

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Σχολή Χρηματοοικονομικής και Στατιστικής



Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

Διαγράμματα ελέγχου διαμέσου

Μάστορας Θρασύβουλος

Διπλωματική Εργασία

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

Πειραιάς
Ιούνιος 2026

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

Τα μέλη της Επιτροπής ήταν:

- Αντζουλάκος Δημήτριος (Επιβλέπων), Αναπληρωτής Καθηγητής
- Ευαγγελάρας Χαράλαμπος, Αναπληρωτής Καθηγητής
- Ρακιτζής Αθανάσιος, Αναπληρωτής Καθηγητής

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

UNIVERSITY OF PIRAEUS

School of Finance and Statistics



Department of Statistics and Insurance Science

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

Median Control Charts

Thrasyvoulos Mastoras

MSc Dissertation

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

Piraeus, Greece

June 2026

Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω πρωτίστως τον κ. Δημήτριο Αντζουλάκο, επιβλέποντα καθηγητή μου, για την καθοδήγηση και την βοήθεια που μου παρείχε καθ' όλη την διάρκεια εκπόνησης αυτής της εργασίας διότι αποδείχτηκαν πολύτιμες. Ακόμα , οφείλω να ευχαριστήσω και τους ανθρώπους που με στήριξαν τόσο από την οικογένεια μου όσο και από το φιλικό μου περιβάλλον.

Περίληψη

Τα διαγράμματα ελέγχου είναι ένα σημαντικό και ισχυρό εργαλείο σε μια προσπάθεια προς μια σταθερότερη παραγωγική διεργασία. Μέσω αυτών των διαγραμμάτων, μπορούν να ελεγχθούν πιθανές αποκλίσεις σε διάφορες παραμέτρους του προϊόντος, οι οποίες αν δεν προσδιοριστούν εγκαίρως, θα μπορούσαν ενδεχομένως να οδηγήσουν σε διεργασία εκτός ελέγχου. Η αποτελεσματική χρήση των διαγραμμάτων ελέγχου έχει ως αποτέλεσμα την παραγωγή προϊόντων βέλτιστης ποιότητας και την ελαχιστοποίηση της παραγωγής μη συμμορφούμενων προϊόντων. Συνήθως, χρησιμοποιούνται δύο τύποι διαγραμμάτων ελέγχου για μεταβλητές. Ένα διάγραμμα για την παρακολούθηση της μέσης τιμής της διεργασίας και ένα για την παρακολούθηση της μεταβλητότητας της διεργασίας.

Το διάγραμμα ελέγχου διαμέσου $\bar{X}$ είναι ένα δημοφιλές διάγραμμα για την παρακολούθηση της μέσης τιμής μιας διαδικασίας. Αν και το $\bar{X}$ διάγραμμα ελέγχου είναι ένα πιο αποδοτικό διάγραμμα, το διάγραμμα ελέγχου διαμέσου αποτελεί μια χρήσιμη εναλλακτική λύση, με ενδιαφέρουσα απόδοση. Τα κύρια πλεονεκτήματα του διαγράμματος ελέγχου διαμέσου είναι ότι είναι απλούστερο από το $\bar{X}$ διάγραμμα ελέγχου και είναι ανθεκτικό σε ακραίες τιμές ή μικρές αποκλίσεις από την κανονικότητα.

Η παρούσα διατριβή χωρίζεται σε τέσσερα κεφάλαια. Στο πρώτο κεφάλαιο γίνεται περιγραφή των διαγραμμάτων ελέγχου για την παρακολούθηση της μέσης τιμής και της μεταβλητότητας μιας διεργασίας. Επιπλέον, παρουσιάζονται βασικές έννοιες όπως το μέσο μήκος ροής καθώς και εκτιμήτριες της μέσης τιμής και της τυπικής απόκλισης που χρησιμοποιούνται στα διαγράμματα ελέγχου. Στο δεύτερο κεφάλαιο μελετώνται εκτενώς τα διαγράμματα ελέγχου διαμέσου τύπου Shewhart, τόσο για τη Φάση II όσο και για τη Φάση I. Εξετάζεται η απόδοσή τους μέσω του μέσου μήκους ροής για διάφορα μεγέθη δειγμάτων και για διαφορετικές μετατοπίσεις της μέσης τιμής της διεργασίας. Στο τρίτο κεφάλαιο μελετώνται αναλυτικά τα διαγράμματα ελέγχου διαμέσου τύπου CUSUM, ενώ στο τέταρτο κεφάλαιο μελετώνται αναλυτικά τα διαγράμματα ελέγχου διαμέσου τύπου EWMA. Εξετάζεται η απόδοσή τους μέσω του μέσου μήκους ροής και παρουσιάζονται αποτελέσματα για διάφορους συνδυασμούς των παραμέτρων τους και διαφορετικές μετατοπίσεις της μέσης τιμής της διεργασίας.

Abstract

Shewhart control charts are a significant and powerful tool in an attempt towards a more stable production process. Through these charts, possible deviations on various product parameters can be controlled, which if not identified in time, could possibly lead to an out of control process. Effective use of Shewhart control charts results in the production of optimum quality products and minimize the production of non-conforming ones. Usually, two types of control charts for variables are compiled. One chart for monitoring the process mean and one for monitoring the process variability.

The median $\bar{X}$ chart is a popular chart for monitoring the mean of a process. Although the mean chart $\bar{\bar{X}}$ is a more efficient chart, the median chart is a useful alternative, with an interesting performance. The main advantages of the median chart are that it is simpler than the $\bar{\bar{X}}$ chart and it is robust against outliers or small deviations from normality.

The present thesis is divided into four chapters. The first chapter describes the control charts for monitoring the mean value and variability of a process. In addition, the average run length is presented as well as mean and variability estimators used in the control charts. In the second chapter, Shewhart-type median control charts are studied in detail, for both Phase II and Phase I. Their performance is evaluated through the average run length for various sample sizes and different shifts in the process mean. In the third chapter, CUSUM-type median control charts are examined analytically, while in Chapter 4 the EWMA-type median control charts are studied in detail. Their performance is evaluated through the average run length and results are presented for various combinations of their parameters and different shifts in the process mean.

Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1	1
Στατιστικός έλεγχος διεργασιών και διαγράμματα ελέγχου	1
1.1 Εισαγωγή.....	1
1.2 Περιγραφή και χρήση ενός διαγράμματος ελέγχου	2
1.3 Διαγράμματα ελέγχου Φάσης I και Φάσης II	5
1.4 Εκτίμηση της μέσης τιμής και της διασποράς.....	5
1.4.1 Η περίπτωση των μεμονωμένων παρατηρήσεων.....	5
1.4.2 Η περίπτωση των δειγμάτων	7
1.5 Διαγράμματα ελέγχου Shewhart για τη μέση τιμή	8
1.5.1 Διάγραμμα ελέγχου Φάσης II	8
1.5.2 Διαγράμματα ελέγχου Φάσης I.....	9
1.6 Διαγράμματα ελέγχου Shewhart για τη διασπορά	12
1.6.1 Διάγραμμα ελέγχου Φάσης II	12
1.6.2 Διάγραμμα ελέγχου Φάσης I.....	13
1.7 Διαγράμματα ελέγχου με μνήμη.....	15
1.7.1 Διάγραμμα ελέγχου CUSUM	15
1.7.2 Διάγραμμα ελέγχου EWMA	19
1.8 Μέσο μήκος ροής.....	22
ΚΕΦΑΛΑΙΟ 2	27
Διαγράμματα ελέγχου διαμέσου τύπου Shewhart	27
2.1 Εισαγωγή.....	27
2.2 Η κατανομή της δειγματικής διαμέσου.....	28
2.3 Το διάγραμμα ελέγχου διαμέσου	31
2.3.1 Φάση II – Γνωστές παράμετροι.....	31
2.3.2 Φάση I – Άγνωστες παράμετροι.....	34
2.4 Απόδοση του διαγράμματος ελέγχου διαμέσου	36
2.5 Εφαρμογή διαγράμματος ελέγχου διαμέσου	41
2.6 Μη παραμετρικό διάγραμμα ελέγχου διαμέσου	45
2.7 Εφαρμογή μη παραμετρικού διαγράμματος ελέγχου διαμέσου	50
ΚΕΦΑΛΑΙΟ 3	57
Διαγράμματα ελέγχου διαμέσου τύπου CUSUM.....	57

3.1 Εισαγωγή.....	57
3.2 Το διάγραμμα ελέγχου διαμέσου τύπου CUSUM.....	58
3.2.1 Φάση II – Γνωστές παράμετροι.....	58
3.2.2 Φάση I – Άγνωστες παράμετροι.....	61
3.3 Απόδοση του διαγράμματος ελέγχου διαμέσου τύπου CUSUM	63
3.4 Εφαρμογή διαγράμματος ελέγχου διαμέσου τύπου CUSUM.....	73
ΚΕΦΑΛΑΙΟ 4	77
Διαγράμματα ελέγχου διαμέσου τύπου EWMA.....	77
4.1 Εισαγωγή.....	77
4.2 Διάγραμμα ελέγχου διαμέσου τύπου EWMA	78
4.2.1 Φάση II – Γνωστές Παράμετροι.....	78
4.2.2 Φάση I – Άγνωστες Παράμετροι	81
4.3 Απόδοση του διαγράμματος ελέγχου διαμέσου τύπου EWMA.....	83
4.4 Εφαρμογή διαγράμματος ελέγχου διαμέσου τύπου EWMA.....	89
Παράρτημα	95
Κ.1 Κώδικας για την εξαγωγή του Πίνακα 2.1.....	95
Κ.2 Κώδικας για την εξαγωγή του Πίνακα 2.2.....	97
Κ.3 Κώδικας για την εξαγωγή του Πίνακα 2.4.....	99
Κ.4 Κώδικας για την εξαγωγή του Πίνακα 3.1.....	101
Κ.5 Κώδικας για την εξαγωγή του Πίνακα 3.3.....	103
Κ.6 Κώδικας για την εξαγωγή του Πίνακα 4.1.....	105
Κ.7 Κώδικας για την εξαγωγή του Πίνακα 4.3.....	107
Κ.8 Κώδικας για την εξαγωγή του Σχήματος 4.5.....	109
Βιβλιογραφία.....	111

Πίνακες

Πίνακας 1.1. Τιμές των σταθερών d_2 , d_3 και c_4	6
Πίνακας 1.2. Δεδομένα για τη διάμετρο της στεφάνης του εμβόλου.....	11
Πίνακας 1.3: Δεδομένα για την επίδειξη διαγράμματος CUSUM: $k = 5$, $h = 5$	17
Πίνακας 1.4. Δεδομένα για την επίδειξη διαγράμματος EWMA: $\lambda = 0.1$, $L = 2.7$	21
Πίνακας 2.1. Τιμές της παραμέτρου K	36
Πίνακας 2.2. Τιμές ARL ($ARL_0 = 370.4$).....	37
Πίνακας 2.3. Τιμές ARL ($ARL_0 = 500$).....	38
Πίνακας 2.4. Τιμές της παραμέτρου K' ($ARL_0 = 370.4$).....	39
Πίνακας 2.5. Προσομοιωμένα δεδομένα Φάσης Ι.....	40
Πίνακας 2.6. Δεδομένα Φάσης Ι.....	42
Πίνακας 2.7. Παράμετροι διαγράμματος ελέγχου προτεραιότητας διαμέσου.....	47
Πίνακας 2.8. Δεδομένα δακτυλίων εμβόλων (Montgomery (2013)).....	48
Πίνακας 3.1. Τιμές h για $ARL_0 = 370.4$	64
Πίνακας 3.2. Τιμές h για $ARL_0 = 500$	64
Πίνακας 3.3. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($ARL_0 = 370.4$).....	65
Πίνακας 3.4. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($ARL_0 = 500$).....	66
Πίνακας 3.5. Τιμές C_i^+ και Q_i^+ διαγραμμάτων CUSUM.....	71
Πίνακας 4.1. Τιμές k για $ARL_0 = 370.4$	84
Πίνακας 4.2. Τιμές k για $ARL_0 = 500$	84
Πίνακας 4.3. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($ARL_0 = 370.4$).....	85
Πίνακας 4.4. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($ARL_0 = 500$).....	86
Πίνακας 4.5. Τιμές $Z_i^{\bar{X}}$ και $Z_i^{\bar{X}}$ διαγραμμάτων EWMA.....	90

Σχήματα

Σχήμα 1.1. Τυπικό διάγραμμα ελέγχου Shewhart.....	3
Σχήμα 1.2: $\bar{X}$ διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.2.....	12
Σχήμα 1.3: CUSUM διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.3.....	18
Σχήμα 1.4. EWMA διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.4.....	22
Σχήμα 1.5. Καμπύλη ARL του $\bar{X}$ διαγράμματος: $L = 3, n = 2$	24
Σχήμα 2.1. Διάγραμμα ελέγχου διαμέσου Φάσης I.....	41
Σχήμα 2.2. Διάγραμμα ελέγχου διαμέσου Φάσης II.....	43
Σχήμα 2.3. S διάγραμμα ελέγχου Φάσης I για την εσωτερική διάμετρο.....	50
Σχήμα 2.4. $\bar{X}$ διάγραμμα ελέγχου Φάσης I για την εσωτερική διάμετρο	51
Σχήμα 2.5. S διάγραμμα ελέγχου Φάσης II για την εσωτερική διάμετρο	52
Σχήμα 2.6. $\bar{X}$ διάγραμμα ελέγχου Φάσης II για την εσωτερική διάμετρο.....	52
Σχήμα 2.7. Διάγραμμα ελέγχου προτεραιότητας για την εσωτερική διάμετρο.....	54
Σχήμα 3.1. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 3, ARL_0 = 370.4$).....	67
Σχήμα 3.2. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 5, ARL_0 = 370.4$).....	68
Σχήμα 3.3. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 7, ARL_0 = 370.4$).....	68
Σχήμα 3.4. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 9, ARL_0 = 370.4$).....	69
Σχήμα 3.5. Άνω μονόπλευρο CUSUM $\bar{X}$ διάγραμμα ελέγχου Φάσης II για την εσωτερική διάμετρο.....	72
Σχήμα 3.6. Άνω μονόπλευρο CUSUM διάγραμμα ελέγχου διαμέσου Φάσης II για την εσωτερική διάμετρο.....	73
Σχήμα 4.1. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 3, ARL_0 = 370.4$).....	87
Σχήμα 4.2. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 5, ARL_0 = 370.4$).....	87
Σχήμα 4.3. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 7, ARL_0 = 370.4$).....	88
Σχήμα 4.4. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 9, ARL_0 = 370.4$).....	88
Σχήμα 4.5. EWMA $\bar{X}$ διάγραμμα ελέγχου για την εσωτερική διάμετρο.....	91
Σχήμα 4.6. EWMA διάγραμμα ελέγχου διαμέσου για την εσωτερική διάμετρο.....	92

ΚΕΦΑΛΑΙΟ 1

Στατιστικός έλεγχος διεργασιών και διαγράμματα ελέγχου

1.1 Εισαγωγή

Κάθε καταναλωτής που επιλέγει ένα προϊόν ή μια υπηρεσία έρχεται αντιμέτωπος με την ανάγκη σύγκρισης μεταξύ εναλλακτικών επιλογών που ικανοποιούν την ίδια απαίτηση. Η τελική του απόφαση επηρεάζεται κυρίως από δύο βασικούς παράγοντες, την ποιότητα και το κόστος.

Η σημασία αυτών των παραγόντων οδήγησε τις επιχειρήσεις στην ανάπτυξη μεθόδων που επιτρέπουν τη διασφάλιση υψηλής ποιότητας με το μικρότερο δυνατό κόστος παραγωγής. Στο πλαίσιο αυτό αναπτύχθηκαν σύγχρονες προσεγγίσεις διοίκησης και ελέγχου της παραγωγικής διαδικασίας, με βασικό στόχο τη συνεχή βελτίωση της ποιότητας των προϊόντων και των υπηρεσιών.

Κεντρικό ρόλο στις διαδικασίες αυτές κατέχει ο Στατιστικός Έλεγχος Ποιότητας (Statistical Quality Control), ο οποίος αποτελεί μία από τις σημαντικότερες μεθόδους παρακολούθησης και βελτίωσης των παραγωγικών διεργασιών. Βασικός του στόχος είναι η έγκαιρη ανίχνευση αποκλίσεων από τις επιθυμητές προδιαγραφές, ώστε να λαμβάνονται διορθωτικές ενέργειες και να διασφαλίζεται η σταθερότητα της διεργασίας (δείτε Μπερσίμης, Ρακιτζής και Σαχλάς (2021), Ταγαρά (2001) και Montgomery (2013)).

Ο Στατιστικός Έλεγχος Ποιότητας περιλαμβάνει ένα σύνολο στατιστικών εργαλείων που εφαρμόζονται σε διαφορετικά στάδια της παραγωγικής διαδικασίας. Ενδεικτικά, διακρίνονται τρεις βασικές κατηγορίες:

- Σχεδιασμός Πειραμάτων (Design of Experiments)
- Στατιστικός Έλεγχος Διεργασιών (Statistical Process Control)

- Δειγματοληψία Αποδοχής (Acceptance Sampling).

Στην παρούσα εργασία το ενδιαφέρον επικεντρώνεται στον Στατιστικό Έλεγχο Διεργασιών.

Η μεταβλητότητα που εμφανίζεται σε μία διεργασία μπορεί να αποδοθεί είτε σε φυσικές αιτίες, οι οποίες είναι εγγενείς στο σύστημα, είτε σε ειδικές αιτίες, οι οποίες οφείλονται σε εξωτερικούς ή μη προβλέψιμους παράγοντες, όπως βλάβες εξοπλισμού ή μεταβολές στις πρώτες ύλες. Μία διεργασία θεωρείται ότι βρίσκεται υπό στατιστικό έλεγχο όταν επηρεάζεται μόνο από φυσικές αιτίες μεταβλητότητας, ενώ η εμφάνιση ειδικών αιτιών υποδηλώνει κατάσταση εκτός ελέγχου.

Για την ανίχνευση τέτοιων μεταβολών χρησιμοποιούνται τα διαγράμματα ελέγχου. Τα κλασικά διαγράμματα ελέγχου Shewhart αποτελούν το πρώτο και πιο διαδεδομένο εργαλείο, παρουσιάζουν όμως περιορισμένη αποτελεσματικότητα στην ανίχνευση μικρών μετατοπίσεων της μέσης τιμής της διεργασίας. Για τον λόγο αυτό αναπτύχθηκαν πιο προηγμένες μέθοδοι, όπως τα διαγράμματα τύπου CUSUM και EWMA, τα οποία είναι περισσότερο ευαίσθητα σε μικρές μεταβολές.

Στην παρούσα εργασία εξετάζονται διαγράμματα ελέγχου που βασίζονται στη δειγματική διάμεσο, η οποία αποτελεί έναν ανθεκτικό εκτιμητή της κεντρικής τάσης, ιδιαίτερα σε περιπτώσεις όπου υπάρχουν ακραίες τιμές ή αποκλίσεις από την κανονικότητα. Στόχος είναι η μελέτη της απόδοσης των διαγραμμάτων αυτών ως προς την ικανότητά τους να ανιχνεύουν μετατοπίσεις της μέσης τιμής της διεργασίας.

Στο παρόν κεφάλαιο παρουσιάζονται συνοπτικά τα κλασικά διαγράμματα ελέγχου Shewhart, CUSUM και EWMA, καθώς και οι θεμελιώδεις έννοιες που τα συνοδεύουν. Στα επόμενα κεφάλαια αναλύονται αναλυτικά τα αντίστοιχα διαγράμματα διαμέσου και εξετάζεται η απόδοσή τους.

1.2 Περιγραφή και χρήση ενός διαγράμματος ελέγχου

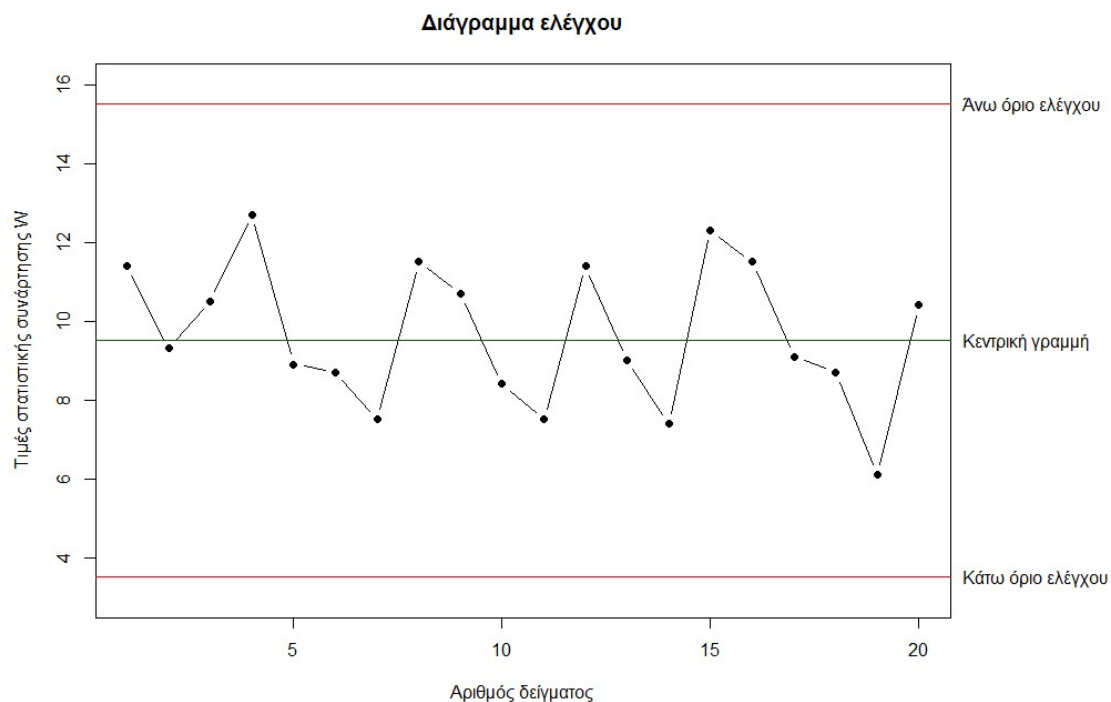
Κατά την παραγωγική διαδικασία μας ενδιαφέρει η παρακολούθηση της συμπεριφοράς μιας κρίσιμης ποσότητας ενός χαρακτηριστικού X (τυχαία μεταβλητή). Για την παρακολούθηση του συγκεκριμένου χαρακτηριστικού επιλέγονται τυχαία δείγματα προϊόντων από την παραγωγή σε διάφορες χρονικές στιγμές τα οποία αντιστοιχούν σε τυχαία δείγματα τιμών του χαρακτηριστικού X , έστω τα $\mathbf{X}_1, \mathbf{X}_2, \dots$. Στη συνέχεια υπολογίζεται η τιμή μιας στατιστικής συνάρτησης $W_i = g(\mathbf{X}_i)$, $i = 1, 2, 3, \dots$ η

οποία εκτιμά ένα κρίσιμο χαρακτηριστικό ποιότητας της X . Έτσι τη συμπεριφορά του κρίσιμου χαρακτηριστικού που μας ενδιαφέρει την παρακολουθούμε μέσα από τις τιμές που παίρνει η στατιστική συνάρτηση W , οι οποίες απεικονίζονται σε ένα ειδικό διάγραμμα που ονομάζεται διάγραμμα ελέγχου. Μέσω του διαγράμματος αυτού παρακολουθείται η συμπεριφορά της διεργασίας στον χρόνο και ελέγχεται αν αυτή παραμένει υπό στατιστικό έλεγχο (δείτε Αντζουλάκος (2010)).

Ένα τυπικό διάγραμμα ελέγχου αποτελείται από τρία βασικά στοιχεία:

- την κεντρική γραμμή (Center Line – CL),
- το άνω όριο ελέγχου (Upper Control Limit – UCL),
- και το κάτω όριο ελέγχου (Lower Control Limit – LCL).

Η κεντρική γραμμή αντιπροσωπεύει συνήθως το μέσο επίπεδο λειτουργίας της διεργασίας όταν αυτή λειτουργεί υπό συνθήκες φυσικής μεταβλητότητας, ενώ τα όρια ελέγχου χρησιμοποιούνται για την ανίχνευση πιθανών μεταβολών ή ειδικών αιτιών μεταβλητότητας.



Σχήμα 1.1. Τυπικό διάγραμμα ελέγχου Shewhart

Στο παραπάνω διάγραμμα παρατηρούμε ότι όλες οι τιμές της στατιστικής συνάρτησης W_i ($i = 1, 2, \dots, 20$) βρίσκονται εντός των ορίων ελέγχου. Στην περίπτωση αυτή θεωρούμε ότι η διεργασία λειτουργεί υπό στατιστικό έλεγχο και επηρεάζεται μόνο

από φυσικές αιτίες μεταβλητότητας. Αντίθετα, όταν τουλάχιστον μία τιμή βρεθεί εκτός των ορίων ελέγχου, τότε υπάρχει ένδειξη ότι έχουν εμφανιστεί ειδικές αιτίες μεταβλητότητας και η διεργασία πιθανόν να βρίσκεται εκτός ελέγχου.

Επιπλέον, σε ένα διάγραμμα ελέγχου δεν εξετάζεται μόνο το αν τα σημεία βρίσκονται εντός των ορίων ελέγχου, αλλά και ο τρόπος με τον οποίο κατανέμονται στο διάγραμμα. Η εμφάνιση συστηματικών ή μη τυχαίων προτύπων, όπως συνεχόμενες αυξήσεις ή μειώσεις, μεγάλος αριθμός σημείων από τη μία πλευρά της κεντρικής γραμμής ή περιοδική συμπεριφορά, μπορεί επίσης να αποτελεί ένδειξη ότι η διεργασία δεν βρίσκεται υπό στατιστικό έλεγχο.

Για την κατασκευή των περισσότερων διαγραμμάτων ελέγχου χρησιμοποιείται συνήθως το μοντέλο ορίων $L\sigma$, σύμφωνα με το οποίο τα όρια ελέγχου δίνονται από τις σχέσεις

$$UCL = \mu_W + L\sigma_W$$

$$CL = \mu_W$$

$$LCL = \mu_W - L\sigma_W$$

όπου μ_W και σ_W είναι η μέση τιμή και η τυπική απόκλιση της στατιστικής συνάρτησης W , αντίστοιχα. Η ποσότητα L εκφράζει την απόσταση των ορίων ελέγχου από την κεντρική γραμμή σε μονάδες τυπικής απόκλισης.

Στην πράξη, ιδιαίτερα συνηθισμένη είναι η επιλογή $L = 3$, οπότε προκύπτουν τα γνωστά όρια ελέγχου τριών σίγμα (3σ limits). Η επιλογή της παραμέτρου L επηρεάζει σημαντικά την απόδοση του διαγράμματος ελέγχου. Μεγάλες τιμές του L αυξάνουν την απόσταση των ορίων ελέγχου από την κεντρική γραμμή και μειώνουν την πιθανότητα εμφάνισης ψευδών συναγερμών. Ωστόσο, στην περίπτωση αυτή αυξάνεται και ο χρόνος που απαιτείται για την ανίχνευση πραγματικών μετατοπίσεων της διεργασίας. Αντίθετα, μικρότερες τιμές του L οδηγούν σε μεγαλύτερη ευαισθησία του διαγράμματος, αλλά αυξάνουν την πιθανότητα λανθασμένων ενδείξεων εκτός ελέγχου.

Τα διαγράμματα ελέγχου αποτελούν ένα από τα σημαντικότερα εργαλεία του Στατιστικού Ελέγχου Διεργασιών, καθώς επιτρέπουν την έγκαιρη ανίχνευση μεταβολών στη συμπεριφορά της διεργασίας και συμβάλλουν στη διατήρηση της ποιότητας των παραγόμενων προϊόντων.

1.3 Διαγράμματα ελέγχου Φάσης I και Φάσης II

Ανάλογα με τη χρονική περίοδο εφαρμογής ενός διαγράμματος ελέγχου αυτά χωρίζονται σε διαγράμματα ελέγχου Φάσης I και Φάσης II.

Φάση I: Σε αυτή τη φάση χρησιμοποιούνται δείγματα που συλλέχθηκαν σε παρελθόντα χρόνο. Σκοπός είναι να ελεγχθεί αν η διεργασία ήταν εντός ελέγχου και αν δεν ήταν να γίνουν τα κατάλληλα βήματα ώστε ο διαχειριστής να τη φέρει εντός στατιστικού ελέγχου. Το αποτέλεσμα αυτής της φάσης είναι να προκύψουν η κεντρική γραμμή και τα όρια ελέγχου τα οποία θα χρησιμοποιηθούν για την παρακολούθηση της διεργασίας στο μέλλον. Αυτή η χρήση των διαγραμμάτων αναφέρεται και ως αναδρομική (retrospective). Πολλές δράσεις και σκέψεις εξελίσσονται στη διάρκεια αυτής της φάσης εκτός από την απλή απεικόνιση κάποιων δεδομένων σε ένα διάγραμμα. Αυτό γιατί ο διαχειριστής πρέπει να μελετήσει σε βάθος την διεργασία έως ότου αποφασίσει αν αυτή βρισκόταν εντός ή εκτός στατιστικού ελέγχου την περίοδο που συλλέχθηκαν τα δεδομένα. Στη συνέχεια τα διαγράμματα μπορούν να χρησιμοποιηθούν για την παρακολούθηση της μελλοντικής συμπεριφοράς της διεργασίας (Φάση II).

Φάση II: Σε αυτή τη φάση χρησιμοποιούνται η κεντρική γραμμή και τα όρια ελέγχου που προέκυψαν από τη Φάση I. Τα διαγράμματα ελέγχου χρησιμοποιούνται για την συνεχή παρακολούθηση της παραγωγικής διαδικασίας και γίνεται έλεγχος κάθε στιγμή αν υπάρχει κάποια αλλαγή στο μέσο επίπεδο των χαρακτηριστικών που μελετάμε που θα μπορούσε να επηρεάσει την ποιότητα του προϊόντος. Για κάθε δείγμα που συλλέγεται ο διαχειριστής έχει την δυνατότητα να καταλάβει εάν η διεργασία παραμένει εντός στατιστικού ελέγχου.

1.4 Εκτίμηση της μέσης τιμής και της διασποράς

1.4.1 Η περίπτωση των μεμονωμένων παρατηρήσεων

Έστω $X = (X_1, X_2, \dots, X_n)$ είναι ένα τυχαίο δείγμα μεγέθους n από ένα πληθυσμό με άγνωστη μέση τιμή μ και άγνωστη διακύμανση σ^2 . Μια αμερόληπτη εκτιμήτρια της μέσης τιμής είναι ο δειγματικός μέσος:

$$\hat{\mu} = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

και μια αμερόληπτη εκτιμήτρια της διακύμανσης είναι η δειγματική διακύμανση

$$\sigma^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Αν υποθέσουμε τώρα ότι το τυχαίο δείγμα προέρχεται από κανονικό πληθυσμό $N(\mu, \sigma^2)$, μπορούμε να βρούμε τις ακόλουθες αμερόληπτες εκτιμήτριες της τυπικής απόκλισης σ ,

$$\hat{\sigma} = \frac{R}{d_2} = \frac{X_{(n)} - X_{(1)}}{d_2}$$

και

$$\hat{\sigma} = \frac{S}{c_4} = \frac{\sqrt{S^2}}{c_4}.$$

Επίσης, για κανονικούς πληθυσμούς ισχύουν τα ακόλουθα:

$$\mu_R = E(R) = \sigma d_2, \quad \sigma_R = \sqrt{V(R)} = \sigma d_3,$$

$$\mu_S = E(S) = \sigma c_4, \quad \sigma_S = \sqrt{V(S)} = \sigma \sqrt{(1 - c_4^2)}.$$

Οι ποσότητες d_2, d_3 και c_4 είναι σταθερές που εξαρτώνται από το μέγεθος του δείγματος n . Στον ακόλουθο πίνακα δίνονται οι τιμές των σταθερών d_2, d_3 και c_4 για μεγέθη δειγμάτων $n = 2, 3, \dots, 10$.

Πίνακας 1.1. Τιμές των σταθερών d_2, d_3 και c_4

n	d_2	d_3	c_4
2	1.12837917	0.85250247	0.79788456
3	1.69256875	0.88836800	0.88622693
4	2.05875075	0.87980820	0.92131773
5	2.32592895	0.86408194	0.93998560
6	2.53441272	0.84803969	0.95153286
7	2.70435675	0.83320534	0.95936879
8	2.84720061	0.81983110	0.96503046
9	2.97002632	0.80783427	0.96931070
10	3.07750546	0.79705067	0.97265927

1.4.2 Η περίπτωση των δειγμάτων

Σε αυτή την περίπτωση έχουμε στη διάθεση μας m τυχαία δείγματα μεγέθους n το καθένα από έναν κανονικό πληθυσμό $X \sim N(\mu, \sigma^2)$ με άγνωστη μέση τιμή και διακύμανση, έστω τα $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$, για $1 \leq i \leq m$.

- **Εκτίμηση της μέσης τιμής**

Έστω $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m$ οι δειγματικοί μέσοι των m δειγμάτων. Θέτουμε

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

Η ποσότητα $\bar{\bar{X}}$ ακολουθεί κατανομή $N(\mu, \frac{\sigma^2}{mn})$ και χρησιμοποιείται ως εκτιμήτρια της ποσότητας μ (αμερόληπτη και συνεπής εκτιμήτρια του μ), δηλαδή $\hat{\mu} = \bar{\bar{X}}$.

- **Εκτίμηση της τυπικής απόκλισης σ (μέθοδος R)**

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

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

Θέτουμε

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

Μια αμερόληπτη εκτιμήτρια του σ είναι η στατιστική συνάρτηση

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

- **Εκτίμηση της τυπικής απόκλισης σ (Μέθοδος S)**

Έστω S_i η δειγματική τυπική απόκλιση του i δείγματος, δηλαδή

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

Θέτουμε

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

Μια αμερόληπτη εκτιμήτρια του σ είναι η στατιστική συνάρτηση

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

- **Εκτίμηση της τυπικής απόκλισης σ (Μέθοδος S^2)**

Έστω $S_1^2, S_2^2, \dots, S_m^2$ οι δειγματικές διασπορές των m δειγμάτων, δηλαδή

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

Θέτουμε

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

Μια εκτιμήτρια του σ είναι η στατιστική συνάρτηση

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

η οποία αν και δεν είναι αμερόληπτη, χρησιμοποιείται αρκετές φορές (λόγω του ότι έχει μικρότερο μέσο τετραγωνικό σφάλμα από την $\bar{S}/c_4$).

1.5 Διαγράμματα ελέγχου Shewhart για τη μέση τιμή

Στην παρούσα ενότητα παρουσιάζονται τα διαγράμματα ελέγχου Shewhart που χρησιμοποιούνται για την παρακολούθηση της μέσης τιμής μιας διεργασίας, τόσο στη Φάση II όσο και στη Φάση I. Τα διαγράμματα αυτά βασίζονται στις δειγματικές μέσες τιμές και χρησιμοποιούνται για την ανίχνευση πιθανών μετατοπίσεων του μέσου της διεργασίας.

1.5.1 Διάγραμμα ελέγχου Φάσης II

Έστω ότι το χαρακτηριστικό X των προϊόντων μιας διεργασίας ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ με γνωστές παραμέτρους. Θεωρούμε ανεξάρτητα τυχαία δείγματα μεγέθους n , $X_i = (X_{i1}, X_{i2}, \dots, X_{in}), i \geq 1$, και ορίζουμε τη δειγματική μέση τιμή

$$\bar{X}_i = \frac{X_{i1} + X_{i2} + \dots + X_{in}}{n}, \quad i = 1, 2, \dots$$

Το διάγραμμα ελέγχου που χρησιμοποιείται για την παρακολούθηση της μέσης τιμής της διεργασίας είναι το $\bar{X}$ διάγραμμα ελέγχου, στο οποίο απεικονίζονται οι δειγματικές μέσες τιμές $\bar{X}_i$. Στην περίπτωση ορίων ελέγχου 3σ , τα όρια ελέγχου και η κεντρική γραμμή δίνονται από τις σχέσεις

$$UCL = \mu + A\sigma,$$

$$CL = \mu,$$

$$LCL = \mu - A\sigma,$$

$$\text{όπου } A = \frac{3}{\sqrt{n}}.$$

Η διεργασία θεωρείται εντός ελέγχου όταν οι τιμές των $\bar{X}_i$ βρίσκονται εντός των ορίων ελέγχου και δεν παρουσιάζουν κάποια συστηματική ή μη τυχαία συμπεριφορά.

