



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Διαμεταπτυχιακό - Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών:

Μαθηματική Προτυποποίηση στις Σύγχρονες Τεχνολογίες και την Οικονομία

Μεταπτυχιακή Διπλωματική Εργασία

Παραμετρικά και Μη Παραμετρικά Διαγράμματα Ελέγχου και Εφαρμογές

Ανδρέας Κτενίδης

Επιβλέπων: Κουκουβίνος Χρήστος, Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2011

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	3
Abstract.....	4
Ευχαριστίες.....	5

ΚΕΦΑΛΑΙΟ 1

1.1 Εισαγωγή.....	6
1.1.1 Η Έννοια της Ποιότητας και οι Διαστάσεις της	6
1.1.2 Ιστορική Αναδρομή της Ποιότητας.....	7
1.2 Συνιστώσες του Στατιστικού Ελέγχου Ποιότητας.....	11
1.3 Τα 7 Κυριότερα Εργαλεία του Στατιστικού Ελέγχου Διεργασιών.....	12
1.4 Περιγραφή και Χρήση ενός Διαγράμματος Ελέγχου.....	13
1.5 Τύποι Διαγραμμάτων Ελέγχου.....	17
1.6 Κατασκευή ενός Διαγράμματος Ελέγχου.....	17
1.6.1 Προειδοποιητικά Όρια Ελέγχου και Μέγεθος Δείγματος.....	19
1.7 Μέτρα Απόδοσης ενός Διαγράμματος Ελέγχου.....	19
1.8 Άλλα Μέτρα Απόδοσης.....	21
1.9 Ταξινόμηση Διαγραμμάτων Ελέγχου.....	22
1.10 Βασικά Είδη Διαγραμμάτων Ελέγχου.....	23
1.10.1 Διαγράμματα Ελέγχου Τύπου Shewhart.....	24
1.10.2 Διαγράμματα Ελέγχου Τύπου CUSUM.....	25
1.10.3 Διαγράμματα Ελέγχου Τύπου EWMA.....	26
1.11 Διαγράμματα Ελέγχου Ιδιοτήτων.....	27
1.11.1 Διαγράμματα Ελέγχου p και np	27
1.11.2 Διαγράμματα Ελέγχου c	28
1.11.3 Διάγραμμα Ελέγχου u	29

ΚΕΦΑΛΑΙΟ 2

2.1 Εισαγωγή.....	30
2.2 Σημαντικά Μη Παραμετρικά Διαγράμματα Ελέγχου.....	30
2.2.1 Η επίδραση της μη κανονικότητας στα \bar{X} – διαγράμματα.....	32
2.2.2 Προσημικά διαγράμματα ελέγχου για την καταγραφή και τη παρακολούθηση του μέσου.....	34
2.2.3 Ένα προσημικό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση της διακύμανσης μιας διεργασίας.....	37
2.2.4 Σχέδια συντόμευσης δειγματοληψίας.....	39
2.3 Προσημικό διάγραμμα CUSUM για την παρακολούθηση του μέσου.....	41
2.4 Άλλες εκδοχές μη παραμετρικών διαγραμμάτων ελέγχου.....	45
2.5 Συγκρίσεις διαγραμμάτων ελέγχου.....	46
2.6 Παρατηρήσεις και συμπεράσματα.....	51

ΚΕΦΑΛΑΙΟ 3

3.1 Σημαντικά μη παραμετρικά διαγράμματα ελέγχου με χρήση μιας εμπειρικής κατανομής αναφοράς.....	53
3.2 Μη παραμετρικά διαγράμματα ελέγχου με χρήση της διαμέσου.....	59
3.3 Ένα νέο μη παραμετρικό διάγραμμα ελέγχου.....	63
3.3.1 Μια παραλαγή του νέου διαγράμματος ελέγχου.....	71
3.3.2 Αριθμητικές συγκρίσεις του νέου διαγράμματος ελέγχου.....	72
3.4 Μη παραμετρικά διαγράμματα ελέγχου με χρήση ρών και βαθμολογικών συναρτήσεων τύπου Wilcoxon.....	76
3.5 Νέα μη-παραμετρικά διαγράμματα ελέγχου.....	77
3.5.1 Αριθμητικά αποτελέσματα των νέων διαγραμμάτων ελέγχου.....	82

Εφαρμογές

A. Παραμετρικά και μη-παραμετρικά διαστήματα πρόβλεψης βασισμένα κατά την φάση II των διαγραμμάτων ελέγχου των επαναλαμβανόμενων δεδομένων βιολογικής ανάλυσης	
1. Το πρόβλημα.....	89
2. Εξετάζοντας τη συμπεριφορά των ιστορικών τρεξιμάτων.....	91
3. Τα διαγράμματα ελέγχου μιας χρονικής στιγμής δεδομένων.....	92
3.1. Φάση I παραμετρικά διαγράμματα Shewhart.....	92
3.2. Φάση II παραμετρικού διαγράμματος ελέγχου που χρησιμοποιεί ένα διάστημα πρόβλεψης	93
3.3. Φάση II των μη-παραμετρικών διαγραμμάτων.....	97
4. Μια περιπτωσιολογική μελέτη.....	97
5. Μια εναλλακτική αμφίπλευρη σύγκριση.....	100
6. Επίλογος.....	101
B. Η χρήση της μεθόδου του διαγράμματος ποιότητας ελέγχου για την ανάλυση της μεταβλητότητας του PEF του επαγγελματικού άσθματος	
1. Εισαγωγή.....	102
2. Εργαλεία και μέθοδοι.....	103
3. Αποτελέσματα.....	107
4. Παρατηρήσεις.....	108
5. Συμπεράσματα.....	110
Βιβλιογραφία.....	111

Περίληψη

Σκοπός της μεταπτυχιακής εργασίας είναι η παρουσίαση των παραμετρικών και μη παραμετρικών διαγραμμάτων ελέγχου καθώς και η παρουσίαση κάποιων εφαρμογών.

Στην εποχή μας η ποιότητα των προϊόντων είναι σίγουρα ένας πολύ σημαντικός παράγοντας που επηρεάζει την επιλογή μας. Αυτό το κριτήριο είναι που οδήγησε στον Στατιστικό Έλεγχο Διεργασιών (ΣΕΔ), μέσω των εργαλείων του οποίου παρέχεται η δυνατότητα προσδιορισμού της ικανότητας μιας διαδικασίας να λειτουργεί και να παράγει σύμφωνα με τις απαιτούμενες προδιαγραφές, δηλαδή εκείνες που ικανοποιούν τους καταναλωτές ώστε να εμπιστεύονται το προϊόν. Ένα από τα εργαλεία αυτά είναι και το διάγραμμα ελέγχου το οποίο είναι ένα εύχρηστο και ισχυρό εργαλείο για παρακολούθηση σε πραγματικό χρόνο της ποιότητας παραγωγής καθώς και της πορείας παραγωγικών διεργασιών.

Στη παρούσα εργασία περιλαμβάνονται τρία κεφάλαια.

- Στο πρώτο κεφάλαιο παρουσιάζονται τα παραμετρικά διαγράμματα ελέγχου.
- Στο δεύτερο κεφάλαιο παρουσιάζονται τα μη-παραμετρικά διαγράμματα ελέγχου.
- Τέλος, στο τρίτο κεφάλαιο παρουσιάζονται δυο εφαρμογές ιατρικής και βιολογικής φύσεως, οι οποίες χρησιμοποιούν τα παραμετρικά και μη-παραμετρικά διαγράμματα ελέγχου για την ανάλυση των δεδομένων.

Abstract

The aim of this postgraduate work is the presentation of the parametric and non-parametric control charts, as well as the presentation of some medical and biological applications.

In our days, the quality of products is a very important factor that influences our choices. This criterion led to the creation of the Statistical Control of Activities. Several Statistical Control tools provide the possibility to determine the faculty of the operation and the production of a process, according to the required specifications which should satisfy the consumers, so that they entrust the product. One of these tools, is the control chart which is easy to use, and it is a powerful tool for following-up in real time the quality of production, as well as the progress of productive activities.

In the present work includes two chapters and a unit including the description of two applications.

- In the first chapter the parametric control charts are presented.
- In the second chapter the non-parametric control charts are presented.
- Finally, a unit which describes two applications, a medical and a biological one, is presented using parametric and non-parametric control charts.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κ. Χρήστο Κουκουβίνο, Καθηγητή του Ε.Μ.Π. και επιβλέποντα της διπλωματικής εργασίας μου, αφενός μεν για την διδασκαλία του μαθήματος και αφετέρου για την καθοδήγηση του προς την επιλογή του θέματος και τελικώς την περαίωση της εργασίας.

Ιδιαίτερες ευχαριστίες θέλω να εκφράσω και στην υποψήφια διδάκτορα Παρπούλα Χριστίνα για τον χρόνο και την διάθεση που κατέβαλε για την επιτυχή εκπόνηση της παρούσας εργασίας.

Κεφάλαιο 1

1.1 Εισαγωγή

Κάθε άνθρωπος που αποφασίζει να αγοράσει ένα προϊόν (ή υπηρεσία) αντιμετωπίζει το πρόβλημα της επιλογής ανάμεσα σε ομοειδή προϊόντα που ικανοποιούν την ίδια ανάγκη. Η επιλογή του προϊόντος από τον καταναλωτή καθορίζεται κυρίως από δύο παράγοντες, την ποιότητά του και την τιμή του. Η αναγνώριση αυτών των δύο κύριων παραγόντων οι οποίοι καθορίζουν τις αγοραστικές συνήθειες των καταναλωτών οδήγησε τις επιχειρήσεις στην αναζήτηση επιχειρηματικών τακτικών που θα οδηγήσουν στη βέλτιστη δυνατή ποιότητα των παραγόμενων προϊόντων με το ελάχιστο δυνατό κόστος.

Η αναζήτηση αυτή οδήγησε σε μια σειρά παρεμβάσεων στο επιχειρηματικό περιβάλλον οι οποίες τελικά διαμόρφωσαν νέα πλαίσια λειτουργίας και διοίκησης των επιχειρήσεων με τελικό στόχο και σκοπό τη διαρκή βελτίωση της ποιότητας. Η στρατηγική / φιλοσοφία που επιζητεί τη συνεχή βελτίωση στην ποιότητα εκτέλεσης όλων των διεργασιών, προϊόντων και υπηρεσιών, σε μια επιχείρηση ονομάζεται Διοίκηση Ολικής Ποιότητας (*Total Quality Management*).

Σύμφωνα με αυτή τη φιλοσοφία “η διοίκηση πρέπει να κατευθύνει και το εργατικό δυναμικό να ενεργεί με στόχο τη βελτίωση της ποιότητας σε κάθε επίπεδο λειτουργίας της επιχείρησης”. Σημαντικό ρόλο στη διαμόρφωση των νέων πλαισίων λειτουργίας και διοίκησης των επιχειρήσεων διαδραματίζει ο

Διεθνής Οργανισμός Τυποποίησης (*ISO, International Standards Organization*) ο οποίος με τα πρότυπα ISO καθορίζει τις διαδικασίες που πρέπει να υιοθετηθούν από μια επιχείρηση με τελικό σκοπό τη συνεχή προαγωγή και διασφάλιση της ποιότητας.

Σημαντικό ρόλο στις διαδικασίες αυτές κατέχει ο Στατιστικός Έλεγχος Ποιότητας (*Statistical Quality Control*). Ο Στατιστικός Έλεγχος Ποιότητας αποτελεί την παλαιότερη και γνωστότερη μέθοδο ελέγχου παραγωγικών διεργασιών για τη βελτίωση της ποιότητας των παραγόμενων προϊόντων. Ένας από τους βασικούς στόχους του είναι η έγκαιρη ανακάλυψη μη συμμορφωμένων με τις προδιαγραφές παραγόμενων προϊόντων η οποία σηματοδοτεί τη λήψη διορθωτικών ενεργειών για την απομάκρυνση των αιτιών που είναι υπεύθυνες για τις αποκλίσεις, συμβάλλοντας έτσι στη διατήρηση της ποιότητας των προϊόντων. Θα μπορούσαμε να πούμε ότι ο Στατιστικός Έλεγχος Ποιότητας επηρεάζει σημαντικά αποφάσεις σχετικές με τις προδιαγραφές, την παραγωγή και τον έλεγχο των παραγόμενων προϊόντων μιας επιχείρησης.

1.1.1 Η έννοια της ποιότητας και οι διαστάσεις της

Η έννοια της ποιότητας μπορεί να οριστεί με πολλούς τρόπους. Οι περισσότεροι άνθρωποι έχουν μια διαισθητική άποψη για το τι είναι ποιότητα και τι συνδέουν με τα επιθυμητά χαρακτηριστικά / διαστάσεις που πρέπει να έχει ένα προϊόν (ή μια υπηρεσία). Οι διαστάσεις αυτές, βάσει των οποίων ο καταναλωτής αξιολογεί τα

προϊόντα, καθορίζουν την τελική ποιότητα ενός προϊόντος και σχετίζονται με την ποιότητα του σχεδιασμού και τη διαδικασία παραγωγής του προϊόντος. Οι βασικές διαστάσεις της ποιότητας παρουσιάζονται αναλυτικά στον ακόλουθο πίνακα

Πίνακας 1.1. Οι Διαστάσεις της ποιότητας

ΑΠΟΔΟΣΗ	Το προϊόν κάνει την δουλειά για την οποία προορίζεται Κάνει τη δουλειά καλύτερα από άλλα ομοειδή προϊόντα;
ΑΞΙΟΠΙΣΤΙΑ	Το προϊόν χρειάζεται συχνή επισκευή;
ΔΙΑΡΚΕΙΑ	Η διάρκεια ζωής του προϊόντος είναι μεγάλη;
ΕΠΙΣΚΕΥΗ	Πόσο γρήγορη και οικονομική είναι η επισκευή του προϊόντος στην περίπτωση που εμφανιστεί βλάβη;
ΑΙΣΘΗΤΙΚΗ	Πόσο ικανοποιητικό είναι από άποψη εμφάνισης (χρώμα, σχήμα, περιτύλιγμα, κτλ.) το προϊόν;
ΔΥΝΑΤΟΤΗΤΕΣ	Ποιες είναι οι επιπρόσθετες δυνατότητες του προϊόντος;
ΦΗΜΗ ΕΠΙΧΕΙΡΗΣΗΣ	Ποια είναι η φήμη της εταιρείας, καλή ή κακή;
ΣΥΜΜΟΡΦΩΣΗ ΜΕ ΤΙΣ ΠΡΟΔΙΑΓΡΑΦΕΣ	Το προϊόν κατασκευάστηκε σύμφωνα με τις προδιαγραφές που έθεσε ο σχεδιαστής του;

Σύμφωνα με τα όσα αναφέρθηκαν για τις διαστάσεις της ποιότητας μπορούμε να ορίσουμε ότι ποιότητα είναι η προσαρμογή των χαρακτηριστικών ενός προϊόντος (ή υπηρεσίας) στις απαιτήσεις του καταναλωτή / χρήστη, δηλαδή ότι ποιότητα σημαίνει καταλληλότητα προς χρήση (παραδοσιακός ορισμός της ποιότητας, *Joseph Juran, 1974*). Έτσι η ποιότητα μπορεί να “μετρηθεί” μέσω της ικανοποίησης που προσφέρει το προϊόν στον καταναλωτή. Δυστυχώς ο παραπάνω ορισμός έχει συσχετιστεί περισσότερο με την έννοια της συμμόρφωσης των παραγόμενων προϊόντων με τις προδιαγραφές που πρέπει να πληρούν τα παραγόμενα προϊόντα, και λιγότερο με την έννοια της ποιότητας του σχεδιασμού τους.

Για παράδειγμα όλα τα αυτοκίνητα αποσκοπούν στην ικανοποίηση της ανάγκης ασφαλούς μεταφοράς των ανθρώπων, ωστόσο διαφέρουν σε μέγεθος, εμφάνιση, απόδοση, εφόδια, κτλ. Οι διαφορές οφείλονται σε διαφορετικές σχεδιαστικές μεθόδους που εξαρτώνται από τον τύπο των υλικών που χρησιμοποιούνται στην παραγωγή, τις προδιαγραφές των εξαρτημάτων, το σχεδιασμό και την ανάπτυξη νέων κινητήρων, κτλ. Ένας νεότερος ορισμός ορίζει την ποιότητα ως αντιστρόφως ανάλογη της μεταβλητότητας των χαρακτηριστικών της παραγωγικής διαδικασίας που προσδιορίζουν την ποιότητα του προϊόντος.

1.1.2 Ιστορική αναδρομή της ποιότητας

Ο Στατιστικός Έλεγχος Ποιότητας (*Statistical Quality Control*) είναι ένα κρίσιμο εργαλείο της επιχείρησης που επηρεάζει αποφάσεις σχετικές με τις προδιαγραφές, την παραγωγική διαδικασία και τον έλεγχο των παραγόμενων προϊόντων.

Για την αποτελεσματική χρήση του Στατιστικού Ελέγχου Ποιότητας απαιτείται η ενσωμάτωσή του σε ένα πλαίσιο λειτουργίας και διοίκησης της επιχείρησης που έχει ως στόχο τη διαρκή βελτίωση της ποιότητας σε όλα τα επίπεδα της επιχείρησης, γνωστού ως Διοίκησης Ολικής Ποιότητας (*Total Quality Management* ή

Total Quality Assurance). Θα μπορούσαμε να πούμε ότι ο Στατιστικός Έλεγχος Ποιότητας αποτέλεσε τον προάγγελο της Διοίκησης Ολικής Ποιότητας. Η ιστορία της ποιότητας ξεκινά στις αρχές του 20ου αιώνα με τη δημιουργία των πρώτων εργαστηρίων για πρότυπα στη Μεγάλη Βρετανία και με τη γραμμή παραγωγής στην αυτοκινητοβιομηχανία Ford στο Highland Park στις ΗΠΑ (1905). Τα βήματα στη συνέχεια ήταν αλματώδη και σήμερα η ποιότητα έχει αναδειχθεί σε κύριο μέλημα τόσο της βιομηχανίας όσο και των υπηρεσιών. Στον ακόλουθο πίνακα παρουσιάζονται τα κυριότερα ιστορικά σημεία στην ανάπτυξη της ποιότητας από το 1900 μέχρι σήμερα.

Πίνακας 1.2. Ιστορική αναδρομή της ποιότητας

1901: Τα πρώτα εργαστήρια προτύπων (standards) ιδρύονται στη Μ. Βρετανία
1907: Η AT&T Bell Laboratories αρχίζει τη συστηματική επιθεώρηση και έλεγχο προϊόντων και υλικών
1919: Η Ένωση Τεχνικών Επιθεωρητών ιδρύεται στην Αγγλία η οποία αργότερα μετονομάζεται σε Ινστιτούτο Διασφάλισης της Ποιότητας
1920: Στα εργαστήρια της AT&T Bell Laboratories ιδρύεται τμήμα ποιότητας
1924: Ο W. A. Shewhart εισάγει τα διαγράμματα ελέγχου σε ένα Technical Report στην AT&T Bell
1928: Το δειγματοληπτικό u963 σχέδιο αποδοχής σωρού αναπτύσσεται από τους Dodge και Romig
1931: Ο W.A. Shewhart εκδίδει το περιοδικό Economic Control of Quality of Manufactured Product
1932: Ο W.A. Shewhart δίνει διαλέξεις σε στατιστικές μεθόδους στην παραγωγή και στα διαγράμματα ελέγχου στο Πανεπιστήμιο του Λονδίνου
1938: Ο W.E. Deming προσκαλεί τον Shewhart για σεμινάρια στα διαγράμματα ελέγχου στο U.S. Department of Agriculture Graduate School
1940: Το Υπουργείο Πολέμου των Η.Π.Α. εκδίδει ένα οδηγό για την ανάλυση δεδομένων με χρήση διαγραμμάτων ελέγχου
1946: Ιδρύεται η American Society for Quality Control (ASQC)
1946: Ο W.E. Deming προσκαλείται στην Ιαπωνία για να δώσει σεμινάρια Στατιστικού Ποιοτικού Ελέγχου
1948: Ο Καθηγητής G. Taguchi αρχίζει τη μελέτη των πειραματικών σχεδιασμών
1950: Ο W.E. Deming αρχίζει την εκπαίδευση ανώτατων στελεχών βιομηχανιών της Ιαπωνίας
1950: Ο K. Ishikawa εισάγει το διάγραμμα αιτίου-αποτελέσματος (cause and effect diagram)
1954: Ο J. M. Juran προσκαλείται από την Ιαπωνία για να δώσει διαλέξεις σε θέματα διοίκησης και βελτίωσης της ποιότητας. Ο E.S. Page εισάγει το διάγραμμα ελέγχου CUSUM
1959: Ο S. Roberts εισάγει το διάγραμμα ελέγχου EWMA. Ξεκινά η έκδοση του επιστημονικού περιοδικού Technometrics
1960: Η ιδέα των κύκλων ποιότητας εισάγεται από τον K. Ishikawa
1969: Τα περιοδικά Quality Progress και Journal of Quality Technology αρχίζουν να εκδίδονται
1975: Εμφανίζονται τα πρώτα βιβλία σε σχεδιασμό πειραμάτων
1989: Αρχίζει να εκδίδεται το περιοδικό Quality Engineering. Η Motorola εισάγει την έννοια six-sigma.
1990: Σταδιακή αύξηση της ζήτησης στην βιομηχανία για πιστοποίηση κατά ISO 9000
1997: Η προσέγγιση six-sigma της Motorola υιοθετείται και από άλλες βιομηχανίες
2000: Γίνεται η δεύτερη αναθεώρηση της σειράς προτύπων ISO 9000

2003: Ενισχύετε η δέσμευση, η υπευθυνότητα και η επίγνωση των προμηθευτών με τις επιχειρήσεις σχετικά με την ποιότητα.
2005: Θεσπίζεται από τους Filho και Cezar η εφαρμογή διαδικασιών βελτίωσης του Συστήματος Διοίκησης Ποιότητας
2006: Θεσπίζεται το μοντέλο αξιολόγησης προγράμματος στην τριτοβάθμια εκπαίδευση

Στην Ιαπωνία, μετά το τέλος του πολέμου η ανέλιξη νεαρών ατόμων στην ιεραρχία των επιχειρήσεων που ήταν δεκτικά σε νέες ιδέες οδήγησε διακεκριμένους Αμερικανούς επιστήμονες στο να επισκεφθούν την περίοδο εκείνη την Ιαπωνία, αφού οι ιδέες τους ήταν πιθανό να βρουν πρόσφορο έδαφος αποδοχής και εφαρμογής των. Ένας από αυτούς ήταν ο *Edward Deming* (θεωρείται από πολλούς ως ο πατέρας της Διοίκησης Ολικής Ποιότητας) ο οποίος είναι υπεύθυνος για την ποιοτική επανάσταση που επιτεύχθηκε στην Ιαπωνία μετά το Β΄ Παγκόσμιο Πόλεμο που τελικά κατέληξε στη διεθνή κυριαρχία της Ιαπωνικής βιομηχανίας. Το κεντρικό σημείο της φιλοσοφίας του *Deming* ήταν η μείωση της μεταβλητότητας στην παραγωγική διαδικασία με ιδιαίτερη έμφαση στη στατιστική επιμέλεια και έλεγχο. Τα 14 βασικά σημεία της φιλοσοφίας του *Deming* παρουσιάζονται στον ακόλουθο πίνακα

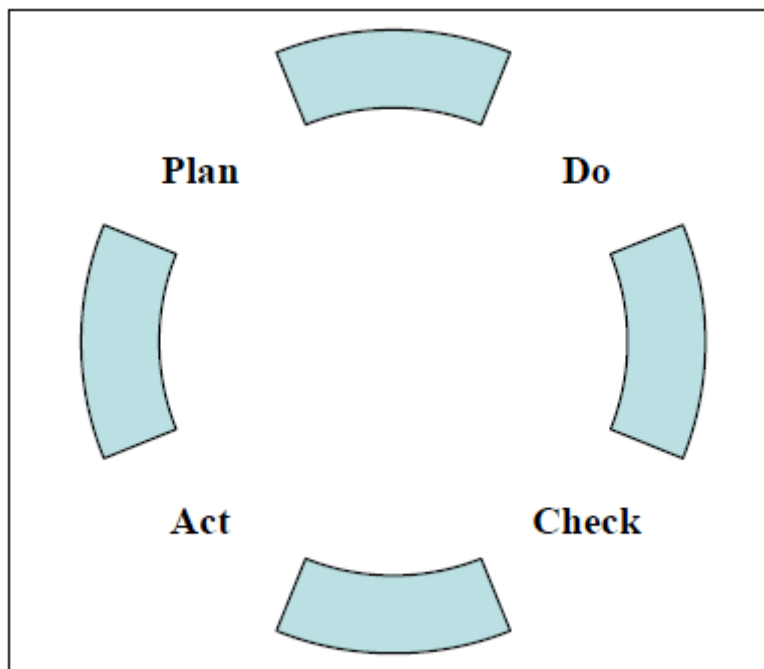
Πίνακας 1.3. Οι 14 κανόνες του *Deming*

Κανόνας 1: Δημιούργησε το κλίμα για ένα σταθερό πρόγραμμα που να αποβλέπει στη συνεχή βελτίωση των προϊόντων και των υπηρεσιών
Κανόνας 2: Υιοθέτησε τη νέα φιλοσοφία της απόρριψης της κακής δουλειάς, των ελαττωματικών και των μη ικανοποιητικών υπηρεσιών
Κανόνας 3: Μην βασίζεστε στον τελικό έλεγχο όλης της παραγωγής για την επίτευξη υψηλής ποιότητας
Κανόνας 4: Σταματήστε τη συνήθεια αγοράς προμηθειών με τη χαμηλότερη τιμή
Κανόνας 5: Προσπαθήστε να βελτιώνεται χωρίς διακοπή όλα τα συστήματα παραγωγής και υπηρεσιών
Κανόνας 6: Θεσπίστε ένα πρόγραμμα σύγχρονης εκπαίδευσης και εφαρμόστε το σε όλους τους εργαζόμενους
Κανόνας 7: Υιοθετήστε σύγχρονους τρόπους επίβλεψης
Κανόνας 8: Εξαλείψτε το φόβο. Μη διστάζετε να ρωτάτε και να αναφέρετε προβλήματα
Κανόνας 9: Εξαφανίστε τα εμπόδια μεταξύ των τμημάτων της επιχείρησης. Η από κοινού δουλειά μεταξύ των διαφόρων τμημάτων της επιχείρησης είναι απαραίτητο στοιχείο για την ανάπτυξη της ποιότητας
Κανόνας 10: Εξαλείψτε τους στόχους και τα συνθήματα για μηδέν ελαττωματικά προϊόντα
Κανόνας 11: Εξαφανίστε τους αριθμητικούς στόχους
Κανόνας 12: Εξαλείψτε τα εμπόδια που αφαιρούν από τον εργαζόμενο το δικαίωμα να νιώθει υπερήφανος για την δουλειά του
Κανόνας 13: Εφαρμόστε ένα πρόγραμμα συνεχούς ενημέρωσης και εκπαίδευσης για όλους τους εργαζομένους
Κανόνας 14: Δημιουργήστε ένα ανώτατο κλιμάκιο διοίκησης το οποίο θα εργαστεί σθεναρά για την επίτευξη των πρώτων 13 κανόνων

Στόχος της Διοίκησης Ολικής Ποιότητας είναι η συνεχής βελτίωση στην ποιότητα εκτέλεσης όλων των διεργασιών, προϊόντων και υπηρεσιών, σε μια επιχείρηση.

Αυτό επιτυγχάνεται με τέσσερα βήματα τα οποία επαναλαμβάνονται το ένα μετά το άλλο και όταν ολοκληρωθεί το τελευταίο βήμα η διαδικασία ξεκινά πάλι από την αρχή. Τα βήματα αυτά φαίνονται στον ακόλουθο τροχό του Deming

Σχήμα 1.1. Ο τροχός του Deming



Προγραμματίσει <i>Plan</i>	Προσδιόρισε τους στόχους.
Κάνει <i>Do</i>	Εφάρμοσε το σχεδιασμό. Έλεγξε και ρύθμισε δείκτες που εμφανίζουν κάποιο πρόβλημα Εκπαίδευσε τους εργαζομένους.
Εξέτασε <i>Check</i>	Εξέτασε τα αποτελέσματα του σχεδίου που εφαρμόστηκε και προσδιόρισε αν υπάρχουν προβλήματα
Δράσει <i>Act</i>	Ενέργησε ώστε να εξαλείψεις όποια προβλήματα βρέθηκαν στο προηγούμενο βήμα. Τυποποίησε τη διαδικασία αν το σχέδιο ήταν επιτυχημένο. Σχεδίασε μελλοντικές βελτιώσεις.

Ένας δεύτερος ακαδημαϊκός, εξίσου γνωστός για τη συμβολή του στην παραγωγή προϊόντων υψηλής ποιότητας στην Ιαπωνία, είναι ο *Joseph Juran*. Όπως και ο *Deming*, υποστήριξε τη μείωση της μεταβλητότητας στις διαδικασίες ως τον πιο ενδεδειγμένο τρόπο για τη βελτίωση της ποιότητας. Κεντρική αρχή της φιλοσοφίας του είναι η έμφαση που δίνει στις δράσεις της διοίκησης για την βελτίωση της ποιότητας. Υποστήριξε ότι τα συστήματα που ελέγχονται από τη διεύθυνση είναι κυρίως υπεύθυνα (80%) για τη βελτίωση της ποιότητας ενώ ένα μικρό μέρος (20%) οφείλεται στην απόδοση του εργατικού δυναμικού.

