\mathbf{Z}_t} = \left\{ r \left[1 - (1-r)^{2t} \right] / (2-r) \right\} \Sigma$. Οι Lowry, Woodall, Champ and Rigdon (1992) προτείνουν ως άνω όριο το $h = 8.79$ για $p = 2$, $r = 0.10$ και $ARL_0 = 200$. Στο Σχήμα 3.10 απεικονίζεται το διάγραμμα $MEWMA$. Παρατηρηρούμε ότι η διεργασία είναι εντός ελέγχου.



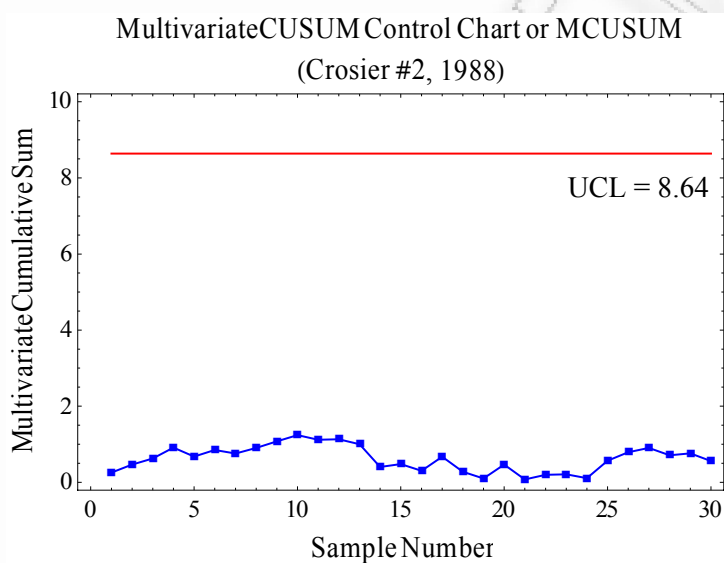
Σχήμα 3.10: Διάγραμμα $MEWMA$ Φάσης I

Για τη γρήγορη ανίχνευση μικρών ή μεσαίων μετατοπίσεων του μέσου, προτείνεται η κατασκευή των διαγραμμάτων ελέγχου τύπου $CUSUM$. Θα κατασκευάσουμε δύο πολυμεταβλητά διαγράμματα $CUSUM$, το δεύτερο διάγραμμα του Crosier (1988) και το πρώτο διάγραμμα των Pignatiello and Runger (1990). Ως άνω όριο ελέγχου θα χρησιμοποιήσουμε το $UCL = 8.64$ και στα δύο διαγράμματα. Την τιμή αυτή προτείνουν οι Mahmoud and Maravelekis (2011) για $K = 0.25$ και $p = 2$, ώστε $ARL_0 = 200$. Οι τιμές των στατιστικών συναρτήσεων των παραπάνω διαγραμμάτων δίνονται στον Πίνακα 3.4.

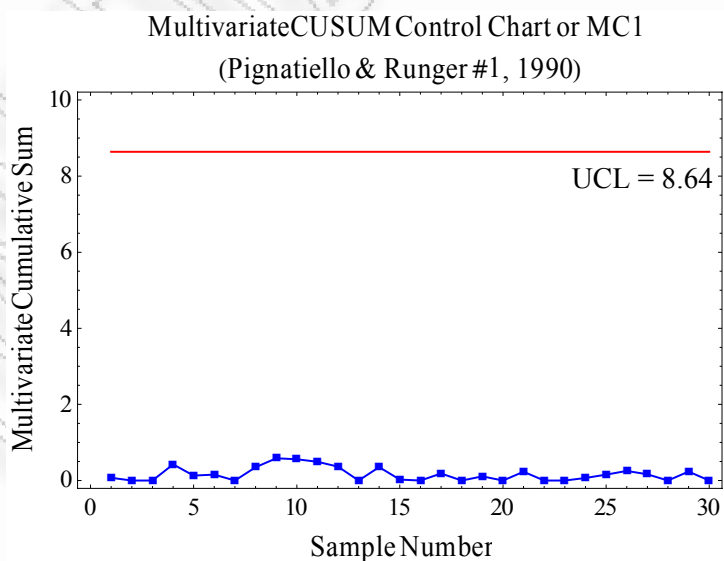
Οι στατιστικές συναρτήσεις υπολογίζονται μέσω των τύπων (3.4.2) και (3.4.3), αντίστοιχα. Ωστόσο, επειδή έχουμε δείγματα θα αντικαταστήσουμε το \mathbf{X}_i με $\bar{\mathbf{X}}_i$. Επιπλέον, χρησιμοποιούμε

$$\boldsymbol{\mu} = \begin{pmatrix} 99.9629 \\ 10.021 \end{pmatrix} \text{ και } \boldsymbol{\Sigma} = \begin{pmatrix} 3.85238 & 1.57 \\ 1.57 & 1.0019 \end{pmatrix}.$$

Παρακάτω απεικονίζονται τα δύο διαγράμματα *MCUSUM*, από τα οποία γίνεται φανερό ότι η διεργασία βρίσκεται εντός στατιστικού ελέγχου.



Σχήμα 3.11: Δεύτερο διάγραμμα *MCUSUM* του Crosier Φάσης I



Σχήμα 3.12: Πρώτο διάγραμμα *MCUSUM* των Pignatiello και Runger Φάσης I

Χρονική Στιγμή (k)	Δειγματικοί Μέσοι		Crosier				Pignatiello και Runger				
	\bar{X}		S_k	C_k	MC_k	n_k	C_k	$\ C_k\ $	MC_k		
0	-	-	0.000	0.000	-	-	0	-	-	-	-
1	100.529	10.337	0.452	0.253	1.245	0.257	1	0.566	0.316	0.321	0.071
2	98.566	9.249	-0.830	-0.456	2.055	0.466	2	-0.831	-0.456	0.466	0.000
3	99.548	9.805	-1.128	-0.609	2.676	0.626	1	-0.415	-0.216	0.225	0.000
4	99.283	10.090	-1.689	-0.505	3.785	0.913	1	-0.680	0.069	0.671	0.421
5	100.354	10.201	-1.185	-0.297	2.878	0.678	2	-0.289	0.249	0.627	0.127
6	99.738	10.085	-1.310	-0.216	3.547	0.851	3	-0.514	0.313	0.907	0.157
7	99.837	9.876	-1.323	-0.333	3.180	0.757	4	-0.640	0.168	0.783	0.000
8	99.421	9.481	-1.741	-0.815	3.751	0.904	1	-0.542	-0.540	0.598	0.348
9	99.910	9.692	-1.693	-1.079	4.425	1.078	2	-0.595	-0.869	1.085	0.585
10	99.690	9.796	-1.868	-1.239	5.052	1.240	3	-0.868	-1.094	1.307	0.557
11	100.782	10.199	-0.992	-1.003	4.575	1.117	4	0.049	-0.916	1.490	0.490
12	100.138	10.017	-0.773	-0.953	4.632	1.131	5	0.127	-0.920	1.616	0.366
13	99.276	9.921	-1.371	-0.989	4.123	1.000	6	-0.560	-1.020	1.346	0.000
14	100.487	10.560	-0.729	-0.388	1.803	0.401	1	0.524	0.539	0.603	0.353
15	99.684	9.890	-0.889	-0.457	2.107	0.480	2	0.245	0.408	0.527	0.027
16	100.132	10.172	-0.593	-0.252	1.422	0.303	3	0.414	0.559	0.682	0.000
17	99.113	9.662	-1.317	-0.558	2.851	0.672	1	-0.850	-0.359	0.434	0.184
18	100.616	10.336	-0.538	-0.197	1.321	0.277	2	-0.197	-0.044	0.117	0.000
19	100.548	10.139	0.029	-0.048	0.638	0.100	1	0.585	0.118	0.359	0.109
20	99.061	9.541	-0.767	-0.464	2.048	0.464	2	-0.317	-0.362	0.419	0.000
21	100.579	10.493	-0.081	0.005	0.539	0.075	1	0.616	0.472	0.483	0.233
22	99.962	9.827	-0.062	-0.143	1.102	0.198	2	0.615	0.278	0.317	0.000
23	100.259	10.110	0.180	-0.041	1.068	0.211	1	0.296	0.089	0.160	0.000
24	100.020	10.234	0.149	0.108	0.673	0.109	1	0.057	0.213	0.317	0.067
25	101.049	10.453	1.109	0.485	2.448	0.568	2	1.143	0.645	0.654	0.154
26	100.346	10.380	1.379	0.780	3.312	0.791	3	1.526	1.004	1.004	0.254
27	100.340	10.182	1.639	0.878	3.757	0.905	4	1.903	1.165	1.166	0.166
28	99.339	9.904	0.931	0.698	3.003	0.711	5	1.280	1.048	1.091	0.000
29	100.610	10.071	1.454	0.689	3.186	0.758	1	0.647	0.050	0.485	0.235
30	99.671	9.927	1.043	0.534	2.424	0.561	2	0.355	-0.044	0.362	0.000

Πίνακας 3.4: Τιμές στατιστικών συναρτήσεων των διαγραμμάτων $MCUSUM$.

3.5.2 Διαγράμματα Ελέγχου Τύπου Shewhart για Ιδιότητες

Για την παραγωγή των δεδομένων που θα χρησιμοποιηθούν στην κατασκευή των πολυμεταβλητών διαγραμμάτων ελέγχου για ιδιότητες, θεωρούμε μία πολυμεταβλητή διεργασία με τέσσερα ποιοτικά χαρακτηριστικά ($p=4$) για ιδιότητες, έστω $\mathbf{X}_k = (\mathbf{X}_{1k}, \mathbf{X}_{2k}, \mathbf{X}_{3k}, \mathbf{X}_{4k})'$, $k=1, 2, \dots, m$. Υποθέτουμε ότι τα δεδομένα προέρχονται από την πολυδιάστατη κατανομή Poisson με μέση τιμή το διάνυσμα $\boldsymbol{\mu} = (1, 1, 1, 1)'$ και συντελεστή συσχέτισης μεταξύ οποιονδήποτε δύο μεταβλητών $\rho = 0.75$. Επιπλέον, θεωρούμε ότι οι μετρήσεις κάθε χαρακτηριστικού προέρχονται από $m=30$ (προκαταρκτικά) δείγματα μεγέθους $n=1$. Οι μετρήσεις αυτές δίνονται στον Πίνακα 3.5.

Πρώτα, θα κατασκευάσουμε το πολυδιάστατο διάγραμμα np ή διαφορετικά το διάγραμμα MNP των Lu, Xie, Goh and Lai (1998), που θεωρείται γενίκευση του μονοδιάστατου διαγράμματος np . Στην περίπτωση αυτή, το \mathbf{X}_{jk} αντιπροσωπεύει τις μετρήσεις των ελαττωματικών προϊόντων (*counts of nonconforming units*) του j χαρακτηριστικού του k δείγματος, δηλαδή το C_{jk} για $j=1, 2, 3, 4$ και $k=1, 2, \dots, m$. Στο διάγραμμα MNP απεικονίζεται η στατιστική συνάρτηση $X = \sum_{j=1}^p (C_j / \sqrt{p_j})$, που παριστάνει το ζυγισμένο άθροισμα των ελαττωματικών προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα. Οι τιμές της συνάρτησης X δίνονται στον Πίνακα 3.5. Για να υπολογίσουμε τα όρια ελέγχου του διαγράμματος, θα πρέπει να εκτιμήσουμε το διάνυσμα του ποσοστού των ελαττωματικών προϊόντων (*fraction nonconforming vector*) $\bar{\mathbf{P}} = (\bar{p}_1, \bar{p}_2, \bar{p}_3, \bar{p}_4)$ και τον πίνακα συσχέτισης $\bar{\boldsymbol{\Sigma}} = [\bar{\delta}_{jl}]_{4 \times 4}$ της διεργασίας από τους τύπους που ακολουθούν

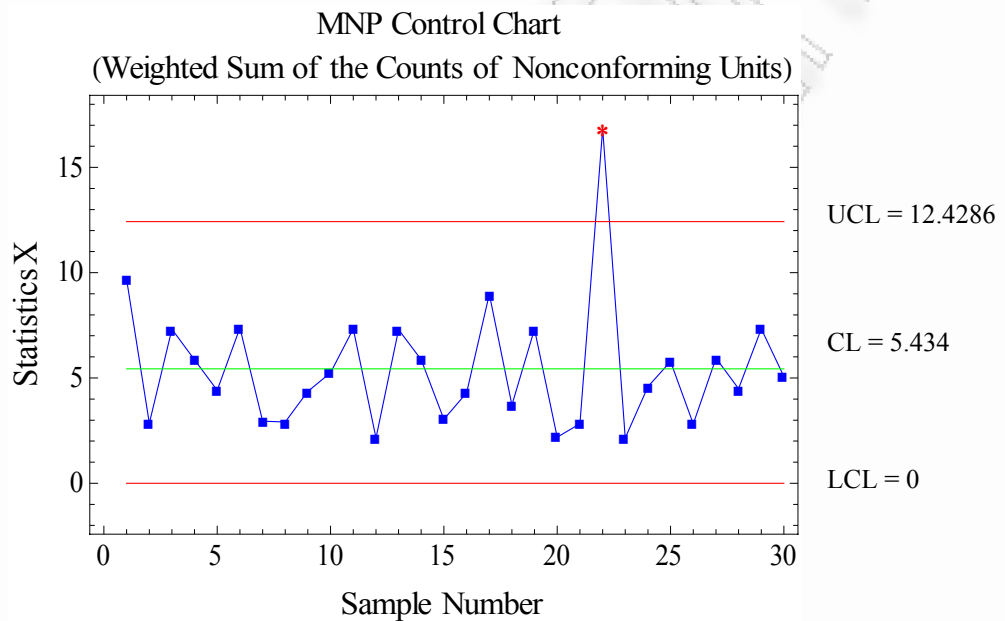
$$\bar{p}_j = \frac{\sum_{k=1}^m C_{jk}}{n \times m}, \quad j=1, 2, 3, 4, \quad k=1, 2, \dots, 30, \quad \bar{\delta}_{jl} = \frac{\text{Cov}(C_l, C_j)}{\sqrt{V(C_l) \times V(C_j)}} \quad \text{για } j \neq l \quad \text{και } \bar{\delta}_{jj} = 1 \quad \text{για } j=l.$$

$$\begin{aligned} \text{Συνεπώς, προκύπτει ότι } \bar{\delta}_{12} = \bar{\delta}_{21} = 0.368017, \quad \bar{\delta}_{13} = \bar{\delta}_{31} = 0.430924, \quad \bar{\delta}_{14} = \bar{\delta}_{41} = 0.597335, \\ \bar{\delta}_{23} = \bar{\delta}_{32} = 0.218535, \quad \bar{\delta}_{24} = \bar{\delta}_{42} = 0.263374, \quad \bar{\delta}_{34} = \bar{\delta}_{43} = 0.594, \quad \bar{\delta}_{11} = \bar{\delta}_{22} = \bar{\delta}_{33} = \bar{\delta}_{44} = 1 \quad \text{και} \\ \bar{p}_1 = \frac{59}{1 \times 30} = 0.96667, \quad \bar{p}_2 = \frac{61}{1 \times 30} = 2.03333, \quad \bar{p}_3 = \frac{47}{1 \times 30} = 1.56667, \quad \bar{p}_4 = \frac{55}{1 \times 30} = 0.83333. \end{aligned}$$

Τα δοκιμαστικά όρια ελέγχου Φάσης I, που υπολογίζονται από τον τύπο (3.3.2) είναι

$$\begin{cases} UCL = 1 \cdot \sum_{j=1}^4 \sqrt{\bar{p}_j} + 3 \sqrt{1 \cdot \left\{ \sum_{j=1}^4 (1 - \bar{p}_j) + 2 \sum_{l < j} \bar{\delta}_{lj} \sqrt{(1 - \bar{p}_l)(1 - \bar{p}_j)} \right\}} = 12.4286 \\ CL = 1 \cdot \sum_{j=1}^4 \sqrt{\bar{p}_j} = 5.434 \\ LCL = 1 \cdot \sum_{j=1}^4 \sqrt{\bar{p}_j} - 3 \sqrt{1 \cdot \left\{ \sum_{j=1}^4 (1 - \bar{p}_j) + 2 \sum_{l < j} \bar{\delta}_{lj} \sqrt{(1 - \bar{p}_l)(1 - \bar{p}_j)} \right\}} = -1.56063 \end{cases}$$

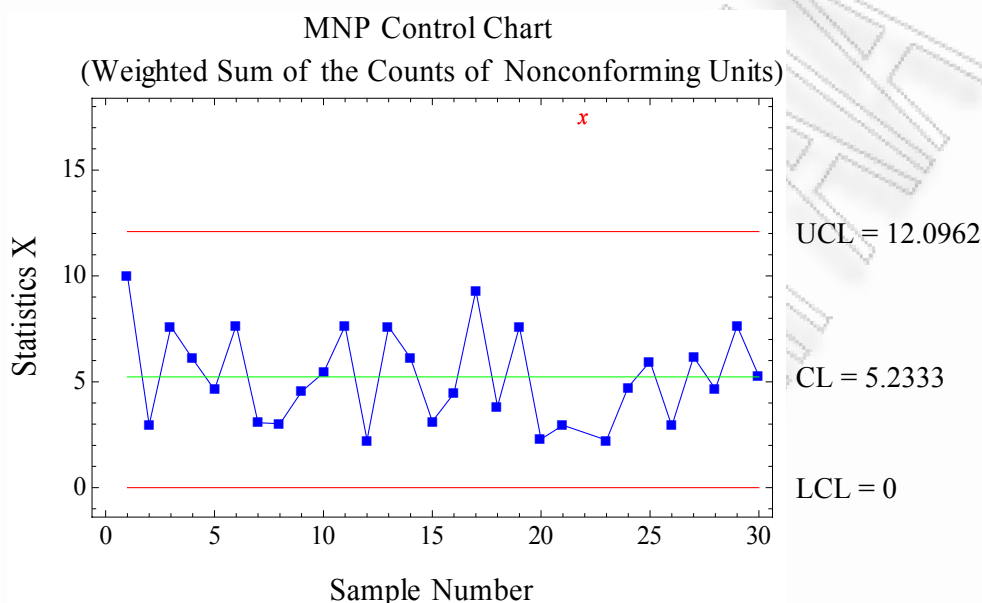
και αφού $LCL < 0$, θέτουμε $LCL = 0$.



Σχήμα 3.13: Διάγραμμα *MNP* Φάσης I για το ζυγισμένο άθροισμα των ελαττωματικών προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 30$)

Από το παραπάνω σχήμα φαίνεται ότι τα σημεία που αντιστοιχούν στο δείγμα 22 βρίσκεται εκτός των ορίων ελέγχου, οπότε θα πρέπει να ελέγξουμε τις ειδικές αιτίες μεταβλητότητας που πιθανόν να έθεσαν τη διεργασία εκτός στατιστικού ελέγχου. Υποθέτουμε ότι και στο σημεία αυτό ανακαλύφθηκαν ειδικές αιτίες μεταβλητότητας και ως αποτέλεσμα τα δοκιμαστικά όρια θα πρέπει να υπολογιστούν εκ νέου, χρησιμοποιώντας τα υπόλοιπα 29 δείγματα. Τα νέα όρια είναι $LCL = -1.62962$, $CL = 5.2333$ και $UCL = 12.0962$. Επειδή $LCL < 0$, θέτουμε $LCL = 0$. Στο Σχήμα 3.14 απεικονίζεται το διάγραμμα *MNP* μόνο για τα 29 δείγματα. Παρατηρούμε ότι τώρα όλα τα σημεία βρίσκονται εντός των ορίων. Πρακτικά αυτό σημαίνει ότι τα νέα όρια μπορούν να χρησιμοποιηθούν για τη μελλοντική

παρακολούθηση της διεργασίας. Προφανώς, διαφοροποιούνται οι τιμές της στατιστικής συνάρτησης X του παρακάτω διαγράμματος.



Σχήμα 3.14: Διάγραμμα *MNP* Φάσης I για το ζυγισμένο άθροισμα των ελαττωματικών προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 29$)

Στη συνέχεια, θα κατασκευάσουμε το διάγραμμα *MP* των Chiu and Kuo (2007), που θεωρείται γενίκευση των μονοδιάστατων μονοδιάστατων c και u . Στην περίπτωση αυτή, το X_{jk} αντιπροσωπεύει τον αριθμό των ελαττωμάτων ή μη συμμορφωμένων προϊόντων του j χαρακτηριστικού του k δείγματος για $j=1, 2, 3, 4$ και $k=1, 2, \dots, m$. Στο διάγραμμα απεικονίζεται η στατιστική συνάρτηση

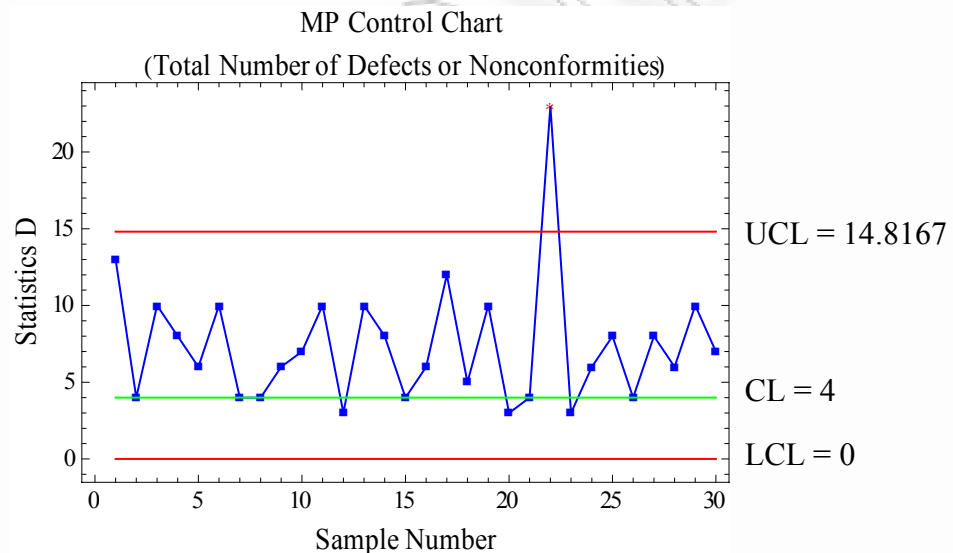
$$D = \sum_{j=1}^p X_j,$$

η οποία παριστάνει το άθροισμα ή το συνολικό αριθμό ελαττωμάτων ή μη συμμορφωμένων προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα. Για να υπολογίσουμε τα όρια ελέγχου του διαγράμματος, θα πρέπει να γνωρίζουμε το μέσο αριθμό ελαττωμάτων ή μη συμμορφώσεων για κάθε χαρακτηριστικό $\lambda_j = Var(X_j)$ για $j=1, 2, 3, 4$, και το συντελεστή συσχέτισης $\rho_{jl} = Cov(X_j, X_l)$ μεταξύ των χαρακτηριστικών X_j και X_l , για $j, l=1, 2, 3, 4$ και $j \neq l$. Για τα δεδομένα που κατασκευάσαμε έχουμε $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$

και $\rho = \rho_{12} = \rho_{21} = \rho_{13} = \rho_{31} = \rho_{14} = \rho_{41} = \rho_{23} = \rho_{32} = \rho_{24} = \rho_{42} = \rho_{34} = \rho_{43} = 0.75$. Οπότε, τα δοκιμαστικά όρια ελέγχου Φάσης I, που υπολογίζονται από τον τύπο (3.3.8) είναι

$$\begin{cases} UCL = \sum_{j=1}^4 \lambda_j + 3\sqrt{\sum_{j=1}^4 \lambda_j + 2\sum_{j<l} \rho_{jl}\sqrt{\lambda_j\lambda_l}} = 14.8167 \\ CL = \sum_{j=1}^p \lambda_j = 4 \\ LCL = \sum_{j=1}^4 \lambda_j - 3\sqrt{\sum_{j=1}^4 \lambda_j + 2\sum_{j<l} \rho_{jl}\sqrt{\lambda_j\lambda_l}} = -6.81665 \end{cases}$$

και αφού $LCL < 0$, θέτουμε $LCL = 0$.



Σχήμα 3.15: Διάγραμμα *MP* Φάσης I για το συνολικό αριθμό ελαττωμάτων ή μη συμμορφωμένων προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 30$)

Από το παραπάνω σχήμα φαίνεται ότι το σημείο που αντιστοιχεί στο δείγμα 22 είναι εκτός των ορίων ελέγχου. Υποθέτουμε ανακαλύφθηκαν ειδικές αιτίες μεταβλητότητας στο σημείο αυτό και κατά συνέπεια τα δοκιμαστικά όρια ελέγχου θα πρέπει να υπολογιστούν εκ νέου, χρησιμοποιώντας τα υπόλοιπα 29 δείγματα. Για το σκοπό αυτό, χρειάζεται να εκτιμήσουμε τις ποσότητες λ_j , $j=1,2,3,4$ και ρ_{jl} που προκύπτουν από τα 29 δείγματα. Θα πρέπει να αναφέρουμε ότι οι τιμές της στατιστικής συνάρτησης D , οι οποίες φαίνονται στον Πίνακα 3.5, δεν αλλάζουν.

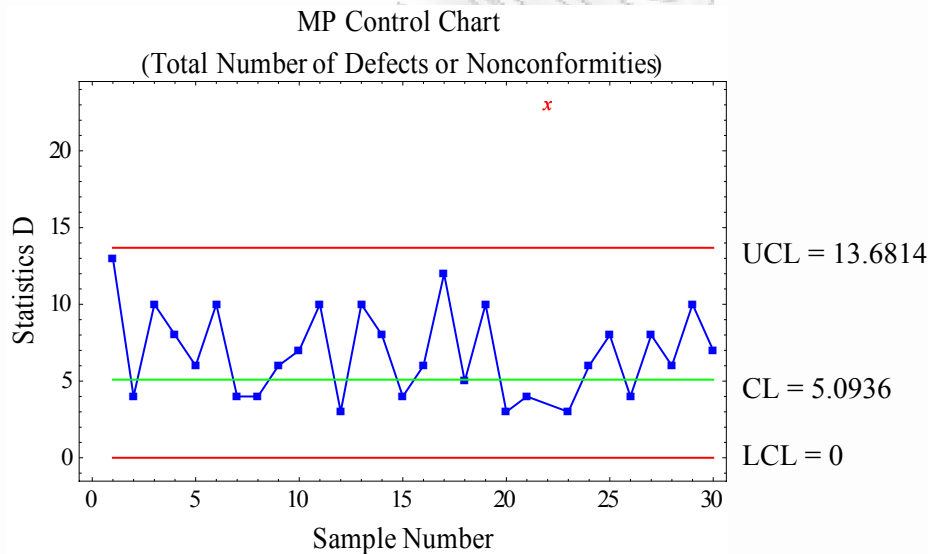
Χρονική Στιγμή (k)	Ποιοτικά Χαρακτηριστικά				Στατιστικές Συναρτήσεις		
	X_{1k}	X_{2k}	X_{3k}	X_{4k}	$X (m = 30)$	$X (m = 29)$	D
1	2	3	4	4	9.67995	10.0292	13
2	1	2	0	1	2.85420	2.96263	4
3	3	3	2	2	7.31805	7.60217	10
4	2	3	2	1	5.86643	6.07879	8
5	3	0	1	2	4.41526	4.62207	6
6	2	2	2	4	7.38079	7.66709	10
7	1	1	1	1	2.95185	3.06424	4
8	1	2	1	0	2.91458	3.01455	4
9	2	2	1	1	4.36621	4.53793	6
10	1	2	3	1	5.25100	5.42632	7
11	4	2	2	2	7.32984	7.63662	10
12	1	1	0	1	2.15291	2.24301	3
13	3	2	1	4	7.29493	7.59993	10
14	3	2	2	1	5.87822	6.11324	8
15	0	1	2	1	3.03771	3.13139	4
16	1	3	0	2	4.29403	4.45156	6
17	2	2	4	4	8.97866	9.30955	12
18	1	2	2	0	3.71352	3.83578	5
19	3	3	2	2	7.31805	7.60217	10
20	1	0	0	2	2.19017	2.29269	3
21	2	2	0	0	2.82872	2.94739	4
22	8	5	4	6	16.83810	17.5315	23
23	1	1	0	1	2.15291	2.24301	3
24	1	0	3	2	4.58698	4.75638	6
25	0	5	1	2	5.78247	5.95796	8
26	1	3	0	0	2.81693	2.91294	4
27	2	1	2	3	5.94095	6.17816	8
28	1	1	2	2	4.48933	4.65478	6
29	4	2	2	2	7.32984	7.63662	10
30	2	3	1	1	5.06749	5.25756	7
Σύνολο	59	61	47	55			

Πίνακας 3.5: Τιμές στατιστικών συναρτήσεων X και D των διαγραμμάτων MNP και MP

Οι εκτιμήσεις του μέσου αριθμού ελαττωμάτων ή μη συμμορφωμένων προϊόντων για κάθε χαρακτηριστικό είναι $\bar{\lambda}_1 = 1.11823$, $\bar{\lambda}_2 = 1.20936$, $\bar{\lambda}_3 = 1.33005$ και $\bar{\lambda}_4 = 1.43596$, ενώ οι συντελεστές συσχέτισης μεταξύ δύο ποιοτικών χαρακτηριστικών είναι $\bar{\rho}_{12} = \bar{\rho}_{21} = 0.047$, $\bar{\rho}_{13} = \bar{\rho}_{31} = 0.245$, $\bar{\rho}_{14} = \bar{\rho}_{41} = 0.333$, $\bar{\rho}_{23} = \bar{\rho}_{32} = 0.055$, $\bar{\rho}_{24} = \bar{\rho}_{42} = 0.01$ και $\bar{\rho}_{34} = \bar{\rho}_{43} = 0.50$. Τα νέα όρια που προκύπτουν είναι

$$\begin{cases} UCL = \sum_{j=1}^4 \bar{\lambda}_j + 3 \sqrt{\sum_{j=1}^4 \bar{\lambda}_j + 2 \sum_{j<l} \bar{\rho}_{jl} \sqrt{\bar{\lambda}_j \bar{\lambda}_l}} = 13.6814 \\ CL = \sum_{j=1}^4 \bar{\lambda}_j = 5.0936 \\ LCL = \sum_{j=1}^4 \bar{\lambda}_j - 3 \sqrt{\sum_{j=1}^4 \bar{\lambda}_j + 2 \sum_{j<l} \bar{\rho}_{jl} \sqrt{\bar{\lambda}_j \bar{\lambda}_l}} = -3.49426 \end{cases}$$

και αφού $LCL < 0$, θέτουμε $LCL = 0$.



Σχήμα 3.16: Διάγραμμα *MP* Φάσης I για το συνολικό αριθμό ελαττωμάτων ή μη συμμορφωμένων προϊόντων όλων των ποιοτικών χαρακτηριστικών σε κάθε δείγμα ($m = 29$)

Από το Σχήμα 3.15, γίνεται φανερό ότι όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου και δε δημιουργούν κάποιο συστηματικό πρότυπο. Ως εκ τούτου η διεργασία βρίσκεται εντός στατιστικού ελέγχου. Συνεπώς, τα όρια του παραπάνω διαγράμματος μπορούν να χρησιμοποιηθούν στο διάγραμμα Φάσης II με σκοπό την παρακολούθηση μελλοντικών δειγμάτων της διεργασίας.

3.6 Ανίχνευση της Εκτός Ελέγχου Μεταβλητής

Ο Jackson (1991) αναφέρει ότι σε κάθε διεργασία πολυμεταβλητού ποιοτικού ελέγχου θα πρέπει να

- Δοθεί μια απάντηση στο ερώτημα: «Είναι η διεργασία εντός ελέγχου;»
- Προσδιοριστεί το ολικό Σφάλμα Τύπου I.
- Λαμβάνονται υπ' όψιν οι σχέσεις μεταξύ των μεταβλητών.
- Απαντηθεί το ερώτημα: «Εάν η διεργασία είναι εκτός στατιστικού ελέγχου, σε ποια από τις υπό εξέταση μεταβλητές οφείλεται το πρόβλημα;»

Η τέταρτη ερώτηση έχει κεντρίσει ιδιαίτερα το ενδιαφέρον των ερευνητών τα τελευταία χρόνια. Όταν ένα μονοδιάστατο διάγραμμα ελέγχου δίνει σήμα εκτός ελέγχου, τότε η απάντηση στο παραπάνω ερώτημα μπορεί να δοθεί εύκολα, αφού το μονοδιάστατο διάγραμμα σχετίζεται με μία μόνο μεταβλητή. Ωστόσο, στην περίπτωση ενός πολυδιάστατου διαγράμματος ελέγχου η απάντηση δεν είναι εύκολο να δοθεί. Αυτό συμβαίνει διότι το πολυδιάστατο διάγραμμα εμπλέκεται με περισσότερες από μία μεταβλητές, οι οποίες είναι συνήθως συσχετισμένες. Συνεπώς, η αιτία του συναγεμού μπορεί να οφείλεται σε μία μόνο μεταβλητή, στη σχέση μεταξύ δύο ή περισσότερων μεταβλητών ή ακόμη και σε κάποιο συνδυασμό των δύο προηγούμενων περιπτώσεων.

3.6.1 Χρήση Μονοδιάστατων Διαγραμμάτων με Τυπικά ή Bonferroni Όρια Ελέγχου

Για να βρεθεί τι προκάλεσε την εκτός ελέγχου κατάσταση μιας διεργασίας, έχουν προταθεί πολλές μέθοδοι. Η πρώτη προφανής μέθοδος που προτάθηκε ήταν για κάθε μέτρηση να κατασκευαστεί ένα ξεχωριστό μονοδιάστατο διάγραμμα ελέγχου. Στην περίπτωση αυτή, η διεργασία τίθεται εκτός στατιστικού ελέγχου εάν σε κάποιο από τα διαγράμματα αυτά σημάνει συναγεμός. Ωστόσο, η μέθοδος αυτή έχει μειονεκτήματα. Πρώτον, το κόστος κατασκευής πολλών μονοδιάστατων διαγραμμάτων ελέγχου μπορεί να είναι αρκετά υψηλό, ιδιαίτερα όταν ο αριθμός των μεταβλητών μιας διεργασίας είναι μεγάλος. Δεύτερον, ο μεγάλος αριθμός διαγραμμάτων αυξάνει την πιθανότητα λανθασμένου συναγεμού (Σφάλμα Τύπου I). Όταν οι μετρήσεις είναι ανεξάρτητες και η πιθανότητα σφάλματος τύπου I σε κάθε διάγραμμα είναι

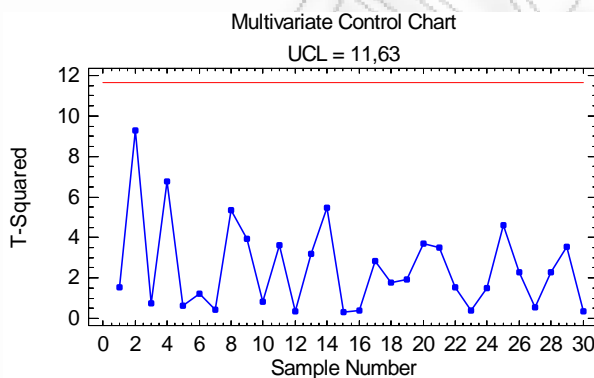
$$P(\text{σφάλματος τύπου I}) = a ,$$

τότε η πραγματική (από κοινού) πιθανότητα σφάλματος τύπου I της διεργασίας είναι

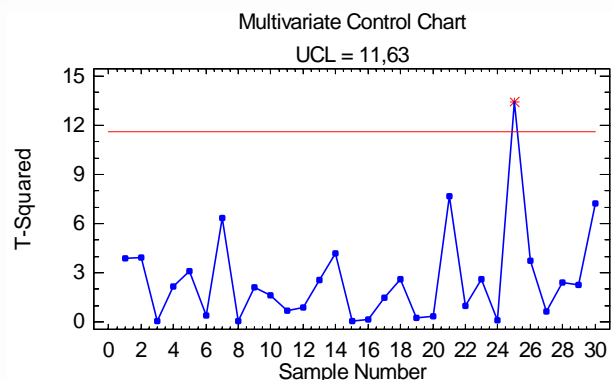
$$a' = 1 - (1 - a)^p ,$$

όπου $(1 - a)^p$ είναι η πιθανότητα όλα τα σημεία να βρίσκονται εντός των ορίων ελέγχου σε όλα τα διαγράμματα (Montgomery (1997)). Όταν, όμως, οι μετρήσεις είναι εξαρτημένες ο υπολογισμός του σφάλματος τύπου I γίνεται πιο δύσκολος. Προφανώς, σε αυτή την περίπτωση η από κοινού πιθανότητα του σφάλματος τύπου I θα είναι πολύ μεγαλύτερη από $1 - (1 - a)^p$ (Oktaý Firat and Aricigil (2000)). Οπότε, η συσχέτιση μεταξύ των μεταβλητών πρέπει να λαμβάνεται υπ' όψιν.

Θα περιγράψουμε την προαναφερθείσα μέθοδο με το παράδειγμα της Ενότητας 5.1, στο οποίο θα μετατοπίσουμε εσκεμμένα το μέσο του πρώτου χαρακτηριστικού κατά $\delta = 0.3$. Σκοπός μας είναι να διαπιστώσουμε εάν η συγκεκριμένη μέθοδος θα ανιχνεύσει σωστά το χαρακτηριστικό εκείνο που ευθύνεται για τη μετατόπιση του μέσου της διεργασίας. Αρχικά, κατασκευάζουμε το διάγραμμα T^2 του Hotelling πριν και μετά την μετατόπιση του μέσου του πρώτου χαρακτηριστικού, δηλαδή της αντοχής στη θραύση μιας υφαντικής ίνας, με σκοπό να δούμε σε τι κατάσταση βρίσκεται η διεργασία.

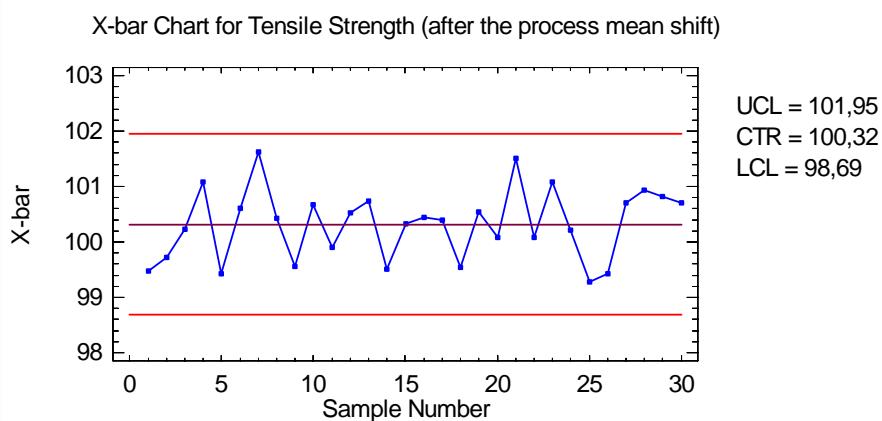


Σχήμα 3.17: Διάγραμμα T^2 του Hotelling Φάσης I πριν τη μετατόπιση του μέσου της αντοχής στη θραύση για $a = 0.0027$

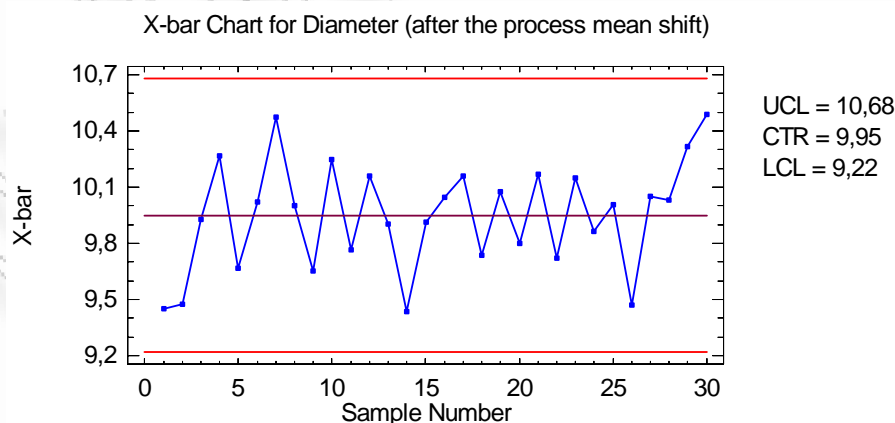


Σχήμα 3.18: Διάγραμμα T^2 του Hotelling Φάσης I μετά τη μετατόπιση του μέσου της αντοχής στη θραύση για $a = 0.0027$

Από το Σχήμα 3.17 παρατηρούμε ότι πριν αλλάξουμε το μέσο της αντοχής στη θραύση, η διεργασία βρισκόταν εντός ελέγχου. Αντίθετα, μετά τη μετατόπιση του μέσου, η διεργασία δίνει εκτός ελέγχου σήμα στο δεύτερο δείγμα. Αυτό σημαίνει ότι το διάγραμμα T^2 ανίχνευσε τη μετατόπιση αυτή. Στη συνέχεια, θα προχωρήσουμε με την κατασκευή ξεχωριστών διαγραμμάτων ελέγχου \bar{X} για κάθε χαρακτηριστικό. Μετά τη μετατόπιση του μέσου της διεργασίας, εξετάζουμε εάν τα δύο χαρακτηριστικά προέρχονται από ανεξάρτητες κανονικές κατανομές. Εφαρμόζοντας τον έλεγχο του Shapiro-Wilk, καταλήγουμε στο συμπέρασμα ότι κάθε χαρακτηριστικό ακολουθεί την κανονική κατανομή (οι τιμές των p_values για την αντοχή στη θραύση και τη διάμετρο είναι 0.946 και 0.672, αντίστοιχα).



Σχήμα 3.19: Διάγραμμα \bar{X} Φάσης I για την αντοχή στη θραύση της υφαντικής ίνας μετά τη μετατόπιση του μέσου της διεργασίας



Σχήμα 3.20: Διάγραμμα \bar{X} Φάσης I για τη διάμετρο της υφαντικής ίνας μετά τη μετατόπιση του μέσου της διεργασίας

Από τα Σχήματα 3.19 και 3.20, παρατηρούμε ότι και στα δύο διαγράμματα όλα τα σημεία είναι εντός των ορίων ελέγχου. Οπότε, θα μπορούσε κανείς να πει και ότι η διεργασία είναι εντός στατιστικού ελέγχου. Όμως, όπως διαπιστώσαμε από το διάγραμμα T^2 του Σχήματος 3.18 αυτό δεν ισχύει.