1.5.2 Διαγράμματα ελέγχου Φάσης I

Στη Φάση I οι παράμετροι μ και σ είναι άγνωστες και επομένως πρέπει να εκτιμηθούν από προκαταρκτικά δείγματα. Για τον σκοπό αυτό λαμβάνονται m ανεξάρτητα δείγματα μεγέθους n , τα $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in}), i = 1, 2, \dots, m$ και υποθέτουμε ότι τα δείγματα αυτά προέρχονται από διεργασία που βρίσκεται υπό στατιστικό έλεγχο.

Η μέση τιμή της διεργασίας εκτιμάται μέσω της ποσότητας

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

ενώ για την εκτίμηση της τυπικής απόκλισης χρησιμοποιούνται συνήθως οι ακόλουθες εκτιμήτριες:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}, \quad \hat{\sigma} = \frac{\bar{S}}{c_4}, \quad \hat{\sigma} = \sqrt{\bar{S}^2}.$$

Ανάλογα με τη μέθοδο εκτίμησης της τυπικής απόκλισης προκύπτουν διαφορετικά $\bar{X}$ διαγράμματα ελέγχου.

Με τη μέθοδο R , τα όρια ελέγχου δίνονται από τις σχέσεις

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

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

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

Με τη μέθοδο S , τα όρια ελέγχου είναι

$$\begin{aligned}UCL &= \bar{\bar{X}} + A_3\bar{S}, \\CL &= \bar{\bar{X}}, \\LCL &= \bar{\bar{X}} - A_3\bar{S}.\end{aligned}$$

Τέλος, όταν χρησιμοποιείται η δειγματική διασπορά S^2 , προκύπτουν τα όρια

$$\begin{aligned}UCL &= \bar{\bar{X}} + A\sqrt{\bar{S}^2}, \\CL &= \bar{\bar{X}}, \\LCL &= \bar{\bar{X}} - A\sqrt{\bar{S}^2}.\end{aligned}$$

Στις παραπάνω σχέσεις ισχύει ότι

$$A_2 = \frac{3}{d_2\sqrt{n}}, \quad A_3 = \frac{3}{c_4\sqrt{n}}, \quad A = \frac{3}{\sqrt{n}}.$$

Τα παραπάνω διαγράμματα χρησιμοποιούνται για την παρακολούθηση πιθανών μεταβολών της μέσης τιμής της διεργασίας και αποτελούν τα βασικότερα διαγράμματα ελέγχου τύπου Shewhart για ποσοτικά χαρακτηριστικά.

- **Εφαρμογή**

Στον ακόλουθο πίνακα δίνονται οι μετρήσεις ενός ποιοτικού χαρακτηριστικού που περιγράφεται από συνεχή τυχαία μεταβλητή και αφορά την εσωτερική διάμετρο μιας στεφάνης εμβόλου. Τα δεδομένα προέκυψαν από τη λήψη 10 τυχαίων δειγμάτων μεγέθους $n = 2$ από την παραγωγική διεργασία.

Πίνακας 1.2. Δεδομένα για τη διάμετρο της στεφάνης του εμβόλου

Δείγμα	1 ^η Μέτρηση	2 ^η Μέτρηση	Μέση τιμή
1	9.84422	9.62656	9.73539
2	9.80879	9.93767	9.87323
3	10.50880	9.37680	9.94280
4	9.94629	10.2645	10.10540
5	9.55296	10.09280	9.82288
6	9.58023	9.71789	9.64906
7	9.40171	10.15210	9.77691
8	9.59285	9.95854	9.77570
9	9.54142	9.62176	9.58159
10	10.66530	10.23660	10.45095

Έστω ότι, όταν η διεργασία βρίσκεται υπό συνθήκες φυσικής μεταβλητότητας, το χαρακτηριστικό X ακολουθεί κανονική κατανομή με μέση τιμή $\mu = 10$ και τυπική απόκλιση $\sigma = 0.25$.

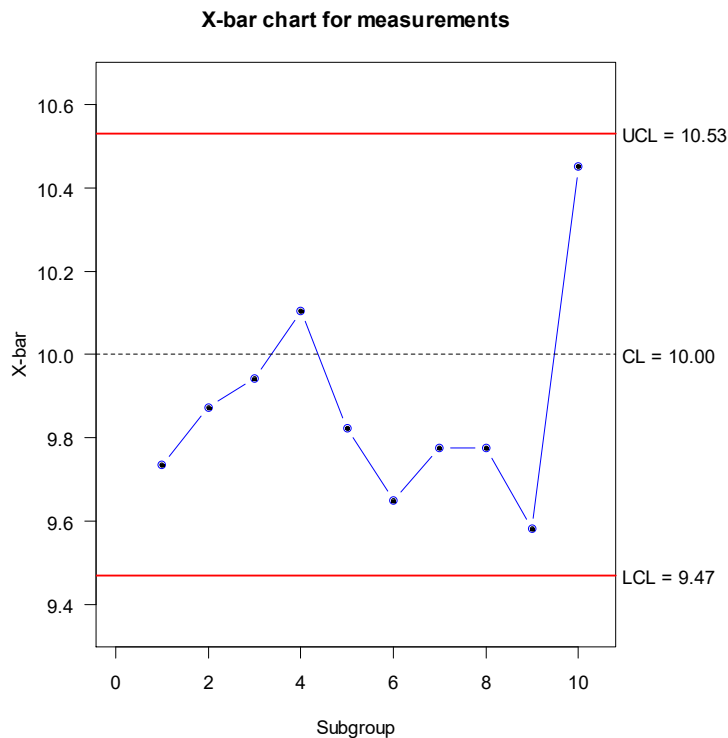
Για το $\bar{X}$ διάγραμμα ελέγχου Φάσης II με όρια ελέγχου 3σ , έχουμε:

$$UCL = \mu_W + 3\sigma_W = 10 + 3 \frac{0.25}{\sqrt{2}} = 10.5303$$

$$CL = \mu_W = 10$$

$$LCL = \mu_W - 3\sigma_W = 10 - 3 \frac{0.25}{\sqrt{2}} = 9.4697$$

Στο αντίστοιχο $\bar{X}$ διάγραμμα ελέγχου απεικονίζονται οι δειγματικές μέσες τιμές των 10 δειγμάτων, μαζί με την κεντρική γραμμή και τα όρια ελέγχου.



Σχήμα 1.2: $\bar{X}$ διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.2

Παρατηρούμε ότι όλες οι δειγματικές μέσες τιμές βρίσκονται εντός των ορίων ελέγχου. Επομένως, δεν προκύπτει ένδειξη μετατόπισης της μέσης τιμής της διεργασίας και μπορούμε να θεωρήσουμε ότι η διεργασία παραμένει εντός στατιστικού ελέγχου.

1.6 Διαγράμματα ελέγχου Shewhart για τη διασπορά

Σε αυτή την ενότητα παρουσιάζονται τα διαγράμματα ελέγχου τύπου Shewhart για τη διασπορά στην περίπτωση των δειγμάτων, για τη Φάση I και για τη Φάση II.

1.6.1 Διάγραμμα ελέγχου Φάσης II

Έστω ότι το χαρακτηριστικό X των παραγόμενων προϊόντων ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ με γνωστές παραμέτρους και ότι $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in}), i \geq 1$ είναι τυχαίο δείγμα μεγέθους n από τη διεργασία. Για την παρακολούθηση της διασποράς της διεργασίας έχουν προταθεί διαφορετικά διαγράμματα ελέγχου, ανάλογα με τη στατιστική συνάρτηση που χρησιμοποιείται κάθε φορά.

Αρχικά θέτουμε

$$W_i = R_i = X_{i(n)} - X_{i(1)}, \quad i = 1, 2, \dots$$

Είναι γνωστό ότι $E(R_i) = \sigma d_2$ και $\sqrt{V(R_i)} = \sigma d_3$ όπου οι σταθερές d_2 και d_3 εξαρτώνται αποκλειστικά από το μέγεθος του δείγματος n . Το διάγραμμα ελέγχου που χρησιμοποιείται για την παρακολούθηση της διασποράς μέσω των δειγματικών εκτάσεων είναι το διάγραμμα R , στο οποίο απεικονίζονται οι τιμές των R_i . Για όρια ελέγχου 3σ , τα όρια του διαγράμματος δίνονται από τις σχέσεις

$$UCL = (d_2 + 3d_3)\sigma,$$

$$CL = d_2\sigma,$$

$$LCL = (d_2 - 3d_3)\sigma.$$

Σε περιπτώσεις όπου το κάτω όριο ελέγχου είναι αρνητικό, λαμβάνεται ίσο με μηδέν.

Ένα εναλλακτικό διάγραμμα ελέγχου βασίζεται στη δειγματική τυπική απόκλιση

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

Για τη στατιστική αυτή ισχύει $E(S_i) = \sigma c_4$ και $\sqrt{V(S_i)} = \sigma \sqrt{1 - c_4^2}$ όπου η σταθερά c_4 εξαρτάται επίσης από το μέγεθος του δείγματος. Το αντίστοιχο διάγραμμα ελέγχου είναι το διάγραμμα S , στο οποίο απεικονίζονται οι τιμές των δειγματικών τυπικών αποκλίσεων S_i . Τα όρια ελέγχου 3σ δίνονται από τους τύπους

$$UCL = (c_4 + 3\sqrt{1 - c_4^2}) \sigma,$$

$$CL = c_4 \sigma,$$

$$LCL = (c_4 - 3\sqrt{1 - c_4^2}) \sigma.$$

Αν το κάτω όριο ελέγχου προκύπτει αρνητικό, θεωρείται ίσο με μηδέν.

Τέλος, για την παρακολούθηση της διασποράς μπορεί να χρησιμοποιηθεί και η δειγματική διασπορά

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

Γνωρίζουμε ότι $E(S_i^2) = \sigma^2$ και ότι η ποσότητα $\frac{(n-1)S_i^2}{\sigma^2}$ ακολουθεί κατανομή χ^2 με $n - 1$ βαθμούς ελευθερίας. Το διάγραμμα ελέγχου S^2 χρησιμοποιεί όρια πιθανότητας και τα όρια ελέγχου δίνονται από τους τύπους

$$UCL = \frac{\sigma^2}{n-1} \chi_{n-1; \alpha/2}^2,$$

$$CL = \sigma^2,$$

$$LCL = \frac{\sigma^2}{n-1} \chi_{n-1; 1-\alpha/2}^2,$$

όπου $\chi_{n; \alpha}^2$ συμβολίζει το άνω ποσοστιαίο σημείο της κατανομής χ^2 με $n - 1$ βαθμούς ελευθερίας.

1.6.2 Διάγραμμα ελέγχου Φάσης I

Εδώ η τυπική απόκλιση είναι άγνωστη επομένως πρέπει να εκτιμηθεί. Για τον σκοπό αυτό λαμβάνονται m ανεξάρτητα τυχαία δείγματα μεγέθους n , $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $i = 1, 2, \dots, m$ τα οποία υποτίθεται ότι προέρχονται από διεργασία που βρίσκεται εντός στατιστικού ελέγχου.

Για την εκτίμηση της τυπικής απόκλισης χρησιμοποιούνται συνήθως οι ακόλουθες εκτιμήτριες:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}, \quad \hat{\sigma} = \frac{\bar{S}}{c_4}, \quad \hat{\sigma} = \sqrt{\bar{S}^2}.$$

Ανάλογα με την εκτιμήτρια που χρησιμοποιείται και τη στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα, προκύπτουν διαφορετικά διαγράμματα ελέγχου για την παρακολούθηση της διασποράς της διεργασίας. Στην περίπτωση του διαγράμματος R , στο διάγραμμα απεικονίζονται οι δειγματικές εκτάσεις R_i , ενώ τα όρια ελέγχου δίνονται από τους τύπους

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

$$CL = \bar{R},$$

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

όπου

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

Αντίστοιχα, στο διάγραμμα S απεικονίζονται οι δειγματικές τυπικές αποκλίσεις S_i , με όρια ελέγχου

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

$$CL = \bar{S},$$

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

όπου

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

Τέλος, στο διάγραμμα S^2 απεικονίζονται οι δειγματικές διασπορές S_i^2 . Τα όρια ελέγχου πιθανότητας δίνονται από

$$UCL = \frac{\bar{S}^2}{n-1} \chi_{n-1; \alpha/2}^2,$$

$$CL = \bar{S}^2,$$

$$LCL = \frac{\bar{S}^2}{n-1} \chi_{n-1; 1-\alpha/2}^2,$$

όπου $\chi_{n; \alpha}^2$ συμβολίζει το άνω ποσοστιαίο σημείο της κατανομής χ^2 με $n-1$ βαθμούς ελευθερίας.

1.7 Διαγράμματα ελέγχου με μνήμη

Στα διαγράμματα ελέγχου που παρουσιάστηκαν στις προηγούμενες ενότητες, οι τιμές που απεικονίζονται στο διάγραμμα βασίζονται αποκλειστικά στις παρατηρήσεις του τρέχοντος δείγματος, χωρίς να χρησιμοποιούνται πληροφορίες από προηγούμενα δείγματα. Για τον λόγο αυτό τα διαγράμματα ελέγχου τύπου Shewhart χαρακτηρίζονται ως διαγράμματα ελέγχου χωρίς μνήμη. Τα διαγράμματα αυτά είναι ιδιαίτερα αποτελεσματικά στην ανίχνευση μεγάλων μετατοπίσεων της μέσης τιμής της διεργασίας.

Στην περίπτωση όπου στόχος είναι η ανίχνευση μικρότερων μετατοπίσεων, χρησιμοποιούνται τα διαγράμματα ελέγχου με μνήμη. Στην κατηγορία αυτή ανήκουν κυρίως τα διαγράμματα τύπου CUSUM και EWMA καθώς και τα διαγράμματα κινούμενου μέσου (Moving Average – MA). Τα διαγράμματα αυτά χρησιμοποιούνται κυρίως στη Φάση II, όταν δηλαδή οι παράμετροι της διεργασίας είναι γνωστές ή έχουν ήδη εκτιμηθεί από προηγούμενα δεδομένα.

Ο χαρακτηρισμός «με μνήμη» οφείλεται στο γεγονός ότι κάθε σημείο του διαγράμματος δεν εξαρτάται μόνο από το τρέχον δείγμα αλλά και από πληροφορίες προηγούμενων δειγμάτων. Με τον τρόπο αυτό επιτυγχάνεται καλύτερη απόδοση στην ανίχνευση μικρών και σταδιακών μετατοπίσεων της διεργασίας σε σύγκριση με τα κλασικά διαγράμματα Shewhart.

Τα τελευταία χρόνια τα διαγράμματα ελέγχου με μνήμη έχουν μελετηθεί εκτενώς στη βιβλιογραφία, καθώς παρουσιάζουν σημαντικά πλεονεκτήματα στην παρακολούθηση της παραγωγικής διαδικασίας. Ιδιαίτερα, εμφανίζουν αυξημένη ευαισθησία στην ανίχνευση μικρών μεταβολών της μέσης τιμής, οι οποίες συχνά δεν μπορούν να εντοπιστούν έγκαιρα από τα διαγράμματα χωρίς μνήμη.

Στις επόμενες ενότητες παρουσιάζονται συνοπτικά τα βασικά χαρακτηριστικά των διαγραμμάτων ελέγχου τύπου CUSUM και EWMA, τα οποία αποτελούν βασικό αντικείμενο της παρούσας εργασίας.

1.7.1 Διάγραμμα ελέγχου CUSUM

Έστω μία διεργασία με εντός ελέγχου μέση τιμή μ_0 και τυπική απόκλιση σ . Από τη διεργασία λαμβάνονται διαδοχικές παρατηρήσεις X_t , $t \geq 1$, και στόχος είναι η ανίχνευση

μετατοπίσεων της μέσης τιμής της μορφής $\mu_1 = \mu_0 \pm \delta\sigma$, $\delta > 0$. Στα διαγράμματα ελέγχου τύπου CUSUM, τα οποία εισήχθησαν από τον Page (1954), χρησιμοποιούνται οι ακόλουθες σωρευτικές στατιστικές συναρτήσεις:

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

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

Οι ποσότητες S_0^+ και S_0^- ονομάζονται τιμές εκκίνησης (headstart values), ενώ η ποσότητα $K \geq 0$ ονομάζεται τιμή αναφοράς (reference value) και δίνεται συνήθως από τη σχέση

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

Η στατιστική συνάρτηση S_t^+ χρησιμοποιείται για την ανίχνευση ανοδικών μετατοπίσεων της μέσης τιμής της διεργασίας, ενώ η S_t^- για την ανίχνευση καθοδικών μετατοπίσεων. Μεγάλες θετικές τιμές της S_t^+ αποτελούν ένδειξη αύξησης της μέσης τιμής, ενώ μεγάλες αρνητικές τιμές της S_t^- υποδηλώνουν μείωση της μέσης τιμής της διεργασίας.

Η απόφαση για το αν η διεργασία βρίσκεται εκτός ελέγχου βασίζεται στη σύγκριση των παραπάνω ποσοτήτων με ένα προκαθορισμένο όριο H , το οποίο ονομάζεται διάστημα απόφασης (decision interval) και δίνεται από τη σχέση

$$H = h\sigma.$$

Σήμα εκτός ελέγχου διεργασίας παίρνουμε όταν

$$S_t^+ > H \quad \text{ή} \quad S_t^- < -H$$

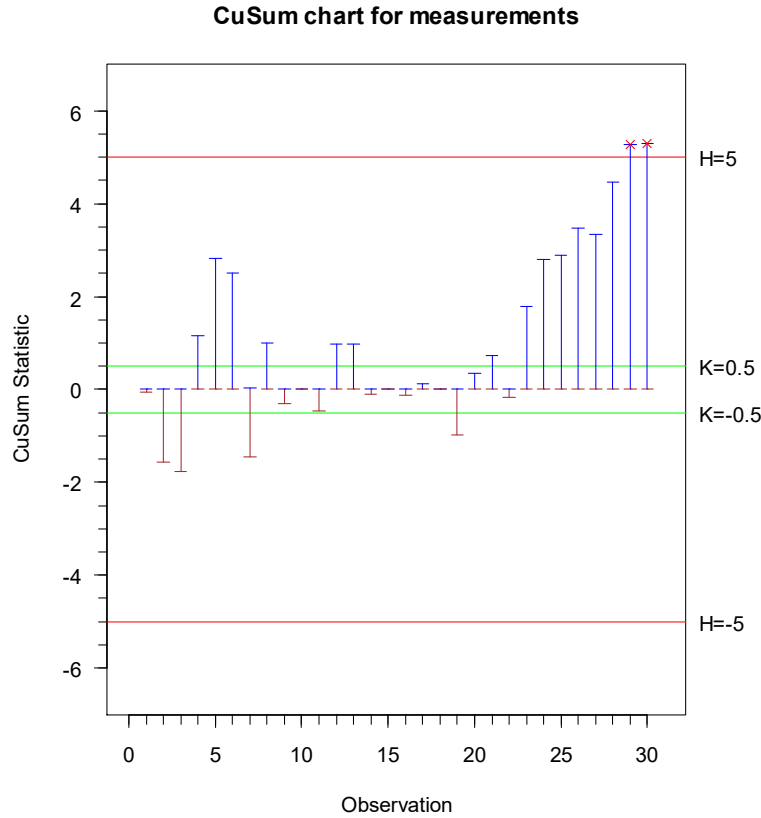
οπότε θεωρούμε ότι έχει πραγματοποιηθεί μετατόπιση της μέσης τιμής της διεργασίας σε υψηλότερο ή χαμηλότερο επίπεδο αντίστοιχα.

Για την επίδειξη του διαγράμματος ελέγχου τύπου CUSUM χρησιμοποιούνται 30 παρατηρήσεις από κανονική κατανομή (βλ. Montgomery (2013)). Οι πρώτες 20 παρατηρήσεις προέρχονται από την κατανομή $N(10,1)$, ενώ οι τελευταίες 10 από την κατανομή $N(11,1)$, ώστε να προσομοιωθεί μετατόπιση της μέσης τιμής της διεργασίας. Στον Πίνακα 1.2 παρουσιάζονται τα δεδομένα και οι αντίστοιχες τιμές των ποσοτήτων S_t^+ και S_t^- , για $K = 0.5$ και $H = 5$ (δηλαδή $k = 0.5$ και $h = 5$).

Πίνακας 1.3: Δεδομένα για την επίδειξη διαγράμματος CUSUM: $k = 5, h = 5$

Δείγμα	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 τα δεδομένα του Πίνακα 1.3 προκύπτει το ακόλουθο διάγραμμα ελέγχου CUSUM.



Σχήμα 1.3: CUSUM διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.3

Από το παραπάνω σχήμα παρατηρούμε ότι

$$S_{29}^+ > 5$$

και συνεπώς η διεργασία θεωρείται εκτός ελέγχου λόγω μετατόπισης της μέσης τιμής της σε υψηλότερο επίπεδο.

Στην περίπτωση όπου χρησιμοποιούνται δείγματα μεγέθους $n > 1$, αντί των μεμονωμένων παρατηρήσεων X_t χρησιμοποιούνται οι δειγματικές μέσες τιμές $\bar{X}_t$, ενώ η ποσότητα σ αντικαθίσταται από $\sigma/\sqrt{n}$. Έτσι, οι αντίστοιχες στατιστικές συναρτήσεις γράφονται ως

$$S_t^+ = \max(0, \bar{X}_t - (\mu_0 + K) + S_{t-1}^+), \quad S_0^+ \geq 0,$$

$$S_t^- = \min(0, \bar{X}_t - (\mu_0 - K) + S_{t-1}^-), \quad S_0^- \leq 0.$$

όπου

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

και $k, h > 0$ είναι σταθερές σχεδιασμού του διαγράμματος.

1.7.2 Διάγραμμα ελέγχου EWMA

Έστω μία διεργασία με εντός ελέγχου μέση τιμή μ_0 και τυπική απόκλιση σ . Από τη διεργασία λαμβάνονται διαδοχικές μεμονωμένες παρατηρήσεις X_t , $t \geq 1$, και στόχος είναι η ανίχνευση μετατοπίσεων της μέσης τιμής της μορφής $\mu_1 = \mu_0 \pm \delta\sigma$, $\delta > 0$. Στα διαγράμματα ελέγχου εκθετικά σταθμισμένου κινούμενου μέσου, δηλαδή στα διαγράμματα EWMA (Exponentially Weighted Moving Average), τα οποία εισήχθησαν από τον Roberts (1959), απεικονίζεται η στατιστική συνάρτηση

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

Για την εφαρμογή του διαγράμματος απαιτείται ο καθορισμός της αρχικής τιμής Z_0 , η οποία συνήθως λαμβάνεται ίση με την εντός ελέγχου μέση τιμή, δηλαδή $Z_0 = \mu_0$. Εφαρμόζοντας αναδρομικά τον παραπάνω τύπο, προκύπτει ότι

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

Η μέση τιμή και η διακύμανση της στατιστικής συνάρτησης Z_t είναι αντίστοιχα

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

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

Συνεπώς, σύμφωνα με τη λογική κατασκευής των διαγραμμάτων τύπου Shewhart, στο διάγραμμα ελέγχου EWMA απεικονίζονται οι τιμές της Z_t , με όρια ελέγχου

$$\begin{aligned} UCL &= \mu_0 + L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2t}]}, \\ CL &= \mu_0, \\ LCL &= \mu_0 - L\sigma \sqrt{\frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2t}]}. \end{aligned}$$

Παρατηρούμε ότι τα παραπάνω όρια ελέγχου είναι μεταβλητά ως προς τον χρόνο t . Ωστόσο, καθώς το t αυξάνεται, η ποσότητα $(1-\lambda)^{2t}$ τείνει σχετικά γρήγορα στο μηδέν, με αποτέλεσμα τα όρια ελέγχου να σταθεροποιούνται στις τιμές

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

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

Αν και τα διαγράμματα EWMA αναπτύχθηκαν αρχικά για μεμονωμένες παρατηρήσεις, μπορούν να προσαρμοστούν άμεσα και στην περίπτωση δειγμάτων μεγέθους n . Σε αυτή την περίπτωση η ποσότητα X_t αντικαθίσταται από τη δειγματική μέση τιμή $\bar{X}_t$, ενώ η τυπική απόκλιση σ αντικαθίσταται από $\sigma/\sqrt{n}$. Έτσι, η στατιστική συνάρτηση γίνεται

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

με όρια ελέγχου

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

$$CL = \mu_0,$$

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

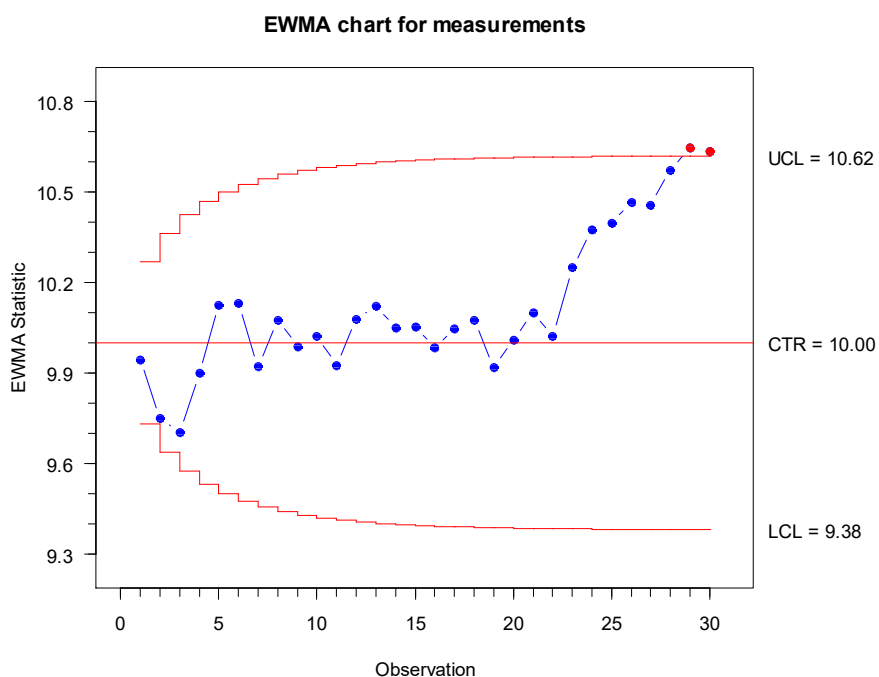
Τα διαγράμματα ελέγχου τύπου EWMA χρησιμοποιούνται, όπως και τα διαγράμματα τύπου CUSUM, για την ανίχνευση μικρών μετατοπίσεων της μέσης τιμής της διεργασίας. Ένα βασικό πλεονέκτημά τους είναι ότι εμφανίζουν μικρότερη ευαισθησία σε αποκλίσεις από την κανονικότητα των παρατηρήσεων, γεγονός που τα καθιστά ιδιαίτερα χρήσιμα στην περίπτωση μεμονωμένων παρατηρήσεων.

Για την επίδειξη ενός διαγράμματος EWMA χρησιμοποιούνται τα δεδομένα του Πίνακα 1.3, με $Z_0 = \mu_0 = 10$, $\sigma = 1$, $\lambda = 0.1$ και $L = 2.7$. Στον Πίνακα 1.4 δίνονται οι τιμές της στατιστικής συνάρτησης Z_t , οι οποίες απεικονίζονται στο αντίστοιχο διάγραμμα ελέγχου EWMA.

Πίνακας 1.4. Δεδομένα για την επίδειξη διαγράμματος EWMA: $\lambda = 0.1, L = 2.7$

Δείγμα	X_t	Z_t	Δείγμα	X_t	Z_t
1	9.45	9.45	16	9.37	9.98426
2	7.99	9.495	17	10.62	10.0478
3	9.29	9.0355	18	10.31	10.074
4	11.66	9.992	19	8.52	9.9186
5	12.16	10.253	20	10.84	10.0108
6	10.18	10.307	21	10.9	10.0997
7	8.04	9.2167	22	9.33	10.0227
8	11.46	10.755	23	12.29	10.2495
9	9.2	9.8796	24	11.5	10.3745
10	10.34	10.232	25	10.6	10.3971
11	9.03	9.2384	26	11.08	10.4654
12	11.47	10.9785	27	10.38	10.4568
13	10.51	10.1216	28	11.62	10.5731
14	9.4	10.0495	29	11.31	10.6468
15	10.08	10.0525	30	10.52	10.6341

Το διάγραμμα ελέγχου EWMA δίνεται στο ακόλουθο σχήμα.



Σχήμα 1.4. EWMA διάγραμμα ελέγχου Φάσης II για τα δεδομένα του Πίνακα 1.4

Στο συγκεκριμένο διάγραμμα παρατηρούμε ότι μετά το 23ο δείγμα εμφανίζεται σταδιακή μετατόπιση της μέσης τιμής της διεργασίας σε υψηλότερο επίπεδο, η οποία

γίνεται αντιληπτή στο 29ο δείγμα, όπου προκύπτει ένδειξη εκτός ελέγχου διεργασίας. Σημειώνεται ότι τα σταθεροποιημένα όρια ελέγχου, δηλαδή τα όρια για $t \rightarrow \infty$, είναι

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

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

1.8 Μέσο μήκος ροής

Μία σημαντική έννοια που σχετίζεται με τα διαγράμματα ελέγχου είναι το μέσο μήκος ροής ARL (Average Run Length). Έστω T η τυχαία μεταβλητή που εκφράζει το πλήθος των σημείων ή δειγμάτων που πρέπει να σχεδιαστούν σε ένα διάγραμμα ελέγχου μέχρι να εμφανιστεί για πρώτη φορά ένδειξη ότι η διεργασία βρίσκεται εκτός ελέγχου. Η τυχαία μεταβλητή T ονομάζεται μήκος ροής (Run Length) του διαγράμματος. Το μέσο μήκος ροής ορίζεται ως η αναμενόμενη τιμή της τυχαίας μεταβλητής T , δηλαδή

$$ARL = E(T).$$

Η ποσότητα ARL εκφράζει τον αναμενόμενο αριθμό σημείων που πρέπει να απεικονιστούν στο διάγραμμα μέχρι να ληφθεί για πρώτη φορά ένδειξη ότι η διεργασία είναι εκτός ελέγχου.

Στα διαγράμματα ελέγχου τύπου Shewhart, έστω ότι η στατιστική συνάρτηση W που απεικονίζεται στο διάγραμμα έχει εντός ελέγχου μέση τιμή μ , τυπική απόκλιση σ και συνάρτηση κατανομής $F_0(\cdot)$. Η πιθανότητα α να εμφανιστεί ένα σημείο του διαγράμματος εκτός των ορίων ελέγχου όταν η διεργασία βρίσκεται υπό στατιστικό έλεγχο είναι

$$\alpha = 1 - P(LCL \leq W \leq UCL) = 1 - F_0(\mu + L\sigma) + F_0(\mu - L\sigma).$$

Το εντός ελέγχου μέσο μήκος ροής είναι

$$ARL_0 = \frac{1}{\alpha}$$

καθώς το μήκος ροής ακολουθεί γεωμετρική κατανομή με πιθανότητα επιτυχίας α . Για μία διεργασία που βρίσκεται υπό στατιστικό έλεγχο επιθυμούμε μεγάλες τιμές του ARL_0 ,

ώστε να μειώνεται η πιθανότητα εμφάνισης λανθασμένων ενδείξεων εκτός ελέγχου ή αλλιώς η πιθανότητα εμφάνισης ψευδών συναγερμών.

Στην περίπτωση που η διεργασία βρίσκεται εκτός ελέγχου και η συνάρτηση κατανομής της στατιστικής συνάρτησης W είναι η $F_1(\cdot)$, η πιθανότητα $1 - \beta$ να εμφανιστεί σημείο εκτός των ορίων ελέγχου είναι

$$1 - \beta = 1 - P(LCL \leq W \leq UCL) = 1 - F_1(\mu + L\sigma) + F_1(\mu - L\sigma).$$

Η πιθανότητα να βρεθεί ένα σημείο του διαγράμματος εντός των ορίων ελέγχου στην εκτός ελέγχου κατάσταση είναι ίση με β . Το εκτός ελέγχου μέσο μήκος ροής δίνεται από τη σχέση

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

Για μία διεργασία που βρίσκεται εκτός ελέγχου επιθυμούμε μικρές τιμές του ARL_1 , έτσι ώστε να μειώνεται ο αριθμός των δειγμάτων και συνεπώς ο χρόνος που απαιτείται μέχρι να εντοπιστεί η μεταβολή της διεργασίας.

Στην ειδική περίπτωση όπου η στατιστική συνάρτηση W που απεικονίζεται στο διάγραμμα ακολουθεί κανονική κατανομή, $W \sim N(\mu, \sigma^2)$, το εντός ελέγχου μέσο μήκος ροής δίνεται από τη σχέση

$$ARL_0 = \frac{1}{\alpha} = \frac{1}{2\Phi(-L)}.$$

Όταν χρησιμοποιούνται όρια ελέγχου 3σ , δηλαδή για $L = 3$, προκύπτει

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

Η τιμή αυτή σημαίνει ότι όταν η διεργασία βρίσκεται υπό στατιστικό έλεγχο, αναμένεται κατά μέσο όρο μία λανθασμένη ένδειξη εκτός ελέγχου ανά περίπου 370 σημεία του διαγράμματος.

Αν κατά τη διάρκεια της λειτουργίας της διεργασίας η μέση τιμή της στατιστικής συνάρτησης W μετατοπιστεί από την τιμή μ στην τιμή $\mu^* = \mu + \delta\sigma$, ενώ η διακύμανση παραμένει σταθερή, τότε το εκτός ελέγχου μέσο μήκος ροής δίνεται από τη σχέση

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{2 - \Phi(L - \delta) - \Phi(L + \delta)}.$$

Η παραπάνω σχέση δείχνει ότι όσο αυξάνεται το μέγεθος της μετατόπισης δ , τόσο μειώνεται η τιμή του ARL_1 , δηλαδή τόσο ταχύτερα το διάγραμμα ανιχνεύει την εκτός ελέγχου κατάσταση.

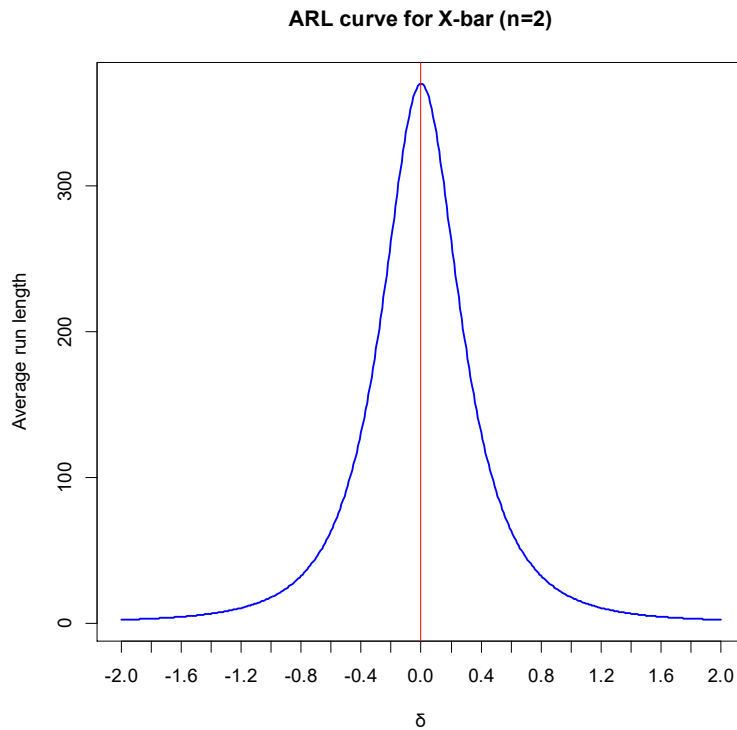
Στην περίπτωση του $\bar{X}$ διαγράμματος ελέγχου έχουμε, δηλαδή η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα είναι η $W = \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, έχουμε ότι

$$ARL_0 = \frac{1}{\alpha} = \frac{1}{2\Phi(-L)},$$

και αν έχουμε μετατόπιση της μέσης τιμής μ στην τιμή $\mu^* = \mu + \delta\sigma$,

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{2 - \Phi(L - \delta\sqrt{n}) - \Phi(L + \delta\sqrt{n})}.$$

Στο παρακάτω σχήμα παρουσιάζεται η καμπύλη του μέσου μήκους ροής για διαφορετικές μετατοπίσεις της μέσης τιμής της διεργασίας. Παρατηρούμε ότι η μέγιστη τιμή του ARL εμφανίζεται όταν η διεργασία βρίσκεται υπό στατιστικό έλεγχο, ενώ όσο αυξάνεται η μετατόπιση της μέσης τιμής, το ARL μειώνεται σημαντικά, γεγονός που σημαίνει ότι το διάγραμμα εντοπίζει ταχύτερα την αλλαγή στη συμπεριφορά της διεργασίας.