Ο *Philip Crosby* θεωρείται ο τρίτος μεγάλος θεωρητικός της Διοίκησης Ολικής Ποιότητας αν και η επιρροή του δεν ήταν τόσο μεγάλη όσο των *Deming* και *Juran*. Το κεντρικό σημείο της φιλοσοφίας του ήταν ότι η ποιότητα πρέπει να κτίζεται και όχι να επιθεωρείται. Η εφαρμογή των ιδεών του *Crosby* επαφίεται κυρίως σε αλλαγές στη συμπεριφορά και νοοτροπία του εργατικού δυναμικού της επιχείρησης παρά στη χρήση του Στατιστικού Ελέγχου Ποιότητας. Αναγνώρισε την τεράστια σπατάλη που παρουσιαζόταν κατά τις διαδικασίες επιθεώρησης και επιδιόρθωσης ελαττωματικών προϊόντων και την ανάγκη για μείωση των δαπανών αυτών. Ο *Crosby* ήταν υπέρμαχος του προγράμματος “Μηδέν Ελαττωματικά” (*Zero Defects Program*), το οποίο έχει δημιουργήσει πολλές παρανοήσεις και διαφωνίες. Ο *Crosby* πιστεύει ότι η αποδοχή ορισμένου ποσοστού ελαττωματικών στην παραγωγή είναι απαράδεκτη και ότι κάθε εργαζόμενος έχει την ευθύνη να κάνει την οποιαδήποτε δουλειά σωστά από την αρχή, να προλαμβάνει λάθη, σε αντίθεση με τους *Deming* και *Juran* που πιστεύουν ότι οποιεσδήποτε ατέλειες οφείλονται στο σχεδιασμό των συστημάτων παραγωγής και όχι στους εργαζομένους.

Αξίζει να σημειωθεί ότι παρά τις διαφορές μεταξύ των τριών (*Deming*, *Juran*, *Crosby*), εντούτοις προκύπτουν και κάποια κοινά σημεία στις θεωρίες τους όπως

1. η επιθεώρηση δεν αποτελεί την απάντηση στη βελτίωση της ποιότητας,
2. η ανάμειξη και καθοδήγηση από την ανώτερη διεύθυνση είναι ζωτικής σημασίας για τη δημιουργία της απαραίτητης κουλτούρας αφοσιωμένης στην ποιότητα, και
3. ένα πρόγραμμα για ποιότητα απαιτεί προσπάθειες από ολόκληρο τον οργανισμό και μακροπρόθεσμη δέσμευση καθώς επίσης και την απαραίτητη επένδυση για επιμόρφωση.

1.2 Συνιστώσες του στατιστικού ελέγχου ποιότητας

Ο Στατιστικός Έλεγχος Ποιότητας αποτελείται από ένα σύνολο μεθόδων στατιστικής ανάλυσης δεδομένων. Το σύνολο αυτό μπορεί να χωριστεί σε τρία βασικά υποσύνολα που το καθένα περιέχει στατιστικές μεθόδους προσανατολισμένες σε διαφορετικές φάσεις της παραγωγικής διαδικασίας.

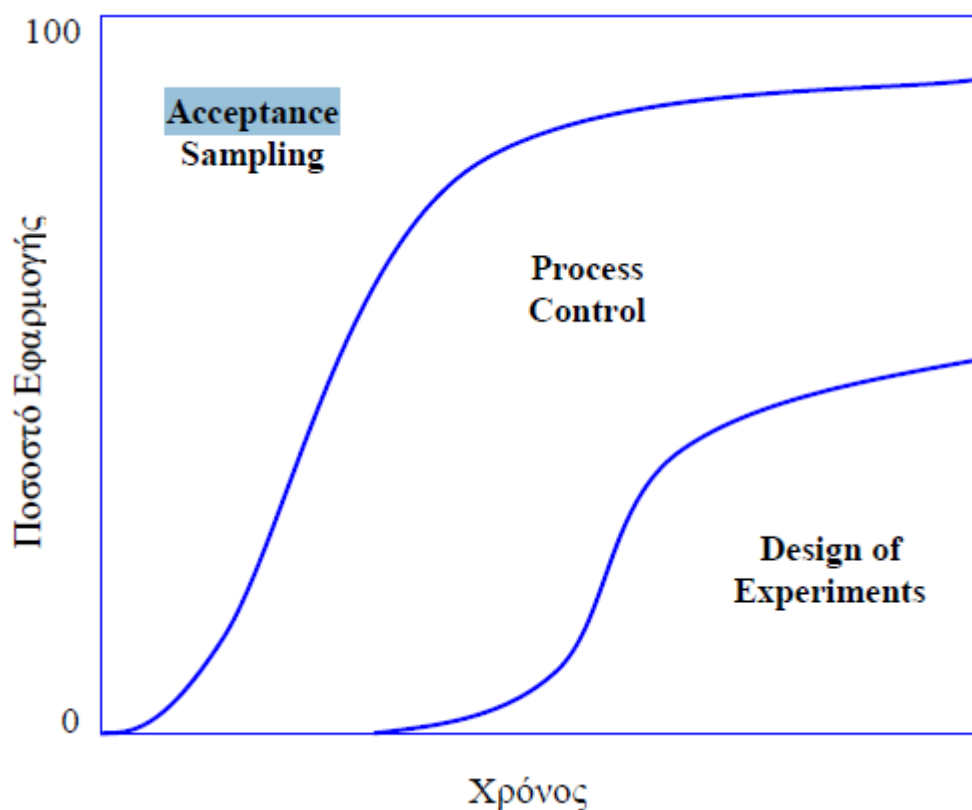
Τα τρία υποσύνολα είναι τα ακόλουθα:

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

Ο Σχεδιασμός και η Ανάλυση Πειραμάτων περιέχει όλες εκείνες τις στατιστικές τεχνικές οι οποίες μας βοηθούν στην ανακάλυψη της επίδρασης που έχουν τα διάφορα επίπεδα των παραγόντων (μεταβλητών) που επηρεάζουν τις ποιοτικές παραμέτρους του τελικού προϊόντος και συνεπώς διαδραματίζει σημαντικό ρόλο

στη βέλτιστη σχεδίαση της παραγωγικής διεργασίας. Ο Στατιστικός Έλεγχος Διεργασιών περιέχει στατιστικές τεχνικές που είναι απαραίτητες για τον έλεγχο της παραγωγικής διεργασίας κατά την διάρκεια της παραγωγής των προϊόντων. Η Δειγματοληψία Αποδοχής περιέχει στατιστικές τεχνικές (δειγματοληπτικές) που είναι απαραίτητες για να αποφασίσουμε αν μια συγκεκριμένη παρτίδα (σωρός) προϊόντων θα γίνει δεκτή ή θα απορριφθεί. Τα μοντέρνα συστήματα διασφάλισης της ποιότητας συνήθως εστιάζονται σε θέματα Στατιστικού Ελέγχου Διεργασιών και Σχεδιασμού Πειραμάτων παρά στη Δειγματοληψία Αποδοχής αφού αυτή δεν ανατροφοδοτεί (*feedback*) το σύστημα με κρίσιμες πληροφορίες που θα μπορούσαν να οδηγήσουν στη βελτίωση της ποιότητας. Η τυπική διαχρονική χρήση των τριών παραπάνω συστατικών του Στατιστικού Έλεγχου Ποιότητας στις περισσότερες επιχειρήσεις φαίνεται στο ακόλουθο σχήμα

Σχήμα 1.2. Διαχρονική χρήση των τριών συνιστωσών του στατιστικού ελέγχου ποιότητας



1.3 Τα 7 κυριότερα εργαλεία του στατιστικού ελέγχου διεργασιών

Για να ικανοποιεί ένα προϊόν το χρήστη πρέπει να παράγεται σύμφωνα με μια “σταθερή επαναλαμβανόμενη” διεργασία. Η διεργασία πρέπει να είναι ικανή να λειτουργεί με μικρή μεταβλητότητα γύρω από κάποιες τιμές στόχους που έχουν τεθεί στα ποιοτικά χαρακτηριστικά που πρέπει να διακρίνει το τελικό προϊόν.

Ο Στατιστικός Έλεγχος Διεργασιών είναι μια συλλογή εργαλείων που είναι χρήσιμα για την επίβλεψη της σταθερότητας μιας διεργασίας και τη βελτίωση της ικανότητάς της (*capability*) μέσω της μείωσης της μεταβλητότητάς της.

Ο Στατιστικός Έλεγχος Διεργασιών μπορεί να εφαρμοσθεί σε κάθε διαδικασία. Τα επτά κυριότερα εργαλεία που χρησιμοποιεί είναι τα ακόλουθα:

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

Τα παραπάνω στατιστικά εργαλεία αναφέρονται και ως “*the magnificent seven*”.

Από τα παραπάνω εργαλεία, το διάγραμμα ελέγχου αποτελεί ίσως το βασικότερο εργαλείο στο στατιστικό έλεγχο ποιότητας για την παρακολούθηση μιας παραγωγικής διαδικασίας. Όπως είναι γνωστό σε κάθε παραγωγική διαδικασία, όσο καλά σχεδιασμένη και προσεκτικά εκτελεσμένη είναι, εμφανίζεται πάντοτε ένα ποσό έμφυτης ή φυσικής μεταβλητότητας. Η μεταβλητότητα αυτή είναι αποτέλεσμα πολλών μικρών και συνήθως μη-σημαντικών παραγόντων. Στα πλαίσια του στατιστικού ελέγχου ποιότητας η παρουσία αυτής της μεταβλητότητας μόνο έχει ως αποτέλεσμα τη λειτουργία της διεργασίας υπό συνθήκες φυσικής μεταβλητότητας ή ισοδύναμα λέμε ότι η διεργασία είναι εντός στατιστικού ελέγχου (για συντομία εντός ελέγχου). Αν όμως υπάρχουν άλλες αιτίες μεταβλητότητας (π.χ λανθασμένα ρυθμισμένες μηχανές, απειρία χειριστών) τότε η διεργασία εμφανίζει αυξημένη μεταβλητότητα και λειτουργεί σε μη-αποδεκτό επίπεδο. Αυτές οι αιτίες μεταβλητότητας ονομάζονται ειδικές αιτίες μεταβλητότητας. Μια διεργασία που λειτουργεί παρουσία ειδικών αιτιών μεταβλητότητας λέμε ότι βρίσκεται εκτός στατιστικού ελέγχου (για συντομία εκτός ελέγχου).

1.4 Περιγραφή και χρήση ενός διαγράμματος ελέγχου

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 παρακολούθηση της κρίσιμης ποσότητας βασίζεται σε μετρήσεις του χαρακτηριστικού X (τυχαία μεταβλητή), που προκύπτουν από την επιλογή τυχαίων δειγμάτων προϊόντων από την παραγωγή σε διαφορετικές χρονικές στιγμές, έστω τα X_1, X_2, \dots, X_n χρησιμοποιώντας τα τυχαία δείγματα X_1, X_2, \dots, X_n υπολογίζουμε την τιμή $W_i = g(X_i), i = 1, 2, \dots, n$ μιας κατάλληλης στατιστικής συνάρτησης (τυχαίας μεταβλητής) που εκτιμά (συνήθως αμερόληπτη εκτιμήτρια) την κρίσιμη ποσότητα που μας ενδιαφέρει (π.χ. μέση τιμή ή διακύμανση της X). Έτσι η (διαχρονική) παρακολούθηση της συμπεριφοράς της κρίσιμης ποσότητας επιτυγχάνεται με την παρακολούθηση των τιμών που λαμβάνει η στατιστική συνάρτηση W στα διάφορα δείγματα.

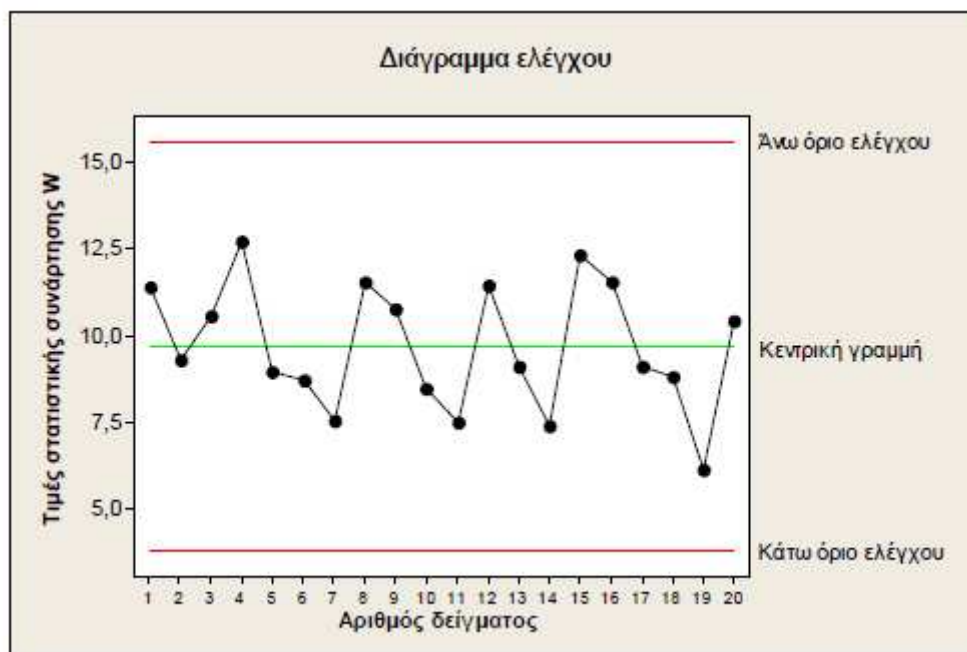
Για παράδειγμα ας υποθέσουμε ότι ενδιαφερόμαστε να παρακολουθήσουμε τη συμπεριφορά της μέσης τιμής της διαμέτρου X των κυλίνδρων που παράγει μια μηχανή. Για το σκοπό μπορούμε να επιλέξουμε τυχαία δείγματα μεγέθους n ($n \geq 1$) κυλίνδρων από την παραγωγή της μηχανής σε διαφορετικά χρονικά διαστήματα και να χρησιμοποιήσουμε τη στατιστική συνάρτηση

$$W_i = g(X_i) = (X_{i1} + X_{i2} + \dots + X_{in})/n$$

(η οποία είναι αμερόληπτη εκτιμήτρια του μέσου της X) για την παρακολούθηση της συμπεριφοράς της μέσης τιμής.

Ένα τυπικό διάγραμμα ελέγχου Shewhart (τα διαγράμματα αυτά τα εισήγαγε το 1924 ο W. A. Shewhart) είναι μια γραφική παράσταση με την ακόλουθη μορφή

Σχήμα 13. Τυπικό διάγραμμα ελέγχου τύπου Shewhart

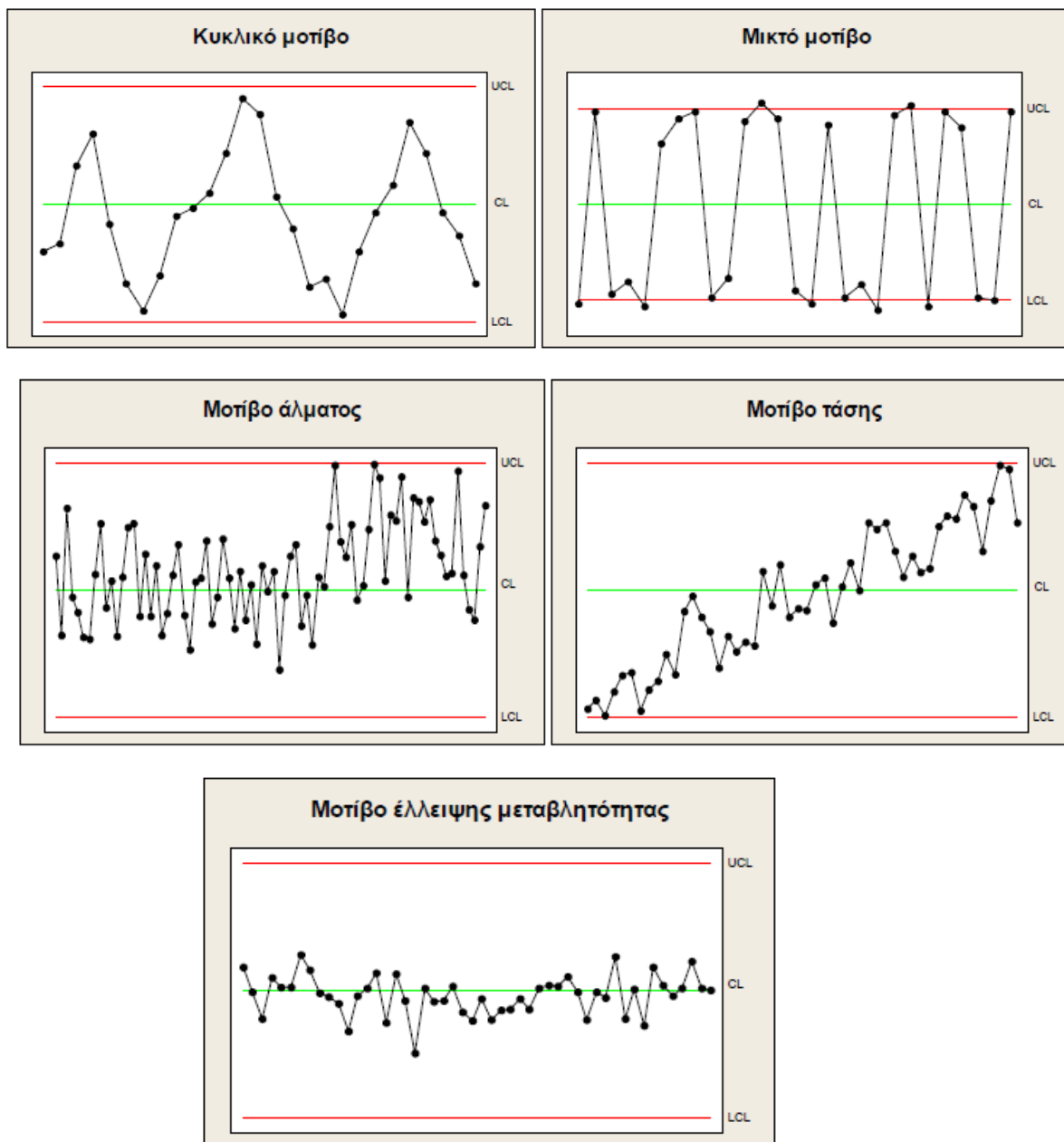


Στο παραπάνω σχήμα, εκτός από τις παρατηρούμενες τιμές της W που έχουν παρασταθεί με σημεία (κουκκίδες) τα οποία έχουν συνδεθεί με μια τεθλασμένη γραμμή, έχουν σχεδιαστεί και άλλες τρεις γραμμές. Η κεντρική γραμμή (*center line, CL*) ή μέσο επίπεδο της διεργασίας παριστάνει συνήθως τη μέση τιμή (*mean value*) της W όπως αυτή προκύπτει από τη λειτουργία μιας εντός ελέγχου διεργασίας ή την τιμή στόχο του προϊόντος. 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 (για την ακρίβεια εντός στατιστικού 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 οπότε αντιμετωπίζουμε κατάσταση συναγερμού (*alarm*) 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.

Θα πρέπει να σημειωθεί 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. Οι βασικότερες μορφές προτύπων ή μοτίβων (*patterns*) μη τυχαίου τρόπου συμπεριφοράς μιας χρονοσειράς σημείων είναι τα πρότυπα κύκλων (*cycles*), τα μικτά πρότυπα (*mixtures*), τα πρότυπα αλμάτων (*shifts in the process level*), τα πρότυπα τάσεων (*trends*), και τα πρότυπα έλλειψης μεταβλητότητας (*stratification*) τα οποία περιγράφονται οπτικά στο ακόλουθο σχήμα

Σχήμα 1.4 Τα κυριότερα μη τυχαία μοτίβα



Στα κυκλικά ή περιοδικά πρότυπα (*cycles*) εμφανίζεται μια περιοδικότητα στη διάταξη των σημείων με αποτέλεσμα να σχηματίζονται επαναλαμβανόμενοι κύκλοι. Πιθανές αιτίες είναι οι περιοδικές διακυμάνσεις των περιβαλλοντικών συνθηκών και συνθηκών παραγωγής όπως η θερμοκρασία, η ηλεκτρική τάση, εναλλαγή εργατοτεχνιτών, εναλλαγή μηχανών, κτλ.

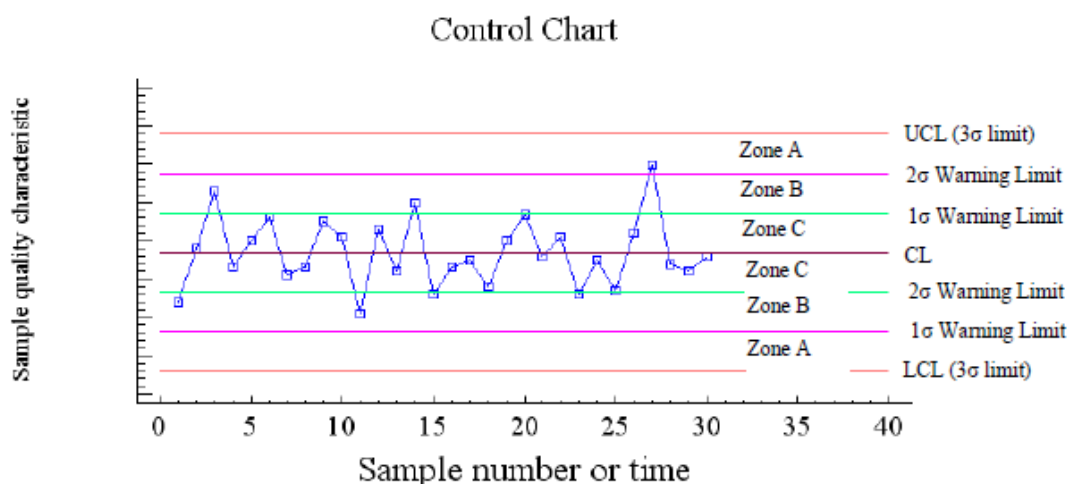
Τα μικτά πρότυπα (*mixture*) φαίνεται να προκύπτουν από δύο διαφορετικές κατανομές του χαρακτηριστικού X μεταξύ των οποίων παλινδρομεί η διεργασία.

Τα πρότυπα αλμάτων (*shift in the process level*) οφείλονται συνήθως σε νέους εργαζόμενους, νέες μεθόδους παραγωγής, νέες μηχανές, διαδοχική επεξεργασία παρτίδων πρώτων υλών διαφορετικής ποιότητας, κτλ.

Τα πρότυπα τάσεων (*trend*) οφείλονται συνήθως σε παράγοντες όπως η σταδιακή φθορά εργαλείων, η κόπωση των εργαζομένων, η παρουσία επιθεωρητή, κτλ.

Στα πρότυπα έλλειψης μεταβλητότητας (*stratification*) τα σημεία του διαγράμματος κινούνται με τεχνητό τρόπο γύρω από την κεντρική γραμμή χωρίς να υπάρχει κάποια αξιοσημείωτη μεταβλητότητα. Οφείλονται συνήθως σε εσφαλμένη σχεδίαση των ορίων ελέγχου. Σε πολλές περιπτώσεις για να κάνουμε περισσότερο ευαίσθητο ένα διάγραμμα ελέγχου ως προς την ικανότητά του να ανιχνεύει πιο γρήγορα εκτός ελέγχου διεργασίες, εκτός από τη σχεδίαση των ορίων ελέγχου, σχεδιάζουμε επίσης και προειδοποιητικά όρια (*warning limits*) εσωτερικά των ορίων ελέγχου όπως δείχνει το παρακάτω σχήμα

Σχήμα 1.5 Ζώνες A, B και C σε ένα διάγραμμα ελέγχου

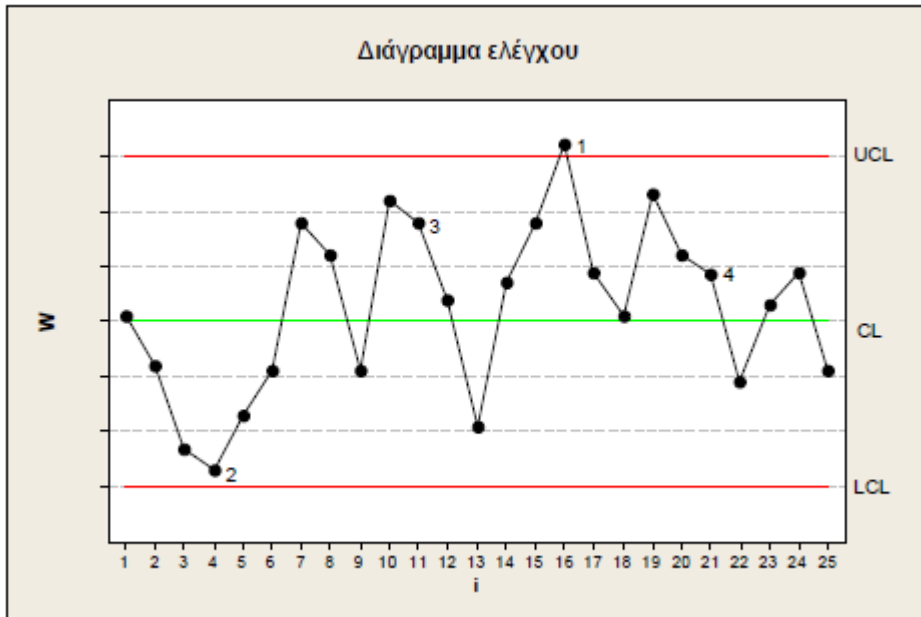


Τα προειδοποιητικά όρια χρησιμοποιούνται μαζί με κάποιους “κανόνες” που περιγράφουν την εμφάνιση ειδικών μοτίβων σε ένα διάγραμμα ελέγχου. Στην περίπτωση που εμφανιστεί το μοτίβο που περιγράφει ο κανόνας τότε θεωρούμε ότι η διεργασία είναι εκτός ελέγχου χωρίς απαραίτητα να έχουμε κάποιο σημείο του διαγράμματος εκτός των ορίων ελέγχου (*UCL* και *LCL*). Οι σημαντικότεροι κανόνες που χρησιμοποιούνται για την ευαισθητοποίηση ενός διαγράμματος ελέγχου είναι οι ακόλουθοι:

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

Οι πρώτοι τέσσερις κανόνες είναι γνωστοί ως *Western Electric Rules*. Στο ακόλουθο διάγραμμα ελέγχου έχουν σημειωθεί τα σημεία που “χτυπούν” για πρώτη φορά ο καθένας από τους *Western Electric Rules*.

Σχήμα 1.6 Western Electric Rules



Έτσι, στο σημείο 4 χτύπησε ο Κανόνας 2, στο σημείο 11 χτύπησε ο Κανόνας 3, στο σημείο 16 χτύπησε ο Κανόνας 1 και στο σημείο 21 χτύπησε ο Κανόνας 4.

1.5 Τύποι διαγραμμάτων ελέγχου

Υπάρχουν δύο βασικές κατηγορίες διαγραμμάτων ελέγχου ανάλογα με το είδος της μεταβλητής που περιγράφει το ποιοτικό χαρακτηριστικό του προϊόντος που ενδιαφερόμαστε να παρακολουθήσουμε. Έτσι έχουμε τα διαγράμματα ελέγχου για μεταβλητές (control charts for variables) όταν η απεικονιζόμενη ποσότητα είναι μια συνεχής τυχαία μεταβλητή, και τα διαγράμματα ελέγχου για διακριτά χαρακτηριστικά (control charts for attributes) όταν η απεικονιζόμενη ποσότητα είναι μια διακριτή τυχαία μεταβλητή.

1.6 Κατασκευή ενός διαγράμματος ελέγχου

Στην παρούσα παράγραφο θα περιγράψουμε τη γενική διαδικασία κατασκευής ενός διαγράμματος ελέγχου και θα τονίσουμε τα ιδιαίτερα χαρακτηριστικά αυτής. Έστω ότι επιλέγουμε τυχαία δείγματα του προϊόντος σε διαφορετικές χρονικές στιγμές της παραγωγής στα οποία αντιστοιχούν τυχαία δείγματα τιμών της μεταβλητής X έστω τα X_1, X_2, \dots, X_n . Έστω $W_i = g(X_i), i = 1, 2, \dots, n$ οι τιμές της στατιστικής συνάρτησης $W = g(X)$ που εκτιμά την κρίσιμη ποσότητα (π.χ μέση τιμή, διακύμανση, ποσοστό συμφωνούντων με τις προδιαγραφές προϊόντων κτλ.) που μας ενδιαφέρει να παρακολουθήσουμε. Το γενικό μοντέλο ορίων για την κατασκευή ενός διαγράμματος ελέγχου περιγράφεται ως ακολούθως

Μοντέλο ορίων L σίγμα

$$\begin{aligned}UCL &= \mu_{\bar{W}} + L\sigma_{\bar{W}} \\Center\ Line &= \mu_{\bar{W}} \\LCL &= \mu_{\bar{W}} - L\sigma_{\bar{W}}\end{aligned}$$

όπου $\mu_{\bar{w}}$, $\sigma_{\bar{w}}$ δηλώνουν τη μέση τιμή και την τυπική απόκλιση της στατιστικής συνάρτησης W η οποία απεικονίζεται στο διάγραμμα ελέγχου, ενώ η ποσότητα L δηλώνει τη απόσταση των ορίων ελέγχου από την κεντρική γραμμή σε μονάδες τυπική απόκλισης. Η πιο συνηθισμένη τιμή του L είναι το 3 και σε αυτή την περίπτωση θα αναφερόμαστε στην κατασκευή διαγράμματος ελέγχου τριών σίγμα. Αν ένα σημείο βρεθεί εκτός των ορίων ελέγχου τότε έχουμε ένδειξη εκτός ελέγχου διεργασίας και θα πρέπει να γίνει έρευνα για την ανεύρεση ειδικών αιτιών μεταβλητότητας οι οποίες, αν βρεθούν, θα πρέπει να απομονωθούν.