Σύμφωνα με αυτά που αναφέραμε παραπάνω, η πιθανότητα είτε το p_1 είτε το p_2 να ξεπεράσει τα 3σ όρια ελέγχου είναι 0.0027. Όμως, η πιθανότητα και τα δύο χαρακτηριστικά να βρεθούν εκτός ορίων ταυτόχρονα, όταν και τα δύο χαρακτηριστικά βρίσκονται εντός ελέγχου, είναι $0.0027 \times 0.0027 = 0.00000729$. Παρατηρούμε ότι η πιθανότητα αυτή είναι αρκετά μικρότερη από 0.0027. Επιπλέον, η πιθανότητα τα δύο χαρακτηριστικά να βρεθούν εντός ορίων ταυτόχρονα, όταν η διεργασία είναι πραγματικά εντός ελέγχου, είναι $0.9973 \times 0.9973 = 0.99460729$. Είναι φανερό ότι το σφάλμα τύπου I και η από κοινού πιθανότητα ένα σημείο να βρεθεί εντός των ορίων δεν είναι ίσα με τις αντίστοιχες τιμές που προκύπτουν από τη χρήση δύο ξεχωριστών διαγραμμάτων ελέγχου. Για αυτό το λόγο, η χρήση της παραπάνω μεθόδου μπορεί να είναι παραπλανητική. Για να είναι πιο δίκαιη η σύγκριση του πολυδιάστατου διαγράμματος με τα μονοδιάστατα, θα πρέπει η πιθανότητα σφάλματος τύπου I του διαγράμματος T^2 να είναι τουλάχιστον $a' = 1 - (1 - a)^p = 1 - (1 - 0.0027)^2 = 0.005393$ (Montgomery (1997)).

Ο Alt (1985) πρότεινε τη χρήση p μονοδιάστατων διαγραμμάτων ελέγχου με Bonferroni όρια ελέγχου, έναντι των τυπικών ορίων. Σύμφωνα με τη μέθοδο αυτή, τα Bonferroni όρια υπολογίζονται ως εξής

$$\begin{cases} LCL = \mu_i - z_{1-a/(2p)} \cdot \frac{\sigma_i}{\sqrt{n}} \\ UCL = \mu_i + z_{1-a/(2p)} \cdot \frac{\sigma_i}{\sqrt{n}} \end{cases}$$

Το πλεονέκτημα χρήσης των παραπάνω ορίων είναι ότι σε κάθε διάγραμμα η πιθανότητα τα σημεία της στατιστικής συνάρτησης να απεικονίζονται εκτός των ορίων ελέγχου, καθώς η διεργασία είναι εντός στατιστικού ελέγχου, είναι ίση με a/p και όχι a . Συνεπώς, ο αριθμός των λανθασμένων συναγεργμών μειώνεται. Ωστόσο, ο Alt τονίζει το γεγονός ότι τα p μονοδιάστατα διαγράμματα δεν πρέπει να χρησιμοποιούνται στη θέση του διαγράμματος χ^2 .

Οι Doganaksoy, Faltin and Tucker (1991) βελτίωσαν την τεχνική του Alt (1985) συνδυάζοντας διάφορες μεθόδους που κατέταξαν τις μεταβλητές με σειρά προτεραιότητας, ώστε να διερευνηθούν χρησιμοποιώντας Bonferroni όρια ελέγχου. Η μέθοδος αυτή θα

παρουσιαστεί αναλυτικά παρακάτω. Αντίθετα, οι Hayter and Tsui (1994) επέκτειναν την ιδέα του Alt με τα Bonferroni όρια ελέγχου, παρουσιάζοντας μία διαδικασία για ακριβή ταυτόχρονα διαστήματα ελέγχου για το μέσο κάθε μεταβλητής, χρησιμοποιώντας προσομοίωση. Έστω ότι a είναι η πιθανότητα ο μέσος να βρεθεί εκτός των ορίων ελέγχου, ενώ η διεργασία είναι εντός στατιστικού ελέγχου. Σύμφωνα με τη μέθοδο αυτή, για γνωστό πίνακα διακύμανσης-συνδιακύμανσης Σ και για δεδομένη πιθανότητα a , ο πειραματιστής αξιολογεί πρώτα το κρίσιμο σημείο $C_{R,a}$, όπου R είναι ο πίνακας συσχέτισης που λαμβάνεται από τον πίνακα Σ . Οι Hayter and Tsui παρουσίασαν πίνακες για την επιλογή του κατάλληλου κρίσιμου σημείου $C_{R,a}$. Στη συνέχεια, εάν συμβολίσουμε με $\mathbf{X}' = (X_1, X_2, \dots, X_p)$ κάθε παρατήρηση, τότε ο πειραματιστής κατασκευάζει p διαστήματα εμπιστοσύνης της μορφής $(X_i - \sigma_i \cdot C_{R,a}, X_i + \sigma_i \cdot C_{R,a})$, ένα για κάθε χαρακτηριστικό, όπου σ_i είναι η τυπική απόκλιση του i χαρακτηριστικού. Η μεθοδολογία αυτή διασφαλίζει ότι μπορεί να επιτευχθεί η συνολική πιθανότητα a για το σφάλμα τύπου I. Η διεργασία θεωρείται ότι είναι εντός ελέγχου, εφόσον κάθε διάστημα περιέχει την αναμενόμενη τιμή μ_{0i} . Κατά συνέπεια, η μεταβλητή (ή μεταβλητές), της οποίας (ή των οποίων) το αντίστοιχο διάστημα δεν περιέχει το μ_{0i} , είναι υπεύθυνη (ή υπεύθυνες) για το εκτός ελέγχου σήμα. Όταν

$$M = \max [|X_i - \mu_{0i}| / \sigma_i] > C_{R,a}, \quad i = 1, 2, \dots, p,$$

τότε η διεργασία δίνει σήμα εκτός ελέγχου. Γραφικά απεικονίζεται συνήθως η στατιστική συνάρτηση M για κάθε πολυμεταβλητή παρατήρηση.

3.6.2 Χρήση Ελλειπτικής Περιοχής Ελέγχου

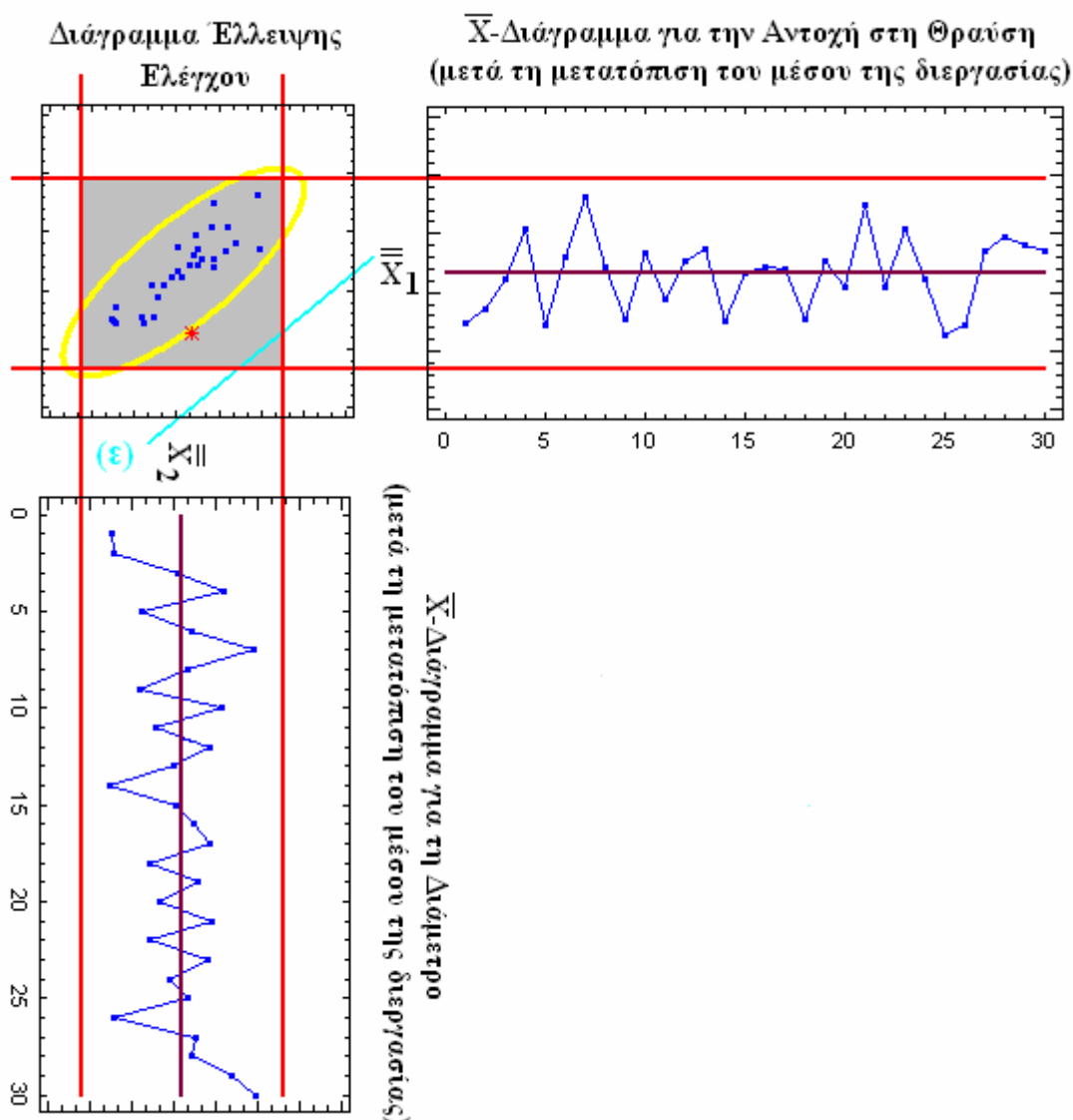
Οι Alt (1985) και Jackson (1991) πρότειναν μια μέθοδο, στην περίπτωση που έχουμε δύο μόνο μεταβλητές. Η μέθοδος αυτή βασίζεται στην κατασκευή ενός διαγράμματος που είναι γνωστό ως **διάγραμμα έλλειψης ελέγχου** (*control ellipse*). Για να δούμε πως κατασκευάζεται το διάγραμμα αυτό, υπενθυμίζουμε ότι $\chi_0^2 = (\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2$, εάν $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Στην περίπτωση που έχουμε δύο χαρακτηριστικά ισχύει

$$\chi_0^2 = \frac{1}{\sigma_{11} \cdot \sigma_{22} - \sigma_{12}^2} \cdot \left[\sigma_{22} (X_1 - \mu_1)^2 + \sigma_{11} (X_2 - \mu_2)^2 - 2\sigma_{12} (X_1 - \mu_1)(X_2 - \mu_2) \right],$$

όπου σ_{11} , σ_{22} και σ_{12} ($\sigma_{12} = \sigma_{21}$) θεωρούμε ότι είναι γνωστά. Η ποσότητα χ_0^2 ($\chi_0^2 \sim \chi_2^2$) μπορεί να αποτελέσει τη βάση κατασκευής ενός διαγράμματος για τους μέσους μ_1 και μ_2 . Εάν ο μέσος της διεργασίας παραμένει στις τιμές μ_1 και μ_2 , τότε οι τιμές της ποσότητας χ_0^2 θα πρέπει να είναι μικρότερες από το άνω όριο ελέγχου $UCL = \chi_2^2(a)$, όπου $\chi_2^2(a)$ είναι το άνω a -ποσοστιαίο σημείο της χ^2 -κατανομής με δύο βαθμούς ελευθερίας. Εάν, έστω ένας από τους δύο μέσους μετατοπισθεί σε νέα θέση, τότε η πιθανότητα η στατιστική συνάρτηση χ_0^2 να ξεπεράσει το άνω όριο ελέγχου αυξάνεται.

Σε ένα διάγραμμα έλλειψης ελέγχου, οι μετρήσεις των δύο χαρακτηριστικών απεικονίζονται με σημεία, όπως ακριβώς και σε ένα διάγραμμα διασποράς. Επιπλέον, η έλλειψη που σχεδιάζεται αντιστοιχεί στο άνω όριο ελέγχου που αναφέραμε προηγουμένως, υπό την έννοια ότι δεσμεύει το 99.73% της αναμενόμενης κατανομής των παρατηρήσεων. Στην πραγματικότητα, η έλλειψη αυτή βασίζεται σε μια πολυμεταβλητή κανονική κατανομή. Το κέντρο βάρους της τοποθετείται στο δειγματικό μέσο (μ_1, μ_2) , ενώ το μήκος και ο προσανατολισμός των αξόνων της υπολογίζονται από το δειγματικό πίνακα διακύμανσης (Polhemus (2005)).

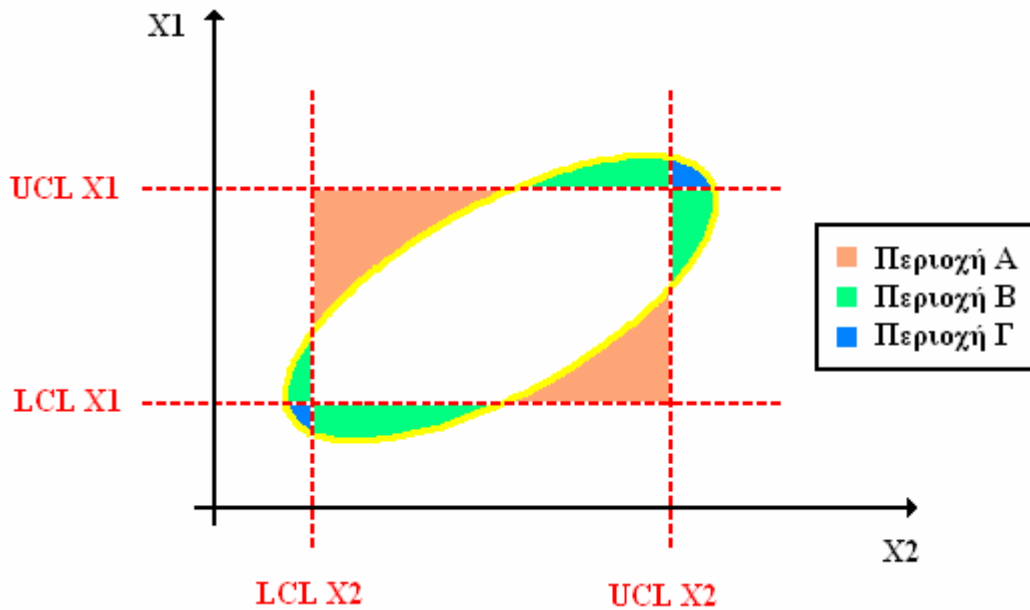
Πιο συγκεκριμένα, εάν τα δύο χαρακτηριστικά είναι ασυσχέτιστα (δηλαδή $\sigma_{12} = 0$), τότε ο κύριος άξονας της έλλειψης είναι παράλληλος με την ευθεία (ε), η οποία απεικονίζεται στο Σχήμα 3.21 και στην περίπτωση δειγμάτων περνά από το σημείο $(\bar{\bar{X}}_2, 0)$ και το συμμετρικό του σημείου $(0, \bar{\bar{X}}_1)$ ως προς την ευθεία $x = \bar{\bar{X}}_1$. Θέτοντας το χ_0^2 ίσο με $\chi_2^2(a)$, το ζεύγος $(\bar{\bar{X}}_1, \bar{\bar{X}}_2)$ δίνει κάποια τιμή στη στατιστική συνάρτηση χ_0^2 . Με βάση την τιμή αυτή εάν το ζεύγος $(\bar{\bar{X}}_1, \bar{\bar{X}}_2)$ σχεδιάζεται στο εσωτερικό της έλλειψης, η διεργασία είναι εντός ελέγχου, ενώ εάν το ζεύγος $(\bar{\bar{X}}_1, \bar{\bar{X}}_2)$ σχεδιάζεται στο εξωτερικό της έλλειψης, η διεργασία τίθεται εκτός ελέγχου. Αντίθετα, εάν τα χαρακτηριστικά είναι συσχετισμένα (δηλαδή $\sigma_{12} \neq 0$), τότε ο κύριος άξονας της έλλειψης δεν είναι πλέον παράλληλος με την ευθεία (ε). Εάν κάποιο από τα σημεία βρεθεί εκτός της έλλειψης, υπάρχει ένδειξη ότι παρουσιάστηκε κάποια συστηματική αιτία (Montgomery (1997)).



Σχήμα 3.21: Διάγραμμα έλλειψης των δύο συσχετισμένων ποιοτικών χαρακτηριστικών

Χρησιμοποιώντας το παράδειγμα που εφαρμόσαμε παραπάνω στην κατασκευή ξεχωριστών διαγραμμάτων ελέγχου, θα κατασκευάσουμε και το διάγραμμα έλλειψης. Παρατηρώντας προσεκτικά το Σχήμα 3.21, θα διαπιστώσουμε ότι ένα σημείο συμπεριφέρεται διαφορετικά σε σχέση με τα υπόλοιπα. Το σημείο αυτό, ενώ βρίσκεται μέσα στην από κοινού περιοχή ελέγχου (περιοχή με γκρι χρώμα), δεν είναι εντός της έλλειψης και σημειώνεται με κόκκινο. Η ύπαρξη του σημείου αυτού υποδεικνύει ότι κάποια συστηματική αιτία είναι παρούσα και θα πρέπει να προβούμε σε διορθωτικές ενέργειες. Θα πρέπει να υπογραμμίσουμε το γεγονός ότι τα ξεχωριστά μονοδιάστατα διαγράμματα ελέγχου δεν ήταν

ικανά να εντοπίσουν τη συστηματική αιτία αυτού του είδους, η οποία είναι πιθανώς υπεύθυνη για την εκτός ελέγχου κατάσταση της διεργασίας. Αυτό συνέβη, διότι τα ξεχωριστά διαγράμματα δεν έλαβαν υπ' όψιν τους την ύπαρξη συσχέτισης μεταξύ των δύο χαρακτηριστικών. Στο συγκεκριμένο παράδειγμα, ο βαθμός συσχέτισης των δυο χαρακτηριστικών είναι 0.8029.

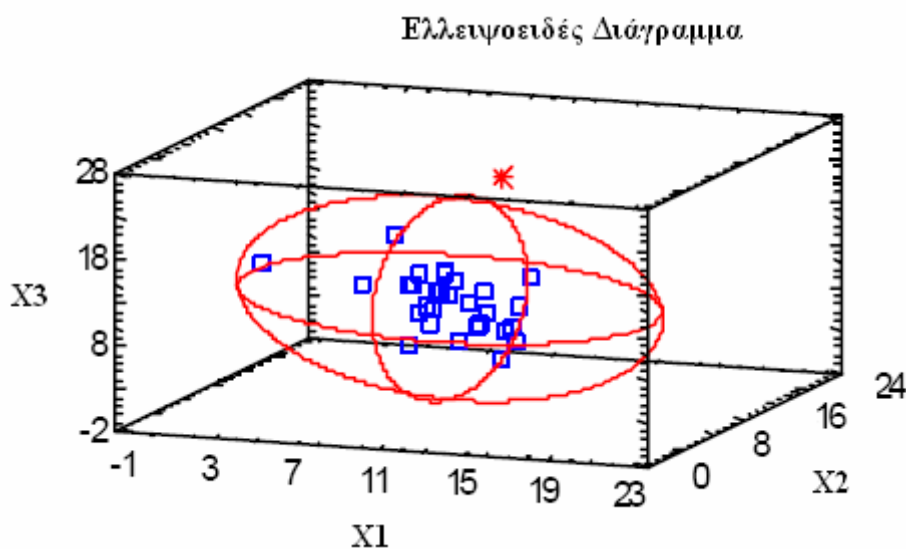


Σχήμα 3.22: Ελλειπτική και ορθογώνια περιοχή ελέγχου

Συγκρίνοντας τις δύο παραπάνω μεθόδους, θα λέγαμε ότι η χρήση ξεχωριστών διαγραμμάτων ελέγχου είναι όντως παραπλανητική. Αυτό γίνεται πιο ξεκάθαρο από το Σχήμα 3.22. Η ορθογώνια περιοχή που απεικονίζεται είναι η περιοχή ελέγχου που προκύπτει από τη χρήση μονοδιάστατων διαγραμμάτων και η έλλειψη αναπαριστά την περιοχή που προκύπτει βάσει της δεύτερης μεθόδου που αναφέραμε και λαμβάνει υπ' όψιν τη συσχέτιση των χαρακτηριστικών. Θα διαπιστώσουμε ότι η πραγματική περιοχή ελέγχου καθορίζεται από την έλλειψη και ως εκ τούτου η διεργασία είναι εκτός στατιστικού ελέγχου εάν υπάρχει ζεύγος σημείων (\bar{X}_1, \bar{X}_2) έξω από την ελλειπτική περιοχή. Ωστόσο, εάν χρησιμοποιηθεί η ορθογώνια περιοχή ελέγχου, μπορεί να καταλήξουμε λανθασμένα στο συμπέρασμα ότι τόσο το χαρακτηριστικό X_1 όσο και το X_2 είναι εντός ελέγχου (Περιοχή Α), ή ότι ένα από τα δύο είναι εντός ελέγχου ενώ το άλλο είναι εκτός (Περιοχή Β) ή ακόμη ότι και τα δύο

χαρακτηριστικά είναι εκτός ελέγχου (Περιοχή Γ). Το μέγεθος των περιοχών αυτών, καθώς και τα αντίστοιχα σφάλματά τους, επηρεάζονται από το συντελεστή συσχέτισης των δύο χαρακτηριστικών. Επιπλέον, όπως έχουμε ήδη αναφέρει, η πιθανότητα η διεργασία να είναι εντός ελέγχου είναι ακριβώς $1 - \alpha$ βάσει της έλλειψης ελέγχου, ενώ η αντίστοιχη πιθανότητα βάσει της ορθογώνιας περιοχής ελέγχου είναι τουλάχιστον $1 - \alpha$ (Alt and Smith (1988)).

Παρόλα αυτά, και το διάγραμμα έλλειψης έχει δύο σημαντικά μειονεκτήματα. Το πρώτο είναι ότι η χρονική σειρά των απεικονιζόμενων σημείων χάνεται και το δεύτερο είναι ότι η κατασκευή του διαγράμματος αυτού είναι δύσκολη για περισσότερα από δύο χαρακτηριστικά (Montgomery (1997)). Για παράδειγμα, όταν έχουμε τρία χαρακτηριστικά η περιοχή ελέγχου περιγράφεται από ένα ελλειψοειδές (Σχήμα 3.23). Το γεγονός ότι οι δύο μέθοδοι που αναφέραμε παραπάνω παρουσιάζουν αδυναμίες, καθιστά τη χρήση τους περιορισμένη. Ωστόσο, οι Mader, Glycenser and Prins (1996) χρησιμοποίησαν την ελλειπτική περιοχή ελέγχου ως τεχνική διαδικασία παρακολούθησης της βαθμονόμησης παροχής ενέργειας.



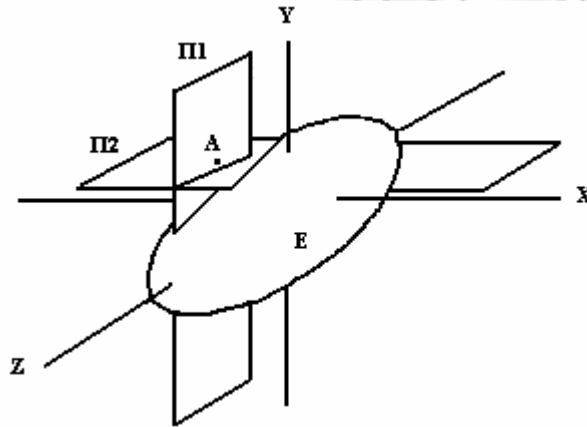
Σχήμα 3.23: Ελλειψοειδές διάγραμμα για τρία χαρακτηριστικά (StatPoint, Inc. (2005))

Εναλλακτικά, οι Chua and Montgomery (1992) πρότειναν μία διαγνωστική μέθοδο, η οποία ενσωματώνει τα χαρακτηριστικά του διαγράμματος ελέγχου έλλειψης, αλλά είναι ελεύθερη από κάθε συμπληρωματικό σφάλμα τύπου I και II. Με τον τρόπο αυτό, διατηρεί τα παραδοσιακά σφάλματα τύπου I και II. Η μέθοδος αυτή είναι γνωστή ως **μέθοδος**

υπερεπιπέδου (*hyperplane method*) και είναι ιδιαίτερα χρήσιμη όταν η μετατόπιση της διεργασίας είναι μικρή. Η εξίσωση για μια υπερ-ελλειψοειδή περιοχή ελέγχου είναι η

$$\mathbf{X}'\mathbf{S}^{-1}\mathbf{X} = T_{ref},$$

όπου \mathbf{X} είναι το διάνυσμα παρατηρήσεων, \mathbf{S}^{-1} είναι ο αντίστροφος του δειγματικού πίνακα συνδιακύμανσης και T_{ref} είναι η τιμή αναφοράς T του Hotelling. Αν $\mathbf{X}'\mathbf{S}^{-1}\mathbf{X} < T_{ref}$, η παρατήρηση θα βρίσκεται εντός του υπερ-ελλειψοειδούς και κατά συνέπεια εντός ελέγχου. Αντίθετα, εάν $\mathbf{X}'\mathbf{S}^{-1}\mathbf{X} > T_{ref}$, η παρατήρηση θα βρίσκεται εκτός του υπερ-ελλειψοειδούς και συνεπώς εκτός ελέγχου.



Σχήμα 3.24: Μέθοδος υπερ-επιπέδου στον τρισδιάστατο χώρο
(Chua and Montgomery (1992))

Στο Σχήμα 3.24 απεικονίζεται η μέθοδος υπερ-επιπέδου στον τρισδιάστατο χώρο και δύο επίπεδα ($\Pi 1$ και $\Pi 2$) κάθετα μεταξύ τους. Τα επίπεδα $\Pi 1$ και $\Pi 2$ είναι παράλληλα στα συντεταγμένα επίπεδα YZ και XZ , αντίστοιχα. Θεωρούμε ότι υπάρχει ένα σημείο του επιπέδου $\Pi 1$, έστω A , το οποίο είναι εκτός ελέγχου. Επιπλέον, στο σχήμα υπάρχει και μία ελλειπτική περιοχή ελέγχου που δημιουργείται από την τομή του επιπέδου $\Pi 1$ με το ελλειψοειδές. Η ελλειπτική περιοχή μπορεί να χρησιμοποιηθεί σαν ένα διάγραμμα έλλειψης. Ως αποτέλεσμα, ένα σημείο που δε βρίσκεται στην περιοχή αυτή, θεωρείται ότι είναι πάντα εκτός ελέγχου. Η ίδια διαδικασία εφαρμόζεται και στο επίπεδο $\Pi 2$, οπότε δημιουργείται ένα

ακόμη διάγραμμα έλλειψης. Κατά συνέπεια, για p μεταβλητές θα υπάρχουν $p(p-1)/2$ διαγράμματα έλλειψης.

Η εξίσωση του υπερ-επιπέδου έχει την ακόλουθη μορφή

$$\mathbf{d} \cdot (\mathbf{X} - \mathbf{b}) = 0,$$

όπου \mathbf{d} είναι το διάνυσμα της κατεύθυνσης του υπερ-επιπέδου, \mathbf{X} είναι το διάνυσμα παρατήρησης και \mathbf{b} είναι μία παρατήρηση που βρίσκεται στο υπερ-επίπεδο. Για να κατασκευάσουμε ένα υπερ-επίπεδο, το οποίο έχει διάνυσμα κατεύθυνσης παράλληλο σε έναν από τους άξονες αναφοράς, θα πρέπει το διάνυσμα κατεύθυνσης \mathbf{d} να έχει τη μορφή

$$\mathbf{d} = (0, 0, \dots, i, \dots, 0, 0),$$

όπου $i=1$ υποδεικνύει τον επιθυμητό άξονα αναφοράς. Για να τμήσουμε το υπερ-ελλειψοειδές με το υπερ-επίπεδο, λύνουμε το σύστημα των $p-1$ εξισώσεων που ακολουθούν

$$\begin{cases} \mathbf{X}'\mathbf{S}^{-1}\mathbf{X} = T_{ref} \\ \mathbf{d} \cdot (\mathbf{X} - \mathbf{b}) = 0 \end{cases}$$

αποκλείοντας τις δύο μεταβλητές ελέγχου που έχουν τεθεί ως στόχοι. Σκοπός της όλης διαδικασίας είναι να μειώσουμε τις $p-1$ εξισώσεις σε μία εξίσωση δευτέρου βαθμού με δύο αγνώστους. Στόχος είναι οι πολλαπλοί όροι που παράγονται να ελαττωθούν. Η τελική εξίσωση μειώνεται περαιτέρω σε μια τυπική εξίσωση έλλειψης με τη μέθοδο συμπλήρωσης τετραγώνων. Τα βήματα που απαιτούνται συνοψίζονται ως εξής

Βήμα 1^ο: Αντικαθιστούμε τις $p-2$ εξισώσεις των υπερ-επιπέδων με την εξίσωση του υπερ-ελλειψοειδούς, υπό την προϋπόθεση ότι τα διανύσματα των κατευθύνσεων που έχουν οριστεί ως στόχοι είναι παράλληλα στους επιθυμητούς άξονες αναφοράς.

Βήμα 2^ο: Διαγωνιοποιούμε το τετραγωνικό τμήμα της εξίσωσης του πρώτου βήματος με ένα μετασχηματισμένο πίνακα \mathbf{C} , όπου $\det(\mathbf{C}) = 1$.

Βήμα 3^ο: Μειώνουμε την εξίσωση του δεύτερου βήματος σε μια τυπική εξίσωση έλλειψης με τη μέθοδο συμπλήρωσης τετραγώνων.

Όπως αναφέραμε, υπό κανονικές συνθήκες, η μέθοδος υπερ-επιπέδων παρέχει $p(p-1)/2$ διαγράμματα έλλειψης. Ωστόσο, είναι πιθανό η εξίσωση της έλλειψης να είναι εκφυλισμένη ή κενή, το οποίο σημαίνει ότι η εκτός ελέγχου παρατήρηση είναι πολύ μακριά από την περιοχή

ελέγχου του υπερ-ελλειψοειδούς. Σε αυτή την περίπτωση, ο αριθμός των διαγραμμάτων έλλειψης μειώνεται σε κάποιο βαθμό, ανάλογα με το πόσο μακριά εντοπίζεται η παρατήρηση.

3.6.3 Χρήση της Διάσπασης της Στατιστικής Συνάρτησης T^2

Οι Mason, Tracy and Young (1995) πρότειναν τη χρήση της μεθόδου διάσπασης του T^2 (T^2 Decomposition) για την ανίχνευση των μεταβλητών που προκάλεσαν το εκτός ελέγχου σήμα. Η μέθοδος είναι γνωστή ως διάσπαση *MYT* (*MYT decomposition*). Η βασική ιδέα της μεθόδου αυτής είναι η διάσπαση της στατιστικής συνάρτησης T^2 σε ανεξάρτητα μέρη, καθένα από τα οποία αντικατοπτρίζει τη συμβολή μιας ανεξάρτητης μεταβλητής. Το κύριο μειονέκτημα της μεθόδου είναι ότι η διάσπαση της συνάρτησης T^2 σε p ανεξάρτητες T^2 συνιστώσες δεν είναι μοναδική, διότι υπάρχουν $p!$ διαφορετικοί εξαρτημένοι συνδυασμοί.

Σύμφωνα με την παραπάνω διαδικασία, η διάσπαση του T^2 του Hotelling μπορεί να γίνει με βάση τον ακόλουθο τύπο

$$T^2 = T_1^2 + T_{2,1}^2 + T_{3,1,2}^2 + \dots + T_{p-1,2,\dots,p-1}^2 = T_1^2 + \sum_{j=1}^{p-1} T_{j,1,2,\dots,j-1}^2. \quad (3.6.1)$$

Ο πρώτος όρος του παραπάνω τύπου αποτελεί το μη δεσμευμένο T^2 του Hotelling (*unconditional Hotelling's T^2*) για την πρώτη μεταβλητή του \mathbf{X} και υπολογίζεται ως εξής

$$T_1^2 = \left(\frac{X_1 - \bar{X}_1}{S_1} \right)^2,$$

όπου \bar{X}_1 και S_1 είναι η μέση τιμή και η τυπική απόκλιση της μεταβλητής X_1 , αντίστοιχα. Οι άλλοι όροι του τύπου (3.6.1), οι οποίοι είναι γνωστοί ως δεσμευμένοι όροι (*conditional terms*), υπολογίζονται μέσω του τύπου

$$T_{j,1,2,\dots,j-1}^2 = \frac{(X_j - \bar{X}_{j,1,2,\dots,j-1})^2}{S_{j,1,2,\dots,j-1}^2}, \quad j = 1, 2, \dots, p,$$

όπου $\bar{X}_{j,1,2,\dots,j-1} = \bar{X}_j + \mathbf{b}_j^t (\mathbf{X}_i^{(j-1)} - \bar{\mathbf{X}}^{(j-1)})$, $\mathbf{X}_i^{(j-1)}$ είναι το $(j-1)$ διάνυσμα που δεν περιλαμβάνει την j μεταβλητή, \bar{X}_j είναι ο δειγματικός μέσος της j μεταβλητής, $\mathbf{b}_j = [S_{\mathbf{X}\mathbf{X}}^{-1} S_{\mathbf{X}X}]$ είναι ένα $(j-1)$ -διάστατο διάνυσμα που εκτιμά τους συντελεστές

παλινδρόμησης της j μεταβλητής, η οποία παλινδρομείται στις πρώτες $(j-1)$ μεταβλητές, $S_{j-1,2,\dots,j-1}^2 = S_X^2 - S_{XX}' S_{XX}^{-1} S_{XX}$. Η ποσότητα $T_{j-1,2,\dots,j-1}^2$ ακολουθεί την κατανομή $((n+1)/n) \times F_{1, n-1}$ και μπορεί να χρησιμοποιηθεί για να ελέγξουμε εάν η j μεταβλητή συμμορφώνεται με τη σχέση των άλλων μεταβλητών, όπως καθορίζεται από τα ιστορικά δεδομένα.

Όπως έχουμε ήδη αναφέρει, η σειρά των p συνιστωσών του διανύσματος \mathbf{X} δεν είναι μοναδική και η σειρά αυτή που δόθηκε παραπάνω αναπαριστά μόνο μία από τις $p!$ διαφορετικές σειρές των συνιστωσών αυτών. Κάθε σειρά δημιουργεί την ίδια ολική τιμή του T^2 , αλλά παρέχει ένα μοναδικό διαχωρισμό του T^2 σε p ορθογώνιους όρους. Εξαιρώντας τις επαναλήψεις, υπάρχουν $p \times 2^{(p-1)}$ διακριτές συνιστώσες μεταξύ των $p \times p!$ δυνατών όρων που πρέπει να αξιολογηθούν για πιθανή συνεισφορά σε ένα σήμα.

Οι Mason, Tracy and Young (1997) πρότειναν έναν αλγόριθμο υπολογισμού, ο οποίος μειώνει αρκετά τις πράξεις που απαιτεί η προηγούμενη διαδικασία. Η μέθοδος αυτή αναπτύχθηκε για την περίπτωση των μεμονωμένων παρατηρήσεων, όμως μπορεί εύκολα να γενικευτεί και στην περίπτωση των ορθολογικών υποομάδων. Ο αλγόριθμος είναι ο εξής:

Βήμα 1^ο: Υπολογίζουμε τη μεμονωμένη στατιστική συνάρτηση T^2 για κάθε συνιστώσα του διανύσματος \mathbf{X} . Αφαιρούμε τις μεταβλητές, των οποίων οι παρατηρήσεις παράγουν στατιστικά σημαντικό T_i^2 . Οι παρατηρήσεις των μεταβλητών αυτών είναι εκτός ελέγχου και δεν είναι απαραίτητο να ελέγξουμε πως σχετίζονται με τις άλλες μεταβλητές. Έχοντας αφαιρέσει τις σημαντικές μεταβλητές, το πλήθος των αρχικών μεταβλητών μειώνεται και μπορούμε να ελέγξουμε το υπο-διάνυσμα των k υπόλοιπων μεταβλητών για σήμανση συναγερμού. Εάν δε λαμβάνουμε σήμα, έχουμε εντοπίσει την πηγή του προβλήματος.

Βήμα 2^ο: (Το βήμα αυτό είναι προαιρετικό, αλλά χρήσιμο για μεγάλο p). Εξετάζουμε τη δομή της συσχέτισης των υπόλοιπων μεταβλητών και αφαιρούμε κάθε μεταβλητή που δεν έχει ισχυρή συσχέτιση (0.3 ή λιγότερο) με όλες τις άλλες μεταβλητές. Η συνεισφορά της μεταβλητής, που ανήκει σε αυτή την κατηγορία, μπορεί να μετρηθεί με την T_i^2 συνιστώσα.

Βήμα 3^ο: Εάν ένα σήμα παραμένει στο υπο-διάνυσμα των k μεταβλητών που δε διαγράφηκαν, υπολογίζουμε όλους τους $T_{i,j}^2$ όρους. Στη συνέχεια αφαιρούμε από

τη διαδικασία όλα τα ζευγάρια των μεταβλητών (X_i, X_j) που έχουν στατιστικά σημαντικό $T_{i,j}^2$ όρο. Αυτό υποδεικνύει ότι υπάρχει κάποιο πρόβλημα στη σχέση μεταξύ των δύο μεταβλητών. Όταν αυτό συμβεί, οι υπό εξέταση μεταβλητές μειώνονται επιπλέον. Εξετάζουμε όλες τις μεταβλητές που αφαιρούνται για να δούμε εάν είναι υπαίτιες για το σήμα. Υπολογίζουμε το T^2 για το υπο-διάγραμμα που έμεινε. Εάν δεν υπάρχει σήμα, η πηγή του προβλήματος εντοπίζεται στις σχέσεις μεταξύ των δύο μεταβλητών και στις μεμονωμένες μεταβλητές εκείνες που ήταν εκτός ελέγχου.

Βήμα 4^ο: Εάν το υπο-διάγραμμα των υπόλοιπων μεταβλητών δίνει σήμα εκτός ελέγχου, υπολογίζουμε όλους τους $T_{i,j,k}^2$ όρους. Αφαιρούμε όλους τους τριπλούς συνδυασμούς (X_i, X_j, X_k) των μεταβλητών που έδωσαν στατιστικά σημαντικά αποτελέσματα και ελέγχουμε το υπο-διάγραμμα που μένει για σήμα.

Βήμα 5^ο: Συνεχίζουμε να υπολογίζουμε τις επιδράσεις μεγαλύτερων τάξεων μέχρι να μη μείνει καμία μεταβλητή στο υπο-διάγραμμα. Η χειρότερη περίπτωση θα ήταν να υπολογιστούν όλοι οι μοναδικοί όροι.

Γενικά, η στατιστική συνάρτηση T^2 που σχετίζεται με μία παρατήρηση πολυμεταβλητού προβλήματος είναι μια συνάρτηση των υπολοίπων που προκύπτουν από τις γραμμικές παλινδρομήσεις μεταξύ των μεταβλητών της διεργασίας. Τα κατάλοιπα αυτά περιέχονται στους δεσμευμένους T^2 όρους της ορθογώνιας διάσπασης της στατιστικής συνάρτησης. Οι Mason and Young (1999) έδειξαν ότι τα μεγάλα κατάλοιπα σε ένα από αυτά τα μοντέλα οφείλονται στις λανθασμένες προδιαγραφές του μοντέλου. Υποστήριξαν, επίσης, ότι εάν βελτιωθούν οι προδιαγραφές του μοντέλου τη στιγμή που τα δεδομένα κατασκευάζονται, αυξάνεται η ευαισθησία της συνάρτησης T^2 με σκοπό την ανίχνευση του σήματος.

Πολλές φορές, η παραγωγικότητα μιας βιομηχανικής διαδικασίας εξαρτάται από τα μηχανήματα που αλλάζουν κατά τη διάρκεια του χρόνου. Αυτές οι αλλαγές μπορεί να μην είναι σταθερές και να συμβαίνουν σταδιακά. Παρόλο που οι αλλαγές σε κάθε στάδιο φαίνεται να μην είναι σημαντικές, όταν εξετάζονται-παρακολουθούνται σε διάφορα στάδια, μπορεί να είναι ουσιαστικές. Αυτό έχει ως αποτέλεσμα, τυπικές διαδικασίες ποιοτικού ελέγχου να μη λειτουργούν σωστά, όταν υπάρχουν σταδιακές αλλαγές και μάλιστα οι παρατηρήσεις είναι συσχετισμένες. Οι Mason, Tracy and Young (1996) παρουσίασαν μία εναλλακτική

διαδικασία ελέγχου για την ανίχνευση της διεργασίας κάτω από αυτές τις προϋποθέσεις. Η μεθοδολογία που πρότειναν βασίζεται στη διπλή διάσπαση της T^2 συνάρτησης του Hotelling.

Οι Doganaksoy, Faltin and Tucker (1991) πρότειναν μία μέθοδο που βασίζεται στη χρήση της μονοδιάστατης διαδικασίας t -κατάταξης και στη χρήση p μη δεσμευμένων T^2 όρων. Η στατιστική συνάρτηση που χρησιμοποιείται είναι η επόμενη

$$t = \frac{\bar{X}_{i, new} - \bar{X}_{i, ref}}{\sqrt{S_{ii} \left(\frac{1}{n_{new}} + \frac{1}{n_{ref}} \right)}}$$

όπου $\bar{X}_{i, new}$ είναι ο μέσος του νέου δείγματος, $\bar{X}_{i, ref}$ είναι ο μέσος του δείγματος αναφοράς, S_{ii} είναι η εκτίμηση της διακύμανσης της i μεταβλητής του δείγματος αναφοράς, n_{new} είναι το μέγεθος του νέου δείγματος και n_{ref} είναι το μέγεθος του δείγματος αναφοράς. Τα βήματα του αλγόριθμου των Doganaksoy, Faltin and Tucker (1991) είναι τα ακόλουθα

Βήμα 1^ο: Διεξάγουμε έναν T^2 έλεγχο με προκαθορισμένο επίπεδο σημαντικότητας α . Εάν σημαίνει συναγεμμός, τότε προχωράμε στο επόμενο βήμα.

Βήμα 2^ο: Για κάθε μεταβλητή υπολογίζουμε το μικρότερο επίπεδο σημαντικότητας K_{ind} , το οποίο θα δώσει ένα μεμονωμένο διάστημα εμπιστοσύνης για $\mu_{i, ref} - \mu_{i, new}$, με $i = 1, 2, \dots, p$, που περιέχει το μηδέν, όπου $\mu_{i, ref}$ και $\mu_{i, new}$ είναι τα μέσα διανύσματα των πληθυσμών από τους οποίους το δείγμα αναφοράς και το νέο δείγμα επιλέγεται, αντίστοιχα. Για το K_{ind} , υποθέτουμε ότι t είναι η τιμή της μονοδιάστατης στατιστικής συνάρτησης t για μία μεταβλητή και $T(t; d)$ είναι η αθροιστική συνάρτηση της t -κατανομής με d βαθμούς ελευθερίας. Συνεπώς, θα ισχύει $K_{ind} = \left| 2T(t; n_{ref} - 1) - 1 \right|$.

Βήμα 3^ο: Απεικονίζουμε το K_{ind} για κάθε μεταβλητή σε μία (0-1) κλίμακα. Παρατηρούμε ότι οι μεταβλητές με μεγάλες K_{ind} τιμές είναι αυτές με τις μεγάλες μονοδιάστατες τιμές της στατιστικής συνάρτησης t , οι οποίες πρέπει να εξεταστούν προσεκτικά, διότι είναι πιθανό να ανήκουν στις συνιστώσες που υπέστησαν αλλαγή. Εάν οι ενδείξεις των ιδιαίτερα ύποπτων μεταβλητών είναι επιθυμητές, τότε συνεχίζουμε.

