



# ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΔΜΠΣ ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΣΕ ΣΥΓΧΡΟΝΕΣ  
ΤΕΧΝΟΛΟΓΙΕΣ ΚΑΙ ΣΤΑ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΑ

## Διαγράμματα ελέγχου και τεχνικές δειγματοληψίας

Ιωάννης Καραχώστας

ΑΜ: 09319008

*Επιβλέπων Καθηγητής*

Χρήστος Κουκουβίνος

Καθηγητής ΕΜΠ

Αθήνα, 2022





# NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL SCIENCES  
MSC MATHEMATICAL MODELING IN MODERN TECHNOLOGIES AND FINANCIAL  
ENGINEERING

## Control charts and sampling techniques

Ioannis Karakostas

*Supervising Professor*

Dr Christos Koukouvinos

NTUA Professor

Athens, 2022





## Περίληψη

Από την εισαγωγή της ιδέας των διαγραμμάτων ελέγχου για την παρακολούθηση διεργασιών από τον Walter Shewhart, η χρήση τους έχει καθιερωθεί σε μεγάλο εύρος εφαρμογών, ιδιαίτερα στη βιομηχανία. Αυτό οφείλεται στην ικανότητά τους να ανιχνεύουν μετατοπίσεις στον μέσο μιας διεργασίας, λαμβάνοντας πληροφορίες από δείγματα μίας ή περισσότερων παρατηρήσεων, βελτιώνοντας με αυτόν τον τρόπο την ποιότητά της.

Στο 1ο Κεφάλαιο της διπλωματικής εργασίας αυτής γίνεται παρουσίαση της απλούστερης μορφής διαγραμμάτων ελέγχου, των διαγραμμάτων τύπου Shewhart. Δύο άλλες κατηγορίες διαγραμμάτων παρουσιάζονται στη συνέχεια, τα διαγράμματα τύπου CUSUM και EWMA, στο 2ο και 3ο Κεφάλαιο αντίστοιχα, κοινό χαρακτηριστικό των οποίων είναι η εξάρτηση από παρελθοντικές παρατηρήσεις. Σε κάθε περίπτωση πάντως, θεωρούμε ότι ισχύει η υπόθεση της κανονικότητας των δεδομένων.

Η εργασία επικεντρώνεται, στη συνέχεια, στα διαγράμματα EWMA, στις ιδιότητές τους και στη μέτρηση της απόδοσής τους. Η αξιολόγησή τους γίνεται μέσω του μέσου μήκους ροής, το οποίο υπολογίζεται σε περιβάλλον R, είτε με χρήση της βιβλιοθήκης "spr" είτε με κώδικα που στηρίζεται στην προσέγγιση της μαρκοβιανής αλυσίδας. Στο 4ο Κεφάλαιο γίνεται παρουσίαση τεχνικών βελτίωσης των υπόψη διαγραμμάτων για γρηγορότερη αντίδραση σε μετατοπίσεις του μέσου, ενώ στο 5ο Κεφάλαιο παρουσιάζεται ένα τροποποιημένο διάγραμμα EWMA, με καλή ανιχνευτική ικανότητα για εύρος που περιλαμβάνει από μικρές και βαθμιαίες έως μεγάλες και απότομες μετατοπίσεις. Προς αυτή την κατεύθυνση σχεδιάστηκαν και τα συνδυαστικά διαγράμματα, ορισμένα εκ των οποίων παρουσιάζονται στο 6ο Κεφάλαιο. Πάντως, οι διάφορες εκδοχές των διαγραμμάτων EWMA που παρουσιάζονται προκύπτουν από επέμβαση είτε στη δομή του στατιστικού είτε στα όρια ελέγχου.

Ένας άλλος τρόπος βελτίωσης της απόδοσης των διαγραμμάτων ελέγχου είναι με επέμβαση στις πληροφορίες που λαμβάνονται από τα δείγματα. Έτσι, στο 7ο Κεφάλαιο, παρουσιάζονται διάφορες τεχνικές δειγματοληψίας, με κάποια μορφή ταξινόμησης στις παρατηρήσεις. Περιγράφονται αναλυτικά οι τεχνικές RSS και MRSS, καθώς και τα διαγράμματα EWMA που προκύπτουν από τη χρήση των αντίστοιχων ταξινομημένων δεδομένων. Επίσης, συγκρίνονται ως προς την ανιχνευτική τους ικανότητα. Αναφορά γίνεται και σε ορισμένες τροποποιήσεις των παραπάνω τεχνικών, συγκεκριμένα των DRSS, MDRSS, DMRSS και ODRSS, οι οποίες επίσης συγκρίνονται ως προς την ανιχνευτική τους ικανότητα.

Τέλος, αναδεικνύοντας το ζήτημα της αυτοσυσχέτισης των παρατηρήσεων, στο 8ο Κεφάλαιο παρουσιάζεται μια πρόσφατη εναλλακτική κατηγορία δειγματοληψιών που

αντιμετωπίζουν το πρόβλημα αυτό (s-skip, mixed) και προτείνεται περαιτέρω μελέτη της επίδοσης ενός διαγράμματος EWMA που βασίζεται στις τεχνικές αυτές.

**Λέξεις Κλειδιά**— Διαγράμματα ΣΕΔ, Διαγράμματα Ελέγχου με Κινητούς Μέσους και Εκθετικά Βάρη, Τεχνικές Δειγματοληψίας

## Abstract

Since the introduction of the idea of control charts for monitoring processes by Walter Shewhart, their use has been established in a wide variety of applications, especially in the industrial domain. This is due to their ability to detect shifts in the process mean, by getting information through samples of one or more observations, thus improving its quality.

In Chapter 1 of this thesis the simplest form of control charts, namely Shewhart control charts, are presented. Two additional categories, the CUSUM and the EWMA charts, are subsequently presented in Chapters 2 and 3 respectively, which share a common characteristic, dependence on past observations. In every case, we consider that our data are normally distributed.

Moving on, the thesis focuses on EWMA control charts, their properties and their performance measurement. Their evaluation is done through the average run length, which is calculated in R, using either the “spc” library or a markov chain based code. In Chapter 4, techniques that improve the EWMA charts, regarding faster response to mean shifts, are presented, while in Chapter 5, a modified EWMA chart is presented, which can detect shifts ranging from small and gradual to large and sudden shifts. It is towards this direction that the combined control charts were designed, some of which are presented in Chapter 6. In any case, the different versions of EWMA charts derive from modifying either the formula of the statistic or the control limits.

Another way of enhancing control charts is by interfering with the sample information. Thus, in Chapter 7, various sampling techniques are presented, with some form of ranking on the observations. The RSS and MRSS techniques are analyzed, as well as the EWMA charts that derive from the relevant ranked data. They are also compared in terms with their ARL performance. Moreover, reference is made to some modifications of the above sampling techniques, namely DRSS, MDRSS, DMRSS and ODRSS, which are compared in terms with their ARL performance as well.

Finally, pointing out the issue of autocorrelation of the observations, in Chapter 8 we present a recent alternative category of sampling techniques which faces this problem (s-skip, mixed samples) and propose further study of the performance of an EWMA scheme based on such techniques.

**Keywords**— SPC Charts, EWMA, Sampling Techniques



# Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κ. Χρήστο Κουκουβίνο, Καθηγητή του Ε.Μ.Π. και επιβλέποντα της διπλωματικής μου εργασίας, για την εισαγωγή μου στον κόσμο της Στατιστικής μέσω της διδασκαλίας του και για την παρότρυνσή του να επιλέξω θέμα πάνω στα Διαγράμματα Ελέγχου.

Επίσης, ιδιαίτερες ευχαριστίες θα ήθελα να εκφράσω στην υποψήφια διδάκτορα Αγγελική Λάππα για τον χρόνο που διάθεσε για διορθώσεις και καθοδήγηση, συμβάλλοντας στην ολοκλήρωση της παρούσας εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω και τον φίλο και απόφοιτο μεταπτυχιακό του παρόντος ΔΠΜΣ, Σπύρο Παρασκευά, για τη βοήθεια που μου προσέφερε σε θέματα προγραμματισμού.



# Περιεχόμενα

Λίστα Σχημάτων	xv
Λίστα Πινάκων	xix
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Διαγράμματα ελέγχου . . . . .	1
1.2 Ιστορική αναδρομή . . . . .	1
1.3 Διαγράμματα τύπου Shewhart . . . . .	4
1.3.1 Διαστήματα Εμπιστοσύνης και Όρια Ελέγχου . . . . .	6
1.4 Τύποι σφαλμάτων . . . . .	7
1.4.1 Σφάλμα Τύπου I . . . . .	8
1.4.2 Σφάλμα Τύπου II . . . . .	8
1.4.3 Ισχύς Στατιστικού Ελέγχου . . . . .	8
1.5 Μέτρηση Απόδοσης Διαγραμμάτων Ελέγχου . . . . .	9
1.5.1 Μέσο Μήκος Ροής (ARL) . . . . .	9
1.5.2 Τυπική Απόκλιση Μήκους Ροής (SDRL) . . . . .	10
1.5.3 Διάμεσος Μήκους Ροής (MRL) . . . . .	10
1.5.4 Μέση Τετραγωνική Επιπλέον Απώλεια (AEQL) . . . . .	11
1.5.5 Μέσος Λόγος Μέσου Μήκους Ροής (ARARL) . . . . .	11
1.5.6 Δείκτης Σύγκρισης Απόδοσης (PCI) . . . . .	11
1.5.7 Ποσοστιαία Σημεία Μήκους Ροής . . . . .	12
1.6 Άλλοι τύποι διαγραμμάτων ελέγχου . . . . .	13
<b>2 Σωρευτικό Διάγραμμα (CUSUM)</b>	<b>15</b>
2.1 Εισαγωγή . . . . .	15
2.2 Κεντρική Ιδέα Διαγραμμάτων CUSUM . . . . .	15
2.3 Εφαρμογή με Διάγραμμα CUSUM . . . . .	16
2.4 Tabular CUSUM . . . . .	17
2.5 Βέλτιστη επιλογή των $K$ και $H$ . . . . .	18
2.6 Τυποποιημένο Διάγραμμα CUSUM . . . . .	19
2.7 CUSUM Διάγραμμα για Έλεγχο Μεταβλητότητας (Scale CUSUM) . . . . .	20
2.8 Διάγραμμα FIR-CUSUM . . . . .	20

2.8.1	Η ιδέα της Γρήγορης Αρχικής Αντίδρασης . . . . .	20
2.8.2	Παράδειγμα με εντός ελέγχου διεργασία . . . . .	21
2.8.3	Παράδειγμα με εκτός ελέγχου διεργασία . . . . .	22
<b>3</b>	<b>Διάγραμμα Ελέγχου με Κινητούς Μέσους με Εκθετικά Βάρη (EWMA)</b>	<b>23</b>
3.1	Το Στατιστικό EWMA . . . . .	23
3.2	Διάγραμμα EWMA με ασυμπτωτικά όρια . . . . .	26
3.3	Διάγραμμα EWMA με χρονικά μεταβαλλόμενα όρια . . . . .	28
3.4	Υπολογισμός ARL για διαγράμματα EWMA . . . . .	32
3.4.1	Μέθοδος Επίλυσης Συστήματος Εξισώσεων . . . . .	34
3.4.2	Μέθοδος Extrapolation . . . . .	35
3.4.3	Μέθοδος Μαρκοβιανής αλυσίδας . . . . .	36
3.4.4	Σύγκριση Μεθόδων – Συμπεράσματα . . . . .	40
3.5	Ιδιότητες ARL διαγραμμάτων EWMA . . . . .	43
3.6	Βέλτιστες παράμετροι $\lambda$ και $L$ . . . . .	47
<b>4</b>	<b>Μέθοδος Γρήγορης Αρχικής Απόκρισης (FIR-EWMA)</b>	<b>53</b>
4.1	Διαγράμματα Ελέγχου FIR – EWMA . . . . .	53
4.1.1	Επιλογή βέλτιστης τιμής της παραμέτρου $f$ . . . . .	56
4.2	FIR με αρχική τιμής εκκίνησης . . . . .	58
4.2.1	Επιλογή της τιμής εκκίνησης . . . . .	60
4.2.2	Διαγράμματα EWMA με Ασυμπτωτικά Όρια Ελέγχου και με Head Start	61
4.2.3	Διαγράμματα EWMA με Χρονικά Μεταβαλλόμενα Όρια Ελέγχου και με Head Start . . . . .	62
4.2.4	Διαγράμματα FIR-EWMA με Head Start . . . . .	63
4.3	Τροποποιημένο FIR-EWMA . . . . .	64
<b>5</b>	<b>Τροποποιημένο Διάγραμμα EWMA</b>	<b>69</b>
5.1	Εισαγωγή . . . . .	69
5.2	Στατιστικό Τροποποιημένου EWMA . . . . .	69
5.3	Διάγραμμα Τροποποιημένου EWMA . . . . .	73
5.4	Εφαρμογή Τροποποιημένου EWMA . . . . .	73
5.5	Υπολογισμός ARL τροποποιημένου EWMA . . . . .	74
5.6	Συμπέρασμα . . . . .	77
<b>6</b>	<b>Συνδυαστικά Διαγράμματα EWMA</b>	<b>83</b>
6.1	Εισαγωγή . . . . .	83
6.2	Διαγράμματα Διασποράς EWMA . . . . .	83
6.3	Συνδυαστικά διαγράμματα EWMA Μέσου – Διασποράς . . . . .	84
6.3.1	Διαγράμματα Ελέγχου Μέσου – Διασποράς EWMA . . . . .	84
6.3.2	Στατιστικό για Διαγράμματα Μέσου – Διασποράς EWMA . . . . .	84
6.3.3	Διαγράμματα Ελέγχου Μέσου - Διασποράς EWMA . . . . .	85
6.3.4	ARL Διαγραμμάτων Μέσου – Διασποράς EWMA . . . . .	85



6.3.5	Ανιχνευτική ικανότητα Διαγραμμάτων Μέσου - Διασποράς EWMA . . .	88
6.3.6	Εφαρμογή Διαγράμματος Μέσου - Διασποράς EWMA . . . . .	89
6.4	Συνδυαστικά διαγράμματα Shewhart–EWMA . . . . .	92
6.4.1	Στατιστικό των διαγραμμάτων Shewhart – EWMA . . . . .	92
6.4.2	Διάγραμμα ελέγχου Shewhart – EWMA . . . . .	92
6.4.3	Υπολογισμός ARL για τα διαγράμματα Shewhart – EWMA . . . . .	93
6.4.4	Ανιχνευτική ικανότητα Διαγραμμάτων Shewhart – EWMA . . . . .	93
6.4.5	Εφαρμογή Διαγράμματος Shewhart – EWMA . . . . .	94
6.5	Συμπέρασμα . . . . .	96
<b>7</b>	<b>Διαγράμματα EWMA με ταξινομημένες δειγματοληψίες</b>	<b>99</b>
7.1	Εισαγωγή στις Τεχνικές Δειγματοληψίας . . . . .	99
7.2	Τεχνική RSS (Ranked Set Sampling) . . . . .	100
7.2.1	Εκτιμητές Τεχνικής RSS . . . . .	103
7.2.2	Διάγραμμα Ελέγχου EWMA με τεχνική RSS . . . . .	105
7.2.3	Ανιχνευτική ικανότητα Διαγραμμάτων RSS-EWMA . . . . .	108
7.2.4	Εφαρμογή Τεχνικής EWMA σε RSS δεδομένα . . . . .	113
7.2.5	Διάγραμμα EWMA με Ατελή Ταξινόμηση Δειγματοληψίας RSS (IRSS)	114
7.2.6	Ανιχνευτική ικανότητα Διαγραμμάτων IRSS-EWMA . . . . .	117
7.3	Τεχνική MRSS (Median Ranked Set Sampling) . . . . .	118
7.3.1	Εκτιμητές Τεχνικής MRSS . . . . .	119
7.3.2	Διάγραμμα Ελέγχου EWMA με τεχνική MRSS . . . . .	121
7.3.3	Ανιχνευτική ικανότητα διαγραμμάτων MRSS-EWMA . . . . .	122
7.3.4	Εφαρμογή Τεχνικής EWMA σε MRSS δεδομένα . . . . .	123
7.3.5	Διάγραμμα EWMA με Ατελή Ταξινόμηση Δειγματοληψίας MRSS (IMRSS)	124
7.3.6	Ανιχνευτική ικανότητα διαγραμμάτων IMRSS - EWMA . . . . .	125
7.4	Τροποποιήσεις Τεχνικών RSS . . . . .	125
7.4.1	Τεχνική DRSS (Double Ranked Set Sampling) . . . . .	126
7.4.2	Τεχνική MDRSS (Median Double Ranked Set Sampling) . . . . .	129
7.4.3	Τεχνική DMRSS . . . . .	131
7.4.4	Διάγραμμα EWMA με τεχνικές MDRSS και DMRSS . . . . .	133
7.4.5	Ανιχνευτική ικανότητα διαγράμματος EWMA με τεχνικές MDRSS και DMRSS . . . . .	134
7.4.6	Εφαρμογή διαγράμματος EWMA με τεχνικές MDRSS και DMRSS . .	135
7.5	Τεχνική δειγματοληψίας ODRSS . . . . .	137
7.5.1	Εκτιμητές Τεχνικής ODRSS . . . . .	139
7.5.2	Διάγραμμα Ελέγχου EWMA με τεχνική ODRSS . . . . .	141
7.5.3	Ανιχνευτική ικανότητα Διαγραμμάτων EWMA-ODRSS . . . . .	142
7.5.4	Διάγραμμα EWMA με Ατελή Ταξινόμηση Δειγματοληψίας ODRSS (OIDRSS) . . . . .	142
7.5.5	Ανιχνευτική ικανότητα Διαγραμμάτων EWMA-OIDRSS . . . . .	143

<b>8</b>	<b>Σύνοψη - Πρόταση</b>	<b>145</b>
8.1	Το πρόβλημα της αυτοσυσχέτισης . . . . .	145
8.2	Τεχνικές δειγματοληψίας για την αντιμετώπιση της αυτοσυσχέτισης . . . . .	146
8.3	Πρόταση . . . . .	147
	<b>Παραρτήματα</b>	<b>149</b>
<b>A</b>	<b>Δεδομένα</b>	<b>150</b>
<b>B</b>	<b>Κώδικας υπολογισμού <i>ARL</i> σε <i>R</i></b>	<b>155</b>
	<b>Βιβλιογραφία</b>	<b>169</b>

# Ακρωνύμια

**ΣΕΔ** Στατιστικός Έλεγχος Διεργασιών. 1

**ACUSUM** Adaptive Cumulative Sum. 3

**AEQL** Average Extra Quadratic Loss. 11

**AEWMA** Adaptive Exponentially Weighted Moving Average. 2

**ARARL** Average Ratio of Average Run Length. 11, 12

**ARL** Average Run Length. 3, 7, 9–12, 35, 40, 42–44, 46–48, 56, 57, 64, 65, 75, 85–88, 93, 94, 108, 110, 112, 116, 117, 122, 125, 134, 142, 143

**ARL<sub>0</sub>** In-control Average Run Length. 7, 9, 10, 18, 21, 26, 28, 43, 47, 48, 50, 55, 56, 58, 69, 73, 75, 77, 88, 93, 113, 134, 136, 142, 143

**ARL<sub>1</sub>** Out-of-control Average Run Length. 7, 10, 21, 26–30, 33, 34, 44, 47–50, 55, 57, 58, 64, 66, 69, 74, 89, 113, 114

**BLUEs** Best Linear Unbiased Estimators. 137

**BLUEs-ODRSS** Best Linear Unbiased Estimators based on Ordered Double Ranked Set Sampling. 137, 140, 141

**BLUEs-ORSS** Best Linear Unbiased Estimators based on Ordered Ranked Set Sampling. 137

**BLUEs-OS** Best Linear Unbiased Estimators based on Simple Random Sampling. 137

**BLUEs-RSS** Best Linear Unbiased Estimators based on Ranked Set Sampling. 137

**CDF** Cumulative Density Function. 32

**CL** Center Line. 5, 6

**CS-EWMA** Cumulative Sum of Exponentially Weighted Moving Averages. 3

**CSEWMA** Combined Shewhart Exponentially Weighted Moving Average. 92–96

**CUSUM** Cumulative Sum. 1–3, 13, 15–17, 21, 23, 47, 96

**CUSUM-EWMA** Combined Cumulative Sum Exponentially Weighted Moving Average. 96

**DMRSS** Double Median Ranked Set Sampling. 3, 99, 124, 131–133, 136

**DMRSS-EWMA** Double Median Ranked Set Sampling - Exponentially Weighted Moving Average. 136

**DRSS** Double Ranked Set Sampling. 99, 126, 127, 138

**EWMA** Exponentially Weighted Moving Average. 1–4, 12, 13, 23–33, 35–38, 41, 43, 44, 46, 47, 53–56, 58, 60, 61, 63, 69–75, 77, 78, 83–85, 88, 89, 91–94, 96, 99, 105, 107, 109, 113, 114, 116, 122, 123, 133, 136, 141, 142, 145, 147

**FIR** Fast Initial Response. 2, 21, 22, 47, 53, 54, 56, 57, 63, 69

**FIR-CUSUM** Fast Initial Response Cumulative Sum. 21, 47, 53, 56, 58

**FIR-EWMA** Fast Initial Response Exponentially Weighted Moving Average. 3, 54–58, 63, 64, 66

**HEWMA** Hybrid Exponentially Weighted Moving Average. 3

**hs** Head Start. 59–61

**iid** Independent and Identically Distributed. 145

**IMRSS-EWMA** Imperfect Median Ranked Set Sampling - Exponentially Weighted Moving Average. 125

**IRSS** Imperfect Ranked Set Sampling. 116

**IRSS-EWMA** Imperfect Ranked Set Sampling - Exponentially Weighted Moving Average. 116, 117

**ISO** International Organization for Standardization. 4

**LCL** Lower Control Limit. 4, 6, 7, 32, 36, 61, 74

**MDRSS** Median Double Ranked Set Sampling. 3, 99, 124, 129, 130, 133, 134, 136

**MDRSS-EWMA** Median Double Ranked Set Sampling - Exponentially Weighted Moving Average. 136

**MFIR-EWMA** Modified Fast Initial Response Exponentially Weighted Moving Average. 65, 66

**Mixed EWMA-CUSUM** Mixed Cumulative Sum and Exponentially Weighted Moving Average. 3

**MRL** Median Run Length. 10

**MRSS** Median Ranked Set Sampling. 3, 99, 118–125, 129, 132, 136

**MRSS-EWMA** Median Ranked Set Sampling - Exponentially Weighted Moving Average. 122–124, 136

**ODRSS** Ordered Double Ranked Set Sampling. 4, 99, 124, 137, 139–142

**ODRSS-EWMA** Ordered Double Ranked Set Sampling - Exponentially Weighted Moving Average. 143

**OIDRSS** Ordered Imperfect Double Ranked Set Sampling. 4, 142

**OIDRSS-EWMA** Ordered Imperfect Double Ranked Set Sampling - Exponentially Weighted Moving Average. 143

**OIRSS** Ordered Imperfect Ranked Set Sampling. 3, 4

**ORSS** Ordered Ranked Set Sampling. 3, 4, 137

**PCI** Performance Comparison Index. 11, 12

**PDF** Probability Density Function. 32, 36, 87

**RL** Run Length. 12, 38, 93

**RSS** Ranked Set Sampling. 3, 4, 99, 100, 104, 105, 114, 116, 118, 121–126, 129, 132, 133, 136–138

**RSS-EWMA** Ranked Set Sampling - Exponentially Weighted Moving Average. 105–108, 113, 114, 117, 122, 123

**Scale CUSUM** Scale Cumulative Sum. 20

**SDRL** Standard Deviation of Run Length. 3, 10, 134

**SPC** Statistical Process Control. 1, 4

**SRS** Simple Random Sampling. 3, 4, 23, 92, 99, 100, 122, 125, 137

**SRS-EWMA** Simple Random Sampling - Exponentially Weighted Moving Average. 108

**Tabular CUSUM** Tabular Cumulative Sum. 17–19, 21, 22

**UCL** Upper Control Limit. 4, 6, 7, 32, 36, 61, 74



# Λίστα Σχημάτων

1.1	Διάγραμμα Ελέγχου . . . . .	5
1.2	Όρια Διαγράμματος Ελέγχου . . . . .	6
1.3	Υπό συνθήκη αναμενόμενη καθυστέρηση $D_T$ και υπό συνθήκη ποσοστιαία καθυστέρηση $D_{\alpha,T}$ για διάγραμμα ( $\lambda=0.1$ ) με σταθερά όρια ( $c = 3$ ) και χρονικά μεταβαλλόμενα όρια ( $c = 3.0055$ ίδιο $ARL_0$ ) και μία μετατόπιση της τάξης του $\delta=1$ . . . . .	13
2.1	Διάγραμμα Ελέγχου Shewhart- $\bar{X}$ διάγραμμα . . . . .	17
2.2	Καμπύλη στατιστικού CUSUM . . . . .	17
2.3	Διάγραμμα Tabular CUSUM για τα δεδομένα του Πίνακα A.1 . . . . .	19
2.4	Διάγραμμα Scale CUSUM για τα δεδομένα του Πίνακα A.1 . . . . .	21
2.5	Διάγραμμα CUSUM με και χωρίς FIR για τα δεδομένα του Πίνακα A.2 . . . . .	21
2.6	Διάγραμμα CUSUM με και χωρίς FIR για τα δεδομένα του Πίνακα A.3 . . . . .	22
3.1	Διάγραμμα ελέγχου EWMA με ασυμπτωτικά όρια για τα δεδομένα του Πίνακα A.1 . . . . .	28
3.2	Διάγραμμα ελέγχου EWMA με χρονικά μεταβαλλόμενα όρια για τα δεδομένα του Πίνακα A.1 . . . . .	30
3.3	Διάγραμμα ελέγχου EWMA με χρονικά μεταβαλλόμενα όρια για τα δεδομένα του Πίνακα A.1 ( $\lambda=0.00008$ , $L = 1.27926$ ) . . . . .	32
3.4	Διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια για διάφορες τιμές της σταθεράς εξομάλυνσης $\lambda$ . . . . .	33
3.5	Το διάστημα μεταξύ των LCL και UCL χωρίζεται σε $g-1$ καταστάσεις πλάτους $w$ , ξεκινώντας από το κάτω όριο και προσθέτοντας κάθε φορά $w$ . . . . .	37
3.6	Διάγραμμα Ακρίβειας υπολογισμού του ARL με τις διάφορες μεθόδους για την εντός ελέγχου περίπτωση του διαγράμματος $\bar{X} - S^2$ EWMA . . . . .	42
3.7	Ακρίβεια υπολογισμού ARL (Extrapolation) σε σχέση με τους κόμβους τετραγωνισμού για διάφορες τιμές της σταθεράς εξομάλυνσης $\lambda$ . Τα ανωτέρω αποτελέσματα υπολογίστηκαν για $L = 3$ . . . . .	43
3.8	Απεικόνιση των ARL των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια ελέγχου . . . . .	45

3.9	Πιθανότητα συμμετοχής κάθε παρατήρησης στη διαμόρφωση του ARL διαγράμματος EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια, για διεργασία εντός ελέγχου . . . . .	46
3.10	Πιθανότητα συμμετοχής κάθε παρατήρησης στη διαμόρφωση του ARL διαγράμματος EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια, για διεργασία εκτός ελέγχου . . . . .	47
3.11	Διαγράμματα Μέσου Μήκους Ροής σε σχέση την Παράμετρο Εξομάλυνσης $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των ασυμπτωτικών ορίων και με $ARL_0=370$ . . . . .	48
3.12	Διαγράμματα Μέσου Μήκους Ροής σε σχέση την Παράμετρο Εξομάλυνσης $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των χρονικά μεταβαλλόμενων ορίων και με $ARL_0=370$ . . . . .	49
3.13	Διαγράμματα Μέσου Μήκους Ροής σε σχέση την Παράμετρο Εξομάλυνσης $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των χρονικά μεταβαλλόμενων ορίων και με $ARL_0=500$ . . . . .	50
3.14	Διαγράμματα Μέσου Μήκους Ροής σε σχέση το μέγεθος της μετατόπισης, για διάφορες τιμές της παραμέτρου $\lambda$ , για την περίπτωση των ασυμπτωτικών ορίων και με $ARL_0=370$ . . . . .	51
3.15	Διαγράμματα Παραμέτρου $L$ συναρτήσει της παραμέτρου εξομάλυνσης $\lambda$ , για τις χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια για $ARL_0=370$ και $ARL_0=500$	51
4.1	Διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια σε σχέση με όρια FIR-EWMA . . . . .	54
4.2	Διάγραμμα FIR-EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	56
4.3	Διάγραμμα FIR-EWMA με βέλτιστο $f$ με τα δεδομένα του Πίνακα A.1 . . . . .	58
4.4	Διαγράμματα FIR-EWMA για βέλτιστους συνδυασμούς $\lambda$ , $L$ και $f$ με τα δεδομένα του Πίνακα A.1 . . . . .	59
4.5	Διαγράμματα EWMA (ασυμπτωτικά όρια) με $hs$ για τα δεδομένα του Πίνακα A.1 . . . . .	62
4.6	Διαγράμματα EWMA (χρονικά μεταβαλλόμενα όρια) με $hs$ για τα δεδομένα του Πίνακα A.1 . . . . .	62
4.7	Διαγράμματα FIR-EWMA με $hs$ για τα δεδομένα του Πίνακα A.1 . . . . .	64
4.8	Διαγράμματα FIR-EWMA με $hs$ για διάφορες τιμές παραμέτρων για τα δεδομένα του Πίνακα A.1 . . . . .	66
4.9	Διάγραμμα MFIR-EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	67
5.1	Διάγραμμα στατιστικού τροποποιημένου EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	71
5.2	Διάγραμμα στατιστικού τροποποιημένου EWMA για τα δεδομένα του Πίνακα A.5 . . . . .	72
5.3	Βοηθητικά διαγράμματα στο Σχήμα 5.2: 2ο μέρος (αριστερά) και 1ο-3ο μέρος (δεξιά) του στατιστικού του τροποποιημένου EWMA . . . . .	72



5.4	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	74
5.5	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	74
5.6	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.1 . . . . .	75
5.7	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.5 . . . . .	75
5.8	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.5 . . . . .	76
5.9	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.5 . . . . .	76
5.10	Τροποποιημένο EWMA για τα δεδομένα του Πίνακα A.4 . . . . .	77
5.11	Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 1 (αριστερά) και 2 παρατηρήσεις (δεξιά) . . . . .	78
5.12	Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 3 (αριστερά) και 4 παρατηρήσεις (δεξιά) . . . . .	78
5.13	Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 5 (αριστερά) και 6 παρατηρήσεις (δεξιά) . . . . .	79
5.14	Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 7 (αριστερά) και 8 παρατηρήσεις (δεξιά) . . . . .	79
5.15	Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 9 (αριστερά) και 10 παρατηρήσεις (δεξιά) . . . . .	79
6.1	Διάγραμμα EWMA για μέσο . . . . .	91
6.2	Διάγραμμα EWMA για τη διασπορά . . . . .	91
6.3	Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα A.1 ( $n=4$ ) . . . . .	97
6.4	Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα A.1 ( $n=1$ ) . . . . .	97
6.5	Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα A.1 ( $n=2$ ) . . . . .	98
7.1	Παράδειγμα δειγματοληψίας RSS . . . . .	101
7.2	Σύνολο παρατηρήσεων (αριστερά) - Παρατηρήσεις 1ου κύκλου (πάνω δεξιά) - Σετ No.1 1ου κύκλου (κάτω δεξιά) . . . . .	103
7.3	Επιλογή παρατήρησης από ταξινομημένα RSS σετ . . . . .	103
7.4	ARL διαγραμμάτων RSS-EWMA για $\lambda=0.14$ και τρεις διαφορετικές διασπορές των τιμών δειγματοληψίας . . . . .	111
7.5	ARL διαγραμμάτων RSS-EWMA για $\lambda=0.35$ και τρεις διαφορετικές διασπορές των τιμών δειγματοληψίας . . . . .	112
7.6	Διάγραμμα RSS-EWMA για τα δεδομένα των Πινάκων A.6 και A.7 . . . . .	114
7.7	Σύγκριση διαγραμμάτων RSS-EWMA και EWMA . . . . .	115
7.8	Σύγκριση διαγραμμάτων RSS-EWMA και EWMA . . . . .	115
7.9	Σύγκριση διαγραμμάτων RSS-EWMA και EWMA . . . . .	116
7.10	Σύγκριση διαγραμμάτων RSS-EWMA και EWMA . . . . .	116
7.11	Παράδειγμα Τεχνικής Δειγματοληψίας MRSS για περιττό πλήθος σετ . . . . .	120
7.12	Παράδειγμα Τεχνικής Δειγματοληψίας MRSS για άρτιο πλήθος σετ . . . . .	121
7.13	Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	124
7.14	Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	124
7.15	Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	125

7.16 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	125
7.17 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	126
7.18 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	126
7.19 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	127
7.20 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	127
7.21 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	128
7.22 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	128
7.23 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	129
7.24 Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA . . . . .	129
7.25 Δεδομένα εφαρμογής διαγράμματος EWMA με τεχνικές SRS MRSS, MDRSS και DMRSS . . . . .	137
7.26 Σύγκριση Διαγράμματος MRSS-EWMA ( $\lambda=0.1, L=2.820$ ) με Διάγραμμα SRS- EWMA ( $\lambda=0.1, L=2.820$ ) . . . . .	138
7.27 Σύγκριση Διαγράμματος MDRSS-EWMA ( $\lambda=0.1, L=2.819$ ) με Διάγραμμα SRS-EWMA ( $\lambda=0.1, L=2.820$ ) . . . . .	138
7.28 Σύγκριση Διαγράμματος DMRSS-EWMA ( $\lambda=0.1, L=2.819$ ) με Διάγραμμα SRS-EWMA ( $\lambda=0.1, L=2.820$ ) . . . . .	139
8.1 Τεχνικές δειγματοληψίας για την αντιμετώπιση της αυτοσυσχέτισης . . . . .	147

# Λίστα Πινάκων

1.1	Σημαντικότεροι κανόνες ευαισθητοποίησης διαγραμμάτων ελέγχου . . . . .	7
1.2	Δυνατοί συνδυασμοί λήψης απόφασης – Κατάσταση διεργασίας . . . . .	8
3.1	Ασυμπτωτικά Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda, L$ ) για $ARL_0=370$ , για ανίχνευση διάφορων μετατοπίσεων $\delta$ , όταν συμβαίνουν στην 1η παρατήρηση ( $q = 1$ ) . . . . .	27
3.2	Χρονικά Μεταβαλλόμενα Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda, L$ ) για $ARL_0=370$ , για ανίχνευση διάφορων μετατοπίσεων $\delta$ , όταν συμβαίνουν στην 20η παρατήρηση ( $q = 19$ ) . . . . .	29
3.3	Χρονικά Μεταβαλλόμενα Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda, L$ ) για $ARL_0=370$ , για ανίχνευση διάφορων μετατοπίσεων $\delta$ , όταν συμβαίνουν στην 1η παρατήρηση ( $q = 1$ ) . . . . .	31
3.4	Σύγκριση Μεθόδων Υπολογισμού ARL . . . . .	40
3.5	Σύγκριση Μεθόδων Υπολογισμού ARL . . . . .	41
3.6	Σύγκριση Μεθόδων Υπολογισμού ARL . . . . .	41
3.7	Σύγκριση Μεθόδων Υπολογισμού ARL . . . . .	42
3.8	Τιμές ARL για το απλό EWMA με $ARL_0=500$ . . . . .	44
3.9	Τιμές ARL για $\lambda=0.1, 0.25$ και $0.5$ , και με $L=3$ για χρονικά μεταβαλλόμενα και σταθερά ασυμπτωτικά όρια, όπου $\delta$ =μετατόπιση του μέσου της διεργασίας, $fix =$ Σταθ. – Ασυμπ. όρια, $vacl$ =Χρον. Μεταβαλλ. όρια. . . . .	45
3.10	Ζεύγη τιμών παραμέτρων για $ARL_0=370$ . . . . .	49
3.11	Δυνατοί Συνδυασμοί Ενδεχομένων – Τελικής Ενέργειας . . . . .	52
4.1	Θεωρητικές τιμές ARL για διάφορες τιμές $f$ . . . . .	57
4.2	Θεωρητικές τιμές ARL για διάφορες τιμές $f$ . . . . .	57
4.3	Θεωρητικές τιμές ARL για διαφορετικούς συνδυασμούς $\lambda, L$ και $f$ . . . . .	58
4.4	Υπολογισμός της τιμής της ποσότητας $(1 - \lambda)^j$ , καθώς το $j$ αυξάνει . . . . .	60
4.5	Αποτελέσματα ARL διαγράμματος FIR-EWMA ( $f=0.5$ ) με και χωρίς $hs$ . . . . .	65
5.1	Ανάλυση του στατιστικού τροποποιημένου EWMA στα μέρη του για τα δεδομένα του Πίνακα A.1 . . . . .	80
5.2	Ανάλυση του στατιστικού τροποποιημένου EWMA στα μέρη του για τα δεδομένα του Πίνακα A.5 . . . . .	81

5.3	Ανιχνευτική ικανότητα διαγράμματος τροποποιημένου EWMA για διάφορες τιμές των παραμέτρων . . . . .	82
5.4	Ανιχνευτική ικανότητα διαγράμματος τροποποιημένου EWMA για διάφορες τιμές των παραμέτρων . . . . .	82
6.1	Ανιχνευτική Ικανότητα Μονόπλευρου Διαγράμματος Μέσου – Διασποράς EWMA	88
6.2	Ανιχνευτική Ικανότητα Δίπλευρου Διαγράμματος Μέσου – Διασποράς EWMA	89
6.3	Δεδομένα για Διάγραμμα Μέσου – Διασποράς EWMA . . . . .	90
6.4	Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA . . . . .	94
6.5	Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA . . . . .	95
6.6	Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA . . . . .	95
6.7	Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA . . . . .	96
7.1	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων SRS-EWMA και RSS-EWMA	109
7.2	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.35$ και $L=2.94863109$	
7.3	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.25$ και $L=2.9015$	110
7.4	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.24$ και $L=2.89326110$	
7.5	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.2$ και $L=2.8632$	111
7.6	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.14$ και $L=2.7895$	111
7.7	Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για $\lambda=0.14$ και $L=2.7895$	112
7.8	Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για $n=2$ , $\lambda=0.12$ και $L=2.75$ . . . . .	117
7.9	Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για $n=3$ , $\lambda=0.12$ και $L=2.75$ . . . . .	117
7.10	Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για $n=4$ , $\lambda=0.12$ και $L=2.75$ . . . . .	117
7.11	Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για $n=5$ , $\lambda=0.12$ και $L=2.75$ . . . . .	118
7.12	Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για $n=6$ , $\lambda=0.12$ και $L=2.75$ . . . . .	118
7.13	Ανιχνευτική Ικανότητα Διαγραμμάτων SRS, RSS και MRSS για $n=3$ και $n=4$	123
7.14	Ανιχνευτική Ικανότητα Διαγραμμάτων SRS, RSS και MRSS για $n=5$ και $n=6$	123
7.15	Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για $n=2$ , $\lambda=0.12$ και $L=2.75$ . . . . .	130
7.16	Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για $n=3$ , $\lambda=0.12$ και $L=2.75$ . . . . .	130
7.17	Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για $n=4$ , $\lambda=0.12$ και $L=2.75$ . . . . .	131
7.18	Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για $n=5$ , $\lambda=0.12$ και $L=2.75$ . . . . .	131
7.19	Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για $n=6$ , $\lambda=0.12$ και $L=2.75$ . . . . .	132

7.20	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Χρονικά μεταβαλλόμενα όρια, $n=5$ , $ARL_0=370$ . . . .	134
7.21	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Χρονικά μεταβαλλόμενα όρια, $n=5$ , $ARL_0=500$ . . . .	135
7.22	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Σταθερά όρια, $n=5$ , $ARL_0=500$ . . . . .	136
7.23	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA-ODRSS με κλασικά EWMA - Χρονικά μεταβαλλόμενα όρια, $n=5$ , $ARL_0=500$ . . . . .	143
7.24	Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA-OIDRSS με EWMA-ODRSS - Χρονικά μεταβαλλόμενα όρια, $n=5$ , $ARL_0=500$ . . . . .	144
A.1	Δεδομένα Montgomery . . . . .	150
A.2	Δεδομένα εντός ελέγχου διεργασίας . . . . .	151
A.3	Δεδομένα εκτός ελέγχου διεργασίας . . . . .	151
A.4	Απλό παράδειγμα EWMA για $\lambda=0.1$ . . . . .	152
A.5	Δεδομένα . . . . .	152
A.6	Δεδομένα για παράδειγμα RSS-EWMA (1) . . . . .	153
A.7	Δεδομένα για παράδειγμα RSS-EWMA (2) . . . . .	154



# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Διαγράμματα ελέγχου

Τα διαγράμματα ελέγχου είναι εργαλεία του Στατιστικού Ελέγχου Διεργασιών (ΣΕΔ) (Statistical Process Control, SPC), που βρίσκουν εφαρμογή σε διαδικασίες μέτρησης ποιότητας. Σαν ορισμό για την ποιότητα θεωρούμε την προσαρμογή στη χρήση. Τα χαρακτηριστικά ενός προϊόντος που παράγεται από μια παραγωγική διεργασία πρέπει να συμπίπτουν με τις προδιαγραφές βάσει των οποίων σχεδιάστηκε. Οι παραγωγικές διεργασίες μέσω των οποίων παράγεται το προϊόν θα πρέπει να έχουν πολύ μικρή μεταβλητότητα γύρω από τις ονομαστικές διαστάσεις του (ποιοτικά χαρακτηριστικά).

Υπό αυτή την έννοια, τα διαγράμματα ελέγχου χρησιμοποιούνται ιδιαίτερος στον καθορισμό ειδικών αιτιών πρόκλησης διακυμάνσεων, έτσι ώστε να ληφθούν εγκαίρως μέτρα για την εξάλειψή τους. Τα διαγράμματα με τα οποία θα ασχοληθούμε σ' αυτή την εργασία ανήκουν στην κατηγορία των διαγραμμάτων ελέγχου μεταβλητών, αυτών δηλαδή που παρακολουθούν ποιοτικά χαρακτηριστικά που μετριοούνται σε αριθμητική κλίμακα.

Από τα πρώτα και βασικότερα διαγράμματα ελέγχου αυτού του είδους είναι το διάγραμμα Shewhart, το οποίο χρησιμοποιείται για ανίχνευση μεγάλων μετατοπίσεων. Δύο εναλλακτικές των διαγραμμάτων Shewhart είναι τα Αθροιστικά διαγράμματα ελέγχου (Cumulative Sum, CUSUM) και τα διαγράμματα ελέγχου με Κινητούς Μέσους και Εκθετικά Βάρη (Exponentially Weighted Moving Average, EWMA), τα οποία έχουν τη δυνατότητα να ανιχνεύουν μικρές και μεσαίες αλλαγές στην ποιότητα.

### 1.2 Ιστορική αναδρομή

Ο Στατιστικός Έλεγχος Διεργασιών (ΣΕΔ) έχει στόχο να ανιχνεύσει τις ασυνήθιστες μεταβολές στις παραμέτρους των διεργασιών παραγωγής όσο το δυνατόν νωρίτερα. Η έννοια ενός διαγράμματος ελέγχου ως εργαλείο για το στόχο αυτό προτάθηκε για πρώτη φορά από

τον Walter A. Shewhart στις 16 Μαΐου του 1924 [11], ως ένα μονοσέλιδο υπόμνημα προς τον προϊστάμενό του στην εταιρία Bell Telephone όπου εργάζονταν. Σήμερα χρησιμοποιούνται εξελιγμένες τεχνικές στατιστικής παρακολούθησης διεργασιών που βασίζονται σε διαγράμματα όπως τα διαγράμματα EWMA και CUSUM. Τα διαγράμματα αυτά είναι με τέτοιο τρόπο κατασκευασμένα ώστε, σε αντίθεση με τα αντίστοιχα τύπου Shewhart, να λαμβάνεται υπόψη η παρελθοντική πληροφορία από προγενέστερες παρατηρήσεις.

Το διάγραμμα CUSUM για παρακολούθηση του μέσου μιας διεργασίας προτάθηκε από τον Page (1954) [42], ενώ το διάγραμμα EWMA εισήχθη από τον Roberts (1959) [47]. Και τα δύο διαγράμματα είναι περισσότερο ευαίσθητα στην ανίχνευση μικρών μετατοπίσεων στο μέσο της διεργασίας απ' ό,τι το διάγραμμα Shewhart. Για αυτό το λόγο, τα διαγράμματα EWMA και CUSUM χρησιμοποιούνται ευρέως στις χημικές και μεταποιητικές βιομηχανίες, όπου μικρές διαταραχές στην εκάστοτε διεργασία συνεπάγονται σοβαρές οικονομικές συνέπειες. Ένα μειονέκτημα και των δύο διαγραμμάτων είναι ότι έχουν ασυμπτωτικά όρια ελέγχου και είναι λιγότερο ευαίσθητα στην αναγνώριση προβλημάτων που εμφανίζονται κατά την έναρξη της διαδικασίας παρακολούθησης.

Για να αυξηθεί η ευαισθησία του διαγράμματος CUSUM απέναντι σε προβλήματα κατά την έναρξη (μετατόπιση του μέσου κατά τις πρώτες μετρήσεις), οι Lucas και Crosier (1982) [34] πρότειναν ένα εργαλείο γρήγορης απόκρισης (Fast Initial Response, FIR). Το εργαλείο αυτό βελτιώνει την απόδοση του διαγράμματος CUSUM, ειδικά κατά την εκκίνηση ή την επανεκκίνηση μετά από ένα σήμα ειδοποίησης εκτός ελέγχου που δίνεται από το διάγραμμα. Αργότερα, οι Lucas και Saccucci (1990) [35] εφάρμοσαν την ιδέα της γρήγορης αρχικής απόκρισης, προκειμένου να βελτιώσουν την απόδοση του διαγράμματος EWMA στην ανίχνευση προβλημάτων κατά την εκκίνηση. Οι Rhoads et al. (1996) [46] εισήγαγαν τη χρήση δύο μονόπλευρων διαγραμμάτων EWMA, τα οποία βασίζονται σε χρονικά μεταβαλλόμενα όρια με μία μη μηδενική τιμή εκκίνησης, δείχνοντας ότι αυτός ο συνδυασμός διαγραμμάτων EWMA επιτρέπει μία γρήγορη απόκριση στην ανίχνευση προβλημάτων κατά την εκκίνηση.

Στη συνέχεια, ο Steiner (1999) [49] επέκτεινε τη δουλειά τους και πρότεινε ένα απλό διάγραμμα EWMA, το οποίο βασίζεται σε μία FIR ρύθμιση. Η ρύθμιση αυτή στενεύει περαιτέρω τα όρια ελέγχου, ειδικά κατά την εκκίνηση, και αυξάνει την ανιχνευτική ικανότητα του διαγράμματος EWMA για πρώιμες αλλαγές στο μέσο της διεργασίας. Με διαγράμματα ελέγχου που βασίζονται σε εργαλεία FIR ασχολήθηκαν μετέπειτα και οι Knoth (2005) [28] και Chiu (2009) [15].

Πέρα από τα παραπάνω, έχουν υπάρξει περαιτέρω σημαντικές εξελίξεις και βελτιώσεις στα σχέδια των διαγραμμάτων EWMA και CUSUM. Οι Capizzi και Masarotto (2003) [13] εισήγαγαν ένα προσαρμοστικό EWMA (Adaptive Exponential Weighted Moving Average, AEWMA) για να μετριάσουν την αδυναμία ανίχνευσης ταυτόχρονα μικρών και μεγάλων μετατοπίσεων από το διάγραμμα EWMA, καθώς και την καθυστέρηση επαναφοράς σε εντός ελέγχου κατάσταση. Αυτό επιτεύχθηκε με τη χρήση βαρών στις παρελθοντικές παρατηρήσεις μέσω κατάλληλης συνάρτησης για το τρέχον σφάλμα. Αντιστοίχως, οι Jiang et al. (2008)



[25] πρότειναν τα προσαρμοστικά CUSUM (Adaptive Cumulative Sum, ACUSUM) εφαρμόζοντας την ίδια λογική.

Οι Li et al. (2010) [33] μελέτησαν τις περιπτώσεις όπου η περιγραφή του ποιοτικού χαρακτηριστικού μιας διεργασίας μέσω παραμετρικής κατανομής είναι αμφίβολη. Στο πλαίσιο αυτό, εισήγαγαν μια νέα κατηγορία διαγραμμάτων, τα μη-παραμετρικά διαγράμματα ελέγχου (Nonparametric Control Charts), προτείνοντας τα μη-παραμετρικά ανάλογα των CUSUM και EWMA.

Κατόπιν, οι Abbas et al. (2012) πρότειναν ένα μικτό EWMA-CUSUM διάγραμμα ελέγχου, το Mixed EWMA-CUSUM [3], ιδιαίτερα ευαίσθητο σε μικρές μετατοπίσεις του μέσου, καθώς και ένα αντίστοιχο διάγραμμα, το CS-EWMA [2], όπου το στατιστικό βασίζεται στο σωρευτικό άθροισμα των εκθετικά κατανομημένων κινητών μέσων. Το τελευταίο παρουσίασε καλύτερη απόδοση, σε σύγκριση με τα απλά CUSUM και EWMA, στην ανίχνευση τόσο θετικών όσο και αρνητικών μεταβολών της μεταβλητότητας. Την ίδια χρονιά, ακολούθησε ο Haq [20] με ένα υβριδικό EWMA διάγραμμα (Hybrid EWMA, HEWMA) συνδυάζοντας δύο EWMA και πετυχαίνοντας ανίχνευση μικρότερων μετατοπίσεων από το μικτό EWMA-CUSUM διάγραμμα.

Βασικό εργαλείο στη σύγκριση των διαφόρων τύπων διαγραμμάτων αποτελεί ο υπολογισμός, κυρίως μέσω προσομοίωσης Monte Carlo, του Μέσου Μήκους Ροής (Average Run Length, ARL) καθώς και της τυπικής απόκλισης του Μήκους Ροής (Standard Deviation of Run Length, SDRL), τα οποία θα αναλυθούν παρακάτω. Επεκτείνοντας τη δουλειά των Abbas et al. (2011) [1], οι Khoo et al. (2016) [26] πρότειναν τη χρήση μαρκοβιανής αλυσίδας (Markov Chain) στους υπόψη υπολογισμούς.

Όλες οι παραπάνω εκδοχές των διαγραμμάτων ελέγχου βασίζονται στην υπόθεση ότι τα δείγματα μιας διεργασίας λαμβάνονται με την τεχνική της Απλής Τυχαίας Δειγματοληψίας (Simple Random Sampling, SRS). Ωστόσο, μια πιο αποτελεσματική τεχνική προτάθηκε από τον McIntyre (1952) [36], η τεχνική δειγματοληψίας καταταγμένων συνόλων (Ranked Set Sampling, RSS), αξιοποιώντας την επιπλέον πληροφορία που προέρχεται από την ταξινόμηση των δειγμάτων πριν την πραγματική μέτρηση. Με την τεχνική αυτή, βελτιώνεται σημαντικά η εκτίμηση του πληθυσμιακού μέσου και μειώνεται η διασπορά της εκτίμησης.

Οι Abujiya et al. εφάρμοσαν διάφορες τεχνικές δειγματοληψίας σε διαγράμματα ελέγχου όπως: συνδυαστικά διαγράμματα Shewhart-EWMA με τεχνική RSS (2012) [8], απλά διαγράμματα CUSUM και EWMA με τεχνική RSS (2013) [4], απλά διαγράμματα EWMA, FIR-EWMA, διπλά και τριπλά EWMA, μικτά EWMA-CUSUM με τεχνική Median RSS (MRSS), Median Double RSS (MDRSS) και Double Median RSS (DMRSS) (2013) [6], συνδυαστικά διαγράμματα Shewhart-EWMA με τεχνική Double Median RSS (DMRSS) (2014) [5]. Οι Haq et al. (2015) [23] πρότειναν την εφαρμογή της τεχνικής Ordered RSS (ORSS) στα διαγράμματα EWMA, συμπεριλαμβανομένης και της περίπτωσης της ατελούς ταξινόμησης των δειγμάτων Ordered Imperfect RSS (OIRSS). Έδειξαν ότι, όσον αφορά την ανίχνευση τυχαίων μεταβολών στη μεταβλητότητα της διεργασίας, τα διαγράμματα αυτά έχουν καλύτερη

απόδοση από τα διαγράμματα EWMA με τυχαία δειγματοληψία. Την ίδια χρονιά, δουλεύοντας σε διαγράμματα MaxEWMA (διαγράμματα με ιδιαίτερη ικανότητα στην ανίχνευση μεταβολών τόσο στον μέσο όσο και στη μεταβλητότητα μιας διεργασίας), πρότειναν την εφαρμογή της τεχνικής Ordered Double RSS (ODRSS) καθώς και της ατελούς εκδοχής της Ordered Imperfect Double RSS (OIDRSS), δείχνοντας πάλι ότι υπερτερούν των αντίστοιχων διαγραμμάτων με δειγματοληψία SRS, ORSS και OIRSS.

Τέλος, όσον αφορά στην τυποποίηση των διαγραμμάτων ελέγχου, μιας και συναντούνται σε πλήθος εφαρμογών στη μηχανική, τα οικονομικά, την ιατρική και όχι μόνο, από το 2016 ο οργανισμός ISO έχει εκδώσει το πρότυπο ISO 7870-6:2016 για τα διαγράμματα EWMA.

### 1.3 Διαγράμματα τύπου Shewhart

Στη Στατιστική, γενικότερα, καλούμαστε να έρθουμε σε συμπεράσματα για πληθυσμούς που περιγράφονται από κατανομές με παραμέτρους, τις οποίες δεν γνωρίζουμε. Μπορούμε, ωστόσο, να αντλήσουμε δεδομένα (παρατηρήσεις) από τυχαία δείγματα ώστε να μπορέσουμε να εκτιμήσουμε τις εκάστοτε παραμέτρους. Έστω κάθε δείγμα αποτελείται από  $n$  παρατηρήσεις, που συμβολίζονται:

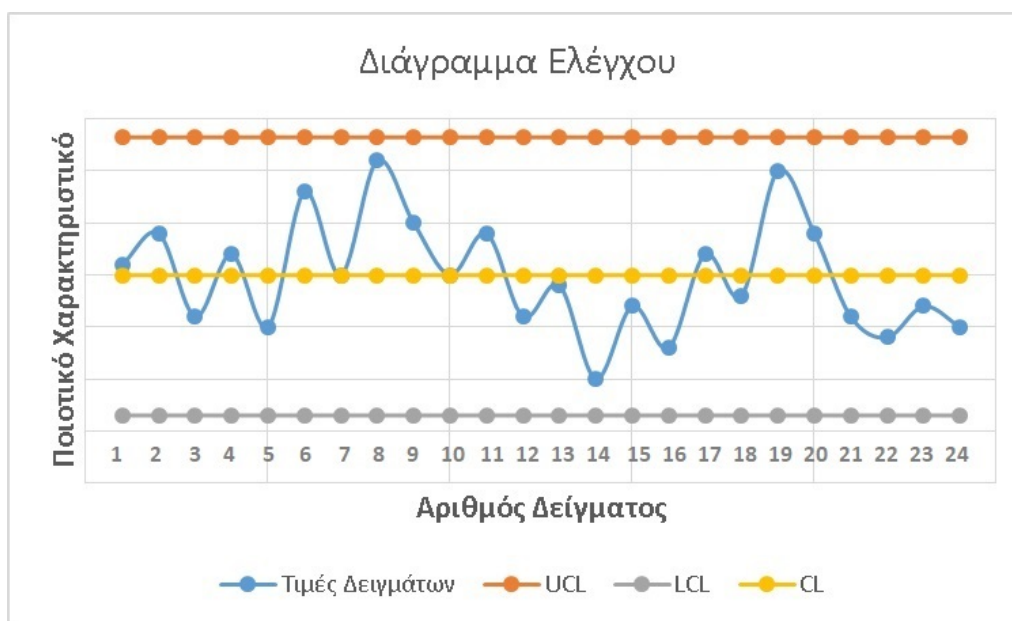
$$X_1, X_2, \dots, X_i, \dots, X_n \text{ (με τιμές } x_1, x_2, \dots, x_i, \dots, x_n \text{ αντίστοιχα)}$$

Κάθε παρατήρηση  $X_i$  αποτελεί τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας  $f_i(x_i)$ , που αν είναι κοινή, δηλαδή  $f_1(x_1) = f_2(x_2) = \dots = f_n(x_n) = f(x)$ , τότε το δείγμα  $X_1, X_2, \dots, X_n$  λέγεται τυχαίο. Η εκτίμηση μιας παραμέτρου επιτυγχάνεται με την εκτέλεση αριθμητικών πράξεων επί των παρατηρήσεων  $x_1, x_2, \dots, x_n$  και εκφράζεται με έναν τύπο. Τον τύπο αυτόν καλούμε **στατιστική συνάρτηση** ή **στατιστικό**. Απλό παράδειγμα αποτελεί η εκτίμηση του πληθυσμιακού μέσου  $\mu$  με το στατιστικό  $\bar{X} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$  (δειγματικός μέσος).

Μπορούμε, λοιπόν, να θεωρήσουμε τις μονάδες που παράγονται από μια διαδικασία παραγωγής ως τον πληθυσμό που χαρακτηρίζεται από συγκεκριμένες παραμέτρους (πληθυσμιακός μέσος  $\mu$ , πληθυσμιακή διασπορά  $\sigma^2$  κλπ). Οι παραγόμενες αυτές μονάδες προφανώς θα παρουσιάζουν διακυμάνσεις, οι οποίες κατατάσσονται σε δύο κατηγορίες: σε εκείνες που οφείλονται σε κοινές ή τυχαίες αιτίες (φυσική μεταβλητότητα) και σε εκείνες που οφείλονται σε ειδικές ή προσδιορισμένες αιτίες (ειδική μεταβλητότητα). Στην πρώτη περίπτωση λέμε ότι η διεργασία βρίσκεται **εντός στατιστικού ελέγχου** ή ότι λειτουργεί σε ευσταθή κατάσταση (stable state) και δεν απαιτούνται περαιτέρω ενέργειες. Στη δεύτερη περίπτωση, λέμε ότι η διεργασία βρίσκεται **εκτός στατιστικού ελέγχου** ή ότι λειτουργεί σε ασταθή κατάσταση (unstable state) και απαιτούνται διορθωτικές ενέργειες.

Το SPC περιλαμβάνει εργαλεία που μας βοηθούν να διακρίνουμε σε ποια από τις παραπάνω δύο κατηγορίες βρισκόμαστε. Ένα από αυτά είναι τα **όρια ελέγχου**, συγκεκριμένα το κάτω όριο (Lower Control Limit, LCL) και το άνω όριο (Upper Control Limit, UCL). Τοποθετούνται συνήθως στο συν ή μείον τρεις φορές την τυπική απόκλιση του στατιστικού

που αναπαρίσταται πάνω και κάτω μιας γραμμής, που τοποθετείται στον μέσο του στατιστικού και ονομάζεται κεντρική γραμμή (Central Line, CL). Μια γραφική αναπαράσταση ενός τυπικού διαγράμματος ελέγχου (Control Chart), καλούμενου και διαγράμματος τύπου Shewhart, παρουσιάζεται στο Σχήμα 1.1, το οποίο περιέχει στον οριζόντιο άξονα δειγματικούς αριθμούς (π.χ. 1=πρώτο δείγμα, 2=δεύτερο δείγμα κλπ) και στον κατακόρυφο άξονα κάποιο ποιοτικό χαρακτηριστικό (ή ένα στατιστικό αυτού).

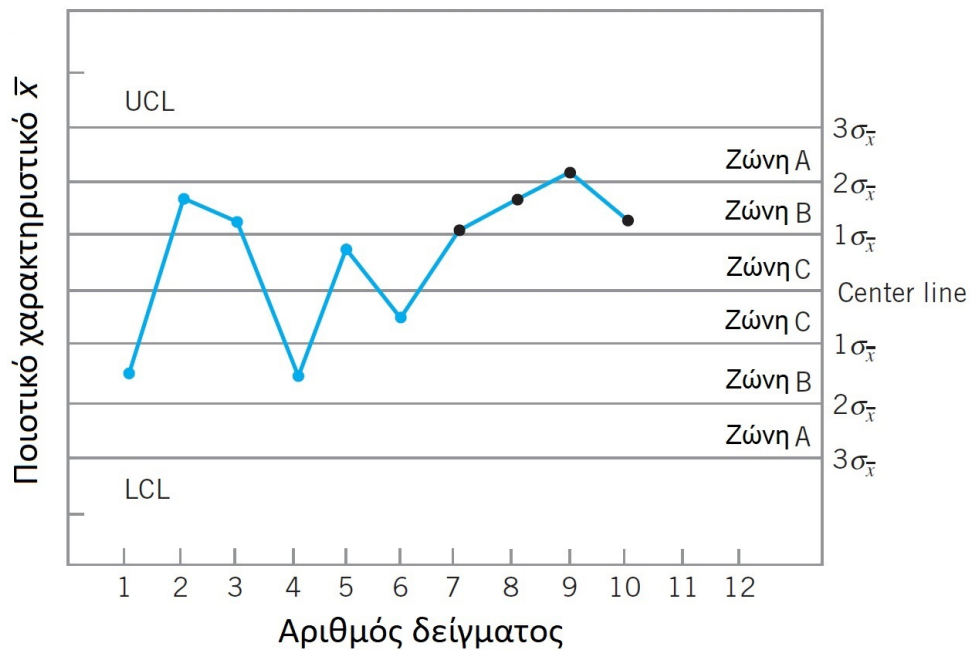


Σχήμα 1.1: Διάγραμμα Ελέγχου

Ακόμη και στην περίπτωση που όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου αλλά συμπεριφέρονται με ένα συστηματικό ή μη τυχαίο τρόπο, τότε και αυτό αποτελεί ένδειξη ότι η διεργασία είναι εκτός ελέγχου. Οι βασικότερες μορφές προτύπων ή μοτίβων (patterns) μη τυχαίου τρόπου συμπεριφοράς μίας χρονοσειράς σημείων είναι τα πρότυπα κύκλων (cycles), τα μιστά πρότυπα (mixtures), τα πρότυπα αλμάτων (shifts in the process level), τα πρότυπα τάσεων (trends) και τα πρότυπα έλλειψης μεταβλητότητας (stratification).

Για να γίνει περισσότερο ευαίσθητο ένα διάγραμμα ελέγχου (να ανιχνεύει πιο γρήγορα εκτός ελέγχου διεργασίες), δύναται επιπλέον των ορίων ελέγχου να σχεδιαστούν και προειδοποιητικά όρια (warning limits) εσωτερικά των ορίων ελέγχου, όπως φαίνεται στο Σχήμα 1.2. Η σχεδίαση των προειδοποιητικών ορίων ελέγχου προϋποθέτει την ανάπτυξη διαγραμμάτων ελέγχου χρησιμοποιώντας το μοντέλο των ορίων  $3\sigma$ .

Τα προειδοποιητικά όρια χρησιμοποιούνται μαζί με ορισμένους “κανόνες” που περιγράφουν την εμφάνιση ειδικών μοτίβων σε ένα διάγραμμα ελέγχου. Σε περίπτωση που εμφανιστεί το μοτίβο που περιγράφει ο κανόνας, τότε θεωρούμε ότι η διεργασία είναι εκτός ελέγχου χωρίς απαραίτητα να έχουμε κάποιο σημείο εκτός των ορίων ελέγχου. Οι σημαντικότεροι κανόνες που χρησιμοποιούνται για την ευαισθητοποίηση ενός διαγράμματος ελέγχου, όπως αναφέρονται από τον Montgomery (2013) [37], παρουσιάζονται στον Πίνακα 1.1.



Σχήμα 1.2: Όρια Διαγράμματος Ελέγχου

### 1.3.1 Διαστήματα Εμπιστοσύνης και Όρια Ελέγχου

Παραπάνω είδαμε (Σχήμα 1.1 και Σχήμα 1.2) ότι η καμπύλη του στατιστικού εμφανίζεται ανάμεσα στα δύο βασικά όρια, τα LCL και UCL. Ο ορισμός των ορίων αυτών, μαζί με την CL, είναι απαραίτητος για την κατασκευή ενός διαγράμματος ελέγχου.

Έστω ότι το στατιστικό που μας ενδιαφέρει είναι κανονικά κατανομημένο με γνωστό μέσο (ή τιμή στόχο)  $\mu_0$  και γνωστή τυπική απόκλιση  $\sigma_0$ . Τότε τα CL, LCL και UCL για δίπλευρο διάγραμμα ελέγχου Shewhart  $\bar{X}$  έχουν ως εξής:

$$UCL = \mu_0 + L \frac{\sigma_0}{\sqrt{n}} \quad (1.1)$$

$$CL = \mu_0 \quad (1.2)$$

$$LCL = \mu_0 - L \frac{\sigma_0}{\sqrt{n}} \quad (1.3)$$

Στις Σχέσεις (1.1) και (1.3) συμβολίζουμε με  $n$  το μέγεθος δείγματος, ενώ το  $L$  αντιπροσωπεύει τον συντελεστή των ορίων ελέγχου, δηλαδή καθορίζει το εύρος του διαστήματος εμπιστοσύνης μεταξύ των ορίων LCL και UCL. Είναι λογικό, στην περίπτωση που η διεργασία είναι εντός ελέγχου, να θέλουμε να συμπεριλαμβάνονται όσο το δυνατόν περισσότερες τιμές μεταξύ του LCL και UCL. Συγκεκριμένα, επιλέγονται με βάση τον βαθμό εμπιστοσύνης  $(1-\alpha)$  έτσι ώστε:

$$P(LCL \leq Z \leq UCL) = 1 - \alpha \quad (1.4)$$

Πίνακας 1.1: Σημαντικότεροι κανόνες ευαισθητοποίησης διαγραμμάτων ελέγχου

A/A	Κανόνας
1	1 ή περισσότερα σημεία εκτός των ορίων ελέγχου
2	2 από 3 συνεχόμενα σημεία σε μία από τις Ζώνες A
3	4 από 5 συνεχόμενα σημεία πέραν της Ζώνης C
4	8 συνεχόμενα σημεία στην ίδια μεριά της κεντρικής γραμμής
5	6 συνεχόμενα σημεία σε αύξουσα ή φθίνουσα διάταξη
6	15 συνεχόμενα σημεία στην ολική Ζώνη C
7	14 συνεχόμενα σημεία σε εναλλασσόμενη μορφή “πάνω-κάτω”
8	8 συνεχόμενα σημεία εκτός της ολικής Ζώνης C
9	Οποιαδήποτε ασυνήθιστη ή μη τυχαία ακολουθία σημείων
10	1 ή περισσότερα σημεία κοντά στα όρια ελέγχου

Από την παραπάνω Σχέση (1.4) μπορούμε να υπολογίσουμε τα LCL και UCL, έτσι ώστε ένα προκαθορισμένο ποσοστό των δυνατών τιμών να βρίσκονται μεταξύ τους. Π.χ. αν επιλέξουμε πιθανότητα  $\alpha=0.0027$  (επίπεδο σημαντικότητας), αυτό σημαίνει ότι το 99,73% των μετρήσεων που παίρνουμε θα βρίσκεται στο διάστημα μεταξύ LCL και UCL. Σ' αυτήν την περίπτωση, υπολογίζεται ότι η τιμή του  $L$  είναι 3.

Επιλέγοντας, δηλαδή, επίπεδο σημαντικότητας  $\alpha$ , προκαθορίζουμε άμεσα το ποσοστό των μετρήσεων που είναι εκτός των αποδεκτών τιμών και έμμεσα:

α. Το μέγιστο αποδεκτό ποσοστό που οδηγεί σε Σφάλμα Τύπου I, και συνεπώς το ρυθμό εσφαλμένου σήματος, δηλαδή το κάθε πόσες μετρήσεις αναμένουμε να πάρουμε εκτός ελέγχου σήμα, ενώ η διεργασία είναι εντός ελέγχου. Ο ρυθμός αυτός ονομάζεται Μέσο Μήκος Ροής για την εντός ελέγχου κατάσταση (in-control ARL) και συμβολίζεται με  $ARL_0$ .

β. Την ταχύτητα ανιχνευτικής ικανότητας του διαγράμματος, δηλαδή το πόσο γρήγορα θα ανιχνευτεί ένα εκτός ελέγχου σήμα, όταν η διεργασία βρίσκεται πραγματικά εκτός ελέγχου, που καθορίζεται από το Μέσο Μήκος Ροής για την εκτός ελέγχου κατάσταση (out-of-control ARL) και συμβολίζεται με  $ARL_1$ .

Γίνεται φανερό ότι αν τα όρια επιλεγούν με τέτοιο τρόπο, έτσι ώστε η μεταξύ τους περιοχή να είναι αρκετά μικρή, το διάγραμμά μας θα δίνει εκτός ελέγχου σήμα (ενώ η διεργασία είναι εντός ελέγχου) πολύ συχνά, ενώ αν τα όρια είναι πολύ μεγάλα, το διάγραμμά μας θα αργεί να δώσει εκτός ελέγχου σήμα και θα έχει πολύ μικρή ανιχνευτική ικανότητα.

## 1.4 Τύποι σφαλμάτων

Κατά τον δειγματοληπτικό έλεγχο σε μια παραγωγική διαδικασία είναι δυνατόν να διαπράξουμε δύο είδη εσφαλμένων εκτιμήσεων.

Έστω ότι παίρνουμε δείγματα από μία παραγωγική διαδικασία. Αν οι τιμές βρίσκονται

εντός προδιαγραφών, η παραγωγή συνεχίζεται. Αν, όμως, βρεθούν τιμές εκτός ελέγχου, τότε η παραγωγή διακόπτεται. Τα δυνατά αποτελέσματα του υπόψη ελέγχου μπορούμε να τα συνοψίσουμε στον Πίνακα 1.2.

Πίνακας 1.2: Δυνατοί συνδυασμοί λήψης απόφασης – Κατάσταση διεργασίας

	Εντός Ελέγχου (πραγματική κατάσταση)	Εκτός Ελέγχου (πραγματική κατάσταση)
Απόφαση για την παραγωγή		
Συνέχιση παραγωγής	Ορθή απόφαση	Σφάλμα Τύπου II (πιθανότητα β)
Διακοπή παραγωγής	Σφάλμα Τύπου I (πιθανότητα α)	Ορθή απόφαση (πιθανότητα γ) Ισχύς Στατιστικού Ελέγχου

### 1.4.1 Σφάλμα Τύπου I

Σφάλμα Τύπου I (Type I error) ονομάζεται το σφάλμα που γίνεται, όταν διακόπτεται η παραγωγή (το διάγραμμα ελέγχου ανίχνευσε εκτός ελέγχου σήμα), ενώ στην πραγματικότητα είναι εντός ελέγχου. Την πιθανότητα να κάνουμε σφάλμα τύπου I τη συμβολίζουμε ως εξής:

$$\alpha = P(\text{διακοπή παραγωγής} / \text{παραγωγή εντός ελέγχου}) \quad (1.5)$$

Η πιθανότητα α (επίπεδο σημαντικότητας) φανερώνει ότι το ενδεχόμενο το διάγραμμα ελέγχου να μας πληροφορήσει εσφαλμένα (δηλαδή ότι πρέπει να σταματήσουμε την παραγωγή, ενώ η διεργασία βρίσκεται εντός ελέγχου) μπορεί να συμβεί α φορές στις εκατό.

### 1.4.2 Σφάλμα Τύπου II

Σφάλμα Τύπου II (Type II error) ονομάζεται το σφάλμα που γίνεται, όταν συνεχίζεται η παραγωγή (το διάγραμμα ελέγχου δεν ανίχνευσε ακόμη κάποιο εκτός ελέγχου σήμα), ενώ στην πραγματικότητα η διεργασία είναι εκτός ελέγχου. Την πιθανότητα να κάνουμε σφάλμα τύπου II τη συμβολίζουμε ως εξής:

$$\beta = P(\text{συνέχιση παραγωγής} / \text{παραγωγή εκτός ελέγχου}) \quad (1.6)$$

Δηλαδή η πιθανότητα β είναι ο κίνδυνος συνέχισης της παραγωγής, ενώ πρέπει να διακοπεί.

### 1.4.3 Ισχύς Στατιστικού Ελέγχου

Εκτός από τις παραπάνω πιθανότητες α και β, μπορούμε να υπολογίσουμε και την πι-

θανότητα να σταματήσουμε την παραγωγή, όταν αυτή πράγματι βρίσκεται εκτός ελέγχου. Η πιθανότητα αυτή ονομάζεται δύναμη ή ισχύς του στατιστικού ελέγχου, και συμβολίζεται ως εξής:

$$\gamma = 1 - \beta = P(\text{διακοπή παραγωγής} / \text{παραγωγή εκτός ελέγχου}) \quad (1.7)$$

Όσο πιο μεγάλη είναι η τιμή της ισχύος του ελέγχου, τόσο πιο σωστή θα είναι η απόφασή μας. Στον έλεγχο στατιστικών υποθέσεων επιδιώκεται η σχετική ελαχιστοποίηση του  $\alpha$  και η σχετική μεγιστοποίηση της δύναμης  $\gamma$ .

## 1.5 Μέτρηση Απόδοσης Διαγραμμάτων Ελέγχου

Καθώς υπάρχουν διαγράμματα ελέγχου διαφορετικής σχεδίασης, είναι επόμενο για την αξιολόγησή τους να υπάρχει δυνατότητα μέτρησης της απόδοσής τους. Τα εργαλεία που συμβάλλουν σ' αυτόν το σκοπό παρουσιάζονται στη συνέχεια.

### 1.5.1 Μέσο Μήκος Ροής (ARL)

Από την πιθανότητα  $\alpha$  να βρεθεί ένα σημείο του διαγράμματος ελέγχου εκτός των ορίων ελέγχου (ενώ η διεργασία είναι στην πραγματικότητα εντός), υπολογίζουμε το ARL του διαγράμματος, δηλαδή το εντός ελέγχου  $ARL_0$ , από την παρακάτω σχέση:

$$ARL_0 = \frac{1}{\alpha} \quad (1.8)$$

Σημείωση: Αν η τυχαία μεταβλητή  $X$  ακολουθεί τη γεωμετρική κατανομή με πιθανότητα επιτυχίας  $p$  ( $X \sim G(p)$ ), δηλαδή:

$$P(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, 3, \dots \quad (1.9)$$

τότε

$$E(X) = \frac{1}{p}, \quad V(X) = \frac{1-p}{p^2} \quad (1.10)$$

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

Για μια διεργασία που βρίσκεται εντός ελέγχου και στην περίπτωση που χρησιμοποιούμε 3σ όρια ελέγχου με κανονική κατανομή για την απεικονιζόμενη στατιστική συνάρτηση, έχουμε ότι το εντός ελέγχου μέσο μήκος ροής  $ARL_0$  είναι ίσο με:

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

Προφανώς στην πράξη θέλουμε να έχουμε μεγάλο  $ARL_0$ .

