

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ  
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**Πολυμεταβλητός Στατιστικός  
Έλεγχος Ποιότητας**

**Αγησίλαος Χ. Σταθάκης**

**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  
Οκτώβρης 2006



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ  
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**Πολυμεταβλητός Στατιστικός  
Έλεγχος Ποιότητας**

**Αγησίλαος Χ. Σταθάκης**

**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  
Οκτώβρης 2006

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. 13<sup>η</sup> /08.05.2006 συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Δημήτριος Αντζουλάκος
- Μαρία Κατέρη
- Μιχαήλ Μπούτσικας

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

**UNIVERSITY OF PIRAEUS**



**DEPARTMENT OF STATISTICS  
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN  
APPLIED STATISTICS**

**Multivariate Statistical  
Process Control**

By

**Agisilaos Ch. Stathakis**

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  
October 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΡΑΙΑ

*Στους γονείς μου  
Χαράλαμπο και Ελένη  
Τα αδέρφια μου  
Παναγιώτη και Ελισάβετ*

## Ευχαριστίες

Θέλω να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Δημήτριο Αντζουλάκο για το θέμα που μου εμπιστεύτηκε καθώς και για τις πολύτιμες συμβουλές και οδηγίες του σε όλα τα στάδια της μελέτης αυτής. Επίσης, θα ήθελα να ευχαριστήσω τον κ. Σωτήριο Μπερσίμη για το πολύτιμο υλικό που μου παρείχε. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την υποστήριξή της, έτσι ώστε να πραγματοποιήσω άλλο ένα όνειρό μου καθώς και το Πανεπιστήμιο Πειραιά για τις υψηλής ποιότητας σπουδές που μου έδωσε.

## Περίληψη

Στην εποχή της οικονομικής μεγέθυνσης και των υψηλών προδιαγραφών που διανύουμε, η ανάγκη για την παραγωγή ποιοτικών προϊόντων κρίνεται κάτι περισσότερο από επιτακτική. Οι βιομηχανίες αντιμετωπίζουν την έντονη ανταγωνιστικότητα από την παγκοσμιοποιημένη αγορά και η παραγωγή προϊόντων χαμηλών προδιαγραφών δεν τους δίνει περιθώρια μακροχρόνιας επιβίωσης. Γι' αυτό οι εταιρείες πρέπει να εφαρμόζουν στρατηγικές ανάπτυξης και να υιοθετούν μεθόδους που τους διασφαλίζουν ποιοτικό παραγόμενο προϊόν στα πλαίσια της Διοίκησης Ολικής Ποιότητας - ΔΟΠ (*Total Quality Control, TQC*). Οι εταιρείες στοχεύουν, όχι μόνο να πιστοποιήσουν (*ISO, International Standards Organization*) το παραγόμενο προϊόν τους έτσι ώστε να αυξήσουν την διαπραγματευτική τους ικανότητα με τους πελάτες τους, αλλά και να βελτιστοποιήσουν την παραγωγική τους διαδικασία.

Ο Στατιστικός Έλεγχος Ποιότητας (*Statistical Process Control, SPC*) είναι ένα από τα πιο αξιόπιστα εργαλεία για την παρακολούθηση του παραγόμενου προϊόντος, διασφαλίζοντας ότι οι διεργασίες βρίσκονται υπό έλεγχο ενώ παράλληλα εξασφαλίζει στις επιχειρήσεις την μακροχρόνια επιβίωσή τους και ειδικότερα την κερδοφορία τους. Επίσης, ο Πολυμεταβλητός Στατιστικός Έλεγχος Ποιότητας (*Multivariate Statistical Process Control, MSPC*) δίνει την δυνατότητα παρακολούθησης δύο ή περισσότερων μεταβλητών ενός προϊόντος ταυτόχρονα. Η ανάγκη χρησιμοποίησης του *MSPC* προέκυψε από τη διαπίστωση ότι η ποιότητα ενός προϊόντος μπορεί να σχετίζεται με περισσότερα του ενός ποιοτικά και μετρήσιμα χαρακτηριστικά. Από τις διάφορες τεχνικές του *MSPC* προκύπτουν διαγράμματα, των οποίων η ανάλυση και παρακολούθηση οδηγούν σε διαπιστώσεις των οποίων η παρουσία θα ήταν διαφορετικά πολύ δύσκολη έως και αδύνατη. Στην παρούσα μελέτη θα ασχοληθούμε με τα πολυμεταβλητά διαγράμματα Shewhart, MCUSUM και MEWMA ως προέκταση των μονοδιάστατων που όπως φαίνεται είναι περισσότερο διαδεδομένα σήμερα .



## Abstract

In the era of economic magnification, the need for high quality products is crucial. Industries are confronted with intense competition from the universalized market and the produce of poor quality products, which doesn't allow them long-term viability. That is why companies must utilize strategic growth and foster methods which will guarantee high quality products within the framework of *Total Quality Control (TQC)*. The companies targets, not only should certify (*ISO, International Standards Organization*) their manufactured negotiable products, to increase negotiable skills with their clients but also to improve their productive procedure.

*Statistical Process Control (SPC)* is one of the most reliable tools for the supervision of manufactured products guaranteeing that production is under control and at the same time, ensure the companies long term survival and especially for profitable business. Also, the *Multivariate Statistical Process Control (MSPC)* gives the opportunity of overseeing two or more products simultaneously. The need for use of *MSPC* arose from the ascertainment that the quality of one product could relate to more than one qualitative countable characteristics. From different techniques of *MSPC* there appear to outline the analysis and observation lead to the ascertainment of which the presence would be extremely difficult if not impossible.

## ΠΕΡΙΕΧΟΜΕΝΑ

### ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> Στατιστικός Έλεγχος Ποιότητας

1.1	Εισαγωγή στον Στατιστικό Έλεγχο Ποιότητας.....	1
1.2	Στατιστικός Έλεγχος Διεργασιών - Διαγράμματα Ελέγχου.....	2
1.3	Διάγραμμα Ελέγχου Shewhart.....	3
1.3.1	Διαγράμματα Ελέγχου για τη Μέση Τιμή.....	7
1.3.2	Διαγράμματα Ελέγχου για τη Διασπορά.....	13
1.3.3	Διαγράμματα Ελέγχου για Μεμονωμένες Παρατηρήσεις.....	17
1.3.4	Διαγράμματα Ελέγχου για το Ποσοστό..... και τον Αριθμό των Ελαττωματικών Προϊόντων.....	20
1.3.5	Διαγράμματα Ελέγχου για τον Αριθμό και το Μέσο Αριθμό των Ελαττωμάτων.....	22
1.5	Διαγράμματα Ελέγχου CUSUM.....	25
1.6	Διαγράμματα Ελέγχου EWMA.....	30

### ΚΕΦΑΛΑΙΟ 2<sup>ο</sup> Πολυμεταβλητά Διαγράμματα Shewhart

2.1	Πολυμεταβλητός Στατιστικός Έλεγχος Ποιότητας.....	36
2.2	Η $p$ -διάστατη κανονική κατανομή.....	37
2.2	Διαγράμματα Ελέγχου για το διάνυσμα των Μέσων Τιμών.....	39
2.3.1	Μέγεθος δείγματος $n > 1$ .....	39
2.3.2	Μεμονωμένες Παρατηρήσεις.....	43
2.4	Διαγράμματα Ελέγχου για τη Διασπορά.....	45

## ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> Πολυμεταβλητά Διαγράμματα CUSUM

3.1	Ακολουθιακοί έλεγχοι λόγου πιθανοφάνειας.....	50
3.2	Το σχέδιο του Crosier.....	53
3.3	Το σχέδιο των J. Pignatiello και G. Runger's.....	55

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> Πολυμεταβλητά Διαγράμματα EWMA

4.1	Διαγράμματα Ελέγχου MEWMA.....	57
4.2	ARL για το MEWMA.....	59

## ΚΕΦΑΛΑΙΟ 5<sup>ο</sup> Εφαρμογή.....

Βιβλιογραφία.....	71
-------------------	----

## Κατάλογος Πινάκων

1	Συσσωρευμένα αθροίσματα για $K=0.5$	27
2	Δεδομένα για την επίδειξη ενός διαγράμματος ελέγχου EWMA	32
3	Μέσο μήκος ροής για διαγράμματα ελέγχου EWMA	34
4	Μετρήσεις παραγόμενου προϊόντος	62 - 63
5	Περιγραφικά στατιστικά	63

## Κατάλογος Σχημάτων

1	Τυπικό διάγραμμα ελέγχου Shewhart	4
2	Διάγραμμα ελέγχου CUSUM	28
3	Διάγραμμα ελέγχου EWMA	33
4	Διάγραμμα Ελέγχου για την $Z_1$	64
5	Διάγραμμα Ελέγχου για την $Z_2$	64
6	Διάγραμμα Ελέγχου για την $Z_3$	65
7	Πολυμεταβλητό Διάγραμμα Ελέγχου για τις $Z_1, Z_2$ και $Z_3$	65
8	Συσχέτιση $Z_1, Z_2$ και $Z_3$	66
9	Διάγραμμα Ελέγχου $T^2$ για $Z_1, Z_2$	66
10	Διάγραμμα Ελέγχου $T^2$ για $Z_1, Z_3$	67
11	Διάγραμμα Ελέγχου $T^2$ για $Z_2, Z_3$	67
12	Διάγραμμα Ελέγχου για τις παρατηρήσεις της $Z_1$	68
13	Διάγραμμα Ελέγχου για τις παρατηρήσεις της $Z_2$	68
14	Διάγραμμα Ελέγχου για τις παρατηρήσεις της $Z_3$	69
15	Διάγραμμα Ελέγχου για τις $Z_1, Z_2$ και $Z_3$	69
16	Πολυμεταβλητό Διάγραμμα Ελέγχου MCUSUM	70
17	Πολυμεταβλητό Διάγραμμα Ελέγχου MEWMA	70

# ΚΕΦΑΛΑΙΟ 1<sup>ο</sup>

## 1.1 Στατιστικός Έλεγχος Ποιότητας

Ο Στατιστικός Έλεγχος Ποιότητας (*Statistical Quality Control*) αναπτύχθηκε ραγδαία τις τελευταίες δεκαετίες με κύριο στόχο την αποτελεσματικότερη παραγωγική διαδικασία. Ουσιαστικά είναι το εργαλείο της Διοίκησης μιας επιχείρησης για να ελέγχει αν το προϊόν που παράγει είναι σύμφωνο με τις προδιαγραφές που έχει ορίσει. Από την παραπάνω πρόταση καταλαβαίνουμε ότι αποτελεί μέρος των επιχειρήσεων για συνεχή βελτίωσή του παραγόμενου προϊόντος τους. Για να επιτύχει κάθε παραγωγική μονάδα τον παραπάνω στόχο χρησιμοποιεί τον Στατιστικό Έλεγχο Ποιότητας στα πλαίσια της Διοίκησης Ολικής Ποιότητας (*Total Quality Management*) που εφαρμόζει, με απώτερο σκοπό την αύξηση της κερδοφορίας της και γενικότερα την επίτευξη όλων των στόχων του στρατηγικού σχεδιασμού που έχει οριοθετήσει.

Η ολοένα ανταγωνιστικότερη αγορά που καλούνται οι επιχειρήσεις να αντιμετωπίσουν, καθώς και οι αυξημένες απαιτήσεις των καταναλωτών αναγκάζουν τις εταιρείες να υιοθετούν πρότυπα ποιότητας που πιστοποιούν το παραγόμενο προϊόν τους. Ουσιαστικά είναι ο τρόπος με τον οποίο η επιχείρηση που διαπραγματεύεται τόσο με πελάτες όσο και με εταιρείες δίνει τις απαραίτητες εγγυήσεις για το προϊόν της. Σήμερα για όλους τους κλάδους παραγωγής υπάρχουν τα αντίστοιχα πρότυπα που ολοένα και περισσότερο φαίνονται να τα εμπιστεύονται οι παραγωγοί, διότι, η διασφάλιση της ποιότητας τους εξασφαλίζει τεράστια πλεονεκτήματα. Το πρότυπο που φαίνεται να έχει επικρατήσει στην Ευρώπη για τη βιομηχανία είναι το ISO 9001:2000 (*International Standards Organization*), του Διεθνούς Οργανισμού Τυποποίησης. Το πρότυπο αυτό διασφαλίζει μέσω των διεργασιών που περιέχει την παραγωγή ποιοτικών προϊόντων ή καλύτερα παραγόμενων προϊόντων συμμορφούμενων με τις προδιαγραφές. Ο Στατιστικός Έλεγχος Ποιότητας από μόνος του αλλά ιδιαίτερα στο πρότυπο αυτό βρίσκει σιγά, σιγά την θέση που του αξίζει και η χρησιμότητά του σε κάθε τμήμα Ποιοτικού Ελέγχου μιας εταιρείας κρίνεται πλέον κάτι περισσότερο από αναγκαία.

Ειδικότερα, θα μπορούσαμε να αναφέρουμε ότι ο Στατιστικός Έλεγχος Ποιότητας αποτελείται από ένα σύνολο μεθόδων ανάλυσης στατιστικών δεδομένων.

Το σύνολο αυτό μπορεί να χωριστεί σε τρία βασικά υποσύνολα που το καθένα περιέχει στατιστικές μεθόδους προσανατολισμένες σε διαφορετικές φάσεις της παραγωγικής διαδικασίας και είναι οι:

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

## 1.2 Στατιστικός Έλεγχος Διεργασιών και Διαγράμματα Ελέγχου

Η αξιόπιστη παραγωγή ενός προϊόντος προϋποθέτει μια σταθερή επαναλαμβανόμενη παραγωγική διεργασία, με όσο το δυνατόν μικρότερη μεταβλητότητα από τις τιμές στόχους που έχουν δοθεί στα ποιοτικά χαρακτηριστικά του τελικού προϊόντος. Ο Στατιστικός Έλεγχος Διεργασιών (*SPC*) είναι ένα σύνολο από εργαλεία, που χρησιμοποιούνται στην επίβλεψη της σταθερότητας μιας διεργασίας και τη βελτίωση της ικανότητάς της (*capability*) μέσω της μείωσης της μεταβλητότητάς της. Τα κυριότερα εργαλεία του *SPC* και γνωστά ως *'the magnificent seven'* είναι τα εξής (Δ. Αντζουλάκος, 2003):

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

Από τις παραπάνω τεχνικές τα διαγράμματα ελέγχου αποτελούν την κυριότερη μέθοδο που αναπτύσσεται στην παρούσα μελέτη.

Στην παραγωγική διαδικασία το παραγόμενο προϊόν θα πρέπει να είναι ομοιογενές και συμμορφούμενο με τις προδιαγραφές. Στην πράξη όμως ποτέ δύο προϊόντα δεν είναι ίδια και αυτό γιατί από την μία υπάρχει η φυσική μεταβλητότητα και από την άλλη η ειδική μεταβλητότητα που τα συνοδεύει. Η φυσική

μεταβλητότητα διαφοροποιεί το προϊόν και οφείλεται σε μη ελέγξιμους παράγοντες (π.χ. θερμοκρασία περιβάλλοντος). Λόγω των μικρών μεταβολών (στα όρια προδιαγραφών) που προκαλεί δεν θεωρείται επιλήψιμη και θεωρούμε την διεργασία μας εντός ελέγχου ή ότι λειτουργεί σε ευσταθή κατάσταση. Αντίθετα, οι ειδικές αιτίες μεταβλητότητας δημιουργούν αστοχίες στο παραγόμενο προϊόν και προκαλούνται κυρίως, είτε από λάθος χειρισμούς ανθρώπων, είτε από την ύπαρξη ελαττωματικών πρώτων υλών και μηχανών ή ακόμα και από λάθος ρύθμιση αυτών. Λόγω των σοβαρών προβλημάτων που δημιουργεί η ύπαρξή τους, πρέπει να ανιχνεύονται άμεσα. Σε αντίθετη περίπτωση το κόστος παραγωγής του μη συμμορφούμενου προϊόντος γίνεται κάτι περισσότερο από ζημιόγONO. Όταν έχουμε ειδικές αιτίες μεταβλητότητας τότε λέμε ότι η διεργασία μας είναι εκτός ελέγχου ή ότι λειτουργεί σε ασταθή κατάσταση. Τα διαγράμματα ελέγχου μας δίνουν τη δυνατότητα να εντοπίζουμε τις ειδικές αιτίες μεταβλητότητας και να επεμβαίνουμε έγκαιρα στην παραγωγική διαδικασία ώστε να την βελτιώνουμε. Έτσι, πραγματοποιείται ο ποιοτικός έλεγχος που πρωταρχικό στόχο του έχει την συνεχή βελτίωση.

Σε κάθε παραγωγική διαδικασία υπάρχουν και τα όρια προδιαγραφών (*specification limits*) των ποιοτικών χαρακτηριστικών του προϊόντος τα οποία και έχουν καθοριστεί στη φάση του σχεδιασμού του. Αυτά αποτελούνται από την τιμή στόχο (*target value T*), το κάτω και το άνω όριο προδιαγραφών (*lower and upper specification limits, LSL and USL*) όπου εντός αυτών πρέπει να βρίσκονται οι τιμές του ποιοτικού χαρακτηριστικού ώστε να είναι αποδεκτό.

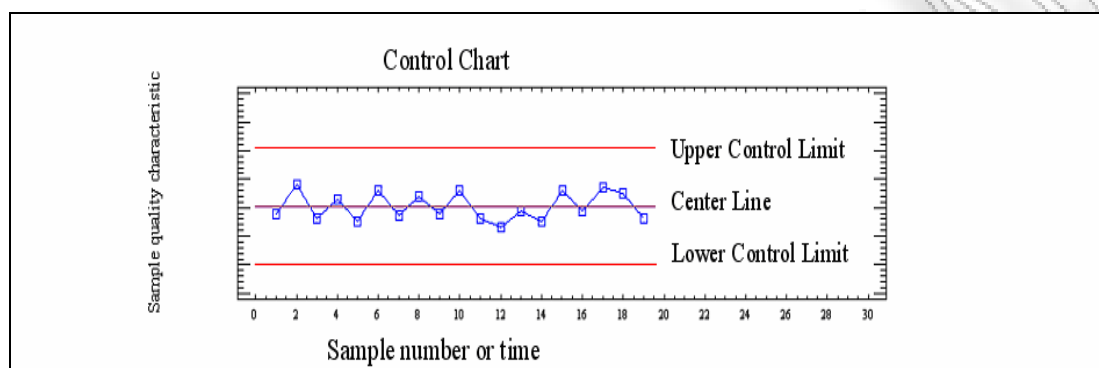
### 1.3 Shewhart

Τα διαγράμματα αυτά μας εξασφαλίζουν την συνεχή παρακολούθηση μιας κρίσιμης ποσότητας ενός ποιοτικού χαρακτηριστικού στην παραγωγική διαδικασία. Έτσι, επιλέγουμε από την παραγωγή τυχαία δείγματα (όχι απαραίτητα μεγέθους  $n > 1$ ) από τα οποία παίρνουμε τιμές για την τυχαία μεταβλητή  $X$ , έστω ότι παίρνουμε αποτελέσματα  $X_1, X_2, \dots$ . Έπειτα, χρησιμοποιώντας τα δείγματα υπολογίζουμε την τιμή  $W_i = g(X_i)$ ,  $i = 1, 2, \dots$  μιας στατιστικής συνάρτησης (τυχαίας μεταβλητής) που εκτιμά (συνήθως αμερόληπτη εκτιμήτρια) την κρίσιμη ποσότητα που μας ενδιαφέρει (π.χ. μέση τιμή, ποσοστό ελαττωματικών προϊόντων, αριθμό ελαττωμάτων, κτλ).



Όταν η κρίσιμη ποσότητα δίνει τιμές  $W_1, W_2, \dots$  εντός των ορίων ελέγχου λέμε ότι η διεργασία μας είναι εντός ελέγχου και λειτουργεί υπό συνθήκες φυσικής μεταβλητότητας (Δ. Αντζουλάκος, 2003).

Ένα τυπικό διάγραμμα ελέγχου Shewhart είναι μια γραφική παράσταση με την ακόλουθη μορφή



**Διάγραμμα 1:** Τυπικό διάγραμμα ελέγχου Shewhart

Στο Διάγραμμα 1, εκτός από τις παρατηρούμενες τιμές  $W_i$ , που έχουν απεικονιστεί με σημεία (τετραγωνάκια) τα οποία έχουν συνδεθεί με μια τεθλασμένη γραμμή, έχουν σχεδιαστεί και άλλες τρεις γραμμές. Η κεντρική γραμμή (*center line*, *CL*) ή μέσο επίπεδο της διεργασίας παριστάνει συνήθως τη μέση τιμή (*mean value*) της  $W$  όπως προκύπτει από τη λειτουργία μιας εντός (στατιστικού) ελέγχου διεργασίας, δηλαδή μιας διεργασίας που λειτουργεί μόνο με την παρουσία φυσικής μεταβλητότητας (*chance causes of variation*). Οι δύο άλλες γραμμές που εμφανίζονται στο παραπάνω διάγραμμα ονομάζονται άνω και κάτω όριο ελέγχου του διαγράμματος (*upper and lower control limit*, *UCL and LCL*). Όσο οι τιμές της  $W$  βρίσκονται εντός των ορίων ελέγχου και η συμπεριφορά τους είναι “τυχαία” μπορούμε να υποθέσουμε ότι η διεργασία παραμένει εντός ελέγχου και δεν χρειάζεται να προβούμε σε κάποια διορθωτική ενέργεια. Αν όμως κάποιο σημείο βρεθεί εκτός των ορίων ελέγχου λέμε ότι υπάρχει ένδειξη ότι η διεργασία είναι εκτός ελέγχου και πρέπει να προχωρήσουμε σε έρευνα για να ανακαλύψουμε τις ειδικές αιτίες μεταβλητότητας (*assignable causes of variation*) που είναι υπεύθυνες για αυτή τη συμπεριφορά και αν κριθεί απαραίτητο να προβούμε σε διορθωτικές ενέργειες. Ωστόσο, θα πρέπει να σημειώσουμε ότι ακόμη και στην περίπτωση που όλα τα σημεία του διαγράμματος βρίσκονται εντός των ορίων ελέγχου αλλά

συμπεριφέρονται με ένα συστηματικό ή μη τυχαίο τρόπο τότε και αυτό αποτελεί ένδειξη ότι η διεργασία είναι εκτός ελέγχου. Ως (ακραίο) παράδειγμα μπορούμε να αναφέρουμε την περίπτωση όπου όλα τα σημεία στο Διάγραμμα 1 βρίσκονται μεταξύ την κεντρικής γραμμής και του κάτω ορίου ελέγχου.

Για τον έλεγχο μίας παραγωγικής διεργασίας και την κατασκευή του διαγράμματος ελέγχου υπάρχουν δύο τεχνικές. Η τεχνική της **Φάσης I** (*Phase I*), όπου χρησιμοποιούνται αναδρομικά στοιχεία του παρελθόντος. Η τεχνική μας βοηθάει να διαχειριζόμαστε καλύτερα στοιχεία μελλοντικά εφόσον καταφέρουμε έπειτα από την επεξεργασία των δεδομένων μας να φέρουμε την διεργασία εντός ελέγχου. Ενώ, στην τεχνική της **Φάσης II** (*Phase II*), τα διαγράμματα ελέγχου φτιάχνονται από δεδομένα άμεσα προερχόμενα από την παραγωγή. Ενώ στην Φάση I ο διαχειριστής προσπαθεί να εντοπίσει αιτίες έτσι ώστε να βελτιώσει στο μέλλον την παραγωγή, στην Φάση II επεμβαίνει άμεσα και διορθωτικά όποτε αυτό χρειαστεί.

Για την κατασκευή διαγραμμάτων ελέγχου χρησιμοποιείται το μοντέλο ορίων σίγμα (*sigma limits model*) (Δ. Αντζουλάκος, 2003)

#### **Sigma limits model**

$$\begin{aligned} UCL &= \mu_w + L\sigma_w \\ \text{Center Line} &= \mu_w \\ LCL &= \mu_w - L\sigma_w \end{aligned}$$

όπου,  $\mu_w$  και  $\sigma_w$  είναι η μέση τιμή και η τυπική απόκλιση της στατιστικής συνάρτησης  $W$  και η ποσότητα  $L$  εκφράζει την απόσταση των ορίων ελέγχου από την κεντρική γραμμή σε μονάδες τυπικής απόκλισης (συνήθως  $L=3$ ).