Το μοντέλο ορίων σίγμα για την κατασκευή ενός διαγράμματος ελέγχου προτάθηκε από τον Walter A. Shewhart (1931) και ονομάζεται διάγραμμα ελέγχου τύπου Shewhart. Αποτελεί το πιο ευρέως χρησιμοποιούμενο διάγραμμα ελέγχου στην πράξη λόγω της απλότητας στην κατασκευή και ερμηνεία του.

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

Μοντέλο ορίων πιθανότητας a

$$\begin{aligned}UCL &= \mu_{\bar{W}} + Z_{a/2}\sigma_{\bar{W}} = w_{a/2} \\Center\ Line &= \mu_{\bar{W}} \\LCL &= \mu_{\bar{W}} - Z_{a/2}\sigma_{\bar{W}} = w_{1-(a/2)}\end{aligned}$$

με w_a συμβολίζεται το άνω a 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 $a/2$ (a είναι η πιθανότητα εσφαλμένου συναγερμού). Για $a/2 = 0.001$ (δηλαδή μοντέλο με όρια πιθανότητας 0.002), έχουμε ότι $Z_{0,001} = 3.09024$.

Σημειώνουμε πως στην περίπτωση που η κατανομή της W δεν είναι κανονική (ή δεν ισχύουν οι προϋποθέσεις του Κεντρικού Οριακού Θεωρήματος) μπορούμε αντί των ποσοστιαίων σημείων της τυπικής κανονικής κατανομής να χρησιμοποιήσουμε ποσοστιαία σημεία της κατανομής W .

1.6.1 Προειδοποιητικά όρια ελέγχου και μέγεθος δείγματος

Εκτός από τα άνω και κάτω όρια ελέγχου του διαγράμματος ελέγχου, έχουν οριστεί και εξωτερικά προειδοποιητικά όρια ελέγχου (outer warning limits, OWL) καθώς και εσωτερικά όρια ελέγχου (inner warning limits, IWL). Σε μοντέλο ορίων 3σ τα εξωτερικά προειδοποιητικά όρια ελέγχου σχεδιάζονται σε απόσταση 2σ από την κεντρική γραμμή του διαγράμματος και στοχεύουν στην αύξηση της ευαισθησίας αυτού. Έτσι αν παρατηρήσουμε ένα ή περισσότερα σημεία μεταξύ των εξωτερικών προειδοποιητικών ορίων και των ορίων ελέγχου, προς τη μια πλευρά της κεντρικής γραμμής, τότε αυτό λαμβάνεται ως ένδειξη ότι η διεργασία μπορεί να είναι εκτός ελέγχου. Επίσης, σε μοντέλο ορίων 3σ τα εσωτερικά προειδοποιητικά όρια σχεδιάζονται σε απόσταση 1σ από την κεντρική γραμμή του διαγράμματος. Έτσι το διάγραμμα ελέγχου χωρίζεται σε τρεις ζώνες, προς κάθε πλευρά της κεντρικής γραμμής, και οι ζώνες αυτές χρησιμοποιούνται για τον ορισμό διαφόρων κανόνων των οποίων η χρήση αποσκοπεί στο να κάνουν πιο ευαίσθητο ένα διάγραμμα ελέγχου στην ανίχνευση της παρουσίας ειδικών αιτιών μεταβλητότητας. Ένα τέτοιο διάγραμμα ελέγχου απεικονίζεται στο σχήμα 1.5.

Ωστόσο, για την ανάπτυξη ενός διαγράμματος ελέγχου είναι απαραίτητο να επιλέξουμε το μέγεθος του δείγματος n που θα χρησιμοποιήσουμε, αφού λόγοι κόστους αλλά και λειτουργίας της ίδιας της διεργασίας δεν επιτρέπουν σε πολλές περιπτώσεις την εξέταση όλων των παραγόμενων προϊόντων. Επίσης θα πρέπει να καθοριστεί και ο ρυθμός με τον οποίο θα συλλέγονται τα δείγματα. Γίνεται άμεσα αντιληπτό ότι οι επιλογές των παραμέτρων αυτών που παίζουν σημαντικό ρόλο στην απόδοση του διαγράμματος ενώ επηρεάζουν και επηρεάζονται από τον παράγοντα κόστος. Η γενική πολιτική που ακολουθείται είναι να συλλέγουμε μικρά μεγέθη δειγμάτων αρκετά συχνά (ιδιαίτερα σε περιπτώσεις μαζικής παραγωγής προϊόντων ή όταν εμφανίζονται συχνά ειδικές αιτίες μεταβλητότητας) ενώ τα μεγάλα μεγέθη δειγμάτων λιγότερο συχνά. Γενικά, μεγάλα μεγέθη δειγμάτων κάνουν πιο εύκολη την ανίχνευση οποιασδήποτε μεταβολής στην τιμή της κρίσιμης ποσότητας που παρακολουθούμε ενώ ένας ταχύς ρυθμός συλλογής δειγμάτων επιτρέπει την έγκαιρη ανίχνευση της μεταβολής. Τα τελευταία 20 χρόνια έχει γίνει μια προσπάθεια είτε το μέγεθος δείγματος είτε ο ρυθμός δειγματοληψίας ή και μεταβλητού μεγέθους δείγματος, δηλαδή η χρήση διαγραμμάτων ελέγχου με προσαρμοσμένες παραμέτρους.

Έτσι, εκτός της τιμής του L θα πρέπει να καθοριστεί η τιμή του n καθώς και η συχνότητα δειγματοληψίας (sampling rate). Σε περιπτώσεις που η επιλογή των παραμέτρων του διαγράμματος γίνεται με στατιστικά κριτήρια θα αναφερόμαστε σε στατιστικό σχεδιασμό των διαγραμμάτων ελέγχου (Statistical Design Control Charts) ενώ αν γίνεται με οικονομικά κριτήρια (Economical Design Control Charts).

1.7 Μέτρα απόδοσης ενός διαγράμματος ελέγχου

Μια άλλη έννοια που σχετίζεται με τα διαγράμματα ελέγχου είναι το μέσο μήκος ροής (ή μέσο μήκος διαδρομής) του διαγράμματος (*average run length, ARL*) που ορίζεται με τη σχέση

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

όπου p συμβολίζει την πιθανότητα να βρεθεί ένα σημείο του διαγράμματος ελέγχου εκτός των ορίων ελέγχου.

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

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

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

τότε

$$E(X) = \frac{1}{p} \quad \text{και} \quad V(X) = \frac{1}{p^2}$$

Για μια διεργασία που βρίσκεται εντός ελέγχου και στην περίπτωση που χρησιμοποιούμε 3σ όρια ελέγχου με κανονική κατανομή για την απεικονιζόμενη στατιστική συνάρτηση W , έχουμε ότι το εντός ελέγχου μέσο μήκος ροής ARL_{in} (*in-control average run length*) είναι ίσο με

$$ARL_{in} = \frac{1}{\alpha} = \frac{1}{0,0027} \approx 370,4$$

όπου α είναι η πιθανότητα να βρεθεί ένα σημείο εκτός των ορίων ελέγχου ενώ η διεργασία βρίσκεται εντός ελέγχου και το χαρακτηριστικό X του προϊόντος ακολουθεί κανονική κατανομή με, εντός ελέγχου, μέση τιμή μ και διασπορά σ^2 . Για μια διεργασία που βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου επιπέδου της διεργασίας, το εκτός ελέγχου μέσο μήκος ροής ARL_{out} (*out-of-control average run length*) είναι ίσο με

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

αφού η πιθανότητα να βρεθεί σε αυτή την περίπτωση ένα σημείο του διαγράμματος ελέγχου εντός των ορίων ελέγχου είναι ίσο με β , οπότε βρίσκεται εκτός των ορίων ελέγχου με πιθανότητα $1 - \beta$. Στη πράξη είναι επιθυμητή μια μικρή τιμή του ARL_{out} αφού το ARL_{out} δηλώνει τον αναμενόμενο αριθμό δειγμάτων που πρέπει να ληφθούν για να εντοπιστεί η αλλαγή στο μέσο επίπεδο της διεργασίας από τη στιγμή που η μετατόπιση αυτή συμβεί. Σημαντικό είναι να αναφερθεί πως αυξάνοντας το μέγεθος του δείγματος n η πιθανότητα β μειώνεται και συνεπώς βελτιώνεται η ικανότητα του διαγράμματος ελέγχου να ανιχνεύει εγκαίρως τις μετατοπίσεις του μέσου επιπέδου της διεργασίας. Με βάση τη μετατόπιση που θέλουμε να ανιχνεύσουμε, είναι δυνατό να επιλέξουμε το κατάλληλο μέγεθος δείγματος έτσι ώστε να έχουμε συγκεκριμένη τιμή για το ARL_{out} .

Αν αντί των 3σ ορίων είχαμε επιλέξει L όρια, τότε θα βλέπαμε πως καθώς αυξάνει η τιμή του L μειώνεται η πιθανότητα εσφαλμένης ένδειξης εκτός ελέγχου διεργασίας. Ενώ αντίθετα αυξάνεται η πιθανότητα λανθασμένης ένδειξης ότι η διεργασία βρίσκεται εντός ελέγχου ενώ στην πραγματικότητα βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου της διεργασίας.

Τα τελευταία χρόνια ωστόσο η χρήση του ARL ως μέτρου για την περιγραφή της απόδοσης μιας διεργασίας έχει υποστεί κριτική, αφού το ARL που παρατηρείται στην πράξη διαφέρει συνήθως αρκετά από το θεωρητικό ARL λόγω του ότι η κατανομή του μήκους ροής είναι μη-συμμετρική (παρουσιάζει δεξιά ουρά) και

συνεπώς η μέση τιμή δεν μπορεί να χρησιμοποιηθεί ως αντιπροσωπευτικό μέτρο κεντρικής τάσης της κατανομής (ιδιαίτερα για μικρές τιμές του p).

Μια συνήθης πρακτική είναι να συνοδεύεται η τιμή του ARL με την τυπική απόκλιση της κατανομής του μήκους ροής $SDRL$ (standard deviation run length). Στην περίπτωση που η κατανομή του μήκους ροής του διαγράμματος είναι γεωμετρική η τιμή του $SDRL$ ταυτίζεται με την τιμή του ARL . Επιπλέον η γνώση της κατανομής του μήκους ροής, η οποία με βάση τα όσα έχουμε δει έως τώρα είναι γεωμετρική κατανομή, μας επιτρέπει τον υπολογισμό και τη χρήση ποσοστιαίων σημείων προκειμένου να έχουμε περισσότερη πληροφορία σχετικά με την απόδοση του διαγράμματος.

Θεωρώντας ένα μοντέλο ορίων 3σ που η κατανομή της W είναι κανονική με γνωστές παραμέτρους έχουμε τον ακόλουθο πίνακα για μια εντός ελέγχου διεργασία.

Πίνακας 1.4: Πίνακας στατιστικών μέτρων για την κατανομή του μήκους ροής

α	ARL	SDRL	Q_1	M	Q_3
0,0027	370,4	370,4	107	257	513

Από το πίνακα προκύπτει ότι κατά μέσο όρο αναμένονται περίπου 370 δείγματα μέχρι να βρεθεί σημείο εκτός των ορίων ελέγχου όταν η διεργασία είναι εντός ελέγχου. Από την τιμή της διαμέσου M ο διαχειριστής του διαγράμματος γνωρίζει ότι η πιθανότητα να περιμένω περίπου 257 δείγματα μέχρι να βρεθεί ένα σημείο εκτός των ορίων ελέγχου είναι τουλάχιστον 50%. Επιπλέον από την τιμή Q_3 έχουμε ότι η πιθανότητα να μη δώσει σημείο το διάγραμμα ένδειξη εκτός ελέγχου διεργασίας μετά το 513 απεικονιζόμενο σημείο είναι το πολύ 25%. Έτσι λοιπόν σε πρακτικές εφαρμογές η χρήση των ποσοστιαίων σημείων της κατανομής του μήκους ροής μας παρέχει περισσότερες πληροφορίες σχετικά με την απόδοση του διαγράμματος ελέγχου.

1.8 Άλλα μέτρα απόδοσης

Στην προηγούμενη παράγραφο αναφερθήκαμε στο βασικότερο μέτρο απόδοσης ενός διαγράμματος ελέγχου και είδαμε τα κύρια χαρακτηριστικά του. Στη παρούσα παράγραφο θα αναφερθούμε σε εναλλακτικά μέτρα απόδοσης ενός διαγράμματος ελέγχου. Ένα μέτρο που χρησιμοποιείται συχνά είναι ο μέσος χρόνος σήματος ATS (average time to signal) που ορίζεται από τη σχέση

$$ATS = ARL \times \tau$$

όπου τ συμβολίζει το χρόνο που μεσολαβεί για τη λήψη δυο διαδοχικών δειγμάτων. Έτσι η ποσότητα ATS δηλώνει το μέσο χρόνο που απαιτείται για να δώσει το διάγραμμα ελέγχου ένδειξη εκτός ελέγχου διεργασίας. Αν ο ρυθμός δειγματοληψίας είναι μεταβλητός τότε ορίζεται μια νέα κατηγορία διαγραμμάτων ελέγχου, τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας (adaptive sampling rate control charts).

Σε αυτή την περίπτωση το τ δεν είναι σταθερό αλλά τυχαία μεταβλητή και ο μέσος χρόνος σήματος δίνεται από τη σχέση

$$ATS = ARL \times E(\tau)$$

Είναι σημαντικό να αναφέρουμε ότι τα ARL , ATS που έχουμε αναφέρει έως τώρα υπολογίζουν την απόδοση του διαγράμματος θεωρώντας ότι η μεταβολή στην εντός ελέγχου τιμή της παραμέτρου συμβαίνει ακριβώς τη χρονική στιγμή που αρχίζει η παρακολούθηση της διεργασίας. Σε αυτή την περίπτωση μιλάμε για μέτρα ARL και ATS αρχικής κατάστασης. Σε πολλές περιπτώσεις όμως η διεργασία ξεκινάει ευρισκόμενη εντός ελέγχου και η μεταβολή στην τιμή της υπό παρακολούθηση παραμέτρου συμβαίνει σε κάποιο τυχαίο χρονικό σημείο μεταξύ της λήψης δυο διαδοχικών δειγμάτων.

Σε αυτή την περίπτωση για τον υπολογισμό της απόδοσης ενός διαγράμματος με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο ATS σταθερής κατάστασης το οποίο στην βιβλιογραφία αναφέρεται ως προσαρμοσμένος μέσος χρόνος σήματος (adjusted average time to signal $AATS$). Αναλόγως στην περίπτωση των διαγραμμάτων με σταθερό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο ARL σταθερής κατάστασης.

Γενικότερα συγκρίσεις μεταξύ των διαφορετικών διαγραμμάτων ελέγχου χωρίς μεταβλητό ρυθμό δειγματοληψίας έχει επικρατήσει να γίνονται χρησιμοποιώντας το ARL αρχικής κατάστασης ενώ για τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το $AATS$.

1.9 Ταξινόμηση διαγραμμάτων ελέγχου

Έχουμε ήδη αναφέρει μια πρώτη διάκριση των διαγραμμάτων ελέγχου που βασίζεται στην κατανομή του ποιοτικού χαρακτηριστικού που μας ενδιαφέρει. Έτσι, αν αυτή είναι συνεχής έχουμε διαγράμματα ελέγχου για μεταβλητές (control charts for variables) ενώ αν είναι διακριτή έχουμε τα διαγράμματα ελέγχου για διακριτά χαρακτηριστικά (control charts for attributes). Έτσι λοιπόν μπορούμε να διακρίνουμε τα διαγράμματα ελέγχου στα εξής:

- Αν από την παραγωγική διεργασία λαμβάνονται δείγματα μετρήσεων μεγέθους μεγαλύτερου της μονάδας μιλάμε για διαγράμματα ελέγχου για ομάδες (control charts for rational subgroups), ενώ αν λαμβάνονται δείγματα μετρήσεων μεγέθους ίσου με την μονάδα μιλάμε για διαγράμματα ελέγχου για μεμονωμένες παρατηρήσεις (control charts for individual observations)
- Αν οι μετρήσεις που λαμβάνονται κάθε χρονική στιγμή t εξαρτώνται από μετρήσεις που ελήφθησαν σε προηγούμενες χρονικές στιγμές τότε αναφερόμαστε σε διαγράμματα ελέγχου για αυτοσυσχετιζόμενες διεργασίες (control charts for auto correlated processes), ενώ αν οι μετρήσεις που λαμβάνονται σε κάθε χρονική στιγμή είναι ανεξάρτητες από μετρήσεις που ελήφθησαν σε προηγούμενες χρονικές στιγμές τότε αναφερόμαστε σε διαγράμματα ελέγχου για ασυσχετίστες διεργασίες (control charts for uncorrelated processes).

- Αν οι μετρήσεις που λαμβάνονται αφορούν ένα ποιοτικό χαρακτηριστικό χρησιμοποιούμε μονομεταβλητά διαγράμματα ελέγχου (univariate control charts) ενώ αν οι μετρήσεις αναφέρονται σε περισσότερα ποιοτικά χαρακτηριστικά χρησιμοποιούμε πολυμεταβλητά διαγράμματα ελέγχου (multivariate control charts).
- Αν οι μετρήσεις που λαμβάνονται ακολουθούν μια γνωστή κατανομή τότε μιλάμε για παραμετρικά διαγράμματα ελέγχου ενώ στην αντίθετη περίπτωση μιλάμε για μη παραμετρικά διαγράμματα ελέγχου με τα οποία και θα ασχοληθούμε εκτενέστερα στα επόμενα κεφάλαια.
- Αν το διάγραμμα ελέγχου στοχεύει στην παρακολούθηση της μέσης τιμής του ποιοτικού χαρακτηριστικού μιλάμε για διάγραμμα ελέγχου για τη μέση τιμή, ενώ αν στοχεύει στην παρακολούθηση της διασποράς του μιλάμε για διάγραμμα ελέγχου για τη διασπορά.
- Αν οι μετρήσεις (δείγματα) λαμβάνονται σε σταθερούς χρόνους μιλάμε για διαγράμματα ελέγχου σταθερού χρόνου. ενώ στην αντίθετη περίπτωση μιλάμε για διαγράμματα ελέγχου μεταβλητού χρόνου.
- Αν το μέγεθος δείγματος είναι σταθερό τότε μιλάμε για διαγράμματα ελέγχου σταθερού δείγματος ενώ στην αντίθετη περίπτωση μιλάμε για διαγράμματα ελέγχου μεταβλητού μεγέθους δείγματος.

1.10 Βασικά είδη διαγραμμάτων ελέγχου

Τα βασικότερα είδη διαγραμμάτων ελέγχου που χρησιμοποιούνται στην πράξη είναι τα διαγράμματα ελέγχου τύπου Shewhart, CUSUM και EWMA.

Πριν αναφερθούμε στα τρία είδη διαγραμμάτων πρέπει να τονίσουμε ότι για τον έλεγχο μια παραγωγικής διαδικασίας με χρήση διαγραμμάτων ελέγχου υπάρχουν δυο φάσεις, η Φάση I και η Φάση II οι οποίες εξηγούνται παρακάτω:

Φάση I: Σε αυτή τη φάση τα διαγράμματα ελέγχου χρησιμοποιούνται αναδρομικά για να ελέγξουν αν η διεργασία ήταν εντός ή εκτός ελέγχου εξετάζοντας δείγματα που συλλέχθηκαν σε παρελθοντικό χρόνο. Σκοπός τους είναι να βοηθήσουν το διαχειριστή της διεργασίας να φέρει τη διεργασία εντός στατιστικού ελέγχου. Όταν αυτό επιτευχθεί τα διαγράμματα ελέγχου που προκύπτουν (κεντρική γραμμή και όρια ελέγχου) είναι κατάλληλα για μελλοντική παρακολούθηση της συμπεριφοράς της διεργασίας. Αυτή η χρήση των διαγραμμάτων αναφέρεται και αναδρομική (retrospective). Γενικά πολύ περισσότερες δράσεις και σκέψεις εξελίσσονται κατά την διάρκεια αυτής της φάσης πλην της απλής διαγραμματοποίησης κάποιων δεδομένων. Ο διαχειριστής της διεργασίας μελετά σε βάθος τη διεργασία και αυτό γιατί δεν είναι εύκολο κανείς να αποφασίσει αν η διεργασία ήταν εντός ή εκτός στατιστικού ελέγχου κατά τη χρονική περίοδο που τα δεδομένα συλλέχθηκαν.

Φάση II: Σε αυτή τη φάση τα διαγράμματα ελέγχου χρησιμοποιούνται προκειμένου να ελέγξουμε συνεχώς αν η διαδικασία παραμένει εντός ελέγχου. Ο διαχειριστής έχει στα χέρια του ένα πολύτιμο εργαλείο μέσω του οποίου είναι δυνατό 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 λαμβάνεται ένα δείγμα από τη διεργασία ο διαχειριστής παίρνει μια απάντηση στο ερώτημα εάν η διεργασία παραμένει εντός ελέγχου. Παράλληλα αδιαφορεί για τον τρόπο με τον οποίο προέκυψαν τα όρια ελέγχου και η κεντρική γραμμή του διαγράμματος που χρησιμοποιεί.

1.10.1 Διαγράμματα ελέγχου τύπου Shewhart

Τα διαγράμματα ελέγχου τύπου Shewhart προτάθηκαν από τον Walter A. Shewhart (1931) ο οποίος εργαζόταν ως στατιστικός στα εργαστήρια της εταιρείας Bell και αποτελούν τα πιο γνωστά αλλά και πιο συχνά χρησιμοποιούμενα διαγράμματα ελέγχου.

Αποτελούνται από δυο όρια ελέγχου, το άνω όριο ελέγχου UCL και το κάτω όριο ελέγχου LCL , και την κεντρική γραμμή του διαγράμματος CL η οποία έχει μέση τιμή ίση με το μέσο επίπεδο της διεργασίας χωρίς την παρουσία ειδικής αιτίας μεταβλητότητας (εντός ελέγχου διεργασία). Για την κατασκευή ενός διαγράμματος Shewhart συλλέγονται δείγματα μεγέθους $n \geq 1$, υπολογίζεται η τιμή κατάλληλης στατιστικής συνάρτησης W η οποία απεικονίζεται σε ένα διάγραμμα όπου τα όρια ελέγχου απέχουν απόσταση L το καθένα (σε μονάδες τυπικής απόκλισης) από την κεντρική γραμμή. Όταν ένα σημείο βρεθεί εκτός των ορίων ελέγχου τότε έχουμε ένδειξη εκτός ελέγχου διεργασίας.

Η πρόταση του Shewhart ήταν για $L = 3$ αν και έκτοτε προτάθηκαν διάφορες επιλογές για την τιμή του L με σκοπό να καλύψουν τις απαιτήσεις στην πράξη. Το διάγραμμα τύπου Shewhart για την παρακολούθηση της μέσης τιμής μιας παραγωγικής διεργασίας, το οποίο βασίζεται σε 3σ όρια ελέγχου με μετρήσεις που προέρχονται από κανονική κατανομή με γνωστές παραμέτρους είναι το πιο γνωστό και ευρέως χρησιμοποιούμενο διάγραμμα της κατηγορίας αυτής και αποτελεί μέτρο σύγκρισης με όλα τα νέα προτεινόμενα διαγράμματα παρακολούθησης της μέσης τιμής. Θα αναφερόμαστε σε αυτό ως το σύνθητες ή το τυπικό \bar{X} διάγραμμα ελέγχου Shewhart.

Το βασικό μειονέκτημα των διαγραμμάτων ελέγχου τύπου Shewhart είναι ότι δεν είναι ευαίσθητα στην ανίχνευση μικρών ή και μεσαίων μετατοπίσεων στην τιμή της υπό παρακολούθηση παραμέτρου που καθορίζει την ποιότητα των παραγόμενων προϊόντων ή προσφερόμενων υπηρεσιών. Έτσι, έχουν προταθεί διαγράμματα τα οποία χρησιμοποιούν σύνθετους κανόνες για την ανακήρυξη μιας διεργασίας εκτός ελέγχου. Ειδικότερα, ένας τρόπος να αυξηθεί η ευαισθησία του διαγράμματος είναι η χρήση προειδοποιητικών ορίων ελέγχου και κανόνων διακοπής οι οποίοι βασίζονται στη θεωρία ροών (run rules). Τέτοιοι κανόνες προτάθηκαν για πρώτη φορά από τη Western Electric Company (1956) με σκοπό την αύξηση της ευαισθησίας του \bar{X} διαγράμματος και χρησιμοποιούνται πολύ συχνά στην πράξη. Αξίζει να αναφέρουμε ότι για την ανάπτυξη ενός διαγράμματος ελέγχου τύπου Shewhart αρκεί η γνώση της κατανομής της απεικονιζόμενης στατιστικής συνάρτησης W . Έτσι, χρησιμοποιώντας την επιθυμητή τιμή για το L μπορούμε να αναπτύξουμε διαγράμματα τύπου Shewhart για την παρακολούθηση της μέσης τιμής και της διασποράς του υπό παρακολούθηση χαρακτηριστικού οποιαδήποτε και αν είναι (συνεχής ή διακριτή) η κατανομή των αρχικών παρατηρήσεων.

1.10.2 Διαγράμματα ελέγχου τύπου CUSUM

Τα διαγράμματα τύπου CUSUM προτάθηκαν από τον Page (1954). Σε αυτά απεικονίζεται η στατιστική συνάρτηση

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

με σκοπό την ανίχνευση μετατοπίσεων στη μέση τιμή της παραγωγικής διεργασίας σε υψηλότερο επίπεδο. Η τιμή μ_0 δηλώνει την εντός ελέγχου μέση τιμή ενώ με X_t συμβολίζεται η τιμή της μεμονωμένης παρατήρησης που λαμβάνεται τη χρονική στιγμή t . Επίσης απεικονίζεται και η ποσότητα

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

με σκοπό την ανίχνευση μετατοπίσεων στη μέση τιμή της παραγωγικής διεργασίας σε χαμηλότερο επίπεδο. Οι τιμές των ποσοτήτων S_0^+ και S_0^- ονομάζονται τιμές εκκίνησης (headstart values). Η ποσότητα $K \geq 0$ ονομάζεται τιμή αναφοράς (reference value) και η συνήθης τιμή της προκύπτει από τη σχέση

$$K = \delta\sigma/2 = |\mu_1 - \mu_0| \frac{1}{2}$$

όπου η τιμή μ_1 δηλώνει μια εκτός ελέγχου μέση τιμή. Η ποσότητα S_t^+ συσσωρεύει τις αποκλίσεις των παρατηρήσεων X_t από την ποσότητα $\mu_0 + K$ από τη στιγμή που θα εμφανιστεί θετική απόκλιση και μπορεί να θεωρηθεί κατάλληλη για τον έλεγχο της υπόθεσης

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

αφού μεγάλες θετικές τιμές της S_t^+ οδηγούν σε αποδοχή της υπόθεσης H_1^+ . Ανάλογα η ποσότητα S_t^- συσσωρεύει τις αποκλίσεις των παρατηρήσεων X_t από την ποσότητα $\mu_0 - K$ από τη στιγμή που θα εμφανιστεί αρνητική απόκλιση και μπορεί να θεωρηθεί κατάλληλη για τον έλεγχο της υπόθεσης

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

αφού μεγάλες αρνητικές τιμές της S_t^- οδηγούν σε αποδοχή της υπόθεσης H_1^- . Για το ποια από τις δύο εναλλακτικές υποθέσεις H_1^+ και H_1^- θα αποδεχθούμε ή όχι σε κάθε βήμα της διαδικασίας η απόφαση μας θα εξαρτηθεί από το αν ισχύει η σχέση $S_t^+ > H$ ή η σχέση $S_t^- < H$, όπου H μια θετική σταθερά ($H > 0$). Η ποσότητα H ονομάζεται διάστημα απόφασης (decision interval).

Η πιο συνηθισμένη τιμή για το διάστημα απόφασης είναι η $H = h\sigma$, $h = 4,5$. Είναι προφανές λοιπόν πως έχουμε ένδειξη εκτός ελέγχου διεργασίας (μετατόπιση της μέσης τιμής της παραγωγικής διεργασίας σε υψηλότερο ή χαμηλότερο επίπεδο) αν γίνει αποδεκτή μια εκ των εναλλακτικών υποθέσεων H_1^+ και H_1^- , αντίστοιχα. Τα διαγράμματα ελέγχου τύπου CUSUM αποτελούν εναλλακτική πρόταση έναντι των διαγραμμάτων τύπου Shewhart στην ανίχνευση μικρών ή και μεσαίων μετατοπίσεων στην τιμή της υπό παρακολούθηση παραμέτρου. Εμπίπτουν σε μια γενική κατηγορία διαγραμμάτων τα οποία ονομάζονται διαγράμματα ελέγχου με

μνήμη που το κύριο χαρακτηριστικό τους είναι η ενσωμάτωση πληροφοριών από τις τιμές των προηγούμενων παρατηρήσεων στην απεικονιζόμενη ποσότητα. Ειδικότερα, τα διαγράμματα ελέγχου τύπου CUSUM χαρακτηρίζονται ως διαγράμματα με ομοιόμορφη μνήμη αφού δίνουν την ίδια βαρύτητα σε όλες τις προηγούμενες παρατηρήσεις.

1.10.3 Διαγράμματα Ελέγχου Τύπου EWMA

Τα διαγράμματα EWMA προτάθηκαν από τον Roberts (1959) και μελετήθηκαν με λεπτομέρεια από τους Crowder (1987, 1989) και τους Lucas και Saccucci (1990). Στα διαγράμματα τύπου EWMA απεικονίζεται η ποσότητα

$$Z_t = \lambda X_t + (1 - \lambda)X_{t-1}, \quad X_0 = \mu$$

με σκοπό την παρακολούθηση της μέσης τιμής μιας παραγωγικής διεργασίας. Με X_t συμβολίζεται η τιμή της μεμονωμένης παρατήρησης που λαμβάνεται τη χρονική στιγμή t ενώ η τιμή μ δηλώνει τη μέση τιμή της κατανομής της X_t . Σε αντιστοιχία με τα διαγράμματα τύπου CUSUM η τιμή της ποσότητας X_0 ονομάζεται τιμή εκκίνησης ενώ η σταθερά λ ($0 \leq \lambda \leq 1$) καθορίζει το βάρος που δίνουμε στην τιμή του πιο πρόσφατος δείγματος.

