



**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
Σ ΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**ΕΠΙΔΡΑΣΗ ΤΗΣ ΕΚΤΙΜΗΣΗΣ ΤΩΝ
ΠΑΡΑΜΕΤΡΩΝ ΣΤΙΣ ΙΔΙΟΤΗΤΕΣ ΤΩΝ
ΔΙΑΓΡΑΜΜΑΤΩΝ ΕΛΕΓΧΟΥ
ΓΙΑ ΤΗ ΜΕΣΗ ΤΙΜΗ**

Γεώργιος Ν. Ψαράκης

Πειραιάς
Φεβρουάριος 2012

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΑΙΑ

Στην οικογένειά μου

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής μου εργασίας, θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή μου, κ. Δημήτριο Αντζουλάκο, για την πολύτιμη βοήθεια που μου προσέφερε, για όλο τον πολύτιμο χρόνο του που μου αφιέρωσε και για την χρήσιμη καθοδήγηση αλλά και τις υποδείξεις του καθ' όλη την διάρκεια της εργασίας και της συγγραφής της. Ακόμη θα ήθελα να ευχαριστήσω θερμά όλους τους καθηγητές μου για τον εμπλουτισμό των γνώσεων μου κατά την διάρκεια των σπουδών μου. Στο σημείο αυτό θα ήθελα επίσης να ευχαριστήσω βαθύτατα όλη την οικογένεια μου, αλλά και κάθε έναν ξεχωριστά, γιατί εκτός από αφανείς ήρωες και συμπαραστάτες της προσπάθειας αυτής, με στήριξαν με όποιο τρόπο εφικτό μπορούσαν και μου προσέφεραν αμέριστη βοήθεια, όμως πάνω από όλα μου συμπαραστάθηκαν ως σωστοί και υποδειγματικοί γονείς.

Περίληψη

Στα διαγράμματα ελέγχου Φάσης I τα όρια ελέγχου υπολογίζονται χρησιμοποιώντας εκτιμήσεις των άγνωστων παραμέτρων της κατανομής του χαρακτηριστικού που μας ενδιαφέρει, οι οποίες προκύπτουν από προκαταρκτικά δείγματα. Στα διαγράμματα ελέγχου Φάσης II, στατιστικές συναρτήσεις νέων δειγμάτων συγκρίνονται με τα όρια ελέγχου του διαγράμματος για να συμπερασματολογήσουμε γύρω από το κατά πόσο η παραγωγική διεργασία λειτουργεί σε ευσταθή κατάσταση ή όχι.

Η πλειονότητα των μελετών που έχουν γίνει για την απόδοση των διαγραμμάτων ελέγχου στη Φάση II βασίζεται στην υπόθεση ότι οι εντός στατιστικού ελέγχου παράμετροι της κατανομής του χαρακτηριστικού που μας ενδιαφέρει είναι γνωστές και ίσες με τις εκτιμώμενες τιμές των παραμέτρων που έχουν προκύψει στη Φάση I. Ωστόσο, όταν οι εκτιμώμενες παράμετροι χρησιμοποιούνται στη θέση των πραγματικών τιμών των παραμέτρων η απόδοση του διαγράμματος ελέγχου αναμένεται να είναι διαφορετική από αυτή των αντίστοιχων διαγραμμάτων ελέγχου με γνωστές παραμέτρους λόγω της μεταβλητότητας των εκτιμώμενων παραμέτρων.

Σκοπός της διπλωματικής εργασίας είναι η μελέτη της επίδρασης της εκτίμησης των παραμέτρων στις ιδιότητες των διαγραμμάτων ελέγχου Φάσης II για τη μέση τιμή μιας παραγωγικής διεργασίας.

Abstract

In Phase I control charts, control limits are calculated using estimates of unknown parameters of the distribution of the attribute of interest, arising from historical reference samples. In Phase II control charts, statistical functions of new samples are compared with the control chart limits to draw a conclusion around whether the manufacturing process operates in a stable condition or not.

The majority of studies conducted on the performance of control charts in Phase II based on the assumption that statistical control parameters of the distribution of the characteristic of interest are known and equal to the estimated parameter values obtained in Phase I. However, when the estimated parameters are used in place of actual parameter values, the performance of control chart will be different from that of the corresponding control charts with known parameters due to the variability of estimated parameters.

The main purpose of the dissertation is to study the effect of parameter estimation on the properties of control charts of Phase II, for monitoring the mean of a manufacturing process.

Περιεχόμενα

Κατάλογος Πινάκων	x
Κατάλογος Διαγραμμάτων	xii
1. Εισαγωγή	1
1.1 Διαγράμματα ελέγχου Shewhart	1
1.2 Φάση I και Φάση II	4
1.3 Εκτίμηση της μέσης τιμής και της διασποράς στα διαγράμματα ελέγχου Shewhart για μεταβλητές	5
1.3.1 Η περίπτωση των μεμονωμένων παρατηρήσεων	5
1.3.2 Η περίπτωση των δειγμάτων	8
1.4 Διαγράμματα ελέγχου για τη μέση τιμή	10
1.5 Διαγράμματα ελέγχου για τη διασπορά	12
1.6 Διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις	20
1.7 Μέγεθος και συχνότητα δειγματοληψίας	23
1.8 Το μέσο μήκος ροής ως μέτρο απόδοσης των διαγραμμάτων ελέγχου	25
2. Επίδραση της εκτίμησης των παραμέτρων στα διαγράμματα ελέγχου για τη μέση τιμή	30
2.1 Εισαγωγή	30
2.2 Η κατανομή του μήκους ροής κατά την εκτίμηση των παραμέτρων των διαγραμμάτων ελέγχου	31
2.2.1 Δεσμευμένη κατανομή του μήκους ροής	32
2.2.2 Περιθώρια κατανομή τού μήκους ροής	34
2.3 Προσδιορισμός της άγνωστης κατανομής του μήκους ροής	35
2.3.1 Ιδιότητες της κατανομής του μήκους ροής	36

2.3.2	Υπολογισμός των τιμών ARL και SD κατά την εκτίμηση του σ στα διαγράμματα ελέγχου για τη μέση τιμή	39
2.3.3	Επίδραση των εκτιμώμενων παραμέτρων στις τιμές ARL και $SDRL$ στα \bar{X} διαγράμματα ελέγχου	48
2.4	Συνοπτική ανασκόπηση σχετικής βιβλιογραφίας	50
3.	Επίδραση της εκτίμησης των παραμέτρων στα διαγράμματα ελέγχου για τη διασπορά	59
3.1	Εισαγωγή	59
3.2	Η κατανομή του μήκους ροής στα R , S και S^2 διαγράμματα ελέγχου με εκτιμώμενο σ	61
3.3	Εμβάθυνση στα S διαγράμματα ελέγχου με εκτιμώμενο σ	70
3.4	S διάγραμμα ελέγχου με όρια ελέγχου πιθανότητας και εκτιμώμενο σ	82
3.5	X διάγραμμα ελέγχου με εκτιμώμενο σ	92
4.	Συμπεράσματα – Προτάσεις	96
Βιβλιογραφία		

Κατάλογος Πινάκων

2.1	Μέσο μήκος ροής και τυπική απόκλιση του μήκους ροής για \bar{X} διαγράμματα ελέγχου για διάφορες εκτιμήσεις του σ (εντός ελέγχου διεργασία)	46
2.2	Μέσο μήκος ροής και τυπική απόκλιση του μήκους ροής για \bar{X} διαγράμματα ελέγχου για διάφορες εκτιμήσεις του σ (εκτός ελέγχου διεργασία, $n=5$)	47
2.3	Εκατοστημόρια και μέσο μήκος ροής στα \bar{X} Shewhart διαγράμματα ελέγχου με εκτιμώμενες παραμέτρους	55
2.4	Συνάρτηση κατανομής του μήκους ροής για εντός ελέγχου διεργασία με άγνωστη διασπορά που εκτιμάται από τη Σχέση (6), $n = 5, k = 3, \delta = 0$	57
2.5	Τιμές του ARL στα \bar{X} διαγράμματα ελέγχου με άγνωστη διασπορά που εκτιμάται από την Σχέση (6), $n = 5, k = 3$	57
3.1	Μέση τιμή και τυπική απόκλιση του μήκους ροής για R, S, S^2 διαγράμματα ελέγχου για διάφορες τιμές των m, n, b	68
3.2	Τιμές του συντελεστή συσχέτισης για διάφορες τιμές των m και n	73
3.3	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 5$)	75
3.4	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 10$)	76
3.5	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 20$)	77
3.6	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 50$)	78
3.7	Συντελεστής συσχέτισης για διάφορες τιμές των m και n για το S διάγραμμα ελέγχου με όρια ελέγχου πιθανότητας	85

3.8	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου ($a=0.002, n=5$)	86
3.9	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου ($a=0.002, n=10$)	87
3.10	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου ($a=0.002, n=20$)	88
3.11	Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου ($a=0.002, n=50$)	89
4.1	Τιμές των ARL και $SDRL$ για το X διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις για διάφορες τιμές του N	94

Κατάλογος Διαγραμμάτων

1.1	Τυπικό διάγραμμα ελέγχου Shewhart	2
1.2	Χαρακτηριστικές καμπύλες για $W \square N(10, 0.25^2/n)$, $n = 2, 5$	24
1.3	Γραφική παράσταση του μέσου μήκους ροής για $W \square N(10, 0.25^2/n)$, $n = 2, 5$	29
2.1	Ιστόγραμμα πυκνότητας σχετικών συχνοτήτων για την $\bar{R}/d_2(4)\sigma$ και προσαρμοσμένη πυκνότητα $f(w; 13.9259, 1.0181)$	44
2.2	Ιστόγραμμα πυκνότητας σχετικών συχνοτήτων για την $\bar{S}/c_4(4)\sigma$ και προσαρμοσμένη πυκνότητα $f(w; 14.2745, 1.0177)$	44
2.3	Συνάρτηση κατανομής του μήκους ροής για εντός ελέγχου διεργασία ($\delta = 0$) στην περίπτωση των άγνωστων παραμέτρων για $m = 30, 100, 500$, $n = 5$, $a = 0.0027$	53
2.4	Συνάρτηση κατανομής του μήκους ροής για εκτός ελέγχου διεργασία ($\delta = 0.5$) στην περίπτωση των άγνωστων παραμέτρων για $m = 30, 100, 500$, $n = 5$, $a = 0.0027$	54
3.1	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 5$, $m = 20$, $b = 1$	64
3.2	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 5$, $m = 75$, $b = 1$	64
3.3	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 10$, $m = 20$, $b = 0.8$	65
3.4	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 10$, $m = 75$, $b = 0.8$	65
3.5	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 10$, $m = 20$, $b = 1.2$	66

3.6	Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n = 10, m = 75, b = 1.2$	66
3.7	Συνάρτηση κατανομής του μήκους ροής για $n = 5$	80
3.8	Συνάρτηση κατανομής του μήκους ροής για $n = 10$	81
3.9	Συνάρτηση κατανομής του μήκους ροής για $n = 50$	81
3.10	Συνάρτηση κατανομής του μήκους ροής για $n = 5$ ($a = 0.002$)	91
3.11	Συνάρτηση κατανομής του μήκους ροής για $n = 50$ ($a = 0.002$)	92
4.1	Τιμές των ARL και $SDRL$ για το X διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις για διάφορες τιμές του n	95

ΚΕΦΑΛΑΙΟ 1

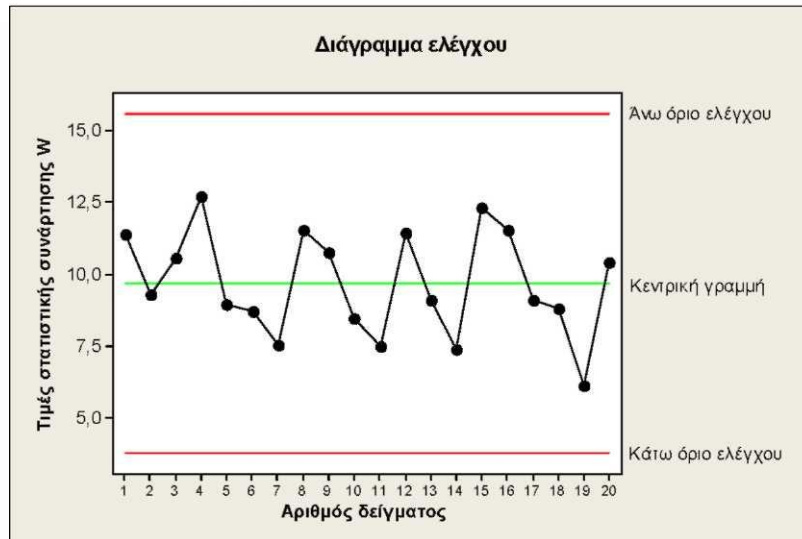
Εισαγωγή

1.1 Διαγράμματα ελέγχου Shewhart

Στις παραγωγικές διεργασίες μας ενδιαφέρει η παρακολούθηση της συμπεριφοράς μιας κρίσιμης ποσότητας ενός (μετρήσιμου) χαρακτηριστικού X (τυχαία μεταβλητή) των προϊόντων που παράγονται (για παράδειγμα το χαρακτηριστικό X μπορεί να είναι μήκος, βάρος, όγκος προϊόντων). Η διαδικασία παρακολούθησης της κρίσιμης ποσότητας βασίζεται σε μετρήσεις του χαρακτηριστικού X (τυχαία μεταβλητή), όπως προκύπτουν από την επιλογή τυχαίων δειγμάτων προϊόντων από την παραγωγή σε διαφορετικές χρονικές στιγμές στα οποία αντιστοιχούν τυχαία δείγματα τιμών του χαρακτηριστικού X , έστω τα $\mathbf{X}_1, \mathbf{X}_2, \dots$. Χρησιμοποιώντας τα τυχαία δείγματα $\mathbf{X}_1, \mathbf{X}_2, \dots$ υπολογίζουμε την τιμή $W_t = g(\mathbf{X}_t)$, $t=1, 2, \dots$, μιας κατάλληλης στατιστικής συνάρτησης (τυχαίας μεταβλητής) που εκτιμά (συνήθως αμερόληπτη εκτιμήτρια) την κρίσιμη ποσότητα που μας ενδιαφέρει. Έτσι η (διαχρονική) παρακολούθηση της συμπεριφοράς της κρίσιμης ποσότητας επιτυγχάνεται με την παρακολούθηση των τιμών που λαμβάνει η στατιστική συνάρτηση W στα διάφορα δείγματα.

Για παράδειγμα ας υποθέσουμε ότι ενδιαφερόμαστε να παρακολουθήσουμε την συμπεριφορά της μέσης τιμής μ της διαμέτρου X των κυλίνδρων που παράγει μια μηχανή. Για το σκοπό αυτό επιλέγονται τυχαία δείγματα μεγέθους n κυλίνδρων από την παραγωγή της μηχανής σε διαφορετικά χρονικά διαστήματα και μπορούμε να χρησιμοποιήσουμε την στατιστική συνάρτηση $W_t = g(\mathbf{X}_t) = (X_{t_1} + X_{t_2} + \dots + X_{t_n}) / n$ (η οποία είναι αμερόληπτη εκτιμήτρια του μέσου της κατανομής της X) για την παρακολούθηση της συμπεριφοράς της μέσης τιμής.

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



Διάγραμμα 1.1 Τυπικό διάγραμμα ελέγχου Shewhart

Στο Διάγραμμα 1.1, εκτός από τις διαδοχικές παρατηρούμενες τιμές της W_i , που έχουν παρασταθεί με κουκκίδες τα οποία έχουν συνδεθεί με μια τεθλασμένη γραμμή, έχουν σχεδιαστεί και άλλες τρεις γραμμές. Η κεντρική γραμμή (*center line, CL*) ή μέσο επίπεδο της διεργασίας παριστάνει συνήθως τη μέση τιμή (*mean value*) της W όπως αυτή προκύπτει από τη λειτουργία μιας εντός (στατιστικού) ελέγχου διεργασίας, δηλαδή μιας διεργασίας που λειτουργεί μόνο με την παρουσία φυσικής μεταβλητότητας. Οι δύο ακραίες γραμμές που εμφανίζονται ονομάζονται άνω και κάτω όρια ελέγχου (*upper and lower control limits, UCL and LCL*). Όσο οι τιμές (σημεία, δεδομένα) της W εμφανίζονται εντός των ορίων ελέγχου και η συμπεριφορά τους είναι “τυχαία” μπορούμε να υποθέσουμε ότι η διεργασία παραμένει εντός ελέγχου και δεν χρειάζεται να προβούμε σε κάποια διορθωτική ενέργεια. Αν όμως κάποιο σημείο βρεθεί εκτός των ορίων ελέγχου λέμε ότι υπάρχει ένδειξη ότι η διεργασία είναι εκτός ελέγχου και πρέπει να προχωρήσουμε σε έρευνα για να ανακαλύψουμε τις ειδικές αιτίες μεταβλητότητας που είναι υπεύθυνες για αυτή τη συμπεριφορά και αν κριθεί απαραίτητο να προβούμε σε διορθωτικές ενέργειες. Παρολαυτά, θα πρέπει να σημειώσουμε ότι 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.

Στο ακόλουθο πλαίσιο δίνεται ένα γενικό μοντέλο, το μοντέλο ορίων σίγμα (*sigma limits model*), για την κατασκευή ενός διαγράμματος ελέγχου

Μοντέλο ορίων L σίγμα

$$UCL = \mu_w + L\sigma_w$$

$$\text{Center Line} = \mu_w$$

$$LCL = \mu_w - L\sigma_w$$

Το μ_w και το σ_w δηλώνουν τη μέση τιμή και την τυπική απόκλιση της στατιστικής συνάρτησης W που απεικονίζεται στο διάγραμμα ελέγχου, η οποία εκτιμά την κρίσιμη ποσότητα ενός προϊόντος που θέλουμε να παρακολουθήσουμε, ενώ ο αριθμός L δηλώνει την απόσταση των ορίων ελέγχου από την κεντρική γραμμή σε μονάδες τυπικής απόκλισης. Συνήθως $L = 3$ οπότε ομιλούμε για διαγράμματα ελέγχου Shewhart με 3σ όρια ελέγχου (*three sigma control limits*).

Για μεγάλες τιμές του L η απόσταση των ορίων ελέγχου από την κεντρική γραμμή μεγαλώνει και έτσι μειώνεται ο κίνδυνος (ρίσκο, πιθανότητα) να βρεθεί ένα σημείο του διαγράμματος εκτός των ορίων ελέγχου. Σε αναλογία με την ορολογία που χρησιμοποιείται στους ελέγχους στατιστικών υποθέσεων, λέμε ότι μειώνεται η πιθανότητα του σφάλματος τύπου I (ή ρίσκο α), αλλά ταυτόχρονα αυξάνεται η πιθανότητα του σφάλματος τύπου II (ή ρίσκο β). Για μικρές τιμές του L έχουμε τα αντίθετα αποτελέσματα.

Για όρια ελέγχου 3σ και κανονική κατανομή της στατιστικής συνάρτησης W , όταν η διεργασία είναι εντός ελέγχου, η πιθανότητα να πάρει η W τιμή εκτός των ορίων ελέγχου, είναι ίση με $\alpha = 0.0027$. Το γεγονός αυτό σημαίνει ότι ένας λανθασμένος συναγερμός συμβαίνει 27 φορές ανα 10000 σημεία του διαγράμματος ελέγχου. Επίσης για εντός ελέγχου διεργασίες, η πιθανότητα να βρεθεί σημείο πέραν του LCL (UCL) είναι ίση με 0.00135.

Εκτός από το μοντέλο ορίων σίγμα, για την κατασκευή διαγραμμάτων ελέγχου Shewhart, υπάρχει και το μοντέλο ορίων πιθανότητας (*probability limits model*) που παρουσιάζεται στο ακόλουθο πλαίσιο (μοντέλο ορίων πιθανότητας α) για κανονική (ή προσεγγιστικά) κανονική κατανομή της στατιστικής συνάρτησης W που απεικονίζεται σε αυτό

Μοντέλο ορίων πιθανότητας α

$$UCL = \mu_W + z_{\alpha/2}\sigma_W = w_{\alpha/2}$$

$$\text{Center Line} = \mu_W$$

$$LCL = \mu_W - z_{\alpha/2}\sigma_W = w_{1-(\alpha/2)}$$

(με w_α συμβολίζεται το άνω α ποσοστιαίο σημείο της W). Για εντός ελέγχου διεργασίες (με κανονική κατανομή για την W) έχουμε ότι η πιθανότητα εσφαλμένου συναγερμού σε κάθε πλευρά των ορίων ελέγχου είναι ίση με $\alpha / 2$ (α είναι η πιθανότητα εσφαλμένου συναγερμού).

Στα διαγράμματα ελέγχου Shewhart διακρίνουμε δυο μεγάλες κατηγορίες ανάλογα με το αν το χαρακτηριστικό X είναι συνεχής ή διακριτή τυχαία μεταβλητή. Αν η τυχαία μεταβλητή X είναι συνεχής με μέση τιμή μ και διακύμανση σ^2 , τότε υπάρχουν διαγράμματα ελέγχου Shewhart για την παρακολούθηση της μέσης τιμής και της διασποράς της X . Στην περίπτωση που η τυχαία μεταβλητή X είναι διακριτή υπάρχουν διαγράμματα ελέγχου Shewhart για την παρακολούθηση του ποσοστού (και του αριθμού) των ελαττωματικών προϊόντων που αποδίδει η παραγωγική διεργασία, καθώς επίσης και για τον αριθμό (και τον μέσο αριθμό) των ελαττωμάτων (ατελειών) σε μια μονάδα ελέγχου.

1.2 Φάση I και Φάση II

Για τον έλεγχο μιας παραγωγικής διεργασίας υπάρχουν στην βιβλιογραφία δύο φάσεις (phases), η Φάση I και η Φάση II.

Φάση I (Phase I): Σε αυτήν τη φάση συλλέγεται ένα σύνολο δεδομένων από τη διεργασία που αναλύονται όλα μαζί ταυτοχρόνως, για να καθοριστεί αν η διεργασία ήταν εντός ή εκτός ελέγχου κατά τη χρονική περίοδο συλλογής των δεδομένων.

Σε αυτήν τη φάση τα διαγράμματα ελέγχου βοηθούν τον διαχειριστή της διαδικασίας να “φέρει” τη διεργασία εντός στατιστικού ελέγχου με τη χρησιμοποίηση δοκιμαστικών ορίων

ελέγχου. Όταν αυτό επιτευχθεί τα διαγράμματα ελέγχου που προκύπτουν (κεντρική γραμμή και όρια ελέγχου) είναι κατάλληλα για την παρακολούθηση της μελλοντικής συμπεριφοράς της διεργασίας (Φάση II). Αυτή η χρήση των διαγραμμάτων ελέγχου αναφέρεται και ως αναδρομική (retrospective). Γενικά πολύ περισσότερες δράσεις και σκέψεις εξελίσσονται κατά την διάρκεια αυτής της φάσης πλην της απλής διαγραμματοποίησης κάποιων δεδομένων. Ο διαχειριστής της διαδικασίας μελετά σε βάθος τη διεργασία και αυτό γιατί δεν είναι εύκολο να αποφασίσει αν η διεργασία ήταν εντός ή εκτός στατιστικού ελέγχου κατά τη χρονική περίοδο που συλλέχθηκαν τα δεδομένα.

Φάση II (Phase II): Σε αυτή τη φάση τα διαγράμματα ελέγχου χρησιμοποιούνται προκειμένου να ελέγχουμε συνεχώς αν η διαδικασία παραμένει εντός ελέγχου. Στη φάση αυτή ο διαχειριστής έχει στα χέρια του ένα πολύτιμο εργαλείο μέσω του οποίου είναι δυνατόν να παρακολουθεί συνεχώς την παραγωγική διεργασία και να ανιχνεύει εγκαίρως μια πιθανή αλλαγή στο μέσο επίπεδο των χαρακτηριστικών που καθορίζουν την ποιότητα του παραγόμενου προϊόντος. Δηλαδή σε κάθε χρονική περίοδο που ένα δείγμα λαμβάνεται από την διεργασία ο διαχειριστής δίνει άμεσα μια απάντηση στο ερώτημα “παραμένει η διεργασία εντός ελέγχου;”. Σε αυτήν τη φάση ο διαχειριστής αδιαφορεί για τον τρόπο με τον οποίο το μέσο επίπεδο της διεργασίας είχε εκτιμηθεί, ή αν αυτό ήταν εκ των προτέρων γνωστό.

1.3 Εκτίμηση της μέσης τιμής και της διασποράς στα διαγράμματα ελέγχου Shewhart για μεταβλητές

1.3.1 Η περίπτωση των μεμονωμένων παρατηρήσεων

Έστω $\mathbf{X} = (X_1, X_2, \dots, X_n)$ ένα τυχαίο δείγμα μεγέθους n από πληθυσμό με μέση τιμή μ και διακύμανση σ^2 . Ένας αμερόληπτος εκτιμητής της μέσης τιμής μ είναι ο δειγματικός μέσος

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

και ένας αμερόληπτος εκτιμητής της διακύμανσης σ^2 είναι η δειγματική διακύμανση

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

Ας υποθέσουμε τώρα ότι το τυχαίο δείγμα (X_1, X_2, \dots, X_n) προέρχεται από κανονικό πληθυσμό $N(\mu, \sigma^2)$. Τότε μπορούμε να βρούμε αμερόληπτους εκτιμητές της τυπικής απόκλισης σ που είναι οι ακόλουθοι

$$\sigma = \frac{R}{d_2(n)} = \frac{X_{(n)} - X_{(1)}}{d_2(n)} = \frac{\max\{X_1, X_2, \dots, X_n\} - \min\{X_1, X_2, \dots, X_n\}}{d_2(n)}$$

και

$$\sigma = \frac{S}{c_4(n)} = \frac{\sqrt{S^2}}{c_4(n)}$$

όπου η ποσότητα $d_2(n)$ είναι σταθερά που εξαρτάται μόνο από το μέγεθος του δείγματος n και ορίζεται ως ακολούθως

$$d_2 = d_2(n) = E(Z_{(n)} - Z_{(1)})$$

με $Z_{(n)}$ και $Z_{(1)}$ να είναι η μέγιστη και η ελάχιστη τιμή τυχαίου δείγματος (Z_1, \dots, Z_n) από την τυπική κανονική κατανομή $N(0,1)$.

Επίσης η ποσότητα $c_4(n)$ είναι σταθερά που εξαρτάται από το μέγεθος του δείγματος n και ορίζεται από την ακόλουθη σχέση

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

1.3.2 Η περίπτωση των δειγμάτων

Έστω ότι έχουμε στην διάθεσή μας m ανεξάρτητα τυχαία δείγματα μεγέθους n το καθένα, τα $X_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $1 \leq i \leq m$, από ένα κανονικό πληθυσμό $X \sim N(\mu, \sigma^2)$ με μέση τιμή μ και διακύμανση σ^2 .

- **Εκτίμηση του μ**

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

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

Η ποσότητα $\bar{\bar{X}}$ ακολουθεί κατανομή $X \sim N\left(\mu, \frac{\sigma^2}{mn}\right)$ και χρησιμοποιείται ως εκτίμηση της ποσότητας μ (αμερόληπτος και συνεπής εκτιμητής του μ), δηλαδή

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

- **Εκτίμηση του σ (Μέθοδος R)**

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

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

και

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

Τότε

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

Η ποσότητα $\sigma = \frac{\bar{R}}{d_2}$ είναι αμερόληπτος εκτιμητής της ποσότητας σ , και χρησιμοποιείται ως εκτίμησή της.

- **Εκτίμηση του σ (Μέθοδος S)**

Έστω S_i η ποσότητα που ορίζεται από την σχέση

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

και

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

Τότε

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

Η ποσότητα $\sigma = \frac{\bar{S}}{c_4}$ είναι αμερόληπτος εκτιμητής της ποσότητας σ και χρησιμοποιείται ως εκτίμησή της.

- **Εκτίμηση του σ (Μέθοδος S^2)**

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

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

Η ποσότητα $\sqrt{S^2}$ όπου

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

αν και δεν είναι αμερόληπτος εκτιμητής του σ , χρησιμοποιείται αρκετές φορές (λόγω του ότι έχει μικρότερη διακύμανση από την εκτίμηση $\sigma = \frac{\bar{S}}{c_4}$) ως εκτίμηση της ποσότητας σ , δηλαδή

$$\sigma = \sqrt{S^2}.$$

1.4 Διαγράμματα ελέγχου για τη μέση τιμή

Έστω ότι η κατανομή του χαρακτηριστικού X των προϊόντων που παράγονται ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ με μ και σ γνωστά. Έστω επίσης ότι έχουμε στην διάθεση μας ανεξάρτητα τυχαία δείγματα μεγέθους n το καθένα, τα $\mathbf{X}_i = (X_{i_1}, X_{i_2}, \dots, X_{i_n})$, $i \geq 1$. Για κάθε τυχαίο δείγμα $\mathbf{X}_i = (X_{i_1}, X_{i_2}, \dots, X_{i_n})$ υπολογίζουμε το δειγματικό μέσο

$$W_i = \bar{X}_i = \frac{X_{i_1} + X_{i_2} + \dots + X_{i_n}}{n}$$

ο οποίος είναι αμερόληπτος εκτιμητής της μέσης τιμής μ του χαρακτηριστικού X και ακολουθεί την κανονική κατανομή $N(\mu, \frac{\sigma^2}{n})$. Στο \bar{X} διάγραμμα ελέγχου Φάσης II απεικονίζονται οι τιμές W_i . Τα 3σ όρια ελέγχου και η κεντρική γραμμή του διαγράμματος ελέγχου για την παρακολούθηση της μέσης τιμής μ του χαρακτηριστικού X , δίνονται στο ακόλουθο πλαίσιο ($Xbar$ ή \bar{X} διάγραμμα ελέγχου)

\bar{X} διάγραμμα ελέγχου με όρια ελέγχου 3σ	
UCL	$= \mu + \frac{3}{\sqrt{n}}\sigma$
$Center Line$	$= \mu$
LCL	$= \mu - \frac{3}{\sqrt{n}}\sigma$

Τα παραπάνω όρια ελέγχου του διαγράμματος είναι γνωστά ως όρια ελέγχου Φάσης II.