Βήμα 4^ο: Υπολογίζουμε το διάστημα εμπιστοσύνης $K_{Bonf} = (p + K_{sim} - 1) / p$, το οποίο αποδίδει το επιθυμητό ονομαστικό διάστημα εμπιστοσύνης K_{sim} των ταυτόχρονων διαστημάτων εμπιστοσύνης Bonferroni για $\mu_{i,ref} - \mu_{i,new}$, $i = 1, 2, \dots, p$ (οι συγγραφείς έδωσαν οδηγίες για τη σωστή επιλογή του K_{sim}).

Βήμα 5^ο: Οι συνιστώσες που έχουν $a^{ind} > a^{bonf}$, ταξινομούνται ως εκείνες που είναι πιο πιθανές να αλλάξουν.

Ο Murphy (1987) πρότεινε μία μέθοδο, η οποία πηγάζει από τη διαχωριστική ανάλυση (*discriminant analysis*). Το βασικό της μειονέκτημα είναι ότι όσο μεγαλύτερος είναι ο αριθμός των ποιοτικών χαρακτηριστικών, τόσο μεγαλύτερη ασάφεια εισάγεται στη διαδικασία αναγνώρισης, με αποτέλεσμα να οδηγούμαστε σε λανθασμένα αποτελέσματα. Συνεπώς, όταν ο αριθμός των μεταβλητών είναι μεγάλος, η χρήση της μεθόδου περιορίζεται.

Η μέθοδος του Murphy χρησιμοποιεί την ολική τιμή του T^2 και τη συγκρίνει με την τιμή του T_*^2 που βασίζεται σε ένα υποσύνολο μεταβλητών. Αυτή η διαγνωστική μέθοδος ενεργοποιείται από ένα εκτός ελέγχου σήμα που δίνει ένα διάγραμμα T^2 . Ο Murphy χώρισε το δειγματικό μέσο $\bar{\mathbf{X}}$ σε δύο υπο-διανύσματα $\bar{\mathbf{X}}_{*1}$ και $\bar{\mathbf{X}}_{*2}$, όπου το p_1 -διάστατο διάνυσμα $\bar{\mathbf{X}}_{*1}$ είναι υποσύνολο των $p = p_1 + p_2$ μεταβλητών, οι οποίες είναι ύποπτες για το εκτός ελέγχου σήμα. Οπότε, η ολική τετραγωνική απόσταση δίνεται από τον τύπο

$$T_p^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

και η μειωμένη απόσταση που αντιστοιχεί στο υποσύνολο των p μεταβλητών είναι η εξής:

$$T_*^2 = n(\bar{\mathbf{X}}_{*1} - \boldsymbol{\mu}_{01})' \boldsymbol{\Sigma}_0^{-1} (\bar{\mathbf{X}}_{*1} - \boldsymbol{\mu}_{01}).$$

Αποδεικνύεται ότι υπό τη μηδενική υπόθεση, η διαφορά $D = T_p^2 - T_*^2$ ακολουθεί τη χ^2 -κατανομή με p_1 β.ε. και το διάνυσμα $\bar{\mathbf{X}}_{*1}$ ακολουθεί p_1 -διάστατη κατανομή με μέσο $\boldsymbol{\mu}_{01}$ και πίνακα διακύμανσης συνδιακύμανσης $\boldsymbol{\Sigma}_{01}$.

Ο Murphy πρότεινε τον επόμενο προς-τα-εμπρός αλγόριθμο επιλογής (*Forward Selection Algorithm, FSA*). Για κάθε $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_p)'$ εκτελούμε τα επόμενα βήματα

Βήμα 1^ο: Διεξάγουμε έναν T^2 έλεγχο με προκαθορισμένο επίπεδο σημαντικότητας α (ο Murphy συστήνει το επίπεδο σημαντικότητας να ανήκει στο διάστημα

$0.1 \leq a \leq 0.2$, όταν $D_{p-i} \sim \chi_{p-i}^2$). Εάν σημάνει συναγερμός, τότε προχωράμε στο επόμενο βήμα.

Βήμα 2^ο: Υπολογίζουμε τα p μεμονωμένα $T_1^2(X_i)$, το οποίο είναι ισοδύναμο με το να εφαρμόσουμε p μεμονωμένα διαγράμματα ελέγχου. Επίσης, υπολογίζουμε τις p διαφορές $D_{p-1}(i) = [T_p^2 - T_1^2(X_i)]$, επιλέγουμε το $\min(D_{p-1}(i)) = D_{p-1}(r)$ και ελέγχουμε αυτήν την ελάχιστη διαφορά. Εάν η $D_{p-1}(r)$ δεν είναι σημαντική, τότε μόνο η r -οστή μεταβλητή χρήζει προσοχής. Αντίθετα, εάν η $D_{p-1}(r)$ είναι σημαντική συνεχίζουμε στο Βήμα 3.

Βήμα 3^ο: Υπολογίζουμε τις $p-1$ διαφορές $D_{p-2}(r, j) = [T_p^2 - T_2^2(X_r, X_j)]$, $1 \leq r \neq j \leq p$. Επιλέγουμε το $\min(D_{p-2}(r, j)) = D_{p-2}(r, s)$ και ελέγχουμε αυτήν την ελάχιστη διαφορά. Εάν η $D_{p-2}(r, s)$ δεν είναι σημαντική, τότε η r -οστή και η s -οστή μεταβλητή χρήζουν προσοχής. Αντίθετα, εάν η $D_{p-2}(r, s)$ είναι σημαντική συνεχίζουμε στο Βήμα 4.

Βήμα 4^ο: Όμοιο με το Βήμα 3, αλλά ελέγχουμε τη διαφορά $D_{p-3}(r, s, t)$.

⋮

Βήμα p : Εάν ο τελικός $D_{p-(p-1)}$ έλεγχος είναι στατιστικά σημαντικός, τότε όλες οι p μεταβλητές χρήζουν προσοχής.

Οι Chua and Montgomery (1992) παρουσίασαν ένα σχέδιο ελέγχου που περιλαμβάνει τρία βασικά βήματα, το διάγραμμα ελέγχου *MEWMA* των Lowry, Woodall, Champ and Rigdon (1992), έναν προς-τα-πίσω αλγόριθμο επιλογής (*Backward Selection Algorithm, BSA*) και τη μέθοδο του υπερ-επιπέδου, που έχουμε ήδη αναφέρει. Το διάγραμμα *MEWMA* χρησιμοποιείται για την ανίχνευση μετατοπίσεων μιας πολυμεταβλητής διεργασίας, ο *BSA* αλγόριθμος για την ανίχνευση των μεταβλητών που είναι πραγματικά εκτός ελέγχου, όταν η διεργασία τίθεται εκτός ελέγχου, και η μέθοδος υπερ-επιπέδου παρέχει ένα σύνολο ελλειπτικών διαγραμμάτων ως εργαλείο διάγνωσης των μεταβλητών της διεργασίας. Η μεθοδολογία αυτή έχει ως εξής:

1. Μία πολυμεταβλητή παρατήρηση απεικονίζεται σε ένα διάγραμμα *MEWMA*.

2. Εάν η παρατήρηση είναι εντός ελέγχου, τότε το σχέδιο ελέγχου επιστρέφει στην αρχή και ελέγχει την επόμενη παρατήρηση. Διαφορετικά, ελέγχει τον αριθμό των μεταβλητών της διεργασίας.
3. Εάν ο αριθμός των μεταβλητών της διεργασίας είναι μεγαλύτερος από πέντε, τότε ο BSA αλγόριθμος παρακάμπτεται και χρησιμοποιείται η μέθοδος υπερ-επιπέδου. Διαφορετικά, εφαρμόζεται ο BSA αλγόριθμος για την παρατήρηση.
4. Ο BSA αλγόριθμος επιλέγει το σύνολο των εκτός ελέγχου μεταβλητών, στο οποίο εφαρμόζεται η μέθοδος του υπερ-επιπέδου.
5. Η μέθοδος υπερ-επιπέδου παράγει τα απαραίτητα διαγράμματα έλλειψης για διάγνωση.
6. Με βάση την παραπάνω διάγνωση, προβαίνουμε σε διορθωτικές κινήσεις και το σχέδιο ελέγχου γυρνάει στην αρχή και ελέγχει την επόμενη παρατήρηση.

Ο αλγόριθμος των Chua and Montgomery (1992) είναι παρόμοιος με αυτόν του Murphy (1987) με τη διαφορά ότι είναι ένας προς τα πίσω αλγόριθμος επιλογής. Ο BSA αλγόριθμος χωρίζει τις p μεταβλητές σε δύο υποσύνολα p_1 και p_2 , όπου $p = p_1 + p_2$. Επίσης, ο πίνακας συνδιακύμανσης χωρίζεται σε δύο μέρη με p_1 και p_2 μεταβλητές, αντίστοιχα. Για όλες τις μεταβλητές p υπολογίζεται η συνάρτηση $T_p^2 = (\mathbf{X} - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X} - \bar{\mathbf{X}})$, όπου \mathbf{X} είναι το δειγματικό διάνυσμα παρατηρήσεων, $\bar{\mathbf{X}}$ είναι το δειγματικό μέσο διάνυσμα και \mathbf{S}^{-1} είναι ο αντίστροφος του δειγματικού πίνακα συνδιακύμανσης. Αντίστοιχα, για τις p_1 μεταβλητές υπολογίζεται η ποσότητα $T_{p_1}^2 = (\mathbf{X}_{p_1} - \bar{\mathbf{X}}_{p_1})' \mathbf{S}_{p_1}^{-1} (\mathbf{X}_{p_1} - \bar{\mathbf{X}}_{p_1})$, όπου \mathbf{X}_{p_1} είναι το δειγματικό διάνυσμα παρατηρήσεων των p_1 μεταβλητών, $\bar{\mathbf{X}}_{p_1}$ είναι το δειγματικό μέσο διάνυσμα των p_1 μεταβλητών και $\mathbf{S}_{p_1}^{-1}$ είναι ο αντίστροφος του δειγματικού πίνακα συνδιακύμανσης των p_1 μεταβλητών. Υπολογίζεται η διαφορά $D = T_p^2 - T_{p_1}^2$, η οποία ακολουθεί την κατανομή με p_2 β.ε.. Τα βήματα του αλγόριθμου είναι τα επόμενα

Βήμα 1^ο: Διεξάγουμε έναν T_p^2 έλεγχο για τις p μεταβλητές και συγκρίνουμε με το K_p χρησιμοποιώντας τις p μεταβλητές ως κριτήριο. Εάν $T_p^2 < K_p$, η διεργασία είναι εντός ελέγχου, ενώ εάν $T_p^2 \geq K_p$, η διεργασία είναι εκτός στατιστικού ελέγχου και προχωράμε στο επόμενο βήμα.

Βήμα 2^ο: Εφαρμόζουμε τους T_{p-1}^2 ελέγχους και υπολογίζουμε τις διαφορές $D_{p-1} = T_p^2 - T_{p-1}^2$. Εάν $D_{p-1} \leq K_{p-1}$, ο έλεγχος δεν είναι σημαντικός, ενώ εάν $D_{p-1} > K_{p-1}$, ο έλεγχος είναι σημαντικός. Η ένωση του συνόλου των μεταβλητών για όλους τους ελέγχους που είναι στατιστικά σημαντικοί, αποτελεί ένα νέο κριτήριο. Διαφορετικά, διατηρούμε το κριτήριο του πρώτου βήματος. Με $D(p)$ συμβολίζουμε την ελάχιστη τιμή της D και συνεχίζουμε στο επόμενο βήμα.

Βήμα 3^ο: Εφαρμόζουμε τους T_{p-2}^2 ελέγχους και υπολογίζουμε τις διαφορές $D_{p-2} = T_p^2 - T_{p-2}^2$. Εάν $D_{p-2} \leq K_{p-2}$, ο έλεγχος δεν είναι σημαντικός, ενώ εάν $D_{p-2} > K_{p-2}$, είναι σημαντικός. Η ένωση του συνόλου των μεταβλητών για όλους τους ελέγχους που είναι στατιστικά σημαντικοί, αποτελεί ένα νέο κριτήριο. Διαφορετικά, διατηρούμε το κριτήριο του δεύτερου βήματος. Με $D(p-1)$ συμβολίζουμε την ελάχιστη τιμή της διαφοράς D και συνεχίζουμε στο επόμενο βήμα.

⋮

Βήμα p : Εφαρμόζουμε T_1^2 ελέγχους. Εάν $D_1 \leq K_1$, ο έλεγχος δεν είναι σημαντικός, ενώ εάν $D_1 > K_1$, ο έλεγχος είναι σημαντικός. Με $D(1)$ συμβολίζουμε την ελάχιστη τιμή της διαφοράς D και συνεχίζουμε στο επόμενο βήμα.

Βήμα $p+1$: Εάν όλοι οι έλεγχοι από το Βήμα 2 έως το Βήμα p είναι σημαντικοί, τότε όλες οι p μεταβλητές είναι εκτός ελέγχου. Εάν $D_i \leq K_j$, όπου $i=1, 2, \dots, p-1$ και $j=i+1, i+2, \dots, p$ με $i < j$, και το σύνολο των μεταβλητών της διαφοράς $D(i)$ είναι υποσύνολο του συνόλου των μεταβλητών της διαφοράς $D(j)$, τότε το σύνολο των μεταβλητών για την $D(i)$ είναι το εκτός ελέγχου σύνολο μεταβλητών. Διαφορετικά, αυξάνουμε το i κατά 1 και επαναλαμβάνουμε αυτή τη διαδικασία ελέγχου.

Μία ακόμα μέθοδος, που βασίζεται στο T^2 , είναι η προς-τα-κάτω διαδικασία (*Step-Down Procedure*) του Timm (1996), η οποία χρησιμοποιεί πεπερασμένους ελέγχους τομής (*Finite Intersections Tests, FIT*). Η μέθοδος αυτή, υποθέτει ότι υπάρχει μία προκαθορισμένη σειρά μεταξύ των μέσων των p μεταβλητών και ότι ελέγχει διαδοχικά υποσύνολα

χρησιμοποιώντας αυτή τη σειρά για να ορίσει τη διάταξη. Εάν και το T^2 είναι βέλτιστο για την ανίχνευση μετατόπισης του μέσου διανύσματος, δεν είναι βέλτιστο όταν η αλλαγή συμβαίνει σε ένα υποσύνολο μεταβλητών. Σύμφωνα με τον Timm, όταν κάτι τέτοιο συμβεί, το πιο αποτελεσματικό είναι να χρησιμοποιήσουμε πεπερασμένο έλεγχο τομής.

Η διαδικασία είναι εντός ελέγχου, όταν η υπόθεση $H_0: \mu_i = \mu_{0i}, i=1, 2, \dots, p$ είναι αληθής, όπου μ_i είναι ο μέσος της i μεταβλητής. Προφανώς, για να ελέγξουμε κάθε μία από τις p υποθέσεις μεμονωμένα, είναι απαραίτητη η χρήση της *FIT* διαδικασίας. Ωστόσο, θα πρέπει πρώτα να ορίσουμε μία στατιστική συνάρτηση ελέγχου του λόγου πιθανοφάνειας για κάθε στοιχειώδη υπόθεση H_i και στη συνέχεια να προσδιορίσουμε την από κοινού κατανομή της συνάρτησης αυτής. Η στατιστική συνάρτηση υπολογίζεται από τον τύπο

$$Z_i^2 = \frac{(X_i - \mu_{0i})^2}{\sigma_{ii}}.$$

Η από κοινού κατανομή της Z_i^2 ακολουθεί τη μη κεντρική χ^2 -κατανομή με ένα βαθμό ελευθερίας. Η πολυμεταβλητή διεργασία θα είναι εντός στατιστικού ελέγχου εάν $Z_i^2 < \chi_{p,1-a}^2$, όπου $P(Z_i^2 < \chi_{p,1-a}^2 | H_0) = 1 - a$, ενώ εκτός στατιστικού ελέγχου εάν $\max(Z_i^2) > \chi_{p,1-a}^2$, όπου $P(\max(Z_i^2) > \chi_{p,1-a}^2) = a$. Εναλλακτικά, αφού η μη κεντρική χ^2 -κατανομή είναι ειδική περίπτωση της πολυδιάστατης κατανομής t με άπειρους βαθμούς ελευθερίας, η διεργασία τίθεται εκτός ελέγχου όταν $P(\max|T_i| = |X_i - \mu_{0i}| / \sqrt{\sigma_{ii}} > T_{p,1-a}^2) = a$. Ως αποτέλεσμα, έχουμε $P(|T_i| \leq T_{p,1-a}^2 | H_0) = 1 - a$. Για κάθε μεταβλητή μπορεί να κατασκευαστεί ένα ταυτόχρονο διάστημα εμπιστοσύνης της μορφής $X_i - T_{p,1-a}^2 \times \sqrt{\sigma_{ii}} \leq \mu_i \leq X_i + T_{p,1-a}^2 \times \sqrt{\sigma_{ii}}$. Εάν τα σύνολα εμπιστοσύνης d_i δεν περιέχουν το μ_{0i} για $i=1, 2, \dots, p$, η διεργασία τίθεται εκτός στατιστικού ελέγχου.

Σύμφωνα με τους Runger, Alt and Montgomery (1996), η συνεισφορά μιας μεταβλητής σε ένα εκτός ελέγχου σήμα μπορεί να μετρηθεί από την ελάχιστη τιμή της στατιστικής συνάρτησης T^2 , η οποία υπολογίζεται αλλάζοντας μία μεταβλητή. Οι μεταβλητές που επιβαρύνονται με μεγάλες αλλαγές, συνεισφέρουν περισσότερο στην ενεργοποίηση του συναγερμού. Οι συγγραφείς θεωρούν ότι το διάνυσμα παρατήρησης από μία διεργασία σε κάθε χρονική στιγμή είναι το \mathbf{Y} με συνιστώσες $Y_i, i=1, 2, \dots, p$. Για ευκολία υποθέτουν ότι

το \mathbf{Y} έχει γνωστό μέσο διάνυσμα $\mathbf{0}$ και γνωστό πίνακα διακύμανσης-συνδιακύμανσης Σ . Το διάγραμμα T^2 , κατά τη στιγμή ανωμαλίας, συμβολίζεται ως εξής

$$T^2 = \mathbf{Y}'\Sigma^{-1}\mathbf{Y}.$$

Εάν με UCL συμβολίσουμε το άνω όριο ελέγχου του διαγράμματος T^2 , τότε τη στιγμή που $T^2 > UCL$, ενεργοποιείται ο συναγερμός. Η μεθοδολογία των Runger, Alt and Montgomery (1996) είναι γνωστή ως μέθοδος *RAM*. Στην πραγματικότητα, οι συγγραφείς χρησιμοποιούν τρεις προσεγγίσεις ώστε να αναπτύξουν μία διαγνωστική τεχνική για το διάγραμμα T^2 και δημιουργούν την ίδια στατιστική συνάρτηση για κάθε προσέγγιση.

Σύμφωνα με την πρώτη τεχνική, θεωρούμε ότι \mathbf{e}_i είναι το μοναδιαίο διάνυσμα στην κατεύθυνση του i άξονα συνιστώσεων. Για να μετρήσουμε τη συνεισφορά του \mathbf{Y} με το T^2 , προσδιορίζουμε το c_i , ώστε να ελαχιστοποιηθεί η ακόλουθη ποσότητα:

$$T_i^2 = \left(\mathbf{Y} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i' \Sigma^{-1} \mathbf{e}_i}} \right)' \Sigma^{-1} \left(\mathbf{Y} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i' \Sigma^{-1} \mathbf{e}_i}} \right) \quad (3.6.2)$$

Εάν το c_j είναι μεγάλο σε σχέση με τα άλλα c_i , τότε απαιτείται μια μεγάλη τροποποίηση του Y_j για την ελαχιστοποίηση του T^2 , γεγονός που υποδηλώνει ότι το Y_j συνεισφέρει σημαντικά στο T^2 .

Εάν $\mathbf{Z} = \Sigma^{-1/2}\mathbf{Y}$ και $\mathbf{v}_i = \Sigma^{-1/2}\mathbf{e}_i / \left(\mathbf{e}_i' \Sigma^{-1} \mathbf{e}_i \right)^{1/2}$, τότε \mathbf{v}_i είναι ένα μοναδιαίο διάνυσμα στην κατεύθυνση του $\Sigma^{-1/2}\mathbf{e}_i$. Συνεπώς, ο τύπος (3.6.2) μπορεί να πάρει τη μορφή

$$T_i^2 = (\mathbf{Z} - c_i \mathbf{v}_i)' (\mathbf{Z} - c_i \mathbf{v}_i)$$

όπου c_i είναι η Ευκλείδεια απόσταση κατά μήκος του μοναδιαίου διανύσματος \mathbf{v}_i που ελαχιστοποιεί το T^2 . Η τιμή του c_i που ελαχιστοποιεί τον προηγούμενο τύπο, θεωρείται ότι είναι η κλίση της ευθείας παλινδρόμησης του \mathbf{Z} στο \mathbf{v}_i . Έτσι προκύπτει ότι

$$c_i = \mathbf{v}_i' \mathbf{Z} = \mathbf{e}_i' \Sigma^{-1} \mathbf{Y} / \left(\mathbf{e}_i' \Sigma^{-1} \mathbf{e}_i \right)^{1/2},$$

οπότε προφανώς το c_i είναι ανάλογο του $\mathbf{e}_i' \Sigma^{-1} \mathbf{Y}$.

Κατά τους Healy (1987), Pignatiello and Runger (1990) και Hawkins (1993), το c_i ισούται με την προτεινόμενη στατιστική συνάρτηση για την ανίχνευση μιας μετατόπισης του μέσου της διεργασίας κατά μήκος του διανύσματος \mathbf{e}_i . Μία απλή πολυμεταβλητή διαγνωστική τεχνική ελέγχου είναι να συγκρίνουμε τα σχετικά μεγέθη του $|c_i|$ (ή του c_i^2) για $i = 1, 2, \dots, p$. Εάν το $|c_i|$ είναι μεγάλο, τότε η συνιστώσα μέτρησης Y_i είναι μακριά από το μεγαλύτερο μέρος των ιστορικών δεδομένων και είναι αυτή που συμβάλλει κυρίως στο T^2 . Γεωμετρικά, το $|c_i|$ είναι το μήκος της ορθής προβολής του διανύσματος \mathbf{Z} στο \mathbf{v}_i . Οπότε, θα ισχύει $|c_i| = \|\mathbf{P}_i \mathbf{Z}\|$, όπου \mathbf{P}_i είναι η ορθή προβολή του πίνακα στο διάνυσμα \mathbf{v}_i . Οι γεωμετρικές αναλογίες απλοποιούνται θεωρώντας τα \mathbf{Z} και \mathbf{v}_i , επειδή μετά την τροποποίηση των μεταβλητών αυτών, το T_i^2 είναι η τετραγωνική Ευκλείδεια απόσταση του \mathbf{Z} από το μηδενικό διάνυσμα.

Η δεύτερη διαγνωστική τεχνική σχετίζεται με την προσέγγιση που οδηγεί στο c_i . Έστω ότι η ελάχιστη τιμή του T_i^2 μπορεί να επιτευχθεί αλλάζοντας το Y_i με c_i . Ορίζουμε τον τύπο

$$T_i^2 = \left(\mathbf{Y} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i' \boldsymbol{\Sigma}^{-1} \mathbf{e}_i}} \right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{Y} - \frac{c_i \mathbf{e}_i}{\sqrt{\mathbf{e}_i' \boldsymbol{\Sigma}^{-1} \mathbf{e}_i}} \right) = (\mathbf{Z} - c_i \mathbf{v}_i)' (\mathbf{Z} - c_i \mathbf{v}_i). \quad (3.5.3)$$

Τότε ο παραπάνω τύπος (3.5.3) είναι το άθροισμα τετραγώνων των υπολειμμάτων του μοντέλου παλινδρόμησης. Συνεπώς,

$$T_i^2 = \mathbf{Z}' (\mathbf{I} - \mathbf{P}_i) \mathbf{Z} = \mathbf{Z}' \mathbf{Z} - \mathbf{Z}' \mathbf{v}_i \mathbf{v}_i' \mathbf{Z} = \mathbf{Z}' \mathbf{Z} - c_i^2. \quad (3.5.4)$$

Για το λόγο αυτό, το Y_i συνεισφέρει αρκετά στο T^2 , εάν η τιμή του T^2 μπορεί να ελαττωθεί ουσιαστικά από μία τροποποίηση του Y_i . Ως αποτέλεσμα, εάν η τιμή της μετρικής συνάρτησης $D_i = T^2 - T_i^2$ είναι μεγάλη, τότε το Y_i συμβάλλει σημαντικά στο T^2 . Από τον τύπο $T^2 = \mathbf{Y}' \boldsymbol{\Sigma}^{-1} \mathbf{Y} = \mathbf{Z}' \mathbf{Z}$ και τον τύπο (3.5.4) συνεπάγεται ότι $D_i = c_i^2$, οπότε η μετρική D_i είναι ισοδύναμη με τη μετρική c_i^2 .

Τέλος, στην τρίτη διαγνωστική τεχνική, θεωρούμε την τυποποιημένη μεταβλητή $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2} \mathbf{Y}$ και ελέγχουμε την υπόθεση $H_0: \boldsymbol{\mu} = \mathbf{0}$ έναντι της εναλλακτικής $H_1: \boldsymbol{\mu} \neq \mathbf{0}$. Σύμφωνα με τη μηδενική υπόθεση, για όλα τα διανύσματα \mathbf{v} ισχύει $\mathbf{v}' \boldsymbol{\mu} = 0$. Η T^2

στατιστική συνάρτηση είναι το μέγιστο των συναρτήσεων που χρησιμοποιούνται για τον έλεγχο της υπόθεσης $H_0 : \mathbf{v}'\boldsymbol{\mu} = 0$ έναντι της $H_1 : \mathbf{v}'\boldsymbol{\mu} \neq 0$. Οπότε,

$$T^2 = \max_{\mathbf{v}} (\mathbf{v}'\mathbf{Z})^2 / (\mathbf{v}'\mathbf{v}).$$

Η μέγιστη τιμή μπορεί να επιτευχθεί σε κάθε πολλαπλάσιο του διανύσματος $\mathbf{v} = \mathbf{Z}$. Μία μεμονωμένη παρατήρηση Y_i είναι σημαντική, εάν το διάνυσμα που ελαχιστοποιεί το T^2 είναι στην κατεύθυνση του \mathbf{e}_i . Συνεπώς, αφού κλιμακωθούν τα διανύσματα στη μονάδα μήκους, ένα μέτρο συνεισφοράς της Y_i στο T^2 είναι

$$\frac{\mathbf{Z}'(\boldsymbol{\Sigma}^{-1/2}\mathbf{e}_i)}{\|\mathbf{Z}\| \|\boldsymbol{\Sigma}^{-1/2}\mathbf{e}_i\|} = \frac{\mathbf{e}_i'\boldsymbol{\Sigma}^{-1}\mathbf{Y}}{\sqrt{\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}}\sqrt{\mathbf{e}_i'\boldsymbol{\Sigma}^{-1}\mathbf{e}_i}}. \quad (3.5.5)$$

Επειδή, στον παραπάνω τύπο (3.5.5), η ποσότητα $\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}$ είναι μία σταθερά στο i , αυτό το μέτρο συνεισφοράς είναι ισοδύναμο με το μέτρο που λάβαμε παραπάνω.

Επειδή, η δομή συσχέτισης χρησιμοποιείται στον υπολογισμό του D_i , η μετρική D_i είναι αρκετά διαφορετική από τον υπολογισμό των t -ελέγχων (t -tests). Η μετρική αυτή είναι ισοδύναμη με έναν t -έλεγχο για την Y_i , στην ειδική περίπτωση που όλες οι συνιστώσες του \mathbf{Y} είναι ανεξάρτητες. Επειδή οι μεταβλητές της κύριας συνιστώσας είναι ανεξάρτητες, ο t -έλεγχος είναι κατάλληλος για τις μεταβλητές αυτές. Εάν μερικές τουλάχιστον από αυτές τις μεταβλητές είναι ερμηνεύσιμες, τότε οι Runger, Alt and Montgomery (1996) συστήνουν να υπολογιστεί η συνεισφορά κάθε μεταβλητής. Στην ειδική περίπτωση που μία ειδική αιτία προκαλεί ίσες μετατοπίσεις σε κάθε μεταβλητή, κανένα από τα D_i δεν είναι σχετικά μεγάλο και ως εκ τούτου όλες οι μεταβλητές συνεισφέρουν στο σήμα.

Θα εφαρμόσουμε τη διάσπαση των Runger, Alt and Montgomery (1996) στα δεδομένα που δημιουργήσαμε για την αντοχή στη θραύση και τη διάμετρο μιας υφαντικής ίνας, έχοντας αλλάξει το μέσο του πρώτου χαρακτηριστικού. Σκοπός μας είναι να διαπιστώσουμε εάν η T^2 διάσπαση θα αναγνωρίσει σωστά ποια από τις δύο μεταβλητές συνεισφέρει στο εκτός ελέγχου σήμα. Στον πίνακα που ακολουθεί, το T^2 διασπάται στις συνιστώσες που αναλογούν σε κάθε χαρακτηριστικό.

Από τον Πίνακα 3.6 παρατηρούμε ότι $D_1 = T^2 - T_1^2 = 13.3887$ και $D_2 = T^2 - T_2^2 = 13.4396$, όπου T_1^2 και T_2^2 είναι η τιμή του T^2 χωρίς να χρησιμοποιεί τις τιμές του πρώτου και του

δεύτερου χαρακτηριστικού, αντίστοιχα. Συνεπώς, η διάμετρος, η οποία είναι η κυρίαρχη μεταβλητή στο δείγμα 25, δε συνεισφέρει καθόλου στο εκτός ελέγχου σήμα, διότι $D_2 = T^2$. Αντίθετα, η αντοχή στη θραύση μειώνει αρκετά το T^2 από μόνη της. Το γεγονός αυτό υποδηλώνει ότι το πρώτο χαρακτηριστικό συνεισφέρει, ώστε το εικοστό πέμπτο δείγμα να βρεθεί εκτός των ορίων ελέγχου. Καταλήγουμε στο συμπέρασμα ότι η διάσπαση του T^2 ανίχνευσε σωστά το χαρακτηριστικό που μετατοπίστηκε.

T-Squared Decomposition

Relative Contribution to T-Squared Signal

<i>Subgroup</i>	<i>T-Squared</i>	<i>Tensile Strength_Shifted</i>	<i>Diameter</i>
25	13.4396	13.3887	13.4396

Πίνακας 3.6: Συνεισφορά κάθε χαρακτηριστικού στο εκτός ελέγχου σήμα του T^2

3.6.4 Διάγραμμα Ελέγχου *Minimax*, Διάγραμμα Επιλογής-Αιτίας και Προσαρμοσμένες Μεταβλητές Παλινδρόμησης

Οι Serúlveda and Nachlas (1997) παρουσίασαν ένα νέο πολυμεταβλητό διάγραμμα που ονόμασαν *Minimax*. Η ονομασία αυτή προέκυψε από γεγονός ότι το διάγραμμα ελέγχου παρακολουθεί το μέγιστο (*maximum*) τυποποιημένο δειγματικό μέσο και τον ελάχιστο (*minimum*) τυποποιημένο δειγματικό μέσο των δειγμάτων που επιλέγονται από μία πολυμεταβλητή διεργασία. Θεωρούμε ότι τα δεδομένα κατανέμονται κανονικά και ο πίνακας διακύμανσης-συνδιακύμανσης είναι γνωστός και σταθερός στην πάροδο του χρόνου. Από μία p -διάστατη διεργασία επιλέγονται δείγματα μεγέθους n . Υποθέτουμε ότι ο μέσος της διεργασίας μετατοπίζεται απότομα και ο νέος μέσος για την i μεταβλητή, που υπέστη την αλλαγή, είναι $\mu_{1i} = \mu_{0i} + \delta_i \cdot \sigma_{0i}$ και κατά συνέπεια $\delta_i = (\mu_{1i} - \mu_{0i}) / \sigma_{0i}$. Εάν $\delta_i = 0$, τότε η διεργασία είναι εντός ελέγχου. Σύμφωνα με το διάγραμμα *Minimax*, για να αποφασίσουμε πότε η διεργασία είναι εντός ελέγχου και πότε εκτός, ελέγχουμε σε κάθε δείγμα την υπόθεση $H_0 : \delta = \mathbf{0}$ έναντι της εναλλακτικής $H_1 : \delta \neq \mathbf{0}$.

Η βασική ιδέα πίσω από το διάγραμμα *Minimax* είναι να τυποποιήσουμε όλους τους p μέσους και παρακολουθούμε το μέγιστο $Z_{[p]}$ και το ελάχιστο $Z_{[1]}$ των τυποποιημένων δειγματικών μέσων. Οι ποσότητες αυτές υπολογίζονται αντίστοιχα ως εξής

$$\left. \begin{aligned} Z_{[p]} &= \max_i (Z_i) \\ Z_{[1]} &= \min_i (Z_i) \end{aligned} \right\}, i = 1, 2, \dots, p,$$

όπου

$$Z_i = \frac{\bar{X}_i - \mu_{0i}}{\sigma_{ii} / \sqrt{n}}.$$

Ως αποτέλεσμα, με το να παρακολουθούμε το μέγιστο και το ελάχιστο των τυποποιημένων δειγματικών μέσων, ένα εκτός ελέγχου σήμα είναι άμεσα συνδεδεμένο με την αντίστοιχη εκτός ελέγχου μεταβλητή. Οι Serúlveda and Nachlas (1997) μελέτησαν τις στατιστικές ιδιότητες και την επίδοση του ARL του διαγράμματος *Minimax*.

Οι Wade and Woodall (1993) ανέπτυξαν μία μέθοδο που αποτελείται από δύο εξαρτημένα βήματα. Το πρώτο βήμα της μεθόδου είναι να απεικονίσουμε τη μεταβλητή X_1 και το δεύτερο βήμα είναι να παρακολουθήσουμε την εξερχόμενη ποιότητα X_2 , μετά την προσαρμογή για την εισερχόμενη ποιότητα X_1 . Η μέθοδος αυτή χρησιμοποιεί τη σχέση που υπάρχει μεταξύ της X_2 και της X_1 , όπου ένα απλό μοντέλο παλινδρόμησης φαίνεται να είναι πολύ χρήσιμο. Πιο συγκεκριμένα, χρησιμοποιούμε ένα διάγραμμα για τη X_1 και ένα για τη $Z = X_2 - \hat{X}_2$, όπου \hat{X}_2 είναι η εκτίμηση του X_2 που βασίζεται στη γραμμή παλινδρόμησης. Για το λόγο αυτό, οι Z_i είναι ανεξάρτητες κανονικές τυχαίες μεταβλητές, όταν οι X_1 και X_2 μεταβλητές κατανομούνται κανονικά. Κατά συνέπεια, εάν στη διαδικασία υπάρχουν ειδικές (ή συστηματικές) ελεγχόμενες αιτίες, τότε η κατανομή των Z_i δεν είναι κανονική για κάποιες τιμές του i .

Όπως έχουμε ήδη αναφέρει, ο Hawkins (1991, 1993) όρισε ένα σύνολο προσαρμοσμένων μεταβλητών παλινδρόμησης, στο οποίο παλινδρόμησε κάθε μεταβλητή σε όλες τις υπόλοιπες. Επίσης, πρότεινε την απεικόνιση κάθε μεταβλητής Z_i σε ένα διάγραμμα τύπου *CUSUM*, διότι δεν είναι γνωστό ποια από τις p μεταβλητές είναι εκτός ελέγχου. Η στατιστική συνάρτηση που χρησιμοποιεί ο Hawkins και εμπλέκει p προσαρμοσμένες μεταβλητές, σχετίζεται με τη στατιστική συνάρτηση των Mason, Tracy and Young (1995).

Οι Kalagonda and Kulkarni (2003) παρουσίασαν μία διαγνωστική μέθοδο παλινδρόμησης χρησιμοποιώντας ψευδο-μεταβλητές (*dummy variables*), η οποία ανιχνεύει τους παράγοντες

(όπως ο μέσος ή/και σχέση της μετατόπισης) που είναι υπεύθυνοι για το εκτός ελέγχου σήμα και ταυτόχρονα προσδιορίζει την κατεύθυνση της μετατόπισης. Η μέθοδος αυτή ονομάζεται **τεχνική D** (*D-technique*).

Αρχικά, θα αναπτύξουμε την τεχνική D , όταν μετατοπίζεται ο μέσος. Υποθέτουμε ότι είναι διαθέσιμες m παρατηρήσεις του διανύσματος \mathbf{X} από μία εντός ελέγχου διεργασία και n παρατηρήσεις από την τρέχουσα διεργασία. Ορίζουμε την ψευδο-μεταβλητή D ως εξής

$$D = \begin{cases} 1, & \text{εάν μία παρατήρηση προέρχεται από την τρέχουσα διεργασία} \\ 0, & \text{εάν μία παρατήρηση προέρχεται από την εντός ελέγχου διεργασία.} \end{cases}$$

Αφού υπάρχουν m παρατηρήσεις από την εντός ελέγχου διεργασία, έχουμε m τιμές του $D=0$ και σε αντιστοιχία με τις n παρατηρήσεις της τρέχουσας διεργασίας, έχουμε n τιμές του $D=1$. Στην πραγματικότητα, το D προσδιορίζει την κατάσταση μιας παρατήρησης. Συμβολίζοντας την κεντραρισμένη στο μηδέν παρατήρηση X_j με x_j , τα δεδομένα από την εντός ελέγχου και την τρέχουσα διεργασία αναπαριστώνται, αντίστοιχα, ως εξής

$$(X_{ik}, 0, x_{jk}), \quad k=1, 2, \dots, m \quad \text{και} \quad (X_{ir}, 1, x_{jr}), \quad r=1, 2, \dots, n, \quad i \neq j=1, 2.$$

Τα σύνολα δεδομένων μπορούν να αναπαρασταθούν από κοινού ως εξής

$$(X_{iq}, D_q, x_{jq}), \quad q=1, 2, \dots, m+n, \quad i \neq j=1, 2.$$

Στη συνέχεια, προσαρμόζουμε τα από κοινού δεδομένα στο ακόλουθο ενιαίο πολλαπλό μοντέλο παλινδρόμησης του X_j στο x_j ($i \neq j=1, 2$)

$$X_i = \mu_i^0 + \tau_i \cdot D + \beta_{ij}^0 \cdot x_j + \varepsilon, \quad i \neq j=1, 2, \quad (3.6.6)$$

όπου τ_i είναι μία επιπλέον παράμετρος που θα συζητήσουμε παρακάτω. Για τις δύο διαφορετικές τιμές του D προκύπτουν δύο μοντέλα παλινδρόμησης, τα οποία είναι

$$\begin{cases} D=0: & X_i = \mu_i^0 + \beta_{ij}^0 \cdot x_j + \varepsilon, & i \neq j=1, 2 \\ D=1: & X_i = (\mu_i^0 + \tau_i) + \beta_{ij}^0 \cdot x_j + \varepsilon, & i \neq j=1, 2. \end{cases} \quad (3.6.7)$$

Ο όρος $\tau_i = \mu_i^c - \mu_i^0$ μετρά την τιμή της μετατόπισης του μέσου κατά τη διάρκεια των δύο περιόδων, όπου μ_i^c είναι ο μέσος της τρέχουσας διεργασίας και μ_i^0 ο μέσος της εντός ελέγχου διεργασίας. Όταν στον τύπο (3.6.6) το $\tau_i = 0$, παρατηρούμε ότι οι δύο εξισώσεις παλινδρόμησης του τύπου (3.6.7) είναι ίδιες και κατά συνέπεια η διεργασία παραμένει εντός

ελέγχου. Αντίθετα, εάν για κάποιο από τα i ισχύει $\tau_i \neq 0$, η διεργασία μετατοπίζεται σε ένα εκτός ελέγχου στάδιο και η αντίστοιχη (ή οι αντίστοιχες) μεταβλητές είναι υπεύθυνη (ή υπεύθυνες) για την κατάσταση αυτή της διεργασίας.

Επειδή, όμως, δε γνωρίζουμε εκ των προτέρων ποια μεταβλητή είναι υπεύθυνη για την εκτός ελέγχου κατάσταση της διεργασίας, είναι απαραίτητος ο έλεγχος της υπόθεσης $H_{0i} : \tau_i = 0, i = 1, 2$ ή ισοδύναμα της υπόθεσης $H_{0i} : \mu_i^c = \mu_i^0, i = 1, 2$. Η απαιτούμενη στατιστική συνάρτηση είναι η ακόλουθη

$$t = \frac{\hat{\tau}_i}{\sqrt{V(\hat{\tau}_i)}}, \quad i = 1, 2, \quad (3.6.8)$$

όπου $\hat{\tau}_i$ είναι η εκτίμηση ελαχίστων τετραγώνων του τ_i . Υπό τη μηδενική υπόθεση, ο τύπος (3.6.8) ακολουθεί την t -κατανομή με $m+n-4$ β.ε.. Κάθε μία από αυτές τις τιμές μπορεί να συγκριθεί με το ποσοστιαίο σημείο της κατανομής t_{m+n-4} για ένα δεδομένο επίπεδο σημαντικότητας α . Κάθε τιμή του τύπου (3.6.8), που είναι στατιστικά σημαντική, υποδηλώνει ότι ο συντελεστής τ_i διαφέρει από το μηδέν και κατά συνέπεια ο μέσος της αντίστοιχης μεταβλητής της τρέχουσας διεργασίας διαφέρει σημαντικά από τον εντός ελέγχου μέσο.

Στη συνέχεια, θα αναπτύξουμε την τεχνική D για την ανίχνευση μετατοπίσεων εξαιτίας των σχέσεων που αναπτύσσονται μεταξύ των μεταβλητών. Θεωρούμε ότι στο διάνυσμα \mathbf{X} υφίσταται αλλαγή μόνο η σχέση μεταξύ των μεταβλητών. Το μοντέλο παλινδρόμησης έχει την ακόλουθη μορφή

$$X_i = \mu_i^0 + \beta_{ij}^c \cdot x_j + \varepsilon, \quad i \neq j = 1, 2, \quad (3.6.9)$$

όπου ο συντελεστής β_{ij}^c είναι η κλίση της τρέχουσας διεργασίας. Για να ανιχνεύσουμε μία αλλαγή στη σχέση του X_i με το X_j , είναι απαραίτητος ο έλεγχος της υπόθεσης $H_0 : \beta_{ij}^c = \beta_{ij}^0, i \neq j = 1, 2$. Το κοινό σύνολο δεδομένων έχει τη μορφή

$$(X_{iq}, x_{jq}, D_q \cdot x_{jq}), \quad q = 1, 2, \dots, m+n, \quad i \neq j = 1, 2.$$

Το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης του X_j στο x_j και στο Dx_j για $i \neq j = 1, 2$ έχει τη μορφή

$$X_i = \mu_i^0 + \beta_{ij}^0 \cdot x_j + \gamma_{ij} \cdot (D \cdot x_j) + \varepsilon, \quad i \neq j = 1, 2. \quad (3.6.10)$$

Είναι φανερό ότι όταν $D = 0$, ο τύπος (3.6.10) παίρνει τη μορφή

$$X_i = \mu_i^0 + \beta_{ij}^0 \cdot x_j + \varepsilon, \quad i \neq j = 1, 2$$

και όταν $D = 1$, παίρνει τη μορφή του τύπου (3.6.9) με $\beta_{ij}^c = \beta_{ij}^0 + \gamma_{ij}$. Αντί να ελέγξουμε εάν $\beta_{ij}^c = \beta_{ij}^0$, μπορούμε ισοδύναμα να ελέγξουμε εάν $\gamma_{ij} = 0$. Η απαιτούμενη στατιστική συνάρτηση είναι η ακόλουθη

$$t = \frac{\hat{\gamma}_{ij}}{\sqrt{V(\hat{\gamma}_{ij})}}, \quad i = 1, 2. \quad (3.6.11)$$

Εάν η τιμή της ποσότητας (3.6.11) είναι στατιστικά σημαντική, τότε υπάρχει αλλαγή στη σχέση μεταξύ των μεταβλητών σε σύγκριση με το εντός ελέγχου στάδιο.