Σχήμα 1.5. Καμπύλη ARL του $\bar{X}$ διαγράμματος: $L = 3, n = 2$

Παρότι το *ARL* αποτελεί το πιο διαδεδομένο μέτρο αξιολόγησης της απόδοσης ενός διαγράμματος ελέγχου, η χρήση του έχει δεχθεί αρκετή κριτική τα τελευταία χρόνια. Αυτό συμβαίνει επειδή το μήκος ροής T ακολουθεί γεωμετρική κατανομή και συνεπώς η μέση τιμή της δεν είναι πάντοτε πλήρως αντιπροσωπευτική της πραγματικής συμπεριφοράς του διαγράμματος, ιδιαίτερα όταν η πιθανότητα εμφάνισης σήματος είναι μικρή και η διασπορά της κατανομής είναι μεγάλη.

ΚΕΦΑΛΑΙΟ 2

Διαγράμματα ελέγχου διαμέσου τύπου Shewhart

2.1 Εισαγωγή

Στο παρόν κεφάλαιο εξετάζεται το διάγραμμα ελέγχου της διαμέσου τύπου Shewhart, το οποίο αποτελεί μια εναλλακτική μορφή του κλασικού διαγράμματος ελέγχου μέσης τιμής ($\bar{X}$). Το διάγραμμα αυτό χρησιμοποιείται για την παρακολούθηση της κεντρικής τάσης μιας διεργασίας και βασίζεται στη δειγματική διάμεσο, η οποία, σε αντίθεση με τον μέσο, επηρεάζεται πολύ λιγότερο από ακραίες τιμές ή αποκλίσεις από την κανονικότητα. Για τον λόγο αυτό θεωρείται ιδιαίτερα χρήσιμο σε περιπτώσεις όπου τα δεδομένα προέρχονται από διεργασίες με ασυμμετρίες ή παρουσία θορύβου.

Η βασική ιδέα παραμένει ίδια με τα κλασικά διαγράμματα ελέγχου τύπου Shewhart, όπου για κάθε δείγμα που λαμβάνεται, υπολογίζεται μια στατιστική συνάρτηση και ελέγχεται αν η τιμή της βρίσκεται εντός των προκαθορισμένων ορίων ελέγχου. Αν κάποια τιμή υπερβεί τα όρια αυτά, θεωρούμε ότι η διεργασία έχει μεταβληθεί και απαιτείται έλεγχος των ειδικών αιτιών μεταβλητότητας.

Στο παρόν κεφάλαιο αρχικά γίνεται μια συνοπτική παρουσίαση θεωρητικών αποτελεσμάτων που αφορούν τη διάμεσο ενός δείγματος. Στη συνέχεια εισάγεται το παραμετρικό διάγραμμα ελέγχου διαμέσου όπου αρχικά μελετάται η Φάση II, στην οποία υποθέτουμε ότι οι παράμετροι της διεργασίας (η μέση τιμή και η τυπική απόκλιση) είναι γνωστές. Αυτή η προσέγγιση μάς επιτρέπει να κατανοήσουμε τη θεωρητική συμπεριφορά του διαγράμματος όταν όλα τα στοιχεία είναι υπό έλεγχο. Έπειτα, εξετάζεται η Φάση I, όπου οι παράμετροι εκτιμώνται από περιορισμένο πλήθος δειγμάτων, κάτι που εισάγει επιπλέον αβεβαιότητα στα όρια ελέγχου. Επίσης μελετάται

η απόδοση του διαγράμματος ελέγχου διαμέσου μέσω του μέσου μήκους ροής, τόσο για γνωστές όσο και για εκτιμημένες παραμέτρους. Τα αποτελέσματα που προκύπτουν συγκρίνονται με εκείνα του άρθρου των Castagliola & Figueiredo (2013), πάνω στο οποίο βασίζεται η παρούσα ανάλυση.

Στη συνέχεια παρουσιάζεται το μη παραμετρικό διάγραμμα ελέγχου διαμέσου για το οποίο γίνεται αντίστοιχη μελέτη. Τα κύρια άρθρα για αυτό το διάγραμμα είναι οι εργασίες των Chakraborti et al. (2001, 2004).

2.2 Η κατανομή της δειγματικής διαμέσου

Ας θεωρήσουμε $n \geq 2$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές $X_1, X_2, \dots, X_n$ με κοινή συνάρτηση κατανομής $F(x)$, και ας συμβολίσουμε με $X_{(k)}$ την k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$). Έτσι μπορούμε να γράψουμε ότι

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

Από τη θεωρία των διατεταγμένων τυχαίων μεταβλητών έχουμε ότι η συνάρτηση κατανομής $F_{X_{(k)}}(x) = P(X_{(k)} \leq x)$ της τυχαίας μεταβλητής $X_{(k)}$ δίνεται από τον τύπο (δείτε Ahsanullah et al. (2013) Κεφάλαιο 2, και Arnold et al. (1992))

$$F_{X_{(k)}}(x) = \sum_{r=k}^n \binom{n}{r} [F(x)]^r [1 - F(x)]^{n-r}, \quad x \in R.$$

Η παραπάνω συνάρτηση κατανομής μπορεί να γραφτεί στη μορφή

$$F_{X_{(k)}}(x) = k \binom{n}{k} \int_0^{F(x)} t^{k-1} (1-t)^{n-k} dt = I_{F(x)}(k, n-k+1) \quad (2.1)$$

όπου

$$I_x(a, b) = \frac{1}{B(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$$

είναι η κανονικοποιημένη μη πλήρης Βήτα συνάρτηση (regularized incomplete Beta function), και

$$B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

είναι η συνήθης Βήτα συνάρτηση.

Στο παρόν σημείο υπενθυμίζουμε ότι η συνάρτηση πυκνότητας $f(x)$ της κατανομής Βήτα με παραμέτρους $a > 0$ και $b > 0$ (συμβολισμός $Beta(a, b)$), δίνεται από τον τύπο

$$f(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, \quad 0 \leq x \leq 1.$$

Πρόταση 2.1. Έστω $U_1, U_2, \dots, U_n$ τυχαίο δείγμα μεγέθους n από την κατανομή $U(0, 1)$. Τότε η k -οστή διατεταγμένη τυχαία μεταβλητή $U_{(k)}$ έχει την κατανομή Βήτα με παραμέτρους k και $n - k + 1$, δηλαδή $U_{(k)} \sim Beta(k, n - k + 1)$.

Απόδειξη. Έστω $F(u) = u$ ($0 \leq u \leq 1$), η κοινή συνάρτηση κατανομής των τυχαίων μεταβλητών $U_i, 1 \leq i \leq n$. Από τη σχέση (2.1) προκύπτει ότι

$$\begin{aligned} F_{U_{(k)}}(u) &= \frac{n!}{(k-1)!(n-k)!} \int_0^u t^{k-1} (1-t)^{n-k} dt \\ &= \frac{1}{\frac{\Gamma(k)!\Gamma(n-k+1)!}{\Gamma(n+1)}} \int_0^u t^{k-1} (1-t)^{(n-k+1)-1} dt \\ &= \int_0^u \frac{1}{B(k, n-k+1)} t^{k-1} (1-t)^{(n-k+1)-1} dt. \end{aligned}$$

Παραγωγίζοντας παίρνουμε ότι

$$f_{U_{(k)}}(u) = \frac{1}{B(k, n-k+1)} u^{k-1} (1-u)^{(n-k+1)-1}, \quad 0 \leq u \leq 1.$$

Συνεπώς $U_{(k)} \sim Beta(k, n - k + 1)$. ■

Πρόταση 2.2. Έστω $X_1, X_2, \dots, X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με κοινή συνάρτηση κατανομής $F(\cdot)$, και έστω $X_{(k)}$ η k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$). Τότε η τυχαία μεταβλητή $F(X_{(k)})$ έχει την κατανομή Βήτα με παραμέτρους k και $n - k + 1$, δηλαδή

$$F(X_{(k)}) \sim Beta(k, n - k + 1).$$

Απόδειξη. Έστω

$$U_i = F(X_i), \quad 1 \leq i \leq n.$$

Από γνωστό αποτέλεσμα της θεωρίας πιθανοτήτων έχουμε ότι οι τυχαίες μεταβλητές $U_1, U_2, \dots, U_n$ αποτελούν τυχαίο δείγμα μεγέθους n από την κατανομή $U(0,1)$. Επίσης είναι εύκολο να διαπιστωθεί ότι

$$U_{(k)} = F(X_{(k)}), \quad 1 \leq k \leq n.$$

Έτσι από την Πρόταση 2.1 προκύπτει άμεσα ότι

$$U_{(k)} = F(X_{(k)}) \sim \text{Beta}(k, n - k + 1). \quad \blacksquare$$

Ορισμός 2.1. Έστω $X_1, X_2, \dots, X_n$ ένα τυχαίο δείγμα από κάποια κατανομή και έστω $X_{(k)}$ η k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$). Η διάμεσος $\tilde{X}_n$ του δείγματος ορίζεται από τη σχέση

$$\tilde{X}_n = \begin{cases} X_{(m+1)} = X_{(n+1)/2}, & \text{αν } n = 2m + 1 \\ \frac{X_{(m)} + X_{(m+1)}}{2} = \frac{X_{n/2} + X_{n/2+1}}{2}, & \text{αν } n = 2m. \end{cases}$$

Χωρίς βλάβη της γενικότητας από εδώ και στο εξής θα υποθέσουμε ότι ο αριθμός n είναι περιττός, δηλαδή $n = 2m + 1$.

Πρόταση 2.3. Έστω $X_1, X_2, \dots, X_n$ ένα τυχαίο δείγμα από την κατανομή $N(\mu, \sigma^2)$ και $n = 2m + 1$. Τότε

$$\tilde{X}_n = \mu + \sigma \Phi^{-1}(Y)$$

όπου $Y \sim \text{Beta}(m + 1, m + 1)$ και $\Phi(z)$ η συνάρτηση κατανομής της τυπικής κανονικής κατανομής $N(0,1)$.

Απόδειξη. Έστω $F(x)$ η συνάρτηση κατανομής της κατανομής $N(\mu, \sigma^2)$. Από τον ορισμό της διαμέσου και από την Πρόταση 2.2 προκύπτει ότι

$$F(\tilde{X}_n) = F(X_{(m+1)}) \sim \text{Beta}(m + 1, m + 1).$$

Αφού $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ παίρνουμε άμεσα ότι

$$F(\tilde{X}_n) = \Phi\left(\frac{\tilde{X}_n - \mu}{\sigma}\right) \sim \text{Beta}(m + 1, m + 1).$$

Συνεπώς αν $Y \sim \text{Beta}(m + 1, m + 1)$ τότε μπορούμε να γράψουμε (υπό την έννοια της κατανομής) ότι

$$\Phi\left(\frac{\tilde{X}_n - \mu}{\sigma}\right) = Y$$

ή ισοδύναμα

$$\tilde{X}_n = \mu + \sigma \Phi^{-1}(Y). \quad \blacksquare$$

Πρόταση 2.4. Έστω $X_1, X_2, \dots, X_n$ ένα τυχαίο δείγμα από την κατανομή $N(\mu, \sigma^2)$ και $n = 2m + 1$. Τότε

$$F_{\tilde{X}_n}(x) = P(\tilde{X}_n \leq x) = F_B\left(\Phi\left(\frac{x - \mu}{\sigma}\right) \mid m + 1, m + 1\right)$$

όπου $F_B(x \mid a, b)$ είναι η συνάρτηση κατανομής της κατανομής $Beta(a, b)$.

Απόδειξη. Χρησιμοποιώντας την Πρόταση 2.3 παίρνουμε

$$\begin{aligned} F_{\tilde{X}_n}(x) &= P(\tilde{X}_n \leq x) = P\left(\frac{\tilde{X}_n - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(\Phi^{-1}(Y) \leq \frac{x - \mu}{\sigma}\right) \\ &= P\left(Y \leq \Phi\left(\frac{x - \mu}{\sigma}\right)\right) = F_B\left(\Phi\left(\frac{x - \mu}{\sigma}\right) \mid m + 1, m + 1\right) \end{aligned}$$

όπου $F_B(x \mid a, b)$ είναι η συνάρτηση κατανομής της κατανομής $Beta(a, b)$. ■

2.3 Το διάγραμμα ελέγχου διαμέσου

Το διάγραμμα ελέγχου της διαμέσου χρησιμοποιείται για την παρακολούθηση της μέσης τιμής μιας διεργασίας, εναλλακτικά του κλασικού διαγράμματος της μέσης τιμής ($\bar{X}$). Αντί να βασίζεται στη δειγματική μέση τιμή, χρησιμοποιεί τη δειγματική διάμεσο, η οποία αποτελεί πιο ανθεκτικό μέτρο κεντρικής τάσης, καθώς δεν επηρεάζεται σημαντικά από ακραίες τιμές ή αποκλίσεις από την κανονικότητα.

Σκοπός του διαγράμματος είναι να εντοπίζει εγκαίρως πιθανές μετατοπίσεις στο κεντρικό επίπεδο της διεργασίας, δηλαδή αλλαγές στη μέση τιμή της κατανομής των τιμών του χαρακτηριστικού που παρακολουθείται. Η χρήση της διαμέσου το καθιστά πιο αξιόπιστο διάγραμμα παρακολούθησης σε περιπτώσεις όπου τα δεδομένα παρουσιάζουν θόρυβο, ασυμμετρίες ή μεμονωμένες ακραίες τιμές. Τα αποτελέσματα που δίνονται για το διάγραμμα ελέγχου διαμέσου βασίζονται στις εργασίες των Castagliola & Figueiredo (2013), Montgomery, D. C. (2013) και Chakraborti et al. (2004).

2.3.1 Φάση II – Γνωστές παράμετροι

Ας υποθέσουμε ότι η εντός ελέγχου (In Control, IC) κατανομή των μετρήσεων που παρακολουθούμε είναι η κανονική κατανομή με μέση τιμή μ_0 και διασπορά σ_0^2 . Από τη

διεργασία επιλέγουμε τυχαία δείγματα μεγέθους n σε τακτά χρονικά διαστήματα. Έστω $\mathbf{X}_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n})$ το δείγμα που πήραμε στην i -οστή δειγματοληψία ($i = 1, 2, \dots$). Σύμφωνα με τα παραπάνω, όταν η διεργασία λειτουργεί υπό συνθήκες φυσικής μεταβλητότητας, δηλαδή όταν έχουμε μια εντός ελέγχου διεργασία, ισχύει ότι $X_{i,j} \sim N(\mu_0, \sigma_0^2)$ για $i = 1, 2, \dots$ και $j = 1, 2, \dots, n$. Στην περίπτωση που η μέση τιμή της κανονικής κατανομής μετατοπιστεί στη θέση $\mu_1 = \mu_0 + \delta\sigma_0$ όπου $\delta \neq 0$, δηλαδή όταν έχουμε μια εκτός ελέγχου διεργασία (Out of Control, OOC), τότε $X_{i,j} \sim N(\mu_0 + \delta\sigma_0, \sigma_0^2)$. Η περίπτωση $\delta = 0$ αντιστοιχεί σε μια εντός ελέγχου διεργασία.

Η δειγματική διάμεσος του i -οστού δείγματος ορίζεται από τη σχέση

$$\tilde{Y}_i = \begin{cases} X_{i,((n+1)/2)}, & \text{αν } n \text{ περιττός,} \\ \frac{X_{i,(n/2)} + X_{i,(n/2+1)}}{2}, & \text{αν } n \text{ άρτιος,} \end{cases}$$

όπου $X_{i,(k)}$ είναι η k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$) του δείγματος $(X_{i,1}, X_{i,2}, \dots, X_{i,n})$. Συνήθως, χωρίς απώλεια της γενικότητας των αποτελεσμάτων, υποθέτουμε ότι το μέγεθος δείγματος n είναι περιττός αριθμός για να απλοποιηθεί ο υπολογισμός της διαμέσου.

Για την κατασκευή του διαγράμματος της διαμέσου ($\tilde{X}$ διάγραμμα ελέγχου) θα χρησιμοποιηθεί ένα δίπλευρο διάγραμμα ελέγχου όπου τα όρια ελέγχου θα είναι τα

$$LCL = \mu_0 - K\sigma_0,$$

$$UCL = \mu_0 + K\sigma_0,$$

με $K > 0$. Η τιμή της σταθεράς K καθορίζεται με τέτοιο τρόπο έτσι ώστε να επιτυγχάνεται μια προκαθορισμένη τιμή για το ARL_0 . Η πιθανότητα να δώσει σήμα εκτός ελέγχου διεργασίας το i -οστό δείγμα, δίνεται από τη σχέση

$$p = P(\tilde{Y}_i \leq LCL) + P(\tilde{Y}_i \geq UCL) = F_{\tilde{Y}_i}(LCL) + 1 - F_{\tilde{Y}_i}(UCL),$$

όπου $F_{\tilde{Y}_i}(y)$ είναι η αθροιστική συνάρτηση κατανομής της τυχαίας μεταβλητής $\tilde{Y}_i$ ($i \geq 1$). Σύμφωνα με την Πρόταση 2.4, όταν η μέση τιμή είναι ίση με $\mu_1 = \mu_0 + \delta\sigma_0$, έχουμε ότι

$$F_{\tilde{Y}_i}(y) = F_B\left(\Phi\left(\frac{y - \mu_1}{\sigma_0}\right) \mid \frac{n+1}{2}, \frac{n+1}{2}\right) = F_B\left(\Phi\left(\frac{y - \mu_0}{\sigma_0} - \delta\right) \mid \frac{n+1}{2}, \frac{n+1}{2}\right)$$

όπου $F_B(x \mid a, b)$ είναι η συνάρτηση κατανομής της κατανομής $Beta(a, b)$.

Συνεπώς

$$p = F_B \left(\Phi \left(\frac{LCL - \mu_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + 1 - F_B \left(\Phi \left(\frac{UCL - \mu_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) \\ = F_B \left(\Phi(-K - \delta) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + 1 - F_B \left(\Phi(K - \delta) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right).$$

Για τη συνάρτηση κατανομής $F_B(x | a, a)$ της κατανομής $Beta(a, a)$, ισχύει ότι (ιδιότητα συμμετρίας)

$$F_B(x | a, a) = 1 - F_B(1 - x | a, a).$$

Επίσης

$$\Phi(K - \delta) = 1 - \Phi(-K + \delta).$$

Χρησιμοποιώντας τις δύο παραπάνω σχέσεις παίρνουμε τελικά ότι

$$p = F_B \left(\Phi(-(K + \delta)) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + F_B \left(\Phi(-(K - \delta)) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right).$$

Φυσικά για $\delta = 0$, προκύπτει η πιθανότητα σφάλματος τύπου I, δηλαδή

$$p = \alpha = 2F_B \left(\Phi(-K) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right).$$

Η κατανομή του μήκους ροής (run length, RL) στο διάγραμμα ελέγχου διαμέσου, όταν οι παράμετροι μ_0 και σ_0 είναι γνωστές, ακολουθεί τη γεωμετρική κατανομή με παράμετρο p (p είναι η πιθανότητα να εμφανιστεί σήμα εκτός ελέγχου διεργασίας). Η συνάρτηση πιθανότητας και η αθροιστική συνάρτηση κατανομής του μήκους ροής δίνονται από τις σχέσεις

$$f_{RL}(l | n, K, \delta) = P(RL = l | n, K, \delta) = p(1 - p)^{l-1}, \quad l = 1, 2, \dots,$$

$$F_{RL}(l | n, K, \delta) = P(RL \leq l | n, K, \delta) = 1 - (1 - p)^l, \quad l = 1, 2, \dots$$

Το μέσο μήκος ροής ($ARL = E(RL)$) και η τυπική απόκλιση του μήκους ροής ($SDRL = \sqrt{V(RL)}$) δίνονται από τις σχέσεις

$$ARL(n, K, \delta) = \frac{1}{p},$$

$$SDRL(n, K, \delta) = \frac{\sqrt{1 - p}}{p}.$$

2.3.2 Φάση I – Άγνωστες παράμετροι

Οι εντός ελέγχου παράμετροι της διεργασίας, δηλαδή η πραγματική μέση τιμή μ_0 και η τυπική απόκλιση σ_0 , συνήθως δεν είναι γνωστές. Για τη δημιουργία ενός διαγράμματος ελέγχου διαμέσου Φάσης I, αυτές οι παράμετροι εκτιμώνται από ένα σύνολο δεδομένων που αποτελείται από m ανεξάρτητα δείγματα μεγέθους n . Οι πιο συνηθισμένες εκτιμήτριες για τη μέση τιμή και την τυπική απόκλιση μιας διεργασίας, όταν χρησιμοποιούνται διαγράμματα μέσης τιμής είναι οι

$$\hat{\mu}_0 = \frac{1}{m} \sum_{i=1}^m \bar{X}_i = \bar{\bar{X}},$$

$$\hat{\sigma}_0 = \frac{1}{c_4} \left(\frac{1}{m} \sum_{i=1}^m S_i \right) = \frac{\bar{S}}{c_4},$$

όπου $\bar{X}_i$ και S_i είναι ο δειγματικός μέσος και η δειγματική τυπική απόκλιση του i -δείγματος (δείτε Παράγραφο 1.4.2). Ωστόσο, στο διάγραμμα διαμέσου χρησιμοποιούνται πιο ανθεκτικές εκτιμήτριες, οι οποίες βασίζονται σε διατεταγμένες στατιστικές συναρτήσεις. Τέτοιες εκτιμήτριες είναι οι

$$\hat{\mu}'_0 = \frac{1}{m} \sum_{i=1}^m \tilde{X}_i,$$

$$\hat{\sigma}'_0 = \frac{1}{d_2} \left(\frac{1}{m} \sum_{i=1}^m R_i \right),$$

όπου $\tilde{X}_i$ είναι η δειγματική διάμεσος και $R_i = X_{i,(n)} - X_{i,(1)}$ είναι το εύρος του i -δείγματος. Με τη χρήση αυτών των εκτιμητριών, τα όρια ελέγχου του διαγράμματος ελέγχου διαμέσου Φάσης II με εκτιμώμενες παραμέτρους γίνονται

$$LCL' = \hat{\mu}'_0 - K \hat{\sigma}'_0,$$

$$UCL' = \hat{\mu}'_0 + K \hat{\sigma}'_0.$$

Επειδή οι εκτιμήτριες $\hat{\mu}'_0$ και $\hat{\sigma}'_0$ είναι τυχαίες μεταβλητές, τα όρια ελέγχου είναι πλέον και αυτά τυχαίες μεταβλητές.

Η πιθανότητα να δώσει σήμα εκτός ελέγχου διεργασίας το i -οστό δείγμα, δίνεται από τη σχέση

$$\begin{aligned}
p &= F_{\bar{Y}_i}(LCL') + 1 - F_{\bar{Y}_i}(UCL') \\
&= F_B \left(\Phi \left(\frac{LCL' - \mu_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + 1 - F_B \left(\Phi \left(\frac{UCL' - \mu_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) \\
&= F_B \left(\Phi \left(\frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} - K \frac{\hat{\sigma}'_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + 1 \\
&\quad - F_B \left(\Phi \left(\frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} + K \frac{\hat{\sigma}'_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) \\
&= F_B \left(\Phi \left(\frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} - K \frac{\hat{\sigma}'_0}{\sigma_0} - \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) \\
&\quad + F_B \left(\Phi \left(-\frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} - K \frac{\hat{\sigma}'_0}{\sigma_0} + \delta \right) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right).
\end{aligned}$$

Έστω οι τυχαίες μεταβλητές

$$U = \frac{\hat{\mu}'_0 - \mu_0}{\sigma_0}, \quad W = \frac{\hat{\sigma}'_0}{\sigma_0}.$$

Τότε η πιθανότητα να δώσει σήμα εκτός ελέγχου διεργασίας το i -οστό δείγμα, γράφεται στη μορφή

$$p = F_B \left(\Phi(U - KW - \delta) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right) + F_B \left(\Phi(-U - KW + \delta) \middle| \frac{n+1}{2}, \frac{n+1}{2} \right).$$

Η παραπάνω πιθανότητα είναι στην πραγματικότητα δεσμευμένη πιθανότητα αφού για να υπολογιστεί χρειάζεται να γνωρίζουμε τις τιμές των U, W . Αν η από κοινού συνάρτηση πυκνότητας του ζεύγους (U, W) είναι η $f_{(U,W)}(u, w|m, n)$, τότε η (μη δεσμευμένη) η συνάρτηση πιθανότητας και η συνάρτηση κατανομής του μήκους ροής δίνονται από τους τύπους

$$\begin{aligned}
f_{RL}(l|m, n, K, \delta) &= \int_{-\infty}^{+\infty} \int_0^{+\infty} p(1-p)^{l-1} f_{(U,W)}(u, w|m, n) dw du, \\
F_{RL}(l|m, n, K, \delta) &= 1 - \int_{-\infty}^{+\infty} \int_0^{+\infty} (1-p)^l f_{(U,W)}(u, w|m, n) dw du,
\end{aligned}$$

αντίστοιχα. Για το μέσο μήκος ροής του διαγράμματος έχουμε

$$ARL(m, n, K, \delta) = \int_{-\infty}^{+\infty} \int_0^{+\infty} \frac{1}{p} f_{(U,W)}(u, w|m, n) dw du,$$

και για την τυπική απόκλιση του έχουμε

$$SDRL(m, n, K, \delta) = \sqrt{E2RL(m, n, K, \delta) - ARL^2(m, n, K, \delta)},$$

όπου

$$E2RL(m, n, K, \delta) = \int_{-\infty}^{+\infty} \int_0^{+\infty} \frac{2-p}{p^2} f_{(U,W)}(u, w|m, n) dw du.$$

Όταν έχουμε δείγμα από κανονική κατανομή, η δειγματική διάμεσος και το εύρος, αν και είναι εξαρτημένες τυχαίες μεταβλητές, ωστόσο είναι ασυσχέτιστες. Καθώς το πλήθος των δειγμάτων m αυξάνεται, λόγω του Κεντρικού Οριακού Θεωρήματος προκύπτει ότι οι στατιστικές συναρτήσεις $\hat{\mu}_0'$ και $\hat{\sigma}_0'$ έχουν ασυμπτωτική κατανομή τη διδιάστατη κανονική κατανομή και είναι ασυμπτωτικά ανεξάρτητες. Έτσι, η από κοινού συνάρτηση πυκνότητας του ζεύγους (U, W) μπορεί να προσεγγιστεί σύμφωνα με τη σχέση

$$f_{(U,W)}(u, w|m, n) \cong f_U(u|m, n)f_W(w|m, n).$$

Συνεπώς είναι εφικτός ο υπολογισμός του ARL και του $SDRL$ αρκεί να γνωρίζουμε τις (ασυμπτωτικές) κατανομές των τυχαίων μεταβλητών U και W . Ικανοποιητικές προσεγγίσεις τους έχουν προταθεί από τους Castagliola & Figueiredo (2013).

2.4 Απόδοση του διαγράμματος ελέγχου διαμέσου

Η αξιολόγηση της απόδοσης ενός διαγράμματος ελέγχου βασίζεται συνήθως στο μέσο μήκος ροής του το οποίο εκφράζει τον μέσο αριθμό δειγμάτων που απαιτούνται μέχρι να δοθεί ένδειξη εκτός ελέγχου διεργασίας. Ένα αποτελεσματικό διάγραμμα ελέγχου πρέπει να έχει μεγάλο ARL όταν η διεργασία βρίσκεται εντός ελέγχου, ώστε να αποφεύγονται ψευδείς συναγερμοί, και μικρό ARL όταν συμβαίνουν πραγματικές μετατοπίσεις στη διεργασία έτσι ώστε να εντοπίζονται γρήγορα αυτές οι αλλαγές.

Στην παρούσα ενότητα εξετάζεται η απόδοση του διαγράμματος ελέγχου της διαμέσου στη Φάση II και στη Φάση I. Για διάφορες τιμές του ARL_0 (370.4, 500), του μεγέθους του δείγματος n (3, 5, 7, 9), και του αριθμού των δειγμάτων m Φάσης I (5, 10, 20, 40, 80, ∞) υπολογίζεται το ARL για διάφορες τιμές του δ που καθορίζει τη μετατόπιση της μέσης τιμής. Ο υπολογισμός του ARL είναι ακριβής στη Φάση II (είναι η

περίπτωση $m = \infty$ στους Πίνακες 2.2 και 2.3), ενώ στη Φάση I υπολογίζεται μέσω προσομοίωσης (συνολικά 10^5 προσομοιώσεις).

Ο Πίνακας 2.1 δίνει τιμές της παραμέτρου K που είναι απαραίτητη για να υπολογιστούν τα όρια ελέγχου Φάσης II. Ο κώδικας στην R που παράγει τις τιμές της παραμέτρου K δίνεται στο Παράρτημα.

Πίνακας 2.1. Τιμές της παραμέτρου K

Μέγεθος δείγματος n	$ARL_0 = 370.4$	$ARL_0 = 500$
3	2.026338	2.088636
5	1.619279	1.669069
7	1.385449	1.42794
9	1.229683	1.267303
11	1.116606	1.150697
13	1.029785	1.061173
15	0.960445	0.9896792

Έτσι προκύπτουν οι δύο παρακάτω πίνακες.

Πίνακας 2.2. Τιμές ARL ($ARL_0 = 370.4$)

δ	$n = 3, K = 2.026338$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	51036.8	1985.6	687.3	484.1	419.3	370.4
0.1	48937.1	1981.7	643.3	448.9	385.9	335.2
0.2	43163.2	1565.4	534.1	361.5	303.7	258.3
0.3	35071.2	1223.1	402.3	263.1	217.1	182.6
0.5	18310.5	697.7	187.7	121.8	100.9	85.7
0.7	7224.3	233.3	82.3	55.9	47.8	41.6
1.0	1228.1	61.0	26.4	19.9	17.8	16.0
1.5	55.9	7.2	4.9	4.2	4.0	4.7
2.0	6.2	2.9	2.4	2.2	2.1	2.1
δ	$n = 5, K = 1.619279$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	1579.8	609.5	447.7	400.0	383.5	370.4
0.1	1541.6	571.8	404.5	358.2	337.2	317.7
0.2	1344.1	443.7	313.4	262.3	239.1	218.3
0.3	888.4	331.0	210.8	170.1	152.7	137.4
0.5	461.2	134.0	81.8	65.6	59.4	53.9
0.7	167.9	50.2	32.5	27.2	25.1	23.2
1.0	40.5	13.4	10.2	9.0	8.6	8.1
1.5	3.4	2.5	2.2	2.1	2.1	2.4
2.0	1.6	1.4	1.4	1.3	1.3	1.3
δ	$n = 7, K = 1.385449$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	747.2	464.9	397.4	379.4	373.7	370.4
0.1	698.8	419.7	352.9	325.7	313.3	301.4
0.2	590.9	321.3	248.9	217.2	202.2	187.2
0.3	410.8	212.7	151.8	127.6	117.6	107.6
0.5	172.3	72.3	50.6	43.1	40.0	37.1
0.7	51.1	24.8	18.7	16.6	15.7	14.8
1.0	10.7	6.8	5.8	5.4	5.2	5.0
1.5	1.8	1.6	1.5	1.5	1.5	1.7
2.0	1.2	1.1	1.1	1.1	1.1	1.1
δ	$n = 9, K = 1.229683$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	561.6	411.6	376.2	368.8	368.5	370.4
0.1	515.2	362.9	323.0	305.2	295.4	286.3
0.2	405.0	261.2	210.6	186.3	174.6	162.6
0.3	270.0	158.1	118.2	101.0	93.9	86.8
0.5	90.9	46.9	34.9	30.7	28.9	27.2
0.7	26.5	15.3	12.4	11.3	10.8	10.4
1.0	5.8	4.3	3.9	3.7	3.6	3.5
1.5	1.4	1.3	1.3	1.2	1.2	1.3
2.0	1.0	1.0	1.0	1.0	1.0	1.0

Πίνακας 2.3. Τιμές ARL ($ARL_0 = 500$)

δ	$n = 3, K = 2.088636$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	878028	3539.9	1017.6	679.1	576.2	500
0.1	105036	3679.2	949.6	627.8	528.2	450.1
0.2	141134	2725.7	78.3	500.8	411.3	342.7
0.3	325233	2123.2	582.3	359.7	290.0	239.0
0.5	45350	1244.8	263.2	161.6	131.0	109.3
0.7	30730	370.0	111.5	72.0	60.5	51.8
1.0	90025	89.4	33.9	24.7	21.6	19.2
1.5	159	12.3	7.3	6.1	5.7	5.3
2.0	12	3.3	2.6	2.4	2.3	2.2
δ	$n = 5, K = 1.669069$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	2762.0	897.9	625.7	548.3	521.4	500
0.1	2699.9	847.1	563.0	488.7	455.9	425.6
0.2	2364.8	644.4	432.3	353.5	318.6	287.7
0.3	1482.9	477.5	286.6	225.5	200.1	178.1
0.5	767.4	186.6	107.4	84.1	75.4	67.8
0.7	265.3	66.9	41.1	33.7	30.9	28.4
1.0	62.8	16.7	12.2	10.7	10.1	9.5
1.5	5.4	3.4	3.0	2.8	2.7	2.7
2.0	1.7	1.5	1.4	1.4	1.4	1.4
δ	$n = 7, K = 1.42794$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	1158.7	660.8	547.9	517.0	506.8	500
0.1	1081.0	594.7	484.6	441.2	421.5	402.7
0.2	920.1	451.3	337.5	289.6	267.2	245.2
0.3	626.0	294.7	202.1	166.9	152.5	138.2
0.5	256.8	96.2	64.7	54.2	50.0	46.1
0.7	71.0	31.4	23.0	20.1	19.0	17.8
1.0	13.6	8.0	6.7	6.2	5.9	5.7
1.5	2.4	2.0	1.9	1.8	1.8	1.8
2.0	1.2	1.2	1.1	1.1	1.1	1.1
δ	$n = 9, K = 1.267303$					
	$m = 5$	$m = 10$	$m = 20$	$m = 40$	$m = 80$	$m = \infty$
0.0	835.8	575.6	515.0	500.7	498.7	500
0.1	763.0	505.7	439.9	411.3	395.9	381.6
0.2	599.9	360.2	282.4	246.4	229.2	211.9
0.3	392.4	214.4	155.3	130.6	120.6	110.8
0.5	126.8	60.6	43.9	38.1	35.7	33.4
0.7	35.0	18.8	14.9	13.5	12.8	12.2
1.0	7.0	5.0	4.4	4.1	4.0	3.9
1.5	1.6	1.5	1.4	1.4	1.4	1.4
2.0	1.1	1.1	1.0	1.0	1.0	1.0

Όπως φαίνεται από τα αποτελέσματα, η απόδοση του διαγράμματος ελέγχου διαμέσου εξαρτάται σε μεγάλο βαθμό από τον αριθμό των δειγμάτων m που χρησιμοποιούνται στη Φάση I για την εκτίμηση των παραμέτρων. Για μικρές τιμές του m , το ARL αυξάνεται σημαντικά σε σχέση με τη Φάση II, γεγονός που σημαίνει μειωμένη ικανότητα ανίχνευσης μετατοπίσεων στη μέση τιμή της διεργασίας. Αντίθετα, καθώς το m αυξάνεται, οι τιμές του ARL προσεγγίζουν αυτές της Φάσης II, όπως αναμένεται θεωρητικά.