Για μια διεργασία που βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου της, το εκτός ελέγχου μέσο μήκος ροής  $ARL_1$  είναι ίσο με:

$$ARL_1 = \frac{1}{1 - \beta} \quad (1.11)$$

Το  $\beta$  είναι ουσιαστικά η πιθανότητα ένα σημείο του διαγράμματος ελέγχου να είναι εντός των ορίων ελέγχου (ενώ η διεργασία είναι στην πραγματικότητα εκτός), άρα η πιθανότητα να είναι εκτός των ορίων ελέγχου (δηλαδή εκτός ελέγχου σήμα στο διάγραμμα) είναι ίση με πιθανότητα  $1 - \beta$ .

Προφανώς στην πράξη θέλουμε να έχουμε μικρό  $ARL_1$ , καθώς αυτό δηλώνει τον αναμενόμενο αριθμό δειγμάτων που πρέπει να ληφθούν, προκειμένου να εντοπιστεί μια μετατόπιση στο μέσο επίπεδο της διεργασίας, από τη στιγμή που η μετατόπιση αυτή θα συμβεί. Επισημαίνεται ότι, όσον αφορά τη χρονική στιγμή που λαμβάνει χώρα μια μετατόπιση σε σχέση με τη χρονική στιγμή παρακολούθησης μιας διεργασίας, διακρίνονται οι εξής περιπτώσεις: η μηδενική κατάσταση (zero-state), όπου κατά την εκκίνηση της παρακολούθησης έχει συμβεί η μετατόπιση, και σταθερή κατάσταση (steady-state), όπου αρχικά η διεργασία είναι εντός ελέγχου και σε κάποιο σημείο εμφανίζεται εκτός ελέγχου (μετατόπιση). Σε κάποιους από τους υπολογισμούς  $ARL$  που θα δούμε σε επόμενα κεφάλαια, θα δούμε την επίδραση των παραπάνω καταστάσεων στα αποτελέσματα των υπολογισμών.

Παρά την δημοτικότητά του, το  $ARL$  έχει δεχθεί έντονη αμφισβήτηση, καθώς αυτό που παρατηρείται στην πράξη, διαφέρει συνήθως αρκετά από το θεωρητικό  $ARL$  λόγω της ασυμμετρίας που εμφανίζει η κατανομή του μήκους ροής, με αποτέλεσμα η μέση τιμή του να μην μπορεί να θεωρηθεί ως αντιπροσωπευτικό μέτρο κεντρικής τάσης της κατανομής, ιδιαίτερα για μικρές τιμές του  $p$  [10] [12].

### 1.5.2 Τυπική Απόκλιση Μήκους Ροής (SDRL)

Η απόδοση ενός διαγράμματος ελέγχου εκφράζεται και μέσω της τυπικής απόκλισης των μετρούμενων μηκών ροής (Standard Deviation of Run Length, SDRL), δηλαδή των δειγματικών σημείων που πρέπει να σχεδιαστούν πριν εμφανιστεί ένα εκτός ελέγχου σήμα. Χρησιμοποιείται συχνά σαν ένας συμπληρωματικός δείκτης αποτελεσματικότητας της διεργασίας και υπολογίζεται ως εξής:

$$SDRL = \sqrt{Var(RL)} \quad (1.12)$$

### 1.5.3 Διάμεσος Μήκους Ροής (MRL)

Κατ' αναλογία με τον ορισμό του  $ARL$ , η διάμεσος του μήκους ροής (Median Run Length, MRL) είναι ο διάμεσος των δειγματικών σημείων που πρέπει να σχεδιαστούν στο



διάγραμμα πριν εμφανιστεί ένα εκτός ελέγχου σήμα, δηλαδή

$$MRL = \text{median}(RL) \quad (1.13)$$

#### 1.5.4 Μέση Τετραγωνική Επιπλέον Απώλεια (AEQL)

Η μέση τετραγωνική επιπλέον απώλεια (Average Extra Quadratic Loss, AEQL) αποτελεί ένα συνολικό μέτρο απόδοσης του διαγράμματος ελέγχου πάνω σε όλες τις μετατοπίσεις της τυπικής απόκλισης της διεργασίας. Η AEQL ορίζεται ως ο σταθμισμένος μέσος ARL πάνω σε ένα εύρος μετατοπίσεων  $\delta_{min} < \delta < \delta_{max}$  στην τυπική απόκλιση της διεργασίας. Οι ποσότητες  $\delta_{min}$  και  $\delta_{max}$  είναι τα άνω και κάτω όρια των μετατοπίσεων, αντίστοιχα. Η AEQL, δηλαδή, μετράει την συνολική αποτελεσματικότητα του διαγράμματος ελέγχου σε ένα εύρος μετατοπίσεων [45]. Δίνεται από την σχέση:

$$AEQL = \frac{1}{\delta_{max} - \delta_{min}} \int_{\delta_{min}}^{\delta_{max}} \delta^2 ARL(\delta) d\delta \quad (1.14)$$

όπου  $ARL(\delta)$  είναι η τιμή ARL ενός συγκεκριμένου διαγράμματος στην μετατόπιση  $\delta$ . Όσο πιο μικρή είναι η τιμή AEQL του διαγράμματος ελέγχου, τόσο καλύτερη είναι η συνολική στατιστική απόδοση του διαγράμματος.

#### 1.5.5 Μέσος Λόγος Μέσου Μήκους Ροής (ARARL)

Ο μέσος λόγος του ARL (Average Ratio of ARL, ARARL) είναι μία εξίσωση ολοκληρώματος που μετράει την συνολική απόδοση ενός διαγράμματος ελέγχου σε ένα ευρύ φάσμα μετατοπίσεων από  $\delta_{min}$  έως  $\delta_{max}$ , και υπολογίζεται από την σχέση [52]:

$$ARARL = \frac{1}{\delta_{max} - \delta_{min}} \int_{\delta_{min}}^{\delta_{max}} \frac{ARL(\delta)}{ARL(\delta)_{benchmark}} d\delta \quad (1.15)$$

όπου  $\delta_{min}$  και  $\delta_{max}$  είναι οι ελάχιστες και οι μέγιστες μετατοπίσεις, αντίστοιχα, στην τυπική απόκλιση της διεργασίας και  $ARL(\delta)$  και  $ARL(\delta)_{benchmark}$  είναι οι τιμές ARL ενός συγκεκριμένου διαγράμματος ελέγχου και ενός διαγράμματος σημείου αναφοράς (Benchmark Control Chart) στη μετατόπιση  $\delta$ , αντίστοιχα. Το διάγραμμα ελέγχου αναφοράς είναι εκείνο με το μικρότερο εκτός ελέγχου ARL.

#### 1.5.6 Δείκτης Σύγκρισης Απόδοσης (PCI)

Ο δείκτης σύγκρισης απόδοσης (Performance Comparison Index, PCI) είναι το ποσοστό της AEQL ενός συγκεκριμένου διαγράμματος ελέγχου με την AEQL ενός διαγράμματος σημείου αναφοράς και ορίζεται από την σχέση [41]:

$$PCI = \frac{AEQL}{AEQL_{benchmark}} d\delta \quad (1.16)$$

Ένα διάγραμμα ελέγχου με ελάχιστο ARARL δίνει μικρότερες τιμές AEQL και PCI και είναι πιο αποτελεσματικό στην ανίχνευση αλλαγών στην διεργασία.

### 1.5.7 Ποσοστιαία Σημεία Μήκους Ροής

Σύμφωνα με τον Knoth (2015) [30], η χρήση του ARL και μόνο, που υπενθυμίζουμε ότι πρόκειται για μέσο όρο, για την αξιολόγηση των διαγραμμάτων ελέγχου ακολουθείται από περιορισμούς. Μία άλλη προσέγγιση για την αξιολόγησή τους αποτελεί η παρακολούθηση της επίδρασης μεταβολών της διεργασίας πάνω σε κατάλληλα ορισμένα ποσοστιαία μήκη ροής των διαγραμμάτων. Έτσι, ορίζουμε τα ακόλουθα:

α. Ποσοστιαίο Σημείο Μήκους Ροής Μηδενικής Κατάστασης  $\mathcal{Q}_\alpha$  (zero-state quantile RL), που ορίζεται από τις παρακάτω σχέσεις:

$$\mathcal{Q}_\alpha = \inf\{i \geq 1 : P_\delta(RL \leq i) \geq \alpha\} \quad (1.17)$$

$$P_\delta(N \leq \mathcal{Q}_\alpha - 1) < \alpha \quad (1.18)$$

$$P_\delta(N \leq \mathcal{Q}_\alpha) \geq \alpha \quad (1.19)$$

όπου  $N$  ο χρόνος που εμφανίζεται σήμα εκτός ελέγχου (ισοδύναμο του RL και  $\delta$  η πραγματική τιμή της πιθανής μετατόπισης του μέσου.

β. Υπό Συνθήκη Υστέρηση Ποσοστιαίου Σημείου Μήκους Ροής  $\mathcal{D}_{\alpha,\tau}$  (conditional quantile delay), που ορίζεται από την παρακάτω σχέση:

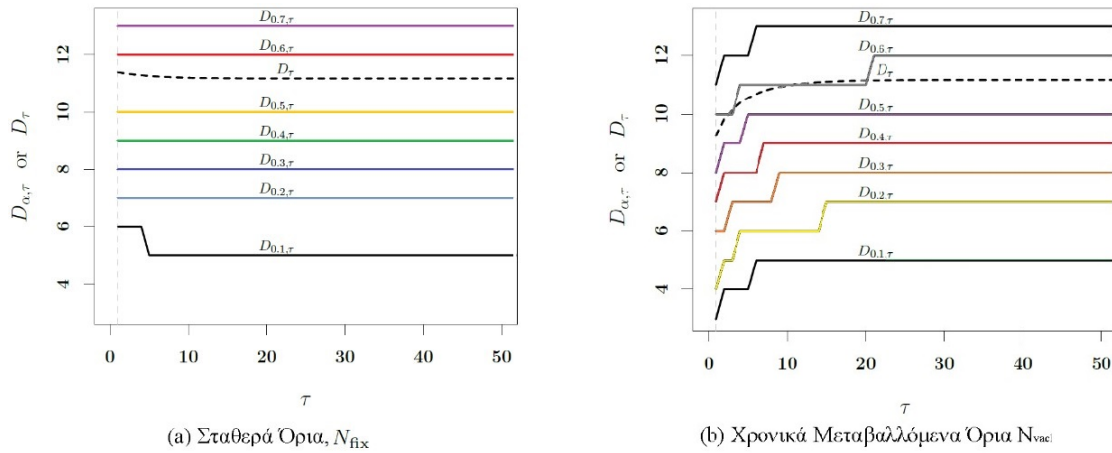
$$\mathcal{D}_{\alpha,\tau} = \inf\{i \geq 1 : P_\tau(N - \tau + 1 \leq i | N \geq \tau) \geq \alpha\} \quad (1.20)$$

όπου  $\tau$  είναι το σημείο αλλαγής (change point), δηλαδή το σημείο που χωρίζει την ακολουθία των σημείων του διαγράμματος ελέγχου στο εντός ελέγχου υποσύνολο, όπου  $i < \tau$ , και στο εκτός ελέγχου υποσύνολο, όπου  $i \geq \tau$ .

γ. Υστέρηση Ποσοστιαίου Σημείου Μήκους Ροής Σταθερής Κατάστασης  $\mathcal{D}_\alpha$  (steady-state quantile delay), που ορίζεται από την παρακάτω σχέση:

$$\mathcal{D}_\alpha = \lim_{\tau \rightarrow \infty} \mathcal{D}_{\alpha,\tau} \quad (1.21)$$

Ισχύει ότι το  $\mathcal{Q}_\alpha$  είναι ίσο με  $\mathcal{D}_{\alpha,\infty}$  για κάθε  $\delta \neq 0$ . Επισημαίνεται ότι  $\delta = 0$  αν και μόνο αν  $\tau = \infty$ . Το Σχήμα 1.3 απεικονίζει τη συμπεριφορά του υπόψη μέτρου, για διάφορες τιμές του  $\alpha \in \{0.1, 0.2, \dots, 0.7\}$  και για σημείο αλλαγής  $1 \leq \tau \leq 50$ , κατά την αξιολόγηση δύο διαφορετικών διαγραμμάτων EWMA (ενός με σταθερά και ενός με χρονικά μεταβαλλόμενα όρια, που θα αναλυθούν στο 3ο Κεφάλαιο).



Σχήμα 1.3: Υπό συνθήκη αναμενόμενη καθυστέρηση  $D_\tau$  και υπό συνθήκη ποσοστιαία καθυστέρηση  $D_{\alpha,\tau}$  για διάγραμμα ( $\lambda=0.1$ ) με σταθερά όρια ( $c = 3$ ) και χρονικά μεταβαλλόμενα όρια ( $c = 3.0055$  ίδιο  $ARL_0$ ) και μία μετατόπιση της τάξης του  $\delta=1$

## 1.6 Άλλοι τύποι διαγραμμάτων ελέγχου

Στον τύπο των διαγραμμάτων ελέγχου που είδαμε στο παρόν κεφάλαιο, των επονομαζόμενων διαγραμμάτων Shewhart, αποφασίζουμε για την κατάσταση μιας διεργασίας με βάση μόνο τις πιο πρόσφατες πληροφορίες, εκτός και αν χρησιμοποιηθούν συμπληρωματικοί κανόνες.

Ωστόσο, έχουν αναπτυχθεί τύποι διαγραμμάτων ελέγχου που βασίζονται, με διαφορετικούς τρόπους, στις προγενέστερες παρατηρήσεις μαζί με τις τρέχουσες. Αυτό το χαρακτηριστικό τα καθιστά πιο αποτελεσματικά στην ανίχνευση μικρών και μέτριων μετατοπίσεων. Τέτοια διαγράμματα, όπως αναφέρθηκε πιο πάνω, είναι τα διαγράμματα CUSUM και EWMA. Στα επόμενα κεφάλαια θα παρουσιαστούν και οι δύο τύποι διαγραμμάτων, αλλά θα γίνει εστίαση στα διαγράμματα ελέγχου EWMA και τις τροποποιήσεις τους.



## Κεφάλαιο 2

# Σωρευτικό Διάγραμμα (CUSUM)

### 2.1 Εισαγωγή

Τα διαγράμματα CUSUM χρησιμοποιούνται για την παρακολούθηση και ανίχνευση μικρών μετατοπίσεων της τάξης του  $0.5\sigma$  ως  $1\sigma$  ενός χαρακτηριστικού ποιότητας  $\mu_0$ . Όπως επισημάνθηκε ήδη, η βασική διαφορά στη σχεδιάσή του σε σχέση με το διάγραμμα Shewhart έγκειται στο γεγονός ότι το τελευταίο λαμβάνει υπόψη κάθε φορά μόνο μία παρατήρηση, ενώ το διάγραμμα CUSUM λαμβάνει υπόψη την ακολουθία των σημείων ή των παρατηρήσεων.

### 2.2 Κεντρική Ιδέα Διαγραμμάτων CUSUM

Υποθέτουμε ότι έχουμε μια κανονική διαδικασία  $N(\mu_0, \sigma^2)$  και μπορούμε να πάρουμε  $N$  δείγματα, το καθένα αποτελούμενο από  $n$  παρατηρήσεις, δηλαδή:

$$X_{i1}, X_{i2}, \dots, X_{in} \text{ (με τιμές } x_{i1}, x_{i2}, \dots, x_{in} \text{ αντίστοιχα), } i = 1, 2, \dots, N.$$

Για λόγους απλούστευσης, θεωρούμε ότι λαμβάνουμε μεμονωμένες παρατηρήσεις ( $n = 1$ , κάθε δείγμα αποτελείται από μία παρατήρηση), δηλαδή έχουμε μια κανονική μεταβλητή  $X_i$  με μέση τιμή  $\mu_0$  και τυπική απόκλιση  $\sigma$ . Αν θεωρήσουμε ως  $C_i$  την διαφορά (απόσταση) κάθε παρατήρησης από μια τιμή στόχο  $\mu_0$ , τότε έχουμε ότι  $C_i = X_i - \mu_0$ .

Αν οι εν λόγω διαφορές είναι μικρές, ένα διάγραμμα Shewhart δεν θα έδινε εκτός ελέγχου σήματα (όπως θα δούμε σε παράδειγμα παρακάτω). Γενικώς, τα διαγράμματα Shewhart δύνανται να ανιχνεύσουν μεταβολές της τάξης του  $3\sigma$ .

Ωστόσο, αν τις μικρές διαφορές τις αθροίσουμε, δηλαδή θεωρήσουμε ως  $C_i$  τα συσσωρευμένα αθροίσματα των διαφορών κάθε παρατήρησης από το  $\mu_0$ , τότε θα προκύψει μια ποσότητα όλο και μεγαλύτερη. Άρα, για  $i = 1, 2, \dots, N$  θα έχουμε:

$$C_1 = X_1 - \mu_0, C_2 = X_1 - \mu_0 + X_2 - \mu_0, \dots$$

που οδηγεί στον τύπο του στατιστικού για τα διαγράμματα CUSUM:

$$C_i = \sum_{j=1}^i (X_j - \mu_0) \quad (2.1)$$

ή

$$C_i = X_i - \mu_0 + C_{i-1}, \quad C_0 = 0 \quad (2.2)$$

Αν η διεργασία είναι εντός ελέγχου, δηλαδή ο μέσος δεν έχει μετατοπιστεί, τότε οι τιμές που θα δίνει το στατιστικό θα είναι σαν ένας τυχαίος περίπατος με μέση τιμή  $\mu_0$  ίση με 0.

Αν όμως μετατοπιστεί ο μέσος, αυξηθεί ή ελαττωθεί, τότε το στατιστικό θα αρχίσει να αυξάνεται ή να μειώνεται αντίστοιχα, οπότε θα εμφανίζεται μία τάση στο διάγραμμα.

Ένα μειονέκτημα των διαγραμμάτων είναι ότι μετά τη όποια διόρθωση, που θα επαναφέρει τον μέσο στην τιμή στόχο, θα χρειαστούν ορισμένες χρονικές περίοδοι για να απεικονίζεται μια εντός ελέγχου διεργασία. Αυτό συμβαίνει διότι για να εξαλειφθεί μια διαφορά  $X_i - \mu_0$  θα απαιτηθούν αρκετές αντίθετου προσήμου διαφορές, ώστε να επανέλθει η καμπύλη του διαγράμματος στο 0.

Άλλο μειονέκτημα των υπόψη διαγραμμάτων είναι ότι δεν μπορεί να ανιχνεύσει τυχαία σφάλματα. Αυτό συμβαίνει γιατί, όταν μια διεργασία παρουσιάζει πολλά τυχαία σφάλματα (μεγάλη διασπορά παρατηρήσεων), εμφανίζει και πολλές τιμές εκατέρωθεν του μέσου, οι οποίες αλληλοεξουδετερώνονται και δεν προκύπτει ένδειξη προβλήματος.

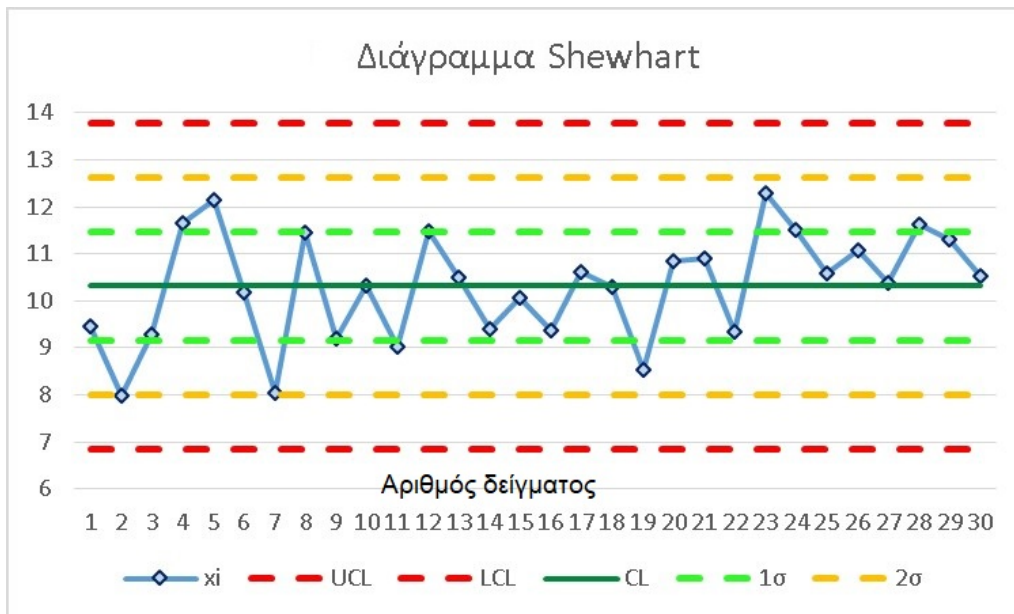
Γενικά, τα διαγράμματα CUSUM είναι αποτελεσματικά στην ανίχνευση συστηματικών σφαλμάτων.

## 2.3 Εφαρμογή με Διάγραμμα CUSUM

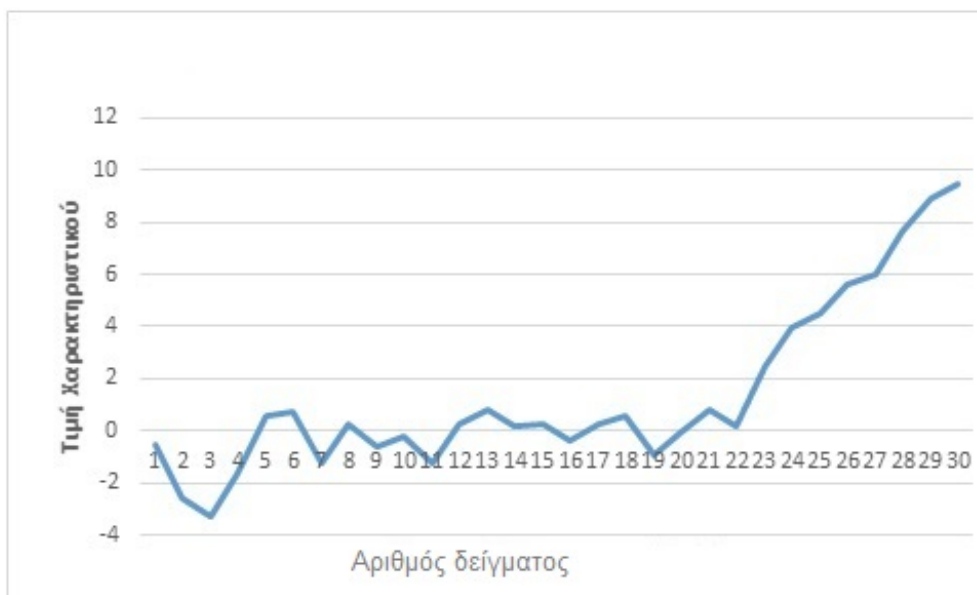
Το παρακάτω παράδειγμα παρουσιάστηκε από τον Montgomery (2013) [37] και δείχνει τη χρησιμότητα του διαγράμματος CUSUM σε σύγκριση με την αδυναμία του διαγράμματος Shewhart στην ανίχνευση μικρών διαδοχικών μετατοπίσεων. Τα δεδομένα (Πίνακας A.1 Παραρτήματος Α) είναι προσομοιωμένα, ώστε οι πρώτες 20 παρατηρήσεις να προέρχονται από κανονική κατανομή  $N(10,1)$ , ενώ στις τελευταίες 10 έχει μετατοπιστεί ο μέσος κατά  $1\sigma$ , δηλαδή τα δεδομένα προέρχονται από κανονική κατανομή  $N(11,1)$ .

Κάνοντας χρήση διαγράμματος Shewhart για να ελέγξουμε αν η διεργασία είναι εντός ελέγχου, θα προκύψει το διάγραμμα 2.1, στο οποίο παρατηρούμε ότι η μετατόπιση του μέσου μετά την 20η παρατήρηση δεν γίνεται αντιληπτή.

Όμως, χρησιμοποιώντας την Σχέση (2.1) για το στατιστικό του διαγράμματος CUSUM, παίρνουμε το γράφημα του Σχήματος 2.2, από το οποίο βλέπουμε ότι υπάρχει μια τάση αύξησης της μέσης τιμής. Ωστόσο, το διάγραμμα είναι ποιοτικό και δείχνει απλά τάση, δεν περιέχει άνω και κάτω όρια ελέγχου.



Σχήμα 2.1: Διάγραμμα Ελέγχου Shewhart- $\bar{X}$  διάγραμμα



Σχήμα 2.2: Καμπύλη στατιστικού CUSUM

## 2.4 Tabular CUSUM

Όρια ελέγχου ανάλογα με αυτά που έχουμε σε ένα διάγραμμα Shewhart εισάγονται από έναν διαφορετικό τύπο διαγραμμάτων CUSUM, τα διαγράμματα Tabular CUSUM. Το Tabular CUSUM ορίζεται από τις ακόλουθες σχέσεις:

$$C_i^+ = \max[0, X_i - (\mu_0 + K) + C_{i-1}^+] \quad (2.3)$$

$$C_i^- = \max[0, (\mu_0 - K) - X_i + C_{i-1}^-] \quad (2.4)$$

$$C_0^+ = C_0^- = 0 \quad (2.5)$$

όπου η σταθερά  $K$  καλείται τιμή αναφοράς (reference value) και δίνεται από τη σχέση:

$$K = k\sigma = \frac{\delta}{2}\sigma \quad (2.6)$$

Στην παραπάνω σχέση το  $k$  είναι μια παράμετρος ίση με  $\delta/2$ , ενώ  $\delta$  είναι το μέγεθος της μετατόπισης που επιθυμούμε να ανιχνεύσουμε, εκφρασμένο σε μονάδες τυπικών αποκλίσεων και δίνεται από τη σχέση:

$$\delta = \frac{|\mu_1 - \mu_0|}{2} \quad (2.7)$$

Η παραπάνω σταθερά  $K$  δεν σχεδιάζεται στο διάγραμμα Tabular CUSUM, απλά συμμετέχει στους υπολογισμούς των στατιστικών  $C_i^+$  και  $C_i^-$ . Όρια ελέγχου εισάγονται μέσω μιας άλλης σταθεράς, του διαστήματος απόφασης (decision interval) που συμβολίζεται με  $H$ . Αυτή η σταθερά καθορίζει το άνω και κάτω όριο και δίνεται από τη σχέση:

$$H = h\sigma \quad (2.8)$$

Συνήθως επιλέγεται  $H = 5\sigma$ , οπότε για  $\sigma = 1$  το κάτω όριο θα είναι το -5 και το άνω όριο το +5.

## 2.5 Βέλτιστη επιλογή των $K$ και $H$

Οι παράμετροι  $K$  και  $H$  επιλέγονται με τέτοιο τρόπο, ώστε να επιτυγχάνεται το επιθυμητό  $ARL_0$ . Στη βιβλιογραφία υπάρχουν πίνακες για τη βέλτιστη επιλογή των παραμέτρων αυτών, ανάλογα με την μετατόπιση που επιθυμούμε να ανιχνεύουμε. Στο παραπάνω παράδειγμα η μετατόπιση ήταν  $\delta = 1$  και  $\sigma = 1$ , οπότε:

$$K = \frac{\delta}{2}\sigma = \frac{1}{2}$$

Έστω  $k = 1/2$  και  $h = 4$  ή  $5$ . Για το  $ARL_0$  υπολογίζεται:

-  $h = 4$  για  $ARL_0=168$

-  $h = 5$  για  $ARL_0=465$

Οπότε, προκειμένου να υπάρχει το ίδιο μέτρο σύγκρισης στη σχεδίαση των διαγραμμάτων ελέγχου (δηλαδή επιθυμητό  $ARL_0=370$ ), υπολογίζεται:

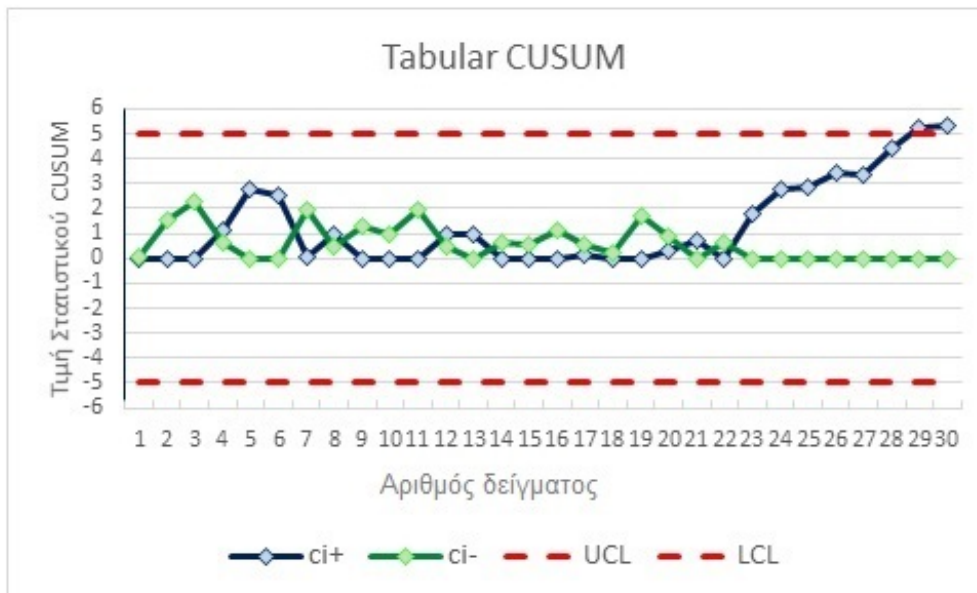
-  $h = 4,77$  για  $ARL_0=370$



Επιλέγοντας  $k = 1/2$  και  $h = 5$  και κάνοντας εφαρμογή του διαγράμματος Tabular CUSUM, προκύπτει ότι:

$$K = k\sigma \Leftrightarrow K = \frac{1}{2}\sigma = \frac{1}{2}$$

Οπότε, για τα δεδομένα του παραπάνω παραδείγματος (Πίνακας Α.1), παίρνουμε το ακόλουθο γράφημα:



Σχήμα 2.3: Διάγραμμα Tabular CUSUM για τα δεδομένα του Πίνακα Α.1

Διακρίνουμε ότι το διάγραμμα δίνει εκτός ελέγχου σήμα, δηλαδή τιμή μεγαλύτερη από 5, στα δείγματα 29 και 30. Άρα, με το διάγραμμα Tabular CUSUM εντοπίστηκε η μετατόπιση του μέσου, κάτι που δεν κατέστη εφικτό με το Shewhart.

## 2.6 Τυποποιημένο Διάγραμμα CUSUM

Τα εν λόγω διαγράμματα προκύπτουν μέσω τυποποίησης της μεταβλητής  $X_i$  πριν τον υπολογισμό των στατιστικών  $C_i^+$  και  $C_i^-$ . Οπότε, για  $1 \leq i \leq m$  θα έχουμε:

$$Y_i = \frac{X_i - \mu_0}{\sigma} \sim N(0, 1) \quad (2.9)$$

$$C_i^+ = \max[0, Y_i - k + C_{i-1}^+] \quad (2.10)$$

$$C_i^- = \max[0, -k - Y_i + C_{i-1}^-] \quad (2.11)$$

$$C_0^+ = C_0^- = 0 \quad (2.12)$$

Για το παράδειγμά μας θα ισχύει  $K = k\sigma = k$ .

Βασικό πλεονέκτημα είναι ότι οι παράμετροι  $K = k$  και  $H = h$  δεν εξαρτώνται από την τυπική απόκλιση  $\sigma$  (στην τυποποιημένη κανονική κατανομή είναι 1), οπότε μπορούμε να έχουμε πολλά διαγράμματα ελέγχου με τις ίδιες τιμές  $K$  και  $H$ .

Άλλο πλεονέκτημα είναι ότι με τη χρήση της μεταβλητής  $Y_i$  οδηγούμαστε ευκολότερα στα διαγράμματα για τον έλεγχο της μεταβλητότητας μιας διεργασίας.

## 2.7 CUSUM Διάγραμμα για Έλεγχο Μεταβλητότητας (Scale CUSUM)

Το διάγραμμα αυτό χρησιμοποιείται για την παρακολούθηση της μεταβλητότητας που ενέχει μια διεργασία. Η κατασκευή του βασίζεται στην ποσότητα [24]:

$$\nu_i = \frac{\sqrt{|Y_i|} - 0.822}{0.349} \sim N(0, 1) \quad (2.13)$$

όπου  $Y_i = \frac{x_i - \mu_0}{\sigma}$

Οι ποσότητες  $\nu_i$  είναι ευαίσθητες περισσότερο σε αλλαγές της διασποράς μιας διεργασίας παρά σε μετατοπίσεις του μέσου. Ακολουθούν την τυποποιημένη κανονική κατανομή  $N(0, 1)$  οπότε, για  $1 \leq i \leq m$ , τα δύο μονόπλευρα Scale CUSUM είναι:

$$S_i^+ = \max[0, \nu_i - k + S_{i-1}^+] \quad (2.14)$$

$$S_i^- = \max[0, -k - \nu_i + S_{i-1}^-] \quad (2.15)$$

με αρχικές τιμές:

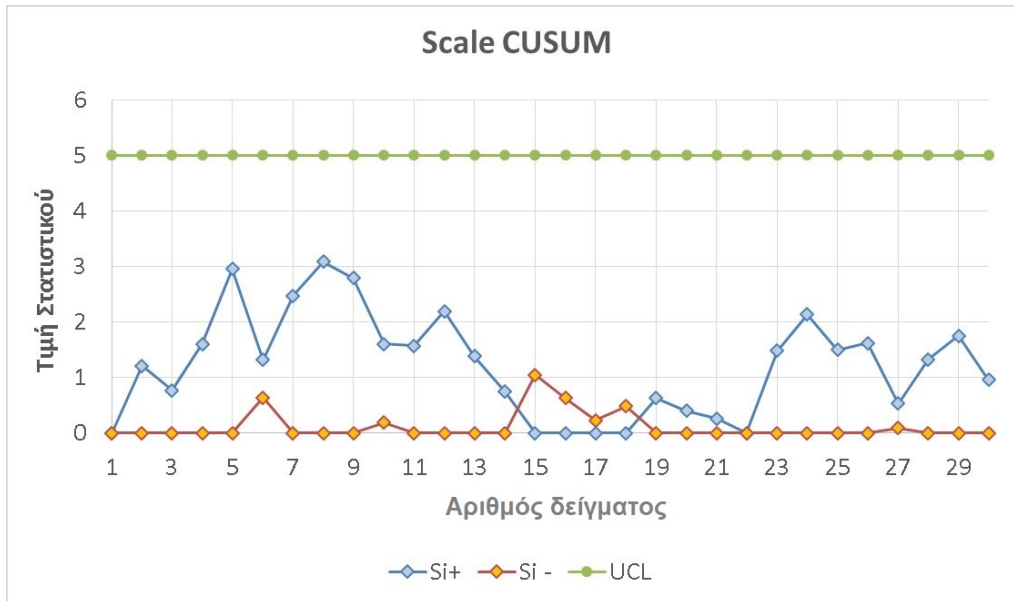
$$S_i^+ = S_i^- = 0 \quad (2.16)$$

Η εφαρμογή των παραπάνω σχέσεων, για τα δεδομένα του παραδείγματός μας (Πίνακας A.1 Παραρτήματος Α), απεικονίζεται στο Σχήμα 2.4. Παρατηρούμε ότι σε όλες τις περιόδους τα αθροίσματα  $S_i^+$  και  $S_i^-$  δεν ξεπερνούν την τιμή του διαστήματος απόφασης  $H = 5$ . Συμπεραίνουμε έτσι ότι η διεργασία είναι εντός ελέγχου όσον αφορά τη μεταβλητότητα.

## 2.8 Διάγραμμα FIR-CUSUM

### 2.8.1 Η ιδέα της Γρήγορης Αρχικής Αντίδρασης

Όταν μια διεργασία είναι εκτός ελέγχου, είναι επιθυμητό να μπορούμε να έχουμε κάποιο σήμα για την εκτός ελέγχου κατάσταση όσο το δυνατόν πιο γρήγορα. Αυτό είναι ιδιαίτερα χρήσιμο όταν διακοπεί μια διεργασία μετά από κάποια ρύθμιση και επανεκκινήθει.



Σχήμα 2.4: Διάγραμμα Scale CUSUM για τα δεδομένα του Πίνακα Α.1

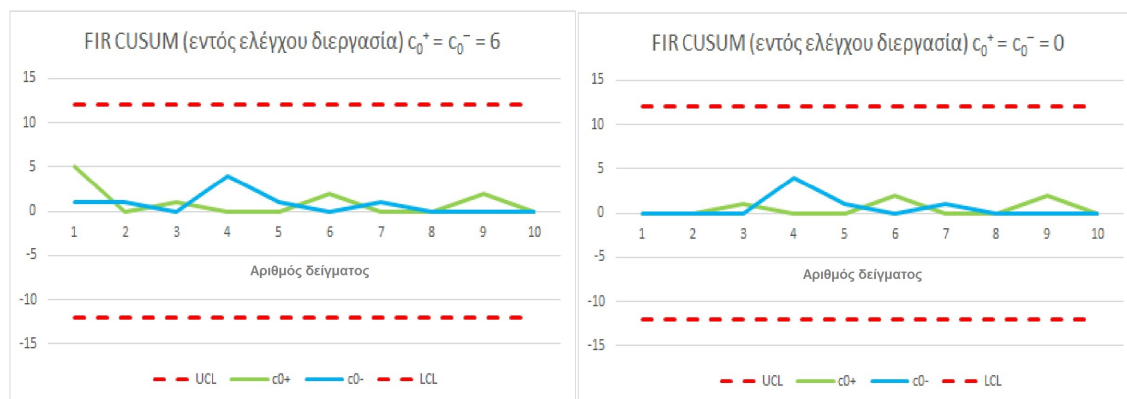
Ένας τρόπος για να επιτευχθεί αυτό είναι με χρήση διαγράμματος CUSUM γρήγορης αρχικής απόκρισης (Fast Initial Response, FIR), το οποίο ορίζεται από τις Σχέσεις (2.10) και (2.11) με τις εξής αρχικές τιμές:

$$C_0^+ = C_0^- = \frac{H}{2} \tag{2.17}$$

Με τον τρόπο αυτό αυξάνεται το  $ARL_0$  και μειώνεται το  $ARL_1$ .

### 2.8.2 Παράδειγμα με εντός ελέγχου διεργασία

Έστω ότι έχουμε τα δεδομένα του Πίνακα Α.2 με  $\mu_0 = 100$ . Επίσης, θεωρούμε  $k = 3$  και  $H = 12$  (το  $H$  αρκετά μεγάλο επειδή τα δεδομένα είναι λίγα). Κατασκευάζουμε, λοιπόν, το διάγραμμα Tabular CUSUM με αρχική τιμή ( $C_0^+ = C_0^- = 6$ ), δηλαδή δημιουργούμε το FIR-CUSUM, όπως φαίνεται στο Σχήμα 2.5:



Σχήμα 2.5: Διάγραμμα CUSUM με και χωρίς FIR για τα δεδομένα του Πίνακα Α.2

Διαπιστώνουμε ότι η διεργασία είναι εντός ελέγχου. Στο ίδιο συμπέρασμα καταλήγουμε και για  $C_0^+ = C_0^- = 0$ , δηλαδή αν πάρουμε το απλό Tabular CUSUM.

### 2.8.3 Παράδειγμα με εκτός ελέγχου διεργασία

Στο παράδειγμα αυτό ο μέσος έχει μετακινηθεί από  $\mu_0 = 100$  στο  $\mu_1 = 105$  και  $\delta = \frac{5}{\sigma}$ . Με την τεχνική FIR ( $C_0^+ = C_0^- = 6$ ), το διάγραμμα δίνει σύντομα εκτός ελέγχου σήμα, από την 3η κιόλας παρατήρηση. Ενώ στη περίπτωση του Tabular CUSUM ( $C_0^+ = C_0^- = 0$ ), βλέπουμε ότι το διάγραμμα δίνει εκτός ελέγχου σήμα από την 6η παρατήρηση. Τα παραπάνω φαίνονται στο Σχήμα 2.6.



Σχήμα 2.6: Διάγραμμα CUSUM με και χωρίς FIR για τα δεδομένα του Πίνακα Α.3

## Κεφάλαιο 3

# Διάγραμμα Ελέγχου με Κινητούς Μέσους με Εκθετικά Βάρη (EWMA)

### 3.1 Το Στατιστικό EWMA

Ένας άλλος τύπος διαγράμματος ελέγχου, με προσεγγιστικά ισοδύναμη συμπεριφορά με το διάγραμμα CUSUM, είναι το διάγραμμα EWMA. Χρησιμοποιείται για εντοπισμό κυρίως μικρών μεταβολών της διεργασίας ( $0.5\sigma$  έως  $2\sigma$ ).

Ένας τρόπος να περιγραφεί η τεχνική EWMA είναι ως η διαδοχική άθροιση της απόστασης κάθε παρατήρησης από μια τιμή στόχο, με την άθροιση αυτή να πραγματοποιείται με διαφορετική βαρύτητα για κάθε παρατήρηση και να είναι συνεχώς μειούμενη για τις παλαιότερες παρατηρήσεις. Αυτό επιτυγχάνεται με τη χρήση μιας σταθεράς εξομάλυνσης.

Ας θεωρήσουμε ότι λαμβάνουμε τυχαία δείγματα  $X_{ij}$ , που είναι οι ανεξάρτητες παρατηρήσεις (SRS). Ο αριθμός των δειγμάτων είναι  $j = 1, 2, \dots, m$  και το κάθε ένα από αυτά μεγέθους  $i = 1, 2, \dots, n$ . Τότε, το στατιστικό EWMA ορίζεται από την παρακάτω σχέση:

$$Z_{(j)} = \lambda \bar{X}_{(j)} + (1 - \lambda)Z_{(j-1)} \quad (3.1)$$

όπου:

$$\bar{X}_{(j)} = \frac{1}{n} \sum_{i=1}^n X_{ij} = \bar{X}_{SRS(j)} \quad (3.2)$$

Οι ανεξάρτητες παρατηρήσεις  $X_{ij}$  θεωρούμε ότι ακολουθούν την κανονική κατανομή  $N(\mu_0, \frac{\sigma^2}{n})$ . Ο μέσος του  $j$ -οστού δείγματος,  $\bar{X}_{(j)}$ , δίνεται από τη Σχέση (3.2).

Το  $\bar{X}_{(j)}$  είναι η τρέχουσα πληροφορία, δηλαδή η μέση τιμή των απλών τυχαίων δειγμάτων (SRS), ενώ το  $Z_{(j-1)}$  είναι η αμέσως προγενέστερη πληροφορία. Το  $\lambda$  λέγεται σταθερά

εξομάλυνσης ή συντελεστής βαρύτητας (weighting factor), καθώς καθορίζει τον βαθμό κατά τον οποίο παλαιότερα δεδομένα συμμετέχουν στον υπολογισμό της τιμής του στατιστικού. Παίρνει τιμές μεταξύ 0 και 1.

Παρατηρούμε ότι η τρέχουσα τιμή του στατιστικού είναι ένας γραμμικός συνδυασμός της προηγούμενης τιμής του στατιστικού και της μετρούμενης πληροφορίας. Όταν το  $\lambda$  είναι μεγάλο δίνει βαρύτητα στις πρόσφατες παρατηρήσεις, ενώ όταν είναι μικρό δίνει βάρος στις παλαιότερες. Συνεπώς, η Σχέση (3.1) είναι ένας αναδρομικός τύπος και μπορεί να αναπτυχθεί αντικαθιστώντας κάθε φορά την τιμή του  $Z$  στο δεύτερο μέλος της εξίσωσης, καταλήγοντας στην ακόλουθη εναλλακτική μορφή του στατιστικού EWMA, τη Σχέση (3.3):

$$Z_{(j)} = \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k \bar{X}_{(j-k)} + (1 - \lambda)^j Z_0 \quad (3.3)$$

Στην παραπάνω σχέση παρατηρούμε ότι, για  $\lambda = 1$ , το διάγραμμα τύπου EWMA γίνεται ισοδύναμο με το διάγραμμα τύπου Shewhart. Επίσης, παρατηρούμε ότι τα βάρη  $\lambda(1 - \lambda)^k$  ελαττώνονται εκθετικά, καθώς οι παρατηρήσεις του δείγματος απομακρύνονται από την τρέχουσα χρονική στιγμή. Για μικρότερες τιμές του  $\lambda$ , το διάγραμμα γίνεται περισσότερο ευαίσθητο σε μικρές μετατοπίσεις.

Η αρχική τιμή για τις προγενέστερες παρατηρήσεις, δηλαδή η  $Z_{(0)}$ , λαμβάνεται ίση με την τιμή στόχο του μέσου  $\mu_0$ . Συνεπώς, η Σχέση (3.3) μπορεί να γραφτεί και ως:

$$Z_{(j)} = \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k \bar{X}_{(j-k)} + (1 - \lambda)^j \mu_0 \quad (3.4)$$

Με βάση την ανωτέρω σχέση μπορούμε να εκτιμήσουμε το μέσο και τη διασπορά του στατιστικού. Οπότε:

$$E[Z_{(j)}] = \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k E[\bar{X}_{(j-k)}] + (1 - \lambda)^j \mu_0 \quad (3.5)$$

και

$$Var[Z_{(j)}] = \lambda^2 \sum_{k=0}^{j-1} (1 - \lambda)^{2k} Var[\bar{X}_{(j-k)}] \quad (3.6)$$

Εάν δεν διαθέτουμε πληροφορίες για την τιμή του μέσου  $\mu_0$ , τότε μπορεί να εκτιμηθεί από το μέσο των προγενέστερων δεδομένων. Υποθέτοντας ότι τα  $X_j$  είναι ανεξάρτητες τυχαίες μεταβλητές με μέσο  $\mu_0$  και διασπορά  $\sigma^2/n$ , όπου  $n$  το μέγεθος δείγματος που χρησιμοποιήθηκε κάθε φορά για τον υπολογισμό των  $X_{(j)}$ , για τη μεν μέση τιμή του στατιστικού EWMA έχουμε:

$$E[Z_{(j)}] = E \left[ \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k \bar{X}_{(j-k)} \right] + E [(1 - \lambda)^j \mu_0] =$$

$$\begin{aligned}
 &= \lambda E \left[ \sum_{k=0}^{j-1} (1-\lambda)^k \bar{X}_{(j-k)} \right] + (1-\lambda)^j \mu_0 = \\
 &= \lambda \sum_{k=0}^{j-1} E \left[ (1-\lambda)^k \bar{X}_{(j-k)} \right] + (1-\lambda)^j \mu_0 = \\
 &= \lambda \sum_{k=0}^{j-1} (1-\lambda)^k E \left[ \bar{X}_{(j-k)} \right] + (1-\lambda)^j \mu_0 = \\
 &= \lambda \sum_{k=0}^{j-1} (1-\lambda)^k \mu_0 + (1-\lambda)^j \mu_0 = \\
 &= \lambda \mu_0 \sum_{k=0}^{j-1} (1-\lambda)^k + (1-\lambda)^j \mu_0 = \\
 &= \lambda \mu_0 \frac{1 - (1-\lambda)^j}{1 - (1-\lambda)} + (1-\lambda)^j \mu_0 = \mu_0
 \end{aligned}$$

οπότε:

$$E[Z_{(j)}] = \mu_0 \quad (3.7)$$

Για τη δε διακύμανση του στατιστικού EWMA έχουμε:

$$\begin{aligned}
 Var[Z_{(j)}] &= \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} Var[\bar{X}_{(j-k)}] = \\
 &= \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} \frac{\sigma^2}{n} = \\
 &= \lambda^2 \frac{\sigma^2}{n} \sum_{k=0}^{j-1} (1-\lambda)^{2k} = \\
 &= \lambda^2 \frac{\sigma^2}{n} (1 + (1-\lambda)^2 + (1-\lambda)^4 + \dots + (1-\lambda)^{2(j-1)}) = \\
 &= \lambda^2 \frac{\sigma^2}{n} \left( \frac{1 - (1-\lambda)^{2j}}{1 - (1-\lambda)^2} \right) = \\
 &= \lambda^2 \frac{\sigma^2}{n} \left( \frac{1 - (1-\lambda)^{2j}}{-\lambda^2 + 2\lambda} \right) = \\
 &= \lambda^2 \frac{\sigma^2}{n} \left( \frac{1 - (1-\lambda)^{2j}}{\lambda(2-\lambda)} \right)
 \end{aligned}$$

οπότε:

$$Var[Z_{(j)}] = \frac{\sigma^2}{n} \frac{\lambda}{2-\lambda} (1 - (1-\lambda)^{2j}) \quad (3.8)$$

όπου  $\sigma^2/n$  η διακύμανση της διαδικασίας, η οποία μπορεί να είναι γνωστή ( $\sigma_0^2/n$ ) ή πρέπει να εκτιμηθεί από τα αρχικά, εντός ελέγχου, δείγματα της διεργασίας.

Τα όρια ελέγχου για ένα διάγραμμα ελέγχου EWMA υπολογίζονται τυπικά με βάση τα όρια  $\pm L \sigma$  όρια, όπου το  $L$  είναι συνήθως ίσο με 3, όπως και στη σχεδίαση των ορίων ελέγχου των διαγραμμάτων Shewhart.

### 3.2 Διάγραμμα EWMA με ασυμπτωτικά όρια

Στο διάγραμμα ελέγχου EWMA, η κεντρική γραμμή και τα όρια δίνονται από τις ακόλουθες σχέσεις:

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2-\lambda)n} (1 - (1-\lambda)^{2j})} \quad (3.9)$$

$$CL = \mu_0 \quad (3.10)$$

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2-\lambda)n} (1 - (1-\lambda)^{2j})} \quad (3.11)$$

Το εύρος των ορίων ελέγχου καθορίζεται από την τιμή της παραμέτρου  $L$ , η οποία εξαρτάται από την επιλογή της σταθεράς εξομάλυνσης  $\lambda$  (γίνεται με τέτοιο τρόπο ώστε να προκύπτει ένα επιθυμητό  $ARL_0$ ). Τα όρια αυτά, όπως περιγράφονται από τις Σχέσεις (3.9) και (3.11), ονομάζονται χρονικά μεταβαλλόμενα όρια των διαγραμμάτων EWMA.

Ωστόσο, για μεγάλες τιμές του  $j$ , αυτά τα όρια συγκλίνουν στα ακόλουθα σταθερά (ασυμπτωτικά) όρια:

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{(2-\lambda)n}} \quad (3.12)$$

$$CL = \mu_0 \quad (3.13)$$

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{(2-\lambda)n}} \quad (3.14)$$

Δηλαδή, ο παράγοντας  $(1 - (1 - \lambda)^{2j})$  στις Σχέσεις (3.9) και (3.11) τείνει στο 1 καθώς προχωρούμε χρονικά (ο αριθμός των δειγμάτων συνεχώς αυξάνεται). Συνεπώς, τα χρονικά μεταβαλλόμενα όρια μετατρέπονται σε σταθερά, με τον ρυθμό σύγκλισης να εξαρτάται από το  $\lambda$ . Επισημαίνεται ότι η σύγκλιση γίνεται πιο αργή για μικρές τιμές του  $\lambda$ .

Μπορούμε εύκολα να κατασκευάσουμε έναν πίνακα (Πίνακας 3.1), όπου για διάφορες τιμές του  $\lambda$ , υπολογίζουμε το αντίστοιχο  $L$ , για δεδομένο επιθυμητό  $ARL_0 = 370$ . Αυτό γίνεται για μια σειρά τιμών μετατοπίσεων που θα θέλαμε να ανιχνεύσουμε. Για τους υπολογισμούς γίνεται χρήση του πακέτου "spc" της R [31], όπου στην εντολή "xewma.arl" που χρησιμοποιείται, το  $q$  συμβολίζει τον αριθμό της παρατήρησης όπου συμβαίνει η μετατόπιση (change point position). Το ζεύγος τιμών παραμέτρων που αντιστοιχεί στο μικρότερο  $ARL_1$  της συγκεκριμένης μετατόπισης που θέλουμε να εντοπίσουμε είναι το βέλτιστο. Οπότε, για



ανίχνευση μετατόπισης π.χ. της τάξης του  $1\sigma$ , οι βέλτιστες τιμές για τα  $\lambda$  και  $L$ , για  $\delta = 1$ , είναι:  $\lambda = 0.14$  και  $L = 2.784641$ .

Πίνακας 3.1: Ασυμπτωτικά Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda$ ,  $L$ ) για  $ARL_0=370$ , για ανίχνευση διάφορων μετατοπίσεων  $\delta$ , όταν συμβαίνουν στην 1η παρατήρηση ( $q = 1$ )

$\lambda$	$L$	$ARL_0$	0.25	0.5	0.75	1	1.25	1.5	1.75
0.01	1.819130	370.00	67.56	30.72	19.78	14.61	11.61	9.65	8.28
0.02	2.134916	370.00	66.71	28.19	17.62	12.82	10.10	8.36	7.14
0.03	2.301887	370.00	68.08	27.07	16.47	11.83	9.25	7.61	6.49
0.04	2.410879	370.00	70.38	26.58	15.77	11.18	8.68	7.12	6.05
0.05	2.489686	370.00	73.15	26.45	15.32	10.73	8.27	6.75	5.72
0.06	2.550249	370.00	76.19	26.55	15.03	10.41	7.97	6.47	5.47
0.07	2.598719	370.00	79.36	26.81	14.85	10.16	7.73	6.25	5.27
0.08	2.638649	370.00	82.62	27.19	14.75	9.98	7.54	6.07	5.10
0.09	2.672265	370.00	85.92	27.67	14.72	9.84	7.38	5.93	4.97
0.10	2.701046	370.00	89.23	28.22	14.73	9.74	7.26	5.80	4.85
0.11	2.726021	370.00	92.54	28.83	14.78	9.66	7.15	5.69	4.75
0.12	2.747933	370.00	95.83	29.50	14.87	9.61	7.07	5.60	4.66
0.13	2.767332	370.00	99.09	30.21	14.99	9.59	7.00	5.53	4.58
0.14	2.784641	370.00	102.33	30.97	15.14	9.58	6.95	5.46	4.52
0.15	2.800184	370.00	105.53	31.76	15.30	9.58	6.91	5.41	4.46
0.16	2.814223	370.00	108.69	32.58	15.49	9.60	6.87	5.36	4.40
0.17	2.826964	370.00	111.82	33.43	15.70	9.63	6.85	5.32	4.36

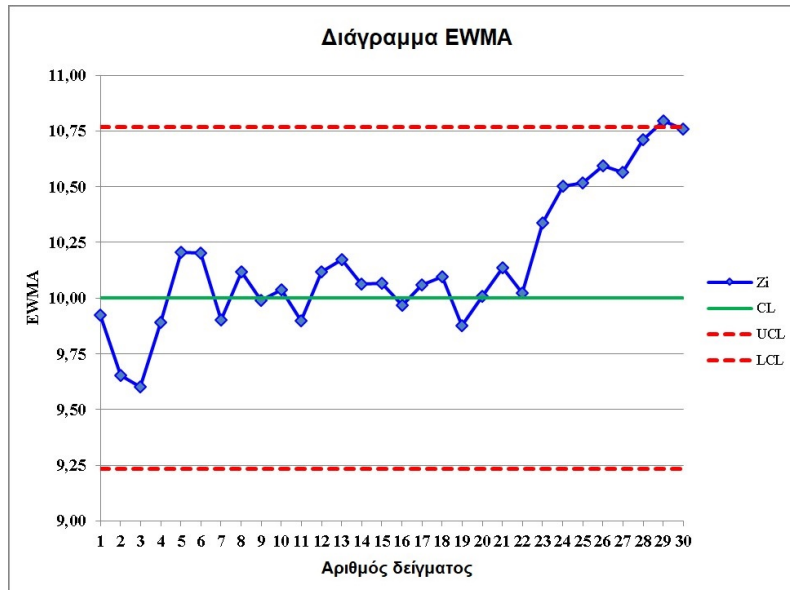
Θα εφαρμόσουμε τις Σχέσεις (3.12), (3.13) και (3.14) (ασυμπτωτικά όρια) στα δεδομένα του Πίνακα A.1, κατασκευάζοντας δηλαδή το διάγραμμα EWMA. Υπενθυμίζουμε ότι τα δεδομένα αυτά προέρχονται από μία διεργασία, η οποία κατά τις πρώτες 20 παρατηρήσεις έχει συγκεκριμένο  $\mu_0$  και κατά τις 10 τελευταίες ο μέσος αυτός έχει μετατοπιστεί κατά 1. Το διάγραμμα βλέπουμε ότι εντοπίζει εκτός ελέγχου διεργασία στην 29η παρατήρηση (Σχήμα 3.2).

Η παραπάνω συμπεριφορά του διαγράμματος συμφωνεί με την αντίστοιχη θεωρητικά αναμενόμενη. Αυτό μπορούμε να το δούμε αναλυτικά ως εξής:

- Υπολογίζουμε το  $ARL_1$  (με χρήση του πακέτου "src" της R [31]) για κάθε παρατήρηση που έχουμε (συνολικά 30):

```
xewma.arl(l=0.14, c=2.784641, mu=1, sided="two", limits="fix", q=20)
```

Οι τιμές που παίρνουμε δίνουν τον θεωρητικό αριθμό δειγμάτων που χρειάζονται, έως να εμφανιστεί εκτός ελέγχου σήμα στο διάγραμμα. Καθώς προστίθενται νέα δεδομένα στον υπολογισμό του στατιστικού, βλέπουμε την πορεία των υπόψη τιμών:



Σχήμα 3.1: Διάγραμμα ελέγχου EWMA με ασυμπτωτικά όρια για τα δεδομένα του Πίνακα Α.1

9.575347	9.520164	9.477735	9.445894	9.422843	9.406595	9.395317	9.387553	9.382232
9.378593	9.376107	9.374410	9.373252	9.372462	9.371923	9.371555	9.371303	9.371132
9.371015	9.370935	9.370881	9.370844	9.370818	9.370801	9.370789	9.370781	9.370776
9.370772	<b>9.370769</b>	9.370768						

- Ξεχωρίζουμε την τιμή του  $ARL_1$  στον δειγματικό αριθμό 29 (29η παρατήρηση), επειδή σ' αυτόν τον αριθμό παρατηρήσαμε εκτός ελέγχου σήμα στο διάγραμμα EWMA με τα πειραματικά δεδομένα του Πίνακα Α.1. Η τιμή αυτή είναι 9.370769. Η θεωρητική ερμηνεία είναι ότι για να δοθεί εκτός ελέγχου σήμα στην 29η παρατήρηση, θα πρέπει η μετατόπιση να έχει ξεκινήσει το αργότερο μετά τη χρονική στιγμή 29 – 9.371 = 19.629, δηλαδή μετά την 20η παρατήρηση.

- Παρατηρούμε ότι στο Σχήμα 3.1 που προκύπτει από τα πειραματικά δεδομένα, στην 20η παρατήρηση εμφανίζεται η πρώτη μετατόπιση προς την κατεύθυνση που τελικά ανιχνεύεται η μετατόπιση.

### 3.3 Διάγραμμα EWMA με χρονικά μεταβαλλόμενα όρια

Αναφέρθηκε ήδη ότι η κεντρική γραμμή και τα όρια για το διάγραμμα EWMA με χρονικά μεταβαλλόμενα όρια, δίνονται από τις Σχέσεις (3.9), (3.10) και (3.11).

Θα εφαρμόσουμε ξανά τις παραπάνω σχέσεις (χρονικά μεταβαλλόμενα όρια) στα δεδομένα του Πίνακα Α.1, ακολουθώντας την ίδια μεθοδολογία με αυτή που χρησιμοποιήθηκε στα ασυμπτωτικά όρια. Έτσι, ομοίως με την περίπτωση των ασυμπτωτικών ορίων, υπολογίζουμε τα ζεύγη τιμών των  $\lambda$  και  $L$  που δίνουν  $ARL_0 = 370$  και επιλέγουμε το βέλτιστο ζεύγος τιμών για τα  $\lambda$  και  $L$  για την ανίχνευση της μετατόπισης που επιθυμούμε. Οπότε, για  $\delta = 1$

βρίσκουμε  $\lambda = 0.14$  και  $L = 2.792883$  (Πίνακας 3.2).

Πίνακας 3.2: Χρονικά Μεταβαλλόμενα Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda, L$ ) για  $ARL_0=370$ , για ανίχνευση διάφορων μετατοπίσεων  $\delta$ , όταν συμβαίνουν στην 20η παρατήρηση ( $q = 19$ )

$\lambda$	$L$	$ARL_0$	0.25	0.5	0.75	1	1.25	1.5	1.75
0.01	2.017113	370	60.82	24.55	14.86	10.60	8.24	6.75	5.74
0.02	2.233469	370	64.35	25.47	15.31	10.88	8.44	6.91	5.86
0.03	2.363788	370	67.02	25.62	15.22	10.76	8.33	6.81	5.77
0.04	2.454496	370	69.80	25.69	15.02	10.55	8.13	6.64	5.62
0.05	2.522615	370	72.76	25.81	14.82	10.32	7.92	6.45	5.45
0.06	2.576279	370	75.87	26.03	14.65	10.11	7.73	6.27	5.29
0.07	2.619982	370	79.08	26.35	14.54	9.93	7.55	6.11	5.14
0.08	2.656451	370	82.35	26.76	14.47	9.78	7.39	5.96	5.01
0.09	2.687456	370	85.64	27.25	14.45	9.66	7.25	5.83	4.89
0.1	2.714208	370	88.95	27.81	14.47	9.57	7.14	5.71	4.78
0.11	2.737568	370	92.25	28.43	14.53	9.50	7.04	5.61	4.68
0.12	2.758167	370	95.54	29.10	14.62	9.45	6.96	5.52	4.60
0.13	2.776483	370	98.80	29.81	14.74	9.42	6.89	5.45	4.52
0.14	2.792883	370	102.03	30.57	14.88	9.41	6.84	5.38	4.46
0.15	2.807657	370	105.23	31.37	15.05	9.42	6.80	5.33	4.40
0.16	2.821035	370	108.39	32.19	15.24	9.44	6.76	5.28	4.34
0.17	2.833205	370	111.52	33.05	15.45	9.47	6.74	5.24	4.30

Στη συνέχεια, παίρνουμε το διάγραμμα EWMA του Σχήματος 3.2, το οποίο εμφανίζει παραπλήσια απόδοση με εκείνη του αντίστοιχου διαγράμματος ασυμπτωτικών ορίων (εντοπίζει την μετατόπιση στην 29η παρατήρηση).

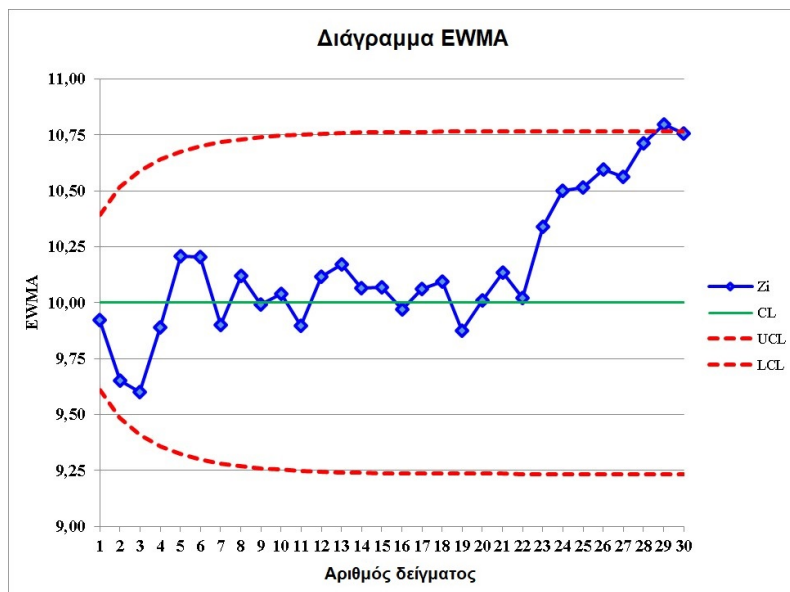
Με ανάλογο τρόπο της περίπτωσης των ασυμπτωτικών ορίων, μπορούμε να δούμε ότι υπάρχει συμφωνία στη συμπεριφορά του διαγράμματος με την αντίστοιχη θεωρητικά αναμενόμενη:

- Υπολογίζουμε το εκτός ελέγχου μέσο μήκος ροής  $ARL_1$  (με χρήση λογισμικού R) για κάθε παρατήρηση που έχουμε (συνολικά 30):

```
for (i in 1:30) {
print(xewma.arl(l=0.14, c=2.782883, mu=1, sided="two", limits="vacl", q=i,
r=40)) }
```

Οι τιμές που παίρνουμε δίνουν τον θεωρητικό αριθμό δειγμάτων που χρειάζονται, έως να δοθεί εκτός ελέγχου σήμα στο διάγραμμα. Καθώς προστίθενται νέα δεδομένα στον υπολογισμό του στατιστικού, βλέπουμε την πορεία των τιμών αυτών:

```
8.042184 8.532194 8.796796 8.964453 9.077401 9.155997 9.211767 9.251842 9.280885
9.302056 9.317554 9.328931 9.337302 9.34347 9.348021 9.35138 9.353861 9.355695
9.357051 9.358053 9.358794 9.359342 9.359748 9.360048 9.36027 9.360434 9.360555
9.360645 9.360711 9.36076
```



Σχήμα 3.2: Διάγραμμα ελέγχου EWMA με χρονικά μεταβαλλόμενα όρια για τα δεδομένα του Πίνακα Α.1

- Ξεχωρίζουμε την τιμή του  $ARL_1$  στον δειγματικό αριθμό 29 (29η παρατήρηση), επειδή σ' αυτόν τον αριθμό παρατηρήσαμε εκτός ελέγχου σήμα στο διάγραμμα EWMA με τα πειραματικά δεδομένα του πίνακα Α.1. Η τιμή αυτή είναι 9.360711. Η θεωρητική ερμηνεία είναι ότι για να δοθεί εκτός ελέγχου σήμα στην 29η παρατήρηση, θα πρέπει η μετατόπιση να έχει ξεκινήσει το αργότερο μετά τη χρονική στιγμή  $29 - 9.361 = 19.639$ , δηλαδή μετά την 20η παρατήρηση.

- Παρατηρούμε ότι στο διάγραμμα EWMA που προκύπτει από τα πειραματικά δεδομένα (Σχήμα 3.2), στην 20η παρατήρηση εμφανίζεται η πρώτη μετατόπιση προς την κατεύθυνση που τελικά ανιχνεύεται η μετατόπιση.

Σε πρώτη ανάγνωση, η απόδοση των δύο διαγραμμάτων (ασυμπτωτικών και χρονικά μεταβαλλόμενων ορίων) φαίνεται να είναι ίδια, καθώς τα υπολογιζόμενα  $ARL_1$  είναι παραπλήσια. Στο διάγραμμα με χρονικά μεταβαλλόμενα όρια, όμως, η μετατόπιση του μέσου λαμβάνει χώρα αρκετά μετά την έναρξη των μετρήσεων. Στον Πίνακα 3.3, παρουσιάζονται τα αποτελέσματα των υπολογισμών για τον καθορισμό του βέλτιστου ζεύγους  $\lambda$  και  $L$ , μόνο που η μετατόπιση συμβαίνει από την πρώτη κιόλας παρατήρηση.

Παρατηρώντας τον Πίνακα 3.3, διαπιστώνουμε ότι για μικρές τιμές του  $\lambda$  (και αντίστοιχα μικρές τιμές του  $L$ ), το  $ARL_1$  γίνεται μικρότερο. Μπορούμε να υπολογίσουμε και για ακόμα μικρότερες τιμές των  $\lambda$  και  $L$ . Συγκεκριμένα, με χρήση του πακέτου "spc" της R [31], για  $\lambda = 0.00008$ :

```
xewma.crit(l=0.00008, L0=370, mu0=0, sided="two", limits="vacl", r=270)
```

υπολογίζεται  $L = 1.27926$ , και το αντίστοιχο  $ARL_1$  για μετατόπιση ίση με 1:

```
xewma.arl(l=0.00008, c=1.27926, mu=1, sided="two", limits="vacl", r=40, q=1)
```

Πίνακας 3.3: Χρονικά Μεταβαλλόμενα Όρια - Επιλογή βέλτιστου συνδυασμού ( $\lambda$ ,  $L$ ) για  $ARL_0=370$ , για ανίχνευση διαφορών μετατοπίσεων  $\delta$ , όταν συμβαίνουν στην 1η παρατήρηση ( $q = 1$ )

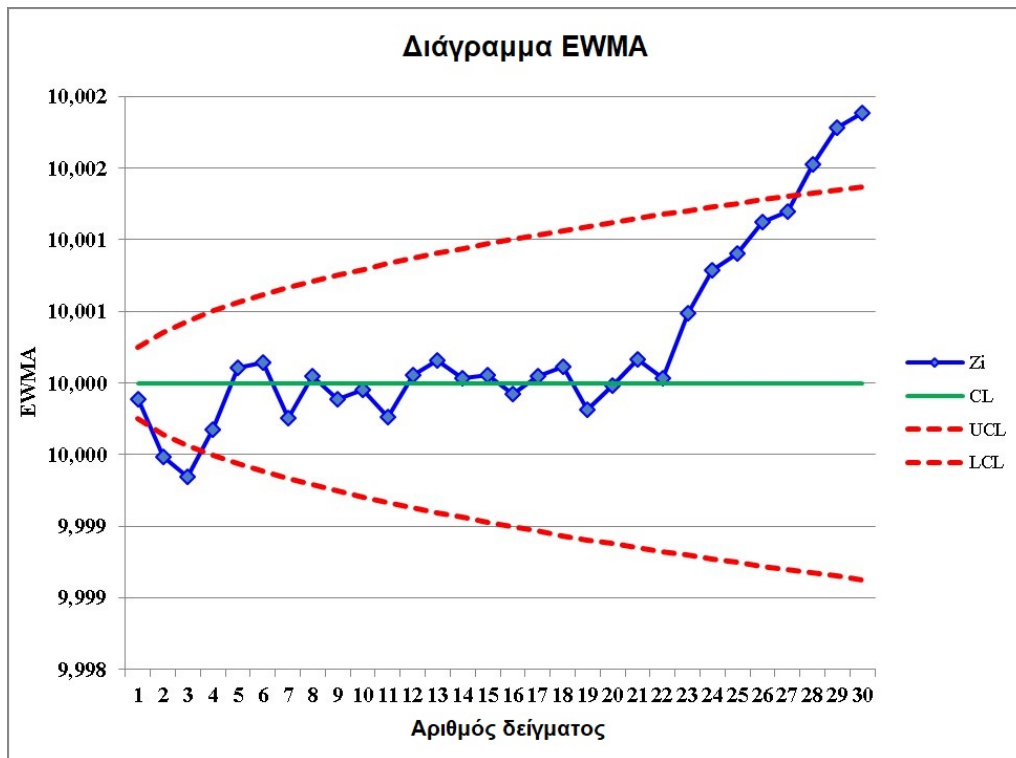
$\lambda$	$L$	$ARL_0$	0.25	0.5	0.75	1	1.25	1.5	1.75
0.01	2.017113	370	44.03	14.70	7.60	4.81	3.41	2.61	2.11
0.02	2.233469	370	52.43	17.47	8.96	5.60	3.93	2.98	2.38
0.03	2.363788	370	58.01	19.16	9.81	6.11	4.27	3.21	2.55
0.04	2.454496	370	62.67	20.40	10.41	6.48	4.51	3.39	2.68
0.05	2.522615	370	66.95	21.42	10.88	6.76	4.70	3.52	2.78
0.06	2.576279	370	71.01	22.33	11.27	6.98	4.85	3.63	2.86
0.07	2.619982	370	74.95	23.19	11.60	7.17	4.97	3.72	2.93
0.08	2.656451	370	78.78	24.03	11.91	7.33	5.08	3.79	2.98
0.09	2.687456	370	82.54	24.86	12.19	7.48	5.17	3.86	3.03
0.10	2.714208	370	86.22	25.70	12.47	7.62	5.26	3.92	3.08
0.11	2.737568	370	89.83	26.55	12.74	7.74	5.33	3.97	3.11
0.12	2.758167	370	93.37	27.42	13.02	7.87	5.40	4.01	3.15
0.13	2.776483	370	96.85	28.30	13.29	7.98	5.47	4.06	3.18
0.14	2.792883	370	100.27	29.20	13.57	8.10	5.53	4.10	3.21
0.15	2.807657	370	103.64	30.11	13.86	8.22	5.59	4.14	3.24
0.16	2.821035	370	106.94	31.05	14.15	8.33	5.65	4.17	3.26
0.17	2.833205	370	110.20	31.99	14.45	8.45	5.71	4.21	3.29

υπολογίζεται  $ARL_1 = 2.659032$ .

Αυτό σημαίνει πως ένα διάγραμμα ελέγχου EWMA που σχεδιάζεται με τα υπόψη  $\lambda$  και  $L$ , θα εμφανίζει καλύτερη ανιχνευτική ικανότητα. Χρησιμοποιώντας δεδομένα του Πίνακα Α.1, κατασκευάζουμε ένα EWMA διάγραμμα ελέγχου (Σχήμα 3.3) με τις παραπάνω τιμές παραμέτρων και βλέπουμε την καλύτερη ανιχνευτική ικανότητα, καθώς το διάγραμμα δίνει εκτός ελέγχου σήμα νωρίτερα.