Τέλος, θα αναπτύξουμε την τεχνική D για την ανίχνευση των αιτιών που είναι υπεύθυνες για την εκτός ελέγχου κατάσταση. Θεωρούμε η μετατόπιση μιας τρέχουσας παρατήρησης \mathbf{X} επηρεάζει και το μέσο $\boldsymbol{\mu}$ και τον πίνακα διακύμανσης $\boldsymbol{\Sigma}$. Το μοντέλο παλινδρόμησης έχει την ακόλουθη μορφή

$$X_i = \mu_i^c + \beta_{ij}^c \cdot x_j + \varepsilon, \quad i \neq j = 1, 2. \quad (3.6.12)$$

Η διάγνωση της εκτός ελέγχου κατάστασης εξαρτάται από τη σύγκριση των αντίστοιχων εξισώσεων παλινδρόμησης όσον αφορά τόσο το συντελεστή τομής, όσο και το συντελεστή κλίσης. Μία τέτοια σύγκριση είναι ισοδύναμη με τον ταυτόχρονο έλεγχο της υπόθεσης

$$H_0 : \mu_i^c = \mu_i^0, \beta_{ij}^c = \beta_{ij}^0, \quad i \neq j = 1, 2.$$

Με τον τρόπο αυτό, μπορούμε να ανιχνεύσουμε μετατόπιση στο μέσο της i μεταβλητής και/ή αλλαγή της σχέσης των μεταβλητών (X_i με X_j).

Το κοινό σύνολο δεδομένων με την πρόσθετη μεταβλητή D έχει τη μορφή

$$(X_{iq}, D_q, x_{jq}, D_q \cdot x_{jq}), \quad q = 1, 2, \dots, m+n, \quad i \neq j = 1, 2.$$

Το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης είναι

$$X_i = \mu_i^0 + \tau_i D + \beta_{ij}^0 \cdot x_j + \gamma_{ij} \cdot (D \cdot x_j) + \varepsilon, \quad i \neq j = 1, 2. \quad (3.6.13)$$

Είναι φανερό ότι όταν $D = 0$, ο τύπος (3.6.13) παίρνει την ακόλουθη μορφή

$$X_i = \mu_i^0 + \beta_{ij}^0 \cdot x_j + \varepsilon, \quad i \neq j = 1, 2 \quad (3.6.14)$$

και όταν $D=1$, παίρνει τη μορφή

$$X_i = (\mu_i^0 + \tau_i) + (\beta_{ij}^0 + \gamma_{ij}) \cdot x_j + \varepsilon, \quad i \neq j = 1, 2,$$

ο οποίος είναι ο τύπος (3.6.12) για $\mu_i^c = \mu_i^0 + \tau_i$ και $\beta_{ij}^c = \beta_{ij}^0 + \gamma_{ij}$. Η παράμετρος γ_{ij} υποδηλώνει αλλαγή στη σχέση της μεταβλητής X_i με τη X_j , αφού $\gamma_{ij} = \beta_{ij}^c - \beta_{ij}^0$, $i \neq j = 1, 2$. Ισοδύναμα, θα μπορούσαμε να ελέγξουμε την υπόθεση

$$H_0 : \tau_i = 0; \gamma_{ij} = 0, \quad i \neq j = 1, 2.$$

Από τη στατιστική σημαντικότητα των εκτιμημένων διαφοροποιημένων παραμέτρων τομής και κλίσης, μπορούμε να διαπιστώσουμε εάν τα σύνολα των γραμμικών παλινδρομήσεων της εντός ελέγχου και της τρέχουσας κατάστασης διαφέρουν.

3.6.5 Ανάλυση Κυρίων Συνιστωσών

Η μέθοδος των κυρίων συνιστωσών (*Principal Components*) μπορεί να χρησιμοποιηθεί για να ερευνήσουμε ποια από τις p μεταβλητές είναι υπεύθυνη για την ενεργοποίηση ενός εκτός ελέγχου σήματος. Παρόλο που έχουν προταθεί διάφορες μέθοδοι, η πιο συνηθισμένη πρακτική, στην περίπτωση όπου τα διαγράμματα ελέγχου T^2 δίνουν ένα εκτός ελέγχου σήμα, είναι να γίνεται χρήση των πρώτων k πιο σημαντικών κύριων συνιστώσεων. Η βασική ιδέα είναι ότι οι πρώτες k κύριες συνιστώσες μπορούν να ερμηνευτούν εύκολα σε σχέση με τις αρχικές μεταβλητές. Για το λόγο αυτό, εάν το διάγραμμα T^2 δώσει ένα εκτός ελέγχου σήμα και, αν για παράδειγμα, το διάγραμμα της δεύτερης κύριας συνιστώσας δώσει επίσης εκτός ελέγχου σήμα, τότε από την ερμηνεία αυτής της συνιστώσας, μπορούμε να έχουμε μία εικόνα για το ποιες μεταβλητές είναι ύποπτες να βρίσκονται εκτός ελέγχου. Ως αποτέλεσμα, οι μεταβλητές μετατρέπονται σε ένα σύνολο ιδιοτήτων. Ωστόσο, το βασικό πρόβλημα που προκύπτει είναι ότι οι κύριες συνιστώσες δεν έχουν πάντα φυσική ερμηνεία.

Οι πρώτοι που πρότειναν τη χρήση κύριων συνιστωσών για την ερμηνεία ενός εκτός ελέγχου σήματος ήταν οι Jackson (1985) και Pignatiello and Runger (1990). Επιπλέον, υποστήριξαν ότι χρησιμοποιώντας μεμονωμένες μεταβλητές και κύριες συνιστώσες με μονοδιάστατα διαγράμματα ελέγχου, οι πληροφορίες σχετικά με την επίδραση της συσχέτισης των μεταβλητών δε χάνονται. Ο Jackson (1991) παρουσίασε μία μεθοδολογία για την παρακολούθηση μιας πολυμεταβλητής διεργασίας με τη χρήση της Ανάλυσης Κυρίων

Συνιστωσών. Σύμφωνα με τη μέθοδο αυτή, για κάθε διάνυσμα παρατηρήσεων, δημιουργούμε τα z -σκορ των κύριων συνιστώσεων και από αυτά υπολογίζουμε το T^2 . Εάν είναι εντός ελέγχου, συνεχίζουμε τη διαδικασία. Αντίθετα, εάν είναι εκτός ελέγχου, εξετάζουμε τα z -σκορ. Αφού οι κύριες συνιστώσες είναι ασυσχέτιστες, μπορούμε να έχουμε μία εικόνα της φύσης της εκτός ελέγχου κατάστασης, η οποία μας οδηγεί στην εξέταση συγκεκριμένων αρχικών παρατηρήσεων.

Οι Tracy, Young and Mason (1992) παρουσίασαν μία μέθοδο για δύο μεταβλητές, όπου η ερμηνεία των κύριων συνιστώσεων έχει νόημα, επεκτείνοντας την τεχνική κατάταξης των Doganaksoy, Faltin and Tucker (1991), καθώς και την τεχνική προσαρμογής της παλινδρόμησης του Hawkins (1991, 1993). Όταν παρακολουθούμε μία διεργασία με ζευγαρωτές παρατηρήσεις (*paired measurements*) σε ένα δείγμα, οι κύριες συνιστώσες του πίνακα συσχέτισης αναπαριστούν τα χαρακτηριστικά, τα οποία μας ενδιαφέρουν για τον έλεγχο της διεργασίας. Ο συντελεστής συσχέτισης των αρχικών μεταβλητών αποτελεί μία συμπληρωτική πληροφορία που περιγράφει την κατάσταση της διεργασίας.

Οι Kourti and MacGregor (1996) παρουσίασαν μία διαφορετική προσέγγιση που βασίζεται στην Ανάλυση Κύριων Συνιστωσών, σύμφωνα με την οποία το T^2 εκφράζεται με τη χρήση κανονικοποιημένων σκορ κύριων συνιστώσεων σε πολυμεταβλητές κανονικές μεταβλητές. Όταν λαμβάνεται ένα εκτός ελέγχου σήμα, τότε ανιχνεύεται το κανονικοποιημένο σκορ με τις μεγαλύτερες τιμές. Στη συνέχεια, χρησιμοποιούμε διαγράμματα συνεισφοράς¹ (*contribution plots*), τα οποία δείχνουν κατά πόσο μία μεταβλητή, που έχει εμπλακεί στον υπολογισμό αυτού του σκορ, συνεισφέρει στο σήμα. Με άλλα λόγια, διαπιστώσουμε ποιες μεταβλητές είναι υπεύθυνες για το εκτός ελέγχου σήμα.

Οι Maravelakis, Bersimis, Panaretos and Psarakis (2002) πρότειναν μία νέα μέθοδο που βασίζεται στην Ανάλυση Κύριων Συνιστωσών. Σύμφωνα με τη μέθοδο αυτή, θεωρούμε ότι $\mathbf{X}_i = (X_{1i}, X_{2i}, \dots, X_{pi})'$ είναι το i διάνυσμα παρατήρησης, που ακολουθεί την p -διάστατη κανονική κατανομή $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, όπου $\boldsymbol{\mu}_0$ είναι το γνωστό $(p \times 1)$ διάνυσμα των μέσων και $\boldsymbol{\Sigma}_0$ είναι ο γνωστός $(p \times p)$ πίνακας διακυμάνσεων-συνδιακυμάνσεων. Για τη διατήρηση της

¹ Περισσότερες πληροφορίες σχετικά με τη χρήση των διαγραμμάτων συνεισφοράς στον πολυμεταβλητό έλεγχο διεργασιών υπάρχουν στη δημοσίευση των Westerhuis, Gurden and Smilde (2000).

διεργασίας υπό έλεγχο, χρησιμοποιούμε το διάγραμμα ελέγχου χ^2 με στατιστική συνάρτηση την $\chi_i^2 = (\mathbf{X}_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_0)$. Όπως γνωρίζουμε, εάν η στατιστική συνάρτηση πάρει τιμές μεγαλύτερες από το $\chi_{p, 1-a}^2$, τότε λαμβάνουμε ένα εκτός ελέγχου σήμα. Το επόμενο βήμα είναι να ανιχνεύσουμε τη μεταβλητή εκείνη που προκάλεσε το πρόβλημα.

Η συνάρτηση $Z_k = u_{1k}X_1 + u_{2k}X_2 + \dots + u_{pk}X_p$ αποτελεί τη μορφή ενός τυπικού μοντέλου Ανάλυσης Κύριων Συνιστωσών, όπου Z_k είναι η k κύρια συνιστώσα, $(u_{1k}, u_{2k}, \dots, u_{pk})'$ είναι το αντίστοιχο k ιδιοδιάνυσμα και X_1, X_2, \dots, X_p είναι οι μεταβλητές της διεργασίας. Το σκορ του διανύσματος \mathbf{X}_i στην k κύρια συνιστώσα είναι $Y_{ki} = u_{1k}X_{1i} + \dots + u_{pk}X_{pi}$. Αφού υποθέτουμε ότι οι μεταβλητές της διεργασίας ακολουθούν πολυδιάστατη κανονική κατανομή, οι κύριες συνιστώσες κατανέμονται κανονικά.

Σκοπός της μεθοδολογίας είναι η χρήση της Ανάλυσης Κύριων Συνιστωσών, όταν δοθεί ένα σήμα στο διάγραμμα χ^2 , για την ανίχνευση της μεταβλητής (ή των μεταβλητών) που είναι υπεύθυνη (ή υπεύθυνες). Για το σκοπό αυτό, οι συγγραφείς ανέπτυξαν δύο διαφορετικές μεθοδολογίες, μία για την περίπτωση που ο πίνακας συσχέτισης έχει μόνο θετικές συσχετίσεις και μία για την περίπτωση που έχει και θετικές και αρνητικές συσχετίσεις.

Η συνήθης τακτική είναι η επιλογή $d \leq p$ στατιστικά σημαντικών κύριων συνιστώσεων. Ωστόσο, όταν ο πίνακας συσχέτισης έχει μόνο θετικές συσχετίσεις, η επιλογή των κύριων συνιστώσεων βασίζεται σε λόγους της μορφής

$$r_{ki} = \frac{(u_{k1} + u_{k2} + \dots + u_{kd}) \cdot X_{ki}}{Y_{1i} + Y_{2i} + \dots + Y_{di}}, \quad (3.6.15)$$

όπου X_{ki} είναι η i τιμή της μεταβλητής X_k , Y_{ji} , $j = 1, 2, \dots, d$, είναι το σκορ της i διανυσματικής παρατήρησης της j κύριας συνιστώσας. Ο αριθμητής του τύπου (3.6.15) παριστάνει το άθροισμα των συνεισφορών της μεταβλητής X_k των d πρώτων κύριων συνιστώσεων της i διανυσματικής παρατήρησης, ενώ ο παρονομαστής είναι το άθροισμα των σκορ της i διανυσματικής παρατήρησης στις d πρώτες κύριες συνιστώσες. Αφού υποθέσαμε ότι οι μεταβλητές ακολουθούν πολυδιάστατη κανονική κατανομή, τα r_{ki} είναι λόγοι δύο συσχετισμένων κανονικών μεταβλητών.

Το σκεπτικό της μεθόδου αυτής, είναι ο υπολογισμός της επίπτωσης που έχει κάθε μία από τις p μεταβλητές στο εκτός ελέγχου σήμα, χρησιμοποιώντας τη συνεισφορά τους στο ολικό σκορ. Προφανώς, η χρήση των d πρώτων κύριων συνιστώσεων έχει ως αποτέλεσμα να χάνεται ένα μέρος της πληροφορίας. Παρόλα αυτά, όταν υπάρχει μέτρια ή μεγάλη συσχέτιση μεταξύ των μεταβλητών, χρησιμοποιούμε ένα πολυμεταβλητό διάγραμμα. Υπό αυτές τις προϋποθέσεις, οι d πρώτες κύριες συνιστώσεις αποδίδουν το μεγαλύτερο μέρος της μεταβλητότητας της διεργασίας.

Για τον υπολογισμό της κατανομής του λόγου των δύο συσχετισμένων μεταβλητών θα βασιστούμε στο αναλυτικό αποτέλεσμα του Hinkley (1969). Αν X_1 και X_2 είναι τυχαίες μεταβλητές που προέρχονται από κανονικές κατανομές με μέσους μ_i , διακυμάνσεις σ_i^2 και συντελεστή συσχέτισης ρ , η συνάρτηση κατανομής του $R = X_1 / X_2$ δίνεται από τον τύπο

$$F(r) = L\left(\frac{\mu_1 - \mu_2 r}{\sigma_1 \sigma_2 a(r)}, -\frac{\mu_2}{\sigma_2}, \frac{\sigma_2 r - \rho \sigma_1}{\sigma_1 \sigma_2 a(r)}\right) + L\left(\frac{\mu_2 r - \mu_1}{\sigma_1 \sigma_2 a(r)}, \frac{\mu_2}{\sigma_2}, \frac{\sigma_2 r - \rho \sigma_1}{\sigma_1 \sigma_2 a(r)}\right), \quad (3.6.16)$$

όπου

$$L(h; k; \gamma) = \frac{1}{2\pi\sqrt{1-\gamma^2}} \int_h^\infty \int_k^\infty \exp\left(-\frac{x^2 - 2\gamma xy + y^2}{2(1-\gamma^2)}\right) dx dy.$$

Στην περίπτωση που όλες οι μεταβλητές είναι θετικά συσχετισμένες, ο Jackson (1991) υποστηρίζει ότι η πρώτη κύρια συνιστώσα είναι ο σταθμισμένος μέσος όλων των μεταβλητών. Συνεπώς, μπορούμε να χρησιμοποιήσουμε μόνο αυτή τη συνιστώσα στην εξαγωγή συμπερασμάτων.

Στην περίπτωση που ο πίνακας συσχέτισης έχει και θετικές και αρνητικές συσχετίσεις, η επιλογή των κύριων συνιστώσεων βασίζεται στους λόγους στις μορφές

$$r_{ki}^* = \frac{(u_{k1} + u_{k2} + \dots + u_{kd}) \cdot X_{ki}}{\bar{Y}_1 + \bar{Y}_2 + \dots + \bar{Y}_d}, \quad (3.6.17)$$

όπου X_{ki} είναι η i τιμή της μεταβλητής X_k , \bar{Y}_j , $j = 1, 2, \dots, d$, είναι το σκορ της j κύριας συνιστώσας, χρησιμοποιώντας στη θέση κάθε X_k , τις εντός ελέγχου τιμές τους. Επιπλέον, το d παριστάνει τον αριθμό των στατιστικά σημαντικών κύριων συνιστώσεων. Ο τύπος (3.6.17) παριστάνει το άθροισμα των συνεισφορών της μεταβλητής X_k των d πρώτων κύριων συνιστώσεων της i διανυσματικής παρατήρησης, διά το άθροισμα των εντός ελέγχου σκορ

των d πρώτων κύριων συνιστώσεων. Αφού ο παρανομαστής του παραπάνω τύπου είναι σταθερός, στην πραγματικότητα υπολογίζουμε την επίδραση κάθε μίας από τις p μεταβλητές στο εκτός ελέγχου σήμα. Όπως έχουμε ήδη αναφέρει, ο αριθμητής του τύπου (3.6.17) κατανέμεται κανονικά, ενώ ο παρανομαστής είναι μία σταθερά. Συνεπώς, τα r_{ki}^* ακολουθούν την κανονική κατανομή. Θα ήταν παράλειψη να μην αναφέρουμε ότι η μέθοδος αυτή δεν μπορεί να εφαρμοστεί, όταν έχουμε τυποποιημένες μεταβλητές, διότι μηδενίζεται ο παρανομαστής του τύπου (3.6.17).

Η μέθοδος των Maravelakis, Bersimis, Panaretos and Psarakis (2002) και για τις δύο περιπτώσεις συνοψίζεται στα ακόλουθα βήματα:

Βήμα 1^ο: Υπολογίζουμε τη στατιστική συνάρτηση χ^2 για την εισερχόμενη παρατήρηση.

Εάν το διάγραμμα δώσει σήμα εκτός ελέγχου, συνεχίζουμε στο επόμενο βήμα.

Βήμα 2^ο: Από τον τύπο (3.6.15) ή (3.6.17) υπολογίζουμε τους λόγους για όλες τις μεταβλητές, ανάλογα με το αν ο πίνακας συνδιακύμανσης έχει μόνο θετικές ή θετικές και αρνητικές συσχετίσεις, αντίστοιχα. Οι λόγοι για κάθε μεταβλητή είναι τόσοι, όσοι ο αριθμός των παρατηρήσεων από την αρχή της διεργασίας. Εάν η προτεινόμενη μέθοδος δε χρησιμοποιείται για πρώτη φορά, υπολογίζουμε τόσοι λόγους για κάθε μεταβλητή, όσοι και ο αριθμός των παρατηρήσεων από το τελευταίο εκτός ελέγχου σήμα έως το τελευταίο εκτός ελέγχου σήμα του πρώτου βήματος. Εναλλακτικά, υπολογίζουμε τους λόγους μόνο για την (τελευταία) παρατήρηση που προκάλεσε το εκτός ελέγχου σήμα.

Βήμα 3^ο: Απεικονίζουμε τους λόγους κάθε μεταβλητής σε ένα διάγραμμα. Στη συνέχεια, στην περίπτωση που ο πίνακας συσχέτισης έχει μόνο θετικές συσχετίσεις, υπολογίζουμε τα a και $1-a$ ποσοστιαία σημεία της κατανομής (3.6.16) με τις κατάλληλες παραμέτρους και τα χρησιμοποιούμε ως άνω όριο (UCL) και κάτω όριο ελέγχου (LCL), αντίστοιχα. Ενώ, στην περίπτωση που ο πίνακας συσχέτισης έχει θετικές και αρνητικές συσχετίσεις, υπολογίζουμε τα a και $1-a$ ποσοστιαία σημεία της κανονικής κατανομής και τα χρησιμοποιούμε ως όρια ελέγχου, με τον ίδιο τρόπο.

Βήμα 4^ο: Ελέγχουμε ποια μεταβλητή ή μεταβλητές προκάλεσαν το εκτός ελέγχου σήμα.

Βήμα 5^ο: Διορθώνουμε το πρόβλημα και συνεχίζουμε τη διαδικασία ξεκινώντας από το πρώτο βήμα.

3.6.6 Γραφικές Μέθοδοι

Οι Maravelakis, Bersimis, Panaretos and Psarakis (2002) πρότειναν και μία γραφική τεχνική, η οποία μπορεί να εφαρμοστεί για τα παραπάνω όρια ελέγχου. Επειδή τα διαγράμματα που αναφέρουν οι συγγραφείς στη μεθοδολογία τους είναι διαγράμματα τύπου Shewhart, μπορούν να ανιχνεύσουν γρήγορα μόνο τις μεγάλες μετατοπίσεις και δεν είναι αποτελεσματικά στις μικρές μετατοπίσεις. Για το λόγο αυτό, πρότειναν την απεικόνιση ενός διαγράμματος *CUSUM*. Για την ανίχνευση ανοδικών μετατοπίσεων, η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα έχει τη μορφή

$$S_n^+ = \max\left(0, S_{n-1}^+ + (X_{pn} + K)\right), \quad (3.6.18)$$

όπου $S_0^+ = 0$, X_{pn} είναι η n -οστή παρατήρηση της μεταβλητής X_p και K είναι η τιμή αναφοράς. Για την ανίχνευση καθοδικών μετατοπίσεων, η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα έχει τη μορφή

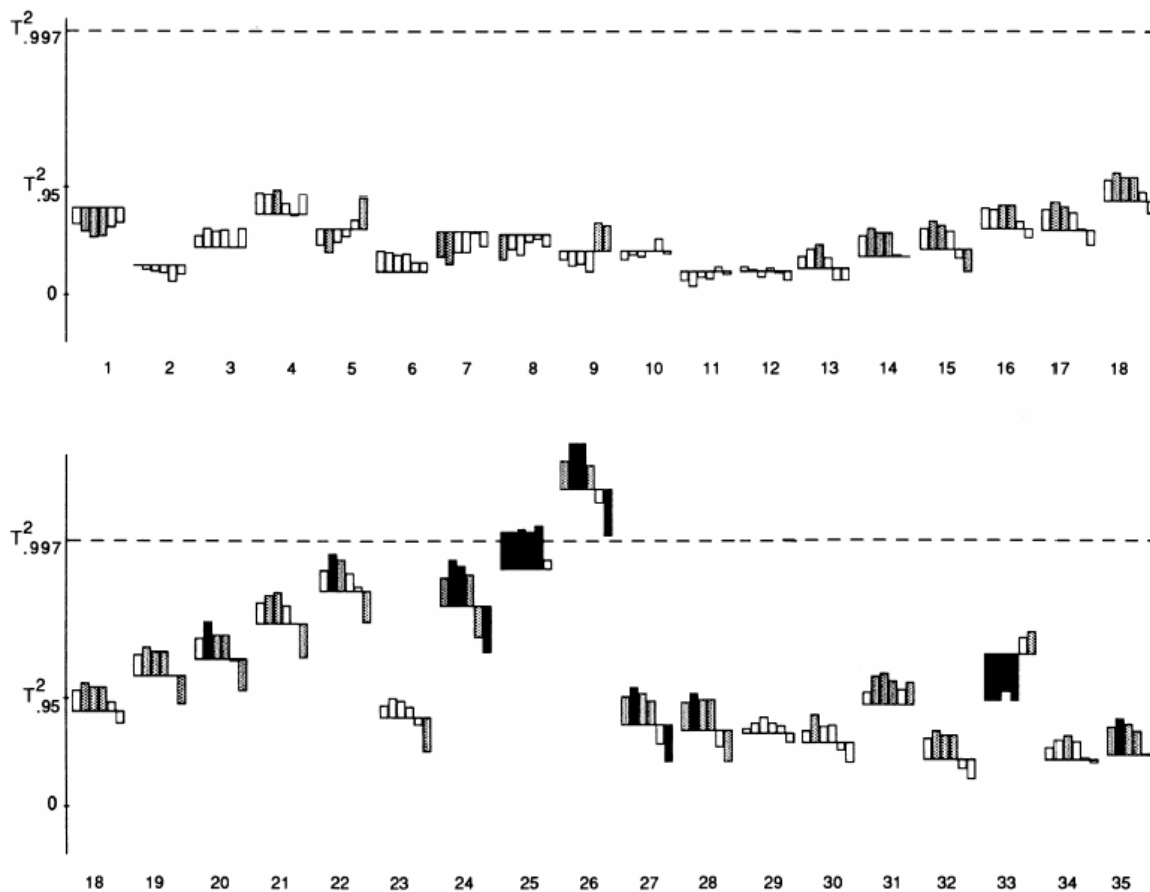
$$S_n^- = \max\left(0, S_{n-1}^- + (K - X_{pn})\right) \quad (3.6.19)$$

με $S_0^- = 0$.

Η γενική ιδέα των διαγραμμάτων *CUSUM* είναι η αξιολόγηση της βέλτιστης τιμής του K , η οποία βασίζεται στην υπόθεση κατανομής. Η τιμή αυτή, χρησιμοποιείται σε συνδυασμό με την τιμή του h , που είναι το όριο ελέγχου, με σκοπό το χαρακτηρισμό της επίδοσης του διαγράμματος βάσει του *ARL*. Ωστόσο, η εφαρμογή αυτής της θεωρίας στα διαγράμματα *CUSUM* είναι δύσκολη. Παρόλα αυτά, χρησιμοποιούμε τους τύπους (3.6.18) και (3.6.19) αποκλειστικά ως γραφικές τεχνικές. Επομένως, η μόνη άγνωστη ποσότητα είναι το K . Στον τύπο (3.6.15), στη θέση κάθε X_{ki} χρησιμοποιούμε την εντός ελέγχου τιμή, τόσο στο αριθμητή, όσο και στον παρονομαστή. Κάτι αντίστοιχο γίνεται και στην περίπτωση του τύπου (3.6.17), αλλά μόνο για τον αριθμητή.

Οι Fuchs and Benjamini (1994) παρουσίασαν το **πολυμεταβλητό διάγραμμα προφίλ** (*Multivariate Profile or MP chart*) για την πρακτικότητα, παρακολούθηση μιας διεργασίας και την ερμηνεία ενός εκτός ελέγχου σήματος, ταυτόχρονα. Στην πραγματικότητα, το διάγραμμα *MP* είναι ένα διάγραμμα διασποράς (*scatterplot*) με σύμβολα, τα οποία χρησιμοποιούνται για δεδομένα μεμονωμένων μεταβλητών. Η θέση των συμβόλων στο διάγραμμα παρέχει πληροφορίες σχετικά με το σύνολο των δεδομένων αυτών. Κάθε σύνολο παρατηρήσεων απεικονίζεται με ένα σύμβολο και αυτό το σύμβολο του διαγράμματος *MP*

επιτρέπει στο χρήστη να έχει μία ξεκάθαρη εικόνα για το μέγεθος και το πρόσημο κάθε μεταβλητής από τις τιμές αναφοράς τους.



Σχήμα 3.25: Διάγραμμα MP για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994))

Για την κατασκευή του διαγράμματος MP , φέρουμε μία οριζόντια γραμμή (*base line*) που θεωρείται η βάση κάθε συμβόλου και δημιουργούμε διαδοχικά μία ράβδο (*bar*) για κάθε μεταβλητή. Κάθε ράβδος μπορεί να απεικονίζεται είτε πάνω από τη βασική γραμμή, εάν η απόκλιση είναι θετική, είτε κάτω από τη γραμμή, εάν η απόκλιση είναι αρνητική. Το μέγεθος της ράβδου εξαρτάται από το μέγεθος της απόκλισης (*deviation*). Στη συνέχεια, θεωρούμε την ποσότητα $d_{ij} = (\bar{X}_{ij} - m_i) / v_i$, όπου $\bar{X}_{ij} = 1/n \sum_l X_{ijl}$, v_i είναι η παράμετρος κλίμακας (*scale factor*), m_i είναι η i τιμή αναφοράς (*reference value*) και X_{ijl} είναι η τιμή της l παρατήρησης της i μεταβλητής της j υποομάδας. Συνεπώς, το μέγεθος της ράβδου είναι ανάλογο του d_{ij} , το οποίο μπορεί να πάρει μέχρι την τιμή 4. Μετά την τιμή αυτή έχει

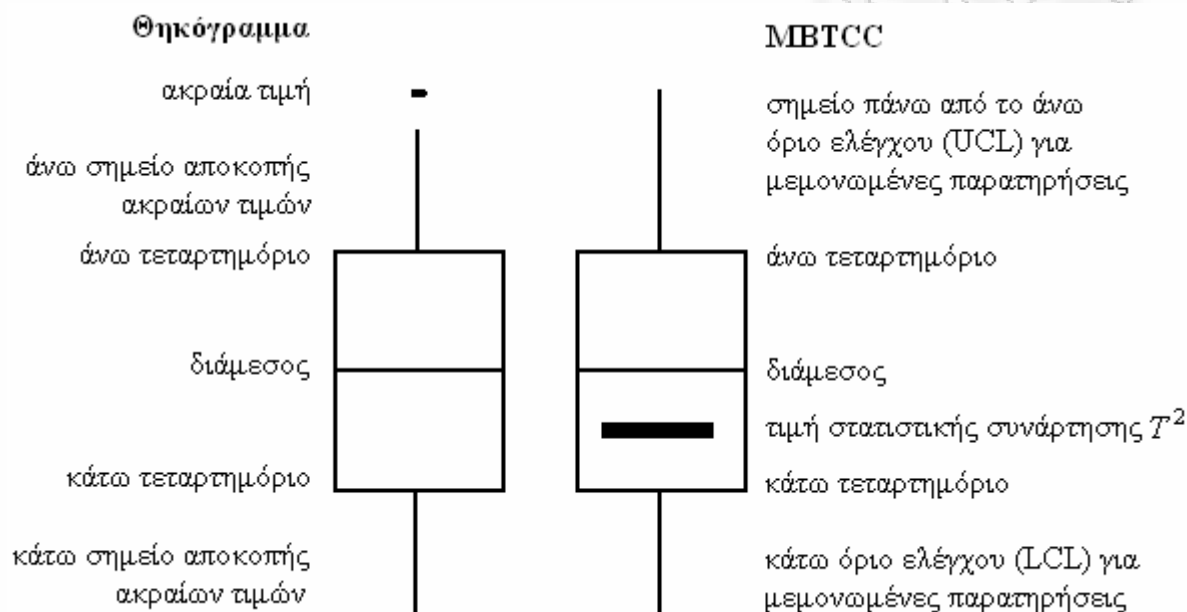
σταθερό μέγεθος (βλ. τις μεταβλητές 2, 3 και 6 του δείγματος 26 στο Σχήμα 3.25). Εάν η τυποποιημένη απόκλιση ξεπεράσει την τιμή 2, η αντίστοιχη ράβδος έχει χρώμα γκρι (βλ. τις μεταβλητές 2, 3 και 4 του πρώτου δείγματος του Σχήματος 3.25). Εάν η τυποποιημένη απόκλιση ξεπεράσει την τιμή 3, η αντίστοιχη ράβδος έχει χρώμα μαύρο (βλ. την τρίτη μεταβλητή του δείγματος 24 στο Σχήμα 3.25). Και εάν όλες οι μέσοι των μεταβλητών είναι ίσοι με τις πρότυπες τιμές τους, το σύμβολο είναι η βασική γραμμή. Για παράδειγμα, το δείγμα 12 βρίσκεται σχεδόν στη βασική γραμμή.

Τα σύμβολα απεικονίζονται σε χρονική σειρά στον οριζόντιο άξονα. Στον κάθετο άξονα, η θέση κάθε συμβόλου καθορίζεται από την πολυμεταβλητή απόκλιση του δειγματικού διανυσματικού μέσου από το τυπικό διάνυσμα μέσω των υπολογίζεται από το T^2 . Ο άξονας κλιμακώνεται έτσι, ώστε η κρίσιμη τιμή που αντιστοιχεί στην τιμή 0.997 (η οποία είναι ισοδύναμη με τον κανόνα των τριών τυπικών αποκλίσεων ενός μονοδιάστατου διαγράμματος) να είναι σε απόσταση μεγέθους ενός συμβόλου από την αρχή. Επιπλέον, σχηματίζεται μία διακεκομμένη γραμμή στο επίπεδο $T^2_{0.997}$. Η κρίσιμη τιμή που αντιστοιχεί στο μηδέν, τοποθετείται στη μισή απόσταση του συμβόλου από το κάτω μέρος. Η οριζόντια γραμμή ενός συμβόλου τοποθετείται κατάλληλα στον κατακόρυφο άξονα, όταν το T^2 είναι μικρότερο από την κρίσιμη τιμή. Αντίθετα, όλες οι παρατηρήσεις που έχουν T^2 μεγαλύτερο από την κρίσιμη τιμή, βρίσκονται στην κορυφή του διαγράμματος και τελείως πέρα από τη διακεκομμένη γραμμή.

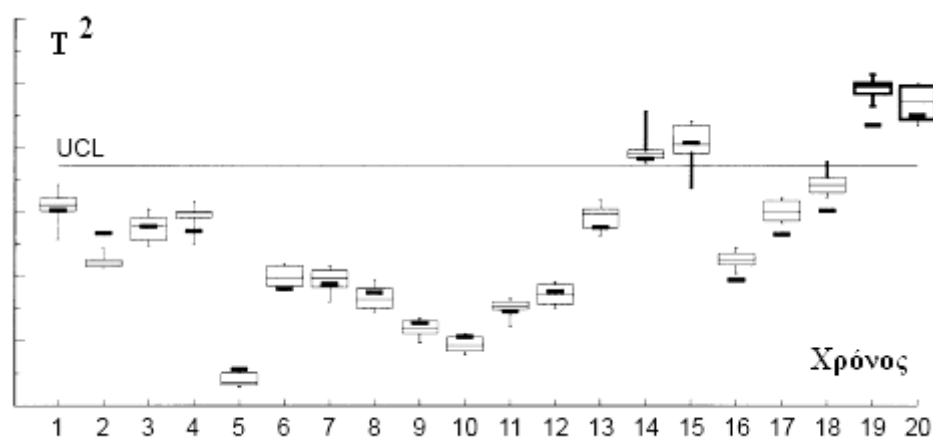
Εάν η διεργασία είναι εντός ελέγχου και στις πρότυπες τιμές, το διάγραμμα εμφανίζεται σαν μία απλή οριζόντια γραμμή, ενώ όταν η διεργασία είναι εκτός στατιστικού ελέγχου, το διάγραμμα MP δίνει τρία προειδοποιητικά οπτικά σημάδια. Το πρώτο είναι ότι το σύμβολο της ομάδας που προκαλεί ένα σήμα εκτός ελέγχου βρίσκεται πάνω από τη γραμμή των παρατηρήσεων που είναι εντός ελέγχου. Το δεύτερο σημάδι είναι ότι όσο μεγαλώνει το μέγεθος των αποκλίσεων, τόσο γίνεται και το μέγεθος του συμβόλου. Τέλος, το τρίτο σημάδι είναι ότι καθώς το σύμβολο γίνεται πιο σκούρο, χρίζει περισσότερης οπτικής επαφής. Αυτά τα προειδοποιητικά σημάδια ισχύουν όταν έχουμε ανοδικό ή καθοδικό σήμα ή ακόμη και στην περίπτωση που η μετατόπιση εμφανίζεται σταθερά σε μία κατεύθυνση.

Το διάγραμμα MP μπορεί εύκολα να χρησιμοποιηθεί για την ανίχνευση της αιτίας μιας μετατόπισης. Αφού όλες οι μεμονωμένες μεταβλητές εμφανίζονται σε μία κοινή κλίμακα μέσα σε μία ομάδα, το διάγραμμα MP παρέχει τη δυνατότητα οπτικής ανίχνευσης πιθανής αλλαγής στις αλληλεπιδράσεις τους. Επιπρόσθετα, όταν οι μεταβλητές αποκλίνουν από τις

αρχικές τιμές τους, πρέπει να εφιστήσουμε την προσοχή μας στη συμπεριφορά των μεμονωμένων μεταβλητών, στο ύψος και στο χρώμα των αντίστοιχων ράβδων. Θα πρέπει να αναφέρουμε ότι οι Fuchs and Kenett (1998) ανέπτυξαν εντολές στο στατιστικό πακέτο *Minitab* για τη δημιουργία διαγραμμάτων *MP* με σκοπό την παρακαλούθηση τόσο του μέσου όσο και της διασποράς της διεργασίας.

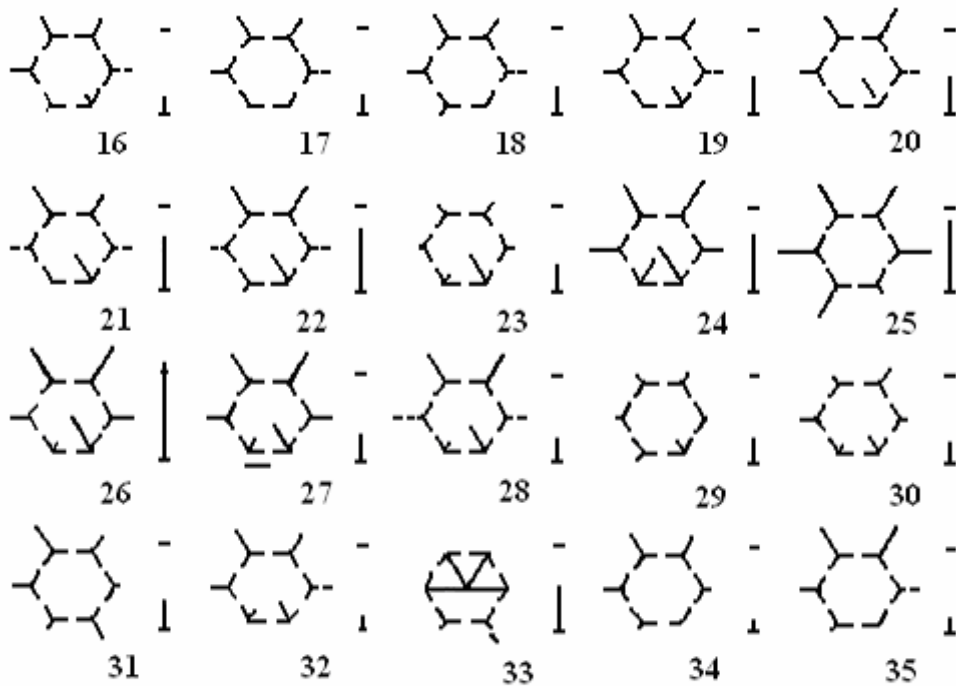


Σχήμα 3.26: Πληροφορίες που περιέχονται σε κάθε θηκόγραμμα τη χρονική στιγμή t (Atienza, Tang and Ang (1998))

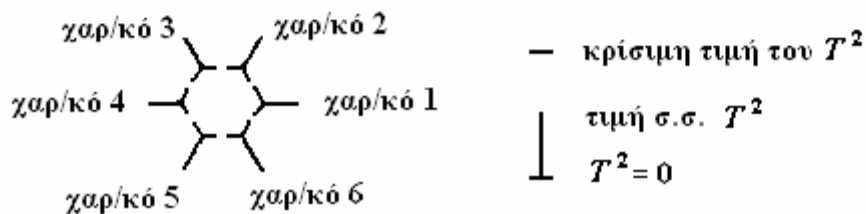


Σχήμα 3.27: Διάγραμμα ελέγχου *MBTCC* (Atienza, Tang and Ang (1998))

Η αριθμητική τιμή του T^2 δεν παρέχει αρκετές πληροφορίες, όπως στην περίπτωση των μονοδιάστατων διαγραμμάτων ελέγχου που οι αποκλίσεις από τις τιμές αναφοράς μετριοούνται σε τυπικά σφάλματα. Για το λόγο αυτό, οι Fuchs and Benjamini (1994) παρουσίασαν σε οριζόντια κλίμακα του διαγράμματος MP τις αντίστοιχες πιθανότητες ουρών αντί των πραγματικών τιμών του T^2 . Οι Atienza, Tang and Ang (1998) πρότειναν τη χρήση θηκόγραμμάτων (*boxplot*) αντί της χρήσης ραβδογραμμάτων (*bar graph*). Ονόμασαν το διάγραμμα αυτό **πολυμεταβλητό διάγραμμα ελέγχου θηκόγραμμα- T^2** (*multivariate boxplot- T^2 control chart or MBTCC*). Οι πληροφορίες που περιέχονται σε κάθε θηκόγραμμα τη χρονική στιγμή t παρουσιάζονται στο Σχήμα 3.26, ενώ η τυπική μορφή του διαγράμματος *MBTCC* δίνεται στο Σχήμα 3.27.



ΥΠΟΜΝΗΜΑ



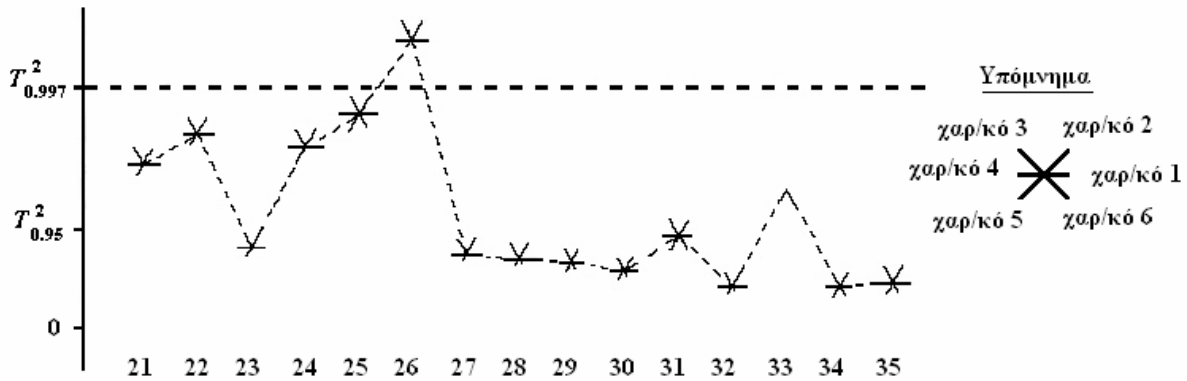
Σχήμα 3.28: Πολυδιάγραμμα για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994))

Η χρήση γραφικών μεθόδων στην ανίχνευση της αιτίας ενεργοποίησης συναγερμού σε μία πολυμεταβλητή διεργασία είχε προταθεί νωρίτερα από τους Blazek, Novic and Scott (1987), οι οποίοι ανέπτυξαν μία μέθοδο που βασίζεται σε ένα **πολυδιάγραμμα** (*polyplot*) και είναι αρκετά χρήσιμη στον προσδιορισμό των εκτός ελέγχου μεταβλητών. Το μήκος του εύρους ενός πολυδιαγράμματος παρέχει πληροφορίες για την τιμή κάθε αρχικής μεταβλητής που σχετίζεται με την τιμή στόχο. Η συνήθης τακτική είναι η χρήση της στατιστικής συνάρτησης T^2 σε ένα πολυδιάγραμμα, ωστόσο οι Lowry and Montgomery (1995) θεωρούν ότι η χρήση της πολυμεταβλητής εκθετικά ζυγισμένης στατιστικής συνάρτησης κινούμενου μέσου (*MEWMA*) θα ήταν πιο αποτελεσματική. Η τυπική μορφή ενός πολυδιαγράμματος δίνεται στο Σχήμα 3.28.