Παρατηρούμε, επίσης, ότι στη Φάση I η χρήση της ίδιας σταθεράς K με αυτή της Φάσης II δεν διατηρεί το επιθυμητό εντός ελέγχου μέσου μήκους ροής. Συγκεκριμένα, όπως παρατηρείται και στους Πίνακες 2.2 και 2.3, το ARL_0 στη Φάση I είναι αρκετά μεγαλύτερο από τις τιμές 370.4 και 500, αντίστοιχα, ιδιαίτερα για μικρές τιμές του m , γεγονός που καθιστά το διάγραμμα πιο συντηρητικό και λιγότερο ευαίσθητο στην ανίχνευση μετατοπίσεων. Για τον λόγο αυτό, προτείνεται η χρήση μιας τροποποιημένης σταθεράς K' , η οποία υπολογίζεται κατάλληλα για κάθε συνδυασμό n και m , ώστε το εντός ελέγχου ARL στη Φάση I να έχει την επιθυμητή τιμή. Με την εισαγωγή του K' αντισταθμίζεται η επιπλέον αβεβαιότητα που προκαλείται από την εκτίμηση των παραμέτρων και εξασφαλίζεται συγκρίσιμη συμπεριφορά του διαγράμματος με την περίπτωση των γνωστών παραμέτρων. Στον Πίνακα 2.4 δίνονται οι τιμές της παραμέτρου K' για διάφορα μεγέθη δειγμάτων n και αριθμό δειγμάτων m για $ARL_0 = 370.4$). Ο κώδικας στην R που παράγει τις τιμές της παραμέτρου K' δίνεται στο Παράρτημα.

Πίνακας 2.4. Τιμές της παραμέτρου K' ($ARL_0 = 370.4$)

Μέγεθος δείγματος n	Αριθμός δειγμάτων m	K'
3	5	1.6390
	10	1.8187
	20	1.9221
	40	1.9751
	80	2.0012
	∞	2.0263
5	5	1.4735
	10	1.5527
	20	1.5901
	40	1.6065
	80	1.6135
	∞	1.6193
7	5	1.3136
	10	1.3575
	20	1.3756
	40	1.3822
	80	1.3844
	∞	1.3854
9	5	1.1904
	10	1.2178
	20	1.2276
	40	1.2302
	80	1.2305
	∞	1.2297

2.5 Εφαρμογή διαγράμματος ελέγχου διαμέσου

Εδώ θα δοθεί ένα ενδεικτικό παράδειγμα εφαρμογής του διαγράμματος ελέγχου διαμέσου στη Φάση I και στη Φάση II, με σκοπό την πρακτική κατανόηση της λειτουργίας του. Θεωρούμε μια διεργασία η οποία όταν βρίσκεται εντός στατιστικού ελέγχου περιγράφεται από μια κανονική κατανομή με μέση τιμή $\mu_0 = 125$ και τυπική απόκλιση $\sigma_0 = 5$. Θέλουμε το διάγραμμα ελέγχου διαμέσου να έχει $ARL_0 = 370.4$ και το μέγεθος των δειγμάτων να είναι $n = 5$. Για τις ανάγκες της Φάσης I θα χρησιμοποιηθούν $m = 10$ προσομοιωμένα τυχαία δείγματα που προέρχονται από την εντός ελέγχου κατανομή. Τα δεδομένα μας δίνονται στον ακόλουθο πίνακα.

Πίνακας 2.5. Προσομοιωμένα δεδομένα Φάσης I

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	$\tilde{X}_i$	R_i
1	121.587	124.921	122.787	126.763	125.366	124.921	5.175
2	123.281	114.066	129.400	128.619	126.099	126.099	15.334
3	127.267	124.383	121.185	126.141	130.597	126.141	9.413
4	123.237	131.633	119.299	132.066	121.989	123.237	12.766
5	137.313	121.800	131.057	128.556	123.987	128.556	15.513
6	126.293	126.945	135.078	123.382	126.893	126.893	11.696
7	127.579	126.662	129.716	126.558	119.553	126.662	10.163
8	138.857	133.571	118.829	125.871	129.916	129.916	20.028
9	125.682	126.830	126.832	122.896	120.502	125.682	6.330
10	118.469	124.876	127.452	116.569	120.852	120.852	10.884
						1258.95961	117.3009131

Οι εκτιμήσεις για τη μέση τιμή και τη διακύμανση είναι

$$\hat{\mu}'_0 = \frac{1}{10} \sum_{i=1}^{10} \tilde{X}_i = \frac{1258.95961}{10} = 125.895961,$$

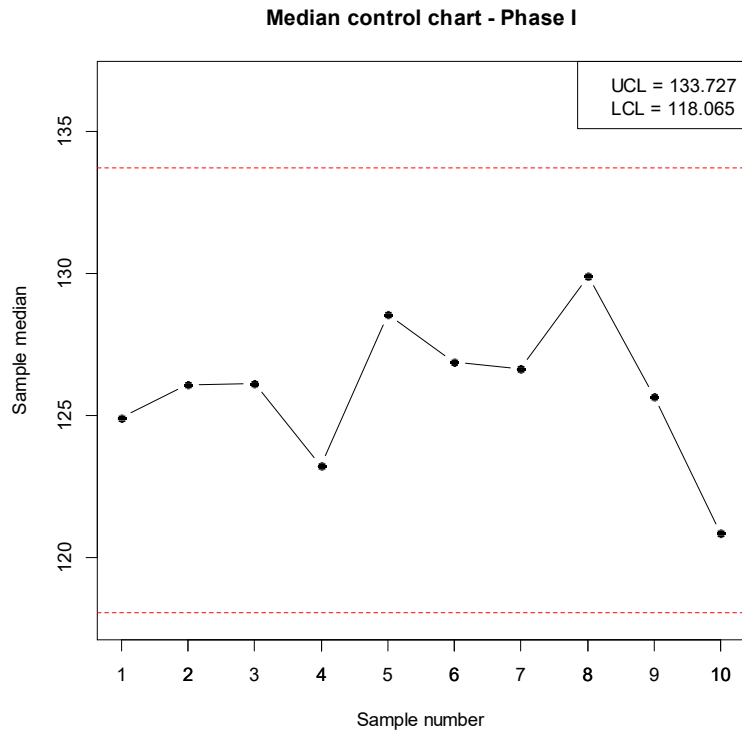
$$\hat{\sigma}'_0 = \frac{1}{d_2} \left(\frac{1}{m} \sum_{i=1}^m R_i \right) = \frac{1}{2.325928947} \left(\frac{117.3009131}{10} \right) = 5.043185573,$$

Με βάση τα αποτελέσματα της προηγούμενης ενότητας, για $ARL_0 = 370.4$ η τιμή της παραμέτρου K' είναι ίση με 1.5527 (δείτε Πίνακα 2.4). Συνεπώς τα όρια ελέγχου του διαγράμματος ελέγχου διαμέσου είναι

$$LCL' = \hat{\mu}'_0 - K' \hat{\sigma}'_0 = 118.0654068,$$

$$UCL' = \hat{\mu}'_0 + K' \hat{\sigma}'_0 = 133.7265153 .$$

Το διάγραμμα ελέγχου Φάσης I είναι το ακόλουθο



Σχήμα 2.1. Διάγραμμα ελέγχου διαμέσου Φάσης I

Παρατηρούμε ότι όλα τα σημεία του διαγράμματος βρίσκονται εντός των ορίων ελέγχου και επομένως δεν προβαίνουμε σε κάποια διορθωτική κίνηση (π.χ., απομάκρυνση κάποιου δείγματος και επανυπολογισμός των ορίων ελέγχου). Συνεπώς στη Φάση II θα χρησιμοποιηθούν τα ίδια όρια ελέγχου.

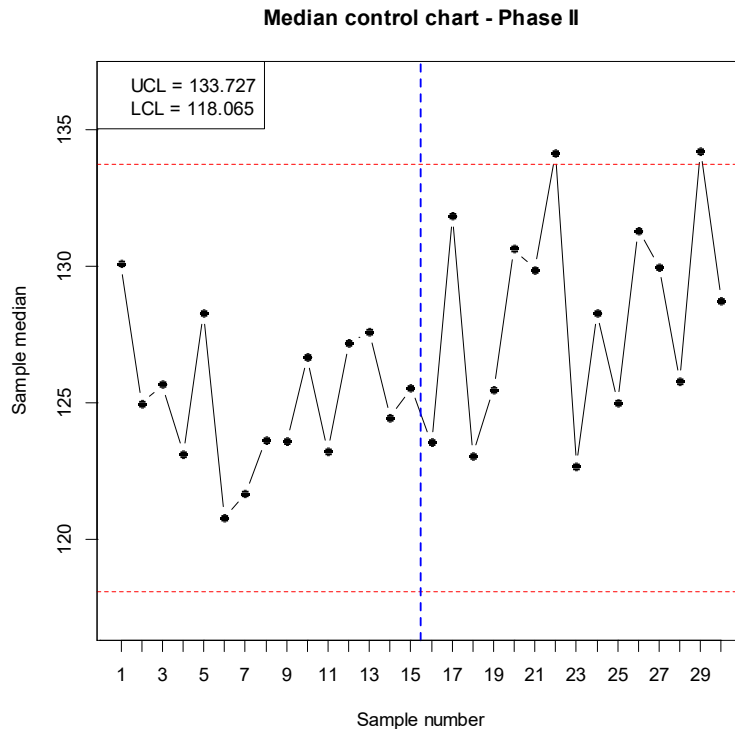
Για τις ανάγκες της Φάσης II πήραμε 30 επιπλέον προσομοιωμένα δείγματα μεγέθους 5. Τα πρώτα 15 είναι από την εντός στατιστικού ελέγχου κανονική κατανομή με μέση τιμή $\mu_0 = 125$ και τυπική απόκλιση $\sigma_0 = 5$, ενώ τα υπόλοιπα από μια εκτός ελέγχου κανονική κατανομή με τυπική απόκλιση $\sigma_1 = \sigma_0 = 5$ και μέση τιμή $\mu_1 = \mu_0 + 0.8\sigma_0 = 129$. Τα δεδομένα δίνονται στον ακόλουθο πίνακα.

Πίνακας 2.6. Δεδομένα Φάσης I

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	$\bar{X}_i$
1	125.751	121.034	130.103	135.545	135.570	130.103
2	129.320	133.072	124.950	122.840	121.661	124.950
3	130.807	125.669	125.715	119.478	124.412	125.669
4	119.299	123.112	125.822	130.685	123.041	123.112
5	128.270	123.947	131.587	121.832	136.418	128.270

6	119.214	117.587	130.204	132.536	120.766	120.766
7	124.745	119.636	126.281	118.781	121.672	121.672
8	125.737	119.909	127.199	120.196	123.623	123.623
9	123.598	130.357	130.566	122.798	115.637	123.598
10	123.787	129.501	135.840	123.727	126.663	126.663
11	120.926	123.600	129.149	122.102	123.206	123.206
12	128.441	137.688	126.600	127.196	126.903	127.196
13	124.664	118.217	127.598	133.007	129.175	127.598
14	124.430	129.473	135.267	122.986	110.216	124.430
15	122.822	119.642	127.777	132.152	125.553	125.553
16	123.535	121.162	122.682	136.372	133.205	123.535
17	125.554	133.469	134.611	129.933	131.847	131.847
18	118.915	126.111	128.695	119.854	123.053	123.053
19	123.037	122.294	125.488	132.513	136.632	125.488
20	124.774	128.704	135.554	135.187	130.656	130.656
21	125.659	131.468	122.152	129.865	136.403	129.865
22	120.238	130.382	144.053	139.174	134.139	134.139
23	131.856	121.719	129.526	122.507	122.651	122.651
24	138.973	119.116	128.298	127.988	133.956	128.298
25	123.028	133.821	134.996	124.988	122.879	124.988
26	131.306	132.575	120.435	128.075	133.787	131.306
27	129.977	138.216	133.517	126.709	122.123	129.977
28	128.025	126.656	125.777	123.071	116.709	125.777
29	134.232	127.968	136.109	129.283	135.413	134.232
30	127.179	127.032	131.323	128.725	134.329	128.725

Το διάγραμμα ελέγχου Φάσης II είναι το ακόλουθο



Σχήμα 2.2. Διάγραμμα ελέγχου διαμέσου Φάσης II

Στο διάγραμμα που προκύπτει παρατηρείται ότι, όσο η διεργασία παραμένει εντός ελέγχου (στα πρώτα 15 δείγματα), οι τιμές της διαμέσου βρίσκονται εντός των ορίων ελέγχου. Μετά όμως από την αύξηση της μέσης τιμής εμφανίζονται σημεία εκτός των ορίων ελέγχου (22^ο και 29^ο), γεγονός που αναδεικνύει την εκτός ελέγχου κατάσταση της διεργασίας. Το αποτέλεσμα αυτό επιβεβαιώνει ότι το διάγραμμα ελέγχου διαμέσου είναι ικανό να ανιχνεύει μεταβολές στη μέση τιμή της διεργασίας.

2.6 Μη παραμετρικό διάγραμμα ελέγχου διαμέσου

Στις προηγούμενες ενότητες η κατασκευή και η αξιολόγηση του διαγράμματος ελέγχου διαμέσου βασίστηκε στην υπόθεση ότι τα δεδομένα που έχουμε στη διάθεσή μας ακολουθούν την κανονική κατανομή. Ωστόσο, στην πράξη η υπόθεση αυτή συχνά δεν επαληθεύεται, καθώς τα δεδομένα ενδέχεται να εμφανίζουν έντονη ασυμμετρία, βαριές ουρές ή γενικότερα να προέρχονται από μια άγνωστη κατανομή. Στις περιπτώσεις αυτές, η χρήση των κλασικών παραμετρικών διαγραμμάτων ελέγχου μπορεί να οδηγήσει σε

σημαντική υποβάθμιση της απόδοσής τους. Για την αντιμετώπιση του προβλήματος αυτού έχουν αναπτυχθεί τα μη παραμετρικά διαγράμματα ελέγχου.

Τα μη παραμετρικά διαγράμματα ελέγχου δεν απαιτούν καμία συγκεκριμένη υπόθεση για τη μορφή της κατανομής των δεδομένων, πέρα από την ανεξαρτησία αυτών και τη συνέχεια της συνάρτησης κατανομής από την οποία προέρχονται. Η βασική τους ιδέα βασίζεται στη χρήση διατεταγμένων στατιστικών συναρτήσεων, οι οποίες έχουν κοινές ιδιότητες ανεξάρτητα από την υποκείμενη κατανομή των δεδομένων. Έτσι, τα όρια ελέγχου μπορούν να προσδιοριστούν με τρόπο που εξασφαλίζει προκαθορισμένη πιθανότητα ψευδούς συναγερμού (FAR, false alarm rate), χωρίς την ανάγκη εκτίμησης παραμέτρων όπως η μέση τιμή ή η διασπορά (δείτε, Chakraborti et al. (2001, 2004)).

Για την κατασκευή ενός μη παραμετρικού (ή ελεύθερου κατανομής) διαγράμματος ελέγχου θεωρούμε ότι έχουμε στη διάθεσή μας ένα εντός ελέγχου τυχαίο δείγμα αναφοράς μεγέθους m , έστω το $\mathbf{X} = (X_1, X_2, \dots, X_m)$, το οποίο προέρχεται από άγνωστη κατανομή με συνεχή συνάρτηση κατανομής $F_X(x) = F(x - \theta)$, όπου θ είναι μια παράμετρος θέσης. Οι m παρατηρήσεις διατάσσονται σε αύξουσα σειρά, έστω

$$X_{(1:m)} \leq X_{(2:m)} \leq \dots \leq X_{(m:m)},$$

όπου $X_{(j:m)}$ είναι η j -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq j \leq m$) του δείγματος $\mathbf{X} = (X_1, X_2, \dots, X_m)$.

Στα μη παραμετρικά διαγράμματα ελέγχου (Φάσης II), τα όρια ελέγχου ορίζονται μέσω κατάλληλης επιλογής διατεταγμένων στατιστικών συναρτήσεων του δείγματος αναφοράς. Συγκεκριμένα, επιλέγονται δύο ακέραιοι δείκτες a και b , με

$$1 \leq a < b \leq m,$$

και τα όρια ελέγχου ορίζονται ως

$$LCL = X_{(a:m)}, \quad UCL = X_{(b:m)}.$$

Για κάθε μελλοντικό δείγμα ελέγχου μεγέθους n , έστω το $\mathbf{Y}^{(h)} = (Y_1^{(h)}, Y_2^{(h)}, \dots, Y_n^{(h)})$, όπου $h \geq 1$, υπολογίζεται η j -οστή διατεταγμένη τυχαία μεταβλητή $Y_{(j:n)}^{(h)}$, $1 \leq j \leq n$, η οποία απεικονίζεται στο διάγραμμα ελέγχου με τα προαναφερθέντα όρια ελέγχου LCL και UCL (για παράδειγμα, η j -οστή διατεταγμένη τυχαία μεταβλητή $Y_{(j:n)}^{(h)}$ θα μπορούσε να είναι η διάμεσος τους δείγματος $\mathbf{Y}^{(h)}$). Έστω $F_Y(y) = F(y - \theta_h)$ η συνάρτηση κατανομής από την οποία προέρχεται το δείγμα $\mathbf{Y}^{(h)}$. Αν η διατεταγμένη τυχαία μεταβλητή $Y_{(j:n)}^{(h)}$ βρίσκεται εντός των ορίων ελέγχου, δηλαδή όταν

$$X_{(a:m)} \leq Y_{(j:n)}^{(h)} \leq X_{(b:m)}$$

τότε θεωρούμε ότι η διεργασία παραμένει εντός ελέγχου κατά τη διάρκεια συλλογής του τυχαίου δείγματος $Y^{(h)}$, δηλαδή $\theta_h = \theta$. Στην αντίθετη περίπτωση θεωρούμε ότι η διεργασία είναι εκτός στατιστικού ελέγχου.

Η λειτουργία του διαγράμματος μπορεί να μελετηθεί και μέσω της στατιστικής συνάρτησης προτεραιότητας (precedence statistic) W_j , η οποία ορίζεται ως ο αριθμός των παρατηρήσεων του τυχαίου δείγματος αναφοράς X που δεν υπερβαίνουν την τιμή $Y_{(j:n)}^{(h)}$. Λόγω της συνεχούς φύσης των δεδομένων μας έχουμε ότι τα ενδεχόμενα

$$X_{(a:m)} \leq Y_{(j:n)}^{(h)} \leq X_{(b:m)}$$

και

$$a \leq W_j \leq b - 1$$

έχουν την ίδια πιθανότητα αφού είναι ίσα.

Όταν η διεργασία βρίσκεται υπό στατιστικό έλεγχο, η ακριβής κατανομή πιθανότητας της στατιστικής συνάρτησης προτεραιότητας W_j ικανοποιεί τη σχέση

$$P_{IC}(W_j = w) = \frac{\binom{j+w-1}{w} \binom{m+n-j-w}{m-w}}{\binom{m+n}{m}}, \quad w = 0, 1, \dots, m,$$

(δείτε Chakraborti & Van der Laan (1996)), η οποία εξαρτάται μόνο από τα μεγέθη m, n και τη θέση j , συνεπώς είναι «ελεύθερη κατανομής».

Έτσι η πιθανότητα το διάγραμμα ελέγχου προτεραιότητας να μη δώσει σήμα εκτός ελέγχου διεργασίας, όταν η διεργασία βρίσκεται υπό έλεγχο, είναι

$$p_0 = P(X_{(a:m)} \leq Y_{(j:n)}^{(h)} \leq X_{(b:m)}) = P(a \leq W_j \leq b - 1) = \sum_{w=a}^{b-1} \frac{\binom{j+w-1}{w} \binom{m+n-j-w}{m-w}}{\binom{m+n}{m}}$$

και επομένως η πιθανότητα να δώσει σήμα εκτός ελέγχου διεργασίας είναι ίση με $1 - p_0$.

Για την επιλογή των ορίων ελέγχου έχουν προταθεί δύο βασικές προσεγγίσεις. Η πρώτη βασίζεται στον καθορισμό μιας επιθυμητής τιμής της πιθανότητας ψευδούς συναγερμού, έστω $FAR = 1 - P_0$. Στην περίπτωση αυτή, οι ακέραιες τιμές a και b επιλέγονται έτσι ώστε, λόγω της διακριτής φύσης της συνάρτησης προτεραιότητας, να ικανοποιείται η σχέση

$$P(a \leq W_j \leq b-1) = \sum_{w=a}^{b-1} \frac{\binom{j+w-1}{w} \binom{m+n-j-w}{m-w}}{\binom{m+n}{m}} \geq P_0$$

προκειμένου να μην ξεπεραστεί η τιμή του FAR. Προφανώς υπάρχουν πολλά ζευγάρια ακεραίων τιμών a και b που ικανοποιούν την παραπάνω σχέση. Από αυτά επιλέγεται εκείνο για το οποίο η πιθανότητα $P(a \leq W_j \leq b-1)$ είναι πιο κοντά στην τιμή P_0 .

Η δεύτερη προσέγγιση είναι να βρεθούν οι ακέραιες τιμές a και b έτσι ώστε να επιτυγχάνεται μιας επιθυμητή τιμή του εντός ελέγχου μέσου μήκους ροής. Ο αναγνώστης παραπέμπεται στους Chakraborti et al. (2001) για τη μελέτη της κατανομή του μήκους ροής του διαγράμματος ελέγχου προτεραιότητας.

Στην ειδική περίπτωση του μη παραμετρικού διαγράμματος ελέγχου προτεραιότητας διαμέσου, θεωρείται συνήθως ότι το μέγεθος των δειγμάτων είναι περιττός αριθμός, έστω $n = 2s + 1$, έτσι ώστε η διάμεσος να είναι μοναδική και να αντιστοιχεί στη διατεταγμένη παρατήρηση $Y_{(s+1:n)}$. Τότε η κατανομή της στατιστικής συνάρτησης προτεραιότητας W_{s+1} είναι συμμετρική, και μια φυσική επιλογή για την τιμή του b είναι η $m - a + 1$. Έτσι η σταθερά a προσδιορίζεται ως ο μεγαλύτερος ακέραιος που ικανοποιεί τη σχέση

$$P(a \leq W_{s+1} \leq m-a) = \sum_{w=a}^{m-a} \frac{\binom{s+w}{w} \binom{m+s-w}{m-w}}{\binom{m+2s+1}{m}} \geq P_0$$

για να επιτευχθεί συγκεκριμένη πιθανότητα ψευδούς συναγερμού, έστω $FAR \leq 1 - P_0$. Έτσι τα όρια ελέγχου προκύπτουν ως

$$LCL = X_{(a:m)}, \quad UCL = X_{(m-a+1:m)}.$$

Για το μη παραμετρικό διάγραμμα ελέγχου προτεραιότητας διαμέσου, στον παρακάτω πίνακα δίνονται τιμές του ζεύγους (a, b) , η τιμή $FAR/2$ και το εντός ελέγχου μέσο μήκος ροής για διάφορες τιμές του $1 - P_0$, του m και του n .

Πίνακας 2.7. Παράμετροι διαγράμματος ελέγχου προτεραιότητας διαμέσου

$1 - P_0$	n	j	m			
			50	100	500	1000
0.00100	5	3	(3,48)	(7,94)	(40,461)	(82,919)
			0.0036	0.00432	0.00477	0.00499
			635.7	213.8	114.5	103.5
	11	6	(7,44)	(15,86)	(83,418)	(167,834)
			0.00465	0.00428	0.00494	0.0497
			642.2	245.0	113.3	108.4
	25	13	(10,41)	(23,78)	(123,375)	(248,743)
			0.00305	0.00402	0.00473	0.00497
			109990.0	510.8	128.3	109.8
0.005	5	3	(2,49)	(5,96)	(31,470)	(64,937)
			0.0015	0.00176	0.00233	0.00246
			5671.0	678.4	242.3	215.1
	11	6	(5,46)	(13,88)	(72,429)	(146,855)
			0.00125	0.00225	0.0024	0.00244
			9503.0	574.5	240.9	219.8
	25	13	(9,42)	(21,80)	(118,383)	(239,762)
			0.00155	0.002	0.00243	0.00245
			44750.0	1488.0	261.0	227.5
0.0027	5	3	(1,50)	(4,97)	(24,476)	(51,950)
			0.0004	0.00102	0.00127	0.00129
			∞	1550.0	460.2	419.5
	11	6	(5,46)	(11,90)	(64,437)	(130,871)
			0.00125	0.00106	0.00131	0.00131
			9503.0	1630.0	456.1	409.8
	25	3	(8,43)	(19,82)	(110,391)	(224,777)
			0.00075	0.00092	0.00127	0.00133
			173700.0	5183.0	526.2	430.2

Για παράδειγμα, ας υποθέσουμε ότι $m = 50$ και $n = 5$, οπότε $s = 2$ και $j = s + 1 = 3$. Για $P_0 = 0.99$, έτσι ώστε το ονομαστικό FAR να είναι 0.01, από τον Πίνακα 2.7 παίρνουμε τις τιμές $a = 3$ και $b = 48$. Συνεπώς

$$LCL = X_{(3:50)}, \quad UCL = X_{(48:50)}.$$

Η ακριβής πιθανότητα ψευδούς συναγερμού για αυτό το διάγραμμα είναι

$$FAR = 1 - \sum_{w=3}^{47} \frac{\binom{2+w}{w} \binom{52-w}{50-w}}{\binom{55}{5}} = 0.007187042$$

Συνεπώς υπάρχει πιθανότητα 0.72% να προκύψει σήμα εκτός ελέγχου διεργασίας ενώ η διεργασία είναι στην πραγματικότητα εντός ελέγχου.

2.7 Εφαρμογή μη παραμετρικού διαγράμματος ελέγχου διαμέσου

Στην παράγραφο αυτή εξετάζεται ένα σύνολο δεδομένων από τον Montgomery (2013) που αναφέρεται στην εσωτερική διάμετρο δακτυλίων εμβόλων (piston rings). Για λόγους πληρότητας αλλά και σύγκρισης τα δεδομένα θα μελετηθούν με το μη παραμετρικό διάγραμμα ελέγχου διαμέσου αλλά και με την κλασική «παραμετρική» μέθοδο των $\bar{X} - S$ διαγραμμάτων ελέγχου. Πρώτα παρουσιάζεται η κλασική μέθοδος μέσω των $\bar{X} - S$ διαγραμμάτων ελέγχου στη Φάση I και II και στη συνέχεια θα εξεταστεί το αντίστοιχο μη παραμετρικό διάγραμμα διαμέσου Φάσης II. Η σύγκριση έχει ενδιαφέρον διότι και οι δύο προσεγγίσεις, όπως θα δούμε στη συνέχεια, οδηγούν στην ίδια ουσιαστικά πληροφορία ως προς τη μετατόπιση της διεργασίας. Ωστόσο η μη παραμετρική μέθοδος το επιτυγχάνει αυτό χωρίς την ανάγκη υπόθεσης συγκεκριμένης κατανομής για τα δεδομένα.

Τα συνολικά δεδομένα της εσωτερικής διαμέτρου των δακτυλίων εμβόλων σε mm δίνονται στον Πίνακα 2.8 και αποτελούνται από 40 υποομάδες (δείγματα) μεγέθους $n = 5$. Στις τρεις τελευταίες στήλες του Πίνακα 2.8 δίνεται ο μέσος, η τυπική απόκλιση και η διάμεσος κάθε υποομάδας.

Πίνακας 2.8. Δεδομένα δακτυλίων εμβόλων (Montgomery (2013))

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	$\bar{X}_i$	S_i	$\tilde{X}_i$
1	74.030	74.002	74.019	73.992	74.008	74.0102	0.0148	74.008
2	73.995	73.992	74.001	74.011	74.004	74.0006	0.0075	74.001
3	73.988	74.024	74.021	74.005	74.002	74.008	0.0147	74.005
4	74.002	73.996	73.993	74.015	74.009	74.003	0.0091	74.002
5	73.992	74.007	74.015	73.989	74.014	74.0034	0.0122	74.007
6	74.009	73.994	73.997	73.985	73.993	73.9956	0.0087	73.994
7	73.995	74.006	73.994	74.000	74.005	74.000	0.0055	74.000
8	73.985	74.003	73.993	74.015	73.988	73.9968	0.0123	73.993
9	74.008	73.995	74.009	74.005	74.004	74.0042	0.0055	74.005
10	73.988	74.000	73.990	74.007	73.995	73.996	0.0063	73.998
11	73.994	73.998	73.994	73.995	73.990	73.9942	0.0029	73.994
12	74.004	74.000	74.007	74.000	73.996	74.0014	0.0042	74.000
13	73.983	74.002	73.998	73.997	74.012	73.9984	0.0105	73.998
14	74.006	73.967	73.994	74.000	73.984	73.9902	0.0153	73.994
15	74.012	74.014	73.998	73.999	74.007	74.006	0.0073	74.007
16	74.000	73.984	74.005	73.998	73.996	73.9966	0.0078	73.998
17	73.994	74.012	73.986	74.005	74.007	74.0008	0.0106	74.005
18	74.006	74.010	74.018	74.003	74.000	74.0074	0.007	74.006
19	73.984	74.002	74.003	74.005	73.997	73.9982	0.0085	74.002
20	74.000	74.010	74.013	74.020	74.003	74.0092	0.008	74.010
21	73.982	74.001	74.015	74.005	73.996	73.9998	0.0122	74.001
22	74.004	73.999	73.990	74.006	74.009	74.0016	0.0074	74.004
23	74.010	73.989	73.990	74.009	74.014	74.0024	0.0119	74.009
24	74.015	74.008	73.993	74.000	74.010	74.0052	0.0087	74.008
25	73.982	73.984	73.995	74.017	74.013	73.9982	0.0162	73.995
26	74.012	74.015	74.030	73.986	74.000	74.0086	0.0165	74.012
27	73.995	74.010	73.990	74.015	74.001	74.0022	0.0103	74.001
28	73.987	73.999	73.985	74.000	73.990	73.9922	0.0069	73.990
29	74.008	74.010	74.003	73.991	74.006	74.0036	0.0075	74.006
30	74.003	74.000	74.001	73.986	73.997	73.9974	0.0067	74.000
31	73.994	74.003	74.015	74.02	74.004	74.0072	0.0103	74.004
32	74.008	74.002	74.018	73.995	74.005	74.0056	0.0084	74.005
33	74.001	74.004	73.990	73.996	73.998	73.9978	0.0053	73.998
34	74.015	74.000	74.016	74.025	74.000	74.0112	0.0109	74.015
35	74.030	74.005	74.000	74.016	74.012	74.0126	0.0115	74.012
36	74.001	73.99	73.995	74.01	74.024	74.004	0.0134	74.001
37	74.015	74.02	74.024	74.005	74.019	74.0166	0.0072	74.019
38	74.035	74.01	74.012	74.015	74.026	74.0196	0.0106	74.015
39	74.017	74.013	74.036	74.025	74.026	74.0234	0.0089	74.025
40	74.010	74.005	74.029	74.000	74.020	74.0128	0.0117	74.010

Για την ανάλυση Φάσης I χρησιμοποιούνται τα πρώτα $m = 25$ δείγματα. Χρησιμοποιώντας αυτά τα δείγματα παίρνουμε

$$\bar{S} = \frac{1}{25} \sum_{i=1}^{25} S_i = \frac{1}{25} (0.231) = 0.00924.$$

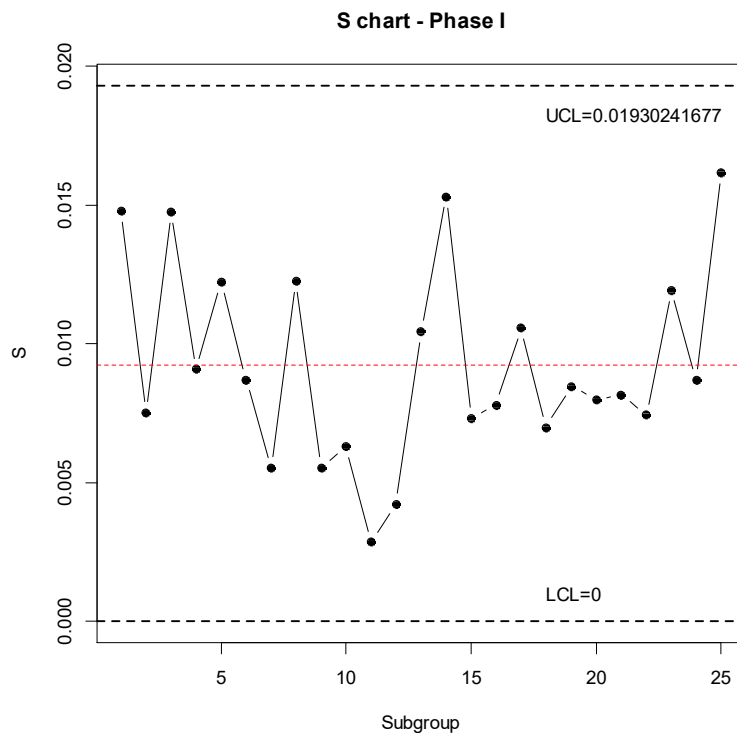
Για $n = 5$ έχουμε ότι $B_3 = 0$ και $B_4 = 2.088997869$. Επομένως, τα όρια ελέγχου του S διαγράμματος ελέγχου για τη Φάση I είναι

$$UCL_S = B_4 \bar{S} = (2.088997869)(0.00924) = 0.019302,$$

$$CL_S = \bar{S} = 0.00924,$$

$$LCL_S = B_3 \bar{S} = 0(0.0094) = 0.$$

Το S διάγραμμα ελέγχου Φάσης I είναι το ακόλουθο



Σχήμα 2.3. S διάγραμμα ελέγχου Φάσης I για την εσωτερική διάμετρο

Η διασπορά βρίσκεται εντός στατιστικού ελέγχου, οπότε προχωρούμε στην κατασκευή του $\bar{X}$ διαγράμματος ελέγχου Φάσης I. Έχουμε ότι

$$\bar{\bar{X}} = \frac{1}{25} \sum_{i=1}^{25} \bar{X}_i = \frac{1}{25} (1850.0294) = 74.001176.$$

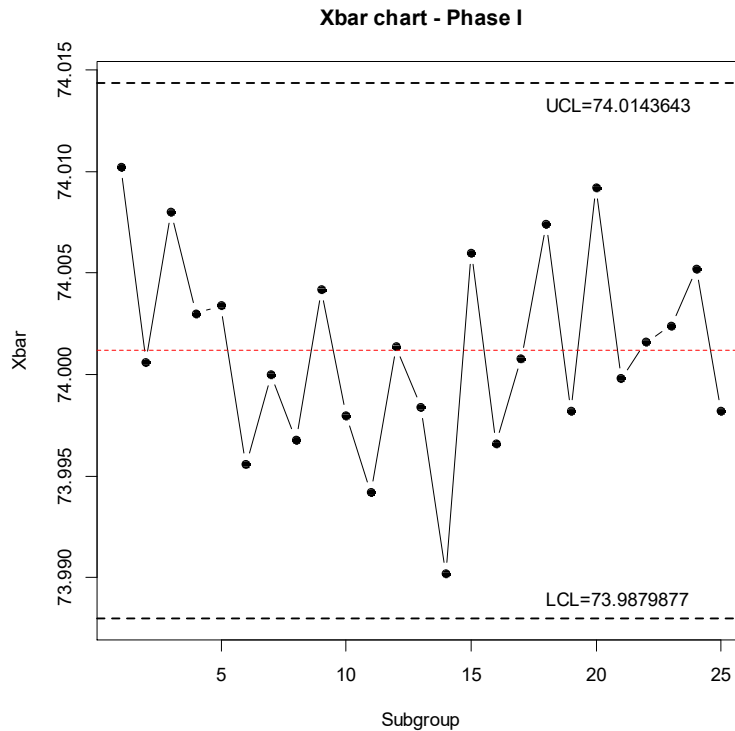
Για $n = 5$ έχουμε ότι $A_3 = 1.42729929$. Επομένως, τα όρια ελέγχου του S διαγράμματος ελέγχου για τη Φάση I είναι

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_3 \bar{S} = 74.001176 + (1.42729929)(0.00924) = 74.014364,$$

$$CL_{\bar{X}} = \bar{\bar{X}} = 74.001176,$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_3 \bar{S} = 74.001176 - (1.42729929)(0.00924) = 73.9879877.$$

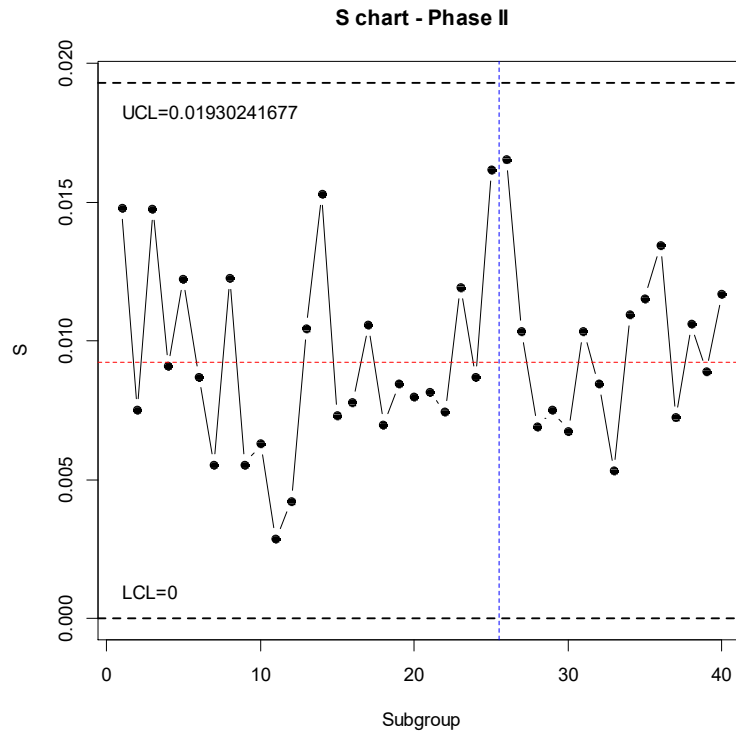
Το $\bar{X}$ διάγραμμα ελέγχου Φάσης I είναι το ακόλουθο



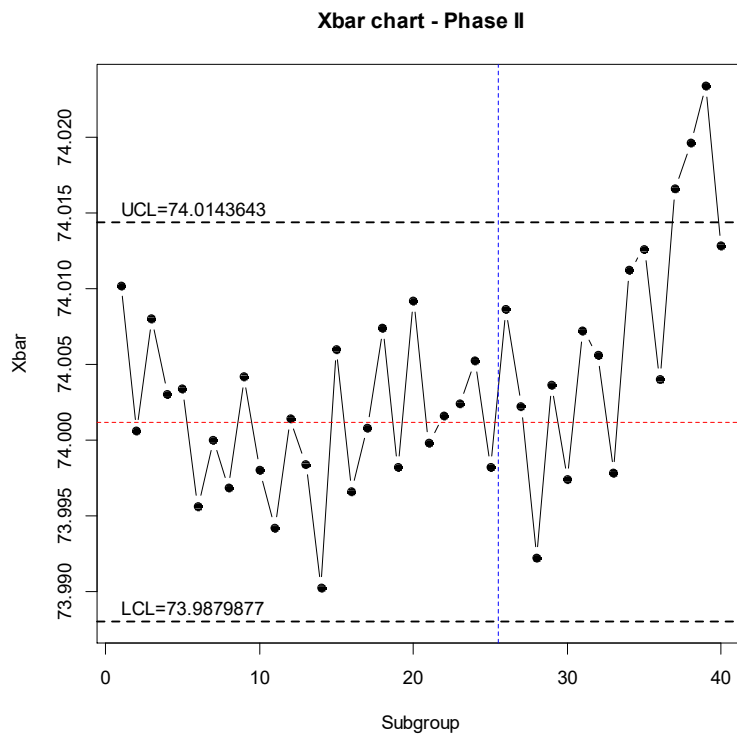
Σχήμα 2.4. $\bar{X}$ διάγραμμα ελέγχου Φάσης I για την εσωτερική διάμετρο

Και στο $\bar{X}$ διάγραμμα ελέγχου όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου. Συνεπώς, δεν απαιτείται αφαίρεση κάποιας υποομάδας ή αναθεώρηση των ορίων και επομένως τα όρια του S και του $\bar{X}$ διαγράμματος ελέγχου Φάσης I θα χρησιμοποιηθούν και στη Φάση II.