Ωστόσο, αυτή η υπεροχή των διαγραμμάτων με χρονικά μεταβαλλόμενα όρια έγκειται μόνον στην περίπτωση που η μετατόπιση εντόπιζεται κοντά στην έναρξη μέτρησης. Όσο αυξάνεται το πλήθος των δειγματοληψιών η ανιχνευτική ικανότητα μειώνεται και τείνει να γίνει ίση με εκείνη των διαγραμμάτων με ασυμπτωτικά όρια για τις τιμές των παραμέτρων που επιλέξαμε. Είδαμε ότι για την περίπτωση των ασυμπτωτικών ορίων με μετατόπιση από την 1η παρατήρηση (Πίνακας 3.1), τα  $\lambda = 0.14$  και  $L = 2.784641$  είναι το βέλτιστο ζεύγος παραμέτρων για ανίχνευση μετατόπιση ίσης με 1, ενώ στην περίπτωση των χρονικά μεταβαλλόμενων ορίων με μετατόπισης από την 20η παρατήρηση (Πίνακας 3.2), το βέλτιστο ζεύγος είναι  $\lambda = 0.14$  και  $L = 2.792883$ . Γενικώς, συστήνεται να επιλέγονται ως βέλτιστες τιμές, ακόμη και για την περίπτωση των χρονικά μεταβαλλόμενων ορίων, εκείνες των ασυμπτωτικών ορίων, θεωρώντας ότι η μετατόπιση συμβαίνει στην έναρξη μέτρησης ( $q = 1$ ).

Μια περίπτωση όπου μπορούμε να δούμε την υπεροχή των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα όρια είναι το παρακάτω παράδειγμα, με τα δεδομένα του Πίνακα Α.4, που αναπαριστούν μία πρωτοεμφανιζόμενη εκτός ελέγχου κατάσταση. Οι τιμές του στατιστικού



Σχήμα 3.3: Διάγραμμα ελέγχου EWMA με χρονικά μεταβαλλόμενα όρια για τα δεδομένα του Πίνακα Α.1 ( $\lambda=0.00008$ ,  $L = 1.27926$ )

εξάγονται από τη Σχέση (3.1), με χρονικά μεταβαλλόμενα όρια ελέγχου από τις Σχέσεις (3.9), (3.10) και (3.11).

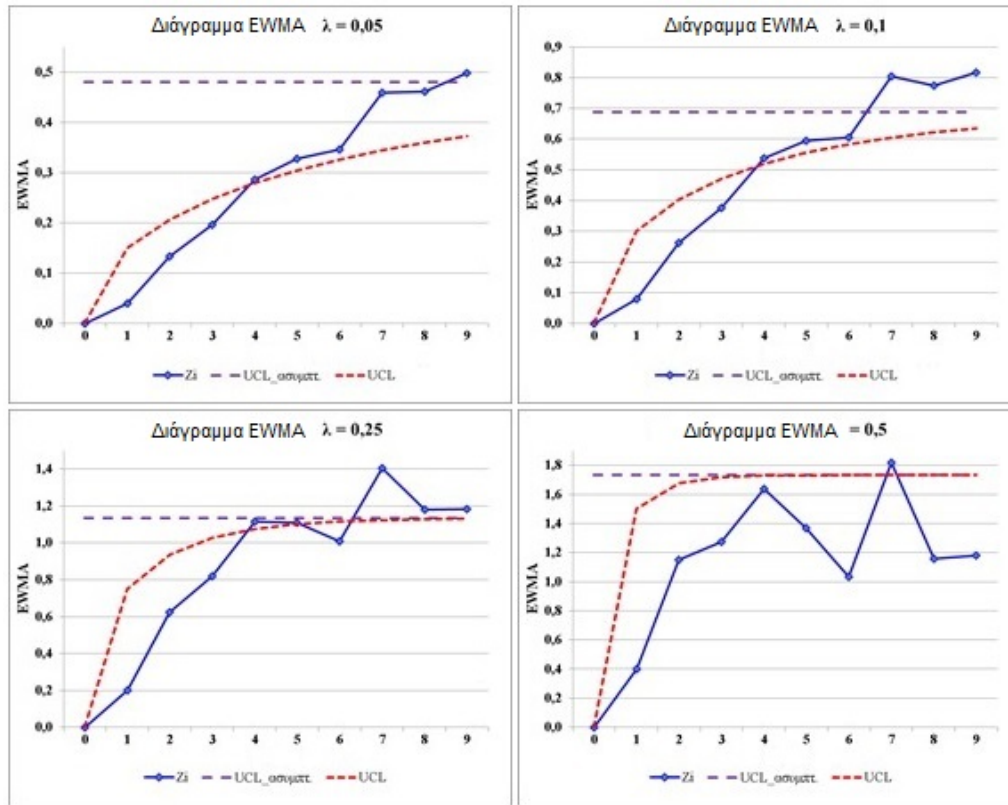
Στο Σχήμα 3.4 απεικονίζονται διαγράμματα EWMA για διάφορες τιμές του  $\lambda$ , με χρονικά μεταβαλλόμενο όριο  $UCL(j)$  καθώς και με ασυμπτωτικό όριο UCL. Παρατηρούμε ότι το πόσο γρήγορα θα δώσει εκτός ελέγχου σήμα κάθε διάγραμμα εξαρτάται τόσο από την τιμή του  $\lambda$ , όσο και από τη χρήση ή όχι των χρονικά μεταβαλλόμενων ορίων. Θα μπορούσε να ειπωθεί ότι η χρήση ασυμπτωτικών ορίων αντί χρονικά μεταβαλλόμενων κάνει τα διαγράμματα λιγότερο ευαίσθητα σε μετατοπίσεις της διεργασίας κατά τις πρώτες παρατηρήσεις.

### 3.4 Υπολογισμός ARL για διαγράμματα EWMA

Έστω οι τυχαίες μεταβλητές  $T_1, T_2, \dots, T_k, \dots$ , των οποίων η συνάρτηση πυκνότητας πιθανότητας (Probability Density Function, PDF) συμβολίζεται ως  $f_T$  και η αθροιστική συνάρτηση πυκνότητας πιθανότητας (Cumulative Density Function, CDF) ως  $F_T$ . Θεωρούμε τα άνω και κάτω όρια UCL και LCL ως δύο σταθερές που ικανοποιούν τη σχέση:

$$LCL < \mu_T < UCL \tag{3.15}$$

Λόγω του ότι οι τυχαίες μεταβλητές  $T_1, T_2, \dots, T_k, \dots$  θεωρούνται ανεξάρτητες, το σφάλμα



Σχήμα 3.4: Διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια για διάφορες τιμές της σταθεράς εξομάλυνσης  $\lambda$

τύπου II για την ακολουθία  $T_1, T_2, \dots, T_k, \dots$  θα δίνεται από τη σχέση:

$$\beta = P[\text{Type II Error}] = \Phi \left[ \frac{UCL - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] - \Phi \left[ \frac{LCL - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] \quad (3.16)$$

ή

$$\beta = \Phi \left[ \frac{UCL - \mu_1}{\frac{\sigma}{\sqrt{n}}} \right] - \Phi \left[ \frac{LCL - \mu_1}{\frac{\sigma}{\sqrt{n}}} \right] \quad (3.17)$$

Συνεπώς το  $ARL_1$  προκύπτει από την σχέση:

$$ARL_1 = \frac{1}{1 - \beta} = \frac{1}{1 - \Phi \left[ \frac{UCL - \mu_1}{\frac{\sigma}{\sqrt{n}}} \right] + \Phi \left[ \frac{LCL - \mu_1}{\frac{\sigma}{\sqrt{n}}} \right]} \quad (3.18)$$

$$ARL_1 = \frac{1}{1 - F_T(UCL) + F_T(LCL)} \quad (3.19)$$

Έστω ότι το  $ARL_Y(y)$  είναι το μέσο μήκος ροής για μια ακολουθία EWMA  $Y_1, Y_2, \dots, Y_k, \dots$ , υποθέτοντας ότι  $Y_0 = y$ , και έστω ότι:

$$ARL_Y = ARL_Y(\mu_T) \quad (3.20)$$

Προφανώς, καθώς κάθε τιμή του στατιστικού EWMA εξαρτάται από τις προγενέστερες

τιμές του, οι τυχαίες μεταβλητές  $Y_1, Y_2, \dots, Y_k, \dots$  δεν μπορούν να θεωρηθούν ανεξάρτητες. Συνεπώς, δεν μπορεί να γίνει χρήση των Σχέσεων (3.18) ή (3.19) για τον υπολογισμό του  $ARL_Y(y)$ . Στη συνέχεια θα παρατεθούν οι κάποιες προσεγγίσεις για τον υπολογισμό του  $ARL_1$ .

### 3.4.1 Μέθοδος Επίλυσης Συστήματος Εξισώσεων

Η προσέγγιση αυτή στηρίζεται στο γεγονός ότι το  $ARL_Y(y)$  πρέπει να ικανοποιεί την ακόλουθη εξίσωση:

$$ARL_Y(y) = 1 + \int_{LCL}^{UCL} \frac{1}{\lambda} f_T \left( \frac{z - (1 - \lambda)y}{\lambda} \right) ARL_Y(z) dz \quad (3.21)$$

Πρόκειται για μια εξίσωση Fredholm δεύτερου είδους:

$$f(y) = g(y) + \int_{z_1}^{z_n} h(y, z) f(z) dz$$

όπου  $h(y, z)$ ,  $g(y)$  είναι γνωστές συναρτήσεις και  $f(z)$  μια άγνωστη συνάρτηση που ικανοποιεί την υπόψη εξίσωση. Στην περίπτωση μας έχουμε:

$$g(y)=1, h(y, z) = f_T((z - (1 - \lambda)y)/\lambda) \text{ και } f(y) = ARL_Y(y)$$

Η αριθμητική εκτίμηση της παραπάνω εξίσωσης Fredholm συνίσταται στην προσέγγιση του τελεστή ολοκλήρωσης με ένα σταθμισμένο άθροισμα:

$$f(y) \approx g(y) + \sum_{i=1}^n w_i h(y, z_i) f(z_i) \quad (3.22)$$

όπου  $z_1, z_2, \dots, z_n$  και  $w_1, w_2, \dots, w_n$  είναι, αντίστοιχα, οι τετμημένες και τα βάρη της μεθόδου τετραγώνων στο  $[z_1, z_n]$ , έτσι ώστε, για παράδειγμα  $z_n$ , η  $n$ -οστή τετμημένη στην οποία αντιστοιχεί το  $w_n$ ,  $n$ -οστό τετραγωνικό σημείο της μεθόδου Gauss-Legendre.

Εάν εφαρμόσουμε την Σχέση (3.22) για τα  $y = z_1, z_2, \dots, z_n$  έχουμε:

$$\begin{cases} f_1 \approx g_1 + w_1 h_{1,1} f_1 + w_2 h_{1,2} f_2 + \dots + w_n h_{1,n} f_n \\ f_2 \approx g_2 + w_1 h_{2,1} f_1 + w_2 h_{2,2} f_2 + \dots + w_n h_{2,n} f_n \\ \dots \\ f_n \approx g_n + w_1 h_{n,1} f_1 + w_2 h_{n,2} f_2 + \dots + w_n h_{n,n} f_n \end{cases} \quad (3.23)$$

όπου  $f_i = f(z_i)$ ,  $g_i = g(z_i)$  και  $h_{i,j} = h(z_i, z_j)$ .



Αυτό το σύστημα εξισώσεων μπορεί να γραφτεί με μορφή πινάκων:

$$\begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix} \cong \begin{pmatrix} g_1 \\ g_2 \\ \dots \\ g_n \end{pmatrix} + \begin{pmatrix} w_1 h_{1,1} f_1 + w_2 h_{1,2} f_2 + \dots + w_n h_{1,n} f_n \\ w_1 h_{2,1} f_1 + w_2 h_{2,2} f_2 + \dots + w_n h_{2,n} f_n \\ \dots \\ w_1 h_{n,1} f_1 + w_2 h_{n,2} f_2 + \dots + w_n h_{n,n} f_n \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix} \quad (3.24)$$

ή πιο συνοπτικά:

$$f \approx g + Hf \quad (3.25)$$

ή

$$ARL_1 = f \approx (I - H)^{-1}g \quad (3.26)$$

### 3.4.2 Μέθοδος Extrapolation

Κατά την προσέγγιση αυτή, θεωρούμε ως  $P(\cdot)$  την πιθανότητα της ακολουθίας των τιμών του στατιστικού EWMA για κατάσταση σταθερών παραμέτρων  $\mu, \sigma, z_0^\mu$ . Για κάθε  $j \in N_0$ , η πιθανότητα να μεταβληθεί η παραπάνω κατάσταση είναι  $P(L > j)$ . Συνεπώς, το μέσο μήκος ροής θα δίνεται από την παρακάτω σχέση:

$$E(L) = \sum_{j=0}^{\infty} P(L > j) \quad (3.27)$$

Σύμφωνα με τους Waldmann (1986) [51] και Gan (1995) [19], χρησιμοποιώντας την Σχέση (3.27) μπορούμε να ορίσουμε τα παρακάτω όρια για το ARL:

$$ARL_J^\pm = \sum_{j=0}^{J-1} P(L > j) + \frac{P(L > J)}{1 - m_J^\pm} \quad (3.28)$$

Οι ποσότητες  $m_J^\pm$ , πάνω στις επιτρεπτές τιμές των τιμών εκκίνησης  $z_0^\mu$ , δίνονται από τις σχέσεις:

$$m_J^+ = \sup \left\{ \frac{P(L > j; z_0^\mu)}{P(L > j-1; z_0^\mu)} \right\} \quad (3.29)$$

$$m_J^- = \inf \left\{ \frac{P(L > j; z_0^\mu)}{P(L > j-1; z_0^\mu)} \right\} \quad (3.30)$$

και συγκλίνουν γρήγορα σε ένα κοινό όριο. Ο υπολογισμός των τιμών του  $P(L > j)$ ,  $j = 1, 2, \dots, J$  γίνεται αριθμητικά με χρήση κανόνων τετραγωνισμού, με τον ακόλουθο επαναληπτικό κανόνα για τα απλά διαγράμματα ελέγχου:

$$P(L_\mu > 1)(z_0^\mu) = \int_{-c^*}^{c^*} f(z_0, z) dz \quad (3.31)$$

$$P(L_\mu > j)(z_0) = \int_{-c^*}^{c^*} P(L > j-1)(z) f(z_0, z) dz \quad (3.32)$$

όπου  $f(\cdot)$  είναι ίσο με:

$$f(z_0, z) = \frac{1}{\lambda} \Phi \left( \frac{z - (1 - \lambda)z_0}{\lambda} \right) \quad (3.33)$$

με  $\Phi(\cdot)$  την PDF της κανονικά κατανομημένης μεταβλητής για μέσο  $\mu$  και διασπορά  $\sigma^2$ .

Τα ολοκληρώματα στον παραπάνω επαναληπτικό κανόνα υπολογίζονται προσεγγιστικά, με αριθμητική μέθοδο. Και σ' αυτή την περίπτωση, γίνεται χρήση της μεθόδου Gauss-Legendre. Έτσι, στο διάστημα ολοκλήρωσης  $[-c, c]$  θα έχουμε  $z_s$  ρίζες ( $z_s$  είναι ο  $s$ -ιστός κόμβος τετραγωνισμού) και  $w_s$  βάρη,  $s = 1, 2, \dots, N$ , έτσι ώστε:

$$P(L > j)(z_r) = \sum_{s=1}^N w_s P(L > j - 1)(z_s) f_\mu(z_r, z_s), \quad r = 1, 2, \dots, N \quad (3.34)$$

Οπότε, από την Σχέση (3.28) υπολογίζονται τα όρια  $ARL_J^+$  και  $ARL_J^-$ . Τελικά, το μέσο μήκος ροής δίνεται από την σχέση:

$$ARL = \frac{ARL_J^+ + ARL_J^-}{2}, \text{ δηλαδή:}$$

$$ARL = \sum_{j=0}^{J-1} P(L > j) + \frac{1}{2\lambda} \sum_{s=1}^N w_s P(L > j - 1)(z_s) \Phi \left( \frac{z_s - (1 - \lambda)z_r}{\lambda} \right) \frac{2 - m_J^+ - m_J^-}{(1 - m_J^+)(1 - m_J^-)} \quad (3.35)$$

Το  $J$  του ανωτέρω τύπου επιλέγεται έτσι ώστε τα  $m_J^+$  και  $m_J^-$  να διαφέρουν κατά λιγότερο από  $10^{-7}$ .

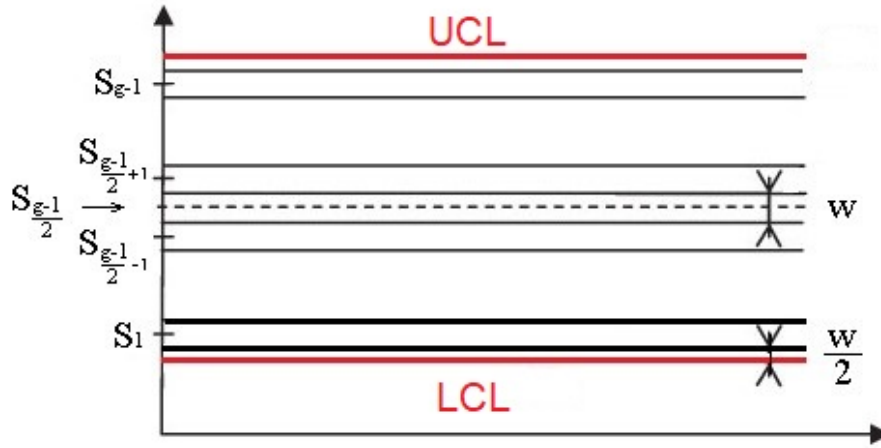
### 3.4.3 Μέθοδος Μαρκοβιανής αλυσίδας

Κατά την προσέγγιση αυτή, η αναμενόμενη τιμή του μήκους ροής για την κατανομή των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα όρια γίνεται με τη μέθοδο της μαρκοβιανής αλυσίδας. Η διαδικασία της λύσης χρησιμοποιεί μη ομογενή μαρκοβιανή αλυσίδα με  $g$  διακριτές καταστάσεις, όπου ο χώρος καταστάσεων μεταξύ των ορίων ελέγχου LCL και UCL χωρίζεται σε  $g - 1$  ξεχωριστές διακριτές καταστάσεις, και η εκτός ελέγχου κατάσταση αντιστοιχεί στην  $g$ -οστή κατάσταση. Ο χώρος εφικτών καταστάσεων προσεγγίζεται μέσω καθορισμού ενός αριθμού διακριτών καταστάσεων και της πιθανότητας μετάβασης από μια οποιαδήποτε κατάσταση σε μια άλλη, για κάθε θεωρούμενη χρονική περίοδο. Χρησιμοποιώντας μεγάλο πλήθος διακριτών καταστάσεων, η προσέγγιση των ιδιοτήτων του μήκους ροής γίνεται περισσότερο ακριβής.

Οι διαφορετικές καταστάσεις ορίζονται ακολούθως:

$$s = (s_1, s_2, \dots, s_{g-1}) = (LCL + w, LCL + 2w, \dots, UCL - 2w, UCL - w) \quad (3.36)$$

όπου  $w = (UCL - LCL)/g$  και UCL και LCL είναι τα ασυμπτωτικά όρια που δίνονται από τις Σχέσεις (3.12) και (3.14). Καθώς το  $g$  αυξάνεται, η προσέγγιση βελτιώνεται.



Σχήμα 3.5: Το διάστημα μεταξύ των LCL και UCL χωρίζεται σε  $g - 1$  καταστάσεις πλάτους  $w$ , ξεκινώντας από το κάτω όριο και προσθέτοντας κάθε φορά  $w$

Τον πίνακα μεταβατικής πιθανότητας για χρονική περίοδο  $t$  τον ορίζουμε με την παρακάτω σχέση:

$$P_t = \begin{bmatrix} tP_{11} & tP_{12} & \dots & tP_{1g} \\ tP_{21} & tP_{22} & \dots & tP_{2g} \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ \cdot & \dots & \cdot & \cdot \\ tP_{g1} & tP_{g2} & \dots & tP_{gg} \end{bmatrix} = \begin{bmatrix} R_t & (I - R_t)\mathbf{1} \\ 0, \dots, 0 & \mathbf{1} \end{bmatrix} \quad (3.37)$$

όπου  $I$  είναι πίνακας  $g \times g$ ,  $\mathbf{1}$  είναι πίνακας στήλη ( $g \times 1$ ) με όλα τα στοιχεία μονάδα και  $tP_{ij}$  είναι η πιθανότητα μετάβασης από την κατάσταση  $s_i$  στην κατάσταση  $s_j$  για περίοδο χρόνου  $t$ . Η τελευταία γραμμή και στήλη αντιστοιχούν στην απορροφητική κατάσταση που αντιπροσωπεύει ένα εκτός ελέγχου σήμα.

Ο πίνακας  $R_t$ , που χρησιμοποιείται για την εξαγωγή των ιδιοτήτων του μήκους ροής των διαγραμμάτων EWMA, ισούται με τον πίνακα της πιθανότητας μετάβασης με γραμμή και στήλη που αντιστοιχεί στην απορροφητική κατάσταση που διαγράφεται.

Λόγω του ότι τα χρονικά μεταβαλλόμενα όρια των Σχέσεων (3.9) και (3.11) προσεγγίζουν τις σταθερές τιμές των Σχέσεων (3.12) και (3.14), οι καταστάσεις πιθανότητας  $tp_{ij}$  συγκλίνουν στις πιθανότητες  $\infty p_{ij}$  και ο πίνακας  $R_t$  συγκλίνει στον πίνακα μετάβασης σε άπειρο χρόνο  $R_\infty$ , καθώς  $t \rightarrow \infty$ . Οι τιμές για τα  $\infty p_{ij}$  μπορούν να καθοριστούν κάνοντας μερικές υποθέσεις για τη διαδικασία. Υποθέτοντας ένα κανονικό μοντέλο με  $X_i \sim N(\mu_x, \sigma_x^2)$  και δεδομένης της τρέχουσας τιμής EWMA, η κατανομή της μελλοντικής τιμής του  $z_{t+1}$  είναι  $N(\lambda\mu_x + (1 - \lambda)z_t, \lambda^2\sigma_x^2)$ . Έτσι, οι πιθανότητες μετάβασης σε άπειρο χρόνο είναι:

$$\infty p_{ij} = Pr\left(s_j - \frac{w}{2} < z < s_j + \frac{w}{2}\right), \quad j = 1, 2, \dots, g - 1 \quad (3.38)$$

$${}_∞p_{ig} = Pr(z > s_{g-1} + \frac{w}{2}) + Pr(z < s_1 - \frac{w}{2}) \quad (3.39)$$

όπου  $z \sim N(\lambda\mu + (1-\lambda)s_i, \lambda^2\sigma^2)$ . Οι τιμές αυτές μπορούν εύκολα να υπολογιστούν για τον καθορισμό των  $P_\infty$  και  $R_\infty$ .

Οι χρονικά εξαρτημένοι πίνακες μετάβασης  $R_t$  μπορούν να καθοριστούν από το  $R_\infty$  αλλάζοντας τις πιθανότητες μετάβασης, οι οποίες οδηγούν σε ένα πρώιμο σήμα. Οι πιθανότητες μετάβασης από αρχικές τιμές (γραμμές) που βρίσκονται εκτός των χρονικά μεταβαλλόμενων ορίων, προς τις τελικές τιμές (στήλες) που οδηγούν σε εκτός ελέγχου σήματα, θέτονται ίσες με μηδέν. Για κάθε τιμή του  $t$ , οι κατάλληλες γραμμές και στήλες ταυτοποιούνται, συγκρίνοντας τα χρονικά μεταβαλλόμενα όρια με τις καταστάσεις από το χώρο καταστάσεων.

Για να καθορίσουμε, δηλαδή, το  $R_t$ , οι πρώτες  $f_1(t)$  και τελευταίες  $f_2(t)$  γραμμές και στήλες του  $R_\infty$  θέτονται σε μηδενικά διανύσματα, όπου  $f_1(t)$  ισούται με τον μεγαλύτερο ακέραιο για τον οποίο ισχύει:

$$s_{f_1} - \frac{w}{2} \leq LCL(t) \quad (3.40)$$

και  $f_2(t)$  είναι ο μικρότερος ακέραιος για τον οποίο ισχύει:

$$s_{f_2} + \frac{w}{2} \geq UCL(t) \quad (3.41)$$

Επίσης, οποιαδήποτε πιθανότητα μετάβασης που επηρεάζεται πάρα πολύ από τη μεταβολή του ορίου ελέγχου τίθεται ίση με μηδέν, σε μια προσπάθεια να αποδοθεί με συνέπεια τιμή στο μήκος ροής μικρότερη από την πραγματική τιμή.

Μία κατάσταση  $s_i$  επηρεάζεται εάν το χρονικά μεταβαλλόμενο όριο είναι είτε πιο κοντά στο μηδέν από ότι το  $s_i$ , είτε εντός  $\frac{w}{2}$  του  $s_i$ . Χρησιμοποιώντας αυτή τη διαδικασία μπορούμε να εκτιμήσουμε τους πίνακες  $R_1, R_2$  κλπ.

Καθορίζοντας το αναμενόμενο μήκος ροής και τη διακύμανση του μήκους ροής, μπορούμε να προχωρήσουμε στη χρήση των πινάκων  $R_t$ . Θέτοντας με RL το μήκος ροής του EWMA, δοθείσας της αρχικής κατάστασης  $s_i$ , έχουμε:

$$Pr(RL \leq t) = \left( I - \prod_{i=1}^t R_i \right) \mathbf{1} \quad (3.42)$$

και

$$Pr(RL = t) = \left( \prod_{i=1}^{t-1} R_i - \prod_{i=1}^t R_i \right) \mathbf{1}, \quad t \geq 1 \quad (3.43)$$

Έτσι, η μέση τιμή του RL και η διασπορά του δίνονται, αντίστοιχα, από τις παρακάτω σχέσεις):

$$E(RL) = \sum_{t=1}^{\infty} t Pr(RL = t) = \sum_{t=1}^{\infty} \left( \prod_{s=1}^t R_s \mathbf{1} \right) \quad (3.44)$$

και

$$Var(RL) = I + \sum_{t=1}^{\infty} \left[ (2t+1) \left( \prod_{s=1}^t R_s \mathbf{1} \right) \right] \quad (3.45)$$

Οι παραπάνω εκφράσεις είναι  $g \times 1$  διανύσματα, τα οποία δίνουν το μέσο μήκος ροής και τη διασπορά από οποιαδήποτε αρχική τιμή ή κατάσταση  $s_i$ . Οι τιμές που αντιστοιχούν σε έναρξη με  $z_0 = \bar{X}$  υπολογίζονται εύκολα, υποθέτοντας ότι τα όρια ελέγχου είναι συμμετρικά γύρω από το  $X$ , οπότε η αντίστοιχη κατάσταση θα είναι  $s_{g/2}$ .

Οι Σχέσεις (3.46) και (3.47) δίνουν τις ροπές του μήκους ροής σε μορφή άπειρων αθροισμάτων τα οποία συγκλίνουν για μεγάλες τιμές του  $t$ . Αυτές οι εκφράσεις μπορούν να απλοποιηθούν σε αυτή την περίπτωση, καθώς τα όρια ελέγχου συγκλίνουν ασυμπτωτικά, και έτσι οι πίνακες πιθανότητας μετάβασης  $R_t$  επίσης συγκλίνουν στο  $R_{\infty}$ , καθώς το  $t$  αυξάνεται. Αντικαθιστώντας όλους τους  $R_t$  πίνακες για μεγάλες τιμές του  $t$  με  $R_{\infty}$ , τα άπειρα αθροίσματα των Σχέσεων (3.44) και (3.45) μπορούν να γραφτούν ως ακολούθως:

$$E(RL) = \left( \sum_{t=1}^{t_{max}-1} \left( \prod_{s=1}^t R_s \right) \right) \mathbf{1} + \prod_{s=1}^{t_{max}} R_s (I - R_{\infty})^{-1} \mathbf{1} \quad (3.46)$$

και

$$Var(RL) = 1 + \sum_{t=1}^{t_{max}-1} [(2t+1) \left( \prod_{s=1}^t R_s \mathbf{1} \right)] + (2t_{max}+1) \left( \prod_{s=1}^{t_{max}} R_s \right) (1 - R_{\infty})^{-1} \mathbf{1} + 2 \left( \prod_{s=1}^{t_{max}} R_s \right) R_{\infty} (1 - R_{\infty})^{-2} \mathbf{1} \quad (3.47)$$

όπου το  $t_{max}$  ισούται με τον αριθμό των χρονικών περιόδων, για τις οποίες χρησιμοποιούνται διαφορετικοί πίνακες πιθανότητας μετάβασης. Για τους υπολογισμούς, το  $t_{max}$  επιλέχθηκε με βάση τα  $\lambda$  και  $g$ , έτσι ώστε ο πίνακας  $R_{t_{max}}$  να είναι δύσκολα διαχωρίσιμος από τον  $R_{\infty}$ . Με αυτό τον τρόπο αυξάνοντας το  $t_{max}$  περαιτέρω, δεν θα έχουμε καμία επίδραση στην ακρίβεια της λύσης. Η ελάχιστη τιμή για το  $t_{max}$  εξάγεται με βάση το γεγονός ότι εάν τα χρονικά μεταβαλλόμενα όρια στο χρόνο  $t_{max}$  διαφέρουν από τα ασυμπτωτικά όρια κατά λιγότερο από  $\frac{w}{2}$ , τότε ο πίνακας  $R_{t_{max}}$  είναι ο ίδιος με τον  $R_{\infty}$ . Λύνοντας τις ανισώσεις:

$$UCL - UCL(t) \leq \frac{w}{2} \quad (3.48)$$

και

$$LCL - LCL(t) \leq \frac{w}{2} \quad (3.49)$$

για την ελάχιστη τιμή του  $t$ , λαμβάνουμε το  $t_{max}$ , τη μικρότερη ακέραια τιμή που είναι μεγαλύτερη από:

$$t_{max} \geq \log \left( \frac{\left( \frac{12nw(2-\lambda)\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}-w}}{36\lambda\sigma^2} \right)}{2\log(1-\lambda)} \right) \quad (3.50)$$

Για υπολογιστική επάρκεια και ακρίβεια, τα  $E(RL)$  και  $Var(RL)$  προτείνεται να υπολογίζονται με χρήση γκαουσιανής απλοποίησης, και όχι χρησιμοποιώντας τον αντίστροφο πίνακα (σημειώνεται ότι στον κώδικα του Παραρτήματος Β οι υπολογισμοί γίνονται με χρήση αντίστροφου πίνακα, όπως αποτυπώνεται στις Σχέσεις (3.46) και (3.47)).

Γενικά, καθώς ο αριθμός διακριτών καταστάσεων  $g$  αυξάνεται, οι τιμές των  $E(RL)$  και  $Var(RL)$  αυξάνονται και προσεγγίζουν περισσότερο τις πραγματικές τιμές. Οι τιμές αυξάνονται, διότι η διαδικασία συνεχώς υποεκτιμά το πραγματικό μήκος ροής. Το μήκος ροής υποεκτιμάται για δύο λόγους:

- Τα απορροφούντα όρια για το  $R_\infty$  είναι στενότερα από ότι τα όρια ελέγχου, καθώς θέτονται σε  $LCL + \frac{w}{2}$  και  $UCL - \frac{w}{2}$ .

- Για το  $R_i$ , οι απορροφούσες πιθανότητες υπολογίζονται συντηρητικά, καθώς όλες οι καταστάσεις, ακόμη και οι περιθώριες που επηρεάζονται από τα όρια ελέγχου, υποθέτουμε ότι οδηγούν σε απορρόφηση.

Τον ρυθμό αύξησης αυτόν μπορούμε να χρησιμοποιήσουμε για να εκτιμήσουμε τις πραγματικές τιμές. Μάλιστα, σύμφωνα με τον Steiner (1999) [49], η χρήση της μεθοδολογίας αυτής δίνει αποτελέσματα που διαφέρουν από την πραγματική τιμή λιγότερο από 1%, εκτός από περιπτώσεις μεγάλων μετατοπίσεων στη διεργασία (όπου το ARL είναι κοντά στη μονάδα).

### 3.4.4 Σύγκριση Μεθόδων – Συμπεράσματα

Η σύγκριση του 2ου τρόπου υπολογισμού του ARL (Extrapolation) με τον 3ο τρόπο υπολογισμού (Markov Chain) απεικονίζεται στους Πίνακες 3.4, 3.5, 3.6 και 3.7, όπου και σημειώνεται η απόκλιση τους. Λαμβάνοντας υπόψη ότι με τον 2ο τρόπο υπολογίζεται η πραγματική τιμή με μία απόκλιση της τάξης του  $10^{-7}$ , αυτή εκλαμβάνεται προσεγγιστικά ίση με την πραγματική τιμή. Συνεπώς, η απόκλιση ισούται κατά προσέγγιση με το τυπικό σφάλμα της μεθόδου Μαρκοβιανής Αλυσίδας.

Πίνακας 3.4: Σύγκριση Μεθόδων Υπολογισμού ARL

	3ος	2ος		3ος	2ος		3ος	2ος	
( $\lambda$ )	0.2	0.2		0.3	0.3		0.2	0.2	
( $L$ )	3	3		3	3		2.7	2.7	
$\delta$			[1]			[1]			[1]
0.00	553.195	554.488	1.29	460.193	462.572	2.38	231.386	232.627	1.24
0.1	407.733	409.017	1.28	371.379	373.568	2.19	182.813	184.041	1.23
0.25	159.377	160.618	1.24	175.283	177.079	1.80	84.424	85.614	1.19
0.5	41.534	42.712	1.18	50.706	52.196	1.49	26.674	27.815	1.14
1	8.751	9.857	1.11	9.793	11.090	1.30	6.662	7.734	1.07
2	1.938	2.917	0.98	1.829	2.967	1.14	1.528	2.497	0.97

[1] Απόκλιση

Από την μελέτη των παραπάνω πινάκων διαπιστώνουμε ότι το σφάλμα της μεθόδου Μαρ-

Πίνακας 3.5: Σύγκριση Μεθόδων Υπολογισμού ARL

	3ος	2ος		3ος	2ος		3ος	2ος	
(λ)	0.4	0.4		0.5	0.5		0.9	0.9	
(L)	3	3		3	3		3	3	
δ			[1]			[1]			[1]
0.00	417.100	419.302	2.20	395.647	396.256	0.61	369.864	370.903	1.04
0.1	354.348	356.436	2.09	347.783	348.439	0.66	348.835	349.867	1.03
0.25	190.687	192.500	1.81	206.935	207.708	0.77	266.317	267.327	1.01
0.5	61.325	62.877	1.55	73.930	74.821	0.89	135.530	136.514	0.98
1	11.579	12.921	1.34	14.469	15.417	0.95	34.306	35.281	0.98
2	1.931	3.065	1.13	2.3096	3.225	0.92	4.170	5.153	0.98

[1] Απόκλιση

Πίνακας 3.6: Σύγκριση Μεθόδων Υπολογισμού ARL

	3ος	2ος		3ος	2ος		3ος	2ος	
(λ)	0.35	0.35		0.34	0.34		0.19	0.19	
(L)	3	3		3	3		3	3	
δ			[1]			[1]			[1]
0.00	436.570	437.366	0.80	440.317	441.739	1.42	569.818	569.140	-0.68
0.1	362.641	363.474	0.83	363.870	365.215	1.34	414.515	414.296	-0.22
0.25	183.937	184.846	0.91	182.120	183.304	1.18	158.381	158.839	0.46
0.5	56.423	57.385	0.96	55.252	56.323	1.07	41.061	41.834	0.77
1	10.961	11.928	0.97	10.733	11.748	1.02	8.913	9.766	0.85
2	2.071	3.010	0.94	1.997	3.000	1.00	2.058	2.914	0.86

[1] Απόκλιση

κοβιανής Αλυσίδας είναι μικρότερο από 1% για μικρές μετατοπίσεις και αυξάνεται στο 30% για μετατοπίσεις της τάξης των 2 μονάδων τυπικής απόκλισης. Διαπιστώνουμε, επίσης, ότι για μικρές τιμές των  $\lambda$  και  $L$  η απόκλιση ελαττώνεται, ενώ για μεγάλες τιμές των παραμέτρων αυτών αυξάνεται.

Επιπρόσθετα, όταν οι μετατοπίσεις είναι μεγάλες, οι τιμές στον πίνακα πιθανοτήτων μετάβασης  $R_t$  γίνονται μικρότερες και οι υπολογισμοί που απαιτούνται για να εξαχθεί το  $E(RL)$  γίνονται πιο επιρρεπείς προς σφάλμα στρογγυλοποίησης. Σαν αποτέλεσμα, η εκτίμηση του  $E(RL)$  ενδέχεται να μην αυξάνεται με την αύξηση του  $g$ . Εάν συμβεί αυτό, χρησιμοποιείται το μεγαλύτερο  $E(RL)$  που έχει εξαχθεί ως εκτίμηση του πραγματικού  $E(RL)_{g=\infty}$ , η οποία ίσως παρουσιάσει σφάλμα της τάξης του 10%. Ένα αντίστοιχο πρόβλημα αναφέρεται από τους Lucas και Crosier (1982) [34].

Από τους Πίνακες 3.4, 3.5, 3.6 και 3.7 παρατηρούμε ότι σε κάθε περίπτωση, με εξαίρεση τις περιπτώσεις για  $\lambda \leq 0.2$ , η τελική απόκλιση είναι περίπου ίση με τη μονάδα.

Το Σχήμα 3.6 αφορά τη συνδυαστική περίπτωση EWMA  $\bar{X} - S^2$  (παρακολούθηση του μέσου και της διασποράς ταυτόχρονα, για κανονικά κατανομημένα δεδομένα) και παρουσιάζεται

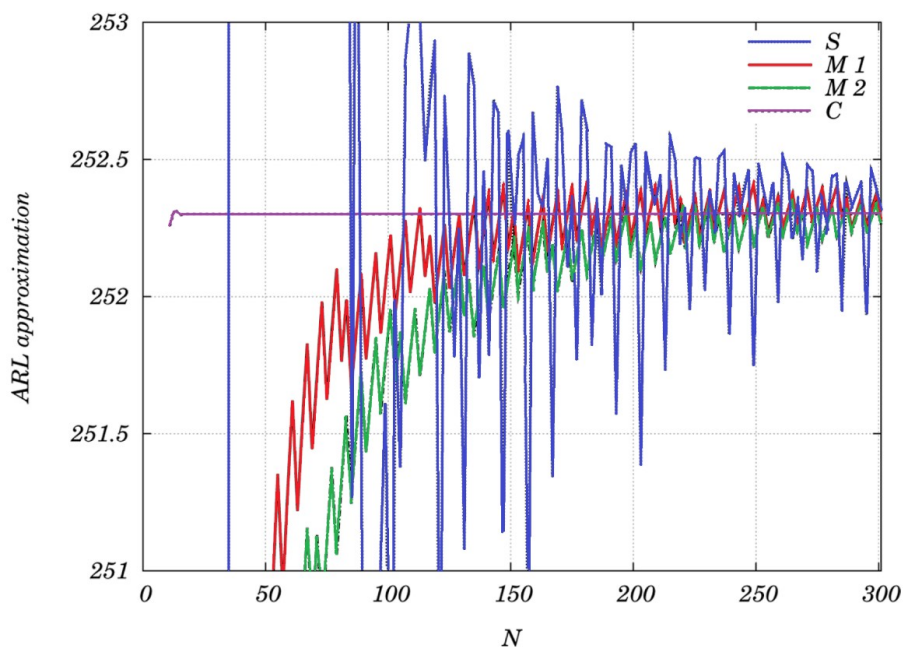
Πίνακας 3.7: Σύγκριση Μεθόδων Υπολογισμού ARL

	Τροπ. <i>Steiner</i> $g=40,$ 80,120	Μέθοδ. 3ος $g=40,$ 100,150	Μέθοδ. 2ος		Τροπ. <i>Steiner</i> $g=100,$ 200,300	Μέθοδ. 3ος $g=50,$ 100,150	Μέθοδ. 2ος	
( $\lambda$ )	0.15	0.15	0.15		0.2	0.2	0.2	
( $L$ )	3	3	3		3	3	3	
$\delta$				[1] / [2]				[1] / [2]
0.00	647.635	647.481	647.04	-0.60/-0.44	554.55	553.2	554.49	1.29/-0.07
0.1	440.866	440.446	440.462	-0.40/0.02	409.08	407.73	409.01	1.28/-0.07
0.25	151.502	150.690	151.282	-0.22/0.59	160.68	159.38	160.62	1.24/-0.06
0.5	38.607	37.637	38.474	-0.13/0.84	42.76	41.53	42.71	1.18/-0.05
1	9.479	8.496	9.468	-0.01/0.97	9.89	8.75	9.86	1.11/-0.03
2	2.788	1.957	2.906	0.12/0.95	2.90	1.94	2.92	0.98/0.02

[1] Απόκλιση Τροποποιημένου *Steiner*

[2] Απόκλιση *Steiner* από 2ο

όπως δημοσιεύτηκε από τον Knoth (2007) [29]. Παρατηρούμε ότι το ARL παρουσιάζει μεγάλη διασπορά για μικρές αλλαγές στη διάσταση του  $N$ . Η πραγματική τιμή του ARL είναι 252.3, με τιμές παραμέτρων  $\lambda_\mu=0.134$  και  $\lambda_\sigma=0.1$ , όπου S=Simpson rule, M1=Αλυσίδα Markov (με τον κανόνα του μέσου σημείου), M2 = Αλυσίδα Markov (Με επίλυση εξισώσεων Sylvester) και C= Συγκριτική Παράθεση.



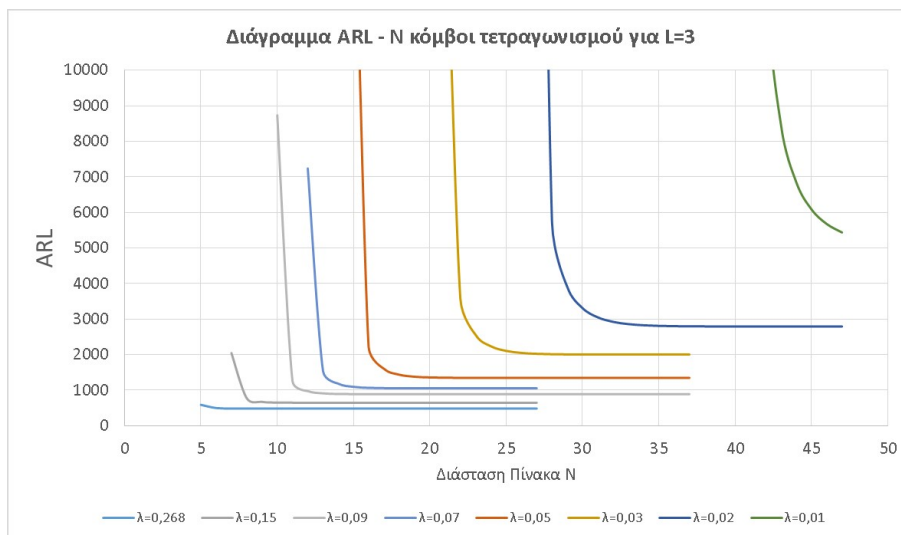
Σχήμα 3.6: Διάγραμμα Ακρίβειας υπολογισμού του ARL με τις διάφορες μεθόδους για την εντός ελέγχου περίπτωση του διαγράμματος  $\bar{X} - S^2$  EWMA

Από το ανωτέρω σχήμα διαπιστώνουμε ότι για μικρές τιμές της διάστασης του πίνακα τετραγωνισμού  $N$ , το ARL παρουσιάζει μεγάλη διασπορά. Κατά τον Knoth (2007) [29], η μέθοδος Μαρκοβιανής Αλυσίδας ενδείκνυται όταν θέλουμε γρήγορα και αρκετά ακριβή απο-



τελέσματα.

Στη συνέχεια εξετάστηκε η επίδραση του πλήθους τετραγωνισμού σε σχέση με την ακρίβεια της λύσης για διαφορετικές τιμές της σταθεράς εξομάλυνσης  $\lambda$ . Από το Σχήμα 3.7 συμπεραίνουμε ότι όσο το  $\lambda$  γίνεται μικρότερο, τόσο η διασπορά του ARL γίνεται μεγαλύτερη και απαιτείται μεγαλύτερο  $N$  προκειμένου να εξασφαλιστεί σωστή προσέγγιση.



Σχήμα 3.7: Ακρίβεια υπολογισμού ARL (Extrapolation) σε σχέση με τους κόμβους τετραγωνισμού για διάφορες τιμές της σταθεράς εξομάλυνσης  $\lambda$ . Τα ανωτέρω αποτελέσματα υπολογίστηκαν για  $L = 3$

### 3.5 Ιδιότητες ARL διαγραμμάτων EWMA

Το διάγραμμα EWMA, όπως και το διάγραμμα Shewhart, ενώ στηρίζεται στην υπόθεση της κανονικότητας των δεδομένων, εντούτοις δεν είναι πολύ ευαίσθητο σ' αυτή. Παρουσιάζεται, δηλαδή, ανθεκτικό και για ποιοτικά χαρακτηριστικά που δεν είναι κανονικά κατανομημένα [37]. Οι Lucas και Saccucci (1990) [35] παρουσίασαν μία μέθοδο για τη μελέτη της κατανομής του μήκους ροής των διαγραμμάτων EWMA, η οποία περιελάμβανε τη μελέτη του ARL για απλά διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια, για διαφορετικούς συνδυασμούς των  $L$  και  $\lambda$ , τα οποία δίνονται στον Πίνακα 3.8 (οι τιμές μετατόπισης  $\delta$  σε  $\sigma/\sqrt{n}$  μονάδες,  $ARL_0 = 500$ ).

Παρατηρώντας τον Πίνακα 3.8 διαπιστώνουμε ότι για μετατόπιση  $\delta > 1.5$  είναι προτιμότερο να επιλέξουμε  $\lambda = 0.25$  αντί για  $\lambda = 0.1$ . Αντίστοιχα για  $\delta \leq 1$  είναι προτιμότερο να επιλέξουμε  $\lambda = 0.1$ , συγκρινόμενο με τις υπόλοιπες τιμές  $\lambda = 0.25$  ή  $\lambda = 0.5$  ή  $\lambda = 0.75$ . Οι υπόψη τιμές ARL επιβεβαιώθηκαν (με ελάχιστη απόκλιση) με τον κώδικα του Παραρτήματος Β, που ακολουθεί τη μεθοδολογία της Μαρκοβιανής Αλυσίδας.

Στη συνέχεια, θα συγκριθούν οι ιδιότητες του μήκους ροής των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα όρια με τα διαγράμματα EWMA με ασυμπτωτικά όρια. Όπως θα

Πίνακας 3.8: Τιμές ARL για το απλό EWMA με  $ARL_0=500$

	$\lambda=0.1$	$\lambda=0.25$	$\lambda=0.5$	$\lambda=0.75$
$\delta$	$L=2.814$	$L=2.998$	$L=3.071$	$L=3.087$
0	499.58	499.836	499.906	500
0.25	106.491	170	255	321
0.5	31.3	48.2	88.8	140
0.75	15.9	20.1	35.9	62.5
1	10.3	11.1	17.5	30.6
1.5	6.09	5.46	6.53	9.90
2	4.36	3.61	3.63	4.54

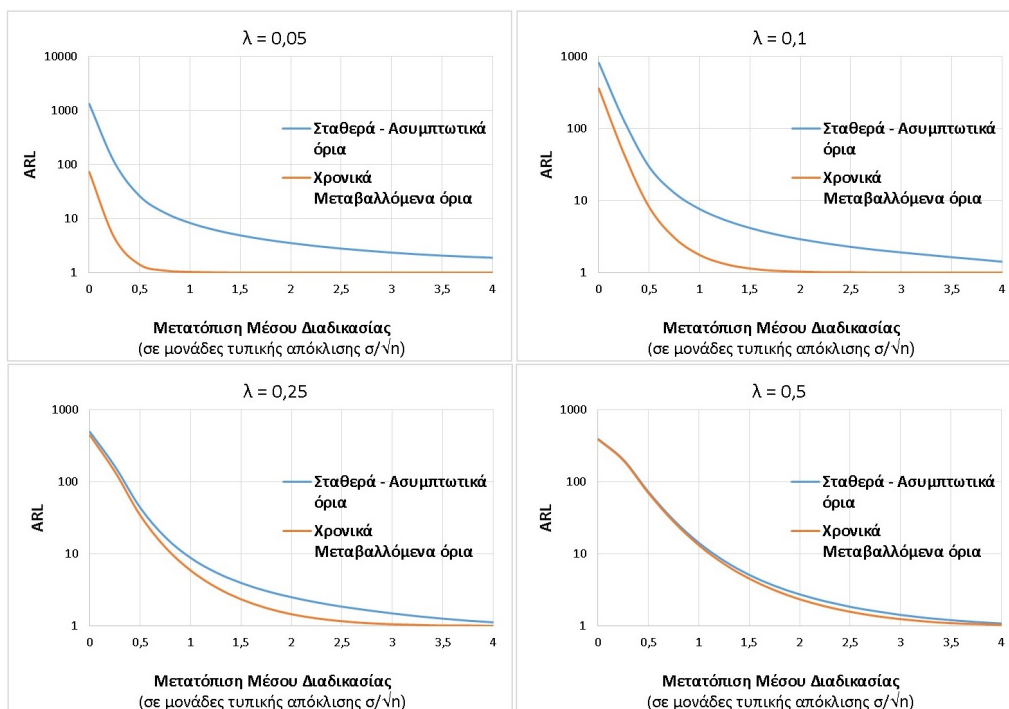
φανεί, όσο η διαδικασία βρίσκεται εντός ελέγχου, τα ARL των διαγραμμάτων με χρονικά μεταβαλλόμενα όρια είναι σχεδόν ίδια με τα ARL των διαγραμμάτων EWMA με ασυμπτωτικά όρια. Ωστόσο, όταν το επίπεδο της αρχικής διαδικασίας είναι εκτός ελέγχου, τα ARL των δύο διαγραμμάτων ενδέχεται να διαφέρουν σημαντικά αναλόγως της τιμής εξομάλυνσης  $\lambda$  του EWMA.

Είναι σημαντικό να ποσοτικοποιηθεί η επίδραση των χρονικά μεταβαλλόμενων ορίων, καθώς τα διαγράμματα EWMA συχνά σχεδιάζονται (επιλογή παραμέτρων  $\lambda$  και  $L$ ) να δίνουν μικρό  $ARL_1$  κάτω από συγκεκριμένες συνθήκες.

Ωστόσο, λόγω του ότι τα χρονικά μεταβαλλόμενα όρια συγκλίνουν στις σταθερές ασυμπτωτικές τιμές, καθώς ο χρόνος αυξάνεται, για μετατοπίσεις οι οποίες συμβαίνουν αργότερα στο χρόνο τα δύο διαγράμματα έχουν παρόμοιες ιδιότητες μήκους ροής. Σαν αποτέλεσμα, τα διαγράμματα ελέγχου EWMA με χρονικά μεταβαλλόμενα όρια ελέγχου μπορούν να σχεδιαστούν με τον ίδιο τρόπο, όπως και τα ασυμπτωτικά όρια των EWMA.

Οι ιδιότητες των μηκών ροής των διαγραμμάτων ελέγχου EWMA με ασυμπτωτικά όρια καθορίστηκαν από τον Crowder (1987) [17] χρησιμοποιώντας μία ολοκληρωτική εξίσωση. Η προσέγγιση αυτή της λύσης με ολοκληρωτική εξίσωση δεν μπορεί να εφαρμοστεί για διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια. Ωστόσο, οι ιδιότητες των μηκών ροής των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα όρια μπορούν να προσεγγιστούν με χρήση μη ομογενούς διακριτής αλυσίδας Markov.

Η επίδραση των χρονικά μεταβαλλόμενων ορίων στην απόδοση ενός διαγράμματος αποτυπώνεται στο Σχήμα 3.8. Τα αποτελέσματα προέκυψαν χρησιμοποιώντας  $L=3$  σαν σταθερό όριο ελέγχου και υποθέτοντας για την εντός ελέγχου κατάσταση μέσο μηδέν και τυπική απόκλιση μονάδα. Στο Σχήμα 3.8, ο οριζόντιος άξονας δίνει την μετατόπιση του πραγματικού μέσου της διεργασίας, που παρατηρείται στην αρχή, και εκφράζεται σε  $\sigma_x$  μονάδες, η τυπική απόκλιση δηλαδή του δειγματικού μέσου. Τα αποτελέσματα που απεικονίζονται αφορούν μόνο θετικές μετατοπίσεις, αλλά λόγω συμμετρικότητας το ίδιο μοτίβο θα εμφανίζεται και για αρνητικές μετατοπίσεις.



Σχήμα 3.8: Απεικόνιση των ARL των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια ελέγχου

Στον Πίνακα 3.9 παρακάτω συνοψίζονται τα αποτελέσματα για τις υπόψη τέσσερις περιπτώσεις (για  $\lambda=0.1, 0.25$  και  $0.5$ ).

Πίνακας 3.9: Τιμές ARL για  $\lambda=0.1, 0.25$  και  $0.5$ , και με  $L=3$  για χρονικά μεταβαλλόμενα και σταθερά ασυμπτωτικά όρια, όπου  $\delta$ =μετατόπιση του μέσου της διεργασίας,  $fix = Σταθ.$  - Ασυμπ. όρια,  $vacl$ =Χρον. Μεταβαλλ. όρια.

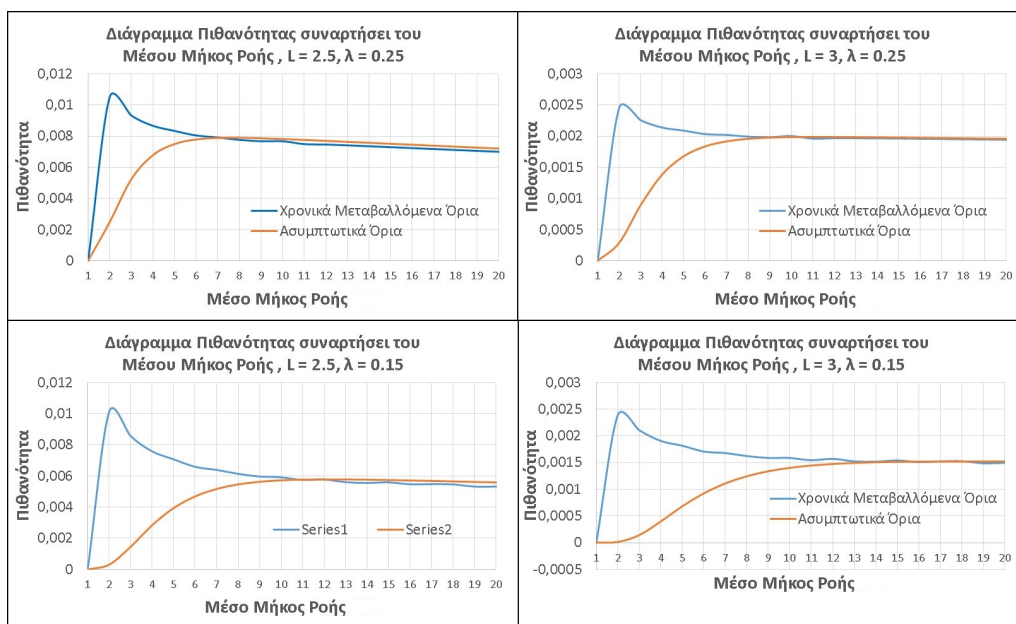
$\lambda$	0.1		0.25		3	
$L$	3		3		3	
$\delta$	<i>fix</i>	<i>vacl</i>	<i>fix</i>	<i>vacl</i>	<i>fix</i>	<i>vacl</i>
0	825.91	363.25	502.90	498.98	397.46	396.26
0.25	129.88	44.68	171.10	169.08	208.54	207.71
0.5	29.57	8.15	48.45	47.30	75.35	74.82
0.75	12.8	3.05	20.16	19.30	31.46	31.07
1	7.65	1.76	11.15	10.40	15.74	15.42
1.25	5.39	1.32	7.39	6.68	9.22	8.93
1.5	4.17	1.14	5.47	4.77	6.11	5.85
1.75	3.41	1.06	4.34	3.66	4.45	4.2
2	2.91	1.03	3.62	2.94	3.47	3.23

Από τον Πίνακα 3.9 διαπιστώνουμε ότι για μικρές τιμές του  $\lambda$  τα δύο διαγράμματα πα-

ρουσιάζουν σημαντική διαφορά στην ανιχνευτική τους ικανότητα. Βλέπουμε π.χ. ότι για μετατόπιση  $0.75 \sigma_x$ , το διάγραμμα με χρονικά μεταβαλλόμενα όρια εντοπίζει την μετατόπιση μετά από 3 περίπου παρατηρήσεις, ενώ εκείνο με τα ασυμπτωτικά όρια δίνει εκτός ελέγχου σήμα μετά από σχεδόν 13 παρατηρήσεις. Η υπόψη διαφορά ελαττώνεται καθώς το  $\lambda$  αυξάνεται (για  $\lambda=0.5$  οι τιμές των δύο ARL τείνουν να ταυτιστούν).

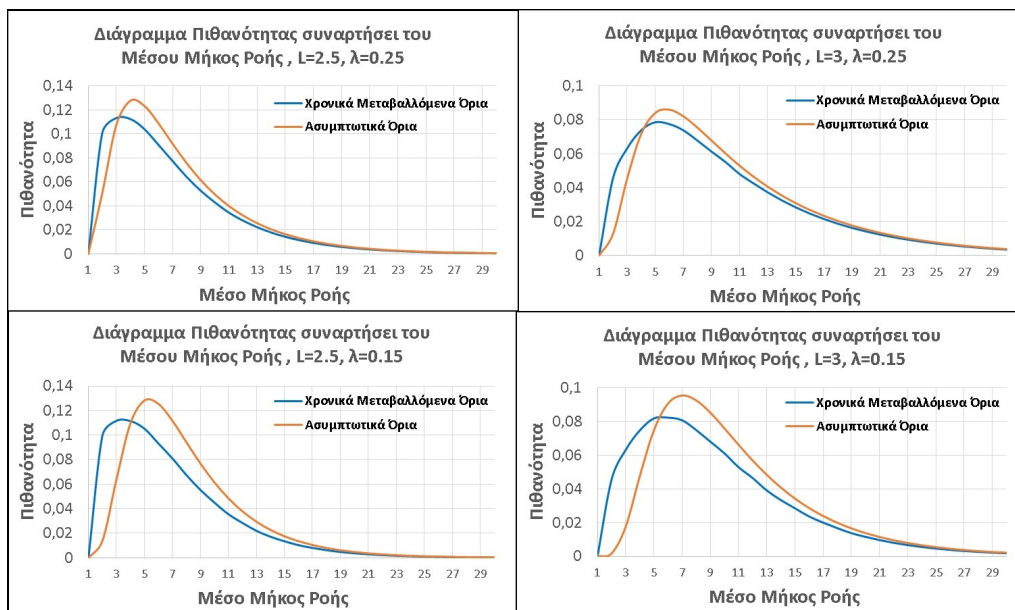
Γενικά, η επίδραση των χρονικά μεταβαλλόμενων ορίων ελέγχου στο ARL των διαγραμμάτων EWMA είναι ουσιαστική, όταν η διαδικασία δεν είναι αρχικά εντός ελέγχου, ειδικά όταν το  $\lambda$  είναι μικρό.

Ενδιαφέρον είναι επίσης να εξεταστεί και η επίδραση των χρονικά μεταβαλλόμενων ορίων στη διασπορά των ARL των διαγραμμάτων EWMA, η οποία μπορεί να υπολογιστεί μέσω των Σχέσεων (3.45) και (3.47). Τα Σχήματα 3.9 και 3.10 απεικονίζουν την πιθανότητα να ανιχνευτεί μετατόπιση συναρτήσει του μήκους ροής των διαγραμμάτων EWMA με χρονικά μεταβαλλόμενα όρια και ασυμπτωτικά όρια, όταν η αρχική διαδικασία βρίσκεται εντός ελέγχου και όταν μετατοπίζεται κατά μία  $\sigma_x$  μονάδα αντίστοιχα (θεωρούμε  $\sigma=1$ ). Τα δεδομένα υπολογίστηκαν με τη βοήθεια του κώδικα του Παραρτήματος Β (μέθοδος μαρκοβιανής αλυσίδας, βήμα  $g=300$ ). Στόχος είναι να απεικονιστεί η τάση των τιμών της πιθανότητας συναρτήσει του ARL.



Σχήμα 3.9: Πιθανότητα συμμετοχής κάθε παρατήρησης στη διαμόρφωση του ARL διαγράμματος EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια, για διεργασία εντός ελέγχου

Όπως φαίνεται στο Σχήμα 3.9, για μικρές τιμές του ARL, η πιθανότητα ανίχνευσης μετατόπισης παρουσιάζει απότομη άνοδο για τα διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια, σε σύγκριση με τα αντίστοιχα με σταθερά όρια. Η πιθανότητα αυτή συγκλίνει και για τις δύο περιπτώσεις καθώς το ARL αυξάνεται. Παρατηρούμε, επίσης, ότι για μικρότερες τιμές της παραμέτρου  $L$ , οι παραπάνω αρχικές άνοδοι εμφανίζονται οξύτερες. Όσον αφορά την



Σχήμα 3.10: Πιθανότητα συμμετοχής κάθε παρατήρησης στη διαμόρφωση του ARL διαγράμματος EWMA με χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια, για διεργασία εκτός ελέγχου

παράμετρο  $\lambda$ , φαίνεται ότι δεν επηρεάζει την οξύτητα της αρχικής ανόδου, αλλά την ταχύτητα σύγκλισης των πιθανοτήτων των δύο περιπτώσεων διαγραμμάτων. Βέβαια, οι τιμές των πιθανοτήτων είναι γενικά πολύ μικρές, οπότε η αρχική άνοδος που παρατηρείται, ήτοι αύξηση πιθανότητας για μικρές τιμές ARL σε εντός ελέγχου κατάσταση, δεν έχει κάποια ουσιαστική επίδραση.

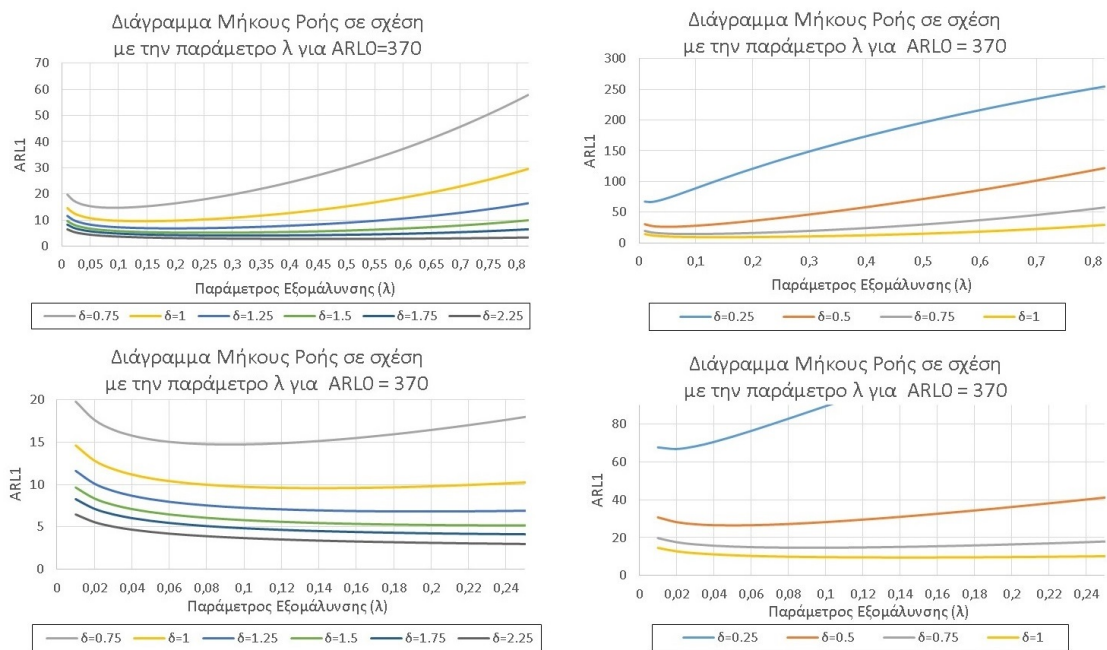
Από την άλλη πλευρά, όπως φαίνεται στο Σχήμα 3.10, καλύτερη πιθανότητα για ανίχνευση μετατόπισης (μικρότερο ARL) εμφανίζει η περίπτωση διαγράμματος EWMA με χρονικά μεταβαλλόμενα όρια. Αυτό είναι προφανώς επιθυμητό σε μια διαδικασία που είναι εκτός ελέγχου από την αρχή.

Από τη σύγκριση των γραφημάτων των Σχημάτων 3.9 και 3.10, λαμβάνοντας υπόψη και τα ανάλογα γραφήματα των διαγραμμάτων CUSUM και FIR-CUSUM [34], προκύπτει ότι η επίδραση των χρονικά μεταβαλλόμενων ορίων είναι παρόμοια με εκείνη που επιτυγχάνεται με το εργαλείο FIR. Σε κάθε περίπτωση πάντως, η επίδρασή τους εμφανίζεται να είναι λιγότερο έντονη από ότι στα διαγράμματα FIR-CUSUM, γεγονός που ενθαρρύνει ένα περαιτέρω στένεμα των ορίων, ιδιαίτερα για περιπτώσεις όπου τα προβλήματα ποιότητας εμφανίζονται κατά έναρξη της διεργασίας.

### 3.6 Βέλτιστες παράμετροι $\lambda$ και $L$

Παραπάνω διαπιστώσαμε ότι, τόσο στην περίπτωση που η διαδικασία βρίσκεται εντός ελέγχου ( $ARL_0$ ), όσο και στην περίπτωση εκτός ελέγχου ( $ARL_1$ ), η επιλογή των παραμέτρων  $\lambda$  και  $L$  επηρεάζεται από το ARL.

Με χρήση εντολών του πακέτου "spc" της R [31], μπορούμε να βρούμε τους συνδυασμούς των παραμέτρων  $\lambda$  και  $L$ , που δίνουν την επιθυμητή τιμή  $ARL_0$ . Αυτή επιλέγεται συνήθως να είναι 370 και 500, μιμούμενοι το αντίστοιχο μέσο μήκος ροής που συναντάται στα κλασικά διαγράμματα, όπου το Σφάλμα Τύπου I είναι 0.0027 και 0.002 αντίστοιχα. Με τον τρόπο αυτό καταρτίστηκαν πίνακες τιμών, τόσο για την περίπτωση ασυμπτωτικών ορίων όσο και για αυτή των μεταβαλλόμενων, βάσει των οποίων δημιουργήθηκαν τα Σχήματα 3.11, 3.12, 3.13 και 3.14.



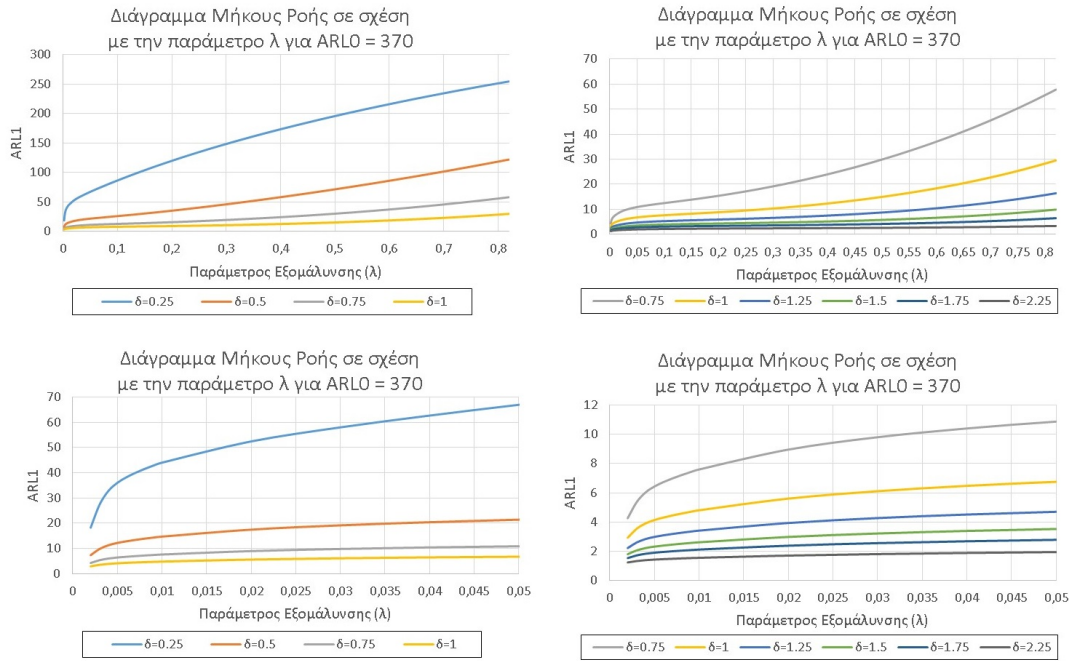
Σχήμα 3.11: Διαγράμματα Μέσου Μήκους Ροής σε σχέση την Παράμετρο Εξομάλυνσης  $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των ασυμπτωτικών ορίων και με  $ARL_0=370$

Από τη σύγκριση των παραπάνω σχημάτων διαπιστώνουμε ότι υπάρχει σημαντική διαφορά μεταξύ της περίπτωσης των ασυμπτωτικών ορίων και των χρονικά μεταβαλλόμενων. Όπως φαίνεται στα Σχήματα 3.11 και 3.14 (ασυμπτωτικά όρια), το  $ARL_1$  μειώνεται καθώς αυξάνεται η τιμή της παραμέτρου  $\lambda$ , εμφανίζει ελάχιστο, και στη συνέχεια αυξάνεται πάλι. Η τιμή του  $\lambda$  όπου εμφανίζεται το ελάχιστο αποτελεί την βέλτιστη επιλογή παραμέτρου για ικανότητα ανίχνευσης συγκεκριμένης μετατόπισης. Από την άλλη πλευρά, στην περίπτωση των χρονικά μεταβαλλόμενων ορίων, όπως φαίνεται στα Σχήματα 3.12 και 3.13, το  $ARL_1$  μειώνεται συνεχώς όσο μειώνεται η τιμή της παραμέτρου  $\lambda$ . Οπότε, αρκεί να μειώσουμε τόσο την τιμή του  $\lambda$ , ώστε να πάρουμε το επιθυμητό  $ARL_1$ .

Στον Πίνακα 3.10 αποτυπώνονται τα ζεύγη τιμών ( $\lambda$ ,  $L$ ), τα οποία υπολογίστηκαν με χρήση του πακέτου "spc" της R [31], για την περίπτωση που το επιθυμητό Σφάλμα Τύπου I είναι 5% ( $ARL_0=370$ ) και το σφάλμα τύπου II κοντά στο 50% ( $ARL_1=2$ ).

Η τιμή του  $ARL$  που προκύπτει εξαρτάται τόσο από το  $\lambda$ , όσο και τους κόμβους τετραγωνισμού που επιλέγουμε. Για πολύ χαμηλές τιμές του  $\lambda$  θα πρέπει να εξασφαλίζουμε ότι το





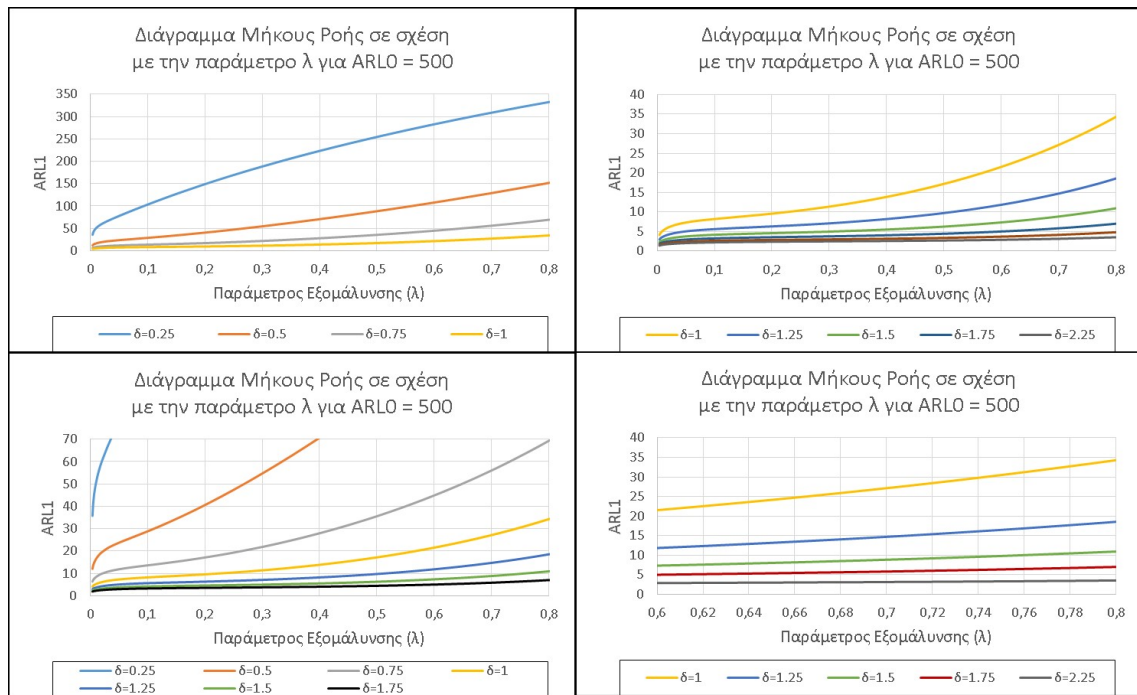
Σχήμα 3.12: Διαγράμματα Μέσου Μήκους Ροής σε σχέση με την Παράμετρο Εξομάλυνσης  $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των χρονικά μεταβαλλόμενων ορίων και με  $ARL_0=370$

Πίνακας 3.10: Ζεύγη τιμών παραμέτρων για  $ARL_0=370$

$\lambda$	$L$	$ARL_0$	$ARL_1$ για $\delta=1$	κόμβοι	Σφάλμα Τύπου II
0.0004	1.404533	370	2.957128	270	0.661834
0.0003	1.37677	370	2.888783	270	0.653833
0.0002	1.34213	370	2.805338	270	0.643537
0.0001	1.292947	370	2.690323	270	0.628297
0.00009	1.286393	370	2.675301	270	0.62621
0.00008	1.2745237	370	2.574740	270	0.611611

$ARL_1$  παραμένει σταθερό για τουλάχιστον δύο διαφορετικές τιμές των κόμβων τετραγωνισμού. Για τα παραπάνω ζεύγη παραμέτρων διαπιστώθηκε ότι η τιμή του  $ARL_1$  σταθεροποιείται πάνω από 250 κόμβους. Για τον υπολογισμό του κάθε ζεύγους τιμών απαιτείται πολύ μεγάλος υπολογιστικός χρόνος, ενδεικτικά αναφέρεται ότι για  $\lambda \leq 0.00001$  απαιτείται χρόνος πολύ μεγαλύτερος από 3 ώρες με ένα μέσο υπολογιστή.