Για την κατασκευή ενός διαγράμματος τύπου EWMA χρειαζόμαστε τη μέση τιμή μ_{Z_t} και την τυπική απόκλιση σ_{Z_t} της ποσότητας Z_t που απεικονίζεται στο διάγραμμα. Για την περίπτωση που οι αρχικές παρατηρήσεις X_t είναι ανεξάρτητες τυχαίες μεταβλητές με εντός ελέγχου μέση τιμή μ_0 και

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

Παρακάτω δίνονται τα όρια ελέγχου και η κεντρική γραμμή ενός διαγράμματος τύπου EWMA για την παρακολούθηση της μέσης τιμής μιας παραγωγικής διεργασίας.

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

$$CL = \mu_0$$

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

Για την κατασκευή ενός διαγράμματος τύπου EWMA απαιτείται ο προσδιορισμός των τιμών των παραμέτρων L και λ . Ο προσδιορισμός αυτός γίνεται συνήθως με στατιστικά κριτήρια. Η τιμή της παραμέτρου L καθορίζει την απόσταση των ορίων ελέγχου από την κεντρική γραμμή. Οι συνηθέστερες τιμές για το λ είναι το $\lambda = 0.05$, $\lambda = 0.10$ και $\lambda = 0.20$

Σημειώνουμε πως τα όρια ελέγχου ονομάζονται χρονομεταβλητά αφού υπάρχει εξάρτηση της τυπικής απόκλισης σ_{Z_t} από το χρόνο. Όμως καθώς το t αυξάνει, τα όρια ελέγχου προσεγγίζουν μια σταθερή τιμή και έτσι καταλήγουμε στα όρια ελέγχου σταθερής κατάστασης τα οποία δίνονται παρακάτω.

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

$$CL = \mu_0$$

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

Σημαντικό είναι επίσης να αναφέρουμε πως στα διαγράμματα τύπου EWMA λόγω του ότι απεικονίζονται σε αυτά σημεία που αποτελούν σταθμισμένο μέσο όρο όλων των προηγούμενων παρατηρήσεων, είναι ιδιαίτερα ανθεκτικά στην παραβίαση της υπόθεσης της κανονικότητας των αρχικών παρατηρήσεων. Έτσι λοιπόν αποτελούν τα πλέον κατάλληλα διαγράμματα για την παρακολούθηση διεργασιών μέσω μεμονωμένων παρατηρήσεων έναντι των αντίστοιχων διαγραμμάτων ελέγχου Shewhart.

1.11 Διαγράμματα ελέγχου ιδιοτήτων

Τα διαγράμματα ελέγχου ιδιοτήτων χρησιμοποιούνται όταν τα δεδομένα μας δεν αφορούν μετρήσιμες μονάδες αλλά ποσότητες με δυο πιθανά αποτελέσματα της μορφής ναι ή όχι. Το χαρακτηριστικό της ποιότητας που μας ενδιαφέρει είναι είτε οι ελαττωματικές μονάδες, είτε τα ελαττώματα μιας μονάδας. Το μέγεθος του δείγματος μπορεί να είναι σταθερό ή και να μεταβάλλεται από δείγμα σε δείγμα. Για να χαρακτηρίσουμε ένα προϊόν ως ελαττωματικό ή μη συμμορφούμενο θα πρέπει τουλάχιστον ένα ποιοτικό χαρακτηριστικό του να βρίσκεται εκτός των ορίων προδιαγραφών, τότε και λέμε ότι το προϊόν αυτό παρουσιάζει τουλάχιστον ένα ελάττωμα. Συνεπώς μπορούμε να πούμε ότι ελαττωματικές χαρακτηρίζονται οι μονάδες που θεωρούνται εντελώς αποτυχημένες, δηλαδή δεν καλύπτουν τις προϋποθέσεις λόγω παρουσίας ελαττωμάτων, ενώ ελαττώματα είναι οι ατέλειες ή τα μη συμμορφούμενα χαρακτηριστικά τα οποία προκαλούν δυσλειτουργία στις μονάδες και τις εμποδίζουν να προσαρμοστούν στις προδιαγραφές. Τέλος αναφέρουμε πως μια μονάδα μπορεί να παρουσιάσει πολλά ελαττώματα και μάλιστα όχι του ίδιου τύπου.

1.11.1 Διαγράμματα Ελέγχου p και np

Το ποιοτικό χαρακτηριστικό που μας ενδιαφέρει στα p και np διαγράμματα ελέγχου είναι αντίστοιχα το ποσοστό και ο αριθμός των ελαττωματικών προϊόντων. Λέγοντας ποσοστό προϊόντων εννοούμε το πηλίκο του αριθμού των ελαττωματικών προϊόντων προς το συνολικό αριθμό των παραγόμενων προϊόντων. Για το σχεδιασμό των διαγραμμάτων αρχικά επιλέγονται m προκαταρκτικά δείγματα

ισομεγέθη ή μη. Πρακτικά τα δείγματα της παραγωγής δεν μπορεί να είναι τα ίδια κάθε μέρα διότι επηρεάζει και ο παράγοντας κόστος κατά τον οποίο μπορεί το μέγεθος του δείγματος μπορεί να μειωθεί για οικονομικούς λόγους ή αντίστοιχα λόγω του ότι η επιθεώρηση των προϊόντων δεν είναι δαπανηρή μπορεί ή διοίκηση να αυξήσει το μέγεθος του δείγματος.

Ας υποθέσουμε ότι το ποσοστό των ελαττωματικών προϊόντων μιας παραγωγικής διαδικασίας είναι p και ότι επιλέγουμε m ανεξάρτητα τυχαία δείγματα με μεγέθη n_1, n_2, \dots, n_m αντίστοιχα. Θα συμβολίσουμε με $X_{ij}, i \geq 1, 1 \leq j \leq n_i$ την τυχαία μεταβλητή με τιμές 1 και 0 ανάλογα με το αν το j προϊόν του i δείγματος είναι ελαττωματικό ή όχι. Για την μεταβλητή X_{ij} έχουμε ότι $X_{ij} \sim B(1, p)$, ενώ για την

$$X_i = X_{i1} + X_{i2} + \dots + X_{in_i}, \quad 1 \leq i \leq m$$

που δηλώνει τον αριθμό των ελαττωματικών προϊόντων στο i δείγμα έχουμε ότι $X_i \sim B(n, p)$.

Επομένως τα όρια ελέγχου του p και np διαγράμματος ελέγχου είναι

Όρια Ελέγχου p διαγράμματος	Όρια Ελέγχου np διαγράμματος
$LCL_p = \bar{p} - 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{\bar{n}}}$	$LCL_{np} = n\bar{p} - 3\sqrt{n\bar{p}(1 - \bar{p})}$
$CL_p = \bar{p}$	$CL_{np} = n\bar{p}$
$UCL_p = \bar{p} + 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{\bar{n}}}$	$UCL_{np} = n\bar{p} + 3\sqrt{n\bar{p}(1 - \bar{p})}$

1.11.2 Διαγράμματα ελέγχου c

Το ποιοτικό χαρακτηριστικό που μας ενδιαφέρει στα c διαγράμματα είναι ο συνολικός αριθμός των ελαττωμάτων σε μια μονάδα επιθεώρησης. Η βασική υπόθεση είναι ότι ο συνολικός αριθμός των ελαττωμάτων μιας μονάδας ακολουθεί την κατανομή Poisson και σύμφωνα με αυτήν την υπόθεση, η πιθανότητα εμφάνισης ελαττώματος σε οποιοδήποτε σημείο μιας μονάδας θα πρέπει να είναι πολύ μικρή. Έτσι ο αριθμός X των ελαττωμάτων που εμφανίζονται σε μια μονάδα επιθεώρησης ακολουθεί την κατανομή Poisson με παράμετρο c . Δηλαδή $X \sim P(c)$ Έτσι λοιπόν μπορούμε να αναπτύξουμε ένα διάγραμμα ελέγχου για την παρακολούθηση του αριθμού των ελαττωμάτων των μονάδων, στο οποίο θα απεικονίζεται η στατιστική συνάρτηση $X_i, i \geq 1$, όπου δηλώνει τον αριθμό των ελαττωμάτων στην i μονάδα επιθεώρησης. Τα όρια ελέγχου δίνονται από τον παρακάτω πίνακα

Όρια Ελέγχου c διαγράμματος
$LCL_p = \bar{c} - 3\sqrt{\bar{c}}$
$CL_p = \bar{c}$
$UCL_p = \bar{c} + 3\sqrt{\bar{c}}$

1.11.3 Διάγραμμα ελέγχου u

Το ποιοτικό χαρακτηριστικό που μας ενδιαφέρει στα u διαγράμματα είναι η αναλογία των ελαττωματικών ανά δείγμα, ή αλλιώς το μέσο αριθμό των ελαττωμάτων ανά μονάδα επιθεώρησης σε κάθε δείγμα. Η βασική υπόθεση που θα κάνουμε είναι ότι ο αριθμός X των ελαττωμάτων σε μια μονάδα επιθεώρησης ακολουθεί την κατανομή Poisson με παράμετρο c . Η βασική διαφορά με το c διάγραμμα ελέγχου είναι ότι στα u διαγράμματα ελέγχου μπορούμε να έχουμε δείγματα μεγέθους μεγαλύτερου της μιας μονάδας επιθεώρησης.

Υποθέτοντας ότι από την παραγωγή επιλέγουμε m ανεξάρτητα τυχαία δείγματα με μεγέθη n_1, n_2, \dots, n_m μονάδες επιθεώρησης αντίστοιχα.

Συμβολίζοντας με $X_{ij}, i \geq 1, 1 \leq j \leq n_i$ την τυχαία μεταβλητή που δηλώνει τον αριθμό των ελαττωμάτων της j μονάδας επιθεώρησης στο i δείγμα. Για την τυχαία μεταβλητή X_{ij} έχουμε ότι $X_{ij} \sim P(c)$ ενώ για την

$$X_i = X_{i1} + X_{i2} + \dots + X_{in_i}, \quad 1 \leq i \leq m$$

που δηλώνει τον αριθμό των ελαττωμάτων στο i δείγμα έχουμε ότι $X_i \sim P(n_i c)$ συνεπώς μπορούμε να αναπτύξουμε ένα διάγραμμα ελέγχου για την παρακολούθηση του μέσου αριθμού των ελαττωμάτων ανά μονάδα επιθεώρησης σε κάθε δείγμα.

Τα όρια ελέγχου του διαγράμματος ελέγχου απεικονίζονται στον παρακάτω πίνακα.

Όρια Ελέγχου u διαγράμματος
$LCL_u = \bar{u} - 3\sqrt{\frac{\bar{u}}{\bar{n}}}$
$CL_u = \bar{u}$
$UCL_u = \bar{u} + 3\sqrt{\frac{\bar{u}}{\bar{n}}}$

Κεφάλαιο 2

2.1 Εισαγωγή

Τα μη παραμετρικά ή ελεύθερα κατανομής διαγράμματα ελέγχου, όπως έχει ήδη αναφερθεί, έχουν ιδιότητες που δεν εξαρτώνται από την κανονικότητα ή κάποια άλλη συγκεκριμένη παραμετρική κατανομή. Αξίζει να σημειωθεί ότι ο όρος «μη παραμετρικά» δεν δηλώνει ότι στο σχεδιασμό των διαγραμμάτων αυτών δεν περιλαμβάνονται παράμετροι και κατά συνέπεια ο όρος «ελεύθερα κατανομής» φαίνεται να προσδιορίζει καλύτερα τη φύση των διαγραμμάτων αυτών. Για την ανάπτυξη και θεωρητική μελέτη των μη παραμετρικών διαγραμμάτων ελέγχου χρησιμοποιούνται κατά βάση οι κλασικές μη παραμετρικές τεχνικές συμπερασματολογίας που αντιμετωπίζουν αποτελεσματικά το πρόβλημα ελέγχου υποθέσεων όταν η κατανομή δεν έχει καθορισθεί.

Το σημαντικότερο πλεονέκτημα των μη παραμετρικών διαγραμμάτων ελέγχου είναι ότι η εφαρμογή τους δεν απαιτεί την υπόθεση κάποιας συγκεκριμένης κατανομής για την υπό μελέτη διεργασία. Για το λόγο αυτό, το εντός ελέγχου μέσο μήκος ροής είναι το ίδιο για όλες τις συνεχείς κατανομές. Επιπλέον, η διακύμανση της διεργασίας δεν χρειάζεται να είναι γνωστή ή να εκτιμηθεί, ενώ επιπλέον τα διαγράμματα αυτά παρουσιάζουν αντοχή στην παρουσία έκτροπων παρατηρήσεων και σε περιπτώσεις που ο υπό μελέτη πληθυσμός ακολουθεί λοξή ή με βαριές ουρές κατανομή. Τα μη παραμετρικά διαγράμματα ελέγχου παρουσιάζονται κατά την εμφάνιση του προβλήματος ανίχνευσης μιας μετατόπισης του μέσου μιας διεργασίας ή μιας αλλαγής στην διακύμανση μιας διεργασίας, όταν τα δείγματα που χρησιμοποιούνται έχουν παρθεί ανά τακτά διαδοχικά χρονικά διαστήματα. Ένα πρόσθετο πλεονέκτημα των μη παραμετρικών διαγραμμάτων ελέγχου είναι ότι για τη δημιουργία ενός διαγράμματος ελέγχου για το μέσο μιας διεργασίας δε χρειάζεται ο υπολογισμός της διακύμανσης της διεργασίας αυτής.

2.2 Σημαντικά μη παραμετρικά διαγράμματα ελέγχου

Για πολλά χρόνια τα διαγράμματα ελέγχου χρησιμοποιήθηκαν για την ανίχνευση μιας μεταβολής του μέσου μιας κατανομής, από παρατηρήσεις που λαμβάνονται ως μια ακολουθία δειγμάτων ανά τακτά διαδοχικά χρονικά διαστήματα από μια διεργασία. Τα περισσότερα διαγράμματα ελέγχου βασίζονται στους μέσους των δειγμάτων και έχουν σχεδιαστεί με βάση την υπόθεση ότι οι παρατηρήσεις του κάθε δείγματος ακολουθούν την κανονική κατανομή και κατά συνέπεια ο δειγματικός μέσος ακολουθεί επίσης την κανονική κατανομή. Αν η κατανομή των παρατηρήσεων δεν είναι κανονική τότε χρησιμοποιείται το κεντρικό οριακό θεώρημα για να δικαιολογήσει την υπόθεση ότι η κατανομή που ακολουθεί ο δειγματικός μέσος είναι κατά προσέγγιση κανονική.

Υπάρχουν, ωστόσο, αρκετά πιθανά μειονεκτήματα των διαγραμμάτων ελέγχου που βασίζονται στους δειγματικούς μέσους. Όταν η πραγματική κατανομή των παρατηρήσεων δεν είναι κοντά στην κανονική κατανομή, τότε, η κατανομή του δειγματικού μέσου μπορεί να είναι αρκετά διαφορετική από την κανονική κατανομή, έτσι ώστε οι υπολογισμοί της κανονικής θεωρίας για τις ιδιότητες του

διαγράμματος να μην είναι πια αξιόπιστοι. Ένα άλλο πρόβλημα είναι ότι η αποτελεσματικότητα του δειγματικού μέσου μπορεί να είναι σχετικά χαμηλή για κάποιες μη κανονικές κατανομές, με την έννοια ότι ο αναμενόμενος χρόνος για την ανίχνευση μιας εκτός ελέγχου κατάστασης είναι μεγαλύτερος για το δειγματικό μέσο απ' ότι για μια άλλη στατιστική παράμετρο. Επιπλέον, τα διαγράμματα ελέγχου βασισμένα σε δειγματικούς μέσους απαιτούν τη γνώση ή τον ακριβή υπολογισμό της διακύμανσης της διεργασίας. Όταν μια διεργασία έχει μόλις ξεκινήσει είναι πολύ πιθανό να μην υπάρχουν αρκετά δεδομένα για τον υπολογισμό μιας καλής εκτίμησης της διακύμανσης.

Έχουν αναπτυχθεί διαγράμματα ελέγχου που χρησιμοποιούν τη στατιστική των προσήμων αντί του δειγματικού μέσου, όταν ο στόχος είναι η ανίχνευση μιας μεταβολής του μέσου της κατανομής των παρατηρήσεων. Ένα διάγραμμα ελέγχου βασισμένο στη στατιστική των προσήμων μπορεί να αναπτυχθεί επίσης για την ανίχνευση μιας αύξησης στη διακύμανση μιας διεργασίας. Τα διαγράμματα ελέγχου των προσήμων (sign control charts) είναι εύκολα στη χρήση και στην κατανόηση τους, με τα οποία και αποφεύγουμε τα μειονεκτήματα που αναφέρθηκαν προηγουμένως για τα αντίστοιχα διαγράμματα ελέγχου βασισμένα στους δειγματικούς μέσους.

Στην περίπτωση των διαγραμμάτων που είναι βασισμένα στα πρόσημα, το *ARL* θα είναι ίδιο για κάθε κατανομή όπου ο μέσος θα ισούται με την τιμή στόχο. Αν η κατανομή είναι συμμετρική τότε ο μέσος και ο διάμεσος θα ταυτίζονται, εφόσον θα υπάρχει μέσος. Τα μη παραμετρικά διαγράμματα ελέγχου λοιπόν, έχουν το πλεονέκτημα ότι η υπόθεση της κανονικότητας δεν είναι αναγκαία για τον υπολογισμό των ορίων ελέγχου. Ένα άλλο πλεονέκτημα είναι ότι τα μη παραμετρικά διαγράμματα ελέγχου είναι πιο αποτελεσματικά από τα διαγράμματα βασισμένα στο δειγματικό μέσο όταν η κατανομή των παρατηρήσεων είναι "heavy-tailed", δηλαδή, όταν οι παρατηρήσεις στην ουρά της κατανομής έχουν μεγαλύτερη πιθανότητα απ' ότι θα είχαν στην κανονική κατανομή. Μια τέτοια κατάσταση είναι πολύ συνήθης σε πολλές εφαρμογές όπου περιστασιακά, ακραίες τιμές φαίνεται να συναντώνται πιο συχνά απ' ότι θα τις συναντούσαμε αν η κατανομή ήταν κανονική. Ένα επιπλέον πλεονέκτημα των μη παραμετρικών διαγραμμάτων ελέγχου είναι ότι η διακύμανση της διεργασίας δε χρειάζεται να είναι γνωστή ούτε χρειάζεται να υπολογιστεί για να εφαρμοστεί το διάγραμμα ελέγχου. Στην πραγματικότητα, τα μη παραμετρικά διαγράμματα ελέγχου για τον έλεγχο του μέσου, δεν επηρεάζονται από τις μεταβολές της διακύμανσης, εφόσον το μ παραμένει σταθερό. Οι Ghosh, Reynolds και Hui (1981) και ο Quesenberry (1993), απέδειξαν ότι χρησιμοποιώντας μικρότερο αριθμό παρατηρήσεων για την εκτίμηση της διακύμανσης, όταν θέτονται τα όρια για το \bar{X} -διάγραμμα, οι ιδιότητες του \bar{X} -διαγράμματος μπορεί να μεταβληθούν σημαντικά. Έτσι λοιπόν, τα μη παραμετρικά διαγράμματα ελέγχου, μπορεί να είναι ιδιαίτερα χρήσιμα κατά την έναρξη μιας διεργασίας, όπου θέλουμε να εφαρμόσουμε ένα διάγραμμα ελέγχου χωρίς όμως να υπάρχουν αρκετά δεδομένα για τη εκτίμηση της διακύμανσης της διεργασίας. Οι συνέπειες μιας εσφαλμένης υπόθεσης για την κανονικότητα της κατανομής μιας διεργασίας είναι πιο σφοδρές στα διαγράμματα ελέγχου για τη διακύμανση απ' ότι για το μέσο της διεργασίας

Τα βασικά χαρακτηριστικά των προσημικών διαγραμμάτων ελέγχου μπορούν να συνοψιστούν ως εξής:

- Τα προσημικά διαγράμματα μπορούν να σχεδιαστούν για να ελέγχουν το μέσο και τη διακύμανση μιας διεργασίας. Αυτά τα διαγράμματα βασίζονται σε πρόσημα υπολογισμένα μέσα από τα δείγματα. Είναι εύκολα στη χρήση και στην κατανόηση. Στην περίπτωση των προσημικών διαγραμμάτων που χρησιμοποιούνται για εντοπισμό, γίνεται μια εκτίμηση του μέσου της διεργασίας, και τοποθετείται ένα (+) όταν η τιμή μιας παρατήρησης υπερβαίνει την εντός ελέγχου τιμή του μέσου. Στα αντίστοιχα διαγράμματα που αφορούν τη διακύμανση, το πρόσημο (+) τοποθετείται όταν η διακύμανση πέσει κάτω από το πρώτο τεταρτημόριο Q_1 ή ανέβει πάνω από το τρίτο τεταρτημόριο Q_3 .
- Τα προσημικά διαγράμματα δεν απαιτούν την υπόθεση της κανονικότητας και επίσης δε χρειάζεται να εκτιμηθεί η διακύμανση της διεργασίας ούτως ώστε να τεθούν τα κατάλληλα όρια ελέγχου. Όταν η κατανομή των δεδομένων είναι άγνωστη, είναι καλύτερο να χρησιμοποιηθεί ένα προσημικό διάγραμμα, δεδομένου ότι το εντός – ελέγχου μέσο μήκος ροής (ARL) δεν εξαρτάται από τη γνώση των παραμέτρων της υποκείμενης κατανομής. Έτσι, τα διαγράμματα αυτά είναι ιδιαίτερα χρήσιμα κατά την εκκίνηση της διεργασίας.
- Τα προσημικά διαγράμματα είναι πιο αποτελεσματικά από τα αντίστοιχα παραμετρικά διαγράμματα όταν η κατανομή της διεργασίας έχει βαριά ουρά.
- Παρόλο που τα προσημικά διαγράμματα απαιτούν δείγματα μεσαίου μεγέθους, μεγαλύτερα συνήθως των 4 ή 5 είναι πιθανό να χρησιμοποιηθούν σχέδια μειωμένων δειγμάτων για να αποκτηθούν δείγματα μικρότερου μεγέθους.

2.2.1 Η επίδραση της μη κανονικότητας στα \bar{X} -διαγράμματα

Για την πληρότητα της ορολογίας και τον ορισμό του προβλήματος, θεωρούμε μια κατάσταση στην οποία έχουμε πάρει από μια διεργασία, διαδοχικά δείγματα από n παρατηρήσεις. Εστω $X_i = X_{i1}, X_{i2}, \dots, X_{in}$ το δείγμα που πήραμε την στιγμή i . Υποθέτουμε ότι η κατανομή των παρατηρήσεων είναι συνεχής με μέσο μ . Στόχος του διαγράμματος ελέγχου είναι η γρήγορη ανίχνευση οποιασδήποτε μετατόπισης του μέσου μ από την συγκεκριμένη τιμή στόχο μ_0 που έχει τεθεί, χρησιμοποιώντας αποκλειστικά τις πληροφορίες των δειγμάτων που υπάρχουν ως εκείνη τη δεδομένη χρονική στιγμή. Ορισμένες φορές η τιμή στόχος μ_0 είναι άγνωστη και στη θέση της χρησιμοποιείται ο εντός ελέγχου διάμεσος του δείγματος Q_2 . Τα διαγράμματα ελέγχου αξιολογούνται και συγκρίνονται με βάση το μέσο μήκος ροής (ARL), που είναι ο αναμενόμενος αριθμός δειγμάτων που απαιτείται από τη διαδικασία για να δώσει το σήμα ότι μια μεταβολή στο μ μόλις συνέβη. Όσο το μ παραμένει ίσο με μ_0 , το ARL πρέπει να είναι μεγάλο ώστε η συχνότητα των λανθασμένων σημάτων να είναι μικρή. Αν όμως το μ μεταβληθεί από την τιμή του μ_0 , το ARL (που υπολογίζεται από τη στιγμή της μεταβολής) πρέπει να είναι μικρό ώστε η μεταβολή να ανιχνευθεί γρήγορα.

Το \bar{X} -διάγραμμα Shewhart βασίζεται στους μέσους των δειγμάτων $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n$ όπου $\bar{X}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$. Αυτό το διάγραμμα δείχνει ότι το μ μετατοπίστηκε από το μ_0 στη πρώτη i -οστή παρατήρηση για την οποία το \bar{X}_i ξεπέρασε $\mu_0 + \alpha_1$, όπου η σταθερά α_1 ισούται με $3\sigma/\sqrt{n}$. Η παράμετρος σ υπολογίζεται συνήθως από τη μέση τυπική απόκλιση του δείγματος S ή το μέσο εύρος του δείγματος \bar{R} , που αποκτήθηκε από έναν αριθμό αρχικών δειγμάτων, όταν η διεργασία ήταν εντός ελέγχου. Το ARL ή $L(\mu)$, εκφραζόμενο ως συνάρτηση του μ , μπορεί να υπολογιστεί από την σχέση

$$L(\mu) = 1/P(|\bar{X}_i| \geq \mu_0 + \alpha_1 | \mu)$$

όσο το α_1 είναι σταθερό και η κατανομή του \bar{X}_i είναι γνωστή.

Ο σχεδιασμός των διαγραμμάτων Shewhart με τη χρήση του \bar{X}_i βασίζεται συνήθως στην υπόθεση ότι η κατανομή των \bar{X}_i είναι κατά προσέγγιση κανονική, ακόμα και αν στην πραγματικότητα η κατανομή των παρατηρήσεων δεν είναι κανονική.

Ένα επιπλέον πρόβλημα με τον σχεδιασμό των \bar{X} -διαγραμμάτων προκύπτει όταν η διακύμανση της διεργασίας είναι άγνωστη και πρέπει να εκτιμηθεί από τα υπάρχοντα δεδομένα. Οι συνέπειες της μη κανονικότητας του \bar{X} έχουν μελετηθεί από τους Burr (1967), Schilling και Nelson (1976), και Bradley (1971, 1973). Ο Hillier (1969), και οι Yang και Hillier (1970) εξέτασαν το πρόβλημα όπου ο α_1 βασίζεται σε μια εκτίμηση της διακύμανσης. Για κατανομές που δεν ακολουθούν την κανονική κατανομή, η κατανομή των \bar{X}_i μπορεί είτε να συγκλίνει πολύ αργά στην κανονική κατανομή, είτε και καθόλου.

Ο πίνακας 1 δίνει τις τιμές ARL για τα \bar{X} -διάγραμμα Shewhart με $n = 10$ και $\alpha_1 = 3\sigma/\sqrt{n}$ για την ομοιόμορφη, κανονική, γάμμα, εκθετική και Cauchy κατανομή.

Πίνακας 2.1: Τιμές $L^+(\mu)$ για τα διαγράμματα Shewhart χρησιμοποιώντας τα \bar{X}_i όταν $n = 10$ και $\alpha_1 = 3\sigma/\sqrt{n}$ για διάφορες κατανομές ($\sigma=1$)

Shift ($\mu - \mu_0$)	Distribution					
	Uniform	Normal	Double Exponential	Gamma	Exponential	Cauchy*
.00	1068.7	740.8	441.9	268.3	148.9	11.7
.25	78.2	73.7	65.8	49.1	38.0	8.8
.50	12.7	12.8	13.2	11.9	11.4	6.0
1.00	1.8	1.8	1.8	1.8	1.9	1.8
2.00	1.0	1.0	1.0	1.0	1.0	1.1

* το σ δεν ορίζεται για την Cauchy κατανομή

Χρησιμοποιείται μόνο το ανώτατο όριο ελέγχου για διάφορες μετατοπίσεις του μ . Οι τιμές του ARL υπολογίστηκαν με την υπόθεση ότι η διακύμανση της κατανομής είναι γνωστή και δεν χρειάστηκε να εκτιμηθεί. Για όλες τις κατανομές, εκτός από την κατανομή Cauchy, η βασική παράμετρος προσαρμόστηκε ώστε να δίνει $\sigma = 1$. Το ARL της ομοιόμορφης κατανομής έχει συμπεριληφθεί στο πίνακα ως παράδειγμα κατανομών που δεν έχουν βαριά ουρά.

Το ARL έχει υπολογιστεί για την εκθετική κατανομή, που είναι μια συμμετρική κατανομή με βαρύτερη ουρά απ' ότι η κανονική κατανομή.

Η κατανομή γάμμα έχει χρησιμοποιηθεί επίσης, αφού είναι μια ασύμμετρη κατανομή για την οποία η κατανομή των \bar{X}_i μπορεί εύκολα να αξιολογηθεί. Το όριο ελέγχου τοποθετείται κατά α_1 μονάδες πάνω από το μέσο. Χρησιμοποιώντας ένα όριο ελέγχου α_1 μονάδες πάνω από το μέσο θα δώσει χαμηλότερες τιμές για το ARL .

Η κατανομή Cauchy (μια t-κατανομή με ένα βαθμό ελευθερίας) χρησιμοποιήθηκε επίσης γιατί είναι μια συμμετρική κατανομή με εξαιρετικά βαριά ουρά και η κατανομή των X_i είναι ίδια με την κατανομή του X_{ij} . Η τιμή της βασικής παραμέτρου της κατανομής επιλέχθηκε έτσι ώστε η κατανομή να έχει πιθανότητα 0.05 να βρίσκεται πάνω από το $\mu + 1.645$ (το ίδιο με την κανονική κατανομή με διακύμανση 1 και μέσο μ).