Όταν λοιπόν χρησιμοποιούμε όρια  $3\sigma$  και το δείγμα μας θεωρούμε ότι ακολουθεί την κανονική κατανομή σε συνθήκες που η διεργασία μας είναι εντός ελέγχου λέμε ότι η πιθανότητα να πάρει η  $W$  τιμή εκτός ορίων ελέγχου είναι ίση με 0.0027 (όσο και το σφάλμα τύπου  $I$ ). Αυτό σημαίνει ότι έχουμε ένα εσφαλμένο συναγερμό 27 φορές στα 10000 σημεία του διαγράμματος ελέγχου.

Μια βασική έννοια που σχετίζεται με τα διαγράμματα ελέγχου είναι το μέσο μήκος ροής *ARL* (*average run length*) του διαγράμματος. Η τυχαία μεταβλητή  $T$  που δηλώνει το πλήθος των σημείων που πρέπει να σχεδιαστούν σε ένα διάγραμμα ελέγχου έως ότου πάρουμε ένδειξη εκτός ελέγχου διεργασίας ονομάζεται μήκος ροής (*run length*) του διαγράμματος. Το μέσο μήκος ροής *ARL* ορίζεται ως ο αναμενόμενος αριθμός των σημείων που πρέπει να σχεδιαστούν στο διάγραμμα έως

όπου λάβουμε για πρώτη φορά ένδειξη εκτός ελέγχου διεργασίας, δηλαδή  $ARL = E(T)$ .

Στα διαγράμματα ελέγχου Shewhart με απεικονιζόμενη ποσότητα τη  $W$  η οποία έχει εντός ελέγχου μέσο  $\mu$ , τυπική απόκλιση  $\sigma$  και συνάρτηση κατανομής  $F_{in}(\cdot)$ , έχουμε ότι η πιθανότητα  $p_{in}$  εμφάνισης ενός σημείου του διαγράμματος εκτός των ορίων ελέγχου είναι ίση με

$$p_{in} = 1 - P(LCL \leq W \leq UCL) = 1 - F_{in}(\mu + L\sigma) + F_{in}(-(\mu - L\sigma))$$

και το εντός ελέγχου μέσο μήκος ροής είναι ίσο με

$$ARL_{in} = \frac{1}{p_{in}}$$

αφού το μήκος ροής  $T$  ακολουθεί τη γεωμετρική κατανομή με πιθανότητα επιτυχίας  $p_{in}$ . Φυσικά για μια διεργασία που βρίσκεται εντός ελέγχου θέλουμε να έχουμε μεγάλη τιμή για το  $ARL_{in}$  έτσι ώστε να μειωθεί ο αριθμός των λανθασμένων ενδείξεων εκτός ελέγχου διεργασίας ή αλλιώς ο αριθμός των λανθασμένων συναγερμών (*false alarms*).

Για μια εκτός ελέγχου διεργασία, όπου η συνάρτηση κατανομή της τυχαίας μεταβλητής  $W$  είναι η  $F_{out}(\cdot)$ , έχουμε ότι η πιθανότητα  $p_{out}$  εμφάνισης ενός σημείου εκτός των ορίων ελέγχου του διαγράμματος είναι ίση με

$$p_{out} = 1 - P(LCL \leq W \leq UCL) = 1 - F_{out}(\mu + L\sigma) + F_{out}(-(\mu - L\sigma))$$

και το εκτός ελέγχου μέσο μήκος ροής είναι ίσο με

$$ARL_{out} = \frac{1}{p_{out}}$$

Φυσικά για μια διεργασία που βρίσκεται εκτός ελέγχου θέλουμε να έχουμε μικρή τιμή για το  $ARL_{out}$ , έτσι ώστε, να μειωθεί ο αριθμός των δειγμάτων (και συνεπώς ο χρόνος) που απαιτούνται για να γίνει αντιληπτό ότι η διεργασία είναι εκτός ελέγχου.

Αν για την τυχαία μεταβλητή  $W$  είναι γνωστό ότι  $W \sim N(\mu, \sigma^2)$  τότε έχουμε ότι το εντός ελέγχου μέσο μήκος ροής είναι ίσο με

$$ARL_{in} = \frac{1}{p_{in}} = \frac{1}{2\Phi(-L)}$$

αφού

$$\begin{aligned}
p_{in} &= 1 - P(LCL \leq W \leq UCL \mid W \sim N(\mu, \sigma^2)) \\
&= 1 - P(\mu - L\sigma \leq W \leq \mu + L\sigma \mid W \sim N(\mu, \sigma^2)) \\
&= 1 - \Phi(L) + \Phi(-L) = 2\Phi(-L).
\end{aligned}$$

Για  $L = 3$  προκύπτει ότι

$$ARL_{in} \cong \frac{1}{0.0027} \cong 370.$$

Αν κατά τη διάρκεια της παραγωγικής διεργασίας η μέση τιμή της τυχαίας μεταβλητής  $W$  μετατοπιστεί από τη θέση  $\mu$  στη θέση  $\mu^* = \mu + \delta\sigma$  (μετατόπιση εκφρασμένη σε μονάδες τυπικής απόκλισης) και η διακύμανσή της δεν αλλάξει έχουμε ότι

$$\begin{aligned}
p_{out} &= 1 - P(LCL \leq W \leq UCL \mid W \sim N(\mu + \delta\sigma, \sigma^2)) \\
&= 1 - P(\mu - L\sigma \leq W \leq \mu + L\sigma \mid W \sim N(\mu + \delta\sigma, \sigma^2)) \\
&= 1 - \Phi(L - \delta) + \Phi(-L - \delta) = 1 - \Phi(L - \delta) - \Phi(-L - \delta)
\end{aligned}$$

και

$$ARL_{out} = \frac{1}{p_{out}} = \frac{1}{2 - \Phi(L - \delta) - \Phi(L + \delta)}.$$

### 1.3.1 Διαγράμματα Ελέγχου για τη Μέση Τιμή

Στην παρούσα παράγραφο θα αναπτύξουμε την κατασκευή του διαγράμματος ελέγχου τύπου Shewhart για την παρακολούθηση της συμπεριφοράς της μέσης τιμής  $\mu$  ενός συνεχούς χαρακτηριστικού  $X$ .

Θέλουμε να δημιουργήσουμε διαγράμματα ελέγχου Φάσης II όπου  $\mu$  και  $\sigma$  μας είναι γνωστά. Ας υποθέσουμε ότι η κατανομή του χαρακτηριστικού  $X$  ακολουθεί την κανονική κατανομή  $N(\mu, \sigma^2)$ . Αν  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$  είναι ένα τυχαίο δείγμα μεγέθους  $n$  από την  $X$  τότε είναι γνωστό ότι ο δειγματικός μέσος

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

ακολουθεί την κατανομή  $N(\mu, \sigma^2/n)$ , και είναι αμερόληπτη εκτιμήτρια της μέσης τιμής  $\mu$  του ποιοτικού χαρακτηριστικού  $X$ . Έτσι για οποιοδήποτε δείγμα μεγέθους  $n$  ο δειγματικός μέσος  $W_i = \bar{X}_i$  παίρνει τιμές στο διάστημα

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

με πιθανότητα  $1 - a$ . Για  $a = 0.0027$ , δηλαδή για  $Z_{a/2} = 3$ , προκύπτει το διάστημα

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

και ο δειγματικός μέσος  $W_i = \bar{X}_i$  παίρνει τιμές σε αυτό το διάστημα με πιθανότητα 99.73%. Επομένως, χρησιμοποιώντας ένα διάγραμμα ελέγχου στο οποίο θα απεικονίζεται η τιμή του δειγματικού μέσου  $W_i = \bar{X}_i$  στα δείγματα που επιλέγουμε από την παραγωγή με

$$LCL = \mu_{W_i} - 3\sigma_{W_i} = \mu - 3 \frac{\sigma}{\sqrt{n}}, \quad CL = \mu_{W_i} = \mu, \quad UCL = \mu_{W_i} + 3\sigma_{W_i} = \mu + 3 \frac{\sigma}{\sqrt{n}}$$

και με την υπόθεση ότι η διακύμανση (ή γενικότερα η διασπορά) του χαρακτηριστικού  $X$  σε όλη τη διαδικασία παραμένει σταθερή μπορούμε να πούμε τα εξής: (α) εφόσον τα σημεία του διαγράμματος βρίσκονται εντός των ορίων ελέγχου θεωρούμε ότι η διεργασία είναι εντός ελέγχου οπότε η μέση τιμή  $\mu$  του χαρακτηριστικού  $X$  δεν έχει αλλάξει (μετατοπιστεί) και επομένως το 99.73% των σημείων του διαγράμματος ελέγχου θα βρίσκονται εντός των ορίων ελέγχου, και (β) στην περίπτωση που ένα σημείο του διαγράμματος βρεθεί εκτός των ορίων του διαγράμματος και επειδή η πιθανότητα αυτού του ενδεχομένου είναι πολύ μικρή (0.0027), τότε υπάρχει ένδειξη ότι η διεργασία είναι εκτός ελέγχου λόγω μετατόπισης της μέσης τιμής του χαρακτηριστικού  $X$ .

Τα όρια του διαγράμματος ελέγχου για τη μέση τιμή του χαρακτηριστικού  $X$  συνοψίζονται στο ακόλουθο πλαίσιο

**Πλαίσιο 1 ( $\bar{X}$  διάγραμμα)**

**Phase II control limits ( $3\sigma$ )**

$$\begin{aligned} UCL &= \mu + A\sigma \\ \text{Center Line} &= \mu \\ LCL &= \mu - A\sigma \end{aligned}$$

όπου

$$A = \frac{3}{\sqrt{n}}$$

Τα παραπάνω όρια ελέγχου του διαγράμματος είναι γνωστά ως όρια ελέγχου **Φάσης II** (*Phase II or on-line control phase or control to standard phase*) επειδή δεν αντιμετωπίζουμε καμία δυσκολία για την εύρεση των ορίων ελέγχου και της κεντρικής γραμμής αφού η μέση τιμή και η διακύμανση του χαρακτηριστικού  $X$  θεωρούνται γνωστά.

Θέλουμε να δημιουργήσουμε Διαγράμματα ελέγχου **Φάσης I** όπου  $\mu$  και  $\sigma$  μας είναι άγνωστα. Στην πράξη οι ποσότητες  $\mu$ ,  $\sigma$  είναι άγνωστες οπότε πρέπει να εκτιμηθούν. Για το σκοπό αυτό παίρνουμε  $m$  **ανεξάρτητα** προκαταρκτικά **τυχαία** δείγματα μεγέθους  $n$  το καθένα  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$ ,  $1 \leq i \leq m$ , για να εκτιμήσουμε τις ποσότητες  $\mu$  και  $\sigma$ , υποθέτοντας ότι η επιλογή των προκαταρκτικών δειγμάτων έγινε όταν η διεργασία ήταν εντός ελέγχου (τα  $\mu$  και  $\sigma$  σταθερά καθόλη τη διάρκεια της δειγματοληψίας). Οι εκτιμήσεις  $\hat{\mu}$  και  $\hat{\sigma}$  προκύπτουν συνήθως από 20 έως 25 προκαταρκτικά δείγματα μεγέθους 4 έως 6.

Στην ανάλυση που ακολουθεί υποθέτουμε ότι το χαρακτηριστικό  $X$  ακολουθεί κατανομή  $N(\mu, \sigma^2)$  ( $\mu$ ,  $\sigma$  άγνωστα) και ότι έχουμε  $m$  το πλήθος ανεξάρτητα τυχαία δείγματα μεγέθους  $n$  το καθένα.

- **Εκτίμηση του  $\mu$**

Έστω,  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$  οι δειγματικοί μέσοι των  $m$  δειγμάτων και ας θέσουμε

$$\bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_m}{m} = \frac{\sum_{i=1}^m \sum_{j=1}^n X_{ij}}{mn}$$

Η ποσότητα  $\bar{\bar{X}}$  ακολουθεί κατανομή  $N(\mu, \sigma^2 / nm)$  και χρησιμοποιείται ως εκτίμηση της ποσότητας  $\mu$  (αμερόληπτη και συνεπής εκτιμήτρια του  $\mu$ ), δηλαδή

$$\hat{\mu} = \bar{\bar{X}}$$

Αξίζει να σημειώσουμε ότι

$$E(\bar{\bar{X}}) = \mu, \quad V(\bar{\bar{X}}) = \sigma^2 / nm$$

ανεξάρτητα από την κατανομή του χαρακτηριστικού  $X$ .

- **Εκτίμηση του  $\sigma$  (Μέθοδος  $R$ )**

Έστω,  $R_1, R_2, \dots, R_m$  τα εύροι των  $m$  δειγμάτων, δηλαδή:

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

Μπορεί να δειχθεί ότι

$$\mu_{R_i} = E(R_i) = \sigma d_2$$

όπου  $d_2$  είναι μια σταθερά που εξαρτάται από το μέγεθος του δείγματος  $n$ . Θέτοντας

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

έχουμε ότι  $E(\bar{R}) = \sigma d_2$ , οπότε η ποσότητα  $\bar{R} / d_2$  είναι αμερόληπτη εκτιμήτρια της ποσότητας  $\sigma$ , και χρησιμοποιείται ως εκτίμηση της ποσότητας  $\sigma$ , δηλαδή

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

- **Εκτίμηση του  $\sigma$  (Μέθοδος  $S$ )**

Έστω,  $S_1, S_2, \dots, S_m$  οι ποσότητες που ορίζονται από τη σχέση

$$S_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}, \quad 1 \leq i \leq m.$$

Αν και η ποσότητα  $S_i^2$  είναι αμερόληπτη (και συνεπής) εκτιμήτρια της διακύμανσης  $\sigma^2$ , η ποσότητα  $S_i = \sqrt{S_i^2}$  δεν αποτελεί αμερόληπτη εκτιμήτρια της τυπικής απόκλισης  $\sigma$ . Μπορεί να δειχθεί ότι:

$$\mu_{S_i} = E(S_i) = \sigma c_4$$

όπου  $c_4$  είναι μια σταθερά που εξαρτάται από το μέγεθος του δείγματος  $n$ . Θέτοντας

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

έχουμε ότι  $E(\bar{S}) = \sigma c_4$ , οπότε η ποσότητα  $\bar{S} / c_4$  είναι αμερόληπτη εκτιμήτρια της ποσότητας  $\sigma$ , και χρησιμοποιείται ως εκτίμηση της ποσότητας  $\sigma$ , δηλαδή

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

- **Εκτίμηση του  $\sigma$  (Μέθοδος  $S^2$ )**

Έστω,  $S_1^2, S_2^2, \dots, S_m^2$  οι ποσότητες που ορίζονται από τη σχέση

$$S_i^2 = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2, \quad 1 \leq i \leq m$$

για τις οποίες είναι γνωστό ότι

$$\mu_{S_i^2} = E(S_i^2) = \sigma^2.$$

Η ποσότητα  $\sqrt{S^2}$  όπου

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

αν και δεν είναι αμερόληπτη εκτιμήτρια του  $\sigma$  χρησιμοποιείται αρκετές φορές (λόγω του ότι έχει μικρότερη διακύμανση από την  $\bar{S}/c_4$ ) ως εκτίμηση της ποσότητας  $\sigma$ , δηλαδή

$$\hat{\sigma} = \sqrt{\overline{S^2}}$$

- **Όρια ελέγχου**

Χρησιμοποιώντας ως εκτίμηση του  $\mu$  την ποσότητα  $\hat{\mu} = \bar{\bar{X}}$  και ως εκτίμηση του  $\sigma$  την ποσότητα  $\hat{\sigma} = \bar{R}/d_2$  τότε το Πλαίσιο 1 παίρνει την ακόλουθη μορφή

**Πλαίσιο 2 ( $\bar{X}$  διάγραμμα)**

**Phase I control limits ( $3\sigma$ )**

**R Method**

$$\begin{aligned} UCL &= \bar{\bar{X}} + A_2 \bar{R} \\ \text{Center Line} &= \bar{\bar{X}} \\ LCL &= \bar{\bar{X}} - A_2 \bar{R} \end{aligned}$$

όπου η ποσότητα  $A_2$  είναι ίση με

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



Χρησιμοποιώντας ως εκτίμηση του  $\mu$  την ποσότητα  $\hat{\mu} = \bar{\bar{X}}$  και ως εκτίμηση του  $\sigma$  την ποσότητα  $\hat{\sigma} = \bar{S} / c_4$  τότε το Πλαίσιο 1 παίρνει την ακόλουθη μορφή

<b>Πλαίσιο 3 (<math>\bar{X}</math> διάγραμμα)</b>	
<b>Phase I control limits (3<math>\sigma</math>)</b>	
<b>S Method</b>	
<i>UCL</i>	= $\bar{\bar{X}} + A_3\bar{S}$
<i>Center Line</i>	= $\bar{\bar{X}}$
<i>LCL</i>	= $\bar{\bar{X}} - A_3\bar{S}$

όπου

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

Χρησιμοποιώντας ως εκτίμηση του  $\mu$  την ποσότητα  $\hat{\mu} = \bar{\bar{X}}$  και ως εκτίμηση του  $\sigma$  την ποσότητα  $\hat{\sigma} = \sqrt{S^2}$  τότε το Πλαίσιο 1 παίρνει την ακόλουθη μορφή

<b>Πλαίσιο 4 (<math>\bar{X}</math> διάγραμμα)</b>	
<b>Phase I control limits (3<math>\sigma</math>)</b>	
<b>S<sup>2</sup> Method</b>	
<i>UCL</i>	= $\bar{\bar{X}} + A\sqrt{S^2}$
<i>Center Line</i>	= $\bar{\bar{X}}$
<i>LCL</i>	= $\bar{\bar{X}} - A\sqrt{S^2}$

### 1.3.2 Διαγράμματα Ελέγχου για τη Διασπορά

Στην προηγούμενη παράγραφο αναφέραμε ότι στην περίπτωση που ένα σημείο του διαγράμματος ελέγχου για τη μέση τιμή του χαρακτηριστικού  $X$  βρεθεί εκτός των ορίων του διαγράμματος τότε υπάρχει ένδειξη ότι η διεργασία είναι εκτός ελέγχου με την προϋπόθεση ότι η διασπορά του ποιοτικού χαρακτηριστικού  $X$  είχε παραμείνει σταθερή σε όλη τη διάρκεια της διαδικασίας. Στην παρούσα παράγραφο θα αναπτύξουμε την κατασκευή διαγραμμάτων τύπου *Shewhart* για την παρακολούθηση της συμπεριφοράς της διασποράς ενός συνεχούς ποιοτικού χαρακτηριστικού  $X$ .

Στην ανάλυση που ακολουθεί υποθέτουμε ότι το χαρακτηριστικό  $X$  ακολουθεί κατανομή  $N(\mu, \sigma^2)$  και ότι έχουμε  $m$  το πλήθος ανεξάρτητα τυχαία δείγματα από τη  $X$  τα  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$ ,  $1 \leq i \leq m$ , μεγέθους  $n$  το καθένα.

- **$R$  διάγραμμα ελέγχου**

Θέλουμε να φτιάξουμε διαγράμματα ελέγχου Φάσης II όπου  $\sigma$  γνωστό. Έστω  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$  ένα τυχαίο δείγμα μεγέθους  $n$  από τη  $X$  και ας θέσουμε  $W_i = R_i = X_{i(n)} - X_{i(1)}$ . Είναι γνωστό ότι

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

όπου οι ποσότητες  $d_2$  και  $d_3$  εξαρτώνται από το μέγεθος  $n$  του δείγματος.

Συνεπώς ένα διάγραμμα ελέγχου για τη διασπορά του ποιοτικού χαρακτηριστικού  $X$  μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι το εύρος  $R_i$  των δειγμάτων που ως γνωστό είναι ένα μέτρο διασποράς της  $X$ . Το μοντέλο με όρια  $3\sigma$  θα έχει τη μορφή

$$LCL = \mu_{R_i} - 3\sigma_{R_i} = (d_2 - 3d_3)\sigma, \quad CL = \mu_{R_i} = d_2\sigma,$$

$$UCL = \mu_{R_i} + 3\sigma_{R_i} = (d_2 + 3d_3)\sigma.$$

Θέτοντας

$$D_1 = d_2 - 3d_3, \quad D_2 = d_2 + 3d_3$$

προκύπτει το ακόλουθο πλαίσιο

**Πλαίσιο 5 (R διάγραμμα)**

**Phase II control limits (3σ)**

$$\begin{aligned}UCL &= D_2\sigma \\ \text{Center Line} &= d_2\sigma \\ LCL &= D_1\sigma\end{aligned}$$

Επειδή,  $D_1 < 0$  για  $n \leq 6$ , σε αυτές τις περιπτώσεις θέτουμε  $D_1 = 0$ .

• **S διάγραμμα ελέγχου**

Έστω  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$  είναι ένα τυχαίο δείγμα μεγέθους  $n$  από τη  $X$  και ως θέσουμε

$$S_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}, \quad 1 \leq i \leq m.$$

Γνωρίζουμε ότι

$$\mu_{S_i} = E(S_i) = \sigma c_4, \quad \sigma_{S_i} = \sqrt{V(S_i)} = \sigma \sqrt{1 - c_4^2}$$

Συνεπώς, ένα διάγραμμα ελέγχου για τη διασπορά του χαρακτηριστικού  $X$  μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι οι δειγματικές τυπικές αποκλίσεις  $S_i$  που είναι το πιο σύνηθες μέτρο διασποράς της  $X$ .

Το μοντέλο με όρια  $3\sigma$  θα έχει τη μορφή

$$\begin{aligned}LCL &= \mu_{S_i} - 3\sigma_{S_i} = (c_4 - 3\sqrt{1 - c_4^2})\sigma, \\ CL &= \mu_{S_i} = c_4\sigma, \dots UCL = \mu_{S_i} + 3\sigma_{S_i} = (c_4 + 3\sqrt{1 - c_4^2})\sigma.\end{aligned}$$

Θέτοντας

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

προκύπτει το ακόλουθο πλαίσιο

**Πλαίσιο 6 (S διάγραμμα)**

**Phase II control limits (3σ)**

$$\begin{aligned}UCL &= B_6\sigma \\ \text{Center Line} &= c_4\sigma \\ LCL &= B_5\sigma\end{aligned}$$

- Επειδή,  $B_5 < 0$  για  $n \leq 5$ , σε αυτές τις περιπτώσεις θέτουμε  $B_5 = 0$ .

- $S^2$  διάγραμμα ελέγχου

Έστω,  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$  είναι ένα τυχαίο δείγμα μεγέθους  $n$  από τη  $X$  και ας θέσουμε

$$S_i^2 = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2, \quad 1 \leq i \leq m.$$

Είναι γνωστό ότι

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

Επιπρόσθετα,

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

οπότε

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

Συνεπώς ένα διάγραμμα ελέγχου για τη διασπορά του χαρακτηριστικού  $X$  μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι η δειγματική διακύμανση  $S_i^2$  που είναι μέτρο διασποράς της  $X$ , το οποίο περιγράφεται στο ακόλουθο πλαίσιο

<b>Πλαίσιο 7 (<math>S^2</math> διάγραμμα)</b>	
<b>Phase II control limits (<math>a/2</math>)</b>	
$UCL$	$= \frac{\sigma^2}{n-1} \chi_{n-1;a/2}^2$
$Center\ Line$	$= \sigma^2$
$LCL$	$= \frac{\sigma^2}{n-1} \chi_{n-1;1-a/2}^2$

Πρέπει να σημειώσουμε ότι τα παραπάνω όρια ελέγχου είναι όρια πιθανότητας (*probability limits*) αφού δεν συνηθίζεται η ανάπτυξη διαγραμμάτων ελέγχου ορίων  $3\sigma$  για την ποσότητα  $S^2$ .

Θέλουμε να φτιάξουμε Διαγράμματα ελέγχου Φάσης I όπου  $\sigma$  άγνωστο. Στην πράξη η ποσότητα  $\sigma$  είναι άγνωστη οπότε πρέπει να εκτιμηθεί παίρνοντας προκαταρκτικά ανεξάρτητα τυχαία δείγματα.

Στην ανάλυση που ακολουθεί υποθέτουμε ότι το χαρακτηριστικό  $X$  ακολουθεί κατανομή  $N(\mu, \sigma^2)$  και ότι έχουμε  $m$  το πλήθος ανεξάρτητα τυχαία δείγματα από τη  $X$  τα  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$ ,  $1 \leq i \leq m$ , μεγέθους  $n$  το καθένα.