Επίσης, όσον αφορά τα διαγράμματα με ασυμπτωτικά όρια, καταλήγουμε στο συμπέρασμα ότι το Σφάλμα Τύπου II παραμένει πολύ μεγάλο όσο η ανιχνευτική ικανότητα ( $ARL_1$ ), ακόμη και για μεγάλες μετατοπίσεις, παραμένει σχετικά μέτρια (τιμές μεγαλύτερες του 2.80. Από την άλλη πλευρά, στην περίπτωση των διαγραμμάτων με χρονικά μεταβαλλόμενα όρια, μπορούμε με κατάλληλη επιλογή της παραμέτρου  $\lambda$  (της τάξης του 0.00008) να πετύχουμε ικανοποιητική ανιχνευτική ικανότητα για μετατόπιση  $\delta=1$  ( $ARL_1=2.575$  όπως φαίνεται στον Πίνακα 3.10) με μικρό Σφάλμα Τύπου II.



Σχήμα 3.13: Διαγράμματα Μέσου Μήκους Ροής σε σχέση την Παράμετρο Εξομάλυνσης  $\lambda$ , για διάφορες μετατοπίσεις, για την περίπτωση των χρονικά μεταβαλλόμενων ορίων και με  $ARL_0=500$

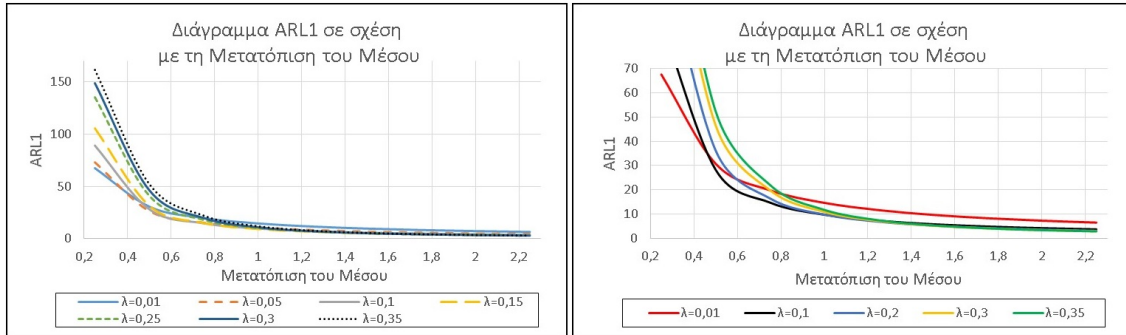
Τέλος, όπως παρατηρούμε στο Σχήμα 3.15, για μικρές τιμές της παραμέτρου  $\lambda$ , οι τιμές του  $L$  των δύο περιπτώσεων διαγραμμάτων εμφανίζουν μια απόκλιση, που μετά από μια τιμή που εκτιμάται μεταξύ  $\lambda=0.2$  με  $\lambda=0.3$ , σχεδόν παύει να υφίσταται.

Συνοψίζοντας, τα βήματα που πρέπει να ακολουθούνται για τον προσδιορισμό του βέλτιστου ζεύγους παραμέτρων  $\lambda$  και  $L$  αποτυπώνονται στον ακόλουθο αλγόριθμο:

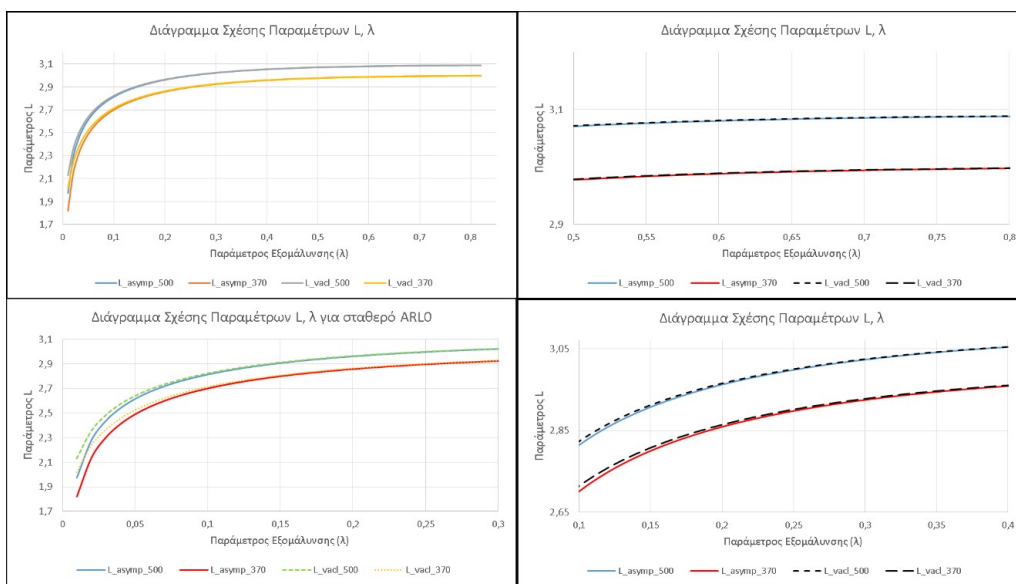
- α. Καθορισμός του μέγιστου επιτρεπτού Σφάλματος Τύπου I, με συνήθη τιμή:  $\alpha=0.0027$ .
- β. Υπολογισμός του επιθυμητού  $ARL_0$  για το Σφάλμα Τύπου I από τη σχέση  $ARL_0=\frac{1}{\alpha}$ . Για  $\alpha=0.0027$  οπότε προκύπτει:  $ARL_0=370$ .
- γ. Καθορισμός της μετατόπισης  $\delta$  που θέλουμε να ανιχνεύει το διάγραμά μας.
- δ. Υπολογισμός (επιλέγοντας ένα εύρος τιμών για το  $\lambda$ ) των ζευγών παραμέτρων  $\lambda$  και  $L$ , που δίνουν το επιθυμητό  $ARL_0$  (χρήση κώδικα: επιλογή ασυμπτωτικών ή χρονικά μεταβαλλόμενων ορίων).
- ε. Υπολογισμός του  $ARL_1$ , για τη μετατόπιση  $\delta$  που θέλουμε να ανιχνεύει το διάγραμά μας, για κάθε ζεύγος παραμέτρων  $\lambda$  και  $L$  που υπολογίστηκε από το παραπάνω βήμα.
- στ. Επιλογή του ζεύγους τιμών που δίνει μικρότερο  $ARL_1$  για την επιλεγμένη μετατόπιση.

Τα Σφάλματα Τύπου I και II για την περίπτωσή μας αναλύονται στον Πίνακα 3.11.





Σχήμα 3.14: Διαγράμματα Μέσου Μήκους Ροής σε σχέση το μέγεθος της μετατόπισης, για διάφορες τιμές της παραμέτρου  $\lambda$ , για την περίπτωση των ασυμπτωτικών ορίων και με  $ARL_0=370$



Σχήμα 3.15: Διαγράμματα Παραμέτρου  $L$  συναρτήσει της παραμέτρου εξομάλυνσης  $\lambda$ , για τις χρονικά μεταβαλλόμενα και ασυμπτωτικά όρια για  $ARL_0=370$  και  $ARL_0=500$

Συνεπώς, με  $t \in [1, \infty)$ , για τα Σφάλματα Τύπου I και II ισχύουν οι παρακάτω σχέσεις αντίστοιχα:

$$\alpha = P[ Z(t) > UCL(t) \text{ και } Z(t) < LCL(t) ] = \frac{1}{ARL_0}, \text{ με } Z \sim N(\mu_0, \sigma^2) \quad (3.51)$$

και

$$\beta = P[ LCL(t) < Z(t) < UCL(t) ] \Rightarrow$$

$$\beta = 1 - P[ Z(t) > UCL(t) \text{ και } Z(t) < LCL(t) ] = \frac{1}{ARL_0}, \text{ με } Z \sim N(\mu_1, \sigma^2) \quad (3.52)$$

Πίνακας 3.11: Δυνατοί Συνδυασμοί Ενδεχομένων – Τελικής Ενέργειας

$\delta$	$Z > UCL    Z < LCL$	Ενέργεια βάσει αποτελέσματος στατιστικού	Είναι η ενέργεια σωστή ; (σύγκριση $\mu_0$ με $\mu_1$ )	Είδος σφάλματος
0	NAI $\Rightarrow$ Σήμα εκτός	Διακοπή παραγωγής	Όχι (διεργασία εντός)	Τύπου I
0	OXI $\Rightarrow$ Σήμα εντός	Συνέχιση παραγωγής	Ναι	
1	NAI $\Rightarrow$ Σήμα εκτός	Διακοπή παραγωγής	Ναι	
1	OXI $\Rightarrow$ Σήμα εντός	Συνέχιση παραγωγής	Όχι (διεργασία εκτός)	Τύπου II

## Κεφάλαιο 4

# Μέθοδος Γρήγορης Αρχικής Απόκρισης (FIR-EWMA)

### 4.1 Διαγράμματα Ελέγχου FIR – EWMA

Παρατηρήσαμε ότι τα διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια, όταν συγκρίνονται με τα αντίστοιχα διαγράμματα με ασυμπτωτικά όρια, παρουσιάζουν όμοιες ιδιότητες με τα διαγράμματα FIR-CUSUM. Η χρήση, ωστόσο, των πρώτων διαγραμμάτων διαφέρει από τη χρήση των τελευταίων, καθώς η ρύθμιση των διαγραμμάτων ελέγχου EWMA διορθώνει μόνο τα όρια ελέγχου, έτσι ώστε να λάβουν υπόψη τη χρονικά εξαρτώμενη φύση της στατιστικής συνάρτησης EWMA της Σχέσης (3.5).

Για γρήγορη αντίδραση σε μετατόπιση του μέσου, οι Lucas και Saccucci (1990) [35] πρότειναν τη χρήση δύο μονόπλευρων διαγραμμάτων EWMA με αρχική κατάσταση διαφορετική από το μηδέν. Κατόπιν, οι Rhoads et al (1996) [46], υιοθετώντας την παραπάνω προσέγγιση, εισήγαγαν χρονικά μεταβαλλόμενα όρια σε κάθε μονόπλευρο διάγραμμα. Και οι δύο αυτές μέθοδοι έχουν το επιθυμητό αποτέλεσμα, δημιουργώντας διάγραμμα περισσότερο ευαίσθητο σε πρωτοεμφανιζόμενα προβλήματα, χρησιμοποιούν όμως δύο διαγράμματα EWMA για την ολοκλήρωση του έργου που προγενέστερα ήταν εφικτό με ένα μόνο διάγραμμα.

Μια διαφορετική προσέγγιση προτάθηκε από τον Steiner (1999) [49], με την οποία διατηρείται η απλότητα ενός απλού διαγράμματος EWMA. Συγκεκριμένα, το εργαλείο FIR, στα διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια, εφαρμόζεται μέσω περαιτέρω στενέματος τους για τις πρώτες μονάδες του δείγματος. Αυτό πραγματοποιείται με τη βοήθεια της παρακάτω σχέσης:

$$FIR_{adj} = 1 - (1 - f)^{1+\alpha(j-1)} \quad (4.1)$$

Με αυτή τη ρύθμιση το εργαλείο FIR καθιστά τα όρια ελέγχου για το πρώτο δείγμα ( $j = 1$ ) ένα κλάσμα της αρχικής απόστασης από το σημείο εκκίνησης.

Η επίδραση της χρήσης του εργαλείου FIR ελαττώνεται με το χρόνο για να διασφαλίσουμε ότι οι ιδιότητες του μήκους ροής του EWMA θα παραμείνουν αμετάβλητες. Μία εύλογη ρύθμιση θα ήταν να ορίσουμε την παράμετρο ρύθμισης  $\alpha$  έτσι ώστε η επίδραση FIR να έχει μικρή επίδραση μετά την παρατήρηση 20 και η παράσταση  $FIR_{adj}$  στην παρατήρηση 20 να είναι 0.99. Αυτό θα πρέπει να αρκεί για να ανίχνευση προβλημάτων ποιότητας που συμβαίνουν στην αρχή των μετρήσεων. Συνεπώς, για  $FIR_{adj}=0.99$  και  $j=20$ , η Σχέση (4.1) θα δώσει:

$$0.99 = 1 - (1 - f)^{1+19\alpha} \Rightarrow$$

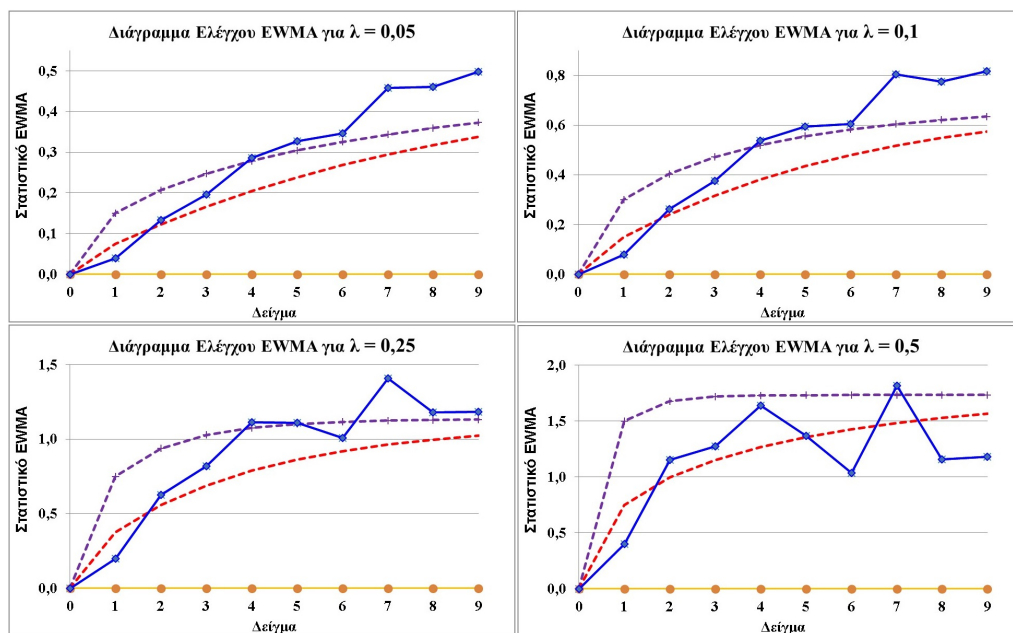
$$\Rightarrow 10^{-2} = (1 - f)^{1+19\alpha} \Rightarrow -2 = (1 + 19\alpha)\log(1 - f) \Rightarrow$$

$$\alpha = \frac{\left(-\frac{2}{\log(1-f)} - 1\right)}{19} \quad (4.2)$$

Μια συνηθισμένη τιμή για το  $f$  είναι  $f=0.5$ , που δίνει  $\alpha=0.3$ . Εφαρμόζοντας τον παράγοντα ρύθμισης  $FIR_{adj}$  της Σχέσης (4.1) στις Σχέσεις (3.9) και (3.11), προκύπτουν τα όρια ελέγχου FIR-EWMA ως εξής:

$$\mu_x \pm L\sigma_x(1 - (1 - f)^{1+\alpha(j-1)})\sqrt{\frac{\lambda[1 - (1 - \lambda)^{2j}]}{(2 - \lambda)n}} \quad (4.3)$$

Χρησιμοποιώντας τα δεδομένα του Πίνακα Α.4 του παραρτήματος, μπορούμε να δούμε την επίδραση της χρήσης των ορίων της Σχέσης (4.3), με  $f=0.5$  και  $\alpha=0.3$ , στο Σχήμα 4.1. Το πλεονέκτημα του επιπρόσθετου στενέματος των ορίων ελέγχου για την ανίχνευση προβλημάτων ποιότητας είναι προφανές: για διάφορες τιμές του  $\lambda$ , το διάγραμμα FIR-EWMA δίνει σήμα σε λίγες μόλις παρατηρήσεις, ειδικά για μεγάλες τιμές του  $\lambda$ .



Σχήμα 4.1: Διαγράμματα EWMA με χρονικά μεταβαλλόμενα όρια σε σχέση με όρια FIR-EWMA

Το παραπάνω πλεονέκτημα όμως μειώνεται όταν η μετατόπιση του μέσου της διαδικασίας δεν συμβεί κοντά στην έναρξη των μετρήσεων. Τότε το διάγραμμα FIR-EWMA θα έχει μεγαλύτερα  $ARL_1$  από ότι το κλασικό διάγραμμα EWMA. Αυτό συμβαίνει, διότι τα διαγράμματα FIR-EWMA απαιτούν μεγαλύτερες τιμές για το όριο ελέγχου  $L$  από ότι τα παραδοσιακά διαγράμματα EWMA, για την επίτευξη  $ARL_0=370$ . Με τη χρήση του πακέτου "spc" της R [31] (εντολές `xewma.arl` και `xewma.crit`), μπορεί εύκολα να επιβεβαιωθεί το παραπάνω γεγονός, υπολογίζοντας την τιμή του  $L$ , ώστε να προκύπτει  $ARL_0=370$ , τόσο για την μια όσο και για την άλλη περίπτωση διαγραμμάτων. Από τα αποτελέσματα φαίνεται η μεγαλύτερη τιμή του  $L$  στην περίπτωση των διαγραμμάτων FIR-EWMA ( $L=2.888008 > 2.792883$ ).

```
xewma.crit(l=0.14, 370, sided="two", limits="Steiner", r=50)
2.888008
xewma.crit(l=0.14, 370, sided="two", limits="vacl", r=50)
2.792883
```

και μπορούμε να επιβεβαιώνουμε ότι για αυτές τις τιμές των  $\lambda$  και  $L$  προκύπτει  $ARL_0=370$ :

```
xewma.arl(l=0.14, c=2.888008, mu=0, sided="two", limits="Steiner", r=50)
369.9995
xewma.arl(l=0.14, c=2.792883, mu=0, sided="two", limits="vacl", r=50)
369.9998
```

Στη συνέχεια, υπολογίζουμε το  $ARL_1$  για μετατόπιση  $\delta=1$ , από όπου φαίνεται ότι για μετατόπιση στο μέσο, που υπάρχει από την έναρξη μέτρησης, το διάγραμμα FIR-EWMA έχει πολύ μικρότερη τιμή  $ARL_1$ , δηλαδή καλύτερη ανιχνευτική ικανότητα από την αντίστοιχη του διαγράμματος με χρονικά μεταβαλλόμενα όρια. Αυτό μπορούμε να το δούμε παρακάτω:

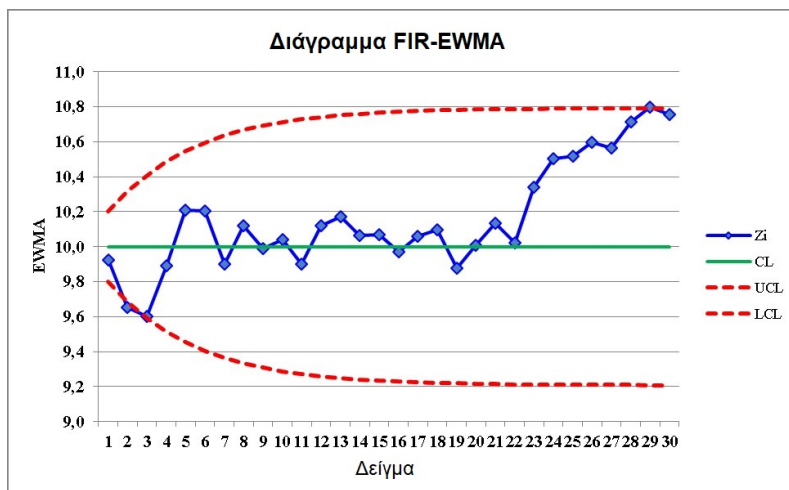
```
xewma.arl(l=0.14, c=2.888008, mu=1, sided="two", limits="Steiner", r=50)
4.609751
xewma.arl(l=0.14, c=2.792883, mu=1, sided="two", limits="vacl", r=50)
8.100764
```

Τέλος, γνωρίζοντας ότι η μετατόπιση στο μέσο συμβαίνει μετά την 20η παρατήρηση, υπολογίζουμε τα νέα  $ARL_1$ , λαμβάνοντας υπόψη αυτό το δεδομένο. Από τα αποτελέσματα διαπιστώνουμε ότι το διάγραμμα FIR-EWMA εμφανίζει χειρότερη ανιχνευτική ικανότητα, δηλαδή μεγαλύτερο  $ARL_1$ , καθώς  $ARL_{1vacl}=9.41$ , ενώ  $ARL_{1STEINER}=9.93$ .

```
xewma.arl(l=0.14, c=2.888008, mu=1, sided="two", limits="Steiner", r=50, q=20)
9.925329
xewma.arl(l=0.14, c=2.792883, mu=1, sided="two", limits="vacl", r=50, q=20)
9.414313
```

Ας εφαρμόσουμε τη μέθοδο, δηλαδή την Σχέση (4.3), με  $f=0.5$ , στα δεδομένα του Πίνακα A.1 του παραρτήματος. Ως παράμετροι  $\lambda$  και  $L$  επιλέγονται οι τιμές  $\lambda=0.14$  και  $L=2.888008$ , που αποτελούν τις βέλτιστες για ανίχνευση μετατόπισης  $\delta=1$  και  $ARL_0=370$ .

Στο διάγραμμα του Σχήματος 4.2 που προκύπτει παρατηρούμε ότι η μετατόπιση του μέσου



Σχήμα 4.2: Διάγραμμα FIR-EWMA για τα δεδομένα του Πίνακα Α.1

κατά 2.1 μονάδες στη 2η παρατήρηση, ανιχνεύεται αμέσως, καθώς το διάγραμμα δίνει αμέσως εκτός ελέγχου σήμα. Κάτι ανάλογο δεν συνέβη στην αντίστοιχη περίπτωση του κλασικού διαγράμματος EWMA (Σχήμα 3.2). Προφανώς, το συγκεκριμένο σετ δεδομένων αναδεικνύει μόνο το πλεονέκτημα της μεθόδου, και όχι το μειονέκτημα που θα ήταν η ανίχνευση της μετατόπισης του μέσου (που συμβαίνει στην 20η παρατήρηση) αργότερα από την 29η παρατήρηση. Η θεωρητική ανιχνευτική ικανότητα του κλασικού EWMA διαγράμματος είναι μεν καλύτερη, για ύστερες μετατοπίσεις ( $q = 20$ ), από αυτή του FIR-EWMA (9.414313 έναντι 9.92529), οι τιμές της ωστόσο είναι αρκετά κοντινές που το κέρδος δεν είναι εμφανές στη συγκεκριμένη περίπτωση δεδομένων.

#### 4.1.1 Επιλογή βέλτιστης τιμής της παραμέτρου $f$

Η συνηθισμένη τιμή για την παράμετρο FIR είναι  $f=0.5$ . Ωστόσο, με τη βοήθεια κώδικα παρόμοιου με αυτόν που χρησιμοποιείται στο πακέτο "spc" της R [31], υπολογίστηκαν οι θεωρητικές τιμές ARL για διαφορετικές τιμές του  $f$ . Η επιλογή του βέλτιστη  $f$  εξαρτάται από την επιλογή τιμών για τις παραμέτρους  $\lambda$  και  $L$ . Γενικά, εφαρμόζουμε την πρακτική κατά την οποία επιλέγουμε το επιθυμητό  $ARL_0=370$  και για συγκεκριμένο ζεύγος παραμέτρων  $\lambda$  και  $L$  βρίσκουμε (με δοκιμές) την τιμή του  $f$  που αντιστοιχεί στο επιθυμητό  $ARL_0$ . Τα αποτελέσματα παρουσιάζονται στους Πίνακες 4.1 και 4.2, όπου βλέπουμε ότι για το ζεύγος  $\lambda=0.25$ ,  $L=3$  η βέλτιστη επιλογή του είναι η  $f=0.47316$ , ενώ για το ζεύγος  $\lambda=0.1$ ,  $L=3$ , η βέλτιστη τιμή είναι  $f=0.2525$ .

Ένα πρώτο συμπέρασμα είναι ότι για να έχουμε ένα ουσιαστικό όφελος από το εργαλείο FIR, θα πρέπει το  $f$  να είναι σχετικά μικρό. Η χρήση  $f=0.5$  αντιστοιχεί σε προσαρμογή των χρονικά μεταβαλλόμενων ορίων στο μισό της αρχικής περιόδου, μιμούμαστε δηλαδή το 50% προβάδισμα έναρξης ( $hs$ ) των διαγραμμάτων FIR-CUSUM.

Ας εφαρμόσουμε τη μέθοδο (Σχέση (4.3), με  $f=0.2525$ ) πάλι στα δεδομένα του Πίνακα Α.1 του παραρτήματος. Ως παράμετροι  $\lambda$  και  $L$  επιλέγονται οι τιμές  $\lambda=0.1$  και  $L=3$ . Το

Πίνακας 4.1: Θεωρητικές τιμές ARL για διάφορες τιμές  $f$

$\mu_x/\sigma_{\bar{x}}$	$\lambda = 0.25 \quad L = 3$						
	$f = 0.4$	$f = 0.47316$	$f = 0.5$	$f = 0.6$	$f = 0.7$	$f = 0.8$	$f = 1$
0.00	324.359	370.002	384.404	427.695	457.004	475.998	498.977
0.25	103.117	118.453	123.396	138.721	149.756	157.501	169.077
0.50	24.543	28.647	30.036	34.642	38.390	41.417	47.303
0.75	8.420	9.951	10.497	12.434	14.196	15.788	19.297
1.00	4.078	4.815	5.090	6.119	7.137	8.129	10.400
1.50	1.861	2.123	2.228	2.650	3.117	3.616	4.773
2.00	1.309	1.426	1.476	1.688	1.944	2.237	2.937
3.00	1.038	1.061	1.072	1.125	1.204	1.312	1.617
4.00	1.003	1.005	1.006	1.014	1.029	1.056	1.163

Πίνακας 4.2: Θεωρητικές τιμές ARL για διάφορες τιμές  $f$

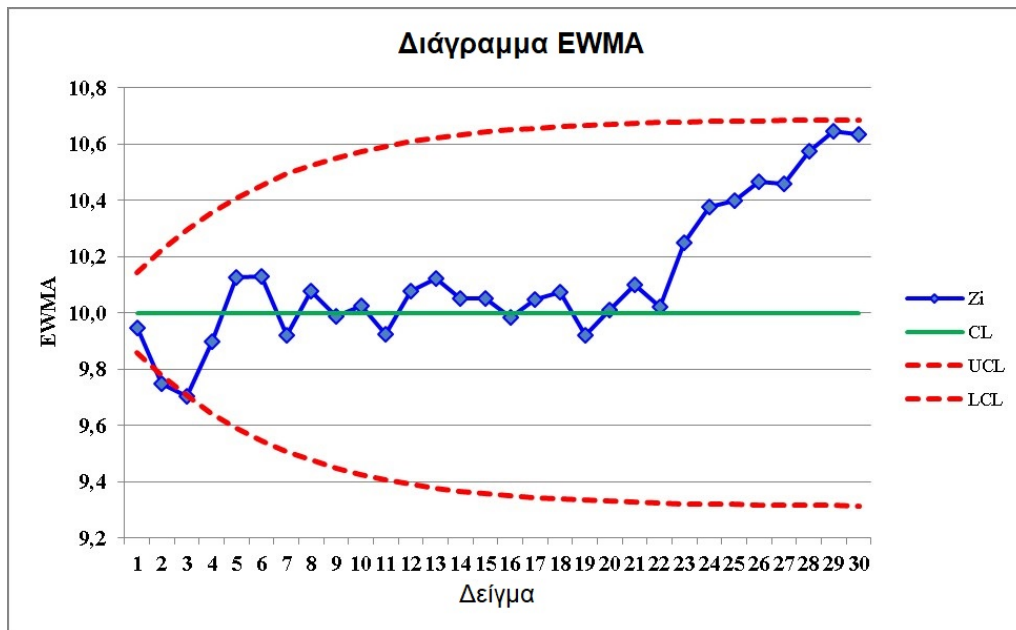
$\mu_x/\sigma_{\bar{x}}$	$\lambda = 0.10 \quad L = 3$						
	$f = 0.2525$	$f = 0.4$	$f = 0.5$	$f = 0.6$	$f = 0.7$	$f = 0.8$	$f = 1$
0.00	370.0053	564.280	659.298	726.048	770.084	797.806	828.626
0.25	59.610	91.369	107.406	119.09	127.188	132.638	140.245
0.50	13.212	20.335	24.228	27.312	29.681	31.486	34.761
0.75	5.250	8.005	9.678	11.144	12.398	13.468	15.644
1.00	2.846	4.198	5.117	6.004	6.836	7.609	9.250
1.50	1.475	1.916	2.283	2.694	3.136	3.595	4.614
2.00	1.148	1.320	1.489	1.701	1.954	2.238	2.903
3.00	1.013	1.038	1.072	1.126	1.205	1.312	1.614
4.00	1.001	1.003	1.006	1.014	1.029	1.056	1.162

διάγραμμα που προκύπτει απεικονίζεται στο Σχήμα 4.3.

Αξίζει να σημειωθεί ότι στην περίπτωση αυτή, σε σύγκριση με το αντίστοιχο διάγραμμα ελέγχου με  $f=0.5$  (Σχήμα 4.2), η μετατόπιση του μέσου (που συμβαίνει στην 20η παρατήρηση) δεν γίνεται αντιληπτή. Αυτό αποτελεί το μειονέκτημα χρήσης εργαλείων FIR στα εν λόγω διαγράμματα.

Σε γενικές γραμμές, το διάγραμμα FIR-EWMA με βέλτιστο  $f$  είναι καλύτερο από τις προγενέστερες προσεγγίσεις. Για παράδειγμα, με  $\lambda=0.25$ ,  $f=0.47316$  και μέση μετατόπιση μία μονάδα τυπικής απόκλισης, το προτεινόμενο διάγραμμα FIR-EWMA εντοπίζει τη μετατόπιση σε 4.815 παρατηρήσεις. Αντίστοιχα, με  $\lambda=0.1$ , μέση μετατόπιση μία μονάδα τυπικής απόκλισης και  $f=0.2525$ , το προτεινόμενο διάγραμμα FIR-EWMA απαιτεί μόνο 2.846 παρατηρήσεις για να δώσει εκτός ελέγχου σήμα. Δηλαδή για μικρότερο  $\lambda$  και  $f$  πετυχαίνουμε μικρότερο  $ARL_1$ . Επίσης, από τη σύγκριση των στηλών των δύο των πινάκων που αφορούν στα  $f=0.47316$  και  $f=0.2525$ , διαπιστώνουμε ότι η μείωση των εκτός ελέγχου ARL εμφανίζεται να είναι μεγαλύτερη, όταν το  $\lambda$  είναι μικρότερο.

Ο βέλτιστος συνδυασμός  $\lambda$ ,  $L$  και  $f$  μπορεί να βρεθεί υπολογίζοντας τα αντίστοιχα ARL,



Σχήμα 4.3: Διάγραμμα FIR-EWMA με βέλτιστο  $f$  με τα δεδομένα του Πίνακα Α.1

όπως απεικονίζεται στον Πίνακα 4.3.

Πίνακας 4.3: Θεωρητικές τιμές ARL για διαφορετικούς συνδυασμούς  $\lambda$ ,  $L$  και  $f$

$\lambda$	0.0001		0.01		0.01	
$L$	1.660853		2.1		2.193302	
$f$	0.543		0.69754		0.5	
$\mu_x/\sigma_{\bar{x}}$	$q = 1$	$q = 20$	$q = 1$	$q = 20$	$q = 1$	$q = 20$
0.00	370.082	-	370.002	-	370.002	-
1.00	2.218	2.611	3.686	6.646	3.010	6.956
2.00		-	1.407	3.372	1.241	3.5

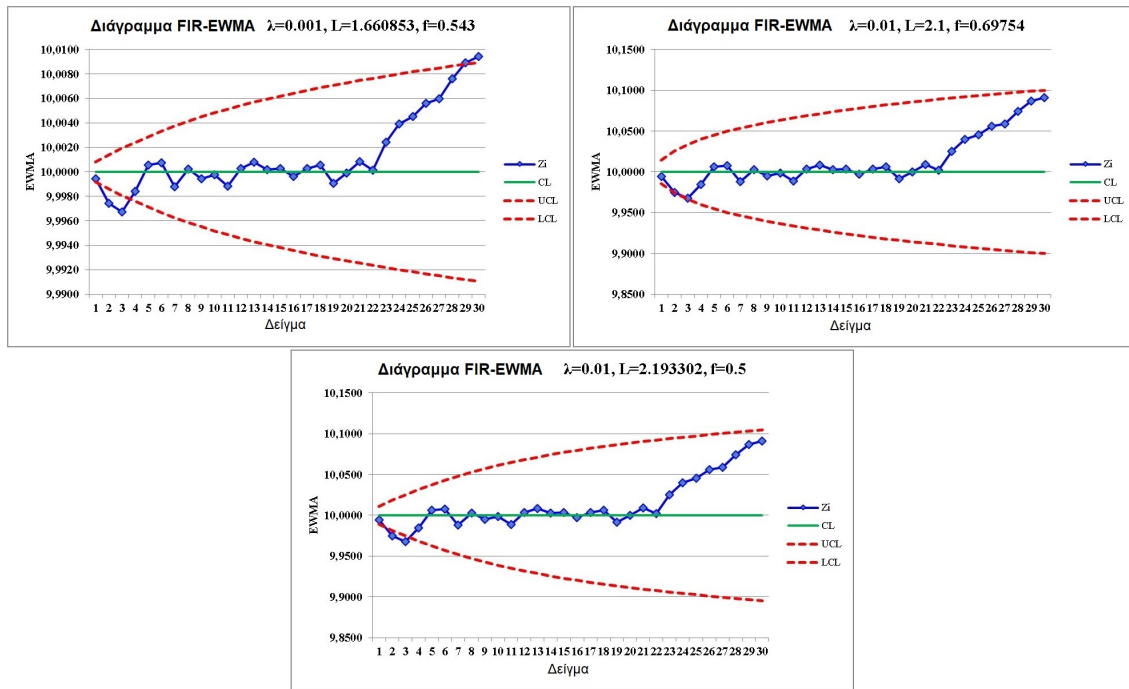
Τα διαγράμματα FIR-EWMA που προκύπτουν με τους παραπάνω συνδυασμούς παραμέτρων  $\lambda$ ,  $L$  και  $f$  απεικονίζονται στο Σχήμα 4.4.

Από τον Πίνακα 4.3 διαπιστώνουμε ότι αυξάνοντας την παράμετρο  $L$  και μειώνοντας την παράμετρο  $f$  μπορούμε να διατηρήσουμε σταθερό  $ARL_0$ , άρα και να πετύχουμε την επιθυμητή τιμή για το  $ARL_0$ . Επίσης, μπορούμε να πετύχουμε μικρότερο  $ARL_1$  μειώνοντας την τιμή του  $f$ . Παρατηρούμε, όμως, ότι αν η μετατόπιση δεν συμβεί κοντά στην έναρξη, όπως στα δεδομένα του παραδείγματος όπου η μετατόπιση λαμβάνει χώρα κατά την 20η παρατήρηση ( $q=20$ ), τότε το  $ARL_1$  αυξάνεται.

## 4.2 FIR με αρχική τιμής εκκίνησης

Σε αντιστοιχία με τη μέθοδο γρήγορης αρχικής απόκρισης των διαγραμμάτων FIR-CUSUM, ανάλογη τεχνική μπορούμε να εφαρμόσουμε και στα διαγράμματα EWMA, τοποθε-





Σχήμα 4.4: Διαγράμματα FIR-EWMA για βέλτιστους συνδυασμούς  $\lambda$ ,  $L$  και  $f$  με τα δεδομένα του Πίνακα Α.1

τώντας αρχική τιμή (προβαδίσμα) κατά την εκκίνηση (head start,  $hs$ ).

Έτσι, αν στη Σχέση (3.3), αντί για  $Z_0=\mu_0$  θέσουμε  $Z_0=\mu_0 + hs$ , όπου  $hs$  η επιθυμητή τιμή προβαδίσματος, τότε προκύπτει:

$$Z_j = \lambda \sum_{k=0}^{j-1} (1-\lambda)^k \bar{X}_{j-k} + (1-\lambda)^j Z_0 \Rightarrow$$

$$Z_{j(\mu\epsilon\ hs)} = \lambda \sum_{k=0}^{j-1} (1-\lambda)^k \bar{X}_{j-k} + (1-\lambda)^j (\mu_0 + hs) \Rightarrow$$

$$Z_{j(\mu\epsilon\ hs)} = Z_{j(\chi\omega\rho\rho\rho\varsigma\ hs)} + (1-\lambda)^j hs \quad (4.4)$$

Από την παραπάνω σχέση προκύπτει ότι η πρωτοποθέτηση τιμής φέρνει τις πρώτες παρατηρήσεις πιο κοντά στα όρια ελέγχου, διότι στην ποσότητα που είχαμε πριν  $Z_{j(\chi\omega\rho\rho\rho\varsigma\ hs)}$ , προστίθεται η ποσότητα  $(1-\lambda)^j hs$  (για θετικές μετατοπίσεις, αντίστοιχα εφαρμόζουμε  $-hs$  για αρνητικές μετατοπίσεις). Καθώς προχωράμε σε δειγματοληψίες ή αλλιώς καθώς αυξάνεται το  $j$ , η ποσότητα  $(1-\lambda)^j hs$  συνεχώς μειώνεται. Επομένως, η επίδραση του  $hs$  γίνεται ολοένα και μικρότερη. Στον Πίνακα 4.4 υπολογίστηκαν οι πρώτες 30 παρατηρήσεις για  $\lambda=0.14$ . Διαπιστώνουμε ότι η τιμή του  $hs$  φθάνει στο 1% της αρχικής τιμής του, δηλαδή στο 0.01 μετά από περίπου 30 παρατηρήσεις.

Συνοπτικά, ισχύει ότι:

$$\text{καθώς } j \uparrow \Rightarrow (1-\lambda)^j hs \downarrow$$

Πίνακας 4.4: Υπολογισμός της τιμής της ποσότητας  $(1 - \lambda)^j$ , καθώς το  $j$  αυξάνει

Δείγμα	$(1 - \lambda)^j$	Δείγμα	$(1 - \lambda)^j$	Δείγμα	$(1 - \lambda)^j$
1	0.86	11	0.190319	21	0.042118
2	0.7396	12	0.163675	22	0.036221
3	0.636056	13	0.14076	23	0.03115
4	0.547008	14	0.121054	24	0.026789
5	0.470427	15	0.104106	25	0.023039
6	0.404567	16	0.089531	26	0.019813
7	0.347928	17	0.076997	27	0.01704
8	0.299218	18	0.066217	28	0.014654
9	0.257327	19	0.056947	29	0.012602
10	0.221302	20	0.048974	30	0.010838

Συνεπώς, στην περίπτωση που η διαδικασία είναι εκτός ελέγχου από την έναρξη μετρήσεων, τότε με την υπόψη τεχνική θα επιταχυνθεί η εμφάνιση εκτός ελέγχου σήματος.

Παρατηρούμε επίσης, ότι με μείωση της τιμής της παραμέτρου  $\lambda$ , η ποσότητα  $(1 - \lambda)^j$  αυξάνεται. Συνεπώς, για μικρότερες τιμές του  $\lambda$  η αρχική επίδραση του  $hs$  γίνεται μεγαλύτερη. Καθώς όμως, αυξάνεται η αρχική τιμή του  $hs$  απαιτούνται περισσότερες δειγματοληψίες για να μειωθεί η επίδρασή του στο 1%. Έτσι, για παράδειγμα για  $\lambda=0.01$  απαιτούνται 458 παρατηρήσεις για να φθάσει η επίδραση  $hs$  σε ποσοστό 1% της αρχικής τιμής του.

Συνεπώς, το  $\lambda$  επηρεάζει τη διάρκεια επίδρασης του  $hs$ . Συνοπτικά, μπορούμε να πούμε ότι:

$$\lambda \downarrow \Rightarrow (1 - \lambda)^j hs \uparrow \Rightarrow \text{Διάρκεια επίδρασης } hs \uparrow$$

Συνεπώς, η τεχνική εφαρμογής head start σε διαγράμματα EWMA δεν ενδείκνυται για περιπτώσεις που η παράμετρος  $\lambda$  είναι σχετικά μικρή, καθώς αν εφαρμοστεί επηρεάζει το στατιστικό για πολλές παρατηρήσεις και η έννοια της γρήγορης αρχικής απόκρισης δεν έχει νόημα.

#### 4.2.1 Επιλογή της τιμής εκκίνησης

Μια λογική τιμή του  $hs$  είναι η τοποθέτησή του στο μισό της απόστασης μεταξύ της κεντρικής γραμμής και του ανώτερου ορίου ελέγχου (ή του κατώτερου ορίου για αρνητικές μετατοπίσεις) για την πρώτη δειγματοληψία, όπου τα όρια είναι στενότερα, Lucas και Saccucci (1990) [35]. Δηλαδή:

- Θετική τιμή προβαδίσματος:  $hs = \frac{UCL - \mu_0}{2}$

- Αρνητική τιμή προβαδίσματος:  $hs = \frac{LCL - \mu_0}{2}$

Η θετική τιμή προβαδίσματος επιταχύνει την εμφάνιση εκτός ελέγχου σήματος όταν η μετατόπιση του μέσου γίνεται προς μεγαλύτερες τιμές, ενώ η αρνητική τιμή προβαδίσματος

λειτουργεί αντίστοιχα για μετατόπιση του μέσου που γίνεται προς μικρότερες τιμές.

Η επιλογή της τιμής του  $hs$  γίνεται με γνώμονα το διάγραμμα που χρησιμοποιούμε, καθώς πρέπει να ληφθεί υπόψη ο τρόπος της χρονικής μεταβολής των ορίων ελέγχου και να υπολογιστούν τα UCL και LCL. Ορίζουμε, λοιπόν, τις ακόλουθες ποσότητες, που θα χρησιμοποιηθούν σε διάφορους τύπους διαγραμμάτων EWMA (Knoth [28], 2005):

$$c^* = L\sqrt{\frac{2}{2-\lambda}} \quad (4.5)$$

$$l_n = \sqrt{1 - (1-\lambda)^{2n}} \quad (4.6)$$

$$Z_j = \lambda \sum_{k=0}^{j-1} (1-\lambda)^k \bar{X}_{j-k} + (1-\lambda)^j \mu_0 + (1-\lambda)^j Z_0 \quad (4.7)$$

## 4.2.2 Διαγράμματα EWMA με Ασυμπτωτικά Όρια Ελέγχου και με Head Start

Σε αυτή την περίπτωση ισχύει η συνθήκη της παρακάτω σχέσης:

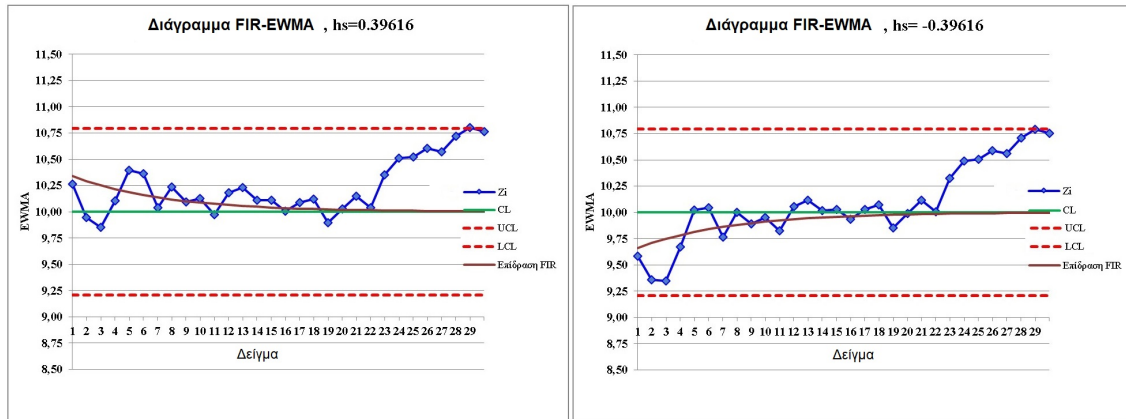
$$(fir) : \max \left\{ |Z_i^u - \mu_0|, |Z_i^l - \mu_0| \right\} > c^* \quad (4.8)$$

Οι Lucas και Saccucci (1990) [35] πρότειναν την κάτωτι τιμή προβαδίσματος για την περίπτωση των εν λόγω διαγραμμάτων ( $n = 1$ ):

$$hs = z_0^u = -z_0^l = \frac{c^*}{2} = \frac{L}{2} \sqrt{\frac{2}{2-\lambda}} \quad (4.9)$$

Εφαρμόζοντας τις Σχέσεις (4.8) και (4.9) στα δεδομένα του Πίνακα A.1 του παραρτήματος, παίρνουμε τα παρακάτω διαγράμματα ( $\lambda=0.14$ ,  $L=2.792883$ ). Το αριστερό γράφημα αφορά στην περίπτωση θετικής τιμής προβαδίσματος  $Z_0^u$  και το δεξί γράφημα αφορά στην περίπτωση αρνητικής τιμής προβαδίσματος  $Z_0^l$ . Στα διαγράμματα αναπαρίσταται με μια γραμμή η επίδραση του  $hs$ , από την οποία φαίνεται ότι η επίδρασή του είναι σημαντική μόνο κατά τις πρώτες παρατηρήσεις.

Διαπιστώνουμε ότι, στην περίπτωση των ασυμπτωτικών ορίων, η χρήση  $hs$  χρησιμοποιώντας τη Σχέση (4.9) δεν βελτιώνει στον επιθυμητό βαθμό την ανιχνευτική ικανότητα του διαγράμματος, καθώς δεν εντοπίζει την μετατόπιση του μέσου στη 2η παρατήρηση.



Σχήμα 4.5: Διαγράμματα EWMA (ασυμπτωτικά όρια) με  $hs$  για τα δεδομένα του Πίνακα Α.1

### 4.2.3 Διαγράμματα EWMA με Χρονικά Μεταβαλλόμενα Όρια Ελέγχου και με Head Start

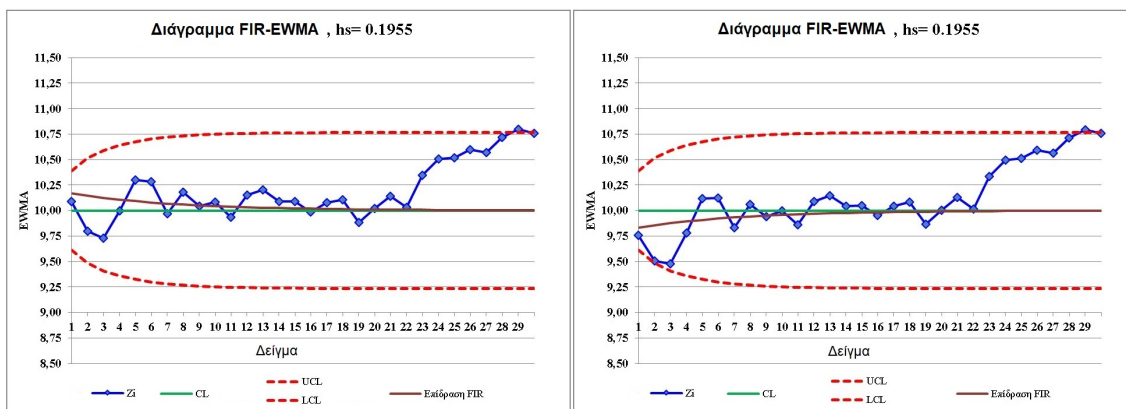
Σε αυτή την περίπτωση ισχύει η συνθήκη της Σχέσης (4.10):

$$(fvacl) : \max \left\{ |Z_n^u|, |Z_n^l| \right\} > c^* l_n \quad (4.10)$$

Για αυτά τα διαγράμματα οι Rhoads et al. (1996) [46] πρότειναν την ακόλουθη τιμή προβαδίσματος:

$$hs = z_0^u = -z_0^l = l_1 \frac{c^*}{2} = \frac{L}{2} \sqrt{\frac{2}{2-\lambda}} \sqrt{1 - (1-\lambda)^2} \quad (4.11)$$

Εφαρμόζοντας τις Σχέσεις (4.10) και (4.11) για τα δεδομένα του Πίνακα Α.1 του παραρτήματος, παίρνουμε τα παρακάτω διαγράμματα ( $\lambda=0.14$ ,  $L=2.792883$ ). Το αριστερό γράφημα αφορά στην περίπτωση θετικής τιμής προβαδίσματος  $Z_0^u$ , ενώ το δεξί γράφημα αφορά στην περίπτωση αρνητικής τιμής προβαδίσματος  $Z_0^l$ .



Σχήμα 4.6: Διαγράμματα EWMA (χρονικά μεταβαλλόμενα όρια) με  $hs$  για τα δεδομένα του Πίνακα Α.1

Παρατηρούμε ότι, στην περίπτωση αρνητικής τιμής προβαδίσματος, το διάγραμμα δίνει

2 πολύ οριακές τιμές στην 2η και 3η παρατήρηση, χωρίς να δίνεται εκτός ελέγχου σήμα. Αντίθετα, το εκτός ελέγχου σήμα δίνεται στην 29η παρατήρηση και στις δύο περιπτώσεις (θετική και αρνητική τιμή προβαδίσματος), γεγονός που είναι αναμενόμενο, καθώς στην 29η παρατήρηση η επίδραση της τιμής προβαδίσματος έχει σχεδόν μηδενιστεί και το διάγραμμα EWMA με χρονικά μεταβαλλόμενα όρια θα εντόπιζε ούτως ή άλλως την μετατόπιση του μέσου στην 29η παρατήρηση.

#### 4.2.4 Διαγράμματα FIR-EWMA με Head Start

Όπως έχουμε ήδη σχολιάσει, τα διαγράμματα FIR-EWMA αποτελούν εργαλείο FIR και ως εκ τούτου δεν συνίσταται η συνδυαστική χρήση τους με τεχνική FIR με αρχική τιμή εκκίνησης, καθώς το ίδιο αποτέλεσμα, ως επί τω πλείστω, είναι δυνατόν επιτευχθεί με κατάλληλη επιλογή παραμέτρων. Ωστόσο, θα εξεταστεί και η περίπτωση αυτή.

Η συνθήκη που πρέπει να ικανοποιείται για τα υπόψη διαγράμματα, σύμφωνα με τον Steiner (1999) [49], είναι η εξής:

$$(fadj) : |Z_n - \mu_0| > c^* l_n (1 - (1 - f)^{1+\alpha(n-1)}), \quad f = 0.5, \alpha = 0.3 \quad (4.12)$$

Εφαρμόζοντας head start στα υπόψη διαγράμματα και ακολουθώντας την ίδια λογική με εκείνη που ακολουθήσαμε στα διαγράμματα με χρονικά μεταβαλλόμενα όρια, η τιμή προβαδίσματος μπορεί να είναι η ακόλουθη:

$$hs = z_0^u = -z_0^l = l_1 \frac{c^*}{2} (1 - (1 - f)^{1+\alpha(n-1)})$$

και συνεπώς:

$$hs = \frac{L}{2} f \sqrt{\frac{2}{2-\lambda}} \sqrt{1 - (1 - \lambda)^2} \quad (4.13)$$

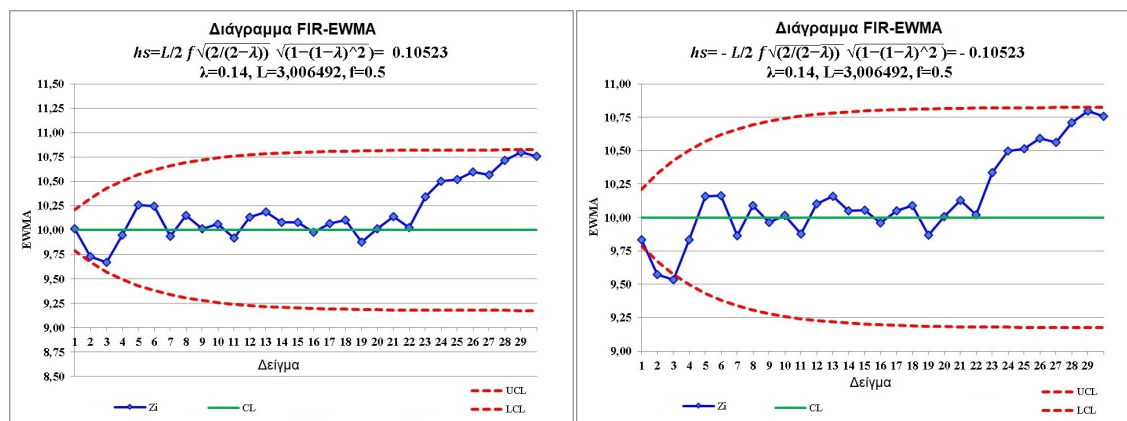
Αν εφαρμόσουμε τις Σχέσεις (4.12) και (4.13) στα δεδομένα του Πίνακα Α.1 του παρατήματος, προκύπτουν τα διαγράμματα του Σχήματος 4.7, όπου βλέπουμε ότι η μεγάλη μετατόπιση κατά την 2η παρατήρηση εντοπίζεται άμεσα από το διάγραμμα. Επίσης, παρατηρούμε ότι το διάγραμμα δεν δύναται να εντοπίσει την μετατόπιση στο μέσο μετά την 20η παρατήρηση.

Από θεωρητική πλευρά, η παραπάνω συμπεριφορά μπορεί να επιβεβαιωθεί και χρήση του στατιστικού πακέτου "spc" της R [31] (εντολές `xewma.arl` και `xewma.crit`). Για τους υπολογισμούς χρειάζεται να γίνει ο παρακάτω μετασχηματισμός:

$$hs = \frac{hs^*}{\sqrt{\frac{\lambda}{2-\lambda}}}$$

όπου  $hs^*$  είναι η τιμή εκκίνησης που επιθυμούμε.

Έτσι, μπορούμε να υπολογίσουμε την ανιχνευτική ικανότητα για μια μετατόπιση μέσου



Σχήμα 4.7: Διαγράμματα FIR-EWMA με  $hs$  για τα δεδομένα του Πίνακα Α.1

που λαμβάνει χώρα από την πρώτη παρατήρηση ( $q=1$ ) ή μετά από αρκετές παρατηρήσεις. Για παράδειγμα, για  $\mu_1=1$ , ένα διάγραμμα FIR-EWMA με  $l=0.14$ ,  $f=0.5$  και  $hs=0.73687$  δίνει  $ARL_1=2.168092$  αν η μετατόπιση γίνεται στην αρχή, ενώ δίνει  $ARL_1=10.64668$  αν η μετατόπιση γίνεται στην 20η παρατήρηση.

Στη συνέχεια, στον Πίνακα 4.4 παρουσιάζονται τιμές του ARL για διάφορες τιμές εκκίνησης ( $hs=0.5$ ,  $hs=0,73687$  και  $hs=1$ ), για μετατοπίσεις  $\mu_0=0$  ή  $\mu_1=1$  που συμβαίνουν από την πρώτη παρατήρηση ( $q=1$ ) ή από την 20η.

Από τον παραπάνω πίνακα μπορούμε να διαλέξουμε τις βέλτιστες τιμές των  $\lambda$  και  $L$  για ανίχνευση μετατόπισης  $\mu_1=1$  και, με χρήση των δεδομένων του Πίνακα Α.1, να πάρουμε τα παρακάτω διαγράμματα FIR-EWMA ( $f=0.5$ ).

Στα διαγράμματα του Σχήματος 4.8 παρατηρούμε ότι, όταν οι μετατοπίσεις συμβαίνουν κοντά στην έναρξη των μετρήσεων, η επίδραση του head start είναι σημαντική και βελτιώνει την ανιχνευτική ικανότητα του διαγράμματος.

### 4.3 Τροποποιημένο FIR-EWMA

Μία ακόμα τεχνική γρήγορης αρχικής απόκρισης, η οποία παρέχει καλύτερα αποτελέσματα, παρουσιάστηκε από τους Haq et al. (2014) [21]. Η υπόψη τεχνική βασίζεται στην ακόλουθη σχέση:

$$MFIR_{adj} = (1 - (1 - f)^{1+\alpha(j-1)})^{1+\frac{1}{j}} \quad (4.14)$$

Από τη Σχέση (4.14) προκύπτουν οι σχέσεις των ορίων ελέγχου, δηλαδή:

$$\mu_x \pm L\sigma_x (1 - (1 - f)^{1+\alpha(j-1)})^{1+\frac{1}{j}} \sqrt{\frac{\lambda[1 - (1 - \lambda)^{2j}]}{(2 - \lambda)n}} \quad (4.15)$$

Εφαρμόζοντας τις Σχέσεις (4.15) στα δεδομένα του Πίνακα Α.1, προκύπτει το διάγραμμα του Σχήματος 4.9, στο οποίο φαίνεται ότι η πρώιμη μετατόπιση στο μέσο κατά 2.1 μονάδες ανιχνεύεται αμέσως, ενώ η ύστερη μετατόπιση μετά από 20 παρατηρήσεις δεν ανιχνεύεται

Πίνακας 4.5: Αποτελέσματα ARL διαγράμματος FIR-EWMA ( $f=0.5$ ) με και χωρίς  $hs$

		$\lambda = 0.10$			$\lambda = 0.13$			
$hs^*$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$
0	2.816648	370	4.432	10.077	2.87304	370	4.570	9.931
0.5	2.904413	370	2.368	10.523	2.934248	370	2.697	10.277
0.73687	3.00371	370	1.765	11.075	3.004479	370	2.0716	10.703
1	3.151898	370	1.360	11.959	3.111002	370	1.594	11.390
		$\lambda = 0.14$			$\lambda = 0.15$			
$hs^*$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$
0	2.888008	370	4.610	<b>9.925</b>	2.901543	370	4.648	9.938
0.5	2.943088	370	2.794	10.250	2.886858	370	2.887	<b>10.243</b>
0.73687	3.006492	370	2.075	10.715	3.008921	370	2.262	10.616
1	3.103034	370	1.674	11.289	3.09682	370	1.755	11.22
		$\lambda = 0.16$			$\lambda = 0.17$			
$hs^*$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$
0	2.913843	370	4.686	9.967	2.925074	370	4.723	10.010
0.5	2.959171	370	2.975	10.256	2.966472	370	3.060	10.285
0.73687	3.011614	370	2.334	<b>10.536</b>	3.014463	370	2.334	10.620
1	3.091967	370	1.836	11.179	3.088182	370	1.917	<b>11.162</b>
		$\lambda = 0.18$			$\lambda = 0.19$			
$hs^*$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$	$L$	$q = 1$ $\mu_0 = 0$	$q = 1$ $\mu_1 = 1$	$q = 20$ $\mu_1 = 1$
0	2.93537	370	4.760	10.067	2.94484	370	4.797	10.137
0.73687	3.017393	370	2.496	10.649	3.020348	370	2.615	10.694
1	3.085239	370	1.998	11.167	3.082968	370	2.078	11.190

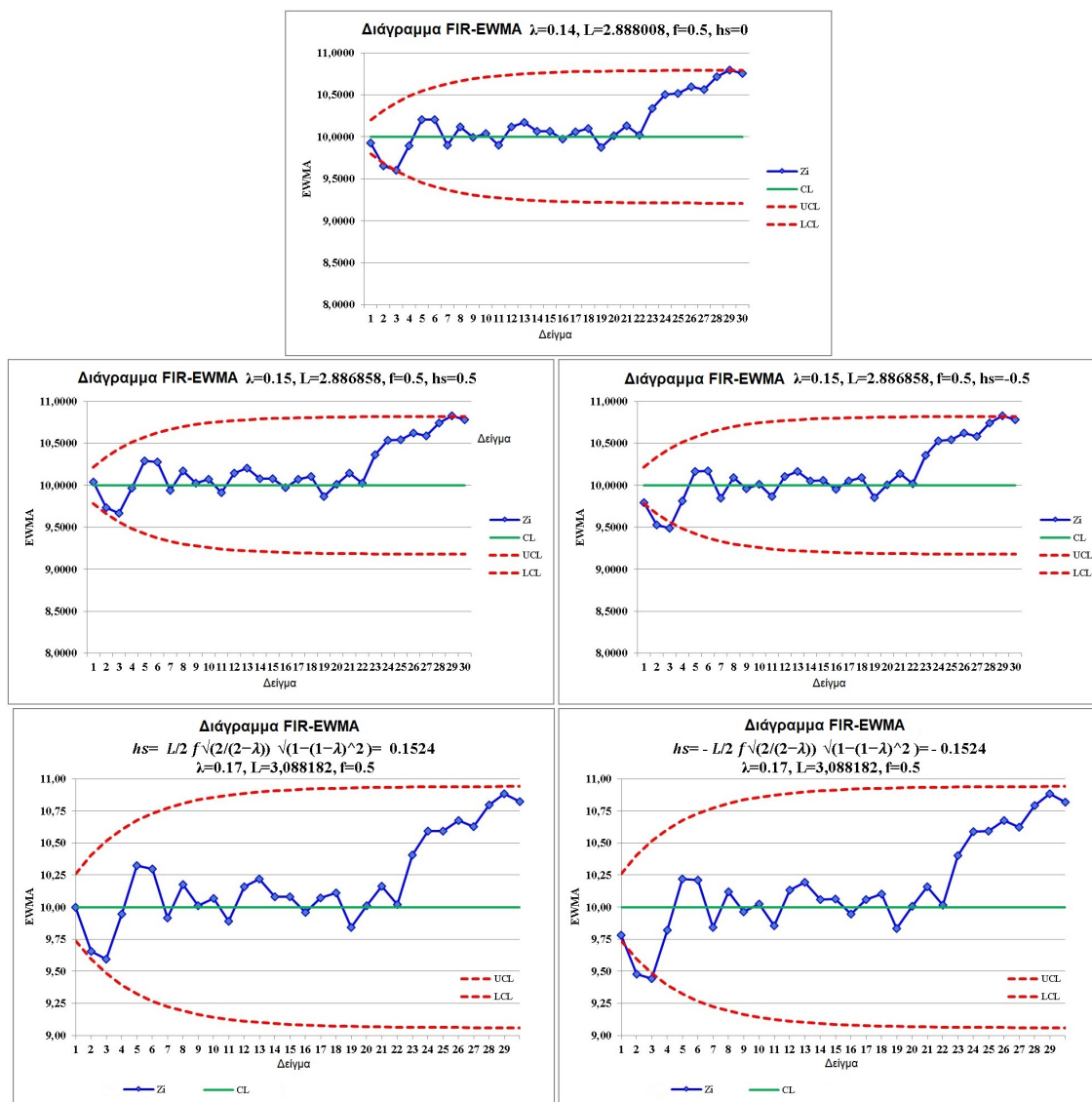
οριακά.

Με τη βοήθεια κώδικα παρόμοιου με αυτόν που χρησιμοποιείται στο πακέτο "spc" της R [31], υπολογίστηκε το θεωρητικό ARL για διάγραμμα MFIR-EWMA με τις παραμέτρους του παραπάνω παραδείγματος (για  $\lambda=0.14$ ,  $L=3.04547$ ). Έτσι, υπολογίζεται:

- για  $q=1$   $ARL_1$  (για  $q=1$ )=1.128829
- για  $q=20$  βρίσκουμε  $ARL_1$  (για  $q=20$ )=10.8394.

Υπάρχει, δηλαδή, συμφωνία της πειραματικής συμπεριφοράς του διαγράμματος με την θεωρητικά αναμενόμενη. Ισχύει, πάντως, πάλι πως ό,τι κερδίζουμε σε ανιχνευτική ικανότητα κοντά στην έναρξη των μετρήσεων, χάνεται αν οι μετατοπίσεις συμβούν σε ύστερη χρονική

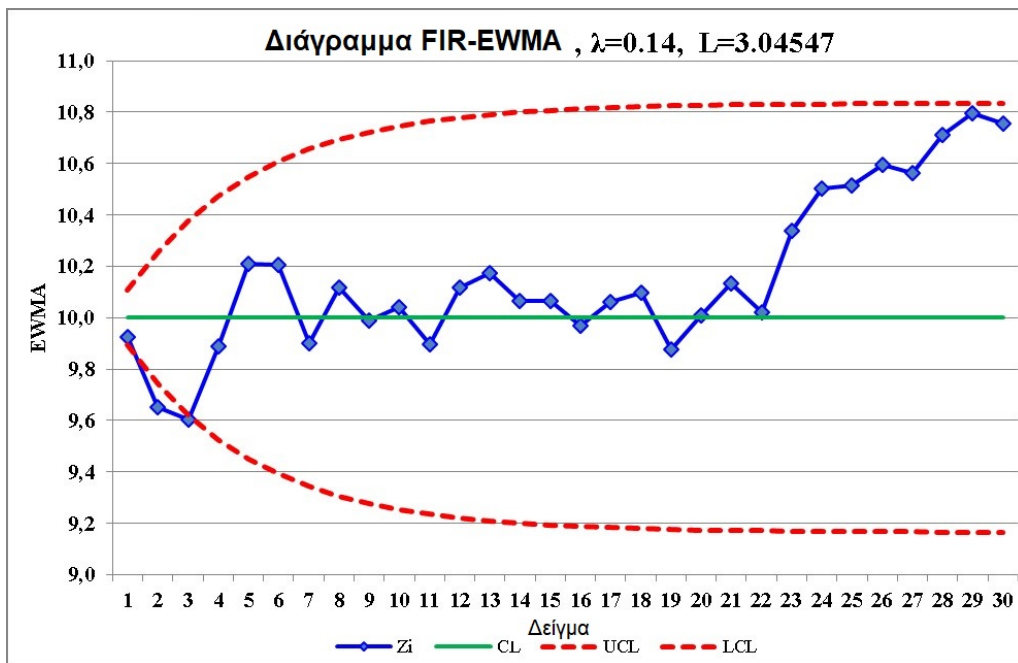




Σχήμα 4.8: Διαγράμματα FIR-EWMA με  $h_s$  για διάφορες τιμές παραμέτρων για τα δεδομένα του Πίνακα Α.1

στιγμή. Το  $ARL_1$  της MFIR-EWMA τεχνικής είναι καλύτερο για πρώιμες μετατοπίσεις, ενώ για την μετέπειτα μετατοπίσεις υστερεί συγκριτικά με την απλή FIR-EWMA τεχνική.





Σχήμα 4.9: Διάγραμμα MFIR-EWMA για τα δεδομένα του Πίνακα Α.1



# Κεφάλαιο 5

## Τροποποιημένο Διάγραμμα EWMA

### 5.1 Εισαγωγή

Είδαμε μέχρι τώρα ότι η απόδοση των απλών διαγραμμάτων EWMA βελτιώνεται με τη χρήση χρονικά μεταβαλλόμενων ορίων, αλλά και μέσω τεχνικών FIR που επιφέρουν περαιτέρω στένεμα των ορίων. Οι τεχνικές αυτές βελτιώνουν την απόδοση του  $ARL_1$ , αλλά με κόστος τη μείωση του  $ARL_0$ . Προκειμένου να διατηρήσουμε σταθερό το  $ARL_0$ , οδηγούμαστε σε μείωση της ανιχνευτικής ικανότητας, δηλαδή του  $ARL_1$ , για εντοπισμό μετατοπίσεων σε ύστερες χρονικές στιγμές. Δηλαδή, τα όρια είναι πιο στενά στις πρώτες παρατηρήσεις και πιο πλατιά στις ύστερες. Ένα ακόμη βασικό μειονέκτημα της μεθόδου EWMA είναι η αδυναμία γρήγορου εντοπισμού μεγάλων και απότομων μετατοπίσεων.

### 5.2 Στατιστικό Τροποποιημένου EWMA

Ένα διάγραμμα που διατηρεί τα επιθυμητά χαρακτηριστικά των διαγραμμάτων EWMA και, ταυτόχρονα, εμφανίζει ευαισθησία σε μεγάλες και απότομες αλλαγές, παρουσιάστηκε από τους Patel και Divecha (2011) [43]. Συγκεκριμένα, πρότειναν το εξής στατιστικό:

$$Z_j = \lambda X_j + (1 - \lambda)Z_{j-1} + X_j - X_{j-1} \quad (5.1)$$

όπου  $j = 1, 2, \dots, m$  είναι ο αύξων αριθμός των δειγματοληψιών,  $\lambda$  είναι η σταθερά εξομάλυνσης με  $0 < \lambda \leq 1$  και η τιμή εκκίνησης είναι  $X_0 = \mu_0 = Z_0$ .

Η Σχέση (5.1) είναι αναδρομικός τύπος. Συνεπώς, μπορεί να αναπτυχθεί αντικαθιστώντας κάθε φορά την τιμή του  $Z_j$  στο δεύτερο μέλος της εξίσωσης, καταλήγοντας στην

ακόλουθη εναλλακτική μορφή του στατιστικού του Τροποποιημένου EWMA:

$$\begin{aligned}
 Z_j &= \lambda X_j + (1 - \lambda)Z_{j-1} + A_j = \\
 &= \lambda X_j + (1 - \lambda) [\lambda X_{j-1} + (1 - \lambda)Z_{j-2} + A_{j-1}] + A_j = \\
 &= \lambda X_j + \lambda(1 - \lambda)X_{j-1} + (1 - \lambda)^2 Z_{j-2} + (1 - \lambda)A_{j-1} + A_j = \\
 &= \lambda X_j + \lambda(1 - \lambda)X_{j-1} + (1 - \lambda)^2 [\lambda X_{j-2} + (1 - \lambda)Z_{j-3} + A_{j-2}] \\
 &\quad + (1 - \lambda)A_{j-1} + A_j = \\
 &= \lambda X_j + \lambda(1 - \lambda)X_{j-1} + \lambda(1 - \lambda)^2 X_{j-2} + (1 - \lambda)^3 Z_{j-3} + (1 - \lambda)^2 A_{j-2} \\
 &\quad + (1 - \lambda)^1 A_{j-1} + (1 - \lambda)^0 A_j = \dots
 \end{aligned}$$

όπου  $A_j = X_j - X_{j-1}$ , δηλαδή η διαφορά μεταξύ της τιμής τελευταίου και προ-τελευταίου δείγματος, που αποτελεί ένδειξη του πόσο απότομα εμφανίζεται η μετατόπιση στη διεργασία.

Συνεχίζοντας, είναι προφανές ότι μπορούμε να εκφράσουμε το στατιστικό  $Z_j$  ως εξής:

$$Z_j = \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k X_{j-k} + \sum_{k=0}^{j-1} (1 - \lambda)^k A_{(j-k)} + (1 - \lambda)^j Z_0 \quad (5.2)$$

Συνεπώς, καταλήγουμε σε σχέση ανάλογη της Σχέσης (3.3), με τον επιπλέον όρο που αφορά στην ποσότητα μέτρησης των απότομων αλλαγών του μέσου:

$$\sum_{k=0}^{j-1} (1 - \lambda)^k A_{(j-k)}$$

Παρατηρώντας τη Σχέση (5.2), διαπιστώνουμε ότι το στατιστικό απαρτίζεται από τους τρεις παρακάτω όρους:

α.  $\sum_{k=0}^{j-1} (1 - \lambda)^k \lambda \bar{X}_{(j-k)}$ , ο οποίος αντιπροσωπεύει τις προγενέστερες παρατηρήσεις, στις οποίες δίνεται συνεχώς μειούμενη σημασία. Έμφαση δίνεται στις πιο πρόσφατες παρατηρήσεις.

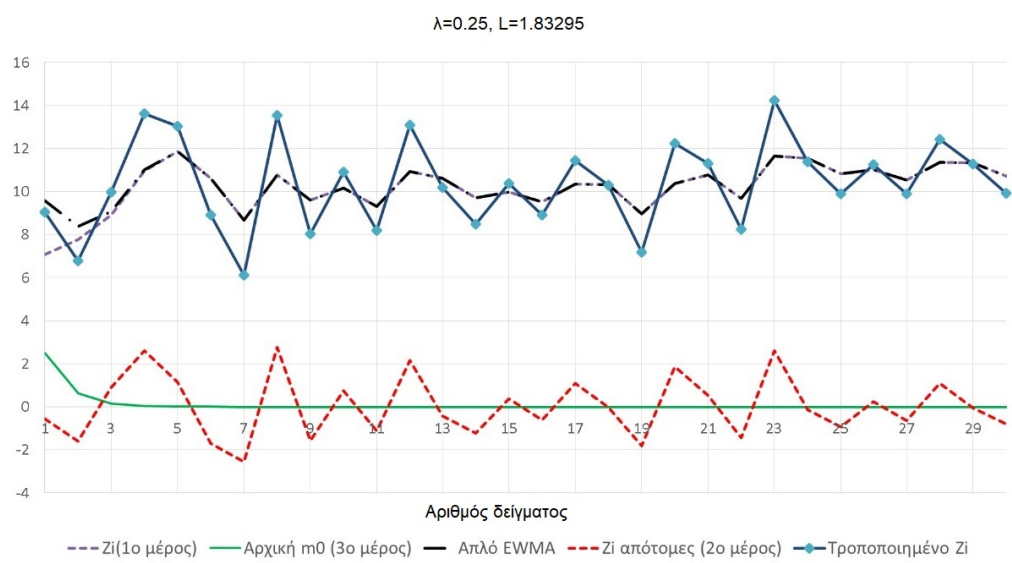
β.  $\sum_{k=0}^{j-1} (1 - \lambda)^k A_{(j-k)}$ , ο οποίος αντιπροσωπεύει τις απότομες μεταβολές, καθώς η τιμή του είναι σημαντική μόνον όταν παρατηρούνται μεγάλες μεταβολές σε κοντινές παρατηρήσεις. Έμφαση δίνεται στις πιο πρόσφατες απότομες μεταβολές και συνεχώς μειούμενη σημασία στις προγενέστερες.

γ.  $(1 - \lambda)^j Z_0$ , ο οποίος αντιπροσωπεύει την υπόθεση ότι ο μέσος είναι ίσος με την τιμή στόχο. Η συμμετοχή του στην τιμή του στατιστικού μειώνεται καθώς αυξάνεται ο αριθμός των δειγμάτων. Στο παράδειγμα του Πίνακα 5.1 (για μια τιμή  $\lambda=0.25$ ), μετά από 16 παρατηρήσεις καταλήγει σε ποσοστό 1% περίπου της αρχικής του τιμής.