Από τον πίνακα 2.1 βλέπουμε ότι για τις κατανομές με βαριά ουρά, η πραγματική τιμή του ARL τείνει να είναι μικρότερη από αυτήν που θα υπολογίζαμε αν το \bar{X} ακολουθούσε κανονική κατανομή. Από την άλλη, για την ομοιόμορφη κατανομή η οποία δεν έχει βαριά ουρά οι τιμές του ARL είναι μεγαλύτερες από ότι όταν έχουμε κανονική κατανομή. Η διαφορά μεταξύ της πραγματικής τιμής ARL και της τιμής ARL για την κανονική κατανομή είναι ιδιαιτέρως έντονη στις εντός ελέγχου καταστάσεις, όταν η κατανομή έχει πολύ βαριά ουρά. Αυτό σημαίνει ότι για τις κατανομές με πολύ βαριές ουρές, θα υπάρχουν συχνές λανθασμένες ειδοποιήσεις, ενώ η διεργασία θα εκτελείται κανονικά. Για παράδειγμα, όταν οι παρατηρήσεις ακολουθούν μια εκθετική κατανομή, το ARL θα είναι μόνο 148.9 όταν $\mu = \mu_0$, και μακροπρόθεσμα, οι λανθασμένες προειδοποιήσεις θα είναι πέντε φορές περισσότερες από ότι όταν θα είχαμε τιμή $ARL = 740.8$ και κανονική κατανομή. Σε πολλές περιπτώσεις, το μέγεθος των δειγμάτων είναι μικρότερο του 10 και το σφάλμα που δημιουργείται χρησιμοποιώντας την κατά προσέγγιση κανονική κατανομή για τις τιμές των X_i μπορεί να είναι μεγαλύτερο από την αντίστοιχη περίπτωση που υποδεικνύει ο πίνακας 2.1.

2.2.2 Προσημικά διαγράμματα ελέγχου για την καταγραφή και την παρακολούθηση του μέσου

Ας υποθέσουμε ότι παρακολουθούμε μια παραγωγική διαδικασία και έστω $X_{i1}, X_{i2}, \dots, X_{in}$, $i = 1, 2, \dots, n$ δείγματα μεγέθους n , που λαμβάνουμε από τη διεργασία (X_{ij} είναι η j -οστή παρατήρηση του i -οστού δείγματος που λαμβάνεται από την παραγωγή) με σκοπό να ανιχνεύσουμε αν παραμένει εντός ελέγχου ή έχει μετατοπισθεί σε εκτός ελέγχου κατάσταση. Ο Amin (1995) προκειμένου να κατασκευάσουν ένα μη παραμετρικό διάγραμμα ελέγχου τύπου Shewhart χρησιμοποίησαν τη προσημική στατιστική συνάρτηση (sign test statistic)

$$SN_i = \sum_{j=1}^n \text{sign}(X_{ij} - \mu_0), \text{ όπου}$$

$$\text{sign}(x) = \begin{cases} 1 & , \text{αν } x > 0 \\ 0 & , \text{αν } x = 0 \\ -1 & , \text{αν } x < 0 \end{cases}$$

και μ_0 είναι η εντός ελέγχου τιμή της διαμέσου της διεργασίας. Αν T_i είναι η συνήθης προσημική συνάρτηση (usual sign statistic) που μετρά το πλήθος των παρατηρήσεων των δειγμάτων που είναι μεγαλύτερες ή ίσες από τη διάμεσο μ_0 , τότε είναι γνωστό ότι

$$SN_i = 2T_i - n.$$

Δεδομένου ότι κατανομή της T_i είναι η διωνυμική με παραμέτρους n και $p = 0.5$, έπεται ότι η εντός ελέγχου κατανομή της SN_i είναι συμμετρική γύρω από το μηδέν και συνεπώς τα όρια ελέγχου του προσημικού διαγράμματος (sign control chart) που προκύπτει θα είναι $LCL = -c$ και $UCL = c$ αντίστοιχα, όπου c είναι ένας θετικός ακέραιος αριθμός.

Στην περίπτωση που το $|SN_i| \geq \alpha_2$ όπου $\alpha_2 > 0$ είναι μια σταθερά τότε η μέθοδος μας ειδοποιεί ότι μια μετατόπιση έχει πραγματοποιηθεί στο μέσο. Το ARL για την αμφίπλευρη περίπτωση είναι

$$L(\mu) = 1/P(|SN_i| \geq \alpha_2)$$

με την πιθανότητα να υπολογίζεται εύκολα για οποιαδήποτε κατανομή. Στην μονόπλευρη περίπτωση το ARL δίνεται από

$$L^+(\mu) = 1/P(SN_i \geq \alpha_2)$$

Το ARL στην εντός ελέγχου κατάσταση θα έχει την ίδια τιμή για όλες τις κατανομές με μέσο μ_0 .

Από τη στιγμή που υποτίθεται πως η κατανομή των παρατηρήσεων είναι συνεχής, $P(X_{ij} - \mu_0 = 0) = 0$ η περίπτωση κατά την οποία $sign(X_{ij} - \mu_0) = 0$ θα πρέπει θεωρητικά να μην εμφανιστεί ποτέ. Στην πραγματικότητα βέβαια, κάποιες παρατηρήσεις μπορεί να στρογγυλοποιηθούν, και περιστασιακά κάποια μηδενικά ίσως να εμφανιστούν. Όσο τα μηδενικά δεν εμφανίζονται συχνά, είναι ασφαλές να υπολογίσουμε το SN_i όπως έχει οριστεί. Η εφαρμογή των προσημικών τεστ είναι πολύ εύκολη αφού η αριθμητική αξία των παρατηρήσεων δε χρειάζεται να οριστεί επ' ακριβώς παρά μόνο όταν αυτές βρίσκονται πάνω και κάτω από το μ_0 . Το τεστ πρέπει να είναι ιδιαίτερα χρήσιμο όταν τα δεδομένα για την κατανομή των παρατηρήσεων είναι λίγα, ή όταν είναι γνωστό ότι η κατανομή δεν είναι κανονική. Από τη στιγμή που τα μεγέθη των παρατηρήσεων δεν χρησιμοποιούνται, η SN_i θα είναι πιο αναποτελεσματική συγκριτικά με την \bar{X}_i όταν η κατανομή των παρατηρήσεων είναι σχεδόν κανονική. Από την άλλη πλευρά αυτό μπορεί να αντιστραφεί για κατανομές με βαριές ουρές.

Οι μεγαλύτερες πιθανές εντός ελέγχου τιμές για το ARL σε ένα συμμετρικό μονόπλευρο και αμφίπλευρο έλεγχο είναι $2n$ και $2n - 1$ αντίστοιχα, όταν $p = 1/2$. Αν το δείγμα n είναι μετρίου μεγέθους μπορεί να είναι δύσκολο να προσδιοριστεί, έστω και προσεγγιστικά, μια τιμή για το $L^+(\mu_0)$. Στην πραγματικότητα, η μεγαλύτερη δυνατή τιμή για το $L^+(\mu_0)$ είναι 1024, όταν το $n = 10$ και $\alpha_2 = 10$. Η αποτελεσματικότητα των διαγραμμάτων Shewhart για την ανίχνευση μιας μεταβολής σε μια παράμετρο που μας ενδιαφέρει, μπορεί να βελτιωθεί αν μέσα στα όρια ελέγχου ορίσουμε κάποια επιπλέον όρια προειδοποίησης. Για τα διαγράμματα Shewhart που χρησιμοποιούν τα \bar{X}_i , τα προειδοποιητικά όρια είναι $\mu_0 \pm w_1$ όπου w_1 είναι μια σταθερά $0 \leq w_1 < a_1$. Μια προειδοποίηση πραγματοποιείται όταν r συνεχόμενα σημεία πέσουν στην περιοχή μεταξύ $\mu_0 - a_1$ και $\mu_0 - w_1$ ή αν κάποιο σημείο βρεθεί εκτός των ορίων ελέγχου.

Οι ιδιότητες των \bar{X} –διαγραμμάτων με όρια προειδοποίησης μελετήθηκαν από τους Page (1962), Weindling, Littauer, και Oliveira (1979), και Champ και Woodall (1987).

Στην περίπτωση που το SN_i χρησιμοποιείται στα διαγράμματα Shewhart, τα όρια προειδοποίησης είναι ακέραιοι $\pm w_2$ όπου $0 \leq w_2 < a_2$. Μια προειδοποίηση πραγματοποιείται όταν r συνεχόμενα σημεία ικανοποιούν την ανίσωση $w_2 \leq SN_i < a_2$ ή την ανίσωση $-a_2 < SN_i < -w_2$ ή όταν οποιοδήποτε σημείο βρίσκεται εκτός των ορίων ελέγχου που ορίζετε από το a_2 . Το ARL ενός μονόπλευρου διαγράμματος με όρια ελέγχου και προειδοποίησης προς τα θετικά (Page, 1962) δίνονται από την παρακάτω σχέση

$$L^+(\mu) = \frac{1 - p_1^r}{1 - p_1 - p_0(1 - p_1^r)}$$

με $p_0 = P\{SN_i < w_2 | \mu\}$ και $p_1 = P\{w_2 \leq SN_i < a_2 | \mu\}$.
Το ARL των αμφίπλευρων διαγραμμάτων είναι

$$L(\mu) = \frac{L^-(\mu)L^+(\mu)}{L^-(\mu) + L^+(\mu)}$$

με $L^-(\mu)$ να είναι το ARL του διαγράμματος με τα όρια ελέγχου και τα προειδοποιητικά όρια να βρίσκονται στα αρνητικά του διαγράμματος και υπολογίζεται από τη σχέση

$$L^-(\mu) = \frac{1 - p_1^r}{1 - p_1 - p_0(1 - p_1^r)}$$

με $p_0 = P\{SN_i > -w_2 | \mu\}$ και $p_1 = P\{-a_2 \leq SN_i < -w_2 | \mu\}$.
Ο πίνακας 2.2 που ακολουθεί δίνει τις τιμές του $L^+(\mu_0)$ για διάφορες τιμές των a_2 , w_2 και r . Στα προσημικά διαγράμματα για τον έλεγχο του μέσου, η τιμή του n είναι 10. Τιμές του ARL για άλλες τιμές του n μπορούν να κατασκευαστούν με τις μεθόδους που παρουσιάστηκαν. Επειδή το $L^+(\mu_0)$ για τα διαγράμματα Shewhart που χρησιμοποιούν το SN_i εξαρτάται από την κατανομή των παρατηρήσεων, μόνο μέσω του p_1 οι τιμές του $L^+(\mu_0)$ είναι ίδιες για κάθε κατανομή με $p = 1/2$. Αν το w_2 είναι κοντά στο a_2 και το r είναι σχετικά μεγάλο, η εισαγωγή των ορίων προειδοποίησης θα έχει μικρές επιπτώσεις στο $L^+(\mu_0)$ αλλά θα μπορούσε να μειώσει σημαντικά το $L^+(\mu)$ για μικρές μεταβολές του μ .

Πίνακας 2.2: Τιμές του $L^+(\mu_0)$ για τα διαγράμματα Shewhart χρησιμοποιώντας τη προσημική στατιστική

w_2	$a_2 = 8$				$a_2 = 10$			
	0	2	4	6	2	4	6	8
2	4.1	9.2	30.2	79.4	9.6	38.6	269.2	933.7
3	7.9	23.0	70.1	92.4	27.8	194.7	890.3	1023.0
4	13.5	44.7	88.4	93.1	73.0	593.7	1015.8	1023.0
5	21.2	66.9	92.3	93.1	175.4	911.2	1023.6	1024.0
6	31.0	81.5	93.0	93.1	364.4	1002.0	1024.0	1024.0
7	42.3	88.5	93.1	93.1	609.8	1020.3	1024.0	1024.0

Ο πίνακας 2.3 που ακολουθεί συγκρίνει τρεις διεργασίες, μια χωρίς όρια προειδοποίησης και δυο με όρια προειδοποίησης. Το ARL εκφράζεται ως συνάρτηση του p . Η τιμή του μ που δίνει συγκεκριμένη τιμή στο p εξαρτάται από την κατανομή των παρατηρήσεων. Η παρουσία των μηδενικών αλλάζει την κατανομή των SN_i , αλλά οι επιπτώσεις δε θα πρέπει να είναι μεγάλες εφόσον η πιθανότητα εμφάνισης του μηδενός είναι μικρή. Έτσι, στη προκειμένη περίπτωση, τα ίδια όρια ελέγχου μπορούν να χρησιμοποιηθούν χωρίς να επηρεάζουν σημαντικά τις ιδιότητες της διεργασίας. Στην περίπτωση εμφάνισης αρκετών μηδενικών, το ένα μηδενικό μπορεί να μετρηθεί ως μείον (-) και το άλλο ως συν (+).

Πίνακας 2.3: Τιμές του $L^+(\mu)$ σαν συνάρτηση του p για τα διαγράμματα Shewhart χρησιμοποιώντας την προσημική στατιστική όταν το $n = 10$

p	0.50	0.60	0.70	0.80	0.90	0.95
$a_2=10$	1024.0	165.4	35.4	9.3	2.9	1.7
$a_2=10, w_2=4, r=6$	1002.8	127.8	19.5	5.9	2.7	1.7
$a_2=10, w_2=6, r=4$	1015.8	151.1	25.9	6.3	2.5	1.6

2.2.3 Ένα προσημικό διάγραμμα ελέγχου τύπου Shewhart για την παρακολούθηση της διακύμανσης μιας διεργασίας

Όταν χρειαστεί να παρακολουθήσουμε τη διακύμανση και τον μέσο μιας κατανομής, τότε χρησιμοποιείται ένα διάγραμμα για τη διακύμανση όπως τα διαγράμματα του R , S , ή S^2 (Wetherill και Brown 1991). Το R διάγραμμα είναι λιγότερο αποτελεσματικό από το αντίστοιχο S^2 διάγραμμα όταν η κατανομή είναι κανονική. Το S^2 διάγραμμα για την καταγραφή μιας αύξησης στη διακύμανση μιας διεργασίας, δίνει σήμα αν το S^2 υπερβεί το όριο ελέγχου $\alpha_3 = \sigma^2 x_a^2$, όπου το x_a^2

υποδεικνύει τα άνω α -ποσοστιαία σημεία με $n - 1$ βαθμούς ελευθερίας, και σ^2 να είναι μια εκτίμηση της διακύμανσης της διεργασίας.

Είναι σύνηθες να υπολογίζουμε το σ μιας διεργασίας και μετά να χρησιμοποιούμε το κατάλληλο ποσοστιαίο σημείο της χ^2 κατανομής για να θέσουμε το όριο ελέγχου. Το ARL για το S^2 διάγραμμα εξαρτάται σημαντικά από τη χρήση σωστών σταθερών για τη δημιουργία των ορίων ελέγχου. Προς το παρόν δεν υπάρχει κάποια μελέτη σχετικά με τα όρια ελέγχου των διαγραμμάτων ελέγχου της διακύμανσης, για την περίπτωση όπου η κατανομή δεν είναι κανονική.

Οι συνέπειες της μη κανονικότητας είναι μεγαλύτερες στα διαγράμματα ελέγχου της διακύμανσης, απ' ό,τι στην περίπτωση των διαγραμμάτων του μέσου. Η διακύμανση του S^2 δίδεται από τον τύπο

$$V(S^2) = \frac{2\sigma^4}{(n-1)} \left(1 + \frac{\gamma_2(n-1)}{2n}\right)$$

όπου γ_2 είναι ο συντελεστής κύρτωσης. Οι Wetherill και Brown (1991) επισημαίνουν ότι η κύρτωση της αρχικής κατανομής μπορεί να επηρεάσει σημαντικά τη διακύμανση του S^2 και η επίδραση αυτή δε διορθώνεται με μια απλή αύξηση του μεγέθους του δείγματος.

Τα μη παραμετρικά διαγράμματα για τη διακύμανση χρειάζονται σε αυτές τις περιπτώσεις, αλλά δεν υπάρχουν αρκετές δημοσιεύσεις πάνω στο θέμα αυτό. Είναι δυνατόν να υιοθετηθούν μη παραμετρικά τεστ για την ισότητα δυο διακυμάνσεων (Lehmann, 1975) για τη χρήση τους στα μη παραμετρικά διαγράμματα ελέγχου για τη διακύμανση. Τα διαγράμματα ελέγχου που χρησιμοποιούν στατιστικά τεστ για τη σύγκριση δυο διακυμάνσεων, απαιτούν την απόκτηση ενός αρχικού δείγματος μεγέθους m όταν η διεργασία είναι εντός ελέγχου. Σε κάθε χρονικό σημείο i αποκτάται ένα δείγμα μεγέθους n και τελικά δημιουργείται ένα ενοποιημένο δείγμα μεγέθους $m + n$. Οι παρατηρήσεις στο ενοποιημένο δείγμα ταξινομούνται σε αύξουσα σειρά, από τη μικρότερη στη μεγαλύτερη, και βάση αυτών υπολογίζονται κάποια στατιστικά μεγέθη. Τέτοιες διαδικασίες όμως είναι αρκετά πολύπλοκες και φαίνεται απίθανο να δημιουργηθεί κατ' αυτόν τον τρόπο μια διαγραμματική διαδικασία για την παρακολούθηση της διακύμανσης που θα είναι απλή στη χρήση της.

Μια άλλη προσέγγιση στο πρόβλημα αυτό είναι να προσαρμόσουμε το Westenberg's Two-Sample Interquartile Range Test Bradley (1968). Το τεστ αυτό βασίζεται στην ενοποίηση δυο δειγμάτων, S_1 και S_2 , σε ένα δείγμα και μετά, μετριέται ο αριθμός των παρατηρήσεων που ανήκουν στο δείγμα S_1 και βρίσκονται πάνω από το 3^ο τεταρτημόριο Q_3 , ή κάτω από το 1^ο τεταρτημόριο Q_1 , όπου τα Q_1 και Q_3 βασίζονται στο ενοποιημένο δείγμα. Στις εφαρμογές των διαγραμμάτων ελέγχου, τα Q_1 και Q_3 πρέπει να ορίζονται από τους μηχανικούς της διεργασίας ή να υπολογίζονται με τα δεδομένα της διεργασίας, όταν η διεργασία βρίσκεται εντός ελέγχου.

Ας υποθέσουμε το $U_{ij} = 1$ αν $X_{ij} < Q_1$ ή $X_{ij} > Q_3$, το $U_{ij} = 0$ αν $X_{ij} = Q_1$ ή $X_{ij} = Q_3$ και $U_{ij} = -1$ αν $Q_1 < X_{ij} < Q_3$ όπου X_{ij} είναι η j -οστή παρατήρηση του i -οστού δείγματος. Η προσημική στατιστική συνάρτηση του i -οστού δείγματος είναι $U_i = \sum_{j=1}^n U_{ij}$ και η τυχαία μεταβλητή $V_i = (U_i + n)/2$ ακολουθεί τη δωδυμική κατανομή με παραμέτρους n και p , με $p = P(X_{ij} > Q_3 \text{ ή } X_{ij} < Q_1)$.

Μια ειδοποίηση θα πραγματοποιηθεί όταν $V_i \geq c$ όπου V_i είναι ο αριθμός των παρατηρούμενων X_{ij} που υπερβαίνουν το Q_3 και των παρατηρούμενων X_{ij} που βρίσκονται κάτω από το Q_1 στο δείγμα i . Αυτό είναι ένα σχετικά εύκολο μη παραμετρικό διάγραμμα ελέγχου για το οποίο χρειάζεται να ορίσουμε τα Q_1 και Q_3 , τις εντός ελέγχου τιμές των ανώτατων και κατώτατων τεταρτημόριων.

2.2.4 Σχέδια συντόμευσης δειγματοληψίας

Σε ορισμένες περιπτώσεις είναι δυνατό να ληφθούν παρατηρήσεις διαδοχικά αφότου έχει ήδη αποκτηθεί ένα δείγμα από τη διεργασία. Για παράδειγμα, στη βιομηχανία παραγωγής νάιλον, είναι πολύ σύνηθες να λαμβάνεται από την παραγωγική γραμμή ένα δείγμα ινών νάιλον μεγέθους n , το οποίο αναλύεται στα εργαστήρια. Τα προσημικά διαγράμματα για την παρακολούθηση του μέσου και της διακύμανσης μιας διεργασίας, προσφέρονται πολύ καλά για τα σχέδια συντόμευσης δειγματοληψίας, όπως αυτά που περιέγραψαν οι Williams, Looney, και Pekrs (1985), όπου τα σχέδια συντόμευσης δειγματοληψίας χρησιμοποιήθηκαν στον οικονομικό σχεδιασμό των np – διαγραμμάτων. Στην έρευνά τους έλαβαν υπόψη τους μόνο δειγματοληπτικά σχέδια στα οποία είναι δυνατό να σταματήσει η λήψη των παρατηρήσεων και να δοθεί σήμα ότι η διεργασία είναι εκτός ελέγχου, προτού να έχουν αποκτηθεί όλες οι n παρατηρήσεις. Είναι επίσης δυνατό να σταματήσει νωρίτερα και να αποφασιστεί ότι κανένα σήμα δεν έλαβε μέρος.

Για να περιγράψουμε τη συντομευμένη δειγματοληψία, θεωρούμε ένα μονόπλευρο προσημικό διάγραμμα Shewhart για την παρακολούθηση του μέσου της διεργασίας, όπου ένα σήμα πραγματοποιείται όταν το $T_i > c$, όπου T_i ο αριθμός των θετικών πρόσημων στο i – οστό δείγμα μεγέθους n . Αν σε οποιαδήποτε στιγμή της δειγματοληψίας αποκτηθούν $(n - c + 1)$ παρατηρήσεις κάτω του μ_0 (π.χ. αν αποκτηθούν $n - c + 1$ αρνητικά πρόσημα), δεν υπάρχει λόγος να συνεχίσουμε να λαμβάνουμε περαιτέρω παρατηρήσεις από το δείγμα για να περιμένουμε ένα σήμα από τη διεργασία μας, καθώς είναι αδύνατο να αποκτήσουμε τιμή $T_i > c$. Η ίδια ιδέα εφαρμόζεται αν θέλουμε να σταματήσουμε νωρίτερα την ειδοποίηση της διεργασίας. Αν σε κάποια στιγμή της δειγματοληψίας, παρατηρηθούν c θετικά πρόσημα, τότε μπορεί να δοθεί ένα σήμα από τη διεργασία, χωρίς να χρειαστούν επιπλέον παρατηρήσεις. Ο μέσος αριθμός παρατηρήσεων για μια ειδοποίηση (Average number of observations ANOS) για ένα μονόπλευρο προσημικό διάγραμμα σε μια συντομευμένη δειγματοληψία είναι $(ARL) \cdot E(\eta)$ όπου η είναι το απαραίτητο μέγεθος του δείγματος για να παρθεί μια απόφαση. Η αναμενομένη τιμή του η σε ένα μονόπλευρο έλεγχο δίνεται από την εξίσωση

$$E^+(\eta) = \sum_{k=c}^n k \binom{k-1}{k-c} p^c (1-p)^{k-c} + \sum_{k=n-c+1}^n k \binom{k-1}{k-(n-c+1)} p^{k-(n-c+1)} (1-p)^{n-c+1}$$

όπου p είναι η πιθανότητα απόκτησης μιας παρατήρησης άνω μέσου που αντιστοιχεί στην εντός ελέγχου κατάσταση. Η παραπάνω εξίσωση για $p = 0.5$ μπορεί να απλοποιηθεί στην εντός-ελέγχου περίπτωση

$$E^+(\eta) = \sum_{k=c}^n k \binom{k-1}{k-c} 0,5^k + \sum_{k=n-c+1}^n k \binom{k-1}{k-(n-c+1)} 0,5^k$$

Το αναμενόμενο μέγεθος του δείγματος σε ένα αμφίπλευρο έλεγχο είναι

$$E(\eta) = \sum_{k=c}^n k \binom{k-1}{k-c} \{p^c(1-p)^{k-c} + (1-p)^c p^{k-c}\} + \sum_{k=2(n-c+1)}^n k \binom{k-1}{k-(n-c+1)} \{p^{k-(n-c+1)}(1-p)^{n-c+1} + (1-p)^{k-(n-c+1)} p^{n-c+1}\}$$

Παρόμοια, η παραπάνω εξίσωση μπορεί να απλοποιηθεί για την εντός ελέγχου περίπτωση για $p = 0.5$ ως εξής

$$E(\eta) = \sum_{k=c}^n k \binom{k-1}{k-c} 0,5^{k-1} + \sum_{k=2(n-c+1)}^n k \binom{k-1}{k-(n-c+1)} 0,5^{k-1}$$

Το μέγεθος της δειγματοληψίας μπορεί να μειωθεί σημαντικά με τη χρήση της συντομευμένης δειγματοληψίας. Για παράδειγμα, σε ένα μονόπλευρο έλεγχο με $n = 10$ και $c = 10$, το ARL έχει μια εντός ελέγχου τιμή $(0.5)^{-10} = 1024$ και ο μέσος αριθμός παρατηρήσεων ($ANOS$) χωρίς τη συντομευμένη δειγματοληψία είναι $(ARL)(n) = 10,240$. Με τη συντομευμένη δειγματοληψία, $E(\eta|\mu_0) = 1,99805$ και το $ANOS$ είναι $(ARL)(E(\eta)) = 2046$, που είναι 20% λιγότερο από τον αριθμό των παρατηρήσεων που απαιτούνται για την εντός ελέγχου περίπτωση χωρίς τη συντομευμένη δειγματοληψία. Τα παραπάνω εφαρμόζονται επίσης στα προσημικά διαγράμματα Shewhart για την παρακολούθηση της διακύμανσης μιας διεργασίας.

Τέλος, η εντός ελέγχου τιμή στο μονόπλευρο έλεγχο είναι $E^+(\eta) < 2$, όπου ένα σήμα δίνεται όταν όλες οι παρατηρήσεις είναι μεγαλύτερες του εντός ελέγχου μέσου, π.χ. όταν το $c = n$. Παρομοίως, μπορεί να δειχθεί ότι όταν η διεργασία είναι εντός ελέγχου, $E(\eta) < 3$ για τον αμφίπλευρο έλεγχο με $c = n$. Όταν $c < n$ το εντός ελέγχου μέγεθος του δείγματος είναι μεγαλύτερο του 2 και του 3 για την μονόπλευρη και αμφίπλευρη περίπτωση αντίστοιχα. Προτείνεται να λαμβάνονται υπόψη και το ARL και το $ANOS$ όταν σχεδιάζονται διαγράμματα ελέγχου με συντομευμένη δειγματοληψία.

2.3 Προσημικό διάγραμμα CUSUM για τη παρακολούθηση του μέσου

Το παραμετρικό διάγραμμα CUSUM (Page, 1954) για την ανίχνευση μεταβολών στον μέσο μιας διεργασίας βασίζεται στο άθροισμα των μετατοπίσεων από την τιμή στόχο. Μη παραμετρικά διαγράμματα CUSUM μπορούν να δημιουργηθούν με τη χρήση μη παραμετρικών στατιστικών ελέγχων, όπως είναι ο προσημικός έλεγχος SN_i ή η προσημική βαθμολογική στατιστική συνάρτηση Wilcoxon (Wilcoxon signed-rank statistic) αντί του \bar{X}_i . Τα μη παραμετρικά διαγράμματα ελέγχου που χρησιμοποιούν προσημικές βαθμολογικές συναρτήσεις (signed-rank statistics) είναι ιδιαίτερα δημοφιλή. Στην εργασία του Reynolds (1975), μελετώνται διαγράμματα που βασίζονται σε προσημικούς διαδοχικούς βαθμούς (signed sequential ranks) των παρατηρήσεων. Ο Miller (1970) εισήγαγε έναν έλεγχο που χρησιμοποιεί την προσημική βαθμολογική στατιστική συνάρτηση του Wilcoxon, ενώ οι McGilchrist και Woodyer (1975) πρότειναν ένα διάγραμμα CUSUM για τη μελέτη της διαμέσου της κατανομής βροχοπτώσεων που βασίζεται σε προσημικούς βαθμούς. Οι Bakir και Reynolds (1979) πρότειναν ένα διάγραμμα CUSUM για την παρακολούθηση του μέσου μιας διεργασίας, που βασίζεται στην ιδέα της βαθμολόγησης (ranking) των παρατηρήσεων εντός του δείγματος, στο οποίο ανήκουν. Η συγκεκριμένη ιδέα είχε αρχικά εφαρμοσθεί από τους Wilcoxon (1963) και Van der Laan (1966) για την κατασκευή ενός ελέγχου για δύο διαδοχικά δείγματα.

Οι Bakir και Reynolds (1979) ερεύνησαν τα διαγράμματα CUSUM χρησιμοποιώντας τη προσημική βαθμολογική στατιστική συνάρτηση Wilcoxon. Χρησιμοποιώντας το συσσωρευτικό άθροισμα σαν μια αλυσίδα Markov, ανέπτυξαν μια απλή μέθοδο για την εκτίμηση του ARL μιας κατανομής, με την προϋπόθεση ότι η κατανομή της προσημικής βαθμολογικής στατιστικής συνάρτησης Wilcoxon είναι γνωστή. Οι τιμές των ARL στην εργασία τους είναι ο αναμενόμενος αριθμός των παρατηρήσεων και όχι των δειγμάτων, που χρειάζονται για να δοθεί ένα σήμα.