Έστω ότι έχουμε στη διάθεσή μας m δείγματα μεγέθους n τα $X_i = (X_{i_1}, X_{i_2}, \dots, X_{i_n})$, $i = 1, 2, \dots, m$. Στην Φάση I όπου οι παράμετροι μ και σ είναι άγνωστες, για τον υπολογισμό των ορίων ελέγχου και της κεντρικής γραμμής του διαγράμματος χρησιμοποιούμε ως εκτίμηση του μ την ποσότητα $\mu = \bar{\bar{X}}$ και ως εκτίμηση του σ την ποσότητα $\sigma = \frac{\bar{R}}{d_2}$. Τότε το \bar{X} διάγραμμα ελέγχου για την μέση τιμή με όρια ελέγχου 3σ της Φάσης I παίρνει την ακόλουθη μορφή:

**\bar{X} διάγραμμα ελέγχου με όρια
ελέγχου 3σ Φάσης I
Μέθοδος R**

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

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

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

όπου η σταθερά A_2 είναι ίση με

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

Χρησιμοποιώντας ως εκτίμηση του μ την ποσότητα $\mu = \bar{\bar{X}}$ και ως εκτίμηση του σ την ποσότητα $\sigma = \frac{\bar{S}}{c_4}$ τότε το \bar{X} διάγραμμα ελέγχου για την μέση τιμή με όρια ελέγχου 3σ της Φάσης

I παίρνει την ακόλουθη μορφή:

**\bar{X} διάγραμμα ελέγχου με όρια
ελέγχου 3σ Φάσης I
Μέθοδος S**

$$UCL = \bar{\bar{X}} + A_3 \bar{S}$$

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

$$LCL = \bar{\bar{X}} - A_3 \bar{S}$$

όπου η σταθερά A_3 είναι ίση με

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

Χρησιμοποιώντας ως εκτίμηση του μ την ποσότητα $\mu = \bar{\bar{X}}$ και ως εκτίμηση του σ την ποσότητα $\sqrt{S^2}$ τότε το \bar{X} διάγραμμα ελέγχου για την μέση τιμή με όρια ελέγχου 3σ της Φάσης II παίρνει την ακόλουθη μορφή:

\bar{X} διάγραμμα ελέγχου με όρια ελέγχου 3σ Φάσης I Μέθοδος S^2	
UCL	$= \bar{\bar{X}} + \frac{3}{\sqrt{n}} \sqrt{S^2}$
$Center\ Line$	$= \bar{\bar{X}}$
LCL	$= \bar{\bar{X}} - \frac{3}{\sqrt{n}} \sqrt{S^2}$

1.5 Διαγράμματα ελέγχου για τη διασπορά

Στην παρούσα παράγραφο θα αναπτύξουμε την κατασκευή διαγραμμάτων ελέγχου για την παρακολούθηση της συμπεριφοράς της διασποράς ενός συνεχούς ποιοτικού χαρακτηριστικού X . Υποθέτουμε ότι το χαρακτηριστικό X ακολουθεί κατανομή $N(\mu, \sigma^2)$ με σ γνωστό. Αν $X_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $i \geq 1$, είναι ένα τυχαίο δείγμα μεγέθους n από την X , τότε για τη στατιστική συνάρτηση

$$W_i = R_i = X_{i(n)} - X_{i(1)} = \max \{X_{i1}, X_{i2}, \dots, X_{in}\} - \min \{X_{i1}, X_{i2}, \dots, X_{in}\}$$

είναι γνωστό ότι

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

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

Συνεπώς ένα διάγραμμα ελέγχου για την παρακολούθηση της διασποράς του ποιοτικού χαρακτηριστικού X μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι το εύρος R_i των δειγμάτων που ως γνωστό είναι ένα μέτρο διασποράς της X (το διάγραμμα ονομάζεται R διάγραμμα ελέγχου). Το μοντέλο με όρια ελέγχου 3σ θα έχει τη μορφή

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

Θέτοντας

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

προκύπτει το ακόλουθο πλαίσιο

**R διάγραμμα ελέγχου με όρια
ελέγχου 3σ Φάσης II**

$$UCL = D_2\sigma$$

$$Center Line = d_2\sigma$$

$$LCL = D_1\sigma$$

Όμως για την παρακολούθηση της διασποράς του χαρακτηριστικού X μπορούμε επίσης να απεικονίσουμε στο διάγραμμα τη στατιστική συνάρτηση

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

για την οποία ισχύει ότι

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

Συνεπώς ένα διάγραμμα ελέγχου για την παρακολούθηση της διασποράς του χαρακτηριστικού X μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι οι δειγματικές τυπικές αποκλίσεις S_i που είναι το πιο συνηθισμένο μέτρο διασποράς της X (το διάγραμμα ονομάζεται S διάγραμμα ελέγχου). Το μοντέλο με όρια 3σ θα έχει τη μορφή

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

$$CL = \mu_{S_i} = c_4\sigma$$

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

Θέτοντας

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

προκύπτει το ακόλουθο πλαίσιο

**S διάγραμμα ελέγχου με όρια
ελέγχου 3σ Φάσης II**

$$UCL = B_6\sigma$$

$$Center\ Line = c_4\sigma$$

$$LCL = B_5\sigma$$

Μπορούμε εύκολα να αναπτύξουμε διάγραμμα ελέγχου με όρια πιθανότητας α χρησιμοποιώντας τη σχέση

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

ή ισοδύναμα τη σχέση

$$\left(\chi_{n-1;1-\alpha/2}^2 = (\chi_{n-1;1-\alpha/2})^2\right)$$

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

$$\left(UCL = \sigma\sqrt{\frac{\chi_{n-1;\alpha/2}^2}{n-1}} , LCL = \sigma\sqrt{\frac{\chi_{n-1;1-\alpha/2}^2}{n-1}} \right)$$

Τέλος για την παρακολούθηση της διασποράς του χαρακτηριστικού X μπορούμε να χρησιμοποιήσουμε επίσης τη στατιστική συνάρτηση

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

για την οποία ισχύει ότι

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

Έτσι

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

και

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

Συνεπώς ένα διάγραμμα ελέγχου για την παρακολούθηση της διασποράς του χαρακτηριστικού X μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι η δειγματική διακύμανση S_i^2 που είναι μέτρο διασποράς της X (το διάγραμμα ονομάζεται S^2 διάγραμμα ελέγχου). Το S^2 διάγραμμα ελέγχου περιγράφεται στο ακόλουθο πλαίσιο

**S^2 διάγραμμα ελέγχου με όρια ελέγχου
Πιθανότητας α Φάσης II**

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

$$\text{Center Line} = \sigma^2$$

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

Για την ανάπτυξη διαγραμμάτων ελέγχου Φάσης I για την παρακολούθηση της διασποράς της διεργασίας (σ άγνωστο) επιλέγονται m ανεξάρτητα προκαταρκτικά τυχαία δείγματα μεγέθους n το καθένα $X_i = (X_{i1}, X_{i2}, \dots, X_{in})$, $1 \leq i \leq m$. Τα προκαταρκτικά δείγματα είναι συνήθως 20 έως 25 μεγέθους 4 έως 6 και υποθέτουμε ότι το χαρακτηριστικό X ακολουθεί κατανομή $N(\mu, \sigma^2)$.

Χρησιμοποιώντας ως εκτίμηση του σ την ποσότητα $\sigma = \frac{\bar{R}}{d_2}$ τα όρια ελέγχου και η κεντρική γραμμή του R διαγράμματος ελέγχου Φάσης I με όρια 3σ δίνονται από τις σχέσεις

$$LCL = \mu_{R_i} - 3\sigma_{R_i} = (d_2 - 3d_3)\sigma = \frac{(d_2 - 3d_3)\bar{R}}{d_2} = \left(1 - 3\frac{d_3}{d_2}\right)\bar{R}$$

$$CL = \mu_{R_i} = d_2\sigma = d_2 \frac{\bar{R}}{d_2} = \bar{R}$$

$$UCL = \mu_{R_i} + 3\sigma_{R_i} = (d_2 + 3d_3)\sigma = \frac{(d_2 + 3d_3)\bar{R}}{d_2} = \left(1 + 3\frac{d_3}{d_2}\right)\bar{R}$$

Έτσι το προηγούμενο R διάγραμμα ελέγχου Φάσης II παίρνει την ακόλουθη μορφή:

<p>R διάγραμμα ελέγχου με όρια ελέγχου 3σ Φάσης I</p>
$UCL = D_4 \bar{R}$ $Center\ Line = \bar{R}$ $LCL = D_3 \bar{R}$

όπου

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

Χρησιμοποιώντας ως εκτίμηση του σ την ποσότητα $\sigma = \frac{\bar{S}}{c_4}$ τα όρια ελέγχου και η κεντρική γραμμή του S διαγράμματος ελέγχου Φάσης I με όρια 3σ δίνονται από τις σχέσεις

$$LCL = \mu_{s_i} - 3\sigma_{s_i} = (c_4 - 3\sqrt{1-c_4^2})\sigma = (c_4 - 3\sqrt{1-c_4^2})\frac{\bar{S}}{c_4} = \left(1 - \frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{S}$$

$$CL = \mu_{s_i} = c_4\sigma = c_4\frac{\bar{S}}{c_4} = \bar{S}$$

$$UCL = \mu_{s_i} + 3\sigma_{s_i} = (c_4 + 3\sqrt{1-c_4^2})\sigma = (c_4 + 3\sqrt{1-c_4^2})\frac{\bar{S}}{c_4} = \left(1 + \frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{S}$$

Έτσι το προηγούμενο S διάγραμμα ελέγχου Φάσης II παίρνει την ακόλουθη μορφή:

**S διάγραμμα ελέγχου με όρια
ελέγχου 3σ Φάσης I**

$$\begin{aligned} UCL &= B_4 \bar{S} \\ \text{Center Line} &= \bar{S} \\ LCL &= B_3 \bar{S} \end{aligned}$$

όπου

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

Χρησιμοποιώντας ως εκτίμηση του σ^2 την ποσότητα $\sigma^2 = \bar{S}^2$, είναι εύκολο να διαπιστωθεί ότι τα όρια ελέγχου και η κεντρική γραμμή του S^2 διαγράμματος ελέγχου με όρια πιθανότητας α δίνονται στο ακόλουθο πλαίσιο

**S^2 διάγραμμα ελέγχου με όρια
ελέγχου Πιθανότητας α Φάσης I**

$$\begin{aligned} UCL &= \frac{\bar{S}^2}{n-1} \chi_{n-1; \alpha/2}^2 \\ \text{Center Line} &= \bar{S}^2 \\ LCL &= \frac{\bar{S}^2}{n-1} \chi_{n-1; 1-\alpha/2}^2 \end{aligned}$$

1.6 Διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις

Σε ορισμένες περιπτώσεις το μέγεθος του δείγματος είναι ίσο με 1 (αυτόματη επιθεώρηση παραγόμενων προϊόντων, μικρός ρυθμός παραγωγής, κτλ.). Σε αυτή την περίπτωση χρησιμοποιούμε διαγράμματα ελέγχου για μεμονωμένες ή ατομικές παρατηρήσεις (individual observations). Ας υποθέσουμε ότι η κατανομή του χαρακτηριστικού X των προϊόντων που παράγονται ακολουθεί κανονική κατανομή $N(\mu, \sigma^2)$ με μ και σ γνωστά και ότι έχουμε στη διάθεσή μας ανεξάρτητες παρατηρήσεις $X_i, i \geq 1$ του χαρακτηριστικού X . Για την παρακολούθηση της μέσης τιμής μ κατασκευάζουμε ένα διάγραμμα ελέγχου στο οποίο απεικονίζονται οι μεμονωμένες παρατηρήσεις X_i ($W_i = X_i, i \geq 1$). Το διάγραμμα αυτό ονομάζεται X διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις, και τα όρια ελέγχου του παρουσιάζονται στο ακόλουθο πλαίσιο:

<p style="text-align: center;">X διάγραμμα ελέγχου με όρια ελέγχου</p> <p style="text-align: center;">3σ Φάσης II</p>
<p style="text-align: center;">$UCL = \mu + 3\sigma$</p> <p style="text-align: center;">$Center\ Line = \mu$</p> <p style="text-align: center;">$LCL = \mu - 3\sigma$</p>

Για τη συμπεριφορά της διασποράς του χαρακτηριστικού X χρησιμοποιούμε το κινούμενο εύρος (*moving range*) των μεμονωμένων παρατηρήσεων που ορίζεται από τη σχέση:

$$MR_i = |X_i - X_{i-1}| = \max\{X_{i-1}, X_i\} - \min\{X_{i-1}, X_i\} \quad i \geq 2.$$

Για το κινούμενο εύρος MR_i , που είναι το εύρος δύο παρατηρήσεων, ισχύει ότι

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

Συνεπώς ένα διάγραμμα ελέγχου για τη διασπορά του χαρακτηριστικού X μπορεί να βασιστεί σε ένα διάγραμμα όπου η απεικονιζόμενη ποσότητα θα είναι το κινούμενο εύρος MR_i των παρατηρήσεων. Το διάγραμμα ελέγχου με όρια 3σ θα έχει όρια ελέγχου και κεντρική γραμμή που δίνονται από τις σχέσεις:

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

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

MR διάγραμμα ελέγχου με όρια ελέγχου 3σ Φάσης II	
	$UCL = D_2\sigma$
	$Center\ Line = d_2\sigma$
	$LCL = D_1\sigma$

Στην περίπτωση που οι ποσότητες μ και σ είναι άγνωστες πρέπει να εκτιμηθούν. Αν έχουμε στην διάθεσή μας m ανεξάρτητες παρατηρήσεις X_1, X_2, \dots, X_m από το χαρακτηριστικό X η εκτίμηση του μέσου δίνεται από την σχέση:

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

Για την εκτίμηση του σ , θέτοντας

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

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

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

Επομένως το X διάγραμμα ελέγχου της Φάσης I για μεμονωμένες παρατηρήσεις είναι το ακόλουθο:

X διάγραμμα ελέγχου με όρια ελέγχου 3σ Φάσης I
$UCL = \bar{X} + 3 \frac{\overline{MR}}{d_2}$ $Center\ Line = \bar{X}$ $LCL = \bar{X} - 3 \frac{\overline{MR}}{d_2}$

ενώ το MR διάγραμμα ελέγχου Φάσης I είναι το ακόλουθο:

MR διάγραμμα ελέγχου με όρια ελέγχου 3σ Φάσης I
$UCL = D_4 \overline{MR}$ $Center\ Line = \overline{MR}$ $LCL = D_3 \overline{MR}$

όπου D_3 και D_4 είναι σταθερές που υπολογίζονται για $n = 2$.

1.7 Μέγεθος δείγματος και συχνότητα δειγματοληψίας

Κατά τον σχεδιασμό ενός διαγράμματος ελέγχου πρέπει να καθοριστεί το μέγεθος των δειγμάτων των προϊόντων και η συχνότητα δειγματοληψίας. Γενικά τα μεγάλα μεγέθη δειγμάτων κάνουν πιο έγκαιρη ανίχνευση μικρών μετατοπίσεων του μέσου επιπέδου της διεργασίας. Πράγματι χρησιμοποιώντας 3σ όρια ελέγχου, υποθέτοντας ότι η διεργασία είναι εντός ελέγχου και η κατανομή της απεικονιζόμενης στατιστικής συνάρτησης είναι κανονική, τότε ανεξαρτήτως του μεγέθους του δείγματος, το διάγραμμα ελέγχου δίνει (εσφαλμένη) ένδειξη ότι η διεργασία είναι εκτός ελέγχου (σφάλμα τύπου I) με πιθανότητα $\alpha = 0.0027$.

Πράγματι, για το \bar{X} διάγραμμα ελέγχου Φάσης II έχουμε ότι

$$\alpha = P\left(\bar{X} < \mu - 3\frac{\sigma}{\sqrt{n}} \text{ ή } \bar{X} > \mu + 3\frac{\sigma}{\sqrt{n}} \mid X \square N(\mu, \sigma^2)\right) = 2(1 - \Phi(3)) = 0.0027$$

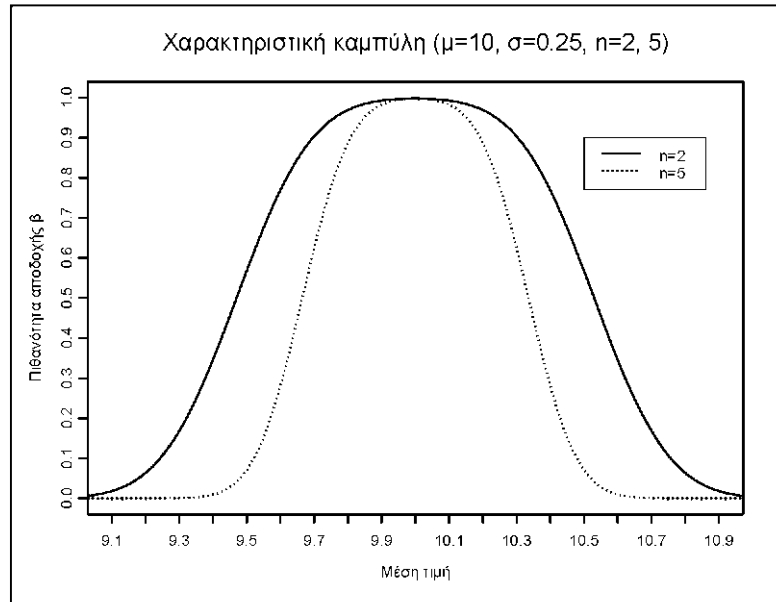
και η πιθανότητα α δεν εξαρτάται από το μέγεθος του δείγματος n .

Αντίστοιχα με την έννοια του σφάλματος τύπου II, έχουμε ότι η πιθανότητα β (ή ρίσκο β) να μας δώσει το διάγραμμα ένδειξη (εσφαλμένη) ότι η διαδικασία είναι εντός ελέγχου ενώ στην πραγματικότητα είναι εκτός ελέγχου λόγω μετατόπισης του μέσου επιπέδου της διεργασίας από μ σε μ^* είναι ίση με

$$\beta = P\left(\mu - 3\frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + 3\frac{\sigma}{\sqrt{n}} \mid X \square N(\mu^*, \sigma^2)\right) = \Phi\left(\sqrt{n}\frac{\mu - \mu^*}{\sigma} + 3\right) - \Phi\left(\sqrt{n}\frac{\mu - \mu^*}{\sigma} - 3\right)$$

Η πιθανότητα β εξαρτάται από το μ^* και το μέγεθος του δείγματος n , δηλαδή $\beta = \beta(\mu^*, n)$. Επιπρόσθετα όσο αυξάνεται το μέγεθος του δείγματος η πιθανότητα β μειώνεται και συνεπώς αυξάνεται η ικανότητα του διαγράμματος ελέγχου να ανιχνεύει τις μετατοπίσεις του μέσου επιπέδου της διεργασίας. Στο ακόλουθο σχήμα δίνεται η χαρακτηριστική καμπύλη (operating

curve), δηλαδή η γραφική παράσταση της συνάρτησης $\beta = \beta(\mu^*, n)$ για $n=2$ και $n=5$ και $X \sim N(10, 0.25^2)$.



Διάγραμμα 1.2 Χαρακτηριστικές καμπύλες για $W \sim N(10, 0.25^2 / n)$, $n = 2, 5$.

Έτσι αν το μέσο επίπεδο της διεργασίας από $\mu = \mu_w = 10$ μετατοπιστεί στη θέση $\mu^* = 10.5$ τότε για μέγεθος δείγματος ίσο με 2 η πιθανότητα να δώσει ένδειξη το διάγραμμα ελέγχου ότι η διεργασία είναι εντός ελέγχου ενώ στην πραγματικότητα είναι εκτός ελέγχου είναι ίση με 0.568113, ενώ για μέγεθος δείγματος ίσο με 5 η πιθανότητα είναι ίση με 0.0704921 που είναι αρκετά ικανοποιητική

$$n = 2: \quad \mu = 10 \rightarrow \mu^* = 10.3, \quad \beta = \Phi(1.30294) - \Phi(-4.69706) = 0.903702$$

$$n = 5: \quad \mu = 10 \rightarrow \mu^* = 10.3, \quad \beta = \Phi(0.376118) - \Phi(-5.68328) = 0.624271$$

Επομένως για να ανιχνεύει το διάγραμμα ελέγχου σε ικανοποιητικό βαθμό (μικρή πιθανότητα για το σφάλμα τύπου II) μικρές μετατοπίσεις του μέσου επιπέδου της διεργασίας κρίνεται αναγκαία η αύξηση του μεγέθους του δείγματος. Συνεπώς το κατάλληλο μέγεθος του δείγματος εξαρτάται σε μεγάλο βαθμό από το είδος της μετατόπισης (μικρής ή μεγάλης) του μέσου επιπέδου της διεργασίας που θέλουμε να ανιχνεύσουμε.

Σχετικά με τη συχνότητα της δειγματοληψίας η επιθυμητή κατάσταση θα ήταν να παίρνουμε δείγματα πολύ συχνά και προφανώς η βέλτιστη κατάσταση θα ήταν να ελέγχουμε όλα τα παραγόμενα προϊόντα. Όμως αυτές οι πολιτικές δειγματοληψίας συνεπάγονται μεγάλο οικονομικό κόστος.

Γενικά η πολιτική που ακολουθείται σε τέτοιου είδους προβλήματα είναι μικρά μεγέθη δειγμάτων αρκετά συχνά (ιδιαίτερα σε περιπτώσεις μαζικής παραγωγής προϊόντων ή όταν εμφανίζονται συχνά ειδικές αιτίες μεταβλητότητας) παρά μεγάλα μεγέθη δειγμάτων λιγότερο συχνά.

1.8 Το μέσο μήκος ροής ως μέτρο απόδοσης των διαγραμμάτων ελέγχου

Το μήκος ροής (Run Length) ενός διαγράμματος ελέγχου είναι μια τυχαία μεταβλητή, έστω T , που παριστάνει τον αριθμό των σημείων του διαγράμματος μέχρι να εμφανιστεί σήμα εκτός ελέγχου διεργασίας.

Αν τα σημεία του διαγράμματος είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές και τα όρια ελέγχου είναι γνωστές σταθερές, όπως στην περίπτωση των διαγραμμάτων τύπου Shewhart Φάσης II, τότε το μήκος ροής, δηλαδή η τ.μ. T , ακολουθεί τη γεωμετρική κατανομή με παράμετρο p , που παριστάνει την πιθανότητα ένα σημείο του διαγράμματος να βρεθεί εκτός των ορίων ελέγχου.

Επομένως η τυχαία μεταβλητή T ακολουθεί την γεωμετρική κατανομή, δηλαδή $T \sim G(p)$, και τότε

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

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

Το μέσο μήκος ροής (ARL: Average Run Length) είναι το πιο διαδεδομένο μέτρο απόδοσης των διαγραμμάτων ελέγχου κυρίως λόγω της ελκυστικά διαισθητικής ερμηνείας του. Γι' αυτόν τον

λόγο χρησιμοποιείται συχνά κατά τον σχεδιασμό και την σύγκριση της απόδοσης των διαγραμμάτων ελέγχου. Το μέσο μήκος ροής είναι ο αναμενόμενος αριθμός των σημείων που πρέπει να σχεδιαστούν στο διάγραμμα μέχρι να εμφανιστεί ένα σημείο εκτός των ορίων ελέγχου, δηλαδή $ARL = E(T)$ και φυσικά δίνεται από τη σχέση

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

(όπου p είναι η πιθανότητα ένα σημείο να βρεθεί εκτός των ορίων ελέγχου).

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

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

Το μέσο μήκος ροής μιας διεργασίας που βρίσκεται εντός στατιστικού ελέγχου συμβολίζεται με ARL_{in} (in-control average run length) και δίνεται από τη σχέση

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

Προφανώς στην πράξη επιθυμούμε να έχουμε μεγάλο ARL_{in} , προκειμένου να μειωθεί ο αριθμός των λανθασμένων ενδείξεων εκτός ελέγχου διεργασίας.

Ένα τυπικό \bar{X} διάγραμμα ελέγχου τύπου Shewhart με 3-σίγμα όρια ελέγχου που βασίζεται σε γνωστές παραμέτρους έχει εντός ελέγχου μέσο μήκος ροής 370, δηλαδή $ARL = 370$ υπό την υπόθεση ότι τα δεδομένα ακολουθούν την κανονική κατανομή.

Δηλαδή όταν η απόσταση των ορίων ελέγχου είναι $L=3$, τότε

$$ARL_{in} = \frac{1}{0.0027} = 370.4.$$

Επομένως, προκύπτει ότι μπορούμε να αναμένουμε σήμα εκτός ελέγχου διεργασίας, κατά μέσο όρο, μια φορά σε κάθε 370 σημεία του διαγράμματος.

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

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

Το μέσο μήκος ροής μια διεργασίας που βρίσκεται εκτός στατιστικού ελέγχου ορίζεται ως ARL_{out} (out of control average length) και δίνεται από τη σχέση

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

Στην πράξη επιθυμούμε να έχουμε μικρή τιμή του ARL_{out} καθώς αυτό το μέτρο δηλώνει τον αριθμό των δειγμάτων που πρέπει να ληφθούν για να εντοπιστεί η αλλαγή στο μέσο επίπεδο διεργασίας από την στιγμή που η μετατόπιση αυτή συνέβη.

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

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

αφού

$$p_{in} = 1 - P(LCL \leq W \leq UCL | W \square N(\mu, \sigma^2))$$

$$= 1 - P(\mu - L\sigma \leq W \leq \mu + L\sigma | W \square N(\mu, \sigma^2))$$

$$= 1 - \Phi(L) + \Phi(-L) = 2\Phi(-L)$$

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

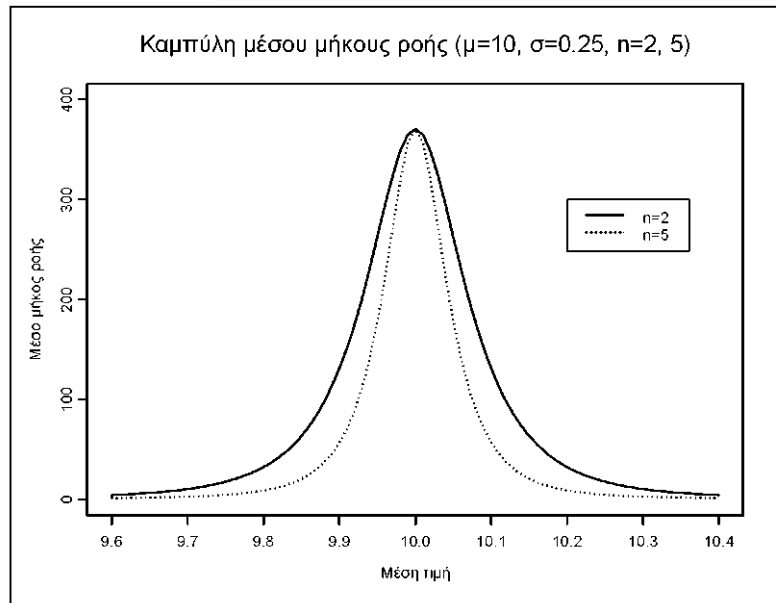
$$\begin{aligned}
 p_{out} &= 1 - P(LCL \leq W \leq UCL | W \sim N(\mu + \delta\sigma, \sigma^2)) \\
 &= 1 - P(\mu - L\sigma \leq W \leq \mu + L\sigma | W \sim N(\mu + \delta\sigma, \sigma^2)) \\
 &= 1 - \Phi(L - \delta) + \Phi(-L - \delta) = 1 - \Phi(L - \delta) - \Phi(-L - \delta)
 \end{aligned}$$

και

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

Επίσης με ανάλογο τρόπο γίνεται ο υπολογισμός του ARL στα μονόπλευρα διαγράμματα ελέγχου.

Για παράδειγμα, στο ακόλουθο διάγραμμα δίνεται η γραφική παράσταση του μέσου μήκους ροής ARL_{out} για $X \sim N(10, 0.25^2)$, $W = \bar{X}$, $n = 2$ και $n = 5$.



Διάγραμμα 1.3 Γραφική παράσταση του μέσου μήκους ροής για $W \square N(10, (0.25^2 / n))$, $n = 2, 5$

Η χρήση του ARL ως μέτρου για την περιγραφή της απόδοσης μιας διεργασίας έχει υποστεί αρκετή κριτική τα τελευταία χρόνια γιατί το ARL που παρατηρείται στην πράξη διαφέρει συνήθως αρκετά από το “θεωρητικό” ARL (είτε είναι αρκετά μεγαλύτερο είτε είναι αρκετά μικρότερο). Αυτό συμβαίνει επειδή η κατανομή του μήκους ροής T είναι πολύ ασυμμετρική (ως γεωμετρική) και επομένως η μέση τιμή της δεν μπορεί να θεωρηθεί ως αντιπροσωπευτικό μέτρο κεντρικής τάσης, αφού η διακύμανσή της γίνεται αρκετά μεγάλη όσο το p μικραίνει. Γι’ αυτόν τον λόγο είναι απαραίτητο η μέση τιμή της T να συνοδεύεται από την διακύμανση της T ή ακόμα και τα ποσοστιαία σημεία της.

ΚΕΦΑΛΑΙΟ 2

Επίδραση της εκτίμησης των παραμέτρων στα διαγράμματα ελέγχου για την μέση τιμή

2.1 Εισαγωγή

Οι περισσότερες μελέτες που έχουν πραγματοποιηθεί και περιλαμβάνουν την ανάπτυξη και τη χρήση των διαγραμμάτων ελέγχου στη Φάση II βασίζονται στην υπόθεση ότι οι εντός στατιστικού ελέγχου παράμετροι της διεργασίας είναι γνωστές. Για παράδειγμα, όταν ένα ποιοτικό χαρακτηριστικό ακολουθεί την κανονική κατανομή, τότε οι παράμετροι της διεργασίας που μας ενδιαφέρουν είναι η εντός ελέγχου μέση τιμή, μ , και η εντός ελέγχου τυπική απόκλιση, σ . Η υπόθεση ότι οι εντός ελέγχου τιμές των παραμέτρων είναι γνωστές απλοποιεί την ανάπτυξη και κατά συνέπεια τους υπολογισμούς διαφόρων ποσοτήτων που αφορούν τα διαγράμματα ελέγχου.

Στην πράξη όμως οι παράμετροι είναι σπάνια γνωστές και τα όρια ελέγχου συνήθως βασίζονται στις εκτιμώμενες παραμέτρους. Όταν οι εκτιμήσεις χρησιμοποιούνται στη θέση των γνωστών παραμέτρων, η μεταβλητότητα των εκτιμητών μπορεί να επηρεάσει την απόδοση των διαγραμμάτων και αυτό έχει σαν συνέπεια η απόδοση αυτή να διαφέρει από την απόδοση των διαγραμμάτων που σχεδιάστηκαν με γνωστές παραμέτρους.

Επιπρόσθετα, η ενδεχόμενη αδυναμία χρησιμοποίησης περιορισμένου αριθμού δεδομένων ή ακόμα και η χρησιμοποίηση πληροφοριών από τα δείγματα που δεν είναι καθόλου αντιπροσωπευτικές για την διεργασία για να καθοριστούν τα όρια ελέγχου έχει ήδη επισημανθεί από πολλούς ερευνητές. Ο Shewhart (1939) επεσήμανε ότι: “Στην πλειονότητα των πρακτικών περιπτώσεων, το πιο δύσκολο κομμάτι από όλα είναι να επιλέξουμε το δείγμα που πρόκειται να χρησιμοποιηθεί ως η βάση για τον προσδιορισμό των ορίων ελέγχου”. Και αυτό διότι, συγκεντρώνοντας ένα αντιπροσωπευτικό δείγμα επαρκούς μεγέθους, τότε και μόνο τότε θα είμαστε αρκετά σίγουροι για την ορθότητα των ορίων ελέγχου.