- **R διάγραμμα ελέγχου**

Χρησιμοποιώντας ως εκτίμηση του  $\sigma$  την ποσότητα  $\hat{\sigma} = \bar{R} / d_2$  τότε το Πλαίσιο 5 παίρνει την ακόλουθη μορφή

<b>Πλαίσιο 8 (R διάγραμμα)</b>		
<b>Phase I control limits (3<math>\sigma</math>)</b>		
$UCL$	$=$	$D_4 \bar{R}$
$Center\ Line$	$=$	$\bar{R}$
$LCL$	$=$	$D_3 \bar{R}$

όπου

$$D_3 = 1 - 3 \frac{d_3}{d_2}, \quad D_4 = 1 + 3 \frac{d_3}{d_2}$$

Επειδή,  $D_3 < 0$  για  $n \leq 6$ , σε αυτές τις περιπτώσεις θέτουμε  $D_3 = 0$ .

- **S διάγραμμα ελέγχου**

Χρησιμοποιώντας ως εκτίμηση του  $\sigma$  την ποσότητα  $\hat{\sigma} = \bar{S} / c_4$  τότε το Πλαίσιο 6 παίρνει την ακόλουθη μορφή

<b>Πλαίσιο 9 (S διάγραμμα)</b>		
<b>Phase I control limits (3<math>\sigma</math>)</b>		
$UCL$	$=$	$B_4 \bar{S}$
$Center\ Line$	$=$	$\bar{S}$
$LCL$	$=$	$B_3 \bar{S}$

όπου

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

Επειδή,  $B_3 < 0$  για  $n \leq 5$ , σε αυτές τις περιπτώσεις θέτουμε  $B_3 = 0$ .

- **$S^2$  διάγραμμα ελέγχου**

Χρησιμοποιώντας ως εκτίμηση του  $\sigma^2$  την ποσότητα  $\hat{\sigma}^2 = \bar{S}^2$  το Πλαίσιο 7 παίρνει την ακόλουθη μορφή

<b>Πλαίσιο 10 (<math>S^2</math> διάγραμμα)</b>	
<b>Phase I control limits (<math>\alpha/2</math>)</b>	
$UCL$	$= \frac{\bar{S}^2}{n-1} \chi_{n-1; \alpha/2}^2$
$Center\ Line$	$= \bar{S}^2$
$LCL$	$= \frac{\bar{S}^2}{n-1} \chi_{n-1; 1-\alpha/2}^2$

Σημειώσουμε πάλι ότι τα παραπάνω όρια ελέγχου είναι όρια πιθανότητας (*probability limits*).

### 1.3.3 Διαγράμματα Ελέγχου για Μεμονωμένες Παρατηρήσεις

Σε ορισμένες περιπτώσεις το μέγεθος του δείγματος είναι ίσο με 1 (αυτόματη επιθεώρηση παραγόμενων προϊόντων, μικρός ρυθμός παραγωγής, κτλ.). Σε αυτή την περίπτωση ομιλούμε για διαγράμματα ελέγχου για **μεμονωμένες ή ατομικές παρατηρήσεις** (*individual observations*). Σε ότι ακολουθήσει στην παρούσα παράγραφο υποθέτουμε ότι το χαρακτηριστικό  $X$  ακολουθεί κατανομή  $N(\mu, \sigma^2)$  και ότι έχουμε στη διάθεσή μας  $m$  ανεξάρτητες παρατηρήσεις  $X_1, X_2, \dots, X_m$  από την  $X$  για την ανάλυση Φάσης  $I$  (Δ. Αντζουλάκος, 2003).

Θέλουμε να φτιάξουμε διαγράμματα ελέγχου Φάσης  $II$  όπου  $\mu$  και  $\sigma$  είναι γνωστά. Σε αυτή την περίπτωση το Πλαίσιο 1 παίρνει τη μορφή

**Πλαίσιο 11 (X διάγραμμα)****Phase II control limits (3σ)**

$$\begin{aligned}UCL &= \mu + 3\sigma \\ \text{Center Line} &= \mu \\ LCL &= \mu - 3\sigma\end{aligned}$$

αφού  $n=1$ , και ομιλούμε για  $X$  διάγραμμα ελέγχου αφού στο διάγραμμα απεικονίζονται οι μεμονωμένες παρατηρήσεις  $X_i$  ( $W_i = X_i, i \geq 1$ ).

Για τη συμπεριφορά της διασποράς του χαρακτηριστικού  $X$  το αντίστοιχο  $R$  διάγραμμα ελέγχου δεν μπορεί να χρησιμοποιηθεί επειδή για  $n=1$  δεν έχει νόημα η ποσότητα  $R$ . Σε αυτή την περίπτωση χρησιμοποιούμε το **κινούμενο εύρος** (*moving range*) των μεμονωμένων παρατηρήσεων που ορίζεται από τη σχέση

$$MR_i = |X_i - X_{i-1}|, \quad 2 \leq i \leq m.$$

Προφανώς, για το κινούμενο εύρος  $MR_i$  ισχύει ότι

$$\mu_{MR_i} = E(MR_i) = \sigma d_2, \quad \sigma_{MR_i} = \sqrt{V(MR_i)} = \sigma d_3.$$

Συνεπώς, ένα διάγραμμα ελέγχου για τη διασπορά του χαρακτηριστικού  $X$  μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι το κινούμενο εύρος  $MR_i$  των παρατηρήσεων. Το μοντέλο με όρια  $3\sigma$  θα έχει τη μορφή

$$LCL = \mu_{MR_i} - 3\sigma_{MR_i} = (d_2 - 3d_3)\sigma, \quad CL = \mu_{MR_i} = d_2\sigma,$$

$$UCL = \mu_{MR_i} + 3\sigma_{MR_i} = (d_2 + 3d_3)\sigma.$$

Για  $D_1 = d_2 - 3d_3$ ,  $D_2 = d_2 + 3d_3$  προκύπτει το ακόλουθο πλαίσιο

**Πλαίσιο 12 (MR διάγραμμα)****Phase II control limits (3σ)**

$$\begin{aligned}UCL &= D_2\sigma \\ \text{Center Line} &= d_2\sigma \\ LCL &= D_1\sigma\end{aligned}$$

Θέλουμε να φτιάξουμε διαγράμματα ελέγχου Φάσης  $I$  όπου  $\mu$  και  $\sigma$  είναι άγνωστα. Στην περίπτωση που οι ποσότητες  $\mu$  και  $\sigma$  είναι άγνωστες πρέπει να εκτιμηθούν. Η εκτίμηση του μέσου δίνεται από τη σχέση

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

Η ποσότητα  $\bar{X}$  ακολουθεί κατανομή  $N(\mu, \sigma^2/m)$  και χρησιμοποιείται ως εκτίμηση της ποσότητας  $\mu$  (αμερόληπτη και συνεπής εκτιμήτρια του  $\mu$ ), δηλαδή

$$\hat{\mu} = \bar{X}$$

Για την εκτίμηση του  $\sigma$ , θέτοντας

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

έχουμε ότι  $E(\overline{MR}) = \sigma d_2$ , οπότε η ποσότητα  $\overline{MR}/d_2$  είναι αμερόληπτη εκτιμήτρια της ποσότητας  $\sigma$ , και χρησιμοποιείται ως εκτίμηση της ποσότητας  $\sigma$ , δηλαδή

$$\hat{\sigma} = \frac{\overline{MR}}{d_2}$$

Χρησιμοποιώντας τις παραπάνω εκτιμήσεις το Πλαίσιο 11 παίρνει την ακόλουθη μορφή

**Πλαίσιο 13 (X διάγραμμα)**  
**Phase I control limits (3σ)**

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

$$Center\ Line = \bar{X}$$

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

ενώ το Πλαίσιο 12, για

$$D_3 = 1 - 3 \frac{d_3}{d_2}, \quad D_4 = 1 + 3 \frac{d_3}{d_2}$$

παίρνει τη μορφή

**Πλαίσιο 14 (MR διάγραμμα)**  
**Phase I control limits (3σ)**

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

$$Center\ Line = \overline{MR}$$

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



### 1.3.4 Διαγράμματα Ελέγχου για το Ποσοστό και τον Αριθμό των Ελαττωματικών Προϊόντων

Στην παρούσα παράγραφο θα αναπτύξουμε την κατασκευή διαγράμματος τύπου *Shewhart* για την παρακολούθηση της συμπεριφοράς του ποσοστού  $p$  των ελαττωματικών ή μη συμμορφούμενων προϊόντων μιας παραγωγικής διεργασίας.

Ας υποθέσουμε ότι το ποσοστό των ελαττωματικών προϊόντων είναι γνωστό και ίσο με  $p$  και ότι από την παραγωγή επιλέγουμε ανεξάρτητα τυχαία δείγματα μεγέθους  $n$  προϊόντων το καθένα. Ας συμβολίσουμε με  $X_{ij}$ ,  $i \geq 1$ ,  $1 \leq j \leq n$ , την τυχαία μεταβλητή με τιμές 1 και 0 ανάλογα με το αν το αντίστοιχο προϊόν  $(i, j)$  είναι ελαττωματικό ή όχι. Για την τυχαία μεταβλητή  $X_{ij}$  έχουμε ότι  $X_{ij} \sim B(1, p)$ , ενώ για την τυχαία μεταβλητή  $X_i = X_{i1} + X_{i2} + \dots + X_{in}$  που δηλώνει τον αριθμό των ελαττωματικών προϊόντων στο  $i$  δείγμα έχουμε ότι  $X_i \sim B(n, p)$ .

Για την τυχαία μεταβλητή

$$W_i = p_i = \frac{X_i}{n}, \quad i \geq 1$$

που δηλώνει το ποσοστό των ελαττωματικών προϊόντων στο  $i$  δείγμα ισχύει ότι

$$\mu_{W_i} = p, \quad \sigma_{W_i}^2 = \frac{p(1-p)}{n}, \quad i \geq 1.$$

Επομένως, μπορούμε να αναπτύξουμε ένα διάγραμμα ελέγχου για την παρακολούθηση του ποσοστού των ελαττωματικών προϊόντων στο οποίο θα απεικονίζεται η τιμή της  $W_i = p_i = X_i/n$  στα διάφορα δείγματα που επιλέγουμε από την παραγωγή. Τα όρια του διαγράμματος ελέγχου συνοψίζονται στο ακόλουθο πλαίσιο

#### Πλαίσιο 15 ( $p$ διάγραμμα )

#### Phase II control limits ( $3\sigma$ )

$$\begin{aligned} UCL &= p + 3\sqrt{\frac{p(1-p)}{n}} \\ \text{Center Line} &= p \\ LCL &= p - 3\sqrt{\frac{p(1-p)}{n}} \end{aligned}$$

Η κατασκευή ενός διαγράμματος ελέγχου για την παρακολούθηση του αριθμού των ελαττωματικών προϊόντων επιτυγχάνεται με τη βοήθεια της τυχαίας μεταβλητής  $X_i$ . Στο διάγραμμα ελέγχου θα απεικονίζεται η τιμή της  $X_i (= W_i)$  στα διάφορα δείγματα που επιλέγουμε από την παραγωγή, και τα όρια του διαγράμματος ελέγχου δίνονται στο ακόλουθο πλαίσιο

<b>Πλαίσιο 16 (<math>np</math> διάγραμμα )</b>	
<b>Phase II control limits (<math>3\sigma</math>)</b>	
$UCL$	$= np + 3\sqrt{np(1-p)}$
$Center Line$	$= np$
$LCL$	$= np - 3\sqrt{np(1-p)}$

Όταν το ποσοστό των ελαττωματικών προϊόντων της διεργασίας είναι άγνωστο πρέπει να εκτιμηθεί από τις πληροφορίες που θα μας δώσουν  $m$  ανεξάρτητα προκαταρκτικά δείγματα μεγέθους  $n$  το καθένα, έστω τα  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$ ,  $1 \leq i \leq m$ . Θέτοντας

$$p_i = \frac{X_i}{n} = \frac{X_{i1} + X_{i2} + \dots + X_{in}}{n}, \quad 1 \leq i \leq m$$

$$\bar{P} = \frac{p_1 + p_2 + \dots + p_m}{m} = \frac{X_1 + X_2 + \dots + X_m}{mn} = \frac{\sum_{i=1}^m \sum_{j=1}^n X_{ij}}{mn}$$

έχουμε ότι  $E(\bar{P}) = p$  αφού  $\sum_{i=1}^m \sum_{j=1}^n X_{ij} \sim B(nm, p)$ . Η ποσότητα  $\bar{P}$  χρησιμοποιείται ως εκτίμηση της ποσότητας  $p$ , δηλαδή

$$\hat{p} = \bar{P}$$

Συνεπώς, το Πλαίσιο 15 παίρνει τη μορφή

**Πλαίσιο 17 (  $p$  διάγραμμα )**

**Phase I control limits (3 $\sigma$ )**

$$\begin{aligned}UCL &= \bar{P} + 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}} \\Center\ Line &= \bar{P} \\LCL &= \bar{P} - 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}}\end{aligned}$$

ενώ το Πλαίσιο 16 παίρνει τη μορφή

**Πλαίσιο 18 (  $np$  διάγραμμα )**

**Phase I control limits (3 $\sigma$ )**

$$\begin{aligned}UCL &= n\bar{P} + 3\sqrt{n\bar{P}(1-\bar{P})} \\Center\ Line &= n\bar{P} \\LCL &= n\bar{P} - 3\sqrt{n\bar{P}(1-\bar{P})}\end{aligned}$$

### 1.3.5 Διαγράμματα Ελέγχου για τον Αριθμό και το Μέσο Αριθμό των Ελαττωμάτων

Στην παρούσα παράγραφο θα αναπτύξουμε την κατασκευή διαγράμματος τύπου Shewhart για τον (συνολικό) αριθμό των ελαττωμάτων σε μια μονάδα ελέγχου και για το μέσο αριθμό των ελαττωμάτων ανά μονάδα ελέγχου. Η βασική υπόθεση που θα κάνουμε είναι ότι ο (συνολικός) αριθμός των ελαττωμάτων (πιθανόν διαφορετικών τύπων) σε μια μονάδα ελέγχου ακολουθεί την κατανομή *Poisson*.

Ας υποθέσουμε ότι ο αριθμός  $X$  των ελαττωμάτων που εμφανίζονται σε μια μονάδα ελέγχου ακολουθεί την κατανομή Poisson με παράμετρο  $c$ .

Είναι γνωστό ότι  $\mu_X = \sigma_X^2 = c$ . Συνεπώς μπορούμε να αναπτύξουμε ένα διάγραμμα ελέγχου για την παρακολούθηση του **αριθμού των ελαττωμάτων των μονάδων ελέγχου** στο οποίο θα απεικονίζεται η τιμή της  $W_i = X_i$  στις διάφορες μονάδες ελέγχου που επιλέγουμε από την παραγωγή. Τα όρια του διαγράμματος ελέγχου συνοψίζονται στο ακόλουθο πλαίσιο

**Πλαίσιο 19 ( c διάγραμμα )****Phase II control limits (3σ)**

$$\begin{aligned}UCL &= c + 3\sqrt{c} \\ \text{Center Line} &= c \\ LCL &= c - 3\sqrt{c}\end{aligned}$$

Έστω τώρα ότι από την παραγωγή επιλέγουμε ανεξάρτητα τυχαία δείγματα μεγέθους  $n$  μονάδων το καθένα. Ας συμβολίσουμε με  $X_{ij}$ ,  $i \geq 1$ ,  $1 \leq j \leq n$ , την τυχαία μεταβλητή που δηλώνει τον αριθμό των ελαττωμάτων της αντίστοιχης μονάδος ελέγχου  $(i, j)$ . Για την τυχαία μεταβλητή  $X_{ij}$  έχουμε ότι  $X_{ij} \sim P(c)$ , ενώ για την τυχαία μεταβλητή  $X_i = X_{i1} + X_{i2} + \dots + X_{in}$  που δηλώνει το συνολικό αριθμό των ελαττωμάτων στο  $i$  δείγμα έχουμε ότι  $X_i \sim P(nc)$ .

Για την τυχαία μεταβλητή  $U_i = X_i / n$  που δηλώνει το μέσο αριθμό των ελαττωμάτων ανά μονάδα ελέγχου στο  $i$  δείγμα έχουμε ότι

$$\mu_{U_i} = c, \quad \sigma_{U_i}^2 = \frac{c}{n}, \quad i \geq 1.$$

Συνεπώς, μπορούμε να αναπτύξουμε ένα διάγραμμα ελέγχου για την παρακολούθηση του **μέσου αριθμού των ελαττωμάτων ανά μονάδα ελέγχου** σε κάθε δείγμα στο οποίο θα απεικονίζεται η τιμή της  $U_i = X_i / n$  στα διάφορα δείγματα που επιλέγουμε από την παραγωγή. Τα όρια του διαγράμματος ελέγχου συνοψίζονται στο ακόλουθο πλαίσιο

**Πλαίσιο 20 (u διάγραμμα )****Phase II control limits (3σ)**

$$\begin{aligned}UCL &= c + 3\sqrt{\frac{c}{n}} \\ \text{Center Line} &= c \\ LCL &= c - 3\sqrt{\frac{c}{n}}\end{aligned}$$

Σχετικά με το  $c$  διάγραμμα όταν η παράμετρος  $c$  της κατανομής Poisson είναι άγνωστη πρέπει να εκτιμηθεί από τις πληροφορίες που θα μας δώσουν  $m$  προκαταρκτικές μονάδες ελέγχου. Ας συμβολίσουμε με  $X_i$  τον αριθμό των ελαττωμάτων της  $i$  μονάδος ελέγχου,  $1 \leq i \leq m$ . Θέτοντας

$$\bar{C} = \frac{X_1 + X_2 + \dots + X_m}{m}, \quad 1 \leq i \leq m$$

έχουμε ότι  $E(\bar{C}) = c$ . Η ποσότητα  $\bar{C}$  χρησιμοποιείται ως εκτίμηση της ποσότητας  $c$ , δηλαδή

$$\hat{c} = \bar{C}$$

Συνεπώς το Πλαίσιο 19 παίρνει τη μορφή

<b>Πλαίσιο 21 ( <math>c</math> διάγραμμα )</b>	
<b>Phase I control limits (3<math>\sigma</math>)</b>	
$UCL$	$= \bar{C} + 3\sqrt{\bar{C}}$
$Center\ Line$	$= \bar{C}$
$LCL$	$= \bar{C} - 3\sqrt{\bar{C}}$

Στην περίπτωση του  $u$  διαγράμματος η παράμετρος  $c$  της κατανομής Poisson θα εκτιμηθεί από τις πληροφορίες που θα μας δώσουν  $m$  ανεξάρτητα προκαταρκτικά δείγματα μεγέθους  $n$  μονάδων ελέγχου το καθένα, έστω τα  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$ ,  $1 \leq i \leq m$ . Θέτοντας

$$U_i = \frac{X_i}{n} = \frac{X_{i1} + X_{i2} + \dots + X_{in}}{n}, \quad 1 \leq i \leq m$$

και

$$\bar{U} = \frac{U_1 + U_2 + \dots + U_m}{m} = \frac{X_1 + X_2 + \dots + X_m}{mn} = \frac{\sum_{i=1}^m \sum_{j=1}^n X_{ij}}{mn}$$

έχουμε ότι  $E(\bar{U}) = c$  αφού  $\sum_{i=1}^m \sum_{j=1}^n X_{ij} \sim P(mnc)$ . Η ποσότητα  $\bar{U}$  χρησιμοποιείται ως εκτίμηση της ποσότητας  $c$ , δηλαδή:

$$\hat{c} = \bar{U}$$

Συνεπώς το Πλαίσιο 20 παίρνει τη μορφή

### Πλαίσιο 22 (u διάγραμμα )

#### Phase I control limits (3σ)

$$UCL = \bar{U} + 3\sqrt{\frac{\bar{U}}{n}}$$

$$\text{Center Line} = \bar{U}$$

$$LCL = \bar{U} - 3\sqrt{\frac{\bar{U}}{n}}$$

## 1.5 Διαγράμματα Ελέγχου CUSUM

Μέχρι τώρα έχουμε διαπιστώσει ότι στα διαγράμματα ελέγχου που αναφέραμε απεικονίζονται τιμές που βασίζονται στις μετρήσεις κάθε δείγματος χωρίς να λαμβάνουν υπόψη τους μετρήσεις από προηγούμενα δείγματα. Για το λόγο αυτό χαρακτηρίζουμε τα διαγράμματα ελέγχου Shewhart ως διαγράμματα ελέγχου χωρίς μνήμη (*control charts without memory*). Τα διαγράμματα Shewhart μας φανερώνουν μεγάλες μετατοπίσεις του μέσου και διαγραμματικά τις εντοπίζουμε γρήγορα για τιμές μεγαλύτερες του  $2\sigma$ .

Τα διαγράμματα ελέγχου τύπου CUSUM έχουν μη περιορισμένη και ομοιόμορφη μνήμη αφού λαμβάνουν πληροφορία από όλα τα προηγούμενα δείγματα και το καθένα από αυτά έχει την ίδια βαρύτητα. Χρησιμοποιούνται για να εντοπίζουν γρήγορα μικρές μετατοπίσεις του μέσου και ανήκουν στην κατηγορία των διαγραμμάτων με μνήμη (*control charts with memory*).

Ας θεωρήσουμε μια εντός ελέγχου διεργασία η οποία παράγει προϊόντα στα οποία η τιμή  $X$  ενός ποιοτικού χαρακτηριστικού τους που θέλουμε να παρακολουθήσουμε έχει μέση τιμή  $\mu_0$  και τυπική απόκλιση  $\sigma$ . Από τη διεργασία λαμβάνονται μεμονωμένες παρατηρήσεις  $X_i$ ,  $i \geq 1$ , και μας ενδιαφέρει να ανιχνεύσουμε μετατοπίσεις του μέσου της διεργασίας της μορφής  $\mu_1 = \mu_0 \pm \delta\sigma$  ( $\delta > 0$ ). Τα δίπλευρα (συμμετρικά) διαγράμματα ελέγχου CUSUM (*two-sided CUSUMS*) με διαστήματα απόφασης εισήχθησαν από τον Page (1954). Στα

διαγράμματα αυτά απεικονίζονται ταυτοχρόνως οι δύο ακόλουθες στατιστικές συναρτήσεις

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

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

Οι τιμές των ποσοτήτων  $S_0^+$  και  $S_0^-$  ονομάζονται τιμές εκκίνησης (*headstart values*). Η ποσότητα  $K$  ( $\geq 0$ ) ονομάζεται τιμή αναφοράς (*reference value*) και η πιο συνήθης τιμή της δίνεται από τη σχέση

$$K = \frac{\delta\sigma}{2} = \frac{|\mu_1 - \mu_0|}{2} = k\sigma, \quad k = \delta/2.$$

Η ποσότητα  $S_t^+$  συσσωρεύει τις αποκλίσεις των παρατηρήσεων  $X_t$  από την ποσότητα  $\mu_0 + K$  από τη στιγμή που θα εμφανιστεί θετική απόκλιση και μπορεί να θεωρηθεί κατάλληλη για τον έλεγχο της υπόθεσης

$$H_0 : \mu = \mu_0 \quad \text{—} \quad H_1^+ : \mu = \mu_1 = \mu_0 + \delta\sigma, \quad \delta > 0$$

αφού μεγάλες θετικές τιμές της  $S_t^+$  οδηγούν στην αποδοχή της υπόθεσης  $H_1^+$ .

Ανάλογα, η ποσότητα  $S_t^-$  συσσωρεύει τις αποκλίσεις των παρατηρήσεων  $X_t$  από την ποσότητα  $\mu_0 - K$  από τη στιγμή που θα εμφανιστεί αρνητική απόκλιση και συνεπώς η ποσότητα  $S_t^-$  μπορεί να θεωρηθεί κατάλληλη για τον έλεγχο της υπόθεσης

$$H_0 : \mu = \mu_0 \quad \text{—} \quad H_1^- : \mu = \mu_1 = \mu_0 - \delta\sigma, \quad \delta > 0$$

αφού μικρές αρνητικές τιμές της  $S_t^-$  οδηγούν στην αποδοχή της υπόθεσης  $H_1^-$ .

Για το ποια από τις δύο εναλλακτικές υποθέσεις  $H_1^+$  και  $H_1^-$  θα αποδεχθούμε ή όχι σε κάθε βήμα της διαδικασίας η απόφασή μας θα εξαρτηθεί από το αν ισχύει η σχέση  $S_t^+ > H$  ή η σχέση  $S_t^- < -H$ , όπου  $H$  μια θετική σταθερά ( $H > 0$ ). Η πιο συνήθης τιμή για το διάστημα απόφασης είναι η

$$H = h\sigma, \quad h = 4,5.$$

Η ποσότητα  $H$  ονομάζεται διάστημα απόφασης (*decision interval*). Φυσικά αν μια από τις δύο εναλλακτικές υποθέσεις  $H_1^+$  ή  $H_1^-$  γίνει αποδεκτή τότε θεωρούμε ότι η διεργασία είναι εκτός ελέγχου λόγω μετατόπισης του μέσου της διεργασίας σε υψηλότερο ή χαμηλότερο επίπεδο αντίστοιχα.