Αν R_{ij} είναι ο βαθμός της απόλυτης τιμής X_{ij} της κάθε παρατήρησης μεταξύ όλων των απόλυτων τιμών των παρατηρήσεων του i - οστού δείγματος, για $j = 1, 2, \dots, n$ και $i = 1, 2, \dots, n$ τότε η Wilcoxon προσημική βαθμολογική στατιστική

συνάρτηση που απεικονίζεται στο διάγραμμα των Bakir και Reynolds (1979) δίνεται στον ακόλουθο τύπο

$$SR_i = \sum_{j=1}^n \text{sign}(X_{ij})R_{ij}$$

Η διαδικασία που προτείνεται, χρησιμοποιεί το συσσωρευμένο άθροισμα των SR_1, SR_2, \dots τα οποία είναι μια ακολουθία ανεξάρτητων προσημικών βαθμολογικών συναρτήσεων τύπου Wilcoxon και εκφράζουν τις διαφορές μεταξύ των αθροισμάτων των βαθμών των X_{ij} που αντιστοιχούν σε θετικές και αρνητικές τιμές αντίστοιχα.

Τα διαγράμματα CUSUM που παρουσιάζονται χρησιμοποιούν ένα σύνολο από προσημικούς ελέγχους.

Ο μονόπλευρος έλεγχος για την ανίχνευση μιας απόκλισης από τον εντός ελέγχου μέσο, δίνει σήμα για το πρώτο t όταν

$$\sum_{i=1}^t (SN_i - k) - \min_{0 \leq u \leq t} \sum_{i=1}^u (SN_i - k) \geq h \quad (1)$$

όπου $h > 0$ και $k > 0$ είναι παράμετροι της διεργασίας. Ο μονόπλευρος έλεγχος για την ανίχνευση μιας αρνητικής απόκλισης δίνει σήμα στο πρώτο t όταν

$$\max_{0 \leq u \leq t} \sum_{i=1}^u (SN_i + k) - \sum_{i=1}^t (SN_i + k) \geq h \quad (2)$$

Οι αντίστοιχοι αμφίπλευροι έλεγχοι δίνουν σήμα στο πρώτο t για το οποίο ένας από τους παραπάνω μονόπλευρους ελέγχους δίνει σήμα.

Ένας εναλλακτικός τρόπος εφαρμογής των διαγραμμάτων CUSUM εμπεριέχει τη χρήση της V-μάσκας. Αν k και n είναι μη αρνητικοί ακέραιοι, τότε ο παραπάνω μονόπλευρος θετικός έλεγχος ισοδυναμεί με μια Μαρκοβιανή αλυσίδα διακριτού χρόνου $\{SN_t^*, t = 0, 1, 2, \dots\}$ με το χώρο καταστάσεων το υποσύνολο $\{0, 1, 2, \dots, h\}$, όπου $SN_0^* = 0$ και

$$SN_t^* = \min \{h, \max \{0, SN_{t-1}^* + \{SN_t - k\}\}\}$$

όπου η κατάσταση h είναι μια απορροφητική κατάσταση, με την έννοια της απορρόφησης του σήματος μιας κατάστασης.

Το ARL για τα διαγράμματα CUSUM που χρησιμοποιούν τον έλεγχο SN_i μπορούν να προσδιοριστούν από το μέσο αριθμό απορρόφησης για την κατάσταση h . Έστω $\underline{m}' = (m_0, m_1, \dots, m_{h-1})$ όπου m_j είναι ο μέσος αριθμός απορρόφησης δεδομένου ότι η αλυσίδα ξεκίνησε αρχικά από την κατάσταση j . Αν το διάγραμμα CUSUM ξεκινά με $S_0^* = 0$ τότε το ARL είναι m_0 .

Αν το Q είναι ένας $h \times h$ πίνακας μετάβασης πιθανοτήτων για μια μη απορροφητική κατάσταση της Μαρκοβιανής αλυσίδας, I είναι ένας $h \times h$ πίνακας ταυτοτήτων και $\underline{1}$ είναι ένα διάνυσμα $h \times 1$ με όλα τα στοιχεία να είναι μονάδα, τότε γνωρίζουμε ότι το \underline{m} δίνεται από τη σχέση $(I - Q)^{-1} \underline{1}$. Οι πιθανότητες μετάβασης για την Μαρκοβιανή αλυσίδα μπορούν εύκολα να υπολογιστούν από τη στιγμή που η κατανομή του SN_i μπορεί να αποκτηθεί μέσω της διωνυμικής κατανομής.

Η τιμή του ARL για το διάγραμμα CUSUM που χρησιμοποιεί το SN_i εξαρτάται από την τιμή των παραμέτρων h και k . Μια προσέγγιση για την επιλογή των h και k είναι να επιλέξουμε τις παραμέτρους με τιμές που ελαχιστοποιούν το $L^+(\mu_1)$, διατηρώντας μια συγκεκριμένη τιμή για το $L^+(\mu_0)$, όπου μ_1 είναι μια τιμή του μ που θεωρείται ως μια σημαντική μεταβολή. Η βέλτιστη τιμή για το k είναι $k = \frac{1}{2} E[SN_i | \mu_1]$ (Reynolds, 1975). Χρησιμοποιώντας αυτή την τιμή για το k , η τιμή για το h πρέπει να επιλεγεί ώστε να επιτευχθεί η επιθυμητή τιμή για το $L^+(\mu_0)$.

Ο πίνακας 2.4 δίνει τις βέλτιστες τιμές του k για διάφορες τιμές των $\mu_1 - \mu_0$ όταν το $n = 10$ για την ομοιόμορφη, κανονική, διπλή εκθετική, Cauchy και γάμμα κατανομή.

Οι τιμές είναι ίδιες με αυτές που χρησιμοποιούνται στα \bar{X} –διαγράμματα ούτως ώστε, οι τιμές των $\mu_1 - \mu_0$ να μπορούν να θεωρηθούν σε μονάδες της τυπικής απόκλισης. Εκτός της Cauchy κατανομής, οι τιμές του k δε διαφέρουν πολύ στις υπόλοιπες κατανομές. Ο προσδιορισμός του μ_1 δεν είναι τόσο ακριβής στην πραγματικότητα, έτσι, χρησιμοποιώντας τις τιμές του k για την κανονική κατανομή δε θα οδηγήσει σε μεγάλα σφάλματα. Οι τιμές του k για την κανονική κατανομή, στρογγυλοποιημένες στον πλησιέστερο ακέραιο, δίνονται στην τελευταία στήλη του πίνακα 2.4.

Πίνακας 2.4: Βέλτιστες τιμές του k για τα διαγράμματα CUSUM χρησιμοποιώντας την στατιστική προσημική συνάρτηση για $n = 10$ και $\sigma = 1$

Shift ($\mu_1 - \mu_0$)	Distribution					Rounded Normal Values
	Uniform	Normal	Double Exponential	Cauchy	Gamma	
.25	0.72	0.99	1.49	1.32	1.09	1
.50	1.44	1.91	2.53	1.74	2.20	2
1.00	2.88	3.41	3.78	2.09	4.11	3
2.00	5.00	4.77	4.70	2.29	5.00	5
3.00	5.00	4.98	4.93	2.36	5.00	5

Ο πίνακας 2.5 δίνει τις τιμές του $L^+(\mu_0)$ για διάφορες τιμές των h και k όταν το $n = 10$. Παρατηρούμε ότι όταν το k είναι ζυγός ακέραιος ο χώρος καταστάσεων αποτελείται από ζυγούς ακέραιους και το ARL για μονές τιμές του h θα είναι το ίδιο με το ARL για τον επόμενο ζυγό ακέραιο. Είναι σημαντικό να αναφέρουμε πως στα διαγράμματα Shewhart που χρησιμοποιούν τη στατιστική SN_i είναι απαραίτητο το δείγμα n να είναι επαρκούς μεγέθους ώστε να έχουμε μια μεγάλη τιμή για το $L^+(\mu_0)$. Παρόλα αυτά, για τα διαγράμματα CUSUM, το μέγεθος n του δείγματος μπορεί να είναι μικρότερο, καθώς η διεργασία βασίζεται σε ένα άθροισμα στατιστικών ελέγχων από μεμονωμένα δείγματα και η τιμή του h μπορεί να επιλεγεί αρκετά υψηλή για να δώσει αντίστοιχα υψηλές τιμές στο $L^+(\mu_0)$. Το μειονέκτημα των μικρών δειγμάτων για τα διαγράμματα CUSUM που χρησιμοποιούν τη στατιστική SN_i είναι ότι η διεργασία δεν μπορεί να δώσει κάποιο σήμα μετά από ένα μόνο δείγμα αν το $n < h + k$.

Πίνακας 2.5: Οι τιμές του ARL για τα αμφίπλευρα προσημικά και \bar{X}_i διαγράμματα για διάφορες κατανομές όταν $n = 10$ και $L(\mu_0) = 512$ και $\sigma = 1$

h	k	1	2	3	4	5	6	7
3		5.6	14.4	17.8	78.0	92.4	930.0	1023.0
4		9.1	14.4	44.6	78.0	521.5	930.0	50083.8
5		11.8	36.8	67.0	464.9	875.6	45525.1	
6		16.9	36.8	148.0	464.9	4756.3	45525.1	
7		22.8	91.6	282.5	3166.5	23143.0		
8		30.1	91.6	519.0	3166.5	41364.0		
9		39.9	216.3	1074.2	17931.9			
10		51.1	216.3	1886.4	17931.9			
11		65.0	499.5	3663.9				
12		83.7	499.5	6968.4				
13		105.8	1147.8	13030.2				
14		133.3	1147.8	25236.3				
15		166.8	2623.3	47128.0				
16		208.3	2623.3	89762.1				

Ένα παρόμοιο διάγραμμα CUSUM μπορεί να αναπτυχθεί για τη διακύμανση μιας διεργασίας, αν αντικαταστήσουμε τη στατιστική συνάρτηση SN_i με την V_i στις σχέσεις (1) και (2) της ίδιας παραγράφου. Το πρόβλημα της επιλογής της κατάλληλης τιμής του k χρειάζεται μια ξεχωριστή διαδικασία στην περίπτωση αυτή. Έτσι λοιπόν για την παρακολούθηση του μέσου μιας διεργασίας, ας θεωρήσουμε ότι λαμβάνουμε διαδοχικά και ανεξάρτητα τυχαία δείγματα $X_{i1}, X_{i2}, \dots, X_{in}$, και έστω ότι $|X_{i1} - \mu_0|, |X_{i2} - \mu_0|, \dots, |X_{in} - \mu_0|$ είναι οι απόλυτες τιμές των αποκλίσεων των παρατηρήσεων από τη διάμεσο του i -οστού δείγματος και R_{ij} ο βαθμός της ποσότητας $|X_{ij} - \mu_0|$, για $j = 1, 2, \dots, n$, μεταξύ των n απόλυτων αποκλίσεων (από τη διάμεσο), όπου μ_0 είναι η εντός ελέγχου τιμή της διαμέσου. Ο Bakir (2004) χρησιμοποίησε την ακόλουθη μορφή προσημικής βαθμολογικής συνάρτησης Wilcoxon

$$SR'_i = \sum_{j=1}^n \text{sign}(X_{ij} - \mu_0) R_{ij}, i = 1, 2, \dots$$

για την κατασκευή ενός μη παραμετρικού διαγράμματος ελέγχου τύπου Shewhart, και υπολόγισε το εντός ελέγχου μέσο μήκος ροής (ARL_{in}) και την πιθανότητα λανθασμένου συναγερμού (FAR) με τη βοήθεια της κατανομής της συνάρτησης Wilcoxon W^+ που εκφράζει το συνολικό άθροισμα των βαθμών μόνο για τις θετικές αποκλίσεις ($X_{ij} - \mu_0$). Πιο συγκεκριμένα, αποδεικνύεται ότι ισχύει η σχέση

$$SR' = 2W^+ - \frac{n(n+1)}{2}$$

οπότε προκύπτει ότι η συνάρτησης SR' είναι συμμετρική γύρω από το μηδέν. Συνεπώς η εντός ελέγχου κατανομή της SR' υπολογίζεται από την κατανομή της W^+ υπό τη μηδενική υπόθεση (null distribution).

Οι Chakraborti & Eryilmaz (2007) πρότειναν απλές παραλλαγές του διαγράμματος του Bakir (2004), εφαρμόζοντας σε αυτό κανόνες ροών τύπου 2-από-2. Πιο συγκεκριμένα, κατασκεύασαν τα διαγράμματα 2-από-2 KL και 2-από-2 DR, χρησιμοποιώντας τη στατιστική συνάρτηση SR' . Το πρώτο διάγραμμα

δίνει σήμα ότι η διεργασία είναι εκτός ελέγχου αν οι τιμές της SR' των δύο πιο πρόσφατων δειγμάτων, βρίσκονται πάνω ή κάτω από τα όρια ελέγχου, ενώ στο δεύτερο διάγραμμα το σήμα παράγεται ακόμη και αν η μια τιμή βρίσκεται πάνω και η δεύτερη κάτω από τα όρια ελέγχου.

2.4 Άλλες εκδοχές μη παραμετρικών διαγραμμάτων ελέγχου

Οι Amin & Searcy (1991) πρότειναν ένα μη παραμετρικό διάγραμμα ελέγχου τύπου EWMA που βασίζεται στη βαθμολόγηση των παρατηρήσεων εντός του δείγματος με τη χρήση της προσημικής στατιστικής συνάρτησης (grouped signed-ranked statistic) ή αλλιώς GRS-EWMA. Συγκεκριμένα, η στατιστική συνάρτηση που απεικονίζεται στο διάγραμμα δίνεται στην ακόλουθη σχέση

$$Z_i = \lambda SR_i + (1 - \lambda)Z_{i-1}, 0 < \lambda < 1,$$

όπου η αρχική τιμή Z_0 είναι η τιμή στόχος που έχει καθορισθεί.

Η διεργασία θεωρείται να είναι εκτός ελέγχου κάθε φορά που κάποιο Z_i είτε ξεπερνάει το άνω όριο ελέγχου (UCL) είτε βρίσκεται κάτω από το κάτω όριο ελέγχου (LCL). Τα όρια ελέγχου δίνονται από την σχέση $\mu_0 \pm L$. Επιπλέον όσο αναφορά το EWMA μη παραμετρικό διάγραμμα ελέγχου, αυτό έχει εκτιμηθεί και συγκριθεί με βάση το ARL . Επιλέχθηκαν η ομοιόμορφη, η κανονική, η διπλή εκθετική, η γάμμα, και η Cauchy κατανομή. Τα όρια ελέγχου για το κανονικό \bar{X} – EWMA και για το GRS – EWMA επιλέχθηκαν έτσι ώστε η συχνότητα των σημείων που βρίσκονται έξω από τα όρια ελέγχου να είναι ίδια και για τις διαδικασίες όταν η διεργασία είναι εντός ελέγχου.

Η προσθήκη των προειδοποιητικών ορίων ελέγχου βελτιώνουν την απόδοση του διαγράμματος.

Προτείνεται το διάγραμμα ελέγχου ότι το διάγραμμα ελέγχου για την διακύμανση μπορεί να χρησιμοποιείται μαζί την GRS – EWMA διαδικασία.

Όσο αναφορά το ARL , αυτό επηρεάζεται από τις τιμές της παραμέτρου λ του προτεινόμενου μη παραμετρικού διαγράμματος ελέγχου GRS – EWMA.

Έτσι λοιπόν θα μπορούσαμε να πούμε πως η μέθοδος του GRS – EWMA είναι μια καλή εναλλακτική προσέγγιση μη παραμετρικού διαγράμματος ελέγχου.

Επίσης στην εργασία του McDonald (1990), προτείνεται ένα διάγραμμα ελέγχου CUSUM που βασίζεται στους διαδοχικούς βαθμούς (sequential ranks) των παρατηρήσεων, οι οποίοι ορίστηκαν αρχικά στην εργασία του Parent (1965) μέσω του τύπου

$$R_i = 1 + \sum_{j=1}^{i-1} I(X_j < X_i)$$

όπου $I(\cdot)$ είναι η συνήθης δείκτης συνάρτηση.

Το διάγραμμα στηρίζεται στις συναρτήσεις $U_i = \frac{R_i}{i+1}, i = 1, 2, \dots$, οι οποίες είναι ανεξάρτητες τυχαίες μεταβλητές και ακολουθούν ομοιόμορφη κατανομή (σε εντός

ελέγχου διεργασία) στο $(\frac{1}{i+1}, \frac{2}{i+1}, \dots, \frac{i}{i+1})$. Το σήμα για εκτός ελέγχου κατάσταση δίνεται όταν ικανοποιείται η ανίσωση $T_i \geq h$, όπου $T_i = \max(T_{i-1} + U_i - k, 0)$ και h, k είναι οι παράμετροι σχεδιασμού του διαγράμματος.

Αξίζει να αναφερθεί ότι οι Bhattacharya & Frierson (1981) πρότειναν ένα μη παραμετρικό διάγραμμα ελέγχου τύπου CUSUM, χρησιμοποιώντας σταθμισμένα αθροίσματα των διαδοχικών βαθμών R_i των παρατηρήσεων X_1, X_2, \dots, X_n που λαμβάνονται από την παραγωγική διαδικασία. Πιο συγκεκριμένα, αν εισάγουμε τις ποσότητες Z_i

$$Z_i = \frac{R_i - (i + 1)/2}{i}$$

τότε στο παραπάνω διάγραμμα απεικονίζεται μια κανονικοποιημένη μορφή των ποσοτήτων $S_k = \sum_{i=1}^k Z_i$, $1 \leq k \leq n$ που δίνεται ως ακολούθως

$$S_n(t) = \sqrt{\frac{12}{n}} \{S_{[nt]} + (nt - [nt])Z_{[nt]+1}\}, \quad 0 \leq t \leq 1.$$

Τέλος όσο αναφορά τους διαδοχικούς ελέγχους και υπολογισμούς, υπάρχουν προβλήματα που αναφέρονται ως “change-point” ή γενικότερα “detection” προβλήματα. Οι Bhattacharya & Frierson (1981) εισήγαγαν το εξής πρόβλημα. Ας υποθέσουμε ότι έχουμε μια ακολουθία από ανεξάρτητες τυχαίες μεταβλητές X_1, X_2, \dots, X_N των οποίων η κατανομή αλλάζει από F σε G μετά τις πρώτες $[N\theta]$ παρατηρήσεις, όπου θ είναι μια άγνωστη παράμετρος. Το ζήτημα στο πρόβλημα που τέθηκε είναι να βρεθεί το άγνωστο σημείο αλλαγής “change-point” χωρίς πολλούς λανθασμένους συναγερμούς και χωρίς να υπάρξει καμία υπόθεση παραμετρικού μοντέλου για τις F και G . Προτείνεται και σε αυτό το πρόβλημα ένα μη παραμετρικό διάγραμμα ελέγχου το οποίο βασίζεται στα αθροίσματα βαρών διαδοχικών βαθμών και την ασυμπτωτική συμπεριφορά των συσσωρευμένων αθροισμάτων των διαδοχικών βαθμών, υποθέτοντας ότι οι μικρές αλλαγές στη κατανομή γίνονται μετά από ένα μεγάλο αριθμό παρατηρήσεων. Στη βιβλιογραφία προτείνονται διάφορες εφαρμογές στο προαναφερθέν πρόβλημα, Zacks (1991), Huskova και Sen (1989), Siegmund (1994), Siegmund και Venkataraman (1995).

2.5 Συγκρίσεις διαγραμμάτων ελέγχου

Οι συγκρίσεις των διαγραμμάτων που παρουσιάζονται έγιναν βάση των τιμών ARL για τα μη παραμετρικά και τα παραμετρικά διαγράμματα βασισμένα στις συναρτήσεις \bar{X}_i και S_i^2 για διάφορες κατανομές. Όταν $\mu = \mu_0$, η τιμή ARL για το προσημικό διάγραμμα ελέγχου του μέσου μιας διεργασίας είναι ανεξάρτητη από την υποκείμενη κατανομή, ενώ από την άλλη, το προσημικό διάγραμμα ελέγχου της διακύμανσης απαιτεί τον υπολογισμό των Q_1, Q_3 όταν η διεργασία είναι εντός ελέγχου. Τα όρια στα παραμετρικά διαγράμματα προσαρμόζονται για κάθε κατανομή ώστε κατά τη σύγκρισή τους, η τιμές ARL να συμπίπτουν. Όταν χρησιμοποιήθηκε η προσομοίωση για τον υπολογισμό των ARL , οι τιμές των $L^+(\mu_0)$ ήταν κατά προσέγγιση ίσες. Κρατώντας το $L^+(\mu_0)$ σταθερό και συγκρίνοντας τις τιμές του $L^+(\mu)$ για $\mu > \mu_0$ ή $\sigma > \sigma_0$, μπορούμε να αξιολογήσουμε την αποτελεσματικότητα των διαγραμμάτων στην ανίχνευση κάποιων μεταβολών.

Βασικό πλεονέκτημα των προσημικών διαγραμμάτων είναι η ευκολία υπολογισμού της σωστής εντός ελέγχου τιμής του ARL για κάθε κατανομή, ενώ τα παραμετρικά διαγράμματα ελέγχου εξισώθηκαν με τα αντίστοιχα προσημικά διαγράμματα για λόγους σύγκρισης.

Ο πίνακας 2.6 που παρουσιάζεται παρακάτω δίνει τις τιμές του ARL για τον αμφίπλευρο έλεγχο Shewhart και ο πίνακας 2.7 δίνει τις τιμές του $L^+(\mu)$ για τους ελέγχους Shewhart που χρησιμοποιούν τη στατιστική συνάρτηση SN_i και \bar{X}_i για διάφορες κατανομές με $n = 10$. Οι κατανομές προσαρμόστηκαν στην ίδια κλίμακα, εκτός από την κατανομή Cauchy, και οι διακυμάνσεις ορίστηκαν ίσες με ένα. Οι τιμές του $L^+(\mu)$ έχουν εφαρμογή και σε περιπτώσεις που η διακύμανση δεν είναι ένα, αλλά οι τιμές του a_1 και $\mu - \mu_0$ είναι σε μονάδα μέτρησης αντίστοιχη της τυπικής απόκλισης. Από τους πίνακες 2.6 και 2.7 φαίνεται ότι οι έλεγχοι που χρησιμοποιούν τη στατιστική συνάρτηση \bar{X}_i είναι περισσότερο αποτελεσματικοί από αυτούς που χρησιμοποιούν την SN_i στις περιπτώσεις της κανονικής και ομοιόμορφης με ελαφριά ουρά κατανομής. Παρόλα αυτά, για την ανίχνευση μικρών μεταβολών για κατανομές με βαριά ουρά όπως η διπλή εκθετική κατανομή, Cauchy και γάμμα, οι έλεγχοι που χρησιμοποιούν τη στατιστική συνάρτηση SN_i είναι περισσότερο αποτελεσματικοί. Όταν οι κατανομές είναι ασύμμετρες, τα μη παραμετρικά διαγράμματα είναι πιο αποτελεσματικά. Για την ανίχνευση μεγάλων μεταβολών, η SN_i δεν είναι τόσο αποτελεσματική όσο η \bar{X}_i για την διπλή εκθετική, αλλά είναι πιο αποτελεσματική για την Cauchy.

Πίνακας 2.6: Οι τιμές του ARL για τον αμφίπλευρο προσημικό διάγραμμα και για το \bar{X}_i -διάγραμμα Shewhart όταν $n = 10, L(\mu_0) = 512$ και $\sigma=1$

Shift ($\mu - \mu_0$)	Normal		Double Exponential		Gamma ₁		Gamma ₂		Gamma ₃	
	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i
	$a_2=10$	$a_1=.945$	$a_2=10$	$a_1=.979$	$a_2=10$	$a_1=1.198$ $a_1=.762$	$a_2=10$	$a_1=1.288$ $a_1=.675$	$a_2=10$	$a_1=1.412$ $a_1=0.555$
0	512.0	512.0	512.0	512.0	512.0	512.0	512.0	512.0	512.0	512.0
.25	166.0	94.5	75.4	135.9	167.2	137.7	164.1	157.0	156.1	184.8
.50	40.0	15.4	17.0	24.3	43.4	25.6	43.4	31.1	41.8	40.1
1.00	5.6	1.9	3.7	2.3	7.2	2.6	7.5	3.1	7.8	3.8
2.00	1.3	1.0	1.4	1.0	1.6	1.0	1.7	1.2	1.9	1.0

Gamma₁: παράμετρος σχήματος=8

Gamma₂: παράμετρος σχήματος=4

Gamma₃: παράμετρος σχήματος=2

Πίνακας 2.7: Οι τιμές του *ARL* για τον μονόπλευρο προσημικό διάγραμμα και για το \bar{X}_i –διάγραμμα *Shewhart* για διάφορες κατανομές όταν $n = 10, L^+(\mu_0) = 1024$ και $\sigma=1$

Shift ($\mu - \mu_0$)	Uniform		Normal		Double Exponential		Cauchy*		Gamma2	
	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i
	$a_2=10$	$a_1=.945$	$a_2=10$	$a_1=.979$	$a_2=10$	$a_1=1.05$	$a_2=10$	$a_1=84.9$	$a_2=10$	$a_1=1.12$
0	1024.0	1024.0	1024.0	1024.0	1024.0	1024.0	1024.0	1024.0	1024.0	1024.0
.25	265.9	75.9	169.0	94.9	75.5	137.9	19.4	1021.0	143.1	157.2
.50	81.1	12.4	49.0	15.4	17.0	24.3	5.3	1018.0	26.6	31.1
1.00	10.7	1.8	5.6	1.9	3.7	2.3	2.3	1011.9	2.5	3.0
2.00	1.0	1.0	1.3	1.0	1.4	1.5	1.5	999.9	1.0	1.0

* το σ δεν έχει οριστεί για την κατανομή Cauchy

Ο πίνακας 2.8 που ακολουθεί είναι παρόμοιος με τον πίνακα 2.7 με τη μόνη διαφορά ότι στον πίνακα 2.8 χρησιμοποιούνται όρια προειδοποίησης και το $L^+(\mu_0) = 593.7$ και για τις δυο διεργασίες. Τα όρια προειδοποίησης ορίζονται έτσι ώστε το p_0 και το p_1 να είναι ίδια και στις δυο διεργασίες. Πάλι, η στατιστική συνάρτηση \bar{X}_i είναι πιο αποτελεσματική από την SN_i για την ομοιόμορφη και κανονική κατανομή, και λιγότερο αποτελεσματική για την κατανομή Cauchy. Για την διπλή εκθετική κατανομή, η \bar{X}_i είναι πιο αποτελεσματική για μεγάλες μεταβολές από ότι για μικρές. Για την κατανομή γάμμα δεν υπάρχει ιδιαίτερη διαφορά στην αποτελεσματικότητα των δυο αυτών ελέγχων.

Πίνακας 2.8: Οι τιμές του *ARL* για το προσημικό διάγραμμα και για το \bar{X}_i –διάγραμμα *Shewhart* με προειδοποιητικά όρια για διάφορες κατανομές όταν $n = 10, L(\mu_0) = 593,7$ και $\sigma=1$

Shift ($\mu - \mu_0$)	Uniform		Normal		Double Exponential		Cauchy*		Gamma2	
	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i
	$a_2=10$ $w_2=4$ $r=4$	$a_1=.945$ $w_1=.303$ $r=4$	$a_2=10$ $w_2=4$ $r=4$	$a_1=.979$ $w_1=.299$ $r=4$	$a_2=10$ $w_2=4$ $r=4$	$a_1=1.05$ $w_1=.292$ $r=4$	$a_2=10$ $w_2=4$ $r=4$	$a_1=84.9$ $w_1=.435$ $r=4$	$a_2=10$ $w_2=4$ $r=4$	$a_1=1.12$ $w_1=.296$ $r=4$
0	593.7	593.7	593.7	593.7	593.7	593.7	593.7	593.7	593.7	593.7
.25	116.5	31.2	82.8	33.0	21.7	34.8	6.7	145.2	44.6	41.2
.50	23.5	6.4	7.9	6.9	6.1	7.1	3.2	18.6	8.4	7.9
1.00	4.6	1.7	1.6	1.6	2.7	2.1	2.1	5.9	2.2	2.4
2.00	1.0	1.0	1.3	1.0	1.3	1.0	1.5	4.6	1.0	1.0

* το σ δεν έχει οριστεί για την κατανομή Cauchy

Στον πίνακα 2.9 που ακολουθεί οι τιμές των *ARL* για τα μη παραμετρικά διαγράμματα *Shewhart* για τη διακύμανση, συγκρίνονται με τα αντίστοιχα διαγράμματα S^2 για την κανονική, γάμμα και διπλή εκθετική κατανομή για $n = 7$. Οι *ARL* τιμές των S^2 διαγραμμάτων για μη κανονικές κατανομές υπολογίστηκαν από εξομοιώσεις.

Στον πίνακα 2.9 στο (*) φαίνεται ότι η εντός ελέγχου τιμή του ARL για τον έλεγχο S^2 αλλάζει από 128 για την κανονική κατανομή, σε 39.1 και 4.9 για την γάμμα και διπλή εκθετική κατανομή αντίστοιχα. Η τυπική απόκλιση σ είναι σωστά υπολογισμένη, αλλά τα όρια ελέγχου βασίζονται σε σταθερές που προϋποθέτουν κανονικότητα. Οι τιμές του ARL στις περιπτώσεις όπου το $\sigma > \sigma_0$ δείχνουν ότι ο έλεγχος S^2 είναι πιο αποτελεσματικός από τον αντίστοιχο μη παραμετρικό, αλλά μόνο στην περίπτωση όπου η εντός ελέγχου τιμή του ARL είναι σωστή. Οι αριθμοί στο (**) δίνουν τις τιμές ANOS όταν χρησιμοποιείται συντομευμένη δειγματοληψία. Ο μη παραμετρικός έλεγχος με συντομευμένη δειγματοληψία έχει το πλεονέκτημα ότι χρησιμοποιεί λιγότερες παρατηρήσεις κατά μέσο όρο όταν η διεργασία είναι εντός ελέγχου. Στην περίπτωση της κανονικής κατανομής, μια εντός ελέγχου τιμή του ARL ίση με 128, χρειάζεται κατά μέσο όρο 254 παρατηρήσεις για το προσημικό διάγραμμα αντί για 896 παρατηρήσεις για το διάγραμμα S^2 .