Ο εμπειρικός κανόνας που χρησιμοποιείται κυρίως για τα \bar{X} διαγράμματα ελέγχου είναι να λαμβάνονται $m=20$ έως $m=30$ προκαταρκτικά δείγματα μεγέθους $n=4$ έως $n=6$ στην Φάση I (Montgomery, 2005). Οι κανόνες όμως αυτοί βασίζονται κυρίως σε εμπειρικά δεδομένα και είναι συνήθως ανεπαρκείς. Η ανεπάρκεια των κανόνων αυτών για την επιλογή του αριθμού m και του

μεγέθους n των δειγμάτων οδήγησε στο συμπέρασμα ότι η επίδραση της εκτίμησης των παραμέτρων πρέπει να λαμβάνεται σοβαρά υπ' όψιν κατα τον σχεδιασμό ενός διαγράμματος ελέγχου.

2.2 Η κατανομή του μήκους ροής κατά την εκτίμηση των παραμέτρων των διαγραμμάτων ελέγχου

Στην περίπτωση που η κατανομή του μήκους ροής είναι γεωμετρική, τότε το μέσο μήκος ροής ARL (και συνεπώς η πιθανότητα ένα σημείο να βρεθεί εκτός των ορίων ελέγχου) χαρακτηρίζει πλήρως την κατανομή του μήκους ροής και είναι το αποδεκτό μέτρο απόδοσης.

Για την απλοποίηση συγκρίσεων αλλά και για μια καλύτερη εικόνα της κατανομής του μήκους ροής, συνήθως χρησιμοποιούνται συνοπτικές τιμές της κατανομής του μήκους ροής με σκοπό την αποτελεσματική εκτίμηση της απόδοσης των διαγραμμάτων ελέγχου με εκτιμώμενες παραμέτρους.

Για παράδειγμα χρησιμοποιούνται συχνά το μέσο μήκος ροής (ARL), η τυπική απόκλιση του μήκους ροής (SD ή $SDRL$) αλλά και τα ποσοστημόρια της κατανομής του μήκους ροής.

Επομένως, αν η τ.μ. T που παριστάνει το μήκος ροής ακολουθεί τη γεωμετρική κατανομή με παράμετρο $p = 0.0027$ (υπό την προϋπόθεση, ότι τα δεδομένα ακολουθούν κανονική κατανομή και έχουμε 3σ όρια ελέγχου στη Φάση II) η οποία παριστάνει την πιθανότητα ένα σημείο του διαγράμματος να βρεθεί εκτός των ορίων ελέγχου, τότε η μέση τιμή της T και η τυπική της απόκλιση αντίστοιχα θα είναι

$$E(T) = \frac{1}{p} = 370.4 \quad , \quad SD(T) = \frac{\sqrt{1-p}}{p} = 369.9 .$$

Επιπρόσθετα, η δεύτερη ροπή της της T είναι η ακόλουθη

$$E(T^2) = \frac{2-p}{p^2}.$$

Όταν τα όρια ελέγχου του διαγράμματος είναι άγνωστα και πρέπει να εκτιμηθούν τότε η κατανομή του μήκους ροής δεν είναι η γεωμετρική αλλά είναι άγνωστη. Αυτό έχει σαν συνέπεια ότι όταν οι παράμετροι της κατανομής στη Φάση I, δηλαδή η μέση τιμή μ και η τυπική απόκλιση σ πρέπει να εκτιμηθούν, τότε δεν μπορούμε να υπολογίσουμε το μέσο μήκος ροής ARL και την τυπική του απόκλιση $SDRL$.

Γι' αυτόν τον λόγο αρκετοί συγγραφείς έχουν πραγματοποιήσει μελέτες προκειμένου να υπολογίσουν προσεγγιστικά αυτές τις τιμές κατά την εκτίμηση των παραμέτρων στη Φάση I. Η ακρίβεια των προσεγγίσεων μελετάται κυρίως μέσω μεθόδων προσομοίωσης.

Στην παρούσα παράγραφο θα αναλυθεί η άγνωστη κατανομή του μήκους ροής κατά την εκτίμηση της τυπικής απόκλισης σ όταν αυτή εκτιμάται με την μέθοδο R και την μέθοδο S , δηλαδή μέσω των ποσοτήτων $\sigma = \frac{\bar{R}}{d_2}$ και $\sigma = \frac{\bar{S}}{c_4}$ αντίστοιχα.

Επιπλέον στη συνέχεια θα εξετάσουμε ποια είναι επίδραση της εκτίμησης της τυπικής απόκλισης σ στις τιμές ARL και $SDRL$, καθώς και τον αριθμό αλλά και το μέγεθος δειγμάτων που προτείνονται προκειμένου να ελαχιστοποιηθεί η επίδραση αυτή.

2.2.1 Δεσμευμένη κατανομή του μήκους ροής

Όταν οι παράμετροι της διεργασίας είναι άγνωστες και οι εκτιμήσεις που προκύπτουν από τη Φάση I χρησιμοποιούνται για την κατασκευή των διαγραμμάτων ελέγχου, οι ιδιότητες του μήκους ροής, συμπεριλαμβανομένων και του μέσου μήκους ροής, πρέπει να ερμηνεύονται πολύ προσεκτικά. Πριν συγκεντρωθούν τα δεδομένα στη Φάση I, το μήκος ροής ενός διαγράμματος ελέγχου είναι εξαρτημένο από τις τυχαίες εκτιμώμενες παραμέτρους. Έτσι το μήκος ροής είναι τυχαία μεταβλητή που δηλώνεται ως T και αν υποθέσουμε ότι οι άγνωστες εντός ελέγχου μέση τιμή και τυπική απόκλιση της διεργασίας, μ_0 και σ_0 , πρόκειται να εκτιμηθούν χρησιμοποιώντας ως δειγματικές εκτιμήσεις, τις V και W , των οποίων οι παρατηρούμενες τιμές δηλώνονται ως v και w αντίστοιχα. Λόγω της εξάρτησης της τυχαίας μεταβλητής T από τις τυχαίες μεταβλητές V και

W , είναι συχνά χρήσιμο να θεωρούμε την δεσμευμένη συνάρτηση πιθανότητας, $f_{T|V,W}(t|v,w) = P(T=t|v,w)$ που υποδηλώνει ότι το μήκος ροής έχει διαφορετική κατανομή για κάθε δυνατή τιμή των $V = v$ και $W = w$. Όμοια οι ροπές και τα ποσοστημόρια αυτής της δεσμευμένης κατανομής παίρνουν διαφορετικές τιμές για κάθε τιμή των v και w . Για παράδειγμα, η δεσμευμένη ροπή της T , $E_T(T|v,w)$ είναι τυχαία μεταβλητή η οποία εξαρτάται από τις τιμές v και w . Μόλις τα δεδομένα της Φάσης I συγκεντρωθούν και οι τιμές v και w έχουν ήδη παρατηρηθεί, η δεσμευμένη συνάρτηση πιθανότητας της κατανομής του μήκους ροής του διαγράμματος θα είναι $f_{T|V,W}(t|v,w)$ ενώ το μέσο μήκος ροής του διαγράμματος θα είναι $E_T(T|v,w)$.

Δυστυχώς, είναι συχνά αδύνατο στην πράξη να πληροφορηθεί ο διαχειριστής (π.χ σε μια βιομηχανία) πως ένα διάγραμμα ελέγχου που έχει κατασκευαστεί με εκτιμώμενες παραμέτρους θα αποδώσει χρησιμοποιώντας τη δεσμευμένη κατανομή. Σε αυτήν την περίπτωση, προκειμένου να υπολογιστεί η απόδοση του διαγράμματος, κανείς πρέπει να γνωρίζει την μέση τιμή μ_0 και την τυπική απόκλιση σ_0 , εκτός από τις τιμές των v και w . Φυσικά, αν κάποιος είχε γνωστές τις τιμές των παραμέτρων, τότε οι εκτιμήσεις των παραμέτρων δεν θα χρειάζονταν να υπολογιστούν, και κατα συνέπεια η εκτίμηση της απόδοσης του διαγράμματος θα γίνονταν μέσω των παραδοσιακών μεθόδων για γνωστές παραμέτρους.

Παρολαυτά, θα μπορούσε κάποιος να υποθέσει περιπτώσεις αυτής της δεσμευμένης κατανομής, έτσι ώστε να αποκτήσει μια εικόνα για την καλύτερη και τη χειρότερη περίπτωση της απόδοσης των διαγραμμάτων ελέγχου με εκτιμώμενες παραμέτρους.

Για παράδειγμα, οι Jones et al. (2001) ξαναέγραψαν τη στατιστική συνάρτηση του διαγράμματος και την κατανομή του μήκους ροής τύπου EWMA με εκτιμώμενες παραμέτρους σε όρους των τυχαίων μεταβλητών χρησιμοποιώντας τους τύπους:

$$Z_0 = \sqrt{m} \frac{V - \mu_0}{\sigma_0 / \sqrt{n}} \quad \text{και} \quad Z_1 = \frac{W}{\sigma_0}$$

όπου m είναι ο αριθμός των δειγμάτων και n είναι ο αριθμός των παρατηρήσεων για κάθε υποομάδα. Η τυχαία μεταβλητή Z_0 παριστάνει την τυποποιημένη διαφορά μεταξύ της εντός ελέγχου μέσης τιμής και της εκτιμώμενης μέσης τιμής, ενώ η Z_1 παριστάνει τον λόγο της εκτιμώμενης τυπικής απόκλισης προς την εντός ελέγχου τυπική απόκλιση. Επίσης, μπορούν να θεωρηθούν υποθετικές τιμές για την Z_0 και την Z_1 και να αντικατασταθούν στη δεσμευμένη

κατανομή του μήκους ροής, η οποία μπορεί να γραφτεί εκ νέου ως $f_{T|Z_0, Z_1}(t|z_0, z_1) = P(T = t | Z_0 = z_0, Z_1 = z_1)$. Για παράδειγμα είναι δυνατόν να μελετηθεί η απόδοση ενός διαγράμματος κατασκευασμένου με το δειγματικό μέσο να υποεκτιμά ή υπερεκτιμά τη μέση τιμή μ_0 κατά ένα, δύο ή τρία τυπικά σφάλματα δηλαδή κατά 1σ, 2σ ή 3σ. Όμοια μπορεί να εκτιμηθεί η απόδοση ενός διαγράμματος κατασκευασμένου με τη δειγματική τυπική απόκλιση που υποεκτιμά ή υπερεκτιμά την τυπική απόκλιση σ_0 κατά 1,5σ ή 2σ. Αυτή η προσέγγιση χρησιμοποιήθηκε από τους Hawkins & Olwell (1998, p. 159-160), Jones et al. (2001, 2004) και Shu (2004).

Γενικότερα, η δεσμευμένη κατανομή του μήκους ροής επιτρέπει στον διαχειριστή να αποκτήσει συγκεκριμένες πληροφορίες για τα σενάρια της καλύτερης και της χειρότερης περίπτωσης εκτίμησης και κάποιος πλέον μπορεί να γνωρίζει μέσω αυτής πόσο άσχημα μπορεί να αποδώσει ένα διάγραμμα ελέγχου με εκτιμώμενες παραμέτρους. Ένα σημαντικό μειονέκτημα όμως είναι ότι η δεσμευμένη κατανομή συχνά είναι περισσότερο ασύμμετρη προς τα δεξιά από τη γεωμετρική κατανομή, όπως συμβαίνει συνήθως στις περιπτώσεις της εκτίμησης των παραμέτρων. Για αυτόν τον λόγο είναι αναγκαίο να εφοδιάζεται το μέσο μήκος ροής με την τυπική απόκλιση του μήκους ροής και διάφορα άνω και κάτω ποσοστιαία σημεία προκειμένου να εκτιμώνται αποτελεσματικά αυτά τα διαγράμματα.

2.2.2 Περιθώρια κατανομή του μήκους ροής

Σε αντίθεση με τη δεσμευμένη κατανομή του μήκους ροής η περιθώρια κατανομή του μήκους ροής, με συνάρτηση πιθανότητας $f_T(t)$, λαμβάνει υπ' όψιν την τυχαία μεταβλητότητα που εισάγεται στη διεργασία μέσω της εκτίμησης παραμέτρων και δεν απαιτεί προηγούμενη γνώση των παρατηρούμενων εκτιμητών. Η περιθώρια κατανομή του μήκους ροής είναι ο "μέσος όρος" της δεσμευμένης κατανομής του μήκους ροής σε όλες τις δυνατές τιμές των εκτιμήσεων των παραμέτρων και μπορεί να χρησιμοποιηθεί πριν συγκεντρωθούν τα δεδομένα της Φάσης I, έτσι ώστε να ληφθούν αποφάσεις σχετικά με τις απαιτήσεις του μεγέθους δείγματος για να επιτευχθεί η επιθυμητή απόδοση. Η περιθώρια κατανομή του μήκους ροής που υπολογίζεται ολοκληρώνοντας τη δεσμευμένη κατανομή πάνω στο εύρος των εκτιμώμενων παραμέτρων και αν οι τυχαίες μεταβλητές V και W είναι ανεξάρτητες, δίνεται από τη σχέση

$$f_T(t) = \int_0^{\infty} \int_{-\infty}^{\infty} f_{T|V,W}(t|v,w) f_v(v) f_w(w) dv dw.$$

Πολλοί συγγραφείς, ανάμεσα στους οποίους και οι Burrroughs et al. (1993,1995), Chen (1997), Chakraborti (2000) και Jones et al. (2001) έχουν χρησιμοποιήσει την περιθώρια κατανομή του μήκους ροής και τις ροπές της για να εκτιμήσουν την απόδοση των διαγραμμάτων ελέγχου με εκτιμώμενες παραμέτρους. Όταν η περιθώρια κατανομή του μήκους ροής $f_T(t)$, οι ροπές της ή τα ποσοστημόρια της χρησιμοποιούνται για την μέτρηση της απόδοσης διαγραμμάτων ελέγχου, πρέπει να λαμβάνεται σοβαρά υπ' όψιν ότι η περιθώρια κατανομή δεν περιγράφει την απόδοση ενός συγκεκριμένου διαγράμματος ελέγχου, αλλά είναι ο μέσος όρος όλων των δυνατών διαγραμμάτων κατασκευασμένων χρησιμοποιώντας δειγματικές εκτιμήσεις από ένα εντός ελέγχου δείγμα Φάσης Ι. Επομένως, η περιθώρια κατανομή δεν μπορεί να μας πληροφορήσει πως ένα συγκεκριμένο διάγραμμα θα αποδώσει αλλά αντίθετα μετράει πως υπολογίζονται τα όρια ελέγχου στα διαγράμματα χρησιμοποιώντας ένα σχέδιο απόδοσης της διεργασίας κατά μέσο όρο. Επιπλέον η περιθώρια κατανομή του μήκους ροής όπως και η δεσμευμένη, είναι περισσότερο ασύμμετρη προς τα δεξιά από ότι η γεωμετρική κατανομή.

2.3 Προσδιορισμός της άγνωστης κατανομής του μήκους ροής

Ο Quesenberry (1993) προσέγγισε αρχικά το θέμα της άγνωστης κατανομής του μήκους ροής μέσω μελέτης προσομοίωσης επισημαίνοντας την ανεπάρκεια των ήδη γνωστών εμπειρικών κανόνων για την επιλογή του μεγέθους και του αριθμού των δειγμάτων που απαιτούνται κατά την εκτίμηση των παραμέτρων.

Στην μελέτη του Chen (1997), η οποία βασίζεται στο άρθρο του Quesenberry (1993) εξετάζεται εκτενώς η άγνωστη κατανομή του μήκους ροής, καθώς και το μέσο μήκος ροής για τα \bar{X} διαγράμματα ελέγχου για τη μέση τιμή κατά την εκτίμηση της τυπικής απόκλισης σ . Συγκεκριμένα μελετώνται οι ιδιότητες της κατανομής όπως η μέση τιμή, η διακύμανση και η τυπική της απόκλιση. Στη συνέχεια τα αποτελέσματα χρησιμοποιούνται προκειμένου να μελετηθεί η καταλληλότητα των ευρέως διαδεδομένων εμπειρικών κανόνων για την επιλογή του αριθμού m και του μεγέθους n των δειγμάτων καθώς και να εξετασθεί η σοβαρότητα της επίδρασης που είχε η εκτίμηση της τυπικής απόκλισης σ . Το συγκεκριμένο θέμα έχει μελετηθεί από πολλούς ερευνητές

(δείτε για παράδειγμα Ghosh et al. (1981), Ng and Case (1992), Del Castillo (1996a,b), Burroughs et al. (1993), Chakraborti (2000,2007), Wu et al. (2002)).

2.3.1 Ιδιότητες της κατανομής του μήκους ροής

Μας ενδιαφέρει η παρακολούθηση της μέσης τιμής του ποιοτικού χαρακτηριστικού X μέσω ενός \bar{X} διαγράμματος ελέγχου. Έστω X_{ij} , $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, παρατηρήσεις που δηλώνουν δεδομένα που συλλέχθηκαν σε παρελθοντικό χρόνο και έστω Y_{ij} , $i = 1, 2, 3, \dots$ και $j = 1, 2, \dots, n$ παρατηρήσεις που αφορούν τωρινά ή μελλοντικά δεδομένα. Για τα δεδομένα X_{ij} , υποθέτουμε ότι $X_{ij} \square N(\mu, \sigma^2)$ ενώ για τα δεδομένα Y_{ij} ισχύει ότι $Y_{ij} \square N(\mu + a\sigma, b^2\sigma^2)$. Για τον χειρισμό της εντός και εκτός ελέγχου διεργασίας χρησιμοποιούμε τις σταθερές a και b . Ειδικότερα όταν $a=0$ και $b=1$, τότε η διεργασία είναι εντός στατιστικού ελέγχου, ενώ σε οποιαδήποτε άλλη περίπτωση θεωρούμε ότι η διεργασία είναι εκτός ελέγχου, είτε λόγω μετατόπισης της μέσης τιμής μ είτε λόγω μετατόπισης της διακύμανσης σ^2 .

Σε ένα \bar{X} διάγραμμα ελέγχου Φάσης II με εκτιμώμενα όρια ελέγχου, θεωρούμε το ενδεχόμενο

F_i : ο i -οστός δειγματικός μέσος \bar{Y}_i βρίσκεται εκτός των ορίων ελέγχου (δηλαδή είτε πάνω από το άνω όριο ελέγχου $UCL = \bar{\bar{X}} + 3 \frac{\sigma}{\sqrt{n}}$, είτε κάτω από το κάτω όριο ελέγχου $LCL = \bar{\bar{X}} - 3 \frac{\sigma}{\sqrt{n}}$).

Επειδή $\bar{X} \square N\left(\mu, \frac{\sigma^2}{mn}\right)$ και για τον δειγματικό μέσο ισχύει $\bar{Y}_i \square N\left(\mu + a\sigma, \frac{b^2\sigma^2}{n}\right)$, τότε για δεδομένη τιμή $\bar{\bar{X}} = \bar{x}$ και για δοθέν σ έχουμε για την πιθανότητα $p = P(F_i | \bar{x}, \sigma)$:

$$\begin{aligned} p &= P(F_i | \bar{x}, \sigma) = P(\bar{Y}_i < LCL \text{ ή } \bar{Y}_i > UCL | \bar{x}, \sigma) \\ &= 1 - P(LCL \leq \bar{Y}_i \leq UCL | \bar{x}, \sigma) \end{aligned}$$

$$\begin{aligned}
&= 1 - P\left(\bar{x} - 3\frac{\sigma}{\sqrt{n}} \leq \bar{Y}_i \leq \bar{x} + 3\frac{\sigma}{\sqrt{n}} \mid Y \sim N\left(\mu + \alpha\sigma, \frac{b^2\sigma^2}{n}\right)\right) \\
&= 1 - \Phi\left(\frac{\bar{x} + 3\frac{\sigma}{\sqrt{n}} - \mu - \alpha\sigma}{\frac{b\sigma}{\sqrt{n}}}\right) + \Phi\left(\frac{\bar{x} - 3\frac{\sigma}{\sqrt{n}} - \mu - \alpha\sigma}{\frac{b\sigma}{\sqrt{n}}}\right) \\
&= 1 - \Phi\left(\frac{1}{b\sqrt{m}}\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{mn}}}\right) + \frac{3\sigma}{b\sigma} - \frac{\alpha}{b}\sqrt{n}\right) + \Phi\left(\frac{1}{b\sqrt{m}}\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{mn}}}\right) - \frac{3\sigma}{b\sigma} - \frac{\alpha}{b}\sqrt{n}\right) \\
&= 1 - \Phi\left(\frac{z}{b\sqrt{m}} + \frac{3}{b}u - \frac{a}{b}\sqrt{n}\right) + \Phi\left(\frac{z}{b\sqrt{m}} - \frac{3}{b}u - \frac{a}{b}\sqrt{n}\right) \\
&= h(z, u; a, b)
\end{aligned}$$

όπου $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{mn}}}$ και $u = \frac{\sigma}{\sigma}$.

Η τυχαία μεταβλητή $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{mn}}}$ έχει την τυποποιημένη κανονική κατανομή.

Ας συμβολίσουμε με $f_U(u)$ την συνάρτηση πυκνότητας της τυχαίας μεταβλητής $U = \frac{\sigma}{\sigma}$.

Επίσης έχουμε ότι τα ενδεχόμενα $\{F_i\}$ για δοθέντα \bar{x} και σ είναι ανεξάρτητα.

Αφού οι τ.μ \bar{X} και σ είναι ανεξάρτητες τότε για το μέσο μήκος ροής T του διαγράμματος έχουμε

$$E(T) = E\left\{E(T|\bar{X}, \sigma)\right\}$$

$$= \int_{-\infty}^{+\infty} \int_0^{+\infty} E(T) f_Z(z) f_U(u) dz du$$

$$= \int_{-\infty}^{+\infty} \int_0^{+\infty} \frac{1}{h(z, u; a, b)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right) f_U(u) dz du,$$

$$E(T^2) = E\left\{E(T^2|\bar{X}, \sigma)\right\}$$

$$= \int_{-\infty}^{+\infty} \int_0^{+\infty} E(T^2) f_Z(z) f_U(u) dz du$$

$$= \int_{-\infty}^{+\infty} \int_0^{+\infty} \left(\frac{2 - h(z, u; a, b)}{h^2(z, u; a, b)}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2\right) f_U(u) dz du$$

Στους παραπάνω τύπους η $f_z(z)$ είναι η συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής.

Προφανώς για την τυπική απόκλιση του μήκους ροής έχουμε

$$SD(T) = \sqrt{E(T^2) - E^2(T)}.$$

2.3.2 Υπολογισμός των τιμών ARL και SD κατά την εκτίμηση του σ στα διάγραμματα ελέγχου για τη μέση τιμή

Προκειμένου να υπολογίσουμε τις μέσες τιμές $E(T)$ και $E(T^2)$ αλλά και την τυπική απόκλιση $SD(T)$ είναι απαραίτητο να γνωρίζουμε την σ.π.π. $f_U(u)$, όταν η τυπική απόκλιση σ εκτιμάται με την μέθοδο R και την μέθοδο S , δηλαδή μέσω των αμερόληπτων εκτιμητριών $\sigma = \frac{\bar{R}}{d_2}$ και $\sigma = \frac{\bar{S}}{c_4}$ αντίστοιχα. Η σ.π.π. $f_U(u)$ είναι το αποτέλεσμα της συνέλιξης μιας γνωστής συνάρτησης πυκνότητας πιθανότητας, αλλά η μορφή της $f_U(u)$ είναι αρκετά πολύπλοκη για αριθμητικούς υπολογισμούς.

Ο Chen (1997) θεώρησε ότι επειδή οι δύο παραπάνω εκτιμήτριες της τυπικής απόκλισης σ δίνουν παραπλήσιες εκτιμήσεις, εστίασε στην προσέγγιση της σ.π.π. $f_U(u)$ βασιζόμενος στην μελέτη του Patnaik (1950).

Στο σημείο αυτό αξίζει να σημειωθεί ότι ο Chen (1997) θεώρησε και έναν τρίτο τρόπο εκτίμησης της τυπικής απόκλισης σ χρησιμοποιώντας την κοινή (pooled) δειγματική τυπική απόκλιση που δίνεται από τη σχέση

$$S_p = \left(\frac{1}{m} \sum_{i=1}^m S_i^2 \right)^{\frac{1}{2}} = \left(\frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \right)^{\frac{1}{2}}$$

Η εκτίμηση της τυπικής απόκλισης σ που χρησιμοποίησε είναι η $\sigma = S_p / c_4 [m(n-1)+1]$.

Ο Patnaik (1950) παρουσίασε με σαφήνεια τα βήματα της μεθόδου που πρέπει να ακολουθήσει κάποιος αν επιθυμεί να προσεγγίσει την κατανομή του μέσου εύρους με εκείνης της κατανομής $\frac{c\chi_\nu}{\sqrt{\nu}}$, όπου c και ν είναι σταθερές και χ_ν είναι μια τ.μ χ με ν βαθμούς ελευθερίας (δείτε επίσης Li et al. (1990))

Μέσω της μεθόδου του Patnaik (1950), είναι δυνατόν να προσεγγιστεί η σ.π.π. $f_U(u)$ της τ.μ $U = \frac{\sigma}{\sigma}$ τόσο για την περίπτωση που $U = \frac{\bar{R}}{d_2(n)\sigma}$ όσο και στην περίπτωση που $U = \frac{\bar{S}}{c_4(n)\sigma}$.

Σημειώνουμε εδώ ότι η διακύμανση της $U = \frac{\sigma}{\sigma}$ για τις περιπτώσεις όπου $U = \frac{\bar{R}}{d_2(n)\sigma}$ (μέθοδος R) και $U = \frac{\bar{S}}{c_4(n)\sigma}$ στην περίπτωση που το δείγμα (X_1, \dots, X_n) προέρχεται από πληθυσμό $N(\mu, \sigma^2)$ υπολογίζονται ως ακολούθως:

- Διακύμανση της $U = \frac{\bar{R}}{d_2(n)\sigma}$

$$\text{Var}\left(\frac{\sigma}{\sigma}\right) = \text{Var}\left(\frac{\bar{R}}{d_2(n)\sigma}\right) = \frac{1}{d_2^2(n)} \text{Var}\left(\frac{R_1 + \dots + R_m}{m\sigma}\right)$$

$$= \frac{1}{d_2^2(n)} \frac{1}{m} \text{Var}\left(\frac{R}{\sigma}\right) = \frac{1}{md_2^2(n)} \text{Var}\left(\frac{X_{(n)} - X_{(1)}}{\sigma}\right)$$

$$= \frac{1}{md_2^2(n)} \text{Var}\left(\frac{(X_{(n)} - \mu) - (X_{(1)} - \mu)}{\sigma}\right)$$

$$= \frac{1}{md_2^2(n)} \text{Var}(Z_{(n)} - Z_{(1)})$$

$$= \frac{d_3^2(n)}{md_2^2(n)}$$

όπου R το εύρος ενός κανονικού δείγματος μεγέθους n .

- Διακύμανση της $U = \frac{\bar{S}}{c_4(n)\sigma}$

$$\text{Var}\left(\frac{\sigma}{\sigma}\right) = \text{Var}\left(\frac{\bar{S}}{c_4(n)\sigma}\right) = \frac{1}{c_4^2(n)} \text{Var}\left(\frac{S_1 + \dots + S_m}{m\sigma}\right)$$

$$= \frac{1}{c_4^2(n)} \frac{1}{m^2} m \text{Var}\left(\frac{S}{\sigma}\right) = \frac{1}{c_4^2(n)m\sigma^2} \text{Var}(S)$$

$$= \frac{\sigma^2 - \sigma^2 c_4^2(n)}{\sigma^2 c_4^2(n)m}$$

$$= \frac{1 - c_4^2(n)}{c_4^2(n)m}$$

όπου $\text{Var}(S) = E(S^2) - (E(S))^2 = \sigma^2 - \sigma^2 c_4^2$ είναι η διακύμανση της στατιστικής συνάρτησης S .

Θέτουμε

$$M_1 = \text{Var} \left(\frac{\bar{R}}{d_2(n)\sigma} \right) = \frac{d_3^2(n)}{md_2^2(n)}$$

$$M_2 = \text{Var} \left(\frac{\bar{R}}{d_2(n)\sigma} \right) = \frac{1-c_4^2(n)}{c_4^2(n)m}$$

όπου M_1 και M_2 είναι γνωστές σταθερές που εξαρτώνται από τις τιμές m και n .

Επιθυμούμε να υπολογίσουμε την σ.π.π της τυχαίας μεταβλητής $\frac{c\chi_\nu}{\sqrt{\nu}}$, προκειμένου μέσω αυτής να προσεγγίσουμε την σ.π.π $f_U(u)$ όταν $U = \frac{\bar{R}}{d_2(n)\sigma}$ και όταν $U = \frac{\bar{S}}{c_4(n)\sigma}$.

Σύμφωνα με τον Patnaik, για οποιαδήποτε θετική σταθερά M θέτουμε

$$r = \left\{ -2 + 2\sqrt{1+2M} \right\}^{-1} \quad (1)$$

$$t = M + \frac{1}{16r^3} \quad (2)$$

$$\nu = \left\{ -2 + 2\sqrt{1+2t} \right\}^{-1} \quad (3)$$

$$c = 1 + \frac{1}{4\nu} + \frac{1}{32\nu^2} - \frac{5}{128\nu^3} \quad (4)$$

Η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής $\frac{c\chi_\nu}{\sqrt{\nu}}$ δίνεται τότε από τον τύπο

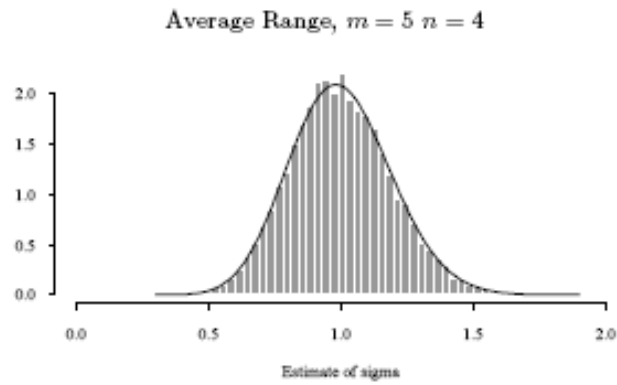
$$f(u; \nu, c) = \frac{2}{c} \frac{(\nu/2)^{\frac{\nu}{2}}}{\Gamma(\nu/2)} \left(\frac{u}{c}\right)^{\nu-1} \exp\left[-\frac{\nu}{2}\left(\frac{u}{c}\right)^2\right], \quad 0 < u < \infty \quad (5)$$

Επομένως, για να προσεγγίσουμε την πυκνότητα $f_U(u)$ της $U = \frac{\sigma}{\sigma} = \frac{\bar{R}}{d_2(n)\sigma}$ βάσει της Σχέσης (5), αρκεί να αντικαταστήσουμε τη σταθερά M στη Σχέση (1) με την σταθερά $M_1 = \frac{d_3^2(n)}{md_2^2(n)}$.