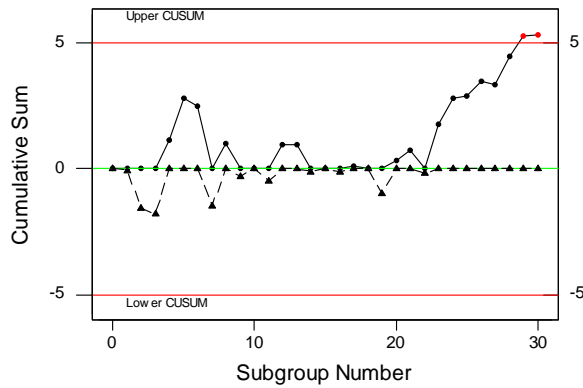
Για τα δεδομένα του Πίνακα 1 έχουμε τα ακόλουθα αποτελέσματα για  $K = |\mu_1 - \mu_0|/2 = |10 - 11|/2 = 0.5$  (επίσης  $k = 0.5, \sigma = 1$ ).

**ΠΙΝΑΚΑΣ 1:** Συσσωρευμένα αθροίσματα για  $K = 0.5$

Παρατήρηση	$X_t$	$X_t - (\mu_0 + K)$	$S_t^+$	$X_t - (\mu_0 - K)$	$S_t^-$
1	9.45	-1.05	0	-0.05	-0.05
2	7.99	-2.51	0	-1.51	-1.56
3	9.29	-1.21	0	-0.21	-1.77
4	11.66	1.16	1.16	2.16	0
5	12.16	1.66	2.82	2.66	0
6	10.18	-0.32	2.5	0.68	0
7	8.04	-2.46	0.04	-1.46	-1.46
8	11.46	0.96	1	1.96	0
9	9.2	-1.3	0	-0.3	-0.3
10	10.34	-0.16	0	0.84	0
11	9.03	-1.47	0	-0.47	-0.47
12	11.47	0.97	0.97	1.97	0
13	10.51	0.01	0.98	1.01	0
14	9.4	-1.1	0	-0.1	-0.1
15	10.08	-0.42	0	0.58	0
16	9.37	-1.13	0	-0.13	-0.13
17	10.62	0.12	0.12	1.12	0
18	10.31	-0.19	0	0.81	0
19	8.52	-1.98	0	-0.98	-0.98
20	10.84	0.34	0.34	1.34	0
21	10.9	0.4	0.74	1.4	0
22	9.33	-1.17	0	-0.17	-0.17
23	12.29	1.79	1.79	2.79	0
24	11.5	1	2.79	2	0
25	10.6	0.1	2.89	1.1	0
26	11.08	0.58	3.47	1.58	0
27	10.38	-0.12	3.35	0.88	0
28	11.62	1.12	4.47	2.12	0
29	11.31	0.81	5.28	1.81	0
30	10.52	0.02	5.3	1.02	0

Για  $H = 5\sigma = 5$  (επίσης  $h = 5$ ), το διάγραμμα ελέγχου CUSUM ( $S_t^+$  και  $S_t^-$  στο ίδιο διάγραμμα) είναι το ακόλουθο





**Διάγραμμα 2:** Διάγραμμα ελέγχου CUSUM για τα δεδομένα του Πίνακα 1 για  $K = 0.5$  και  $H = 5$

Από το παραπάνω διάγραμμα προκύπτει ότι  $S_{29}^+ > 5$  και συνεπώς η διεργασία είναι εκτός ελέγχου λόγω μετατόπισης του μέσου της διεργασίας σε υψηλότερο επίπεδο.

Στη γενική περίπτωση, τόσο η τιμή αναφοράς  $K$  όσο και το διάστημα απόφασης  $H$  μπορούν να διαφέρουν σε ένα δίπλευρο διάγραμμα ελέγχου CUSUM. Τότε ομιλούμε για μη συμμετρικό διάγραμμα ελέγχου CUSUM. Επίσης, οι τιμές εκκίνησης  $S_0^+$  και  $S_0^-$  δεν είναι υποχρεωτικό να είναι ίσες με το 0. Όταν  $0 < S_0^+ < H$  ή/και  $-H < S_0^- < 0$  ομιλούμε για εφαρμογή της μεθόδου της άμεσης αρχικής αντίδρασης (*fast initial response*). Συνεπώς, στη γενική περίπτωση σε ένα δίπλευρο διάγραμμα ελέγχου CUSUM απεικονίζονται οι ποσότητες

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

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

όπου

$$K^+ > 0, \quad H^+ > 0, \quad K^- < 0, \quad H^- < 0.$$

Το παραπάνω διάγραμμα δίνει σήμα εκτός ελέγχου διεργασίας τη χρονική στιγμή  $t$  αν  $S_t^+ > H^+ > 0$  ή  $S_t^- < H^- < 0$ . Φυσικά αν μας ενδιαφέρει να ανιχνεύσουμε μετατοπίσεις του μέσου σε υψηλότερο (χαμηλότερο) επίπεδο στο διάγραμμα ελέγχου CUSUM θα απεικονιστεί μόνο η ποσότητα  $S_t^+$  ( $S_t^-$ ) και το

διάστημα απόφασης  $H^+$  ( $H^-$ ). Σε αυτή την περίπτωση ομιλούμε για μονόπλευρα διαγράμματα ελέγχου συσσωρευμένων αθροισμάτων (*one-sided CUSUM*).

Στη βιβλιογραφία αναφέρονται επίσης και τα τυποποιημένα διαγράμματα ελέγχου CUSUM όπου χρησιμοποιούνται οι τυποποιημένες τιμές

$$Y_t = \frac{X_t - \mu_0}{\sigma}, \quad t \geq 1.$$

Σε αυτή την περίπτωση είναι εύκολο να διαπιστωθεί ότι σε ένα δίπλευρο συμμετρικό διάγραμμα ελέγχου CUSUM απεικονίζονται οι στατιστικές συναρτήσεις:

$$S_t^+ = \max[0, Y_t - k + S_{t-1}^+]$$

$$S_t^- = \min[0, Y_t + k + S_{t-1}^-]$$

και αν  $S_t^+ > h$  ( $S_t^- < -h$ ) τότε ο μέσος της διεργασίας έχει μετατοπιστεί σε υψηλότερο (χαμηλότερο) επίπεδο.

Τα διαγράμματα ελέγχου CUSUM μπορούν να χρησιμοποιηθούν και στην περίπτωση που δεν έχουμε μεμονωμένες παρατηρήσεις αλλά δείγματα μεγέθους  $n > 1$ . Σε αυτή την περίπτωση η ποσότητα  $X_t$  θα πρέπει να αντικατασταθεί με την ποσότητα  $\bar{X}_t$  (δηλαδή με το μέσο του  $t$  δείγματος) και η ποσότητα  $\sigma$  με την ποσότητα  $\sigma/\sqrt{n}$ . Έτσι, στο δίπλευρο συμμετρικό διάγραμμα ελέγχου CUSUM απεικονίζονται οι στατιστικές

$$S_t^+ = \max[0, \bar{X}_t - (\mu_0 + K) + S_{t-1}^+]$$

$$S_t^- = \min[0, \bar{X}_t - (\mu_0 - K) + S_{t-1}^-]$$

όπου

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

## 1.6 Διαγράμματα Ελέγχου EWMA

Ας θεωρήσουμε μια εντός ελέγχου διεργασία η οποία παράγει προϊόντα στα οποία η τιμή  $X$  ενός ποιοτικού χαρακτηριστικού τους που θέλουμε να παρακολουθήσουμε έχει μέση τιμή  $\mu_0$  και τυπική απόκλιση  $\sigma$ . Από τη διεργασία λαμβάνονται ανεξάρτητες μεμονωμένες παρατηρήσεις  $X_i$ ,  $i \geq 1$ , και μας ενδιαφέρει να ανιχνεύσουμε μετατοπίσεις του μέσου της διεργασίας της μορφής  $\mu_1 = \mu_0 \pm \delta\sigma$ . Στα διαγράμματα ελέγχου εκθετικά σταθμισμένου κινητού μέσου ή απλά διαγράμματα ελέγχου EWMA (*exponentially weighted moving average*) τα οποία εισήχθησαν από τον Roberts (1959), απεικονίζεται η στατιστική

$$Z_t = (1 - \lambda)Z_{t-1} + \lambda X_t, \quad t \geq 1, \quad 0 < \lambda \leq 1.$$

Για την εκκίνηση του παραπάνω σχεδίου EWMA απαιτείται ο καθορισμός της τιμής εκκίνησης  $Z_0$  η οποία συνήθως λαμβάνεται ίση με  $\mu_0$  ( $Z_0 = \mu_0$ ), ή γενικότερα λαμβάνεται ίση με μια τιμή στόχο (*target value*)  $T$  ( $Z_0 = T$ ). Χρησιμοποιώντας διαδοχικά τον παραπάνω τύπο παίρνουμε

$$Z_t = (1 - \lambda)^t Z_0 + \lambda \sum_{i=1}^t (1 - \lambda)^{t-i} X_i$$

Παρατηρούμε ότι η ποσότητα  $Z_t$  αποτελεί ένα σταθμισμένο μέσο των παρατηρήσεων  $Z_0, X_1, X_2, \dots, X_{t-1}, X_t$  με αντίστοιχα βάρη  $(1 - \lambda)^t, \lambda(1 - \lambda)^{t-1}, \lambda(1 - \lambda)^{t-2}, \dots, \lambda(1 - \lambda)^1, \lambda$  (τα βάρη αθροίζουν στη μονάδα). Τα βάρη  $\lambda(1 - \lambda)^{t-i}$  των ποσοτήτων  $X_i$  φθίνουν γεωμετρικά καθώς προχωρούμε από την παρατήρηση  $X_t$  προς την παρατήρηση  $X_1$ , και συνεπώς η τιμή της παραμέτρου  $\lambda$  απεικονίζει τη σπουδαιότητα που δίνουμε στις νέες (μεγάλη τιμή για το  $\lambda$ ) ή στις παλαιότερες (μικρή τιμή για το  $\lambda$ ) παρατηρήσεις. Για αυτό το λόγο, ο Roberts (1959) ονόμασε τα διαγράμματα που βασίζονται στη ποσότητα  $Z_t$  ως διαγράμματα ελέγχου γεωμετρικού κινητού μέσου (*geometric moving average control charts*). Στις μέρες μας βέβαια έχει επικρατήσει ο όρος διαγράμματα ελέγχου EWMA.

Για το μέσο και τη διακύμανση της στατιστικής συνάρτησης  $Z_t$  έχουμε ότι ( $Z_0 = \mu_0$ )

και

$$\mu_{Z_t} = \mu_0$$

και

$$\sigma_{Z_t}^2 = \text{Var}\left((1-\lambda)^t Z_0 + \lambda \sum_{i=1}^t (1-\lambda)^{t-i} X_i\right) = \lambda^2 \sigma^2 \sum_{i=1}^t [(1-\lambda)^2]^{t-i} = \sigma^2 \left(\frac{\lambda}{2-\lambda}\right) [1 - (1-\lambda)^{2t}]$$

Συνεπώς, λαμβάνοντας υπόψη τη φιλοσοφία κατασκευής ενός διαγράμματος ελέγχου Shewhart, στο διάγραμμα ελέγχου EWMA απεικονίζεται η τιμή της στατιστικής συνάρτησης  $Z_t$  και τα όρια ελέγχου θα δίνονται από το Πλαίσιο 23

<b>Πλαίσιο 23</b>	
$UCL$	$= \mu_{Z_t} + L\sigma_{Z_t} = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2t}]}$
$Center Line$	$= \mu_{Z_t} = \mu_0$
$LCL$	$= \mu_{Z_t} - L\sigma_{Z_t} = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} [1 - (1-\lambda)^{2t}]}$

Αξίζει να σημειώσουμε ότι για  $\lambda = 1$  το διάγραμμα ελέγχου EWMA ανάγεται στο αντίστοιχο διάγραμμα ελέγχου Shewhart. Επίσης παρατηρούμε ότι τα όρια ελέγχου του παραπάνω διαγράμματος είναι μεταβλητά ( $\lambda \neq 1$ ). Ωστόσο η ποσότητα  $(1-\lambda)^{2t}$  τείνει στο μηδέν καθώς το  $t$  αυξάνει, οπότε σε σχετικά σύντομο χρονικό διάστημα τα όρια ελέγχου σταθεροποιούνται και δίνονται από το Πλαίσιο 24

<b>Πλαίσιο 24</b>	
$UCL$	$= \mu_{Z_t} + L\sigma_{Z_t} = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda}}$
$LCL$	$= \mu_{Z_t} - L\sigma_{Z_t} = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda}}$

Αν και η μέθοδος των διαγραμμάτων ελέγχου EWMA αναπτύχθηκε για μεμονωμένες παρατηρήσεις μπορεί να τροποποιηθεί άμεσα έτσι ώστε να καλύψει και την περίπτωση όπου έχουμε δείγματα μεγέθους  $n > 1$ . Σε αυτή την περίπτωση η ποσότητα  $X_t$  θα πρέπει να αντικατασταθεί με την ποσότητα  $\bar{X}_t$  (δηλαδή με το μέσο

του  $t$  δείγματος) και η ποσότητα  $\sigma$  με την ποσότητα  $\sigma/\sqrt{n}$ . Συνεπώς, στο διάγραμμα ελέγχου απεικονίζεται η ακόλουθη ποσότητα

$$Z_t = (1 - \lambda)Z_{t-1} + \lambda \bar{X}_t, \quad 0 < \lambda \leq 1$$

με όρια ελέγχου

**Πλαίσιο 25**

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

$$Center\ Line = \mu_{Z_t} = \mu_0$$

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

Τα διαγράμματα ελέγχου EWMA χρησιμοποιούνται, όπως και τα διαγράμματα ελέγχου CUSUM, όταν θέλουμε να εντοπίσουμε μικρές μετατοπίσεις του μέσου της διεργασίας. Το πλεονέκτημα των διαγραμμάτων ελέγχου EWMA έναντι των CUSUM είναι ότι δεν είναι ευαίσθητα στην υπόθεση της κανονικότητας των παρατηρήσεων και για το λόγο αυτό είναι ιδανικά στην περίπτωση που έχουμε μεμονωμένες παρατηρήσεις.

Για την επίδειξη ενός διαγράμματος ελέγχου EWMA θα χρησιμοποιήσουμε τα δεδομένα του Πίνακα 1. Για  $Z_0 = \mu_0 = 10$ ,  $\sigma = 1$ ,  $\lambda = 0.1$  και  $L = 2.7$  προκύπτει ο ακόλουθος πίνακας 2 για τις τιμές  $Z_t$  που θα απεικονιστούν στο διάγραμμα:

**ΠΙΝΑΚΑΣ 2:** Δεδομένα για την επίδειξη ενός διαγράμματος ελέγχου EWMA

Παρατήρηση	$X_t$	$Z_t$	Παρατήρηση	$X_t$	$Z_t$
1	9.45	9.4500	16	9.37	9.98426
2	7.99	9.4950	17	10.62	10.0478
3	9.29	9.0355	18	10.31	10.0740
4	11.66	9.9920	19	8.52	9.9186
5	12.16	10.2530	20	10.84	10.0108
6	10.18	10.3070	21	10.90	10.0997
7	8.04	9.2167	22	9.33	10.0227
8	11.46	10.755	23	12.29	10.2495
9	9.20	9.8796	24	11.50	10.3745
10	10.34	10.232	25	10.60	10.3971
11	9.03	9.2384	26	11.08	10.4654
12	11.47	10.9785	27	10.38	10.4568

Παρατήρηση	$X_t$	$Z_t$	Παρατήρηση	$X_t$	$Z_t$
13	10.51	10.1216	28	11.62	10.5731
14	9.40	10.0495	29	11.31	10.6468
15	10.08	10.0525	30	10.52	10.6341

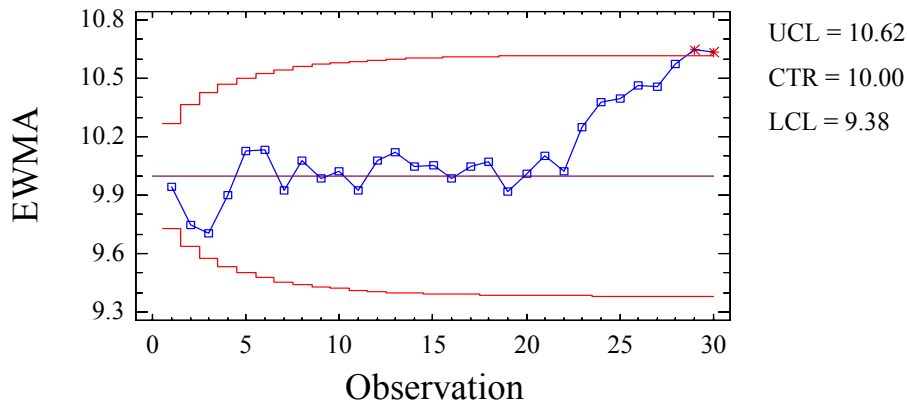
Τα σταθεροποιημένα όρια ελέγχου ( $t \rightarrow \infty$ ) είναι ίσα με

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda}} = 10 + 2.7 \cdot 1 \cdot \sqrt{\frac{0.1}{(2-0.1)}} = 10.62$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda}} = 10 - 2.7 \cdot 1 \cdot \sqrt{\frac{0.1}{(2-0.1)}} = 9.38$$

και το διάγραμμα ελέγχου EWMA είναι το ακόλουθο

EWMA Chart for X



**Διάγραμμα 3:** Διάγραμμα ελέγχου EWMA για τα δεδομένα του Πίνακα 2

Στο συγκεκριμένο διάγραμμα παρατηρούμε ότι μετά την 23<sup>η</sup> παρατήρηση διαφαίνεται μια μετατόπιση του μέσου της διεργασίας σε υψηλότερο επίπεδο η οποία γίνεται αντιληπτή στην 29<sup>η</sup> παρατήρηση.

Βασικό μειονέκτημα των διαγραμμάτων ελέγχου EWMA είναι ότι δεν είναι εύκολη η ερμηνεία τους, σε σχέση μάλιστα με τα διαγράμματα ελέγχου τύπου *Shewhart*. Όταν οι τιμές EWMA βρίσκονται προς τη μία μεριά της κεντρικής τιμής και η μετατόπιση του μέσου πραγματοποιείται από την άλλη μεριά, τότε υπάρχει καθυστέρηση στον εντοπισμό της μετατόπισης.

Παρ' όλα τα μειονεκτήματα των διαγραμμάτων ελέγχου EWMA, η εφαρμογή τους, ιδιαίτερα την τελευταία δεκαετία, είναι ευρέως διαδεδομένη, γεγονός που τα καθιστά ένα από τα πλέον χρήσιμα εργαλεία του Στατιστικού Ποιοτικού Ελέγχου.

Ο σχεδιασμός ενός διαγράμματος ελέγχου EWMA, δηλαδή, η βέλτιστη επιλογή των παραμέτρων  $\lambda$  και  $L$ , εξαρτάται σχεδόν αποκλειστικά στο μέσο μήκος ροής  $ARL$  του διαγράμματος. Γενικά, η τιμή του  $ARL$  πρέπει να είναι μεγάλη όταν η διεργασία λειτουργεί κοντά στην τιμή στόχο  $T$ , ενώ πρέπει να είναι μικρή όταν συμβαίνει το αντίθετο.

Ακολούθως δίνουμε ένα ενδεικτικό πίνακα με τιμές για το μέσο μήκος ροής σε διαγράμματα ελέγχου EWMA όπου η μετατόπιση του μέσου είναι εκφρασμένη σε μονάδες τυπικής απόκλισης ( $\delta$ ) για διάφορες τιμές των παραμέτρων  $\lambda$  και  $L$ . Ο πίνακας έχει κατασκευαστεί με τέτοιο τρόπο ώστε το εντός ελέγχου μέσο μήκος ροής  $ARL_0$  να είναι ίσο με 500 (δείτε Montgomery (2005)).

**ΠΙΝΑΚΑΣ 3:** Μέσο μήκος ροής για διαγράμματα ελέγχου EWMA

$\delta$	$L=3.054$ $\lambda=0.40$	$L=2.998$ $\lambda=0.25$	$L=2.962$ $\lambda=0.20$	$L=2.814$ $\lambda=0.10$	$L=2.615$ $\lambda=0.05$
<b>0.00</b>	500	500	500	500	500
<b>0.25</b>	224	170	150	106	84.1
<b>0.50</b>	71.2	48.2	41.8	31.3	28.8
<b>0.75</b>	28.4	20.1	18.2	15.9	16.4
<b>1.00</b>	14.3	11.1	10.5	10.3	11.4
<b>1.50</b>	5.9	5.5	5.5	6.1	7.1
<b>2.00</b>	3.5	3.6	3.7	4.4	5.2
<b>2.50</b>	2.5	2.7	2.9	3.4	4.2
<b>3.00</b>	2.0	2.3	2.4	2.9	3.5
<b>4.00</b>	1.4	1.7	1.9	2.2	2.7

Για την επιλογή των  $\lambda$ ,  $L$  καθορίζουμε πρώτα το εντός ελέγχου μέσο μήκος ροής και τη μετατόπιση του μέσου της διεργασίας που θέλουμε να ανιχνεύσουμε και κατόπιν επιλέγουμε τα  $\lambda$  και  $L$  από διαθέσιμους πίνακες με κριτήριο το μικρότερο εκτός ελέγχου μέσο μήκος ροής.

Στην πράξη επιλέγεται το  $\lambda$  έτσι ώστε  $0.05 \leq \lambda \leq 0.25$ , με πιο δημοφιλείς επιλογές τις  $\lambda = 0.05$ ,  $\lambda = 0.10$  και  $\lambda = 0.20$ , ενώ για το  $L$  έχουμε συνήθως ότι  $L = 3$ . Θα πρέπει να σημειώσουμε ότι τα διαγράμματα ελέγχου EWMA είναι ανώτερα των διαγραμμάτων CUSUM για ανίχνευση μεγάλων μετατοπίσεων του μέσου, ιδιαίτερα για  $\lambda > 0.10$ .

Για τον υπολογισμό του μέσου μήκους ροής *ARL* στα διαγράμματα ελέγχου για τα *EWMA* έχουν προταθεί δύο μεθοδολογίες (α) η μέθοδος των ολοκληρωτικών εξισώσεων και (β) και η μέθοδος των Μαρκοβιανών αλυσίδων. Οι δύο αυτές μεθοδολογίες έχουν προταθεί και για τον υπολογισμό του μέσου μήκους ροής *ARL* στα διαγράμματα ελέγχου *EWMA*.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΡΑΙΑΣ



## ΚΕΦΑΛΑΙΟ 2<sup>ο</sup>

### Πολυμεταβλητός Στατιστικός έλεγχος Ποιότητας

#### 2.1 Πολυμεταβλητά Διαγράμματα Ελέγχου Shewhart

Ο πρώτος που ασχολήθηκε με τον Πολυμεταβλητό Στατιστικό Έλεγχο Ποιότητας (*Multivariate Statistical Process Control, MSPC*) ήταν ο *Harold Hotelling* (1947), ο οποίος εφάρμοσε τη διαδικασία σε δεδομένα που συνέλλεξε από τον δεύτερο Παγκόσμιο Πόλεμο σχετικά με βομβαρδισμένες περιοχές. Ακολούθησαν μελέτες από άλλους συγγραφείς συμπεριλαμβανομένων των *Hicks* (1955), *Jackson* (1956, 1959, 1985), *Crosier* (1988), *Hawkins* (1991, 1993), *Lowry – Montgomery* (1995), *Pignatiello – Runger* (1990), *Tracy- Yang – Mason* (1992), *Montgomery – Wadsworth* (1972) και ο *Alt* (1985).

Σε πολλές περιπτώσεις η παρακολούθηση ενός μόνο χαρακτηριστικού κρίνεται αναξιόπιστη αφού δεν λαμβάνει υπόψη τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών. Η παρακολούθηση κάθε ενός διαγράμματος ελέγχου ανεξάρτητα, αναμφίβολα δεν μας αξιώνει για το κατά πόσο οι μεταβλητές που διαχειριζόμαστε συσχετίζονται. Αντίστοιχα μπορούμε να πούμε ότι κάθε μία μεταβλητή, ο καταναλωτής δεν την εκλαμβάνει σαν διαφορετική αλλά όλες σαν μία κατά την επιλογή της αγοράς ενός ποιοτικού προϊόντος. Σε καμία περίπτωση οι μεταβλητές δεν είναι ασυσχέτιστες και αυτό οδήγησε τους *MacGregor* και *Kourti* να μιλάνε με αμφιβολία για την ικανότητα των μονοδιάστατων διαγραμμάτων ελέγχου όσον αφορά τον προσδιορισμό της ποιότητας ενός παραγόμενου προϊόντος.