Ο 1ος και ο 3ος όρος δίνουν την τιμή του απλού στατιστικού EWMA. Ο 1ος όρος έχει συνεχώς αυξανόμενη τιμή και ο 3ος συνεχώς μειούμενη τιμή, ώστε στην εντός ελέγχου περίπτωση να διατηρείται σταθερό το στατιστικό. Μία ένδειξη του τρόπου με τον οποίο μεταβάλλονται δίνεται στα Σχήματα 5.1 και 5.2.

Η συμπεριφορά όλων των όρων του τροποποιημένου στατιστικού EWMA, για τα δεδομένα του Πίνακα Α.1, παρατίθεται στον Πίνακα 5.1.

Με βάσει τα δεδομένα αυτά, προκύπτει το διάγραμμα του Σχήματος 5.1, το οποίο προέκυψε για  $\lambda=0.25$  και  $L=1.83295$ .



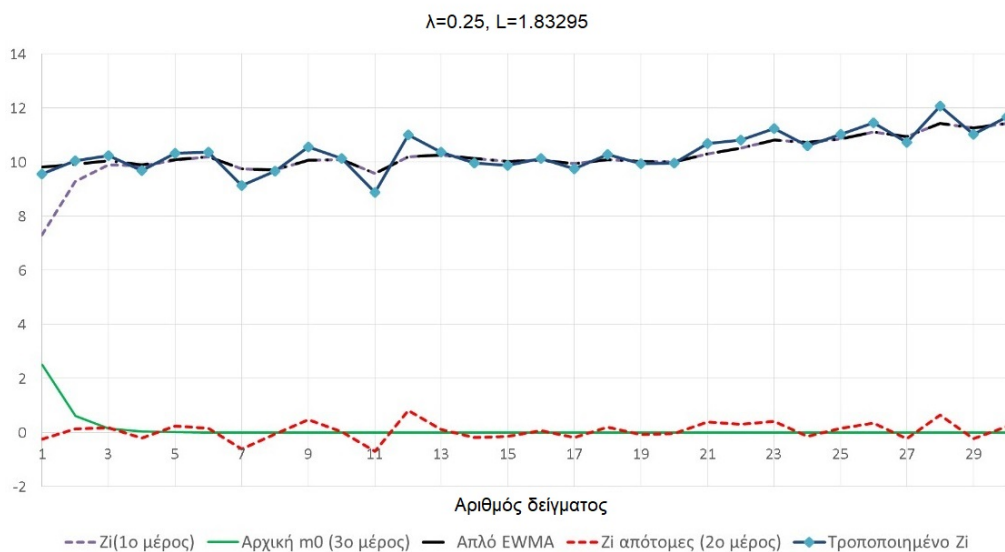
Σχήμα 5.1: Διάγραμμα στατιστικού τροποποιημένου EWMA για τα δεδομένα του Πίνακα Α.1

Από το διάγραμμα του Σχήματος 5.1 διαπιστώνουμε ότι η τιμή του στατιστικού διαμορφώνεται κατά κύριο λόγο από την τιμή του 2ου Μέρους, που μετράει τις απότομες μετατοπίσεις. Μετά την 20η παρατήρηση παρατηρείται μια τάση αύξησης της τιμής του στατιστικού, λόγω της μετατόπισης του μέσου του στατιστικού κατά μία μονάδα.

Καθώς, όμως, οι απότομες μετατοπίσεις είναι αυτές που κατά κύριο λόγο οδηγούν σε εκτός ελέγχου τιμές του στατιστικού, θα χρησιμοποιήσουμε δεδομένα που στηρίζονται στην ίδια λογική με αυτά του Πίνακα Α.1 αλλά προέρχονται από κατανομή με μικρότερη διασπορά τιμών (δεν οδηγούμαστε σε εκτός ελέγχου σήματα αποκλειστικά από τις απότομες μετατοπίσεις του μέσου). Στα δεδομένα αυτά, που παρατίθενται στον Πίνακα Α.5, στις πρώτες 20 παρατηρήσεις ο μέσος του στατιστικού παραμένει σταθερός αλλά με μικρή διασπορά τιμών (σε σχέση με τα δεδομένα του Πίνακα Α.1), ενώ μετά την 20η παρατήρηση αρχίζει η μετατόπιση του μέσου κατά μία μονάδα μέχρι την 30η παρατήρηση.

Η συμπεριφορά όλων των μερών του τροποποιημένου στατιστικού EWMA για τα εν λόγω δεδομένα φαίνεται στον Πίνακα 5.2.

Εφαρμόζοντας στα υπόψη δεδομένα το στατιστικό του τροποποιημένου διαγράμματος EWMA (πάλι για τιμές  $\lambda=0.25$  και  $L=1.83295$ ), παίρνουμε το διάγραμμα του Σχήματος 5.2 και τα βοηθητικά διαγράμματα του Σχήματος 5.3.



Σχήμα 5.2: Διάγραμμα στατιστικού τροποποιημένου EWMA για τα δεδομένα του Πίνακα Α.5



Σχήμα 5.3: Βοηθητικά διαγράμματα στο Σχήμα 5.2: 2ο μέρος (αριστερά) και 1ο-3ο μέρος (δεξιά) του στατιστικού του τροποποιημένου EWMA

Παρατηρώντας τα παραπάνω διαγράμματα συμπεραίνουμε ότι:

α. Το σχήμα του διαγράμματος του στατιστικού του τροποποιημένου EWMA ακολουθεί το 2ο μέρος του στατιστικού, που αφορά στις απότομες αλλαγές.

β. Η μέση τιμή του στατιστικού επηρεάζεται από το 1ο και το 3ο μέρος του στατιστικού του τροποποιημένου EWMA, που αντιπροσωπεύει το απλό EWMA, η οποία δίνει συνεχώς μεγαλύτερες τιμές και οδηγεί στην ανίχνευση της μετατόπισης του μέσου (μετά την 20η παρατήρηση).

γ. Το 2ο μέρος του στατιστικού του τροποποιημένου EWMA δίνει μεγάλες τιμές όταν κάποια παρατήρηση δώσει τιμή μετατόπισης απότομα μεγάλη. Σε τέτοια περίπτωση, ο όρος αυτός θα δώσει εκτός ελέγχου σήμα πριν τη συμπλήρωση της μέγιστης βαθμιαίας μετατόπισης από το 1ο και 3ο μέρος του στατιστικού.

### 5.3 Διάγραμμα Τροποποιημένου EWMA

Ξεκινώντας από τη Σχέση (5.1), αποδεικνύεται ότι ο μέσος και η διακύμανση του στατιστικού  $Z_j$  του τροποποιημένου EWMA δίνονται από τις παρακάτω σχέσεις:

$$E[Z_j] = \mu \quad (5.3)$$

και

$$Var[Z_j] = \left[ \frac{\lambda}{2-\lambda} + \frac{2\lambda(1-\lambda)}{2-\lambda} \right] \sigma^2 \quad (5.4)$$

Συνεπώς, τα άνω και κάτω όρια ελέγχου των τροποποιημένων διαγραμμάτων EWMA υπολογίζονται από τις σχέσεις:

$$UCL = \mu_0 + L\sigma \sqrt{\frac{\lambda}{2-\lambda} + \frac{2\lambda(1-\lambda)}{2-\lambda}} \quad (5.5)$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{\lambda}{2-\lambda} + \frac{2\lambda(1-\lambda)}{2-\lambda}} \quad (5.6)$$

όπου  $\mu_0$  είναι ο μέσος,  $\sigma^2$  είναι η διακύμανση,  $\lambda$  και  $L$  είναι οι σταθερές του Τροποποιημένου EWMA. Το διάγραμμα μετατρέπεται σε διάγραμμα Shewhart για διεργασίες υψηλής αυτοσυσχέτισης, για τιμές των παραμέτρων ίσες με  $\lambda=1$ ,  $L=3$ .

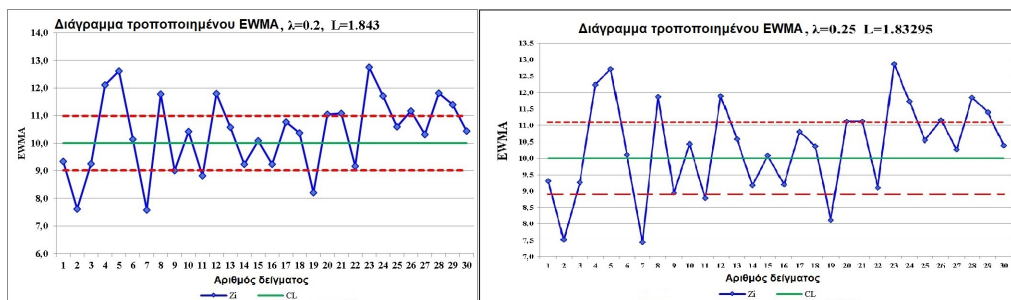
Το υπόψη τροποποιημένο διάγραμμα EWMA είναι κατάλληλο για εντοπισμό ενδείξεων μικρών μετατοπίσεων στο μέσο μιας διεργασίας, όπως επίσης και για ανίχνευση απότομων αλλαγών σε μια αυτοσυσχετιζόμενη διαδικασία. Η διεργασία είναι εντός ελέγχου, εφόσον οι τιμές σχεδιάζονται εντός των ορίων ελέγχου. Σε περίπτωση που ένα σημείο απεικονιστεί εκτός των ορίων ελέγχου, αυτό αποτελεί ένδειξη ότι η διεργασία βρίσκεται εκτός ελέγχου και απαιτούνται ενέργειες για τον εντοπισμό των αιτιών που είναι υπεύθυνες για αυτή τη συμπεριφορά.

### 5.4 Εφαρμογή Τροποποιημένου EWMA

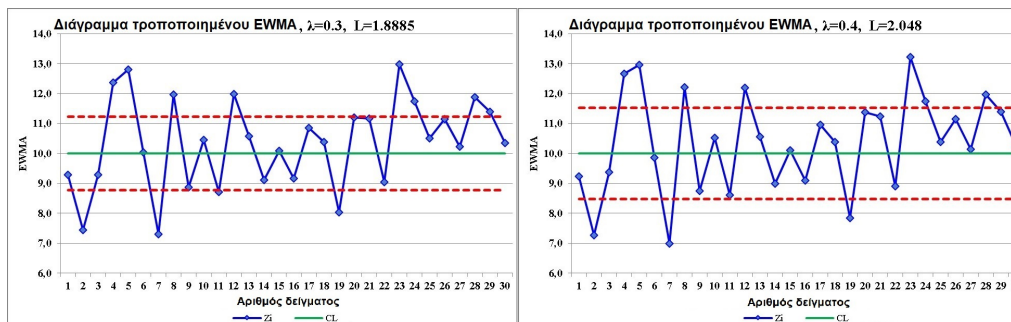
Αν στις παραπάνω Σχέσεις (5.4), (5.5) και (5.6) εφαρμοστούν τα δεδομένα του Πίνακα A.1 και του Πίνακα A.5, θα προκύψουν τα διαγράμματα που παρατίθενται παρακάτω. Οι τιμές που επιλέχθηκαν για τις παραμέτρους είναι τέτοιες ώστε να προκύπτει  $ARL_0=370$ .

Παρατηρώντας τα παραπάνω διαγράμματα, διαπιστώνουμε ότι εντοπίζεται άμεσα κάθε μεγάλη και απότομη μετατόπιση, όπως φαίνεται από τα Σχήματα 5.4, 5.5 και 5.6, ενώ παράλληλα για μικρές τιμές του  $\lambda$  διατηρείται η ικανότητα ανίχνευσης των μικρών και διαδοχικών μετατοπίσεων, όπως φαίνεται από τα Σχήματα 5.7, 5.8 και 5.9.

Έτσι, για τα δεδομένα του Πίνακα A.1 η μετατόπιση κατά 2.1 μονάδες στην 2η παρατήρηση εντοπίζεται αμέσως. Όπως επίσης οι μετατοπίσεις κατά 1.66 και 2.16 στην 4η και 5η παρατήρηση προς τις θετικές τιμές του στατιστικού εντοπίζονται και πάλι αμέσως. Η ευελιξία



Σχήμα 5.4: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.1



Σχήμα 5.5: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.1

του υπόψη διαγράμματος φαίνεται ακόμη περισσότερο στις παρατηρήσεις 7 και 8, στις οποίες παρουσιάζεται η εξής ιδιαιτερότητα: οι τιμές είναι εκτός των ζητούμενων ορίων και είναι προς την αντίθετη κατεύθυνση μεταξύ τους, γεγονός που είναι πολύ δύσκολο να γίνει αντιληπτό με τα απλά διαγράμματα EWMA. Ωστόσο, με το υπόψη διάγραμμα εντοπίζονται άμεσα οι αντίστοιχες μετατοπίσεις του μέσου.

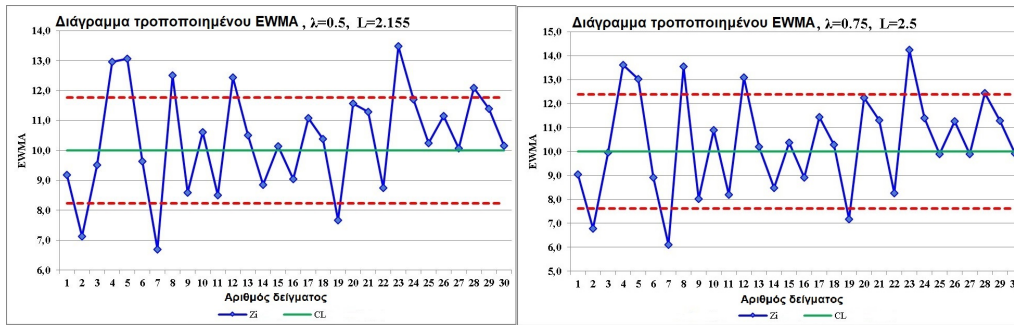
Επίσης, παρατηρούμε ότι για μικρές τιμές της παραμέτρου εξομάλυνσης  $\lambda$  διατηρείται το πλεονέκτημα της μνήμης των προγενέστερων παρατηρήσεων, όπως φαίνεται από τα γραφήματα στα Σχήματα 5.4 και 5.6, ενώ για τις μεγάλες τιμές του  $\lambda$  το στατιστικό μηδενίζεται για μικρές αλλαγές στο μέσο, δηλαδή σβήνονται πολύ εύκολα από τη μνήμη του, όπως φαίνεται στο τελευταίο γράφημα, στα Σχήματα 5.6 και 5.9 για τις παρατηρήσεις 25, 27 και 30.

Τα ανωτέρω αποτελέσματα δείχνουν την ανωτερότητα των τροποποιημένων διαγραμμάτων EWMA έναντι των απλών, καθώς εντοπίζουν πολύ γρήγορα τη μετατόπιση του μέσου σε κάθε περίπτωση.

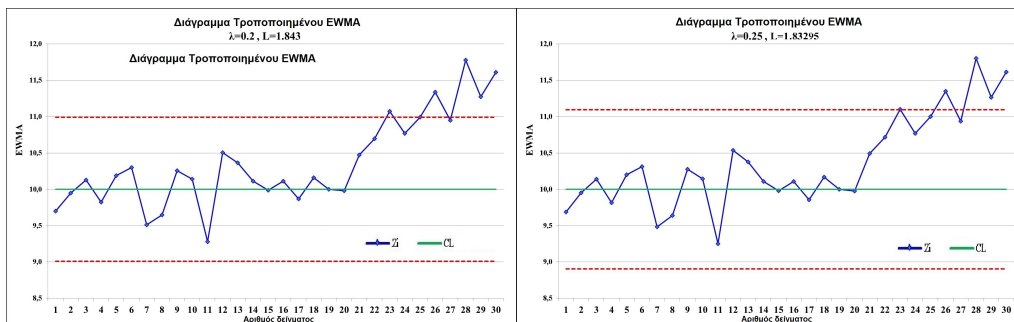
## 5.5 Υπολογισμός ARL τροποποιημένου EWMA

Το επόμενο βήμα είναι να προσδιοριστεί η ανιχνευτική ικανότητα των υπόψη διαγραμμάτων, υπολογίζοντας τις θεωρητικά αναμενόμενες τιμές του  $ARL_1$  για τις διάφορες τιμές μετατόπισης. Οι υπολογισμοί πραγματοποιήθηκαν με τη χρήση του κώδικα του Παραρτήματος (μέθοδος μαρκοβιανής αλυσίδας με βήμα  $g = 30$ ), επιλέγοντας για όρια UCL και LCL αυτά που δίνονται από τις Σχέσεις (5.5) και (5.6).





Σχήμα 5.6: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.1



Σχήμα 5.7: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.5

Επίσης, ως μέσος και διασπορά, για τον υπολογισμό της μετάβασης από την κατάσταση  $k$  με τιμή  $S[k]$  στην κατάσταση  $l$  με τιμή  $S[l]$ , χρησιμοποιήθηκαν οι ακόλουθες σχέσεις, Patel και Divecha (2011) [43]:

$$Mean_{S[l]} = \lambda\mu_1 + (1 - \lambda)S[k] + \mu_1 - \mu_0 \quad (5.7)$$

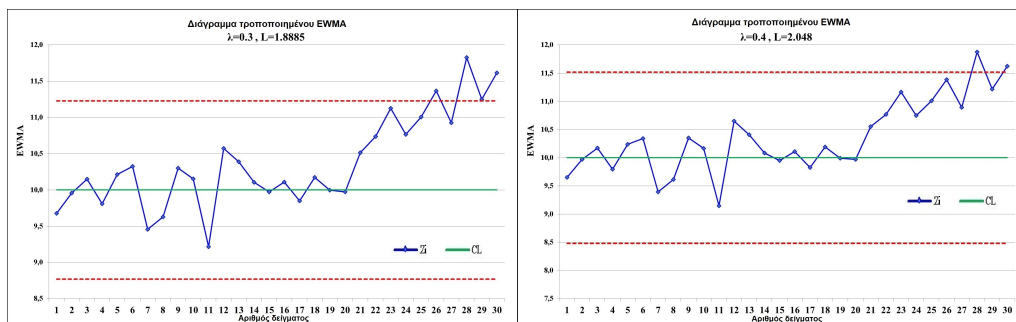
$$Standard\ Deviation = \lambda\sigma \quad (5.8)$$

Στους Πίνακες 5.3 και 5.4 παρατίθενται οι υπολογισμένες με την υπόψη μέθοδο τιμές ARL, για τιμές παραμέτρων  $\lambda$  και  $L$  ώστε ανά περίπτωση να προκύπτει  $ARL_0=370$ .

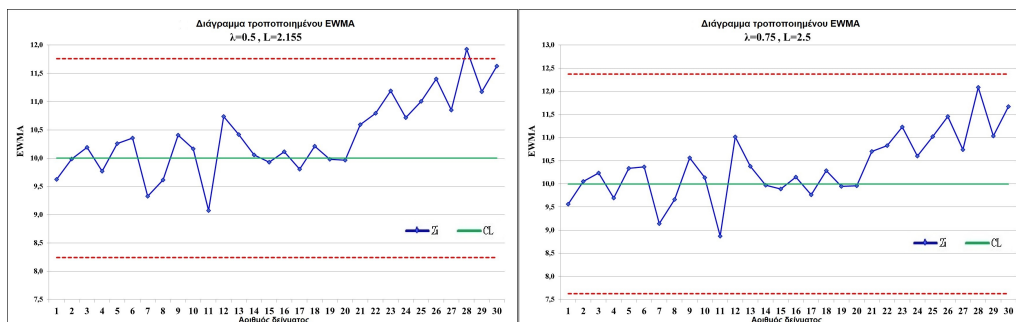
Από τους Πίνακες 5.3 και 5.4 και τα γραφήματα των Σχημάτων 5.4 έως 5.9, διαπιστώνουμε ότι η θεωρητική ανιχνευτική ικανότητα των διαγραμμάτων συμφωνεί με την αντίστοιχη που προκύπτει από την εφαρμογή του Τροποποιημένου EWMA για τα δεδομένα των Πινάκων Α.1 και Α.5 του Παραρτήματος Α.

Ας θεωρήσουμε την περίπτωση με παραμέτρους  $\lambda=0.25$  και  $L=1.83295$  που απεικονίζεται στο Σχήμα 5.8. Η μετατόπιση κατά 0.4 στην 21η παρατήρηση μπορεί να ανιχνευτεί μετά από 2.966963 παρατηρήσεις (υπολογίζεται με τον κώδικα του παραρτήματος). Αυτό σημαίνει ότι μετρώντας ως 1η την 21η παρατήρηση, εφόσον διατηρηθεί η υπόψη μετατόπιση, μπορεί να ανιχνευτεί στην 23η παρατήρηση.

Η μετατόπιση του μέσου στην 22η παρατήρηση διατηρείται, καθώς  $\delta=0.6>0.4$ . Για να



Σχήμα 5.8: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.5



Σχήμα 5.9: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.5

ανιχνευτεί η νέα μετατόπιση απαιτούνται 1.691226 παρατηρήσεις, γεγονός που σημαίνει ότι εφόσον διατηρηθεί η μετατόπιση, θα ανιχνευτεί στην 23η παρατήρηση, γεγονός που επαληθεύεται, καθώς στην 23η παρατήρηση έχουμε εκτός ελέγχου σήμα στο δεξί διάγραμμα του Σχήματος 5.7.

Στην 23η παρατήρηση διατηρείται η μετατόπιση του  $\delta$  που πρωτοανιχνεύτηκε στην 21η παρατήρηση, συνεπώς ικανοποιούνται οι ικανές και αναγκαίες συνθήκες ώστε να δοθεί εκτός ελέγχου σήμα.

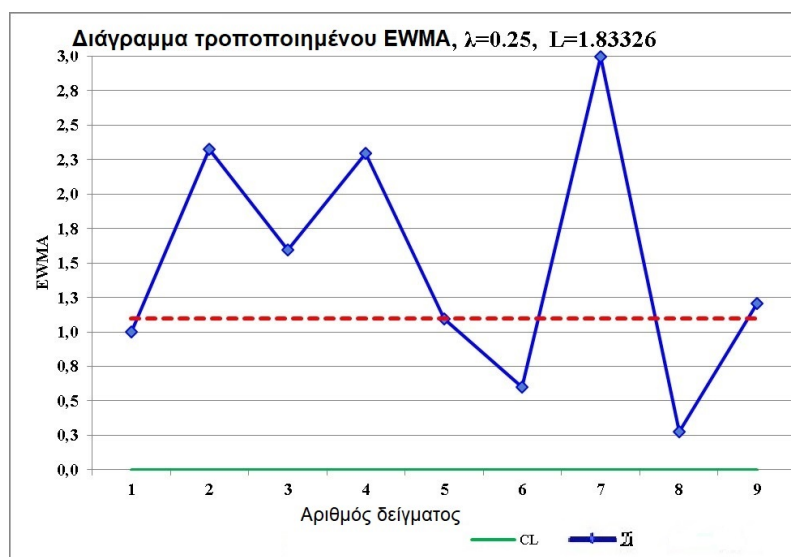
$$\delta=0.93 > 0.6 \text{ (22ης παρατήρησης)} > 0.4 \text{ (21ης παρατήρησης)}$$

Αντίστοιχα, για τα δεδομένα του Πίνακα Α.1, η μετατόπιση του μέσου κατά 0.55 στην 1η παρατήρηση δεν ανιχνεύεται άμεσα, καθώς όπως φαίνεται και από τις θεωρητικές τιμές για την εν λόγω περίπτωση των τιμών των παραμέτρων θα χρειαζόταν περίπου 2 παρατηρήσεις για να ανιχνευτεί η υπόψη μετατόπιση.

Η μεγάλη μετατόπιση στη 2η παρατήρηση ανιχνεύεται άμεσα, λόγω της απότομης αλλαγής στην τιμή του μέσου που προκαλεί. Το ίδιο ισχύει και για τις παρατηρήσεις 4, 5, 7 και 8, οι οποίες δίνουν άμεσα εκτός ελέγχου σήμα κυρίως λόγω της απότομης μετατόπισης που προκαλούν στο μέσο, καθώς όλες είναι μεγαλύτερες από  $\delta=0.88$ . Παρατηρούμε επίσης ότι για τις υπόψη παρατηρήσεις η διακύμανση είναι πολύ μεγάλη οδηγώντας στην πλειονότητα των περιπτώσεων σε εκτός ελέγχου σήμα.

Ας θεωρήσουμε, επίσης, μια άλλη περίπτωση δεδομένων, αυτή του Πίνακα Α.4 (πρωτοεμφανιζόμενη εκτός ελέγχου κατάσταση). Όπως παρατηρούμε στο διάγραμμα του Σχήματος

5.10, στην 1η παρατήρηση δεν δίνεται εκτός ελέγχου σήμα. Αυτό είναι σύμφωνο με την αντίστοιχη θεωρητική τιμή, η οποία δείχνει ότι θα χρειαζόταν ακόμη μία παρατήρηση με ίδια μετατόπιση για να δώσει εκτός ελέγχου σήμα.



Σχήμα 5.10: Τροποποιημένο EWMA για τα δεδομένα του Πίνακα Α.4

Ενδιαφέρον παρουσιάζει η 5η παρατήρηση, στην οποία παίρνουμε οριακά εκτός ελέγχου σήμα, ενώ η μετατόπιση είναι ίση με 1.1 δηλαδή μεγαλύτερη από την ανιχνευτική ικανότητα των υπόψη διαγραμμάτων και λογικά θα έπρεπε να ανιχνεύεται άμεσα. Αυτό οφείλεται στο γεγονός ότι ενώ έχουμε πάρει τρία εκτός ελέγχου σήματα στην 2η, 3η και 4η παρατήρηση, εντούτοις συνεχίσαμε τις δειγματοληψίες, οδηγώντας έτσι στο υπόψη σφάλμα. Με άλλα λόγια, όταν υιοθετούμε μια εκτός ελέγχου διεργασία ως κανονικότητα, τότε το διάγραμμα δυσκολεύεται να αντιληφθεί τις μετατοπίσεις που σχεδιάστηκε να ανιχνεύει.

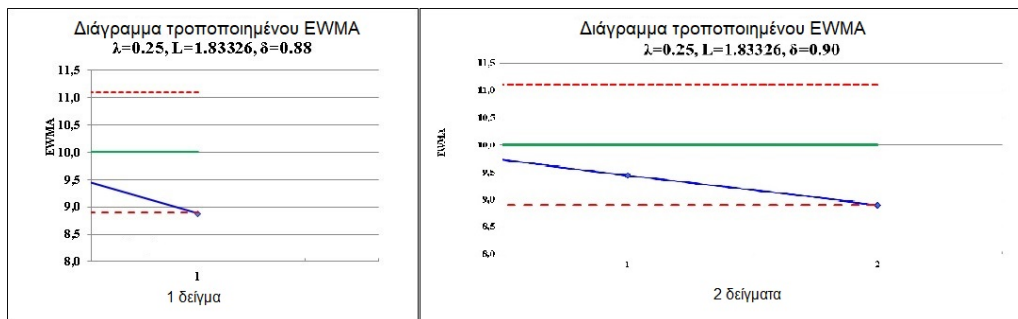
## 5.6 Συμπέρασμα

Τα διαγράμματα τροποποιημένου EWMA εμφανίζονται το ίδιο ικανά στο να ανιχνεύουν μεγάλες και απότομες αλλαγές μετατοπίσεις, μεγαλύτερες από μια σταθερή τιμή, η οποία υπολογίστηκε και δίνεται στους παραπάνω πίνακες για τις διάφορες τιμές του  $\lambda$ , όπως επίσης και μικρές και βαθμιαίες μετατοπίσεις.

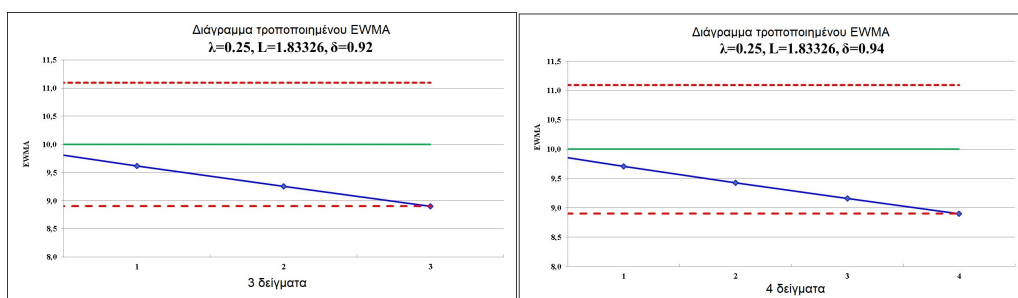
Επίσης, όσες περισσότερες παρατηρήσεις απαιτούνται προκειμένου να μετατοπιστεί ο μέσος κατά ένα συγκεκριμένο μέγεθος, τόσο πιο πολύ μειώνεται η ανιχνευτική ικανότητα των υπόψη διαγραμμάτων. Αυτό είναι λογικό και συναντάται και στα υπόλοιπα διαγράμματα EWMA.

Έτσι, για παράδειγμα, για την περίπτωση που  $\lambda=0.25$  και  $L=1.83295$ , το διάγραμμα έχει  $ARL_0=370$  και μπορεί να εντοπίσει άμεσα απότομες μετατοπίσεις με  $\delta \geq 0.88$  από το μέσο. Αν υποθέσουμε ότι αντί για μία παρατήρηση είχαμε περισσότερες εντός των οποίων συμπληρώνεται η αντίστοιχη μετατόπιση, τότε η ανιχνευτική ικανότητα του διαγράμματος ελαττώνεται

και δεν μπορεί να ανιχνεύσει άμεσα την ίδια μετατόπιση που πριν ήταν εφικτό με μικρότερο πλήθος παρατηρήσεων. Επίσης, διαπιστώνουμε ότι θα χρειαζόταν να παρουσιαστεί μεγαλύτερη μετατόπιση για να καταστεί εφικτό να ανιχνευτεί άμεσα ή μεγαλύτερο πλήθος παρατηρήσεων. Αυτό φαίνεται ξεκάθαρα στα παρακάτω θεωρητικά παραδείγματα στα οποία αναπαρίσταται το πλήθος των παρατηρήσεων που θα χρειαζόταν για να ανιχνευτεί άμεσα η μετατόπιση που αναγράφεται κάθε φορά, επιμερίζοντας τη συνολική μετατόπιση σε κάθε μία παρατήρηση. Έτσι, αν η μετατόπιση  $\delta=0.88$  συμπληρωθεί στην 2η παρατήρηση, τότε δεν γίνεται άμεσα αντιληπτή. Θα χρειαζόταν λίγες ακόμη παρατηρήσεις ή αντί για  $\delta=0.88$  η μετατόπιση να ήταν μεγαλύτερη και συγκεκριμένα μεγαλύτερη ίση με  $\delta=0.90$ . Έτσι, αντίστοιχα για μετατοπίσεις που συμπληρώνονται στην 3η, 4η, 5η παρατήρηση κτλ, τότε η μετατόπιση που είναι εφικτό να ανιχνευτεί άμεσα είναι ίση, όπως φαίνεται από τα διαγράμματα, με  $\delta=0.92$ ,  $\delta=0.94$ ,  $\delta=0.96$  κτλ.

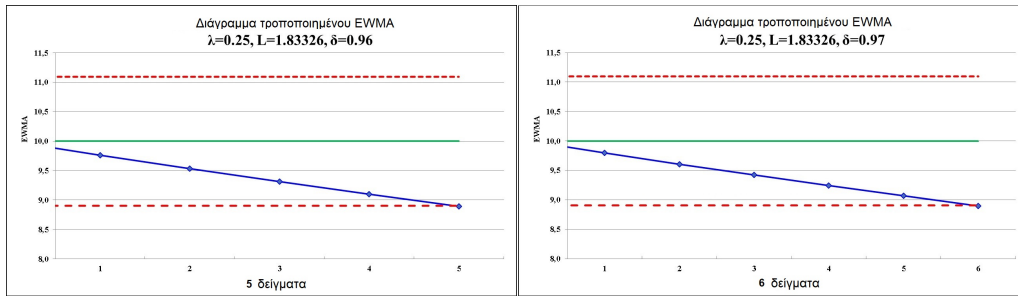


Σχήμα 5.11: Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 1 (αριστερά) και 2 παρατηρήσεις (δεξιά)

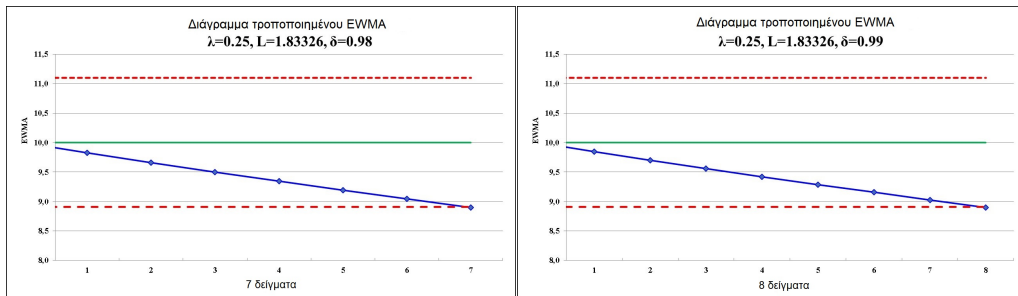


Σχήμα 5.12: Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 3 (αριστερά) και 4 παρατηρήσεις (δεξιά)

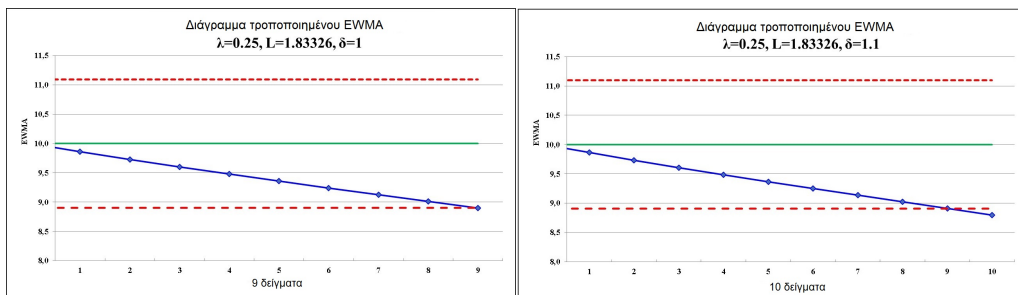
Συμπερασματικά, το τροποποιημένο διάγραμμα EWMA αποτελεί ιδανικό εργαλείο για την ανίχνευση μικρών και βαθμιαίων μετατοπίσεων αλλά και μεγάλων και απότομων. Ωστόσο, πρέπει να λαμβάνουμε υπόψη την ιδιαιτερότητά του να αδυνατεί να ανιχνεύσει ένα εύρος μετατοπίσεων, στην περίπτωση που δεν ληφθεί υπόψη το πρώτο εκτός ελέγχου σήμα.



Σχήμα 5.13: Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 5 (αριστερά) και 6 παρατηρήσεις (δεξιά)



Σχήμα 5.14: Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 7 (αριστερά) και 8 παρατηρήσεις (δεξιά)



Σχήμα 5.15: Τροποποιημένο EWMA για βαθμιαία μετατόπιση του μέσου σε 9 (αριστερά) και 10 παρατηρήσεις (δεξιά)

Πίνακας 5.1: Ανάλυση του στατιστικού τροποποιημένου EWMA στα μέρη του για τα δεδομένα του Πίνακα Α.1

A/A	$X_j$	$Z_j$ (1ο μέρος)	Αρχική $\mu_0$ (3ο μέρος)	Απλό $EWMA$	$Z_j$ απότομα (2ο μέρος)	$Z_j$ τροποποιημένο
	10	0	10	0	10	0
1	9.45	2.363	7.500	9.863	-0.550	9.313
2	7.99	3.769	5.625	9.394	-1.873	7.522
3	9.29	6.183	4.219	10.402	-0.104	9.264
4	11.66	6.777	3.164	9.941	2.292	12.233
5	12.16	8.123	2.373	10.496	2.219	12.715
6	10.18	8.637	1.780	10.417	-0.316	10.101
7	8.04	8.488	1.335	9.823	-2.377	7.446
8	11.46	9.231	1.001	10.232	1.637	11.869
9	9.20	9.223	0.751	9.974	-1.032	8.942
10	10.34	9.502	0.563	10.066	0.366	10.431
11	9.03	9.384	0.422	9.807	-1.036	8.771
12	11.47	9.906	0.317	10.222	1.663	11.886
13	10.51	10.057	0.238	10.294	0.288	10.582
14	9.40	9.893	0.178	10.071	-0.894	9.176
15	10.08	9.939	0.134	10.073	0.009	10.082
16	9.37	9.797	0.100	9.897	-0.703	9.194
17	10.62	10.003	0.075	10.078	0.723	10.801
18	10.31	10.080	0.056	10.136	0.232	10.368
19	8.52	9.690	0.042	9.732	-1.616	8.116
20	10.84	9.977	0.032	10.009	1.108	11.117
21	10.90	10.208	0.024	10.232	0.891	11.123
22	9.33	9.988	0.018	10.006	-0.902	9.105
23	12.29	10.564	0.013	10.577	2.284	12.861
24	11.50	10.798	0.010	10.808	0.923	11.731
25	10.60	10.748	0.008	10.756	-0.208	10.548
26	11.08	10.831	0.006	10.837	0.324	11.161
27	10.38	10.718	0.004	10.723	-0.457	10.266
28	11.62	10.944	0.003	10.947	0.897	11.844
29	11.31	11.035	0.002	11.038	0.363	11.401
30	10.52	10.907	0.002	10.908	-0.518	10.391

Πίνακας 5.2: Ανάλυση του στατιστικού τροποποιημένου EWMA στα μέρη του για τα δεδομένα του Πίνακα Α.5

A/A	$X_j$	$Z_j$ (1ο μέρος)	Αρχική $\mu_0$ (3ο μέρος)	Απλό <i>EWMA</i>	$Z_j$ απότομα (2ο μέρος)	$Z_j$ τροποποιημένο
	10	0	10	10	0	10
1	9.75	2.363	7.500	9.938	-0.250	9.688
2	9.95	3.769	5.625	9.941	0.012	9.953
3	10.1	6.183	4.219	11.047	0.159	10.140
4	9.85	6.777	3.164	9.948	-0.130	9.817
5	10.15	8.123	2.373	9.998	0.202	10.201
6	10.25	8.637	1.780	10.061	0.252	10.313
7	9.6	8.488	1.335	9.946	-0.461	9.485
8	9.7	9.231	1.001	9.884	-0.246	9.639
9	10.2	9.223	0.751	9.963	0.316	10.279
10	10.11	9.502	0.563	10.000	0.147	10.147
11	9.4	9.384	0.422	9.850	-0.600	9.250
12	10.4	9.906	0.317	9.988	0.550	10.537
13	10.3	10.057	0.238	10.066	0.312	10.378
14	10.1	9.893	0.178	10.074	0.034	10.109
15	10	9.939	0.134	10.056	-0.074	9.981
16	10.1	9.797	0.100	10.067	0.044	10.111
17	9.9	10.003	0.075	10.025	-0.167	9.858
18	10.14	10.080	0.056	10.054	0.115	10.169
19	10.01	9.690	0.042	10.043	-0.044	9.999
20	9.99	9.977	0.032	10.030	-0.053	9.977
21	10.4	10.208	0.024	10.122	0.370	10.493
22	10.6	9.988	0.018	10.242	0.478	10.719
23	10.93	10.564	0.013	10.414	0.688	11.102
24	10.7	10.798	0.010	10.485	0.286	10.772
25	10.9	10.748	0.008	10.589	0.415	11.004
26	11.2	10.831	0.006	10.742	0.611	11.353
27	10.9	10.718	0.004	10.781	0.158	10.940
28	11.6	10.944	0.003	10.986	0.819	11.805
29	11.22	11.042	0.002	11.044	0.234	11.279
30	11.55	11.169	0.002	11.171	0.506	11.676

Πίνακας 5.3: Ανιχνευτική ικανότητα διαγράμματος τροποποιημένου EWMA για διάφορες τιμές των παραμέτρων

Τροποποιημένο Διάγραμμα EWMA με $ARL_0=370$					
		$L$	1.4619	1.61427	1.83295
		$\lambda$	0.05	0.1	0.25
Μέση Μετατόπιση	Απότομη μετατόπιση άμεσα ανιχνεύσιμη		$\delta \geq 0.40$	$\delta \geq 0.62$	$\delta \geq 0.88$
$\delta=0$	-		370	370	370
$\delta=0.25$	-		1.547	2.507	6.418
$\delta=0.5$	-		Άμεσα	1.258	2.152
$\delta=0.75$	-		Άμεσα	Άμεσα	1.280
$\delta=1.0$	-		Άμεσα	Άμεσα	Άμεσα
$\delta=1.25$	-		Άμεσα	Άμεσα	Άμεσα
$\delta=1.5$	-		Άμεσα	Άμεσα	Άμεσα
$\delta=1.75$	-		Άμεσα	Άμεσα	Άμεσα
$\delta=2$	-		Άμεσα	Άμεσα	Άμεσα
$\delta=2.1$	-		Άμεσα	Άμεσα	Άμεσα

Πίνακας 5.4: Ανιχνευτική ικανότητα διαγράμματος τροποποιημένου EWMA για διάφορες τιμές των παραμέτρων

Τροποποιημένο Διάγραμμα EWMA με $ARL_0=370$						
		$L$	1.8882	1.99514	2.1059	2.4471
		$\lambda$	0.3	0.4	0.5	0.75
Μέση Μετατόπιση	Απότομη μετατόπιση άμεσα ανιχνεύσιμη		$\delta \geq 0.95$	$\delta \geq 1.06$	$\delta \geq 1.15$	$\delta \geq 1.33$
$\delta=0$	-		370	370	370	370
$\delta=0.25$	-		8.810	16.699	29.727	83.691
$\delta=0.5$	-		2.548	3.657	5.491	16.560
$\delta=0.75$	-		1.429	1.791	2.308	5.282
$\delta=1.0$	-		Άμεσα	1.129	1.353	2.408
$\delta=1.25$	-		Άμεσα	Άμεσα	Άμεσα	1.376
$\delta=1.5$	-		Άμεσα	Άμεσα	Άμεσα	0.907
$\delta=1.75$	-		Άμεσα	Άμεσα	Άμεσα	Άμεσα
$\delta=2$	-		Άμεσα	Άμεσα	Άμεσα	Άμεσα
$\delta=2.1$	-		Άμεσα	Άμεσα	Άμεσα	Άμεσα



# Κεφάλαιο 6

## Συνδυαστικά Διαγράμματα EWMA

### 6.1 Εισαγωγή

Είδαμε ότι επεμβαίνοντας στη μαθηματική δομή του στατιστικού κατέστη δυνατό να βελτιωθεί η συνολική απόδοση των διαγραμμάτων EWMA, τόσο όταν έχουμε μικρές και βαθμιαίες μετατοπίσεις του μέσου, όσο και όταν έχουμε απότομες μεταβολές του.

Ένας τρόπος μέτρησης των απότομων αλλαγών στο μέσο μιας διαδικασίας είναι ο προσδιορισμός της διασποράς των παρατηρήσεων. Ένα διάγραμμα EWMA που σχεδιάστηκε για τον σκοπό αυτό είναι το διάγραμμα διασποράς EWMA, το οποίο παρουσιάζεται στη συνέχεια, και μέσω του οποίου θα οδηγηθούμε στα συνδυαστικά διαγράμματα μέσου – διασποράς EWMA.

### 6.2 Διαγράμματα Διασποράς EWMA

Ο μαθηματικός ορισμός του στατιστικού του διαγράμματος Διασποράς EWMA δίνεται από την ακόλουθη σχέση:

$$S_j^2 = \lambda(x_j - \mu) + (1 - \lambda)S_{j-1}^2 \quad (6.1)$$

παίρνει αρχικές τιμές:

$$S_0^2 = \sigma^2 \quad (6.2)$$

και έχει μέση τιμή:

$$E[S_j^2] = \sigma^2 \quad (6.3)$$

όπου  $j=1,2, \dots, m$  ο αριθμός των δειγμάτων (υποομάδων), αποτελούμενων από  $i=1,2, \dots, n$  παρατηρήσεις.

Σημειώνεται ότι η ποσότητα  $S_j^2/\sigma^2$  ακολουθεί την κατανομή  $X_\nu^2$ , δηλαδή:

$$\frac{S_j^2}{\sigma^2} \sim X_{\frac{2-\lambda}{\lambda}}^2 = X_\nu^2, \text{ όπου } \nu = \frac{2-\lambda}{\lambda} \quad (6.4)$$

Η χρησιμότητα των υπόψη διαγραμμάτων είναι πολύ σημαντική, ιδιαίτερα όταν συνδυάζονται με άλλα διαγράμματα που στοχεύουν στην παρακολούθηση του μέσου.

## 6.3 Συνδυαστικά διαγράμματα EWMA Μέσου – Διασποράς

### 6.3.1 Διαγράμματα Ελέγχου Μέσου – Διασποράς EWMA

Εστω  $X_{ij}$  είναι μία ακολουθία από υποομάδες (δείγματα) ανεξάρτητων και κανονικά κατανομημένων δεδομένων. Κάθε υποομάδα  $i$  αποτελείται από  $n$  παρατηρήσεις,  $X_{i1}, \dots, X_{in}$ . Επίσης, υποθέτουμε ότι  $n > 1$ , οπότε παρακάτω θα αναφερόμαστε στη συνήθη δειγματική διακύμανση  $S^2$ .

Συνήθως, υποθέτοντας κανονικότητα δεδομένων, μας ενδιαφέρει η ανίχνευση μετατοπίσεων και προς τις δύο κατευθύνσεις του μέσου  $\mu$  καθώς και αυξήσεις της διασποράς  $\sigma^2$ . Ωστόσο, βασισμένοι στους Morais και Pacheco (2006) [38], παρακάτω θα επιχειρήσουμε να ανιχνεύσουμε και μειώσεις της διασποράς, προκειμένου να ανιχνεύονται ελλείποντα σήματα. Οπότε, η ακόλουθη τεχνική συνδυάζει δίπλευρα διαγράμματα μέσου με μονόπλευρα προς τα πάνω και δίπλευρα διαγράμματα διακύμανσης αντίστοιχα. Οι τιμές των παραμέτρων, οι οποίες περιγράφουν τη σταθερή απόδοση της υπό παρακολούθηση διεργασίας, είναι οι  $\mu_0$  και  $\sigma_0^2$ , και όσο συμπίπτουν με τις πραγματικές τιμές των παραμέτρων, η διαδικασία θεωρείται εντός ελέγχου.

### 6.3.2 Στατιστικό για Διαγράμματα Μέσου – Διασποράς EWMA

Τα διαγράμματα  $\bar{X} - S^2$  EWMA έχουν τα ακόλουθα στατιστικά για τις δύο ακολουθίες  $Z_j^\mu$  και  $Z_j^\sigma$ , Knoth (2007) [29]:

$$Z_j^\mu = (1 - \lambda_\mu)Z_{j-1}^\mu + \lambda_\mu \bar{X}_j \quad (6.5)$$

$$Z_j^\sigma = (1 - \lambda_\sigma)Z_{j-1}^\sigma + \lambda_\sigma S_j^2 \quad (6.6)$$

όπου

$$\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij} \quad (6.7)$$

$$S_j^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \quad (6.8)$$

$$Z_0^\mu = z_0^\mu \quad (6.9)$$

$$Z_0^\sigma = z_0^\sigma \quad (6.10)$$

Οι παράμετροι σχεδίασης  $\lambda_\mu$  και  $\lambda_\sigma$  επιλέγονται από το διάστημα  $(0,1]$  για να διασφαλίσουμε μια συγκεκριμένη συμπεριφορά σήματος, όπου μικρές τιμές αυξάνουν την ανιχνευτική ισχύ για μικρές αλλαγές και αντίστροφα. Οι τιμές εκκίνησης  $Z_0^\mu$  και  $Z_0^\sigma$  συνήθως επιλέγονται οι  $\mu_0$  και  $\sigma_0^2$  αντίστοιχα. Τότε, εφόσον η διεργασία είναι εντός ελέγχου, οι αναμενόμενες τιμές και για τις δύο EWMA διαδικασίες είναι ίσες με τις τιμές στόχους τους.

### 6.3.3 Διαγράμματα Ελέγχου Μέσου - Διασποράς EWMA

Τα όρια του διαγράμματος μέσου EWMA έχουν οριστεί ήδη από τις Σχέσεις (3.12), (3.13) και (3.14). Μένει λοιπόν να οριστούν τα όρια του διαγράμματος διασποράς EWMA. Κατά αντιστοιχία λοιπόν με τα απλά διαγράμματα EWMA για την περίπτωση των δίπλευρων διαγραμμάτων ελέγχου, τα άνω και κάτω όρια του διαγράμματος διασποράς δίνονται από τις ακόλουθες σχέσεις, όπως ορίστηκαν από Knoth (2005) [27]:

$$UCL = \sigma_0^2 + L_\sigma \sqrt{\frac{\lambda_\sigma}{2 - \lambda_\sigma}} \sqrt{\frac{2}{n-1}} \sigma_0^2 \quad (6.11)$$

$$CL = \sigma_0^2 \quad (6.12)$$

$$LCL = \sigma_0^2 - L_\sigma \sqrt{\frac{\lambda_\sigma}{2 - \lambda_\sigma}} \sqrt{\frac{2}{n-1}} \sigma_0^2 \quad (6.13)$$

Πολύ συχνά το κάτω όριο ορίζεται ίσο με 0, καθώς στην περίπτωση που δεν ενδιαφερόμαστε για την ανίχνευση ελλειπόντων σημάτων, ο ορισμός του σε κάποια τιμή διαφορετική του 0 δεν έχει κάποια άλλη πρακτική αξία.

### 6.3.4 ARL Διαγραμμάτων Μέσου – Διασποράς EWMA

Το συνδυαστικό διάγραμμα μέσου - διασποράς ανιχνεύει εκτός ελέγχου σήμα όταν ένα από τα απλά διαγράμματα (μέσου ή διασποράς δηλαδή) δίνει εκτός ελέγχου σήμα. Συνεπώς, το συνολικό ARL ισούται με:

$$ARL = \min\{ARL_\mu, ARL_\sigma\} \quad (6.14)$$

$$ARL_{\mu} = \min \left\{ j \in N : |Z_j^{\mu} - \mu_0| > L_{\mu}^* = L_{\mu} \sqrt{\frac{\lambda_{\mu}}{2 - \lambda_{\mu}} \frac{\sigma_0}{\sqrt{n}}} \right\} \quad (6.15)$$

$$ARL_{\sigma} = \min \left\{ j \in N : |Z_j^{\sigma} - \sigma_0| > L_{\sigma}^* = L_{\sigma} \sqrt{\frac{\lambda_{\sigma}}{2 - \lambda_{\sigma}}} \sqrt{\frac{2}{n-1}} \sigma_0^2 \right\} \quad (6.16)$$

Τα όρια σήματος στη δεξιά πλευρά των  $ARL_{\mu}$  και  $ARL_{\sigma}$  είναι κανονικοποιημένα με την ασυμπτωτική τυπική απόκλιση του στατιστικού διαγράμματος  $Z_j^{\mu}$  και  $Z_j^{\sigma}$ , αντίστοιχα. Οι τιμές  $c_{\mu}$  και  $c_{\sigma}$  ονομάζονται κρίσιμες τιμές. Για τις δίπλευρες υποθέσεις διακύμανσης πρέπει να εισάγουμε ένα περαιτέρω όριο εξαιτίας της μη συμμετρικής συμπεριφοράς του  $S_j^2$  και συνεπακόλουθα  $Z_j^{\sigma}$ .

Το προαναφερθέν ARL θα μπορούσε να περιγραφεί ως η αναμενόμενη τιμή του ARL. Για να είμαστε ακριβείς, για σταθερές τιμές των αποκαλούμενων τιμών εκκίνησης  $z_0^{\mu}$  και  $z_0^{\sigma}$  και για δοσμένες πραγματικές τιμές των  $\mu$  και  $\sigma$ , οι οποίες υποθέτουμε ότι είναι ζωτικές για τη συνολική διαδικασία υπολογισμού του μέσου όρου, θα ισχύει:

$$ARL = E[ARL] = E[L | \mu, \sigma, z_0^{\mu}, z_0^{\sigma}] \quad (6.17)$$

Φυσικά, όπως με την περίπτωση αμφιδιακύμανσης της μαρκοβιανής αλυσίδας όπως και με την προσέγγιση ολοκληρωτικών εξισώσεων θα μπορούσε να εφαρμοστεί για το πλήρες διδι-άστατο μοντέλο. Και οι δύο περιπτώσεις καταλήγουν σε ένα γραμμικό σύστημα εξισώσεων μεγάλων διαστάσεων, εάν χρειάζεται μια λογική ακρίβεια.

Η τεχνική υπολογισμού του ARL που ακολουθεί, αποτελεί προσέγγιση που προτάθηκε από τον Knoth (2007) [29], ο οποίος εφάρμοσε ιδέες των Waldmann (1986) [51] και Gan (1995) [19].

Πρώτα, οι παράμετροι  $\mu_0$  και  $\sigma_0^2$  τίθενται στο 0 και στο 1, αντίστοιχα, για λόγους απλότητας. Συμβολίζουμε ως  $P(\cdot)$  τη μέτρηση της πιθανότητας για τις σειρές μας για ένα σταθερό σύνολο παραμέτρων  $\mu, \sigma, z_0^{\mu}$  και  $z_0^{\sigma}$ . Τότε, λαμβάνοντας υπόψη την ανεξαρτησία των  $\bar{X}$  και του  $S^2$ , για οποιοδήποτε  $j \in N$  ισχύει:

$$P(L > j) = P(L_{\mu} > j)P(L_{\sigma} > j) \quad (6.18)$$

Χρησιμοποιώντας την σχέση:

$$E(L) = \sum_{j=0}^{\infty} P(L > j) \quad (6.19)$$

και τη γεωμετρική συμπεριφορά ουράς της κατανομής  $L$  μας επιτρέπουν να καθορίσουμε

τα ακόλουθα όρια για το ARL:

$$ARL_J^\pm = \sum_{j=0}^{J-1} P(L > j) + \frac{P(L > J)}{1 - m_J^\pm} \quad (6.20)$$

Η επιλογή του σημείου μετάβασης  $J$  στηρίζεται στην απόσταση μεταξύ των ορίων  $ARL_J^\pm$  και θα μπορούσε να είναι σχετικά μικρή για μικρές αποστάσεις. Οι σταθερές  $m_J^\pm$  ορίστηκαν από τους Waldmann (1986) και Gan (1995), συμβολίζουν τα *sup* και *inf* του ακόλουθου κλάσματος:

$$m_J^+ = \sup \left\{ \frac{P(L > j; z_0^\mu, z_0^\sigma)}{P(L > j-1; z_0^\mu, z_0^\sigma)} \right\} \quad (6.21)$$

$$m_J^- = \inf \left\{ \frac{P(L > j; z_0^\mu, z_0^\sigma)}{P(L > j-1; z_0^\mu, z_0^\sigma)} \right\} \quad (6.22)$$

επί των επιτρεπόμενων πεδίων ορισμού για τις τιμές εκκίνησης  $z_0^\mu$  και  $z_0^\sigma$ . Οι τιμές  $m_J^\pm$  συγχλίνουν γρήγορα σε ένα κοινό όριο. Ο Waldmann (1986) [51] και αργότερα ο Gan (1995) [19], χρησιμοποίησαν κανόνες τετραγωνισμού, προκειμένου να υπολογίσουν αριθμητικά τις τιμές  $P(L > j)$  για  $j = 1, 2, \dots, J$  με τη βοήθεια του ακόλουθου επαναληπτικού κανόνα για τα απλά διαγράμματα ελέγχου:

$$P(L_\mu > 1)(z_0^\mu) = \int_{-c_\mu^*}^{c_\mu^*} f(z_0^\mu, z) dz \quad (6.23)$$

$$P(L_\mu > 1)(z_0^\mu) = \int_{-c_\mu^*}^{c_\mu^*} P(L_\mu > j-1)(z) f_\mu(z_0^\mu, z) dz \quad (6.24)$$

Ως εκ τούτου, το  $f(\cdot)$  συμβολίζει την σχετική μετάβαση Kernel η οποία για το διάγραμμα μέσου είναι ίση με:

$$f_\mu(z_0^\mu, z) = \frac{1}{\lambda_\mu u} \Phi \left( \frac{z - (1 - \lambda_\mu)z_0^\mu}{\lambda_\mu} \right) \quad (6.25)$$

όπου το  $\Phi(\cdot)$  συμβολίζει την PDF της κανονικά κατανεμημένης μεταβλητής για το μέσο  $\mu$  και τη διασπορά  $\sigma^2$ . Για το διάγραμμα  $S^2$  πρέπει να χρησιμοποιήσουμε την παρακάτω σχέση (και φυσικά διαφορετικά όρια στο ολοκλήρωμα):

$$f_\sigma(z_0^\sigma, z) = \frac{1}{\lambda_\sigma} x^2 \left( \frac{n-1}{\sigma^2} \left[ \frac{z - (1 - \lambda_\sigma)z_0^\sigma}{\lambda_\sigma} \right] \right) \frac{n-1}{\sigma^2} \quad (6.26)$$

όπου το  $x^2(\cdot)$  δηλώνει την pdf της  $x^2$  τυχαίας κατανεμημένης μεταβλητής με  $n-1$  βαθμούς ελευθερίας. Ο Gan (1995) [19] θεώρησε το  $\ln S^2$  και κατέληξε σε μια διαφορετική έκδοση της μετάβασης Kernel του διαγράμματος διακύμανσης και προσέγγισε τα ολοκληρώματα και στις δύο ακολουθίες επανάληψης  $P(L_\mu > j)$  και  $P(L_\sigma > j)$  με Gauss-Legendre quadratures.

Στους Knoth και Schmid (2002) [32] και οι δύο επαναληπτικές διαδικασίες προσεγγίστηκαν με όρους που προέρχονται από τη μέθοδο της μαρκοβιανής αλυσίδας. Αντί να εφαρμόζουμε την προσέγγιση σε ένα συγκεκριμένο  $J$ , το συνολικό πρόβλημα μετασχηματίζεται σε ένα

σύστημα εξισώσεων πίνακα Sylvester, το οποίο επιλύεται εφαρμόζοντας μια αριθμητική βιβλιοθήκη. Δυστυχώς, η ακρίβεια χάνει επηρεαζόμενη από το κομμάτι  $S^2$  και συνεπακόλουθα, χρειάζονται μεγάλες διαστάσεις στον πίνακα για να πάρουμε μεγαλύτερη ακρίβεια.

### 6.3.5 Ανιχνευτική ικανότητα Διαγραμμάτων Μέσου - Διασποράς EWMA

Ο υπολογισμός του ARL των διαγραμμάτων EWMA μέσου - διασποράς μπορεί να γίνει με τη βοήθεια της βιβλιοθήκης "spe" της R [31], τρέχοντας την εντολή:

```
xsewma.arl(lx=0.14, cx=2.78465, ls=0.14, csu=2.78465, csl=0.3, df=4, mu=2, sigma=0.3, sided="two", Nx=60, Ns=60)
```

Έτσι, μπορούμε να υπολογίσουμε τη θεωρητική συμπεριφορά των υπόψη διαγραμμάτων για τις περιπτώσεις:

- α. Δίπλευρης ανίχνευσης ως προς μέσο και διασπορά (*sided = "two"*) και
- β. Δίπλευρης ανίχνευσης ως προς μέσο και μονόπλευρης ως προς διασπορά (*sided = "upper"*).

Οι τιμές του ζεύγους των παραμέτρων προσδιορίστηκαν έτσι ώστε να έχουμε  $ARL_0=370$ , για διασπορά  $\sigma_0^2=1$ . Τα αποτελέσματα δίνονται στους Πίνακες 6.1 και 6.2.

Πίνακας 6.1: Ανιχνευτική Ικανότητα Μονόπλευρου Διαγράμματος Μέσου - Διασποράς EWMA

$S^2$	$\lambda_\mu = \lambda_\sigma = 0.14$		$L_\mu = L_\sigma = 2.78465$		$LCL_\sigma = 0.3$			
	0.25	0.5	0.75	1	1.25	1.5	1.75	2
$\delta$								
0.00	24523646638	44783849	7333.56	370	90.38	24.01	9.94	6.02
0.25	39369471	16673	369.9	102.33	51.38	21.07	9.63	5.94
0.5	46162.6	120.5	20.06	30.97	23.84	15.68	8.82	5.73
0.75	25.84	20.06	17.09	15.14	13.69	11.36	7.77	5.42
1	10.24	10.16	9.9	9.56	9.23	8.51	6.72	5.04
2	3.76	3.74	3.81	3.86	3.92	3.95	3.87	3.56
2.5	3	2.97	2.99	3.03	3.08	3.12	3.12	3.00

Από τους Πίνακες 6.1 και 6.2 φαίνεται ότι για το ζεύγος παραμέτρων που επιλέχθηκε, δηλαδή για  $\lambda_\mu = \lambda_\sigma = 0.14$  και για  $L_\mu = L_\sigma = 2.78465$ , η ανιχνευτική ικανότητα των διαγραμμάτων βελτιώνεται σημαντικά για τις διάφορες μετατοπίσεις. Έτσι, για παράδειγμα για μετατόπιση  $\delta=1$  και διασπορά  $S^2=1.5$  η μετατόπιση ανιχνεύεται σε 8.51 παρατηρήσεις, έναντι των 9.56 παρατηρήσεων που απαιτούνται στην απλή περίπτωση, στην οποία  $S^2=1$ . Παρατηρούμε ότι όσο αυξάνεται η διασπορά, τόσο πιο γρήγορα δύναται να ανιχνευτεί η μετατόπιση από το εν λόγω διάγραμμα. Έτσι, για  $S^2=2$ , απαιτούνται 5.04 παρατηρήσεις για να ανιχνευτεί η ίδια μετατόπιση ( $\delta=1$ ).

Πίνακας 6.2: Ανιχνευτική Ικανότητα Δίπλευρου Διαγράμματος Μέσου – Διασποράς EWMA

$S^2$	$\lambda_\mu = \lambda_\sigma = 0.14$		$L_\mu = L_\sigma = 2.78465$		$LCL_\sigma = 0.3$			
	0.25	0.5	0.75	1	1.25	1.5	1.75	2
$\delta$								
0.00	9.6	17.72	1483.32	370	90.38	24.01	9.94	6.02
0.25	9.6	17.72	313.09	102.33	51.38	21.07	9.63	5.94
0.5	9.6	17.53	46.22	30.97	23.84	15.68	8.82	5.73
0.75	9.6	14.88	17.08	15.14	13.69	11.36	7.77	5.42
1	9.6	9.98	9.9	9.56	9.23	8.51	6.72	5.04
2	9.6	3.74	3.81	3.86	3.92	3.95	3.87	3.56
2.5	9.6	2.97	2.99	3.03	3.08	3.12	3.12	3.00

Στην περίπτωση των μονόπλευρων και για όσο η διασπορά είναι μικρότερη της αναμενόμενης τιμής, δηλαδή  $\sigma_0^2 < 1$ , παρατηρούμε ότι για μικρές μετατοπίσεις,  $\delta=0.25$  και  $\delta=0.5$  αυξάνεται ιδιαίτερα το  $ARL_1$ . Αυτό οφείλεται στο γεγονός ότι δεν αποτελεί μέλημα, από τη σχεδίαση των υπόψη διαγραμμάτων, η ανίχνευση μετατοπίσεων με μικρή διασπορά μιας και είναι μονόπλευρα.

Αντίθετα στην περίπτωση των δίπλευρων διαγραμμάτων παρατηρούμε ότι η τιμή που βρίσκεται στο κελί  $\delta=0$  και  $S^2=1$  αποτελεί τη βέλτιστη συνθήκη, καθώς όσο αυξάνεται η μετατόπιση και για όσο διαφοροποιείται η διασπορά από τη θεωρητική τιμή  $S^2=1$ , η αλλαγή γίνεται αντιληπτή σε λιγότερες παρατηρήσεις από τις 370 παρατηρήσεις. Εξάριση αποτελεί η περίπτωση που δεν έχουμε μετατόπιση ( $\delta=0$ ) και η διασπορά είναι ίση με  $S^2=0.75$ . Ωστόσο, δεν δημιουργεί ιδιαίτερο πρόβλημα, καθώς όταν δεν έχουμε μετατόπιση, ακόμη και η διασπορά να είναι ελαφρώς μικρότερη της αναμενόμενης και να αργήσουμε σχετικά να το αντιληφθούμε δεν μας απασχολεί. Στην προκειμένη περίπτωση είναι ίσο με 1483 παρατηρήσεις.

### 6.3.6 Εφαρμογή Διαγράμματος Μέσου - Διασποράς EWMA

Προκειμένου να διαπιστωθεί στην πράξη με ένα παράδειγμα η ανιχνευτική ικανότητα των εν λόγω διαγραμμάτων, κατασκευάστηκαν τα διαγράμματα που ακολουθούν (Σχήμα 6.1 και Σχήμα 6.2) με τα δεδομένα του Πίνακα 6.3.

Οι τιμές των 30 παρατηρήσεων του Πίνακα 6.3 (τα  $X_j$  δηλαδή) είναι οι ίδιες με αυτές της περίπτωσης του Πίνακα A.1. Η διαφορά είναι ότι η κάθε παρατήρηση αναλύθηκε σε 4 υποπαρατηρήσεις (τα  $x_{1j}$ ,  $x_{2j}$ ,  $x_{3j}$  και  $x_{4j}$  δηλαδή), από τις οποίες προκύπτει ως μέση τιμή, που αποτελούν τις επιμέρους παρατηρήσεις της κάθε παρτίδας (δείγματος). Αυτό έγινε ώστε να συμφωνεί η δειγματοληψία με τη σχεδίαση της μεθόδου και να είναι δυνατή η μέτρηση της διασποράς. Επιπρόσθετα, στον υπόψη Πίνακα υπολογίστηκαν και παρατίθενται οι τιμές του στατιστικού EWMA για κάθε μία από τις δύο ακολουθίες, του μέσου ( $Z_j^\mu$ ) και της διασποράς ( $Z_j^\sigma$ ) δηλαδή.

Τα δεδομένα έχουν δομηθεί έτσι ώστε από την 15η παρατήρηση και μετά να παρουσι-

άζουν μεγάλη διασπορά, η οποία διατηρείται μέχρι την 30η παρατήρηση. Θυμίζουμε ότι από την 21η παρατήρηση και μετά παρατηρείται βαθμιαία μετατόπιση στο μέσο, η οποία στην 29η παρατήρηση συμπληρώνει συνολική μετατόπιση μονάδα ( $\delta=1$ ). Η συμπεριφορά που έχουν τα δεδομένα συναντάται αρκετά συχνά στη μηχανολογία και για το λόγο αυτό μπορούμε να πούμε ότι αποτελούν πραγματικά δεδομένα. Συγκεκριμένα, πριν την εμφάνιση μιας μετατόπισης στο μέσο, ξεκινάει να υπάρχει έντονη διασπορά, η οποία ουσιαστικά ευθύνεται για τη μετέπειτα μετατόπιση του μέσου.

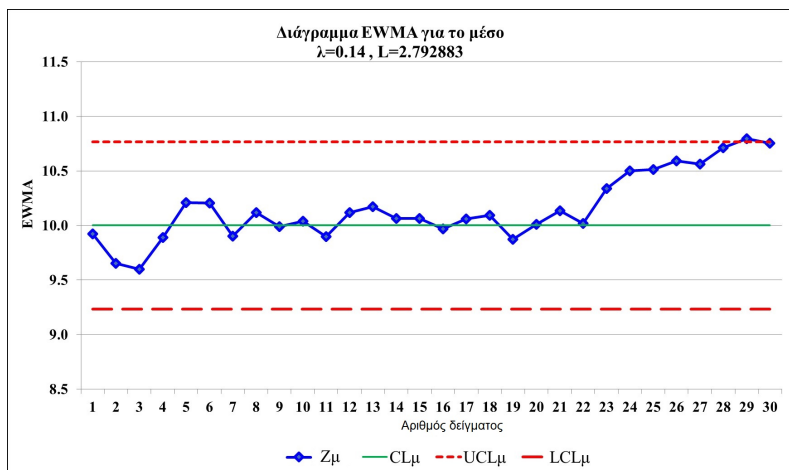
Πίνακας 6.3: Δεδομένα για Διάγραμμα Μέσου – Διασποράς EWMA

$j$	$x_{1j}$	$x_{2j}$	$x_{3j}$	$x_{4j}$	$X_j$	$S_j^2$	$Z_j^\mu$	$Z_j^\sigma$
1	8.80	9.05	9.75	10.20	9.45	0.412	9.923	0.9176333
2	7.99	8.80	7.15	8.00	7.99	0.454	9.652	0.8527107
3	10.00	9.00	10.00	8.14	9.29	0.805	9.600	0.8460172
4	10.40	12.68	11.50	12.14	11.68	0.961	9.891	0.8620868
5	12.90	11.40	12.00	12.32	12.16	0.392	10.208	0.7962886
6	11.50	10.00	9.50	9.70	10.18	0.823	10.204	0.7999582
7	7.10	9.20	8.45	7.40	8.04	0.936	9.900	0.8189516
8	10.79	11.40	12.00	11.65	11.46	0.260	10.119	0.7407077
9	8.91	10.13	8.70	9.05	9.20	0.407	9.990	0.6940108
10	11.32	10.10	9.70	10.25	10.34	0.479	10.039	0.6638474
11	8.49	8.00	10.00	9.64	9.03	0.888	9.898	0.6952976
12	11.89	11.00	12.40	10.59	11.47	0.679	10.118	0.6929973
13	10.43	11.20	9.45	10.95	10.51	0.600	10.173	0.6799625
14	9.50	10.65	9.00	8.45	9.40	0.878	10.065	0.7077344
15	11.05	11.15	10.00	8.10	10.08	2.004	10.066	0.8892349
16	8.60	8.26	11.30	9.30	9.37	1.852	9.968	1.0239614
17	11.50	12.50	8.94	9.55	10.62	2.759	10.060	1.2668236
18	12.42	9.05	10.35	9.40	10.31	2.290	10.094	1.4100356
19	11.05	8.05	7.50	7.47	8.52	2.922	9.873	1.6216488
20	10.00	9.45	11.20	12.70	10.84	2.076	10.008	1.6852055
21	9.10	12.20	11.35	10.95	10.90	1.712	10.133	1.6889101
22	7.90	8.65	11.00	9.75	9.33	1.824	10.020	1.707846
23	12.85	10.65	12.95	12.70	12.29	1.202	10.337	1.6370684
24	12.80	11.25	11.85	10.10	11.50	1.278	10.500	1.5868455
25	12.40	9.90	10.50	9.60	10.60	1.580	10.514	1.5858871
26	12.65	10.05	10.70	10.90	11.08	1.234	10.593	1.5366462
27	10.00	8.85	11.85	10.80	10.38	1.608	10.562	1.5465658
28	12.45	9.53	12.70	11.80	11.62	2.085	10.710	1.6219839
29	11.52	9.65	11.55	12.50	11.31	1.424	10.794	1.5943268
30	11.30	9.56	12.10	9.10	10.52	2.014	10.755	1.6531137

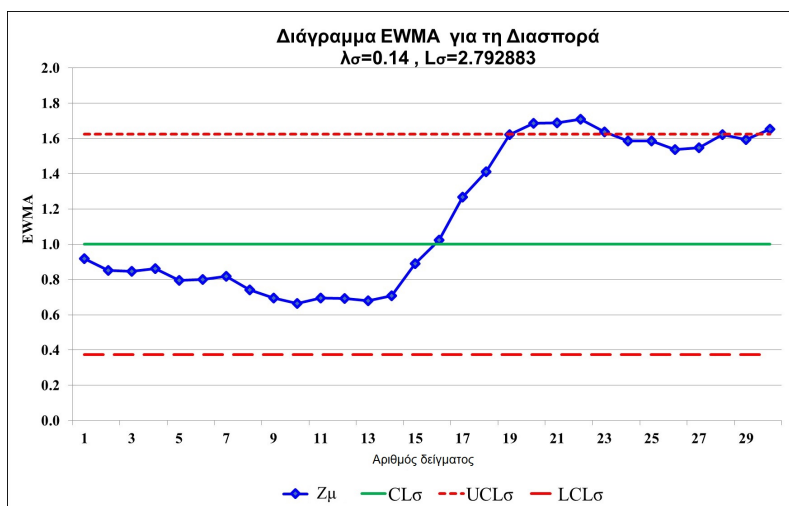
Αναπαριστώντας τα ανωτέρω αποτελέσματα σε δύο διαγράμματα, παίρνουμε τα Σχήματα 6.1 και 6.2.

Από τα υπόψη σχήματα διαπιστώνουμε ότι η μετατόπιση στο διάγραμμα του μέσου δίνεται αντιληπτή στην 29η παρατήρηση, ενώ η έντονη διασπορά που ξεκινάει από την 15η παρατήρηση





Σχήμα 6.1: Διάγραμμα EWMA για μέσο



Σχήμα 6.2: Διάγραμμα EWMA για τη διασπορά

γίνεται αντιληπτή οριακά στην 20η παρατήρηση.

Αυτό συμφωνεί με την θεωρητική ανιχνευτική ικανότητα των διαγραμμάτων των Πινάκων 6.1 και 6.2, σύμφωνα με την οποία για διασπορά ίση με  $\sigma^2=2$ , απαιτούνται κάτι λιγότερο από 6 παρατηρήσεις. Συνεπώς, ξεκινώντας να μετράμε σαν 1η από την 15η παρατήρηση που πρωτοεμφανίζεται η έντονη διασπορά ίση με  $\sigma^2=2$ , αυτή μπορεί να γίνει αντιληπτή όχι νωρίτερα από την 20η παρατήρηση, γεγονός που επαληθεύεται από το διάγραμμα.

Σαν γενικό συμπέρασμα προκύπτει ότι η ανιχνευτική ικανότητα των διαγραμμάτων μέσου – διακύμανσης EWMA βελτιώνεται σημαντικά σε σχέση με τα απλά διαγράμματα EWMA στην περίπτωση που παρατηρείται αύξηση της διασποράς των παρατηρήσεων.