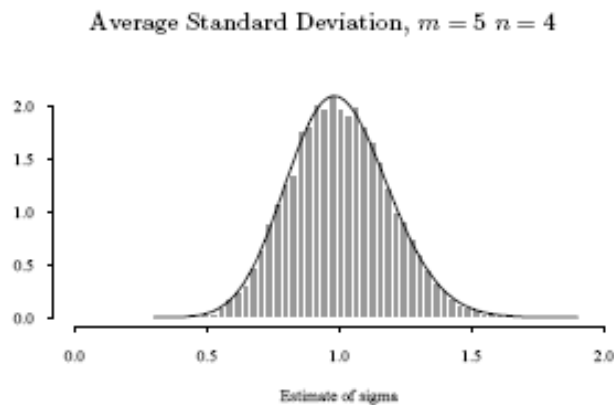
Ομοίως, για να προσεγγίσουμε την πυκνότητα της $f_U(u)$ της $U = \frac{\sigma}{\sigma} = \frac{\bar{S}}{c_4(n)\sigma}$ βάσει της Σχέσης (5), αρκεί να αντικαταστήσουμε την σταθερά M στην Σχέση (1) με την σταθερά $M_2 = \frac{1-c_4^2(n)}{c_4^2(n)m}$.

Η ακρίβεια των παραπάνω προσεγγίσεων στις πυκνότητες των τ.μ. $\frac{\bar{R}}{d_2(n)\sigma}$ και $\frac{\bar{S}}{c_4(n)\sigma}$ αναλύεται μέσω της μεθόδου προσομοίωσης. Συγκεκριμένα, θέτοντας $\sigma=1$, $m=5$ και $n=4$ και για 10.000 τέτοια δείγματα θα υπολογίζουμε 10.000 εκτιμήσεις της άγνωστης παραμέτρου σ με βάση τους δυο τρόπους εκτίμησης της τυπικής απόκλισης, $\sigma = \frac{\bar{R}}{d_2}$ και $\sigma = \frac{\bar{S}}{c_4}$.

Ενδεικτικά είναι τα ακόλουθα ιστογράμματα.



Διάγραμμα 2.1 Ιστόγραμμα πυκνότητας σχετικών συχνοτήτων για την $\frac{\bar{R}}{d_2(4)\sigma}$ και προσαρμοσμένη πυκνότητα $f(w; 13.9259, 1.0181)$



Διάγραμμα 2.2 Ιστόγραμμα πυκνότητας σχετικών συχνοτήτων για την $\frac{\bar{S}}{c_4(4)\sigma}$ και προσαρμοσμένη πυκνότητα $f(w; 14.2745, 1.0177)$

Από τα παραπάνω ιστογράμματα συμπεραίνουμε ότι η προσέγγιση της Σχέσης (5) στις πυκνότητες είναι αρκετά καλή. Επίσης όσο αυξάνονται οι τιμές των m και n η προσέγγιση γίνεται ακόμα καλύτερη.

Στην συνέχεια παρουσιάζουμε πίνακες από τον Chen (1997) για το μέσο μήκος ροής ARL και την τυπική απόκλιση του μήκους ροής SD όταν η εκτίμηση της τυπικής απόκλισης σ , γίνεται μέσω των εκτιμητριών $\sigma = \frac{\bar{R}}{d_2}$, $\sigma = \frac{\bar{S}}{c_4}$ και $\sigma = S_p / c_4 [m(n-1)+1]$.

Ο πρώτος πίνακας αναφέρεται σε εντός ελέγχου διεργασία ($a=0$ και $b=1$) και ο δεύτερος σε εκτός ελέγχου διεργασία εξαιτίας της μετατόπισης της μέσης τιμής ($a \neq 0$ και $b=1$), ή εξαιτίας της μετατόπισης της διασποράς ($a=0$ και $b \neq 1$), ή ακόμα εξαιτίας της μετατόπισης και της μέσης τιμής και της διασποράς ($a \neq 0$ και $b \neq 1$).

Επιπλέον στις τιμές που παρουσιάζονται στον πίνακα για κάθε συνδυασμό των m και n , η πρώτη γραμμή αφορά στην εκτιμήτρια $\sigma = \frac{\bar{R}}{d_2}$, η δεύτερη στην $\sigma = \frac{\bar{S}}{c_4}$ και η τρίτη στην $\sigma = S_p / c_4 [m(n-1)+1]$.

m	ARL							SD						
	n							n						
	4	5	6	7	8	9	10	4	5	6	7	8	9	10
5	3071	1581	1024	775	643	564	512	45928	23152	12429	7225	4587	3179	2381
	2890	1426	904	679	562	493	448	43493	20332	9937	5284	3153	2125	1585
	2223	1191	806	630	534	475	435	33683	15740	7840	4319	2691	1887	1451
10	879	627	524	470	437	414	397	7277	2945	1728	1257	1021	881	790
	848	600	499	446	413	391	375	6684	2583	1497	1085	879	756	676
	754	564	481	435	406	386	371	4947	2146	1344	1012	836	727	657
20	520	454	420	400	387	378	371	1303	893	728	639	584	547	520
	513	445	411	391	377	368	361	1252	850	687	600	546	510	483
	492	435	405	387	375	366	359	1115	798	660	584	535	502	478
30	455	418	398	385	377	371	366	820	651	572	526	496	475	459
	451	413	392	379	371	365	360	800	631	551	504	474	453	437
	440	407	389	377	369	363	359	749	608	538	496	468	449	434
50	415	395	384	377	372	368	366	588	515	477	454	438	427	418
	413	392	380	373	368	365	362	580	505	466	442	426	415	406
	407	389	379	372	367	364	361	559	495	460	438	423	412	404
75	398	385	378	373	370	368	366	502	460	437	422	413	406	400
	397	383	376	371	368	365	364	497	454	430	416	405	398	392
	393	381	375	370	367	365	363	486	448	427	413	404	396	391
100	390	381	376	372	370	368	367	465	435	419	408	401	396	392
	389	380	374	371	368	366	365	461	431	414	403	396	390	386
	387	378	373	370	368	366	365	453	427	411	401	394	389	385
200	380	375	373	371	370	369	368	414	400	393	388	385	382	380
	379	375	372	370	369	368	367	412	399	391	386	382	379	377
	378	374	371	370	369	368	367	409	397	389	385	381	379	377
300	376	373	372	371	370	369	369	399	390	385	382	380	378	377
	376	373	371	370	369	369	368	398	389	384	380	378	376	375
	375	373	371	370	369	369	368	396	387	383	380	377	376	374
500	374	372	371	370	370	370	369	387	382	379	377	376	375	374
	374	372	371	370	370	369	369	386	381	378	376	375	374	373
	373	372	371	370	370	369	369	385	380	377	376	374	373	373
1000	372	371	371	370	370	370	370	378	376	374	373	373	372	372
	372	371	371	370	370	370	370	378	375	374	373	372	372	371
	372	371	370	370	370	370	370	377	375	374	373	372	372	371
∞	370							370						

Πίνακας 2.1 Μέσο μήκος ροής και τυπική απόκλιση του μήκους ροής για \bar{X} διαγράμματα ελέγχου για διάφορες εκτιμήσεις του σ (εντός ελέγχου διεργασία)

m	ARL							SD						
	a/b							a/b						
	.3/1	.6/1	.9/1	0/1.2	0/1.4	.5/1.5	1/2	.3/1	.6/1	.9/1	0/1.2	0/1.4	.5/1.5	1/2
5	789.4	133.7	19.2	148.3	42.5	13.5	3.1	14322.8	3482.2	335.4	709.3	106.6	27.5	3.2
	716.4	123.8	18.4	141.4	41.4	13.3	3.1	12584.5	3065.1	297.5	643.7	100.0	26.3	3.2
	604.8	108.7	17.0	130.3	39.7	13.0	3.1	9753.1	2385.1	235.7	536.1	89.0	24.3	3.1
10	254.7	42.6	9.5	101.6	35.3	11.1	3.0	1354.0	181.1	21.9	216.5	54.0	14.9	2.7
	245.3	41.6	9.4	99.6	34.9	11.0	3.0	1197.6	164.9	20.8	204.0	52.2	14.7	2.6
	232.9	40.3	9.2	96.8	34.4	10.9	2.9	1008.6	145.0	19.3	187.6	49.9	14.2	2.6
20	156.2	28.4	7.5	88.6	32.9	10.1	2.9	334.0	49.1	9.9	125.8	39.9	11.3	2.4
	154.0	28.1	7.4	87.8	32.7	10.1	2.9	320.5	47.8	9.7	122.7	39.4	11.2	2.4
	151.2	27.8	7.4	86.1	32.5	10.1	2.9	304.5	46.2	9.6	118.8	38.6	11.1	2.4
30	133.9	25.3	7.0	85.4	32.2	9.8	2.9	221.7	35.5	8.1	107.4	36.4	10.4	2.4
	132.7	25.1	7.0	84.9	32.1	9.8	2.9	216.3	34.9	8.0	105.6	36.1	10.3	2.4
	131.3	25.0	6.9	84.2	32.0	9.8	2.9	210.1	34.3	7.9	103.5	35.7	10.3	2.4
50	118.6	23.2	6.6	83.2	31.7	9.6	2.8	160.6	28.0	7.0	95.1	33.9	9.7	2.3
	118.0	23.1	6.6	82.9	31.7	9.6	2.8	158.5	27.7	6.9	94.2	33.7	9.7	2.3
	117.4	23.0	6.6	82.5	31.6	9.6	2.8	156.0	27.4	6.9	93.1	33.5	9.6	2.3
75	111.8	22.3	6.5	82.2	31.5	9.5	2.8	136.8	25.0	6.5	89.7	32.8	9.4	2.3
	111.4	22.2	6.5	82.0	31.5	9.5	2.8	135.6	24.8	6.5	89.1	32.6	9.4	2.3
	111.0	22.2	6.5	81.8	31.4	9.5	2.8	134.2	24.7	6.4	88.4	32.5	9.3	2.3
100	108.6	21.8	6.4	81.7	31.4	9.4	2.8	126.2	23.6	6.3	87.1	32.2	9.2	2.3
	108.3	21.8	6.4	81.6	31.4	9.4	2.8	125.4	23.5	6.3	86.7	32.1	9.2	2.3
	108.0	21.7	6.4	81.4	31.4	9.4	2.8	124.5	23.4	6.2	86.2	32.0	9.2	2.3
200	103.9	21.2	6.3	81.1	31.3	9.4	2.8	111.8	21.7	6.0	83.5	31.4	9.0	2.3
	103.8	21.2	6.3	81.0	31.3	9.4	2.8	111.5	21.7	5.9	83.3	31.4	9.0	2.3
	103.7	21.1	6.3	81.0	31.2	9.3	2.8	111.1	21.7	5.9	83.0	31.3	9.0	2.3
300	102.4	21.0	6.3	80.9	31.2	9.3	2.8	107.4	21.2	5.9	82.3	31.1	8.9	2.3
	102.4	21.0	6.3	80.9	31.2	9.3	2.8	107.2	21.1	5.8	82.2	31.1	8.9	2.3
	102.3	20.9	6.3	80.8	31.2	9.3	2.8	106.9	21.1	5.8	82.0	31.1	8.9	2.3
500	101.2	20.8	6.2	80.7	31.2	9.3	2.8	103.8	20.7	5.8	81.3	30.9	8.8	2.3
	101.2	20.8	6.2	80.7	31.2	9.3	2.8	103.8	20.7	5.8	81.3	30.9	8.8	2.3
	101.2	20.8	6.2	80.7	31.2	9.3	2.8	103.7	20.7	5.8	81.2	30.9	8.8	2.3
1000	100.4	20.7	6.2	80.6	31.1	9.3	2.8	101.4	20.4	5.7	80.6	30.8	8.8	2.3
	100.4	20.7	6.2	80.6	31.1	9.3	2.8	101.4	20.4	5.7	80.6	30.8	8.8	2.3
	100.3	20.7	6.2	80.6	31.1	9.3	2.8	101.4	20.4	5.7	80.6	30.8	8.8	2.3
∞	99.5	20.6	6.2	80.5	31.1	9.3	2.8	99.0	20.1	5.7	80.0	30.6	8.8	2.3

Πίνακας 2.2 Μέσο μήκος ροής και τυπική απόκλιση του μήκους ροής για \bar{X} διαγράμματα ελέγχου για διάφορες εκτιμήσεις του σ (εκτός ελέγχου διεργασία, $n=5$)

Στη συνέχεια θα εξετάσουμε την επίδραση της εκτίμησης της τυπικής απόκλισης σ στις τιμές ARL και SD για τα \bar{X} διαγράμματα ελέγχου για τη μέση τιμή αλλά και τον αριθμό και το μέγεθος των δειγμάτων που προτείνονται και είναι κατάλληλα προκειμένου να ελαχιστοποιηθεί η επίδραση αυτή.

2.3.3 Επίδραση των εκτιμώμενων παραμέτρων στις τιμές ARL και SD στα \bar{X} διαγράμματα ελέγχου

Στην παρούσα παράγραφο θα αναλύσουμε τα συμπεράσματα που προκύπτουν από τους πίνακες της προηγούμενης παραγράφου. Από τους δύο πίνακες προκύπτουν τα ακόλουθα:

Η εκτίμηση της τυπικής απόκλισης σ έχει αξιοσημείωτη επίδραση στις τιμές των ARL και SD των \bar{X} διαγραμμάτων ελέγχου για τη μέση τιμή. Παρατηρούμε ότι η επίδραση αυτή είναι αρκετά έντονη όταν $m < 20$, παραμένει σχετικά μεγάλη όταν $20 \leq m < 50$, και μειώνεται όταν $50 \leq m \leq 500$, ενώ η επίδραση είναι πολύ μικρότερη όταν $m > 500$. Φυσικά με την περίπτωση $m = \infty$, εννοούμε την περίπτωση για την οποία τα διαγράμματα σχεδιάζονται με γνωστές παραμέτρους, και τότε η τιμή του μέσου μήκους ροής ARL είναι ίση με 370 για την διεργασία που βρίσκεται εντός στατιστικού ελέγχου.

Επιπρόσθετα, καθώς ο αριθμός των δειγμάτων m αυξάνεται, τότε το μέσο μήκος ροής ARL πλησιάζει στην οριακή του τιμή πολύ πιο γρήγορα από ότι η τυπική του απόκλιση SD . Για παράδειγμα για $m=10$ και $n=6$ είναι $ARL=524$ και $SD=1728$, ενώ όταν ο αριθμός των δειγμάτων m φτάσει π.χ τα 75 τότε $ARL=378$ που είναι πιο “κοντά” στην τιμή 370 από ότι είναι η τιμή του $SD=437$.

Ένα χρήσιμο συμπέρασμα που προκύπτει από τον πρώτο πίνακα είναι ότι ανάμεσα στις τρεις εκτιμήτριες του σ , η εκτιμήτρια $\sigma = S_p / c_4 [m(n-1)+1]$ αποδίδει ομοιομόρφως καλύτερα από την $\sigma = \frac{\bar{S}}{c_4}$, και η $\sigma = \frac{\bar{S}}{c_4}$ αποδίδει ομοιομόρφως καλύτερα από την $\sigma = \frac{\bar{R}}{d_2}$ από την άποψη του SD (είναι πιο “κοντά” στην οριακή του τιμή). Αυτό παρολαυτά δεν ισχύει για το ARL όταν το μέγεθος των δειγμάτων n είναι μεγάλο.

Ειδικότερα από τον δεύτερο πίνακα που αναφέρεται σε εκτός ελέγχου διεργασίες έχουμε τα εξής :

Αν και τα αποτελέσματα έχουν παρουσιαστεί μόνο για $n=5$, τα αποτελέσματα για μεγαλύτερες τιμές του n είναι ποιοτικά τα ίδια.

Επίσης, προκειμένου να αποδώσει ένα \bar{X} διάγραμμα ελέγχου κατασκευασμένο με εκτιμώμενες παραμέτρους παρόμοια όπως αποδίδει στην περίπτωση των γνωστών παραμέτρων, είναι απαραίτητο να ληφθούν τουλάχιστον $m=100$ δείγματα όταν το $n=5$, και θα πρέπει $m \geq 50$ όταν το $n=10$. Αυτό φυσικά δεν είναι πάντα εύκολο στην πράξη, αλλά οπωσδήποτε είναι ένα αποτέλεσμα που πρέπει να λαμβάνεται σοβαρά υπ' όψιν κατά τον σχεδιασμό των \bar{X} διαγραμμάτων ελέγχου.

Επιπλέον, η απόδοση των διαγραμμάτων είναι καλύτερη στην περίπτωση που χρησιμοποιείται η εκτιμήτρια $\sigma = S_p / c_4 [m(n-1)+1]$ και γι' αυτόν τον λόγο η συγκεκριμένη εκτίμηση θα πρέπει να προτιμάται, εκτός εάν η ανάγκη για απλοποίηση ορισμένες φορές καθιστά απαραίτητη την χρησιμοποίηση της εκτιμήτριας $\sigma = \frac{\bar{R}}{d_2}$.

Όταν ο συνολικός αριθμός των μετρήσεων που πρέπει να ληφθεί για να εκτιμηθούν τα όρια ελέγχου του διαγράμματος είναι καθορισμένος (π.χ. όταν το γινόμενο mn είναι μια σταθερά) τότε η αύξηση του n και η μείωση του m μειώνει την τυπική απόκλιση του μήκους ροής. Για παράδειγμα όταν $mn=100$ και χρησιμοποιείται η εκτιμήτρια $\sigma = S_p / c_4 [m(n-1)+1]$, η τυπική απόκλιση είναι $SD=808$ για $m=20$ και $n=5$, ενώ για $m=10$, $n=10$ είναι $SDRL=663$ όπου και οι δυο περιπτώσεις δίνουν αποδεκτή τιμή για το ARL .

Συμπερασματικά, η επίδραση της εκτίμησης των παραμέτρων είναι πιο σοβαρή για μικρές μετατοπίσεις στην διεργασία παρά για μεγαλύτερες μετατοπίσεις και επίσης ότι η επίδραση είναι μεγαλύτερη στις τιμές του SD συγκριτικά με αυτές του ARL . Γι' αυτόν τον λόγο απαιτείται ένα μεγάλο μέγεθος δείγματος. Σε παρόμοια αποτελέσματα κατέληξε και ο Chakraborti (2000) με την διαφορά ότι εκείνος πρότεινε πολύ μεγαλύτερα μεγέθη δειγμάτων (500 με 1000) από ότι ο Quesenberry (1993) και ο Chen (1997).

Επιπρόσθετα, οι εντός και εκτός ελέγχου τιμές των ARL και SD είναι μεγαλύτερες όταν οι παράμετροι εκτιμώνται από ότι στην περίπτωση των γνωστών παραμέτρων, ανεξάρτητα από την επιλογή της εκτιμήτριας της τυπικής απόκλισης σ .

Η χρησιμοποίηση της εκτίμησης του σ μέσω της S_p προτιμάται καθώς βάσει αυτής της εκτιμήτριας η συμπεριφορά του διαγράμματος προσεγγίζει καλύτερα εκείνη του διαγράμματος των γνωστών παραμέτρων.

Σε αυτό το σημείο αξίζει να τονιστεί να σημειωθεί ότι τα συμπεράσματα έχουν νόημα με μια μέση λογική και δεν ισχύουν για συγκεκριμένα \bar{X} διαγράμματα ελέγχου.

2.4 Συνοπτική ανασκόπηση σχετικής βιβλιογραφίας

Οι Prochan & Savage (1960) εξέτασαν την επίδραση του αριθμού m και του μεγέθους n των δειγμάτων στην απόδοση του \bar{X} διαγράμματος ελέγχου, από τη σκοπιά της πιθανότητας να βρεθεί ο δειγματικός μέσος \bar{X} εκτός των ορίων ελέγχου ενώ στην πραγματικότητα είμαστε εντός ελέγχου όταν η άγνωστη παράμετρος της τυπικής απόκλισης σ εκτιμάται μέσω του εύρους \bar{R}_m και της S_p . Για δεδομένες τιμές του μεγέθους n των δειγμάτων παρείχαν τιμές για τον αριθμό m των δειγμάτων που χρειάζονται έτσι ώστε να υπάρχει συγκεκριμένη τιμή της προηγούμενης πιθανότητας στην περίπτωση που η διεργασία είναι εντός ελέγχου.

Παρολαυτά, οι Prochan και Savage (1960) στην μελέτη τους δεν έλαβαν υπ' όψιν τη συσχέτιση μεταξύ του ενδεχομένου ο δειγματικός μέσος \bar{X}_i του δείγματος i να υπερβαίνει το άνω όριο ελέγχου UCL και του ενδεχομένου ο δειγματικός μέσος ενός άλλου δείγματος j να υπερβαίνει το άνω όριο ελέγχου UCL . Γι' αυτόν τον λόγο τα αποτελέσματα της συγκεκριμένης μελέτης είναι περιορισμένης χρήσης.

Ο Hillier (1969) αντιμετώπισε το παραπάνω πρόβλημα των εκτιμώμενων παραμέτρων στην περίπτωση των \bar{X} και R διαγραμμάτων ελέγχου. Συγκεκριμένα, παρουσίασε μια μέθοδο για να υπολογίσει την πιθανότητα σφάλματος τύπου I στην περίπτωση του \bar{X} διαγράμματος

χρησιμοποιώντας το εύρος R προκειμένου να εκτιμήσει την τυπική απόκλιση της διεργασίας. Ωστόσο, η αδυναμία της μεθόδου ήταν η ίδια με τους Prochan και Savage (1960).

Οι Ghosh et al (1981) στη μελέτη τους έδωσαν τύπους για τον υπολογισμό της κατανομής του μήκους ροής των \bar{X} διαγραμμάτων ελέγχου με άγνωστη διακύμανση. Ο σκοπός της μελέτης τους ήταν να εξετάσουν τα ήδη γνωστά κριτήρια που έχουν προταθεί και που υπολογίζουν την απόδοση ενός διαγράμματος ελέγχου από την άποψη της καταλληλότητάς τους με ειδική αναφορά στα Shewhart \bar{X} διαγράμματα για τον έλεγχο της μέσης τιμής όταν η διασπορά της διεργασίας είναι άγνωστη.

Επίσης, ένα ακόμα σημαντικό στοιχείο στο άρθρο τους αναφέρεται στην δημιουργία ενός οικονομικού μοντέλου που εκτιμά την επίδραση στα καθαρά έσοδα και στα κόστη τα οποία σχετίζονται με την κατασκευή ενός διαγράμματος ελέγχου. Την ιδέα για την δημιουργία των οικονομικών μοντέλων πρωτοξεκίνησε ο Duncan (1956) και στη συνέχεια την ακολούθησαν και άλλοι συγγραφείς (Chiu & Cheung (1977)).

Ο Quesenberry (1993) ασχολήθηκε με την επίδραση της εκτίμησης του μέσου της διεργασίας και της τυπικής απόκλισης στα όρια ελέγχου των διαγραμμάτων ελέγχου Shewhart για τον μέσο, τόσο για δείγματα όσο και για μεμονωμένες παρατηρήσεις. Ειδικότερα απέδειξε ότι

$$\text{Corr}(\bar{X}_i - UCL, \bar{X}_j - LCL) = \frac{\text{Var}(UCL)}{\text{Var}(\bar{X}_i - UCL)} = \left[1 + m \left(1 + \frac{9(1 - c_4^2)}{c_4^2} \right)^{-1} \right]^{-1}$$

το οποίο δηλώνει ότι αναμφίβολα υπάρχει συσχέτιση μεταξύ των τυχαίων μεταβλητών $\bar{X}_i - UCL$ και $\bar{X}_j - LCL$. Στη συνέχεια, συμπέρανε ότι για το \bar{X} διάγραμμα ελέγχου απαιτούνται περίπου $\frac{400}{n-1}$ δείγματα για την εκτίμηση των παραμέτρων προκειμένου τα εκτιμώμενα όρια ελέγχου να συμπεριφέρονται με περίπου ίδιο τρόπο όπως τα “θεωρητικά” όρια ελέγχου στην περίπτωση των γνωστών παραμέτρων. Στην περίπτωση των μεμονωμένων παρατηρήσεων έδειξε ότι χρειάζονται 300 παρατηρήσεις για τα εκτιμώμενα όρια ελέγχου έτσι ώστε να συμπεριφέρονται όπως αυτά της περίπτωσης των γνωστών παραμέτρων. Το διάγραμμα ελέγχου το οποίο χρησιμοποίησε στη

δεύτερη περίπτωση ήταν το X διάγραμμα με τη διασπορά να εκτιμάται από το κινούμενο εύρος (moving range).

Μια άλλη προσέγγιση πάνω στο θέμα για την επίδραση της εκτίμησης των παραμέτρων πρότεινε ο Chakraborti (2007). Ο Chakraborti θεώρησε ότι τα παραδοσιακά μέτρα για την απόδοση ενός διαγράμματος ελέγχου, όπως είναι το μέσο μήκος ροής ARL και η τυπική απόκλιση $SDRL$, δεν παρέχουν πλήρη και ικανοποιητική εικόνα για την απόδοση των διαγραμμάτων καθώς η κατανομή του μήκους ροής είναι ασύμμετρη προς τα δεξιά. Η ανάλυση της κατανομής του μήκους ροής και η εύρεση των ποσοστιαίων σημείων της στα Shewhart \bar{X} διαγράμματα ελέγχου πραγματοποιείται στην περίπτωση που η μέση τιμή και η διακύμανση είναι άγνωστες. Φυσικά, υπάρχει και ένας γενικότερος προβληματισμός ο οποίος στηρίζεται στο ότι η σημαντική πληροφορία που μπορεί να αντληθεί μέσω της απόδοσης ενός διαγράμματος ελέγχου είναι πιθανόν να χαθεί αν εστιάσει κανείς σε υπερβολικό βαθμό στο μέσο μήκος ροής ARL . Γι' αυτόν τον λόγο ο συγγραφέας πρότεινε ότι τα εκατοστημόρια (percentiles) της κατανομής του μήκους ροής, συμπεριλαμβανομένης και της διαμέσου του μήκους ροής παρέχουν πολύ καλύτερη ένδειξη για την απόδοση του διαγράμματος ελέγχου.

Ο Chakraborti (2007) υπολόγισε την αθροιστική συνάρτηση κατανομής του μήκους ροής όταν η μέση τιμή μ εκτιμάται από την αμερόληπτη εκτιμήτρια $\bar{\bar{X}}$ και η διασπορά μέσω της σχέσης

$$S = \sqrt{S^2} = \sqrt{\frac{1}{m} \sum_{i=1}^m S_i^2} \quad (\text{Φάση I}).$$

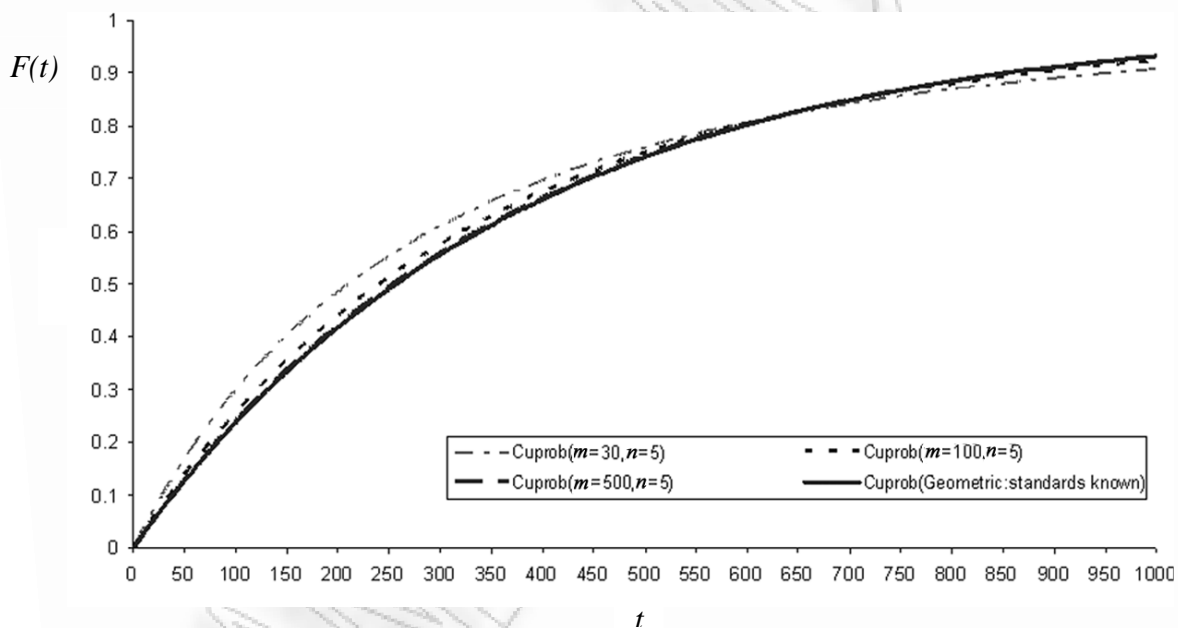
Μέσω της συνάρτησης κατανομής του μήκους ροής είναι δυνατή η μελέτη των ιδιοτήτων του διαγράμματος ελέγχου. Για παράδειγμα κανείς θα μπορούσε να υπολογίσει το $100p$ εκατοστημόριο που εξ' ορισμού είναι ο μικρότερος θετικός ακέραιος αριθμός t στον οποίο η συνάρτηση κατανομής είναι τουλάχιστον ίση με p . Σύμφωνα με τον Chakraborti (2007) η συνάρτηση κατανομής του μήκους ροής δίνεται από τη σχέση

$$P(T \leq t) = 1 - \int_{-\infty}^{\infty} \int_0^{\infty} \left[\Phi \left(-\delta\sqrt{n} + \frac{z_{a/2}}{\sqrt{\nu}} \sqrt{y} + \frac{z}{\sqrt{m}} \right) - \Phi \left(-\delta\sqrt{n} - \frac{z_{a/2}}{\sqrt{\nu}} \sqrt{y} + \frac{z}{\sqrt{m}} \right) \right]^t f_{\chi^2_{\nu}}(y) dy \phi(z) dz$$

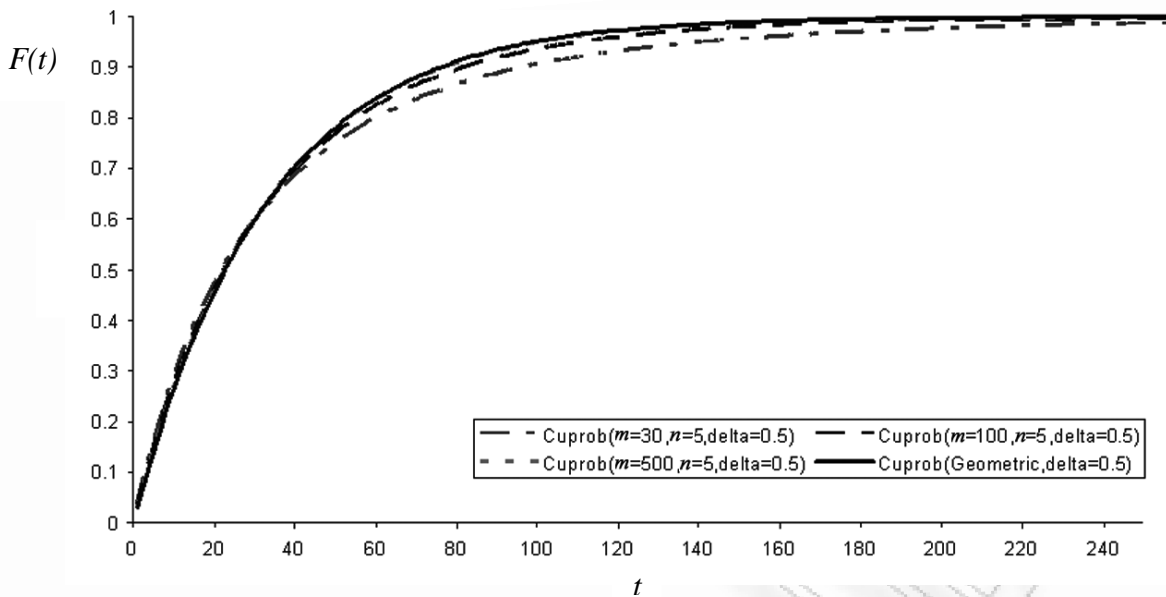
, $t = 1, 2, \dots$

όπου $\delta > 0$ είναι η μετατόπιση της μέσης τιμής σε μονάδες απόκλισης σ , $\nu = m(n-1)$ είναι οι βαθμοί ελευθερίας, a είναι το ποσοστό των λανθασμένων συναγερμών (False Alarm Rate: FAR), $f_{\chi^2_{\nu}}$ είναι η σ.π.π της χ^2 κατανομής με ν β.ε., και ϕ είναι η σ.π.π της τυπικής κανονικής κατανομής.