Πίνακας 2.9: Οι τιμές του ARL για τον αμφίπλευρο προσημικό διάγραμμα και για το S^2 –διάγραμμα Shewhart όταν $n = 7, L^+(\mu_0) \approx 128$

σ_1/σ_0	Normal		Gamma		Double Exponential		Normal (n=2)
	$V_{\bar{c}} = \frac{\bar{c}}{7}$	$S^2_{a_3=2.906}$	$V_{\bar{c}} = \frac{\bar{c}}{7}$	$S^2_{a_3=3.86^{(1)}}$	$V_{\bar{c}} = \frac{\bar{c}}{7}$	$S^2_{a_3=9.0^{(1)}}$	$S^2_{a_3=7.08}$
(*)		(128.0)		(39.1)		(4.9)	
(**)	(254.0)	(896.0)	(248.0)	(878.5)	(254.0)	(886.0)	(256.0)
1.0	128.0	128.0	128.0	125.5	128.0	126.6	128.0
1.1	74.9	39.4	72.6	50.0	82.4	60.3	64.1
1.2	48.7	16.8	44.4	24.5	57.0	33.4	37.5
1.3	34.1	8.9	29.0	14.5	41.8	20.6	24.6
1.4	25.4	5.7	20.1	9.6	32.0	13.5	17.4
1.5	19.8	3.9	14.7	6.8	25.4	9.5	13.1
1.6	15.9	3.0	11.1	5.3	20.8	7.0	10.4
1.7	13.2	2.4	9.1	4.4	17.4	5.6	8.5
1.8	11.2	2.0	7.7	3.7	14.8	4.5	7.2
1.9	9.7	1.8	6.7	3.2	12.9	3.8	6.2
2.0	8.6	1.6	6.0	2.8	11.3	3.2	5.5

(1) Η σταθερά a_3 και το ARL υπολογίστηκαν από εξομοιώσεις

(*) Εντός ελέγχου τιμές ARL όταν τα όρια ελέγχου βασίζονται υπόθεση της κανονικότητας

(**) Εντός ελέγχου τιμές ANOS για κάθε κατανομή

Ο μονόπλευρος προσημικός έλεγχος Shewhart με $c = n$ έχει ένα εντός ελέγχου αναμενόμενο μέγεθος δείγματος $E(\eta) < 2$. Αυτή είναι η περίπτωση όπου οι περισσότερες παρατηρήσεις λαμβάνονται όταν η διεργασία βρίσκεται εντός ελέγχου, και είναι λογικό να συγκρίνουμε το μη παραμετρικό διάγραμμα διακύμανσης με ένα αντίστοιχο S^2 διάγραμμα έτσι ώστε και τα δυο να έχουν τα ίδια εντός ελέγχου μέσο αριθμό παρατηρήσεων (ANOS). Ο έλεγχος S^2 με $n = 2$ είναι ελαφρώς πιο αποτελεσματικός από τον μη παραμετρικό, όπως φαίνεται από την τελευταία στήλη του πίνακα 2.9.

Ο πίνακας 2.10 συγκρίνει τις τιμές ARL για το \bar{X} διάγραμμα Shewhart και το προσημικό διάγραμμα με συντομευμένη δειγματοληψία, για την κανονική και γάμμα κατανομή. Και τα δυο διαγράμματα προσαρμόστηκαν ώστε να έχουν τις ίδιες τιμές ARL και ANOS. Αποδεικνύεται ότι το προσημικό διάγραμμα είναι πιο αποτελεσματικό από το \bar{X} διάγραμμα. Ισχύει ότι $E(\eta|\mu) > E(\eta|\mu_0)$ για $\mu \neq \mu_0$.

Το προσημικό διάγραμμα είναι σαφώς καλύτερο από το \bar{X} διάγραμμα όταν η κατανομή δεν είναι κανονική, όπως στην περίπτωση της ελαφρώς ασύμμετρης κατανομής γάμμα με παράμετρο κλίμακας 4.

Πίνακας 2.10: Οι τιμές του ARL για τον αμφίπλευρο προσημικό διάγραμμα με συντομευμένη δειγματοληψία και για το \bar{X} διάγραμμα για $\sigma=1$

Shift ($\mu - \mu_0$)	$SN_i^{(1)} \bar{X}_i$ n=10 n=3 c=10 Normal		$SN_i^{(2)} \bar{X}_i$ n=14 n=6 c=13 Normal		$SN_i^{(1)} \bar{X}_i$ n=10 n=3 c=10 Gamma ⁽³⁾		$SN_i^{(2)} \bar{X}_i$ n=14 n=6 c=13 Gamma ⁽³⁾	
	0.00	512.0	512.0	546.1	546.1	512.0	512.0	546.1
0.25	166.0	246.0	125.7	160.4	164.1	423.9	124.0	273.1
0.50	40.0	77.7	24.2	34.2	43.4	184.2	26.5	77.2
1.00	5.6	11.6	3.1	4.0	7.5	39.0	4.1	9.3
2.00	1.3	1.6	1.0	1.0	1.7	3.4	1.5	1.1

(1): $E(\eta|\mu_0) \leq 3$

(2): $E(\eta|\mu_0) \leq 5.5$

(3): Γάμμα με παράμετρο κλίμακας 4

Οι τιμές ARL για το διάγραμμα CUSUM που χρησιμοποιεί SN_i και \bar{X}_i δίνονται στον πίνακα 2.11 που ακολουθεί. Οι τιμές του k επιλέχθηκαν ώστε να είναι οι βέλτιστες τιμές στην περίπτωση μιας μεταβολής μεγέθους μιας μονάδας. Η τιμή του $L^+(\mu_0)$ για την SN_i είναι 1074.2. Για τον λόγο αυτό δεν είναι δυνατό να προσαρμόσουμε την τιμή του h στο διάγραμμα CUSUM ώστε να έχουμε $L^+(\mu_0) = 1074.2$. Η αποτελεσματικότητα του SN_i σχετικά με το \bar{X}_i στα διαγράμματα CUSUM είναι αντίστοιχη με την αποτελεσματικότητα του SN_i σχετικά με το \bar{X} διάγραμμα Shewhart. Μια ανωμαλία είναι ότι το SN_i είναι πιο αποτελεσματικό από το \bar{X}_i για μικρές μεταβολές σε κανονικά δεδομένα. Βρέθηκε ότι το $L^+(\mu)$ για $\mu > \mu_0$ είναι ευαίσθητο στην επιλογή των παραμέτρων h και k που χρησιμοποιούνται για την επίτευξη της συγκεκριμένης τιμής του $L^+(\mu_0)$. Χρησιμοποιώντας ένα μικρότερο k και ένα μεγαλύτερο h , η απόδοση του \bar{X}_i για μικρές μεταβολές βελτιώνεται αισθητά, παρόλο που η απόδοση για τις μεγάλες μεταβολές μειώνεται. Τα αποτελέσματα δείχνουν ότι τα διαγράμματα CUSUM που χρησιμοποιούν το SN_i είναι περισσότερο αποτελεσματικά από τα διαγράμματα Shewhart που χρησιμοποιούν το SN_i , είτε χρησιμοποιούνται όρια προειδοποίησης είτε όχι.

Πίνακας 2.11: Οι τιμές του ARL CUSUM χρησιμοποιώντας τη στατιστική συνάρτηση SN_i και \bar{X}_i για διάφορες κατανομές όταν $n = 10, L^+(\mu_0) = 1074,2$ και $\sigma = 1$

Shift ($\mu - \mu_0$)	Uniform		Normal		Double Exponential		Cauchy		Gamma	
	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i	SN_i	\bar{X}_i
	h=9 k=3	h=.415 k=.55	h=9 k=3	h=.451 k=.55	h=9 k=3	h=.530 k=.55	h=9 k=3	h=15.62 k=.55	h=9 k=3	h=.61 k=.55
0	1074.2	1091.8	1074.2	1074.31	1074.2	1053.0	1074.2	1074.2	1076.5	1048.6
.25	94.8	57.3	46.2	68.76	16.0	103.0	5.2	581.1	36.2	104.5
.50	17.3	7.8	8.6	8.95	4.8	11.5	2.8	211.1	6.3	13.9
1.00	3.8	1.6	2.9	1.67	2.5	1.8	2.2	33.9	2.3	2.1
2.00	2.0	1.0	2.0	1.00	2.0	1.0	2.0	11.4	2.0	1.0

2.6 Παρατηρήσεις και συμπεράσματα

Από τα παραπάνω αποτελέσματα που παρουσιάστηκαν αποδεικνύεται ότι, αν η κατανομή των δεδομένων πλησιάζει την κανονική και το μέγεθος του δείγματος δεν είναι μικρό, τότε η κατανομή του \bar{X} είναι σχεδόν κανονική και οι τιμές του ARL που υπολογίζονται είναι σωστές. Επιπλέον, το \bar{X} είναι αρκετά αποτελεσματικό στην ανίχνευση των μεταβολών του μ . Οι συνέπειες της μη κανονικότητας στις τιμές του ARL στα \bar{X} διαγράμματα είναι σημαντικές. Στα διαγράμματα ελέγχου, οι συνέπειες αυτές είναι πιο σοβαρές από ότι στα διαγράμματα θέσης.

Στις περιπτώσεις όπου η κατανομή των δεδομένων έχει βαριά ουρά, οι μη παραμετρικές διεργασίες βασισμένες στα πρόσημα, έχουν το πλεονέκτημα του σταθερού ARL όταν οι διεργασίες είναι εντός ελέγχου και είναι πολύ αποτελεσματικές στην ανίχνευση μεταβολών του μ .

Είναι λοιπόν λογικό να χρησιμοποιούνται διαγράμματα ελέγχου βάση προσημών όταν η κατανομή έχει βαριά ουρά. Παρατηρούμε ότι τα προσημικά τεστ μπορούν να εφαρμοστούν για τον έλεγχο του μέσου αντί του διαμέσου, αλλά η εντός ελέγχου τιμή ARL δεν θα είναι με ακρίβεια υπολογισμένη όπως και στην εντός ελέγχου τιμή του ARL για τα R διαγράμματα η οποία δεν συμβαδίζει με την θεωρητική τιμή.

Τα μη παραμετρικά διαγράμματα είναι πιο χρηστικά για την ανίχνευση μικρών μεταβολών του μ όταν το μέγεθος του δείγματος είναι σχετικά μεγάλο $n > 7$. Ο Duncan (1974) υπογραμμίζει ότι για την ανίχνευση μικρών μεταβολών προτιμώνται δείγματα μεγέθους μεγαλύτερου των τεσσάρων ή πέντε. Τα διαγράμματα ελέγχου που βασίζονται στην προσημική στατιστική απαιτούν δείγματα μετρίου μεγέθους ($n > 7$) προκειμένου να αποκτηθεί μια ικανοποιητική εντός ελέγχου τιμή για το ARL . Αυτή είναι μόνο η περίπτωση όπου δε χρησιμοποιείται συντομευμένη δειγματοληψία. Σε πολλές εφαρμογές είναι δυνατό να χρησιμοποιήσουμε συμτομευμένη δειγματοληψία η οποία θα συμβάλλει στη μείωση του αναμενόμενου αριθμού των παρατηρήσεων που απαιτούνται.

Ένα επιπλέον πλεονέκτημα των μη παραμετρικών διαγραμμάτων ελέγχου για την παρακολούθηση του μέσου μιας διεργασίας είναι ότι η διακύμανση της διεργασίας δεν χρειάζεται να είναι γνωστή ώστε να δημιουργηθούν τα διαγράμματα ελέγχου. Τα προβλήματα που ίσως να παρουσιάζονται σε περίπτωση λανθασμένης εκτίμησης της διακύμανσης μπορούν πολύ εύκολα να λυθούν. Όσο το μ παραμένει ίσο με μ_0 , τα μη παραμετρικά διαγράμματα δεν επηρεάζονται από μεταβολές της διακύμανσης της διεργασίας. Αν είναι απαραίτητο να ελέγξουμε τη διακύμανση και τον μέσο μιας κατανομής, τότε τα μη παραμετρικά διαγράμματα, όπως και τα \bar{X} διαγράμματα, πρέπει να χρησιμοποιηθούν σε συνδυασμό με ένα διάγραμμα για τη διακύμανση. Τα διαγράμματα που χρησιμοποιούν το εύρος του δείγματος ή την τυπική απόκλιση έχουν παραδοσιακά χρησιμοποιηθεί μαζί με τα \bar{X} διαγράμματα, αλλά τα διαγράμματα αυτά είναι πολύ ευαίσθητα στην υπόθεση της κανονικότητας. Το μη παραμετρικό διάγραμμα για τη διακύμανση είναι πολύ χρήσιμο για την παρακολούθηση της διακύμανσης της διεργασίας όταν η κατανομή των παρατηρήσεων δεν είναι κανονική ή/και ασύμμετρη. Τα αποτελέσματα της μη κανονικότητας στις τιμές των ARL για τα διαγράμματα ελέγχου για τη διακύμανση είναι πιο σοβαρά από ότι για τα διαγράμματα θέσης. Μέχρι στιγμής δεν έχουν δημοσιευθεί ιδιότητες των διαγραμμάτων ελέγχου για τη διακύμανση και δεν υπάρχουν προτάσεις για την προσαρμογή των σταθερών για τα όρια ελέγχου των παραμετρικών διαγραμμάτων όταν η κατανομή δεν είναι κανονική.

Όσον αφορά την ευκολία χρήσης, τα μη παραμετρικά διαγράμματα πρέπει να είναι το ίδιο εύκολα στην εφαρμογή όσο τα αντίστοιχα παραμετρικά. Στην πραγματικότητα, τα προσημικά τεστ είναι τα πιο εύκολα στη χρήση. Συμπερασματικά, τα μη παραμετρικά διαγράμματα είναι μια ελκυστική εναλλακτική στις περιπτώσεις που δε μπορούμε να υποθέσουμε κανονικότητα.

Κεφάλαιο 3

3.1 Σημαντικά μη παραμετρικά διαγράμματα ελέγχου με χρήση μιας εμπειρικής κατανομής αναφοράς

Μια αξιόλογη ιδέα στην κατασκευή και το σχεδιασμό ενός διαγράμματος ελέγχου είναι να υποθέσουμε ότι είναι διαθέσιμο ένα αρχικό δείγμα αναφοράς (reference sample) ή μια κατανομή αναφοράς (reference distribution), τα οποία αφορούν τη διεργασία που μας ενδιαφέρει όταν αυτή βρίσκεται εντός ελέγχου. Οι Park & Reynolds (1987) πρότειναν διαγράμματα ελέγχου τύπου Shewhart και CUSUM για την παρακολούθηση της παραμέτρου θέσης μιας παραγωγικής διαδικασίας, για περιπτώσεις που η εντός ελέγχου τιμή της παραμέτρου δεν είναι γνωστή. Τα συγκεκριμένα διαγράμματα στηρίζονται στις συναρτήσεις γραμμικής τοποθέτησης (linear placement statistics) που έχουν εισαχθεί στην εργασία των Orban & Wolfe (1982) για τη σύγκριση δειγμάτων που λαμβάνονται από τη διεργασία με ένα σταθερό δείγμα που έχει συλλεχθεί όταν η διεργασία βρισκόταν εντός ελέγχου. Οι Hackl & Ledolter (1991) χρησιμοποίησαν την έννοια των κανονικοποιημένων βαθμών (standardized ranks) R_i των παρατηρήσεων X_i που ορίζονται στην ακόλουθη σχέση

$$R_i = 2(F_0(X_i) - \frac{1}{2})$$

όπου F_0 είναι η αθροιστική συνάρτηση κατανομής για την εντός ελέγχου διεργασία. Ωστόσο, αν η F_0 δεν είναι γνωστή, τότε ο κανονικοποιημένος βαθμός παίρνει τη μορφή

$$R'_i = \frac{2(R_i^* - (n + 1)/2)}{n}$$

όπου ένα δείγμα αναφοράς μεγέθους $(n - 1)$ είναι διαθέσιμο όταν η διεργασία είναι εντός ελέγχου. Επιπλέον, συμβολίζουμε με R_i^* το βαθμό της i -οστής παρατήρησης ως προς το δείγμα αναφοράς, έτσι ώστε

$$R_i^* = 1 + \sum_{j=1}^{n-1} I(X_i > X_j)$$

Η στατιστική συνάρτηση που απεικονίζεται στο προτεινόμενο διάγραμμα τύπου EWMA δίνεται από τον τύπο

$$T_i = (1 - \lambda)T_{i-1} + \lambda R_i$$

όπου T_0 είναι συνήθως το μηδέν και $\lambda \in (0,1]$.

Ο Yaschin (1992) μελέτησε την κατανομή του μήκους ροής ενός μη παραμετρικού διαγράμματος ελέγχου CUSUM, σε περιπτώσεις που η κατανομή της διεργασίας είναι άγνωστη. Συγκεκριμένα, για τη σημειακή εκτίμηση διαφόρων χαρακτηριστικών του μήκους ροής, πρότεινε την αντικατάσταση της πραγματικής

κατανομής των δεδομένων, με μια εμπειρική κατανομή ενός δείγματος αναφοράς. Πρόσθετα, ο Bakir (2006) κατασκεύασε διαγράμματα ελέγχου για την παρακολούθηση του κέντρου μιας διεργασίας, βασιζόμενος σε προσημικές βαθμολογικές συναρτήσεις, οι οποίες αντί να βαθμολογούν τις παρατηρήσεις του δείγματος που εξέρχεται από τη διεργασία, βαθμολογούν συναρτήσεις αυτών των παρατηρήσεων. Οι συγκεκριμένες προσημικές βαθμολογικές συναρτήσεις ονομάζονται signed rank-like statistics και οφείλουν την ονομασία τους στον Moses (1963). Αν δ είναι η διάμεσος του δείγματος αναφοράς που λαμβάνεται από την διεργασία όταν βρίσκεται εντός ελέγχου και R_{ij} είναι ο βαθμός της απόλυτης τιμής της απόκλισης $|X_{ij} - \delta|$ ανάμεσα στις αποκλίσεις $|X_{i1} - \delta|, |X_{i2} - \delta|, \dots, |X_{in} - \delta|$ τότε το προτεινόμενο διάγραμμα του Bakir (2006) απεικονίζει την ακόλουθη στατιστική συνάρτηση

$$\varphi_i^* = \sum_{j=1}^n \text{sign}(X_{ij} - \delta) R_{ij}^*$$

Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα μη παραμετρικά διαγράμματα ελέγχου που βασίζονται σε ένα δείγμα αναφοράς και χρησιμοποιούν διατεταγμένες παρατηρήσεις του δείγματος αυτού στον προσδιορισμό των ορίων ελέγχου αλλά και στον κανόνα απόφασης για το αν η διεργασία κρίνεται ότι βρίσκεται εντός ή εκτός στατιστικού ελέγχου. Οι Willemain & Runger (1996) πρότειναν το σχεδιασμό διαγραμμάτων ελέγχου με τη βοήθεια μιας εμπειρικής κατανομής αναφοράς (empirical reference distribution), η οποία αναφέρεται στα δεδομένα της διεργασίας όταν αυτή λειτουργεί κάτω από κανονικές συνθήκες, δηλαδή με την παρουσία μόνο της φυσικής μεταβλητότητας. Ας θεωρήσουμε ότι X_1, X_2, \dots, X_m είναι ένα δείγμα αναφοράς μεγέθους m που λαμβάνεται από την εντός ελέγχου διεργασία και έστω F η αθροιστική συνάρτηση κατανομής του. Τα όρια ελέγχου του προτεινόμενου διαγράμματος είναι δύο διατεταγμένες παρατηρήσεις $X_{(k)}, X_{(b+k)}$ του παραπάνω δείγματος, όπου $0 \leq k \leq m$ και $1 \leq b \leq m - k + 1$ (σύμβαση: $X_{(0)} = -\infty$ και $X_{(m+1)} = \infty$). Για την παρακολούθηση της παραγωγικής διαδικασίας, λαμβάνουμε μεμονωμένες παρατηρήσεις Y από αυτήν και τις συγκρίνουμε με τα όρια ελέγχου. Η πιθανότητα να βρεθεί μια παρατήρηση Y μεταξύ των ορίων ελέγχου δίνεται από την ακόλουθη σχέση

$$p = P(X_{(k)} \leq Y \leq X_{(b+k)}) = F(X_{(b+k)}) - F(X_{(k)}).$$

Δεδομένου ότι οι διατεταγμένες παρατηρήσεις είναι τυχαίες μεταβλητές, είναι προφανές ότι και η πιθανότητα p θα είναι μια τυχαία μεταβλητή και συγκεκριμένα η p θα ακολουθεί Βήτα κατανομή. Για το προτεινόμενο διάγραμμα των Willemain και Runger (1996), δεδομένων των ορίων ελέγχου που έχουν επιλεγεί, το μήκος ροής N ακολουθεί γεωμετρική κατανομή με παράμετρο $(1 - p)$. Συνεπώς, το αναμενόμενο μέσο μήκος ροής όταν η διεργασία είναι εντός ελέγχου, δεδομένου ότι γνωρίζουμε την πιθανότητα p , δίνεται από την ακόλουθη σχέση

$$ARL_{in} = \frac{1}{1 - p}$$

Καθώς επίσης όντας η ποσότητα ARL_{in} τυχαία μεταβλητή η συνάρτηση πυκνότητας δίνεται από την σχέση

$$h(ARL_{in}) = \frac{m!}{(b-1)!(m-b)!} \cdot \frac{(ARL_{in}-1)^{b-1}}{(ARL_{in})^{m+1}}, ARL_{in} \geq 1$$

με μέση τιμή

$$E(ARL_{in}) = \frac{m}{m-b}$$

Χρησιμοποιώντας την τελευταία σχέση μπορούμε να πούμε ότι, αν χρησιμοποιούμε δείγμα αναφοράς μεγέθους m , τότε η μέγιστη τιμή του αναμενόμενου εντός ελέγχου μέσου μήκους ροής με συμμετρικά όρια ελέγχου είναι ίση με $m/4$. Για το λόγο αυτό, προκειμένου να πετύχουμε στην πράξη την επιθυμητή τιμή για το ARL_{in} σε ένα συμμετρικό διάγραμμα ελέγχου, χρειαζόμαστε τουλάχιστον $4E(ARL_{in})$ παρατηρήσεις για την εύρεση του κατάλληλου και ελάχιστου δυνατού μεγέθους δείγματος για την ανάπτυξη ασφαλών στατιστικών συμπερασμάτων. Με τη βοήθεια της συνάρτησης πυκνότητας παρουσιάζονται στον πίνακα 3.1 διάφοροι συνδυασμοί των παραμέτρων m, k, b του προτεινόμενου διαγράμματος που παράγουν $ARL_{in} \cong 370.4$

Είναι σημαντικό να αναφερθεί ότι οι Ion & Klaassen (2005) πρότειναν ένα μη παραμετρικό διάγραμμα ελέγχου τύπου Shewhart, το οποίο αποτελεί ειδική περίπτωση του διαγράμματος των Willemain & Runger (1996) με την έννοια ότι κατασκευάζεται με τυχαία επιλογή μεταξύ των σχεδιασμών

$$(k = 0, b = n - n_a) \text{ και } (k = 0, b = n - na + 1),$$

όπου $n_a = [(n+1)a]$ είναι το ακέραιο μέρος του αριθμού $(n+1)a$. Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα μη παραμετρικά διαγράμματα ελέγχου που βασίζονται στις ελεύθερες κατανομής ποσότητες $P_{(r)}$ που ονομάζονται στατιστικές συναρτήσεις προτεραιότητας (precedence statistics). Τα συγκεκριμένα διαγράμματα στηρίζονται σε ένα εντός ελέγχου δείγμα αναφοράς (reference sample) για την εκτίμηση των ορίων ελέγχου, ενώ από την παραγωγική διαδικασία εξετάζονται διαδοχικά και ανεξάρτητα τυχαία δείγματα (test samples), με τη βοήθεια των οποίων λαμβάνεται η απόφαση για το αν η διεργασία παραμένει εντός ελέγχου ή όχι. Η λήψη των τυχαίων δειγμάτων μεγέθους n αποτελεί μια βασική διαφορά των διαγραμμάτων αυτών σε σχέση με το διάγραμμα των Willemain & Runger (1996), στο οποίο ο κανόνας απόφασης στηρίζεται σε μεμονωμένες παρατηρήσεις.

Πίνακας 3.1. Υπολογισμός του ARL_{in} για διάφορες τιμές των παραμέτρων σχεδιασμού του διαγράμματος των Willemain & Runger (1996)

m	k	b	ARL_m
741	3	736	370.5
1112	4	1105	370.7
1478	5	1473	370.5
3334	10	3315	370.4
5186	15	5157	370.4
7038	20	6999	370.4
8890	25	8841	370.4
10742	30	10683	370.4
14446	40	13667	370.4
18150	50	18051	370.4
21854	60	21735	370.4
29262	80	29103	370.4
36670	100	36471	370.4

Μία από τις πρώτες προσπάθειες διερεύνησης των ιδιοτήτων και της χρήσης των στατιστικών ελέγχων προτεραιότητας (precedence tests), ως μία γρήγορη και αποτελεσματική διαδικασία ελέγχου χρόνων ζωής, οφείλεται στον Epstein (1955). Συγκεκριμένα, ο Epstein εξέτασε την περίπτωση δύο πληθυσμών που προέρχονται από την κανονική κατανομή, και μελέτησε εμπειρικά έναν έλεγχο που βασίζεται στο πλήθος των παρατηρήσεων του ενός πληθυσμού που ξεπερνούν την τιμή της r -οστής διατεταγμένης παρατήρησης του δεύτερου.

Ο Nelson (1963) χρησιμοποίησε τη στατιστική συνάρτηση προτεραιότητας για την κατασκευή ενός ελεύθερου κατανομής ελέγχου χρόνων ζωής που βασίζεται στη διάταξη πρόωρων αποτυχιών και επιτρέπει τη γρήγορη λήψη αποφάσεων ακόμη και σε περιπτώσεις (δεξιά) περικομμένων παρατηρήσεων.

Ας υποθέσουμε X_1, X_2, \dots, X_m και Y_1, Y_2, \dots, Y_n είναι δύο ανεξάρτητα τυχαία δείγματα μεγέθους m και n που προέρχονται από τις κατανομές F_x και F_y αντίστοιχα. Στόχος μας είναι ο έλεγχος της μηδενικής υπόθεσης ότι οι δύο κατανομές είναι ισόνομες, έναντι της εναλλακτικής ότι μία από τις δύο κατανομές είναι στοχαστικά μεγαλύτερη από την άλλη. Πιο συγκεκριμένα, η διατύπωση των δύο υποθέσεων δίνεται ακολούθως

$$H_0 : F_x = F_y \text{ έναντι } H_1 : F_x < F_y.$$

Εάν $X_{1:m}, X_{2:m}, \dots, X_{m:m}$ και $Y_{1:n}, Y_{2:n}, \dots, Y_{n:n}$ είναι οι διατεταγμένες παρατηρήσεις των δύο παραπάνω τυχαίων δειγμάτων, τότε συμβολίζουμε με M_1 το πλήθος των Y - παρατηρήσεων που είναι μικρότερες από την $X_{1:m}$ και με M_i το πλήθος των Y - παρατηρήσεων που βρίσκονται μεταξύ των $X_{i-1:m}$ και $X_{i:m}$, $i = 2, 3, \dots, m$. Αξίζει να σημειωθεί ότι οι ποσότητες M_i σχετίζονται με τις γνωστές exceedance στατιστικές συναρτήσεις υπέρβασης (exceedance statistics).

Η στατιστική συνάρτηση προτεραιότητας η οποία εκφράζει το πλήθος των Y - παρατηρήσεων που προηγούνται της r - οστής διατεταγμένης X - παρατήρησης, ορίζεται από τον ακόλουθο τύπο

$$P_{(r)} = \sum_{i=1}^r M_i$$

Μεγάλες τιμές της συνάρτησης $P_{(r)}$ δίνουν ένδειξη για την απόρριψη της μηδενικής υπόθεσης H_0 έναντι της εναλλακτικής H_1 . Για καθορισμένο επίπεδο σημαντικότητας α , η κρίσιμη περιοχή θα είναι της μορφής $\{s, s + 1, \dots, n\}$ όπου

$$\alpha = P(P_{(r)} \geq s | H_0: F_x = F_y)$$

Υπό τη μηδενική υπόθεση $H_0: F_x = F_y$ τα δύο διαθέσιμα δείγματα X_1, X_2, \dots, X_m και Y_1, Y_2, \dots, Y_n είναι ισόνομα, οπότε το πλήθος των τρόπων επιλογής του Y - δείγματος μεγέθους n από ένα σύνολο $(m + n)$ παρατηρήσεων, είναι ίσο με $\binom{m+n}{n}$. Παρατηρώντας ότι το συμβάν $P_{(r)} = j$ εκφράζει το πλήθος των τρόπων επιλογής j από τις πρώτες $(j + r - 1)$ διατεταγμένες παρατηρήσεις και στη συνέχεια $(n - j)$ από τις υπόλοιπες $(m + n - j - r)$ για τη δημιουργία του Y - δείγματος, μπορούμε να γράψουμε την ακόλουθη σχέση

$$P(P_{(r)} = j | H_0: F_x = F_y) = \frac{\binom{j+r-1}{j} \binom{m+n-j-r}{n-j}}{\binom{m+n}{m}}$$

Στον παρακάτω πίνακα 3.2 φαίνονται για διαφορετικές τιμές των μεγεθών m, n και $r = 1, 2, 3, 4, 5$ η κρίσιμη τιμή s και το ακριβές επίπεδο σημαντικότητας α που επιτυγχάνεται, όσο το δυνατόν πλησιέστερα στο 0,05.