## 6.4 Συνδυαστικά διαγράμματα Shewhart–EWMA

### 6.4.1 Στατιστικό των διαγραμμάτων Shewhart – EWMA

Σε αντιστοιχία με το συνδυαστικό διάγραμμα μέσου – διασποράς που αναπτύχθηκε προηγουμένως, το συνδυαστικό διάγραμμα Shewhart-EWMA (CSEWMA) βασίζεται στο συνδυασμό των δύο αντίστοιχων απλών διαγραμμάτων, του EWMA δηλαδή, όπως αυτό ορίστηκε από τις Σχέσεις (3.9) ως (3.11), και του διαγράμματος Shewhart, όπως αυτό ορίστηκε από τις Σχέσεις (1.1), (1.2) και (1.3).

Το κλασικό διάγραμμα ελέγχου CSEWMA προτείνεται από τους Lucas και Saccucci (1990) [35] και βασίζεται σε απλά τυχαία δείγματα (SRS).

Έστω  $X_{ij}$ , με  $i = 1, 2, \dots, n$  και  $j = 1, 2, \dots, m$ , είναι ανεξάρτητες παρατηρήσεις της  $j$ -οστής υποομάδας μεγέθους  $n$ , προερχόμενες από κανονική κατανομή  $N(\mu, \sigma^2)$ . Τότε, ο μέσος του  $j$ -οστού δείγματος (αποτελούμενου από  $n$  παρατηρήσεις) θα δίνεται από τη σχέση:

$$\bar{X}_{srsj} = \frac{1}{n} \sum_{i=1}^n X_{ij} \quad (6.27)$$

και θα είναι κανονικά κατανεμημένος με μέσο  $\mu$  και διασπορά  $\sigma^2/n$ . Ισοδύναμα, συμβολίζουμε ως εξής:

$$\bar{X}_{srsj} \sim N(\mu, \sigma^2/n)$$

Το διάγραμμα Shewhart χρησιμοποιεί την τρέχουσα πληροφορία βασισμένη στο στατιστικό  $\bar{X}_{srsj}$  για την παρακολούθηση του μέσου της διαδικασίας, ενώ το διάγραμμα EWMA στηρίζεται στο στατιστικό  $Z_j$ , όπως αυτό ορίστηκε με τη Σχέση (3.3), το οποίο ακολουθεί κανονική κατανομή με μέσο και διασπορά ως ακολούθως:

$$Z_j \sim N\left(\mu_0, \frac{\sigma^2}{n} \left(\frac{\lambda}{2-\lambda}\right)\right) \quad (6.28)$$

Σημειώνεται ότι στο υπόψη συνδυαστικό διάγραμμα CSEWMA χρησιμοποιούμε την ασυμπτωτική τιμή της διασποράς.

### 6.4.2 Διάγραμμα ελέγχου Shewhart – EWMA

Προφανώς, στα συνδυαστικά διαγράμματα CSEWMA λαμβάνουμε σήμα για διεργασία εκτός ελέγχου είτε όταν το στατιστικό  $\bar{X}_{srsj}$  πέφτει έξω από τα όρια του διαγράμματος Shewhart  $\bar{X}$ , που δίνονται από τη Σχέση (1.1), είτε όταν το στατιστικό  $Z_j$  πέφτει έξω από τα όρια του διαγράμματος EWMA, που δίνονται από τις Σχέσεις (3.9) και (3.11).

Τα όρια τόσο του Shewhart, όσο και του EWMA είναι συνήθως κεντραρισμένα στην τιμή στόχο, δηλαδή την κεντρική γραμμή  $CL = \mu_0$ . Οι τιμές των παραμέτρων σχεδίασης  $k$  και  $L$  επιλέγονται με βάση την επιλογή της σταθεράς εξομάλυνσης  $\lambda$ , προκειμένου να ικανοποιούνται οι απαιτήσεις για τον έλεγχο της διαδικασίας. Για ένα τυπικό διάγραμμα ελέγχου τύπου

Shewhart, οι Lucas και Saccucci (1990) [35] πρότειναν την επιλογή της παραμέτρου  $k$ , ως ακολούθως:  $3.0 < k < 4.5$ . Σημειώνεται ότι το CSEWMA περιορίζεται στο διάγραμμα Shewhart, καθώς το  $L$  πλησιάζει στο άπειρο, και μεγαλύτερες τιμές του  $k$  παράγουν ένα διάγραμμα EWMA, σύμφωνα με τους Capizzi και Masarotto (2010) [14].

### 6.4.3 Υπολογισμός ARL για τα διαγράμματα Shewhart – EWMA

Το RL του CSEWMA ορίζεται από την κάτωθι σχέση:

$$RL_{csewma} = \min\{RL_{Shewhart}, RL_{ewma}\} \quad (6.29)$$

όπου

$$RL_{Shewhart} = \inf\left\{j : \mu_0 + k\left(\frac{\sigma}{\sqrt{n}}\right) < \bar{X}_{srsj} < \mu_0 - k\left(\frac{\sigma}{\sqrt{n}}\right)\right\} \quad (6.30)$$

και

$$RL_{ewma} = \inf\left\{j : \mu_0 + k\left(\frac{\sigma}{\sqrt{n}}\right)\sqrt{\frac{\lambda}{2-\lambda}} < \bar{X}_{srsj} < \mu_0 - k\left(\frac{\sigma}{\sqrt{n}}\right)\sqrt{\frac{\lambda}{2-\lambda}}\right\} \quad (6.31)$$

Οι Σχέσεις (6.29), (6.30) και (6.31) είναι οι μηδενικής κατάστασης τιμές για τα RL του στατιστικού  $\bar{X}_{srs}$  του διαγράμματος Shewhart και του στατιστικού  $Z_j$  του διαγράμματος EWMA, αντίστοιχα. Επιπρόσθετα, το αναμενόμενο RL του CSEWMA στην μηδενική κατάσταση είναι το ARL εντός ελέγχου, το οποίο είναι:  $E[RL_{csewma}] = ARL_0$ .

### 6.4.4 Ανιχνευτική ικανότητα Διαγραμμάτων Shewhart – EWMA

Η ανιχνευτική ικανότητα των συνδυαστικών διαγραμμάτων CSEWMA υπολογίστηκε με τη βοήθεια του κώδικα του παραρτήματος, λαμβάνοντας υπόψη την υπόθεση της κανονικότητας με  $\bar{X}_{srsj} \sim N(\delta, 1)$ , για διαφορετικούς συνδυασμούς των  $k$ ,  $L$ ,  $\lambda$  και  $\delta$  για ένα εντός ελέγχου  $ARL_0=370$ , όπου η μετατόπιση στο μέσο είναι:

$$\delta = \frac{\sqrt{n}|\mu_1 - \mu_0|}{\sigma_0}$$

και μετρείται σε μονάδες  $\sigma_0/\sqrt{n}$ . Τα  $\mu_0$  και  $\mu_1$  είναι ο εντός ελέγχου και εκτός ελέγχου μέσος, αντίστοιχα. Υπενθυμίζεται ότι ισχύει  $\delta=0$ , όταν η διαδικασία είναι εντός ελέγχου.

Οι τιμές του ARL υπολογίστηκαν για τις εξής τέσσερις επιλογές παραμέτρων:  $\lambda=0.14$  και  $L_{ewma}=2.784641$  με  $n=5$ ,  $\lambda=0.14$  και  $L_{ewma}=2.784641$  με  $n=1$ ,  $\lambda=0.05$  και  $L_{ewma}=2.48969$  με  $n=1$  και  $\lambda=0.14$  και  $L_{ewma}=2.784641$  με  $n=4$ . Τα αποτελέσματα παρουσιάζονται στους Πίνακες 6.4, 6.5, 6.6 και 6.7 αντίστοιχα. Σε κάθε περίπτωση θεωρήθηκε ότι  $L_{Shewhart}=3$ , προκειμένου να διατηρηθεί το  $ARL_0=370$ .

Από τη μελέτη των πινάκων διαπιστώνουμε ότι τα απλά διαγράμματα EWMA έχουν μικρότερα ARL από τα αντίστοιχα διαγράμματα Shewhart για μικρές μετατοπίσεις, ενώ για μεγάλες μετατοπίσεις υπερτερούν τα διαγράμματα Shewhart. Επίσης, όσο αυξάνεται το μέγεθος της υποομάδας των δειγματοληψιών, τόσο μειώνεται η συνεισφορά του διαγράμματος Shewhart στην ανιχνευτική ικανότητα των συνδυαστικών διαγραμμάτων CSEWMA.

Έτσι, για παράδειγμα για τα αποτελέσματα του Πίνακα 6.4, για μέγεθος υποομάδας  $n=5$ , μετατοπίσεις με  $\delta \leq 1$  ανιχνεύονται πιο γρήγορα από τα απλά διαγράμματα EWMA, ενώ μετατοπίσεις μεγαλύτερες από 1 ανιχνεύονται γρηγορότερα από τα διαγράμματα Shewhart.

Αντίστοιχα, αναφορικά με τα αποτελέσματα του Πίνακα 6.5, για μέγεθος υποομάδας  $n=1$ , οι μετατοπίσεις με  $\delta \leq 2$  ανιχνεύονται πιο γρήγορα από τα απλά διαγράμματα EWMA, ενώ μετατοπίσεις μεγαλύτερες από 2 ανιχνεύονται γρηγορότερα από τα διαγράμματα Shewhart.

Πίνακας 6.4: Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA

$L_{Shewhart} = 3, ARL_0 = 370, n = 5$			
	Απλό EWMA	Shewhart	CSEWMA
	$\lambda=0.14$		$\lambda=0.14$
$\delta$	$L=2.784641$	$L_{Shewhart}=3.0$	$L=2.784641$
0	370	370	370
0.25	24.29	133.06	24.29
0.5	7.13	33.36	7.13
0.75	3.75	10.67	3.75
1	2.41	4.27	2.41
1.5	1.27	0.93	0.93
2	0.85	0.15	0.15
3	0.11	0.00021	0.00021
4	0.00025	2.76434e-09	2.76434e-09

Γενικά, μπορούμε να πούμε ότι όσο αυξάνεται το μέγεθος υποομάδας των δειγματοληψιών, τόσο μειώνεται η συνεισφορά του EWMA στο συνδυαστικό CSEWMA.

### 6.4.5 Εφαρμογή Διαγράμματος Shewhart – EWMA

Αν χρησιμοποιήσουμε τα δεδομένα του Πίνακα A.1, μπορούμε να κατασκευάσουμε το διάγραμμα CSEWMA του Σχήματος 6.3. Θεωρούμε συγκεκριμένες τιμές  $\lambda_\mu$  και  $L_\mu$ , καθώς επίσης και μέγεθος δείγματος  $n=4$ . Διαπιστώνουμε ότι οι μετατοπίσεις σε αρκετές παρατηρήσεις, συγκεκριμένα την 2η, 4η, 5η, 7η, 24η, 25η και 28η, γίνονται άμεσα αντιληπτές.

Εάν το μέγεθος δείγματος  $n$  είναι μικρότερο, τότε μειώνεται η ανιχνευτική ικανότητα των διαγραμμάτων. Θα μπορούσαμε να πούμε ότι εφόσον σε κάθε χρονική στιγμή λαμβάνονται λιγότερες δειγματοληψίες, τότε μικρότερο θα είναι και το τελικό πλήθος δειγματοληψιών συνολικά που θα παρθούν μέχρι εκείνη τη χρονική στιγμή που εξετάζεται, γεγονός που είναι

Πίνακας 6.5: Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA

$L_{Shewhart} = 3, ARL_0 = 370, n = 1$			
	Απλό EWMA	Shewhart	CSEWMA
	$\lambda=0.14$		$\lambda=0.14$
$\delta$	$L=2.784641$	$L_{Shewhart}=3.0$	$L=2.784641$
0	370	370	370
0.25	102.33	280.89	102.33
0.5	30.97	155.11	30.97
0.75	15.14	81.16	15.15
1	9.56	43.85	9.56
1.5	5.46	14.9	5.46
2	3.86	6.14	3.86
3	2.52	1.5	1.5
4	1.99	0.35	0.35

Πίνακας 6.6: Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA

$L_{Shewhart} = 3, ARL_0 = 370, n = 1$			
	Απλό EWMA	Shewhart	CSEWMA
	$\lambda=0.05$		$\lambda=0.05$
$\delta$	$L=2.48969$	$L_{Shewhart}=3.0$	$L=2.48969$
0	370	370	370
0.25	73.15	280.89	73.15
0.5	26.45	155.11	26.45
0.75	15.32	81.16	15.32
1	10.73	43.85	10.73
1.5	6.75	14.9	6.75
2	4.98	6.14	4.98
3	3.35	1.5	1.5
4	2.57	0.35	0.35

φυσιολογικό να επηρεάζει την τελική ανιχνευτική ικανότητα του διαγράμματος. Για να πειστούμε γι' αυτό, τα ίδια δεδομένα χρησιμοποιήθηκαν ξανά σε διαγράμματα CSEWMA, με  $n=1$  και  $n=2$  αντίστοιχα. Τα αποτελέσματα παρουσιάζονται στα Σχήματα 6.4 και 6.5.

Από τα διαγράμματα φαίνεται καθαρά το γεγονός που προαναφέρθηκε, ότι δηλαδή όσο μεγαλύτερο είναι το μέγεθος της υποομάδας των δειγματοληψιών, τόσο περισσότερο επιταχύνεται η ανίχνευση των μετατοπίσεων. Συνεπώς, με κατάλληλη επιλογή του  $n$  είναι δυνατόν να προσδιοριστεί επακριβώς το μέγεθος της μετατόπισης που είναι επιθυμητό να ανιχνεύεται άμεσα.

Συμπερασματικά, μπορούμε να πούμε ότι το συνδυαστικό διάγραμμα CSEWMA επιταχύνει την ανίχνευση της μετατόπισης στο μέσο μιας διαδικασίας συνολικά, καθώς εκμεταλλεύεται τα πλεονεκτήματα και των δύο μεθόδων που συμμετέχουν. Δηλαδή, με το διάγραμμα

Πίνακας 6.7: Ανιχνευτική Ικανότητα Συνδυαστικών Διαγραμμάτων CSEWMA

$L_{Shewhart} = 3, ARL_0 = 370, n = 4$			
	Απλό EWMA	Shewhart	CSEWMA
	$\lambda=0.14$		$\lambda=0.14$
$\delta$	$L=2.784641$	$L_{Shewhart}=3.0$	$L=2.784641$
0	370	370	370
0.25	30.47	155.11	30.47
0.5	9.08	43.85	9.08
0.75	4.96	14.9	4.96
1	3.37	6.14	3.37
1.5	2.02	1.5	1.5
2	1.5	0.35	0.35
3	0.8	0.0027	0.0027
4	0.51	5.71694ε-07	5.71694ε-07

Shewhart ανιχνεύονται γρηγορότερα οι μεγάλες μετατοπίσεις, ενώ με το διάγραμμα EWMA ανιχνεύονται ταχύτερα οι μικρότερες μετατοπίσεις.

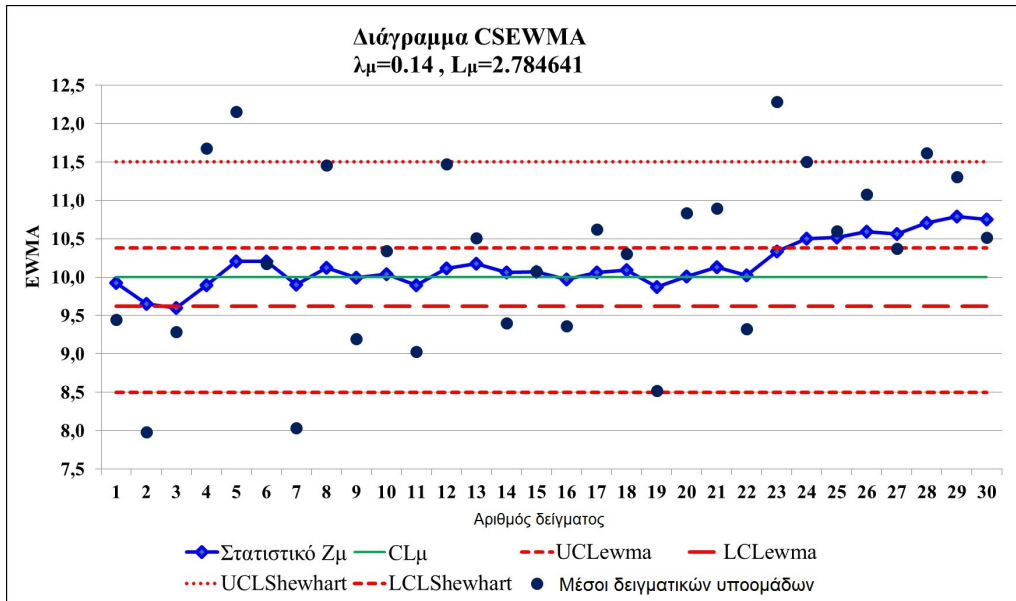
Επίσης, παρατηρούμε ότι το μέγεθος του δείγματος  $n$  συμμετέχει στην ανιχνευτική ικανότητα του διαγράμματος CSEWMA και πρέπει να λαμβάνεται υπόψη, σε συνδυασμό με την επιλογή των παραμέτρων, για τον προσδιορισμό του επιθυμητού μεγέθους ανιχνευτικής ικανότητας.

## 6.5 Συμπέρασμα

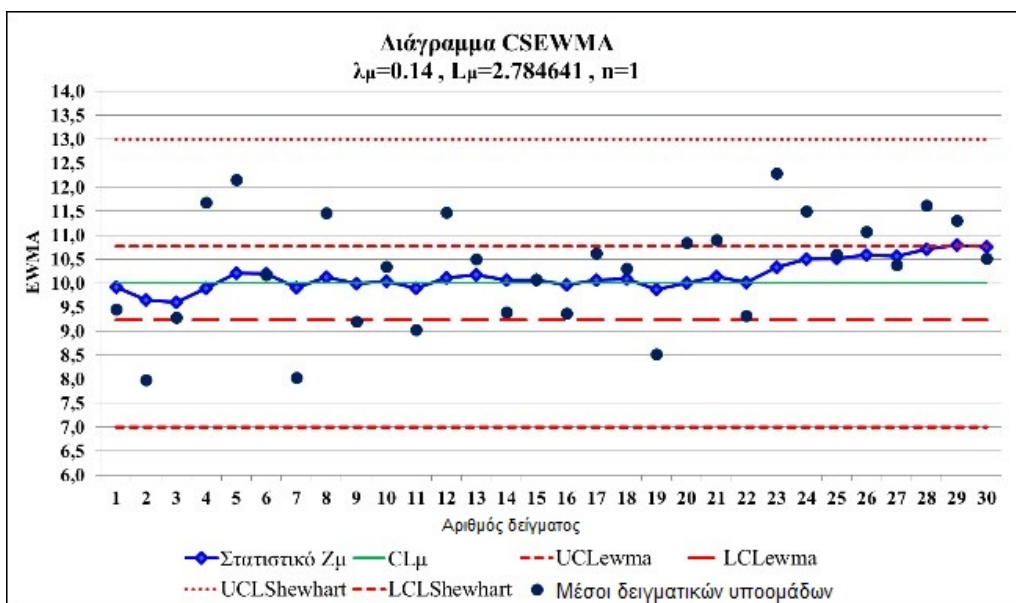
Προφανώς, εκτός από τα συνδυαστικά διαγράμματα που εξετάστηκαν παραπάνω, μπορεί να υπάρξουν και άλλοι συνδυασμοί τύπων διαγραμμάτων, ακόμη και περισσότερων από δύο διαφορετικών. Ενδεικτικά αναφέρεται το συνδυαστικό διάγραμμα CUSUM-EWMA, το οποίο είναι ίδιας λογικής με το CSEWMA.

Μια ενδιαφέρουσα κατηγορία συνδυαστικών διαγραμμάτων είναι αυτή, όπου τα στατιστικά δύο διαφορετικών μεθόδων συνδυάζονται έτσι ώστε το ένα να τροφοδοτεί το άλλο. Έτσι, υπάρχει το διάγραμμα των Abbas et al. (2012) [2], όπου το στατιστικό του EWMA τροφοδοτεί ένα CUSUM διάγραμμα, και το διάγραμμα των Zaman et al. (2015) [53], όπου το στατιστικό του CUSUM τροφοδοτεί ένα EWMA διάγραμμα. Αντίστοιχης λογικής είναι και το διάγραμμα Double EWMA, όπου το στατιστικό από το 1ο EWMA τροφοδοτεί το στατιστικό του 2ου EWMA.

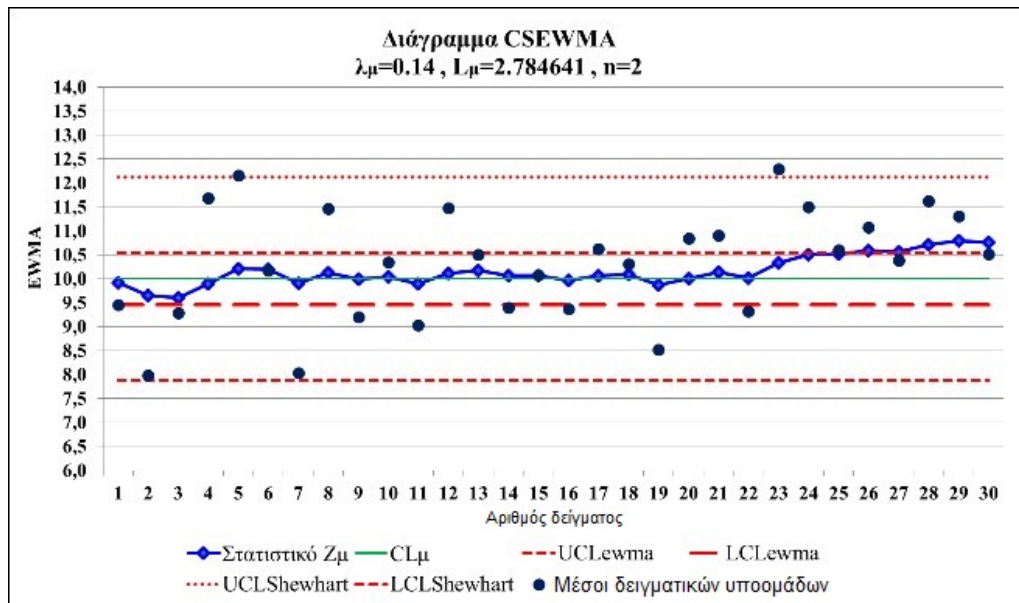
Από την ανάλυση των ανωτέρω διαγραμμάτων συνειδητοποιούμε ότι, με κατάλληλη επιλογή των δύο ή και τριών διαγραμμάτων που συμμετέχουν στα συνδυαστικά διαγράμματα, είναι δυνατόν να βελτιώσουμε την ανιχνευτική ικανότητα για τα μεγέθη μετατοπίσεων που επιθυμούμε να ανιχνεύουμε.



Σχήμα 6.3: Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα Α.1 ( $n=4$ )



Σχήμα 6.4: Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα Α.1 ( $n=1$ )



Σχήμα 6.5: Διάγραμμα CSEWMA για τα δεδομένα του Πίνακα Α.1 ( $n=2$ )



# Κεφάλαιο 7

## Διαγράμματα EWMA με ταξινομημένες δειγματοληψίες

### 7.1 Εισαγωγή στις Τεχνικές Δειγματοληψίας

Είδαμε ότι μπορούμε να βελτιώσουμε τα διαγράμματα EWMA επεμβαίνοντας στις τιμές των παραμέτρων και στη μαθηματική δομή της εξίσωσης τόσο του στατιστικού, όσο και των ορίων ελέγχου. Στις περιπτώσεις που αναλύθηκαν, θεωρώντας ότι λαμβάναμε  $j$  δείγματα ( $j=1, 2, \dots, r$ ) με μέγεθος δείγματος  $i$  δείγματα (με  $i=1, 2, \dots, n$ ), η τιμή της τρέχουσας πληροφορίας  $\bar{X}_j$  προερχόταν από απλές τυχαίες δειγματοληψίες (SRS). Δηλαδή:

$$\bar{X}_{SRS(j)} = \frac{1}{n} \sum_{i=1}^n X_{ij}$$

Η τιμή της τρέχουσας πληροφορίας, ωστόσο, μπορεί να προσδιοριστεί και διαφορετικά. Αυτό γίνεται με επιλογή τεχνικής δειγματοληψίας, κατά την οποία πραγματοποιείται κάποια ταξινόμηση στις μετρήσεις (ranking). Η βασική τεχνική αυτής της κατηγορίας είναι η δειγματοληψία ταξινομημένων συνόλων (RSS), ενώ οι βασικές τροποποιήσεις της είναι:

- H Median Ranked Set Sampling (MRSS)
- H Double Ranked Set Sampling (DRSS)
- H Median Double Ranked Set Sampling (MDRSS)
- H Double Median Ranked Set Sampling (DMRSS)
- H Ordered Double Ranked Set Sampling (ODRSS)

Οι παραπάνω τεχνικές δειγματοληψίας με κατάταξη ουσιαστικά χρησιμοποιούν παραπάνω πληροφορία από ότι η SRS, καθώς ταξινομούν τις μονάδες (παρατηρήσεις), χωρίς όμως να

μετρούν παραπάνω μονάδες από ότι η SRS. Αυτό που επιτυγχάνεται είναι ότι τα σύνολα μονάδων τα οποία μετρούνται είναι πιο ομοιογενή, οδηγώντας σε πιο σταθερά διαγράμματα ελέγχου και σε καλύτερη ανιχνευτική ικανότητα.

## 7.2 Τεχνική RSS (Ranked Set Sampling)

Η ιδέα της τεχνικής δειγματοληψίας RSS πρωτοεισήχθη από τον McIntyre (1952) [36]. Κατά την τεχνική αυτή, από ένα πληθυσμό λαμβάνεται ένα σύνολο  $n^2$  μονάδων δειγματοληψίας, χωρισμένων σε  $n$  υποσύνολα (σετ). Τα δείγματα εντός του κάθε σετ κατατάσσονται με αύξουσα σειρά, χωρίς όμως αυτό να προσμετράται ως πραγματική μέτρηση της μεταβλητής ενδιαφέροντος (χωρίς κόστος δηλαδή), καθώς η κατάταξη αυτή μπορεί να γίνει με απλό τρόπο (π.χ. οπτικά). Μετά την κατάταξη ακολουθεί επιλογή διαφορετικής τάξης μονάδας δειγματοληψίας, δηλαδή: από το 1ο σετ λαμβάνεται το 1ο σε κατάταξη (μικρότερο) δείγμα, από το 2ο σύνολο λαμβάνεται το 2ο σε κατάταξη (2ο μικρότερο) δείγμα κοκ. Η πραγματική μέτρηση γίνεται στο σύνολο  $n$  μονάδων δειγματοληψίας που προκύπτει μ' αυτόν τον τρόπο. Αυτή η διαδικασία επαναλαμβάνεται, την κάθε επανάληψη δε ονομάζουμε κύκλο (για την αρίθμηση των κύκλων χρησιμοποιούμε τον συμβολισμό  $j = 1, 2, \dots, r$ ).

Η μεθόδος RSS μπορεί να περιγραφεί από τον ακόλουθο αλγόριθμο:

1ο βήμα: Τυχαία επιλογή ενός δείγματος μεγέθους  $n^2$  μονάδων από τον πληθυσμό.

2ο βήμα: Διαχωρισμός του δείγματος σε  $n$  σετ, καθένα από αυτά μεγέθους  $n$ .

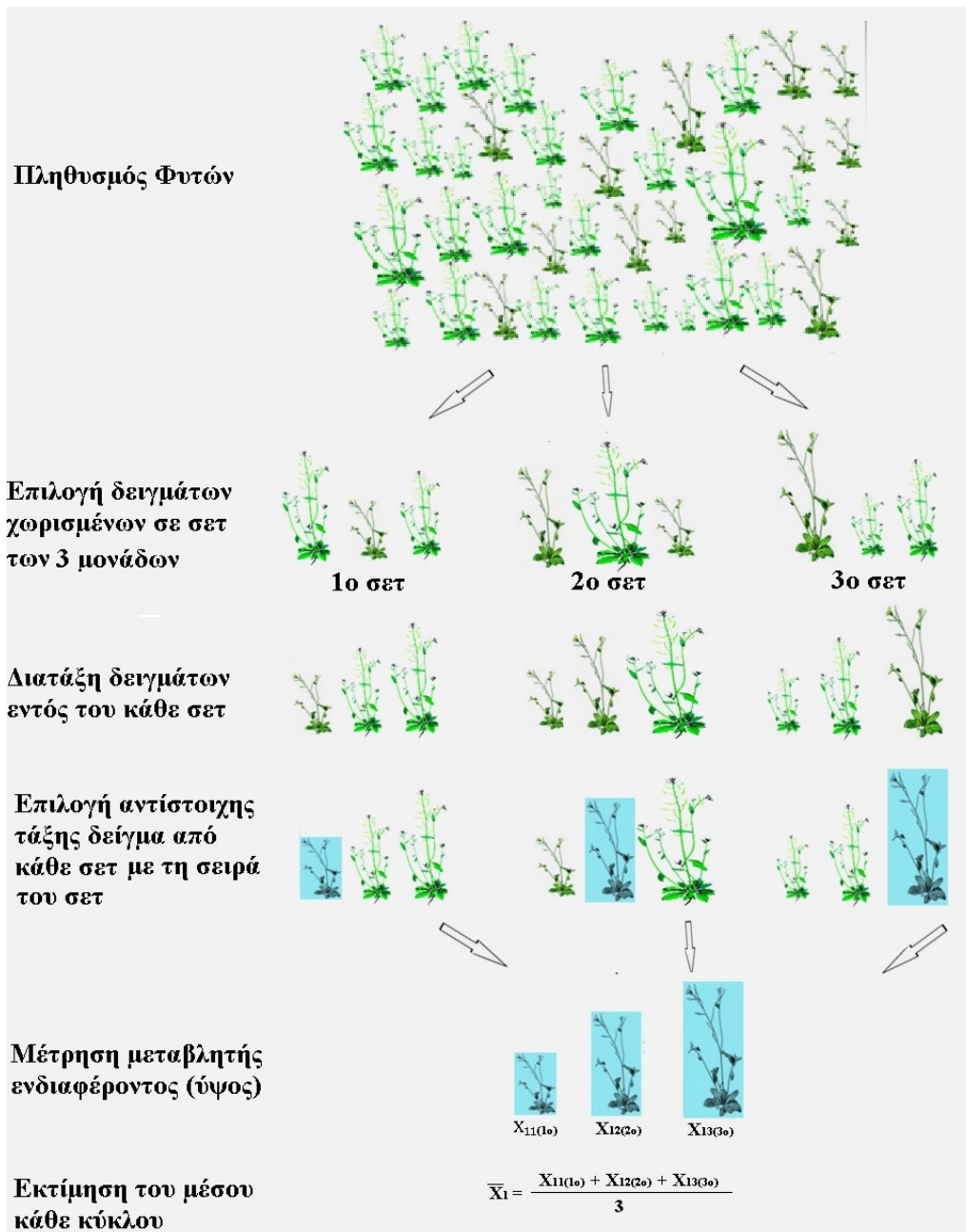
3ο βήμα: Ταξινόμηση των μονάδων εντός κάθε σετ βάσει της μεταβλητής που μας ενδιαφέρει, με απλό τρόπο (χωρίς κοστοβόρα μέτρηση).

4ο βήμα: Επιλογή της μικρότερης ταξινομημένης μονάδας από το 1ο σετ, της 2ης μικρότερης από το 2ο σύνολο κοκ μέχρι να επιλεγεί η μεγαλύτερη ταξινομημένη μονάδα από το τελευταίο σετ.

5ο βήμα: Μέτρηση των  $n$  επιλεγμένων μονάδων και υπολογισμός του μέσου τους (αυτή την ποσότητα θα τροφοδοτήσουμε στο στατιστικό του διαγράμματος ελέγχου).

Η παραπάνω διαδικασία επαναλαμβάνεται  $r$  φορές, δημιουργώντας δηλαδή ένα σύνολο από  $nr$  μονάδες RSS δεδομένων.

Ένα καλό παράδειγμα για την κατανόηση της τεχνικής RSS είναι η δειγματοληψία από πληθυσμό φυτών, όπου το χαρακτηριστικό (μεταβλητή) που μας ενδιαφέρει είναι το ύψος τους, όπως απεικονίζεται στο Σχήμα 7.1. Στο υπόψη παράδειγμα, σε κάθε κύκλο επιλέγουμε τυχαία 9 φυτά ( $n=3$ ) από τον πληθυσμό, η ταξινόμηση των οποίων γίνεται οπτικά.



Σχήμα 7.1: Παράδειγμα δειγματοληψίας RSS

Η διαδικασία που ακολουθείται έχει ως εξής:

1ος Κύκλος δειγματοληψίας

1ο βήμα: Τυχαία επιλογή δείγματος: 9 δειγματικές μονάδες

2ο βήμα: Διαχωρισμός του δείγματος σε 3 σετ, καθένα από αυτά μεγέθους 3:

Μη διατεταγμένο σετ No.1:  $X_{111}$ ,  $X_{112}$ ,  $X_{113}$

Μη διατεταγμένο σετ No.2:  $X_{121}$ ,  $X_{122}$ ,  $X_{123}$

Μη διατεταγμένο σετ Νο.3:  $X_{131}, X_{132}, X_{133}$

3ο βήμα: Ταξινόμηση των μονάδων εντός κάθε σετ:

Διατεταγμένο σετ Νο.1:  $X_{11(1^\circ)} \leq X_{11(2^\circ)} \leq X_{11(3^\circ)}$

Διατεταγμένο σετ Νο.2:  $X_{12(1^\circ)} \leq X_{12(2^\circ)} \leq X_{12(3^\circ)}$

Διατεταγμένο σετ Νο.3:  $X_{13(1^\circ)} \leq X_{13(2^\circ)} \leq X_{13(3^\circ)}$

4ο βήμα: Επιλογή δειγματικών μονάδων συγκεκριμένης κατάταξης από κάθε σετ:

Επιλογή δειγματικής μονάδας 1ης τάξης από το σετ Νο.1:  $X_{11(1^\circ)}$

Επιλογή δειγματικής μονάδας 2ης τάξης από το σετ Νο.2:  $X_{12(2^\circ)}$

Επιλογή δειγματικής μονάδας 3ης τάξης από το σετ Νο.3:  $X_{13(3^\circ)}$

5ο βήμα: Μέτρηση των επιλεγμένων δειγματικών μονάδων και υπολογισμός του μέσο όρου τους:

$$\bar{X}_{RSS1} = \frac{X_{11(1^\circ)} + X_{12(2^\circ)} + X_{13(3^\circ)}}{3}$$

2ος Κύκλος δειγματοληψίας

Επανάληψη βημάτων 1 έως 5

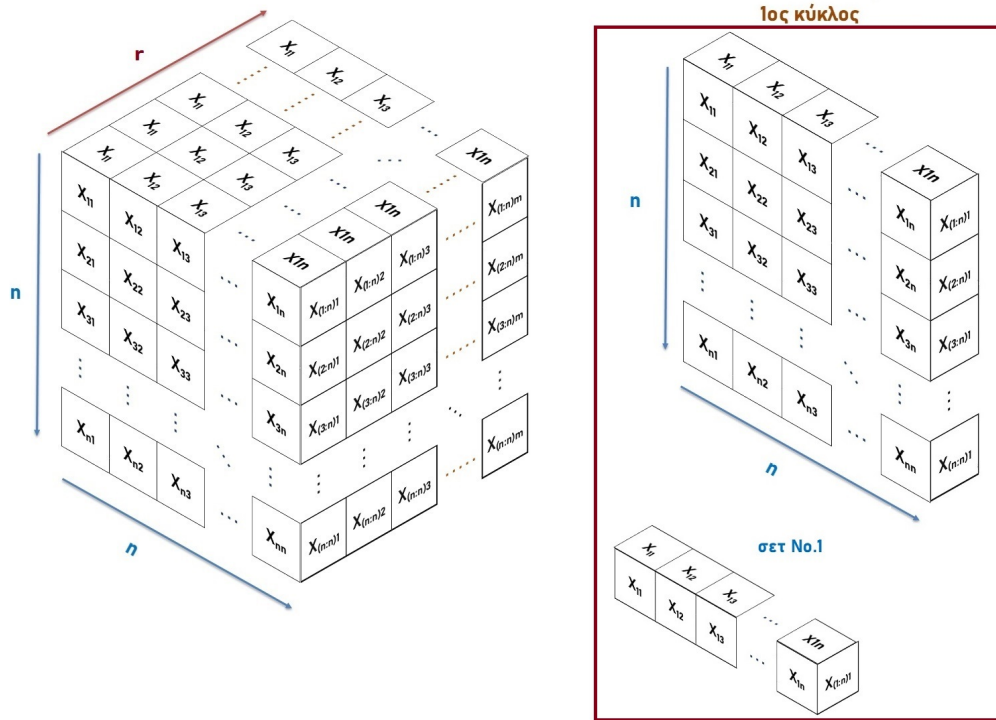
3ος Κύκλος δειγματοληψίας

Επανάληψη βημάτων 1 έως 5

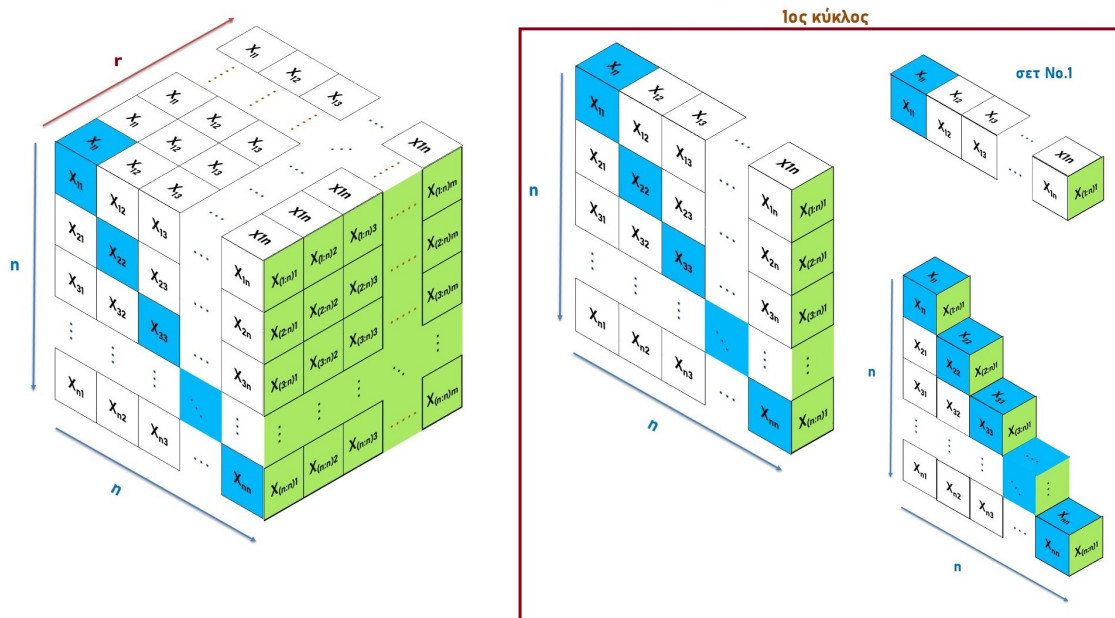
κ.ο.κ

Τέλος, παρατίθενται στη συνέχεια μια σειρά από σχηματικές απεικονίσεις της μεθόδου, οι οποίες βοηθούν στην κατανόηση της μεθοδολογίας μέσω οπτικοποίησής της.

Στο Σχήμα 7.3, με μπλε χρώμα σημειώνεται η θέση (σε ταξινομημένο σετ) που αντιστοιχεί στη δειγματική μονάδα (παρατήρηση) που επιλέγουμε. Η δειγματική μονάδα αυτή, που συμβολίζεται  $X_{(i:n)j}$ , σημειώνεται με πράσινο χρώμα και, όπως θα δούμε στη συνέχεια, είναι το στατιστικό  $i$ -τάξης για το  $i$ -οστό δείγμα του υποσυνόλου μεγέθους  $n$  στο  $j$ -οστό κύκλο.



Σχήμα 7.2: Σύνολο παρατηρήσεων (αριστερά) - Παρατηρήσεις 1ου κύκλου (πάνω δεξιά)  
- Σεν No.1 1ου κύκλου (κάτω δεξιά)



Σχήμα 7.3: Επιλογή παρατήρησης από ταξινομημένα RSS σεν

## 7.2.1 Εκτιμητές Τεχνικής RSS

Έστω η ακολουθία  $X_{(i:n)j}$ , με  $i = 1, 2, \dots, n$  και  $j = 1, 2, \dots, r$ , η οποία δηλώνει το στατιστικό  $i$ -τάξης για το  $i$ -οστό δείγμα του υποσυνόλου μεγέθους  $n$  στο  $j$ -οστό κύκλο.

Πρόκειται δηλαδή για τα  $X_{11(1^\circ)}, X_{12(2^\circ)}, X_{13(3^\circ)}$  του προηγούμενου παραδείγματος. Οι Takahasi και Wakimoto (1968) [50] πρότειναν ένα αμερόληπτο εκτιμητή του πληθυσμιακού μέσου για την τεχνική RSS ως ακολούθως:

$$\bar{X}_{RSS(j)} = \sum_{i=1}^n \frac{X_{(i:n)j}}{n} \quad (7.1)$$

Η δε διακύμανση του  $\bar{X}_{RSS(j)}$  ορίστηκε ως εξής:

$$Var [\bar{X}_{RSS(j)}] = \sigma_{\bar{X}_{RSS(j)}}^2 = \sum_{i=1}^n \frac{\sigma_{(i:n)}^2}{n^2} \quad (7.2)$$

όπου

$$\sigma_{(i:n)}^2 = E[X_{(i:n)} - E(X_{(i:n)})]^2 \quad (7.3)$$

Οπότε, για όλους τους κύκλους δειγματοληψίας ( $j = 1, 2, \dots, r$ ) από την έναρξη ως τον  $r$ -οστό κύκλο, καθώς χρησιμοποιούμε RSS δεδομένα, η αμερόληπτη εκτιμητή του πληθυσμιακού μέσου  $\bar{X}_{RSS}$  δίνεται από τον τύπο:

$$\bar{X}_{RSS} = \frac{1}{r} \sum_{j=1}^r \bar{X}_{RSS(j)} \quad (7.4)$$

Για τη διακύμανση του  $\bar{X}_{RSS}$ , τη διακύμανση δηλαδή της δειγματοληψίας RSS, οι Muttalak και Al-Sabah (2003) [39] πρότειναν τον εξής εκτιμητή:

$$\hat{\sigma}_{\bar{X}_{RSS}} = \left[ \frac{1}{n} \hat{\sigma}_{RSS}^2 - \frac{1}{n^2} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right]^{1/2} \quad (7.5)$$

όπου

$$\hat{\sigma}_{RSS}^2 = \frac{1}{nr - 1} \sum_{i=1}^n \sum_{j=1}^r (X_{(i:n)j} - \bar{X}_{RSS})^2 \quad (7.6)$$

είναι η εκτίμηση της συνολικής διασποράς του πληθυσμού, και

$$\tilde{X}_{(i)} = \frac{1}{r} \sum_{j=1}^r X_{(i:n)j} \quad (7.7)$$

είναι η εκτίμηση για τον πληθυσμιακό μέσο του στατιστικού  $i$ -τάξης.

Από την παραπάνω Σχέση (7.5) βλέπουμε ότι η διασπορά των παρατηρήσεων της τεχνικής RSS αποτελεί συνάρτηση της μετατόπισης του μέσου της διαδικασίας. Άρα, όσο μετατοπίζεται ο μέσος, τόσο μειώνεται η διασπορά  $\hat{\sigma}_{\bar{X}_{RSS}}$  του στατιστικού της δειγματοληψίας. Αυτό έχει ως έμμεσο αποτέλεσμα να στενεύουν τα όρια του διαγράμματος και να παίρνουμε ευκολότερα εκτός ελέγχου σήμα, όταν συμβαίνει να έχουμε ταυτόχρονα μετατόπιση του μέσου και της διασποράς.

Οι παραπάνω σχέσεις της τεχνικής RSS μπορούν να εφαρμοστούν σε όλα τα διαγράμματα

ελέγχου.

## 7.2.2 Διάγραμμα Ελέγχου EWMA με τεχνική RSS

Με τη λογική των διαγραμμάτων EWMA που είδαμε σε προηγούμενο κεφάλαιο, το στατιστικό EWMA για δεδομένα RSS δίνεται από τη σχέση:

$$Z_{(j)} = \lambda \bar{X}_{RSS(j)} + (1 - \lambda)Z_{(j-1)} \quad (7.8)$$

ή ισοδύναμα:

$$Z_{(j)} = \lambda \sum_{i=1}^n \frac{X_{(i:n)j}}{n} + (1 - \lambda)Z_{(j-1)} \quad (7.9)$$

όπου  $i = 1, 2, \dots, n$  και  $j = 1, 2, \dots, r$ . Αναπτύσσοντας προκύπτει:

$$\begin{aligned} Z_{(j)} &= \lambda \bar{X}_{RSS(j)} + (1 - \lambda)[(1 - \lambda)Z_{(j-2)} + \lambda \bar{X}_{RSS(j-1)}] = \\ &= \lambda \bar{X}_{RSS(j)} + (1 - \lambda)^2 Z_{(j-2)} + \lambda(1 - \lambda) \bar{X}_{RSS(j-1)} = \\ &= \lambda \bar{X}_{RSS(j)} + \lambda(1 - \lambda) \bar{X}_{RSS(j-1)} + (1 - \lambda)^2 [(1 - \lambda)Z_{(j-3)} + \lambda \bar{X}_{RSS(j-2)}] = \\ &= \lambda \bar{X}_{RSS(j)} + \lambda(1 - \lambda) \bar{X}_{RSS(j-1)} + \lambda(1 - \lambda)^2 \bar{X}_{RSS(j-2)} + (1 - \lambda)^3 Z_{(j-3)} = \\ &= \dots = \\ &= (1 - \lambda)^j Z_0 + \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k \bar{X}_{RSS(j-k)} \end{aligned}$$

Την αρχική τιμή, δηλαδή την  $Z_0$ , το θεωρούμε ίση με την τιμή στόχο του μέσου  $\mu_0$ . Επομένως, η τιμή του στατιστικού RSS-EWMA είναι:

$$Z_{(j)} = (1 - \lambda)^j \mu_0 + \lambda \sum_{k=0}^{j-1} (1 - \lambda)^k \bar{X}_{RSS(j-k)} \quad (7.10)$$

ή ισοδύναμα:

$$Z_{(j)} = (1 - \lambda)^j \mu_0 + \frac{\lambda}{n} \sum_{k=0}^{j-1} \left( (1 - \lambda)^k \sum_{i=1}^n X_{(i:n)(j-k)} \right) \quad (7.11)$$

Θα χρησιμοποιήσουμε τη Σχέση (7.11) για να εκτιμήσουμε τη μέση τιμή και τη διασπορά του στατιστικού RSS-EWMA. Συνεπώς, για τη μέση τιμή του στατιστικού RSS-EWMA κατά τον  $j$ -ιστό κύκλο δειγματοληψιών ισχύει:

$$E [Z_{(j)}] = E \left[ \frac{\lambda}{n} \sum_{k=0}^{j-1} \left( (1 - \lambda)^k \sum_{i=1}^n X_{(i:n)(j-k)} \right) \right] + E [(1 - \lambda)^j \mu_0] =$$

$$\begin{aligned}
 &= \frac{\lambda}{n} \sum_{k=0}^{j-1} \left( (1-\lambda)^k \sum_{i=1}^n E [X_{(i:n)(j-k)}] \right) + (1-\lambda)^j E [\mu_0] = \\
 &= \frac{\lambda}{n} \sum_{k=0}^{j-1} \left( (1-\lambda)^k \sum_{i=1}^n E [\mu_{(i:n)(j-k)}] \right) + (1-\lambda)^j \mu_0 = \\
 &= \frac{\lambda}{n} \sum_{k=0}^{j-1} \left( (1-\lambda)^k \sum_{i=1}^n \left( \frac{1}{j} \sum_{k=1}^j E [x_{i(i)k}] \right) \right) + (1-\lambda)^j \mu_0 = \\
 &= \frac{\lambda}{rn} \sum_{k=0}^{j-1} \left( (1-\lambda)^k \sum_{i=1}^n \left( \sum_{k=1}^j \mu_0 \right) \right) + (1-\lambda)^j \mu_0 = \\
 &= \frac{\lambda}{rn} \sum_{k=0}^{j-1} \left( (1-\lambda)^k nj \mu_0 \right) + (1-\lambda)^j \mu_0 = \\
 &= \lambda \mu_0 \sum_{k=0}^{j-1} (1-\lambda)^k + (1-\lambda)^j \mu_0 = \\
 &= \lambda \mu_0 \frac{1 - (1-\lambda)^j}{1 - (1-\lambda)} + (1-\lambda)^j \mu_0 = \\
 &= \mu_0 - (1-\lambda)^j \mu_0 + (1-\lambda)^j \mu_0 = \mu_0
 \end{aligned}$$

Άρα, βρίσκουμε ότι:

$$E [Z_{(j)}] = \mu_0 \quad (7.12)$$

Όσον αφορά τη διακύμανση του στατιστικού RSS-EWMA, από τη Σχέση (7.10) προκύπτει ότι:

$$\begin{aligned}
 Var [Z_{(j)}] &= Var [(1-\lambda)^j \mu_0] + Var \left[ \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} \bar{X}_{RSS(j-k)} \right] = \\
 &= Var \left[ \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} \bar{X}_{RSS(j-k)} \right] =
 \end{aligned}$$

Όμως, με χρήση της Σχέσης 7.6 για τη διακύμανση του  $\bar{X}_{RSS}$ , παίρνουμε την παρακάτω σχέση:

$$Var [\bar{X}_{RSS(j)}] = \frac{1}{n} \left[ \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n \left( \tilde{X}_{(i)} - \frac{1}{j} \sum_{k=1}^j \bar{X}_{RSS(k)} \right)^2 \right]^{1/2} \quad (7.13)$$

ή

$$Var [\bar{X}_{RSS(j-k)}] = \frac{1}{n} \left[ \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n \left( \tilde{X}_{(i)} - \frac{1}{j-k} \sum_{l=1}^{j-k} \bar{X}_{RSS(l)} \right)^2 \right]^{1/2} \quad (7.14)$$



Μια παραδοχή που μπορούμε να κάνουμε, προς αποφυγή περαιτέρω πολυπλοκότητας (λόγω εξάρτησης της διακύμανσης από το  $\bar{X}_{RSS}$ ), είναι ότι το  $\bar{X}_{RSS}$  παραμένει σταθερό κατά τη διάρκεια των μετρήσεων. Πρόκειται για την παραδοχή να δεχόμαστε, όταν αναφερόμαστε σε πόσες παρατηρήσεις χρειάζονται για την ανίχνευση μιας μετατόπισης, ότι υπάρχει η ίδια μετατόπιση σε όλες αυτές τις παρατηρήσεις. Οπότε, μη κάνοντας στατιστικό σφάλμα, μπορούμε να δεχτούμε ότι:

$$\frac{1}{r} \sum_{j=1}^r \bar{X}_{RSS(j)} = \frac{1}{r-j} \sum_{l=1}^{r-j} \bar{X}_{RSS(l)} = \bar{X}_{RSS}$$

Συνεπώς, για τη διασπορά του στατιστικού RSS-EWMA κατά τον  $r$ -οστό κύκλο δειγματοληψιών ισχύει:

$$\begin{aligned} Var[Z_{(j)}] &= \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} Var[\bar{X}_{RSS(j-k)}] = \\ &= \lambda^2 \sum_{k=0}^{j-1} (1-\lambda)^{2k} \frac{1}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) = \\ &= \frac{\lambda^2}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) \sum_{k=0}^{j-1} (1-\lambda)^{2k} = \\ &= \frac{\lambda^2}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) \left( \frac{1 - (1-\lambda)^{2j}}{1 - (1-\lambda)^2} \right) = \\ &= \frac{\lambda^2}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) \left( \frac{1 - (1-\lambda)^{2j}}{-\lambda^2 + 2\lambda} \right) = \\ &= \frac{\lambda^2}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) \left( \frac{1 - (1-\lambda)^{2j}}{\lambda(2-\lambda)} \right) \end{aligned}$$

Άρα, βρίσκουμε ότι:

$$Var[Z_{(j)}] = \frac{\lambda}{2-\lambda} (1 - (1-\lambda)^{2j}) \frac{1}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \bar{X}_{RSS})^2 \right) \quad (7.15)$$

Όταν η διεργασία είναι εντός ελέγχου, τότε  $\bar{X}_{RSS} = \mu_0$ , ίσο δηλαδή με την τιμή στόχο. Οπότε:

$$Var[Z_{(j)}] = \frac{\lambda}{2-\lambda} (1 - (1-\lambda)^{2j}) \frac{1}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \mu_0)^2 \right) \quad (7.16)$$

Παρατηρούμε ότι η διακύμανση των διαγραμμάτων RSS-EWMA είναι μικρότερη από ότι αυτή των κλασικών διαγραμμάτων EWMA. Στην περίπτωση μάλιστα που υπάρχει μετατόπιση, δηλαδή οι επιμέρους μετρήσεις  $\tilde{X}_{(i)}$  δώσουν αρκετά διαφορετική τιμή από την τιμή στόχο για το μέσο  $\mu_0$ , τότε η διακύμανση του στατιστικού ελαττώνεται. Αυτό, σε συνδυασμό με την

αντίστοιχη μέτρηση της μετατόπισης από το στατιστικό RSS-EWMA, οδηγεί σε καλύτερη ανιχνευτική ικανότητα. Με την υπόψη τεχνική, δηλαδή, επιτυγχάνεται να συναρτησουμε την ανιχνευτική ικανότητα, πέρα από το στατιστικό, και με τα όρια του διαγράμματος.

Έχοντας ορίσει την κατανομή του στατιστικού RSS-EWMA, δηλαδή τη μέση τιμή και τη διασπορά του, μπορούμε να κατασκευάσουμε το Διάγραμμα Ελέγχου RSS-EWMA. Συγκεκριμένα, η κεντρική γραμμή  $CL_{RSS}$ , όπως επίσης το ανώτερο  $UCL_{RSS}$  και κατώτερο όριο  $LCL_{RSS}$ , δίνονται από τις παρακάτω σχέσεις:

$$LCL_{RSS} = \mu_0 - L \sqrt{\frac{\lambda}{(2-\lambda)}(1 - (1-\lambda)^{2j}) \frac{1}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \mu_0)^2 \right)} \quad (7.17)$$

$$CL_{RSS} = \mu_0 \quad (7.18)$$

$$UCL_{RSS} = \mu_0 + L \sqrt{\frac{\lambda}{(2-\lambda)}(1 - (1-\lambda)^{2j}) \frac{1}{n} \left( \hat{\sigma}_{RSS}^2 - \frac{1}{n} \sum_{i=1}^n (\tilde{X}_{(i)} - \mu_0)^2 \right)} \quad (7.19)$$

### 7.2.3 Ανιχνευτική ικανότητα Διαγραμμάτων RSS-EWMA

Μελέτη της ανιχνευτικής ικανότητας των υπόψη διαγραμμάτων έγινε από τους Abujiya και Lee (2013) [4], όπου υπολογίστηκε η μέση τιμή του ARL για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 500.000 επαναλήψεις. Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στον Πίνακα 7.1 παρακάτω, όπου συγκρίνεται η ανιχνευτική ικανότητα των απλών διαγραμμάτων SRS-EWMA με αυτή των διαγραμμάτων RSS-EWMA. Οι τιμές ARL που παρατίθενται αποτελούν ουσιαστικά τις μέσες τιμές για κάθε μετατόπιση, καθώς λαμβάνεται υπόψη όλο το εύρος των πιθανών τιμών διασποράς των παρατηρήσεων. Παρατηρούμε ότι τα διαγράμματα RSS-EWMA οδηγούν ταχύτερα σε εκτός ελέγχου σήμα.

Με τη βοήθεια του κώδικα του Παραρτήματος Β (υπολογισμός ARL με μέθοδο Μαρκοβιανής Αλυσίδας), θα επιχειρήσουμε να υπολογίσουμε την ανιχνευτική ικανότητα του διαγράμματος RSS-EWMA για διάφορες τιμές της διασποράς των παρατηρήσεων, προκειμένου να μετρηθεί και να προσδιοριστεί εύκολα η επίδραση της αύξησης της διασποράς των παρατηρήσεων στην ανιχνευτική ικανότητα. Συγκεκριμένα, υπολογίζουμε την ανιχνευτική ικανότητα (ARL) για διάφορους συνδυασμούς τιμών διασποράς παρατηρήσεων, τιμών των παραμέτρων και τιμών συνολικής διασποράς του πληθυσμού.

Παραδοχές που γίνονται, λόγω της φύσης του κώδικα που γράφτηκε, είναι ότι η μέση μετατόπιση παραμένει σταθερή από την έναρξη μετρήσεων και ότι, χωρίς περιορισμό της γενικότητας, ο δειγματικός μέσος και η διασπορά ανά σετ δειγματοληψίας παραμένουν σταθερά. Τα αποτελέσματα παρατίθενται στους Πίνακες 7.2, 7.3, 7.4, 7.5, 7.6 και 7.7.

Οι υπολογισμοί έγιναν για μέγεθος υποσυνόλου (σετ) δείγματος  $n=4$ , για έξι ζεύγη

Πίνακας 7.1: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων SRS-EWMA και RSS-EWMA

$\delta$	<i>SRS – EWMA</i> $\lambda=0.12, L=2.75$		<i>RSS – EWMA</i> $\lambda=0.12, L=2.75$			
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
0.00	370.4	370.1	370.0	369.8	370.2	369.9
0.25	98.0	70.8	56.3	47.4	41.3	36.1
0.50	29.7	21.3	17.1	14.4	12.7	11.4
0.75	14.8	11.1	9.2	7.9	7.1	6.5
1.00	9.6	7.4	6.3	5.5	4.9	4.6
1.50	5.6	4.5	3.9	3.4	3.2	2.9
2.00	4.0	3.3	2.9	2.6	2.4	2.2
2.50	3.2	2.6	2.3	2.1	2.0	1.9
3.00	2.6	2.2	2.0	1.9	1.8	1.7
4.00	2.1	1.8	1.6	1.4	1.2	1.1

Πίνακας 7.2: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.35$  και  $L=2.94863$

$\delta$	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
0.00	370.0781237	370.0781237	370.0781237
0.25	37.93277536	47.77374913	49.03427489
0.50	7.17584604	9.33507773	9.62464925
0.75	2.66257696	3.44985958	3.5551698
1.00	1.28510032	1.70083347	1.75497817
1.25	0.63111129	0.91231148	0.94841103
1.50	0.27422023	0.46306541	0.48858097
1.75	0.09902143	0.20791291	0.22402282
2.00	0.02790981	0.07933712	0.08795686

τιμών των παραμέτρων  $\lambda$  και  $L$  και πληθυσμιακή διασπορά  $\sigma_0=1$ . Τα αποτελέσματα αφορούν διάγραμμα EWMA με χρονικά μεταβαλλόμενα όρια. Επίσης θεωρήθηκαν τρεις τιμές της διασποράς των παρατηρήσεων ( $Var_{sample}$ ) ως προς το δειγματικό μέσο, δηλαδή της ποσότητας  $\frac{1}{n} \sum_{i=1}^n (\mu_i - \mu_1)^2$ , προκειμένου να γίνει αντιληπτή η επίδραση της αύξησης της δειγματικής διασποράς. Ως  $Var_{sample}$  εννοείται η διασπορά των τιμών  $\mu_{i:i}$ , δηλαδή των μετρήσεων που γίνονται στα δείγματα που λαμβάνονται, σε σχέση με τη μέση τιμή τους  $\mu_1$ , ήτοι:

$$Var_{sample} = \frac{1}{n} \sum_{i=1}^n (\mu_i - \mu_1)^2$$

Από τους υπόψη πίνακες συμπεραίνουμε ότι όσο αυξάνεται η διασπορά των παρατηρήσεων κάθε σετ ως προς το δειγματικό μέσο ( $Var_{sample}$ ), τόσο βελτιώνεται η ανιχνευτική ικανότητα του διαγράμματος. Μπορούμε να δούμε και γραφικά το συμπέρασμα αυτό, ενδεικτικά για δύο

Πίνακας 7.3: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.25$  και  $L=2.9015$

	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
$\delta$			
0.00	370.0719807	370.0719807	370.0719807
0.25	29.88930827	37.4835155	38.465627
0.50	6.55034855	8.2495456	8.4741631
0.75	2.66529749	3.4077556	3.5038295
1.00	1.26376551	1.6929531	1.7491595
1.25	0.64128595	0.897268	0.9314599
1.50	0.32499937	0.4942554	0.516536
1.75	0.14411022	0.2611031	0.2769927
2.00	0.05079209	0.1203158	0.1308763

Πίνακας 7.4: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.24$  και  $L=2.89326$

	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
$\delta$			
0.00	370.0062032	370.0062032	370.0062032
0.25	29.3281349	36.7099114	37.6649822
0.50	6.5591692	8.2378066	8.459133
0.75	2.7390412	3.4534083	3.547049
1.00	1.4380001	1.8351422	1.886682
1.25	0.8007368	1.0786803	1.1137581
1.50	0.418245	0.6273833	0.6542625
1.75	0.1907202	0.3380774	0.3580349
2.00	0.0706402	0.1605172	0.173942

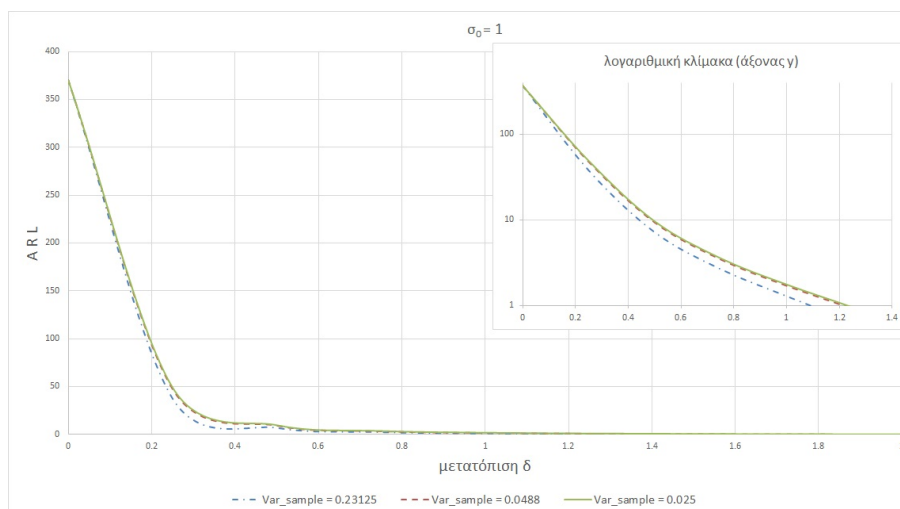
από τις τιμές της παραμέτρου  $\lambda$  που χρησιμοποιήσαμε στους παραπάνω υπολογισμούς, έστω  $\lambda=0.14$  και  $\lambda=0.35$ , και να πάρουμε δύο διαγράμματα των Σχημάτων 7.4 και 7.5 αντίστοιχα, που στο καθένα συγκρίνεται το ARL που αντιστοιχεί σε διάφορες τιμές της  $Var_{sample}$ .

Πίνακας 7.5: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.2$  και  $L=2.8632$

	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
$\delta$			
0.00	370.0330833	370.0330833	370.0330833
0.25	26.09906957	32.5622798	33.401122
0.50	6.19689992	7.6872967	7.8829288
0.75	2.6814587	3.3692528	3.4575952
1.00	1.34320492	1.7629959	1.8169742
1.25	0.68420444	0.9656268	1.0019926
1.50	0.32468321	0.5164988	0.5420351
1.75	0.1340248	0.2549507	0.2720813
2.00	0.04498493	0.1106654	0.1209788

Πίνακας 7.6: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.14$  και  $L=2.7895$

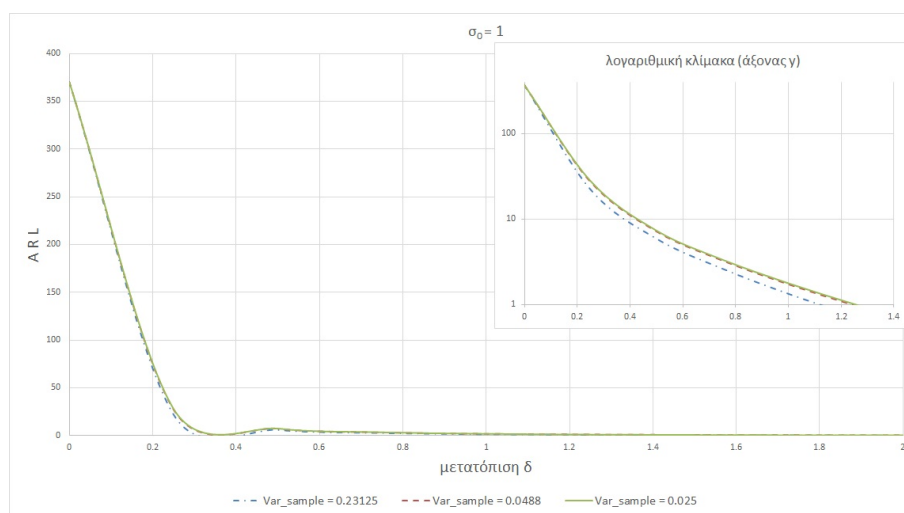
	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
$\delta$			
0.00	370.0152417	370.0152416	370.0152416
0.25	22.1745385	27.3241903	27.9930593
0.50	5.86100157	7.1654465	7.3343801
0.75	2.60458057	3.2654532	3.3495898
1.00	1.32067776	1.7177533	1.7694497
1.25	0.7223019	0.9742577	1.0071166
1.50	0.38862822	0.5710144	0.5943284
1.75	0.18141821	0.317072	0.3350186
2.00	0.06724664	0.1529387	0.1656227



Σχήμα 7.4: ARL διαγραμμάτων RSS-EWMA για  $\lambda=0.14$  και τρεις διαφορετικές διασπορές των τιμών δειγματοληψίας

Πίνακας 7.7: Ανιχνευτικής Ικανότητα Διαγράμματος RSS-EWMA για  $\lambda=0.14$  και  $L=2.7895$

	$Var_{sample}$ 0.23125	$Var_{sample}$ 0.0488	$Var_{sample}$ 0.025
$\delta$			
0.00	370.0152417	370.0152416	370.0152416
0.25	22.1745385	27.3241903	27.9930593
0.50	5.86100157	7.1654465	7.3343801
0.75	2.60458057	3.2654532	3.3495898
1.00	1.32067776	1.7177533	1.7694497
1.25	0.7223019	0.9742577	1.0071166
1.50	0.38862822	0.5710144	0.5943284
1.75	0.18141821	0.317072	0.3350186
2.00	0.06724664	0.1529387	0.1656227



Σχήμα 7.5: ARL διαγραμμάτων RSS-EWMA για  $\lambda=0.35$  και τρεις διαφορετικές διαφορές των τιμών δειγματοληψίας

Στους υπολογισμούς ARL που πραγματοποιήσαμε μέχρι τώρα, θεωρήσαμε τη διασπορά του πληθυσμού  $\sigma_0=1$ . Μπορούμε, πάντως, να κάνουμε υπολογισμούς κάνοντας χρήση του κώδικα του Παραρτήματος Β, και για τιμές πληθυσμιακής διασποράς ( $\sigma_0$ ) διαφορετικές από 1. Συλλέχθηκαν, έτσι, δεδομένα ανάλογα των Πινάκων 7.2 ως και 7.7, τα οποία κρίθηκε σκόπιμο να μην παρατεθούν. Τα συμπεράσματα, ωστόσο, από το σύνολο των αποτελεσμάτων είναι τα εξής:

- Περίπτωση όπου η πληθυσμιακή διασπορά είναι  $\sigma_0=1$ :

- Για μικρές μετατοπίσεις, όσο μειώνεται το  $\lambda$  τόσο αυξάνεται η ανιχνευτική ικανότητα. Αυτό αφορά για μετατοπίσεις μικρότερες ή ίσες με  $\delta=0.5$ .

- Αντίθετα, για  $\delta \geq 0.75$  (μεσαίες και μεγάλες μετατοπίσεις), η μείωση του  $\lambda$  δεν οδηγεί, αρχικά τουλάχιστον, σε μείωση της ανιχνευτικής ικανότητας των διαγραμμάτων.

- Περίπτωση όπου η πληθυσμιακή διασπορά  $\sigma_0$  είναι διαφορετική από 1:

- Για πληθυσμιακή διασπορά της τάξης  $\sigma_0=0.5$  παρατηρούμε ότι πάλι για μικρές μετατοπίσεις, μικρότερες ή ίσες με  $\delta=0.5$ , η ανιχνευτική ικανότητα αυξάνεται όσο μειώνεται το  $\lambda$ . Αυτό φαίνεται όμως να ισχύει μόνο για μικρές τιμές δειγματικής διασποράς ( $Var_{sample}=0.025$  και  $0.05$ ). Σε αντίστοιχα αποτελέσματα καταλήξαμε και για άλλες τιμές  $\sigma_0 < 1$ .

- Αυξάνοντας την τιμή της πληθυσμιακής διασποράς,  $\sigma_0=3$ , παρατηρούμε ότι για όλες τις μετατοπίσεις η ανιχνευτική ικανότητα αυξάνεται όσο μειώνεται το  $\lambda$ .

Για μικρές τιμές του  $\lambda$  η αύξηση της δειγματικής διασποράς οδηγεί σε ταχύτερη ανιχνευτική ικανότητα, γεγονός που επαληθεύει τη σχεδίαση του υπόψη τύπου καθώς οδηγεί σε βελτίωση της ανιχνευτικής ικανότητας των διαγραμμάτων EWMA. Στο αντίθετο αποτέλεσμα καταλήγουμε για μεγάλες τιμές του  $\lambda$ , όπως παρατηρούμε από τους Πίνακες 7.2 και 7.3. Έτσι, συγκεκριμένα, για μετατοπίσεις μικρότερες ίσες με 1, η αύξηση της διασποράς, καθώς πηγαίνουμε από τη στήλη 1 στην στήλη 3, για παράδειγμα για  $\delta=1$ , το  $ARL_1=1.950442$  στη συνέχεια για μεγαλύτερη διασπορά γίνεται  $ARL_1=1.885664$  και με περαιτέρω αύξηση της δειγματικής διασποράς το μέσο μήκος ροής ελαττώνεται ακόμη περισσότερο σε  $ARL_1=1.790467$ . Ιδιαίτερο ενδιαφέρον παρουσιάζει το RSS-EWMA για μεγάλες τιμές του  $\lambda$  και για μετατοπίσεις μεγαλύτερες από 1.25, καθώς για μεσαίες τιμές της δειγματικής διασποράς το  $ARL_1$  παρουσιάζει μέγιστο, ενώ για μικρές και μεγάλες τιμές της δειγματικής διασποράς το  $ARL_1$  είναι μικρότερο και μάλιστα για μικρότερες τιμές της δειγματικής διασποράς το  $ARL_1$  είναι μικρότερο. Αυτό δεν είναι επιθυμητό, καθώς τα διαγράμματα RSS-EWMA σχεδιάστηκαν με τη λογική η μεγαλύτερη διασπορά των τιμών της δειγματικής διασποράς να οδηγεί σε καλύτερη ανιχνευτική ικανότητα.

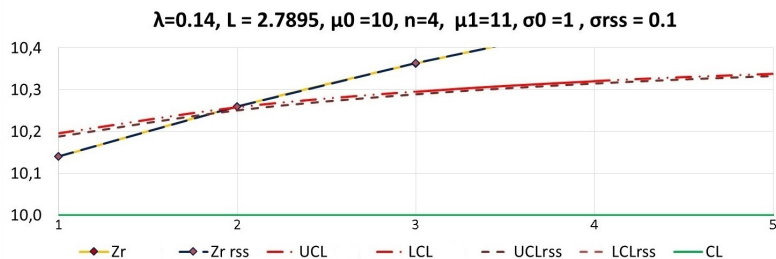
#### 7.2.4 Εφαρμογή Τεχνικής EWMA σε RSS δεδομένα

Οι σχέσεις στις οποίες θα εφαρμόσουμε τα δεδομένα, προκειμένου να προκύψει διάγραμμα RSS-EWMA είναι οι Σχέσεις (7.17) και (7.19). Για να υπάρχει μέτρο σύγκρισης με τα προηγούμενα παραδείγματα, θα χρησιμοποιήσουμε τα δεδομένα των Πινάκων A.6 και A.7 του Παραρτήματος A. Στα δεδομένα αυτά, τα  $\bar{X}_{RSS(j)}$  που υπολογίζονται είναι ίδια με τα  $X_i$  του Πίνακα A.1, προερχόμενα από σετ δειγματοληψίας τεσσάρων παρατηρήσεων ( $n=4$ ). Στους Πίνακες A.6 και A.7 οι παρατηρήσεις σε κάθε σετ είναι ταξινομημένες.

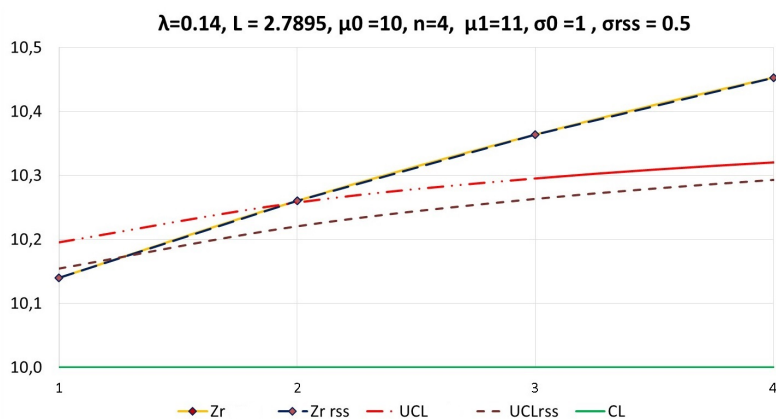
Κατασκευάζουμε, λοιπόν, το διάγραμμα RSS-EWMA που φαίνεται στο Σχήμα 7.6, επιλέγοντας παραμέτρους  $\lambda=0.14$  και  $L=2.7895$  (οι οποίες δίνουν  $ARL_0=370$ ).