Για τις ανάγκες της Φάσης II θα χρησιμοποιηθούν τα υπόλοιπα 15 δείγματα (από το 26^ο έως το 40^ο). Τα διαγράμματα ελέγχου Φάσης II είναι τα ακόλουθα.



Σχήμα 2.5. S διάγραμμα ελέγχου Φάσης II για την εσωτερική διάμετρο



Σχήμα 2.6. $\bar{X}$ διάγραμμα ελέγχου Φάσης II για την εσωτερική διάμετρο

Αρχικά εξετάζεται το S διάγραμμα ελέγχου Φάσης II. Από το Σχήμα 2.5 προκύπτει ότι οι 15 νέες τιμές S_i ($i = 25, 26, \dots, 40$) παραμένουν εντός των ορίων ελέγχου, γεγονός που υποδηλώνει ότι η διασπορά της διεργασίας παραμένει εντός στατιστικού ελέγχου.

Στη συνέχεια εξετάζεται το $\bar{X}$ διάγραμμα ελέγχου Φάσης II. Όταν οι μέσοι των νέων υποομάδων τοποθετηθούν στο διάγραμμα, παρατηρείται ότι στο 37^ο δείγμα ο δειγματικός μέσος υπερβαίνει το άνω όριο ελέγχου για πρώτη φορά. Αυτό αποτελεί ένδειξη ότι η διεργασία δεν βρίσκεται πλέον υπό στατιστικό έλεγχο και συμπεραίνουμε ότι υπάρχει μετατόπιση της μέσης τιμής της διεργασίας σε υψηλότερο επίπεδο. Με τον τρόπο αυτό, η «παραμετρική» προσέγγιση μέσω των διαγραμμάτων $\bar{X}$ και S οδηγεί στο συμπέρασμα ότι στη Φάση II η διασπορά της διεργασίας παραμένει σταθερή, αλλά η μέση τιμή της παρουσιάζει μετατόπιση. Το αποτέλεσμα αυτό θα χρησιμοποιηθεί στη συνέχεια ως σημείο σύγκρισης με τη μη παραμετρική προσέγγιση που ακολουθεί.

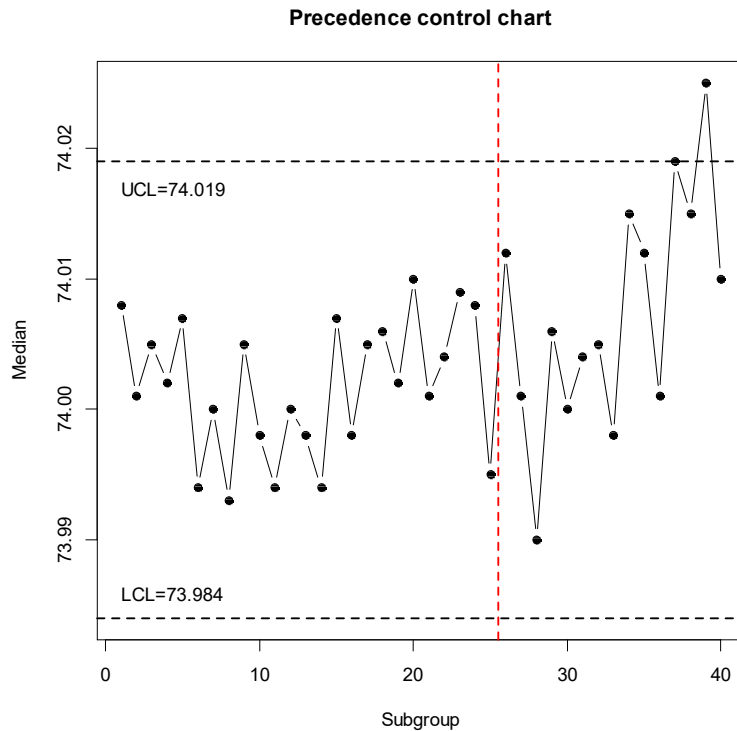
Για το μη παραμετρικό διάγραμμα ελέγχου προτεραιότητας (δείτε Chakraborti (2004)) ακολουθούμε την ίδια στρατηγική για τη Φάση I και τη Φάση II. Επιλέγουμε ως επιθυμητό επίπεδο πιθανότητας ψευδούς συναγερμού $FAR = 0.0027$. Με βάση το δείγμα αναφοράς των $125 = 25 \times 5$ παρατηρήσεων των πρώτων 25 δειγμάτων, για μέγεθος δείγματος $n = 5$ οπότε $s = 2$, προκύπτει ότι $a = 5$ και $b = 121$. Συνεπώς τα όρια ελέγχου αντιστοιχούν στην 5^η διατεταγμένη τιμή για το κάτω όριο ελέγχου και στην 121^η διατεταγμένη τιμή του δείγματος αναφοράς για το άνω όριο ελέγχου. Έτσι

$$LCL = X_{(5:125)} = 73.984, \quad UCL = X_{(121:125)} = 74.019.$$

Με αυτά τα όρια ελέγχου η ακριβής πιθανότητα ψευδούς συναγερμού είναι ίση με

$$FAR = 1 - \sum_{w=5}^{120} \frac{\binom{2+w}{w} \binom{127-w}{125-w}}{\binom{130}{125}} = 0.001865061.$$

Το διάγραμμα ελέγχου προτεραιότητας είναι το ακόλουθο



Σχήμα 2.7. Διάγραμμα ελέγχου προτεραιότητας για την εσωτερική διάμετρο

Η μορφή του διαγράμματος ελέγχου προτεραιότητας δεν έχει μεγάλες διαφορές με εκείνη του παραμετρικού $\bar{X}$ διαγράμματος. Η διάμεσος του 37^{ου} δείγματος είναι ίση με το άνω όριο ελέγχου, ενώ το διάγραμμα δίνει σήμα εκτός ελέγχου διεργασίας στο 39^ο δείγμα.

Η σύγκριση των δύο διαγραμμάτων δείχνει ότι και οι δύο προσεγγίσεις οδηγούν σε παρόμοια συμπεράσματα σχετικά με τη μετατόπιση της διεργασίας. Ωστόσο, το μη παραμετρικό διάγραμμα παρουσιάζει δύο σημαντικά πλεονεκτήματα. Πρώτον, δεν απαιτεί καμία υπόθεση για τη μορφή της κατανομής των δεδομένων. Δεύτερον, η πιθανότητα ψευδούς συναγερμού είναι καθορισμένη, ενώ στο παραμετρικό διάγραμμα $\bar{X}$ η αντίστοιχη πιθανότητα είναι μόνο προσεγγιστική και βασίζεται στην υπόθεση ότι τα δεδομένα ακολουθούν κανονική κατανομή.

Για τον λόγο αυτό, τα μη παραμετρικά διαγράμματα μπορούν να αποτελέσουν μια καλύτερη εναλλακτική λύση των κλασικών παραμετρικών διαγραμμάτων ελέγχου στην πράξη, ιδιαίτερα σε περιπτώσεις όπου η υπόθεση κανονικότητας δεν μπορεί να θεωρηθεί ασφαλής.

ΚΕΦΑΛΑΙΟ 3

Διαγράμματα ελέγχου διαμέσου τύπου CUSUM

3.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο παρουσιάστηκαν τα διαγράμματα ελέγχου τύπου Shewhart για την παρακολούθηση της παραμέτρου θέσης μιας διεργασίας με τη χρήση της δειγματικής διαμέσου. Ωστόσο, όπως είναι γνωστό, τα διαγράμματα αυτά βασίζονται αποκλειστικά στην πληροφορία του τρέχοντος δείγματος, γεγονός που περιορίζει την ικανότητά τους να ανιχνεύουν μικρές και σταδιακές μετατοπίσεις της παραμέτρου θέσης της διεργασίας.

Για τον λόγο αυτό, έχουν αναπτυχθεί διαγράμματα ελέγχου που αξιοποιούν πληροφορίες από διαδοχικά δείγματα, τα λεγόμενα διαγράμματα ελέγχου με μνήμη. Στην κατηγορία αυτή ανήκουν τα διαγράμματα τύπου CUSUM (Cumulative Sum), τα οποία βασίζονται στη συσσώρευση των αποκλίσεων από την τιμή στόχο και έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην ανίχνευση μικρών μετατοπίσεων.

Στην κλασική τους μορφή, τα διαγράμματα CUSUM βασίζονται στον δειγματικό μέσο. Παρ' όλα αυτά, η χρήση του δειγματικού μέσου ενδέχεται να δημιουργήσει προβλήματα σε περιπτώσεις όπου τα δεδομένα περιέχουν ακραίες τιμές ή δεν ακολουθούν την κανονική κατανομή. Για τον λόγο αυτό, και σε συνέχεια της προσέγγισης που ακολουθήθηκε στο προηγούμενο κεφάλαιο, στο παρόν κεφάλαιο εξετάζεται η χρήση της δειγματικής διαμέσου σε ένα διάγραμμα ελέγχου τύπου CUSUM.

Στόχος του κεφαλαίου είναι η ανάπτυξη του διαγράμματος αυτού, καθώς και η μελέτη της συμπεριφοράς του τόσο στην περίπτωση γνωστών όσο και εκτιμημένων παραμέτρων της διεργασίας. Η ανάπτυξη του παρόντος κεφαλαίου βασίστηκε στο

άρθρο των Castagliola et al. (2019) (δείτε επίσης Yang et al. (2010) και Nazir et al. (2013)).

3.2 Το διάγραμμα ελέγχου διαμέσου τύπου CUSUM

Εδώ παρουσιάζεται το διάγραμμα ελέγχου διαμέσου τύπου CUSUM, το οποίο χρησιμοποιείται για την παρακολούθηση της παραμέτρου θέσης μιας διεργασίας. Η ανάλυση του διαγράμματος γίνεται αρχικά υπό την υπόθεση γνωστών παραμέτρων της διεργασίας, περίπτωση η οποία αντιστοιχεί σε λειτουργία Φάσης II. Στη συνέχεια εξετάζεται η περίπτωση όπου οι παράμετροι είναι άγνωστες και εκτιμώνται από δεδομένα Φάσης I, γεγονός που επηρεάζει σημαντικά τη συμπεριφορά του διαγράμματος.

3.2.1 Φάση II – Γνωστές παράμετροι

Έστω ότι η εντός ελέγχου κατανομή των μετρήσεων που παρακολουθούμε είναι η κανονική κατανομή με μέση τιμή μ_0 και διασπορά σ_0^2 . Από τη διεργασία επιλέγουμε τυχαία δείγματα μεγέθους n σε τακτά χρονικά διαστήματα. Έστω $\mathbf{X}_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n})$ το δείγμα που πήραμε στην i -οστή δειγματοληψία ($i = 1, 2, \dots$). Όταν έχουμε μια εντός ελέγχου διεργασία, ισχύει ότι $X_{i,j} \sim N(\mu_0, \sigma_0^2)$ για $i = 1, 2, \dots$ και $j = 1, 2, \dots, n$. Στην περίπτωση που η μέση τιμή μετατοπιστεί στη θέση $\mu_1 = \mu_0 + \delta\sigma_0$, $\delta \neq 0$, τότε $X_{i,j} \sim N(\mu_0 + \delta\sigma_0, \sigma_0^2)$ και έχουμε μια εκτός ελέγχου διεργασία (για $\delta = 0$ έχουμε μια εντός ελέγχου διεργασία).

Υπενθυμίζουμε ότι στο κλασικό άνω μονόπλευρο διάγραμμα ελέγχου CUSUM για την ανίχνευση αύξησης της μέσης τιμής η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα είναι της μορφής

$$C_i^+ = \max(0, C_{i-1}^+ + \bar{X}_i - \mu_0 - K_C^+),$$

όπου $\bar{X}_i$ ($i = 1, 2, \dots$) είναι ο δειγματικός μέσος του i -οστού δείγματος, $C_0^+ = c_0^+ \geq 0$, και K_C^+ είναι μια σταθερά αναφοράς. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$C_i^+ \geq H_C^+,$$

όπου H_C^+ είναι το άνω όριο ελέγχου.

Αντίστοιχα, στο κάτω μονόπλευρο διάγραμμα ελέγχου CUSUM για την ανίχνευση μείωσης της μέσης τιμής η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα είναι της μορφής

$$C_i^- = \min(0, C_{i-1}^- + \bar{X}_i - \mu_0 + K_C^-),$$

όπου $C_0^- = c_0^- \leq 0$. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$C_i^- \leq -H_C^-$$

όπου $-H_C^-$ είναι το κάτω όριο ελέγχου.

Οι σταθερές K_C^+ , K_C^- και τα όρια H_C^+ , H_C^- καθορίζονται κατά τη σχεδίαση του διαγράμματος (π.χ. από την επιθυμητή τιμή του εντός ελέγχου ARL).

Το διάγραμμα ελέγχου διαμέσου τύπου CUSUM προκύπτει από την προαναφερθείσα διαδικασία αντικαθιστώντας απλά το δειγματικό μέσο με τη δειγματική διάμεσο. Υπενθυμίζουμε ότι η δειγματική διάμεσος του i -οστού δείγματος ορίζεται από τη σχέση

$$\tilde{Y}_i = \begin{cases} X_{i,((n+1)/2)}, & \text{αν } n \text{ περιττός,} \\ \frac{X_{i,(n/2)} + X_{i,(n/2+1)}}{2}, & \text{αν } n \text{ άρτιος,} \end{cases}$$

όπου $X_{i,(k)}$ είναι η k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$) του δείγματος $(X_{i,1}, X_{i,2}, \dots, X_{i,n})$. Καθώς ο υπολογισμός της διαμέσου είναι απλούστερος για περιττό n , χωρίς βλάβη της γενικότητας θεωρείται στη συνέχεια ότι το n είναι περιττός αριθμός.

Στο άνω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM απεικονίζεται η στατιστική συνάρτηση

$$U_i^+ = \max(0, U_{i-1}^+ + \tilde{Y}_i - \mu_0 - K_U^+),$$

όπου $U_0^+ = u_0^+ \geq 0$ και K_U^+ είναι μια σταθερά αναφοράς. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$U_i^+ \geq H_U^+, \quad H_U^+ > 0,$$

όπου H_U^+ είναι το άνω όριο ελέγχου.

Αντίστοιχα, στο κάτω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM απεικονίζεται η στατιστική συνάρτηση

$$U_i^- = \min(0, U_{i-1}^- + \tilde{Y}_i - \mu_0 + K_U^-),$$

όπου $U_0^- = u_0^- \leq 0$ και K_U^- είναι μια σταθερά αναφοράς. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$U_i^- \leq -H_U^-, \quad H_U^- > 0,$$

όπου $-H_U^-$ είναι το κάτω όριο ελέγχου.

Για την ανάπτυξη τυποποιημένων διαγραμμάτων ελέγχου διαμέσου τύπου CUSUM, αρχικά τυποποιούμε τις παρατηρήσεις μας, δηλαδή εργαζόμαστε με τις παρατηρήσεις $Z_{i,j} = (X_{i,j} - \mu_0)/\sigma_0$. Η διάμεσος του i -οστού δείγματος τότε είναι ίση με

$$\tilde{Z}_i = Z_{i,((n+1)/2)}$$

(θεωρούμε ότι το n είναι περιττός αριθμός).

Στο τυποποιημένο άνω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM απεικονίζεται η στατιστική συνάρτηση

$$Q_i^+ = \max(0, Q_{i-1}^+ + \tilde{Z}_i - k_Q^+),$$

όπου $Q_0^+ = q_0^+ \geq 0$ και k_Q^+ είναι μια σταθερά αναφοράς. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$Q_i^+ \geq h_Q^+, \quad h_Q^+ > 0,$$

όπου h_Q^+ είναι το άνω όριο ελέγχου.

Αντίστοιχα, στο κάτω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM απεικονίζεται η στατιστική συνάρτηση

$$Q_i^- = \min(0, Q_{i-1}^- + \tilde{Z}_i + k_Q^-),$$

όπου $Q_0^- = q_0^- \leq 0$ και k_Q^- είναι μια σταθερά αναφοράς. Το διάγραμμα δίνει εκτός ελέγχου σήμα όταν

$$Q_i^- \leq -h_Q^-, \quad h_Q^- > 0,$$

όπου $-h_Q^-$ είναι το κάτω όριο ελέγχου.

3.2.2 Φάση Ι – Άγνωστες παράμετροι

Στην πράξη, οι εντός ελέγχου τιμές των παραμέτρων της διεργασίας μ_0 και σ_0 συνήθως δεν είναι γνωστές και πρέπει να εκτιμηθούν από δεδομένα Φάσης Ι. Έστω ότι είναι διαθέσιμα m ανεξάρτητα δείγματα μεγέθους n το καθένα, δηλαδή τα

$$X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n}), \quad i = 1, 2, \dots, m.$$

Τα δείγματα αυτά θα πρέπει να σχηματίζονται από παρατηρήσεις που λαμβάνονται διαδοχικά στον χρόνο, ώστε να επιτρέπεται η εκτίμηση της μεταβλητότητας μεταξύ των δειγμάτων. Παράλληλα, οι παρατηρήσεις εντός κάθε δείγματος πρέπει να λαμβάνονται την ίδια χρονική στιγμή από μία σταθερή διεργασία, ή όσο το δυνατόν πιο κοντά χρονικά, ώστε να εξασφαλίζεται η ανεξαρτησία τους και να αποτυπώνεται η μεταβλητότητα εντός του δείγματος.

Οι πιο συνηθισμένες εκτιμήτριες για τη μέση τιμή και την τυπική απόκλιση μιας κανονικής διεργασίας, είναι οι

$$\hat{\mu}_0 = \frac{1}{m} \sum_{i=1}^m \bar{X}_i = \bar{\bar{X}}, \quad \hat{\sigma}_0 = \frac{1}{c_4} \left(\frac{1}{m} \sum_{i=1}^m S_i \right) = \frac{\bar{S}}{c_4},$$

αντίστοιχα, όπου $\bar{X}_i$ και S_i είναι ο δειγματικός μέσος και η δειγματική τυπική απόκλιση του i –δείγματος.

Στα διαγράμματα που βασίζονται στη διάμεσο χρησιμοποιούνται πιο ανθεκτικές εκτιμήτριες, οι οποίες βασίζονται σε διατεταγμένες στατιστικές συναρτήσεις. Τέτοιες εκτιμήτριες είναι οι

$$\hat{\mu}'_0 = \frac{1}{m} \sum_{i=1}^m \tilde{X}_i, \\ \hat{\sigma}'_0 = \frac{1}{d_2} \left(\frac{1}{m} \sum_{i=1}^m R_i \right),$$

(δείτε Castagliola et al. (2019)) όπου $\tilde{X}_i$ είναι η δειγματική διάμεσος και $R_i = X_{i,(n)} - X_{i,(1)}$ είναι το εύρος του i –δείγματος.

Οι τυποποιημένες μορφές των μονόπλευρων διαγραμμάτων ελέγχου διαμέσου τύπου CUSUM με εκτιμημένες παραμέτρους δίνονται από τις σχέσεις

$$G_i^- = \min \left(0, G_{i-1}^- + \frac{\tilde{Y}_i - \hat{\mu}'_0}{\hat{\sigma}'_0} + k_G^- \right),$$

$$G_i^+ = \max \left(0, G_{i-1}^+ + \frac{\tilde{Y}_i - \hat{\mu}'_0}{\hat{\sigma}'_0} - k_G^+ \right),$$

με αρχικές τιμές $G_0^- = g_0^- \leq 0$ και $G_0^+ = g_0^+ \geq 0$, και k_G^-, k_G^+ είναι σταθερές αναφοράς. Το άνω διάγραμμα ελέγχου δίνει σήμα όταν

$$G_i^+ \geq h_G^+$$

και το κάτω όταν

$$G_i^- \leq -h_G^-.$$

Οι παραπάνω σχέσεις μπορούν να γραφούν ισοδύναμα ως

$$G_i^- = \min \left(0, G_{i-1}^- + \frac{\frac{\tilde{Y}_i - \mu_0}{\sigma_0} + \frac{\mu_0 - \hat{\mu}'_0}{\sigma_0}}{\frac{\hat{\sigma}'_0}{\sigma_0}} + k_G^- \right),$$

$$G_i^+ = \max \left(0, G_{i-1}^+ + \frac{\frac{\tilde{Y}_i - \mu_0}{\sigma_0} + \frac{\mu_0 - \hat{\mu}'_0}{\sigma_0}}{\frac{\hat{\sigma}'_0}{\sigma_0}} - k_G^+ \right).$$

Ορίζοντας τις τυχαίες μεταβλητές

$$V = \frac{\hat{\mu}'_0 - \mu_0}{\sigma_0}, \quad W = \frac{\hat{\sigma}'_0}{\sigma_0},$$

παίρνουμε ότι

$$G_i^+ = \max \left(0, G_{i-1}^+ + \frac{\frac{\tilde{Y}_i - \mu_0}{\sigma_0} - V}{W} - k_G^+ \right),$$

$$G_i^- = \min \left(0, G_{i-1}^- + \frac{\frac{\tilde{Y}_i - \mu_0}{\sigma_0} - V}{W} + k_G^- \right).$$

Έστω $\hat{f}_{RL}(\ell)$, $\hat{F}_{RL}(\ell)$, $\hat{\nu}_1$ και $\hat{\nu}_2$ η δεσμευμένη συνάρτηση πιθανότητας, η δεσμευμένη συνάρτηση κατανομής, η πρώτη και η δεύτερη δεσμευμένη καθοδική ροπή του μήκους ροής (RL) όταν είναι γνωστές οι τιμές των τυχαίων μεταβλητών V και W . Έστω επίσης

$f_{V,W}(v, w | m, n)$ η από κοινού συνάρτηση πυκνότητας πιθανότητας των V και W . Τότε σύμφωνα με τους Castagliola et al. (2019) ισχύει ότι

$$\begin{aligned} f_{RL}(\ell) &= \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{V,W}(v, w | m, n) \hat{f}_{RL}(\ell) dw dv, \\ F_{RL}(\ell) &= \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{V,W}(v, w | m, n) \hat{F}_{RL}(\ell) dw dv, \\ ARL &= E(RL) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{V,W}(v, w | m, n) \hat{v}_1 dw dv, \end{aligned}$$

και

$$SDRL = \sqrt{E(RL^2) - ARL^2},$$

όπου

$$E(RL^2) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{V,W}(v, w | m, n) (\hat{v}_1 + \hat{v}_2) dw dv.$$

Όταν έχουμε δείγμα από κανονική κατανομή, η δειγματική διάμεσος και το εύρος, αν και είναι εξαρτημένες τυχαίες μεταβλητές, ωστόσο είναι ασυσχέτιστες. Καθώς το πλήθος των δειγμάτων m αυξάνεται, λόγω του Κεντρικού Οριακού Θεωρήματος προκύπτει ότι οι στατιστικές συναρτήσεις $\hat{\mu}'_0$ και $\hat{\sigma}'_0$ έχουν ασυμπτωτική κατανομή τη διδιάστατη κανονική κατανομή και είναι ασυμπτωτικά ανεξάρτητες. Έτσι, η από κοινού συνάρτηση πυκνότητας του ζεύγους (V, W) μπορεί να προσεγγιστεί σύμφωνα με τη σχέση

$$f_{(V,W)}(v, w | m, n) \cong f_V(v | m, n) f_W(w | m, n).$$

Συνεπώς είναι εφικτός ο υπολογισμός του ARL και του $SDRL$ αρκεί να γνωρίζουμε τις (ασυμπτωτικές) κατανομές των τυχαίων μεταβλητών V και W . Ικανοποιητικές προσεγγίσεις τους έχουν προταθεί από τους Castagliola et al. (2019).

3.3 Απόδοση του διαγράμματος ελέγχου διαμέσου τύπου CUSUM

Η απόδοση του διαγράμματος ελέγχου διαμέσου τύπου CUSUM αξιολογείται ως γνωστόν μέσω της κατανομής του αριθμού των δειγμάτων (RL) μέχρι την εμφάνιση σήματος εκτός ελέγχου διεργασίας. Για τον υπολογισμό των χαρακτηριστικών της κατανομής αυτής, όπως το μέσο μήκος ροής και η τυπική απόκλιση, μπορεί να

χρησιμοποιηθεί, πέραν της προσομοίωσης, η μέθοδος μέσω των αλυσίδων Markov (δείτε Brook & Evans (1972)).

Ακολουθώντας αυτή την κλασική προσέγγιση, η κατανομή του RL για το άνω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM με γνωστές παραμέτρους μπορεί να προσδιοριστεί θεωρώντας μία αλυσίδα Markov με καταστάσεις

$$\{0, 1, \dots, r-1, r\},$$

όπου η κατάσταση r είναι απορροφητική, δηλαδή αντιστοιχεί στην εμφάνιση σήματος εκτός ελέγχου, και οι υπόλοιπες μεταβατικές. Το διάστημα $(0, H_U^+]$ χωρίζεται σε r υποδιαστήματα της μορφής

$$(H_j - \Delta, H_j + \Delta], \quad j = 0, 1, \dots, r-1,$$

όπου

$$H_j = (2j+1)\Delta, \quad \Delta = \frac{H_U^+}{2r}$$

(το H_j είναι η αντιπροσωπευτική τιμή της κατάστασης j της αλυσίδας). Η αλυσίδα εισέρχεται στην μεταβατική κατάσταση j στο δείγμα i όταν

$$U_i^+ \in (H_j - \Delta, H_j + \Delta],$$

ενώ μεταβαίνει στην απορροφητική κατάσταση r όταν το U_i^+ υπερβεί το όριο H_U^+ .

Έστω $\mathbf{Q}$ ο πίνακας μετάβασης διαστάσεων $r \times r$, ο οποίος περιέχει τις πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων της αλυσίδας Markov, δηλαδή

$$\mathbf{Q} = \begin{bmatrix} Q_{0,0} & Q_{0,1} & \cdots & Q_{0,r-1} \\ Q_{1,0} & Q_{1,1} & \cdots & Q_{1,r-1} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{r-1,0} & Q_{r-1,1} & \cdots & Q_{r-1,r-1} \end{bmatrix}.$$

Για $j = 0, 1, \dots, r-1$ και $k = 1, 2, \dots, r-1$ έχουμε ότι

$$\begin{aligned} Q_{j,k} &= P(U_i^+ = H_k \mid U_{i-1}^+ = H_j) \\ &= P(H_k - \Delta < U_i^+ \leq H_k + \Delta \mid U_{i-1}^+ = H_j) \\ &= P(H_k - \Delta < U_{i-1}^+ + \tilde{Y}_i - \mu_0 - K_U^+ \leq H_k + \Delta \mid U_{i-1}^+ = H_j) \\ &= P(H_k - \Delta < H_j + \tilde{Y}_i - \mu_0 - K_U^+ \leq H_k + \Delta) \\ &= P(H_k - H_j - \Delta + K_U^+ + \mu_0 < \tilde{Y}_i \leq H_k - H_j + \Delta + K_U^+ + \mu_0) \\ &= F_{\tilde{Y}}(H_k - H_j + \Delta + K_U^+ + \mu_0) - F_{\tilde{Y}}(H_k - H_j - \Delta + K_U^+ + \mu_0) \end{aligned}$$

όπου $F_{\tilde{Y}}(\cdot)$ είναι η συνάρτηση κατανομής της δειγματικής διαμέσου.

Για την περίπτωση $j = 0, 1, \dots, r-1$ και $k = 0$ έχουμε ότι

$$\begin{aligned}
Q_{j,0} &= P(U_i^+ = H_0 \mid U_{i-1}^+ = H_j) \\
&= P(U_i^+ \leq H_0 + \Delta \mid U_{i-1}^+ = H_j) \\
&= P(U_{i-1}^+ + \tilde{Y}_i - \mu_0 - K_U^+ \leq H_0 + \Delta \mid U_{i-1}^+ = H_j) \\
&= P(H_j + \tilde{Y}_i - \mu_0 - K_U^+ \leq H_0 + \Delta) \\
&= P(\tilde{Y}_i \leq H_0 + \Delta - H_j + K_U^+ + \mu_0) \\
&= P(\tilde{Y}_i \leq -H_j + 2\Delta + K_U^+ + \mu_0) \\
&= F_{\tilde{Y}}(-H_j + 2\Delta + K_U^+ + \mu_0).
\end{aligned}$$

Υπενθυμίζουμε ότι η συνάρτηση κατανομής της δειγματικής διαμέσου δίνεται από τον τύπο

$$F_{\tilde{Y}}(y) = F_B\left(\Phi\left(\frac{y - \mu_0}{\sigma_0} - \delta\right) \mid \frac{n+1}{2}, \frac{n+1}{2}\right),$$

όπου $\Phi(\cdot)$ είναι η συνάρτηση κατανομής της τυπικής κανονικής κατανομής και $F_B(\cdot)$ η συνάρτηση κατανομής της κατανομής Beta με παραμέτρους μορφής $\frac{n+1}{2}$ και $\frac{n+1}{2}$.

Έστω

$$\mathbf{q} = (q_0, q_1, \dots, q_{r-1})$$

το διάνυσμα των αρχικών πιθανοτήτων των μεταβατικών καταστάσεων. Η ακριβής μορφή του διανύσματος $\mathbf{q}$ εξαρτάται από την τιμή $U_0^+ = u_0^+$. Πράγματι

$$q_j = \begin{cases} 0, & U_0^+ = u_0^+ \notin (H_j - \Delta, H_j + \Delta] \\ 1, & U_0^+ = u_0^+ \in (H_j - \Delta, H_j + \Delta]. \end{cases}$$

Σημειώνεται ότι συνήθως $U_0^+ = u_0^+ = 0$, οπότε $\mathbf{q} = (1, 0, \dots, 0)$.

Η συνάρτηση πιθανότητας του RL δίνεται από τον τύπο

$$f_{RL}(\ell) = \mathbf{q} \mathbf{Q}^{\ell-1} \mathbf{c}, \quad \ell = 1, 2, \dots$$

και η συνάρτηση κατανομής από τον τύπο

$$F_{RL}(\ell) = 1 - \mathbf{q} \mathbf{Q}^{\ell} \mathbf{1}^T,$$

όπου $\mathbf{1} = (1, 1, \dots, 1)$ και $\mathbf{c} = (\mathbf{I} - \mathbf{Q}) \mathbf{1}^T$. Επίσης

$$ARL = E(RL) = \mathbf{q} (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}^T,$$

$$E(RL(RL - 1)) = 2 \mathbf{q}(\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1}^T,$$

$$SDRL = \sqrt{E(RL(RL - 1)) + E(RL) - E^2(RL)}.$$

Με βάση τις παραπάνω σχέσεις, είναι δυνατός ο ακριβής υπολογισμός των χαρακτηριστικών της κατανομής του RL μέσω της προσέγγισης με αλυσίδες Markov.