Στην συνέχεια παρουσιάζονται οι γραφικές παραστάσεις που έδωσε ο Chakraborti για την συνάρτηση κατανομής του μήκους ροής όταν $m = 30, 100, 500$, $n = 5$, $a = 0.0027$.



Διάγραμμα 2.3 Συνάρτηση κατανομής του μήκους ροής για εντός ελέγχου διεργασία ($\delta = 0$) στην περίπτωση των άγνωστων παραμέτρων για $m = 30, 100, 500$, $n = 5$, $a = 0.0027$



Διάγραμμα 2.4 Συνάρτηση κατανομής του μήκους ροής για εκτός ελέγχου διεργασία ($\delta = 0.5$) στην περίπτωση των άγνωστων παραμέτρων για $m = 30, 100, 500$, $n = 5$, $a = 0.0027$

Από τα παραπάνω διαγράμματα συμπεραίνουμε ότι η κατανομή του μήκους ροής ενδέχεται να είναι τελείως διαφορετική από τη γεωμετρική κατανομή, εκτός εάν η τιμή του m είναι μεγάλη, ($m = 500$). Στην εντός ελέγχου διεργασία για $m = 30$ και $m = 100$, η συνάρτηση κατανομής βρίσκεται πάνω από την γεωμετρική κατανομή μέχρι περίπου το 82° εκατοστημόριο. Εντούτοις, πέρα από το 82° εκατοστημόριο η συνάρτηση κατανομής βρίσκεται κάτω από την γεωμετρική κατανομή. Το ίδιο συμβαίνει και για την εκτός ελέγχου διεργασία αλλά για το 62° εκατοστημόριο. Η κατανομή του μήκους ροής κατά την εκτίμηση των παραμέτρων και όταν το m είναι σχετικά μικρό δεν επικρατεί της αντίστοιχης κατανομής στην περίπτωση των γνωστών παραμέτρων αλλά ούτε και συγκλίνει σε αυτή, δηλαδή στη γεωμετρική κατανομή. Αντίθετα, όταν $m = 500$ η συνάρτηση κατανομής τους μήκους ροής συγκλίνει σε εκείνη την κατανομή της περίπτωσης των γνωστών παραμέτρων και στις δυο ουρές της κατανομής (δεξιά και αριστερή ουρά της κατανομής).

Ο Chakraborti (2007) παρουσίασε τον ακόλουθο πίνακα που ακολουθεί με εκατοστημόρια του μήκους ροής για την εντός και εκτός ελέγχου διεργασία όταν ο αριθμός των δειγμάτων είναι $m = 20, 30, 50, 100$, το μέγεθος της μετατόπισης της μέσης τιμής είναι $\delta = 0, 0.2, 0.6, 1$, και το μέγεθος δείγματος είναι $n = 5$ και $a = 0.0027$.

m	Shift δ	ARL	Percentiles of the run length distribution: standards unknown												
			5	10	20	25	30	40	50 (MRL)	60	70	75	80	90	95
20	0.0	422.29	12	25	54	71	90	136	194	274	390	473	582	997	1540
	0.2	250.84	7	14	30	39	49	75	108	154	223	272	338	592	931
	0.6	27.25	1	2	4	6	7	10	14	19	26	32	38	63	95
	1.0	5.14	1	1	1	2	2	2	3	4	5	6	7	11	16
30	0.0	398.77	14	29	62	81	101	150	211	291	405	482	584	947	1390
	0.2	224.52	8	15	33	43	54	80	113	157	221	265	323	535	799
	0.6	24.65	1	2	5	6	7	10	14	19	26	30	36	57	83
	1.0	4.91	1	1	1	2	1	2	3	4	5	6	7	11	15
50	0.0	384.19	16	32	69	90	112	163	227	308	418	492	586	908	1274
	0.2	205.03	8	17	35	46	57	84	117	159	218	258	309	486	693
	0.6	22.86	1	1	5	6	7	10	14	19	25	30	35	53	73
	1.0	4.73	1	1	1	2	1	2	3	4	5	6	7	10	14
100	0.0	375.91	18	36	76	98	122	175	241	322	431	501	590	879	1190
	0.2	191.10	9	18	37	48	60	87	120	161	216	252	298	448	613
	0.6	21.66	1	2	5	6	7	11	14	19	25	29	34	50	67
	1.0	4.61	1	1	1	2	2	3	3	4	5	6	7	10	13
Standards Known	0.0	370.37	19	39	83	107	132	189	257	339	446	513	596	852	1109
	0.2	177.72	10	19	40	40	64	91	123	163	214	214	286	409	531
	0.6	20.56	2	3	5	5	8	11	14	19	25	25	33	47	61
	1.0	4.49	1	1	1	1	2	3	3	5	5	6	7	10	12

Πίνακας 2.3 Εκατοστημόρια και μέσο μήκος ροής στα \bar{X} Shewhart διαγράμματα ελέγχου με εκτιμώμενες παραμέτρους

Από τον παραπάνω πίνακα παίρνουμε μια αρκετά ξεκάθαρη εικόνα από την εξέταση των εκατοστημορίων. Για παράδειγμα, στην εντός ελέγχου διεργασία ($\delta = 0$) για $m = 20, 30, 50, 100$ τα 50% εκατοστημόρια (MRL_0), που ουσιαστικά είναι η διάμεσος, είναι αντίστοιχα 194, 211, 227, 241. Συγκρίνοντας τις τιμές αυτές με την τιμή 257 που είναι η διάμεσος του μήκους ροής στην περίπτωση των γνωστών παραμέτρων, έχουμε ότι οι τιμές MRL_0 είναι μικρότερες κατά την εκτίμηση των παραμέτρων κάτι το οποίο σημαίνει περισσότερους λανθασμένους συναγερμούς από ότι αναμενόταν στην περίπτωση των γνωστών παραμέτρων. Αντίστοιχα για τις παραπάνω τιμές του m , οι τιμές του μέσου μήκους ροής ARL_0 (για $\delta = 0$) είναι 422.29, 398.77, 384.19, 375.91 γεγονός που τονίζει την δεξιά ασυμμετρία της κατανομής του μήκους ροής για εντός ελέγχου διεργασίες. Ειδικότερα, η ασυμμετρία που εμφανίζει προς τα δεξιά είναι πιο έντονη για μικρές τιμές του m . Επιπλέον, κατά την εκτίμηση των παραμέτρων οι ουρές της κατανομής του μήκους ροής συμπεριφέρονται διαφορετικά από τη γεωμετρική κατανομή. Έτσι, ενώ για $p \leq 0.82$, τα εκατοστημόρια είναι μικρότερα από εκείνα της περίπτωσης των γνωστών παραμέτρων, ωστόσο για $p > 0.82$ αυτά είναι μεγαλύτερα. Για παράδειγμα στην εντός ελέγχου διεργασία, για $m = 30$ και $n = 5$, το 90° και το 95° είναι αντίστοιχα 947 και 1390, σε σύγκριση με εκείνα των γνωστών παραμέτρων που είναι 852 και 1109. Αντίθετα, το 40° ($p \leq 0.82$) είναι 150 και 189 όταν οι

παράμετροι είναι άγνωστες και όταν είναι γνωστές αντίστοιχα. Για $m=500$ προκαταρκτικά δείγματα, μεγέθους $n=5$ δηλαδή $mn=2500$ σημεία δεδομένα, η κατανομή του μήκους ροής και τα εκατοστημόρια συγκλίνουν σε εκείνα της περίπτωσης των γνωστών παραμέτρων.

Εξάλλου, αυτό επαληθεύει και το γεγονός ότι για ένα μεγάλο αριθμό δεδομένων στη Φάση I, τα εκτιμώμενα όρια ελέγχου του διαγράμματος συμπεριφέρονται παρόμοια με εκείνα τα όρια ελέγχου στην περίπτωση των γνωστών παραμέτρων.

Ο Del Castillo (1996a) έδειξε ότι οι ιδιότητες του μήκους ροής του \bar{X} διαγράμματος ελέγχου με άγνωστη τυπική απόκλιση σ η οποία εκτιμάται μέσω της σχέσης

$$\sigma = \sqrt{S^2} = \sqrt{\frac{1}{m} \sum_{i=1}^m S_i^2} \quad (6)$$

είναι καλύτερες τόσο από την άποψη του ποσοστού των λανθασμένων συναγεργμών όσο και της ικανότητας στην ανίχνευση μετατοπίσεων στο μέσο επίπεδο στην διεργασία από το αντίστοιχο \bar{X} διάγραμμα ελέγχου που το σ εκτιμάται μέσω της σχέσης $\sigma = \frac{\bar{S}}{c_4}$.

Στη συνέχεια παρουσιάζουμε δύο πίνακες που έδωσε ο Del Castillo για τη συνάρτηση κατανομής του μήκους ροής για εντός ελέγχου διεργασία και το μέσο μήκος ροής στα \bar{X} διαγράμματα ελέγχου με άγνωστη τυπική απόκλιση η οποία εκτιμάται χρησιμοποιώντας την Σχέση (6).

Για τη συνάρτηση κατανομής του μήκους ροής ισχύει ο τύπος (δείτε Ghosh et al. (1981))

$$P(T \leq t) = 1 - \frac{1}{\Gamma(\xi/2)} \int_0^{\infty} [\Phi(c\sqrt{u} - \delta\sqrt{n}) - \Phi(-c\sqrt{u} - \delta\sqrt{n})]^t u^{0.5(\xi-2)} e^{-u} du$$

και για το μέσο μήκος ροής έχουμε

$$ARL = \frac{1}{\Gamma(\xi/2)} \int_0^{\infty} \frac{u^{0.5(\xi-2)} e^{-u}}{1 - \Phi(c\sqrt{u} - \delta\sqrt{n}) + \Phi(-c\sqrt{u} - \delta\sqrt{n})} du$$

όπου ξ είναι οι β.ε. στην εκτίμηση του σ , $c = k\sqrt{2/\xi}$ και το k δηλώνει πόσες τυπικές αποκλίσεις απέχουν τα όρια ελέγχου από την κεντρική γραμμή του διαγράμματος. Οι β.ε. ξ ισοδυναμούν με $m(n-1)$ αν η μέση τιμή εκτιμάται από τον δειγματικό μέσο.

Μέσω των συγκεκριμένων πινάκων έγινε σύγκριση με τους αντίστοιχους πίνακες που έδωσε ο Quesenberry (1993) ο οποίος βασίστηκε στην εκτίμηση $\sigma = \frac{\bar{S}}{c_4}$.

t	$m = 5$	$m = 20$	$m = 50$	geometric
20	0.1160	0.0680	0.0590	0.0526
40	0.1980	0.1290	0.1130	0.1025
60	0.2620	0.1840	0.1640	0.1497
80	0.3130	0.2330	0.2110	0.1945
100	0.3560	0.2780	0.2550	0.2369
200	0.4980	0.4520	0.4340	0.4177
300	0.5820	0.5690	0.5630	0.5556
400	0.6390	0.6520	0.6580	0.6609
500	0.6810	0.7140	0.7280	0.7412
600	0.7140	0.7610	0.7820	0.8025
700	0.7400	0.7980	0.8240	0.8493
800	0.7610	0.8270	0.8560	0.8850
900	0.7790	0.8510	0.8820	0.9123
1000	0.7950	0.8700	0.9020	0.9330
2000	0.8770	0.9560	0.9800	0.9955
4000	0.9320	0.9890	0.9980	1.0000
8000	0.9650	0.9980	1.0000	1.0000
16000	0.9830	1.0000	1.0000	1.0000
20000	0.9860	1.0000	1.0000	1.0000

Πίνακας 2.4 Συνάρτηση κατανομής του μήκους ροής για εντός ελέγχου διεργασία με άγνωστη διασπορά που εκτιμάται από τη Σχέση (6), $n = 5, k = 3, \delta = 0$

m	$\delta = 0.0$	$\delta = 0.5$	$\delta = 1.0$	$\delta = 1.5$	$\delta = 2.0$
5	2088.87	75.41	6.10	1.71	1.10
10	747.82	47.00	5.13	1.63	1.08
20	511.30	39.14	4.78	1.59	1.08
30	456.61	37.03	4.68	1.58	1.08
50	418.86	35.49	4.60	1.57	1.07
75	401.70	34.77	4.56	1.57	1.07
∞	370.40	33.42	4.50	1.57	1.07

Πίνακας 2.5 Τιμές του ARL στα \bar{X} διαγράμματα ελέγχου με άγνωστη διασπορά που εκτιμάται από τη Σχέση (6), $n = 5, k = 3$

Ο Del Castillo έπειτα από σύγκριση με τους πίνακες του Quesenberry κατέληξε στο συμπέρασμα ότι οι τιμές του ARL για την εντός ελέγχου διεργασία ($\delta = 0$) είναι μεγαλύτερες από τις αντίστοιχες τιμές που προκύπτουν όταν χρησιμοποιείται ως εκτιμήτρια η $\sigma = \frac{\bar{S}}{c_4}$. Επιπρόσθετα, για την εκτός ελέγχου διεργασία ($\delta > 0$) όλες οι τιμές του ARL είναι μικρότερες από τις αντίστοιχες που προκύπτουν με την προηγούμενη εκτιμήτρια. Γι' αυτόν τον λόγο θα πρέπει να χρησιμοποιείται η S_p για την εκτίμηση της άγνωστης τυπικής απόκλισης αντί της $\sigma = \frac{\bar{S}}{c_4}$ ειδικά όταν ο διαθέσιμος αριθμός των δειγμάτων είναι μικρός. Το αποτέλεσμα αυτό είναι συνέπεια του μικρότερου τετραγωνικού σφάλματος της εκτιμήτριας S_p .

ΚΕΦΑΛΑΙΟ 3

Επίδραση της εκτίμησης των παραμέτρων στα διαγράμματα ελέγχου για τη διασπορά

3.1 Εισαγωγή

Τα R , S και S^2 διαγράμματα ελέγχου είναι ευρέως διαδεδομένα στην πράξη για την παρακολούθηση της διασποράς μιας διεργασίας. Η συνήθης πρακτική στον έλεγχο διεργασιών είναι η εκτίμηση των ορίων ελέγχου που προκύπτουν από προκαταρκτικά δείγματα (Φάση I) και η χρησιμοποίηση των εκτιμήσεων αυτών για την κατασκευή των ορίων ελέγχου στην Φάση II. Ωστόσο ελάχιστες μελέτες έχουν πραγματοποιηθεί πάνω στην κατανομή του μήκους ροής στα διαγράμματα ελέγχου για τη διασπορά όταν τα όρια ελέγχου βασίζονται σε εκτιμώμενες παραμέτρους.

Στο παρόν κεφάλαιο θα αναλύσουμε την επίδραση της εκτίμησης στα διαγράμματα ελέγχου για τη διασπορά με βάση τις μελέτες που έχουν πραγματοποιήσει πάνω στο θέμα αυτό ο Chen (1998) και οι Maravelakis et al. (2002). Επίσης θα ακολουθήσει εμβάθυνση στα S διαγράμματα ελέγχου με εκτιμώμενη τυπική απόκλιση τόσο για 3σ όρια ελέγχου όσο και για όρια ελέγχου πιθανότητας. Τέλος, θα παρουσιάσουμε και θα αναλύσουμε τα διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις με εκτιμώμενο σ .

Ο Chen (1998) θεώρησε την περίπτωση που η τυπική απόκλιση σ είναι άγνωστη. Τότε τα όρια ελέγχου που θα χρησιμοποιηθούν για την παρακολούθηση της διακύμανσης της διεργασίας στη Φάση I είναι τα

$$UCL = U_n \sigma$$

$$CL = \sigma$$

$$LCL = L_n \sigma$$

όπου σ είναι η εκτίμηση της τυπικής απόκλισης σ , που βασίζεται σε m προκαταρκτικά δείγματα μεγέθους n το καθένα, τα X_{ij} , $1 \leq j \leq n$, $1 \leq i \leq m$, και U_n , L_n είναι κατάλληλες σταθερές που κοντρολάρουν την πιθανότητα α του σφάλματος τύπου I.

Για ένα διάγραμμα ελέγχου με τα παραπάνω εκτιμώμενα όρια ελέγχου, θεωρούμε το ενδεχόμενο

F_i : η εκτίμηση της τυπικής απόκλισης σ_i ($1 \leq i \leq m$) βρίσκεται εκτός των ορίων ελέγχου, (δηλαδή βρίσκεται είτε πάνω από το UCL , είτε κάτω από το LCL).

Η πιθανότητα να συμβεί το $\{F_i\}$ ενδεχόμενο υπολογίζεται από τη σχέση

$$P(F_i) = P(\sigma_i < LCL \text{ ή } \sigma_i > UCL).$$

Ορίζουμε επίσης την τυχαία μεταβλητή W που δηλώνει τον αριθμό των δειγμάτων μέχρι να πραγματοποιηθεί το πρώτο F_i ενδεχόμενο.

Προφανώς, η κατανομή της τυχαίας μεταβλητής W δεν είναι η γεωμετρική καθώς τα ενδεχόμενα $\{F_i\}$ δεν είναι ανεξάρτητα μεταξύ τους.

Ο Hillier (1969) μελέτησε το πρόβλημα του καθορισμού των m , n για να παραχθεί συγκεκριμένο σφάλμα τύπου I στο R διάγραμμα. Συγκεκριμένα, παρατήρησε ότι όταν χρησιμοποιούνται τα εκτιμώμενα όρια ελέγχου, οι σταθερές U_n και L_n που προέρχονται από την περίπτωση των γνωστών παραμέτρων, (δηλαδή της γνωστής τυπικής απόκλισης σ), παύουν να δημιουργούν την επιθυμητή πιθανότητα του σφάλματος τύπου I. Υπό την προϋπόθεση όμως της κανονικότητας των δεδομένων, κατάφερε να τροποποιήσει τις σταθερές U_n και L_n για διάφορες τιμές m και n του μεγέθους των δειγμάτων έτσι ώστε να εγγυηθεί η διατήρηση της επιθυμητής πιθανότητας α του σφάλματος τύπου I. Το συγκεκριμένο θέμα έχει μελετηθεί από πολλούς ερευνητές (δείτε για παράδειγμα Proschan and Savage (1960), Hawkins (1987), Quesenberry (1991,1995).

Ανεξαρτήτως όμως των τροποποιήσεων των σταθερών U_n και L_n , η κατανομή της τυχαίας μεταβλητής W που εξαρτάται από την πιθανότητα $P(F_i)$ δεν είναι η γεωμετρική.

Πράγματι, αυτό διαπιστώνεται από την συνδιακύμανση των τυχαίων μεταβλητών $\sigma_i - L_n \sigma$ και $\sigma_j - L_n \sigma$ που ικανοποιεί τη σχέση

$$\begin{aligned} \text{Cov}(\sigma_i - L_n \sigma, \sigma_j - L_n \sigma) &= \\ &= \text{Cov}(\sigma_i, \sigma_j) - L_n \text{Cov}(\sigma_i, \sigma) - L_n \text{Cov}(\sigma_j, \sigma) + L_n^2 \text{Cov}(\sigma, \sigma) \\ &= L_n^2 \text{Cov}(\sigma, \sigma) \\ &= L_n^2 \text{Var}(\sigma) \neq 0 \quad \text{για κάθε } i \neq j. \end{aligned}$$

Σημειώνεται ότι $\text{Cov}(\sigma_i, \sigma_j) = 0$, επειδή οι τυχαίες μεταβλητές σ_i και σ_j είναι ανεξάρτητες μεταξύ τους.

Ο Chen (1998) προκειμένου να μελετήσει τις ιδιότητες της άγνωστης κατανομής της W εστίασε στον υπολογισμό την πιθανότητας $P(F_i)$, όπως θα δούμε στην επόμενη παράγραφο.

3.2 Η κατανομή του μήκους ροής στα R , S και S^2 διαγράμματα ελέγχου με εκτιμώμενο σ

Ο Chen (1998) μελέτησε την κατανομή του μήκους ροής των R , S και S^2 διαγραμμάτων ελέγχου με εκτιμώμενα όρια ελέγχου. Το R διάγραμμα βασίζεται στην εκτιμήτρια $\bar{R} = \frac{\bar{R}}{d_2}$ του σ ,

το S διάγραμμα βασίζεται στην εκτιμήτρια $\bar{S} = \frac{\bar{S}}{c_4}$ ενώ το S^2 διάγραμμα στην εκτιμήτρια

$$S_p = \sqrt{\bar{S}^2}.$$

Βασιζόμενος στην μέθοδο του Patnaik (1950) ο Chen μελέτησε την επίδραση της εκτίμησης της τυπικής απόκλισης στην κατανομή του μήκους ροής, υπολογίζοντας τη συνάρτηση κατανομής του μήκους ροής, τη μέση τιμή και την τυπική της απόκλιση και συγκρίνοντας τις τιμές αυτές με

τις αντίστοιχες της περίπτωσης των γνωστών παραμέτρων (όταν δηλαδή όταν το σ θεωρείται γνωστό).

Μας ενδιαφέρει η παρακολούθηση της διασποράς του ποιοτικού χαρακτηριστικού X μέσω ενός \bar{X} διαγράμματος ελέγχου. Έστω X_{ij} , $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, δείγματα που δηλώνουν δεδομένα που συλλέχθηκαν σε παρελθοντικό χρόνο, και έστω Y_{ij} , $i = 1, 2, 3, \dots$ και $j = 1, 2, \dots, n$ δείγματα που αφορούν τωρινά ή μελλοντικά δεδομένα. Για τα δεδομένα X_{ij} , υποθέτουμε ότι $X_{ij} \sim f((x - \mu)/\sigma)/\sigma$ ενώ για τα δεδομένα Y_{ij} ισχύει ότι $Y_{ij} \sim f((y - \mu)/(b\sigma))/(b\sigma)$. Η σταθερά b χρησιμοποιείται για τον χειρισμό της εντός και εκτός ελέγχου διεργασίας. Ειδικότερα όταν $b=1$, τότε η διεργασία είναι εντός στατιστικού ελέγχου, ενώ σε οποιαδήποτε άλλη περίπτωση θεωρούμε ότι η διεργασία είναι εκτός ελέγχου, και η διακύμανση σ^2 έχει μετατοπιστεί.

Επίσης, έστω σ η εκτίμηση της τυπικής απόκλισης σ που βασίζεται σε όλα τα X_{ij} δείγματα, έτσι ώστε η τ.μ $U = \frac{\sigma}{\sigma}$ να είναι ανεξάρτητη των μ και σ . Έστω ότι για τη συνάρτηση πυκνότητας της τ.μ U ισχύει $U \sim h(u; m, n)$. Έστω σ_i είναι μια εκτίμηση της ποσότητας $b\sigma$ που βασίζεται στο i -οστό δείγμα Y_{ij} , $j = 1, 2, \dots, n$, τέτοια ώστε η τ.μ $T_i = \sigma_i/(b\sigma)$ να είναι ανεξάρτητη των μ και $b\sigma$.

Η συνάρτηση κατανομής της τ.μ T_i συμβολίζεται με $P(T_i \leq t) = G(t; b, n)$. Για δοθέν σ έχουμε ότι τα ενδεχόμενα F_i είναι ανεξάρτητα και για την πιθανότητα $p = P(F_i | \sigma)$ έχουμε ότι

$$p = P(F_i | \sigma) = P(\sigma_i < L_n \sigma | \sigma) + P(\sigma_i > U_n \sigma | \sigma)$$

$$= P\left(\frac{\sigma_i}{b\sigma} < L_n \frac{\sigma}{b\sigma} \mid \sigma\right) + \left(\frac{\sigma_i}{b\sigma} > U_n \frac{\sigma}{b\sigma} \mid \sigma\right)$$

$$= G(L_n u / b; b, n) + 1 - G(U_n u / b; b, n)$$

$$= l(u; b, n)$$

(1)

$$\text{όπου } u = \frac{\sigma}{\sigma}$$

Η δεσμευμένη συνάρτηση κατανομής της τ.μ. W δίνεται από τον τύπο

$$P(W \leq w | \sigma) = 1 - \{1 - l(u; b, n)\}^w, \quad w = 1, 2, 3, \dots \quad (2)$$

Η σ.κ της τ.μ W δίνεται από τον τύπο

$$P(W \leq w) = 1 - \int_0^{+\infty} \{1 - l(u; b, n)\}^w h(u; m, n) du \quad (3)$$

Για τις δύο πρώτες ροπές της τ.μ W έχουμε

$$E(W) = \int_0^{+\infty} \frac{1}{l(u; b, n)} h(u; m, n) du \quad (4)$$

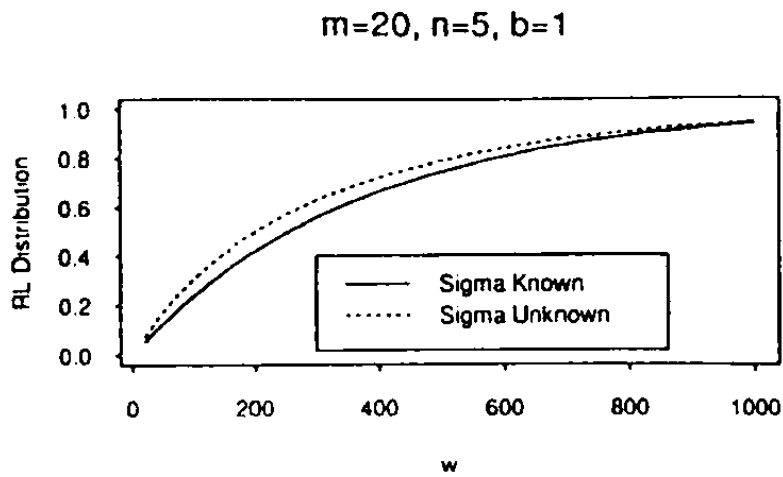
και

$$E(W^2) = \int_0^{+\infty} \frac{2 - l(u; b, n)}{l^2(u; b, n)} h(u; m, n) du \quad (5)$$

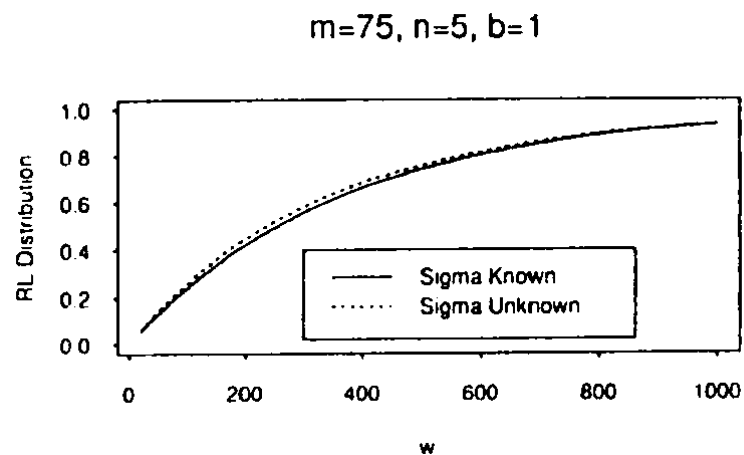
Προφανώς για την τυπική απόκλιση της τ.μ W ισχύει

$$SD(W) = \sqrt{E(W^2) - E^2(W)} \quad (6)$$

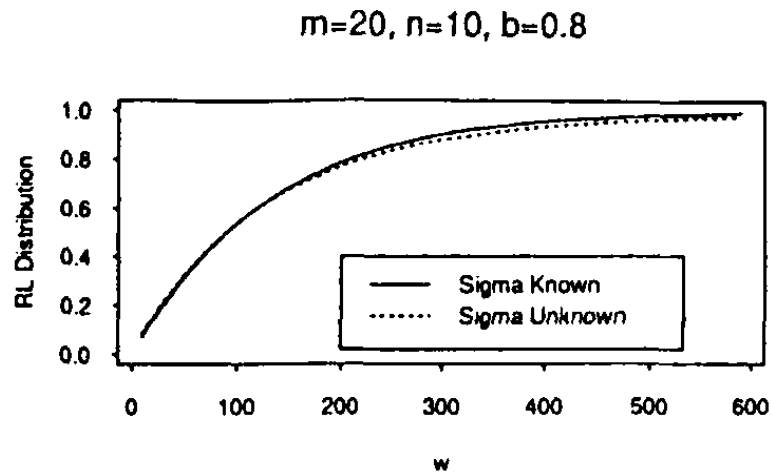
Στη συνέχεια παρουσιάζονται ορισμένες γραφικές παραστάσεις του Chen (1998) για τη συνάρτηση κατανομής του μήκους ροής των S διαγραμμάτων ελέγχου για διάφορες τιμές των m , n και b τόσο για εντός όσο και εκτός ελέγχου διεργασία. Σημειώνεται ότι δίνονται τα πιο αντιπροσωπευτικά γραφήματα δηλαδή εκείνα στα οποία γίνεται φανερό ότι καθώς ο αριθμός των δειγμάτων m αυξάνεται η συνάρτηση κατανομής συγκλίνει σε εκείνη της περίπτωσης των γνωστών παραμέτρων.