Σχήμα 7.7: Σύγκριση διαγραμμάτων RSS-EWMA και EWMA



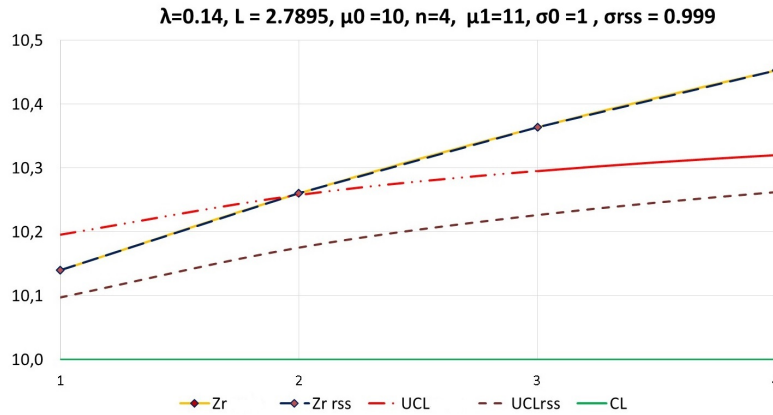
Σχήμα 7.8: Σύγκριση διαγραμμάτων RSS-EWMA και EWMA

πραγματοποιείται κάποια ταξινόμηση (κατάταξη με αύξουσα σειρά βάσει μιας μεταβλητής  $X$ ) με απλό και χωρίς κόστος τρόπο. Στην πραγματικότητα όμως, η ταξινόμηση αυτή ενδέχεται να μην είναι απόλυτα ακριβής, να περιέχει κάποια σφάλματα, δηλαδή, που συνεπακόλουθα μεταβιβάζονται στη μεταβλητή  $Y$ . Αυτό ονομάζεται ατελής ταξινόμηση (Imperfect Ranking). Λόγω του ότι η ταξινόμηση πραγματοποιείται στην μεταβλητή  $Y$ , η ακρίβεια της διαδικασίας εξαρτάται από τη συσχέτιση  $\rho$  μεταξύ των μεταβλητών  $X$  και  $Y$ , με τις ακραίες περιπτώσεις της τέλει ταξινόμησης και της απλής τυχαίας δειγματοληψίας να αντιστοιχούν στις τιμές  $\rho=1$  και  $\rho=0$  αντίστοιχα.

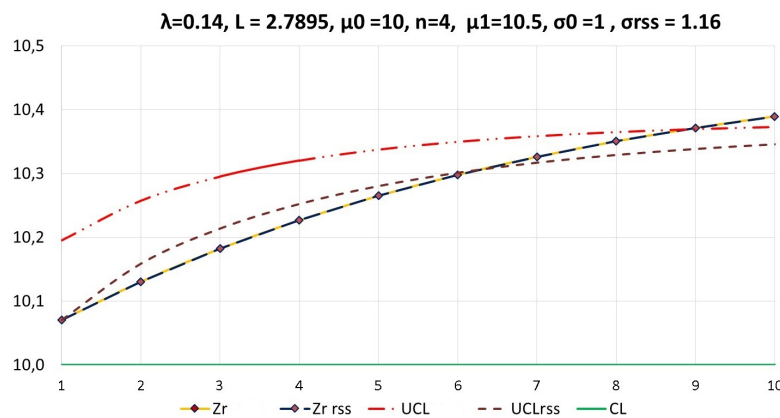
Το μαθηματικό υπόβαθρο για την υπόψη περίπτωση παρέχεται από τους Muttalak και Al-Sabah (2003) [39]. Έτσι, θεωρώντας το  $(X, Y)$  ένα διάνυσμα μιας διμεταβλητής κανονικής κατανομής και υποθέτοντας ότι η παλινδρόμηση των  $X$  και  $Y$  είναι γραμμική, θα ισχύει:

$$X_{[i:n]j} = \mu_x + \frac{\rho\sigma_x}{\sigma_y} (Y_{[i:n]j} - \mu_y) + \epsilon_{ij} \quad (7.20)$$

όπου  $Y_{[i:n]j}$  και  $X_{[i:n]j}$  είναι η  $i$ -οστή μικρότερη τιμή του  $Y$  και η αντίστοιχη τιμή του  $X$ , τις οποίες παίρνουμε από το  $i$ -οστό σετ του  $j$ -οστού κύκλου ( $i=1, 2, \dots, n$  και  $j=1, 2, \dots, r$ ). Οι όροι  $\mu_x, \sigma_x, \mu_y$  και  $\sigma_y$  είναι οι μέσοι και οι τυπικές αποκλίσεις των μεταβλητών  $X$  και  $Y$  αντίστοιχα,  $\rho$  είναι η συσχέτιση μεταξύ τους και  $\epsilon_{ij}$  είναι ο όρος σφάλματος. Επισημαίνεται ότι  $Y$  και  $\epsilon_{ij}$  είναι ανεξάρτητες μεταβλητές, με το  $\epsilon_{ij}$  να έχει μέσο 0 και διασπορά  $\sigma_{\epsilon}^2 = (1 - \rho^2)$ .



Σχήμα 7.9: Σύγκριση διαγραμμάτων RSS-EWMA και EWMA



Σχήμα 7.10: Σύγκριση διαγραμμάτων RSS-EWMA και EWMA

Για τον μέσο της μεταβλητής ενδιαφέροντος  $X$ , με ταξινόμηση που βασίζεται στη συνακόλουθη μεταβλητή  $Y$  (μέθοδος IRSS), ο αμερόληπτος εκτιμητής του πληθυσμιακού μέσου  $\mu$  μπορεί να γραφεί ως εξής:

$$\bar{X}_{IRSS(j)} = \frac{1}{n} \sum_{i=1}^n X_{[i:n]j}, \quad j = 1, 2, \dots, r \quad (7.21)$$

Για τη δε διακύμανση του  $\bar{X}_{IRSS(j)}$  θα ισχύει η ακόλουθη σχέση:

$$Var[\bar{X}_{IRSS(j)}] = \frac{\sigma_x^2}{n} \left[ (1 - \rho^2) + \frac{\rho^2}{n} \sum_{i=1}^n \sigma_{z(i:n)}^2 \right] \quad (7.22)$$

όπου  $\sigma_{z(i:n)}^2$  είναι η διακύμανση του στατιστικού  $i$ -οστής τάξης σε ένα δείγμα μεγέθους  $n$  που πάρθηκε από κανονική κατανομή με τεχνική RSS.

Οι παραπάνω σχέσεις προφανώς αναφέρονται σε έναν, τον  $j$ -οστό, κύκλο δειγματοληψίας. Περαιτέρω ανάλυση για τον υπολογισμό του στατιστικού EWMA με τεχνική IRSS (υπολογισμός  $E[Z_{(j)}]$  και  $Var[Z_{(j)}]$ ) δεν θα γίνει. Θα ακολουθήσει όμως μια παράσταση της συμπεριφοράς του ARL των διαγραμμάτων IRSS-EWMA.

### 7.2.6 Ανιχνευτική ικανότητα Διαγραμμάτων IRSS-EWMA

Μελέτη της ανιχνευτικής ικανότητας των υπόψη διαγραμμάτων, όπως είδαμε και παραπάνω για της περίπτωση των διαγραμμάτων RSS-EWMA, έγινε από τους Abujiya και Lee (2013) [4], όπου υπολογίστηκε η μέση τιμή του ARL για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 500.000 επαναλήψεις. Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στους Πίνακες 7.8, 7.9, 7.10, 7.11, 7.12 παρακάτω, όπου συγκρίνεται η ανιχνευτική ικανότητα των διαγραμμάτων IRSS-EWMA για διάφορες τιμές της παραμέτρου συσχέτισης  $\rho$ .

Πίνακας 7.8: Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για  $n=2$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	95.9	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.2	94.9	29.2	14.6	9.5	5.5	4.0	3.1	2.6	2.0
0.50	372.7	89.8	27.6	14.0	9.1	5.3	3.8	3.0	2.5	2.0
0.75	372.7	82.6	24.8	12.8	8.4	5.0	3.6	2.9	2.4	1.9
1.00	370.1	70.8	21.3	11.1	7.4	4.5	3.3	2.6	2.2	1.8

Πίνακας 7.9: Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για  $n=3$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	96.3	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.0	94.1	28.8	14.5	9.4	5.5	3.9	3.1	2.6	2.0
0.50	372.7	87.1	26.5	13.5	8.8	5.2	3.7	3.0	2.5	2.0
0.75	372.6	75.3	22.6	11.7	7.7	4.6	3.4	2.7	2.3	1.9
1.00	369.2	56.4	17.0	9.1	6.2	3.8	2.8	2.3	2.0	1.6

Πίνακας 7.10: Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για  $n=4$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	96.2	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	372.9	93.5	28.6	14.5	9.4	5.5	3.9	3.1	2.6	2.0
0.50	371.5	85.2	25.9	13.2	8.6	5.1	3.7	2.9	2.4	2.0
0.75	371.8	70.5	21.1	11.0	7.3	4.4	3.2	2.6	2.2	1.8
1.00	369.8	47.4	14.4	7.9	5.5	3.4	2.6	2.1	1.9	1.4

Πίνακας 7.11: Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για  $n=5$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	370.6	96.1	29.7	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.4	93.0	28.5	14.4	9.3	5.5	3.9	3.1	2.6	2.0
0.50	371.0	83.6	25.3	13.0	8.5	5.0	3.6	2.9	2.4	2.0
0.75	372.9	67.2	20.2	10.6	7.1	4.3	3.2	2.5	2.2	1.8
1.00	369.4	40.9	12.7	7.1	5.0	3.2	2.4	2.0	1.8	1.2

Πίνακας 7.12: Ανιχνευτικής Ικανότητα Διαγράμματος IRSS-EWMA για  $n=6$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	372.0	96.0	29.6	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	373.2	92.9	28.4	14.3	9.3	5.5	3.9	3.1	2.6	2.0
0.50	372.6	82.9	25.1	12.9	8.4	5.0	3.6	2.9	2.4	2.0
0.75	372.3	64.9	19.4	10.3	6.9	4.2	3.1	2.5	2.1	1.8
1.00	369.9	36.1	11.4	6.5	4.6	2.9	2.2	1.9	1.7	1.1

### 7.3 Τεχνική MRSS (Median Ranked Set Sampling)

Η ιδέα της τεχνικής δειγματοληψίας MRSS αρχικά παρουσιάστηκε από τον Muttlak (1997) [40], με σκοπό τον περιορισμό των σφαλμάτων κατά την ταξινόμηση. Κατά την τεχνική αυτή, όπως και στην RSS, από ένα πληθυσμό λαμβάνεται ένα σύνολο  $n^2$  μονάδων δειγματοληψίας, χωρισμένων σε  $n$  υποσύνολα (σετ). Αντίστοιχα, τα δείγματα εντός του κάθε σετ κατατάσσονται με αύξουσα σειρά με απλό τρόπο (π.χ. οπτικά). Μετά την κατάταξη ακολουθεί επιλογή διαφορετικής τάξης μονάδας δειγματοληψίας, όπως περιγράφεται στη συνέχεια. Η πραγματική μέτρηση γίνεται στο σύνολο  $n$  μονάδων δειγματοληψίας που προκύπτει μ' αυτόν τον τρόπο. Αυτή η διαδικασία επαναλαμβάνεται, με την κάθε επανάληψη να ονομάζουμε κύκλο.

Η μεθόδος MRSS μπορεί να περιγραφεί από τον ακόλουθο αλγόριθμο:

1ο βήμα: Τυχαία επιλογή ενός δείγματος μεγέθους  $n^2$  μονάδων από τον πληθυσμό.

2ο βήμα: Διαχωρισμός του δείγματος σε  $n$  σετ, καθένα από αυτά μεγέθους  $n$ .

3ο βήμα: Ταξινόμηση των μονάδων εντός κάθε σετ βάσει της μεταβλητής που μας ενδιαφέρει, με απλό τρόπο (χωρίς κοστοβόρα μέτρηση).

4ο βήμα:

- Αν το μέγεθος  $n$  είναι μονός αριθμός, τότε από κάθε ταξινομημένο σετ επιλέγουμε την παρατήρηση με ταξινόμηση  $(n + 1)/2$ , δηλαδή τη διάμεσο.
- Αν το μέγεθος  $n$  είναι ζυγός αριθμός, τότε από κάθε ταξινομημένο σετ επιλέγουμε:
  - για τις πρώτες  $n/2$  παρατηρήσεις: την παρατήρηση με ταξινόμηση  $n/2$  και
  - για τις επόμενες  $n/2$  παρατηρήσεις: την παρατήρηση με ταξινόμηση  $(n + 2)/2$ .

5ο βήμα: Μέτρηση των  $n$  επιλεγμένων μονάδων και υπολογισμός του μέσου τους (αυτή την ποσότητα θα τροφοδοτήσουμε στο στατιστικό του διαγράμματος ελέγχου).

Η παραπάνω διαδικασία επαναλαμβάνεται  $r$  φορές, δημιουργώντας δηλαδή ένα σύνολο από  $nr$  μονάδες MRSS δεδομένων.

Ένα καλό παράδειγμα για την κατανόηση της τεχνικής MRSS είναι η δειγματοληψία από πληθυσμό φυτών, όπου το χαρακτηριστικό (μεταβλητή) που μας ενδιαφέρει είναι το ύψος τους, όπως απεικονίζεται στα Σχήματα 7.11 και 7.12. Στο υπόψη παράδειγμα, σε κάθε κύκλο επιλέγουμε τυχαία είτε 9 φυτά ( $n=3$ ) είτε 16 φυτά ( $n=4$ ) από τον πληθυσμό, η ταξινόμηση των οποίων γίνεται οπτικά.

### 7.3.1 Εκτιμητές Τεχνικής MRSS

Αν το πλήθος των σετ είναι περιττό, συμβολίζουμε με  $X_{(i:m)j}$  την  $i$ -οστή διάμεσο του  $i$ -οστού σετ στον  $j$ -οστό κύκλο. Πρόκειται δηλαδή για το στατιστικό  $((n + 1)/2)$ -τάξης:  $X_{i(\frac{n+1}{2})}$ , όπου  $i=1, 2, \dots, n$ . Αν το πλήθος των σετ είναι άρτιο, θα έχουμε κατ' αντιστοιχία:

$$X_{i(\frac{n}{2})}, \text{ όπου } i=1, 2, \dots, L = n/2 \text{ (στατιστικό } (n/2)\text{-τάξης)}$$

και

$$X_{i(\frac{n+2}{2})}, \text{ όπου } i= L+1, L+2, \dots, n \text{ (στατιστικό } ((n + 2)/2)\text{-τάξης)}.$$

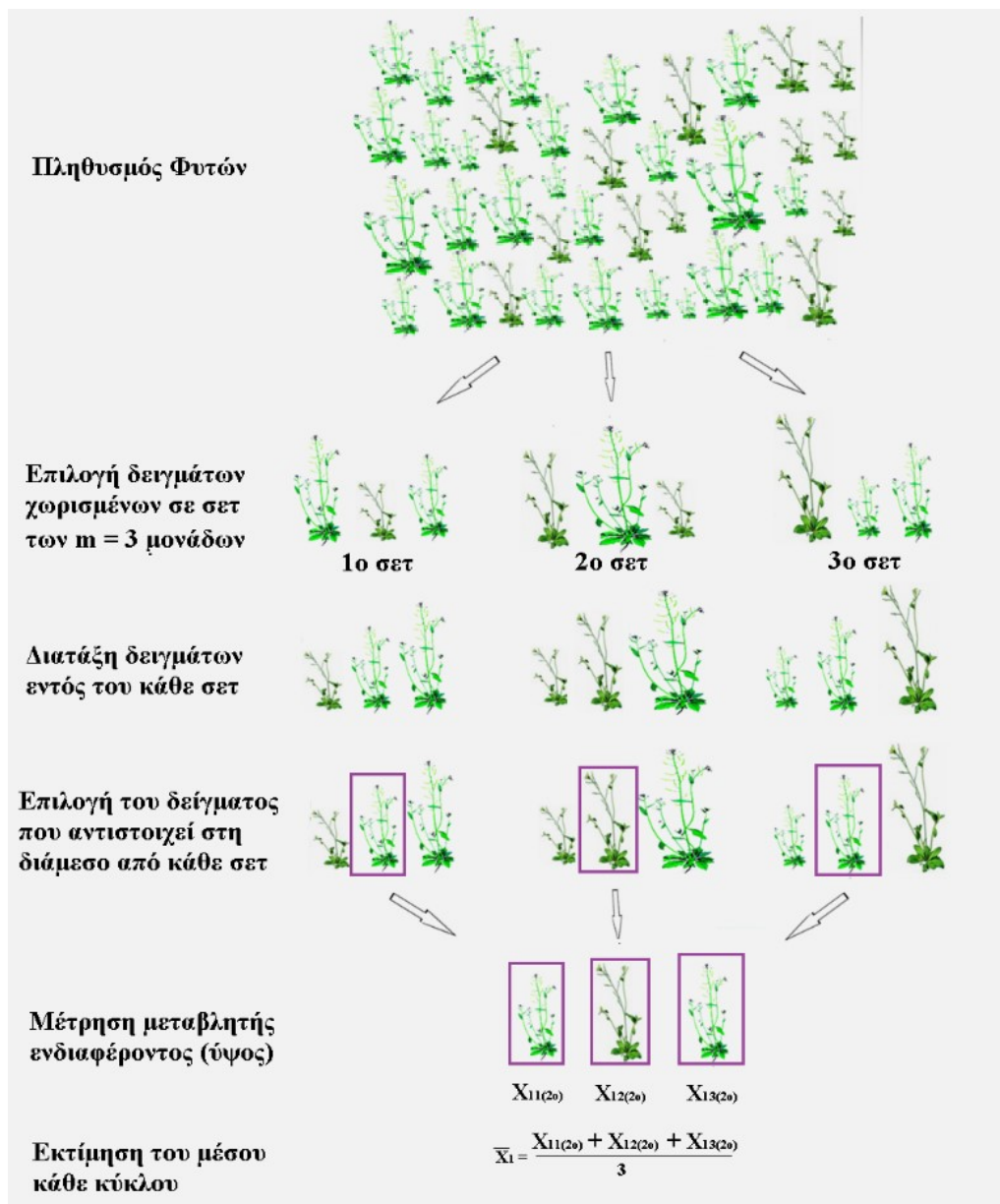
Ο συμβολισμός  $X_{(i:m)j}$  αντί  $X_{(i:n)j}$  (όπως θα αναμενόταν) χρησιμοποιείται απλά και μόνο για να γίνεται αντιληπτό ότι εφαρμόζουμε τεχνική MRSS.

Σύμφωνα με τους Muttlak και Al-Sabah (2003) [39], ο εκτιμητής για τον πληθυσμιακό μέσο στον  $j$ -οστό κύκλο για MRSS δεδομένα δίνεται από την παρακάτω σχέση:

$$\bar{X}_{MRSS(j)} = \frac{1}{n} \sum_{i=1}^n X_{(i:m)j} \quad (7.23)$$

Η δε διακύμανση του  $\bar{X}_{MRSS(j)}$  δίνεται από την εξής σχέση:

$$Var [\bar{X}_{MRSS(j)}] = \frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:m)}^2 \quad (7.24)$$



Σχήμα 7.11: Παράδειγμα Τεχνικής Δειγματοληψίας MRSS για περιττό πλήθος σετ

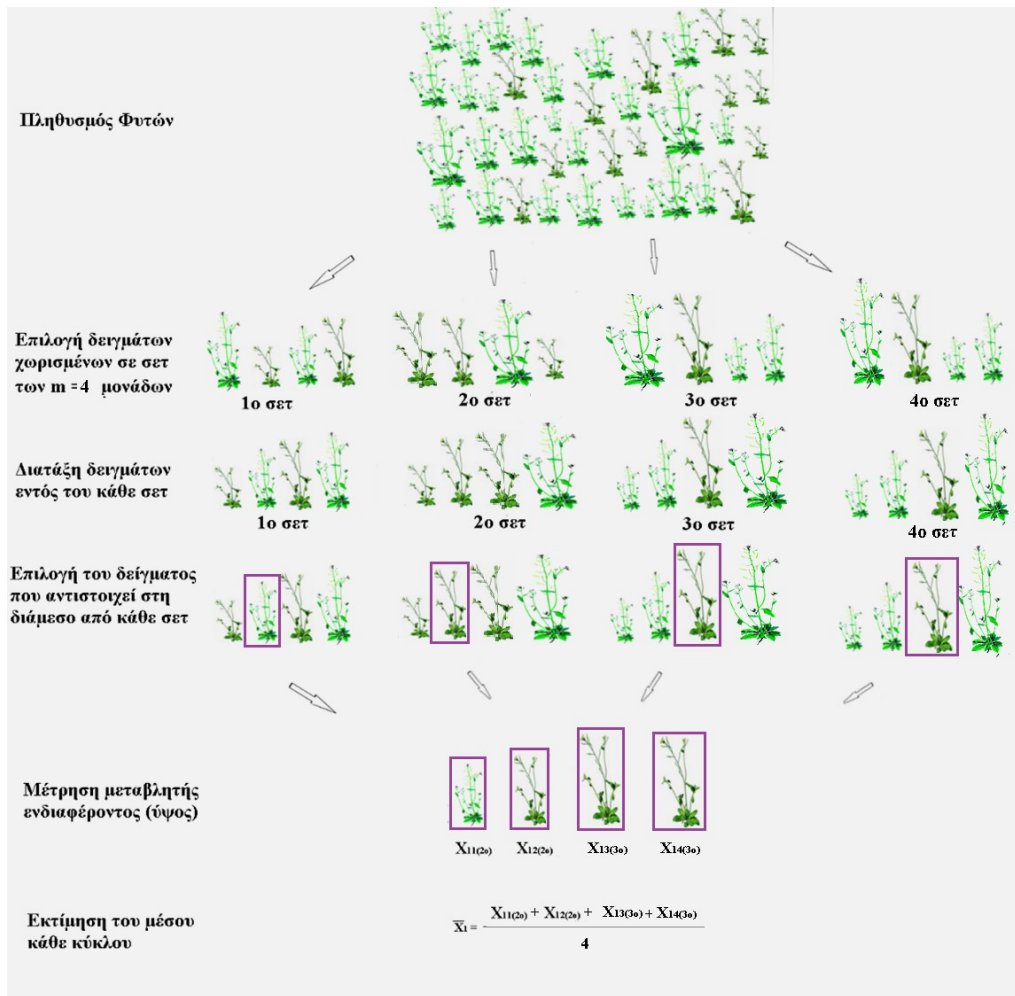
όπου

$$\sigma_{(i:m)}^2 = E[X_{(i:m)} - E(X_{(i:m)})]^2 \quad (7.25)$$

Οπότε, για όλους τους κύκλους δειγματοληψίας ( $j = 1, 2, \dots, r$ ) από την έναρξη ως τον  $r$ -οστό κύκλο, καθώς χρησιμοποιούμε MRSS δεδομένα, η αμερόληπτη εκτιμήτρια του πληθυσμιακού μέσου  $\bar{X}_{MRSS}$  δίνεται από τον τύπο:

$$\bar{X}_{MRSS} = \frac{1}{r} \sum_{j=1}^r \bar{X}_{MRSS(j)} \quad (7.26)$$

Για τη διακύμανση του  $\bar{X}_{MRSS}$ , τη διακύμανση δηλαδή της δειγματοληψίας MRSS, ο



Σχήμα 7.12: Παράδειγμα Τεχνικής Δειγματοληψίας MRSS για άρτιο πλήθος σετ

εκτιμητής που προτείνεται είναι ο εξής:

$$\hat{\sigma}_{\bar{X}_{MRSS}} = \left[ \frac{1}{n(nr - 1)} \sum_{j=1}^r \sum_{i=1}^n (X_{(i:m)j} - \bar{X}_{MRSS})^2 \right]^{1/2} \quad (7.27)$$

Από την παραπάνω Σχέση (7.27) βλέπουμε, κατ' αναλογία με την περίπτωση της τεχνικής RSS, ότι η διασπορά των παρατηρήσεων της τεχνικής MRSS αποτελεί συνάρτηση της μετατόπισης του μέσου της διαδικασίας.

Οι παραπάνω σχέσεις της τεχνικής MRSS μπορούν να εφαρμοστούν σε όλα τα διαγράμματα ελέγχου.

### 7.3.2 Διάγραμμα Ελέγχου EWMA με τεχνική MRSS

Ακολουθώντας την ίδια λογική υπολογισμών με την περίπτωση της τεχνικής RSS, καταλήγουμε στις παρακάτω σχέσεις:

$$E[Z_{(j)}] = 0 \quad (7.28)$$



$$Var[Z_{(j)}] = \frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2j}) \left[ \frac{\hat{\sigma}_{MRSS}^2}{n} - \frac{1}{n(nr-1)} \sum_{j=1}^r \sum_{i=1}^n (X_{(i:m)j} - \bar{X}_{MRSS})^2 \right]^{1/2} \quad (7.29)$$

Έχοντας ορίσει την κατανομή του στατιστικού MRSS-EWMA, δηλαδή τη μέση τιμή και τη διασπορά του, μπορούμε να κατασκευάσουμε το διάγραμμα ελέγχου MRSS-EWMA. Συγκεκριμένα, η κεντρική γραμμή  $CL_{MRSS}$ , όπως επίσης το ανώτερο  $UCL_{MRSS}$  και κατώτερο όριο  $LCL_{MRSS}$ , δίνονται από τις παρακάτω σχέσεις:

$$LCL_{MRSS} = \mu_0 - L \sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2j}) \left[ \frac{\hat{\sigma}_{MRSS}^2}{n} - \frac{1}{n(nr-1)} \sum_{j=1}^r \sum_{i=1}^n (X_{(i:m)j} - \bar{X}_{MRSS})^2 \right]} \quad (7.30)$$

$$CL_{MRSS} = \mu_0 \quad (7.31)$$

$$UCL_{MRSS} = \mu_0 + L \sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2j}) \left[ \frac{\hat{\sigma}_{MRSS}^2}{n} - \frac{1}{n(nr-1)} \sum_{j=1}^r \sum_{i=1}^n (X_{(i:m)j} - \bar{X}_{MRSS})^2 \right]} \quad (7.32)$$

### 7.3.3 Ανιχνευτική ικανότητα διαγραμμάτων MRSS-EWMA

Είναι προφανές από τις παραπάνω σχέσεις ότι τα όρια των διαγραμμάτων MRSS-EWMA είναι συνάρτηση του στατιστικού  $\bar{X}_{MRSS}$ . Σύμφωνα με τους Pongpullpronsak και Sontisamran (2013) [44], το ARL των διαγραμμάτων RSS και MRSS δεν εξαρτάται μόνο από τη μετατόπιση αλλά και από το μέγεθος του δείγματος.

Μικρότερη διασπορά εμφανίζει άλλοτε το στατιστικό  $\bar{X}_{SRS}$ , άλλοτε το στατιστικό  $\bar{X}_{RSS}$  και άλλοτε το στατιστικό  $\bar{X}_{MRSS}$ . Αυτό έχει ως αποτέλεσμα η ανιχνευτική ικανότητα να προβάλλει διαφορετικό τύπο διαγράμματος ανάλογα με το μέγεθος του δείγματος και την μετατόπιση που θέλουμε να εντοπίσουμε. Στους Πίνακες 7.13 και 7.14, για ευκολία ανάγνωσης, η καλύτερη ανιχνευτική ικανότητα ξεχωρίζει με γκρι φόντο.

Σημειώνεται ότι τα ARL που παρατίθενται στους εν λόγω πίνακες (ανιχνευτική ικανότητα) αφορούν διαγράμματα, όπου το στατιστικό που μετράται είναι το  $\bar{X}$ . Το πρώτο (SRS) είναι το διάγραμμα ελέγχου Shewhart. Ωστόσο, ανάλογη συμπεριφορά αναμένεται να παρουσιάζεται και στα διαγράμματα RSS-EWMA και MRSS-EWMA, σε σύγκριση με το απλό διάγραμμα EWMA.



Πίνακας 7.13: Ανιχνευτική Ικανότητα Διαγραμμάτων SRS, RSS και MRSS για  $n=3$  και  $n=4$

$\delta$	Μέγεθος δείγματος $n=3$			Μέγεθος δείγματος $n=4$		
	<i>SRS</i>	<i>RSS</i>	<i>MRSS</i>	<i>SRS</i>	<i>RSS</i>	<i>MRSS</i>
0.00	346.90	345.470	347.886	346.214	341.513	343.138
0.50	161.448	156.354	144.185	161.782	150.134	155.275
1.00	45.601	42.542	46.388	45.752	42.001	42.224
1.50	15.176	15.276	15.192	14.698	14.043	14.724
2.00	6.526	6.466	6.265	6.558	6.490	6.339
2.50	3.293	3.508	3.210	3.233	3.259	3.319
3.00	2.012	2.030	2.120	2.032	1.968	2.000

Πίνακας 7.14: Ανιχνευτική Ικανότητα Διαγραμμάτων SRS, RSS και MRSS για  $n=5$  και  $n=6$

$\delta$	Μέγεθος δείγματος $n=5$			Μέγεθος δείγματος $n=6$		
	<i>SRS</i>	<i>RSS</i>	<i>MRSS</i>	<i>SRS</i>	<i>RSS</i>	<i>MRSS</i>
0.00	348.782	345.915	350.471	349.488	343.289	343.184
0.50	155.978	152.849	150.679	158.048	160.457	142.864
1.00	46.69	45.363	39.822	44.872	42.295	43.596
1.50	15.144	15.65	14.219	14.394	14.471	14.422
2.00	6.611	6.394	5.909	6.726	6.15	6.164
2.50	3.224	3.285	3.177	3.394	3.35	3.176
3.00	2.044	2.004	2.008	2.008	2.116	2.017

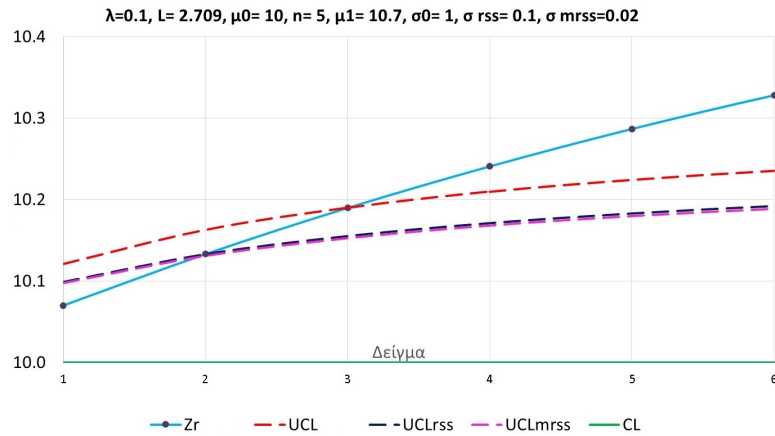
### 7.3.4 Εφαρμογή Τεχνικής EWMA σε MRSS δεδομένα

Κατ' αναλογία με την RSS περίπτωση, θα θεωρήσουμε πάλι μια υποθετική διαδικασία, όπου ο μέσος μετατοπίζεται και σε κάθε κύκλο δειγματοληψίας και τα  $\bar{X}_{RSS}$  και  $\bar{X}_{MRSS}$  είναι τα ίδια (η διασπορά δηλαδή παραμένει σταθερή σε κάθε κύκλο, υπόθεση που απλουστεύει τους υπολογισμούς). Με βάσει τις Σχέσεις (7.17) ως (7.19) και (7.30) ως (7.32), κάνοντας χρήση λογισμικού R, μπορούμε να δούμε γραφικά τη συμπεριφορά των διαγραμμάτων RSS-EWMA και MRSS-EWMA συγκριτικά με τα απλά EWMA καθώς οι διασπορές των εκάστοτε δειγματοληψιών μεταβάλλονται.

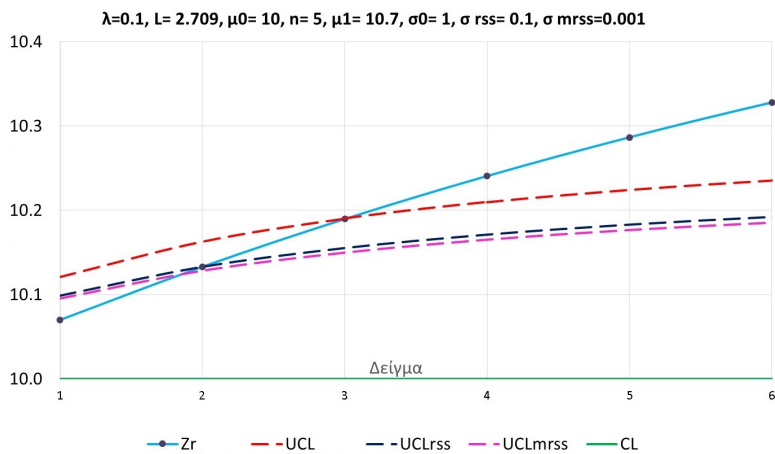
Θα διαπιστώσουμε πάντως πως όσο μικρότερη είναι η διασπορά των παρατηρήσεων της τεχνικής MRSS, τόσο βελτιωμένη ανιχνευτική ικανότητα παρουσιάζει.

Τα διαγράμματα στα Σχήματα 7.13 έως και 7.24 προκύπτουν για διάφορες τιμές των διασπορών  $\sigma_{MRSS}$  και  $\sigma_{RSS}$ , αλλά και για διάφορες τιμές των παραμέτρων EWMA ( $\lambda$  και  $L$ ). Θα παρατηρήσουμε ότι αύξηση της παραμέτρου  $\lambda$  οδηγεί σε μείωση της ανιχνευτικής ικανότητας για όλα τα διαγράμματα.

Συγκρίνοντας τα διαγράμματα MRSS-EWMA, RSS-EWMA και EWMA μπορούμε να πούμε ότι για μεγέθη δειγμάτων  $n=4$  φαίνεται προτιμητέο το διάγραμμα RSS-EWMA. Αντίθε-



Σχήμα 7.13: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.14: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

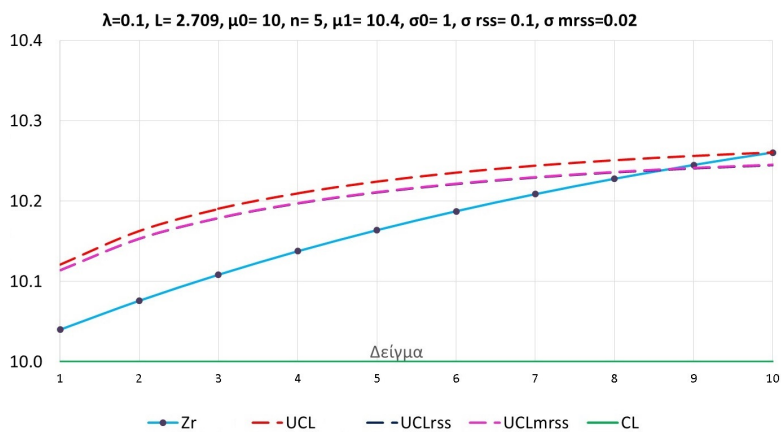
τα, για  $n=5$  παρατηρούμε ότι το διάγραμμα MRSS-EWMA εμφανίζει καλύτερη συμπεριφορά, με την προϋπόθεση ότι η διασπορά  $\sigma_{RSS}$  παραμένει σε χαμηλά επίπεδα.

Γενικά, τα διαγράμματα που θα δούμε στη συνέχεια (MDRSS, DMRSS, ODRSS) στοχεύουν στη μείωση της διασποράς των παρατηρήσεων και συνεπώς στη βελτίωση της ανιχνευτικής ικανότητάς τους.

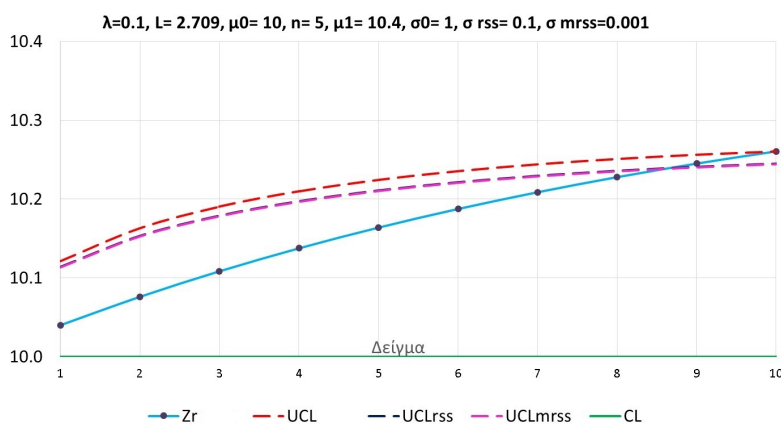
### 7.3.5 Διάγραμμα EWMA με Ατελή Ταξινόμηση Δειγματοληψίας MRSS (IMRSS)

Η λογική της ατελούς ταξινόμησης της δειγματοληψίας MRSS είναι ανάλογη εκείνης που αναπτύχθηκε για την περίπτωση RSS. Το μαθηματικό υπόβαθρο για την υπόψη περίπτωση παρέχεται από τον Muttlak (2003) [39], όπου για τη διακύμανση του  $\bar{X}_{IMRSS(j)}$  θα ισχύει η ακόλουθη σχέση:

$$Var[\bar{X}_{IMRSS(j)}] = \frac{\sigma_x^2}{n} \left[ (1 - \rho^2) + \frac{\rho^2}{n} \sum_{i=1}^n \sigma_{z(i:m)}^2 \right] \quad (7.33)$$



Σχήμα 7.15: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.16: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

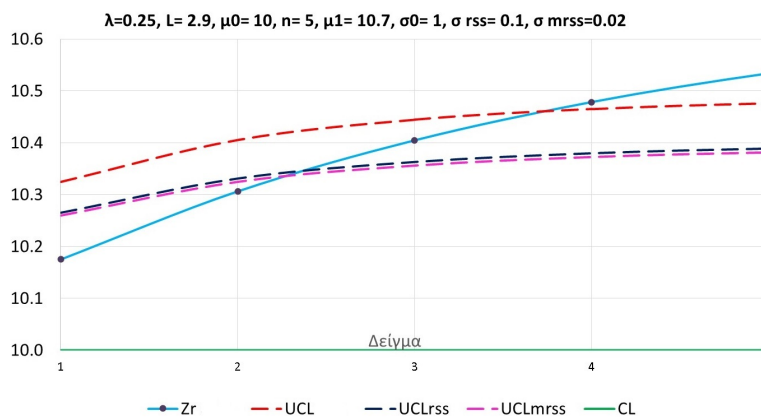
όπου  $\sigma_{z(i:m)}^2$  είναι η διακύμανση του στατιστικού  $i$ -οστής τάξης σε ένα δείγμα μεγέθους  $n$  που πάρθηκε από κανονική κατανομή με τεχνική MRSS.

### 7.3.6 Ανιχνευτική ικανότητα διαγραμμάτων IMRSS - EWMA

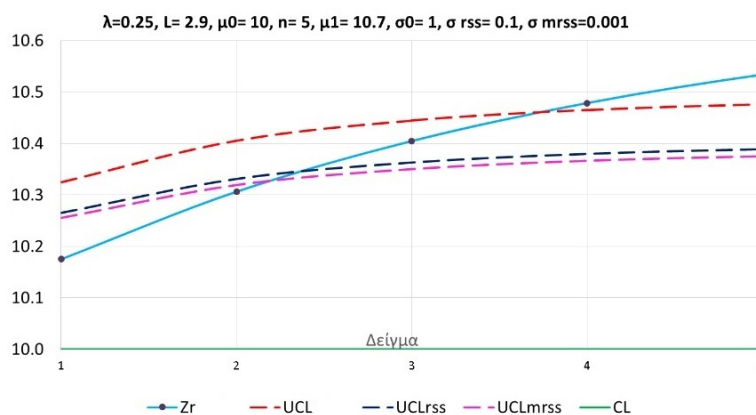
Μελέτη της ανιχνευτικής ικανότητας των υπόψη διαγραμμάτων έγινε από τους Abujiya, Lee και Riaz (2013) [6], όπου υπολογίστηκε το ARL για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 500.000 επαναλήψεις. Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στους Πίνακες 7.15, 7.16, 7.17, 7.18, 7.19 παρακάτω, όπου συγκρίνεται η ανιχνευτική ικανότητα των διαγραμμάτων IMRSS-EWMA για διάφορες τιμές της παραμέτρου συσχέτισης  $\rho$ .

## 7.4 Τροποποιήσεις Τεχνικών RSS

Εκτός από τις παραπάνω τεχνικές δειγματοληψίας που είδαμε ότι υπερέχουν της απλής τυχαίας δειγματοληψίας (SRS), έχουν αναπτυχθεί περαιτέρω τροποποιήσεις των RSS τεχνικών. Αυτές οι τεχνικές, που αναλύονται στη συνέχεια, είναι ακόμη πιο αποτελεσματικές στην εκτίμηση του πληθυσμιακού μέσου της διεργασίας που παρακολουθείται.



Σχήμα 7.17: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.18: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

### 7.4.1 Τεχνική DRSS (Double Ranked Set Sampling)

Η τεχνική δειγματοληψίας DRSS μελετήθηκε από τους Abujiya και Muttlak (2004) [7], ως ένας πιο αποτελεσματικός τρόπος εκτίμησης του πληθυσμιακού μέσου συγκριτικά με την μέθοδο RSS. Μπορεί να περιγραφεί από τον ακόλουθο αλγόριθμο:

1ο βήμα: Τυχαία επιλογή  $n^3$  μονάδων (παρατηρήσεις) από τον πληθυσμό.

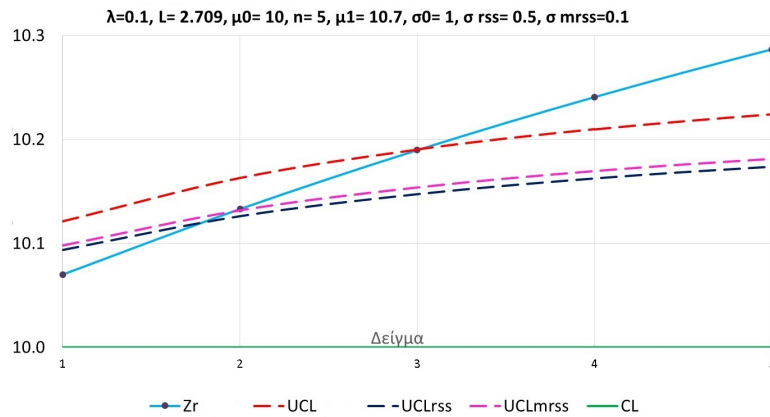
2ο βήμα: Τυχαία επιλογή  $n$  δειγμάτων (σύνολα παρατηρήσεων) από τον πληθυσμό, που το καθένα έχει μέγεθος  $n^2$  μονάδες.

3ο βήμα: Εφαρμογή της διαδικασίας RSS σε κάθε σύνολο, ώστε να προκύψουν  $n$  RSS σετ μεγέθους  $n$  μονάδων το καθένα, χωρίς κοστοβόρα μέτρηση.

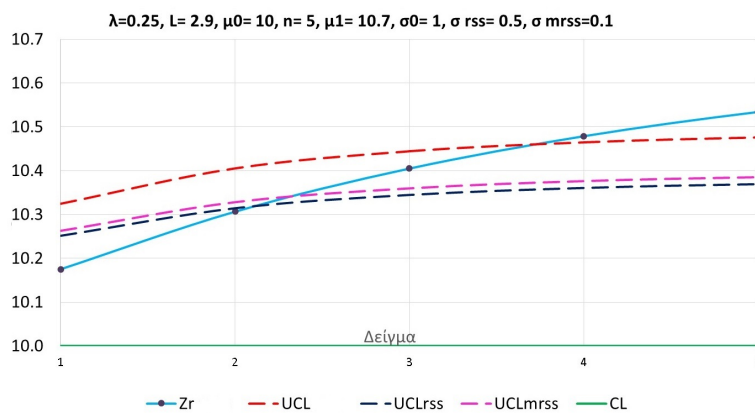
4ο βήμα: Εφαρμογή ξανά της διαδικασίας RSS στα παραπάνω σετ (RSS δεδομένα), ώστε να προκύψουν  $n$  μονάδες.

5ο βήμα: Μέτρηση των  $n$  επιλεγμένων μονάδων, που τις συμβολίζουμε  $Y_{(i:n)j}$  (με το  $j=1, 2, \dots, r$  να αντιστοιχεί στον κύκλο), και υπολογισμός του μέσου τους.

Η παραπάνω διαδικασία επαναλαμβάνεται  $r$  φορές (κύκλοι), δημιουργώντας δηλαδή ένα



Σχήμα 7.19: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.20: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

σύνολο από  $nr$  μονάδες DRSS δεδομένων.

Ας υποθέσουμε ότι η διεργασία που θέλουμε να παρακολουθήσουμε ακολουθεί κανονική κατανομή, με γνωστά  $\mu$  και  $\sigma$ . Τότε, στον  $j$ -οστό κύκλο ο μέσος των  $n$  DRSS δεδομένων θα δίνεται από τη σχέση:

$$\bar{Y}_{DRSS(j)} = \frac{1}{n} \sum_{i=1}^n Y_{(i:n)j} \quad (7.34)$$

ενώ η διασπορά τους θα δίνεται από τη σχέση:

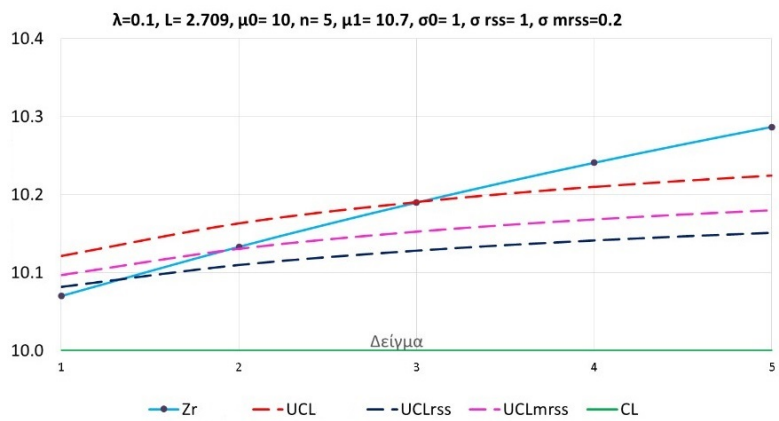
$$\sigma_{\bar{Y}_{DRSS}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:n)}^*{}^2} \quad (7.35)$$

όπου:

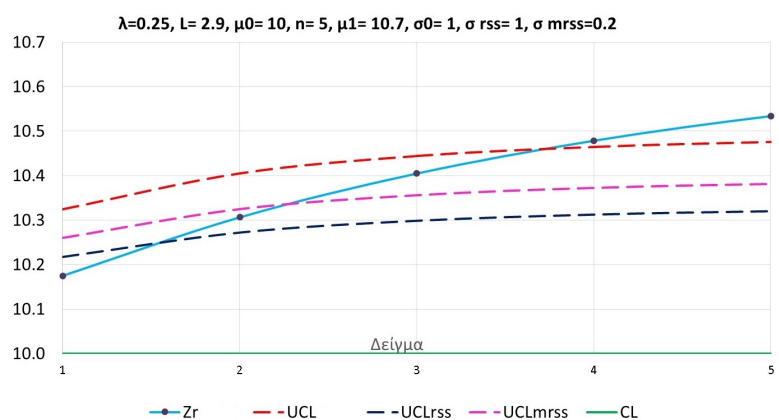
$$\sigma_{(i:n)}^*{}^2 = E [ (Y_{(i:n)} - E[Y_{(i:n)}])^2 ]$$

η διακύμανση του στατιστικού  $i$ -οστής τάξης της υπόψη μεθόδου (υπολογίζεται με αριθμητική ολοκλήρωση).

Καθώς συνήθως τα  $\mu$  και  $\sigma$  μιας διεργασίας είναι άγνωστα, μπορούμε να τα εκτιμήσουμε



Σχήμα 7.21: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.22: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

με τον εξής αμερόληπτο εκτιμητή για τον μέσο:

$$\bar{Y}_{DRSS} = \frac{1}{r} \sum_{j=1}^r \bar{Y}_{DRSS(j)} \quad (7.36)$$

και τον ακόλουθο εκτιμητή για την διασπορά:

$$\hat{\sigma}_{\bar{Y}_{DRSS}} = \sqrt{\frac{1}{n} \left( \hat{\sigma}_{DRSS}^2 - \frac{1}{n} \sum_{i=1}^n (\bar{Y}_{(i)} - \bar{Y}_{DRSS})^2 \right)} \quad (7.37)$$

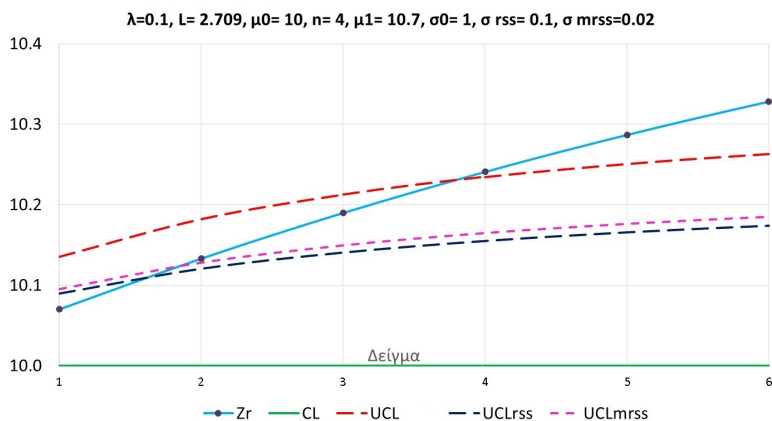
όπου:

$$\hat{\sigma}_{DRSS}^2 = \frac{1}{nr - 1} \sum_{i=1}^n \sum_{j=1}^r (Y_{(i:n)j} - \bar{Y}_{DRSS})^2$$

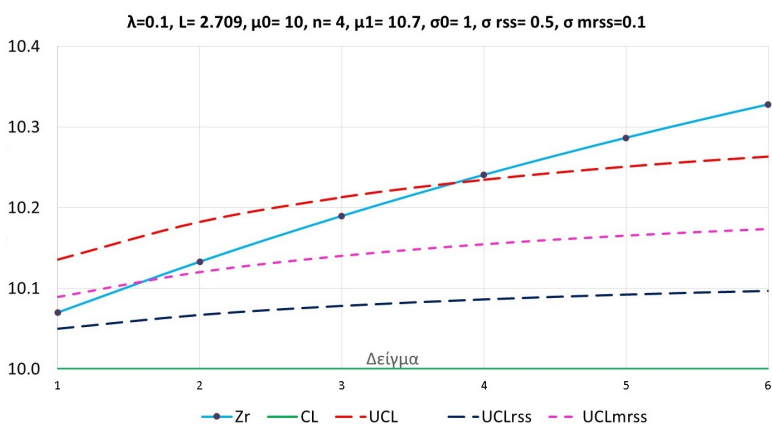
και

$$\bar{Y}_{(i)} = \frac{1}{r} \sum_{j=1}^r Y_{(i:n)j}$$

Με τη βοήθεια υπολογιστικής προσομοίωσης (50.000 επαναλήψεις, δεδομένα από κανονική κατανομή) βρέθηκε ότι η μεροληψία του εκτιμητή της διασποράς είναι αμελητέα για



Σχήμα 7.23: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA



Σχήμα 7.24: Σύγκριση διαγραμμάτων MRSS-EWMA, RSS-EWMA και EWMA

οποιοδήποτε μέγεθος δείγματος  $n$  και για μεγάλο αριθμό κύκλων  $r$ .

### 7.4.2 Τεχνική MDRSS (Median Double Ranked Set Sampling)

Η τεχνική δειγματοληψίας MDRSS μελετήθηκε από τους Abujiya και Muttlac (2004) [7], ως ένας πιο αποτελεσματικός τρόπος εκτίμησης του πληθυσμιακού μέσου συγκριτικά με την μέθοδο RSS. Μπορεί να περιγραφεί από τον ακόλουθο αλγόριθμο:

1ο βήμα: Τυχαία επιλογή  $n$  δειγμάτων (σύνολα παρατηρήσεων) από τον πληθυσμό, που το καθένα έχει μέγεθος  $n^2$  μονάδες.

2ο βήμα: Εφαρμογή της διαδικασίας RSS σε κάθε σύνολο, ώστε να προκύψουν  $n$  RSS σετ μεγέθους  $n$  μονάδων το καθένα, χωρίς κοστοβόρα μέτρηση.

3ο βήμα: Εφαρμογή της διαδικασίας MRSS στα παραπάνω σετ (RSS δεδομένα), ώστε να προκύψουν  $n$  μονάδες.

4ο βήμα: Μέτρηση των  $n$  επιλεγμένων μονάδων, που τις συμβολίζουμε  $Y_{(i:md)j}$  (με το

Πίνακας 7.15: Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για  $n=2$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	95.9	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.2	94.9	29.2	14.6	9.5	5.5	4.0	3.1	2.6	2.0
0.50	372.7	89.8	27.6	14.0	9.1	5.3	3.8	3.0	2.5	2.0
0.75	372.7	82.6	24.8	12.8	8.4	5.0	3.6	2.9	2.4	1.9
1.00	370.1	70.8	21.3	11.1	7.4	4.5	3.3	2.6	2.2	1.8

Πίνακας 7.16: Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για  $n=3$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	96.3	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.0	94.1	28.8	14.5	9.4	5.5	3.9	3.1	2.6	2.0
0.50	372.7	87.1	26.5	13.5	8.8	5.2	3.7	3.0	2.5	2.0
0.75	372.6	75.3	22.6	11.7	7.7	4.6	3.4	2.7	2.3	1.9
1.00	369.2	56.4	17.0	9.1	6.2	3.8	2.8	2.3	2.0	1.6

$j=1, 2, \dots, r$  να αντιστοιχεί στον κύκλο), και υπολογισμός του μέσου τους.

Η παραπάνω διαδικασία επαναλαμβάνεται  $r$  φορές (κύκλοι), δημιουργώντας δηλαδή ένα σύνολο από  $nr$  μονάδες MDRSS δεδομένων.

Ας υποθέσουμε ότι η διεργασία που θέλουμε να παρακολουθήσουμε ακολουθεί κανονική κατανομή, με γνωστά  $\mu$  και  $\sigma$ . Τότε, στον  $j$ -οστό κύκλο ο μέσος των  $n$  MDRSS δεδομένων θα δίνεται από τη σχέση:

$$\bar{Y}_{MDRSS(j)} = \frac{1}{n} \sum_{i=1}^n Y_{(i:md)j} \quad (7.38)$$

ενώ η διασπορά τους θα δίνεται από τη σχέση:

$$\sigma_{\bar{Y}_{MDRSS}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:md)}^2} \quad (7.39)$$

όπου:

$$\sigma_{(i:md)}^2 = E [ (Y_{(i:md)} - E[Y_{(i:md)}])^2 ]$$

η διακύμανση του στατιστικού  $i$ -οστής τάξης της υπόψη μεθόδου (υπολογίζεται με αριθμητική ολοκλήρωση).

Καθώς συνήθως τα  $\mu$  και  $\sigma$  μιας διεργασίας είναι άγνωστα, μπορούμε να τα εκτιμήσουμε



Πίνακας 7.17: Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για  $n=4$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	371.9	96.2	29.5	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	372.9	93.5	28.6	14.5	9.4	5.5	3.9	3.1	2.6	2.0
0.50	371.5	85.2	25.9	13.2	8.6	5.1	3.7	2.9	2.4	2.0
0.75	371.8	70.5	21.1	11.0	7.3	4.4	3.2	2.6	2.2	1.8
1.00	369.8	47.4	14.4	7.9	5.5	3.4	2.6	2.1	1.9	1.4

Πίνακας 7.18: Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για  $n=5$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	370.6	96.1	29.7	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	371.4	93.0	28.5	14.4	9.3	5.5	3.9	3.1	2.6	2.0
0.50	371.0	83.6	25.3	13.0	8.5	5.0	3.6	2.9	2.4	2.0
0.75	372.9	67.2	20.2	10.6	7.1	4.3	3.2	2.5	2.2	1.8
1.00	369.4	40.9	12.7	7.1	5.0	3.2	2.4	2.0	1.8	1.2

με τον εξής αμερόληπτο εκτιμητή για τον μέσο:

$$\bar{Y}_{MDRSS} = \frac{1}{r} \sum_{j=1}^r \bar{Y}_{MDRSS(j)} \quad (7.40)$$

και τον ακόλουθο εκτιμητή για την διασπορά:

$$\hat{\sigma}_{\bar{Y}_{MDRSS}} = \sqrt{\frac{1}{n(n-1)} \left( \sum_{j=1}^r \sum_{i=1}^n (Y_{(i:md)j} - \bar{Y}_{MDRSS})^2 \right)} \quad (7.41)$$

Με τη βοήθεια υπολογιστικής προσομοίωσης (50.000 επαναλήψεις, δεδομένα από κανονική κατανομή) βρέθηκε ότι η μεροληψία του εκτιμητή της διασποράς είναι αμελητέα για οποιοδήποτε μέγεθος δείγματος  $n$  και για μεγάλο αριθμό κύκλων  $r$ .

### 7.4.3 Τεχνική DMRSS

Παρόμοια τεχνική δειγματοληψίας αποτελεί και η DMRSS, η οποία πάλι περιγράφεται από τους Abujiya και Muttalac (2004) [7]. Μπορεί να περιγραφεί από τον ακόλουθο αλγόριθμο:

1ο βήμα: Τυχαία επιλογή  $n$  δειγμάτων (σύνολα παρατηρήσεων) από τον πληθυσμό, που το καθένα έχει μέγεθος  $n^2$  μονάδες.

Πίνακας 7.19: Ανιχνευτική Ικανότητα Διαγράμματος IMRSS-EWMA για  $n=6$ ,  $\lambda=0.12$  και  $L=2.75$

$\rho$	$\delta$									
	0.00	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
0.00	372.0	96.0	29.6	14.9	9.6	5.6	4.0	3.2	2.6	2.1
0.25	373.2	92.9	28.4	14.3	9.3	5.5	3.9	3.1	2.6	2.0
0.50	372.6	82.9	25.1	12.9	8.4	5.0	3.6	2.9	2.4	2.0
0.75	372.3	64.9	19.4	10.3	6.9	4.2	3.1	2.5	2.1	1.8
1.00	369.9	36.1	11.4	6.5	4.6	2.9	2.2	1.9	1.7	1.1

2ο βήμα: Εφαρμογή της διαδικασίας MRSS σε κάθε σύνολο, ώστε να προκύψουν  $n$  MRSS σετ μεγέθους  $n$  μονάδων το καθένα, χωρίς κοστοβόρα μέτρηση.

3ο βήμα: Εφαρμογή ξανά της διαδικασίας MRSS στα παραπάνω σετ (RSS δεδομένα), ώστε να προκύψουν  $n$  μονάδες.

4ο βήμα: Μέτρηση των  $n$  επιλεγμένων μονάδων, που τις συμβολίζουμε  $Y_{(i:dm)j}$  (με το  $j=1, 2, \dots, r$  να αντιστοιχεί στον κύκλο), και υπολογισμός του μέσου τους.

Η παραπάνω διαδικασία επαναλαμβάνεται  $r$  φορές (κύκλοι), δημιουργώντας δηλαδή ένα σύνολο από  $nr$  μονάδες DMRSS δεδομένων.

Ας υποθέσουμε πάλι ότι η διεργασία που θέλουμε να παρακολουθήσουμε ακολουθεί κανονική κατανομή, με γνωστά  $\mu$  και  $\sigma$ . Τότε, στον  $j$ -οστό κύκλο ο μέσος των  $n$  DMRSS δεδομένων θα δίνεται από τη σχέση:

$$\bar{Y}_{DMRSS(j)} = \frac{1}{n} \sum_{i=1}^n Y_{(i:dm)j} \quad (7.42)$$

ενώ η διασπορά τους θα δίνεται από τη σχέση:

$$\sigma_{\bar{Y}_{DMRSS}} = \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sigma_{(i:dm)}^2} \quad (7.43)$$

όπου:

$$\sigma_{(i:dm)}^2 = E [ (Y_{(i:dm)} - E[Y_{(i:dm)}])^2 ]$$

η διακύμανση του στατιστικού  $i$ -οστής τάξης της υπόψη μεθόδου (υπολογίζεται με αριθμητική ολοκλήρωση).

Καθώς συνήθως τα  $\mu$  και  $\sigma$  μιας διεργασίας είναι άγνωστα, μπορούμε να τα εκτιμήσουμε με τον εξής αμερόληπτο εκτιμητή για τον μέσο:

$$\bar{Y}_{DMRSS} = \frac{1}{r} \sum_{j=1}^r \bar{Y}_{DMRSS(j)} \quad (7.44)$$

και τον ακόλουθο εκτιμητή για την διασπορά:

$$\hat{\sigma}_{\bar{Y}_{DMRSS}} = \sqrt{\frac{1}{n(n-1)} \left( \sum_{j=1}^r \sum_{i=1}^n (Y_{(i:dm)j} - \bar{Y}_{DMRSS})^2 \right)} \quad (7.45)$$

Με τη βοήθεια υπολογιστικής προσομοίωσης (50.000 επαναλήψεις, δεδομένα από κανονική κατανομή) βρέθηκε ότι η μεροληψία του εκτιμητή της διασποράς είναι αμελητέα για οποιοδήποτε μέγεθος δείγματος  $n$  και για μεγάλο αριθμό κύκλων  $r$ .

#### 7.4.4 Διάγραμμα EWMA με τεχνικές MDRSS και DMRSS

Η χρήση δεδομένων από τροποποιημένες τεχνικές RSS στην κατασκευή διαγραμμάτων EWMA, και συγκεκριμένα αυτών που εμπλέκουν τη λιγότερο επιρρεπή σε σφάλμα μέτρηση διαμέσου (Median-Ranked δεδομένα), έγινε από τους Abujiya, Lee και Riaz (2013) [6].

Υποθέτουμε ότι ο μέσος της διεργασίας είναι γνωστός ( $\mu_0$ ) και ότι η διακύμανση είναι επίσης γνωστή και σταθερή. Φυσικά, είδαμε παραπάνω ότι μπορούμε έχουμε εκτιμήσεις από τις Σχέσεις (7.40) ( $\bar{Y}_{MDRSS}$ ) και (7.41) ( $\hat{\sigma}_{\bar{Y}_{MDRSS}}$ ) για την τεχνική MDRSS και από τις Σχέσεις (7.44) ( $\bar{Y}_{DMRSS}$ ) και (7.45) ( $\hat{\sigma}_{\bar{Y}_{DMRSS}}$ ) για την τεχνική DMRSS. Στο εξής, τους εκτιμητές  $\bar{Y}_{MDRSS(j)}$  και  $\bar{Y}_{DMRSS(j)}$  θα τους συμβολίζουμε ως  $\bar{X}_{median(j)}$ , ενώ τη διακύμανση του στατιστικού  $i$ -οστής τάξης της εκάστοτε μεθόδου θα τη συμβολίζουμε ως  $\sigma_{\bar{X}_{median}}$ .

Συνεπώς, το στατιστικό της μεθόδου EWMA με Median-Ranked δεδομένα θα δίνεται από τη σχέση:

$$Z_{(j:n)} = \lambda \bar{X}_{median(j)} + (1 - \lambda) Z_{(j-1:n)} \quad (7.46)$$

και η διασπορά του θα δίνεται από τη σχέση:

$$Var[Z_{(j:n)}] = \sigma_{\bar{X}_{median(j)}} \left( \frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2j}) \right) \quad (7.47)$$

Μπορούμε, οπότε, να κατασκευάσουμε το διάγραμμα ελέγχου EWMA με Median-Ranked δεδομένα. Συγκεκριμένα, η κεντρική γραμμή  $CL_{median}$ , όπως επίσης το ανώτερο  $UCL_{median}$  και κατώτερο όριο  $LCL_{median}$ , δίνονται από τις παρακάτω σχέσεις:

$$LCL_{median} = \mu_0 - L \sigma_{\bar{X}_{median}} \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2j})} \quad (7.48)$$

$$CL_{median} = \mu_0 \quad (7.49)$$

$$UCL_{median} = \mu_0 + L \sigma_{\bar{X}_{median}} \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2j})} \quad (7.50)$$

### 7.4.5 Ανιχνευτική ικανότητα διαγράμματος EWMA με τεχνικές MDRSS και DMRSS

Μελέτη της ανιχνευτικής ικανότητας των διαγραμμάτων έγινε από τους Abujiya, Lee και Riaz (2014) [6], όπου υπολογίστηκε η μέση τιμή του ARL για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 100.000 επαναλήψεις και μέγεθος δείγματος  $n=5$ . Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στους Πίνακες 7.20, 7.21 και 7.22 παρακάτω, όπου συγκρίνεται η ανιχνευτική ικανότητα των εν λόγω διαγραμμάτων για διάφορες τιμές των παραμέτρων  $\lambda$ ,  $L$  και της μετατόπισης  $\delta$  του μέσου. Στους πίνακες, πέρα από το ARL, δίνεται και ένα επιπλέον μέγεθος μέτρησης της απόδοσης των διαγραμμάτων, το SDRL, όπως ορίστηκε στο πρώτο κεφάλαιο με τη Σχέση (1.12).

Πίνακας 7.20: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Χρονικά μεταβαλλόμενα όρια,  $n=5$ ,  $ARL_0=370$

	Τεχνική	$\delta$							$AEQL$
	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	
		<i>MDRSS</i> $\lambda=0.05$		<i>L=2.509</i>					
<i>ARL</i>	370.71	13.87	4.86	2.82	2.01	1.57	1.28	1.10	2.287
<i>SDRL</i>	377.63	8.99	2.44	1.17	0.73	0.57	0.46	0.30	
		<i>DMRSS</i> $\lambda=0.05$		<i>L=2.510</i>					
<i>ARL</i>	370.23	9.31	3.42	2.08	1.51	1.18	1.04	1.00	1.869
<i>SDRL</i>	373.91	5.52	1.52	0.76	0.55	0.39	0.19	0.06	
		<i>MDRSS</i> $\lambda=0.10$		<i>L=2.708</i>					
<i>ARL</i>	370.00	15.74	5.23	3.01	2.13	1.66	1.34	1.14	2.401
<i>SDRL</i>	369.46	10.59	2.60	1.23	0.75	0.58	0.48	0.35	
		<i>DMRSS</i> $\lambda=0.1$		<i>L=2.708</i>					
<i>ARL</i>	370.29	10.26	3.65	2.20	1.60	1.23	1.05	1.01	1.922
<i>SDRL</i>	368.82	6.15	1.60	0.79	0.56	0.43	0.23	0.09	
		<i>MDRSS</i> $\lambda=0.25$		<i>L=2.900</i>					
<i>ARL</i>	370.43	21.83	5.80	3.12	2.15	1.66	1.34	1.13	2.455
<i>SDRL</i>	369.20	17.94	3.28	1.35	0.79	0.59	0.49	0.34	
		<i>DMRSS</i> $\lambda=0.25$		<i>L=2.900</i>					
<i>ARL</i>	369.77	12.93	3.84	2.22	1.60	1.23	1.05	1.01	1.940
<i>SDRL</i>	368.53	9.53	1.84	0.83	0.57	0.42	0.22	0.08	
		<i>MDRSS</i> $\lambda=0.50$		<i>L=2.982</i>					
<i>ARL</i>	371.07	38.60	7.48	3.31	2.07	1.55	1.24	1.08	2.511
<i>SDRL</i>	369.82	36.76	5.60	1.84	0.92	0.62	0.45	0.27	
		<i>DMRSS</i> $\lambda=0.50$		<i>L=2.982</i>					
<i>ARL</i>	369.53	20.99	4.35	2.18	1.48	1.15	1.03	1.00	1.948
<i>SDRL</i>	366.55	18.80	2.74	0.99	0.59	0.36	0.16	0.05	

Πίνακας 7.21: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Χρονικά μεταβαλλόμενα όρια,  $n=5$ ,  $ARL_0=500$

	Τεχνική	$\delta$							$AEQL$
	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	
		<i>MDRSS</i> $\lambda=0.05$		$L=2.630$					
<i>ARL</i>	500.74	15.21	5.19	2.99	2.12	1.65	1.34	1.13	2.386
<i>SDRL</i>	502.66	9.68	2.55	1.21	0.75	0.58	0.48	0.34	
		<i>DMRSS</i> $\lambda=0.05$		$L=2.628$					
<i>ARL</i>	500.48	10.05	3.63	2.19	1.59	1.23	1.05	1.01	1.916
<i>SDRL</i>	503.86	5.91	1.58	0.79	0.57	0.42	0.22	0.08	
		<i>MDRSS</i> $\lambda=0.10$		$L=2.819$					
<i>ARL</i>	500.02	17.27	5.58	3.17	2.22	1.73	1.40	1.17	2.498
<i>SDRL</i>	498.01	11.63	2.73	1.28	0.77	0.59	0.51	0.38	
		<i>DMRSS</i> $\lambda=0.10$		$L=2.819$					
<i>ARL</i>	501.07	11.09	3.86	2.30	1.67	1.28	1.07	1.01	1.969
<i>SDRL</i>	497.48	6.68	1.68	0.81	0.57	0.45	0.26	0.10	
		<i>MDRSS</i> $\lambda=0.25$		$L=3.000$					
<i>ARL</i>	499.45	24.97	6.15	3.25	2.23	1.72	1.39	1.16	2.551
<i>SDRL</i>	494.00	20.87	3.50	1.41	0.81	0.60	0.50	0.37	
		<i>DMRSS</i> $\lambda=0.25$		$L=3.000$					
<i>ARL</i>	499.97	14.33	4.05	2.32	1.65	1.27	1.07	1.01	1.985
<i>SDRL</i>	499.62	10.73	1.94	0.85	0.58	0.45	0.25	0.10	
		<i>MDRSS</i> $\lambda=0.50$		$L=3.075$					
<i>ARL</i>	500.14	46.24	8.21	3.48	2.17	1.60	1.28	1.10	2.630
<i>SDRL</i>	495.97	43.79	6.22	1.94	0.96	0.64	0.47	0.30	
		<i>DMRSS</i> $\lambda=0.50$		$L=3.075$					
<i>ARL</i>	499.45	24.75	4.67	2.28	1.53	1.18	1.04	1.00	2.000
<i>SDRL</i>	497.49	22.50	2.96	1.04	0.61	0.39	0.19	0.06	

### 7.4.6 Εφαρμογή διαγράμματος EWMA με τεχνικές MDRSS και DMRSS

Στη συνέχεια θα παρατεθούν τα αποτελέσματα από την εφαρμογή πραγματικών δεδομένων (συγκεκριμένα, πρόκειται για μετρήσεις γεμίσματος μπουκαλιών σε βιομηχανία αναψυκτικών) στα υπόψη διαγράμματα, όπως αυτά δημοσιεύτηκαν από τους Abujiya, Lee και Riaz (2013) [6]. Τα εν λόγω δεδομένα αποτελούνται από 18 τυχαία δείγματα, το καθένα αποτελούμενο από 3 παρατηρήσεις ( $n=3$ ). Οι μέσοι των 18 αυτών τυχαίων δειγμάτων, για διαφορετικές τεχνικές δειγματοληψίας, απεικονίζονται στο Σχήμα 7.25.

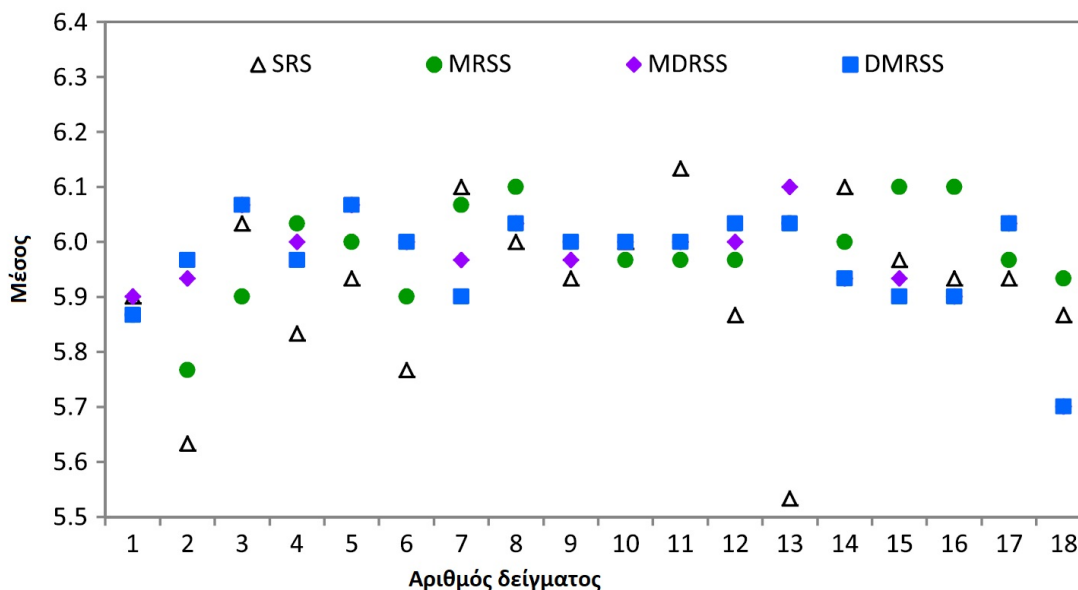
Τα παραπάνω δεδομένα τροποποιήθηκαν, με την πρόσθεση 0.25 μονάδων στους 5 τελευταίους μέσους, έτσι ώστε τα 13 πρώτα δείγματα να είναι εντός στατιστικού ελέγχου ενώ τα

Πίνακας 7.22: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA για δειγματοληψίες MDRSS και DMRSS - Σταθερά όρια,  $n=5$ ,  $ARL_0=500$

	Τεχνική	$\delta$							$AEQL$
	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	
		<i>MDRSS</i>		$\lambda=0.05$	$L=2.614$				
<i>ARL</i>	500.23	18.85	7.98	5.14	3.85	3.14	2.66	2.27	4.465
<i>SDRL</i>	483.48	8.97	2.46	1.24	0.81	0.58	0.53	0.44	
		<i>DMRSS</i>		$\lambda=0.05$	$L=2.614$				
<i>ARL</i>	500.82	13.43	6.02	3.98	3.04	2.46	2.09	1.99	3.669
<i>SDRL</i>	485.32	5.45	1.59	0.84	0.56	0.51	0.29	0.13	
		<i>MDRSS</i>		$\lambda=0.10$	$L=2.814$				
<i>ARL</i>	499.96	18.77	6.92	4.28	3.17	2.55	2.17	1.99	3.793
<i>SDRL</i>	492.16	11.28	2.58	1.22	0.77	0.58	0.40	0.27	
		<i>DMRSS</i>		$\lambda=0.10$	$L=2.814$				
<i>ARL</i>	500.82	12.56	5.08	3.28	2.47	2.07	1.89	1.63	3.007
<i>SDRL</i>	490.95	6.37	1.60	0.80	0.56	0.32	0.32	0.48	
		<i>MDRSS</i>		$\lambda=0.25$	$L=3.000$				
<i>ARL</i>	499.15	25.28	6.45	3.55	2.51	2.01	1.70	1.42	2.945
<i>SDRL</i>	492.51	20.82	3.43	1.36	0.77	0.53	0.50	0.50	
		<i>DMRSS</i>		$\lambda=0.25$	$L=2.999$				
<i>ARL</i>	499.97	14.66	4.36	2.60	1.96	1.57	1.24	1.06	2.209
<i>SDRL</i>	498.34	10.69	1.88	0.81	0.51	0.51	0.43	0.23	
		<i>MDRSS</i>		$\lambda=0.50$	$L=3.075$				
<i>ARL</i>	500.34	46.32	8.27	3.53	2.22	1.65	1.31	1.12	2.680
<i>SDRL</i>	495.94	43.77	6.23	1.93	0.96	0.64	0.48	0.33	
		<i>DMRSS</i>		$\lambda=0.50$	$L=3.076$				
<i>ARL</i>	499.61	24.81	4.71	2.33	1.58	1.21	1.04	1.01	2.024
<i>SDRL</i>	497.53	22.49	2.95	1.03	0.61	0.41	0.21	0.07	

τελευταία 5 να είναι εκτός. Τα στατιστικά EWMA υπολογίστηκαν τόσο για το απλό διάγραμμα όσο και για τα MRSS, MDRSS και DMRSS για  $ARL_0=500$ . Στα Σχήματα 7.26, 7.27 και 7.28 γίνεται σύγκριση κάθε ενός από τα 3 είδη διαγραμμάτων RSS δειγματοληψίας με το κλασσικό.