Στη συνέχεια εξετάζεται η απόδοση του τυποποιημένου άνω διαγράμματος ελέγχου διαμέσου τύπου CUSUM Φάσης II για διάφορους συνδυασμούς των παραμέτρων του (n, ARL_0, k_Q^+, h_Q^+) . Λόγω της συμμετρίας της κανονικής κατανομής η απόδοση του άνω και κάτω διαγράμματος ελέγχου διαμέσου τύπου CUSUM είναι παρόμοια για την ανίχνευση είτε μιας ανοδικής είτε μιας καθοδικής αλλαγής του ίδιου μεγέθους στην μέση τιμή της διαδικασίας. Επομένως, για το άνω και κάτω διάγραμμα ελέγχου με το ίδιο n και ARL_0 έχουμε ότι $k = k_Q^+ = k_Q^-$ και $h = h_Q^+ = h_Q^-$ και θα μας απασχολήσουν μόνο ανοδικές μετατοπίσεις της μέσης τιμής.

Στους Πίνακες 3.1 και 3.2 για δεδομένο μέγεθος δείγματος n (3, 5, 7, 9) και τιμής αναφοράς k (0.05, 0.1, 0.25, 0.5, 0.75, 1), δίνεται η τιμή του διαστήματος h έτσι ώστε να επιτυγχάνεται προκαθορισμένο εντός ελέγχου μέσο μήκος ροής $ARL_0 = 370.4$ και $ARL_0 = 500$, αντίστοιχα. Για τον υπολογισμό του h χρησιμοποιήθηκε η μέθοδος των αλυσίδων Markov με $r = 400$.

Πίνακας 3.1. Τιμές h για $ARL_0 = 370.4$

k	$n = 3$	$n = 5$	$n = 7$	$n = 9$
0.05	8.0136	5.9093	4.7542	4.0111
0.1	6.0062	4.2707	3.3489	2.7701
0.25	3.4375	2.3227	1.7563	1.4117
0.5	1.9499	1.2588	0.9200	0.7179
0.75	1.3081	0.8084	0.5559	0.3986
1	0.9305	0.5107	0.2857	0.1396

Πίνακας 3.2. Τιμές h για $ARL_0 = 500$

k	$n = 3$	$n = 5$	$n = 7$	$n = 9$
0.05	8.9563	6.5629	5.2571	4.4209
0.1	6.5851	4.6561	3.6381	3.0016
0.25	3.6970	2.4907	1.8804	1.5101
0.5	2.0836	1.3443	0.9830	0.7685
0.75	1.3982	0.8696	0.6045	0.4402
1	1.0046	0.5656	0.3315	0.1798

Στη συνέχεια, χρησιμοποιώντας τους παραπάνω συνδυασμούς τιμών (k, h) υπολογίστηκε το (εκτός ελέγχου) μέσο μήκος ροής ARL των διαγραμμάτων ελέγχου CUSUM για διάφορες τιμές της μετατόπισης δ (0, 0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2), Τα αποτελέσματα παρουσιάζονται στους Πίνακες 3.3 και 3.4 για $ARL_0 = 370.4$ και $ARL_0 = 500$, αντίστοιχα. Με έντονη γραφή σημειώνεται η μικρότερη τιμή ARL για κάθε μετατόπιση.

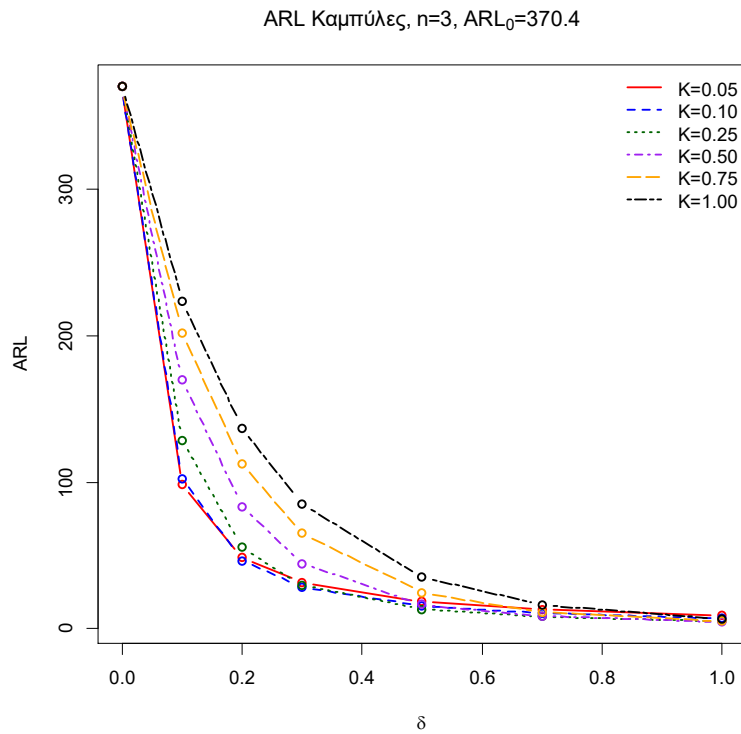
Πίνακας 3.3. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM
($ARL_0 = 370.4$)

δ	$n = 3$					
	$k = 0.05$ $h = 8.0136$	$k = 0.1$ $h = 6.0062$	$k = 0.25$ $h = 3.4375$	$k = 0.5$ $h = 1.9499$	$k = 0.75$ $h = 1.3081$	$k = 1$ $h = 0.9305$
0.0	370.39	370.40	370.41	370.41	370.40	370.41
0.1	98.65	102.49	128.74	170.03	202.00	223.56
0.2	48.65	46.48	55.56	83.50	112.85	136.83
0.3	31.58	28.33	29.65	44.52	65.12	85.15
0.5	18.47	15.60	13.33	16.61	24.63	35.14
0.7	13.06	10.75	8.31	8.56	11.40	16.20
1.0	9.11	7.36	5.31	4.62	5.14	6.52
1.5	6.12	4.89	3.37	2.63	2.46	2.60
2.0	4.66	3.72	2.53	1.89	1.64	1.59
δ	$n = 5$					
	$k = 0.05$ $h = 5.9093$	$k = 0.1$ $h = 4.2707$	$k = 0.25$ $h = 2.3227$	$k = 0.5$ $h = 1.2588$	$k = 0.75$ $h = 0.8084$	$k = 1$ $h = 0.5107$
0.0	370.40	370.40	370.41	370.40	370.40	370.40
0.1	79.08	83.40	112.39	158.45	192.64	208.97
0.2	37.17	35.05	43.89	72.51	102.67	120.80
0.3	23.85	20.90	22.09	36.25	56.52	71.62
0.5	13.86	11.39	9.53	12.35	19.58	27.38
0.7	9.80	7.84	5.90	6.11	8.49	11.94
1.0	6.85	5.39	3.78	3.26	3.65	4.57
1.5	4.64	3.62	2.45	1.87	1.75	1.81
2.0	3.56	2.79	1.92	1.33	1.21	1.20
δ	$n = 7$					
	$k = 0.05$ $h = 4.7542$	$k = 0.1$ $h = 3.3489$	$k = 0.25$ $h = 1.7563$	$k = 0.5$ $h = 0.9200$	$k = 0.75$ $h = 0.5559$	$k = 1$ $h = 0.2857$
0.0	370.39	370.41	370.41	370.40	370.40	370.40
0.1	67.01	71.55	101.94	151.20	183.48	194.15
0.2	30.57	28.56	37.04	65.90	93.84	105.64
0.3	19.49	16.81	17.89	31.46	49.84	59.68
0.5	11.30	9.11	7.52	10.04	16.28	21.39
0.7	7.99	6.28	4.65	4.83	6.78	9.03
1.0	5.61	4.34	3.00	2.56	2.85	3.42
1.5	3.82	2.95	2.00	1.48	1.40	1.43
2.0	2.99	2.24	1.51	1.10	1.07	1.06
δ	$n = 9$					
	$k = 0.05$ $h = 4.0111$	$k = 0.1$ $h = 2.7701$	$k = 0.25$ $h = 1.4117$	$k = 0.5$ $h = 0.7179$	$k = 0.75$ $h = 0.3986$	$k = 1$ $h = 0.1396$
0.0	370.40	370.40	370.40	370.40	370.39	370.40
0.1	58.66	63.29	94.46	145.78	174.14	180.56
0.2	26.22	24.30	32.42	61.19	85.51	92.55
0.3	16.65	14.17	15.18	28.19	44.01	49.88
0.5	9.64	7.67	6.27	8.55	13.77	16.85
0.7	6.83	5.29	3.87	4.03	5.60	6.96
1.0	4.81	3.67	2.52	2.13	2.34	2.67
1.5	3.30	2.50	1.68	1.25	1.21	1.23
2.0	2.58	2.02	1.20	1.03	1.02	1.02

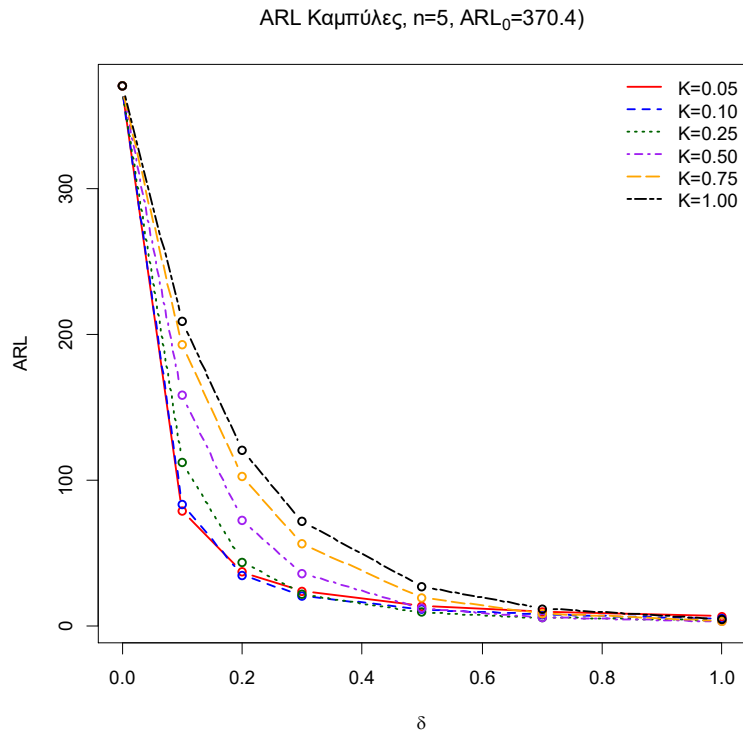
Πίνακας 3.4. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM
($ARL_0 = 500$)

δ	$n = 3$					
	$k = 0.05$ $h = 8.9563$	$k = 0.1$ $h = 6.5851$	$k = 0.25$ $h = 3.6970$	$k = 0.5$ $h = 2.0836$	$k = 0.75$ $h = 1.3982$	$k = 1$ $h = 1.0046$
0.0	500	500	500	499.99	500	500
0.1	115.08	120.71	158.56	217.44	262.80	294.57
0.2	54.91	52.01	64.03	101.62	141.61	175.82
0.3	35.35	31.22	32.87	51.91	78.95	106.64
0.5	20.56	17.04	14.36	18.28	28.14	41.84
0.7	14.51	11.71	8.89	9.17	12.50	18.45
1.0	10.10	8.01	5.66	4.89	5.47	7.09
1.5	6.77	5.31	3.58	2.76	2.59	2.74
2.0	5.15	4.02	2.67	1.99	1.72	1.66
δ	$n = 5$					
	$k = 0.05$ $h = 6.5629$	$k = 0.1$ $h = 4.6561$	$k = 0.25$ $h = 2.4907$	$k = 0.5$ $h = 1.3443$	$k = 0.75$ $h = 0.8696$	$k = 1$ $h = 0.5656$
0.0	500.01	500.01	499.99	500	500	500
0.1	90.94	97.04	137.79	202.36	251.43	276.17
0.2	41.52	38.79	50.18	87.98	129.43	156.09
0.3	26.46	22.82	24.28	42.07	68.83	90.38
0.5	15.31	12.35	10.20	13.50	22.40	32.90
0.7	10.80	8.48	6.28	6.51	9.29	13.70
1.0	7.54	5.82	4.01	3.44	3.88	4.99
1.5	5.09	3.89	2.58	1.96	1.83	1.91
2.0	3.90	3.00	2.01	1.40	1.25	1.23
δ	$n = 7$					
	$k = 0.05$ $h = 5.2571$	$k = 0.1$ $h = 3.6381$	$k = 0.25$ $h = 1.8804$	$k = 0.5$ $h = 0.9830$	$k = 0.75$ $h = 0.6045$	$k = 1$ $h = 0.3315$
0.0	500.01	500	500.01	500	499.99	500
0.1	76.33	82.59	124.64	193.10	240.16	256.70
0.2	33.92	31.38	42.13	79.87	118.87	136.67
0.3	21.50	18.25	19.56	36.43	61.04	75.48
0.5	12.42	9.83	8.02	10.93	18.71	25.80
0.7	8.77	6.76	4.92	5.13	7.44	10.40
1.0	6.14	4.66	3.16	2.69	3.02	3.74
1.5	4.17	3.15	2.09	1.54	1.45	1.49
2.0	3.21	2.40	1.62	1.13	1.08	1.08
δ	$n = 9$					
	$k = 0.05$ $h = 4.4209$	$k = 0.1$ $h = 3.0016$	$k = 0.25$ $h = 1.5101$	$k = 0.5$ $h = 0.7685$	$k = 0.75$ $h = 0.4402$	$k = 1$ $h = 0.1798$
0.0	500	500	500	500	499.99	500
0.1	66.36	72.64	115.27	186.39	228.34	238.54
0.2	28.94	26.57	36.75	74.23	108.67	119.59
0.3	18.28	15.33	16.53	32.64	54.12	63.00
0.5	10.55	8.24	6.66	9.29	15.89	20.31
0.7	7.46	5.68	4.09	4.27	6.15	8.02
1.0	5.24	3.93	2.65	2.24	2.48	2.91
1.5	3.58	2.68	1.78	1.30	1.25	1.26
2.0	2.84	2.08	1.28	1.04	1.02	1.02

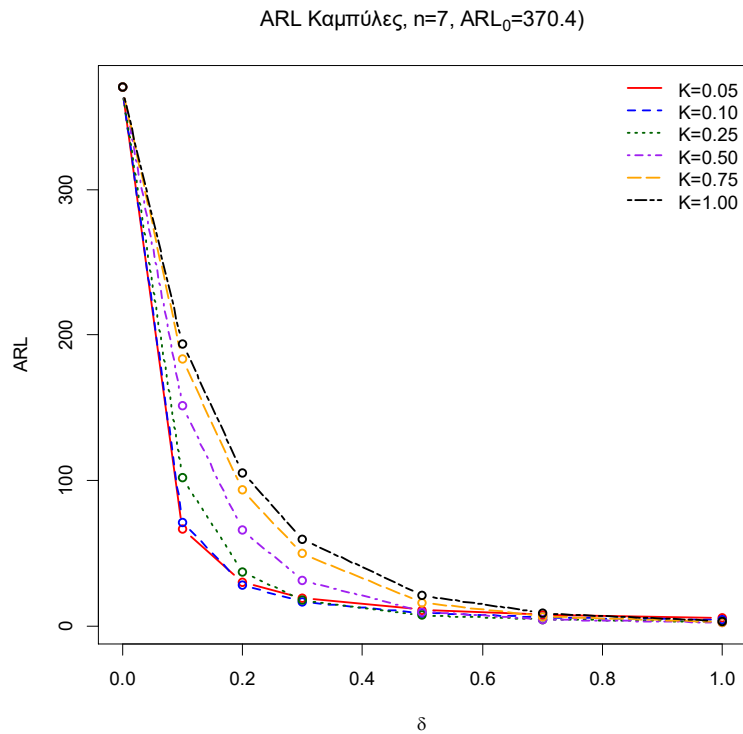
Ειδικότερα, στα Σχήματα 3.1-3.4 απεικονίζονται γραφικά οι τιμές ARL του Πίνακα 3.3 ($ARL_0 = 370.4$) για τις διαφορετικές τιμές του μεγέθους του δείγματος και για $\delta \leq 1$ όπου παρατηρείται η πιο ουσιαστική διαφοροποίηση της συμπεριφοράς του διαγράμματος.



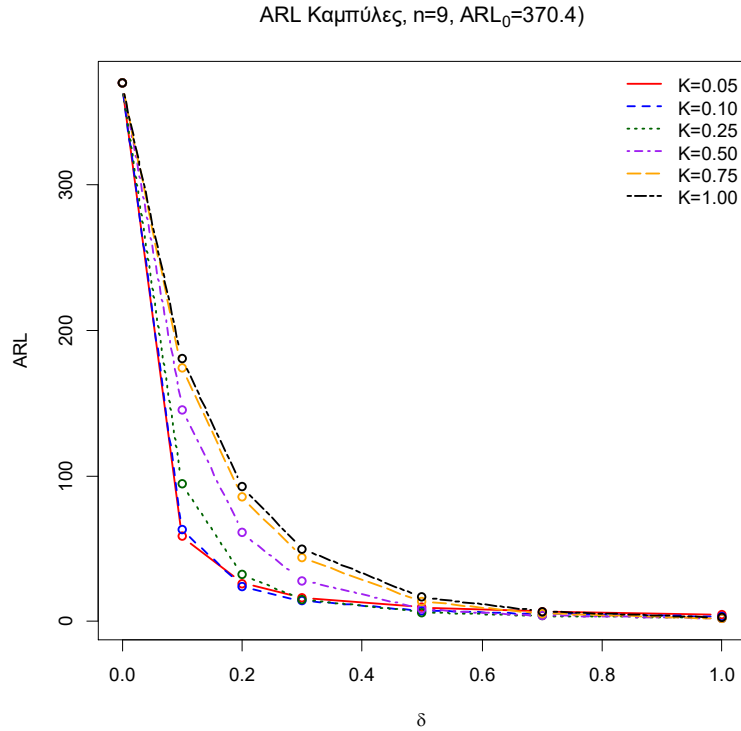
Σχήμα 3.1. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 3, ARL_0 = 370.4$)



Σχήμα 3.2. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 5, ARL_0 = 370.4$)



Σχήμα 3.3. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 7, ARL_0 = 370.4$)



Σχήμα 3.4. Καμπύλες ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου CUSUM ($n = 9, ARL_0 = 370.4$)

Όλες οι καμπύλες ξεκινούν από την ίδια τιμή $ARL_0 = 370.4$ όταν $\delta = 0$. Καθώς η μετατόπιση δ αυξάνεται, το ARL μειώνεται σημαντικά για όλες τις τιμές του k , υποδεικνύοντας ότι το διάγραμμα ανιχνεύει ταχύτερα αποκλίσεις από την κατάσταση εντός ελέγχου. Ωστόσο, η ταχύτητα με την οποία μειώνεται το ARL εξαρτάται έντονα από την τιμή του k , ιδιαίτερα για μικρές και μέτριες μετατοπίσεις.

Πιο συγκεκριμένα, για μικρές τιμές της μετατόπισης ($\delta \leq 0.4$), παρατηρούμε ότι μικρότερες τιμές του k οδηγούν σε μικρότερες τιμές του ARL , γεγονός που δείχνει αυξημένη ευαισθησία του διαγράμματος στην ανίχνευση μικρών μετατοπίσεων. Αντίθετα, για μεγαλύτερες τιμές του k , το ARL παραμένει υψηλότερο στις ίδιες περιοχές του δ , γεγονός που σημαίνει ότι το διάγραμμα είναι λιγότερο ευαίσθητο σε μικρές αποκλίσεις.

Καθώς η μετατόπιση αυξάνεται ($\delta \geq 0.5$), οι διαφορές μεταξύ των καμπυλών μειώνονται σημαντικά και όλες οι τιμές του k οδηγούν σε σχετικά μικρά ARL . Αυτό δείχνει ότι για μεγαλύτερες μετατοπίσεις η επιλογή της σταθεράς k έχει μικρότερη επίδραση στην απόδοση του διαγράμματος, καθώς όλες οι ρυθμίσεις επιτρέπουν γρήγορη ανίχνευση της εκτός ελέγχου κατάστασης.

Τέλος, παρατηρούμε ότι η επίδραση του k είναι μεγαλύτερη για μικρότερα μεγέθη δείγματος, ενώ καθώς το n αυξάνεται, οι καμπύλες συγκλίνουν ταχύτερα και η διαφορά μεταξύ των διαφορετικών τιμών του k μειώνεται. Αυτό υποδηλώνει ότι η επιλογή του k είναι πιο κρίσιμη σε περιπτώσεις μικρών δειγμάτων, ενώ για μεγαλύτερα μεγέθη δείγματος η απόδοση του διαγράμματος γίνεται λιγότερο ευαίσθητη στη συγκεκριμένη παράμετρο.

3.4 Εφαρμογή διαγράμματος ελέγχου διαμέσου τύπου CUSUM

Εδώ θα χρησιμοποιηθεί ξανά το σύνολο δεδομένων του Montgomery (2013) που αναφέρεται στην εσωτερική διάμετρο δακτυλίων εμβόλων (piston rings) για να επιδειχθεί η εφαρμογή του άνω διαγράμματος ελέγχου διαμέσου τύπου CUSUM Φάσης II. Για λόγους πληρότητας αλλά και σύγκρισης τα δεδομένα θα μελετηθούν και με το κλασικό CUSUM $\bar{X}$ διάγραμμα ελέγχου.

Τα συνολικά δεδομένα της εσωτερικής διαμέτρου των δακτυλίων εμβόλων σε mm αποτελούνται από 40 υποομάδες (δείγματα) μεγέθους $n = 5$. Θα υποθέσουμε ότι η εντός ελέγχου κατανομή των δεδομένων είναι κανονική με $\mu_0 = 74$ και $\sigma_0 = 0.01$ (οι τιμές αυτές είναι σχεδόν ίδιες με τις εκτιμήσεις των παραμέτρων αυτών που βρέθηκαν στην Παράγραφο 2.7). Στο κλασικό άνω μονόπλευρο CUSUM $\bar{X}$ διάγραμμα ελέγχου απεικονίζεται στο διάγραμμα η στατιστική

$$C_i^+ = \max(0, C_{i-1}^+ + \bar{X}_i - \mu_0 - K_C^+)$$

όπου $C_0^+ = 0$, $K_C^+ = k \frac{\sigma_0}{\sqrt{n}}$ και $H_C^+ = h \frac{\sigma_0}{\sqrt{n}}$. Για $k = 0.5$ έχουμε ότι

$$K_C^+ = 0.5 \frac{0.01}{\sqrt{5}} = 0.002236068.$$

Για να επιτευχθεί $ARL_0 = 370.4$ θα πρέπει $H_C^+ = 0.01832924$.

Στο τυποποιημένο άνω μονόπλευρο διάγραμμα ελέγχου διαμέσου τύπου CUSUM απεικονίζεται η στατιστική συνάρτηση

$$Q_i^+ = \max(0, Q_{i-1}^+ + \bar{Z}_i - k_Q^+),$$

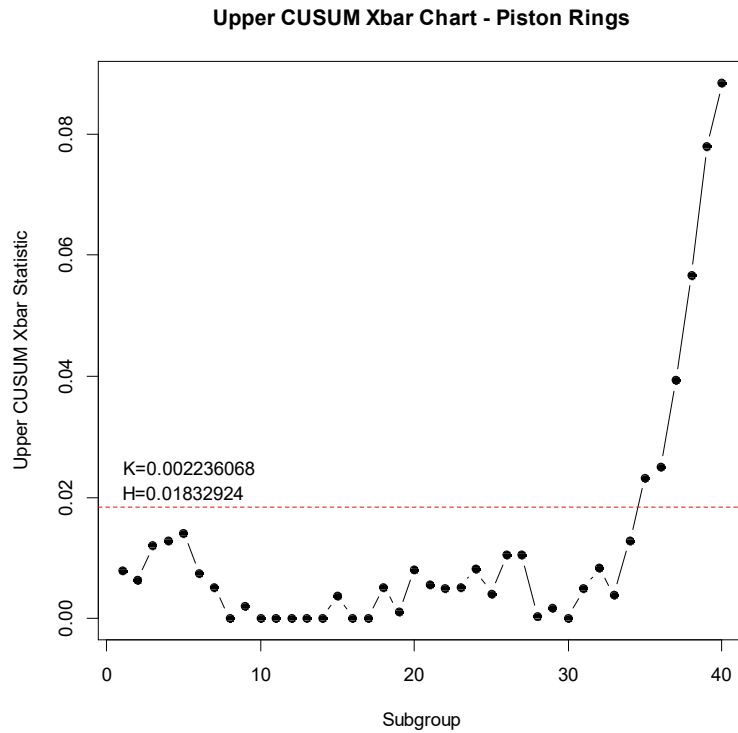
όπου $Q_0^+ = 0$. Για το σχεδιασμό του διαγράμματος επιλέγουμε $k_Q^+ = 0.5$. Για να επιτευχθεί $ARL_0 = 370.4$ θα πρέπει $h_Q^+ = 1.25882$.

Οι τιμές των στατιστικών συναρτήσεων C_i^+ και Q_i^+ δίνονται στον Πίνακα 3.5.

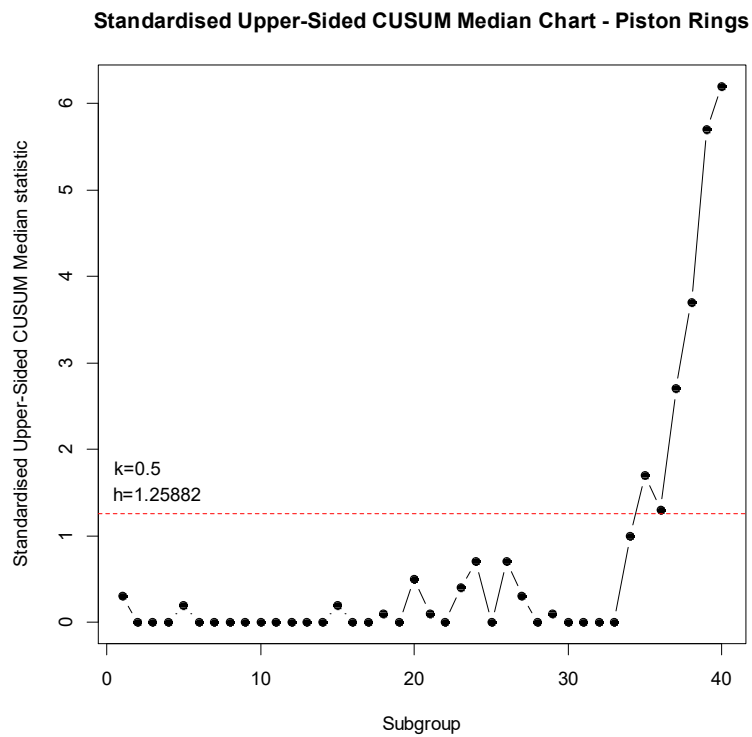
Πίνακας 3.5. Τιμές C_i^+ και Q_i^+ διαγραμμάτων CUSUM

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	$\bar{X}_i$	C_i^+	Q_i^+
1	74.030	74.002	74.019	73.992	74.008	74.0102	0.0080	0.3
2	73.995	73.992	74.001	74.011	74.004	74.0006	0.0063	0
3	73.988	74.024	74.021	74.005	74.002	74.008	0.0121	0
4	74.002	73.996	73.993	74.015	74.009	74.003	0.0129	0
5	73.992	74.007	74.015	73.989	74.014	74.0034	0.0140	0.2
6	74.009	73.994	73.997	73.985	73.993	73.9956	0.0074	0
7	73.995	74.006	73.994	74.000	74.005	74.000	0.0051	0
8	73.985	74.003	73.993	74.015	73.988	73.9968	0.0000	0
9	74.008	73.995	74.009	74.005	74.004	74.0042	0.0020	0
10	73.988	74.000	73.990	74.007	73.995	73.996	0	0
11	73.994	73.998	73.994	73.995	73.990	73.9942	0	0
12	74.004	74.000	74.007	74.000	73.996	74.0014	0	0
13	73.983	74.002	73.998	73.997	74.012	73.9984	0	0
14	74.006	73.967	73.994	74.000	73.984	73.9902	0	0
15	74.012	74.014	73.998	73.999	74.007	74.006	0.0038	0.2
16	74.000	73.984	74.005	73.998	73.996	73.9966	0	0
17	73.994	74.012	73.986	74.005	74.007	74.0008	0	0
18	74.006	74.010	74.018	74.003	74.000	74.0074	0.0052	0.1
19	73.984	74.002	74.003	74.005	73.997	73.9982	0.0011	0
20	74.000	74.010	74.013	74.020	74.003	74.0092	0.0081	0.5
21	73.982	74.001	74.015	74.005	73.996	73.9998	0.0057	0.1
22	74.004	73.999	73.990	74.006	74.009	74.0016	0.0050	1.48×10^{-12}
23	74.010	73.989	73.990	74.009	74.014	74.0024	0.0052	0.4
24	74.015	74.008	73.993	74.000	74.010	74.0052	0.0081	0.7
25	73.982	73.984	73.995	74.017	74.013	73.9982	0.0041	0
26	74.012	74.015	74.030	73.986	74.000	74.0086	0.0105	0.7
27	73.995	74.010	73.990	74.015	74.001	74.0022	0.0104	0.3
28	73.987	73.999	73.985	74.000	73.990	73.9922	0.0004	0
29	74.008	74.010	74.003	73.991	74.006	74.0036	0.0018	0.1
30	74.003	74.000	74.001	73.986	73.997	73.9974	0.0000	0
31	73.994	74.003	74.015	74.02	74.004	74.0072	0.0050	0
32	74.008	74.002	74.018	73.995	74.005	74.0056	0.0083	0
33	74.001	74.004	73.990	73.996	73.998	73.9978	0.0039	0
34	74.015	74.000	74.016	74.025	74.000	74.0112	0.0129	1
35	74.030	74.005	74.000	74.016	74.012	74.0126	0.0232	1.7
36	74.001	73.99	73.995	74.01	74.024	74.004	0.0250	1.3
37	74.015	74.02	74.024	74.005	74.019	74.0166	0.0393	2.7
38	74.035	74.01	74.012	74.015	74.026	74.0196	0.0567	3.7
39	74.017	74.013	74.036	74.025	74.026	74.0234	0.0779	5.7
40	74.010	74.005	74.029	74.000	74.020	74.0128	0.0884	6.2

Τα αντίστοιχα διαγράμματα δίνονται στα Σχήματα 3.5 και 3.6.



Σχήμα 3.5. Άνω μονόπλευρο CUSUM $\bar{X}$ διάγραμμα ελέγχου Φάσης II



Σχήμα 3.6. Άνω μονόπλευρο CUSUM διάγραμμα ελέγχου διαμέσου Φάσης II

Παρατηρούμε ότι και στα δύο διαγράμματα ελέγχου προκύπτει ότι σήμα εκτός ελέγχου διεργασίας παίρνουμε για πρώτη φορά στο 35ο δείγμα. Παρά το γεγονός ότι τα δύο διαγράμματα βασίζονται σε διαφορετικά στατιστικά μέτρα (δειγματικός μέσος και δειγματική διάμεσος), η χρονική στιγμή εμφάνισης του σήματος είναι ίδια. Επομένως, στη συγκεκριμένη περίπτωση και τα δύο διαγράμματα είναι σε θέση να ανιχνεύσουν αποτελεσματικά τη μεταβολή της διεργασίας.

Ωστόσο, παρατηρούμε διαφορές στη μορφή της εξέλιξης της στατιστικής συνάρτησης που απεικονίζεται στο διάγραμμα πριν και μετά το σήμα. Το διάγραμμα που βασίζεται στη δειγματική διάμεσο εμφανίζει πιο ομαλή συμπεριφορά, γεγονός που συνδέεται με τη μικρότερη ευαισθησία της διαμέσου σε ακραίες τιμές. Αντίθετα, το διάγραμμα που βασίζεται στον δειγματικό μέσο μπορεί να παρουσιάζει μεγαλύτερες διακυμάνσεις, γεγονός που οφείλεται στην αυξημένη επίδραση των απομακρυσμένων παρατηρήσεων στις τιμές του.

Συνολικά, τα αποτελέσματα δείχνουν ότι και οι δύο προσεγγίσεις είναι κατάλληλες για την ανίχνευση μεταβολών της διεργασίας, με τη χρήση της διαμέσου να προσφέρει ένα πιο ανθεκτικό εργαλείο σε περιπτώσεις παρουσίας ακραίων τιμών.

ΚΕΦΑΛΑΙΟ 4

Διαγράμματα ελέγχου διαμέσου τύπου EWMA

4.1 Εισαγωγή

Στο παρόν κεφάλαιο εξετάζεται το EWMA διάγραμμα ελέγχου διαμέσου, το οποίο αποτελεί μία εναλλακτική προσέγγιση για την παρακολούθηση μικρών έως μέτριων μεταβολών της διεργασίας.

Η βασική ιδέα του διαγράμματος EWMA είναι η εκθετικά σταθμισμένη συσσώρευση της δειγματικής πληροφορίας, έτσι ώστε να δίνεται μεγαλύτερη βαρύτητα στις πρόσφατες παρατηρήσεις, χωρίς όμως να αγνοείται πλήρως η παλαιότερη πληροφορία. Έτσι, το διάγραμμα είναι ιδιαίτερα αποτελεσματικό στην ανίχνευση μικρών και μέτριων μετατοπίσεων της μέσης τιμής.

Στην κλασική τους μορφή, τα διαγράμματα EWMA για ανίχνευση μετατοπίσεων της θέσης της διεργασίας βασίζονται στον δειγματικό μέσο. Ωστόσο, η χρήση του δειγματικού μέσου ενδέχεται να επηρεάζεται από την παρουσία ακραίων τιμών ή αποκλίσεων από την κανονικότητα. Για τον λόγο αυτό, στο παρόν κεφάλαιο εξετάζεται η χρήση της δειγματικής διαμέσου στο πλαίσιο ενός διαγράμματος ελέγχου τύπου EWMA.

Όπως και στην περίπτωση του διαγράμματος ελέγχου CUSUM, η ανάλυση θα πραγματοποιηθεί τόσο για γνωστές όσο και για εκτιμώμενες παραμέτρους της διεργασίας. Η απόδοση του διαγράμματος θα αξιολογηθεί με βάση την κατανομή του μήκους ροής και ειδικότερα με το μέσο μήκος ροής του αξιοποιώντας την μεθοδολογία των αλυσίδων Markov. Η ανάπτυξη του παρόντος κεφαλαίου βασίστηκε στα άρθρα των Castagliola et al. (2016) και Hu et al. (2019).

4.2 Διάγραμμα ελέγχου διαμέσου τύπου EWMA

Εδώ παρουσιάζεται το διάγραμμα ελέγχου διαμέσου τύπου EWMA, το οποίο χρησιμοποιείται για την παρακολούθηση της παραμέτρου θέσης μιας διεργασίας. Η ανάλυση του διαγράμματος γίνεται αρχικά υπό την υπόθεση γνωστών παραμέτρων της διεργασίας, περίπτωση η οποία αντιστοιχεί σε λειτουργία Φάσης II. Στη συνέχεια εξετάζεται η περίπτωση όπου οι παράμετροι είναι άγνωστες και εκτιμώνται από δεδομένα Φάσης I, γεγονός που επηρεάζει σημαντικά τη συμπεριφορά του διαγράμματος.

4.2.1 Φάση II – Γνωστές Παράμετροι

Έστω ότι η εντός ελέγχου κατανομή των μετρήσεων που παρακολουθούμε είναι η κανονική κατανομή με μέση τιμή μ_0 και διασπορά σ_0^2 . Από τη διεργασία επιλέγουμε τυχαία δείγματα μεγέθους n σε τακτά χρονικά διαστήματα. Έστω $\mathbf{X}_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n})$ το δείγμα που πήραμε στην i -οστή δειγματοληψία ($i = 1, 2, \dots$). Όταν έχουμε μια εντός ελέγχου διεργασία, ισχύει ότι $X_{i,j} \sim N(\mu_0, \sigma_0^2)$ για $i = 1, 2, \dots$ και $j = 1, 2, \dots, n$. Στην περίπτωση που η μέση τιμή μετατοπιστεί στη θέση $\mu_1 = \mu_0 + \delta\sigma_0$, $\delta \neq 0$, τότε $X_{i,j} \sim N(\mu_0 + \delta\sigma_0, \sigma_0^2)$ και έχουμε μια εκτός ελέγχου διεργασία.