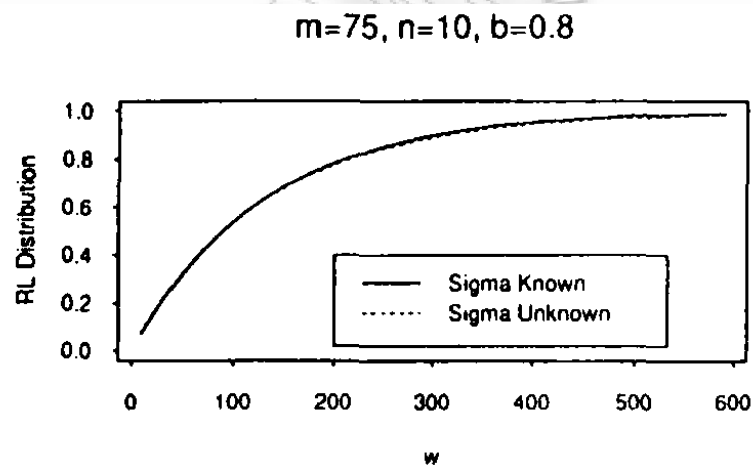
Διάγραμμα 3.1 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=5$, $m=20$, $b=1$



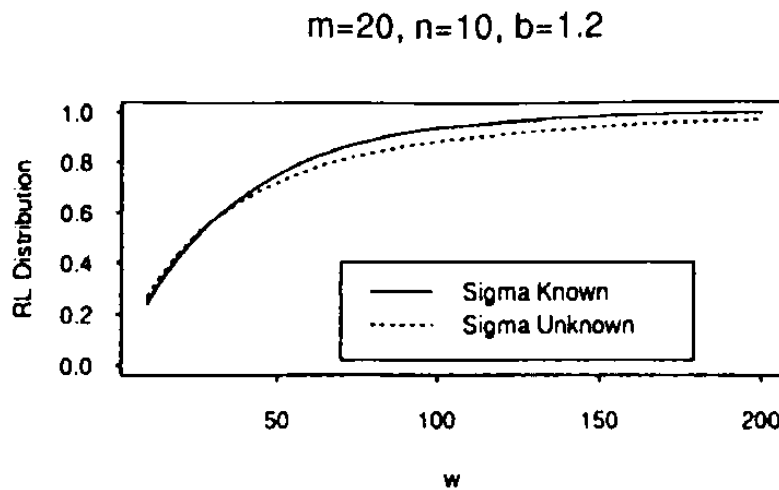
Διάγραμμα 3.2 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=5$, $m=75$, $b=1$



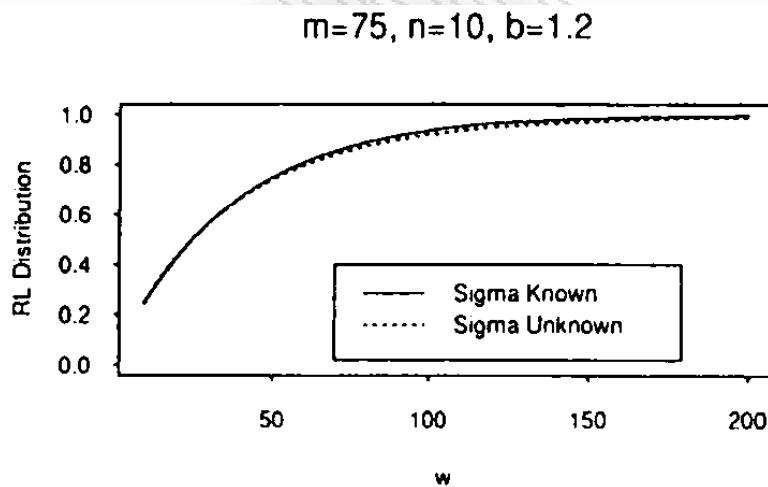
Διάγραμμα 3.3 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=10$, $m=20$, $b=0.8$



Διάγραμμα 3.4 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=10$, $m=75$, $b=0.8$



Διάγραμμα 3.5 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=10$, $m=20$, $b=1.2$



Διάγραμμα 3.6 Συνάρτηση κατανομής του μήκους ροής στο S διάγραμμα ελέγχου για $n=10$, $m=75$, $b=1.2$

Από τα παραπάνω διαγράμματα συμπεραίνουμε ότι όσο μεγαλύτερος είναι ο αριθμός των δειγμάτων και το μέγεθός τους η συνάρτηση κατανομής του μήκους ροής συγκλίνει στην περίπτωση των γνωστών παραμέτρων. Σε αυτό το συμπέρασμα καταλήγουν και στην μελέτη τους οι Castagliola & Maravelakis (2011).

Επιπλέον γίνεται αντιληπτό ότι όταν η διεργασία είναι εντός ελέγχου ($b=1$), η εκτίμηση του σ τείνει να δίνει μεγαλύτερες πιθανότητες για σήμα εκτός ελέγχου διεργασίας κατά μέσο όρο από ότι στην περίπτωση που το σ θεωρείται γνωστό. Όταν η διεργασία είναι εκτός ελέγχου και η διακύμανση έχει μεταβληθεί ($b \neq 1$), τότε τα αποτελέσματα εξαρτώνται από το αν η διακύμανση της διεργασίας έχει αυξηθεί ($b > 1$) ή αν έχει μειωθεί ($b < 1$). Αν η διακύμανση έχει μειωθεί (Διάγραμμα 3.3, 3.4) τότε η εκτίμηση του σ φαίνεται να έχει πολύ μικρή επίδραση στις πιθανότητες για σήμα εκτός ελέγχου διεργασίας κατά μέσο όρο. Αντίθετα, αν η διακύμανση έχει αυξηθεί (Διάγραμμα 3.5, 3.6), η εκτίμηση του σ τείνει να δίνει αρχικά κατά μέσο όρο υψηλότερες πιθανότητες (για σήμα εκτός ελέγχου διεργασίας) και στη συνέχεια τείνει να δίνει μικρότερες πιθανότητες από ότι στην περίπτωση που το σ είναι γνωστό. Επίσης παρατηρούμε ότι η επίδραση της εκτίμησης του σ εξαρτάται από τις τιμές των m και n . Γενικά η επίδραση γίνεται μικρότερη όταν οι τιμές m και n αυξάνονται. Σύμφωνα με τον Chen (1998), για εύρος του n στις περισσότερες εφαρμογές μεταξύ 4 και 10, η επίδραση της εκτίμησης της τυπικής απόκλισης είναι ελάχιστα σημαντική για $m \geq 75$.

Στην συνέχεια παρουσιάζουμε και αναλύουμε τον πίνακα που έδωσε ο Chen (1998) με τις τιμές των ARL και SD που σχετίζονται με τα R, S, S^2 διαγράμματα ελέγχου για τη διασπορά για διάφορες τιμές των m, n και b .

Η περίπτωση με τυπική απόκλιση σ γνωστή δηλώνεται με $m = \infty$.

n	m	Chart	ARL					SD				
			$b = 0.6$	0.8	1.0	1.2	1.4	0.6	0.8	1.0	1.2	1.4
5	20	R	109	312	327	94	24	121	332	366	142	35
		s	109	314	328	87	21	119	333	369	134	30
		s^2	110	317	325	83	20	120	336	366	127	28
	30	R	107	313	339	87	22	114	328	369	120	28
		s	107	314	339	80	19	113	329	371	112	25
		s^2	107	316	337	77	19	113	330	369	107	24
	50	R	105	312	349	81	21	109	322	370	101	24
		s	105	313	350	74	18	108	323	373	92	21
		s^2	105	314	348	72	18	109	323	371	90	20
	75	R	104	311	356	78	20	107	318	371	91	22
		s	104	312	356	71	18	106	318	373	83	19
		s^2	104	312	355	70	18	106	319	371	81	19
	∞	R	102	308	370	72	19	102	307	370	71	19
		s	102	308	370	64	17	102	308	370	64	16
		s^2	102	308	370	64	17	102	308	370	64	16
10	20	R	21	148	319	64	13	23	173	340	91	16
		s	19	144	320	47	9	21	166	345	61	10
		s^2	19	145	319	46	9	21	166	345	67	10
	30	R	20	144	332	60	12	22	160	348	77	14
		s	19	140	333	43	8	20	154	352	57	9
		s^2	19	140	332	42	8	20	155	351	56	9
	50	R	20	140	345	56	12	20	150	355	66	13
		s	18	136	345	41	8	19	145	358	48	8
		s^2	18	137	345	40	8	19	145	358	47	8
	75	R	19	138	352	55	12	20	145	359	61	12
		s	18	135	353	39	8	18	140	362	44	8
		s^2	18	135	352	39	8	18	140	361	43	8
	∞	R	19	135	370	51	11	19	134	370	51	11
		s	18	131	370	37	8	17	131	370	36	7
		s^2	18	131	370	37	8	17	131	370	36	7

Πίνακας 3.1 Μέση τιμή και τυπική απόκλιση του μήκους ροής για R, S, S^2 διαγράμματα ελέγχου για διάφορες τιμές των m, n, b

Από τον παραπάνω πίνακα διαπιστώνουμε ότι η επίδραση της εκτίμησης της τυπικής απόκλισης στα συγκεκριμένα διαγράμματα ελέγχου είναι αξιοσημείωτη. Το πιο σημαντικό χαρακτηριστικό είναι ότι συγκρίνοντας με την περίπτωση του γνωστού σ , η εκτίμηση της τυπικής

απόκλισης μειώνει την τιμή του ARL καθώς και του SD όταν η διεργασία είναι εντός ελέγχου ($b=1$). Αντίθετα όταν η διεργασία είναι εκτός ελέγχου και η τυπική απόκλιση σ της διεργασίας έχει μετατοπιστεί ($b \neq 1$), τότε η εκτίμηση του σ αυξάνει την τιμή του ARL και του SD . Αξίζει να σημειωθεί ότι η μεγαλύτερη διαφορά μεταξύ $m=20$ και $m=\infty$ συμβαίνει στην τιμή του SD όταν η διακύμανση της διεργασίας έχει αυξηθεί ($b > 1$). Αυτό συμβαίνει και όταν $n=5$, που συνεπάγεται ότι η μέθοδος χρησιμοποίησης μικρών δειγμάτων για την κατασκευή των R, S και S^2 διαγραμμάτων ελέγχου για τη διασπορά δεν είναι ικανοποιητική.

Επιπρόσθετα, σύμφωνα με τον Πίνακα 3.1 και στα τρία διαγράμματα το εντός ελέγχου ARL είναι μικρότερο στην περίπτωση της εκτιμώμενης τυπικής απόκλισης, ενώ το εκτός ελέγχου ARL είναι μεγαλύτερο, υποδεικνύοντας ότι οι μετατοπίσεις στη διακύμανση της διεργασίας δεν θα ανιχνευτούν πολύ γρήγορα.

Η χειροτέρευση της απόδοσης γίνεται εντονότερη για αυξήσεις στη διασπορά παρά για μειώσεις, παρόλο που η χειροτέρευση θα ελαττώνεται καθώς το μέγεθος της μετατόπισης θα αυξάνεται. Έτσι, λαμβάνοντας υπ' όψιν και αυτά που προαναφέρθηκαν στο προηγούμενο κεφάλαιο για την επίδραση της εκτίμησης του σ στην μέση τιμή, συμπεραίνεται ότι η επίδραση της εκτίμησης είναι πιο έκδηλη για τα διαγράμματα που χρησιμοποιούνται για την παρακολούθηση της διασποράς από εκείνα που χρησιμοποιούνται για την παρακολούθηση της μέσης τιμής.

Ο εμπειρικός κανόνας που χρησιμοποιείται για την επιλογή ενός δεδομένου αριθμού δειγμάτων ($m=20$ έως $m=30$ προκαταρκτικά δείγματα μεγέθους $n=4$ έως $n=6$ στη Φάση I) στην εκτίμηση της τυπικής απόκλισης είναι ανεπαρκής όπως προκύπτει από την προηγούμενη συζήτηση.

Σύμφωνα με όσα προαναφέρθηκαν για την επίτευξη επαρκούς απόδοσης στην ανίχνευση μετατοπίσεων στην τυπική απόκλιση με $4 \leq n \leq 10$, θα πρέπει να λαμβάνονται τουλάχιστον 75 δείγματα στη Φάση I. Συγκεκριμένα, όταν αρχικά πρόκειται να χρησιμοποιηθούν δοκιμαστικά όρια ελέγχου, ο Chen (1998) πρότεινε ότι πρέπει να λαμβάνονται από τους διαχειριστές 20 με 30 δείγματα μεγέθους 4 έως 5 και αυτά να ανανεώνονται όταν περισσότερα δείγματα γίνουν διαθέσιμα. Αν πρόκειται για σταθερά όρια ελέγχου τα οποία δεν θα μεταβληθούν, τότε πρότεινε ότι πρέπει να λαμβάνονται τουλάχιστον $m=75$ δείγματα μεγέθους $n=5$.

3.3 Εμβάθυνση στα S διαγράμματα ελέγχου με εκτιμώμενο σ

Οι Maravelakis et al. (2002) πραγματοποίησαν παρόμοια μελέτη με εκείνη του Chen (1998) πάνω στην κατανομή του μήκους ροής του S διαγράμματος με όρια ελέγχου 3σ και κατέληξαν σε παρεμφερή συμπεράσματα. Συγκεκριμένα μελέτησαν την επίδραση της εκτίμησης των παραμέτρων της διεργασίας στα διαγράμματα ελέγχου για τη διασπορά επεκτείνοντας τα αποτελέσματα του Chen τόσο για δείγματα όσο και για μεμονωμένες παρατηρήσεις. Μολονότι, οι προτάσεις τους για τα μεγέθη δειγμάτων ήταν διαφορετικές, καθώς πρότειναν μεγαλύτερα μεγέθη δειγμάτων ($m \geq 100$ και $n \geq 20$), αυτό ενδεχομένως να οφείλεται στο γεγονός ότι χρησιμοποίησαν την αμερόληπτη εκτιμήτρια $\sigma = \bar{S}/c_4$ αντί της S_p .

Στο S διάγραμμα ελέγχου Shewhart Φάσης II, θεωρούμε το ενδεχόμενο

A_i : η i -οστή τυπική απόκλιση S_i βρίσκεται εκτός των ορίων ελέγχου (δηλαδή είναι είτε πάνω από το UCL , είτε κάτω από το LCL).

Τότε επειδή τα S_i, S_j είναι ανεξάρτητα για κάθε $i \neq j$, τα ενδεχόμενα A_i και A_j είναι ανεξάρτητα, υποδηλώνοντας ότι η ακολουθία των ενδεχομένων $A_i, i=1,2,\dots$ είναι ακολουθία δοκιμών Bernoulli. Ο μέσος και η τυπική απόκλιση της κατανομής του μήκους ροής, ARL και $SDRL$ αντίστοιχα, της οποίας η κατανομή είναι η γεωμετρική δίνονται φυσικά από τις σχέσεις

$$ARL = \frac{1}{1-\beta}, \quad SDRL = \frac{\sqrt{\beta}}{1-\beta}$$

όπου β είναι το σφάλμα τύπου II και είναι ίσο με $\beta = 1 - P(A_i) = P(LCL \leq S_i \leq UCL)$

Όταν όμως η τυπική απόκλιση είναι άγνωστη, που είναι και η πιο συνηθής περίπτωση (Φάση I) τότε τα όρια ελέγχου πρέπει να εκτιμηθούν από ιστορικά δεδομένα (δείτε Παράγραφο 1.5) και δίνονται ως ακολούθως:

$$UCL = \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \bar{S} = B_4 \bar{S}$$

$$CL = \bar{S}$$

$$LCL = \left(1 - \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \bar{S} = B_3 \bar{S}$$

Για ένα διάγραμμα ελέγχου με τα παραπάνω εκτιμώμενα όρια ελέγχου, θεωρούμε το ενδεχόμενο

B_i : η i -οστή δειγματική τυπική απόκλιση S_i βρίσκεται εκτός των ορίων ελέγχου (δηλαδή είτε πάνω από το UCL , είτε κάτω από το LCL).

Οι προηγούμενοι τύποι για τα ARL και SD δεν μπορούν να εφαρμοστούν εδώ καθώς τα ενδεχόμενα B_i και B_j ($i \neq j$) δεν είναι ανεξάρτητα μεταξύ τους και η κατανομή του μήκους ροής δεν είναι πλέον η γεωμετρική. Αυτό διαπιστώνεται παρακάτω όπως θα δούμε και από την συνδιακύμανση των τ.μ $S_i - UCL$ και $S_j - LCL$.

Συγκεκριμένα, για τη μέση τιμή του άνω εκτιμώμενου ορίου ελέγχου UCL έχουμε

$$E(UCL) = E(B_4 \bar{S}) = B_4 E(\bar{S}) = B_4 E\left(\frac{S_1 + \dots + S_m}{m}\right) = B_4 \frac{1}{m} m \bar{S} = B_4 \bar{S} = UCL$$

ενώ για τη διακύμανση του UCL ισχύει

$$Var(UCL) = Var(B_4 \bar{S}) = B_4^2 Var\left(\frac{S_1 + \dots + S_m}{m}\right) = B_4^2 \frac{1}{m^2} m \sigma^2 (1 - c_4^2) = B_4^2 \frac{\sigma^2 (1 - c_4^2)}{m}$$

Η συνδιακύμανση των τ.μ $S_i - UCL$ και $S_j - LCL$ για $i \neq j$, χρησιμοποιώντας τις παραπάνω σχέσεις δίνεται ως ακολούθως

$$Cov(S_i - UCL, S_j - LCL) =$$

$$= Cov(S_i, S_j) - Cov(S_i, UCL) - Cov(S_j, LCL) + Cov(UCL, LCL)$$

$$= Var(UCL)$$

$$= \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right)^2 \sigma^2 \frac{(1 - c_4^2)}{m} = B_4^2 \frac{\sigma^2 (1 - c_4^2)}{m} \neq 0.$$

Επίσης για τη διακύμανση $Var(S_i - UCL)$ ισχύει

$$\begin{aligned} Var(S_i - UCL) &= Var(S_i) + Var(UCL) = \sigma^2 (1 - c_4^2) + \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right)^2 \sigma^2 \frac{(1 - c_4^2)}{m} \\ &= \left[1 + \frac{\left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right)^2}{m}\right] \sigma^2 (1 - c_4^2) \end{aligned}$$

Επομένως η συσχέτιση μεταξύ των τ.μ $S_i - UCL$ και $S_j - LCL$ υπολογίζεται ως εξής

$$Corr(S_i - UCL, S_j - LCL) =$$

$$= \frac{Cov(S_i - UCL, S_j - LCL)}{\sqrt{Var(S_i - UCL)} \sqrt{Var(S_j - LCL)}} = \frac{Var(UCL)}{Var(S_i - UCL)}$$

$$= \frac{B_4^2 \frac{\sigma^2 (1 - c_4^2)}{m}}{\left(1 + \frac{B_4^2}{m}\right) \sigma^2 (1 - c_4^2)} = \frac{\frac{B_4^2}{m}}{1 + \frac{B_4^2}{m}} = \frac{B_4^2}{m + B_4^2} = \frac{\left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right)^2}{m + \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right)^2}.$$

Από τον τελευταίο τύπο συμπεραίνουμε ότι η συσχέτιση είναι συνάρτηση μόνο των τιμών m και n , καθώς η ποσότητα c_4 είναι σταθερά που εξαρτάται μόνο από το μέγεθος n του δείγματος.

Στον ακόλουθο πίνακα δίνονται τιμές του συντελεστή συσχέτισης για διάφορους συνδυασμούς των m και n .

m	n			
	5	10	20	50
5	0.46581	0.37055	0.30735	0.25370
10	0.30362	0.22741	0.18158	0.14528
20	0.17898	0.12829	0.09986	0.07833
30	0.12689	0.08935	0.06886	0.05362
50	0.08020	0.05560	0.04249	0.03288
100	0.04178	0.02859	0.02171	0.01671
200	0.02133	0.01450	0.01097	0.00843
500	0.00864	0.00585	0.00442	0.00339
1000	0.00434	0.00293	0.00221	0.00170

Πίνακας 3.2 Τιμές του συντελεστή συσχέτισης για διάφορες τιμές των m και n

Από τον πίνακα παρατηρούμε ότι καθώς το μέγεθος δείγματος n και ο αριθμός των δειγμάτων m αυξάνεται, η συσχέτιση μειώνεται. Για σχετικά μικρά ή μεσαία μεγέθη δειγμάτων ($n \leq 20$) απαιτούνται τουλάχιστον $m = 200$ δείγματα έτσι ώστε η συσχέτιση να θεωρείται αμελητέα. Εντούτοις, για μεγαλύτερα μεγέθη δειγμάτων $m = 50$ δείγματα θεωρούνται αρκετά.

Οι Maravelakis et al. (2002) προκειμένου να υπολογίσουν τις δυο πρώτες ροπές της κατανομής του μήκους ροής πραγματοποίησαν μια μελέτη προσομοίωσης βασισμένη σε διάφορες τιμές των m και n . Συγκεκριμένα, οι τιμές m και n που εξετάστηκαν ήταν $m = 5, 10, 20, 50, 100, 200, 500$ και 1000 δείγματα με μέγεθος $n = 5, 10, 20$ και 50 .

Για κάθε συνδυασμό των m, n προσομοιώθηκαν m δείγματα μεγέθους n από μια κανονική κατανομή $N(\mu, \sigma_0^2)$ και υπολογίστηκαν τα όρια ελέγχου UCL και LCL . Στη συνέχεια προσομοιώθηκαν δείγματα από μια κανονική κατανομή $N(\mu, \sigma_1^2)$ μέχρι να ληφθεί μια τιμή πάνω από το UCL ή κάτω από το LCL . Ο αριθμός των δειγμάτων που προσομοιώθηκαν μέχρι ένα δείγμα να οδηγήσει σε μια τιμή εκτός των ορίων ελέγχου αντιστοιχεί σε μια παρατήρηση της

κατανομής του μήκους ροής. Η διαδικασία αυτή επαναλήφθηκε 10000 φορές ώστε να ληφθούν εκτιμήσεις για τις τιμές των *ARL* και *SDRL*. Τα αποτελέσματα παρουσιάζονται στους ακόλουθους πίνακες.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΡΑΙΑ

		σ_1^2 / σ_0^2									
		1		1.2		1.4		1.6		1.8	
m	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	
5	4×10^5	1×10^5	2223.5	7×10^4	594.02	2×10^4	105.38	1353.30	37.41	143.16	
10	2200.10	3×10^4	310.65	2288.70	86.99	330.39	39.29	104.72	21.06	45.34	
20	551.16	1699.70	139.42	297.08	54.88	95.14	27.47	41.75	16.48	21.70	
30	415.06	840.14	112.55	182.48	48.74	74.75	25.79	34.47	15.52	18.54	
50	346.68	545.72	101.62	134.96	43.32	55.43	23.36	27.19	14.73	16.19	
100	298.59	407.05	91.09	106.99	40.75	44.42	22.35	23.56	14.20	14.55	
200	276.08	318.09	85.28	93.97	39.28	41.12	21.55	22.18	13.93	13.92	
500	262.29	275.14	85.20	88.07	38.55	39.24	21.75	21.79	13.94	13.52	
1000	253.76	258.84	84.37	87.67	37.32	37.06	20.97	20.66	13.59	13.14	
∞	249.31	248.81	82.44	81.94	37.72	37.21	21.22	20.71	13.69	13.18	

Πίνακας 3.3 Τιμές των ARL και $SDRL$ για το S διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 5$)

σ_1^2 / σ_0^2										
	1		1.2		1.4		1.6		1.8	
<i>M</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	606.61	1064.81	236.14	634.06	78.31	263.87	29.43	112.67	14.39	41.18
10	538.65	919.10	145.57	329.89	45.83	99.37	19.21	31.64	10.17	13.87
20	461.44	725.80	106.92	175.82	33.86	48.22	15.85	19.75	9.04	10.34
30	430.50	626.79	95.59	137.72	32.34	40.07	15.02	17.08	8.59	9.14
50	389.91	510.09	88.05	106.54	30.29	33.98	14.38	15.65	8.37	8.58
100	359.35	411.69	80.89	88.11	28.79	30.81	13.85	14.24	8.26	8.16
200	344.38	367.08	78.19	82.25	28.46	28.96	13.43	13.31	7.97	7.58
500	334.53	340.97	76.10	76.60	27.45	27.27	13.52	13.14	8.06	7.65
1000	334.56	337.96	75.88	75.93	27.31	27.01	13.50	13.04	7.98	7.48
∞	331.17	330.67	75.66	75.16	27.52	27.01	13.47	12.96	8.00	7.48

σ_1^2 / σ_0^2										
	0.2		0.4		0.6		0.8			
<i>M</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	24.01	37.43	306.28	520.37	1019.6	1274.5	1136.2	1433.6		
10	21.04	25.88	254.28	377.34	1071.7	1253.2	1316.9	1514.7		
20	19.77	22.62	230.60	275.74	1079.4	1207.5	1472.6	1603.4		
30	19.15	20.62	223.33	249.71	1056.5	1155.1	1569.2	1656.9		
50	18.47	19.15	218.20	229.50	1047.3	1106.2	1644.7	1686.9		
100	18.21	18.10	210.57	215.40	1037.5	1061.3	1696.2	1729.9		
200	17.95	17.93	205.32	205.49	1023.1	1027.8	1744.5	1746.7		
500	18.20	17.79	205.59	203.14	1009.1	1022.0	1773.9	1785.0		
1000	17.53	17.28	206.96	204.95	1006.7	1007.9	1768.3	1773.9		
∞	17.90	17.39	206.06	205.56	1011.7	1011.2	1777.2	1776.7		

Πίνακας 3.4 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n=10$)

σ_1^2 / σ_0^2										
	1		1.2		1.4		1.6		1.8	
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	332.72	444.02	121.96	244.63	32.29	78.86	11.46	27.03	5.54	10.33
10	362.96	457.01	92.99	166.56	23.32	45.79	8.71	11.78	4.62	5.39
20	371.24	439.25	75.00	115.39	19.63	25.36	7.87	8.77	4.32	4.20
30	372.32	430.13	68.53	86.90	18.23	21.84	7.67	8.17	4.29	4.12
50	362.66	403.51	63.80	76.74	17.52	18.73	7.49	7.60	4.24	3.97
100	364.01	393.80	60.17	65.57	17.01	17.34	7.36	7.15	4.11	3.58
200	359.00	374.30	59.56	60.31	16.61	16.45	7.15	6.82	4.11	3.59
500	355.18	358.14	59.11	59.61	16.36	16.15	7.13	6.70	4.09	3.56
1000	353.23	353.28	57.59	57.23	16.26	15.79	7.15	6.66	4.08	3.59
∞	356.50	356.00	57.37	56.87	16.39	15.88	7.15	6.63	4.07	3.53

σ_1^2 / σ_0^2									
	0.2		0.4		0.6		0.8		
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	
5	1.32	0.75	11.92	18.83	111.01	190.51	383.43	457.13	
10	1.28	0.64	10.03	12.23	90.20	127.11	423.04	463.05	
20	1.26	0.60	9.21	9.96	80.28	94.68	442.70	473.45	
30	1.26	0.58	8.97	9.20	78.03	88.22	444.96	471.63	
50	1.24	0.54	8.90	8.59	75.70	80.02	451.20	469.84	
100	1.25	0.57	8.68	8.20	73.42	75.19	450.90	455.92	
200	1.23	0.54	8.70	8.27	73.69	74.61	447.50	446.43	
500	1.24	0.55	8.55	8.05	73.62	73.39	441.19	441.96	
1000	1.24	0.56	8.54	8.14	72.08	71.77	445.81	448.29	
∞	1.24	0.54	8.56	8.04	72.91	72.41	449.79	449.29	

Πίνακας 3.5 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 20$)

σ_1^2 / σ_0^2										
	1		1.2		1.4		1.6		1.8	
m	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL
5	263.03	325.32	59.79	125.84	9.49	18.56	3.23	3.96	1.86	1.57
10	304.52	359.74	44.11	76.18	7.69	9.98	2.89	2.87	1.73	1.23
20	328.25	365.28	36.59	49.56	6.91	7.54	2.83	2.52	1.69	1.15
30	340.23	369.51	33.55	39.88	6.65	6.77	2.76	2.37	1.68	1.11
50	345.02	369.81	32.36	35.89	6.64	6.61	2.72	2.24	1.67	1.09
100	355.17	366.97	30.64	31.98	6.37	6.11	2.70	2.20	1.67	1.08
200	357.85	364.35	30.75	30.97	6.39	6.06	2.67	2.09	1.67	1.06
500	362.32	358.59	30.32	30.28	6.38	5.87	2.65	2.10	1.67	1.06
1000	356.30	352.76	30.62	29.97	6.29	5.89	2.67	2.08	1.67	1.05
∞	365.96	365.46	30.23	29.72	6.35	5.83	2.67	2.11	1.66	1.04

σ_1^2 / σ_0^2									
	0.2		0.4		0.6		0.8		
m	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL	
5	1	0	1.25	0.66	8.68	13.45	124.56	199.43	
10	1	0	1.23	0.56	7.20	8.36	110.20	171.69	
20	1	0	1.21	0.51	6.80	7.10	97.84	128.55	
30	1	0	1.20	0.50	6.55	6.53	93.47	110.48	
50	1	0	1.20	0.48	6.51	6.38	89.64	98.31	
100	1	0	1.20	0.48	6.44	6.04	85.98	91.30	
200	1	0	1.19	0.47	6.37	5.91	85.92	88.10	
500	1	0	1.18	0.47	6.34	5.92	85.47	85.74	
1000	1	0	1.18	0.47	6.26	5.82	85.82	85.86	
∞	1	0	1.19	0.48	6.28	5.76	84.25	83.75	

Πίνακας 3.6 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου με εκτιμώμενα όρια ελέγχου ($n = 50$)

Από τους πίνακες Πίνακες 3.3 - 3.6 προκύπτουν ασφαλή συμπεράσματα για διαφορετικά μεγέθη δειγμάτων όταν μετατοπίζεται η διακύμανση της διεργασίας, τόσο προς τα πάνω ($\sigma_1^2 > \sigma_0^2$) όσο και προς τα κάτω ($\sigma_1^2 < \sigma_0^2$). Αξίζει να σημειωθεί ότι για μέγεθος δείγματος $n=5$ η μετατόπιση είναι εφικτή μόνο προς τα πάνω (Πίνακας 3.3). Αυτό συμβαίνει επειδή για $n=5$ το κατώτερο όριο ελέγχου τίθεται ίσο με μηδέν. Έτσι ουσιαστικά το συγκεκριμένο όριο ελέγχου δεν μπορεί ποτέ να παραβιαστεί σε μια μελέτη προσομοίωσης αλλά ούτε και φυσικά στην πραγματικότητα.