Σε ένα πολυδιάγραμμα, κάθε χαρακτηριστικό συμβολίζεται με μία κατακόρυφη γραμμή, η οποία αναπαριστά την τιμή του T^2 , και ένα πολύγωνο. Η κρίσιμη τιμή (*critical value*) της στατιστικής συνάρτησης T^2 συμβολίζεται με ένα οριζόντιο σημάδι πάνω στη γραμμή του T^2 . Κάθε ακτίνα που φέρεται από την κορυφή του πολυγώνου αναπαριστά τις τυποποιημένες τιμές κάθε χαρακτηριστικού, οι οποίες υπολογίζονται από τον τύπο $z_{jk} = (X_{jk} - \mu_j^0) / \sigma_j$ για $j = 1, 2, \dots, p$ και $k = 1, 2, \dots, m$, όπου X_{jk} είναι ο μέσος ή η μεμονωμένη παρατήρηση του j χαρακτηριστικού του k δείγματος, μ_j^0 είναι η τιμή αναφοράς (*reference value*) και σ_j είναι ο συντελεστής κλίμακας που αντιστοιχεί στο j στοιχείο του διαγώνιου πίνακα Σ . Τιμές μικρότερες από το μ_j^0 αναπαρίστανται με ακτίνες που ξεκινούν από την κορυφή και συνεχίζουν προς το κέντρο του πολυγώνου, ενώ τιμές μεγαλύτερες από το μ_j^0 αναπαρίστανται με ακτίνες που ξεκινούν από την κορυφή και συνεχίζουν έξω από το πολύγωνο. Συχνά, η εκτός ελέγχου κατάσταση υποδηλώνεται με τη χρήση διαφορετικού είδους ή χρώματος ακτινών.

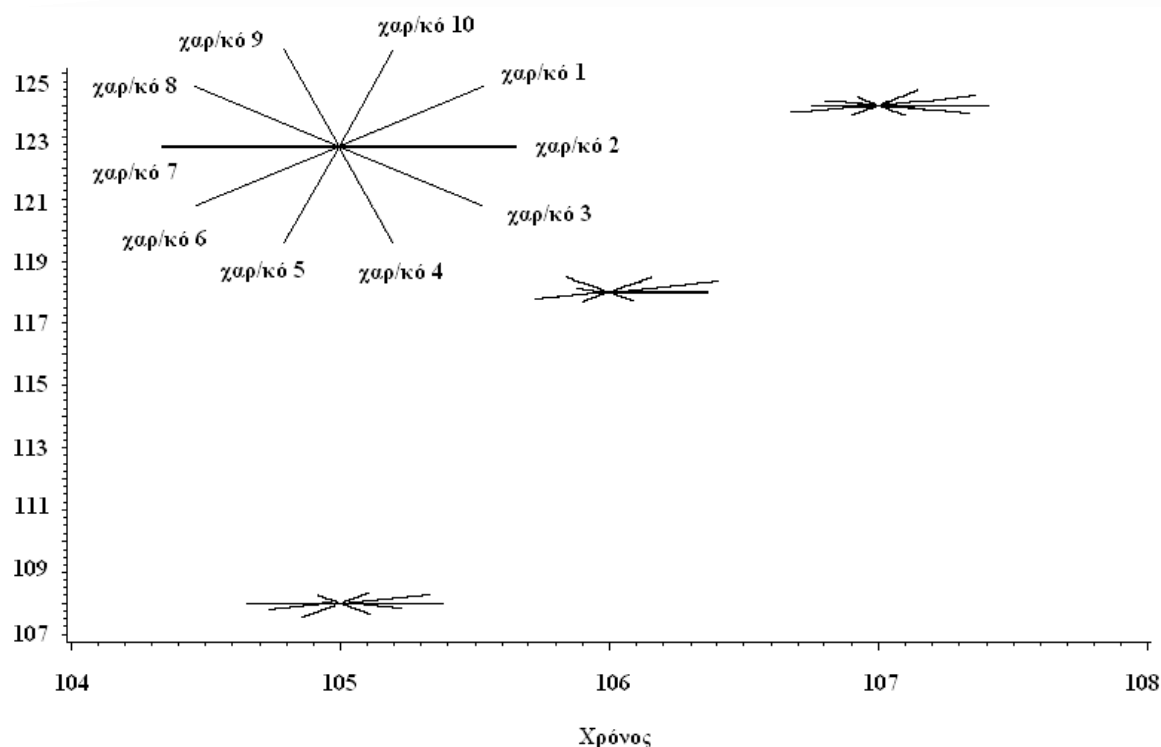
Ένα διάγραμμα, αντίστοιχο του πολυδιαγράμματος, αναπτύχθηκε πρώτη φορά από την εταιρεία SCS και μπορεί να δημιουργηθεί μέσω του στατιστικού πακέτου *Statgraphics*. Στο διάγραμμα αυτό απεικονίζονται οι τιμές του T^2 , όχι με σημεία αλλά με αστέρια. Για το λόγο αυτό το διάγραμμα ονομάζεται **starplot**. Κάθε πολυμεταβλητή παρατήρηση συμβολίζεται με ένα αστέρι, του οποίου οι ακτίνες είναι ανάλογες των αποκλίσεων των μεμονωμένων παρατηρήσεων από τις ελάχιστες τιμές όλων των παρατηρήσεων (ή των συνόλων των παρατηρήσεων). Εάν η διεργασία είναι εντός ελέγχου, τα αστέρια έχουν μεσαίο μέγεθος, ενώ

εάν μία μεταβλητή είναι εκτός ελέγχου με τιμή αρκετά χαμηλότερη από την τιμή στόχο, οι ακτίνες συρρικνώνονται στην ελάχιστη τιμή με σκοπό να τραβήξει οπτικά την προσοχή. Ωστόσο, το starplot δε δίνει ένδειξη στατιστικής σημαντικότητας του μήκους της ακτίνας. Η τυπική μορφή ενός starplot δίνεται στο Σχήμα 3.29.



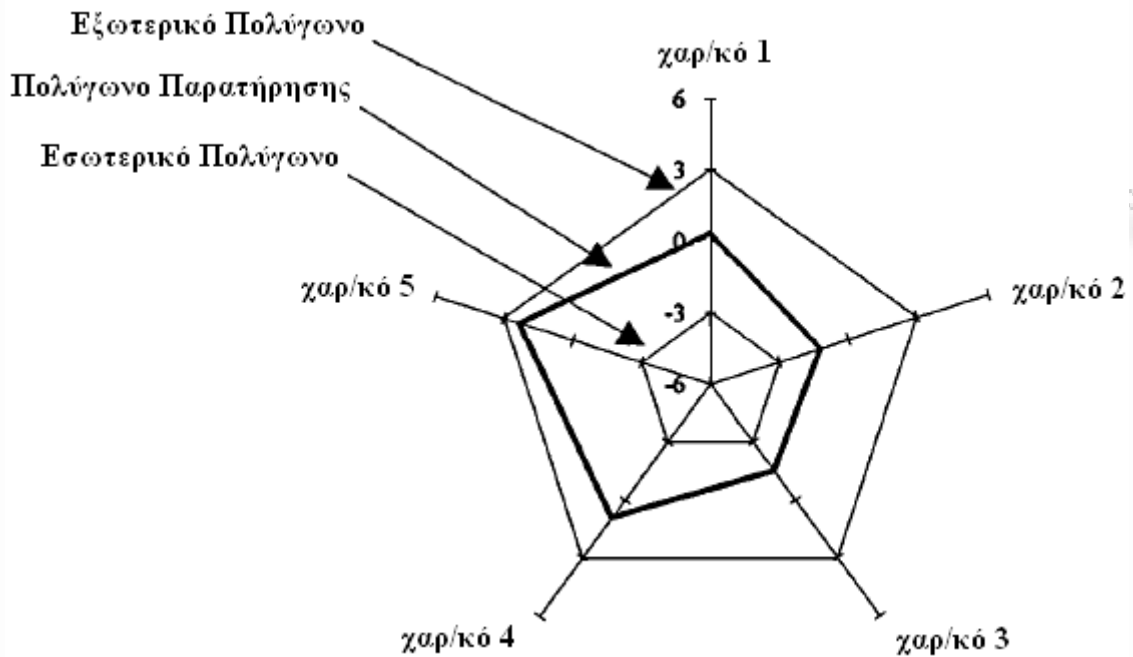
Σχήμα 3.29: StarPlot για έξι ποιοτικά χαρακτηριστικά (Fuchs and Benjamini (1994))

Οι Sparks, Adolphson and Phatak (1997) παρουσίασαν μία μέθοδο για την παρακολούθηση πολυμεταβλητών δεδομένων που βασίζεται στη γραφική αναπαράσταση δύο διαστάσεων του Gabriel (*Gabriel Biplot*). Η συγκεκριμένη μέθοδος μειώνει τις διαστάσεις σε δύο για τον προσδιορισμό της κατάστασης της διεργασίας. Παρόλο που η μέθοδος αυτή μειώνει τις διαστάσεις, χρησιμοποιεί όλα τα δεδομένα για να αποφασίσουμε εάν είμαστε εντός ή εκτός ελέγχου. Επίσης, μπορεί να χρησιμοποιηθεί ως διάγραμμα ελέγχου και στην περίπτωση που η διεργασία είναι εκτός στατιστικού ελέγχου ανιχνεύει την αιτία που προκάλεσε το πρόβλημα αυτό. Επιπλέον, η μέθοδος των Sparks, Adolphson and Phatak ανιχνεύει με ακρίβεια αλλαγές στην τοποθεσία, στη διακύμανση και στη δομή συσχέτισης, εμφανίζοντας συνοπτικά μεγάλο μέρος πληροφορίας. Οι Nottingham, Cook and Zobel (2001) ανέπτυξαν **ακτινωτά διαγράμματα** (*radial plots*) που ενισχύουν τη διαδικασία παρακολούθησης, ανάλυσης και ελέγχου μιας διεργασίας και συνάμα δημιούργησαν κώδικα στο στατιστικό πακέτο *SAS* για την κατασκευή των διαγραμμάτων αυτών. Τα ακτινωτά διαγράμματα κατασκευάζονται με την ίδια λογική που κατασκευάζονται και τα starplots. Ένα τυπικό διάγραμμα ακτίνων παρουσιάζεται στο Σχήμα 3.30 που ακολουθεί.

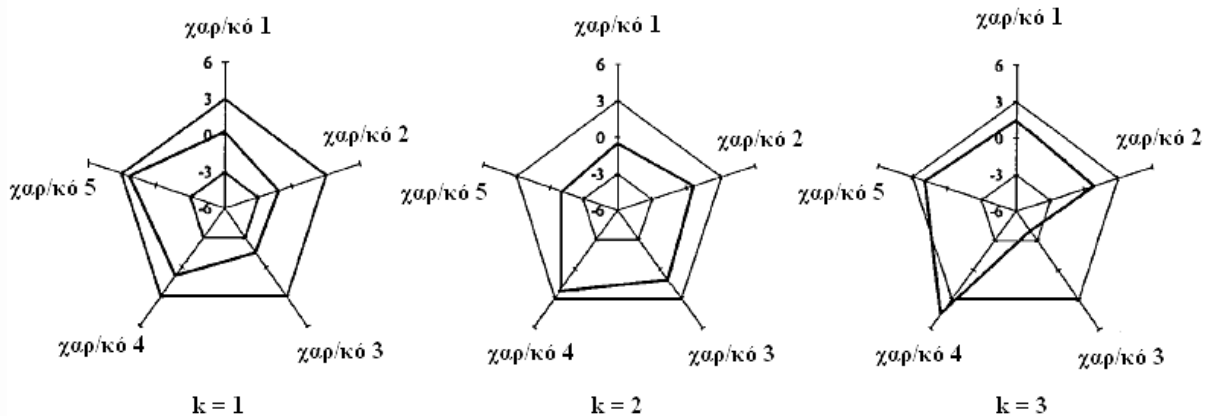


Σχήμα 3.30: Διάγραμμα ακτίνων για δέκα ποιοτικά χαρακτηριστικά
(Nottingham, Cook and Zobel (2001))

Ο Adams (1994) εισήγαγε το **πολυμεταβλητό διάγραμμα ιστού** (*multivariate control web*), το οποίο αποτελεί μία γραφική προσέγγιση που εμφανίζει ταυτόχρονα μονομεταβλητές και πολυμεταβλητές στατιστικές συναρτήσεις. Στην πραγματικότητα, το διάγραμμα αυτό προσαρμόζει ένα αστέρι με σκοπό την αναπαράσταση πολυμεταβλητών δεδομένων και απεικονίζει τα πολυμεταβλητά σημεία ως πολύγωνα (*polygon*). Ένα τυπικό διάγραμμα ιστού απεικονίζεται στο Σχήμα 3.31. Η διεργασία τίθεται εκτός ελέγχου, όταν μία πολυδιάστατη παρατήρηση σχηματίζει ένα πολύγωνο που δεν περιέχεται ολοκληρωτικά στο εξωτερικό πολύγωνο ελέγχου ή δεν περιέχει ολοκληρωτικά το εσωτερικό πολύγωνο ελέγχου (βλ. το τρίτο γράφημα του Σχήματος 3.32). Γενικά, η τεχνική αυτή έχει πολλά πλεονεκτήματα, όπως για παράδειγμα τη δυνατότητα εύκολης ενσωμάτωσης κανόνων ροής στη διαδικασία. Επιπλέον, το διάγραμμα μπορεί να ερμηνευτεί και να κατασκευαστεί εύκολα.



Σχήμα 3.31: Πολυμεταβλητό διάγραμμα ιστού πέντε χαρακτηριστικών (Adams (1994))



Σχήμα 3.32: Πολυμεταβλητό διάγραμμα ιστού πέντε χαρακτηριστικών για τρεις χρονικές περιόδους (Adams (1994))

Τα τελευταία χρόνια παρατηρείται εντατική **χρήση των τεχνητών νευρωνικών δικτύων** (*artificial neural networks*) ως αποτελεσματικό εργαλείο στην ανίχνευση των εκτός ελέγχου σημάτων στα πολυμεταβλητά διαγράμματα ελέγχου. Η τεχνική αυτή επιτρέπει την αυτοματοποίηση της διεργασίας, η οποία εντοπίζει τις μεταβλητές που άλλαξαν κατά τη

διάρκεια της παραγωγικής διαδικασίας, χρησιμοποιώντας νευρωνικά δίκτυα. Πιο συγκεκριμένα, επιλέγονται δείγματα από τη διεργασία με σκοπό τον έλεγχο της. Αυτό μπορεί να επιτευχθεί με τη χρήση ενός πολυμεταβλητού διαγράμματος ελέγχου, συνήθως του T^2 διαγράμματος του Hotelling. Εάν δοθεί εκτός ελέγχου σήμα, χρησιμοποιείται νευρωνικό δίκτυο για την αναγνώριση των μεταβλητών που μετατοπίστηκαν. Η εφαρμογή της παραπάνω τεχνικής έχει αναπτυχθεί από τους: Cheng (1995, 1997), Chang and Aw (1996), Cook and Chiu (1998), Zooriassatine and Tannock (1998), Guh and Hsieh (1999), Guh and Tannock (1999a, 1999b), Ho and Chang (1999), Cook, Zobel and Nottigham (2001), Guh (2003, 2007), Noorossana, Farrokhi and Saghaei (2003), Niaki and Abassi (2005), Aparisi, Avendaño and Sanz (2006), Shao, Wu, Ho and Hsu (2008) και Yu and Xi (2009). Τέλος, οι Alfaro, Alfaro, Gámez and García (2009) πρότειναν την εφαρμογή **ενισχυμένων δένδρων** (*boosting trees*) για την ανίχνευση της μεταβλητής ή των μεταβλητών που ενεργοποίησαν το συναγερμό σε μία πολυμεταβλητή διεργασία.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΑΙΑ

ΚΕΦΑΛΑΙΟ 4

Μη Παραμετρικά Διαγράμματα Ελέγχου

4.1 Εισαγωγή

Σύμφωνα με τους Chakraborti, Van Der Laan and Bakir (2001), Chakraborti and Graham (2008) και Chakraborti (2011), σκοπός του Στατιστικού Ελέγχου Ποιότητας (*Statistical Process Control, SPC*) ή Διεργασιών (*Statistical Quality Control, SQC*) είναι η βελτίωση της ποιότητας των παραγόμενων προϊόντων. Όπως έχουμε αναφέρει, αυτό μπορεί να επιτευχθεί με τον εντοπισμό και τη μείωση των ειδικών ή συστηματικών αιτιών μεταβλητότητας (*special or assignable causes of variation*), στο βαθμό που κάτι τέτοιο είναι εφικτό. Στον Παραμετρικό Στατιστικό Έλεγχο Ποιότητας ή Διεργασιών (*Parametric SPC or SQC*), η βασική υπόθεση που γίνεται είναι ότι οι ειδικές αιτίες μεταβλητότητας ακολουθούν κάποια παραμετρική κατανομή, όπως για παράδειγμα την κανονική κατανομή. Συνεπώς, η επιλογή της στατιστικής συνάρτησης και ο υπολογισμός των ορίων ελέγχου εξαρτάται από τη μορφή και τις ιδιότητες της κατανομής αυτής, οι οποίες θεωρούνται «ακριβείς» (“*exact*”) μόνο όταν η βασική υπόθεση ικανοποιείται.

Παρόλο που η θεωρία του Παραμετρικού Ελέγχου Ποιότητας χρησιμοποιείται στη βιομηχανία, στην πραγματικότητα η κατανομή των αιτιών μεταβλητότητας είτε δεν είναι γνωστή είτε είναι πολύ μακριά από την κανονική. Για το λόγο αυτό, καθίσταται αναγκαία η κατασκευή και η χρήση πιο ευέλικτων (*flexible*) και ανθεκτικών (*robust*) διαγραμμάτων, τα οποία δεν απαιτούν υπόθεση κανονικότητας ή κάποιας άλλης γνωστής κατανομής. Κατά συνέπεια, τα **μη παραμετρικά** ή **ελεύθερα-κατανομής** διαγράμματα ελέγχου (*nonparametric or distribution-free control charts*) αποτελούν χρήσιμες εναλλακτικές των παραμετρικών. Ο όρος «μη παραμετρικά» υποδηλώνει ότι δεν εμπλέκονται παράμετροι στη διεργασία, κάτι το

οποίο δεν ισχύει αφού απλώς τα συγκεκριμένα διαγράμματα ελέγχου δε χρησιμοποιούν τις παραμέτρους (έστω και αν υπάρχουν). Ως εκ τούτου, ο όρος «ελεύθερα-κατανομής» φαίνεται να περιγράφει καλύτερα τι μπορούν να πραγματοποιήσουν τα διαγράμματα αυτά. Ωστόσο, ο πρώτος όρος είναι αυτός που χρησιμοποιείται πιο συχνά.

Τα παραμετρικά διαγράμματα ελέγχου αναπτύσσονται με κατάλληλη προσαρμογή γνωστών συνήθως στατιστικών διαδικασιών, όπως ο έλεγχος υποθέσεων ή τα διαστήματα εμπιστοσύνης. Τα τελευταία χρόνια, ένας μεγάλος αριθμός από αντίστοιχες μεθόδους μη παραμετρικών ελέγχων και διαστημάτων εμπιστοσύνης είναι διαθέσιμο στη διεθνή βιβλιογραφία. Έχει αποδειχθεί ότι αυτές οι μέθοδοι συμπεριφέρονται εξίσου καλά σε σχέση με τις αντίστοιχες παραμετρικές. Συγκεκριμένα, όταν η υποκείμενη κατανομή είναι η κανονική, η απόδοση μερικών μη παραμετρικών ελέγχων ανάλογων των (βέλτιστων) κανονικών ελέγχων μπορεί να φτάσει σε επίπεδο 95.5% (Gibbons and Chakraborti (2003)). Στην πραγματικότητα, κάποιοι μη παραμετρικοί έλεγχοι είναι πιο αποτελεσματικοί-αποδοτικοί, στην περίπτωση των κατανομών με παχιές ουρές, όπως για παράδειγμα της διπλής εκθετικής. Θεωρητικά, οι μη παραμετρικές μέθοδοι είναι λιγότερο αποδοτικές από τις αντίστοιχες παραμετρικές, όταν υπάρχει πλήρης γνώση των κατανομών για τις οποίες οι παραμετρικές μέθοδοι έχουν ειδικά σχεδιαστεί. Στην πράξη, όμως, η γνώση αυτή σπάνια υπάρχει ή είναι διαθέσιμη. Παρόλα αυτά, η αποτελεσματικότητα ενός διαγράμματος συνήθως μπορεί εύκολα να βελτιωθεί με την επεξεργασία περισσότερων παρατηρήσεων. Ένα ακόμη μειονέκτημα των μη παραμετρικών διαγραμμάτων είναι ότι για μικρά μεγέθη δείγματος, κρίνεται απαραίτητη η χρήση ειδικών πινάκων, γεγονός που σήμερα δεν αποτελεί πρόβλημα δεδομένης της παρουσίας των υπολογιστών.

Σύμφωνα με τους Chakraborti, Van Der Laan and Bakir (2001), το βασικό πλεονέκτημα των μη παραμετρικών διαγραμμάτων ελέγχου είναι η ευκολία και η ευελιξία της χρήσης τους, καθώς δε χρειάζεται να υποθέσουμε ότι η υποκείμενη διεργασία ακολουθεί κάποια συγκεκριμένη κατανομή. Επιπλέον, το εντός ελέγχου μήκος ροής (ARL_0) της κατανομής είναι το ίδιο για όλες τις συνεχείς κατανομές. Το ίδιο ισχύει και για το ποσοστό λανθασμένων συναγερμών (*False Alarm Rate, FAR*) και για το λόγο αυτό διαφορετικά μη παραμετρικά διαγράμματα ελέγχου μπορούν να συγκριθούν πιο εύκολα μεταξύ τους. Επίσης, τα μη παραμετρικά διαγράμματα είναι πιθανό να διατηρούν την ιδιότητα της ευρωστίας (*robustness property*) των μη παραμετρικών ελέγχων και διαστημάτων εμπιστοσύνης και ως αποτέλεσμα

είναι πολύ πιο πιθανό να επηρεάζονται λιγότερο από ακραίες τιμές (*outliers*). Ένα ακόμη πλεονέκτημα των διαγραμμάτων αυτών είναι η αποδοτικότητά τους στην ανίχνευση αλλαγών, όταν η πραγματική κατανομή δεν είναι κανονική και κυρίως όταν έχει παχιές ουρές. Τέλος, δεν απαιτείται εκτίμηση της διακύμανσης στην κατασκευή διαγραμμάτων για την παράμετρο θέσης. Είναι γεγονός ότι παρά τα πλεονεκτήματά τους, τα μη παραμετρικά διαγράμματα ελέγχου και οι παραμετρικές τεχνικές δεν είναι ευρέως γνωστά σε αρκετούς ερευνητές.

Στον Παραμετρικό Στατιστικό Έλεγχο Ποιότητας ενδιαφερόμαστε για την κατασκευή διαγραμμάτων για την ανίχνευση αλλαγών στο μέσο της διεργασίας (*process mean*) και/ή στη μεταβλητότητα της διεργασίας (*process variation*), όπως το εύρος (*range*), η διακύμανση (*variance*) και η τυπική απόκλιση (*standard deviation*). Στον Μη Παραμετρικό Έλεγχο, μας ενδιαφέρει η παρακολούθηση του κέντρου (*center*) ή της θέσης (*location*) και/ή της παραμέτρου κλίμακας (*scale parameter*) της κατανομής της διεργασίας. Η παράμετρος θέσης μιας κατανομής συνήθως αντιπροσωπεύεται από τη μέση τιμή (*mean*) ή τη διάμεσο (*median*), ανάλογα με το τι ταιριάζει κάθε φορά καλύτερα. Για παράδειγμα, όταν η υποκείμενη κατανομή είναι ασύμμετρη (*skewed, asymmetric*), επιλέγουμε τη διάμεσο. Προφανώς, όταν η κατανομή είναι συμμετρική, η μέση τιμή και η διάμεσος συμπίπτουν. Συνεπώς, οποιοδήποτε από τα δύο μπορεί να χρησιμοποιηθεί ως μέτρο θέσης (*location measure*).

Στο Μη Παραμετρικό Στατιστικό Έλεγχο Ποιότητας (*nonparametric SPC, NSPC*) υποθέτουμε ότι η υποκείμενη κατανομή ακολουθεί ένα **μοντέλο θέσης** (*location model*) με αθροιστική συνάρτηση κατανομής ή α.κ.σ. (*cumulative distribution function, cdf*) την $F(x-\theta)$, όπου θ είναι η παράμετρος θέσης ή ένα **μοντέλο κλίμακας** (*scale model*) με α.κ.σ. την $F(x/\tau)$, όπου $\tau > 0$ είναι η παράμετρος κλίμακας. Επίσης, πιο γενικά, υποθέτουμε ότι η κατανομή ακολουθεί ένα **μοντέλο θέσης-κλίμακας** (*location-scale model*) με α.κ.σ. την $F[(x-\theta)/\tau]$, όπου θ και τ είναι η παράμετρος θέσης και κλίμακας, αντίστοιχα. Στην περίπτωση που η α.κ.σ. F δεν είναι γνωστή, θεωρούμε ότι είναι συνεχής. Σκοπός είναι ο έλεγχος της παραμέτρου θ ή τ ή και των δύο, βάσει μεμονωμένων μετρήσεων (*individual measurements*) ή τυχαίων δειγμάτων/ορθολογικών υποομάδων (*random samples/rational subgroups*) που συνήθως συλλέγονται σε ισομήκη χρονικά διαστήματα.

Όπως στον Παραμετρικό Έλεγχο Ποιότητας, έτσι και στο Μη Παραμετρικό, υπάρχουν δύο διακριτές Φάσεις. Στη **Φάση I** ή **Αναδρομική Φάση** (*Phase I or Retrospective Phase*), γίνεται εκτίμηση των ορίων ελέγχου, όταν η διεργασία είναι εντός στατιστικού ελέγχου, με

σκοπό τα όρια αυτά να χρησιμοποιηθούν για τη μελλοντική παρακολούθηση της διεργασίας. Στη **Φάση II** ή **Μελλοντική Φάση** ή **Φάση Παρακολούθησης** (*Phase II or Prospective Phase or Monitoring Phase*), χρησιμοποιούνται τα όρια ελέγχου της Φάσης I για να διαπιστώσουμε εάν η διεργασία παραμένει εντός ελέγχου, όταν συλλέγονται μελλοντικά δείγματα. Όταν οι εντός ελέγχου παράμετροι, που ενδιαφερόμαστε να παρακολουθήσουμε, είτε είναι γνωστές (ιστορικά) είτε μπορούν να προσδιοριστούν αναφερόμαστε στην **περίπτωση γνωστών προτύπων** ή **Περίπτωση Κ** (“*standards-known*” case or *Case K*). Όμως, υπάρχουν και περιπτώσεις που δεν υπάρχει αρκετή διαθέσιμη πληροφορία και ως εκ τούτου οι παράμετροι δεν είναι γνωστές και πρέπει να εκτιμηθούν, πράγμα που γίνεται στη Φάση I. Αυτό αναφέρεται ως **περίπτωση άγνωστων προτύπων** ή **Περίπτωση U** (“*standards-unknown*” case or *Case U*). Στη δεύτερη περίπτωση, οι παράμετροι εκτιμώνται από κάποια δεδομένα αναφοράς (*reference data*), όταν η διεργασία είναι εντός ελέγχου. Θα πρέπει να αναφέρουμε ότι στην Περίπτωση U, γίνεται πρώτα ανάλυση Φάσης I και μετά ξεκινά η ανάλυση Φάσης II, ενώ αυτό δε γίνεται στην Περίπτωση Κ. Τέλος, σε κάθε μία από τις περιπτώσεις U και K, υπάρχουν τρεις βασικές κατηγορίες διαγραμμάτων: τα διαγράμματα τύπου Shewhart, *CUSUM* και *EWMA*.

4.2 Μονομεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου

Στα μονομεταβλητά διαγράμματα απεικονίζονται οι μετρήσεις ενός μόνο ποιοτικού χαρακτηριστικού. Είναι γεγονός ότι έχουν αναπτυχθεί περισσότερα διαγράμματα για την παρακολούθηση της θέσης μιας συνεχούς διεργασίας, ενώ ελάχιστα είναι διαθέσιμα για την κλίμακα. Στην περίπτωση Κ, η κατασκευή των διαγραμμάτων για την παράμετρο θέσης βασίζεται σε δύο κυρίως μη παραμετρικούς ελέγχους: τον προσημικό έλεγχο (*sign test* or *SN test*) και τον προσημικό βαθμολογικό έλεγχο του Wilcoxon (*Wilcoxon’s signed-rank test* or *SR test*). Ο προσημικός βαθμολογικός έλεγχος του Wilcoxon είναι σχεδόν τόσο αποτελεσματικός όσο και ο έλεγχος για τη μέση τιμή (*t-test*), ακόμα και αν η κατανομή είναι κανονική. Ωστόσο, ο *SN* έλεγχος είναι πιο αποτελεσματικός από τον *SR*, όταν η κατανομή είναι συμμετρική και έχει παχιές ουρές, όπως η διπλή εκθετική. Επίσης, ο *SN* έλεγχος εφαρμόζεται σε όλα τα ποσοστημόρια (*percentiles*), ενώ ο *SR* έλεγχος προτείνεται μόνο για τη διάμεσο (*50th percentile* or *second quartile Q₂*).

Στην πρόσφατη βιβλιογραφία έχουν προταθεί αρκετά μονομεταβλητά μη παραμετρικά διαγράμματα ελέγχου, τα οποία βασίζονται σε διάφορες μη παραμετρικές μεθόδους. Για παράδειγμα, οι Human, Chakraborti and Smit (2010) παρουσίασαν μια σειρά από μονόπλευρα και δίπλευρα μη παραμετρικά διαγράμματα ελέγχου τύπου Shewhart που βασίζονται στην προσημική στατιστική συνάρτηση και σε κανόνες ροών (*Shewhart-type Sign charts with Runs-type signaling rules*), ενώ οι Chakraborti and Eryilmaz (2007) παρουσίασαν μη παραμετρικά διαγράμματα ελέγχου τύπου Shewhart που βασίζονται στην προσημική βαθμολογική στατιστική συνάρτηση και σε κανόνες ροών (*Shewhart-type Signed-Rank charts with Runs-type signaling rules*). Επιπλέον, οι Amin, Reynolds and Bakir (1995) ανέπτυξαν ένα προσημικό διάγραμμα τύπου *CUSUM* (*CUSUM-type Sign chart*), οι Bakir and Reynolds (1979) ένα προσημικό βαθμολογικό διάγραμμα τύπου *CUSUM* για τη διάμεσο (*CUSUM-type Signed-Rank chart*) και οι Amin and Searcy (1991) ένα διάγραμμα τύπου *EWMA* που βασίζεται στον έλεγχο *SR* (*EWMA-type Signed Rank chart*). Όλα τα παραπάνω διαγράμματα είναι διαγράμματα ελέγχου για την παρακολούθηση της παραμέτρου θέσης και χρησιμοποιούνται στην Περίπτωση Κ.

Αντίστοιχα διαγράμματα Φάσης II έχουν προταθεί και για την Περίπτωση U. Συγκεκριμένα, οι Janacek and Meikle (1997) παρουσίασαν ένα διάγραμμα ελέγχου τύπου Shewhart Φάσης II (*Shewhart-type Precedence chart*) που βασίζεται στους στατιστικούς ελέγχους προτεραιότητας (*precedence tests*), οι Chakraborti, Eryilmaz and Human (2009) πρότειναν ένα αντίστοιχο διάγραμμα που χρησιμοποιεί κανόνες ροής (*Shewhart-type Precedence chart with signaling runs-type rules*), ενώ οι Balakrishnan, Triantafyllou and Koutras (2010) πρότειναν ένα ενισχυμένο διάγραμμα ελέγχου τύπου Shewhart Φάσης II (*Shewhart-type Precedence chart with enhancement*), το οποίο βασίζεται στους στατιστικούς ελέγχους προτεραιότητας. Σύμφωνα με την τεχνική αυτή, με βάση το δείγμα αναφοράς, υπολογίζονται συγκεκριμένες διατεταγμένες παρατηρήσεις, οι οποίες συνιστούν τα όρια ελέγχου του διαγράμματος. Η κατάσταση της διεργασίας κρίνεται από το εάν η παράμετρος θέσης διαδοχικών τυχαίων δειγμάτων της διεργασίας βρίσκεται μεταξύ των ορίων και από τον αριθμό των παρατηρήσεων του δείγματος που βρίσκονται εκτός των ορίων. Το 2009, οι ίδιοι συγγραφείς πρότειναν τρία νέα μη παραμετρικά διαγράμματα ελέγχου Φάσης II που βασίζονται στην προσημική και στην προσημική βαθμολογική στατιστική συνάρτηση του Wilcoxon. Τέλος, οι Chakraborti and Van de Wiel (2008) παρουσίασαν ένα διάγραμμα τύπου Shewhart και ένα διάγραμμα τύπου *EWMA*, τα οποία βασίζονται στον έλεγχο των Mann-

Whitney-Wilcoxon (*Shewhart-type and EWMA-type Mann-Whitney-Wilcoxon Rank-Sum charts*), ενώ οι Li, Tang and Ng (2010) πρότειναν ένα αντίστοιχο διάγραμμα τύπου *CUSUM*.

Επιπρόσθετα, έχουν προταθεί και μη παραμετρικά διαγράμματα ελέγχου Φάσης I για την παρακολούθηση της παραμέτρου θέσης στην Περίπτωση U. Συγκεκριμένα, οι Jones-Farmer, Jordan and Champ (2009) πρότειναν ένα διάγραμμα τύπου Shewhart, το οποίο βασίζεται στον παραμετρικό έλεγχο Kruskal-Wallis και ονομάζεται διάγραμμα mean-rank. Οι Graham, Human and Chakraborti (2010) παρουσίασαν διάγραμμα τύπου Shewhart, το οποίο βασίζεται στον έλεγχο της διαμέσου (*median test*) και ονομάζεται διάγραμμα median. Τέλος, ο Bakir (2010) πρότεινε ένα διάγραμμα Φάσης I για την παρακολούθηση της παραμέτρου κλίμακας. Αξίζει να σημειώσουμε ότι οι Chakraborti, Van der Laan and Bakir (2001), Chakraborti and Graham (2008) και Chakraborti (2011) παρουσίασαν τα μονομεταβλητά μη παραμετρικά διαγράμματα ελέγχου που ήταν διαθέσιμα στη βιβλιογραφία μέχρι το 2000, 2007 και 2010, αντίστοιχα.

Από τα παραπάνω γίνεται φανερό ότι υπάρχει πληθώρα μονομεταβλητών ή μονοδιάστατων μη παραμετρικών διαγραμμάτων ελέγχου, τα οποία έχουν προταθεί στη βιβλιογραφία μέχρι στιγμής. Ωστόσο, στην επόμενη παράγραφο θα παρουσιάσουμε μόνο τα διαγράμματα εκείνα, τα οποία αποτελούν άμεση γενίκευση κάποιων από τα πολυμεταβλητά διαγράμματα που θα αναπτύξουμε στη συνέχεια.

4.2.1 Διαγράμματα Ελέγχου Τύπου Shewhart

Είναι γνωστό ότι τα διαγράμματα τύπου Shewhart είναι πιο αποτελεσματικά σε μεσαίες και μεγάλες μετατοπίσεις. Τα μονόπλευρα διαγράμματα έχουν την κεντρική γραμμή και ένα μόνο άνω ή κάτω όριο ελέγχου. Επίσης, είναι πιο χρήσιμα όταν μας ενδιαφέρει μόνο η κατεύθυνση (ανοδική ή καθοδική) της μετατόπισης της παραμέτρου θέσης (όπως η διάμεσος). Αντίθετα, τα δίπλευρα διαγράμματα έχουν την κεντρική γραμμή, ένα άνω όριο και ένα κάτω όριο ελέγχου και χρησιμοποιούνται συνήθως για την ανίχνευση πιθανής μετατόπισης ή αλλαγής σε οποιαδήποτε κατεύθυνση.

Στην περίπτωση που η τιμή θ_0 της διαμέσου είναι γνωστή ή μπορεί να προσδιοριστεί (Περίπτωση K), οι Amin, Reynolds and Bakir (1995) παρουσίασαν ένα δίπλευρο μη παραμετρικό (μονοδιάστατο) διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση

της διάμεσου (παράμετρος θέσης), το οποίο βασίζεται στον προσημικό έλεγχο (*SN test*) και είναι γνωστό ως **διάγραμμα SN** (*SN chart*). Θεωρούμε ότι $X_{k1}, X_{k2}, \dots, X_{kn}$ είναι το k δείγμα ή η k υποομάδα ($k=1, 2, \dots, m$), που αποτελείται από ανεξάρτητες παρατηρήσεις μεγέθους $n > 1$. Η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα SN ονομάζεται **προσημική συνάρτηση εντός των ομάδων** (*within group sign statistics*) και έχει τη μορφή

$$SN_k = \sum_{i=1}^n \text{sign}(X_{ki} - \theta_0), \quad k = 1, 2, \dots, m, \quad (4.2.1)$$

όπου

$$\text{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

Για τον προσδιορισμό των ορίων ελέγχου, είναι απαραίτητο να βρεθεί η κατανομή της συνάρτησης SN_k . Χρησιμοποιούμε τον τύπο $SN_k = 2T_k - n$, όπου T_k είναι η συνήθης προσημική στατιστική συνάρτηση ελέγχου, η οποία υπολογίζει τον αριθμό των δειγματικών παρατηρήσεων που είναι μεγαλύτερες ή ίσες του θ_0 . Αφού η εντός ελέγχου κατανομή της T_k είναι διωνυμική με παραμέτρους n και 0.5 , συνεπάγεται ότι η εντός ελέγχου κατανομή της SN_k είναι συμμετρική γύρω από το μηδέν. Επομένως, τα όρια ελέγχου τύπου Shewhart και η κεντρική γραμμή του δίπλευρου μη παραμετρικού διαγράμματος SN είναι τα εξής

$$\begin{array}{l} \text{Δίπλευρο Διάγραμμα SN} \\ \text{Τύπου Shewhart} \\ \text{(Περίπτωση K)} \end{array} \begin{cases} LCL = -c \\ CL = 0 \\ UCL = c \end{cases}, \quad (4.2.2)$$

όπου c είναι ένας θετικός ακέραιος αριθμός μεταξύ του 0 και του n ($0 \leq c \leq n$). Η διεργασία τίθεται εκτός στατιστικού ελέγχου εάν κάποιο σημείο της συνάρτησης SN_k βρεθεί εκτός των ορίων ελέγχου, δηλαδή όταν ισχύει $SN_k \leq -c$ ή όταν $SN_k \geq c$. Τυπικά, η σταθερά c προσδιορίζεται για προκαθορισμένο ARL_0 , το οποίο στην Περίπτωση K είναι ίσο με το αντίστροφο του ποσοστού λανθασμένου συναγερμού ($FAR(a)$), δηλαδή $ARL_0 = 1/FAR(a)$. Συνεπώς, χρησιμοποιώντας την συμμετρία της διωνυμικής κατανομής, η σταθερά c είναι ο μικρότερος ακέραιος έτσι ώστε $P_{\theta_0}(SN_k \geq c) \leq a/2$. Παρακάτω παρατίθεται ο Πίνακας 4.1

(Πίνακας 1 των Chakraborti and Graham (2008)), στον οποίο δίνονται διάφορες τιμές της πιθανότητας $P_{\theta_0}(T \geq t)$. Οι τιμές της σταθεράς c προκύπτουν από τον τύπο $c = 2t - n$.

n	5	6	7	8	9	10
t	5	6	7	8	9	10
$P_{\theta_0}(T \geq t)$	0.0312	0.0156	0.0078	0.0039	0.0020	0.0010
$FAR(a)$	0.0624	0.0312	0.0156	0.0078	0.0040	0.0020
ARL_0	16	32	64	128	256	512

Πίνακας 4.1: FAR και ARL_0 του διαγράμματος SN για διάφορες τιμές του n και t

Οι Amin, Reynolds and Bakir (1995) σύγκριναν το μέσο μήκος ροής του κλασσικού διαγράμματος \bar{X} με αυτό του διαγράμματος SN τύπου Shewhart για διάφορα μεγέθη μετατοπίσεων και διαφορετικές κατανομές. Το συμπέρασμα στο οποίο κατέληξαν είναι ότι το διάγραμμα SN είναι πιο αποτελεσματικό, όταν η κατανομή δεν είναι συμμετρική ή είναι συμμετρική με παχιές ουρές. Ενώ, όταν η κατανομή είναι είτε κανονική είτε πλησιάζει την κανονική και έχει λεπτές ουρές, συμβαίνει ακριβώς το αντίθετο. Γενικά, το βασικό πλεονέκτημα του διαγράμματος SN (και κατ' επέκταση όλων των μη παραμετρικών διαγραμμάτων) είναι ότι το FAR και το ARL_0 παραμένουν σταθερά για όλες τις συνεχείς κατανομές. Αυτό δεν ισχύει για κανένα παραμετρικό διάγραμμα με συγκεκριμένη κατανομή συμπεριλαμβανομένου και του διαγράμματος \bar{X} .