Υπάρχουν πολλές περιπτώσεις στις οποίες η ταυτόχρονη παρακολούθηση δύο ή περισσότερων συσχετισμένων ποιοτικών χαρακτηριστικών κρίνεται αναγκαία. Καθώς όμως αυξάνουμε τον αριθμό των ποιοτικών μετρήσιμων μεταβλητών μας, αυξάνεται και η δυσκολία να τις παρακολουθούμε. Γενικά, αν έχουμε  $p$  στατιστικά ανεξάρτητες ποιοτικές μεταβλητές για την παραγωγή ενός ξεχωριστού προϊόντος και το  $\bar{X}$  διάγραμμα έχει  $P\{\text{σφάλμα τύπου I}\} = \alpha$ , τότε η πραγματική πιθανότητα του σφάλματος  $I$  που ενσωματώνει όλη την παραγωγική διαδικασία είναι η  $\alpha' = 1 - (1 - \alpha)^p$  και η πιθανότητα ότι όλες οι  $p$  μεταβλητές θα σχεδιαστούν

ταυτόχρονα μέσα στα όρια ελέγχου όταν η διαδικασία είναι εντός ελέγχου είναι  $P\{\text{όλοι οι } p \text{ μέσοι είναι εντός ελέγχου}\} = (1 - \alpha)^p$ . Αν οι μεταβλητές μας δεν είναι ανεξάρτητες τότε είναι πολύ δύσκολο να υπολογίσουμε την  $\alpha'$  της παραγωγικής διαδικασίας. Η παρακολούθηση συσχετισμένων μεταβλητών είναι το βασικό αντικείμενο μελέτης που επεξεργάζεται ο Πολυμεταβλητός Στατιστικός Έλεγχος Ποιότητας (MSPC) (Montgomery (2005)).

Στην παρούσα μελέτη θα ασχοληθούμε με την επέκταση των μονοδιάστατων διαγραμμάτων ελέγχου Shewhart, CUSUM και EWMA που είδαμε στο πρώτο κεφάλαιο αναπτύσσοντας την πολυδιάστατη μορφή τους. Υπάρχουν και άλλες μέθοδοι στην πολυμεταβλητή ανάλυση του ποιοτικού ελέγχου που δεν αποτελούν αντικείμενο προς ανάλυση στην μελέτη μας.

## 2.2 Η $p$ -διάστατη κανονική κατανομή

Το  $p$ -διάστατο τυχαίο διάνυσμα (δείτε Montgomery (2005))

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

ακολουθεί την  $p$ -διάστατη κανονική κατανομή με παραμέτρους  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$  (συμβολισμός  $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ), όπου

$$\boldsymbol{\mu}' = [\mu_1, \mu_2, \dots, \mu_p]$$

είναι ένα διάνυσμα πραγματικών αριθμών και

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

είναι ένας συμμετρικός θετικά ορισμένος πίνακας, αν η συνάρτηση πυκνότητάς του δίνεται από τον τύπο

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

(με  $|\Sigma|$  συμβολίζουμε την ορίζουσα του πίνακα  $\Sigma$  και με  $\Sigma^{-1}$  τον αντίστροφο του πίνακα  $\Sigma$ ). Μπορεί να δειχθεί ότι

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

Στην ειδική περίπτωση  $p = 2$  προκύπτει η δυσδιάστατη κανονική κατανομή της οποίας η συνάρτηση πυκνότητας δίνεται από τον τύπο

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

όπου

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

και  $\rho\sigma_1\sigma_2 = \sigma_{12}$ . Στην ειδική περίπτωση  $p = 1$  προκύπτει η συνήθης (μονοδιάστατη) κανονική κατανομή.

Έστω, τώρα ότι έχουμε

$$\mathbf{X}_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{ip} \end{bmatrix}, \quad 1 \leq i \leq n,$$

ένα τυχαίο δείγμα μεγέθους  $n$  από την  $p$ -διάστατη κανονική κατανομή με παραμέτρους  $\boldsymbol{\mu}$  και  $\Sigma$ . Ο δειγματικός μέσος  $\bar{\mathbf{X}}$  και ο δειγματικός πίνακας διακυμάνσεων - συνδιακυμάνσεων  $\mathbf{S}$  δίνονται από τους τύπους

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

$$\mathbf{S} = \begin{bmatrix} S_{11}^2 & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22}^2 & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_p^2 \end{bmatrix} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$$

Για τα στοιχεία του πίνακα  $\mathbf{S}$  έχουμε ότι

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

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

Είναι γνωστό ότι

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

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

## 2.3 Διαγράμματα Ελέγχου για το Διάνυσμα των Μέσων Τιμών

### 2.3.1 Μέγεθος δείγματος $n > 1$

Έστω, ότι το ποιοτικό χαρακτηριστικό  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  ακολουθεί την  $p$ -διάστατη κανονική κατανομή με γνωστές παραμέτρους  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$ , και ας υποθέσουμε ότι έχουμε στη διάθεσή μας δείγματα από το ποιοτικό χαρακτηριστικό μεγέθους  $n > 1$ . Η γενική μορφή του  $k$  δείγματος ( $k \geq 1$ ) είναι η ακόλουθη

$$k \text{ δείγμα: } \left\{ \begin{array}{c} \left[ \begin{array}{c} X_{11k} \\ X_{12k} \\ \vdots \\ X_{1pk} \end{array} \right], \left[ \begin{array}{c} X_{21k} \\ X_{22k} \\ \vdots \\ X_{2pk} \end{array} \right], \dots, \left[ \begin{array}{c} X_{n1k} \\ X_{n2k} \\ \vdots \\ X_{npk} \end{array} \right] \end{array} \right\}.$$

Για κάθε δείγμα ορίζουμε την ποσότητα (δειγματικός μέσος)

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

όπου

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

Η στατιστική συνάρτηση

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

Η στατιστική συνάρτηση  $D_k^2$  παριστάνει τη σταθμισμένη απόσταση (απόσταση *Mahalanobis*) του σημείου  $\bar{\mathbf{X}}_k$  από την τιμή  $\boldsymbol{\mu}$ . Μεγάλες τιμές της στατιστικής συνάρτησης  $D_k^2$  δηλώνουν ότι η απόσταση των σημείων  $\bar{\mathbf{X}}_k$  και  $\boldsymbol{\mu}$  είναι μεγάλη με αποτέλεσμα να υιοθετούμε την άποψη ότι η διεργασία βρίσκεται εκτός στατιστικού ελέγχου λόγω μετατόπισης του μέσου  $\boldsymbol{\mu}$  της διεργασίας. Αντίθετα,

μικρές ή μηδενικές τιμές στατιστική συνάρτησης  $D_k^2$  δηλώνουν ότι η διεργασία βρίσκεται εντός στατιστικού ελέγχου.

Συνεπώς, ένα διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\mu$  της διεργασίας μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι η στατιστική  $D_k^2$ . Τα όρια ελέγχου δίνονται στο ακόλουθο Πλαίσιο 26

Πλαίσιο 26 ( $\chi^2$ διάγραμμα ελέγχου)
Όρια Ελέγχου Φάσης II
$UCL = \chi_{p;a}^2$
$LCL = 0$

Το άνω όριο ελέγχου  $UCL$  του  $\chi^2$  διαγράμματος ελέγχου έχει υπολογιστεί έτσι ώστε να διασφαλιστεί ότι το σφάλμα τύπου I να είναι ίσο με  $a$ , δηλαδή

$$P(D_k^2 > UCL) = a.$$

Έτσι, το εντός ελέγχου μέσο μήκος ροής του διαγράμματος είναι ίσο με

$$ARL_m = \frac{1}{a}.$$

Ας υποθέσουμε τώρα ότι το διάνυσμα των μέσων τιμών της διεργασίας έχει μετατοπιστεί από την εντός ελέγχου τιμή  $\mu$  στην εκτός ελέγχου τιμή  $\mu^* = \mu + \delta$ ,  $\delta \neq 0$  (δηλαδή υποθέτουμε ότι  $E(\mathbf{X}) = \mu + \delta$ ), ενώ ο πίνακας διακυμάνσεων - συνδιακυμάνσεων  $\Sigma$  έχει παραμείνει ο ίδιος. Σε αυτή την περίπτωση η κατανομή της στατιστικής συνάρτησης  $D_k^2$  είναι η μη - κεντρική  $\chi^2$  κατανομή με  $p$  βαθμούς ελευθερίας και με παράμετρο μη - κεντρικότητας

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

Έτσι, το εκτός ελέγχου μέσο μήκος ροής υπολογίζεται από τον τύπο

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

όπου το  $\beta$  δηλώνει το σφάλμα τύπου II, δηλαδή

$$1 - \beta = 1 - P(D_k^2 < \chi_{p;a}^2 | \lambda(\mu^*)) = P(D_k^2 > \chi_{p;a}^2 | \lambda(\mu^*)).$$

Ένας τύπος υπολογισμού για το σφάλμα τύπου II είναι ο ακόλουθος

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

Στην ανάπτυξη διαγραμμάτων ελέγχου Φάσης  $I$  οι παράμετροι  $\mu$  και  $\Sigma$  του ποιοτικού χαρακτηριστικού  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  είναι άγνωστοι και πρέπει να εκτιμηθούν. Ας υποθέσουμε ότι έχουμε στη διάθεσή μας  $m$  δείγματα από το ποιοτικό χαρακτηριστικό μεγέθους  $n > 1$ . Ορίζουμε την ποσότητα (ολικός δειγματικός μέσος)

$$\bar{\bar{\mathbf{X}}}' = [\bar{\bar{X}}_1, \bar{\bar{X}}_2, \dots, \bar{\bar{X}}_p]$$

όπου

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

και τον πίνακα (αθροιστικός (pooled) δειγματικός πίνακας διακυμάνσεων – συνδιακυμάνσεων) (δείτε Montgomery (2005))

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

όπου

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

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

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

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

Οι ποσότητες  $\bar{\bar{\mathbf{X}}}$  και  $\bar{\mathbf{S}}$  αποτελούν εκτιμήσεις των παραμέτρων  $\mu$  και  $\Sigma$  του ποιοτικού χαρακτηριστικού. Αντικαθιστώντας τα άγνωστα  $\mu$  και  $\Sigma$  με τα  $\bar{\bar{\mathbf{X}}}$  και  $\bar{\mathbf{S}}$  αντίστοιχα στη στατιστική συνάρτηση  $D_k^2$  προκύπτει η ακόλουθη στατιστική συνάρτηση

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

Λαμβάνοντας υπόψη ότι

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

το διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\mu$  της διεργασίας στη Φάση *I* περιγράφεται στο ακόλουθο Πλαίσιο 27

<b>Πλαίσιο 27 (Hotelling <math>T^2</math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης <i>I</i></b>	
$UCL$	$= \frac{p(m-1)(n-1)}{mn - m - p + 1} F_{p, mn-m-p+1; \alpha}$
$LCL$	$= 0$

Για διαγράμματα ελέγχου Φάσης *II* όπου οι παράμετροι  $\mu$  και  $\Sigma$  του ποιοτικού χαρακτηριστικού  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  έχουν εκτιμηθεί στη Φάση *I* από τις ποσότητες  $\bar{\mathbf{X}}$  και  $\bar{\mathbf{S}}$  χρησιμοποιείται η στατιστική συνάρτηση

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

όπου  $\bar{\mathbf{X}}_f$  είναι ο δειγματικός μέσος ενός μελλοντικού δείγματος. Λαμβάνοντας υπόψη ότι

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

το διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\mu$  της διεργασίας στη Φάση *II* περιγράφεται στο ακόλουθο πλαίσιο 28 (δείτε Montgomery (2005))

<b>Πλαίσιο 28 (Hotelling <math>T^2</math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης <i>II</i></b>	
$UCL$	$= \frac{p(m+1)(n-1)}{mn - m - p + 1} F_{p, mn-m-p+1; \alpha}$
$LCL$	$= 0$

### 2.3.2 Μεμονωμένες Παρατηρήσεις

Έστω, ότι το ποιοτικό χαρακτηριστικό  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  ακολουθεί την  $p$ -διάστατη κανονική κατανομή με γνωστές παραμέτρους  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$ , και ας υποθέσουμε ότι έχουμε στη διάθεσή μας δείγματα από το ποιοτικό χαρακτηριστικό μεγέθους  $n = 1$ . Η γενική μορφή του  $k$  δείγματος ( $k \geq 1$ ) είναι η ακόλουθη (δείτε Montgomery (2005))

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

Η στατιστική συνάρτηση

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

Έτσι, ένα διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\boldsymbol{\mu}$  της διεργασίας είναι εκείνο που η απεικονιζόμενη ποσότητα είναι η στατιστική  $D_k^2$  και τα όρια ελέγχου του διαγράμματος δίνονται στο ακόλουθο Πλαίσιο 29

<b>Πλαίσιο 29 (<math>\chi^2</math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης II</b>	
$UCL$	$= \chi_{p;a}^2$
$LCL$	$= 0$

Στην ανάπτυξη διαγραμμάτων ελέγχου Φάσης I οι παράμετροι  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$  του ποιοτικού χαρακτηριστικού  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  είναι άγνωστοι και πρέπει να εκτιμηθούν. Ας υποθέσουμε ότι έχουμε στη διάθεσή μας  $m$  δείγματα από το ποιοτικό χαρακτηριστικό μεγέθους  $n = 1$ . Ορίζουμε την ποσότητα (δειγματικός μέσος)

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

και τον πίνακα (δειγματικός πίνακας διακυμάνσεων – συνδιακυμάνσεων)



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

Οι ποσότητες  $\bar{\mathbf{X}}$  και  $\mathbf{S}$  αποτελούν εκτιμήσεις των παραμέτρων  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$  του ποιοτικού χαρακτηριστικού. Αντικαθιστώντας τα άγνωστα  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$  με τα  $\bar{\mathbf{X}}$  και  $\mathbf{S}$  αντίστοιχα στη στατιστική συνάρτηση  $D_k^2$  προκύπτει η ακόλουθη στατιστική συνάρτηση

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

Λαμβάνοντας υπόψη ότι

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

το διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\boldsymbol{\mu}$  της διεργασίας στη Φάση *I* περιγράφεται στο ακόλουθο Πλαίσιο 30

<b>Πλαίσιο 30 (Hotelling <math>T^2</math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης <i>I</i></b>	
$UCL$	$= \frac{(m-1)^2}{m} B_{p/2, (m-p-1)/2; \alpha}$
$LCL$	$= 0$

Για διαγράμματα ελέγχου Φάσης *II* όπου οι παράμετροι  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$  του ποιοτικού χαρακτηριστικού  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  έχουν εκτιμηθεί στη Φάση *I* από τις ποσότητες  $\bar{\mathbf{X}}$  και  $\mathbf{S}$  χρησιμοποιείται η στατιστική συνάρτηση

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

όπου  $\mathbf{X}_f$  είναι ένα μελλοντικό δείγμα. Λαμβάνοντας υπόψη ότι:

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

το διάγραμμα ελέγχου για την παρακολούθηση του μέσου  $\boldsymbol{\mu}$  της διεργασίας στη Φάση *II* περιγράφεται στο ακόλουθο Πλαίσιο 31

<b>Πλαίσιο 31 (Hotelling <math>T^2</math> διάγραμμα ελέγχου)</b>
<b>Όρια Ελέγχου Φάσης II</b>
$UCL = \frac{p(m+1)(n-1)}{m(m-p)} F_{p, m-p; \alpha}$
$LCL = 0$

## 2.4 Διαγράμματα ελέγχου για τη Διασπορά

Έστω, ότι το ποιοτικό χαρακτηριστικό  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$  ακολουθεί την  $p$ -διάστατη κανονική κατανομή με γνωστές παραμέτρους  $\boldsymbol{\mu}$  και  $\boldsymbol{\Sigma}$ , και ας υποθέσουμε ότι έχουμε στη διάθεσή μας δείγματα από το ποιοτικό χαρακτηριστικό μεγέθους  $n > 1$ .

Για τον έλεγχο της υπόθεσης

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0 \text{ έναντι } H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$$

ο Alt (1985) πρότεινε τη στατιστική συνάρτηση

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

όπου  $\mathbf{A}_k = (n-1)\mathbf{S}_k$ ,  $\mathbf{S}_k$  είναι ο δειγματικός πίνακας διασποράς – συνδιασποράς του  $k$  δείγματος, και  $tr(\cdot)$  συμβολίζει το ίχνος ενός πίνακα. Στο διάγραμμα ελέγχου για την παρακολούθηση της διασποράς της διεργασίας (ο μέσος  $\boldsymbol{\mu}$  της διεργασίας παραμένει σταθερός) η απεικονιζόμενη ποσότητα είναι η στατιστική συνάρτηση  $W_k$ .

Τα όρια ελέγχου του διαγράμματος δίνονται στο ακόλουθο Πλαίσιο 32

<b>Πλαίσιο 32 (<math>W</math> διάγραμμα ελέγχου)</b>
<b>Όρια Ελέγχου Φάσης II</b>
$UCL = \chi_{p(p+1)/2; \alpha}^2$
$LCL = 0$

Μια άλλη προσέγγιση για την παρακολούθηση της διασποράς μιας διεργασίας βασίζεται στη γενικευμένη δειγματική διακύμανση (*sample generalized variance*) που ορίζεται ως  $|\mathbf{S}|$ , όπου  $\mathbf{S}$  ο  $p \times p$  δειγματικός πίνακας διασποράς συνδιασποράς.

Πιο συγκεκριμένα, στην περίπτωση που το ποιοτικό χαρακτηριστικό είναι της μορφής  $\mathbf{X}' = [X_1, X_2]$  τότε, σύμφωνα με τον Alt (1985), η στατιστική συνάρτηση

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

οπότε

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

Έτσι, ένα διάγραμμα ελέγχου για την παρακολούθηση της διασποράς της διεργασίας είναι εκείνο που η απεικονιζόμενη ποσότητα είναι η στατιστική συνάρτηση  $|\mathbf{S}_k|$  και τα όρια ελέγχου του διαγράμματος δίνονται στο ακόλουθο Πλαίσιο 33

<b>Πλαίσιο 33 (<math> \mathbf{S} </math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης II</b>	
$UCL$	$= \frac{ \boldsymbol{\Sigma}_0  (\chi_{2n-4;a/2}^2)^2}{4(n-1)^2}$
$LCL$	$= \frac{ \boldsymbol{\Sigma}_0  (\chi_{2n-4;1-(a/2)}^2)^2}{4(n-1)^2}$

Ένα διαφορετικό  $|\mathbf{S}|$  διάγραμμα ελέγχου για την παρακολούθηση της διασποράς της διεργασίας βασίζεται στην παρατήρηση ότι στο διάστημα  $E[|\mathbf{S}|] \pm 3\sqrt{V[|\mathbf{S}|]}$  κατανέμεται όλη η πιθανότητα της στατιστικής συνάρτησης  $|\mathbf{S}|$  όπου

$$E[|S|] = b_1 |\Sigma_0|$$

$$V[|S|] = b_2 |\Sigma_0|^2$$

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

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

Τα όρια ελέγχου του διαγράμματος παρουσιάζονται στο Πλαίσιο 33

<b>Πλαίσιο 33 ( S  διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης II</b>	
$UCL$	$=  \Sigma_0  (b_1 + 3b_2^{1/2})$
$CL$	$=  \Sigma_0  b_1$
$LCL$	$=  \Sigma_0  (b_1 - 3b_2^{1/2})$

Σε περίπτωση που το κάτω όριο βγει αρνητικό τότε θέτουμε  $LCL = 0$ .

Μια άλλη προσέγγιση στηρίζεται στην τετραγωνική ρίζα του πίνακα δειγματικών διασπορών - συνδιασπορών  $|S|^{1/2}$  με δύο εκδοχές.

Σύμφωνα με την πρώτη εκδοχή, για δύο μεταβλητές ο *Alt* και ο *Smith* (1988) έδωσαν τα ακόλουθα όρια ελέγχου στο Πλαίσιο 34

<b>Πλαίσιο 34 (<math> S ^{1/2}</math> διάγραμμα ελέγχου)</b>	
<b>Όρια Ελέγχου Φάσης II</b>	
$UCL$	$= \frac{ \Sigma_0 ^{1/2} X_{2n-4, 1-\alpha/2}^2}{2(n-1)}$
$LCL$	$= \frac{ \Sigma_0 ^{1/2} X_{2n-4, \alpha/2}^2}{2(n-1)}$

Για περισσότερα χαρακτηριστικά (μεταβλητές) υπάρχουν άλλες επιλογές (*Anderson* (1985)).

Η δεύτερη εκδοχή πάλι από τον *Alt* και τον *Smith* (1988) βασίζεται στις σχέσεις

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

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

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

Τα όρια ελέγχου παρουσιάζονται στο Πλαίσιο 35

Πλαίσιο 35 ( $ S ^{1/2}$ διάγραμμα ελέγχου)
<b>Όρια Ελέγχου Φάσης II</b>
$UCL =  \Sigma_0 ^{1/2} (b_3 + 3\sqrt{b_1 - b_3^2})$
$CL =  \Sigma_0 ^{1/2} b_3$
$LCL =  \Sigma_0 ^{1/2} (b_3 - 3\sqrt{b_1 - b_3^2})$

Σε περίπτωση που το κάτω όριο βγει αρνητικό θέτουμε  $LCL = 0$ .

Για διαγράμματα ελέγχου Φάσης I, όπου ο πίνακας  $|\Sigma_0|$  είναι άγνωστος, ο *Alt* (1985) χρησιμοποίησε ως αμερόληπτο εκτιμητή του την ποσότητα  $|\bar{S}| + b_1$  όπου  $|\bar{S}| = \frac{1}{m} \sum |S_i|$  και  $S_i$  είναι ο δειγματικός πίνακας διασποράς - συνδιασποράς του  $i$  δείγματος όπου  $i = 1, 2, \dots, m$ . Τα νέα όρια προκύπτουν στο Πλαίσιο 36

**Πλαίσιο 36 ( $|\bar{S}|$  διάγραμμα ελέγχου)****Όρια Ελέγχου Φάσης I**

$$UCL = (|\bar{S}| + b_1)(b_1 + 3b_2^{1/2})$$

$$CL = (|\bar{S}| + b_1)b_1$$

$$LCL = (|\bar{S}| + b_1)(b_1 - 3b_2^{1/2})$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΡΑΙΑ

## ΚΕΦΑΛΑΙΟ 3<sup>ο</sup>

### Πολυμεταβλητά Διαγράμματα CUSUM

#### 3.1 MCUSUM με ακολουθιακούς ελέγχους λόγου πιθανοφάνειας

Ο Healy (1987) εξέτασε τα CUSUM διαγράμματα χρησιμοποιώντας ακολουθιακούς ελέγχους λόγου πιθανοφάνειας (*sequence of sequential probability ratio tests, SPRT*) και τα εφάρμοσε στην πολυμεταβλητή κανονική κατανομή ώστε να προκύψουν τα MCUSUM διαγράμματα.

Ας υποθέσουμε ότι παρακολουθούμε μία ακολουθία από τυχαίες μεταβλητές  $\mathbf{X}_1, \mathbf{X}_2, \dots$  που ακολουθούν  $p$ -διάστατη κανονική κατανομή. Έστω ότι οι πρώτες  $m-1$  παρατηρήσεις  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{m-1}$  έχουν εντός ελέγχου κατανομή  $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$  και οι επόμενες  $\mathbf{X}_m, \mathbf{X}_{m+1}, \dots$  έχουν εκτός ελέγχου κατανομή  $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ . Το πρόβλημα είναι να εντοπιστεί το σημείο  $m$  όπου χρονικά αλλάζει θέση η κατανομή των παρατηρήσεων. Σύμφωνα με τη μεθοδολογία των ακολουθιακών ελέγχων λόγου πιθανοφάνειας έχουμε

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

όπου  $f$  είναι η πυκνότητα πιθανότητας των πολυδιάστατων κανονικών κατανομών. Λογαριθμίζοντας έχουμε

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

Έτσι, το σχήμα CUSUM σύμφωνα με τον *Healy* (1987) είναι το ακόλουθο όπου τα σημεία του διαγράμματος ελέγχου προέρχονται από την στατιστική συνάρτηση

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

όπου

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

και

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

Η μεταβλητή  $\mathbf{a}' \mathbf{X}_t$  ακολουθεί μονοδιάστατη κανονική κατανομή με διασπορά 1.