Έστω $\tilde{Y}_i$ η δειγματική διάμεσος του i -οστού δείγματος, δηλαδή

$$\tilde{Y}_i = \begin{cases} X_{i,((n+1)/2)}, & \text{αν } n \text{ περιττός,} \\ \frac{X_{i,(n/2)} + X_{i,(n/2+1)}}{2}, & \text{αν } n \text{ άρτιος,} \end{cases}$$

όπου $X_{i,(k)}$ είναι η k -οστή διατεταγμένη τυχαία μεταβλητή ($1 \leq k \leq n$) του δείγματος $(X_{i,1}, X_{i,2}, \dots, X_{i,n})$. Στη συνέχεια, χωρίς βλάβη της γενικότητας, θεωρούμε ότι το μέγεθος του δείγματος n είναι περιττός αριθμός, οπότε ισχύει

$$\tilde{Y}_i = X_{i,((n+1)/2)}.$$

Στο διάγραμμα ελέγχου διαμέσου τύπου EWMA απεικονίζεται η στατιστική συνάρτηση

$$Z_i = \lambda \tilde{Y}_i + (1 - \lambda)Z_{i-1}, \quad 0 < \lambda \leq 1,$$

όπου η αρχική τιμή Z_0 τίθεται ίση με την τιμή στόχο, δηλαδή

$$Z_0 = \mu_0.$$

Αν οι παράμετροι μ_0 και σ_0 είναι γνωστές, τότε τα όρια ελέγχου του διαγράμματος ελέγχου διαμέσου τύπου EWMA είναι

$$LCL = \mu_0 - K\sigma_0,$$

$$UCL = \mu_0 + K\sigma_0,$$

όπου $K > 0$ είναι η σταθερά που καθορίζει το εύρος των ορίων ελέγχου.

Η κατανομή του μήκους ροής του διαγράμματος ελέγχου διαμέσου EWMA μπορεί να προσεγγιστεί μέσω της θεωρίας των αλυσίδων Markov, όπως προτάθηκε από τους Brook & Evans (1972). Για τον σκοπό αυτό, το διάστημα $[LCL, UCL]$ χωρίζεται σε $g = 2m + 1$ υποδιαστήματα της μορφής

$$(H_j - \Delta, H_j + \Delta], \quad j = -m, \dots, 0, \dots, m,$$

όπου το πλάτος κάθε υποδιαστήματος είναι

$$2\Delta = \frac{UCL - LCL}{g}.$$

Το σημείο H_j που είναι το μέσο του j υποδιαστήματος και δίνεται από τη σχέση

$$H_j = \frac{LCL + UCL}{2} + 2j\Delta.$$

Κάθε υποδιάστημα αναπαριστά μια μεταβατική κατάσταση της αλυσίδας Markov. Αν

$$H_j - \Delta < Z_i < H_j + \Delta$$

για $j = -m, \dots, 0, \dots, m$, τότε η αλυσίδα Markov εισέρχεται στην κατάσταση j για το δείγμα i . Σε διαφορετική περίπτωση, δηλαδή όταν $Z_i < LCL$ ή $Z_i > UCL$, τότε η αλυσίδα Markov εισέρχεται σε μια απορροφητική κατάσταση.

Έστω Q ο πίνακας μετάβασης διαστάσεων $(2m + 1) \times (2m + 1)$, ο οποίος περιέχει τις πιθανότητες μετάβασης μεταξύ των μεταβατικών καταστάσεων της αλυσίδας Markov, δηλαδή

$$\mathbf{Q} = \begin{bmatrix} Q_{-m,-m} & \cdots & Q_{-m,-1} & Q_{-m,0} & Q_{-m,+1} & \cdots & Q_{-m,+m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q_{-1,-m} & \cdots & Q_{-1,-1} & Q_{-1,0} & Q_{-1,+1} & \cdots & Q_{-1,+m} \\ Q_{0,-m} & \cdots & Q_{0,-1} & Q_{0,0} & Q_{0,+1} & \cdots & Q_{0,+m} \\ Q_{+1,-m} & \cdots & Q_{+1,-1} & Q_{+1,0} & Q_{+1,+1} & \cdots & Q_{+1,+m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q_{+m,-m} & \cdots & Q_{+m,-1} & Q_{+m,0} & Q_{+m,+1} & \cdots & Q_{+m,+m} \end{bmatrix}.$$

Για $j, k = -m, \dots, 0, \dots, m$, έχουμε ότι

$$\begin{aligned} Q_{j,k} &= P(Z_i = H_k \mid Z_{i-1} = H_j) \\ &= P(H_k - \Delta < Z_i \leq H_k + \Delta \mid Z_{i-1} = H_j) \\ &= P(H_k - \Delta < \lambda \tilde{Y}_i + (1 - \lambda)Z_{i-1} \leq H_k + \Delta \mid Z_{i-1} = H_j) \\ &= P(H_k - \Delta < \lambda \tilde{Y}_i + (1 - \lambda)H_j \leq H_k + \Delta) \\ &= P(H_k - \Delta - (1 - \lambda)H_j < \lambda \tilde{Y}_i \leq H_k + \Delta - (1 - \lambda)H_j) \\ &= P\left(\frac{H_k - \Delta - (1 - \lambda)H_j}{\lambda} \leq \tilde{Y}_i \leq \frac{H_k + \Delta - (1 - \lambda)H_j}{\lambda}\right) \\ &= F_{\tilde{Y}}\left(\frac{H_k + \Delta - (1 - \lambda)H_j}{\lambda}\right) - F_{\tilde{Y}}\left(\frac{H_k - \Delta - (1 - \lambda)H_j}{\lambda}\right) \end{aligned}$$

Υπενθυμίζουμε ότι η συνάρτηση κατανομής της δειγματικής διαμέσου δίνεται από τον τύπο

$$F_{\tilde{Y}}(y) = F_{\beta}\left(\Phi\left(\frac{y - \mu_0}{\sigma_0} - \delta\right) \mid \frac{n+1}{2}, \frac{n+1}{2}\right),$$

όπου $\Phi(\cdot)$ είναι η συνάρτηση κατανομής της τυπικής κανονικής κατανομής και $F_{\beta}(\cdot)$ η συνάρτηση κατανομής της κατανομής Beta με παραμέτρους μορφής $\frac{n+1}{2}$ και $\frac{n+1}{2}$.

Έστω

$$\mathbf{q} = (q_{-m}, \dots, q_0, \dots, q_m)$$

το διάνυσμα των αρχικών πιθανοτήτων των μεταβατικών καταστάσεων. Η ακριβής μορφή του διανύσματος $\mathbf{q}$ εξαρτάται από την τιμή Z_0 . Πράγματι

$$q_j = \begin{cases} 0, & Z_0 \notin (H_j - \Delta, H_j + \Delta] \\ 1, & Z_0 \in (H_j - \Delta, H_j + \Delta]. \end{cases}$$

Η συνάρτηση πιθανότητας του μήκος ροής RL δίνεται από τον τύπο

$$f_{RL}(\ell) = \mathbf{q} \mathbf{Q}^{\ell-1} \mathbf{c}, \quad \ell = 1, 2, \dots$$

και η συνάρτηση κατανομής από τον τύπο

$$F_{RL}(\ell) = 1 - \mathbf{q}\mathbf{Q}^\ell \mathbf{1}^T,$$

όπου $\mathbf{1} = (1, 1, \dots, 1)$ και $\mathbf{c} = (\mathbf{I} - \mathbf{Q})\mathbf{1}^T$. Επίσης

$$ARL = E(RL) = \mathbf{q}(\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}^T,$$

$$E(RL(RL - 1)) = 2 \mathbf{q}(\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1}^T,$$

$$SDRL = \sqrt{E(RL(RL - 1)) + E(RL) - E^2(RL)}.$$

4.2.2 Φάση I – Άγνωστες Παράμετροι

Στην πράξη, οι εντός ελέγχου τιμές των παραμέτρων της διεργασίας μ_0 και σ_0 συνήθως δεν είναι γνωστές και πρέπει να εκτιμηθούν από δεδομένα Φάσης I. Έστω ότι είναι διαθέσιμα m ανεξάρτητα δείγματα μεγέθους n το καθένα, δηλαδή τα

$$\mathbf{X}_i = (X_{i,1}, X_{i,2}, \dots, X_{i,n}), \quad i = 1, 2, \dots, m.$$

Υποθέτουμε ότι οι παρατηρήσεις $X_{i,j}$ είναι ανεξάρτητες και ακολουθούν την κανονική κατανομή $N(\mu_0, \sigma_0^2)$. Οι εκτιμητές των παραμέτρων μ_0 και σ_0 δίνονται συνήθως από τους τύπους

$$\hat{\mu}'_0 = \frac{1}{m} \sum_{i=1}^m \tilde{X}_i,$$

$$\hat{\sigma}'_0 = \frac{1}{d_2} \left(\frac{1}{m} \sum_{i=1}^m R_i \right),$$

όπου $\tilde{X}_i$ και $R_i = X_{i,(n)} - X_{i,(1)}$ είναι αντίστοιχα η δειγματική διάμεσος και το εύρος του i -οστού δείγματος, ενώ $d_2 = d_2(n) = E\left(\frac{R_i}{\sigma_0}\right)$ είναι σταθερά που προκύπτει από ειδικούς πίνακες.

Επομένως σε αυτή την περίπτωση τα όρια ελέγχου του διαγράμματος ελέγχου διαμέσου τύπου EWMA δίνονται από τους τύπους

$$LCL = \hat{\mu}'_0 - K \hat{\sigma}'_0,$$

$$UCL = \hat{\mu}'_0 + K \hat{\sigma}'_0.$$

Η τυποποιημένη μορφή των διαγραμμάτων ελέγχου διαμέσου τύπου EWMA με εκτιμημένες παραμέτρους δίνεται από τις σχέσεις

$$Z_i = \lambda \frac{\tilde{Y}_i - \mu_0}{\sigma_0} + (1 - \lambda)Z_{i-1}, \quad 0 < \lambda \leq 1,$$

με αρχική τιμή

$$Z_0 = 0.$$

Τα αντίστοιχα όρια ελέγχου εκφράζονται ως

$$\widehat{LCL} = \frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} - K \frac{\hat{\sigma}'_0}{\sigma_0},$$

$$\widehat{UCL} = \frac{\hat{\mu}'_0 - \mu_0}{\sigma_0} + K \frac{\hat{\sigma}'_0}{\sigma_0}.$$

Ορίζοντας τις τυχαίες μεταβλητές:

$$U = \frac{\hat{\mu}'_0 - \mu_0}{\sigma_0}, \quad V = \frac{\hat{\sigma}'_0}{\sigma_0},$$

τα όρια ελέγχου γράφονται απλούστερα ως

$$\widehat{LCL} = U - KV,$$

$$\widehat{UCL} = U + KV.$$

Έστω $\hat{f}_{RL}(\ell)$, $\hat{F}_{RL}(\ell)$, $\hat{v}_1$ και $\hat{v}_2$ η δεσμευμένη συνάρτηση πιθανότητας, η δεσμευμένη συνάρτηση κατανομής, η πρώτη και η δεύτερη δεσμευμένη καθοδική ροπή του μήκους ροής όταν είναι γνωστές οι τιμές των τυχαίων μεταβλητών U και V . Έστω επίσης $f_{U,V}(u, v \mid m, n)$ η από κοινού συνάρτηση πυκνότητας των U και V . Τότε σύμφωνα με τους Castagliola et al. (2016) ισχύει ότι

$$f_{RL}(\ell) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{U,V}(u, v \mid m, n) \hat{f}_{RL}(\ell) du dv,$$

$$F_{RL}(\ell) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{U,V}(u, v \mid m, n) \hat{F}_{RL}(\ell) du dv,$$

$$ARL = E(RL) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{U,V}(u, v \mid m, n) \hat{v}_1 du dv,$$

και

$$SDRL = \sqrt{E(RL^2) - ARL^2},$$

όπου

$$E(RL^2) = \int_{-\infty}^{+\infty} \int_0^{+\infty} f_{U,V}(u, v \mid m, n) (\hat{v}_1 + \hat{v}_2) du dv.$$

Με τους παραπάνω τύπους μπορεί να μελετηθεί το διάγραμμα ελέγχου διαμέσου τύπου EWMA στη Φάση Ι. Για λεπτομέρειες που αφορούν τους παραπάνω υπολογισμούς με προσεγγιστικές μεθόδους ο αναγνώστης παραπέμπεται στην εργασία των Castagliola et al. (2016).

4.3 Απόδοση του διαγράμματος ελέγχου διαμέσου τύπου EWMA

Στη συνέχεια εξετάζεται η απόδοση του διαγράμματος ελέγχου διαμέσου τύπου EWMA Φάσης ΙΙ για διάφορους συνδυασμούς των παραμέτρων του (n, ARL_0, λ, K) . Λόγω της συμμετρίας της κανονικής κατανομής, η συμπεριφορά του διαγράμματος ως προς την ανίχνευση ανοδικών και καθοδικών μετατοπίσεων ίδιου μεγέθους είναι παρόμοια, οπότε επικεντρωνόμαστε σε ανοδικές μετατοπίσεις της μέσης τιμής της διεργασίας.

Στους Πίνακες 4.1 και 4.2, για δεδομένο μέγεθος δείγματος n (3, 5, 7, 9) και τιμές της παραμέτρου εξομάλυνσης λ (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75), δίνεται η τιμή της παραμέτρου K έτσι ώστε να επιτυγχάνεται προκαθορισμένο εντός ελέγχου μήκος ροής $ARL_0 = 370.4$ και $ARL_0 = 500$, αντίστοιχα. Για τον υπολογισμό των τιμών του K χρησιμοποιήθηκε η μέθοδος των αλυσίδων Markov με κατάλληλη διακριτοποίηση του διαστήματος ελέγχου.

Πίνακας 4.1. Τιμές K για $ARL_0 = 370.4$

λ	$n = 3$	$n = 5$	$n = 7$	$n = 9$
0.05	0.2671	0.2136	0.1830	0.1625
0.1	0.4154	0.3321	0.2845	0.2527
0.2	0.6395	0.5113	0.4378	0.3889
0.3	0.8254	0.6598	0.5649	0.5017
0.4	0.9949	0.7953	0.6808	0.6046
0.5	1.1575	0.9251	0.7919	0.7031
0.75	1.5663	1.2518	1.0711	0.9508

Πίνακας 4.2. Τιμές K για $ARL_0 = 500$

λ	$n = 3$	$n = 5$	$n = 7$	$n = 9$
0.05	0.2806	0.2243	0.1921	0.1707
0.1	0.4328	0.3460	0.2964	0.2633
0.2	0.6626	0.5297	0.4536	0.4029
0.3	0.8533	0.6821	0.5840	0.5186
0.4	1.0273	0.8211	0.7029	0.6242
0.5	1.1942	0.9545	0.8170	0.7253
0.75	1.6148	1.2905	1.1042	0.9801

Στη συνέχεια, χρησιμοποιώντας τους παραπάνω συνδυασμούς τιμών (λ, K) υπολογίστηκε το (εκτός ελέγχου) μέσο μήκος ροής ARL των διαγραμμάτων ελέγχου EWMA για διάφορες τιμές της μετατόπισης δ (0, 0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2). Τα αποτελέσματα παρουσιάζονται στους Πίνακες 4.3 και 4.4 για $ARL_0 = 370.4$ και $ARL_0 = 500$, αντίστοιχα. Με έντονη γραφή σημειώνεται η μικρότερη τιμή ARL για κάθε μετατόπιση.

Πίνακας 4.3. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA
($ARL_0 = 370.4$)

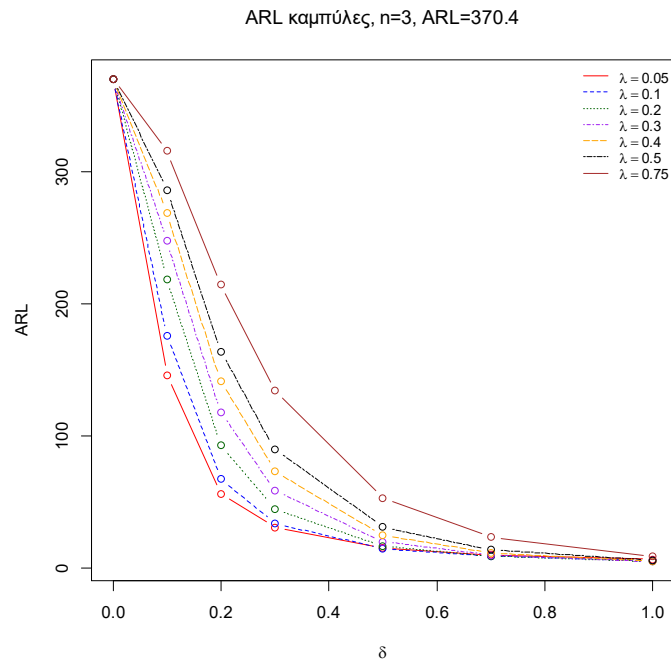
δ	$n = 3$						
	$\lambda = 0.05$ $K = 0.2671$	$\lambda = 0.1$ $K = 0.4154$	$\lambda = 0.2$ $K = 0.6395$	$\lambda = 0.3$ $K = 0.8254$	$\lambda = 0.4$ $K = 0.9949$	$\lambda = 0.5$ $K = 1.1575$	$\lambda = 0.75$ $K = 1.5663$
0.0	370.41	370.4	370.4	370.4	370.4	370.4	370.4
0.1	146.11	176.13	218.75	247.81	269.34	286.16	315.91
0.2	56.36	67.57	93.02	117.93	141.6	164.01	214.9
0.3	30.98	34.09	45.11	58.67	73.73	89.95	134.52
0.5	15.43	14.86	16.69	20.27	25.18	31.39	53.27
0.7	10.19	9.18	9.13	10.07	11.7	14.03	23.76
1.0	6.79	5.84	5.28	5.28	5.58	6.14	9.11
1.5	4.44	3.71	3.15	2.93	2.85	2.85	3.25
2.0	3.36	2.78	2.32	2.11	1.97	1.88	1.82
δ	$n = 5$						
	$\lambda = 0.05$ $K = 0.2136$	$\lambda = 0.1$ $K = 0.3321$	$\lambda = 0.2$ $K = 0.5113$	$\lambda = 0.3$ $K = 0.6598$	$\lambda = 0.4$ $K = 0.7953$	$\lambda = 0.5$ $K = 0.9251$	$\lambda = 0.75$ $K = 1.2518$
0.0	370.4	370.41	370.41	370.41	370.4	370.4	370.4
0.1	110.35	135.36	176.15	207.23	232	252.43	291.01
0.2	40.41	46.46	63.31	81.95	101.25	120.89	170.39
0.3	22.6	23.42	29.14	37.34	47.26	58.72	93.85
0.5	11.67	10.72	11.08	12.63	15.07	18.41	31.52
0.7	7.87	6.87	6.4	6.62	7.24	8.25	13.05
1.0	5.35	4.51	3.92	3.74	3.76	3.91	5.03
1.5	3.57	2.95	2.47	2.26	2.12	2.04	2.04
2.0	2.75	2.25	1.92	1.68	1.52	1.41	1.3
δ	$n = 7$						
	$\lambda = 0.05$ $K = 0.183$	$\lambda = 0.1$ $K = 0.2845$	$\lambda = 0.2$ $K = 0.4378$	$\lambda = 0.3$ $K = 0.5649$	$\lambda = 0.4$ $K = 0.6808$	$\lambda = 0.5$ $K = 0.7919$	$\lambda = 0.75$ $K = 1.0711$
0.0	370.4	370.39	370.41	370.4	370.41	370.4	370.4
0.1	89.2	109.63	146.53	176.98	202.66	224.76	268.87
0.2	32.2	35.66	47.41	61.61	77.18	93.82	139
0.3	18.33	18.24	21.49	26.85	33.75	42.11	69.73
0.5	9.7	8.68	8.51	9.26	10.63	12.62	21.11
0.7	6.63	5.69	5.12	5.09	5.34	5.84	8.52
1.0	4.56	3.81	3.25	3.03	2.96	2.98	3.44
1.5	3.1	2.54	2.14	1.92	1.77	1.67	1.57
2.0	2.35	2.03	1.67	1.4	1.26	1.18	1.11
δ	$n = 9$						
	$\lambda = 0.05$ $K = 0.1625$	$\lambda = 0.1$ $K = 0.2527$	$\lambda = 0.2$ $K = 0.3889$	$\lambda = 0.3$ $K = 0.5017$	$\lambda = 0.4$ $K = 0.6046$	$\lambda = 0.5$ $K = 0.7031$	$\lambda = 0.75$ $K = 0.9508$
0.0	370.41	370.4	370.4	370.39	370.4	370.41	370.39
0.1	75.27	92.06	124.92	153.79	179.22	201.9	249.35
0.2	27.18	29.17	37.71	48.79	61.51	75.6	116.1
0.3	15.71	15.17	17.1	20.8	25.82	32.14	54.22
0.5	8.46	7.44	7.04	7.4	8.23	9.51	15.33
0.7	5.84	4.96	4.36	4.23	4.31	4.58	6.19
1.0	4.05	3.37	2.84	2.62	2.51	2.47	2.64
1.5	2.78	2.28	1.94	1.7	1.54	1.44	1.32
2.0	2.12	1.91	1.45	1.21	1.12	1.07	1.04

Πίνακας 4.4. Τιμές ARL για το άνω διάγραμμα ελέγχου διαμέσου τύπου EWMA
($ARL_0 = 500$)

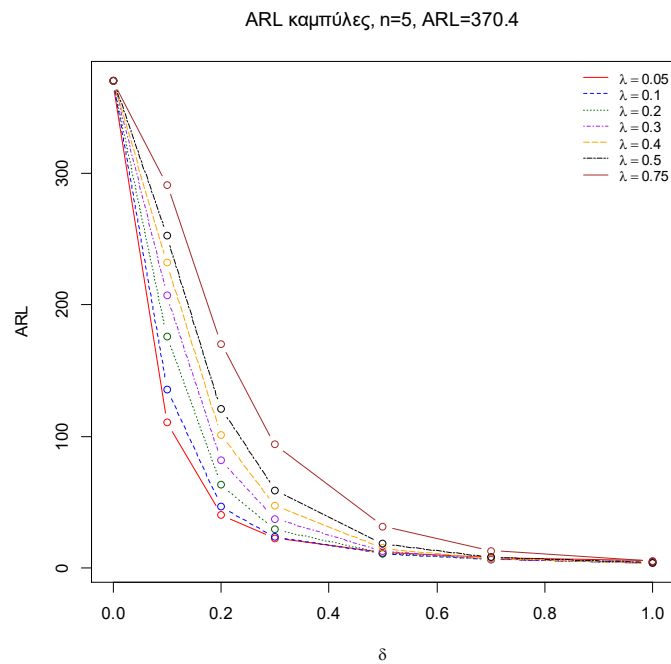
δ	$n = 3$						
	$\lambda = 0.05$ $K = 0.2806$	$\lambda = 0.1$ $K = 0.4328$	$\lambda = 0.2$ $K = 0.6626$	$\lambda = 0.3$ $K = 0.8533$	$\lambda = 0.4$ $K = 1.0273$	$\lambda = 0.5$ $K = 1.1942$	$\lambda = 0.75$ $K = 1.6148$
0.0	500	500	499.99	500	500	500	500
0.1	176.73	220.65	282.53	324.59	355.7	379.97	422.68
0.2	63.65	79.04	113.73	147.86	180.47	211.5	282.27
0.3	33.91	38.22	52.83	70.86	91	112.85	173.35
0.5	16.49	16	18.45	23.05	29.39	37.47	66.25
0.7	10.8	9.73	9.8	11.02	13.09	16.05	28.57
1.0	7.15	6.13	5.55	5.6	6	6.72	10.46
1.5	4.65	3.86	3.27	3.05	2.98	3.01	3.52
2.0	3.51	2.88	2.4	2.18	2.04	1.95	1.92
δ	$n = 5$						
	$\lambda = 0.05$ $K = 0.2243$	$\lambda = 0.1$ $K = 0.346$	$\lambda = 0.2$ $K = 0.5297$	$\lambda = 0.3$ $K = 0.6821$	$\lambda = 0.4$ $K = 0.8211$	$\lambda = 0.5$ $K = 0.9545$	$\lambda = 0.75$ $K = 1.2905$
0.0	500	500	499.99	500	500	500	500
0.1	130.45	166.14	224.1	268.33	303.64	332.79	387.74
0.2	44.79	53.06	75.65	100.76	126.92	153.7	221.67
0.3	24.43	25.72	33.26	43.97	57.04	72.25	119.31
0.5	12.39	11.41	11.99	13.99	17.1	21.37	38.38
0.7	8.31	7.23	6.78	7.09	7.89	9.17	15.26
1.0	5.62	4.71	4.09	3.93	3.97	4.18	5.59
1.5	3.73	3.07	2.55	2.33	2.2	2.13	2.16
2.0	2.87	2.32	1.97	1.74	1.57	1.46	1.34
δ	$n = 7$						
	$\lambda = 0.05$ $K = 0.1921$	$\lambda = 0.1$ $K = 0.2964$	$\lambda = 0.2$ $K = 0.4536$	$\lambda = 0.3$ $K = 0.584$	$\lambda = 0.4$ $K = 0.7029$	$\lambda = 0.5$ $K = 0.817$	$\lambda = 0.75$ $K = 1.1042$
0.0	500	500	500	499.99	499.99	500	500
0.1	103.8	132.53	184.19	226.98	263.18	294.42	356.83
0.2	35.31	40.09	55.68	74.58	95.46	117.9	179.35
0.3	19.68	19.79	24.09	31.03	40.02	50.99	87.63
0.5	10.26	9.18	9.11	10.09	11.82	14.35	25.24
0.7	6.98	5.96	5.38	5.39	5.73	6.37	9.74
1.0	4.78	3.98	3.38	3.16	3.1	3.14	3.74
1.5	3.23	2.64	2.2	1.98	1.83	1.73	1.63
2.0	2.47	2.07	1.73	1.46	1.3	1.21	1.13
δ	$n = 9$						
	$\lambda = 0.05$ $K = 0.1707$	$\lambda = 0.1$ $K = 0.2633$	$\lambda = 0.2$ $K = 0.4029$	$\lambda = 0.3$ $K = 0.5186$	$\lambda = 0.4$ $K = 0.6242$	$\lambda = 0.5$ $K = 0.7253$	$\lambda = 0.75$ $K = 0.9801$
0.0	499.99	500	500	500	499.99	499.99	500
0.1	86.57	109.93	155.45	195.63	231.17	262.97	329.7
0.2	29.58	32.42	43.68	58.31	75.21	94.06	148.72
0.3	16.79	16.35	18.93	23.69	30.18	38.39	67.45
0.5	8.93	7.84	7.48	7.97	9.02	10.64	18.06
0.7	6.14	5.18	4.56	4.45	4.58	4.93	6.96
1.0	4.25	3.51	2.94	2.72	2.61	2.59	2.83
1.5	2.91	2.36	1.99	1.76	1.6	1.49	1.37
2.0	2.19	1.96	1.52	1.25	1.14	1.09	1.04

Ειδικότερα, στα Σχήματα 4.1-4.4 απεικονίζονται γραφικά οι τιμές ARL του Πίνακα 4.3 ($ARL_0 = 370.4$) για τις διαφορετικές τιμές του μεγέθους του δείγματος και για $\delta \leq 1$

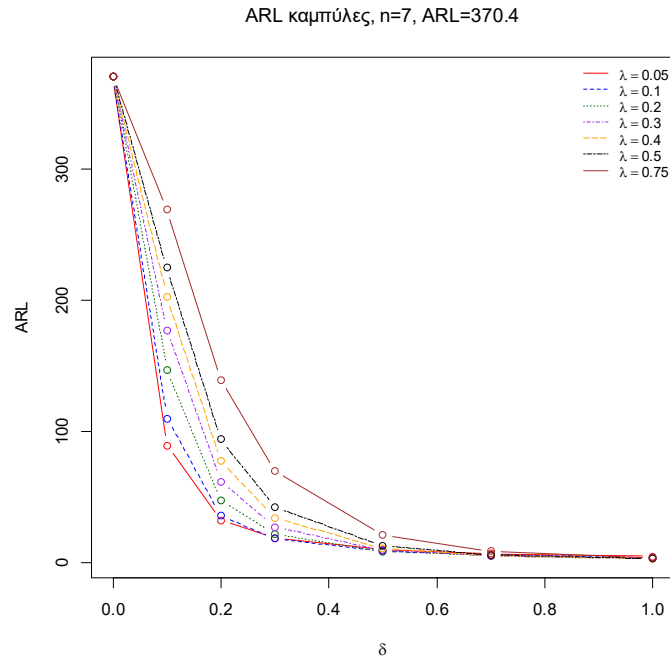
όπου παρατηρείται η πιο ουσιαστική διαφοροποίηση της συμπεριφοράς του διαγράμματος.



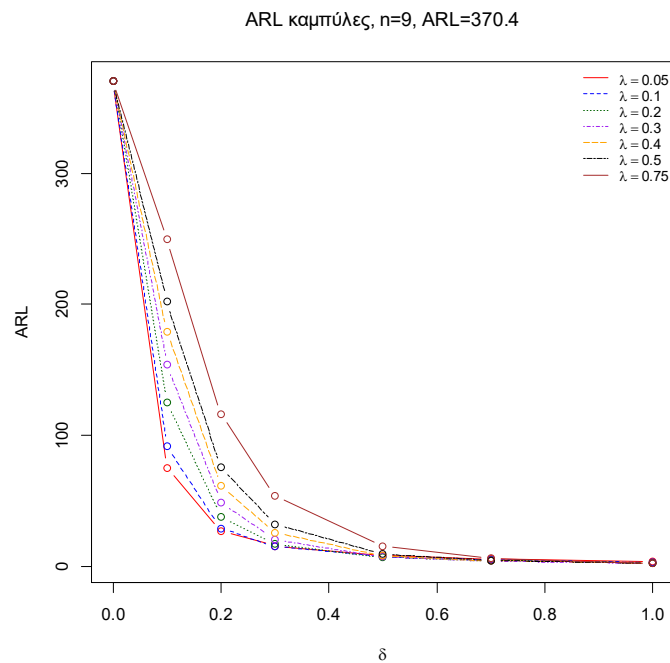
Σχήμα 4.1. Καμπύλες ARL για το διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 3, ARL_0 = 370.4$)



Σχήμα 4.2. Καμπύλες ARL για το διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 5, ARL_0 = 370.4$)



Σχήμα 4.3. Καμπύλες ARL για το διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 7, ARL_0 = 370.4$)



Σχήμα 4.4. Καμπύλες ARL για το διάγραμμα ελέγχου διαμέσου τύπου EWMA ($n = 9, ARL_0 = 370.4$)

Παρατηρούμε ότι η συμπεριφορά των καμπυλών είναι ποιοτικά παρόμοια με εκείνη του διαγράμματος CUSUM. Όλες οι καμπύλες ξεκινούν από την τιμή $ARL_0 = 370.4$ για

$\delta = 0$ και μειώνονται καθώς αυξάνεται η μετατόπιση, άρα υπάρχει πιο γρήγορη ανίχνευση της εκτός ελέγχου κατάστασης. Η επίδραση της παραμέτρου λ είναι πιο έντονη για μικρές μετατοπίσεις, όπου μικρότερες τιμές της οδηγούν σε μικρότερα ARL , ενώ για μεγαλύτερες τιμές της μετατόπισης οι διαφορές μεταξύ των καμπυλών περιορίζονται σημαντικά. Επιπλέον, όπως και στην περίπτωση του CUSUM, η αύξηση του μεγέθους δείγματος n βελτιώνει την απόδοση του διαγράμματος, οδηγώντας σε ταχύτερη μείωση του ARL .

Η συμπεριφορά αυτή ερμηνεύεται από τη φύση του EWMA, όπου η παράμετρος λ καθορίζει τον βαθμό στον οποίο λαμβάνεται υπόψη η προηγούμενη πληροφορία. Μικρότερες τιμές της λ οδηγούν σε εντονότερη εξομάλυνση και μεγαλύτερη συσσώρευση πληροφορίας, ενισχύοντας την ικανότητα ανίχνευσης μικρών μετατοπίσεων, ενώ μεγαλύτερες τιμές καθιστούν το διάγραμμα πιο ευαίσθητο στις πρόσφατες παρατηρήσεις.

4.4 Εφαρμογή διαγράμματος ελέγχου διαμέσου τύπου EWMA

Θα χρησιμοποιηθεί ξανά το σύνολο δεδομένων του Montgomery (2013), που αφορά την εσωτερική διάμετρο δακτυλίων εμβόλων (piston rings), προκειμένου να επιδειχθεί η εφαρμογή του διαγράμματος ελέγχου διαμέσου τύπου EWMA Φάσης II. Για λόγους πληρότητας αλλά και σύγκρισης, τα δεδομένα θα μελετηθούν και με το κλασικό EWMA $\bar{X}$ διάγραμμα ελέγχου.

Υπενθυμίζεται ότι τα δεδομένα της εσωτερικής διαμέτρου των δακτυλίων εμβόλων σε mm αποτελούνται από 40 υποομάδες (δείγματα) μεγέθους $n = 5$. Θα υποθέσουμε ότι η εντός ελέγχου κατανομή των δεδομένων είναι κανονική με $\mu_0 = 74$ και $\sigma_0 = 0.01$ (οι τιμές αυτές είναι σχεδόν ίδιες με τις εκτιμήσεις των παραμέτρων αυτών). Στο κλασικό διάγραμμα ελέγχου τύπου EWMA απεικονίζεται στο διάγραμμα η στατιστική συνάρτηση

$$Z_i^{\bar{X}} = \lambda \bar{X}_i + (1 - \lambda) Z_{i-1}^{\bar{X}},$$

όπου $Z_0^{\bar{X}} = \mu_0$, και τα όρια ελέγχου (χρονομεταβλητά) δίνονται από τις σχέσεις

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

$$CL = \mu_0,$$

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

Για το σχεδιασμό του διαγράμματος επιλέγουμε $\lambda = 0.1$. Για να επιτευχθεί η τιμή $ARL_0 = 370.4$ προκύπτει ότι $L = 2.701461$. Σημειώνεται ότι τα μη χρονομεταβλητά όρια του διαγράμματος είναι

$$UCL = \mu_0 + L \frac{\sigma}{\sqrt{n}} \sqrt{\frac{\lambda}{2-\lambda}} = 74.00277,$$

$$LCL = \mu_0 - L \frac{\sigma}{\sqrt{n}} \sqrt{\frac{\lambda}{2-\lambda}} = 73.99723.$$

Στο διάγραμμα ελέγχου διαμέσου τύπου EWMA απεικονίζεται η στατιστική συνάρτηση

$$Z_i^{\bar{X}} = \lambda \bar{X}_i + (1-\lambda)Z_{i-1}^{\bar{X}},$$

όπου $Z_t^{\bar{X}} = \mu_0$. Για το σχεδιασμό του διαγράμματος επιλέγουμε $\lambda = 0.1$. Για να επιτευχθεί $ARL_0 = 370.4$ προκύπτει ότι $K = 0.3321246$. Τα όρια ελέγχου του διαγράμματος είναι

$$UCL = \mu_0 + K\sigma_0 = 74.003321,$$

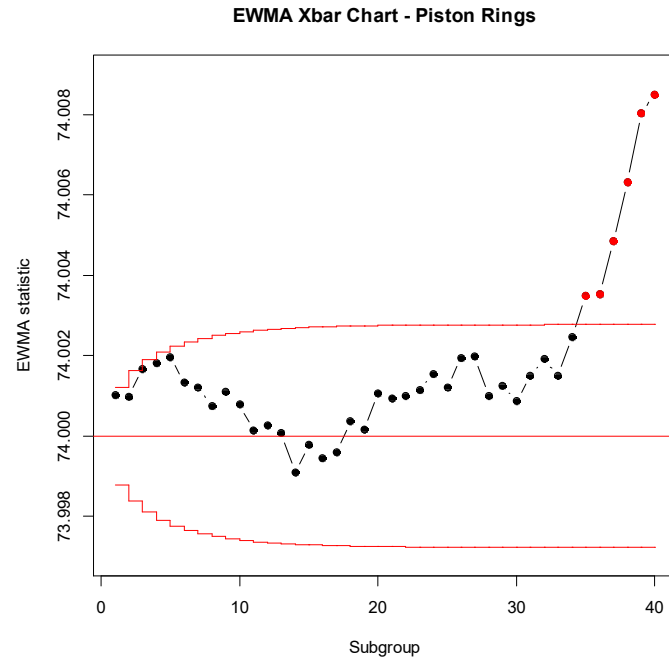
$$LCL = \mu_0 - K\sigma_0 = 73.996679.$$

Οι τιμές των στατιστικών συναρτήσεων που απεικονίζονται στα δύο EWMA διαγράμματα δίνονται στον Πίνακα 4.5.