Ο Bakir (2004) παρουσίασε ένα δίπλευρο μη παραμετρικό (μονοδιάστατο) διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση της διάμεσου (παράμετρος θέσης), το οποίο βασίζεται στον προσημικό βαθμολογικό έλεγχο του Wilcoxon (SR test) και είναι γνωστό ως **διάγραμμα SR** (SR chart). Θεωρούμε ότι $|X_{k1} - \theta_0|, |X_{k2} - \theta_0|, \dots, |X_{kn} - \theta_0|$ είναι οι απόλυτες τιμές των αποκλίσεων των παρατηρήσεων του k -δείγματος από τη γνωστή εντός ελέγχου τιμή θ_0 της διαμέσου (Περίπτωση Κ) και ότι R_{ki} είναι ο βαθμός του $|X_{ki} - \theta_0|$ ανάμεσα σε n απόλυτες αποκλίσεις. Η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα SR συμβολίζεται με SR_k και αναπαριστά τη διαφορά μεταξύ του αθροίσματος των

βαθμών (των απόλυτων τιμών) σε αντιστοιχία με τις θετικές και τις αρνητικές τιμές αποκλίσεις. Η τιμή της υπολογίζεται μέσω του τύπου που ακολουθεί

$$SR_k = \sum_{i=1}^n \text{sign}(X_{ki} - \theta_0) \cdot R_{ki}, \quad k=1, 2, \dots, m. \quad (4.2.3)$$

Η παραπάνω στατιστική συνάρτηση SR συνδέεται γραμμικά με την προσημική βαθμολογική συνάρτηση του Wilcoxon W^+ μέσω του τύπου $SR = 2W^+ - (n(n+1)/2)$, όπου W^+ είναι το άθροισμα των βαθμολογιών των απόλυτων τιμών που αντιστοιχούν στις θετικές αποκλίσεις. Σύμφωνα με τους Gibbons and Chakraborti (2003), η συνάρτηση W^+ είναι ελεύθερη κατανομής. Συνεπώς, ο ίδιο ισχύει και για τη συνάρτηση SR_k και το διάγραμμα SR . Επειδή η εντός ελέγχου κατανομή της SR_k είναι συμμετρική γύρω από το μηδέν, μία λογική επιλογή για το άνω και κάτω όριο είναι $LCL = -UCL$. Τα όρια ελέγχου του διαγράμματος SR είναι

$$\begin{array}{l} \text{Διπλευρο Διάγραμμα } SR \\ \text{Τύπου Shewhart} \\ \text{(Περίπτωση K)} \end{array} \left\{ \begin{array}{l} LCL = -d \\ CL = 0 \\ UCL = d \end{array} \right., \quad (4.2.4)$$

όπου d είναι ένας θετικός ακέραιος. Η τιμή του d επιλέγεται για προκαθορισμένο FAR ή ισοδύναμα για προκαθορισμένο ARL_0 .

UCL	$n = 4$		$n = 5$		$n = 6$	
	ARL_0^+	p_0^+	ARL_0^+	p_0^+	ARL_0^+	p_0^+
10	16	0.06250000	6.40	0.15625	4.57	0.218750
11	∞	0.0	10.67	0.09375	6.40	0.156250
12			10.67	0.09375	6.40	0.156250
13			16.00	0.06250	9.14	0.109375
14			16.00	0.06250	9.14	0.109375
15			32.00	0.03125	12.80	0.781250
16			∞	0.0	12.80	0.781250
17					21.33	0.046875
18					21.33	0.046875
19					32.00	0.031250
20					32.00	0.031250
21					64.00	0.015600
22					∞	0.0

Πίνακας 4.2: Τιμές άνω ορίου ελέγχου του διαγράμματος SR για καθορισμένο p_0^+ και ARL_0^+ σε δείγματα μεγέθους $n = 4, 5$ και 8

UCL	n = 8		UCL	n = 10	
	ARL ₀ ⁺	p ₀ ⁺		ARL ₀ ⁺	p ₀ ⁺
18	8.00	0.12500000	25	8.61	0.116210937
20	10.24	0.09765625	27	10.34	0.096679687
22	13.47	0.07421875	29	12.49	0.080078125
24	18.29	0.05468750	31	15.28	0.065429687
26	25.60	0.00390625	33	18.96	0.052734375
28	36.57	0.02734375	35	23.81	0.041992187
30	51.20	0.01953125	37	31.03	0.032226562
32	85.33	0.01171875	39	40.96	0.024414062
34	128.00	0.00781250	41	53.89	0.018554687
36	256.00	0.00390625	43	73.14	0.013671875
>36	∞	0.0	45	102.40	0.009765625
			47	146.29	0.006835037
			49	204.80	0.004882812
			51	341.33	0.002929687
			53	512.00	0.001953125
			55	1024.00	0.000976562
			>55	∞	0.0

Πίνακας 4.3: Τιμές άνω ορίου ελέγχου του διαγράμματος *SR* για καθορισμένο p_0^+ και ARL_0^+ σε δείγματα μεγέθους $n = 8$ και 10

4.3 Πολυμεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου

4.3.1 Διαγράμματα Ελέγχου Τύπου Shewhart

Ο Boone (2010) πρότεινε ένα πολυμεταβλητό προσημικό διάγραμμα ελέγχου τύπου Shewhart (*multivariate Shewhart type sign control chart or SN² chart*) για την παρακολούθηση του διανύσματος των διαμέσων, το οποίο βασίζεται στην πολυδιάστατη μορφή του προσημικού ελέγχου (*sign test*) που περιγράφεται από τους Puri and Sen (1976) και Hettmansperger (2006). Το διάγραμμα αυτό ονομάζεται SN^2 και αποτελεί γενίκευση του μονοδιάστατου διαγράμματος SN .

Θεωρούμε ότι θέλουμε να παρακολουθήσουμε τις εντός ελέγχου διαμέσους $\theta_j = \theta_{j_0}$ για $j = 1, 2, \dots, p$ των p χαρακτηριστικών, όταν οι εντός ελέγχου τιμές θ_{j_0} , $j = 1, 2, \dots, p$, των

διαμέσων είναι γνωστές. Για κάθε χαρακτηριστικό υπολογίζουμε τη μονοδιάστατη προσημική στατιστική συνάρτηση

$$S_j = \sum_{i=1}^n \text{sign}(X_{ji} - \theta_{j0}), \quad j=1, 2, \dots, p,$$

όπου X_{ji} είναι η i παρατήρηση του j χαρακτηριστικού, n είναι ο συνολικός αριθμός των παρατηρήσεων, θ_{j0} είναι η καθορισμένη εντός ελέγχου τιμή της διαμέσου και

$$\text{sign}(X_{ji} - \theta_{j0}) = \begin{cases} 1, & \text{για } X_{ji} - \theta_{j0} > 0 \\ -1, & \text{για } X_{ji} - \theta_{j0} < 0. \end{cases}$$

Συνεπώς, $S_j = \#(X_{ji} > \theta_{j0}) - \#(X_{ji} < \theta_{j0})$ για $j=1, 2, \dots, p$ και $i=1, 2, \dots, n$. Εάν \mathbf{S} είναι το $(p \times 1)$ διάνυσμα των S_j , τότε η ποσότητα $n^{-1/2}\mathbf{S}$ ακολουθεί ασυμπτωτικά την πολυδιάστατη κανονική κατανομή με μέση τιμή το διάνυσμα $\mathbf{0}$ και πίνακα διακύμανσης-συνδιακύμανσης \mathbf{V} . Η εκτίμηση $\hat{\mathbf{V}} = ((\hat{v}_{jl}))$, $j, l=1, 2, \dots, p$ του πίνακα \mathbf{V} προκύπτει μέσω των τύπων

$$\hat{v}_{jj} = n \quad \text{και} \quad \hat{v}_{jl} = \sum_{i=1}^n \text{sign}(X_{ji} - \theta_{j0}) \text{sign}(X_{li} - \theta_{l0}).$$

Η πολυμεταβλητή προσημική στατιστική συνάρτηση για το k δείγμα ορίζεται από τον τύπο

$$SN_k^2 = \mathbf{S}' \hat{\mathbf{V}}^{-1} \mathbf{S}. \quad (4.3.1)$$

Ο Boone (2010) υποστηρίζει ότι για μεγάλα δειγματικά μεγέθη η οριακή κατανομή (*limiting distribution*) της στατιστικής συνάρτησης SN^2 είναι η χ^2 με p βαθμούς ελευθερίας. Επιπλέον, υπογραμμίζει ότι ο πολυμεταβλητός προσημικός έλεγχος (και κατ' επέκταση το διάγραμμα SN^2) είναι (ασυμπτωτικά) ελεύθερος κατανομής (*distribution-free*) για μεγάλα δείγματα και για τον προσδιορισμό των ορίων ελέγχου χρησιμοποιείται στην πράξη η χ^2 κατανομή. Συνεπώς, τα όρια ελέγχου του διαγράμματος SN^2 είναι

$$\begin{array}{l} \text{Διάγραμμα } SN^2 \\ \text{Τύπου Shewhart} \end{array} \begin{cases} LCL = 0 \\ UCL = \chi_{p; \alpha}^2. \end{cases} \quad (4.3.2)$$

Στην περίπτωση που έχουμε μόνο δύο ποιοτικά χαρακτηριστικά ($p=2$), οι Puri and Sen (1976) και Hettmansperger (2006) έδωσαν στη στατιστική συνάρτηση SN^2 τη μορφή

$$SN^2 = \frac{(C_1 - C_2)^2}{C_1 + C_2} + \frac{(D_1 - D_2)^2}{D_1 + D_2},$$

όπου για $i = 1, 2, \dots, n$ και $\boldsymbol{\theta}' = (\theta_{01}, \theta_{02})$ έχουμε

$$C_1 = \#(X_{1i} \leq \theta_{01}, X_{2i} \leq \theta_{02}), \quad C_2 = \#(X_{1i} > \theta_{01}, X_{2i} > \theta_{02}),$$

$$D_1 = \#(X_{1i} \leq \theta_{01}, X_{2i} \leq \theta_{02}), \quad D_2 = \#(X_{1i} > \theta_{01}, X_{2i} > \theta_{02}).$$

Στην παραπάνω στατιστική συνάρτηση βασίστηκε ο Das (2009) και προσδιόρισε τα όρια ελέγχου του διαγράμματος SN^2 για $p = 2$ ως εξής

$$\text{Διάγραμμα } SN^2 \quad \begin{cases} LCL = 0 \\ UCL = \chi^2_{2; \alpha} \end{cases}$$

Τύπου Shewhart για $p=2$

Προφανώς, το διάγραμμα SN^2 του Boone (2010) αποτελεί επέκταση του διαγράμματος SN^2 του Das (2009). Ο Das σύγκρινε το διάγραμμα που πρότεινε με το διάγραμμα T^2 του Hotelling και κατέληξε στο συμπέρασμα ότι βάσει του ARL_0 το διάγραμμα SN^2 έχει καλύτερη επίδοση από το διάγραμμα T^2 , ενώ βάσει του ARL_1 η επίδοση του διαγράμματος SN^2 είναι ελαφρώς πιο φτωχή σε σχέση με την επίδοση του διαγράμματος T^2 . Ωστόσο, για μετατόπιση μεγαλύτερη από 1.5, οι επιδόσεις των δυο διαγραμμάτων δε διαφέρουν αισθητά. Τέλος, όταν το μέγεθος δείγματος αυξάνεται, βελτιώνεται η επίδοση του διαγράμματος SN^2 για την ανίχνευση συγκεκριμένης μετατόπισης.

Ο Boone (2010) πρότεινε και ένα πολυμεταβλητό προσημικό βαθμολογικό διάγραμμα ελέγχου τύπου Shewhart (*multivariate Shewhart type Wilcoxon signed-rank control chart or SR^2 chart*) για την παρακολούθηση του διανύσματος των διαμέσων, το οποίο βασίζεται στην πολυδιάστατη μορφή του βαθμολογικού προσημικού ελέγχου του Wilcoxon (*Wilcoxon's signed-rank test*) που περιγράφει ο Hettmansperger (2006). Το διάγραμμα αυτό ονομάζεται SR^2 και αποτελεί γενίκευση του μονοδιάστατου διαγράμματος SR .

Έστω ότι θέλουμε να παρακολουθήσουμε τις εντός ελέγχου διαμέσους $\theta_j = \theta_{j0}$ για $j = 1, 2, \dots, p$ των p χαρακτηριστικών, όπου οι εντός ελέγχου τιμές θ_{j0} , $j = 1, 2, \dots, p$, των διαμέσων είναι γνωστές. Ας υποθέτουμε επιπλέον ότι η από κοινού κατανομή (*joint distribution*) των p χαρακτηριστικών είναι διαγώνια συμμετρική (*diagonally symmetric*)

γύρω από το διάνυσμα $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$. Κάτω από αυτή την υπόθεση, οι περιθώριες κατανομές (*marginal distributions*) των μεμονωμένων μεταβλητών είναι συμμετρικές γύρω από τις αντίστοιχες διαμέσους. Όταν η διεργασία είναι εντός στατιστικού ελέγχου, ισχύει $\boldsymbol{\theta} = \boldsymbol{\theta}_0 = (\theta_{10}, \theta_{20}, \dots, \theta_{p0})'$, όπου θ_{j0} είναι η προκαθορισμένη τιμή της διαμέσου του j χαρακτηριστικού. Για κάθε χαρακτηριστικό υπολογίζουμε τη μονοδιάστατη βαθμολογική προσημική στατιστική συνάρτηση του Wilcoxon

$$W_j = \sum_{i=1}^n R(|X_{ji} - \theta_{j0}|) \cdot \text{sign}(X_{ji} - \theta_{j0}), \quad j=1, 2, \dots, p,$$

όπου $R(|X_{ji} - \theta_{j0}|)$ είναι ο βαθμός του $|X_{ji} - \theta_{j0}|$ μεταξύ των $|X_{j1} - \theta_{j0}|, \dots, |X_{jn} - \theta_{j0}|$. Εάν \mathbf{W} είναι το $p \times 1$ διάνυσμα των W_j , τότε η ασυμπτωτική κατανομή της ποσότητας $n^{-3/2} \mathbf{W}_j$ είναι η πολυδιάστατη κανονική κατανομή με μέση τιμή το διάνυσμα $\mathbf{0}$ και πίνακα διακύμανσης-συνδιακύμανσης που εκτιμάται από την ποσότητα $n^{-3} \hat{\mathbf{L}}$. Η εκτίμηση $\hat{\mathbf{L}} = ((\hat{l}_{jl}))$, $j, l=1, 2, \dots, p$ προκύπτει μέσω των τύπων

$$\hat{l}_{jj} = \frac{n(n+1)(2n+1)}{6}, \quad j=1, 2, \dots, p \text{ και}$$

$$\hat{l}_{jl} = \sum_{i=1}^n R(|X_{ji} - \theta_{j0}|) \cdot R(|X_{li} - \theta_{l0}|) \cdot \text{sign}(X_{ji} - \theta_{j0}) \cdot \text{sign}(X_{li} - \theta_{l0}), \quad j, l=1, 2, \dots, p.$$

Η πολυμεταβλητή βαθμολογική προσημική στατιστική συνάρτηση για το k δείγμα, που θα απεικονίζεται στο διάγραμμα SR^2 , ορίζεται από τον τύπο

$$SR_k^2 = \mathbf{W}' \hat{\mathbf{L}}^{-1} \mathbf{W}. \quad (4.3.3)$$

Ο Boone (2010) υποστηρίζει ότι η οριακή κατανομή (*limiting distribution*) της στατιστικής συνάρτησης SR^2 είναι η χ^2 με p βαθμούς ελευθερίας. Όταν η διεργασία είναι εντός ελέγχου, οι διάμεσοι είναι γνωστοί, ενώ ο πίνακας διακύμανσης-συνδιακύμανσης όχι. Για τον υπολογισμό της στατιστικής συνάρτησης SR_k^2 σε κάθε δείγμα χρησιμοποιείται ο συνεπής εκτιμητής του ασυμπτωτικού πίνακα συνδιακύμανσης (*consistent estimator of the asymptotic covariance matrix*) $\hat{\mathbf{L}}$. Επιπλέον, υπογραμμίζει ότι ο πολυμεταβλητός βαθμολογικός προσημικός έλεγχος του Wilcoxon (και κατ' επέκταση το διάγραμμα SR^2) δεν

είναι ελεύθερος κατανομής (*distribution-free*) για πεπερασμένα δείγματα (*finite samples*). Και στην περίπτωση του διαγράμματος ελέγχου SR^2 , η χ^2 κατανομή χρησιμοποιείται για τον προσδιορισμό των ορίων. Συνεπώς, τα όρια ελέγχου του διαγράμματος SR^2 είναι

$$\begin{array}{l} \text{Διάγραμμα } SR^2 \\ \text{Τύπου Shewhart} \end{array} \left\{ \begin{array}{l} LCL = 0 \\ UCL = \chi^2_{p; \alpha} \end{array} \right. \quad (4.3.4)$$

Η Liu (1995) παρουσίασε δύο πολυμεταβλητά διαγράμματα που βασίζονται στο βάθος των δεδομένων (*data depth*). Τα διαγράμματα που πρότείνει ονομάζονται r και Q , τα οποία αποτελούν γενικεύσεις των μονοδιάστατων διαγραμμάτων X και \bar{X} , αντίστοιχα. Σύμφωνα με την Liu, θεωρούμε ότι G είναι η κατανομή p μεταβλητών με Y_1, Y_2, \dots, Y_m τυχαίες παρατηρήσεις, οι οποίες αναπαριστούν το εντός ελέγχου σύνολο δεδομένων αναφοράς (*reference in-control data set*). Στη συνέχεια, συλλέγουμε ένα νέο δείγμα παρατηρήσεων X_1, X_2, \dots, X_n και υποθέτουμε ότι η κατανομή καλείται F (σημείωση: όχι η F -κατανομή). Για να διαπιστώσουμε εάν η διεργασία είναι εκτός ελέγχου, συγκρίνουμε το νέο δείγμα παρατηρήσεων $X_1, X_2, \dots, X_n \sim F$ με το δείγμα αναφοράς $Y_1, Y_2, \dots, Y_m \sim G$.

$H_0 : G = F$ με ποσοστό λανθασμένου συναγερμού α

$H_1 : \text{υπάρχει μετατόπιση της θέσης ή/και αύξηση της κλίμακας από την } G \text{ στην } F.$

Οι στατιστικές συναρτήσεις που χρησιμοποιούνται για τον προσδιορισμό πιθανής διαφοράς μεταξύ των G και F , βασίζονται στην έννοια του βάθους των δεδομένων. Για κάθε σημείο \mathbf{y} στον \mathbb{R}^p , το **απλό βάθος** (*simplicial depth*) του \mathbf{y} σε σχέση με την G ορίζεται ως εξής

$$SD_G(\mathbf{y}) = P_G \left\{ \mathbf{y} \in s[Y_1, Y_2, \dots, Y_{p+1}] \right\},$$

όπου $s[Y_1, Y_2, \dots, Y_{p+1}]$ είναι ένα πολύγωνο με πλευρές Y_1, Y_2, \dots, Y_{p+1} , οι οποίες αποτελούν $p+1$ τυχαίες παρατηρήσεις από την κατανομή G . Η ποσότητα SD_G μετρά πόσο βαθύ (*deep*) ή πόσο κεντρικό (*central*) είναι το σημείο \mathbf{y} σε σχέση με την κατανομή G . Όταν η G δεν είναι γνωστή και δίνεται μόνο το δείγμα Y_1, Y_2, \dots, Y_m , το **απλό δειγματικό βάθος** (*sample simplicial depth*) του \mathbf{y} ορίζεται ως

$$SD_{G_m}(\mathbf{y}) = \binom{m}{p+1}^{-1} \sum_{(*)} I(\mathbf{y} \in s[Y_{i_1}, Y_{i_2}, \dots, Y_{i_{p+1}}])$$

και μετρά πόσο βαθιά (*deep*) είναι το σημείο \mathbf{y} στο σύνολο Y_1, Y_2, \dots, Y_m . Στον παραπάνω τύπο, η δείκτρια συνάρτηση $I(\cdot)$ ορίζεται ως εξής

$$I(A) = \begin{cases} 1, & \text{εάν συμβαίνει το } A \\ 0, & \text{διαφορετικά.} \end{cases}$$

Η συνάρτηση $G_m(\cdot)$ δηλώνει την εμπειρική κατανομή του Y_1, Y_2, \dots, Y_m και το σύμβολο $(*)$ δηλώνει όλα τα δυνατά υποσύνολα του Y_1, Y_2, \dots, Y_m μεγέθους $(p+1)$.

Το βάθος των δεδομένων μπορεί να βασιστεί και στην απόσταση του Mahalanobis, το οποίο μετρά το βάθος του σημείου \mathbf{y} σε σχέση με το πόσο μικρή είναι η τετραγωνική απόστασή του από το μέσο, δεδομένης κατανομής G . Η απόσταση αυτή συμβολίζεται με $MD_G(\mathbf{y})$ και υπολογίζεται μέσω του τύπου που ακολουθεί

$$MD_G(\mathbf{y}) = \frac{1}{1 + (\mathbf{y} - \boldsymbol{\mu}_G)' \boldsymbol{\Sigma}_G^{-1} (\mathbf{y} - \boldsymbol{\mu}_G)},$$

όπου $\boldsymbol{\mu}_G$ και $\boldsymbol{\Sigma}_G$ είναι το μέσο διάνυσμα και ο πίνακας διακύμανσης-συνδιακύμανσης της κατανομής G , αντίστοιχα. Η εμπειρική μορφή της ποσότητας $MD_G(\mathbf{y})$ δίνεται από τον τύπο

$$MD_{G_m}(\mathbf{y}) = \frac{1}{1 + (\mathbf{y} - \bar{\mathbf{Y}})' \mathbf{S}_G^{-1} (\mathbf{y} - \bar{\mathbf{Y}})}, \quad (4.3.5)$$

όπου $\bar{\mathbf{Y}}$ είναι ο δειγματικός μέσος των Y_1, Y_2, \dots, Y_m και \mathbf{S}_G είναι ο δειγματικός πίνακας διακύμανσης-συνδιακύμανσης της κατανομής G .

Για την κατασκευή του διαγράμματος r , διατάσσουμε τα βάθη (*depth*) των Y_1, Y_2, \dots, Y_m σε αύξουσα σειρά και με $Y_{[1]}, Y_{[2]}, \dots, Y_{[m]}$ συμβολίζουμε τη διάταξη αυτή. Το $Y_{[m]}$ θα είναι το πιο κεντρικό σημείο. Όσο μικρότερος είναι ο βαθμός της διάταξης ενός σημείου, τόσο πιο απομακρυσμένο είναι το σημείο αυτό από την κατανομή G . Παρακάτω παραθέτουμε τις στατιστικές συναρτήσεις

$$r_G(\mathbf{y}) = P\{D_G(Y) \leq D_G(\mathbf{y}) | Y \sim G\}$$

και

$$r_{G_m}(\mathbf{y}) = \frac{\#\{Y_j \mid D_{G_m}(Y_j) \leq D_{G_m}(\mathbf{y}), j=1, 2, \dots, m\}}{m}, \quad (4.3.6)$$

με τις οποίες υπολογίζουμε αντίστοιχα τα $\{r_G(X_1), r_G(X_2), \dots\}$ ή τα $\{r_{G_m}(X_1), r_{G_m}(X_2), \dots\}$ εάν είναι διαθέσιμα μόνο τα Y_1, Y_2, \dots, Y_m , και όχι η κατανομή G . Στο διάγραμμα r απεικονίζονται τα $r_G(X_i)$ ή τα $r_{G_m}(X_i)$ με τα ακόλουθα όρια ελέγχου

$$\text{Διάγραμμα } r \begin{cases} CL = 0.5 \\ LCL = a. \end{cases} \quad (4.3.7)$$

Η διεργασία τίθεται εκτός στατιστικού ελέγχου όταν το $r_G(\cdot)$ βρίσκεται κάτω από το a . Επειδή a είναι το ποσοστό λανθασμένου συναγερμού, το οποίο είναι συνήθως κοντά στο μηδέν, το διάγραμμα r έχει μόνο κάτω όριο ελέγχου και όχι άνω.

Συμβολίζοντας με $F_n(\cdot)$ την εμπειρική κατανομή του δείγματος X_1, X_2, \dots, X_n , ορίζουμε

$$Q(G, F) = P\{D_G(Y) \leq D_G(X) \mid Y \sim G, X \sim F\} = E_F\{r_G(X)\},$$

$$Q(G, F_n) = \frac{1}{n} \sum_{i=1}^n r_G(X_i)$$

και

$$Q(G_m, F_n) = \frac{1}{n} \sum_{i=1}^n r_{G_m}(X_i). \quad (4.3.8)$$

Για την κατασκευή του διαγράμματος Q , η Liu (1995) προτείνει την απεικόνιση των μέσων των υποσυνόλων των $r_G(X_i)$ ή των $r_{G_m}(X_i)$. Θεωρούμε ότι κάθε υποσύνολο έχει μέγεθος n .

Οι μέσοι των $r_G(X_i)$ και των $r_{G_m}(X_i)$ υπολογίζονται από τους τύπους $Q(G, F_n^j)$ και $Q(G_m, F_n^j)$, όπου F_n^j η εμπειρική κατανομή των X_i του j υποσυνόλου για $j=1, 2, \dots$.

Στο διάγραμμα Q μπορούν να απεικονιστούν οι ποσότητες $\{Q(G, F_n^1), Q(G, F_n^2), \dots\}$ ή οι ποσότητες $\{Q(G_m, F_n^1), Q(G_m, F_n^2), \dots\}$ εάν μόνο τα Y_1, Y_2, \dots, Y_m είναι διαθέσιμα. Τα όρια ελέγχου του διαγράμματος Q εξαρτώνται από την επιλογή του n . Όταν στο διάγραμμα απεικονίζονται τα $Q(G, F_n^j)$ και η τιμή του n είναι μεγάλη έχουμε

$$\text{Διάγραμμα } Q \begin{cases} CL = 0.5 \\ \text{(για μεγάλο } n) \text{ } LCL = 0.5 - z_a(12n)^{-1/2}, \end{cases}$$

ενώ στην περίπτωση που απεικονίζονται τα $Q(G_m, F_n^j)$ και το n παραμένει μεγάλο έχουμε

$$\text{Διάγραμμα } Q \begin{cases} CL = 0.5 \\ \text{(για μεγάλο } n) \text{ } LCL = 0.5 - z_a \sqrt{\frac{1}{12} \left(\frac{1}{m} + \frac{1}{n} \right)}. \end{cases} \quad (4.3.9)$$

Όταν η τιμή του a είναι μικρή, τα όρια ελέγχου που προκύπτουν είναι τα εξής

$$\text{Διάγραμμα } Q \begin{cases} CL = 0.5 \\ \text{(για μικρό } a) \text{ } LCL = \frac{(n!a)^{1/n}}{n}. \end{cases} \quad (4.3.10)$$

4.3.2 Διαγράμματα Ελέγχου Συσσωρευμένου Αθροίσματος

Η Liu (1995) πρότεινε και ένα τρίτο διάγραμμα ελέγχου, το οποίο ονόμασε S και αποτελεί γενίκευση του μονοδιάστατου παραμετρικού διαγράμματος ελέγχου τύπου *CUSUM*. Για την κατασκευή του διαγράμματος S , η Liu προτείνει την απεικόνιση των ποσοτήτων $S_n(G)$ και $S_n(G_m)$, οι οποίες ορίζονται ως εξής

$$S_n(G) = \sum_{i=1}^n \left[r_G(X_i) - \frac{1}{2} \right] = n \left[Q(G, F_n) - \frac{1}{2} \right]$$

και

$$S_n(G_m) = \sum_{i=1}^n \left[r_{G_m}(X_i) - \frac{1}{2} \right] = n \left[Q(G_m, F_n) - \frac{1}{2} \right]. \quad (4.3.11)$$

Εάν στο διάγραμμα S απεικονίζεται η ποσότητα $S_n(G)$, το κάτω όριο ελέγχου είναι

$$\text{Διάγραμμα } S \begin{cases} LCL = -z_a \sqrt{\frac{n}{12}}, \end{cases}$$

ενώ εάν απεικονίζεται η ποσότητα $S_n(G_m)$, το κάτω όριο ελέγχου είναι

$$\text{Διάγραμμα } S \begin{cases} LCL = -z_a \sqrt{\frac{n^2}{12} \left(\frac{1}{m} + \frac{1}{n} \right)}. \end{cases} \quad (4.3.12)$$

Παρατηρούμε ότι τα παραπάνω όρια ελέγχου δεν είναι ευθείες γραμμές, αλλά καμπύλες. Για το λόγο αυτό, τυποποιούμε όλα τα συσσωρευμένα αθροίσματα, ώστε το κάτω όριο ελέγχου να είναι σταθερό, και στο διάγραμμα απεικονίζονται οι ποσότητες που ακολουθούν

$$S_n^*(G) = \frac{S_n(G)}{\sqrt{\frac{n}{12}}}$$

και

$$S_n^*(G_m) = \frac{S_n(G_m)}{\sqrt{\frac{n^2}{12} \left(\frac{1}{m} + \frac{1}{n} \right)}}, \quad (4.3.13)$$

για $n = 1, 2, \dots$. Το διάγραμμα αυτό ονομάζεται S^* και έχει τα εξής όρια ελέγχου

$$\text{Διάγραμμα } S^* \begin{cases} CL & = 0 \\ LCL & = -z_\alpha \end{cases} \quad (4.3.14)$$

Οι Dai, Zhou and Wang (2011) πρότειναν ένα πολυδιάστατο μη-παραμετρικό διάγραμμα τύπου *CUSUM*, το οποίο βασίζεται στο data depth για την ανίχνευση πιθανής μετατόπισης ή του μέσου ή/και του πίνακα διακύμανσης-συνδιακύμανσης. Θεωρούμε ότι έχουμε n ανεξάρτητες παρατηρήσεις από μία p -διάστατη κατανομή, συμβολικά $X_i \sim F_p(\mu_i, \Sigma_i)$ για $i = 1, 2, \dots, n$. Εάν η διεργασία είναι εντός ελέγχου, τότε $\mu_i = \mu$ και $\Sigma_i = \Sigma$ για όλα τα i . Υποθέτουμε ότι ο μέσος ή/και η διακύμανση μετατοπίζονται μετά την n_1 -οστή παρατήρηση. Η μέση τιμή και η διακύμανση των n_1 πρώτων παρατηρήσεων είναι (μ, Σ) και των $n_2 = n - n_1$ τελευταίων παρατηρήσεων είναι (μ_1, Σ_1) . Όταν $p = 1$, οι μη-παραμετρικοί έλεγχοι που χρησιμοποιούνται για την ανίχνευση αλλαγής του μέσου είναι ο έλεγχος του Mann-Whitney για δύο δείγματα (*Mann-Whitney two-sample test*) και ο έλεγχος αθροίσματος διατάξεων του Wilcoxon (*Wilcoxon rank-sum test*). Σε μεγαλύτερες διαστάσεις ($p > 1$), ο Liu (1995) πρότεινε έναν έλεγχο που θεωρείται γενίκευση του ελέγχου αθροίσματος διατάξεων του Wilcoxon.

Επιπλέον, θεωρούμε ότι έχουμε ένα p -διάστατο διάνυσμα, του οποίου η συνάρτηση κατανομής, όταν η διεργασία είναι εντός ελέγχου, είναι η F . Όταν υπάρχει κάποια ειδική

αιτία μεταβλητότητας, η αλλαγή στη διεργασία συνεπάγεται αλλαγή της θέσης ή/και αύξησης της κλίμακας και χαρακτηρίζεται ως απομάκρυνση από την κατανομή F σε μία συνεχή εκτός ελέγχου κατανομή G . Η απομάκρυνση αυτή έχει ως αποτέλεσμα τη μείωση του data depth. Ορίζουμε τη στατιστική συνάρτηση

$$Q(n_1) = \sum_{j=n_1+1}^n R_{n_1}(j),$$

όπου

$$R_{n_1}(j) = \#\{X_i \mid D_{F_{n_1+1}}(X_i) < D_{F_{n_1+1}}(X_j)\} + \frac{1}{2}\#\{X_i \mid D_{F_{n_1+1}}(X_i) = D_{F_{n_1+1}}(X_j)\}, \quad i = 1, \dots, n_1.$$

Η τυποποιημένη στατιστική συνάρτηση $SQ(n_1)$ ορίζεται από τον τύπο

$$SQ(n_1) = \frac{Q(n_1) - E(SQ(n_1))}{\sqrt{V(SQ(n_1))}}.$$

Καθώς τα n_1 και $n - n_1$ τείνουν στο άπειρο ($n_1 \rightarrow \infty, n - n_1 \rightarrow \infty$), ισχύει το εξής

$$SQ(n_1) = \frac{Q(n_1) - \frac{n_1(n-n_1)}{2}}{\sqrt{\frac{n_1(n-n_1)(n+1)}{12}}} \xrightarrow{D} N(0, 1).$$

Στο διάγραμμα τύπου *CUSUM* που προτείνουν οι Dai, Zhou and Wang (2011) απεικονίζεται το συσσωρευμένο άθροισμα

$$S_i = \max\{0, S_{i-1} - SQ_i - K\}, \quad (4.3.15)$$

όπου $S_0 = 0$ είναι η αρχική τιμή και $K = 2$ είναι η τιμή αναφοράς που προτείνουν οι συγγραφείς, ώστε το διάγραμμα να έχει καλή απόδοση. Εάν το S_i υπερβεί τη σταθερά απόφασης $h_{n,p}$, τότε η διεργασία τίθεται εκτός ελέγχου.

Οι Dai, Zhou and Wang έδωσαν στον Πίνακα 4.4 τιμές στο $h_{n,p}$ για διαφορετικές τιμές του n και του p . Ωστόσο, για ευκολία πρότειναν η τιμή του να υπολογίζεται από τον τύπο

$$h_{n,p} = 1.0936n - 1.4746p,$$

ενώ η θέση της μετατόπισης εκτιμάται από τον ακόλουθο τύπο

$$\hat{t} = \arg \max_{1 < t < n} \{|SQ_t|\},$$

όπου $\arg \max$ υποδηλώνει το σύνολο των σημείων για το οποίο η στατιστική συνάρτηση $|SQ_t|$ λαμβάνει μέγιστη τιμή.

n	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
30	32.14	30.50	29.14	26.54	25.24	23.53	22.98	21.62	20.11	17.92
40	42.26	41.99	38.43	37.61	35.97	35.15	34.19	32.96	32.01	31.05
50	55.11	51.56	48.82	46.64	45.54	46.09	45.00	42.81	41.71	40.62
60	69.06	64.68	63.59	60.85	60.31	57.57	57.03	55.39	53.75	54.29
70	78.90	75.62	72.34	68.78	67.42	66.60	65.23	64.00	63.04	64.14
80	92.03	85.46	83.82	80.00	79.45	77.81	76.99	76.99	75.62	75.35
90	101.80	95.31	92.98	91.48	90.93	88.75	86.28	86.01	84.92	86.01
100	110.60	108.40	130.20	100.70	98.59	98.59	94.21	92.03	93.12	93.94
110	119.30	116.00	115.50	113.90	112.80	111.70	107.80	108.40	107.30	107.80
120	134.60	129.70	127.50	121.50	120.40	119.90	119.30	118.80	118.40	116.00

Πίνακας 4.4: Τιμές του διαστήματος απόφασης $h_{n,p}$ για το πολυμεταβλητό μη-παραμετρικό διάγραμμα τύπου *CUSUM* των Dai, Zhou και Wang (2011)

4.3.3 Διάγραμμα Ελέγχου Εκθετικά Κινούμενου Μέσου

Οι Zou and Tsung (2011) παρουσίασαν ένα πολυμεταβλητό μη παραμετρικό διάγραμμα ελέγχου εκθετικά κινούμενου μέσου, το οποίο ονόμασαν το διάγραμμα *MSEWMA* (*Multivariate Sign EWMA control chart*), προτείνοντας τη χρήση της πολυδιάστατης διαμέσου ισο-μεταβολής (*Affine Equivariant Multivariate median, AEM-median*) των Hettmansperger and Randles (2002). Χρησιμοποιώντας το δείγμα αναφοράς μεγέθους m_0 θα πρέπει να υπολογίσουμε αρχικά την πολυδιάστατη διάμεσο θ_0 και τον μετασχηματισμένο πίνακα \mathbf{A}_0 . Οι ποσότητες θ_0 και \mathbf{A}_0 προκύπτουν από τη λύση των εξισώσεων

$$E\left(\frac{\mathbf{A}(\mathbf{X}-\theta)}{\|\mathbf{A}(\mathbf{X}-\theta)\|}\right) = \mathbf{0} \text{ και } E\left(\frac{\mathbf{A}(\mathbf{X}-\theta)(\mathbf{X}-\theta)' \mathbf{A}'}{\|\mathbf{A}(\mathbf{X}-\theta)\|^2}\right) = \frac{1}{p} \mathbf{I}_p,$$

ενώ οι αντίστοιχες δειγματικές ποσότητες $(\hat{\theta}_0, \hat{\mathbf{A}}_0)$ προσδιορίζονται από τη λύση των εξής δειγματικών εξισώσεων, οι οποίες βασίζονται στις m_0 ιστορικές παρατηρήσεις,

$$\frac{1}{m_0} \sum_{i=-m_0+1}^0 \left(\frac{\mathbf{A}(\mathbf{X}-\theta)}{\|\mathbf{A}(\mathbf{X}-\theta)\|}\right) = \mathbf{0} \text{ και } \frac{1}{m_0} \sum_{i=-m_0+1}^0 \left(\frac{\mathbf{A}(\mathbf{X}-\theta)(\mathbf{X}-\theta)' \mathbf{A}'}{\|\mathbf{A}(\mathbf{X}-\theta)\|^2}\right) = \frac{1}{p} \mathbf{I}_p.$$

Στον τύπο αυτό, \mathbf{A} είναι ένας $(p \times p)$ άνω τριγωνικός θετικά ορισμένος πίνακας με μονάδα στο άνω αριστερό στοιχείο, \mathbf{I}_p είναι ένας $(p \times p)$ ταυτοτικός πίνακας (*identity matrix*) και $\|\cdot\|$ είναι η Ευκλείδεια νόρμα. Εάν το m_0 είναι αρκετά μεγάλο, τότε μπορούμε να χρησιμοποιήσουμε την ασυμπτωτική μορφή $(\hat{\boldsymbol{\theta}}_0, \hat{\mathbf{A}}_0)$ του $(\boldsymbol{\theta}_0, \mathbf{A}_0)$. Αφού εκτιμήσουμε ή προσδιορίσουμε τα $(\boldsymbol{\theta}_0, \mathbf{A}_0)$, μπορούμε να τυποποιήσουμε και μετασχηματίσουμε τις παρατηρήσεις X_i μέσω του μοναδιαίου διανύσματος (*unit vector*) $\mathbf{u}_i, i = 1, 2, \dots$, που ορίζεται ως εξής

$$\mathbf{u}_i = \frac{\mathbf{A}_0(\mathbf{X}_i - \boldsymbol{\theta}_0)}{\|\mathbf{A}_0(\mathbf{X}_i - \boldsymbol{\theta}_0)\|}$$

ARL_0	λ	p								
		2	3	4	5	7	10	15	20	30
200	0.4	6.009	7.920	9.668	11.321	14.448	18.841	25.763	32.367	45.077
	0.2	7.831	9.830	11.674	13.414	16.708	21.329	28.539	35.405	48.534
	0.1	8.043	10.052	11.896	13.636	16.911	21.532	28.730	35.597	48.649
	0.05	7.225	9.177	10.963	12.646	15.819	20.288	27.294	33.994	46.806
	0.025	5.895	7.691	9.345	10.906	13.864	18.066	24.687	31.027	43.233
370	0.4	6.276	8.294	10.125	11.847	15.083	19.628	26.720	33.467	46.345
	0.2	8.567	10.687	12.626	14.448	17.876	22.649	30.118	37.152	50.608
	0.1	9.183	11.303	13.249	15.077	18.511	23.310	30.788	37.870	51.357
	0.05	8.605	10.700	12.607	14.404	17.774	22.472	29.831	36.769	50.032
	0.025	7.399	9.392	11.205	12.918	16.124	20.644	27.725	34.424	47.382
500	0.4	6.390	8.459	10.329	12.083	15.388	19.983	27.151	33.994	47.036
	0.2	8.904	11.074	13.058	14.924	18.409	23.284	30.835	37.965	51.530
	0.1	9.716	11.887	13.877	15.750	19.247	24.147	31.745	38.922	52.567
	0.05	9.265	11.417	13.375	15.216	18.663	23.462	30.931	38.013	51.530
	0.025	8.126	10.198	12.081	13.852	17.165	21.812	29.065	35.956	49.110

Πίνακας 4.5: Τιμές ορίου ελέγχου L του διαγράμματος $MSEWMA$ για διαφορετικές τιμές των λ, p και ARL_0

Στη συνέχεια, ορίζουμε την ακολουθία του εκθετικά ζυγισμένου κινούμενου μέσου ($EWMA$ *sequence*) ως εξής

$$\mathbf{w}_t = (1 - \lambda) \cdot \mathbf{w}_{t-1} + \lambda \cdot \mathbf{u}_t,$$

όπου ως αρχική τιμή λαμβάνεται η $\mathbf{w}_0 = \mathbf{0}$. Στο διάγραμμα *MSEWMA* ενεργοποιείται συναγερμός όταν

$$Q_t = \frac{2-\lambda}{\lambda} p \mathbf{w}'_t \mathbf{w}_t > L,$$

όπου $L > 0$ είναι ένα όριο ελέγχου που καθορίζεται για το επιθυμητό ARL_0 . Προφανώς, η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα *MSEWMA* είναι η ποσότητα Q_t . Οι Zou and Tsung (2011) δίνουν τιμές στο όριο ελέγχου L για διάφορες τιμές των λ , p και ARL_0 , όταν η διεργασία προέρχεται από οποιαδήποτε συνεχή κατανομή.