Η στατιστική συνάρτηση  $S_t$  μπορεί να γραφτεί στη μορφή

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

όπου  $\lambda^2(\boldsymbol{\mu}_1) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$ .

Το MCUSUM θα δώσει **εκτός ελέγχου** σήμα όταν  $S_t \geq H$ , όπου το  $H$  υπολογίζεται κατά τον *Pignatiello* και *Rugner* με προσομοίωση.

Στην περίπτωση που το εκτός ελέγχου διάνυσμα  $\boldsymbol{\mu}_1$  δεν μπορεί να προσδιοριστεί τότε συνίσταται η μεθοδολογία του *Crosier* (1988).

Η παραπάνω προσέγγιση εξετάζει το διάνυσμα των μέσων μιας πολυμεταβλητής κανονικής κατανομής. Παρακάτω θα περιγράψουμε την περίπτωση που χρειάζεται παρακολούθηση ο πίνακας της διασποράς – συνδιασποράς μιας πολυμεταβλητής κανονικής κατανομής.



Έστω, ότι  $\mathbf{X}_t$  είναι μία παρατήρηση στην χρονική στιγμή  $t$ , που προέρχεται από την  $p$ -διάστατη κανονική κατανομή με εντός ελέγχου πίνακα διασποράς - συνδιασποράς  $\Sigma_0$  και εκτός ελέγχου πίνακα  $\Sigma_1 = C \Sigma_0$  και κοινό (εντός ή εκτός ελέγχου) διάνυσμα μέσων  $\mu_0$ . Τότε

$$f_1(\mathbf{X}_t) / f_0(\mathbf{X}_t) = \frac{(2\pi)^{-p/2} |C\Sigma_0|^{-1/2} \exp\left(-0.5(\mathbf{X}_t - \mu_0)' (C\Sigma_0)^{-1} (\mathbf{X}_t - \mu_0)\right)}{(2\pi)^{-p/2} |\Sigma_0|^{-1/2} \exp\left(-0.5(\mathbf{X}_t - \mu_0)' \Sigma_0^{-1} (\mathbf{X}_t - \mu_0)\right)}$$

και λογαριθμίζοντας παίρνουμε

$$\log(f_1(\mathbf{X}_t) / f_0(\mathbf{X}_t)) = -\frac{p}{2} \log C + 0.5(\mathbf{X}_t - \mu_0)' \Sigma_0^{-1} (\mathbf{X}_t - \mu_0) \left(1 - \frac{1}{C}\right).$$

Τελικά προκύπτει η στατιστική συνάρτηση

$$S_t = \max\left[\left(S_{t-1} + (\mathbf{X}_t - \mu_0)' \Sigma_0^{-1} (\mathbf{X}_t - \mu_0) - K\right), 0\right]$$

όπου

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

Αν θέσουμε

$$T_t^2 = (\mathbf{X}_t - \mu_0)' \Sigma_0^{-1} (\mathbf{X}_t - \mu_0)$$

τότε προκύπτει η στατιστική συνάρτηση

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

Το MCUSUM διάγραμμα ελέγχου θα μας δώσει σήμα για εκτός ελέγχου διεργασία όταν  $S_t \geq H$ .

### 3.2 Το σχέδιο του Crosier

Ο *Crosier* (1988) πρότεινε δύο νέα πολυμεταβλητά σχέδια MCUSUM. Το πρώτο στηρίζεται στην τετραγωνική ρίζα της στατιστικής συνάρτησης του *Hotelling-T<sup>2</sup>* ενώ το δεύτερο σχέδιο προκύπτει με κατάλληλη τροποποίηση των μονοδιάστατων CUSUM διαγραμμάτων.

Έστω, ότι η  $\mathbf{X}_t$  είναι μία παρατήρηση την χρονική στιγμή  $t$  που προέρχεται από την  $p$ -διάστατη πολυμεταβλητή κανονική κατανομή με γνωστό πίνακα διασποράς - συνδιασποράς  $\Sigma_0$  και  $\boldsymbol{\mu}_0$ .

Η πρώτη προσέγγιση του *Crosier* χρησιμοποιεί τη στατιστική συνάρτηση

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

Αν θέσουμε

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

τότε

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

όπου  $S_0 \geq 0$  και  $K \geq 0$ .

Η παραπάνω διαδικασία θα μας δώσει σήμα για εκτός ελέγχου διεργασία όταν  $S_t \geq H$ .

Το **δεύτερο** σχέδιο που προτάθηκε από τον *Crosier* έχει καλύτερο *ARL*. Χρησιμοποιεί την στατιστική συνάρτηση

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

όπου

$$\mathbf{S}_t = \mathbf{0}, \text{ αν } C_t \leq K$$

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

και

$$C_t = \sqrt{(\mathbf{S}_{t-1} + \mathbf{X}_t - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\mathbf{S}_{t-1} + \mathbf{X}_t - \boldsymbol{\mu}_0)}$$

Σήμα εκτός ελέγχου παίρνουμε όταν  $Y_t > H$ . Η τιμή του  $H$  επιλέγεται έτσι ώστε να έχουμε προκαθορισμένο εντός ελέγχου *ARL* με την βοήθεια προσομοίωσης. Ο *Crosier* (1988) απέδειξε ότι η κατανομή της  $Y_t$  και το *ARL* εξαρτώνται από το διάνυσμα των μέσων και τον πίνακα διασποράς – συνδιασποράς μόνο μέσω της παραμέτρου της μη-κεντρικότητας. Επίσης, επισήμανε ότι  $K = \lambda(\boldsymbol{\mu}_1)/2$  όπου  $\lambda(\boldsymbol{\mu}_1)$  είναι η τετραγωνική ρίζα της παραμέτρου μη - κεντρικότητας που ελαχιστοποιεί το εκτός ελέγχου *ARL*.

Τα πλεονεκτήματα από την χρήση των MCUSUM συγκριτικά με τα πολυμεταβλητά διαγράμματα Shewhart είναι ότι δίνουν σήμα για εκτός ελέγχου διεργασία για μικρές μετατοπίσεις του μέσου γρηγορότερα.

### 3.3 Το σχέδιο των J. Pignatiello και του G. Runger's

Οι Pignatiello και Runger (1990) εισήγαγαν δύο νέα πολυμεταβλητά CUSUM σχέδια, το MCUSUM #1 (MC#1) και το MCUSUM #2 (MC#2).

Το MC#1 σχέδιο είναι το σχέδιο το οποίο έχει το καλύτερο ARL και έχει ως εξής:

Έστω,  $\mathbf{X}_t$  μία παρατήρηση τη χρονική στιγμή  $t$ , που προέρχεται από την  $p$ -διάστατη κανονική κατανομή με γνωστό πίνακα διασποράς – συνδιασποράς  $\Sigma_0$  και ένα γνωστό μέσο  $\mu_0$ .

Έστω,

$$\mathbf{S}_t = \sum_{i=t-n_t+1}^t (\mathbf{X}_i - \mu_0)$$

όπου

$$n_t = \begin{cases} n_{t-1} + 1, & \text{αν } MC\#1_{t-1} > 0 \\ 1, & \text{διαφορετικά} \end{cases}$$

Ένα πολυμεταβλητό διάγραμμα ελέγχου μπορεί να κατασκευαστεί ορίζοντας

$$C_t = \sqrt{\mathbf{S}_t' \Sigma_0^{-1} \mathbf{S}_t}$$

και

$$MC\#1_t = \max \{C_t - kn_t, 0\}$$

όπου το  $MC\#1_0 = 0$  και το  $k$  επιλέγονται να είναι η μισή απόσταση μεταξύ  $\mu_1$  και  $\mu_0$ ,  $0.5\lambda^2(\mu_1)$ , όπου  $\mu_1$  είναι ένας εκτός ελέγχου μέσος.

Το MC#1 διάγραμμα λειτουργεί σχεδιάζοντας τα σημεία  $MC\#1_t$  με **άνω όριο** ελέγχου  $H$ . Όταν ένα σημείο  $MC\#1_t$  υπερβεί το  $H$  τότε η διεργασία μας είναι εκτός ελέγχου.

Επίσης, ο *Pignatiello* και ο *Runger* (1990) απέδειξαν ότι το *ARL* εξαρτάται από την τετραγωνική ρίζα της παραμέτρου μη-κεντρικότητας και η ποσότητα  $H$  υπολογίζεται με βάση τη μεθοδολογία των αλυσίδων του *Markov*. Επίσης, κατέληξαν στο συμπέρασμα ότι όλα τα πολυμεταβλητά διαγράμματα CUSUM συμπεριφέρονται καλύτερα από τα αντίστοιχα των πολυμεταβλητών *Shewhart* με εξαίρεση την περίπτωση που έχουμε μεγάλες τιμές για την παράμετρο της μη-κεντρικότητας.

Το **MC#2** σχέδιο είναι η δεύτερη προσέγγιση των *Pignatiello* και *Runger* (1990). Έστω,

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

Ένας μονόπλευρος έλεγχος προκύπτει από την σχέση

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

όπου  $MC\#2_0 = 0$  και το  $k$  να επιλέγεται να είναι η μισή απόσταση μεταξύ  $\boldsymbol{\mu}_1$  και  $\boldsymbol{\mu}_0 + p$  δηλαδή

$$k = p + 0.5 \lambda^2(\boldsymbol{\mu}_1)$$

Το **MC#2** διάγραμμα λειτουργεί σχεδιάζοντας τα σημεία  $MC\#2_t$  με **άνω όριο** ελέγχου  $H$ . Όταν ένα σημείο  $MC\#2_t$  υπερβεί το  $H$  τότε η διεργασία μας είναι εκτός ελέγχου. Η ποσότητα  $H$  υπολογίζεται από την μεθοδολογία των αλυσίδων *Markov* με προσομοίωση.

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup>

### Πολυμεταβλητά Διαγράμματα EWMA

#### 4.1 Διαγράμματα Ελέγχου MEWMA

Τα πολυμεταβλητά διαγράμματα EWMA που παρουσιάζονται σε αυτή την ενότητα έχουν προταθεί από τον *Lowry et al.* (1992). Έστω ότι  $\mathbf{X}_t$  ακολουθεί την  $p$ -διάστατη κανονική κατανομή με γνωστό πίνακα διασποράς - συνδιασποράς  $\Sigma_0$  και ένα γνωστό  $\mu_0$ .

Το πολυμεταβλητό EWMA διάγραμμα ελέγχου χρησιμοποιεί τη στατιστική

$$\mathbf{z}_t = \mathbf{R}(\mathbf{X}_t - \mu_0) + (\mathbf{I} - \mathbf{R})z_{t-1}$$

ή ισοδύναμα

$$\mathbf{z}_t = \sum_{j=1}^t \mathbf{R}(\mathbf{I} - \mathbf{R})^{t-j} (\mathbf{X}_j - \mu_0)$$

όπου  $t = 1, 2, 3, \dots$  και  $\mathbf{R}$  είναι  $p \times p$  πίνακας που ορίζεται ως

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

όπου

$$0 \leq r_k \leq 1$$

για  $k = 1, 2, 3, \dots, p$ .

Αν δεν υπάρχει λόγος να σταθμίσουμε με διαφορετικά βάρη τα  $p$  ποιοτικά χαρακτηριστικά που παρακολουθούμε, τότε θέτουμε ότι  $r_1 = r_2 = \dots = r_p = r$ . Η αρχική

τιμή  $\mathbf{z}_0$  συνήθως είναι ίση με τον εντός ελέγχου μέσο. Είναι προφανές ότι αν  $\mathbf{R} = \mathbf{1}$  τότε το πολυμεταβλητό διάγραμμα EWMA είναι ισοδύναμο με το  $T^2$  διάγραμμα ελέγχου (ή το  $\chi^2$  διάγραμμα ελέγχου στην περίπτωση που γνωρίζουμε το διάνυσμα των μέσων και τον πίνακα διασποράς - συνδιασποράς).

Το πολυμεταβλητό EWMA διάγραμμα δίνει ένα εκτός ελέγχου σήμα αν

$$\mathbf{z}_t' \boldsymbol{\Sigma}_{z_t}^{-1} \mathbf{z}_t > h$$

όπου  $\boldsymbol{\Sigma}_{z_t}$  είναι ο πίνακας διασποράς - συνδιασποράς του  $\mathbf{z}_t$ .

Η τιμή της  $h$  καθορίζεται με προσομοίωση έτσι ώστε να επιτευχθεί ένα προκαθορισμένο εντός ελέγχου  $ARL$ .

Ο πίνακας  $\boldsymbol{\Sigma}_{z_t}$  διασποράς - συνδιασποράς του  $\mathbf{z}_t$  υπολογίζεται από την ακόλουθη σχέση

$$\boldsymbol{\Sigma}_{z_t} = \sum_{j=1}^t \left[ \mathbf{R}(\mathbf{I}-\mathbf{R})^{t-j} \boldsymbol{\Sigma}_0 (\mathbf{I}-\mathbf{R})^{t-j} \mathbf{R} \right]$$

και αν  $r_1 = r_2 = \dots = r_p = r$  τότε

$$\boldsymbol{\Sigma}_{z_t} = \frac{r \left[ 1 - (1-r)^{2t} \right]}{2-r} \boldsymbol{\Sigma}_0$$

και για  $t \rightarrow +\infty$  έχουμε ότι

$$\boldsymbol{\Sigma}_{z_t} = \frac{r}{2-r} \boldsymbol{\Sigma}_0$$

## 4.2 ARL για τα MEWMA

Το  $ARL$  στην πολυμεταβλητή ανάλυση των διαγραμμάτων EWMA εξαρτάται αποκλειστικά και μόνο από το  $\mu_0$  των  $\Sigma_0$  και από την τιμή της παραμέτρου μη κεντρικότητας. Για τον υπολογισμό του  $ARL$  έχουν προταθεί οι ακόλουθες μεθοδολογίες:

- Ολοκληρωτική Εξίσωση για το εντός ελέγχου ARL

Εδώ δίνουμε την ολοκληρωτική εξίσωση του *Rigdon* (1995a), υπό την προϋπόθεση ότι ο μέσος της εντός ελέγχου διεργασίας είναι  $\mathbf{0}$  και ο πίνακας διασποράς – συνδιασποράς είναι ο μοναδιαίος  $\mathbf{I}$ . Παρότι αυτή η προϋπόθεση δείχνει λίγο περιοριστική στην πραγματικότητα δεν είναι.

Έστω,  $\mathbf{X}_i, i=1,2,\dots$ , παρατηρήσεις που ακολουθούν την κατανομή  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Αν το πολυμεταβλητό διάγραμμα ελέγχου EWMA με παραμέτρους  $r$  και  $h$  εφαρμοστεί στις  $\mathbf{X}_i$ , τότε το εντός ελέγχου  $ARL$  είναι το ίδιο με αυτό της περίπτωσης των παρατηρήσεων με μέσο  $\mathbf{0}$  και πίνακα διασποράς – συνδιασποράς μοναδιαίο  $\mathbf{I}$ .

Το εντός ελέγχου  $ARL$  ενός πολυμεταβλητού EWMA διαγράμματος, που εφαρμόζεται σε μία διεργασία με μέσο διάνυσμα  $\mathbf{0}$  και πίνακα διασποράς – συνδιασποράς  $\mathbf{I}$ , εξαρτάται από την αρχική τιμή  $\mathbf{z}_0$  μόνο διαμέσου της ποσότητας  $\delta = \mathbf{z}'_0 \mathbf{z}_0$ .

Υποθέτουμε ότι έχουμε ένα πολυμεταβλητό διάγραμμα EWMA με παράμετρο  $r$  που εφαρμόζεται σε μία  $p$ -διάστατη διεργασία με μέσο  $\mathbf{0}$  και πίνακα διασποράς – συνδιασποράς  $\mathbf{I}$ . Έστω,  $L(\delta/h)$  το  $ARL$ , που αντιστοιχεί στις ποσότητες  $\delta$  και  $h$ . Τότε η συνάρτηση  $L$  ικανοποιεί την ολοκληρωτική εξίσωση

$$L(\delta/h) = 1 + \int_0^{hr/(2-r)} L(y/h) \times f(y | z'_0 z_0 = \delta) dy$$

όπου  $f(y | z'_0 z_0 = \delta)$  είναι η συνάρτηση πυκνότητας πιθανότητας της μη κεντρικής  $\chi^2$  κατανομής, με  $p$  βαθμούς ελευθερίας και παράμετρο μη κεντρικότητας  $\{[(1-r)/r]^2 \delta\}$ .



- **Διπλή Ολοκληρωτική Εξίσωση για το εκτός ελέγχου ARL**

Εδώ δίνουμε την διπλή ολοκληρωτική εξίσωση του *Rigdon* (1995b), για τον υπολογισμό του εκτός ελέγχου *ARL* κάτω από την υπόθεση ότι ο εντός ελέγχου μέσος είναι  $\mathbf{0}$  και ότι ο εντός ελέγχου πίνακας διασποράς - συνδιασποράς είναι ο μοναδιαίος  $\mathbf{I}$ .

Υποθέτουμε ότι η παραγωγική διαδικασία ακολουθεί την κατανομή  $N_p(\mathbf{0}, \mathbf{I})$ . Αν ο μέσος μετατοπιστεί από το  $\mathbf{0}$  στο  $\boldsymbol{\mu}$ , τότε το *ARL* εξαρτάται από την  $\mathbf{z}_0$  μόνο διαμέσου των ποσοτήτων  $\alpha = (\mathbf{z}'_0 \mathbf{z}_0)^2$  και  $\beta = \boldsymbol{\mu}' \mathbf{z}_0$ .

Υποθέτουμε ότι  $\mathbf{X}_i, i=1,2,\dots$ , είναι παρατηρήσεις από την κατανομή  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , όπου  $\boldsymbol{\mu} = \mathbf{0}$  όταν η διεργασία είναι εντός ελέγχου. Υποθέτουμε ότι  $\boldsymbol{\mu}' \boldsymbol{\mu} = \delta \neq 0$ .

Τότε για δεδομένα  $\alpha$  και  $\beta$  το *ARL*,  $L(\alpha, \beta)$ , του πολυμεταβλητού EWMA διαγράμματος με παραμέτρους  $r$  και  $h$  ικανοποιούν την ολοκληρωτική εξίσωση

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

όπου  $h(./v, r)$

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

και  $h(./v, \lambda)$  είναι η συνάρτηση πυκνότητας πιθανότητας της μη κεντρικής  $\chi^2$  κατανομής με  $\nu$  βαθμούς ελευθερίας και με παράμετρο μη κεντρικότητας  $\lambda$ . Επίσης,  $R$  είναι η περιοχή για την οποία  $v^2/\delta < u < rh/(2-r)$ .

РАСЧЕТНО ТЕРА

## Κεφάλαιο 5<sup>ο</sup>

### Εφαρμογή

Στο παρόν κεφάλαιο θα δώσουμε μία εφαρμογή η οποία αναφέρεται στην παραγωγική μιας χημικής ουσίας που η ποιότητά της εξαρτάται από τις τρεις ακόλουθες μεταβλητές:

1. Η πίεση  $Z_1$  (*PRESSURE*)
2. Η θερμοκρασία  $Z_2$  (*TEMPERATURE*)
3. Η υγρασία  $Z_3$  (*HUMIDITY*)

Τα δεδομένα δίνονται στον ακόλουθο Πίνακα

**ΠΙΝΑΚΑΣ 4** Μετρήσεις παραγόμενου προϊόντος

$Z_1$	$Z_2$	$Z_3$	$Z_1$	$Z_2$	$Z_3$	$Z_1$	$Z_2$	$Z_3$
84,099	72,140	93,880	81,161	72,860	90,746	88,915	80,288	99,758
81,167	70,147	90,913	83,039	71,780	92,681	79,717	70,380	89,629
85,733	74,563	93,823	87,273	79,328	98,366	86,652	76,202	98,061
88,829	79,997	99,077	86,400	76,093	95,380	84,651	72,157	94,261
88,595	80,465	98,252	87,624	77,921	98,909	85,125	74,536	95,819
90,199	78,859	100,486	86,787	76,888	96,094	83,038	72,542	93,683
78,449	68,708	89,189	80,884	70,972	91,230	83,351	74,335	94,286
84,297	75,071	94,147	81,653	72,307	90,753	87,548	77,279	99,380
88,285	78,517	100,680	87,082	77,291	98,439	87,409	78,181	96,966
81,740	70,725	90,858	85,968	74,245	95,899	86,369	76,466	96,281
82,929	74,023	94,433	82,180	70,972	92,026	87,074	77,791	95,978
79,929	71,070	89,978	84,277	75,571	94,163	89,892	79,474	98,789
79,459	71,130	88,973	85,395	77,307	95,663	85,912	74,699	96,928
82,067	71,685	92,720	86,673	77,008	95,946	86,767	75,182	97,847
82,679	73,695	91,508	85,416	73,292	95,599	90,559	80,894	99,219
78,646	71,039	87,750	82,267	71,927	94,397	83,993	74,131	95,590
83,296	71,874	92,803	90,655	80,674	101,125	87,560	73,829	97,688
83,788	73,823	94,454	86,462	77,584	96,864	89,890	80,818	96,434
85,405	76,280	94,504	85,217	76,305	95,105	88,950	74,409	95,048
83,904	74,024	94,093	87,490	78,599	97,726	87,900	72,863	90,783
84,019	74,386	96,131	87,586	78,315	99,103	89,000	79,962	96,345
83,889	73,569	93,807	83,090	72,360	91,499	90,000	72,068	91,096
89,028	79,787	97,276	82,230	70,942	92,282	90,500	69,478	90,999
84,744	74,330	94,217	88,334	79,716	97,959	89,800	74,996	90,899
84,442	73,511	95,843	81,396	70,116	91,540	91,000	74,797	91,003
83,460	72,130	93,127	80,323	70,273	89,401	91,400	74,025	96,818
90,917	81,048	99,447	87,134	77,199	95,858			
87,597	77,841	97,239	86,915	76,677	95,213			
92,127	82,230	101,160	91,617	81,606	101,268			
83,035	74,393	93,443	89,331	78,419	99,597			
89,984	79,668	98,927	88,912	78,459	99,934			

$Z_1$	$Z_2$	$Z_3$	$Z_1$	$Z_2$	$Z_3$
80,163	69,200	89,378	85,339	75,611	93,559
86,617	73,435	97,032	85,006	74,068	95,754
87,707	76,508	96,374	86,361	74,812	98,247
90,757	81,315	101,556	84,923	75,337	95,423
84,746	74,102	94,476	81,836	71,876	91,879
83,429	73,763	91,340	79,676	68,806	90,447
87,025	78,258	97,664	87,485	77,000	96,459
83,856	73,119	93,943	86,333	75,964	97,681
87,273	77,597	97,403	86,854	78,328	96,672
80,667	70,333	92,405	85,640	76,621	96,247
82,458	71,999	93,431	81,919	72,032	90,561
80,435	70,431	93,048	88,715	79,250	98,196
83,911	74,234	93,698	84,066	73,096	92,601
84,903	74,850	93,851	82,480	71,149	91,444
85,084	73,888	94,791	82,537	72,519	92,201
84,032	73,809	94,462	83,713	75,052	91,185
91,584	79,809	101,500	83,640	74,594	94,307
79,773	70,336	90,373	83,429	75,173	93,336
82,791	72,063	93,248	87,548	76,689	96,987
77,267	68,594	85,040	86,540	75,808	94,492
89,343	77,779	100,244	83,175	73,273	92,803

Τα περιγραφικά στατιστικά μέτρα για τα πρώτα 120 δεδομένα του παραπάνω πίνακα που έχουν προέλθει από μια εντός ελέγχου τρισδιάστατη κανονική κατανομή παρουσιάζονται στον ακόλουθο πίνακα (για τη στατιστική ανάλυση των δεδομένων χρησιμοποιήθηκε το πακέτο *Statgraphics*)

**ΠΙΝΑΚΑΣ 5** Περιγραφικά στατιστικά για τον πίνακα 4

#### Summary Statistics

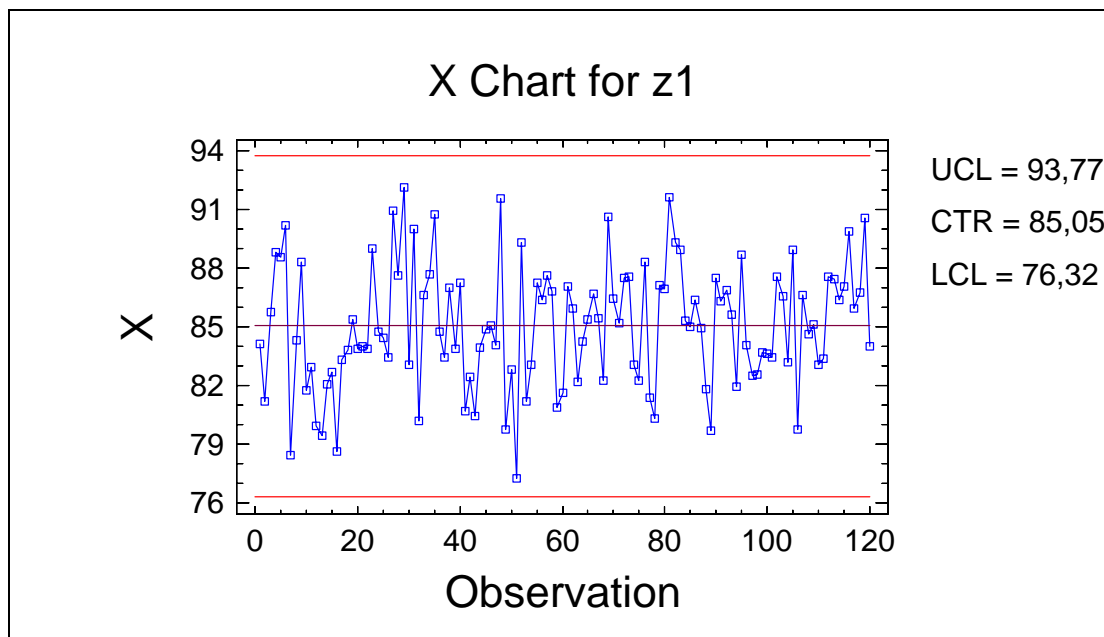
	Z1	Z2	Z3
<b>Count</b>	120	120	120
<b>Average</b>	85,046	75,0063	95,0176
<b>Standard deviation</b>	3,17156	3,2771	3,32435
<b>Coeff. of variation</b>	3,72923%	4,3691%	3,49866%
<b>Minimum</b>	77,267	68,594	85,04
<b>Maximum</b>	92,127	82,23	101,556
<b>Range</b>	14,86	13,636	16,516
<b>Std. skewness</b>	0,0089345	0,63051	-0,58632
<b>Std. kurtosis</b>	-0,998336	-1,80684	-0,641497

#### Correlations

	Z1	Z2	Z3
<b>Z1</b>		0,9483	0,9497
<b>Z2</b>	0,9483		0,8965
<b>Z3</b>	0,9497	0,8965	

Όσον αφορά τη μεταβλητή  $Z_1$  η διεργασία είναι εντός στατιστικού ελέγχου και αυτό φαίνεται στο ακόλουθο σχήμα 4. Όλα τα σημεία συμπεριφέρονται με τυχαίο τρόπο κατά μήκος του διαγράμματος ελέγχου (όρια ελέγχου  $3\sigma$ ) και βρίσκονται εντός των ορίων ελέγχου της διεργασίας μας.