Από τους πίνακες παρατηρούμε ότι για μετατοπίσεις προς τα πάνω, καθώς το m αυξάνεται, οι τιμές των ARL και $SDRL$ μειώνονται και προσεγγίζουν τις θεωρητικές τιμές. Για μετατοπίσεις προς τα κάτω συμβαίνει ακριβώς το ίδιο φαινόμενο αλλά μόνο για $n=50$. Όταν τα μεγέθη δειγμάτων είναι $n=10$ και 20 οι τιμές των ARL και $SDRL$ δεν ακολουθούν κάποια συγκεκριμένη τάση. Στην εντός ελέγχου διεργασία επίσης δεν υπάρχει ένα ξεκάθαρο υπόδειγμα τόσο για το μέσο μήκος ροής όσο και για την τυπική του απόκλιση.

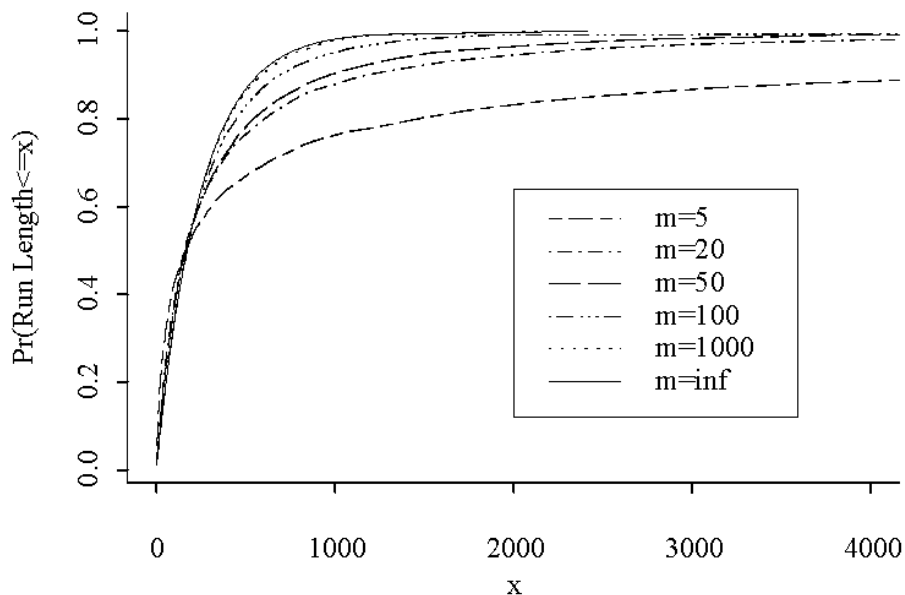
Καθώς ο αριθμός των δειγμάτων m αυξάνεται, η τιμή του ARL προσεγγίζει την θεωρητική του τιμή γρηγορότερα από ότι η τιμή του $SDRL$. Επιπρόσθετα, στην εντός ελέγχου διεργασία καθώς το μέγεθος του δείγματος n αυξάνεται, οι θεωρητικές τιμές της κατανομής του μήκους ροής προσεγγίζουν εκείνες τις τιμές της περίπτωσης της κανονικής κατανομής $N(\mu, \sigma^2)$, οι οποίες είναι $ARL=370.4$ και $SDRL=369.9$. Το ίδιο συμπέρασμα προκύπτει και για την εκτός ελέγχου διεργασία.

Αν πρόκειται να χρησιμοποιηθούν τα S διαγράμματα ελέγχου για ανίχνευση μετατοπίσεων στην διασπορά της διεργασίας, καθίσταται αναγκαία η λήψη δειγμάτων μεγέθους τουλάχιστον 20. Προκειμένου να ελαχιστοποιηθεί η επίδραση της εκτίμησης της τυπικής απόκλισης σ στο S διάγραμμα ελέγχου με όρια ελέγχου 3σ , θα πρέπει δηλαδή οπωσδήποτε να ισχύει $n \geq 20$. Σε αντίθετη περίπτωση, δηλαδή αν το μέγεθος n είναι μικρότερο από την τιμή αυτή, τότε ο διαχειριστής της διεργασίας αναμφισβήτητα θα αντιμετωπίσει έναν αυξημένο αριθμό λανθασμένων συναγερωμών.

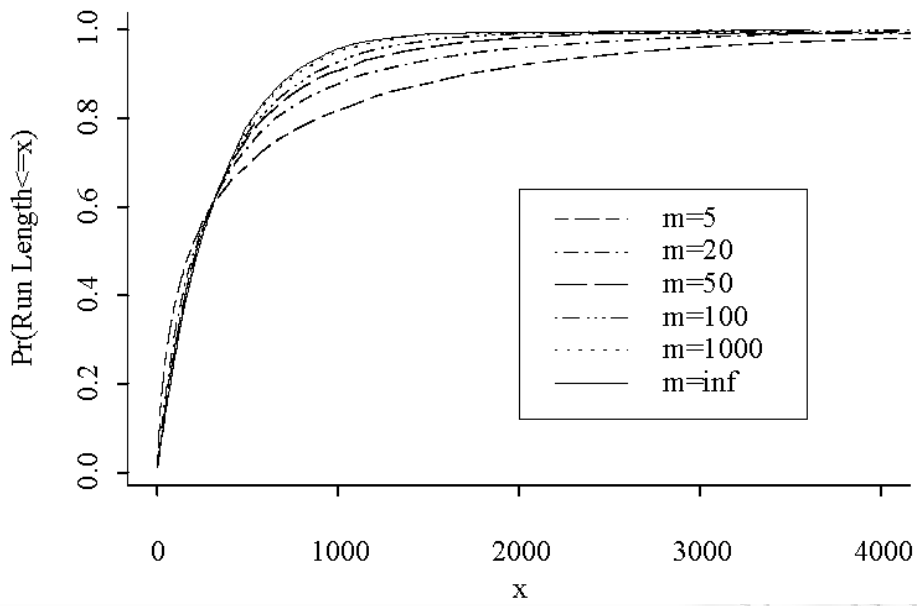
Από τους πίνακες συμπεραίνεται ότι η επίδραση μπορεί να χαρακτηριστεί ιδιαίτερα σοβαρή, ειδικότερα στην εντός ελέγχου διεργασία όταν $m \leq 20$ και για μικρές μετατοπίσεις της διεργασίας. Για τις τιμές $30 \leq m \leq 50$ η επίδραση μετριάζεται σημαντικά, ενώ για τιμές όπου $m \geq 100$, γίνεται αρκετά μικρότερη. Αξίζει να σημειωθεί ότι όταν έχουμε μικρές προς τα κάτω μετατοπίσεις (π.χ $\sigma_1^2 / \sigma_0^2 = 0.2$) για $n \leq 20$, οι τιμές των ARL και $SDRL$ είναι μεγαλύτερες από τις αντίστοιχες εντός ελέγχου τιμές.

Συνεπώς, σε αυτές τις περιπτώσεις χρειάζεται ιδιαίτερη προσοχή στον καθορισμό της επιλογής του τύπου διαγραμμάτων ελέγχου που πρόκειται να χρησιμοποιηθούν. Είναι προτιμότερο να χρησιμοποιούνται διαγράμματα ελέγχου τύπου *CUSUM* ή τύπου *EWMA*, καθώς αυτά τα διαγράμματα είναι πιο ευαίσθητα στην ανίχνευση μικρών μετατοπίσεων σε σύγκριση με τα διαγράμματα Shewhart.

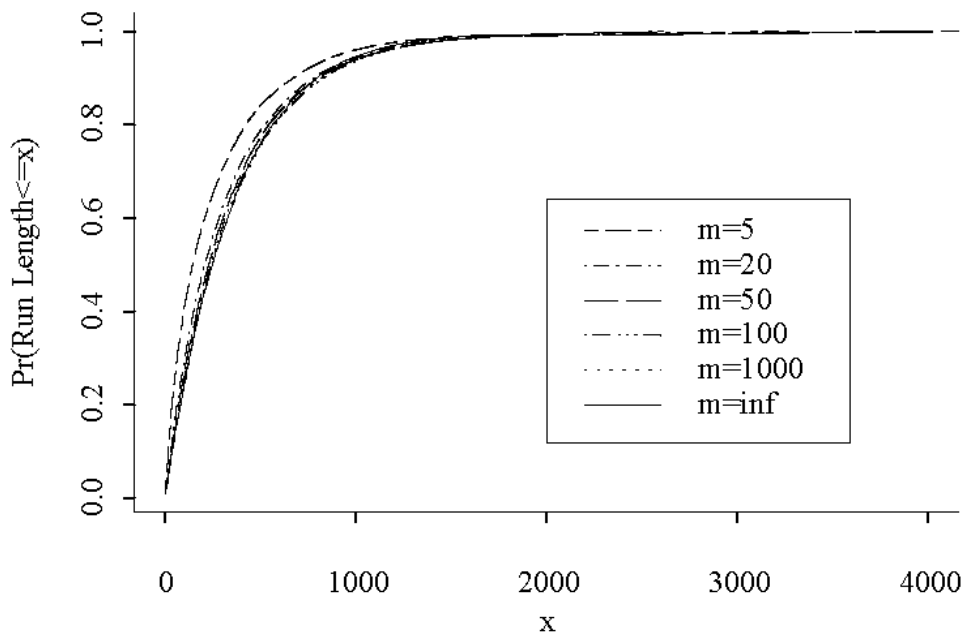
Στην συνέχεια παρουσιάζουμε ορισμένες γραφικές παραστάσεις (Maravelakis et al.) (2002) στις οποίες απεικονίζεται η εμπειρική συνάρτηση κατανομής του μήκους ροής των *S* διαγραμμάτων ελέγχου με όρια ελέγχου 3σ για $n=5$, $n=10$, $n=50$ και για διάφορες τιμές του m .



Διάγραμμα 3.7 Συνάρτηση κατανομής του μήκους ροής για $n=5$



Διάγραμμα 3.8 Συνάρτηση κατανομής του μήκους ροής για $n = 10$



Διάγραμμα 3.9 Συνάρτηση κατανομής του μήκους ροής για $n = 50$

Σε κάθε γραφική παράσταση έχουν σχεδιαστεί 6 διαφορετικοί τύποι γραμμών που αντιπροσωπεύουν τις συναρτήσεις της κατανομής του μήκους ροής για $m = 5, 20, 50, 100, 1000$ και $m = \infty$. Το συμπέρασμα το οποίο προκύπτει είναι ότι καθώς ο αριθμός των δειγμάτων m αυξάνεται στην Φάση I, η εμπειρική συνάρτηση κατανομής προσεγγίζει την θεωρητική κατανομή του μήκους ροής, η οποία είναι η γεωμετρική. Φυσικά για $m = \infty$ οι τιμές των παραμέτρων θεωρούνται γνωστές (Φάση II) και η κατανομή του μήκους ροής είναι η γεωμετρική.

Επιπλέον, από τις γραφικές παραστάσεις γίνεται αντιληπτό ότι καθώς το μέγεθος δείγματος n αυξάνεται, η εμπειρική συνάρτηση κατανομής του μήκους ροής για συγκεκριμένες τιμές του m προσεγγίζει πιο γρήγορα την αντίστοιχη θεωρητική της γεωμετρικής κατανομής. Δηλαδή, για σταθερή τιμή του m , π.χ $m = 50$, η εμπειρική συνάρτηση κατανομής συγκλίνει στη συνάρτηση κατανομής της γεωμετρικής γρηγορότερα με την ταυτόχρονη αύξηση του n , π.χ από $n = 10$ έως $n = 50$.

3.4 S διαγράμματα ελέγχου με όρια ελέγχου πιθανότητας και εκτιμώμενο σ

Οι Maravelakis et al. (2002) εξέτασαν την περίπτωση του S διαγράμματος ελέγχου με όρια ελέγχου πιθανότητας και εκτιμώμενο σ και κατέληξαν σε συμπεράσματα παρόμοια με την περίπτωση των ορίων ελέγχου 3σ .

Στη Φάση II, κατά την οποία η τυπική απόκλιση σ είναι γνωστή τότε όπως έχουμε αναφέρει τα όρια ελέγχου πιθανότητας για το S διάγραμμα είναι γνωστά (S Διάγραμμα ελέγχου με όρια ελέγχου πιθανότητας Φάσης II, βλ. Παρ. 3.1) και δίνονται ως ακολούθως (το $\chi_{n;a}^2$ δηλώνει το άνω α -ποσοστιαίο σημείο της κατανομής χ_n^2)

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

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

Για $a=0.002$, όταν η διασπορά της διεργασίας είναι εντός ελέγχου, τότε η πιθανότητα η τυπική απόκλιση να βρεθεί μεταξύ των ορίων αυτών είναι $P(LCL \leq S_i \leq UCL) = 1 - a = 0.998$ η οποία είναι περίπου ίση με 0.9973, δηλαδή με την πιθανότητα όταν χρησιμοποιούνται τα όρια ελέγχου 3σ .

Αν η τυπική απόκλιση σ είναι άγνωστη, μπορεί να χρησιμοποιηθεί η αμερόληπτη εκτιμήτρια $\sigma = \frac{\bar{S}}{c_4}$ έτσι ώστε να εκτιμηθούν τα όρια ελέγχου πιθανότητας στη Φάση I. Τα όρια ελέγχου (όρια ελέγχου πιθανότητας Φάσης I, S διαγράμματος με βάση την εκτίμηση $\sigma = \frac{\bar{S}}{c_4}$) σε αυτήν την περίπτωση ($a = 0.002$) είναι τα ακόλουθα :

$$UCL = \frac{\bar{S}}{c_4} \sqrt{\frac{\chi_{0.001}^2}{n-1}}$$

$$LCL = \frac{\bar{S}}{c_4} \sqrt{\frac{\chi_{0.999}^2}{n-1}}$$

Όπως στην περίπτωση του S διαγράμματος με όρια ελέγχου 3σ , θα υπολογίσουμε τη συνδιακύμανση των τ.μ $S_i - UCL$ και $S_j - LCL$ εφαρμόζοντας την ίδια μεθοδολογία για να διαπιστώσουμε την μη ανεξαρτησία των ενδεχομένων B_i και B_j και κατά συνέπεια το γεγονός ότι η κατανομή του μήκους ροής δεν είναι η γεωμετρική. Έτσι για την διακύμανση του άνω εκτιμώμενου ορίου ελέγχου UCL έχουμε ότι

$$\begin{aligned} \text{Var}(UCL) &= \text{Var}\left(\frac{\bar{S}}{c_4} \sqrt{\frac{\chi_{0.001}^2}{n-1}}\right) = \frac{1}{c_4^2} \frac{\chi_{0.001}^2}{(n-1)} \text{Var}(\bar{S}) = \frac{\chi_{0.001}^2}{(n-1)c_4^2} \text{Var}\left(\frac{S_1 + \dots + S_m}{m}\right) = \\ &= \frac{\chi_{0.001}^2}{(n-1)c_4^2} \frac{1}{m^2} m\sigma^2 (1 - c_4^2) = \frac{\sigma^2 (1 - c_4^2) \chi_{0.001}^2}{(n-1)c_4^2 m} \end{aligned}$$

Η συνδιακύμανση των τ.μ $S_i - UCL$ και $S_j - LCL$ ($i \neq j$), χρησιμοποιώντας την παραπάνω σχέση, δίνεται από τον τύπο

$$Cov(S_i - UCL, S_j - LCL) = Var(UCL) = \frac{\sigma^2(1-c_4^2)\chi_{0.001}^2}{(n-1)c_4^2m} \neq 0.$$

Επίσης

$$\begin{aligned} Var(S_i - UCL) &= Var(S_i) + Var(UCL) = \sigma^2(1-c_4^2) + \frac{\sigma^2(1-c_4^2)\chi_{0.001}^2}{(n-1)c_4^2m} = \\ &= \sigma^2(1-c_4^2) \left[1 + \frac{\chi_{0.001}^2}{(n-1)c_4^2m} \right] \end{aligned}$$

Επομένως ο συντελεστής συσχέτισης μεταξύ των τ.μ $S_i - UCL$ και $S_j - LCL$ υπολογίζεται ως εξής

$$Corr(S_i - UCL, S_j - LCL) = \frac{Var(UCL)}{Var(S_i - UCL)} = \frac{\frac{\sigma^2(1-c_4^2)\chi_{0.001}^2}{(n-1)c_4^2m}}{\sigma^2(1-c_4^2) \left[1 + \frac{\chi_{0.001}^2}{(n-1)c_4^2m} \right]} = \frac{\chi_{0.001}^2}{\chi_{0.001}^2 + (n-1)c_4^2m}$$

Από τον τελευταίο τύπο συμπεραίνουμε ότι η συσχέτιση είναι συνάρτηση μόνο των τιμών m και n , καθώς η ποσότητα c_4 είναι σταθερά που εξαρτάται από το μέγεθος n του δείγματος.

Στην συνέχεια παρουσιάζουμε έναν πίνακα με τις τιμές του συντελεστή συσχέτισης για διάφορους συνδυασμούς των m και n .

m	n			
	5	10	20	50
5	0.51095	0.39568	0.32137	0.26032
10	0.34314	0.24663	0.19144	0.14964
20	0.20710	0.14066	0.10585	0.08087
30	0.14831	0.09839	0.07315	0.05541
50	0.09460	0.06145	0.04521	0.03400
100	0.04965	0.03170	0.02313	0.01729
200	0.02545	0.01611	0.01170	0.00872
500	0.01034	0.00650	0.00471	0.00351
1000	0.00520	0.00326	0.00236	0.00176

Πίνακας 3.7 Συντελεστής συσχέτισης για διάφορες τιμές των m και n για το S διάγραμμα ελέγχου με όρια ελέγχου πιθανότητας.

Από τον πίνακα προκύπτουν τα ίδια συμπεράσματα όπως και στην περίπτωση των ορίων ελέγχου 3σ . Συγκεκριμένα, καθώς το μέγεθος δείγματος n και ο αριθμός των δειγμάτων m αυξάνεται, η συσχέτιση μειώνεται. Για σχετικά μικρά ή μεσαία μεγέθη δειγμάτων ($n \leq 20$), απαιτούνται $m = 200$ δείγματα έτσι ώστε η συσχέτιση να θεωρείται αμελητέα. Εντούτοις, για μεγαλύτερα μεγέθη δειγμάτων απαιτούνται $m = 50$ δείγματα.

Οι Maravelakis et al. (2002) πραγματοποίησαν μια μελέτη προσομοίωσης για το S διάγραμμα ελέγχου με όρια ελέγχου πιθανότητας βασισμένη σε διάφορες τιμές των m και n , για τον υπολογισμό του ARL και του $SDRL$ παρόμοια με εκείνης της περίπτωσης των 3σ ορίων ελέγχου. Συγκεκριμένα, οι τιμές m και n που εξετάστηκαν ήταν $m = 5, 10, 20, 30, 50, 100, 200, 500$ και 1000 δείγματα με μέγεθος $n = 5, 10, 20$ και 50 . Τα αποτελέσματα παρουσιάζονται στους παρακάτω πίνακες.

σ_1^2 / σ_0^2										
1		1.2		1.4		1.6		1.8		
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	359.97	463.12	267.35	405.54	173.40	312.06	111.11	231.57	71.17	173.00
10	401.46	491.51	268.52	395.19	154.68	263.77	83.88	161.01	47.93	102.77
20	441.09	495.15	254.39	350.22	127.40	199.04	64.92	106.92	36.69	58.40
30	462.04	509.78	247.68	320.05	115.34	164.84	58.02	84.90	33.35	49.45
50	472.24	504.56	239.29	295.97	108.19	137.80	52.48	65.23	30.29	35.47
100	489.90	512.64	229.28	262.50	99.08	115.37	49.79	54.68	28.81	31.03
200	498.35	505.20	221.61	240.21	94.66	102.45	48.20	50.97	27.67	29.10
500	500.93	505.59	216.74	223.58	93.45	95.24	46.06	46.00	28.00	27.60
1000	497.73	503.09	213.01	217.36	92.29	94.50	47.12	47.08	27.31	26.70
∞	500.02	499.52	214.74	214.24	91.78	91.28	46.51	46.01	27.33	26.82

σ_1^2 / σ_0^2									
0.2		0.4		0.6		0.8			
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	
5	59.28	92.90	207.50	279.98	367.15	426.84	423.91	494.16	
10	51.33	60.99	188.80	229.67	383.62	419.59	478.62	506.71	
20	49.25	53.77	178.80	195.31	381.22	406.88	535.06	558.84	
30	47.36	50.56	174.26	182.63	378.10	395.01	551.71	561.82	
50	47.06	47.90	172.37	175.45	374.69	387.19	572.90	579.90	
100	46.45	46.53	170.21	172.41	369.25	373.72	588.21	585.37	
200	44.99	44.60	169.92	169.76	369.89	371.74	595.42	594.84	
500	45.64	45.22	168.09	168.67	364.29	363.97	604.03	601.67	
1000	45.64	44.65	165.86	166.04	364.05	369.30	598.00	601.84	
∞	45.09	44.59	167.40	166.90	366.87	366.37	597.91	597.41	

Πίνακας 3.8 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου ($\alpha = 0.002$, $n = 5$)

σ_1^2 / σ_0^2										
1		1.2		1.4		1.6		1.8		
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	341.44	422.99	217.14	329.46	110.54	218.04	52.43	130.38	25.60	62.05
10	391.03	456.05	208.08	307.21	86.61	155.83	36.61	67.22	17.72	28.51
20	428.95	469.37	194.55	257.65	70.14	106.07	28.50	39.48	14.96	18.23
30	448.41	480.41	187.90	234.75	65.03	88.79	27.33	33.58	14.20	16.15
50	464.28	481.37	178.40	209.85	60.27	72.62	25.81	28.64	13.61	14.78
100	479.05	488.20	169.77	184.28	56.35	61.10	24.52	25.62	13.16	13.69
200	484.86	493.03	166.70	176.09	54.73	56.70	24.26	24.50	12.81	12.66
500	490.54	489.97	161.32	164.74	52.91	53.41	24.02	24.08	13.11	12.82
1000	492.16	480.65	161.60	161.87	53.81	53.11	23.60	23.23	12.70	12.38
∞	500.05	499.55	161.99	161.48	53.44	52.94	23.46	22.95	12.74	12.23

σ_1^2 / σ_0^2									
0.2		0.4		0.6		0.8			
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	
5	5.34	7.41	46.09	81.74	182.01	262.85	339.52	406.60	
10	4.63	4.92	38.28	48.93	162.80	206.09	378.42	422.41	
20	4.40	4.27	35.18	40.33	154.72	181.67	396.79	417.98	
30	4.36	4.10	34.06	37.31	147.37	159.98	400.78	424.06	
50	4.20	3.76	33.36	35.08	144.77	156.54	401.66	421.77	
100	4.27	3.81	32.66	34.24	139.43	140.95	402.89	413.00	
200	4.26	3.73	32.78	33.25	137.34	137.24	400.48	403.36	
500	4.17	3.64	32.44	31.76	136.91	133.59	405.11	405.09	
1000	4.21	3.63	31.95	31.10	133.69	132.26	398.98	400.52	
∞	4.23	3.70	32.13	31.62	136.47	135.97	400.85	400.35	

Πίνακας 3.9 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου ($a = 0.002$, $n = 10$)

σ_1^2 / σ_0^2										
	1		1.2		1.4		1.6		1.8	
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>
5	327.65	381.36	170.43	279.59	56.40	125.43	18.47	47.23	8.07	17.04
10	379.94	415.88	154.94	241.56	40.09	71.48	13.24	19.53	6.45	8.01
20	421.10	434.89	135.60	194.87	33.08	46.55	11.89	14.66	5.93	6.32
30	442.13	451.34	126.95	170.03	30.45	37.64	11.46	12.67	5.81	6.00
50	461.32	467.99	117.98	139.49	29.17	32.83	11.09	11.54	5.69	5.50
100	476.40	478.77	113.42	126.66	27.69	28.82	10.97	10.94	5.57	5.26
200	486.38	486.97	109.86	115.12	27.35	27.82	10.50	10.20	5.50	5.17
500	485.13	488.22	108.31	108.90	26.81	26.19	10.43	10.12	5.41	4.89
1000	494.29	488.10	106.57	108.41	26.63	25.79	10.30	9.78	5.48	4.91
∞	500.01	499.51	106.64	106.14	26.67	26.17	10.42	9.91	5.46	4.93

σ_1^2 / σ_0^2									
	0.2		0.4		0.6		0.8		
m	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	<i>ARL</i>	<i>SDRL</i>	
5	1.17	0.51	7.00	10.62	56.41	103.25	245.11	324.91	
10	1.14	0.43	6.11	7.00	45.91	69.83	247.16	304.65	
20	1.13	0.40	5.71	5.85	40.82	47.93	233.76	273.54	
30	1.13	0.39	5.44	5.32	40.07	44.78	228.37	256.46	
50	1.12	0.36	5.49	5.21	38.59	41.06	223.43	242.44	
100	1.12	0.37	5.36	4.95	37.85	39.05	219.40	225.87	
200	1.11	0.36	5.40	4.92	37.99	38.81	217.81	217.89	
500	1.11	0.36	5.36	4.87	37.41	36.61	213.84	209.40	
1000	1.12	0.37	5.29	4.80	37.69	37.00	213.70	213.52	
∞	1.12	0.36	5.29	4.77	37.44	36.94	215.93	215.43	

Πίνακας 3.10 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου ($\alpha = 0.002$, $n = 20$)

σ_1^2 / σ_0^2										
	1		1.2		1.4		1.6		1.8	
m	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL
5	320.32	380.78	93.28	184.90	13.83	29.14	4.02	5.49	2.10	1.90
10	369.19	410.82	70.91	122.34	10.73	15.32	3.61	4.02	1.93	1.49
20	411.62	433.18	58.04	85.60	9.50	10.86	3.37	3.19	1.93	1.42
30	431.22	447.76	53.56	68.63	8.96	9.54	3.42	3.10	1.90	1.36
50	452.27	459.10	50.77	58.78	8.96	8.95	3.28	2.85	1.89	1.30
100	472.90	472.99	48.14	50.64	8.62	8.59	3.25	2.79	1.88	1.32
200	482.50	481.24	47.71	48.24	8.51	8.24	3.25	2.75	1.86	1.29
500	493.58	498.61	47.47	48.19	8.60	8.11	3.24	2.69	1.85	1.23
1000	490.32	499.04	47.59	47.66	8.56	8.10	3.23	2.66	1.86	1.27
∞	500.01	499.51	47.23	46.73	8.52	8.01	3.22	2.67	1.86	1.27

σ_1^2 / σ_0^2									
	0.2		0.4		0.6		0.8		
m	ARL	SDRL	ARL	SDRL	ARL	SDRL	ARL	SDRL	
5	1	0	1.21	0.62	7.47	11.58	107.80	188.75	
10	1	0	1.19	0.50	6.19	6.96	92.22	141.22	
20	1	0	1.18	0.47	5.86	5.90	79.17	102.14	
30	1	0	1.16	0.45	5.69	5.50	74.86	87.60	
50	1	0	1.16	0.45	5.65	5.46	71.97	78.71	
100	1	0	1.15	0.42	5.64	5.22	69.87	73.40	
200	1	0	1.16	0.43	5.62	5.18	69.61	69.69	
500	1	0	1.15	0.42	5.49	4.95	68.90	70.23	
1000	1	0	1.15	0.42	5.51	5.06	69.27	70.27	
∞	1	0	1.16	0.43	5.48	4.96	68.04	67.54	

Πίνακας 3.11 Τιμές των *ARL* και *SDRL* για το *S* διάγραμμα ελέγχου ($\alpha = 0.002$, $n = 50$)

Από τους πίνακες Πίνακες 3.8 - 3.11 εξάγουμε συμπεράσματα παρόμοια με την ανάλυση της επίδρασης που έχει η εκτίμηση της τυπικής απόκλισης στο S διάγραμμα ελέγχου με όρια ελέγχου 3σ .

Συγκεκριμένα, για μετατοπίσεις προς τα πάνω και καθώς το m αυξάνεται, οι τιμές των ARL και $SDRL$ γενικά μειώνονται και προσεγγίζουν τις θεωρητικές τους τιμές. Για μετατοπίσεις προς τα κάτω, συμβαίνει το ίδιο αλλά για $n=20$ και $n=50$. Όταν τα μεγέθη δειγμάτων είναι $n=10$ και $n=20$ οι τιμές των ARL και $SDRL$ δεν ακολουθούν μια καθορισμένη τάση. Στην εντός ελέγχου διεργασία τόσο το μέσο μήκος ροής όσο και η τυπική του απόκλιση αυξάνονται μέχρι να πλησιάσουν τις θεωρητικές τους τιμές.

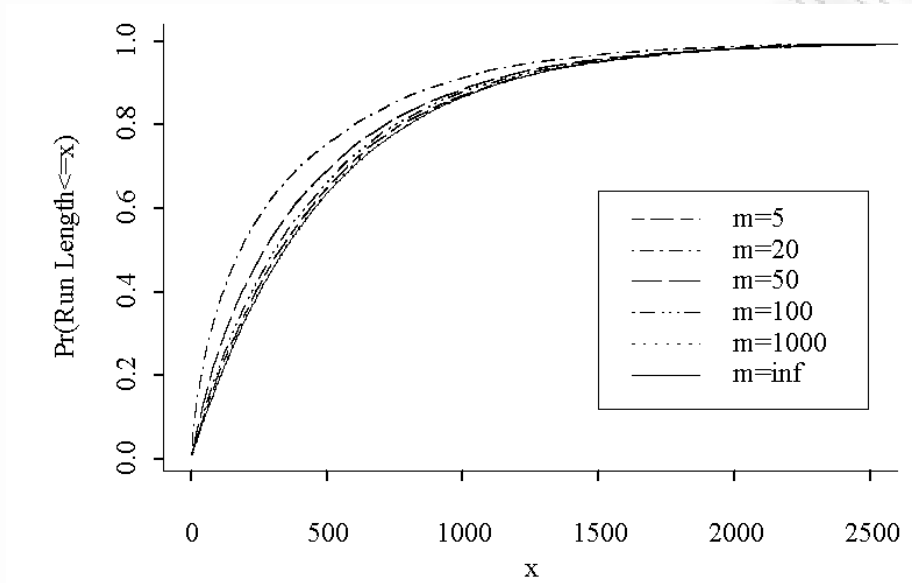
Ως γενικό συμπέρασμα θα μπορούσε να εξαχθεί το γεγονός ότι οι τιμές ARL και $SDRL$ συμπεριφέρονται σχεδόν με τον ίδιο τρόπο εκτός από την περίπτωση που το m αυξάνεται, οπότε η τιμή του ARL πλησιάζει στην θεωρητική του τιμή πιο γρήγορα από ότι η τιμή του $SDRL$.