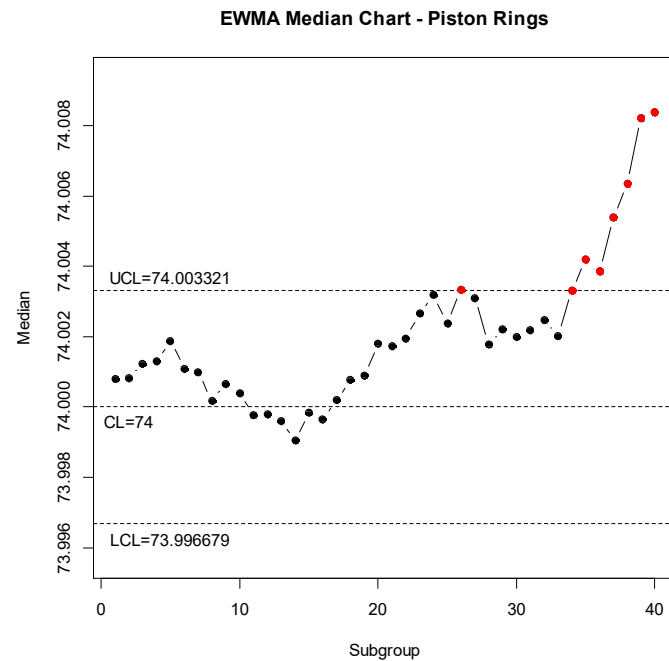
Πίνακας 4.5. Τιμές $Z_i^{\bar{X}}$ και $Z_i^{\tilde{X}}$ διαγραμμάτων EWMA

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	$\bar{X}_i$	$\tilde{X}_i$	$Z_i^{\bar{X}}$	$Z_i^{\tilde{X}}$
1	74.030	74.002	74.019	73.992	74.008	74.0102	74.008	74.00102	74.00080
2	73.995	73.992	74.001	74.011	74.004	74.0006	74.001	74.00098	74.00082
3	73.988	74.024	74.021	74.005	74.002	74.008	74.005	74.00168	74.00124
4	74.002	73.996	73.993	74.015	74.009	74.003	74.002	74.00181	74.00131
5	73.992	74.007	74.015	73.989	74.014	74.0034	74.007	74.00197	74.00188
6	74.009	73.994	73.997	73.985	73.993	73.9956	73.994	74.00133	74.00109
7	73.995	74.006	73.994	74.000	74.005	74.000	74	74.00120	74.00099
8	73.985	74.003	73.993	74.015	73.988	73.9968	73.993	74.00076	74.00019
9	74.008	73.995	74.009	74.005	74.004	74.0042	74.005	74.00110	74.00067
10	73.988	74.000	73.990	74.007	73.995	73.996	73.995	74.00079	74.00040
11	73.994	73.998	73.994	73.995	73.990	73.9942	73.994	74.00013	73.99976
12	74.004	74.000	74.007	74.000	73.996	74.0014	74	74.00026	73.99978
13	73.983	74.002	73.998	73.997	74.012	73.9984	73.998	74.00007	73.99961
14	74.006	73.967	73.994	74.000	73.984	73.9902	73.994	73.99909	73.99905
15	74.012	74.014	73.998	73.999	74.007	74.006	74.007	73.99978	73.99984
16	74.000	73.984	74.005	73.998	73.996	73.9966	73.998	73.99946	73.99966
17	73.994	74.012	73.986	74.005	74.007	74.0008	74.005	73.99959	74.00019
18	74.006	74.010	74.018	74.003	74.000	74.0074	74.006	74.00038	74.00077
19	73.984	74.002	74.003	74.005	73.997	73.9982	74.002	74.00016	74.00090
20	74.000	74.010	74.013	74.020	74.003	74.0092	74.01	74.00106	74.00181
21	73.982	74.001	74.015	74.005	73.996	73.9998	74.001	74.00094	74.00172
22	74.004	73.999	73.990	74.006	74.009	74.0016	74.004	74.00100	74.00195
23	74.010	73.989	73.990	74.009	74.014	74.0024	74.009	74.00114	74.00266
24	74.015	74.008	73.993	74.000	74.010	74.0052	74.008	74.00155	74.00319
25	73.982	73.984	73.995	74.017	74.013	73.9982	73.995	74.00121	74.00237
26	74.012	74.015	74.030	73.986	74.000	74.0086	74.012	74.00195	74.00334
27	73.995	74.010	73.990	74.015	74.001	74.0022	74.001	74.00198	74.00310
28	73.987	73.999	73.985	74.000	73.990	73.9922	73.99	74.00100	74.00179
29	74.008	74.010	74.003	73.991	74.006	74.0036	74.006	74.00126	74.00221
30	74.003	74.000	74.001	73.986	73.997	73.9974	74	74.00087	74.00199
31	73.994	74.003	74.015	74.02	74.004	74.0072	74.004	74.00151	74.00219
32	74.008	74.002	74.018	73.995	74.005	74.0056	74.005	74.00192	74.00247
33	74.001	74.004	73.990	73.996	73.998	73.9978	73.998	74.00150	74.00203
34	74.015	74.000	74.016	74.025	74.000	74.0112	74.015	74.00247	74.00332
35	74.030	74.005	74.000	74.016	74.012	74.0126	74.012	74.00349	74.00419
36	74.001	73.99	73.995	74.01	74.024	74.004	74.001	74.00354	74.00387
37	74.015	74.02	74.024	74.005	74.019	74.0166	74.019	74.00484	74.00538
38	74.035	74.01	74.012	74.015	74.026	74.0196	74.015	74.00632	74.00635
39	74.017	74.013	74.036	74.025	74.026	74.0234	74.025	74.00803	74.00821
40	74.010	74.005	74.029	74.000	74.020	74.0128	74.01	74.00850	74.00839

Τα αντίστοιχα διαγράμματα δίνονται στα Σχήματα 4.5 και 4.6.



Σχήμα 4.5. EWMA $\bar{X}$ διάγραμμα ελέγχου για την εσωτερική διάμετρο



Σχήμα 4.6. EWMA διάγραμμα ελέγχου διαμέσου για την εσωτερική διάμετρο

Παρατηρούμε ότι και τα δύο διαγράμματα ελέγχου τύπου EWMA δίνουν σήμα εκτός ελέγχου διεργασίας, καθώς η αντίστοιχη στατιστική υπερβαίνει το αντίστοιχο άνω όριο ελέγχου. Στο διάγραμμα EWMA $\bar{X}$ σήματα εκτός ελέγχου διεργασίας έχουμε στα σημεία 35-40, ενώ στο διάγραμμα EWMA διαμέσου το πρώτο (οριακό) σήμα εμφανίζεται στο 26^ο σημείο και μετέπειτα στα σημεία 34-40. Επομένως, στη συγκεκριμένη περίπτωση και οι δύο προσεγγίσεις είναι σε θέση να ανιχνεύσουν τη μεταβολή της διεργασίας, με το διάγραμμα EWMA διαμέσου να είναι πιο γρήγορο.

Παρατηρούμε επίσης διαφορές στη μορφή της εξέλιξης των δύο στατιστικών συναρτήσεων. Το διάγραμμα που βασίζεται στη δειγματική διάμεσο εμφανίζει σχετικά πιο ομαλή συμπεριφορά, γεγονός που συνδέεται με τη μικρότερη ευαισθησία της διαμέσου σε ακραίες τιμές. Αντίθετα, το διάγραμμα που βασίζεται στον δειγματικό μέσο παρουσιάζει μεγαλύτερες διακυμάνσεις, γεγονός που οφείλεται στην αυξημένη επίδραση των επιμέρους παρατηρήσεων στις τιμές του.

Συνολικά, τα αποτελέσματα δείχνουν ότι και τα δύο διαγράμματα EWMA είναι κατάλληλα για την ανίχνευση μεταβολών της διεργασίας, με τη χρήση της διαμέσου να προσφέρει ένα πιο ανθεκτικό εργαλείο σε περιπτώσεις παρουσίας ακραίων τιμών ή μεμονωμένων αποκλίσεων.

Παράρτημα

K.1 Κώδικας για την εξαγωγή του Πίνακα 2.1

```
##### PHASE II - Evaluation of K parameter for fixed ARL0

rm(list=ls(all=TRUE))

# Set parameters
n <- 9          # Sample size per subgroup
mu <- 0         # True process mean (in control)
sigma <- 1      # True process std deviation (in control)
set.seed(100)   # For reproducibility

# ARL as a function of K
ARL_COMPUTE <- function(K) {
  p1 <- pbeta(pnorm(- K), shape1=(n+1)/2, shape2=(n+1)/2)
  p2 <- pbeta(pnorm(- K), shape1=(n+1)/2, shape2=(n+1)/2)
  ARL <- 1/(p1+p2)
  return(ARL)
}

# find K as a function of ARL0, ...

find_K <- function(ARL0,K_low,K_high,tolerance,max_iter) {
  iter <- 0      # Number of iterations
  while (iter < max_iter) {
    iter <- iter + 1
    K_mid <- (K_low + K_high) / 2
    estimated_arl <- ARL_COMPUTE(K_mid)
    cat("Iter",iter," :      K=",    K_mid,    ",      Estimated    ARL=",
estimated_arl,"\n")
    if (abs(estimated_arl - ARL0) <= tolerance) {
      return(K_mid)
    } else if (estimated_arl < ARL0) {
      K_low <- K_mid
    } else {
      K_high <- K_mid
    }
  }

  warning("Maximum iterations reached; result may not be accurate.")
  return(K_mid)
}

# Results
tolerance <- 0.001 # Tolerance
max_iter <- 100    # Maximum number of iterations
ARL0 <- 370.4      # ARL0
K_low <- 1         # Lower bound for K
K_high <- 1.4      # Upper bound for K

find_K(ARL0,K_low,K_high,tolerance,max_iter)
```

```
Iter 1 :    K= 1.2 , Estimated ARL= 293.8083
Iter 2 :    K= 1.3 , Estimated ARL= 652.741
Iter 3 :    K= 1.25 , Estimated ARL= 435.1649
Iter 4 :    K= 1.225 , Estimated ARL= 357.0022
Iter 5 :    K= 1.2375 , Estimated ARL= 393.9948
Iter 6 :    K= 1.23125 , Estimated ARL= 375.0056
Iter 7 :    K= 1.228125 , Estimated ARL= 365.8841
Iter 8 :    K= 1.229688 , Estimated ARL= 370.4145
Iter 9 :    K= 1.228906 , Estimated ARL= 368.1417
Iter 10 :    K= 1.229297 , Estimated ARL= 369.2762
Iter 11 :    K= 1.229492 , Estimated ARL= 369.8449
Iter 12 :    K= 1.22959 , Estimated ARL= 370.1295
Iter 13 :    K= 1.229639 , Estimated ARL= 370.272
Iter 14 :    K= 1.229663 , Estimated ARL= 370.3432
Iter 15 :    K= 1.229675 , Estimated ARL= 370.3788
Iter 16 :    K= 1.229681 , Estimated ARL= 370.3967
Iter 17 :    K= 1.229684 , Estimated ARL= 370.4056
Iter 18 :    K= 1.229683 , Estimated ARL= 370.4011
Iter 19 :    K= 1.229682 , Estimated ARL= 370.3989
Iter 20 :    K= 1.229683 , Estimated ARL= 370.4
[1] 1.229683
```

Κ.2 Κώδικας για την εξαγωγή του Πίνακα 2.2

```
# PHASE I
rm(list=ls(all=TRUE))

# Set parameters
n <- 9          # Sample size per subgroup
m <- 10         # Number of Phase I samples
n_sim <- 10^5    # Number of simulations
mu <- 0         # True process mean (in control)
sigma <- 1      # True process std deviation (in control)
K <- 1.229683   # Parameter K
set.seed(100)   # For reproducibility

# Constant d2
install.packages("SixSigma")
library(SixSigma)
options(digits=10)
nmax <- 25
d2_values <- sapply(2:nmax, ss.cc.getd2)
d2 <- d2_values[[n-1]]

# Computation of ARL as a function of delta
ARL_COMPUTE <- function(delta) {
  Med <- rep(0, times=m)
  Range <- rep(0, times=m)
  ARL <- rep(0, times=n_sim)

  for (j in 1:n_sim) {
    for (i in 1:m) {
      vec <- rnorm(n, mu, sigma)
      Med[i] <- sort(vec)[(n+1)/2]
      Range[i] <- diff(range(vec))
    }
    est.mu <- mean(Med)
    est.sigma <- (1/d2)*mean(Range)
    p1 <- pbeta(pnorm((est.mu-mu)/sigma) - K*(est.sigma/sigma) -
      delta, 0, 1), shapel=(n+1)/2, shape2=(n+1)/2)
    p2 <- pbeta(pnorm(-(est.mu-mu)/sigma) - K*(est.sigma/sigma) +
      delta, 0, 1), shapel=(n+1)/2, shape2=(n+1)/2)
    ARL[j] <- 1/(p1+p2)
  }
  return(mean(ARL))
}

# Results

delta <- c(0, 0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.6, 2)
ARL <- sapply(delta, ARL_COMPUTE)

mat <- cbind(delta, ARL)
rownames(mat) <- rep("", times=length(delta))
options(digits=4)
mat
```

delta	ARL
0.0	411.573
0.1	362.902
0.2	261.176
0.3	158.126
0.5	46.863
0.7	15.264
1.0	4.328
1.6	1.279
2.0	1.044

Κ.3 Κώδικας για την εξαγωγή του Πίνακα 2.4

```
# PHASE I
rm(list=ls(all=TRUE))

# Set parameters
n <- 5          # Sample size per subgroup
m <- 5          # Number of Phase I samples
n_sim <- 10^5    # Number of simulations
mu <- 0         # True process mean (in control)
sigma <- 1      # True process std deviation (in control)
delta <- 0
set.seed(100)   # For reproducibility

# Constant d2
library(SixSigma)
options(digits=10)
nmax <- 25
d2_values <- sapply(2:nmax, ss.cc.getd2)
d2 <- d2_values[[n-1]]

# Computation of ARL as a function of K
ARL_COMPUTE <- function(K) {
  Med <- rep(0, times=m)
  Range <- rep(0, times=m)
  ARL <- rep(0, times=n_sim)

  for (j in 1:n_sim) {
    for (i in 1:m) {
      vec <- rnorm(n, mu, sigma)
      Med[i] <- sort(vec)[(n+1)/2]
      Range[i] <- diff(range(vec))
    }
    est.mu <- mean(Med)
    est.sigma <- (1/d2)*mean(Range)
    p1 <- pbeta(pnorm( ((est.mu-mu)/sigma) - K*(est.sigma/sigma) -
      delta, 0, 1 ), shape1=(n+1)/2, shape2=(n+1)/2)
    p2 <- pbeta(pnorm( -((est.mu-mu)/sigma) - K*(est.sigma/sigma) + delta, 0, 1
    ), shape1=(n+1)/2, shape2=(n+1)/2)
    ARL[j] <- 1/(p1+p2)
  }
  return(mean(ARL))
}

# find K as a function of ARL0, ...

find_K <- function(ARL0, K_low, K_high, tolerance, max_iter) {
  iter <- 0          # Number of iterations
  while (iter < max_iter) {
    iter <- iter + 1
    K_mid <- (K_low + K_high) / 2
    estimated_arl <- ARL_COMPUTE(K_mid)
    cat("Iter", iter, ": K=", K_mid, ", Estimated ARL=", estimated_arl, "\n")
    if (abs(estimated_arl - ARL0) <= tolerance) {
      return(K_mid)
    } else if (estimated_arl < ARL0) {
      K_low <- K_mid
    } else {
      K_high <- K_mid
    }
  }
}
```

```

    }

    warning("Maximum iterations reached; result may not be accurate.")
    return(K_mid)
}

# Results
tolerance <- 0.001 # Tolerance
max_iter <- 100    # Maximum number of iterations
ARL0 <- 370.4      # ARL0
K_low <- 1.465     # Lower bound for K
K_high <- 1.49     # Upper bound for K

find_K(ARL0,K_low,K_high,tolerance,max_iter)

Iter 1 : K= 1.4775 , Estimated ARL= 389.8385182
Iter 2 : K= 1.47125 , Estimated ARL= 367.4613765
Iter 3 : K= 1.474375 , Estimated ARL= 377.2516791
Iter 4 : K= 1.4728125 , Estimated ARL= 358.5648144
Iter 5 : K= 1.47359375 , Estimated ARL= 379.4144437
Iter 6 : K= 1.473203125 , Estimated ARL= 369.177214
Iter 7 : K= 1.473398437 , Estimated ARL= 372.6809444
Iter 8 : K= 1.473300781 , Estimated ARL= 373.7784202
Iter 9 : K= 1.473251953 , Estimated ARL= 378.4754267
Iter 10 : K= 1.473227539 , Estimated ARL= 379.1549617
Iter 11 : K= 1.473215332 , Estimated ARL= 386.1032722
Iter 12 : K= 1.473209229 , Estimated ARL= 369.7904274
Iter 13 : K= 1.47321228 , Estimated ARL= 372.5382779
Iter 14 : K= 1.473210754 , Estimated ARL= 354.8276784
Iter 15 : K= 1.473211517 , Estimated ARL= 380.9053694
Iter 16 : K= 1.473211136 , Estimated ARL= 372.1743888
Iter 17 : K= 1.473210945 , Estimated ARL= 372.1370971
Iter 18 : K= 1.47321085 , Estimated ARL= 373.306984
Iter 19 : K= 1.473210802 , Estimated ARL= 365.3535283
Iter 20 : K= 1.473210826 , Estimated ARL= 396.3735638
Iter 21 : K= 1.473210814 , Estimated ARL= 360.3520789
Iter 22 : K= 1.47321082 , Estimated ARL= 385.8337577
Iter 23 : K= 1.473210817 , Estimated ARL= 376.9043303
Iter 25 : K= 1.473210816 , Estimated ARL= 370.7197667
[1] 1.473210816

```

K.4 Κώδικας για την εξαγωγή του Πίνακα 3.1

```

rm(list=ls(all=TRUE))
n <- 3          # Sample size
m0 <- 0         # IC mean
s0 <- 1         # IC standard deviation
delta <- 0      # Shift in mean

# ARL as a function of K and H
r <- 400        # Number of subintervals : I0, I1, I2, ..., Ir-1
Q <- matrix(0, nrow=r, ncol=r)

arl <- function (K, H){
  D <- H/(2*r)
  # First column of Q
  for(j in 1:r){
    y <- -(2*(j-1)+1)*D + 2*D + K + m0
    Q[j,1] <- pbeta(  pnorm( ((y-m0)/s0)-delta , m0, s0 ) ,  shape1=(n+1)/2,
  shape2=(n+1)/2 )
  }
  # Other columns of Q: Qjk
  for(j in 1:r){
    for(k in 2:r) {
      yU <- (2*(k-1)+1)*D - (2*(j-1)+1)*D + D + K + m0
      yL <- (2*(k-1)+1)*D - (2*(j-1)+1)*D - D + K + m0
      A <- pbeta(  pnorm( ((yU-m0)/s0)-delta, m0, s0 ) ,  shape1=(n+1)/2,
  shape2=(n+1)/2 )
      B <- pbeta(  pnorm( ((yL-m0)/s0)-delta, m0, s0 ) ,  shape1=(n+1)/2,
  shape2=(n+1)/2 )
      Q[j,k] <- A - B
    }
  }
  e0 <- rep(0,r); e0[1] <- 1          # Initial probability vector
  l1 <- rep(1,r)                      # Vector of 1's
  Id <- diag(r)                      # Identity r x r matrix
  ARL <- as.vector(e0%%solve(Id-Q)%%l1) # IC ARL
  return(ARL)
}

# Search for H
Find_H <- function(H_low, H_high, tolerance, max_iter) {
  iter <- 0          # Number of iterations
  while (iter < max_iter) {
    iter <- iter + 1
    H_mid <- (H_low + H_high) / 2
    estimated_arl <- arl(K, H_mid)
    cat("Iter",iter," :  H=", H_mid, ", Estimated ARL=", estimated_arl,"\n")
    if (abs(estimated_arl - ARL0) <= tolerance) {
      return(H_mid)
    } else if (estimated_arl < ARL0) {
      H_low <- H_mid
    } else {
      H_high <- H_mid
    }
  }
}

warning("Maximum iterations reached; result may not be accurate.")
return(H_mid)
}

```

```

# Results

tolerance <- 0.01 # Tolerance
max_iter <- 50    # Maximum number of iterations
H_low <- 0
H_high <- 3
ARL0 <- 370.4    # IC ARL
K <- 0.5         # Reference value

(result_H <- Find_H(H_low, H_high, tolerance, max_iter))
result_H

#####
Iter 1 :   H= 1.5 , Estimated ARL= 133.1505
Iter 2 :   H= 2.25 , Estimated ARL= 725.1662
Iter 3 :   H= 1.875 , Estimated ARL= 312.9287
Iter 4 :   H= 2.0625 , Estimated ARL= 476.9378
Iter 5 :   H= 1.96875 , Estimated ARL= 386.4677
Iter 6 :   H= 1.921875 , Estimated ARL= 347.7949
Iter 7 :   H= 1.945312 , Estimated ARL= 366.6306
Iter 8 :   H= 1.957031 , Estimated ARL= 376.4207
Iter 9 :   H= 1.951172 , Estimated ARL= 371.4939
Iter 10 :   H= 1.948242 , Estimated ARL= 369.0544
Iter 11 :   H= 1.949707 , Estimated ARL= 370.2722
Iter 12 :   H= 1.950439 , Estimated ARL= 370.8825
Iter 13 :   H= 1.950073 , Estimated ARL= 370.5772
Iter 14 :   H= 1.94989 , Estimated ARL= 370.4247
Iter 15 :   H= 1.949799 , Estimated ARL= 370.3484
Iter 16 :   H= 1.949844 , Estimated ARL= 370.3865
Iter 17 :   H= 1.949867 , Estimated ARL= 370.4056
[1] 1.949867

```

K.5 Κώδικας για την εξαγωγή του Πίνακα 3.3

```
rm(list=ls(all=TRUE))
H <- 0.5107346 # Decision interval
K <- 1         # Reference value
n <- 5         # Sample size
m0 <- 0        # IC mean
s0 <- 1        # IC standard deviation

r <- 400       # Number of subintervals : I0, I1, I2, ..., Ir-1
(D <- H/(2*r))

FUN.ARL <- function (delta) {

  # Matrix Q
  Q <- matrix(0, nrow=r, ncol=r)
  # First column of Q
  for(j in 1:r){
    y <- -(2*(j-1)+1)*D + 2*D + K + m0
    Q[j,1] <- pbeta(      pnorm(  ((y-m0)/s0)-delta) ,      shapel=(n+1)/2,
    shape2=(n+1)/2 )
  }
  # Other columns of Q: Qjk
  for(j in 1:r){
    for(k in 2:r) {
      yU <- (2*(k-1)+1)*D - (2*(j-1)+1)*D + D + K + m0
      yL <- (2*(k-1)+1)*D - (2*(j-1)+1)*D - D + K + m0
      A <- pbeta(      pnorm(  ((yU-m0)/s0)-delta) ,      shapel=(n+1)/2,
      shape2=(n+1)/2 )
      B <- pbeta(      pnorm(  ((yL-m0)/s0)-delta) ,      shapel=(n+1)/2,
      shape2=(n+1)/2 )
      Q[j,k] <- A - B
    }
  }

  e0 <- rep(0,r); e0[1] <- 1 # Initial probability vector
  l1 <- rep(1,r)             # Vector of 1's
  Id <- diag(r)              # Identity r x r matrix
  ARLout <- as.vector(e0%*%solve(Id-Q)%*%l1)
  return(ARLout)
}

#####
#
# Evaluation of ARLout

delta <- c(0, 0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2)
ARL <- c()
for (i in 1:length(delta)){
  ARL[i] <- FUN.ARL(delta[i])
}

mat <- cbind(delta,ARL)
rownames(mat) <- rep("", nrow(mat))
colnames(mat) <- c("delta","ARL")
round(mat,digits=2)

0.0 370.40
0.1 208.97
0.2 120.80
```

0.3	71.62
0.5	27.38
0.7	11.94
1.0	4.57
1.5	1.81
2.0	1.20

Κ.6 Κώδικας για την εξαγωγή του Πίνακα 4.1

```
# EWMA Median Chart - Find K for given lambda and ARL0

rm(list=ls(all=TRUE))

n <- 3          # Sample size
m0 <- 0         # IC mean
s0 <- 1         # IC standard deviation
delta <- 0      # In-control case
lambda <- 0.1   # Smoothing parameter

r <- 200        # Number of states on each side
g <- 2*r + 1    # Total number of transient states

ARL_EWMA <- function(K) {

  LCL <- m0 - K * s0
  UCL <- m0 + K * s0

  D <- (UCL - LCL) / (2 * g)

  states <- -r:r
  H <- (LCL + UCL)/2 + 2 * states * D

  Q <- matrix(0, nrow=g, ncol=g)

  for (j in 1:g) {
    for (k in 1:g) {

      yU <- (H[k] + D - (1 - lambda) * H[j]) / lambda
      yL <- (H[k] - D - (1 - lambda) * H[j]) / lambda

      A <- pbeta(
        pnorm((yU - m0) / s0) - delta),
        shape1 = (n + 1)/2,
        shape2 = (n + 1)/2
      )

      B <- pbeta(
        pnorm((yL - m0) / s0) - delta),
        shape1 = (n + 1)/2,
        shape2 = (n + 1)/2
      )

      Q[j,k] <- A - B
    }
  }

  q <- rep(0, g)
  q[r + 1] <- 1      # Z0 = mu0, central state

  l1 <- rep(1, g)
  Id <- diag(g)

  ARL <- as.vector(t(q) %*% solve(Id - Q) %*% l1)
  return(ARL)
}

Find_K <- function(K_low, K_high, tolerance, max_iter) {
```

```

iter <- 0

while (iter < max_iter) {

  iter <- iter + 1
  K_mid <- (K_low + K_high) / 2

  estimated_arl <- ARL_EWMA(K_mid)

  cat("Iter", iter, ": K =", K_mid,
      ", Estimated ARL =", estimated_arl, "\n")

  if (abs(estimated_arl - ARL0) <= tolerance) {
    return(K_mid)
  } else if (estimated_arl < ARL0) {
    K_low <- K_mid
  } else {
    K_high <- K_mid
  }
}

warning("Maximum iterations reached; result may not be accurate.")
return(K_mid)
}

tolerance <- 0.01
max_iter <- 50
ARL0 <- 370.4

K_low <- 0.01
K_high <- 1

(result_K <- Find_K(K_low, K_high, tolerance, max_iter))
result_K

#####
Iter 1 : K = 0.505 , Estimated ARL = 1993.046
Iter 2 : K = 0.2575 , Estimated ARL = 38.42929
Iter 3 : K = 0.38125 , Estimated ARL = 212.7398
Iter 4 : K = 0.443125 , Estimated ARL = 600.9638
Iter 5 : K = 0.4121875 , Estimated ARL = 350.8761
Iter 6 : K = 0.4276562 , Estimated ARL = 456.9529
Iter 7 : K = 0.4199219 , Estimated ARL = 399.9349
Iter 8 : K = 0.4160547 , Estimated ARL = 374.4915
Iter 9 : K = 0.4141211 , Estimated ARL = 362.4646
Iter 10 : K = 0.4150879 , Estimated ARL = 368.4221
Iter 11 : K = 0.4155713 , Estimated ARL = 371.4427
Iter 12 : K = 0.4153296 , Estimated ARL = 369.9289
Iter 13 : K = 0.4154504 , Estimated ARL = 370.6849
Iter 14 : K = 0.41539 , Estimated ARL = 370.3067
Iter 15 : K = 0.4154202 , Estimated ARL = 370.4957
Iter 16 : K = 0.4154051 , Estimated ARL = 370.4012
[1] 0.4154051

```

K.7 Κώδικας για την εξαγωγή του Πίνακα 4.3

```
# EWMA Median Chart - ARL for given lambda and K

rm(list=ls(all=TRUE))

K <- 0.4154051      # Control limit coefficient
lambda <- 0.10      # Smoothing parameter
n <- 3              # Sample size
m0 <- 0             # IC mean
s0 <- 1             # IC standard deviation

r <- 200            # Number of states on each side
g <- 2*r + 1        # Total number of transient states

FUN.ARL <- function(delta) {

  LCL <- m0 - K * s0
  UCL <- m0 + K * s0

  D <- (UCL - LCL) / (2 * g)

  states <- -r:r
  H <- (LCL + UCL)/2 + 2 * states * D

  Q <- matrix(0, nrow=g, ncol=g)

  for (j in 1:g) {
    for (k in 1:g) {

      yU <- (H[k] + D - (1 - lambda) * H[j]) / lambda
      yL <- (H[k] - D - (1 - lambda) * H[j]) / lambda

      A <- pbeta(
        pnorm((yU - m0) / s0) - delta,
        shape1 = (n + 1)/2,
        shape2 = (n + 1)/2
      )

      B <- pbeta(
        pnorm((yL - m0) / s0) - delta,
        shape1 = (n + 1)/2,
        shape2 = (n + 1)/2
      )

      Q[j,k] <- A - B
    }
  }

  q <- rep(0, g)
  q[r + 1] <- 1      # Z0 = mu0, central state

  ll <- rep(1, g)
  Id <- diag(g)

  ARLout <- as.vector(t(q) %*% solve(Id - Q) %*% ll)
  return(ARLout)
}

#####
#
```

```

# Evaluation of ARLout

delta <- c(0, 0.1, 0.2, 0.3, 0.5, 0.7, 1, 1.5, 2)

ARL <- c()

for (i in 1:length(delta)) {
  ARL[i] <- FUN.ARL(delta[i])
}

mat <- cbind(delta, ARL)
rownames(mat) <- rep("", nrow(mat))
colnames(mat) <- c("delta", "ARL")

round(mat, digits=2)

delta    ARL
0.0 370.40
0.1 176.13
0.2  67.57
0.3  34.09
0.5  14.86
0.7   9.18
1.0   5.84
1.5   3.71
2.0   2.78

```

Κ.8 Κώδικας για την εξαγωγή του Σχήματος 4.5

```
rm(list=ls(all=TRUE))
library(qcc)
data(pistonrings)

# Data frame
attach(pistonrings)
PR <- qcc.groups(diameter, sample)
Data <- data.frame(X1=PR[,1], X2=PR[,2], X3=PR[,3], X4=PR[,4], X5=PR[,5])

# Parameters
m0 <- 74; s0 <- 0.01; n <- 5; m <- 40; lambda <- 0.1; K <- 0.3321246

# Limits
LCL <- m0 - K*s0
UCL <- m0 + K*s0

# Statistic median
Med <- c()
for(i in 1:m) {
  Med[i] <- median(as.numeric(Data[i,]))
}

# EWMA Median statistic
Zi_tilde <- c()

for(i in 1:m) {
  if(i == 1) {
    Zi_tilde[i] <- lambda*Med[i] + (1-lambda)*m0
  } else {
    Zi_tilde[i] <- lambda*Med[i] + (1-lambda)*Zi_tilde[i-1]
  }
}

# Results
ewma_results <- data.frame(Observation = Data, Median = Med,
                           EWMA_Median = Zi_tilde)
print(ewma_results)

# Plot
plot(1:m, Zi_tilde, pch=19, type="b", col="black",
     main="EWMA Median Chart - Piston Rings",
     xlab="Subgroup",
     ylab="Median",
     ylim=c(LCL - 0.001, max(Zi_tilde, UCL) + 0.001))

abline(h = UCL, col="black", lty=2)
abline(h = LCL, col="black", lty=2)
abline(h = m0, col="black", lty=2)

text(5, UCL + 0.0004, label="UCL=74.003321")
text(5, LCL - 0.0004, label="LCL=73.996679")
text(2, m0 - 0.0004, label="CL=74")

out <- which(Zi_tilde > UCL)
out
points(out, Zi_tilde[out], pch = 19, col = "red", cex = 1)
```


Βιβλιογραφία

Ελληνική

Αντζουλάκος, Δ. (2010). *Στατιστικός Έλεγχος Ποιότητας*, Πανεπιστημιακές Σημειώσεις, ΠΜΣ Εφαρμοσμένης Στατιστικής.

Αντζουλάκος, Δ. (2017). *Ανάλυση Δεδομένων με τη χρήση Στατιστικών Πακέτων*, Πανεπιστημιακές Σημειώσεις, ΠΜΣ Εφαρμοσμένης Στατιστικής.

Μπερσίμης, Σ., Ρακιτζής, Α. & Σαχλάς, Α. (2021). *Στατιστικός Έλεγχος Ποιότητας*, Εκδόσεις Τζιόλας, Θεσσαλονίκη.

Ταγαράς, Γ. (2001). *Στατιστικός Έλεγχος Ποιότητας*, Εκδόσεις Ζήτη, Θεσσαλονίκη.

Ξένη

Ahsanullah, M., Nevzorov, V. B. & Shakil, M. (2013). *An introduction to order statistics*, Springer.

Arnold, B. C., Balakrishnan, N. & Nagaraja, H. N. (1992). *A first course in order statistics*, Wiley.

Brook, D. & Evans, D. A. (1972). An approach to the probability distribution of Cusum run length. *Biometrika*, **59**(3), 539-549.

Castagliola, P. & Figueiredo, F.O. (2013). The median chart with estimated parameters. *European Journal of Industrial Engineering*, **7**(5), 594-614.

Castagliola, P., Figueiredo, F.O. & Maravelakis, P.E. (2019). The CUSUM Median Chart for Known and Estimated Parameters. *REVSTAT - Statistical Journal*, **17**(3), 345-370.

Castagliola, P., Maravelakis, P.E. & Figueiredo, F.O. (2016). The EWMA Median Chart with Estimated Parameters. *IIE Transactions*, **48**(1), 66-74.

Chakraborti, S. & Van Der Laan, P. (1996). Precedence tests and confidence bounds for complete data: an overview and some results. *Statistician*, **45**, 351-369.

Chakraborti, S., Van Der Laan, P. & Bakir, S. T. (2001). Nonparametric control charts: an overview and some results. *J. Qual. Technol.*, **33**, 304-315.

Chakraborti, S., Van Der Laan, P. & Van De Wiel, M. A. (2004). A class of distribution-free control charts. *Applied Statistics*, **53**, 443–462.

Hu, X., Castagliola, P., Zhou, X. & Tang, A. (2019). Conditional Design of the EWMA Median Chart with Estimated Parameters. *Communications in Statistics - Theory and Methods*, **48**(8), 1871–1889.

Montgomery, D.C. (2013). *Introduction to Statistical Quality Control*, 7th edition. New York: John Wiley and Sons.

Nazir, H. Z., Riaz, M., Does, R. J. M. M. & Abbas, N. (2013). Robust cusum control charting. *Quality Engineering*, **25**(3), 211-224.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, **41**, 100-115.

Roberts, S. W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, **42**(1), pp. 97–102.

Yang, L., Pai, S. & Wang, Y.R. (2010). A novel cusum median control chart. *In Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS)*, 3, 1707-1710