4.3.4 Διάγραμμα Ελέγχου Κινούμενου Μέσου που βασίζεται στο βάθος των δεδομένων

Οι Liu, Singh and Teng (2004) παρουσίασαν ένα μη παραμετρικό πολυδιάστατο διάγραμμα ελέγχου κινούμενου μέσου, το οποίο βασίζεται στο data depth. Το διάγραμμα αυτό το ονόμασαν *DDMA* (*Nonparametric Multivariate Data Depth Moving Average chart*). Σύμφωνα με τους συγγραφείς, θεωρούμε ότι ένα δείγμα αναφοράς $\mathbf{Y} = (Y_1, Y_2, \dots, Y_m)$ από την κατανομή F και ένα δείγμα μελλοντικών παρατηρήσεων $\mathbf{X} = (X_1, X_2, \dots, X_n)$. Κατασκευάζουμε το διάγραμμα *DDMA* με σκοπό την παρακολούθηση των κινούμενων μέσων με μήκος q (*moving averages with length q*), δηλαδή των ποσοτήτων $\tilde{X}_q = (X_1 + X_2 + \dots + X_q)/q$, $\tilde{X}_{q+1} = (X_2 + X_3 + \dots + X_{q+1})/q$, ..., $\tilde{X}_n = (X_{n-q+1} + \dots + X_n)/q$. Εάν $\tilde{\mathbf{X}} = (\tilde{X}_q, \dots, \tilde{X}_n)$, τότε το αντίστοιχο δείγμα αναφοράς για την παρακολούθηση των $\tilde{X}_i \in \tilde{\mathbf{X}}$ είναι το $\tilde{\mathbf{Y}} = (\tilde{Y}_q, \dots, \tilde{Y}_m)$, όπου $\tilde{Y}_q = (Y_1 + Y_2 + \dots + Y_q)/q$, $\tilde{Y}_{q+1} = (Y_2 + Y_3 + \dots + Y_{q+1})/q$, ..., $\tilde{Y}_m = (Y_{m-q+1} + \dots + Y_m)/q$. Συνεπώς, για κάθε $\tilde{X}_i \in \tilde{\mathbf{X}}$ υπολογίζουμε τον αντίστοιχο βαθμό (*relative rank*) σε σχέση με το $\tilde{\mathbf{Y}} = (\tilde{Y}_q, \dots, \tilde{Y}_m)$ ως εξής

$$r_{\tilde{G}_{m-q+1}}(\tilde{X}_i) = \frac{\#\{\tilde{Y}_j \mid D_{\tilde{G}_{m-q+1}}(Y_j) < D_{\tilde{G}_{m-q+1}}(X_i), j = q, \dots, m\}}{m - q + 1}, \quad i = q, \dots, n, \quad (4.3.16)$$

όπου \tilde{G}_{m-q+1} είναι η εμπειρική κατανομή (*empirical distribution*) του \tilde{Y} και $D_{\tilde{G}_{m-q+1}}(\cdot)$ είναι το εμπειρικό βάθος (*empirical depth*) που υπολογίζεται σε σχέση με το $\tilde{Y} = (\tilde{Y}_q, \dots, \tilde{Y}_m)$. Στο διάγραμμα *DDMA* απεικονίζονται οι ποσότητες $r_{\tilde{G}_{m-q+1}}(\tilde{X}_i)$ για $i = q, \dots, n$ με όρια ελέγχου

$$\text{Διάγραμμα } DDMA \begin{cases} CL & = & 0.5 \\ LCL & = & a. \end{cases} \quad (4.3.17)$$

Στην ειδική περίπτωση $q=1$, το διάγραμμα *DDMA* είναι στην πραγματικότητα το διάγραμμα r της Liu (1995). Οι Liu, Singh and Teng (2004) σύγκριναν το διάγραμμα *DDMA* με το διάγραμμα Q της Liu (1995) και κατέληξαν στο συμπέρασμα ότι το πρώτο διάγραμμα ελέγχου είναι πιο ευαίσθητο σε μικρές μετατοπίσεις της θέσης. Αυτό γίνεται εύκολα αντιληπτό μέσω του επόμενου παραδείγματος. Υποθέτουμε ότι το μήκος των κινούμενων μέσων είναι $q = 2$. Επιπλέον, υποθέτουμε ότι $Y_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ και $X_i \sim N(\boldsymbol{\mu} + \boldsymbol{\delta}, \boldsymbol{\Sigma})$, για $\boldsymbol{\delta} \neq \mathbf{0}$ και $i = 1, 2$. Οπότε $\tilde{Y} = ((Y_1 + Y_2)/2) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/2)$ και $\tilde{X} = ((X_1 + X_2)/2) \sim N(\boldsymbol{\mu} + \boldsymbol{\delta}, \boldsymbol{\Sigma}/2)$. Μετασηματίζοντας τις μεταβλητές \tilde{X} και \tilde{Y} , ισχύουν τα εξής: $(\sqrt{2}\tilde{Y} - \sqrt{2}\boldsymbol{\mu} + \boldsymbol{\mu}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ και $(\sqrt{2}\tilde{X} - \sqrt{2}\boldsymbol{\mu} + \boldsymbol{\mu}) \sim N(\boldsymbol{\mu} + \sqrt{2}\boldsymbol{\delta}, \boldsymbol{\Sigma})$. Παρατηρούμε ότι η μετατόπιση της παραμέτρου θέσης των αρχικών δεδομένων από $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ σε $N(\boldsymbol{\mu} + \boldsymbol{\delta}, \boldsymbol{\Sigma})$ είναι $\boldsymbol{\delta}$, ενώ η μετατόπιση της θέσης των κινούμενων μέσων \tilde{X} και \tilde{Y} είναι $\sqrt{2}\boldsymbol{\delta}$. Ο πίνακας $\boldsymbol{\Sigma}$ παραμένει σταθερός και στις δύο περιπτώσεις. Συνεπώς, η μετατόπιση της παραμέτρου θέσης στο διάγραμμα *DDMA* είναι $\sqrt{2}\boldsymbol{\delta}$, η οποία είναι μεγαλύτερη από την αντίστοιχη μετατόπιση $\boldsymbol{\delta}$ των διαγραμμάτων r και Q της Liu (1995). Ως αποτέλεσμα, το διάγραμμα *DDMA* είναι πιο ευαίσθητο. Στη γενική περίπτωση που το μήκος των κινούμενων μέσων είναι $q > 1$, η μετατόπιση της παραμέτρου θέσης του διαγράμματος *DDMA* είναι της τάξης $\sqrt{q}\boldsymbol{\delta}$ με σταθερό πίνακα διακύμανσης-συνδιακύμανσης $\boldsymbol{\Sigma}$. Τέλος, οι Liu, Singh and Teng (2004) υποστηρίζουν ότι είναι προτιμότερο να χρησιμοποιούμε το διάγραμμα T^2 του Hotelling, όταν η κατανομή της διεργασίας είναι κανονική.

4.3.5 Άλλα Πολυμεταβλητά Μη Παραμετρικά Διαγράμματα Ελέγχου

Τα τελευταία χρόνια ολοένα και περισσότεροι ερευνητές ασχολούνται με την κατασκευή μη παραμετρικών διαγραμμάτων ελέγχου. Παρακάτω θα αναφέρουμε μερικούς από αυτούς χωρίς να παρουσιάσουμε αναλυτικά το έργο τους, καθώς τα διαγράμματα αυτά δεν αποτελούν το κύριο αντικείμενο μελέτης της παρούσας διπλωματικής εργασίας. Οι Liu and Tang (1996) βασίστηκαν στη μέθοδο Bootstrap και στην προσομοίωση και ανέπτυξαν διαγράμματα ελέγχου για εξαρτημένες και ανεξάρτητες μετρήσεις, των οποίων η κατανομή δεν είναι κανονική. Οι Abu-Shawiesh and Abdullah (2001) ανέπτυξαν ένα ισχυρό (*robust*) διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση της θέσης μιας δισδιάστατης διεργασίας, χρησιμοποιώντας τους εκτιμητές των Hodges-Lehmann και Shamos-Bickel-Lehmann. Οι Chou, Mason and Young (2001) πρότειναν μία μεθοδολογία για τον προσδιορισμό των ορίων ελέγχου στην περίπτωση μεμονωμένων παρατηρήσεων, οι οποίες δεν προέρχονται από πολυδιάστατη κανονική κατανομή. Επιπλέον, οι Qiu and Hawkins (2001) παρουσίασαν ένα πολυμεταβλητό μη παραμετρικό διάγραμμα τύπου *CUSUM*, χρησιμοποιώντας τη σειρά της πληροφορίας μεταξύ και εντός των συνιστωσών των μετρήσεων, καθώς και των εντός ελέγχου μέσω των μετρήσεων. Λίγο αργότερα, οι Qiu and Hawkins (2003) παρουσίασαν ένα αντίστοιχο διάγραμμα που βασίζεται στους βαθμούς των μετρήσεων.

Επιπρόσθετα, οι Sun and Tsung (2003) δημιούργησαν ένα πολυδιάστατο μη παραμετρικό διάγραμμα που βασίζεται στην απόσταση του πυρήνα, δηλαδή στην απόσταση μεταξύ του κέντρου-πυρήνα και του νέου δείγματος που πρόκειται να παρακολουθήσουμε. Οι Zarate and Okogbaa (2002) ανέπτυξαν μία μη παραμετρική μέθοδο για την παρακολούθηση πολλών μεταβλητών ταυτόχρονα, χρησιμοποιώντας την Ανάλυση Κυρίων Συνιστωσών. Οι Stoumbos and Sullivan (2002) εξέτασαν τις επιπτώσεις της μη κανονικότητας στην αποτελεσματικότητα του διαγράμματος *EWMA*. Οι Messaoud, Weihs and Hering (2004) πρότειναν ένα μη παραμετρικό διάγραμμα *EWMA* που βασίζεται στο βάθος των δεδομένων. Οι Thissen, Swierenga, De Weijer, Melssen and Buydens (2005) πρότειναν μία μέθοδο που συνδυάζει το Μη Παραμετρικό Έλεγχο Ποιότητας με μοντέλα συσταδοποίησης. Ο Qiu (2008) παρουσίασε ένα πολυδιάστατο ελεύθερο κατανομής διάγραμμα που βασίζεται στη λογαριθμογραμμική διαδικασία μοντελοποίησης, ενώ οι Zou and Qiu (2009) παρουσίασαν ένα πολυδιάστατο μη παραμετρικό διάγραμμα που βασίζεται στη μέθοδο LASSO, η οποία μπορεί να επιλέξει με

ακρίβεια μη μηδενικούς συντελεστές (πολυδιάστατης) παλινδρόμησης. Τέλος, οι Bush, et al. (2010) πρότειναν το διάγραμμα $k-LINK$, το οποίο είναι ένα μη παραμετρικό διάγραμμα ελέγχου και βασίζεται στον αλγόριθμο συνδέσμων βαθμών (*Linkage Ranking Algorithm*). Επίσης, παρουσίασαν και το διάγραμμα $EWMA$ $k-LINK$ το οποίο αποτελεί την εκθετικά ζυγισμένη εκδοχή κινούμενου μέσου του διαγράμματος $k-LINK$.

Όπως στον Παραμετρικό Έλεγχο Ποιότητας μας ενδιαφέρει η ανίχνευση της μεταβλητής ή των μεταβλητών εκείνων που ενεργοποιούν το συναγερμό σε μία πολυμεταβλητή διεργασία, έτσι και στον Μη Παραμετρικό Έλεγχο Ποιότητας ενδιαφερόμαστε να εντοπίσουμε την αιτία που προκάλεσε ένα εκτός ελέγχου σήμα. Είναι γεγονός ότι δεν έχουν γίνει πολλές μελέτες πάνω σε αυτό το ζήτημα. Ωστόσο, θα ήταν παράλειψη να μην αναφέρουμε ότι οι Zarate and Okogbaa (2004) ανέπτυξαν μία μεθοδολογία για τον εντοπισμό των μεταβλητών που προκάλεσαν τα εκτός ελέγχου σήματα σε ένα μη παραμετρικό διάγραμμα ελέγχου, η οποία βασίζεται στο βάθος των δεδομένων και στην Ανάλυση Κύριων Συνιστωσών. Επιπλέον, οι Jensen, Birch and Woodall (2007) δημιούργησαν μία μεθοδολογία ανίχνευσης των ακραίων τιμών (*outliers*) των πολυδιάστατων δεδομένων πριν τον υπολογισμό των ορίων ελέγχου Φάσης I, ώστε με τις απαραίτητες διορθωτικές κινήσεις τα όρια Φάσης II που θα προκύψουν να είναι πιο αποτελεσματικά.

4.4 Παραδείγματα Πολυμεταβλητών Μη Παραμετρικών Διαγραμμάτων

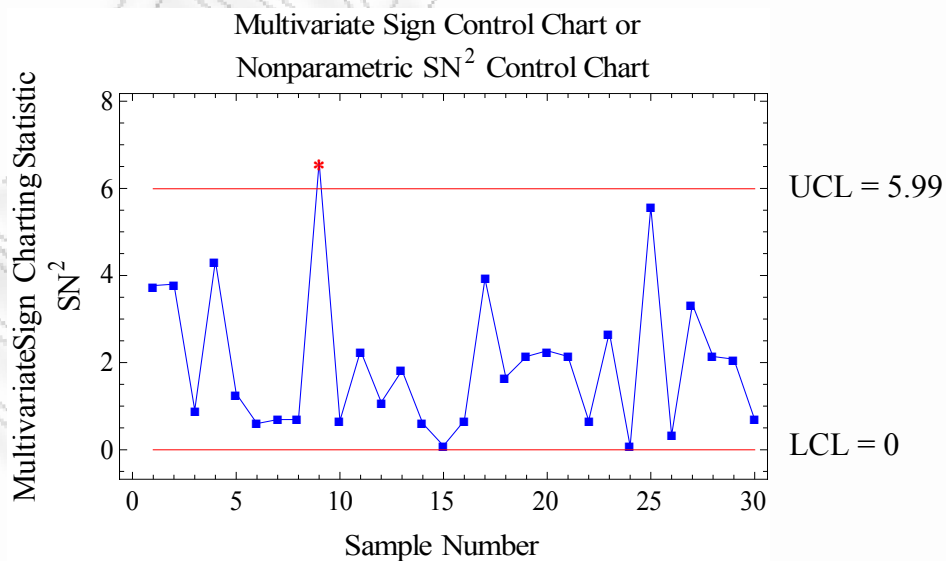
Για την κατασκευή των πολυμεταβλητών μη παραμετρικών διαγραμμάτων SN^2 και SR^2 θα χρησιμοποιήσουμε τα κανονικά δεδομένα της Παραγράφου 3.5 του προηγούμενου κεφαλαίου. Θα πρέπει να αναφέρουμε ότι για την κατασκευή των διαγραμμάτων αυτών, τα δεδομένα δεν είναι απαραίτητο να προέρχονται από την κανονική κατανομή. Ωστόσο, χρησιμοποιούμε τα δεδομένα αυτά για να κατασκευάσουμε το διάγραμμα T^2 του Hotelling και να συγκρίνουμε την αποτελεσματικότητά του με αυτή των διαγραμμάτων SN^2 και SR^2 .

Στον Πίνακα 4.6, παραθέτουμε τις τιμές των στατιστικών συναρτήσεων των διαγραμμάτων SN^2 και SR^2 , οι οποίες υπολογίζονται από τους τύπους (4.3.1) και (4.3.3), αντίστοιχα. Οι τιμές των συναρτήσεων προέκυψαν από τα 30 προκαταρκτικά δείγματα. Στα

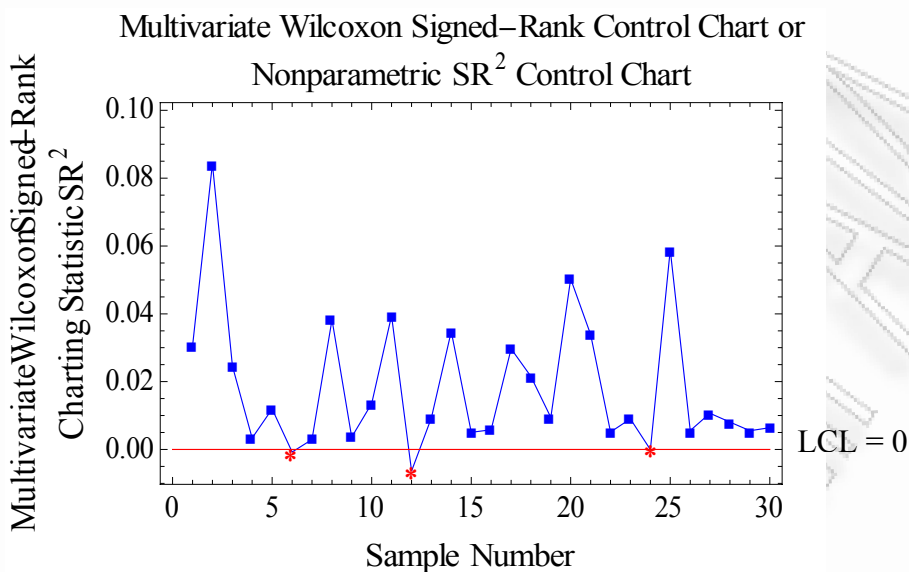
Σχήματα 4.1 και 4.2 απεικονίζονται τα διαγράμματα SN^2 και SR^2 , αντίστοιχα για $a = 0.05$, οπότε τα όρια ελέγχου θα υπολογίζονται από τους τύπους (4.3.1) και (4.3.3) ως εξής

$$\begin{cases} UCL = \chi^2_{p; a} = \chi^2_{2; 0.05} = 5.99 \\ LCL = 0. \end{cases}$$

Από το Σχήμα 4.1 παρατηρούμε ότι κανένα σημείο δεν είναι κάτω από το μηδέν. Ωστόσο, η διεργασία τίθεται εκτός στατιστικού ελέγχου, διότι το ένατο δείγμα βρίσκεται πάνω από το άνω όριο ελέγχου. Αντίθετα, στο Σχήμα 4.2, παρατηρούμε ότι κανένα σημείο δεν ξεπερνά το άνω όριο ελέγχου, επειδή οι τιμές της στατιστικής συνάρτησης SR^2 κυμαίνονται γύρω από το μηδέν. Όμως, τα δείγματα 6, 12 και 24 βρίσκονται κάτω από το κάτω όριο ελέγχου. Συνεπώς, η διεργασία είναι εκτός ελέγχου. Σύμφωνα με το Σχήμα 4.3, η διεργασία τίθεται εκτός ελέγχου λόγω του δεύτερου και του τέταρτου δείγματος. Το συμπέρασμα στο οποίο καταλήγουμε είναι ότι υπάρχουν ειδικές αιτίες μεταβλητότητας και θα πρέπει να προβούμε σε διορθωτικές κινήσεις. Στο σημείο αυτό, θα πρέπει να αναφερθεί ότι το παραμετρικό διάγραμμα ανίχνευσε από το δεύτερο κιάλας δείγμα την αλλαγή του μέσου της διεργασίας. Με τον τρόπο αυτό, επιβεβαιώνεται ότι τα μη παραμετρικά διαγράμματα δεν είναι τόσο γρήγορα στην ανίχνευση ειδικών αιτιών όσο τα αντίστοιχα παραμετρικά, όταν υπάρχει πλήρης γνώση της κατανομής.



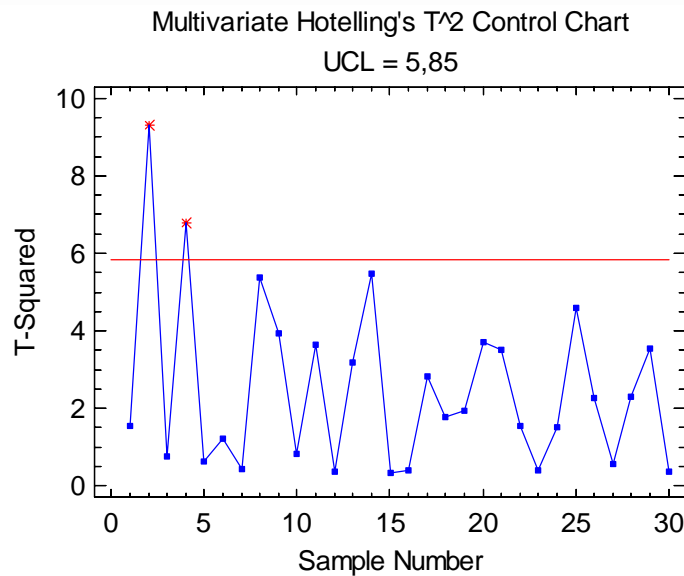
Σχήμα 4.1: Μη παραμετρικό διάγραμμα SN^2 ($a = 0.05$)



Σχήμα 4.2: Μη παραμετρικό διάγραμμα SR^2 ($\alpha = 0.05$)

Αριθμός Δείγματος (k ή t)	Στατιστικές Συναρτήσεις		Αριθμός Δείγματος (k ή t)	Στατιστικές Συναρτήσεις	
	SN^2	SR^2		SN^2	SR^2
1	3.769	0.031	16	0.667	0.006
2	3.800	0.084	17	3.923	0.030
3	0.091	0.025	18	1.667	0.022
4	4.333	0.003	19	2.143	0.009
5	1.286	0.012	20	2.273	0.050
6	0.600	-0.001	21	2.143	0.034
7	0.692	0.003	22	0.667	0.005
8	0.692	0.038	23	2.692	0.009
9	6.600	0.004	24	0.077	0.000
10	0.667	0.014	25	5.571	0.059
11	2.273	0.039	26	0.333	0.006
12	1.091	-0.006	27	3.333	0.011
13	1.818	0.009	28	2.143	0.008
14	0.600	0.035	29	2.077	0.006
15	0.091	0.005	30	0.692	0.006

Πίνακας 4.6: Τιμές στατιστικών συναρτήσεων των διαγραμμάτων SN^2 και SR^2 .



Σχήμα 4.3: Παραμετρικό διάγραμμα T^2 του Hotelling ($\alpha = 0.05$)

Θα ήταν παράλειψη να μην αναφέρουμε ότι οι τιμές της στατιστικής συνάρτησης του διαγράμματος T^2 υπολογίζονται από τον τύπο (3.2.3) και είναι αυτές του Πίνακα 3.3, ενώ τα όρια ελέγχου για $\alpha = 0.05$ υπολογίζονται από τον τύπο (3.2.4) ως εξής

$$\begin{cases} UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} \times F_{p, mn-m-p+1; \alpha} = \frac{2(30-1)(15-1)}{30 \cdot 15 - 30 - 2 + 1} \times F_{2, 419; 0.05} = \frac{812}{419} \times 3.017 = 5.847 \\ LCL = 0. \end{cases}$$

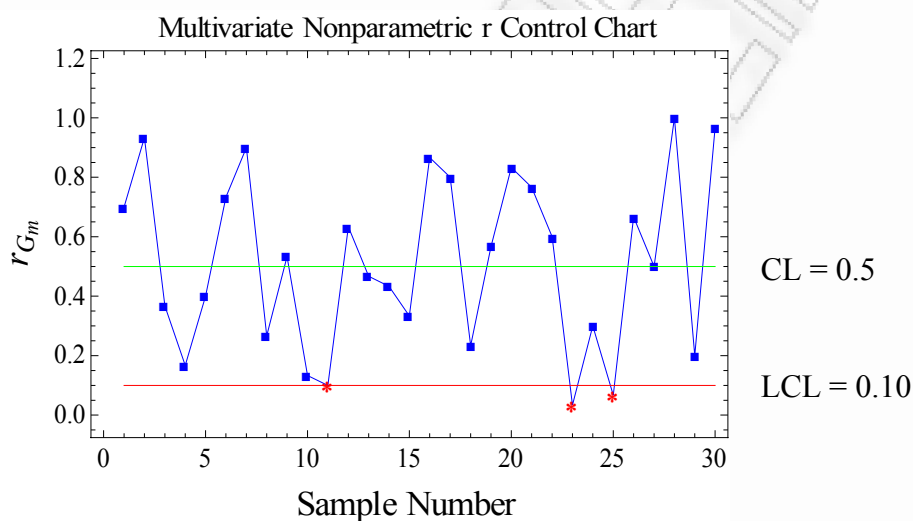
Στη συνέχεια, για την παρουσίαση των υπόλοιπων μη παραμετρικών διαγραμμάτων ελέγχου του κεφαλαίου αυτού, κατασκευάσαμε $m = 20$ τυχαία δείγματα (Y_1, \dots, Y_m) , τα οποία προέρχονται από την κατανομή G , και $n = 10$ τυχαία δείγματα (X_1, \dots, X_n) , τα οποία ακολουθούν μία κατανομή F . Για να συγκρίνουμε τα διαγράμματα αυτά με τη στατιστική συνάρτηση T^2 του Hotelling, θεωρούμε ότι

$$G \sim N\left(\begin{pmatrix} 100 \\ 10 \end{pmatrix}, \begin{pmatrix} 4 & 1.6 \\ 1.6 & 1 \end{pmatrix}\right) \text{ και } F \sim N\left(\begin{pmatrix} 120 \\ 12 \end{pmatrix}, \begin{pmatrix} 4 & 1.6 \\ 1.6 & 1 \end{pmatrix}\right).$$

Για κάθε ένα από τα 30 δείγματα υπολογίζουμε το βάθος τους σε σχέση με την κατανομή G . Το βάθος των δεδομένων που θα χρησιμοποιήσουμε στο συγκεκριμένο παράδειγμα είναι αυτό του Mahalanobis, το οποίο ορίζεται από τον τύπο (4.3.5). Για την κατασκευή των

διαγραμμάτων r , Q , S και S^* είναι απαραίτητος ο υπολογισμός της ποσότητας r_{G_m} μέσω του τύπου (4.3.6) και η απεικόνισή τους στο διάγραμμα r . Το κάτω όριο ελέγχου και η κεντρική γραμμή του διαγράμματος r υπολογίζονται μέσω του τύπου (4.3.7) και για $a = 0.10$ είναι τα ακόλουθα

$$\begin{cases} CL = 0.5 \\ LCL = a = 0.10. \end{cases}$$



Σχήμα 4.4: Μη παραμετρικό διάγραμμα r ($a = 0.10$)

Από το Σχήμα 4.4 παρατηρούμε ότι τα σημεία 23 και 25 βρίσκονται κάτω από το κάτω όριο ελέγχου, ενώ το σημείο 11 βρίσκεται πάνω στο κάτω όριο ελέγχου. Συνεπώς, η διεργασία βρίσκεται εκτός στατιστικού ελέγχου.

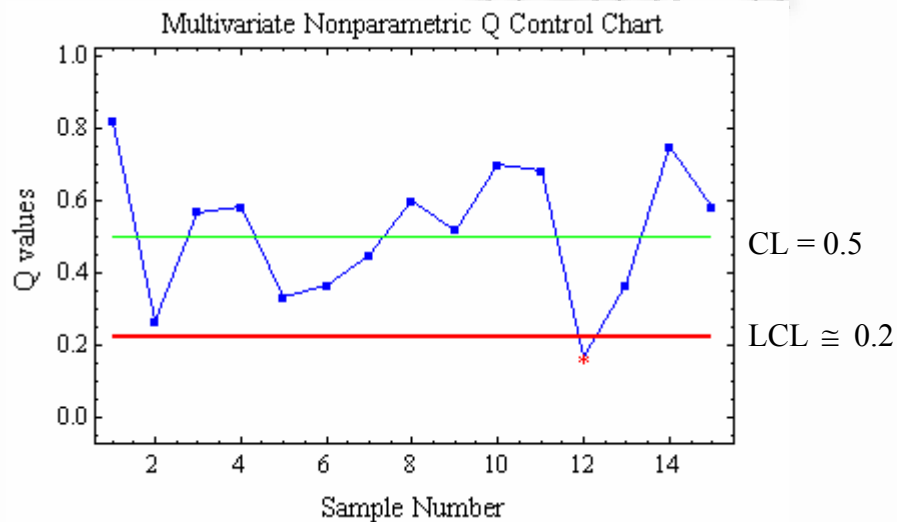
Για την κατασκευή του διαγράμματος Q υπολογίζουμε την $Q(G_m, F_n)$ από τον τύπο (4.3.8). Για μεγάλο n το κάτω όριο ελέγχου και η κεντρική γραμμή υπολογίζονται από τον τύπο (4.3.9), ενώ για μικρό a από τον τύπο (4.3.10). Οι τιμές τους για $a = 0.10$ είναι αντίστοιχα οι εξής

$$\begin{cases} CL = 0.5 \\ LCL_1 = 0.5 - z_a \sqrt{\frac{1}{12} \left(\frac{1}{m} + \frac{1}{n} \right)} = 0.5 - z_{0.10} \sqrt{\frac{1}{12} \left(\frac{1}{m} + \frac{1}{n} \right)} = 0.5 - 1.28 \sqrt{\frac{1}{12} \left(\frac{1}{20} + \frac{1}{2} \right)} = 0.2256 \end{cases}$$

και

$$\begin{cases} CL = 0.5 \\ LCL_2 = \frac{(n!a)^{1/n}}{n} = \frac{(2! \cdot 0.10)^{1/2}}{2} = 0.2236. \end{cases}$$

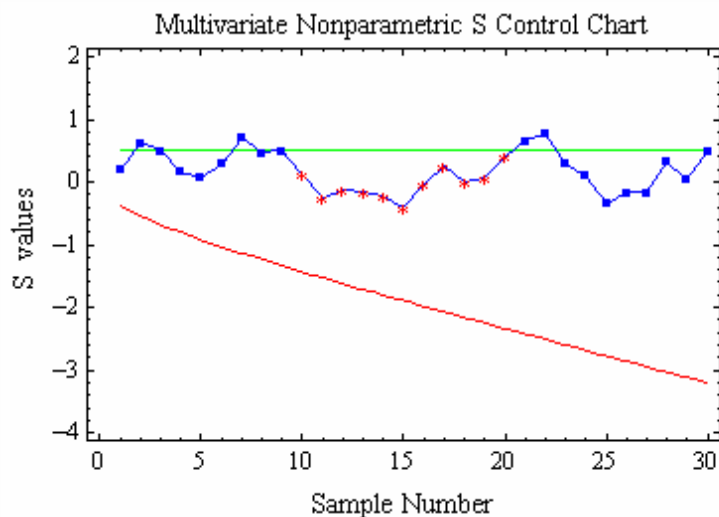
Παρατηρούμε ότι $LCL_1 = LCL_2 \cong 0.2$. Από το Σχήμα 4.5 που ακολουθεί, γίνεται φανερό ότι η διεργασία είναι εκτός ελέγχου, εφόσον το σημείο 12 βρίσκεται κάτω από τα LCL_1 και LCL_2 .



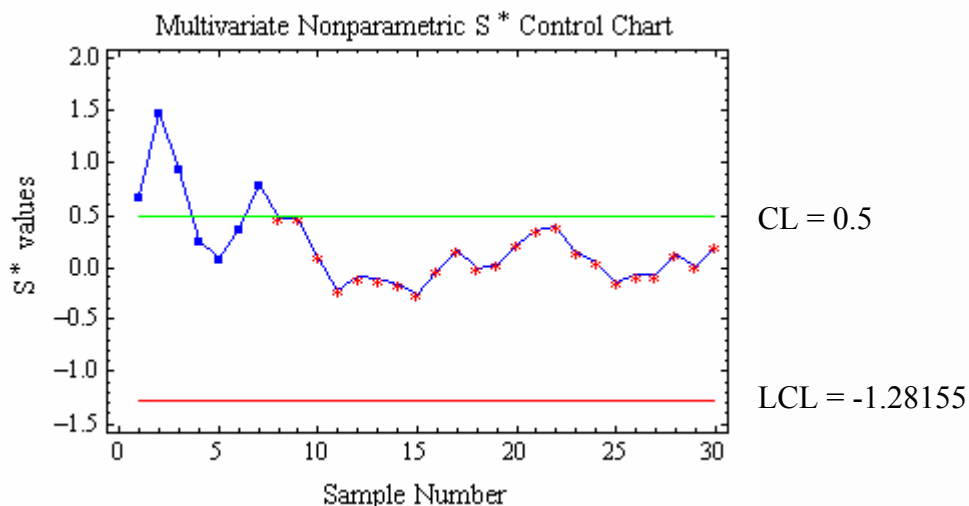
Σχήμα 4.5: Μη παραμετρικό διάγραμμα Q ($a = 0.10$)

Η κατασκευή του διαγράμματος S (συσσωρευμένου αθροίσματος) του Σχήματος 4.6, απαιτεί τον υπολογισμό της ποσότητας $S_n(G_m)$ που γίνεται μέσω του τύπου (4.3.11) και αποτελεί τη στατιστική συνάρτηση του διαγράμματος. Τα όρια ελέγχου του διαγράμματος S είναι αυτά του τύπου (4.3.12), με το κάτω όριο να είναι καμπύλη. Χρησιμοποιώντας την τυποποιημένη ποσότητα $S_n^*(G_m)$ του τύπου (4.3.13), προκύπτει το διάγραμμα S^* του Σχήματος 4.7, το οποίο έχει τα επόμενα όρια ελέγχου (με βάση τον τύπο (4.3.14)).

$$\begin{cases} CL = 0.5 \\ LCL_1 = -z_a = -z_{0.10} = -1.28155. \end{cases}$$



Σχήμα 4.6: Μη παραμετρικό διάγραμμα S τύπου $CUSUM$ ($\alpha = 0.10$)



Σχήμα 4.7: Μη παραμετρικό διάγραμμα S^* τύπου $CUSUM$ ($\alpha = 0.10$)

Από τα Σχήματα 4.6 και 4.7 είναι φανερό ότι κανένα σημείο δε βρίσκεται κάτω από το κάτω όριο ελέγχου. Ωστόσο, η διεργασία δεν είναι εντός στατιστικού ελέγχου, διότι πάνω από οχτώ συνεχόμενα σημεία είναι κάτω από την κεντρική γραμμή. Κατασκευάσαμε και το διάγραμμα τύπου $CUSUM$ των Dai, Zhou και Wang (2011), αλλά οι τιμές της στατιστικής συνάρτησης του τύπου (4.3.15) είναι όλες μηδενικές. Για το λόγο αυτό, δεν είχε νόημα η απεικόνισή του. Στον Πίνακα 4.7 που ακολουθεί, παρουσιάζονται οι τιμές των στατιστικών συναρτήσεων των διαγραμμάτων r , Q , S και S^* .

Δείγμα	Τιμές Χαρακτηριστικών		Βάθος του Mahalanobis	Στατιστικές Συναρτήσεις			
	Αντοχή στη θραύση	Διάμετρος	MD_{G_m}	r_{G_m}	$Q(G_m, F_n)$	$S_n(G_m)$	$S_n^*(G_m)$
Y ₁	97.797	9.009	0.575	0.700	0.8167	0.200	0.682
Y ₂	100.298	10.368	0.864	0.933		0.633	1.502
Y ₃	95.990	8.227	0.277	0.367	0.267	0.500	0.953
Y ₄	97.903	8.414	0.240	0.167		0.167	0.271
Y ₅	97.554	9.646	0.283	0.400	0.567	0.067	0.096
Y ₆	98.481	9.733	0.598	0.733		0.300	0.387
Y ₇	100.010	10.305	0.831	0.900	0.583	0.700	0.825
Y ₈	97.209	8.293	0.275	0.267		0.467	0.508
Y ₉	101.875	11.307	0.391	0.533	0.333	0.500	0.506
Y ₁₀	95.683	8.572	0.218	0.133		0.133	0.126
Y ₁₁	100.135	9.081	0.164	0.100	0.367	-0.267	-0.238
Y ₁₂	101.719	11.048	0.515	0.633		-0.133	-0.113
Y ₁₃	96.841	8.958	0.360	0.467	0.450	-0.167	-0.134
Y ₁₄	102.588	10.889	0.327	0.433		-0.233	-0.178
Y ₁₅	103.271	11.756	0.276	0.333	0.600	-0.400	-0.292
Y ₁₆	100.955	10.467	0.766	0.867		-0.033	-0.023
Y ₁₇	100.998	10.372	0.700	0.800	0.517	0.267	0.179
Y ₁₈	103.503	11.714	0.268	0.233		0.000	0.000
Y ₁₉	102.208	11.253	0.427	0.567	0.700	0.067	0.041
Y ₂₀	98.870	9.822	0.746	0.833		0.400	0.240
X ₁	98.735	9.843	0.638	0.767	0.683	0.667	0.387
X ₂	101.445	10.382	0.478	0.600		0.767	0.430
X ₃	97.113	10.083	0.110	0.033	0.168	0.300	0.163
X ₄	101.333	9.981	0.275	0.300		0.100	0.053
X ₅	104.844	11.746	0.145	0.067	0.367	-0.333	-0.171
X ₆	100.845	10.807	0.568	0.667		-0.167	-0.083
X ₇	101.157	11.135	0.378	0.500	0.750	-0.167	-0.081
X ₈	99.693	9.838	0.938	1.000		0.333	0.157
X ₉	100.971	9.730	0.248	0.200	0.583	0.033	0.015
X ₁₀	99.143	9.848	0.887	0.967		0.500	0.224

Πίνακας 4.7: Τιμές στατιστικών συναρτήσεων των διαγραμμάτων r , Q , S και S^* .

Για την κατασκευή του διαγράμματος κινούμενου μέσου *DDMA* πρέπει να δημιουργήσουμε το διάνυσμα των κινούμενων μέσων $\tilde{\mathbf{X}} = (\tilde{X}_3, \dots, \tilde{X}_{10})$ μήκους $q = 3$ και το αντίστοιχο δείγμα αναφοράς $\tilde{\mathbf{Y}} = (\tilde{Y}_3, \dots, \tilde{Y}_{20})$. Η στατιστική συνάρτηση του διαγράμματος υπολογίζεται από τον τύπο (4.3.16) και είναι ο βαθμός $r_{\tilde{G}_{m-q+1}}(\tilde{X}_i)$ σε σχέση με την εμπειρική κατανομή του $\tilde{\mathbf{Y}}$. Η κατανομή αυτή εκτιμήθηκε από το δείγμα αναφοράς και είναι η

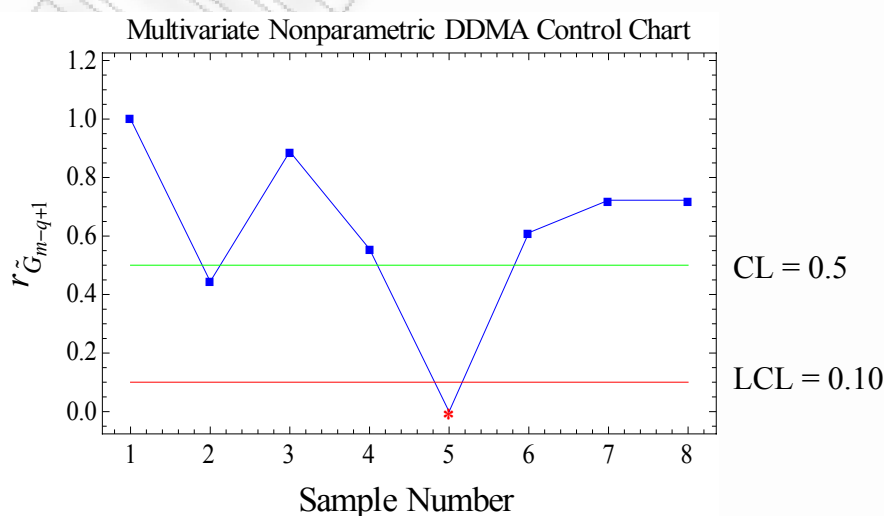
$$\tilde{G}_{m-q+1=18} \sim N\left(\begin{pmatrix} 99.738 \\ 9.791 \end{pmatrix}, \begin{pmatrix} 0.302 & 0.132 \\ 0.132 & 0.072 \end{pmatrix}\right).$$

Τα όρια ελέγχου του διαγράμματος *DDMA* δίνονται από τον τύπο (4.3.17) και είναι τα εξής

$$\begin{cases} CL = 0.5 \\ LCL = a = 0.10. \end{cases}$$

Οι τιμές της στατιστικής συνάρτησης του διαγράμματος T^2 υπολογίζονται από τον τύπο $T_k^2 = (\mathbf{X}_k - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X}_k - \bar{\mathbf{X}})$ με $k = 1, 2, \dots, 30$ και δίνονται στην τελευταία στήλη του Πίνακα 4.8, ενώ τα όρια ελέγχου για $a = 0.10$ υπολογίζονται από τον τύπο (3.2.8) ως εξής

$$\begin{cases} UCL = \frac{(m-1)^2}{m} \times B_{p/2, (m-p+1)/2; a} = \frac{(30-1)^2}{30} \times B_{1, 13.5; 0.10} = \frac{841}{30} \times 0.1568 = 4.396 \\ LCL = 0. \end{cases}$$

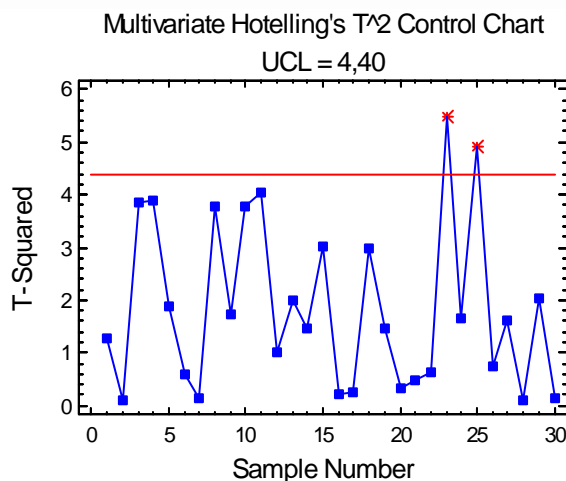


Σχήμα 4.8: Μη παραμετρικό διάγραμμα κινούμενου μέσου *DDMA* ($a = 0.10$)

Από το Σχήμα 4.8 παρατηρούμε ότι το πέμπτο σημείο βρίσκεται κάτω από το κάτω όριο ελέγχου και ως αποτέλεσμα η διεργασία είναι εκτός ελέγχου. Στο Σχήμα 4.9 απεικονίζεται το παραμετρικό διάγραμμα T^2 του Hotelling, το οποίο δίνει σήμα εκτός ελέγχου στα σημεία 23 και 25. Δεδομένου ότι ο μέσος της διεργασίας μετατοπίζεται από το 20^ο δείγμα και μετά, το παραμετρικό διάγραμμα ανιχνεύει την αλλαγή αυτή σχετικά σύντομα.

Δείγμα	Τιμές Χαρακτηριστικών		Διάνυσμα Κινούμενων Μέσων \tilde{X} και Δείγμα Αναφοράς \tilde{Y} για $q = 3$		Βάθος του Mahalanobis	Βαθμός για κάθε $\tilde{X}_i \in \tilde{X}$	Στατιστική Συνάρτηση
	Αντοχή στη θραύση	Διάμετρος			$MD_{\tilde{G}_{m-q+1}}$	$r_{\tilde{G}_{m-q+1}}$	T^2
Y ₁	97.797	9.009	-	-	-	-	1.262
Y ₂	100.298	10.368	-	-	-	-	0.115
Y ₃	95.990	8.227	98.028	9.201	0.093	-	3.849
Y ₄	97.903	8.414	98.064	9.003	0.071	-	3.886
Y ₅	97.554	9.646	97.149	8.762	0.042	-	1.876
Y ₆	98.481	9.733	97.979	9.264	0.087	-	0.580
Y ₇	100.010	10.305	98.682	9.894	0.067	-	0.147
Y ₈	97.209	8.293	98.567	9.444	0.180	-	3.772
Y ₉	101.875	11.307	99.698	9.968	0.980	-	1.727
Y ₁₀	95.683	8.572	98.256	9.391	0.116	-	3.773
Y ₁₁	100.135	9.081	99.231	9.653	0.402	-	4.031
Y ₁₂	101.719	11.048	99.179	9.567	0.263	-	0.999
Y ₁₃	96.841	8.958	99.565	9.696	0.259	-	2.008
Y ₁₄	102.588	10.889	100.383	10.298	0.397	-	1.481
Y ₁₅	103.271	11.756	100.900	10.534	0.176	-	3.008
Y ₁₆	100.955	10.467	102.271	11.037	0.045	-	0.199
Y ₁₇	100.998	10.372	101.741	10.865	0.070	-	0.243
Y ₁₈	103.503	11.714	101.819	10.851	0.065	-	2.967
Y ₁₉	102.208	11.253	102.236	11.113	0.046	-	1.465
Y ₂₀	98.870	9.822	101.527	10.930	0.073	-	0.314
X ₁	98.735	9.843	-	-	-	-	0.463
X ₂	101.445	10.382	-	-	-	-	0.650
X ₃	97.113	10.083	99.098	10.103	0.071	1.000	5.475
X ₄	101.333	9.981	99.964	10.149	0.624	0.444	1.662
X ₅	104.844	11.746	101.097	10.603	0.139	0.889	4.905
X ₆	100.845	10.807	102.341	10.844	0.035	0.556	0.728
X ₇	101.157	11.135	102.282	11.229	0.042	0.000	1.634
X ₈	99.693	9.838	100.565	10.593	0.125	0.611	0.093
X ₉	100.971	9.730	100.607	10.234	0.225	0.722	2.048
X ₁₀	99.143	9.848	99.936	9.805	0.180	0.722	0.154

Πίνακας 4.8: Τιμές στατιστικής συνάρτησης του διαγράμματος $DDMA$.