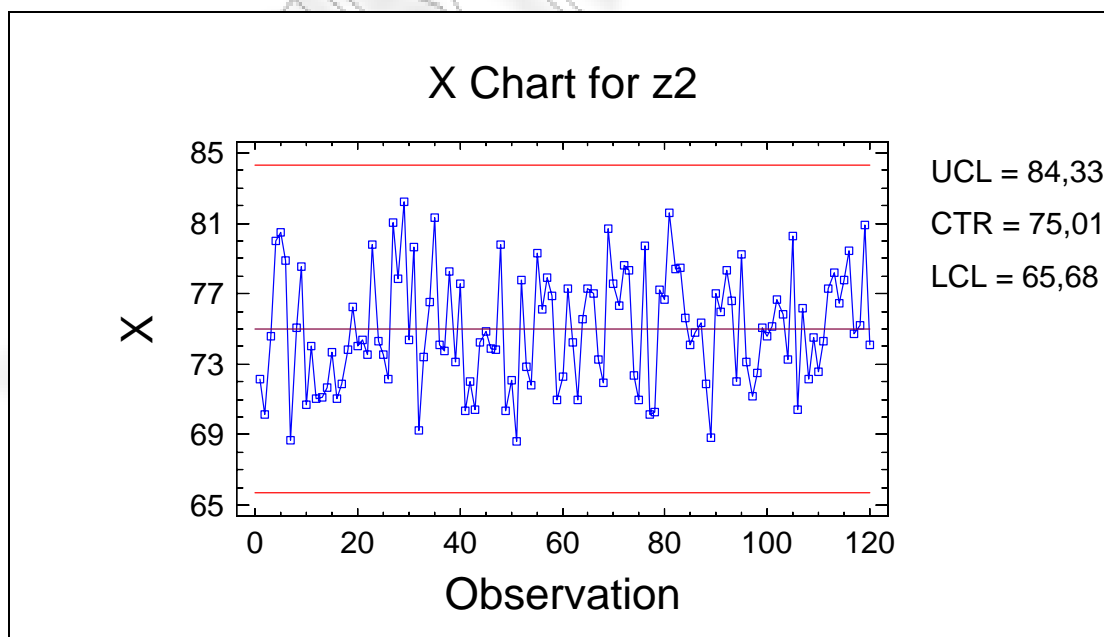
**Σχήμα 4** X διάγραμμα ελέγχου για τη  $Z_1$



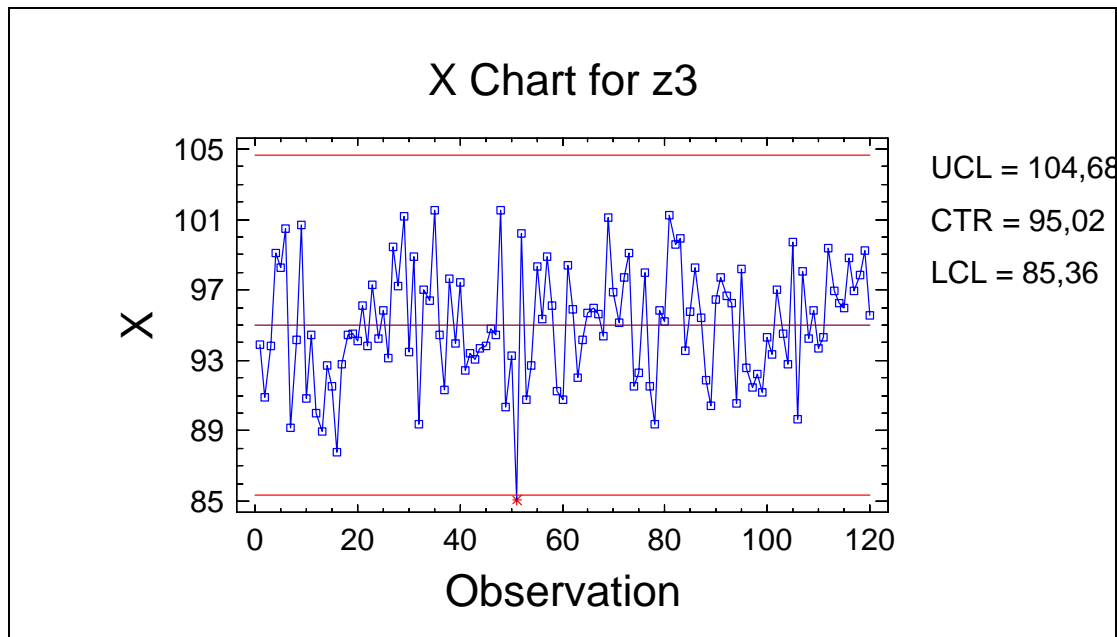
Τα όρια ελέγχου του παραπάνω διαγράμματος ελέγχου έχουν υπολογιστεί χρησιμοποιώντας το Πλαίσιο 13 της Παραγράφου 1.3.3.

Αντίστοιχα για τις μεταβλητές  $Z_2$  και  $Z_3$  έχουμε τα ακόλουθα διαγράμματα.

**Σχήμα 5** X διάγραμμα ελέγχου για την  $Z_2$

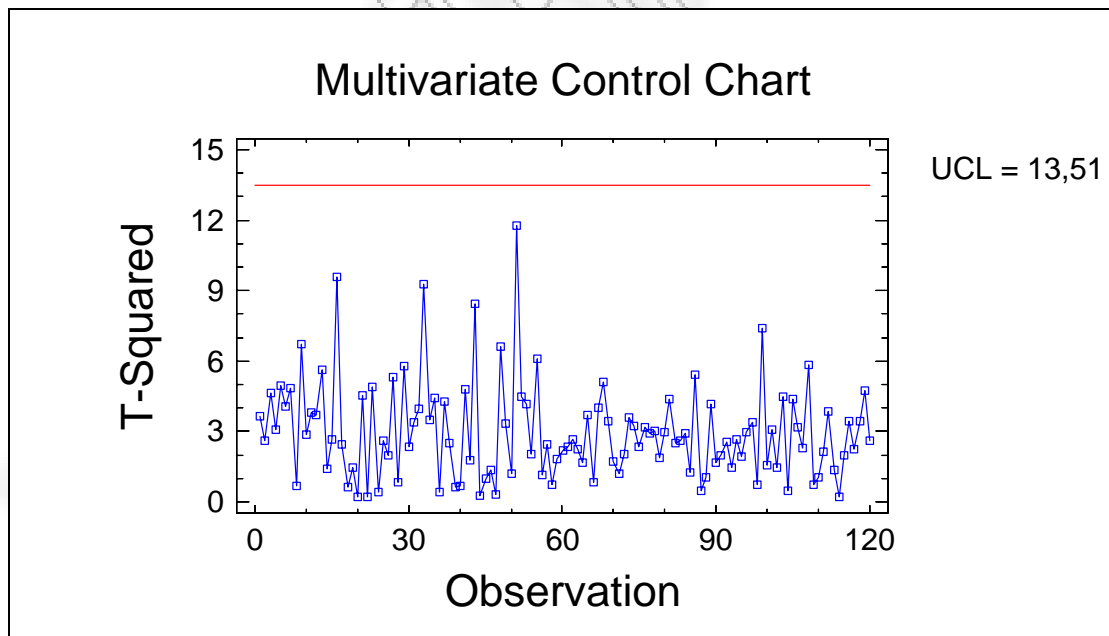


Σχήμα 6  $X$  διάγραμμα ελέγχου για την  $Z_3$



Το πολυμεταβλητό διάγραμμα ελέγχου *Shewhart* που προκύπτει για τις 120 παρατηρήσεις με  $\alpha = 0.0027$  είναι το ακόλουθο

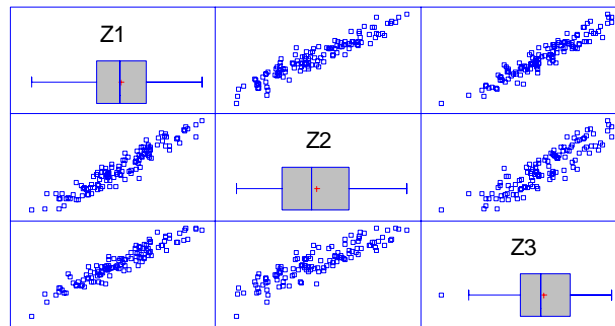
Σχήμα 7  $T^2$  διάγραμμα ελέγχου για τις  $Z_1, Z_2, Z_3$



Το άνω όριο ελέγχου του παραπάνω διαγράμματος υπολογίστηκε με χρήση του Πλαισίου 30 της Παραγράφου 2.3.2. Όλα τα σημεία του διαγράμματος ελέγχου βρίσκονται εντός των ορίων ελέγχου κάτι ήταν αναμενόμενο.

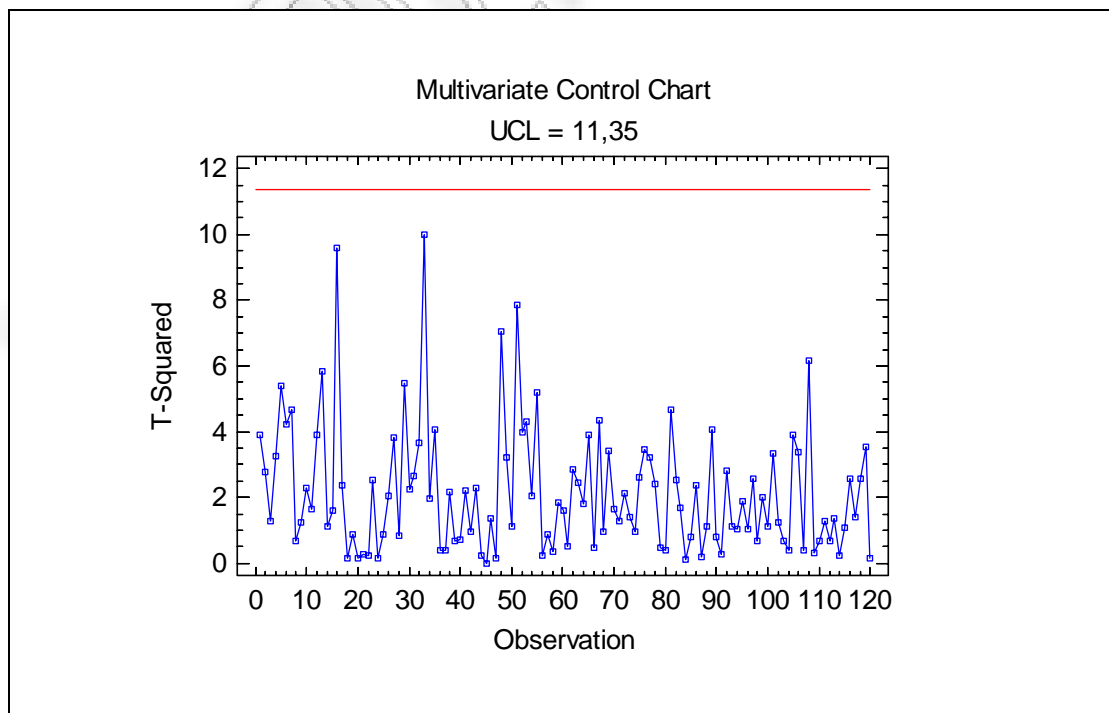
Η έντονη συσχέτιση των μεταβλητών μας φαίνεται αναλυτικά στο ακόλουθο σχήμα

**Σχήμα 8** Διαγράμματα διασποράς για τις  $Z_1$ ,  $Z_2$ ,  $Z_3$

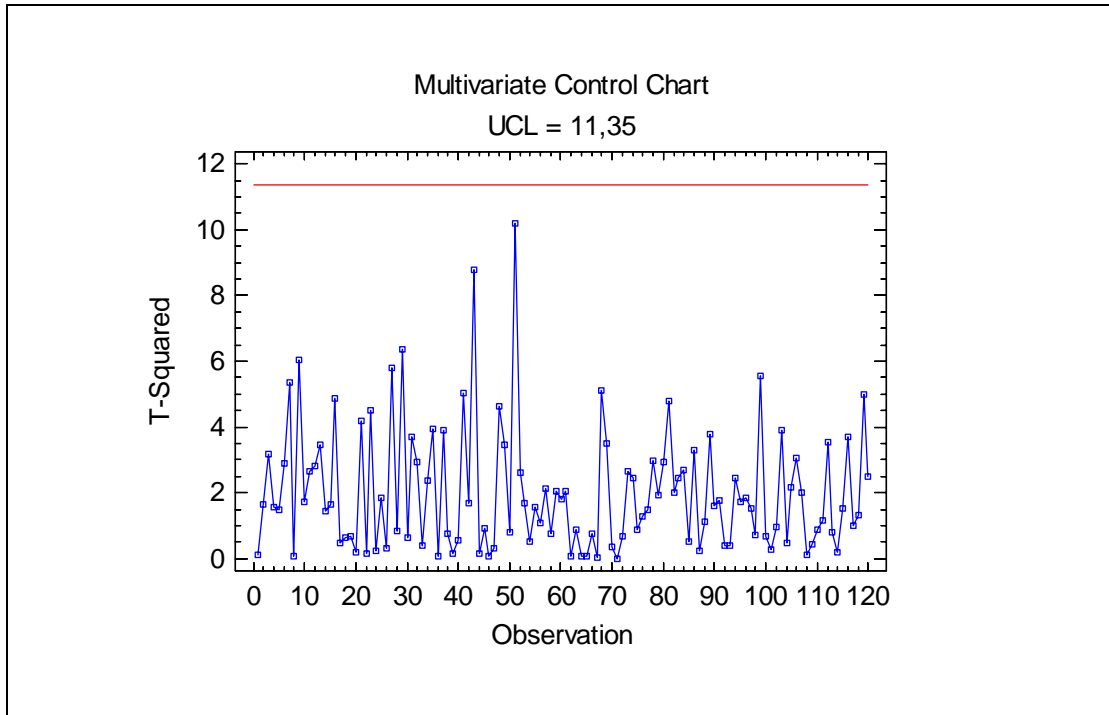


Στην περίπτωση που μελετήσουμε τις μεταβλητές μας ανά δύο έχουμε τα ακόλουθα  $T^2$  διαγράμματα ελέγχου ( $\alpha = 0.0027$  σε κάθε διάγραμμα).

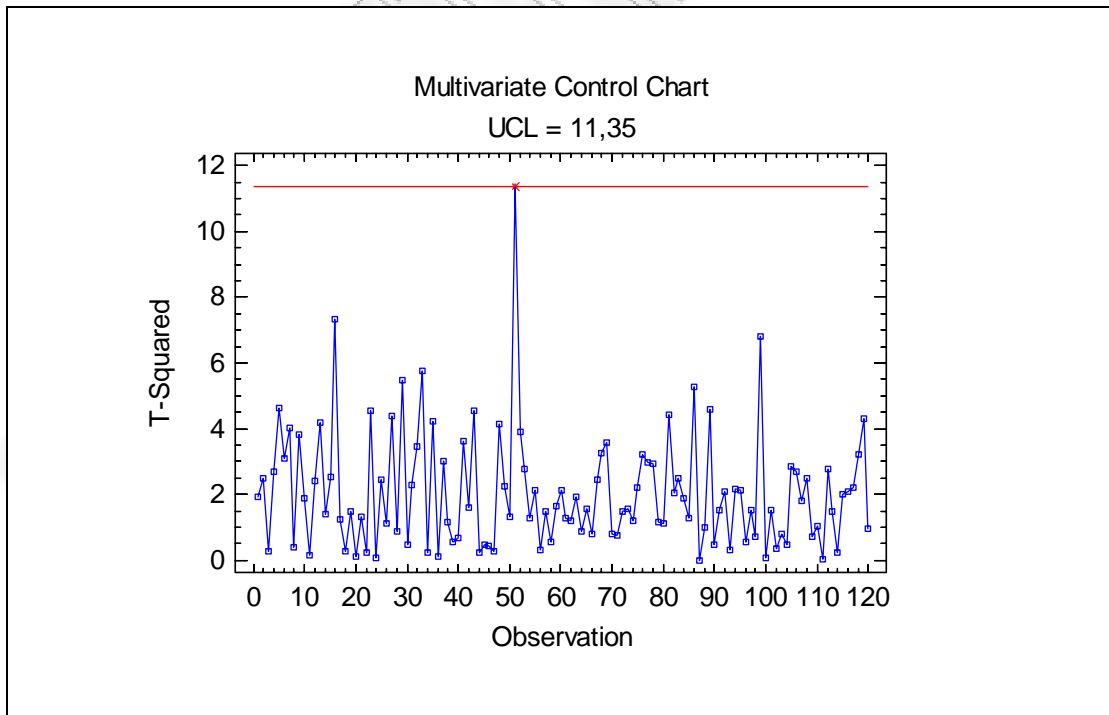
**Σχήμα 9**  $T^2$  διάγραμμα ελέγχου για τις  $Z_1$ ,  $Z_2$



Σχήμα 10  $T^2$  διάγραμμα ελέγχου για τις  $Z_1, Z_3$



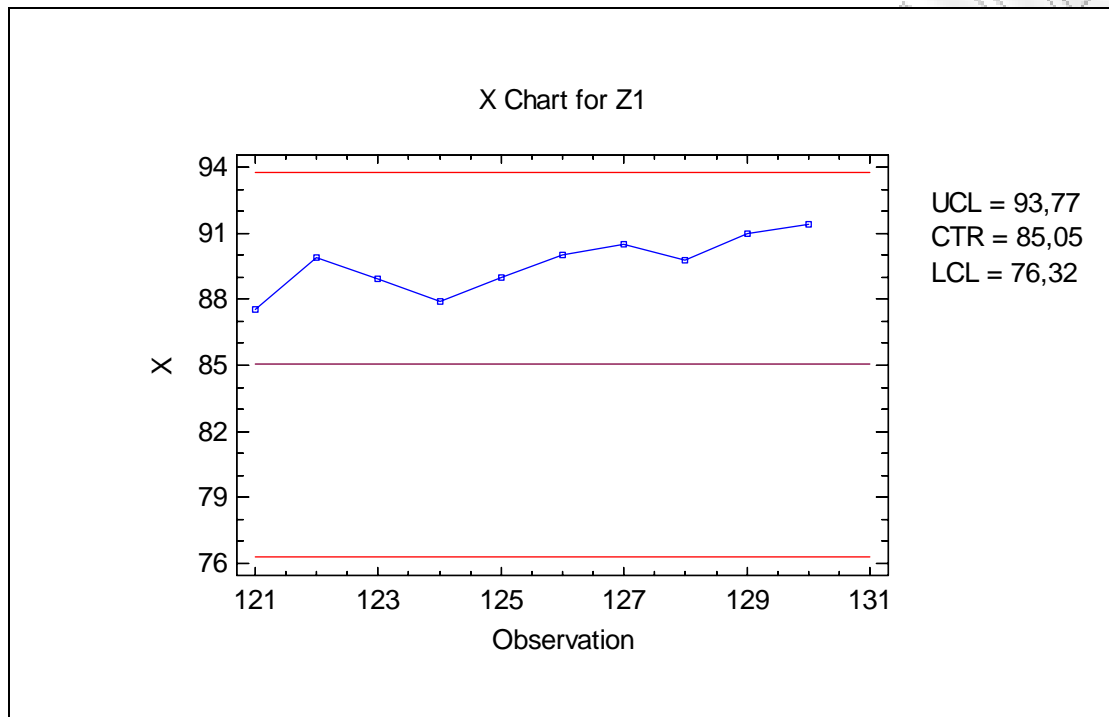
Σχήμα 11  $T^2$  διάγραμμα ελέγχου για τις  $Z_2, Z_3$



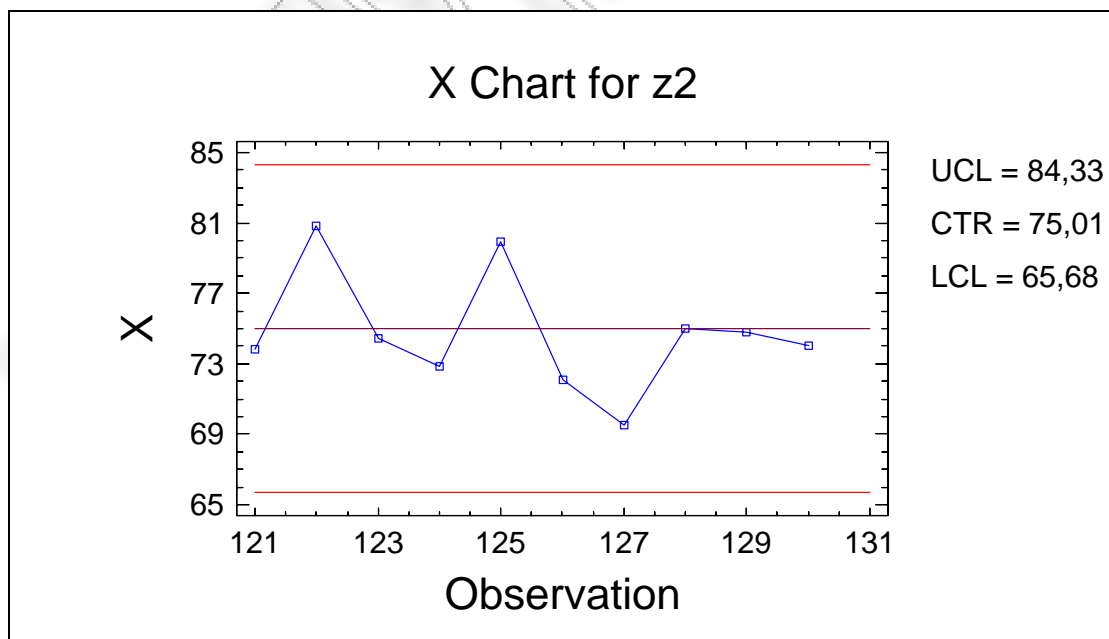


Έστω, τώρα ότι παίρνουμε 10 νέες παρατηρήσεις στις οποίες ο μέσος έχει πλέον μετατοπιστεί (10 τελευταίες γραμμές του Πίνακα 4). Ενδείξεις για την κατεύθυνση προς την οποία έχει μετατοπιστεί δίνουν τα τρία ακόλουθα  $\bar{X}$  διαγράμματα ελέγχου τα οποία όμως δεν δίνουν ένδειξη για εκτός ελέγχου διεργασία.