Θα παρατηρήσουμε, πάντως, ότι το κλασσικό διάγραμμα EWMA αντιλαμβάνεται μόνο μια τάση, χωρίς να δίνει εκτός ελέγχου σήμα, ενώ και τα τρία είδη διαγραμμάτων RSS δειγματοληψίας δίνουν εκτός ελέγχου σήμα αρκετά νωρίς: το MRSS-EWMA από το 16ο δείγμα και τα MDRSS-EWMA και DMRSS-EWMA, δείχνοντας την καλύτερη απόδοση, από το 15ο δείγμα.



Σχήμα 7.25: Δεδομένα εφαρμογής διαγράμματος EWMA με τεχνικές SRS MRSS, MDRSS και DMRSS

## 7.5 Τεχνική δειγματοληψίας ODRSS

Η τεχνική δειγματοληψίας Ordered Double Ranked Set Sampling (ODRSS) εισήχθη ως επέκταση της τεχνικής Ordered Ranked Set Sampling (ORSS). Συγκεκριμένα, όσον αφορά την εύρεση των βέλτιστων γραμμικών αμερόληπτων εκτιμητών - Best Linear Unbiased Estimators (BLUEs) - για την άγνωστη παράμετρο μιας κατανομής (αναφερόμενης στη βιβλιογραφία ως Γενικευμένης Γεωμετρικής Κατανομής ή Generalized Geometric Distribution (GGD)), οι Balakrishnan και Li (2005) [9] είχαν ήδη δείξει ότι βέλτιστοι γραμμικοί αμερόληπτοι εκτιμητές βασισμένοι σε ORSS (BLUEs-ORSS) είναι ομοιόμορφα καλύτεροι από τους αντίστοιχους βασισμένους σε RSS (BLUEs-RSS) και SRS (BLUEs-SRS). Μετέπειτα, οι Haq et al (2014) [22] έδειξαν ότι οι BLUEs-ODRSS είναι ομοιόμορφα καλύτεροι από τους BLUEs-ORSS.

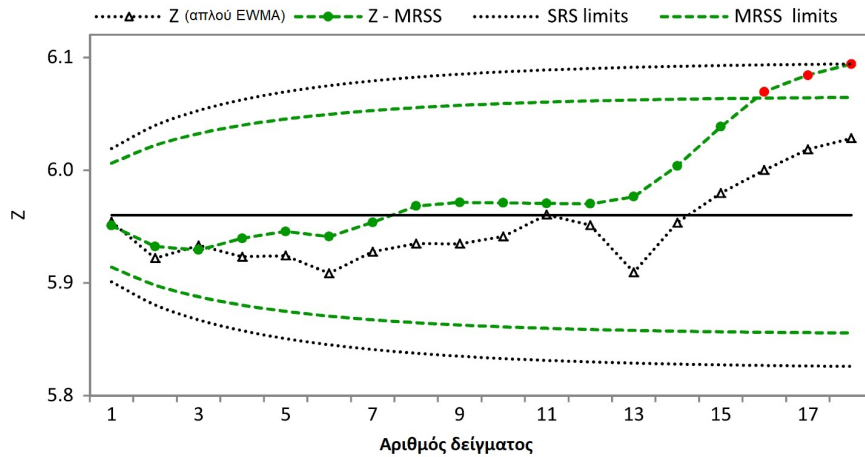
Η τεχνική δειγματοληψίας ODRSS περιγράφεται ως εξής:

- Έστω ότι τα  $S_1, S_2, \dots, S_m$  είναι  $m$  σύνολα (σέτ), μεγέθους  $m^2$  μονάδες το κάθε ένα.
- Χωρίζουμε τις  $m^2$  μονάδες στο  $i$ -οστό σέτ  $S_i$  σε  $m$  υποσύνολα  $s_{ij}$ , μεγέθους  $m$  μονάδων το καθένα, δηλαδή:

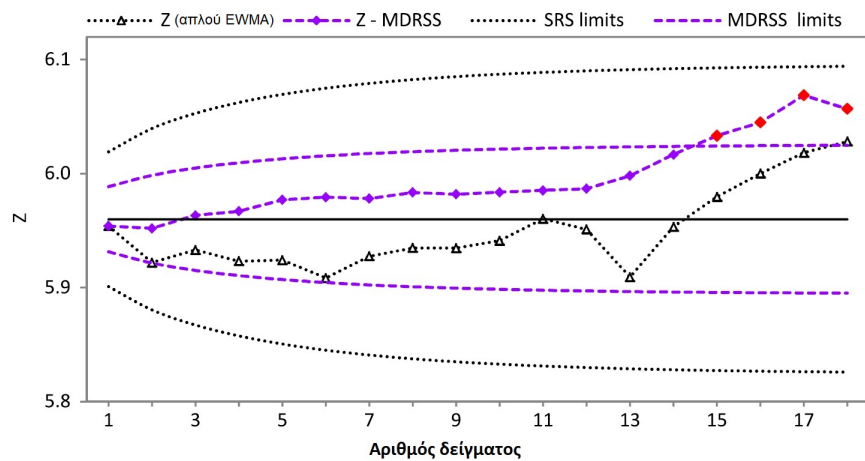
$$S_i = \{s_{ij}\} = \{s_{i1}, s_{i2}, \dots, s_{im}\} \text{ για } j = 1, 2, \dots, m$$

και τις μονάδες του  $j$ -οστού υποσυνόλου  $s_{ij}$  του  $i$ -οστού σέτ  $S_i$  τις συμβολίζουμε ως εξής:

$$s_{ij} = \{Y_{(1)}^{(i)}, Y_{(2)}^{(i)}, \dots, Y_{(m)}^{(i)}\}$$



Σχήμα 7.26: Σύγκριση Διαγράμματος MRSS-EWMA ( $\lambda=0.1, L=2.820$ ) με Διάγραμμα SRS-EWMA ( $\lambda=0.1, L=2.820$ )



Σχήμα 7.27: Σύγκριση Διαγράμματος MDRSS-EWMA ( $\lambda=0.1, L=2.819$ ) με Διάγραμμα SRS-EWMA ( $\lambda=0.1, L=2.820$ )

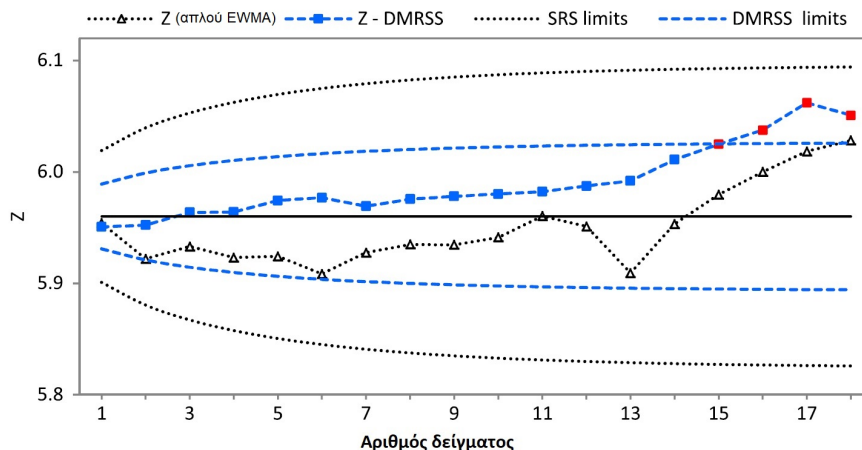
- Εφαρμόζουμε την διαδικασία RSS στα  $m$  παραπάνω υποσύνολα  $s_{ij}$ , λαμβάνοντας έτσι  $m$  ταξινομημένα (ranked) υποσύνολα, μεγέθους  $m$  μονάδων το καθένα.
- Υποθέτουμε ότι το  $i$ -οστό σετ  $S_i^*$  περιέχει το  $i$ -οστό ταξινομημένο υποσύνολο, δηλαδή:

$$S_i^* = \{Y_{(1)}^{(i)}, Y_{(2)}^{(i)}, \dots, Y_{(m)}^{(i)}\} \text{ για } i = 1, 2, \dots, m$$

και εφαρμόζουμε ξανά την διαδικασία RSS στα παραπάνω  $m$  ταξινομημένα σύνολα  $S_i^*$ , λαμβάνοντας με τον τρόπο αυτό ένα διπλά ταξινομημένο δείγμα (DRSS) μεγέθους  $m$ .

Έστω, λοιπόν, ότι  $Z_i^{DRSS}$  είναι το  $i$ -οστό μικρότερο στοιχείο του συνόλου  $S_i^*$ . Θεωρούμε το σύνολο  $(Z_1^{DRSS}, Z_2^{DRSS}, \dots, Z_m^{DRSS})$  μεγέθους  $m$  και κατατάσσοντας τα στοιχεία του συνόλου αυτού από το μικρότερο προς το μεγαλύτερο, παίρνουμε το διατεταγμένο διπλά





Σχήμα 7.28: Σύγκριση Διαγράμματος DMRSS-EWMA ( $\lambda=0.1, L=2.819$ ) με Δι-άγραμμα SRS-EWMA ( $\lambda=0.1, L=2.820$ )

ταξινομημένο σύνολο (ODRSS) μεγέθους  $m$ , δηλαδή το εξής:

$$Z_{(1)}^{ODRSS} \leq Z_{(2)}^{ODRSS}, \leq \dots \leq Z_{(m)}^{ODRSS}$$

### 7.5.1 Εκτιμητές Τεχνικής ODRSS

Έστω το διατεταγμένο διπλά ταξινομημένο σύνολο (ODRSS) μεγέθους  $m$ :

$$Y_{ODRSS} = \left( Y_{(1:m)}^{ODRSS}, Y_{(2:m)}^{ODRSS}, \dots, Y_{(m:m)}^{ODRSS} \right)_{1 \times m}$$

που αποκτήθηκε από μία γενικευμένη οικογένεια κατανομών κλίμακας-θέσης, με  $\sigma$  και  $\mu$  (θετικά) να είναι οι παράμετροι κλίμακας και θέσης αντίστοιχα. Επίσης, έστω ότι η τυποποιημένη διασπορά στο σχέδιο ODRSS είναι:

$$Q_{(r:m)}^{ODRSS} = \frac{Y_{(r:m)}^{ODRSS} - \mu}{\sigma} \text{ για } r = 1, 2, \dots, m \quad (7.51)$$

Η συνάρτηση πυκνότητας πιθανότητας, συνεπώς, του  $Q_{(r:m)}^{ODRSS}$  είναι ανεξάρτητη των  $\mu$  και  $\sigma$ .

Ορίζουμε επίσης:

$$\nu_{(r:m)}^{ODRSS} = E \left[ Q_{(r:m)}^{ODRSS} \right] \text{ για } 1 \leq r \leq m$$

και

$$\theta_{(r,s:m)}^{ODRSS} = Cov \left[ Q_{(r:m)}^{ODRSS}, Q_{(s:m)}^{ODRSS} \right] \text{ για } 1 \leq r < s \leq m$$

Οπότε, μπορούμε να γράψουμε:

$$E [Y_{(r:m)}^{ODRSS}] = \mu + \sigma \nu_{(r:m)}^{ODRSS}$$

και

$$Cov [Y_{(r:m)}^{ODRSS}, Y_{(s:m)}^{ODRSS}] = \sigma^2 \theta_{(r,s:m)}^{ODRSS}$$

Τέλος, ορίζουμε τον βέλτιστο γραμμικό αμερόληπτο εκτιμητή του σχεδίου ODRSS (BLUEs-ODRSS) για το  $\theta = (\mu, \sigma)_{1 \times 2}$ , δηλαδή το:

$$\hat{\theta}_{ODRSS}^{BLUE} = (\hat{\mu}_{ODRSS}^{BLUE}, \hat{\sigma}_{ODRSS}^{BLUE})_{1 \times 2}$$

το οποίο δίνεται από τη σχέση:

$$\hat{\theta}_{ODRSS}^{BLUE} = (B' \Sigma^{-1} B)^{-1} B' \Sigma^{-1} Y_{ODRSS} \quad (7.52)$$

όπου:

$$B = (\mathbf{1}, \nu)_{m \times 2} \text{ και } \Sigma = \begin{pmatrix} \theta_{(r,s:m)}^{ODRSS} \end{pmatrix}$$

$$\mathbf{1} = (1, 1, \dots, 1)_{1 \times m} \text{ και } \nu = (\nu_{(1:m)}^{ODRSS}, \nu_{(2:m)}^{ODRSS}, \dots, \nu_{(m:m)}^{ODRSS})_{1 \times m}$$

Ο πίνακας διακύμανσης-συνδιακύμανσης του  $\hat{\theta}_{ODRSS}^{BLUE}$  δίνεται από την σχέση:

$$Cov [\hat{\theta}_{ODRSS}^{BLUE}] = \sigma^2 (B' \Sigma^{-1} B)^{-1} \quad (7.53)$$

Οι βέλτιστοι γραμμικοί αμερόληπτοι εκτιμητές της τεχνικής ODRSS για τα  $\mu$  και  $\sigma$  (BLUEs-ODRSS) μπορούν να γραφούν ως γραμμικός συνδυασμός του  $Y_{ODRSS}$ , δηλαδή ως εξής:

$$\hat{\mu}_{ODRSS}^{BLUE} = -\nu' \Gamma Y_{ODRSS} \quad (7.54)$$

$$\hat{\sigma}_{ODRSS}^{BLUE} = -\mathbf{1}' \Gamma Y_{ODRSS} \quad (7.55)$$

όπου:

$$\Gamma = \frac{B^{-1}(\mathbf{1}\nu' - \nu\mathbf{1}')}{\Delta} \text{ ο πίνακας συμμετρικότητας-λοξότητας και } \Delta = |B' \Sigma^{-1} B|$$

Επίσης, για τη διασπορά και τη συνδιασπορά κάθε μιας εκ των παραμέτρων (επισημαίνοντας και την περίπτωση της συμμετρικής κατανομής, όπου ισχύει  $\mathbf{1}'\Sigma^{-1}\nu = -\mathbf{1}'\Sigma^{-1}\nu = 0$ , θα έχουμε:

$$Var [\hat{\mu}_{ODRSS}^{BLUE}] = \frac{\sigma^2 (\nu' \Sigma^{-1} \nu)}{\Delta} \quad (7.56)$$

$$Var [\hat{\sigma}_{ODRSS}^{BLUE}] = \sigma^2 (\mathbf{1}' \Sigma^{-1} \mathbf{1})^{-1} \text{ για συμμετρική κατανομή}$$

$$\begin{aligned} Var [\hat{\sigma}_{ODRSS}^{BLUE}] &= \frac{\sigma^2 (\mathbf{1}' \Sigma^{-1} \mathbf{1})}{\Delta} \\ Var [\hat{\sigma}_{ODRSS}^{BLUE}] &= \sigma^2 (\nu' \Sigma^{-1} \nu)^{-1} \text{ για συμμετρική κατανομή} \end{aligned} \quad (7.57)$$

$$\begin{aligned} Cov [\hat{\mu}_{ODRSS}^{BLUE}, \hat{\sigma}_{ODRSS}^{BLUE}] &= -\frac{\sigma^2 (\mathbf{1}' \Sigma^{-1} \nu)}{\Delta} \\ Cov [\hat{\mu}_{ODRSS}^{BLUE}, \hat{\sigma}_{ODRSS}^{BLUE}] &= 0 \text{ για συμμετρική κατανομή} \end{aligned} \quad (7.58)$$

Τέλος, τα BLUEs-ODRSS της Σχέσης (7.54) και Σχέσης (7.55) απλοποιούνται στην εξής μορφή:

$$\hat{\mu}_{ODRSS}^{BLUE} = (\mathbf{1}' \Sigma^{-1} \mathbf{1})^{-1} \mathbf{1}' \Sigma^{-1} Y_{ODRSS} \quad (7.59)$$

$$\hat{\sigma}_{ODRSS}^{BLUE} = (\nu' \Sigma^{-1} \nu)^{-1} \nu' \Sigma^{-1} Y_{ODRSS} \quad (7.60)$$

## 7.5.2 Διάγραμμα Ελέγχου EWMA με τεχνική ODRSS

Έστω  $Y_j$  το ποιοτικό χαρακτηριστικό ενδιαφέροντος, το οποίο υποθέτουμε ότι είναι κανονικά κατανομημένο με μέσο  $\mu_0$  και διασπορά  $\sigma^2$  τη χρονική στιγμή  $j$  ( $Y_j \sim N(\mu_0, \sigma^2)$ ). Χωρίς περιορισμό της γενικότητας, θέτουμε  $\mu_0=0$  και  $\sigma^2=1$ .

Έστω  $\hat{\mu}_{ODRSS}^{BLUE}$  ο βέλτιστος γραμμικός αμερόληπτος εκτιμητής της τεχνικής ODRSS, από το δείγμα μεγέθους  $m$  τη χρονική στιγμή  $j$  ( $j=1, 2, \dots$ ) και έστω  $\hat{\mu}_{ODRSS}^{BLUE}$  μια ακολουθία ανεξάρτητων και όμοια κατανομημένων τυχαίων μεταβλητών.

Τότε, το στατιστικό EWMA με τεχνική δειγματοληψίας ODRSS θα δίνεται από την παρακάτω σχέση:

$$Z_j = \lambda \hat{\mu}_{ODRSS,j}^{BLUE} + (1 - \lambda) Z_{j-1}, \quad Z_0 = \mu_0, \quad 0 < \lambda \leq 1 \quad (7.61)$$

Αποδεικνύεται ότι, για θετικό και ακέραιο  $t$ , η μέση τιμή και η διακύμανση του στατιστικού δίνονται από τις σχέσεις:

$$E [Z_t] = \mu_0 \quad (7.62)$$

$$Var [Z_j] = \frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2j}) \sigma_0^2 (\mathbf{1}' \Sigma^{-1} \mathbf{1})^{-1} \quad (7.63)$$

Συνεπώς, μπορούμε να κατασκευάσουμε το διάγραμμα ελέγχου της μεθόδου, ορίζοντας την κεντρική γραμμή και τα άνω και κάτω όρια ως εξής:

$$LCL = \mu_0 - L\sigma_0 \sqrt{\frac{\lambda}{2 - \lambda} (1 - (1 - \lambda)^{2j}) \sigma_0^2 (\mathbf{1}' \Sigma^{-1} \mathbf{1})^{-1}} \quad (7.64)$$

$$CL = \mu_0 \quad (7.65)$$

$$UCL = \mu_0 + L\sigma_0 \sqrt{\frac{\lambda}{2-\lambda}(1 - (1-\lambda)^{2j})\sigma_0^2(\mathbf{1}'\Sigma^{-1}\mathbf{1})^{-1}} \quad (7.66)$$

### 7.5.3 Ανιχνευτική ικανότητα Διαγραμμάτων EWMA-ODRSS

Μελέτη της ανιχνευτικής ικανότητας των υπόψη διαγραμμάτων έγινε από τους Haq et al (2014) [22], όπου υπολογίστηκε το ARL για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 100.000 επαναλήψεις και μέγεθος δείγματος  $n=5$  (οι συγκεκριμένοι υπολογισμοί έγιναν για  $ARL_0=500$ ). Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στον Πίνακα 7.23, όπου αποτυπώνεται η ανιχνευτική ικανότητα των εν λόγω διαγραμμάτων για διάφορες τιμές των παραμέτρων  $\lambda$ ,  $L$  και της μετατόπισης  $\delta$  του μέσου. Για λόγους σύγκρισης, υπολογίστηκαν και παρατίθενται και οι αντίστοιχες τιμές του ARL για το κλασικό διάγραμμα EWMA, οι υπολογισμοί για τις οποίες έγιναν με χρήση λογισμικού R, θεωρώντας ότι η μετατόπιση γίνεται στην αρχή των μετρήσεων (εντολή *xewma.arl* με  $q=1$ ).

### 7.5.4 Διάγραμμα EWMA με Ατελή Ταξινόμηση Δειγματοληψίας ODRSS (OIDRSS)

Στην παραπάνω δημοσίευση (Haq et al, 2014 [22]) έγινε επίσης μελέτη της επίδρασης της ατελούς ταξινόμησης στην απόδοση των διαγραμμάτων. Στην περίπτωση αυτή, το στατιστικό EWMA με τεχνική δειγματοληψίας ODRSS θα δίνεται από την παρακάτω σχέση:

$$Z_j^* = \lambda \hat{\mu}_{OIDRSS,j}^{BLUE} + (1-\lambda) Z_{j-1}^*, \quad Z_0^* = \bar{\mu}_{OIDRSS}^{BLUE}, \quad 0 < \lambda \leq 1 \quad (7.67)$$

όπου:

$$\bar{\mu}_{OIDRSS}^{BLUE} = \frac{1}{m} \sum_{i=1}^m \hat{\mu}_{OIDRSS,i}^{BLUE}$$

Καθώς είναι ιδιαίτερα δύσκολο να καταλήξουμε σε μαθηματικό τύπο για τους εκτιμητές του μέσου και της διασποράς, η εκτίμησή τους γίνεται μέσω ιστορικών δεδομένων που πάρθηκαν κατά μια περίοδο εντός ελέγχου λειτουργίας της διεργασίας. Οπότε, θεωρώντας τα  $\hat{\mu}_{OIDRSS,i}^{BLUE}$ ,  $i=1, 2, \dots, m$ , ως τις βέλτιστες γραμμικές αμερόληπτες εκτιμήσεις της τεχνικής OIDRSS που πάρθηκαν από  $m$  υποσύνολα μεγέθους  $m$ , όπου:

$$\hat{\mu}_{OIDRSS}^{BLUE} = (\mathbf{1}'\Sigma^{-1}\mathbf{1})^{-1}\mathbf{1}'\Sigma^{-1}Y_{OIDRSS,i}$$

τότε μπορούμε να κατασκευάσουμε το Διάγραμμα Ελέγχου της μεθόδου, ορίζοντας την (εκτιμημένη) κεντρική γραμμή και τα (εκτιμημένα) άνω και κάτω όρια ως εξής:

$$LCL = \bar{\mu}_{OIDRSS}^{BLUE} - L\sigma_{\hat{\mu}_{OIDRSS}^{BLUE}} \sqrt{\frac{\lambda}{2-\lambda}(1 - (1-\lambda)^{2j})} \quad (7.68)$$

Πίνακας 7.23: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA-ODRSS με κλασικά EWMA - Χρονικά μεταβαλλόμενα όρια,  $n=5$ ,  $ARL_0=500$

Τεχνική		$\delta$						
		0.00	0.25	0.50	0.75	1.00	1.50	2.00
<i>ARL</i>	<i>ODRSS</i>		$\lambda=0.05$		$L=2.6402$			
		499.64	18.78	5.84	3.04	1.99	1.22	1.03
<i>ARL</i>	<i>SRS</i>		$\lambda=0.05$		$L=2.6392$			
		500.09	77.77	23.71	11.87	7.31	3.77	2.43
<i>ARL</i>	<i>ODRSS</i>		$\lambda=0.10$		$L=2.8274$			
		500.50	22.44	6.54	3.37	2.18	1.29	1.04
<i>ARL</i>	<i>SRS</i>		$\lambda=0.05$		$L=2.8239$			
		500.04	103.32	28.81	13.61	8.21	4.17	2.66
<i>ARL</i>	<i>ODRSS</i>		$\lambda=0.25$		$L=3.0085$			
		500.36	35.61	7.94	3.79	2.38	1.37	1.06
<i>ARL</i>	<i>SRS</i>		$\lambda=0.05$		$L=3.0071$			
		500.05	169.36	47.39	19.31	10.41	4.77	2.94
<i>ARL</i>	<i>ODRSS</i>		$\lambda=0.50$		$L=3.0910$			
		499.63	69.21	12.53	4.72	2.65	1.41	1.07
<i>ARL</i>	<i>SRS</i>		$\lambda=0.05$		$L=2.6392$			
		500.15	254.59	88.45	35.58	17.18	6.27	3.39

$$CL = \bar{\mu}_{OIDRSS}^{BLUE} \quad (7.69)$$

$$UCL = \bar{\mu}_{OIDRSS}^{BLUE} + L\sigma_{\bar{\mu}_{OIDRSS}^{BLUE}} \sqrt{\frac{\lambda}{2-\lambda}(1-(1-\lambda)^{2j})} \quad (7.70)$$

### 7.5.5 Ανιχνευτική ικανότητα Διαγραμμάτων EWMA-OIDRSS

Θα χρησιμοποιηθούν πάλι τα αποτελέσματα υπολογισμών του  $ARL$  για διάφορες τιμές μετατόπισης με χρήση τεχνικής Monte Carlo για 100.000 επαναλήψεις και μέγεθος δείγματος  $n=5$  (οι συγκεκριμένοι υπολογισμοί έγιναν για  $ARL_0=500$ ), αλλά και για διάφορες τιμές της παραμέτρου συσχέτισης  $\rho$ . Συνοπτικά αποτελέσματα της μελέτης αυτής δίνονται στον Πίνακα 7.24, όπου αποτυπώνεται η ανιχνευτική ικανότητα των διαγραμμάτων ODRSS-EWMA για διάφορες τιμές των παραμέτρων  $\lambda$ ,  $L$  και της μετατόπισης  $\delta$  του μέσου, σε σύγκριση με τις αντίστοιχες που υπολογίστηκαν για διαγράμματα ODRSS-EWMA.

Πίνακας 7.24: Σύγκριση Ανιχνευτικής Ικανότητας Διαγραμμάτων EWMA-OIDRSS με EWMA-ODRSS - Χρονικά μεταβαλλόμενα όρια,  $n=5$ ,  $ARL_0=500$

Τεχνική		$\delta$						
		0.00	0.25	0.50	0.75	1.00	1.50	2.00
<i>OIDRSS</i>		$\lambda=0.05$			$L=2.6402$			
<i>ARL</i>	$\rho=0.95$	499.28	22.59	6.96	3.59	2.33	1.37	1.07
<i>ARL</i>	$\rho=0.85$	502.09	28.78	8.80	4.50	2.86	1.62	1.18
<i>ARL</i>	$\rho=0.70$	497.62	35.84	10.90	5.55	3.49	1.91	1.34
<i>ARL</i>	$\rho=0.50$	502.91	42.55	12.97	6.55	4.09	2.20	1.50
<i>ODRSS</i>		$\lambda=0.05$			$L=2.6402$			
<i>ARL</i>	<i>ODRSS</i>	499.64	18.78	5.84	3.04	1.99	1.22	1.03
<i>OIDRSS</i>		$\lambda=0.10$			$L=2.8274$			
<i>ARL</i>	$\rho=0.95$	499.91	27.42	7.8	3.97	2.54	1.45	1.1
<i>ARL</i>	$\rho=0.85$	503.07	35.56	9.94	5.01	3.16	1.74	1.23
<i>ARL</i>	$\rho=0.70$	504.47	45.56	12.55	6.21	3.87	2.08	1.43
<i>ARL</i>	$\rho=0.50$	503.32	54.82	15.02	7.38	4.55	2.41	1.61
<i>ODRSS</i>		$\lambda=0.10$			$L=2.8274$			
<i>ARL</i>	<i>ODRSS</i>	500.50	22.44	6.54	3.37	2.18	1.29	1.04
<i>OIDRSS</i>		$\lambda=0.25$			$L=3.0085$			
<i>ARL</i>	$\rho=0.95$	506.27	45.13	9.84	4.54	2.82	1.56	1.14
<i>ARL</i>	$\rho=0.85$	510.02	60.71	13.27	5.89	3.53	1.89	1.30
<i>ARL</i>	$\rho=0.70$	511.40	78.73	17.47	7.51	4.41	2.27	1.52
<i>ARL</i>	$\rho=0.50$	507.73	95.15	21.78	9.11	5.26	2.66	1.73
<i>ODRSS</i>		$\lambda=0.25$			$L=3.0085$			
<i>ARL</i>	<i>ODRSS</i>	500.36	35.61	7.94	3.79	2.38	1.37	1.06
<i>OIDRSS</i>		$\lambda=0.50$			$L=3.0910$			
<i>ARL</i>	$\rho=0.95$	509.81	86.35	16.37	5.97	3.23	1.63	1.16
<i>ARL</i>	$\rho=0.85$	522.06	114.37	23.6	8.36	4.32	2.02	1.34
<i>ARL</i>	$\rho=0.70$	526.52	143.52	32.65	11.51	5.72	2.52	1.59
<i>ARL</i>	$\rho=0.50$	525.08	168.52	42	14.94	7.22	3.02	1.86
<i>ODRSS</i>		$\lambda=0.50$			$L=3.0910$			
<i>ARL</i>	<i>ODRSS</i>	499.63	69.21	12.53	4.72	2.65	1.41	1.07

# Κεφάλαιο 8

## Σύνοψη - Πρόταση

### 8.1 Το πρόβλημα της αυτοσυσχέτισης

Είδαμε στο προηγούμενο κεφάλαιο ότι ένας τρόπος βελτίωσης της ανιχνευτικής ικανότητας των διαγραμμάτων ελέγχου, και συγκεκριμένα των EWMA, αποτελεί η χρήση διάφορων τεχνικών δειγματοληψίας που βασίζονται σε κάποιες μορφής ταξινόμηση των παρατηρήσεων των δειγμάτων που λαμβάνουμε.

Θα πρέπει να επισημανθεί, ωστόσο, ότι τα διαγράμματα EWMA που αναφέρθηκαν, όπως και όλα τα διαγράμματα ελέγχου, βασίζονται σε δύο υποθέσεις. Η μία είναι η υπόθεση της κανονικότητας της διεργασίας, που σημαίνει ότι τα δεδομένα που συλλέγονται αποτελούν τυχαίες μεταβλητές της Κανονικής Κατανομής. Η άλλη υπόθεση είναι ότι τα δεδομένα αυτά είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές (independent and identically distributed, iid).

Προφανώς, υπάρχουν διεργασίες που η πρώτη υπόθεση παραβιάζεται, γεγονός που επηρεάζει σημαντικά την απόδοση των διαγραμμάτων. Για την αντιμετώπιση τέτοιων περιπτώσεων, υπάρχει μια ξεχωριστή κατηγορία διαγραμμάτων, τα μη παραμετρικά διαγράμματα ελέγχου, τα οποία μπορούν να χρησιμοποιηθούν ακόμα κι αν η κατανομή που χαρακτηρίζει την υπό εξέταση διεργασία δεν είναι γνωστή. Η μελέτη αυτών των κατηγοριών διεργασιών δεν εξετάστηκε στο πλαίσιο αυτής της διπλωματικής εργασίας.

Ωστόσο, έχει παρατηρηθεί ότι υπάρχουν πολλές περιπτώσεις κανονικών διεργασιών, όπου τα δεδομένα που συλλέγονται είναι αυτοσυσχετιζόμενα. Παραβιάζεται δηλαδή η δεύτερη υπόθεση, καθώς οι παρατηρήσεις δεν είναι ανεξάρτητες χρονικά μεταξύ τους. Π.χ. δύο συνεχόμενες μετρήσεις μπορεί να έχουν παραπλήσια τιμή με θετική αυτοσυσχέτιση λόγω αδράνειας της διεργασίας, όπως παρατηρείται συχνά σε παραγωγικές διεργασίες της χημικής βιομηχανίας και της βιομηχανίας τροφίμων.

Οι τεχνικές δειγματοληψίας που είδαμε στο προηγούμενο κεφάλαιο δεν λαμβάνουν υπόψη την παρουσία αυτοσυσχέτισης ανάμεσα στις παρατηρήσεις. Έχουν αναπτυχθεί, ωστόσο, τε-

χνικές δειγματοληψίας που αντιμετωπίζουν το πρόβλημα αυτό. Οι τεχνικές αυτές δεν βασίζονται στην ίδια λογική με αυτές που αναπτύχθηκαν στο Κεφάλαιο 7, δηλαδή σε ταξινόμηση των παρατηρήσεων των δειγμάτων.

## 8.2 Τεχνικές δειγματοληψίας για την αντιμετώπιση της αυτοσυσχέτισης

Για τη μείωση της επίδρασης της αυτοσυσχέτισης, οι Costa και Castagliola (2011) [16] πρότειναν μια εναλλακτική τεχνική δειγματοληψίας, την οποία αποκάλεσαν τεχνική s-skip. Ας θεωρήσουμε μια παραγωγική διεργασία κατά την οποία συλλέγονται δείγματα μεγέθους  $i = 1, 2, \dots, n$ , σε διαδοχικά χρονικά διαστήματα  $j = 1, 2, 3, \dots, m$ . Η βασική ιδέα της τεχνικής s-skip συνοψίζεται στην παράλειψη κάποιων από τις διαδοχικές παρατηρήσεις, συμβολίζοντας με  $s$  τον αριθμό των παρατηρήσεων που παραλείπονται. Έτσι, η μορφή των δειγμάτων θα έχει ως εξής:  $X_{1,j}, X_{s+2,j}, X_{2s+3,j}, X_{3s+4,j}, \dots$

Μια νέα τεχνική βασισμένη στην ίδια ιδέα προτάθηκε από τους Franco et al (2014) [18], η οποία ονομάστηκε τεχνική mixed samples. Σύμφωνα μ' αυτή την τεχνική, από δύο διαδοχικά δείγματα μεγέθους  $n$ , δηλαδή τα δείγματα:

$X_{1,j-1}, X_{2,j-1}, X_{3,j-1}, \dots, X_{n,j-1}$  της  $j - 1$  χρονικής στιγμής και

$X_{1,j}, X_{2,j}, X_{3,j}, \dots, X_{n,j}$  της  $j$  χρονικής στιγμής

δημιουργείται ένα μεικτό δείγμα μεγέθους  $n$  ως εξής:

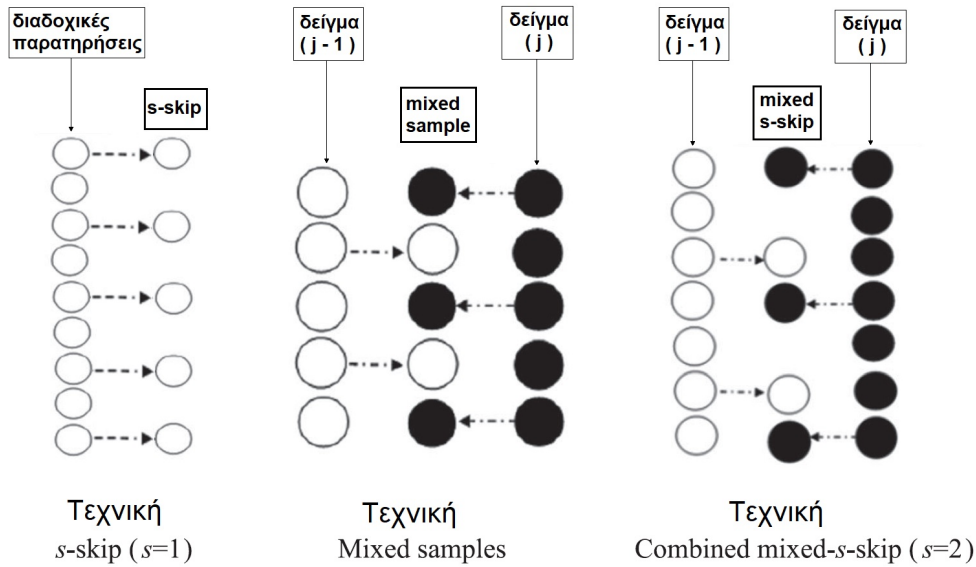
$X_{1,j}, X_{2,j-1}, X_{3,j}, X_{4,j-1}, \dots, X_{n-1,j}, X_{n,j-1}$  αν το  $n$  είναι άρτιο

$X_{1,j}, X_{2,j-1}, X_{3,j}, X_{4,j-1}, \dots, X_{n-1,j-1}, X_{n,j}$  αν το  $n$  είναι περιττό

Τέλος, στηριγμένοι στην παραπάνω τεχνική, οι Shongwe et al (2021) [48] πρότειναν έναν συνδυασμό των δύο προαναφερθέντων τεχνικών, τον οποίο ονόμασαν τεχνική mixed-s-skip. Σύμφωνα μ' αυτή την τεχνική, οι παρατηρήσεις δύο διαδοχικών δειγμάτων μεγέθους  $n$  αναμειγνύονται κατά την τεχνική mixed samples, ενώ παράλληλα κάποιες από τις διαδοχικές παρατηρήσεις παραλείπονται.

Μια συνοπτική σχηματική απεικόνιση και των τριών τεχνικών δειγματοληψίας φαίνεται στο Σχήμα 8.1.





Σχήμα 8.1: Τεχνικές δειγματοληψίας για την αντιμετώπιση της αυτοσυσχέτισης

### 8.3 Πρόταση

Όπως προκύπτει από την προαναφερθείσα βιβλιογραφία, είναι διαθέσιμος ένας οδηγός επιλογής της βέλτιστης τεχνικής δειγματοληψίας για ανίχνευση μετατοπίσεων του μέσου μιας διεργασίας, αρκεί να είναι γνωστό το μέγεθος της μετατόπισης προς ανίχνευση ( $\delta$ ), ο συντελεστής συσχέτισης των παρατηρήσεων ( $\rho$ ) και το μέγεθος των δειγμάτων ( $n$ ).

Ωστόσο, μελέτες πάνω τις υπόψη τεχνικές, καθώς και σύγκριση με άλλες κατηγορίες τεχνικών δειγματοληψίας, έχουν γίνει μόνο για την περίπτωση των απλών διαγραμμάτων Shewhart. Συνεπώς, προτείνεται για μελλοντική έρευνα η εφαρμογή τεχνικών δειγματοληψίας για αντιμετώπιση αυτοσυσχέτισης και σε διαγράμματα EWMA.



# Παραρτήματα

# Παράρτημα Α

## Δεδομένα

Στον Πίνακα Α.1, οι πρώτες 20 παρατηρήσεις προέρχονται από κανονική κατανομή  $N(10,1)$ , ενώ στις επόμενες 10 από κανονική κατανομή  $N(11,1)$  (μετατόπιση μέσου κατά  $1\sigma$ ).

Πίνακας Α.1: Δεδομένα Montgomery

$i$	$X_i$	$i$	$X_i$	$i$	$X_i$
1	9.45	11	9.03	21	10.90
2	7.99	12	11.47	22	9.33
3	9.29	13	10.51	23	12.29
4	11.66	14	9.40	24	11.50
5	12.16	15	10.08	25	10.60
6	10.18	16	9.37	26	11.08
7	8.04	17	10.62	27	10.38
8	11.46	18	10.31	28	11.62
9	9.20	19	8.52	29	11.31
10	10.34	20	10.84	30	10.52

Στον Πίνακα Α.2, οι 10 παρατηρήσεις προέρχονται από εντός ελέγχου διεργασία ( $\mu_0 = 100$ ).

Πίνακας Α.2: Δεδομένα εντός ελέγχου διεργασίας

$i$	$X_i$	$(C_0^+ = C_0^- = 0)$		$(C_0^+ = C_0^- = 6)$	
		$C_i^+$	$C_i^-$	$C_i^+$	$C_i^-$
1	102	0	0	5	1
2	97	0	0	0	1
3	104	1	0	1	0
4	93	0	4	0	4
5	100	0	1	0	1
6	105	2	0	2	0
7	96	0	1	0	1
8	98	0	0	0	0
9	105	2	0	2	0
10	99	0	0	0	0

Στον Πίνακα Α.3, οι 10 παρατηρήσεις προέρχονται από εκτός ελέγχου διεργασία, όπου ο μέσος έχει μετακινηθεί από  $\mu_0 = 100$  στο  $\mu_1 = 105$ .

Πίνακας Α.3: Δεδομένα εκτός ελέγχου διεργασίας

$i$	$X_i$	$(C_0^+ = C_0^- = 0)$		$(C_0^+ = C_0^- = 6)$	
		$C_i^+$	$C_i^-$	$C_i^+$	$C_i^-$
1	107	4	0	10	0
2	102	3	0	9	0
3	109	9	0	15	0
4	98	4	0	10	0
5	105	6	0	12	0
6	110	13	0	19	0
7	101	11	0	17	0
8	103	11	0	17	0
9	110	18	0	24	0
10	104	19	0	25	0

Στον Πίνακα A.4, οι παρατηρήσεις αναπαριστούν μια πρωτοεμφανιζόμενη εκτός ελέγχου κατάσταση.

Πίνακας A.4: Απλό παράδειγμα EWMA για  $\lambda=0.1$

<i>SampleNumber</i>	$\bar{x}_t$	$z_t$	$UCL(t)$
0	-	0	0
1	0.8	0.08	0.30
2	1.9	0.26	0.40
3	1.4	0.38	0.47
4	2.0	0.54	0.52
5	1.1	0.59	0.56
6	0.7	0.61	0.58
7	2.6	0.80	0.60
8	0.5	0.77	0.62
9	1.2	0.82	0.63

Στον Πίνακα A.5, οι παρατηρήσεις προέρχονται από κανονική κατανομή με μικρότερη διασπορά από τις αντίστοιχες του Πίνακα A.1. Στις πρώτες 20 ο μέσος παραμένει σταθερός στην τιμή 10, ενώ στις επόμενες 10 ο μέσος έχει μετατοπιστεί κατά  $1\sigma$ .

Πίνακας A.5: Δεδομένα

A/A	$X_i$	A/A	$X_i$	A/A	$X_i$
1	9.75	11	9.4	21	10.4
2	9.95	12	10.4	22	10.6
3	10.1	13	10.3	23	10.93
4	9.85	14	10.1	24	10.7
5	10.15	15	10	25	10.9
6	10.25	16	10.1	26	11.2
7	9.6	17	9.9	27	10.9
8	9.7	18	10.14	28	11.6
9	10.2	19	10.01	29	11.22
10	10.11	20	9.99	30	11.55

Στους Πίνακες A.6 και A.7, ο κάθε κύκλος  $j$  (με  $j=1, 2, \dots, 30$ ) αποτελείται από 4 σετ δειγματοληψίας, το καθένα με  $n=4$  παρατηρήσεις. Σε κάθε κύκλο, το  $\bar{X}_{RSS j}$  που υπολογίζεται είναι ίδιο με τα  $X_i$  του Πίνακα A.1.

Πίνακας Α.6: Δεδομένα για παράδειγμα RSS-EWMA (1)

$j$	1ο σετ				2ο σετ			
		δειγματοληψίας				δειγματοληψίας		
1	8.80	9.00	9.50	10.00	8.80	9.05	9.60	10.40
2	7.99	8.30	9.00	10.10	8.10	8.80	9.00	10.40
3	10.00	10.25	10.60	10.80	8.50	9.00	10.00	10.60
4	10.40	10.50	10.60	10.70	10.00	12.68	12.70	12.80
5	12.90	12.91	12.93	12.95	11.00	11.40	11.60	12.00
6	11.50	11.70	11.80	12.30	9.60	10.00	10.50	11.20
7	7.10	8.00	8.70	9.80	8.70	9.20	9.50	10.90
8	10.79	10.90	11.50	11.90	9.80	11.40	11.70	12.00
9	8.91	9.20	9.80	10.60	9.60	10.13	10.80	12.30
10	11.32	11.50	12.00	12.30	9.20	10.10	10.50	11.70
11	8.49	8.90	9.30	10.50	7.80	8.00	8.30	9.50
12	11.89	11.95	12.40	12.60	10.20	11.00	11.40	11.60
13	10.43	10.60	10.80	11.20	10.80	11.20	11.50	11.90
14	9.50	9.70	10.00	10.90	10.20	10.65	10.90	11.30
15	11.05	11.20	11.40	12.00	10.90	11.15	11.30	11.70
16	8.60	8.80	9.00	10.00	8.10	8.26	8.50	9.40
17	11.50	11.60	11.90	12.40	12.00	12.50	12.60	12.90
18	12.42	12.50	12.70	12.80	8.40	9.05	9.60	10.00
19	11.05	11.20	11.40	11.80	7.80	8.05	8.50	9.90
20	10.00	10.30	10.50	11.30	9.00	9.45	10.00	12.00
21	9.10	9.40	10.00	11.00	11.20	12.20	12.50	12.80
22	7.90	8.00	8.30	9.30	8.10	8.65	9.30	10.50
23	12.85	12.90	12.91	12.92	10.00	10.65	11.40	12.00
24	12.80	12.82	12.84	12.86	10.90	11.25	11.50	12.50
25	12.40	12.50	12.60	12.70	9.00	9.90	10.40	11.50
26	12.65	12.70	12.75	12.80	9.20	10.05	10.50	11.30
27	10.00	10.40	10.90	11.30	8.10	8.85	9.30	10.50
28	12.45	12.50	12.70	12.80	7.40	9.53	11.00	11.50
29	11.52	11.60	11.80	12.30	8.70	9.65	10.00	11.30
30	11.30	11.50	11.70	12.00	9.10	9.56	10.50	11.00

Πίνακας Α.7: Δεδομένα για παράδειγμα RSS-EWMA (2)

$j$	3ο σετ				4ο σετ			
	δειγματοληψίας				δειγματοληψίας			
1	8.10	9.40	9.75	10.80	7.90	8.50	9.40	10.20
2	7.00	7.10	7.15	9.00	7.10	7.40	7.60	8.00
3	8.40	9.50	10.00	10.40	7.50	7.80	8.00	8.14
4	10.10	11.00	11.50	12.00	10.00	11.00	12.00	12.06
5	11.00	11.60	12.00	12.40	11.00	11.80	12.10	12.32
6	9.00	9.30	9.50	10.90	8.30	9.00	9.50	9.70
7	7.80	8.20	8.45	9.60	7.10	7.20	7.35	7.40
8	9.70	11.40	12.00	12.30	9.90	10.60	11.20	11.65
9	8.10	8.50	8.70	9.80	7.40	8.40	8.90	9.05
10	9.10	9.40	9.70	10.90	9.00	9.50	10.00	10.25
11	9.20	9.70	10.00	10.40	8.80	9.30	9.50	9.64
12	10.20	11.40	12.40	12.60	9.10	10.00	10.30	10.59
13	8.50	9.20	9.45	9.90	9.10	10.20	10.70	10.95
14	8.40	8.80	9.00	10.40	7.80	8.00	8.20	8.45
15	8.20	9.40	10.00	10.90	7.50	7.80	8.00	8.10
16	10.00	10.80	11.30	11.60	8.00	8.70	9.00	9.30
17	8.10	8.40	8.94	9.80	8.00	9.00	9.40	9.55
18	9.50	10.00	10.35	11.00	8.30	8.80	9.10	9.40
19	7.30	7.40	7.50	9.00	7.10	7.20	7.30	7.47
20	10.00	10.90	11.20	11.50	10.40	11.40	12.00	12.70
21	10.00	11.00	11.35	12.00	9.80	10.10	10.50	10.95
22	9.00	10.00	11.00	12.00	8.40	9.20	9.50	9.75
23	11.00	12.00	12.95	12.98	10.30	11.00	12.00	12.70
24	10.30	11.40	11.85	12.50	9.00	9.40	9.70	10.10
25	9.20	10.00	10.50	11.50	8.40	8.80	9.20	9.60
26	9.70	10.40	10.70	11.90	9.30	10.30	10.70	10.90
27	10.10	11.00	11.85	12.50	9.60	10.10	10.40	10.80
28	11.00	11.40	12.70	12.90	10.00	10.80	11.20	11.80
29	10.00	11.00	11.55	12.40	10.40	11.50	12.00	12.50
30	11.20	11.80	12.10	12.40	8.30	8.70	9.00	9.10



## Παράρτημα Β

### Κώδικας υπολογισμού *ARL* σε *R*

```
‘‘{r setup, echo=FALSE, results='hide', warning=FALSE,
  message=FALSE}

# Create the list of required packages
list.of.packages <- c("spc", "pracma")

# Check list of required packages against already
# "installed.packages" or "old.packages" (where an update
# is required) to decypher which (if any) should be
# downloaded and installed
new.packages <- list.of.packages[(!list.of.packages %in%
installed.packages() | list.of.packages %in% old.packages()
[, "Package"])]

# If indeed new packages list is not empty proceed in
# downloading missing packages
if(length(new.packages)) install.packages(new.packages,
repos = "http://cran.us.r-project.org")

# Load all necessary packages in RAM (working environment)
lapply(list.of.packages, require, character.only = TRUE)

# Setup working directory to the path of the source file
# location
setwd(dirname(rstudioapi::getActiveDocumentContext())$path))

# Reproducibility
SEED = 0
```

```

# Setup chunks options
knitr::opts_chunk$set(
  # message = FALSE,
  # warning = FALSE,
  fig.width = 8,
  fig.height = 4.5,
  fig.align = 'center',
  out.width='95%',
  dpi = 200
)

## Select nr of threads left for other than the present
## notebook's task
NR_CORES_LEFT_FOR_OTHER_TASKS = 1

## Set up the # of threads involved in code workload
plan(multiprocess(workers = availableCores()
- NR_CORES_LEFT_FOR_OTHER_TASKS))

rm(list.of.packages,new.packages)
'''

#####

''{r}
# Initialization of statistics

# SAMPLE
## sample <- c(10.66, 10.81, 11.19, ...)
## n = length(sample) # sample size

n <- 4 # sample size

# Process Mean shifts from m0=0 (in-control) to several
# values of m1 (out-of-control)
sigma0 <- 1 # Process Standard Deviation
m0 <- 0 # Process Mean
mu1 <- c(0, 0.1, 0.2, 0.3, 0.4, ...)
'''

```



```

# MOD_EWMA  <- 5 # Modified EWMA
# Shewhart  <- 6 # Shewhart

# Selected type of limits:
limits      <- 2

# Steiner Method values:
fSt = 0.5
aSt = (-2./log10(1. - fSt) - 1.)/19.

# Asymptotic Limits calculation
UCL=as.double()
LCL=as.double()
for (f in 1:length(m1)){
  for (h in 1:length(L)){
    if (limits==1 || limits==2 || limits==3 || limits==4){
      # Asymptotic limits for EWMA (type: 1, 2, 3, 4)
      UCL[(f-1)*length(L)+h] <- m0 + L[h]*sigma[f]*
sqrt(1/n)*sqrt(lamda[h]/(2-lamda[h]))

      LCL[(f-1)*length(L)+h] <- m0 - L[h]*sigma[f]*sqrt(1/n)*
sqrt(lamda[h]/(2-lamda[h]))
    }

    if (limits==5){
      # Asymptotic limits for MODIFIED EWMA
      UCL[(f-1)*length(L)+h] <- m0 + L[h]*sigma[f]*sqrt(1/n)*
sqrt(lamda[h]/(2-lamda[h])) + 2*lamda[h]*
(1-lamda[h])/(2-lamda[h]))

      LCL[(f-1)*length(L)+h] <- m0 - L[h]*sigma[f]*sqrt(1/n)*
sqrt(lamda[h]/(2-lamda[h])) + 2*lamda[h]*
(1-lamda[h])/(2-lamda[h]))
    }

    if (limits==6){
      # Asymptotic limits Shewhart
      UCL[(f-1)*length(L)+h] <- m0 + L[h]*sigma[f]*sqrt(1/n)
      LCL[(f-1)*length(L)+h] <- m0 - L[h]*sigma[f]*sqrt(1/n)
    }
  }
}

```

```

}

# Markov chain method (non-homogeneous, g discrete states)
total_states=30 # number of discrete states (g-1 states
# "in control" and the g-th state "out of control")

reps=3 # for the calculation of ERL (1*total_states,
# 2*total_states and 3*total_states)

ERL=as.double()
ERL_matrix = matrix(data=NA, reps, 1)
'''

#####

'''{r}
st1 = Sys.time()

# 3 values of ERL will be calculated (for g=total_states,
# g=total_states*2 and g=total_states*3)

for (f in 1:length(m1)){
  for (h in 1:length(L)){
    for (pit in 1:reps){
      g=total_states*pit
      states <- g-1
      w <- (UCL[(f-1)*length(L)+h]-LCL[(f-1)*length(L)+h])/g
      S <- matrix(nrow=1, ncol=g-1)
      P <- matrix(nrow=g, ncol=g)
      R <- matrix(nrow=g-1, ncol=g-1)
      tmax<-log(((12*n*w*(2-lamda[h])*sigma0*(lamda[h]/(n*
(2-lamda[h]))))^(1/2)-w)/(36*lamda[h]*(sigma0)^2),10)/
(2*log((1-lamda[h]),10))
      tmax<-round(tmax,digits=0) # number of periods for
# which the transition probability matrices are used

# Calculation of matrix of discrete states S[1x(g-1)]
for(j in 1:dim(S)[2]){
  S[1,j] = LCL[(f-1)*length(L)+h] + w*j
}

```

```

# Calculation of transition probability matrix P[gxg]
if (limits==1 || limits==2 || limits==3 || limits==4){
  for(k in 1:dim(R)[1]){
    for(l in 1:dim(P)[2]){
      P[k,l] = (pnorm(S[l]+w/2, mean=lamda[h]*m1[f]+
        (1-lamda[h])*S[k], sd=lamda[h]*
        sigma[f]*sqrt(1/n), lower.tail=TRUE)
        - pnorm(S[l]-w/2, mean=lamda[h]*
        m1[f]+(1-lamda[h])*S[k], sd=lamda[h]*
        sigma[f]*sqrt(1/n), lower.tail=TRUE))
    }
  }
  for(i in 1:dim(R)[2]){
    P[i,g] = (pnorm(S[g-1]+w/2, mean=lamda[h]*m1[f]+
      (1-lamda[h])*S[i], sd=lamda[h]*
      sigma[f]*sqrt(1/n), lower.tail=FALSE)
      + pnorm(S[1]-w/2, mean=lamda[h]*m1[f]+
      (1-lamda[h])*S[i], sd=lamda[h]*
      sigma[f]*sqrt(1/n), lower.tail=TRUE))
  }
  for(j in 1:dim(R)[2]){
    P[g,j] = 0
  }
  P[g,g] = 1
}

if (limits==5){
  for(k in 1:dim(R)[1]){
    for(l in 1:dim(P)[2]){
      P[k,l] = (pnorm(S[l]+w/2, mean=lamda[h]*m1[f]+
        (1-lamda[h])*S[k]+(m1[f]-m0), sd=
        lamda[h]*sigma[f]*sqrt(1/n),
        lower.tail=TRUE)
        - pnorm(S[l]-w/2, mean=lamda[h]*m1[f]+
        (1-lamda[h])*S[k]+(m1[f]-m0), sd=
        lamda[h]*sigma[f]*sqrt(1/n),
        lower.tail=TRUE))
    }
  }
  for(i in 1:dim(R)[2]){

```

```

P[i,g] = (pnorm(S[g-1]+w/2, mean=lamda[h]*m1[f]+
              (1-lamda[h])*S[i]+(m1[f]-m0), sd=
              lamda[h]*sigma[f]*sqrt(1/n),
              lower.tail=FALSE)
          + pnorm(S[1]-w/2, mean=lamda[h]*m1[f]+
              (1-lamda[h])*S[i]+(m1[f]-m0), sd=
              lamda[h]*sigma[f]*sqrt(1/n),
              lower.tail=TRUE))
}
for(j in 1:dim(R)[2]){
  P[g,j]=0
}
P[g,g]=1
}

if (limits==6){
  for(k in 1:dim(R)[1]){
    for(l in 1:dim(P)[2]){
      P[k,l] = (pnorm(S[1]+w/2,
                    mean=m1[f],
                    sd=sigma[f]*sqrt(1/n),
                    lower.tail=TRUE)
                - pnorm(S[1]-w/2,
                    mean=m1[f],
                    sd=sigma[f]*sqrt(1/n),
                    lower.tail=TRUE))
    }
  }
  for(i in 1:dim(R)[2]){
    P[i,g] = (pnorm(S[g-1]+w/2,
                  mean=m1[f],
                  sd=sigma[f]*sqrt(1/n),
                  lower.tail=FALSE)
              + pnorm(S[1]-w/2,
                  mean=m1[f],
                  sd=sigma[f]*sqrt(1/n),
                  lower.tail=TRUE))
  }
  for(j in 1:dim(R)[2]){
    P[g,j]=0
  }
}

```

```

    P[g,g]=1
  }

  # Construction of matrix R [(g-1)x(g-1)]
  R = P[1:g-1, 1:g-1]

  # Rounding the elements of matrices P and R
  P <- round(P,digits=20)
  R <- round(R,digits=20)

  # Calculation of the time-varying limits of
  # the EWMA diagram

  if (limits==1){
    ewma_UCL <- function(t){
      UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)*
        sqrt(lamda[h]/(2-lamda[h]))
      return(UCLt)
    }
    ewma_LCL <- function(t){
      LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)*
        sqrt(lamda[h]/(2-lamda[h]))
      return(LCLt)
    }
  }

  if (limits==2){
    ewma_UCL <- function(t){
      UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)*
        sqrt(lamda[h]/(2-lamda[h])*(1-(1-lamda[h])^(2*t)))
      return(UCLt)
    }
    ewma_LCL <- function(t){
      LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)*
        sqrt(lamda[h]/(2-lamda[h])*(1-(1-lamda[h])^(2*t)))
      return(LCLt)
    }
  }

  if (limits==3){
    ewma_UCL <- function(t){

```



```

    UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
      (2-lamda[h])*(1-(1-lamda[h])^(2*t)))*(1-(1-fSt)^
      (1.+aSt*(t-1)))
    return(UCLt)
  }
ewma_LCL <- function(t){
  LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
    (2-lamda[h])*(1-(1-lamda[h])^(2*t)))*(1-(1-fSt)^
    (1.+aSt*(t-1)))
  return(LCLt)
}
}

if (limits==4){
ewma_UCL <- function(t){
  UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
    (2-lamda[h])*(1-(1-lamda[h])^(2*t)))*(1-(1-fSt)^
    (1.+aSt*(t-1)))^(1+(1/t))
  return(UCLt)
}
ewma_LCL <- function(t){
  LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
    (2-lamda[h])*(1-(1-lamda[h])^(2*t)))*(1-(1-fSt)^
    (1.+aSt*(t-1)))^(1+(1/t))
  return(LCLt)
}
}

if (limits==5){
ewma_UCL <- function(t){
  UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
    (2-lamda[h]) + 2*lamda[h]*(1-lamda[h])/(2-lamda[h]))
  return(UCLt)
}
ewma_LCL <- function(t){
  LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)*sqrt(lamda[h]/
    (2-lamda[h]) + 2*lamda[h]*(1-lamda[h])/(2-lamda[h]))
  return(LCLt)
}
}
}

```

```

if (limits==6){
  ewma_UCL <- function(t) {
    UCLt <- m0 + L[h]*sigma[f]*sqrt(1/n)
    return(UCLt)
  }
  ewma_LCL <- function(t) {
    LCLt <- m0 - L[h]*sigma[f]*sqrt(1/n)
    return(LCLt)
  }
}

# Calculation of the asymptotic limits of the
# EWMA diagram

sf1 <- function(t){
  sf1 <- ewma_LCL(t)+w/2
  return(sf1)
}
sf2 <- function(t){
  sf2 <- ewma_UCL(t)-w/2
  return(sf2)
}
St <- matrix(nrow=1, ncol=g-1)
Rt <- matrix(nrow=g-1, ncol=g-1)

# Comparing the time-varying control limits with the
# states in the state space
Rt <- function(t){
  Rt <- R
  for (i in 1:dim(R)[2]){
    if (S[i] <= sf1(t)||S[i] >= sf2(t)){
      Rt[i, 1:dim(R)[2]] = 0
      Rt[1:dim(R)[2], i] = 0
    }
  }
  return(Rt)
}
St <- function(t){
  St <- S
  for (i in 1:dim(S)[2]){
    if (S[i] <= sf1(t)||S[i] >= sf2(t)){

```

```

        St[i] = 0
    }
}
return(St)
}

# Calculating the expected run length ERL and Variance

I <- matrix(nrow=g-1, ncol=g-1) # Construction of gxg
# identity matrix
I[1:dim(I)[1], 1:dim(I)[2]] = 0
for (i in 1:g-1){
    I[i, i] = 1
}
ena <- matrix(nrow=g-1, ncol=1) # Construction of gx1
# column vector of ones
ena[1:g-1,1] = 1

ginomeno <- function(t){
    ginomeno <- R
    for (i in 1:t){
        if (i==1){
            ginomeno = Rt(1)
        }
        if (i>1){
            ginomeno = ginomeno(i-1)%*%Rt(i)
        }
    }
    return(ginomeno)
}

ginomeno1 <- function(t){
    ginomeno <- R
    for (i in 1:t){
        if (i==1){
            ginomeno1 = Rt(tmax)
        }
        if (i>1){
            ginomeno1 = ginomeno1%*%ginomeno1
        }
    }
}

```

```

    return(ginomeno1)
}

sum_ginomenon <- function(t){
  sum_ginomenon <- matrix(data=NA,g-1,1)
  for (i in 1:t){
    if (i==1){
      sum_ginomenon=ginomeno(1)%*%ena
    }
    if (i>1){
      sum_ginomenon=sum_ginomenon(i-1)+ginomeno(i)%*%ena
    }
  }
  return(sum_ginomenon)
}

ARL <- sum_ginomenon(tmax-1) +
      (ginomeno(tmax)%*%inv(I-R))%*%ena

ARL1 <- ARL[(g-1)/2, 1] # ERL value corresponding
                       # to z_0 = X_bar_bar

ERL_matrix[pit, 1] = ARL1 # above ERL value placed as
                          # 1st(g=total_states),
                          # 2nd(g=2*total_states)
                          # and 3rd(g=3*total_states)
                          # element of 3x1 matrix
}

# Above, 3 values of ERL have been calculated
# (for g=total_states, g=2*total_states and
# g=3*total_states) and contained in matrix "ERL_matrix"
# As g increases, ERL increases too. Taking advantage of
# the rate of increase, the true value can be estimated
# by fitting the model ERL=ERT(true)+B/g+C/g^2:

agnostoi = matrix(data=NA, reps, 1)
times = matrix(data=NA, reps, reps)
for (i in 1:reps){
  for (j in 1:reps){
    times[i,j]=1/(total_states*i)^(j-1)
  }
}

```

```

    }
  }
  agnostoi = inv(times)**ERL_matrix
  ERL[(f-1)*length(L)+h] = agnostoi[1,1]
}
}
ERL
Sys.time()-st1
'''

#####

'''{r}
# SRS RESULTS PRESENTATION
results <- t(matrix(ERL, nrow=1, ncol=length(mu1)*length(L),
byrow=TRUE))
delta <- rep(mu1, each=length(L))
l.par <- rep(lamda, length(mu1))
L.par <- rep(L, length(mu1))
RESULTS <- cbind(l.par, L.par, results)
rownames(RESULTS) <- delta
RESULTS
'''

#####

'''{r}
# RSS RESULTS PRESENTATION
results <- t(matrix(ERL, nrow=length(Var_sample), ncol=
length(mu1)*length(L), byrow=TRUE))
colnames(results) <- Var_sample
delta <- rep(mu1, each=length(L))
l.par <- rep(lamda, length(mu1))
L.par <- rep(L, length(mu1))
RESULTS <- cbind(l.par, L.par, results)
rownames(RESULTS) <- delta
RESULTS
'''

```



# Βιβλιογραφία

- [1] Abbas, N., Riaz, M., and Does, R. J. M. M. (2011). Enhancing the performance of EWMA Charts. *Quality and Reliability Engineering*, 27:821–833.
- [2] Abbas, N., Riaz, M., and Does, R. J. M. M. (2013a). CS-EWMA Chart for monitoring process dispersion. *Quality and Reliability Engineering*, 29:653–663.
- [3] Abbas, N., Riaz, M., and Does, R. J. M. M. (2013b). Mixed Exponentially Weighted Moving Average–Cumulative Sum Charts for process monitoring. *Quality and Reliability Engineering*, 29:345–356.
- [4] Abujiya, M. R. and Lee, M. H. (2013). The three statistical control charts using Ranked Set Sampling. In *5th International Conference on Modeling, Simulation and Applied Optimization, ICMSAO*, pages 1–6.
- [5] Abujiya, M. R., Lee, M. H., and Riaz, M. (2014a). An Improved Combined Shewhart-EWMA Chart based on Double Median Ranked Set Sampling. *Journal of Computer Science and Computational Mathematics*, 4(2):23–28.
- [6] Abujiya, M. R., Lee, M. H., and Riaz, M. (2014b). Improving the Performance of Exponentially Weighted Moving Average Control Charts. *Quality and Reliability Engineering International*, 30:571–590.
- [7] Abujiya, M. R. and Muttlak, H. (2004). Quality Control Chart for the Mean using Double Ranked Set Sampling. *Journal of Applied Statistics*, 31(10):1185–1201.
- [8] Abujiya, M. R., Riaz, M., and Lee, M. H. (2013). Enhancing the Performance of Combined Shewhart-EWMA Charts. *Quality and Reliability Engineering International*, 29:1093–1106.
- [9] Balakrishnan, N. and Li, T. (2005). Blues of Parameters of Generalized Geometric Distribution Using Ordered Ranked Set Sampling. *Communications in Statistics - Simulation and Computation*, 34(2):253–266.
- [10] Barnard, G. A. (1959). Control Charts and Stochastic Processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 21(2):239–271.

- [11] Best, M. and Neuhauser, D. (2006). Walter A Shewhart, 1924, and the Hawthorne factory. *Quality safety in health care*, 15:142–143.
- [12] Bissell, A. F. (1969). Cusum Techniques for Quality Control. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 18:1–30.
- [13] Capizzi, G. and Masarotto, G. (2003). An Adaptive Exponentially Weighted Moving Average Control Chart. *Technometrics*, 45(3):199–207.
- [14] Capizzi, G. and Masarotto, G. (2010). Combined Shewhart–EWMA control charts with estimated parameters. *Journal of Statistical Computation and Simulation*, 80(7):793–807.
- [15] Chiu, W.-C. (2009). Generally weighted moving average control charts for fast initial response features. *Journal of Applied Statistics*, 36(3):255–275.
- [16] Costa, A. F. B. and Castagliola, P. (2011). Effect of measurement error and autocorrelation on the  $\bar{X}$  chart. *Journal of Applied Statistics*, 38(4):661–673.
- [17] Crowder, S. V. (1987). A Simple Method for studying Run-Length Distributions of Exponentially Weighted Moving Average Charts. *Technometrics*, 29(4):401–407.
- [18] Franco, B. C., Castagliola, P., Celano, G., and Costa, A. F. B. (2014). A new sampling strategy to reduce the effect of autocorrelation on a control chart. *Journal of Applied Statistics*, 41(7):1408–1421.
- [19] Gan, F. F. (1995). Joint Monitoring of Process Mean and Variance Using Exponentially Weighted Moving Average Control Charts. *Technometrics*, 37(4):446–453.
- [20] Haq, A. (2013). A New Hybrid Exponentially Weighted Moving Average Control Chart for Monitoring Process Mean. *Quality and Reliability Engineering*, 29:1015–1025.
- [21] Haq, A., Brown, J., and Moltchanova, E. (2014a). Improved Fast Initial Response Features for Exponentially Weighted Moving Average and Cumulative Sum Control Charts. *Quality and Reliability Engineering International*, 30:697–710.
- [22] Haq, A., Brown, J., and Moltchanova, E. (2014b). New Exponentially Weighted Moving Average Control Charts for Monitoring Process Mean and Process Dispersion. *Quality and Reliability Engineering*, 31:877–901.
- [23] Haq, A., Brown, J., Moltchanova, E., and Al-Omari, A. I. (2015). Improved Exponentially Weighted Moving Average Control Charts for Monitoring Process Mean and Dispersion. *Quality and Reliability Engineering International*, 31:217–237.
- [24] Hawkins, D. M. (1987). Self-Starting CUSUM Charts for Location and Scale. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 36(4):299–316.



- [25] Jiang, W., Shu, L., and Apley, D. W. (2008). Adaptive CUSUM procedures with EWMA-based shift estimators. *IIE Transactions*, 40(10):992–1003.
- [26] Khoo, M. B. C., Castagliola, P., Liew, J. Y., Teoh, W. L., and Maravelakis, P. E. (2016). A Study on EWMA charts with runs rules—the Markov chain approach. *Communications in Statistics - Theory and Methods*, 45(14):4156–4180.
- [27] Knoth, S. (2005a). Accurate ARL computation for EWMA- $S^2$  control charts. *Statistics and Computing*, 15(4):341–352.
- [28] Knoth, S. (2005b). Fast initial response features for EWMA Control Charts. *Statistical Papers*, 46:47–64.
- [29] Knoth, S. (2007). Accurate ARL Calculation for EWMA Control Charts Monitoring Normal Mean and Variance Simultaneously. *Sequential Analysis*, 26(3):251–263.
- [30] Knoth, S. (2015). Run length quantiles of EWMA control charts monitoring normal mean or/and variance. *International Journal of Production Research*, 53(15):4629–4647.
- [31] Knoth, S. (2021). spc: Statistical Process Control – Calculation of ARL and Other Control Chart Performance Measures. <https://CRAN.R-project.org/package=spc>. R package version 0.6.5.
- [32] Knoth, S. and Schmid, W. (2002). Monitoring the mean and the variance of a stationary process. *Statistica Neerlandica*, 56:77–100.
- [33] Li, S.-Y., Tang, L.-C., and Ng, S.-H. (2010). Nonparametric CUSUM and EWMA Control Charts for Detecting Mean Shifts. *Journal of Quality Technology*, 42(2):209–226.
- [34] Lucas, J. M. and Crosier, R. B. (1982). Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM a Head Start. *Technometrics*, 24(3):199–205.
- [35] Lucas, J. M. and Saccucci, M. S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics*, 32(1):1–12.
- [36] McIntyre, G. A. (2005). A Method for Unbiased Selective Sampling, Using Ranked Sets. *The American Statistician*, 59(3):230–232.
- [37] Montgomery, D. C. (2013). *Introduction to Statistical Quality Control*. John Wiley Sons, Inc., Hoboken, USA, 7th edition.
- [38] Morais, M. and Pacheco, A. (2006). *Misleading Signals in Joint Schemes for  $\mu$  and  $\sigma$* , pages 100–122. *Frontiers in Statistical Quality Control 8*. Physica-Verlag HD.
- [39] Muttlak, H. and Al-Sabah, W. (2003). Statistical quality control based on ranked set sampling. *Journal of Applied Statistics*, 30(9):1055–1078.

- [40] Muttlak, H. A. (1998). Median Ranked Set Sampling with concomitant variables and a comparison with Ranked Set Sampling and regression estimators. *Environmetrics*, 9:255–267.
- [41] Ou, Y., Wen, D., Wu, Z., and Boon Chong, M. (2012). A comparison study of effectiveness and robustness of control charts for monitoring process mean. *International Journal of Production Economics*, 135(1):479–490.
- [42] Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 1-2:100–115.
- [43] Patel, A. K. and Divecha, J. (2011). Modified exponentially weighted moving average (EWMA) control chart for an analytical process data. *Journal of Chemical Engineering*, 2(1):12–20.
- [44] Pongpullponsak, A. and Sontisamran, P. (2013). Statistical Quality Control Based on Ranked Set Sampling for Multiple Characteristics. *Chiang Mai Journal of Science*, 40(3):485–498.
- [45] Reynolds, M. and Stoumbos, Z. (2006). Comparisons of Some Exponentially Weighted Moving Average Control Charts for Monitoring the Process Mean and Variance. *Technometrics*, 48(4):550–567.
- [46] Rhoads, T. R., Montgomery, D. C., and Mastrangelo, C. M. (1996). A Fast Initial Response Scheme for the Exponentially Weighted Moving Average Control Chart. *Quality Engineering*, 9(2):317–327.
- [47] Roberts, S. W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics*, 1(3):239–250.
- [48] Shongwe, S. C., Malela-Majika, J.-C., and Castagliola, P. (2021). A combined mixed-skip sampling strategy to reduce the effect of autocorrelation on the  $\bar{X}$  scheme with and without measurement errors. *Journal of Applied Statistics*, 48(7):1243–1268.
- [49] Steiner, S. (1999). EWMA Control Charts with Time-Varying Control Limits and Fast Initial Response. *Journal of Quality Technology*, 31(1):75–86.
- [50] Takahasi, W. (1968). On Unbiased Estimates of the Population Mean Based on the Sample Stratified by Means of Ordering. *Annals of the Institute of Statistical Mathematics*, 20:1–31.
- [51] Waldmann, K.-H. (1986). Bounds for the Distribution of the Run Length of Geometric Moving Average Charts. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 35(2):151–158.
- [52] Wu, Z., Yang, M., Jiang, W., and Khoo, M. B. (2008). Optimization designs of the combined Shewhart-CUSUM control charts. *Computational Statistics and Data Analysis*, 53(2):496–506.

- [53] Zaman, B., Riaz, M., Abbas, N., and Does, R. J. M. M. (2015). Mixed Cumulative Sum–Exponentially Weighted Moving Average Control Charts: An Efficient Way of Monitoring Process Location. *Quality and Reliability Engineering*, 31:1407–1421.