Σχήμα 4.9: Παραμετρικό διάγραμμα T^2 του Hotelling ($\alpha = 0.10$)

Παρατηρούμε ότι στο διάγραμμα r τα δείγματα 23 και 25 (X_3 και X_5) είναι εκτός ελέγχου, ακριβώς όπως και στο διάγραμμα T^2 . Επιπλέον, παρατηρούμε ότι στο διάγραμμα Q ενεργοποιήθηκε συναγερμός για πρώτη φορά τη χρονική στιγμή 12, η οποία σχετίζεται με τα δείγματα 23 και 24 (X_3 και X_4). Τα διαγράμματα Q και T^2 , ανίχνευσαν τη μετατόπιση της διεργασίας την ίδια χρονική στιγμή. Στο διάγραμμα S , τα δείγματα 10 με 20 οδηγούν σε σημεία που βρίσκονται κάτω από την κεντρική γραμμή. Εφόσον, ο μέσος της διεργασίας μετατοπίζεται από το 20^ο δείγμα και μετά, θα μπορούσαμε να πούμε ότι το διάγραμμα αυτό δεν εντόπισε σωστά την αλλαγή στο μέσο. Αντίστοιχα, στο διάγραμμα S^* τα δείγματα 8 με 30 δίνουν σημεία κάτω από την κεντρική γραμμή. Κανονικά, τα διαγράμματα θα έπρεπε να δίνει σήμα εκτός ελέγχου από το 20^ο σημείο και μετά. Επομένως, το διάγραμμα S^* δεν είναι τόσο αξιόπιστο όσο το διάγραμμα T^2 . Τέλος, στο διάγραμμα $DDMA$ το πέμπτο σημείο, το οποίο σχετίζεται με τα δείγματα X_5 , X_6 και X_7 , είναι κάτω από κάτω όριο ελέγχου. Το διάγραμμα $DDMA$ ανιχνεύει τη μετατόπιση λίγο αργότερα σε σχέση με το διάγραμμα T^2 .

Θα επιθυμούσαμε τα μη παραμετρικά διαγράμματα να ανιχνεύσουν την αλλαγή στο μέσο της διεργασίας την ίδια χρονική στιγμή που την ανίχνευσε και το διάγραμμα T^2 ή έστω λίγο αργότερα. Ωστόσο, η αλλαγή αυτή ή άργησε να ανιχνευθεί ή δεν ανιχνεύθηκε σωστά, το οποίο αποτελεί ένδειξη ότι τα μη παραμετρικά διαγράμματα δεν ήταν τόσο αποτελεσματικά με τα δεδομένα που χρησιμοποιήσαμε. Κατά συνέπεια, όταν η κατανομή της διεργασίας είναι κανονική, είναι προτιμότερο να χρησιμοποιούμε παραμετρικά διαγράμματα ελέγχου.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΔΑΙΑ

Βιβλιογραφία

Ελληνική

- Camp, J. M. (2004). *Η Αρχαία Αγορά της Αθήνας: Οι ανασκαφές στην καρδιά της κλασικής πόλης*, Εκδόσεις Miet, Μορφωτικό Ίδρυμα Εθνικής Τραπέζης, Αθήνα (Μετάφραση Κλεώπα, Μ.).
- Αντζουλάκος, Δ. (2010). *Στατιστικός Έλεγχος Ποιότητας*, Σημειώσεις Παραδόσεων του μαθήματος Στατιστικός Έλεγχος Ποιότητας στα πλαίσια του ΠΜΣ στην «Εφαρμοσμένη Στατιστική», Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς, Πειραιάς.
- Βαρουφάκης, Γ. (2005). Ο Έλεγχος Ποιότητας των Προϊόντων και η Προστασία του Καταναλωτή στην Αρχαιότητα, *Αρχαιολογία και Τέχνες*, **95**, 81-87.
- Γεωργακάκος, Γ. Σ. (2002). *Προηγμένα Εργαλεία και Μέθοδοι για τον Έλεγχο Ποιότητας, Τόμος Α: Στατιστικός Έλεγχος Διεργασίας*, Εκδόσεις Ελληνικού Ανοικτού Πανεπιστημίου, Πάτρα.
- Παπαργύρης, Α. και Παπαργύρης, Δ. (2010). *Ποιοτικός Έλεγχος Παραγωγής*, Εκδόσεις Ζήτη, Θεσσαλονίκη.
- Σταθάκης, Α. Χ. (2006). *Πολυμεταβλητός Έλεγχος Ποιότητας*, Διπλωματική Εργασία, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιά.

Ξένη

- Abu-Shawiesh, M. O. and Abdullah, M. B. (2001). A New Robust Bivariate Control Chart for Location, *Communications in Statistics-Simulation and Computation*, **30**, 513-529.
- Adams, B. M. (1994). The Multivariate Control Web, *Quality Engineering*, **6**, 533-545.
- Alfaro, E., Alfaro, J. L., Gámez, M. and García, N. (2009). A Boosting Approach for Understanding Out-of-Control Signals in Multivariate Control Charts, *International Journal of Production Research*, **47**, 6821-6834.
- Alt, F. B. (1985). Multivariate Quality Control, *Encyclopedia of Statistical Sciences*, 110-122, Kotz S., Johnson N. L. and Read C. R. (eds.), John Wiley, New York..
- Alt, F. B. and Smith, N. D. (1988). Multivariate Process Control, *Handbook of Statistics*, **7**, 333-351, Krishnaiah P. R. and Rao C. R. (eds.), Elsevier: Amsterdam, The Netherlands.
- Alt, F. B., Goode, J. J. and Wadsworth, H. M. (1976). Small Sample Probability Limits for the Mean of a Multivariate Normal Process, *30th ASQC Annual Technical Conference Transactions*, **30**, 170-176, Toronto, Canada.
- Amin R. W. and Searcy, A. J. (1991). A Nonparametric Exponentially Weighted Moving Average Control Scheme, *Communications in Statistics-Simulation and Computation*, **20**, 1049-1072.
- Amin, R. W. and Miller, R. W. (1993). A Robustness Study of \bar{X} Charts with Variable Sampling Intervals, *Journal of Quality Technology*, **25**, 36-44.
- Amin, R. W., Reynolds, M. R. and Bakir, S. (1995). Nonparametric Quality Control Charts Based on the Sign Statistic, *Communications in Statistics-Theory and Methods*, **24**, 1597-1623.
- Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York.

- Aparisi, F., Avendaño, G. and Sanz, J. (2006). Techniques to Interpret T^2 Control Chart Signals, *IIE Transactions*, **38**, 647-657.
- Aparisi, F., Jabaioyes, J. and Carrión, A. (1999). Statistical Properties of the $|S|$ Multivariate Control Chart, *Communications in Statistics-Theory and Methods*, **28**, 2671-2686.
- Atienza, O. O., Tang, L. C. and Ang, B. W. (1998). QUALITY NOTES: Simultaneous Monitoring of Univariate and Multivariate SPC Information Using Boxplots, *International Journal of Quality Science*, **3**, 194-204.
- Baardseth, P., Dalen, G. A. and Tandberg, A. (1999). Innovation/Technology transfer to food SMEs, *Trends in Food Science and Technology*, **10**, 234-238.
- Bakir, S. T. (2004). A Distribution-Free Shewhart Quality Control Chart Based on Signed-Ranks, *Quality Engineering*, **16**, 613-623.
- Bakir, S. T. (2010). A Nonparametric Test for Homogeneity of Variances: Application to GPAs of Students Across Academic Majors, *IABR & ITLC Conference Proceedings*, Orlando, FL.
- Balakrishnan, N., Triantafyllou, I. S. and Koutras, M. V. (2009). Nonparametric control charts based on runs and Wilcoxon-type rank-sum statistics, *Journal of Statistical Planning and Inference*, **139**, 3177-3192.
- Balakrishnan, N., Triantafyllou, I. S. and Koutras, M. V. (2010). A Distribution-Free Control Chart Based on Order Statistics, *Communications in Statistics: Theory and Methods*, Vol. **39**, Issue 20, pp. 3652-3677.
- Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall: Englewood Cliffs, NJ.
- Bersimis, S., Psarakis, S. and Panaretos, J. (2007). Multivariate Statistical Process Control Charts: An Overview, *Quality and Reliability Engineering International*, **23**, 517-543.
- Blazek, L. W., Novic, B. and Scott, D. M. (1987). Displaying Multivariate Data Using Polyplots, *Journal of Quality Technology*, **19**, 69-74.
- Boone, J. M. (2010). *Contributions to Multivariate Control Charting: Studies of the Z-chart and four Nonparametric Charts*, Ph. D. Dissertation, University of Alabama.
- Box, G. E. P. and Hunter J. S. (1961). The 2^{k-p} Fractional Factorial Designs Part 1, *Technometrics*, **3**, 311-351.
- Box, G. E. P. and Wilson, K. B. (1951). On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society. Series B (Methodological)*, **13**, 1-45.
- Burr I. W. (1976). *Statistical Quality Methods*, Marcel Dekker, New York.
- Bush, H. M., Chongfuanqprinya, P., Chen, V. C. P., Sukchotrat, T. and Kim, S. B. (2010). Nonparametric Multivariate Control Charts Based on Linking Algorithm, *Quality and Reliability Engineering International*, **26**, 663-675.
- Butler, S. W. and Stefani, J. A. (1994). Supervisory Run-to-Run Control of a Polysilicon Gate etch Using in Situ Ellipsometry, *IEEE Transactions on Semiconductor Manufacturing*, **7**, 193-201.
- Chakraborti, S. (2011). A Nonparametric (Distribution-Free) Quality Control Charts, *Encyclopedia of Statistical Sciences*, 1-27 (DOI: 10.1002/0471667196.ess7150).
- Chakraborti, S. and Eryilmaz, S. (2007). A Nonparametric Shewhart-Type Signed-Rank Control Chart Based on Runs, *Communications in Statistics-Simulation and Computation*, **36**, 335-356.
- Chakraborti, S. and Graham, M. A. (2008). Nonparametric Control Charts, *Encyclopedia of Statistics in Quality and Reliability*, **1**, 415-429, Wiley, New York. (DOI: 10.1002/9780470061572.eqr262).

- Chakraborti, S. and Van de Wiel, M. A. (2008). A Nonparametric Control Chart Based on the Mann-Whitney Statistic. In *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*, 156–172, N. Balakrishnan, E. A. Peña, and M. J. Silvapulle, eds. Institute of Mathematical Statistics, Hayward, CA.
- Chakraborti, S., Eryilmaz, S. and Human, S. W. (2009). A Phase II Nonparametric Control Chart Based on Precedence Statistics with Runs-type Signaling Rules, *Computational Statistics and Data Analysis*, **53**, 1054-1065.
- Chakraborti, S., Van Der Laan, P. and Bakir, S. T. (2001). Nonparametric Control Charts: An Overview and Some Results, *Journal of Quality Technology*, **33**, 304-315.
- Chan, L. K. and Zhang, J. (2001). Cumulative Sum Control Charts for Covariance Matrix, *Statistica Sinica*, **11**, 767-790.
- Chang, S. I. and Aw, C. A. (1996). A Neural Fuzzy Control Chart for Detecting and Classifying Process Mean Shifts, *International Journal of Production Research*, **34**, 2265-2278.
- Chen, G., Cheng S. W. and Xie, H. (2005). A New Multivariate Control Chart for Monitoring Both Location and Dispersion, *Communications in Statistics-Simulation and Computation*, **34**, 203-217.
- Cheng, C. S. (1995). A Multi-layer Neural Network Model for Detecting Changes in the Process Mean, *Computers and Industrial Engineering*, **28**, 51-61.
- Cheng, C. S. (1997). A Neural Network Approach for the Analysis of Control Chart Patterns, *Industrial Journal of Production Research*, **35**, 667-697.
- Chiu, J. E. and Kuo, T. I. (2007). Attribute Control Chart for Multivariate Poisson Distribution, *Communications in Statistics-Theory and Methods*, **37**, 146-158.
- Chou, Y-M, Mason, R. L. and Young, J. C. (2001). The Control Chart for Individual Observations from a Multivariate Non-Normal Distribution, *Communications in Statistics-Theory and Methods*, **30**, 1937-1949.
- Chua, M. K. and Montgomery, D. C. (1992). Investigation of a Control Scheme for Multivariate Quality Control, *Quality and Reliability Engineering International*, **8**, 37-44.
- Cook, D. F. and Chiu, C-C. (1998). Using Radial Basis Function Neural Networks to Recognize Shifts in Correlated Manufacturing Process Parameters, *IIE Transactions*, **30**, 227-234.
- Cook, D. F., Zobel, C. W. and Nottingham, Q. J. (2001). Utilization of Neural Networks for the Recognition of Variance Shifts in Correlated Manufacturing Process Parameters, *International Journal of Production Research*, **39**, 3881-3887.
- Crosier, R. B. (1988). Multivariate Generalizations of Cumulative Sum Quality Control Schemes, *Technometrics*, **30**, 291-303.
- Dai, Y. Zou, C. and Wang, Z. (2011). Multivariate Cusum Control Charts Based on Data Depth for Preliminary Analysis, Working Paper.
- Das, N. (2009). A New Multivariate Non-Parametric Control Chart Based on Sign Test, *Quality Technology and Quantitative Management*, **6**, 155-169.
- Del Castillo, E. and Rajagopal, R. (2002). A Multivariate Double EWMA Process Adjustment Scheme for Drifting Processes, *IIE Transactions*, **34**, 1055-1068.
- Dodge, H. F. and Romig, H. G. (1929). A method of Sampling Inspection, *Bell System Technical Journal*, **8**, 613-631.
- Doganaksoy, N., Faltin, F. W. and Tucker, W. T. (1991). Identification of Out-of-Control Quality Characteristics in a Multivariate Manufacturing Environment, *Communications in Statistics-Theory and Methods*, **22**, 2775-2790.
- Doty, L. A. (1996). *Statistical Process Control*, Industrial Press Inc, New York.

- Elder, R. S., Provost, L. P. and Ecker, O. M. (1981). United States Department of Agriculture CUSUM Acceptance Sampling Procedure, *Journal of Quality Technology*, **13**, 59-64.
- Enns, P. G., Machak, J. A., Spivey, W. A. and Wroblewski, W. J. (1982). Forecasting Applications of an Adaptive Multiple Exponential Smoothing Model, *Management Science*, **28**, 1035-1044.
- Fassó, A. (1999). One-sided MEWMA Control Charts, *Communications in Statistics-Simulation and Computation*, **28**, 381-401.
- Feigenbaum, A. V. (1951). *Total Quality Control*, McGraw-Hill, New York.
- Frisén, M. (2011). On Multivariate Control Charts, *Produção*, **21**, 235-241.
- Fuchs, C. and Benjamini, Y. (1994). Multivariate Profile Charts for Statistical Process Control, *Technometrics*, **36**, 182-195.
- Fuchs, C. and Kenett, R. S. (1998). *Multivariate Quality Control*, Marcel Dekker, New York.
- Gadre, M. P. and Rattihalli, R. N. (2005). Some Group Inspection Based Multi-Attribute Control Charts to Identify Process Deterioration, *Economic Quality Control*, **20**, 191-204.
- Garvin, D. A. (1987). Competing in the Eight Dimensions of Quality, *Harvard Business Review*, **65**, 101-109.
- Geladi, P. and Kowalski, B. (1986). Partial Least-Squares Regression: A tutorial, *Analytica Chimica Acta*, **185**, 1-17.
- Gibbons, J. D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, 4th, Ed., Revised and Expanded. Marcel Dekker, New York, NY.
- Gnanadesikan, M. and Gupta, S. S. (1970). A Selection for Multivariate Normal Distributions in Terms of the generalized Variances, *Technometrics*, **12**, 103-117.
- Golosnoy, V., Schmid, W. and Okhrin, I. (2008). Sequential Monitoring of Optimal Portfolio Weights, *Financial Surveillance*, Frisén M. (eds.), John Wiley & Sons, Ltd, Chichester, UK. (DOI: 10.1002/9780470987179.ch7)
- 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**, 1795-1813.
- Guh, R. S. (2003). Integrating Artificial Intelligence into On-line Statistical Process Control, *Quality and Reliability Engineering International*, **19**, 1-20.
- Guh, R. S. (2007). On-line Identification and Quantification of Mean Shifts in Bivariate Processes Using a Neural Network-based Approach, *Quality and Reliability Engineering International*, **23**, 367-385.
- Guh, R. S. and Hsieh, Y-C. (1999). A Neural Network Based Model for Abnormal Pattern Recognition of Control Charts, *Computers and Industrial Engineering*, **36**, 97-108.
- Guh, R. S. and Tannock, J. D. T. (1999a). A Neural Network Approach to Characterize Pattern Parameters in Process Control Charts, *Journal of Intelligent Manufacturing*, **5**, 449-462.
- Guh, R. S. and Tannock, J. D. T. (1999b). Recognition of Control Chart Concurrent Patterns Using a Neural Network Approach, *International Journal of Production Research*, **37**, 1743-1765.
- Harvey, A. C. (1986). Analysis and Generalization of a Multivariate Exponential Smoothing Model, *Management Science*, **32**, 374-380.
- Hawkins, D. M. (1991). Multivariate Quality Control Based on Regression-Adjusted Variables, *Technometrics*, **33**, 61-75.
- Hawkins, D. M. (1993). Regression Adjustment for Variables in Multivariate Quality Control, *Journal of Quality Technology*, **25**, 170-182.

- Hawkins, D. M. and Wixley, R. A. J. (1986). A Note on the Transformation of Chi-Squared Variables to Normality, *The American Statistician*, **40**, 296-298.
- Hayter, A. J. and Tsui, K-L. (1994). Identification and Quantification in Multivariate Quality Control Problems, *Journal of Quality Technology*, **26**, 197-208.
- Healy, J. D. (1987). A Note on Multivariate CUSUM Procedures, *Technometrics*, **29**, 409-412.
- Hettmansperger, T. P. (2006). Multivariate Location Tests, *Encyclopedia of Statistical Sciences*, John Wiley and Sons. (DOI: 10.1002/0471667196.ess1733.pub2).
- Hettmansperger, T. P. and Randles, R. H. (2002). A Practical Affine Equivariant Multivariate Median, *Biometrika*, **89**, 851-860.
- 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. (1969). On the Ratio of Correlated Normal Random Variables, *Biometrika*, **56**, 635-639.
- Ho, E. S. and Chang, S. I. (1999). An Integrated Neural Network Approach for Simultaneous Monitoring of Process Mean and Variance Shifts-A Comparative Study, *International Journal of Production Research*, **37**, 1881-1901.
- Hoel, P. G. (1937). A Significance Test for Component Analysis, *Annals of Mathematical Statistics*, **8**, 149-158.
- Holgate, P. (1964). Estimation for the Bivariate Poisson Distribution, *Biometrika*, **51**, 241-245.
- Hotelling, H. (1947). Multivariate Quality Control-Illustrated by the Air Testing of Sample Bombsights, *Techniques of Statistical Analysis*, 111-184, Eisenhart, C., Hastay, M. W. and Wallis, W. A. (eds.), McGraw-Hill, New York.
- Hotelling, H. (1951). A Generalized T-Test and Measure of Multivariate Dispersion, Newman J. (eds.), *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, CA, 23-42.
- Howell, J. M. (1952). Statistical Quality Control, *Mathematics Magazine*, **25**, 155-157.
- Human, S. W. and Chakraborti, S. and Smit, C. F. (2010). Nonparametric Shewhart-Type Sign Control Charts Based on Runs, *Communications in Statistics-Theory and Methods*, **39**, 2046-2062.
- Ishikawa, K. (1989). *Introduction to Quality Control*, 3A Corporation, Tokyo, Translated by J.H. Loftus.
- Jackson, J. E. (1985). Multivariate Quality Control, *Communications in Statistics-Theory and Methods*, **14**, 2657-2688.
- Jackson, J. E. (1991). *A User's Guide to Principal Components*, John Wiley and Sons, New York
- Jackson, J. E. and Mudholkan, G. S. (1979). Control Procedures for Residuals Associated with Principal Component Analysis, *Technometrics*, **21**, 341-349.
- Janacek, G. J. and Meikle, S. E. (1997). Control Charts Based on Medians, *Journal of the Royal Statistical Society: Series D (The Statistician)*, **46**, 19-31.
- Jensen, W. A., Birch, J. B. and Woodall, W. H. (2007). High Breakdown Estimation Methods for Phase I Multivariate Control Charts, *Quality and Reliability Engineering International*, **23**, 615-629.
- Johnson, N. L. and Leone, F. C. (1962a). Cumulative Sum Control Charts-Mathematical Principles Applied to Their Construction and Use, Part I, *Industrial Quality Control*, **19**, 15-21.

- Johnson, N. L. and Leone, F. C. (1962b). Cumulative Sum Control Charts-Mathematical Principles Applied to Their Construction and Use, Part II, *Industrial Quality Control*, **19**, 29-36.
- Johnson, N. L. and Leone, F. C. (1962c). Cumulative Sum Control Charts-Mathematical Principles Applied to Their Construction and Use, Part III, *Industrial Quality Control*, **19**, 22-28.
- Jolayemi, J. K. (1991). Approximation to the Convolution of Independent Binomial Variables: An Extension of Theory, *Journal of Nigerian Statistical Association*, **6**, 17-28.
- Jolayemi, J. K. (1994). Convolution of Independent Binomial Variables: An Approximation Method and a Comparative Study, *Computational Statistics and Data Analysis*, **18**, 403-417.
- Jolayemi, J. K. (1999). A Statistical Model for the Design of Multiattribute Control Charts, *Sankhyā: The Indian Journal of Statistics. Series B*, **61**, 351-365.
- Jones, R. H. (1966). Exponential Smoothing for Multivariate Time Series, *Journal of the Royal Statistical Society. Series B (Methodological)*, **28**, 241-251.
- Jones-Farmer, L. A. (2008). Control Charts for Attributes, *Encyclopedia of Statistics in Quality and Reliability*, John Wiley and Sons. (DOI: 10.1002/9780470061572.eqr252).
- 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**, 317-330.
- Juran, J. M. and Gryna, F. M. (1957). *Quality Control Handbook*, McGraw Hill, New York.
- Kalagonda, A. A. and Kulkarni, S. R. (2003). Diagnosis of Multivariate Control Chart Signal Based on Dummy Variable Regression Technique, *Communications in Statistics-Theory and Methods*, **32**, 1665-1684.
- Kemp, C. P. and Kemp, W. A. (1965). Some Properties of the “Hermite” Distribution, *Biometrika*, **52**, 381-394.
- Knoth, S. and Schmid, W. (2002). Monitoring the Mean and the Variance of a Stationary Process, *Statistica Neerlandica*, **56**, 77-100.
- Kourti, T. and MacGregor, J. F. (1996). Multivariate SPC Methods for Process and Product Monitoring, *Journal of Quality Control*, **28**, 409-428.
- Koutras, M. V., Bersimis S. and Antzoulakos, D. L (2006). Improving the Performance of the Chi-square Control Charts via Runs Rules, *Methodology and Computing in Applied Probability*, **8**, 409-426.
- Kramer, H. G. and Schmid, L. V. (1997). EWMA Charts for Multivariate Time Series, *Sequential Analysis: Design Methods and Applications*, **16**, 131-154.
- Kresta, J. V., MacGregor, J. F. and Marlin, T. E. (1991). Multivariate Statistical Monitoring of Process Operating Performance, *The Canadian Journal of Chemical Engineering*, **69**, 35-47.
- Kshirsagar, A. M. (1972). *Multivariate Analysis*, Marcel Dekker, New York.
- Li, S. Y., Tang, L. C., and Ng, S. H. (2010). Nonparametric CUSUM and EWMA Control Charts for Detecting Mean Shifts, *Journal of Quality Technology*, **42**, 209-226.
- Linderman, K. and Love, T. E. (2000a). Economic and Economic Statistical Designs for MEWMA Control Charts, *Journal of Quality Technology*, **32**, 410-417.
- Linderman, K. and Love, T. E. (2000b). Implementing Economic and Economic Statistical Designs for MEWMA Charts, *Journal of Quality Technology*, **32**, 457-463.
- Liu, R. Y. (1995). Control Charts for Multivariate Processes, *Journal of the American Statistical Association*, **20**, 1380-1387.

- Liu, R. Y. and Tang, J. (1996). Control Charts for Dependent and Independent Measurements Based on Bootstrap Methods, *Journal of the American Statistical Association*, **91**, 1694-1700.
- Liu, R. Y., Singh, K. and Teng, J. H. (2004). DDMA-Charts: Nonparametric Multivariate Moving Average Control Charts Based on Data Depth, *Allgemeines Statistisches ARCHIV*, **88**, 235-258.
- Lowry, C. A. and Montgomery, D. C. (1995). A Review of Multivariate Control Charts, *IIE Transactions*, **27**, 800-810.
- Lowry, C. A., Woodall, W. H., Champ, C. W. and Rigdon, S. E. (1992). A Multivariate Exponentially Weighted Moving Average Control Chart, *Technometrics*, **34**, 46-53.
- Lu, X. S., Xie, M., Goh, T. N. and Lai, C. D. (1998). Control Chart for Multivariate Attribute Processes, *International Journal of Production Research*, **36**, 3477-3489.
- Lucas, J. M. (1976). The Design and Use of V-Mask Control Schemes, *Journal of Quality Technology*, **8**, 1-12.
- Lucas, J. M. and Saccucci, M. S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements, *Technometrics*, **32**, 1-12.
- MacGregor, J. F. (1990). A Different View of the Funnel Experiment, *Journal of Quality Technology*, **22**, 255-259.
- MacGregor, J. F. and Koutri, T. (1995). Statistical Process Control of Multivariate Processes, *Control Engineering Practice*, **3**, 403-414.
- MacGregor, J. F., Jaeckle, C., Kiparissides, C. and Koutoudi, M. (1994). Process Monitoring and Diagnosis by Multi-Block PLS Methods, *AIChE Journal*, **40**, 826-838.
- MacGregor, J. F., Nomikos, P. and Kourti, T. (1994). Multivariate statistical process control of batch processes using PCA and PLS, *Preprints IFAC, ADCHEM*, Kyoto, Japan, 525-530.
- Mader, D. P., Glycenfer, J. J. and Prins, J. (1996). An Application in Multivariate Statistical Process Control for Power Supply Calibration, *Quality Engineering*, **9**, 99-106.
- Mahmoud, M. A. and Maravelakis, P. E. (2011). The Performance of Multivariate CUSUM Control Charts with Estimated Parameters, *Journal of Computation and Simulation*, **iFirst**, 1-18. (Forthcoming Article)
- Maravelakis, P. E., Bersimis, S., Panaretos, J. and Psarakis, S. (2002). Identifying the Out-of-Control Variable in a Multivariate Control Chart, *Communications in Statistics-Theory and Methods*, **31**, 2391-2408.
- Margavio, T. M. (1990). *Multivariate Moving Average Control Charts for the Process Mean*, Thesis, Department of Management Science and Statistics, The University of Alabama.
- Margavio, T. M. and Conerly, M. D. (1995). A Comparison of Multivariate Moving Average Control Charts for the Process Mean, *International Journal of Production Research*, **33**, 1313-1321.
- Mason, R. L. and Young, J. C. (1999). Improving the Sensitivity of the T^2 Statistic in Multivariate Process Control, *Journal of Quality Technology*, **31**, 155-165.
- Mason, R. L., Tracy, N. D. and Young, J. C. (1995). Decomposition of T^2 for Multivariate Control Chart Interpretation, *Journal of Quality Technology*, **27**, 99-108.
- Mason, R. L., Tracy, N. D. and Young, J. C. (1996). Monitoring a Multivariate Step Process, *Journal of Quality Technology*, **28**, 39-50.
- Mason, R. L., Tracy, N. D. and Young, J. C. (1997). A Practical Approach for Interpreting Multivariate T^2 Control Chart Signals, *Journal of Quality Technology*, **29**, 396-406.

- Messaoud, A., Weihs, C. and Hering, F. (2004). A Nonparametric Control Chart Based on Data Depth, *Technical Reports*, **2004**, **60**, Technische Universität Dortmund, Sonderforschungsbereich 475: Komplexitätsreduktion in Multivariaten Datenstrukturen.
- Miller, P., Swanson, R. E. and Hecler C. F. (1993). Contribution plots: The missing link in multivariate quality control, *37th Annual Fall Conference ASQC*, Rochester, New York.
- Mitra, A. (2008). Control Charts for the Standard Deviation, *Encyclopedia of Statistics in Quality and Reliability*, John Wiley and Sons. (DOI: 10.1002/9780470061572.eqr251).
- Molnau, W. E., Runger, G.C., Montgomery, D. C., Skinner, K. R., Loredó, E. N. and Prabhu, S. S. (2001). A Program for ARL Calculation for Multivariate EWMA Charts, *Journal of Quality Technology*, **33**, 515-521.
- Montgomery, D. C. (1997). *Introduction to Statistical Quality Control*, Wiley, New York, 3rd Edition.
- Murphy, B. J. (1987). Selecting Out of Control Variables with the T^2 Multivariate Quality Control Procedure, *Journal of the Royal Statistical Society. Series D (The Statistician)*, **36**, 571-581.
- Ngai, C. H. and Zhang, J. (2001). Multivariate Cumulative Sum Control Charts Based on Projection Pursuit, *Statistica Sinica*, **11**, 747-766.
- Niaki, S. T. A. and Abbasi, B. (2005). Fault Diagnosis in Multivariate Control Charts Using Artificial Neural Networks, *Quality and Reliability Engineering International*, **21**, 825-840.
- Niaki, S. T. A. and Abbasi, B. (2007a). Bootstrap Method Approach in Designing Multi-Attribute Control Charts, *The International Journal of Advanced Manufacturing Technology*, **35**, 434-442.
- Niaki, S. T. A. and Abbasi, B. (2007b). Skewness Approach in Multi-Attribute Process Monitoring, *Communications in Statistics-Theory and Methods*, **36**, 2313-2325.
- Niaki, S. T. A. and Abbasi, B. (2008). A Transformation Technique in Designing Multi-Attribute C Control Charts, *Scientia Iranica*, **15**, 125-130.
- Nomikos, P. and MacGregor, J. F. (1994). Monitoring Batch Processes Using Multi-Way Principal Component Analysis, *AIChE Journal*, **40**, 1361-1375.
- Nomikos, P. and MacGregor, J. F. (1995a). Multivariate SPC Charts for Monitoring Batch Processes, *Technometrics*, **37**, 41-59.
- Nomikos, P. and MacGregor, J. F. (1995b). Multi-way partial least squares in monitoring batch processes, *Chemometrics and Intelligent Laboratory Systems*, **30**, 97-108.
- Noorossana, R., Farrokhi, M. and Saghali A. (2003). Using Neural Networks to Detect and Classify Out-of-Control Signals in Autocorrelated Processes, *Quality and Reliability Engineering International*, **19**, 493-504.
- Nottingham, Q. J., Cook, D. F. and Zobel, C. W. (2001). Visualization of Multivariate Data with Radial Plots Using SAS, *Computers and Industrial Engineering*, **41**, 17-35.
- Okhrin, Y. and Schmid, W. (2008). Surveillance of Univariate and Multivariate Nonlinear Time Series, *Financial Surveillance*, Frisén M. (eds.), John Wiley & Sons, Ltd, Chichester, UK. (DOI: 10.1002/9780470987179.ch6)
- Oktay Firat, S. Ü. and Aricigil, Ç. (2000). Multivariate Quality Control: A Historical Perspective, *1st International Joint Symposium on Business Administration*, Gökçeada-Çanakkale, Turkey, 54-67.
- Page, E. S. (1954). Continuous Inspection Schemes, *Biometrika*, **41**, 100-115
- Parr, W. C. (2008). Statistical Quality Control, *Encyclopedia of Statistics in Quality and Reliability*, John Wiley and Sons (DOI: 10.1002/9780470061572.eqr393)

- Patel, H. I. (1973). Quality Control Methods in Multivariate Binomial and Poisson Distributions, *Technometrics*, **15**, 103-112.
- Pignatiello, J. J. and Runger, G. C. (1990). Comparisons of Multivariate CUSUM Charts, *Journal of Quality Technology*, **22**, 173-186.
- Polhemus, N. W. (2005). How to: Apply SPC Techniques to Correlated Variables Using STATGRAPHICS Centurion, *Statgraphics Centurion: How-to Guides*.
- Pongchavalit, C. (2009). A Multivariate Moving Average Control Chart for Photovoltaic Processes, *World Academy of Science, Engineering and Technology: International Journal of Engineering and Physical Science*, **32**, 77-80.
- Puri, M. L. and Sen, P. K. (1976). Nonparametric Methods in Multivariate Analysis, John Wiley, New York.
- Qiu, P. (2008). Distribution-Free Multivariate Process Control Based on Log-Linear Modeling, *IIE Transactions*, **40**, 664-677.
- Qiu, P. and Hawkins, D. (2001). A Rank Based Multivariate CUSUM Procedure, *Technometrics*, **43**, 120-132.
- Qiu, P. and Hawkins, D. (2003). A Nonparametric Multivariate Cumulative Sum Procedure for Detecting Shifts in all Directions, *Journal of the Royal Statistical Society: Series D (The Statistician)*, **52**, 151-164.
- Rakitzis, A. C. and Antzoulakos, D. L. (2011). Chi-square Control Charts with Runs Rules, *Methodology and Computing in Applied Probability*, **13**, 657-669.
- Rao, C. R. and Arthanari, T. S. (1995). Data analysis and statistical thinking for quality and productivity improvement, *International Conference on Statistical Methods and Statistical Computing for Quality and Productivity Improvement*, Seoul, Korea.
- Raveh, A. (1985). On the Use of the Inverse of the Correlation Matrix in Multivariate Data Analysis, *The American Statistician*, **39**, 39-42.
- Reid, R. D. and Sanders, N. R. (2010). *Operations Management: An Integrated Approach*, John Wiley and Sons, USA, 4th Edition.
- Rigdon, S. E. (1995a). A Double-Integral Equation for the Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart, *Statistics and Probability Letters*, **24**, 356-373.
- Rigdon, S. E. (1995b). An Integral Equation for the in-control Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart, *Journal of Statistical Computation and Simulation*, **52**, 351-365.
- Roberts, S. W. (1959). Control Charts Tests Based on Geometric Moving Averages, *Technometrics*, **1**, 239-250.
- Runger, G. C. and Montgomery, D. C. (1997). Multivariate and Univariate Process Control: Geometry and Shift Directions, *Quality and Reliability Engineering International*, **13**, 153-158.
- Runger, G. C. and Prabhu, S. S. (1996). A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Charts, *Journal of the American Statistical Association*, **91**, 1701-1760.
- Runger, G. C. and Testik, M. C. (2004). Multivariate Extensions to Cumulative Sum Control Charts, *Quality and Reliability Engineering International*, **20**, 587-606.
- Runger, G. C., Alt, F. B. and Montgomery, D. C. (1996). Contributors to a Multivariate Statistical Process Control Chart, *Communications in Statistics-Theory and Methods*, **25**, 2203-2213.

- Runger, G. C., Keats, J. B., Montgomery, D. C. and Scranton, R. (1999). Improving the Performance of the Multivariate Exponentially Moving Average Control Chart, *Quality and Reliability International*, **15**, 161-166.
- Ryan, T. P. (1989). *Statistical Methods for Quality Improvement*, John Wiley, New York.
- Sahni, N. S., Aastveit, A. H. and Naes, T. (2005). In-Line Process and Product Control Using Spectroscopy and Multivariate Calibration, *Journal of Quality Technology*, **37**, 1-20.
- Seber, G. A. F. (1984). *Multivariate Observations*, John Wiley, New York.
- Sepúlveda, A. and Nachlas, J. A. (1997). A Simulation Approach to Multivariate Quality Control, *Computers and Industrial Engineering*, **33**, 113-116.
- Shao, Y. E., Wu, C-H., Ho, B-Y. and Hsu, B-S. (2008). A Neural Network-Based Approach to Identifying Out-of-Control Variables for Multivariate Control Charts, *Lecture Notes in Computer Science, Computer Supported Cooperative Work in Design IV*, **5236**, 644-652.
- Shewhart, W. (1931). *Economic Control of Quality of Manufactured Product*, Van Nostrand, New York.
- Snee, R. D. (1990). Statistical Thinking and Its Contribution to Total Quality, *The American Statistician*, **44**, 116-121.
- Sparks, R., Adolphson, A. and Phatak, A. (1997). Multivariate Process Monitoring Using the Dynamic Biplot, *International Statistical Review*, **65**, 325-349.
- Stounbos, Z. G. and Sullivan, J. H. (2002). Robustness to Non-normality of the Multivariate EWMA Control Chart, *Journal of Quality Technology*, **34**, 304-315.
- Sullivan, J. H. and Woodall, W. H. (1998). Adapting Control Charts for the Preliminary Analysis of Multivariate Observations, *Communications in Statistics-Simulation and Computation*, **27**, 953-979.
- Sun, R. and Tsung, F. (2003). A Kernel-Distance-Based Multivariate Control Chart Using Support Vector Methods, *International Journal of Production Research*, **41**, 304-315.
- Testik, M. C. and Borrór, C. M. (2004). Design Strategies for the Multivariate Exponentially Weighted Moving Average Control Chart, *Quality and Reliability Engineering International*, **20**, 571-577.
- Testik, M. C. and Runger, G. C. (2008). Multivariate Cumulative Sum (CUSUM) Chart, *Encyclopedia of Statistics in Quality and Reliability*, John Wiley and Sons (DOI: 10.1002/9780470061572.eqr256)
- Thissen, U., Swierenga, H., de Weijer, A. P., Wehrens, R., Melssen, W. J. and Buydens, M. C. (2005). Multivariate Statistical Process Control Using Mixture Modelling, *Journal of Chemometrics*, **19**, 23-31.
- Timm, N. H. (1996). Multivariate Quality Control Using Finite Intersection Tests, *Journal of Quality Technology*, **28**, 233-243.
- Topalidou, E. and Psarakis, S. (2009). Review of Multinomial and Multiattribute Quality Control Charts, *Quality and Reliability Engineering International*, **25**, 773-804.
- Tracy, N. D., Young, J. C. and Mason, R. L. (1992). Multivariate Control Charts for Individual Observations, *Journal of Quality Technology*, **24**, 88-95.
- Tseng, S. T., Chou, R. J. and Lee, S. P. (2002). A Study on a Multivariate EWMA Controller, *IIE Transactions*, **34**, 541-549.
- Tsung, F., Li, Y. and Jin, M. (2008). Statistical Process Control for Multistage Manufacturing and Service Operations: A Review and Some Extensions, *International Journal of Services Operations and Informatics*, **3**, 191-204.
- Wade, M. R. and Woodall, W. H. (1993). A Review and Analysis of Cause-Selecting Control Charts, *Journal of Quality Technology*, **25**, 161-170.

- Wärnefjord, K. (2004). *Multivariate Quality Control and Diagnosis of Sources of Variation in Assembled Products*, Licentiat Thesis, Göteborg University, Göteborg, Sweden.
- Weisberg, S. (1985). *Applied Linear Regression*, Wiley, New York, 2nd Edition
- Westerhuis, J. A., Gurden, S. P. and Smilde, A. K. (2000). Generalized Contribution Plots in Multivariate Statistical Process Monitoring, *Chemometrics and Intelligent Laboratory Systems*, **51**, 95-114.
- Western Electric. (1956). *Statistical Quality Control Handbook*, AT&T, Chicago, IL.
- Wierda, S. J. (1994). Multivariate Statistical Process Control-Recent Results and Directions for Future Research, *Statistica Neerlandica*, **48**, 147-168.
- Wise, B. M. (1991). *Adapting Multivariate Analysis for Monitoring and Modeling of Dynamic Systems*, Ph.D. Dissertation, University of Washington, Seattle.
- Wise, B. M. and Ricker, N. L. (1991). Recent Advances in Multivariate Statistical Process Control: Improving Robustness and Sensitivity, *IFAC Symposium on Advanced Control of Chemical Processes*, 125-130, Toulouse, France, Pergamon Press.
- Wise, B. M. D. J., Veltkamp, N. L., Ricker, B. R., Kowalski, S. M. Barnes and Arakali V. (1991). Application of Multivariate Statistical Process Control (MSPC) to the West Valley Slurry-Fed Ceramic Melter Process, *Waste Management '91 Proceedings*, **2**, 169-176, Tucson, Arizona.
- Woodall, W. and Ncube, M. M. (1985), Multivariate CUSUM Quality Control Procedures, *Technometrics*, **27**, 285-292.
- Woodall, W. H. (2000). Controversies and Contradictions in Statistical Process Control, *Journal of Quality Technology*, **32**, 341-350.
- Xie, M., Goh, T. N. and Kuralmani, V. (2002). *Statistical Models and Control Charts for High Quality Processes*, Boston: Kluwer Academic Publishers.
- Yang, C. H. and Hillier, F. S. (1970). Mean and Variance Control Chart Limits Based on a Small Number of Subgroups, *Journal of Quality Technology*, **2**, 9-16.
- Yeh, A. B., Huwang, L. and Wu, Y. F. (2004). A Likelihood-Ratio-Based EWMA Control Chart for Monitoring Variability of Multivariate Normal Processes, *IIE Transactions*, **36**, 865-879.
- Yeh, A. B., Lin, D. K. J. and McGrath, R. N. (2006). Multivariate Control Charts for Monitoring Covariance Matrix: A Review, *Quality Technology and Quantitative Management*, **3**, 415-436.
- Yeh, A., Lin, D., Zhou, H. and Venkataramani, C. (2003). A Multivariate Exponentially Weighted Moving Average Control Chart for Monitoring Process Variability, *Journal of Applied Statistics*, **30**, 507-536.
- Yu, J-B. and Xi, L-F. (2009). A Neural Network Ensemble-Based Model for On-line Monitoring and Diagnosis of Out-of-Control Signals in Multivariate Manufacturing Processes, *Expert Systems with Applications*, **36**, 909-921.
- Yumin, L. (1996). An Improvement for MEWMA in Multivariate Process Control, *Computers and Industrial Engineering*, **31**, 779-781.
- Zárate, P. and Okogboa, O. G. (2002). A Non-Parametric Control Chart for Multivariate Processes Using Principal Components Analysis, *IIE Annual Conference and Exhibition*, Orlando, Florida.
- Zárate, P. and Okogboa, O. G. (2004). Methodology for Interpreting Out-of-Control Signals for a Multivariate Control Chart, *IIE Annual Conference and Exhibition*, 309-315, Houston.
- Zorriassatine, F. and Tannock, J. D. T. (1998). A Review of Neural Networks for Statistical Process Control, *Journal of Intelligent Manufacturing*, **9**, 209-224.

- Zou, C. and Qiu, P. (2009). Multivariate Statistical Process Control using LASSO, *Journal of the American Statistical Association*, **104**, 1586-1596.
- Zou, C. and Tsung, F. (2011). A Multivariate Sign EWMA Control Chart, *Technometrics*, **53**, 84-97.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΠΑ