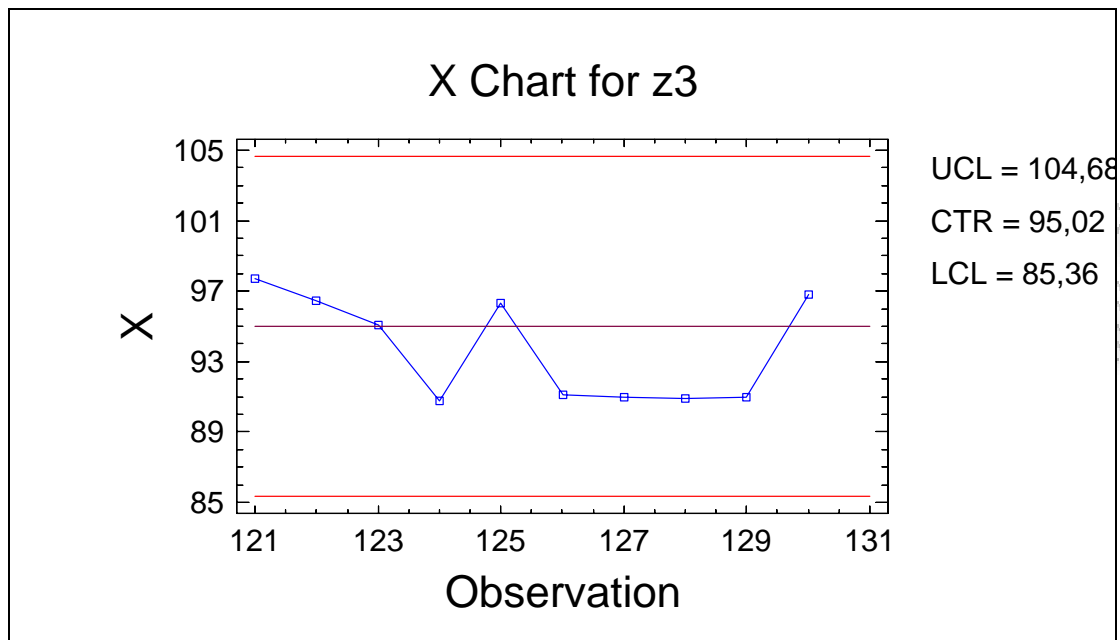
**Σχήμα 12**  $\bar{X}$  διάγραμμα ελέγχου για τη  $Z_1$



**Σχήμα 13**  $\bar{X}$  διάγραμμα ελέγχου για τη  $Z_2$

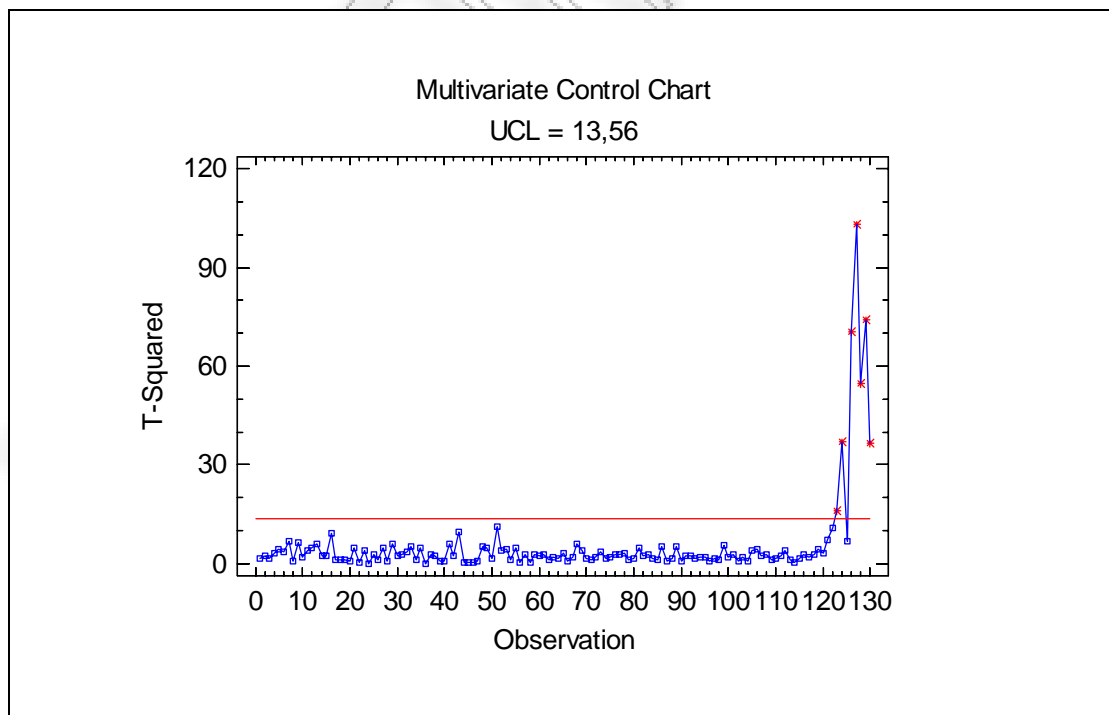


Σχήμα 14  $\bar{X}$  διάγραμμα ελέγχου για τη  $Z_3$



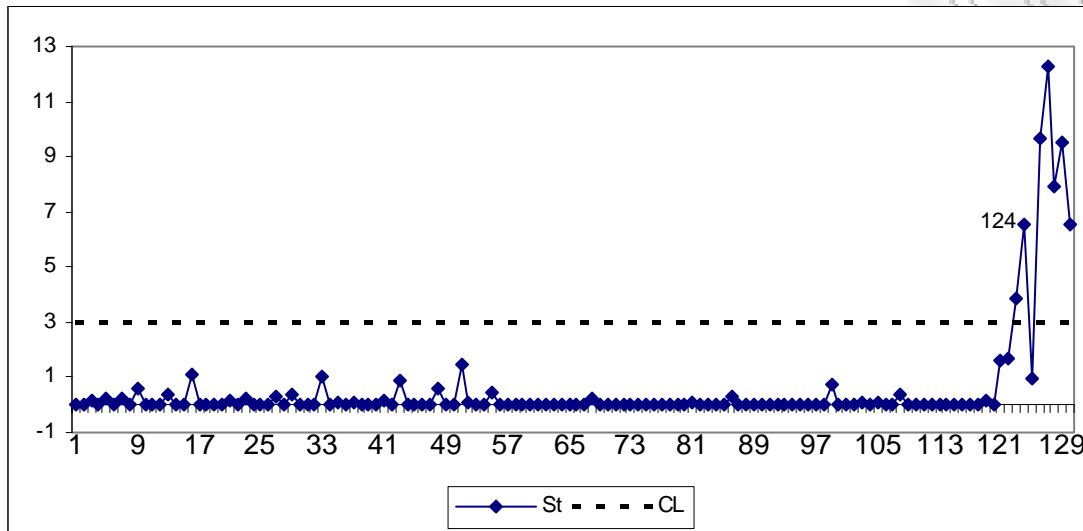
Στο Σχήμα 15 παρατηρούμε ότι το  $T^2$  διάγραμμα ελέγχου εντόπισε την μετατόπιση για πρώτη φορά στην παρατήρηση 123.

Σχήμα 15  $T^2$  διάγραμμα ελέγχου για τις  $Z_1, Z_2, Z_3$



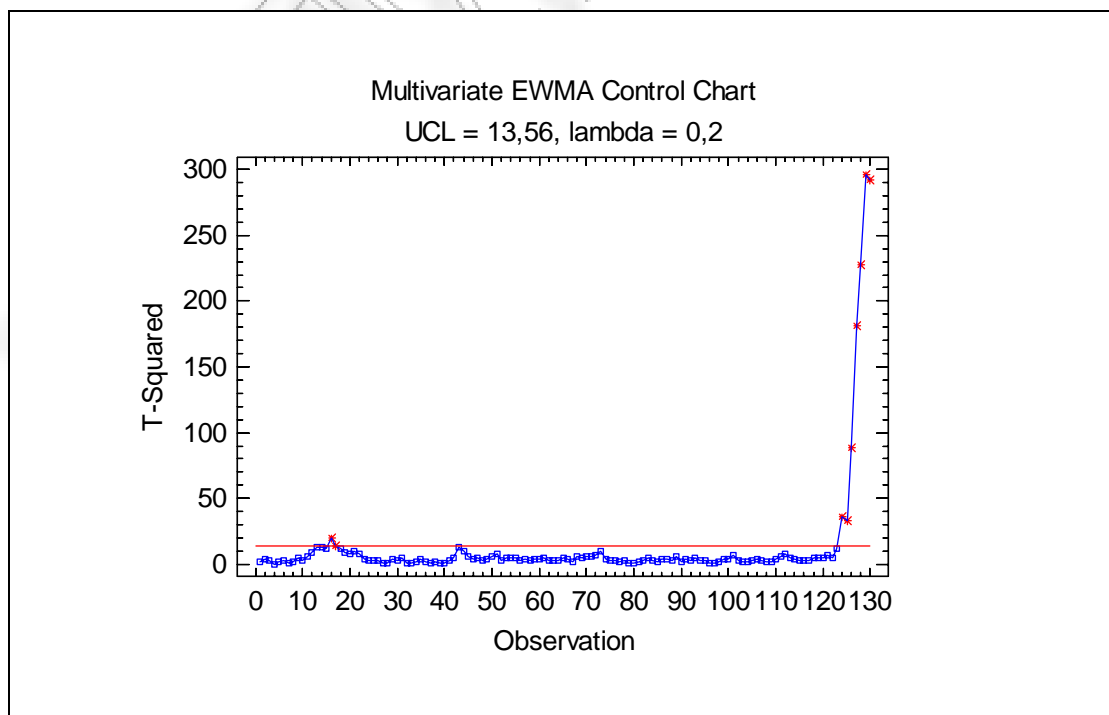
Στο σχήμα 16 δίνεται το διάγραμμα MCUSUM όπου και πήραμε με την βοήθεια του EXCEL και σύμφωνα με το Κεφάλαιο 3 και το τυπολόγιο της σελίδας 50 και 51

**Σχήμα 16** Πολυμεταβλητό Διάγραμμα Ελέγχου MCUSUM



Τέλος, δίνουμε το διάγραμμα MEWMA για όλες της παρατηρήσεις Οι τύποι που χρησιμοποιεί το πρόγραμμα αναφέρονται αναλυτικά στην αρχή του 4<sup>ου</sup> Κεφαλαίου στην σελίδα 57 και 58.

**Σχήμα 17:** Πολυμεταβλητό Διάγραμμα Ελέγχου MEWMA



# ΒΙΒΛΙΟΓΡΑΦΙΑ

## Ελληνική

**Αντζουλάκος, Δ.Α.(2003).** Στατιστικός Έλεγχος Ποιότητας, Σημειώσεις παραδόσεων του μαθήματος του Στατιστικού Ελέγχου Ποιότητας στα πλαίσια του Μ.Π.Σ. στην Εφαρμοσμένη Στατιστική, Πανεπιστήμιο Πειραιώς.

## Ξένα

- Alt, F.B. (1985).** Multivariate Quality Control. *The Encyclopedia of Statistical Sciences*, Kotz S., Johnson, N.L. and Read, C.R. (eds), New York: John Wiley, 110-122.
- Alt, F.B. and Smith, N.D. (1988).** Multivariate Process Control. *Handbook of Statistics*, P.R. Krishnaiah and C.R. Rao (eds), North-Holland: Elsevier Science Publishers B.V., 7, 333-351.
- Anderson, T.W. (1958).** *An Introduction to Multivariate Statistical Analysis*, 2<sup>nd</sup> ed, New York: John Wiley.
- Andrews, D. F.** Plots of high dimensional data, *Biometrics*, **1972**, 28, 125-136.
- Aparisi, F. (1996).** Hotelling's  $T^2$  Control Chart With Adaptive Sample Sizes. *International Journal of Production Research*, 34, 10, 2853-2862.
- Aparisi, F., Jabaloyes, J. and Carrion, A. (1999).** Statistical Properties of the |S| Multivariate Control Chart. *Communications in Statistics - Theory and Methods*, 28, 11, 2671-2686.
- Bersimis, S. (2000).** Multivariate Statistical Process Control. Paper presented at Athens University of Economics and Business, Department of Statistics.
- Boque, R. and Smilde, A.K. (1999).** Monitoring and Diagnosing Batch Processes With Multiway Covariates Regression Models. *AIChE Journal*, 45, 7, 1504-1520.
- Chan, L.K. and Li, G.Y. (1994).** A Multivariate Control Chart for Detecting Linear Trends. *Communications in Statistics - Simulation*, 23, 4, 997-1012.
- Chua, M-K. and Montgomery, D.C. (1992).** Investigation and Characterization of a Control Scheme for Multivariate Quality Control. *Quality and Reliability Engineering International*, 8, 37-44.
- Crosier, R.B. (1988).** Multivariate Generalizations of Cumulative Sum Quality-Control Schemes. *Technometrics*, 30, 3, 291-303.
- Doganaksoy, N., Schmeel, J. and Vandeven, M. (1996).** Process Monitoring with Multiple Product Grades. *Journal of Quality Technology*, 28, 3, 346-355.
- Doganaksoy, N., Faltin, F.W. and Tucker, W.T. (1991).** Identification of Out-of-Control Multivariate Characteristic in a Multivariable Manufacturing Environment. *Communications in Statistics - Theory and Methods*, 20, 9, 2775-2790.
- Fuchs, C. and Benjamini, Y. (1994).** Multivariate Profile Charts for Statistical Process Control. *Technometrics*, 36, 2, 182-195.
- Fuchs, C. and Kenett, R.S. (1998).** *Multivariate Quality Control*, New York: Marcel Dekker, INC.

- Gnanadesikan, (1997).** *Methods for Statistical Data Analysis of Multivariate Observations.* New York: John Wiley.
- Hawkins, D.M. (1991).** Multivariate Quality Control Based on Regression-Adjusted Variables. *Technometrics*, 33, 1, 61-75.
- Hawkins, D.M. (1993).** Regression Adjustment for Variables in Multivariate Quality Control. *Journal of Quality Technology*, 25, 3, 170-182.
- Hayter, A.J. and Tsui, K-L. (1994).** Identification and Quantification in Multivariate Quality Control Problems. *Journal of Quality Technology*, 26, 3, 197-208.
- Healy, J.D. (1987).** A Note on Multivariate CUSUM Procedures. *Technometrics*, 29, 4, 409-412.
- Hicks, C.R. (1955).** Some Applications of Hotelling's  $T^2$ . *Industrial Quality Control*, Vol 11(9), 23-29.
- Hotelling, H. (1947).** Multivariate Quality Control - Illustrated by the Air Testing of Sample Bombsights. In *Techniques of Statistical Analysis*, Eisenhart, C., Hastay, M.W., Wallis, W.A. (eds), New York: MacGraw-Hill, 111-184.
- Jolayemi, J.K. (1999).** A Statistical Model for the Design of Multiattribute Control Charts. *The Indian Journal of Statistics*, 61, B, 2, 351-365.
- Jolayemi, J.K. (1995).** A Power Function Model for Determining Sample Sizes for the Operations of Multivariate Control Charts. *Computational Statistics & Data Analysis*, 20, 633-641.
- Kourti, T. and MacGregor, J.F. (1996).** Multivariate SPC Methods for Process and Product Monitoring. *Journal of Quality Technology*, 28, 4, 409-428.
- Kourti, T., Lee, J. and MacGregor, J.F. (1996).** Experiences With Industrial Applications of Projection Methods for Multivariate SPC. *Computers and Chemical Engineering*, 20, S, S745-S750.
- Kourti, T., Nomikos, P. and MacGregor, J.F. (1996).** Analysis, Monitoring and Fault Diagnosis of Batch Process Using Multiblock and Multiway PLS. *Journal of Production Control*, 5, 4, 277-284.
- Liu, R.Y. (1995).** Control Charts for Multivariate Processes. *Journal of the American Statistical Association*, 90, 432, 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, 436, 1694-1700.
- Lowry, C.A. and Montgomery, D.C (1995).** A Review of Multivariate Control Charts. *IIE Transactions*, 27, 800-810.
- Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992).** A Multivariate EWMA Control Chart. *Technometrics*, 34, 1, 46-53.
- Lu, X.S., Xie, M., Goh, T.N. and Lai, C.D. (1998).** Control Charts for Multivariate Attribute Processes. *International Journal of Production Research*, 36, 12, 3477-3489.
- MacGregor, J.F., Jaeckle, C., Kiparissides, C. and Koutoudi, M. (1994).** Process Monitoring and Diagnosis by Multiblock PLS Method. *AIChE Journal*, 40, 5, 826-838.
- MacGregor, J.F. and Kourti, T. (1995).** Statistical Process Control of Multivariate Processes. *Control Engineering Practice*, 3, 3, 403-414.
- Mader, D.P., Glycenfer, J.J. and Prins, J. (1996).** An Application in Multivariate Statistical Process Control for Power Supply Calibration. *Quality Engineering*, 9, 1, 99-106.

- Maravelakis, P.E. (1998).** CUSUM Procedures in Statistical Process Control. Msc. Thesis. Athens University of Economics and Business Department of Statistics.
- Mason, R.L., Champ, C.W., Tracy, N.D., Wierda, R.J. and Young, J.C. (1997).** Assessment of Multivariate Process Control Techniques. *Journal of Quality Technology*, 29, 2, 140-143.
- Mason, R.L., Tracy, N.D. and Young, J.C. (1995).** Decomposition of T<sup>2</sup> for Multivariate Control Chart Interpretation. *Journal of Quality Technology*, 27, 2, 99-108.
- Mason, R.L., Tracy, N.D. and Young, J.C. (1996).** Monitoring a Multivariate Step Process. *Journal of Quality Technology*, 28, 1, 39-50.
- Mason, R.L., Tracy, N.D. and Young, J.C. (1997).** A Practical Approach for Interpreting Multivariate T<sup>2</sup> Control Chart Signals. *Journal of Quality Technology*, 29, 4, 396-406.
- Mason, R.L. and Young, J.C. (1999).** Improving the Sensitivity of the T<sup>2</sup> Statistic in Multivariate Process Control. *Journal of Quality Technology*, 31, 2, 155-165.
- Monopolis, I.K. (1999).** Exponentially Weighted Moving Averages Processes in Statistical Process Control. Msc. Thesis. Athens University of Economics and Business Department of Statistics.
- Montgomery, D.C. (1991).** *Introduction to Statistical Quality Control*. 2<sup>nd</sup> ed, New York: John Wiley.
- Montgomery, D.C. (2005).** *Introduction to Statistical Quality Control*. 5<sup>th</sup> ed, New York: John Wiley & Sons, Inc.
- Montgomery, D.C., and H.M. Wadsworth, Jr. (1972).** Some Techniques for Multivariate Quality Control Applications, ASQC Technical Conference Transactions, Washington, DC, May, pp. 427-435.
- Murphy, B.J. (1987).** Selecting Out-of-Control Variables with T<sup>2</sup> Multivariate Quality Procedures. *The Statistician*, 36, 571-583.
- Page, E.S. (1954).** Continuous Inspection Schemes, *Biometrics*, 41(1), 100-115.
- Pignatiello, J.J. and Runger, G.C. (1990).** Comparisons of Multivariate CUSUM Charts. *Journal of Quality Technology*, 22, 1, 173-186.
- Prabhu, S.S. and Runger, G.C. (1997).** Designing a Multivariate EWMA Control Chart. *Journal of Quality Technology*, 29, 1, 8-15.
- Rigdon, S.E. (1995).** A double-integral Equation for the Average run length of a mewma Control Chart. *Statistics and Probability Letters*, 24, 365-373.
- Rigdon, S.E. (1995).** An Integral Equation for the In-Control Average Length of a MEWMA Control Chart. *Journal of Statistical Computations and Simulation*, 52, 351-365.
- Runger, G.C. (1996).** Projections and the U<sup>2</sup> Multivariate Control Chart. *Journal of Quality Technology*, 28, 3, 313-319.
- Runger, G.C. (1996).** Multivariate Statistical Process Control for Autocorrelated Processes. *International Journal of Production Research*, 34, 6, 1715-1724.
- Runger, G.C., and Alt, F.B. (1996).** Choosing Principal Components for Multivariate SPC. *Communications in Statistics - Theory and Methods*, 25, 5, 909-922.
- Runger, G.C., Alt, F.B. and Montgomery, D.C. (1996a).** Contributors to a Multivariate SPC Chart signal. *Communications in Statistics - Theory and Methods*, 25, 10, 2203-2213.

- Runger, G.C., Alt, F.B. and Montgomery, D.C. (1996b).** Controlling Multiple Stream Processes with Principal Components. *International Journal of Production Research*, 34, 11, 2991-2999.
- Runger, G.C., Montgomery, D.C., Keats, and Scranton, (1999).** Improving the Performance of a Multivariate EWMA Control Chart. *Quality and Reliability Engineering International*, 15, 161-166.
- Runger, G.C. and Prabhu, S.S. (1996).** A Markov Chain Model for the Multivariate EWMA Control Chart. *Journal of the American Statistical Association*, 91, 436, 1701-1706.
- Shanmugalingam, S. (1982).** On the Analysis of the Ratio of Two Correlated Normal Variables. *The statistician*, 31, 3, 251-258.
- Sparks, R.S. (1992).** Quality Control with Multivariate Data. *Australian Journal of Statistics*, 34, 3, 375-390.
- Sparks, R.S., Adolphson, A. and Phatak, A. (1997).** Multivariate Process Monitoring Using the Dynamic Biplot. *International Statistical Review*, 65, 3, 325-349.
- Tracy, N.D., Young, J.C. and Mason, R.L. (1992).** Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 24, 2, 88-95.
- Tracy, N.D., Young, J.C. and Mason, R.L. (1995).** A Bivariate Control Chart for Paired Measurements. *Journal of Quality Technology*, 27, 4, 370-376.
- Wierda, S.J. (1994).** Multivariate Statistical Process Control - Recent Results and Directions for Future Research. *Statistica Neerlandica*, 48, 2, 147-168.
- Woodall, W.H. and Montgomery, D.C. (1999).** Research Issues and Ideas in Statistical Process Control. *Journal of Quality Technology*, 31, 4, 376-386.
- Woodall, W.H. and Ncube, M.M. (1985).** Multivariate CUSUM Quality Control Procedures. *Technometrics*, 27, 3, 285-292.
- Yumin, L. (1996).** An Improvement for MEWMA in Multivariate Process Control. *Computers and Industrial Engineering*, 31, Y, 779-781.

## Άλλες Προτεινόμενες Βιβλιογραφικές Πηγές

- Bodden, K.M. and Rigdon, S.E. (1999).** A Program for Approximating the In-Control ARL for the MEWMA Chart.
- Bose, N.K. and Liang, P. (1996).** *Neural Networks Fundamentals With Graphs, Algorithms, and Applications*. New York: McGraw-Hill, Inc.
- Chou, Y.-M., Mason, R.L. and Young, J.C. (1999).** Power Comparisons for a Hotelling's  $T^2$  Statistic. *Communications in Statistics - Simulation*, 28, 4, 1031-1050.
- Dayal, B.S., MacGregor, J.F., Taylor, P.A., Kildaw, R. and Marcikic, S. (1994).** Application of Feedforward Neural Networks and PLS Regression for modelling Kappa Number in a continuous Kamyr Digester. *Pulp and Paper Canada*, 95, 1, 26-32.
- Fasso, A. (1999).** One-Sided MEWMA Control Charts. *Communications in Statistics - Theory and Methods*, 28, 381-401
- Fong, D.Y.T. and Lawless, J.F. (1998).** The Analysis of Process Variation Transmission with Multivariate Measurements. *Statistica Sinica*, 8, 151-164.
- 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, 5, 1313-1321.
- Martin, E.B., Morris, A.J., Papazoglou, M.C., Kiparissides, C. (1996).** Batch Process Monitoring for Consistent Production. *Computers and Chemical Engineering*, 20, S, S599-S604.
- Sullivan, J.H. and Woodall, W.H. (1996).** A Comparison of Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 28, 4, 398-408.
- Sullivan, J.H. and Woodall, W.H. (1998).** Adapting Control Charts for the Preliminary Analysis of Multivariate Observations. *Communications in Statistics - Simulation*, 27, 4, 953-979.