Αξίζει να σημειωθεί ότι στην εντός ελέγχου διεργασία απαιτούνται τουλάχιστον $m=200$ δείγματα προκειμένου να ελαχιστοποιηθεί η επίδραση που έχει η εκτίμηση του σ , διαφορετικά ο διαχειριστής θα αντιμετωπίσει πολλούς λανθασμένους συναγερούς. Ωστόσο η τιμή n που παίρνει το μέγεθος δείγματος δεν είναι εξίσου σημαντική όπως ο αριθμός των δειγμάτων στην εντός ελέγχου διεργασία.

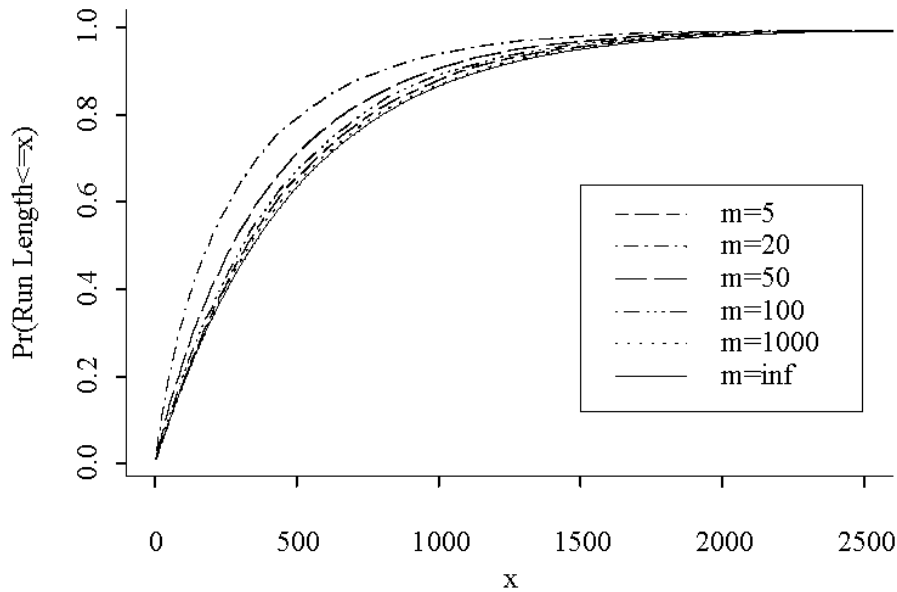
Αντίθετα, στην εκτός ελέγχου διεργασία η τιμή του n είναι σημαντική για την ελαχιστοποίηση της επίδρασης της εκτίμησης της τυπικής απόκλισης. Ειδικότερα, όταν $\sigma_1^2 / \sigma_0^2 = 1.2$ και $m=50$ οι τιμές του ARL για $n=5, 10, 20$ και 50 είναι 239.29, 178.40, 117.98, 50.77 αντίστοιχα. Παρατηρείται λοιπόν δραματική μείωση στο ARL καθώς το n αυξάνεται. Μια παρόμοια κατάσταση συμβαίνει και για μετατοπίσεις προς τα κάτω. Συνεπώς για την ελαχιστοποίηση της επίδρασης για εκτός ελέγχου διεργασίες προτείνονται μεγάλες τιμές του μεγέθους των δειγμάτων, δηλαδή $n \geq 20$.

Επιπρόσθετα, η επίδραση της εκτίμησης είναι σοβαρή για $m \leq 20$, ειδικά για μικρές μετατοπίσεις της διακύμανσης της διεργασίας. Για τιμές $30 \leq m \leq 50$ η επίδραση είναι μέτρια, ενώ για τιμές όπου $m \geq 100$ η επίδραση γίνεται αρκετά μικρή. Ένα ακόμα σημαντικό χαρακτηριστικό που αξίζει να επισημανθεί και αναφέρθηκε και στην περίπτωση των 3σ ορίων ελέγχου, είναι ότι για μικρές προς τα κάτω μετατοπίσεις (όταν $n=5$ και $n=10$, για $m=50$) οι εκτός ελέγχου τιμές των ARL και $SDRL$ είναι μεγαλύτερες από τις αντίστοιχες εντός ελέγχου τιμές. Έτσι σε μια τέτοια περίπτωση είναι προτιμότερο να χρησιμοποιούνται διαγράμματα ελέγχου για ανίχνευση μικρών μετατοπίσεων όπως $CUSUM$ και $EWMA$.

Στην συνέχεια δίνουμε γραφικές παραστάσεις στις οποίες απεικονίζεται η εμπειρική συνάρτηση κατανομής του μήκους ροής των S διαγραμμάτων ελέγχου με όρια ελέγχου πιθανότητας για $n = 5$, $n = 50$ και για διάφορες τιμές του m .



Διάγραμμα 3.10 Συνάρτηση κατανομής του μήκους ροής για $n = 5$ ($\alpha = 0.002$).



Διάγραμμα 3.11 Συνάρτηση κατανομής του μήκους ροής για $n = 50$ ($a = 0.002$).

Από τις παραπάνω γραφικές παραστάσεις της συνάρτησης κατανομής του μήκους ροής των S διαγραμμάτων ελέγχου με όρια ελέγχου πιθανότητας $a = 0.002$ προκύπτουν τα ίδια συμπεράσματα με την ανάλυση της περίπτωσης των S διαγραμμάτων ελέγχου με όρια ελέγχου 3σ . Έτσι, καθώς ο αριθμός των δειγμάτων m αυξάνεται στη Φάση I, η εμπειρική συνάρτηση κατανομής προσεγγίζει τη γεωμετρική κατανομή. Επιπλέον, μια αυξανόμενη τιμή του n συμβάλλει στην γρηγορότερη προσέγγιση της γεωμετρικής κατανομής για συγκεκριμένες τιμές του m .

3.5 X διάγραμμα ελέγχου με εκτιμώμενο σ

Οι Maravelakis et al. (2002) πρότειναν τη χρήση τα διαγραμμάτων ελέγχου για μεμονωμένες παρατηρήσεις για την παρακολούθηση των μετατοπίσεων της διασποράς της διεργασίας. Συγκεκριμένα χρησιμοποίησαν το X διάγραμμα ελέγχου, και όχι το διαγράμμα ελέγχου κινούμενου εύρους (MR), καθώς το τελευταίο δεν συνεισφέρει σημαντικά στην αναγνώριση των εκτός ελέγχου διεργασιών, γεγονός που συμφωνεί και με τις προτάσεις διαφόρων ερευνητών όπως Rigdon et al. (1994) και Sullivan and Woodall (1996).

Όταν η τυπική απόκλιση σ είναι εκ των προτέρων γνωστή (Φάση II), τότε τα όρια ελέγχου είναι γνωστά (δείτε Παράγραφο 1.6).

Στην περίπτωση που η τυπική απόκλιση σ είναι άγνωστη, τότε τα όρια ελέγχου που χρησιμοποιούνται για την παρακολούθηση της διασποράς της διεργασίας στη Φάση I είναι τα

$$UCL = \bar{X} + 3\sigma$$

$$CL = \bar{X}$$

$$LCL = \bar{X} - 3\sigma$$

Συνήθως η ποσότητα που χρησιμοποιείται ως εκτίμηση της τυπικής απόκλισης είναι η αμερόληπτη εκτιμήτρια \overline{MR}/d_2 . Εντούτοις, οι Cryen & Ryan (1990) θεώρησαν ότι η ποσότητα που προτιμάται για την εκτίμηση του σ είναι η αμερόληπτη εκτιμήτρια S/c_4 .

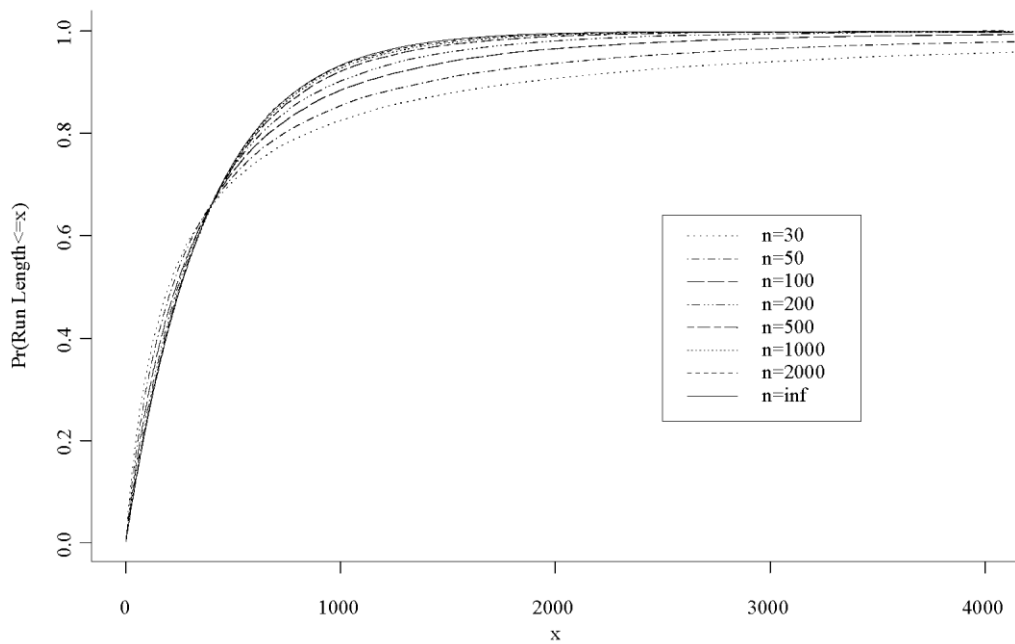
Προκειμένου να μελετήσουν την επίδραση της εκτίμησης της τυπικής απόκλισης σ στο X διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις, οι Maravelakis et al. πραγματοποίησαν μελέτη προσομοίωσης για διάφορες τιμές του αριθμού των παρατηρήσεων N . Ειδικότερα, προσομοιώθηκαν N τιμές από μια κανονική κατανομή $N(\mu, \sigma_0^2)$ και υπολογίστηκαν τα όρια ελέγχου UCL και LCL του X διαγράμματος ελέγχου. Στην συνέχεια χρησιμοποιήθηκε πάλι μέθοδος προσομοίωσης για να πάρουμε τιμές από μια κανονική κατανομή $N(\mu, \sigma_1^2)$ έως ότου ληφθεί μια τιμή πάνω από το UCL ή κάτω από το LCL . Η διαδικασία αυτή επαναλήφθηκε 32000 φορές ώστε να ληφθούν εκτιμήσεις για τις τιμές των ARL και $SDRL$. Τα αποτελέσματα των εκτιμήσεων παρουσιάζονται στον παρακάτω πίνακα. Αξίζει να επισημανθεί στο σημείο αυτό ότι δεν εξετάστηκαν μειώσεις της τυπικής απόκλισης παρά μόνο αυξήσεις, επειδή οι μετατοπίσεις της διασποράς προς τα κάτω δεν είναι δυνατόν να ανιχνευτούν σε ένα διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις, κάτι που θα δούμε στην συνέχεια.

		σ_1^2 / σ_0^2									
		1		1.2		1.4		1.6		1.8	
N	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	ARL	$SDRL$	
30	986.31	5024.83	315.36	1058.44	147.93	439.79	84.36	187.50	53.74	98.54	
50	614.94	1565.00	229.95	476.60	116.69	200.50	69.61	107.23	47.23	66.81	
75	503.75	948.78	202.02	318.54	105.18	150.77	64.51	84.15	43.99	54.87	
100	467.07	770.60	190.53	274.54	100.73	131.39	61.98	75.26	42.78	50.48	
200	413.88	518.65	173.68	205.96	93.86	105.77	58.63	63.56	40.67	42.81	
300	398.94	476.34	167.79	187.69	92.76	100.47	57.93	61.37	41.26	42.29	
500	387.38	429.45	167.90	179.39	90.34	93.58	56.80	58.96	39.69	40.54	
1000	379.32	401.55	162.96	168.50	89.12	91.10	57.03	57.78	39.90	39.85	
2000	372.64	383.71	162.70	166.87	89.45	89.41	56.35	55.82	39.62	39.17	
∞	370.40	369.90	162.08	161.58	89.05	88.55	56.48	55.98	39.45	38.95	

Πίνακας 4.1 Τιμές των ARL και $SDRL$ για το X διάγραμμα ελέγχου για μεμονωμένες παρατηρήσεις για διάφορες τιμές του N

Από τον Πίνακα 4.1 προκύπτει ότι δεν δίνονται αποτελέσματα στις τιμές των ARL και $SDRL$ για μετατοπίσεις της διασποράς προς τα κάτω. Αυτό συμβαίνει επειδή μια φθίνουσα τυπική απόκλιση δεν θα μπορούσε να προκαλέσει ποτέ μια τιμή η οποία να βρίσκεται κάτω από το κατώτερο όριο ελέγχου LCL . Παρατηρούμε επίσης ότι οι τιμές του μέσου μήκους ροής και της τυπικής του απόκλισης γενικά μειώνονται μέχρι να πλησιάσουν τις θεωρητικές τους τιμές. Ένα επιπλέον σημαντικό συμπέρασμα για το μέγεθος δείγματος που πρέπει να τονιστεί είναι ότι απαιτούνται τουλάχιστον 300 παρατηρήσεις για να ελαχιστοποιηθεί η επίδραση της εκτίμησης της τυπικής απόκλισης στα X διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις.

Στην συνέχεια παρουσιάζουμε μια γραφική παράσταση στην οποία απεικονίζεται η εμπειρική συνάρτηση κατανομής του μήκους ροής του X διαγράμματος ελέγχου για μεμονωμένες παρατηρήσεις σε σύγκριση με τη γεωμετρική κατανομή (inf) για $n = 30, 50, 100, 200, 500, 1000$ και 2000.



Διάγραμμα 4.1 Συνάρτηση κατανομής του μήκους ροής του X διαγράμματος ελέγχου για διάφορες τιμές του n

Από την γραφική παράσταση διαπιστώνουμε το γεγονός ότι καθώς το n αυξάνει, η εμπειρική συνάρτηση κατανομής προσεγγίζει την θεωρητική της κατανομή που είναι η γεωμετρική.

ΚΕΦΑΛΑΙΟ 4

4.1 Συμπεράσματα – Προτάσεις

Ανακεφαλαιώνοντας, στην παρούσα διπλωματική εργασία σκοπός ήταν η μελέτη της επίδρασης της εκτίμησης των παραμέτρων στα μονομεταβλητά διαγράμματα ελέγχου τόσο για τη μέση τιμή όσο και για τη διασπορά. Σημαντικό ρόλο στην εξέταση του προβλήματος των άγνωστων παραμέτρων της Φάσης I έπαιξε αναμφισβήτητα η κατανομή του μήκους ροής καθώς και οι τιμές που την συνοδεύουν οι οποίες ήταν κυρίως το μέσο μήκος ροής και η τυπική απόκλιση του μήκους ροής συγκριτικά με τις αντίστοιχες τιμές των γνωστών παραμέτρων της Φάσης II. Πάνω σε αυτό το πλαίσιο μελετήθηκε κυρίως μέσω μεθόδων προσομοίωσης η απόδοση των διαγραμμάτων ελέγχου με άγνωστες παραμέτρους και συγκρίθηκε με την απόδοση των διαγραμμάτων με γνωστές παραμέτρους όπου προέκυψαν σημαντικά συμπεράσματα τόσο για την επιλογή της κατάλληλης εκτιμήτριας αλλά και όσο για την επιλογή του μεγέθους και του αριθμού των δειγμάτων, η λήψη των οποίων είναι απαραίτητη στη Φάση I για την ελαχιστοποίηση της επίδρασης της εκτίμησης των παραμέτρων.

Συμπερασματικά, για την επίτευξη επαρκούς απόδοσης συγκρινόμενης με αυτήν των γνωστών παραμέτρων είναι απαραίτητο να λαμβάνονται περισσότερα δεδομένα στη Φάση I και γενικά όσες περισσότερες παράμετροι εκτιμώνται τόσο μεγαλύτερα μεγέθη δειγμάτων απαιτούνται. Για παράδειγμα στα πολυμεταβλητά διάγραμμα απαιτούνται περισσότερα δεδομένα στη Φάση I σε σύγκριση με τα μονομεταβλητά διαγράμματα ελέγχου.

Έτσι ο Chen καθιστά αναγκαία τη λήψη τουλάχιστον 75 δειγμάτων στη Φάση I με μέγεθος δείγματος τουλάχιστον 5 στα διαγράμματα ελέγχου για την διασπορά (R , S και S^2). Όσον αφορά τα \bar{X} διαγράμματα ελέγχου για τη μέση τιμή, συμπεραίνει ότι η επίδραση της εκτίμησης των παραμέτρων είναι πιο σοβαρή για μικρές μετατοπίσεις στη διεργασία παρά για μεγαλύτερες και επιπλέον ότι η επίδραση είναι μεγαλύτερη στις τιμές του $SDRL$ συγκριτικά με αυτές του ARL .

Αντίθετα, οι Maravelakis et al. (2002) πρότειναν μεγαλύτερα μεγέθη δειγμάτων στα διαγράμματα ελέγχου για τη διασπορά, δηλαδή 100 δείγματα μεγέθους τουλάχιστον 20 καθώς για μικρότερα δείγματα θεώρησαν ότι η επίδραση είναι αρκετά σοβαρή με αποτέλεσμα την πρόκληση εσφαλμένων συναγερμών στη διεργασία, ενώ στα διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις θεώρησαν ότι απαιτούνται τουλάχιστον 300 παρατηρήσεις. Οι προτάσεις αυτές

έρχονται σε αντίθεση φυσικά με τον εμπειρικό κανόνα για την επιλογή του αριθμού των δειγμάτων που είναι $m = 20$ ή $m = 30$ με μεγέθη $n = 4$ ή $n = 5$.

Επιπλέον, για την ανίχνευση μικρών μετατοπίσεων της διεργασίας προτείνεται να χρησιμοποιούνται διαγράμματα περισσότερο ευαίσθητα από τα διαγράμματα ελέγχου *Shewhart* όπως είναι τα διαγράμματα *CUSUM* ή *EWMA*.

Οι ερευνητικές αναζητήσεις στην περιοχή της επίδρασης της εκτίμησης των παραμέτρων στα διαγράμματα ελέγχου δεν εξαντλούνται στο πλαίσιο της παρούσας μεταπτυχιακής εργασίας. Ιδιαίτερο ενδιαφέρον παρουσιάζει η μελέτη της περιθώριας κατανομής του μήκους ροής καθώς στη βιβλιογραφία αρκετοί συγγραφείς ενώ ασχολήθηκαν με το συγκεκριμένο θέμα, δεν έκαναν προτάσεις για την επιλογή των δειγμάτων, επειδή δεν εξέτασαν την περιθώρια κατανομή. Μια ακόμα ερευνητική περιοχή η οποία χρήζει περισσότερης μελέτης από τους ερευνητές αφορά στους ανθεκτικούς εκτιμητές και στη χρήση τους σε έκτροπες μετρήσεις στη Φάση I (δείτε για παράδειγμα *Rocke (1989, 1992), Tatum (1997), Vargas (2003), Davis and Adams (2005)*).

Επιπρόσθετα, οι κανόνες ευαισθητοποίησης ενός διαγράμματος ελέγχου *Shewhart* με εκτιμώμενες παραμέτρους που έχουν χρησιμοποιηθεί από ερευνητές είναι η αιτία για την οποία το διάγραμμα δεν αποδίδει το ίδιο ομοιόμορφα όπως στην περίπτωση των γνωστών παραμέτρων. Εξάλλου, η χρήση πολλών κανόνων ταυτοχρόνως καθιστά δύσκολη την ανάλυση των ιδιοτήτων των διαγραμμάτων ελέγχου με εκτιμώμενες παράμετρους, γι' αυτόν τον λόγο απαιτείται περαιτέρω μελέτη πάνω στο θέμα αυτό (δείτε *Champ and Woodall (1987), Burroughs et al. (1993)*). Επίσης, περισσότερη έμφαση έχει δοθεί πάνω σε διαγράμματα για την παρακολούθηση του μέσου παρά σε διαγράμματα για παρακολούθηση της διασποράς, αλλά αξίζει να σημειωθεί και το γεγονός ότι τα πολυμεταβλητά διαγράμματα ελέγχου για την παρακολούθηση της διασποράς δεν έχουν μελετηθεί επαρκώς.

Τέλος, οι περισσότεροι συγγραφείς συμφωνούν στο γεγονός ότι είναι αναγκαίο να λαμβάνονται όσα το δυνατόν περισσότερα δείγματα στη Φάση I τα οποία εγγυώνται ικανοποιητική απόδοση στη Φάση II. Σε περίπτωση που τα δεδομένα στη Φάση I δεν είναι επαρκή, προτείνονται είτε διορθώσεις στα όρια ελέγχου είτε ανανεώσεις αυτών όταν περισσότερα δείγματα γίνονται διαθέσιμα (δείτε επίσης *Neduraman and Pignatiello (2001), Tsai et al. (2004,2005), Albers and Kallenberg (2004a, 2004c)*).

Επίσης, ο αναγνώστης μπορεί να ανατρέξει στην πιο πρόσφατη βιβλιογραφία (δείτε Castagliola, Celano and Chen (2009), Zhang, Chen and Castagliola (2009), Zhang and Castagliola (2010), Zhang, Castagliola and Wu (2011), Maravelakis and Castagliola (2009,2011)).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΡΑΙΑ

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

Αντζουλάκος Δημήτριος (2007). Στατιστικός Έλεγχος Ποιότητας, Σημειώσεις παραδόσεων του μαθήματος του Στατιστικού Ελέγχου Ποιότητας στα πλαίσια του Π.Μ.Σ στην Εφαρμοσμένη Στατιστική, Πανεπιστήμιο Πειραιώς

Γραφανάκης Δημήτριος (2000). Στατιστικός Έλεγχος Ποιότητας, ΕΑΠ, Πάτρα.

Δαμιανού Χαράλαμπος (1996). Στατιστικός Έλεγχος Ποιότητας και Αξιοπιστία, Εκδόσεις Συμμετρία.

Ταγαράς Γιώργος (2001). Στατιστικός Έλεγχος Ποιότητας, Εκδόσεις Ζήτη.

Ξένη

Albers, W. and Kallenberg, W.C.M. (2004a). Are Estimated Control Charts in Control? , Statistics 38, pp.67-79.

Aroian L.A. and Levene, H. (1950). The effectiveness of Quality Control Charts, Journal of the American Statistical Association 45, pp 520-529.

Burroughs, T. E., Rigdon, S.E., Champ, C.W. (1993). An analysis of Shewhart Charts with Runs Rules When No Standards Are Given, Proceedings of the Quality and Productivity Section of the American Statistical Association, August 8-12, San Francisco, CA, pp. 16-19.

Castagliola, P., Maravelakis P. (2011). A CUSUM Control Chart for Monitoring the Variance when Parameters are Estimated, Journal of Statistical Planning and Inference 141(4), pp. 1463-1478.

Castagliola, P., Celano G., Chen G. (2009). The Exact Run Length Distribution and Design of the S^2 Chart when the In-Control Variance is Estimated, International Journal of Reliability, Quality and Safety Engineering 16(1), pp. 23-28.

Chakraborti, S. (2000). Run Length, Average Run Length and False Alarm Rate of Shewhart \bar{X} Chart: Exact Derivations by Conditioning, Communications in Statistics-Simulation and Computation 29, pp. 61-81.

Chakraborti, S. (2007). Run length distribution and percentiles: The Shewhart \bar{X} chart with unknown parameters, Quality Engineering 19, pp 119-127.

Champ, C.W. and Woodall, W.H. (1987). Exact Results for Shewhart Control Charts with Supplementary Runs Rules, *Technometrics* 29, pp. 393-399.

Chen, G. (1997). The mean and Standard Deviation of the Run Length Distribution of \bar{X} Charts When Control Limits Are Estimated, *Statistica Sinica* 7, pp.789-798.

Chen, G. (1998). The run length distributions of the R, S and S^2 Control Charts when σ Is Estimated, *Canadian Journal of Statistics* 26, pp.311-322.

Chiu, W.K and Cheung, K.C. (1977). An economic study of \bar{X} charts with warning limits, *Journal of Quality Technology* 9, pp. 166-171.

Cryer, J. D. and Ryan, T. P. (1990). The Estimation of Sigma for an X Chart: \overline{MR}/d_2 or s/c_4 ? *Journal of Quality Technology* 1990, 22, pp. 187-192.

Duncan, A.J. (1956). The economic design of \bar{X} charts used to maintain current control of a process, *Journal of American Statistical Association* 51, pp. 228-242.

Davis, C.M. and Adams, B.M. (2005). Robust Monitoring of Contaminated Data, *Journal of Quality Technology* 37, pp. 163-174.

Del Castillo, E. (1996a). Run Length distributions and Economic Design of \bar{X} charts with Unknown Process Variance, *Metrika* 43, pp. 189-201.

Del Castillo, E. (1996c). Evaluation of \bar{X} charts with unknown variance, *Journal of Quality Technology* 28, pp. 116-122.

Eugene, L. Grant & Richard S. Leavenworth (1999), *Statistical Quality Control*, McGraw Hill.

Gosh, B.K., Reynolds, M.R., JR., and Hui, Y. V. (1981). Shewhart Charts with Estimated Process Variance, *Communications in Statistics-Theory and Methods* 18, pp. 1797-1822.

Hawkins, D.M. (1987). Self-Starting CUSUM Charts for Location and Scale, *The Statistician* 36, pp. 299-315.

Hillier, F. (1969). \bar{X} and R Chart Control Limits Based on a Small Number of Subgroups, *Journal of Quality Technology* 1, pp. 17-26.

Jensen, Willis A., Jones-Farmer, L. Allison, Champ, Charles W and Woodall, William H. (2006). Effects of Parameter Estimation on Control Chart Properties: A literature Review, *Journal of Quality Technology* 2006 Vol. 38, No. 4, 2006, pp. 349-364.

Jones, L.A., Champ, C.W. and Rigdon, S.E. (2001). The Performance of Exponentially Weighted Moving Average with Estimated Charts, *Technometrics* 43, pp 156-157.

- Jones, L.A., Champ, C. W. and Rigdon, S.E. (2004). The Run Length Distribution of the CUSUM with Estimated Parameters, *Journal of Quality Technology* 36, pp. 95-108.
- Li, H., Owen, D.B. and Borrego (1990). Lower Confidence Limits on Process Capability Indices Based on Range, *Communications in Statistics – Simulation and Computation* 19(1), pp. 1-24.
- Lowry, C.A., Champ, C.W. and Woodall, W.H. (1995). The performance of Control Charts for Monitoring Process Variation, *Communications in Statistics-Simulation and Computation* 24(2), pp. 409-437.
- Maravelakis, P.E., Panaretos, J., and Psarakis, S. (2002). Effect of Estimation of the Process Parameters on the Control Limits of the Univariate Control Charts for Process Dispersion, *Communication in Statistics-Simulation and Computation* 31, pp. 443-461.
- Maravelakis, P.E., Castagliola P. (2009). An EWMA Chart for Monitoring the Process Standard Deviation when Parameters are Estimated, *Computational Statistics & Data Analysis* 53(7), pp. 2653-2664.
- Montgomery, D.C. (2005). *Introduction to Statistical Quality Control*, 5th Ed. John Wiley.
- Nedumaran, G. and Pignatiello, J.J., JR. (2001). On Estimating \bar{X} Control Chart Limits, *Journal of Quality Technology* 33, pp. 206-212.
- Ng, C.H. and Case, K.E. (1992). Control Limits and the ARL: Some surprises, *First Industrial Engineering Research Conference Proceedings*, pp. 127-129.
- Patnaik, P.B. (1950). The use of mean range as an estimator of variance in statistical tests. *Biometrika* 37, pp. 78-87.
- Prochan, F. and Savage, I.R. (1960). Starting a control chart. *Industrial quality control* 17, pp 12-13.
- Quesenberry, C.P. (1991). SPC Q charts for start-up processes and short or long runs, *Quality Technology* 17, pp. 213-214.
- Quesenberry, C.P. (1993). The effect of sample size on estimated limits for \bar{X} and X control charts, *Journal of Quality Technology* 25, pp. 237-247.
- Quesenberry, C. P. (1995). On Properties of Q charts for variables, *Journal of Quality Technology* (27(3)), pp. 184-203.
- Rigdon, S.E., Cruthis, E.N., and Champ, C.W. (1994). Designs Strategies for Individuals and Moving Range Control Charts, *Journal of Quality Technology* 1994, 26, pp. 274-287.
- Rocke, D.M. (1989). Robust Control Charts. *Technometrics* 31, pp. 173-184

- Rocke, D.M. (1992). \overline{X}_Q and R_Q Charts: Robust Control Charts, *The Statistician* 41, pp. 97-104.
- Shewhart, W. A. (1939). *Statistical Method from the Viewpoint of Quality Control*, Dover Publications, New York, NY.
- Shu, L., Tsung F. and Tsui K.L. (2004). Run Length Performance of Regression Control Charts with Estimated Parameters, *Journal of Quality Technology* 36, pp. 280-292.
- Sullivan, J.H. and Woodall, W.H. (1996). A Control Chart for Preliminary Analysis of Individual Observations, *Journal of Quality Technology* 28, pp. 265-278.
- Tatum, L.G. (1997). Robust Estimation of the Process Standard Deviation for Control Charts, *Technometrics* 39, pp. 127-141.
- Thomas P. Ryan (2000). *Statistical Methods for Quality Improvement*, Second Edition, John Wiley & Sons, Inc.
- Tsai, T.R., Wu, S.J. and Lin, H.C. (2004). An Alternative Control Chart Approach Based on Small Number of Subgroups, *International Journal of Advanced Manufacturing Technology* 26, pp. 1312-1316.
- Tsai, T.R., Lin, J.J., Wu, S. J. and Lin, H.C. (2005). On estimating Control Limits of \overline{X} Chart When the Number of Subgroups is Small, *International Journal of Information and Management Sciences* 15, pp. 61-73.
- Vargas, J.A. (2003). Robust Estimation in Multivariate Control Charts for Individual Observations, *Journal of Quality Technology* 35, pp. 367-376.
- Wu, C., Zhao, Y. and Wang Z. (2002). The Median Absolute Deviations and Their Application to Shewhart \overline{X} Control Charts, *Communication in Statistics-Simulation and Computation* 31, pp. 425-442.
- Zhang, L. and Castagliola P., (2010). Run Rules Xbar Charts when Process Parameters are Unknown, *International Journal of Reliability, Quality and Safety Engineering* 17(4), pp. 381-399.
- Zhang, L., Castagliola P., Wu Z. and M.B.C. Khoo (2011). The synthetic Xbar Chart with Estimated Parameters, *IIE Transactions* 43(9), pp. 676-687.
- Zhang, L., Chen G. and Castagliola P. (2009). On t and EWMA t Charts for Monitoring Changes in the Process Mean, *Quality and Reliability Engineering International* 25(8), pp. 933-945.

РАНЕЕЗНАМО ПЕРПАА

РАНЕЕ НЕ ПЕРПА