Πίνακας 3.2. Κρίσιμες τιμές s και ακριβή επίπεδα σημαντικότητας α για στατιστικούς ελέγχους προτεραιότητας

		$r = 1$		$r = 2$		$r = 3$		$r = 4$		$r = 5$	
n	m	s	α	s	α	s	α	s	α	s	α
10	10	4	0.043	6	0.029	7	0.035	8	0.035	9	0.029
10	15	3	0.052	4	0.064	5	0.062	6	0.053	7	0.042
15	15	4	0.049	6	0.040	7	0.054	8	0.064	10	0.033
15	20	3	0.070	5	0.040	6	0.046	7	0.047	8	0.045
20	20	4	0.053	6	0.046	7	0.064	9	0.041	10	0.048
20	25	4	0.033	5	0.052	6	0.064	8	0.034	9	0.036
25	25	4	0.055	6	0.049	8	0.037	9	0.048	10	0.057
25	30	4	0.037	5	0.061	7	0.038	8	0.045	9	0.049
30	30	4	0.056	6	0.051	8	0.040	9	0.052	10	0.063

Επιπλέον η επιλογή της συνάρτησης $P_{(r)}$ ως στατιστική συνάρτηση που απεικονίζεται σε ένα διάγραμμα, οδηγεί στην κατασκευή ενός μη παραμετρικού διαγράμματος ελέγχου προτεραιότητας (precedence control chart). Πράγματι, η εντός ελέγχου κατανομή της $P_{(r)}$ εξαρτάται μόνο από τις παραμέτρους σχεδιασμού m, n, j και όχι από την κατανομή της διεργασίας. Συνεπώς, ένας κανόνας απόφασης που βασίζεται στην $P_{(r)}$, αλλά και το εντός ελέγχου μήκος ροής, θα είναι ελεύθερα κατανομής, με την προϋπόθεση ότι η κατανομή της διεργασίας είναι συνεχής. Με άλλα λόγια, τα διαγράμματα ελέγχου που στηρίζονται σε στατιστικές συναρτήσεις προτεραιότητας, είναι ανεξάρτητα της κατανομής της διεργασίας, οπότε ανήκουν στην κατηγορία των μη παραμετρικών διαγραμμάτων.

Οι Eilbott & Nadles (1965) μελέτησαν τις ιδιότητες των στατιστικών ελέγχων προτεραιότητας, στις περιπτώσεις που οι κατανομές F_X, F_Y είναι εκθετικές, προσδιορίζοντας την ισχύ τους. Στην εργασία του Shorack (1967) τα αποτελέσματα των Eilbott & Nadles (1965) επεκτείνονται σε μία ευρύτερη κλάση εναλλακτικών υποθέσεων της μορφής $H_A: F_Y(x) = (F_X(x))^\gamma$, όπου γ είναι ένας θετικός ακέραιος αριθμός. Οι παραπάνω υποθέσεις, που έχουν εισαχθεί από τον Lehmann (1953), ονομάζονται εναλλακτικές τύπου Lehmann (Lehmann alternatives) και περιλαμβάνουν την κλάση των εκθετικών και των Weibull κατανομών.

Είναι χρήσιμο να γίνει διάκριση των εναλλακτικών υποθέσεων Lehmann στις κλάσεις:

ι. $H_1: F_Y(x) = (F_X(x))^\gamma$ για $\gamma > 1$ ($\gamma < 1$), η οποία για μικρές τιμές της παραμέτρου γ , αντιστοιχεί σε μετατόπιση προς τα δεξιά (προς τα αριστερά) της κατανομής F_Y ως προς την F_X

- ii. $H_2: F_Y(x) = 1 - (1 - F_X(x))^\gamma$ ($\gamma > 1$ ($\gamma < 1$)), η οποία για μικρές τιμές της παραμέτρου γ , αντιστοιχεί σε μετατόπιση προς τα αριστερά (προς τα δεξιά) της κατανομής F_Y ως προς την F_X .

Η εναλλακτική υπόθεση $H_1: F_Y(x) = (F_X(x))^\gamma$ ανήκει στη γενικότερη κλάση εναλλακτικών $H_A: F_Y(x) < F_X(x)$, για $\gamma > 1$. Η ισχύς του ελέγχου είναι η πιθανότητα απόρριψης της μηδενικής υπόθεσης H_0 , ενώ ισχύει η εναλλακτική υπόθεση. Μια ιδιαίτερα δημοφιλής επιλογή στατιστικής συνάρτησης προτεραιότητας για τον έλεγχο ισονομίας δύο δειγμάτων και συνεπώς για την κατασκευή ενός ελεύθερου κατανομής διαγράμματος ελέγχου, είναι $r = (n + 1) / 2$ με n περιττό. Στη συγκεκριμένη περίπτωση, η παρατήρηση $Y_{(r)}$ του τυχαίου δείγματος που λαμβάνεται από τη διεργασία είναι η διάμεσος, η οποία συγκρίνεται με τα όρια ελέγχου που έχουν εκτιμηθεί από κατάλληλες διατεταγμένες παρατηρήσεις του δείγματος αναφοράς. Οι Janacek & Meikle (1997) πρότειναν την κατασκευή ενός μη παραμετρικού διαγράμματος ελέγχου τύπου Shewhart που βασίζεται στη διάμεσο με όρια που υπολογίζονται με τη βοήθεια ενός εντός ελέγχου δείγματος αναφοράς, ενώ οι Chakraborti (2004) και Balakrishnan (2008a, 2008c) μελέτησαν το παραπάνω διάγραμμα, υπολογίζοντας το μέσο μήκος ροής του με τη βοήθεια μιας τεχνικής δέσμευσης και προτείνοντας ταυτόχρονα πρόσθετες γενικεύσεις του. Τέλος για τον έλεγχο των υποθέσεων που διατυπώνονται ως εξής

$$H_0 : F_x = F_y \text{ έναντι } H_1 : F_x < F_y$$

ή ισοδύναμα για την κατασκευή διαγραμμάτων ελέγχου προτεραιότητας, μπορεί να χρησιμοποιηθεί η στατιστική συνάρτηση Wilcoxon rank-sum, η οποία έχει εισαχθεί στην εργασία του Wilcoxon (1945).

3.2 Μη παραμετρικά διαγράμματα ελέγχου με χρήση της διαμέσου

Το πρόβλημα της ισονομίας δύο ανεξάρτητων δειγμάτων έχει μελετηθεί εκτενώς στη διεθνή βιβλιογραφία. Ένας δημοφιλής και ελευθέρως κατανομής έλεγχος για δύο δείγματα μεγέθους m , n είναι ο από κοινού έλεγχος διαμέσου (joint median test) των Mood (1954) και Westenberg (1948, 1952). Αν τα δείγματα X_1, X_2, \dots, X_m και Y_1, Y_2, \dots, Y_n προέρχονται από τις κατανομές F και G αντίστοιχα, τότε ο συγκεκριμένος έλεγχος βασίζεται στο πλήθος των Y - παρατηρήσεων που είναι μικρότερες (προηγούνται) από τη μεσαία παρατήρηση δ του συνόλου των $(m + n)$ παρατηρήσεων. Η μηδενική υπόθεση $H_0 : F = G$ μπορεί να εξετασθεί με τη βοήθεια των αναλογιών των X - και Y - παρατηρήσεων που βρίσκονται πριν τη διάμεσο δ . Πρόσθετες γενικεύσεις του παραπάνω ελέγχου για διάφορα τεταρτημόρια του από κοινού δείγματος, έχουν προταθεί στην εργασία του Massey (1951). Εναλλακτικά, ο Mathisen (1943) έχει εισάγει έναν έλεγχο που βασίζεται στη διάμεσο για τη σύγκριση των κατανομών δύο δειγμάτων. Συγκεκριμένα το τεστ του Mathisen (1943) (control median test) στηρίζεται στο πλήθος των Y - παρατηρήσεων που είναι μικρότερες από τη διάμεσο του X - δείγματος, με την προϋπόθεση ότι το μέγεθος του X - και Y - δείγματος είναι περιττό και άρτιο αντίστοιχα. Αξίζει να σημειωθεί ότι οι Kimball (1957) και οι συνεργάτες του χρησιμοποίησαν τον παραπάνω έλεγχο για τη μελέτη των χρόνων επιβίωσης πειραματικών μονάδων που ήταν χωρισμένες σε δύο

ομάδες και υποβάλλονταν σε ακτινοβολία, με σκοπό την εξέταση της επίδρασης ενός φαρμάκου ενάντια στον ιονισμό της ακτινοβολίας. Το φάρμακο χορηγήθηκε μόνο στις μονάδες της πειραματικής ομάδας, στην οποία καταγράφονται οι χρόνοι θανάτου μέχρι το θάνατο της μεσαίας μονάδας της ομάδας ελέγχου. Ο αριθμός των θανάτων που συμβαίνουν πριν το θάνατο της μεσαίας μονάδας της ομάδας ελέγχου, αποτελεί το κριτήριο για τη λήψη της απόφασης για το αν το φάρμακο συντελεί στην επιβίωση μιας πειραματικής μονάδας που υποβάλλεται σε ακτινοβολία. Μολονότι ο Bowker (1944) επισήμανε ότι η παραπάνω συνάρτηση δεν θα είναι συνεπής σε ορισμένες περιπτώσεις, όπως για παράδειγμα όταν οι κατανομές των δύο δειγμάτων έχουν τη ίδια διάμεσο, η συνάρτηση κρίνεται συνεπής για εναλλακτικές υποθέσεις μετατόπισης θέσης. Ο Gart (1963) μελέτησε ιδιότητες της διαμέσου ως ελεγχουσυνάρτησης, αποδεικνύοντας ότι είναι ασυμπτωτικά τόσο αποτελεσματική όσο το από κοινού τεστ διαμέσου που περιγράφηκε παραπάνω. Πρόσθετα, ο Gastwirth (1968) επισήμανε ότι η χρήση της διαμέσου σε ελέγχους για το αν δύο δείγματα προέρχονται από την ίδια κατανομή, μειώνει το πλήθος των παρατηρήσεων που απαιτούνται για την εφαρμογή τους. Είναι εύκολο να διαπιστώσει κανείς ότι η διάμεσος του δείγματος αντιστοιχεί στη στατιστική συνάρτηση προτεραιότητας $P_{(r)}$ με $r = \frac{(m+1)}{2}$, m περιττός, που χρησιμοποιεί το άθροισμα των συχνοτήτων M_1, M_2, \dots, M_r , για δύο δείγματα X_1, X_2, \dots, X_m και Y_1, Y_2, \dots, Y_n .

Η χρήση των M_i οδηγεί στην κατασκευή πρόσθετων στατιστικών συναρτήσεων όπως αυτή των Haga (1959) και Hajek και Sidak (1967), οι οποίοι μελέτησαν τις συναρτήσεις $M_{m+1} + M_1^* - (M_{n+1}^* + M_1)$ και $\min(M_{m+1}, M_1^*) - \min(M_{n+1}^*, M_1)$ όπου M_1, M_{m+1} εκφράζουν το πλήθος των Y -παρατηρήσεων που είναι μικρότερες από την πρώτη διατεταγμένη παρατήρηση $Y_{1:n}$ και μεγαλύτερες από τη n -οστή $Y_{n:n}$ αντίστοιχα. Η χρήση της δειγματικής διαμέσου βρίσκει σημαντικές εφαρμογές στο Στατιστικό Έλεγχο Ποιότητας για την παρακολούθηση μιας παραγωγικής διαδικασίας. Οι Janacek και Meikle (1997) πρότειναν την κατασκευή ενός μη παραμετρικού διαγράμματος ελέγχου τύπου Shewhart που απεικονίζει τη διάμεσο του δείγματος που εξετάζεται από τη διεργασία. Ας υποθέσουμε ότι λαμβάνεται ένα δείγμα αναφοράς X_1, X_2, \dots, X_m από την εντός ελέγχου διεργασία, δηλαδή έχουμε ένα τυχαίο δείγμα μεγέθους m που προέρχεται από την εντός ελέγχου (αθροιστική) κατανομή $F_x(x) = F(x)$, όπου το μέγεθος m είναι συνήθως μεγάλο. Στη συνέχεια, εξάγονται από την παραγωγική διαδικασία τυχαία δείγματα (test samples) που είναι ανεξάρτητα μεταξύ τους (και από το δείγμα αναφοράς) και ενδιαφερόμαστε να ελέγξουμε αν η παρατηρούμενη διεργασία παραμένει εντός ελέγχου ή όχι. Με άλλα λόγια, αν Y_1, Y_2, \dots, Y_n είναι ένα τυχαίο δείγμα που λαμβάνεται από διεργασία με αθροιστική συνάρτηση κατανομής $F_y(x) = G(x)$, τότε ο στόχος είναι η γρήγορη ανίχνευση μιας πιθανής αλλαγής στην κατανομή της διεργασίας από $F(x)$ σε $G(x)$, δηλαδή ο έλεγχος της μηδενικής υπόθεσης $H_0: F(x) = G(x)$ έναντι της δίπλευρης εναλλακτικής $H_1: F(x) \neq G(x)$.

Στο διάγραμμα των Janacek και Meikle (1997), αντί να γίνει χρήση των δειγματικών μέσων, υπολογίζεται η δειγματική διάμεσος δ του εξεταζόμενου τυχαίου δείγματος και συγκρίνεται με τα όρια ελέγχου, όπως αυτά έχουν καθορισθεί με βάση το δείγμα αναφοράς. Δεδομένου ότι η εντός ελέγχου κατανομή είναι άγνωστη, χρησιμοποιούνται ως όρια ελέγχου δύο διατεταγμένες παρατηρήσεις του δείγματος

αναφοράς. Σύμφωνα με τους Janacek και Meikle (1997), διατάσσουμε σε αύξουσα σειρά τις παρατηρήσεις του δείγματος αναφοράς, $X_{1:m} \leq X_{2:m} \leq \dots \leq X_{m:m}$, και στη συνέχεια επιλέγουμε δύο συμμετρικά όρια ελέγχου,

$$LCL = X_{a:m} \text{ και } UCL = X_{m-a+1:m}$$

Η παράμετρος σχεδιασμού, $1 \leq a \leq m$, προσδιορίζεται κατά τέτοιο τρόπο ώστε η πιθανότητα λανθασμένου συναγερμού (*FAR*), δηλαδή η πιθανότητα το διάγραμμα να παράγει λανθασμένο σήμα ότι η διεργασία είναι εκτός ελέγχου, να μην υπερβαίνει ένα προκαθορισμένο επίπεδο f , δηλαδή

$$P(LCL < \delta < UCL | H_0) \geq 1 - f$$

Το βασικό πλεονέκτημα του παραπάνω διαγράμματος ελέγχου αποτελεί το γεγονός ότι, εξαιτίας της χρήσης της διαμέσου του εξεταζόμενου τυχαίου δείγματος (αντί της μέσης τιμής του), δεν είναι ευαίσθητο σε πιθανά λάθη μετρήσεων και επιπλέον μπορεί να χρησιμοποιηθεί σε περιπτώσεις που το υπό μελέτη μετρήσιμο χαρακτηριστικό περιλαμβάνει μια υποκειμενική αξιολόγηση ή ακόμη και διατάξιμα δεδομένα. Ο παρακάτω πίνακας 3.3 παρουσιάζει τις πιθανότητες λανθασμένου συναγερμού για το διάγραμμα των Janacek και Meikle (1997), για διάφορες επιλογές των μεγεθών των δύο δειγμάτων m, n και της παραμέτρου σχεδιασμού a .

Πίνακας 3.3: Πιθανότητες λανθασμένου συναγερμού του διαγράμματος των Janacek και Meikle (1997)

		Δείγμα αναφοράς μεγέθους m			
n	a	25	30	35	40
		<i>FAR</i>	<i>FAR</i>	<i>FAR</i>	<i>FAR</i>
5	1	0.0049	0.0031	0.0020	0.0014
	2	0.0196	0.0117	0.0054	0.0054
	3	0.0439	0.0278	0.0186	0.0131
	4	0.0826	0.0528	0.0357	0.0252
	5	0.1359	0.0877	0.0598	0.0425
7	1	0.0020	0.0011	0.0006	0.0004
	2	0.0089	0.0049	0.0029	0.0019
	3	0.0244	0.0137	0.0082	0.0052
	4	0.0517	0.0294	0.0179	0.0115
	5	0.0938	0.0542	0.0335	0.0217
9	1	0.0009	0.0004	0.0002	0.0001
	2	0.0048	0.0024	0.0013	0.0007
	3	0.0149	0.0075	0.0041	0.0024
	4	0.0348	0.0179	0.0099	0.0058
	5	0.0686	0.0360	0.0202	0.0121

Επεκτείνοντας την ιδέα των Janacek και Meikle (1997), ο Chakraborti (2004) μελέτησε μία κλάση μη παραμετρικών διαγραμμάτων ελέγχου τύπου Shewhart, τα οποία χρησιμοποιούν ως όρια ελέγχου κατάλληλα επιλεγμένες διατεταγμένες παρατηρήσεις $X_{\alpha:m}, X_{b:m}, 1 \leq a < b \leq m$, του δείγματος αναφοράς. Αφού πρώτα συλλεγεί ένα τυχαίο δείγμα Y_1, Y_2, \dots, Y_n από τη διεργασία, υπολογίζεται η j -οστή διατεταγμένη παρατήρηση $Y_{j:n}$ και συγκρίνεται με τα όρια ελέγχου

$$LCL = X_{\alpha:m} \text{ και } UCL = X_{b:m}$$

Αν η τιμή της $Y_{j:n}$ βρεθεί μεταξύ των ορίων ελέγχου, τότε η διεργασία κρίνεται ότι είναι εντός ελέγχου, ενώ σε διαφορετική περίπτωση η διεργασία έχει μετατοπιστεί σε κατάσταση εκτός ελέγχου και απαιτείται διερεύνηση για τις αιτίες που προκάλεσαν τη μετατόπιση.

Η πιθανότητα ότι το προτεινόμενο διάγραμμα δεν παράγει σήμα για εκτός ελέγχου μετατόπιση, μπορεί να εκφραστεί με τη βοήθεια της στατιστικής συνάρτησης προτεραιότητας $P_{(j)}$. Συγκεκριμένα, το συμβάν ότι η τιμή της $Y_{j:n}$ βρίσκεται ανάμεσα στις διατεταγμένες παρατηρήσεις $X_{\alpha:m}$ και $X_{b:m}$ είναι ισοδύναμο με το συμβάν ότι η συνάρτηση $P_{(j)}$ παίρνει τιμές μεταξύ a και $b - 1$. Συνεπώς ισχύει η ακόλουθη σχέση

$$p = P(X_{\alpha:m} \leq Y_{j:n} \leq X_{b:m}) = P(a \leq P_{(j)} \leq b - 1)$$

Πράγματι, δεδομένου ότι η κατανομή είναι συνεχής, η ανίσωση $Y_{j:n} \leq X_{b:m}$ ισχύει αν και μόνο αν το πολύ $(b - 1)$ X -παρατηρήσεις είναι μεγαλύτερες από την $Y_{j:n}$, ενώ με παρόμοιο τρόπο διαπιστώνουμε ότι η διάταξη $X_{\alpha:m} \leq Y_{j:n}$ είναι αληθής αν και μόνο αν τουλάχιστον α X -παρατηρήσεις είναι μικρότερες από την $Y_{j:n}$. Αξίζει να σημειωθεί ότι η πιθανότητα $1 - p$ είναι η πιθανότητα το διάγραμμα να δώσει σήμα ότι η διεργασία μετατοπίστηκε σε εκτός ελέγχου κατάσταση. Ο προσδιορισμός των παραμέτρων σχεδιασμού m, n, j γίνεται με τέτοιο τρόπο ώστε η πιθανότητα λανθασμένου συναγερού (FAR) να μην υπερβαίνει ένα προκαθορισμένο επίπεδο f . Η χρήση των αποτελεσμάτων για τις στατιστικές συναρτήσεις προτεραιότητας, οδηγεί στην ακόλουθη συνθήκη, η οποία χρησιμοποιείται για την επιλογή του σχεδιασμού του μη παραμετρικού διαγράμματος ελέγχου.

$$\sum_{w=a}^{b-1} \frac{\binom{j+w-1}{w} \binom{m+n-j-w}{m-w}}{\binom{m+n}{n}} \geq 1 - f$$

Ο παρακάτω πίνακας 3.4 παρουσιάζει τις πιθανότητες λανθασμένου συναγερού για το διάγραμμα του Chakraborti (2004), για διάφορες επιλογές των m, n και των παραμέτρων σχεδιασμού j, a , ώστε να ικανοποιείται η παραπάνω σχέση για $f = 0.01, 0.05$, και 0.0027 .

Πίνακας 3.4: Πιθανότητες λανθασμένου συναγερμού του διαγράμματος Chakraborti (2004) για τον σχεδιασμό ($j = \frac{n+1}{2}, b = m - a + 1$)

		Δείγμα αναφοράς μεγέθους m							
f	n	50		100		500		1000	
		a	FAR	a	FAR	a	FAR	a	FAR
0.01	5	3	0.0036	7	0.0043	40	0.0048	82	0.0050
	11	6	0.0047	15	0.0043	83	0.0049	167	0.0049
	25	13	0.0031	23	0.0040	127	0.0047	258	0.0050
0.005	5	2	0.0015	5	0.0018	31	0.0023	64	0.0025
	11	5	0.0013	13	0.0023	72	0.0024	146	0.0024
	25	9	0.0016	21	0.0020	118	0.0024	239	0.0025
0.0027	5	1	0.0004	4	0.0010	25	0.0013	51	0.0013
	11	5	0.0013	11	0.0011	64	0.0013	130	0.0013
	25	8	0.0008	19	0.0009	110	0.0013	224	0.0013

Από διάφορες αριθμητικές συγκρίσεις που έχουν γίνει στην εργασία του Chakraborti (2004), φαίνεται ότι το διάγραμμα που προτείνεται υπερέρχει ενός \bar{X} –διαγράμματος ελέγχου. Το βασικό μειονέκτημα του συγκεκριμένου διαγράμματος είναι ότι η απόφαση βασίζεται σε μία μόνο διατεταγμένη παρατήρηση (διάμεσος ή κάποιο άλλο ποσοστημόριο).

Το γεγονός αυτό έχει ως πιθανό αποτέλεσμα, ότι ενώ η συγκεκριμένη παρατήρηση μπορεί να βρίσκεται μεταξύ των ορίων ελέγχου, ταυτόχρονα το μεγαλύτερο πλήθος των παρατηρήσεων του εξεταζόμενου δείγματος να βρίσκεται εκτός των ορίων αυτών. Σε αυτές τις περιπτώσεις, το διάγραμμα θα δίνει σήμα ότι η διεργασία βρίσκεται εντός ελέγχου, ενώ στην πραγματικότητα θα υπάρχει ισχυρή ένδειξη για το αντίθετο συμπέρασμα.

3.3 Ένα νέο μη παραμετρικό διάγραμμα ελέγχου

Στη παράγραφο αυτή προτείνεται από τους Balakrishnan, N., Triantafyllou, I.S. & Koutra, M.V. (2010) ένα νέο μη παραμετρικό διάγραμμα ελέγχου τύπου Shewhart, το οποίο λαμβάνει υπόψιν του, όχι μόνο τη θέση της j -οστής διατεταγμένης παρατήρησης του εξεταζόμενου δείγματος, αλλά και το πλήθος των παρατηρήσεων του που βρίσκονται μεταξύ των ορίων ελέγχου.

Έστω ότι ένα δείγμα αναφοράς X_1, X_2, \dots, X_m συλλέγεται από την εντός ελέγχου κατανομή $F(x)$, δηλαδή όταν γνωρίζουμε ότι η παραγωγική διαδικασία είναι εντός ελέγχου. Στη συνέχεια, λαμβάνουμε ένα τυχαίο δείγμα Y_1, Y_2, \dots, Y_n από τη διεργασία και θέλουμε να αποφασίσουμε αν παραμένει εντός ελέγχου ή αν έχει μετατοπιστεί σε μια εκτός ελέγχου κατανομή $G(x)$. Για το σκοπό αυτό, επιλέγουμε δύο συγκεκριμένες διατεταγμένες παρατηρήσεις από το δείγμα αναφοράς και τις χρησιμοποιούμε ως όρια ελέγχου του προτεινόμενου διαγράμματος, όπως φαίνεται παρακάτω

$$LCL = X_{a;m}, UCL = X_{b;m}, \text{ όπου } 1 \leq a < b \leq m$$

Για τη διατύπωση του κανόνα απόφασης του νέου διαγράμματος χρησιμοποιούμε την j -οστή διατεταγμένη παρατήρηση $Y_{j:n}$ σε συνδυασμό με την ακόλουθη στατιστική συνάρτηση

$$R = R(Y_1, Y_2, \dots, Y_n; X_{a;m}, X_{b;m}) = |\{i \in \{1, 2, \dots, n\} : X_{a;m} \leq Y_i \leq X_{b;m}\}|$$

Πιο συγκεκριμένα, η παραγωγική διαδικασία κρίνεται ότι είναι εντός ελέγχου, αν επαληθεύονται οι δύο επόμενες συνθήκες

$$LCL \leq Y_{j:n} \leq UCL \text{ και } R \geq r$$

Οι παράμετροι a, b, j, r είναι παράμετροι σχεδιασμού του προτεινόμενου διαγράμματος. Στη διεθνή βιβλιογραφία, έχουν προταθεί δύο τρόποι προσδιορισμού των παραμέτρων σχεδιασμού ενός διαγράμματος ελέγχου. Σύμφωνα με τον πρώτο τρόπο, αρχικά καθορίζουμε μια συγκεκριμένη πιθανότητα λανθασμένου συναγερμού (FAR) που θέλουμε να πετύχουμε στο διάγραμμα και στη συνέχεια προχωρούμε στην κατασκευή του διαγράμματος. Η πιθανότητα λανθασμένου συναγερμού (FAR) του νέου διαγράμματος, δηλαδή η πιθανότητα το διάγραμμα να παράγει σήμα για εκτός ελέγχου μετατόπιση της διεργασίας, ενώ στην πραγματικότητα βρίσκεται εντός ελέγχου, δίνεται από τη σχέση

$$FAR = 1 - P_c(LCL \leq Y_{j:n} \leq UCL \text{ και } R \geq r)$$

όπου ο δείκτης C δηλώνει ότι η διεργασία είναι εντός ελέγχου, δηλαδή $F = G$. Ο δεύτερος τρόπος προσδιορισμού των παραμέτρων σχεδιασμού του διαγράμματος, είναι η χρήση μιας συγκεκριμένης τιμής του εντός ελέγχου μέσου μήκους ροής ARL_{in} , όπως για παράδειγμα 370 ή 500. Το ARL_{in} του νέου διαγράμματος έχει την ίδια τιμή για όλες τις συνεχείς κατανομές, συνεπώς το νέο διάγραμμα είναι ένα μη παραμετρικό διάγραμμα ελέγχου. Ωστόσο η κατανομή του μήκους ροής N δεν είναι γεωμετρική, συνεπώς δεν μπορούμε να χρησιμοποιήσουμε στη συγκεκριμένη περίπτωση τα γνωστά αποτελέσματα για την κατανομή αυτή. Αυτό έχει ως αποτέλεσμα ότι το ARL_{in} του διαγράμματος δεν θα είναι ίσο με το αντίστροφο της πιθανότητας να παραχθεί σήμα για εκτός ελέγχου μετατόπιση, όπως συμβαίνει για άλλα διαγράμματα ελέγχου. Για το λόγο αυτό, θα προχωρήσουμε στον υπολογισμό της κατανομής του εντός ελέγχου μήκους ροής για το συγκεκριμένο διάγραμμα, χρησιμοποιώντας μια τεχνική δέσμησης που εφαρμόστηκε για πρώτη φορά στην εργασία του Chakraborti (2000). Αν οι παρατηρήσεις Y_1, Y_2, \dots, Y_n του εξεταζόμενου δείγματος ακολουθούν μια συνεχή κατανομή $G(x)$, τότε η πιθανότητα ότι το προτεινόμενο διάγραμμα δεν παράγει σήμα για εκτός ελέγχου μετατόπιση δίνεται στην ακόλουθη σχέση

$$p = p(m, n, a, b, j, r; F; G) = P(X_{a;m} \leq Y_{j:n} \leq X_{b;m} \text{ και } R(Y_1, Y_2, \dots, Y_n; X_{a;m}, X_{b;m}) \geq r)$$

Η παραπάνω πιθανότητα είναι η χαρακτηριστική καμπύλη (operating characteristic curve) του διαγράμματος, ενώ η πιθανότητα να παραχθεί σήμα για εκτός ελέγχου μετατόπιση είναι ίση με $1 - p$. Συνεπώς, υπό τη μηδενική υπόθεση $H_0: F = G$, η πιθανότητα FAR του διαγράμματος είναι ίση με

$$FAR = 1 - p(m, n, a, b, j, r; F, F)$$

Η πιθανότητα λανθασμένου συναγερμού όπως έχει οριστεί παραπάνω αποδεικνύεται να είναι ίση με

$$FAR = 1 - \sum_{r-1 \leq c+d \leq n-1} \frac{\binom{j-c+a-2}{a-1} \binom{m+n-b-d-j}{n-j-d} \binom{b+c+d-a}{c+d+1}}{\binom{m+n}{n}}$$

Αν η διεργασία μετατοπιστεί σε εκτός ελέγχου κατάσταση τότε παρέχεται η πιθανότητα ότι το διάγραμμα ελέγχου δεν παράγει σήμα, γεγονός που δεν είναι το επιθυμητό. Είναι φανερό ότι η συγκεκριμένη πιθανότητα εξαρτάται τόσο από την εντός ελέγχου κατανομή $F(x)$, όσο και από την εκτός ελέγχου $G(x)$. Δεδομένου ότι το τελικό αποτέλεσμα εξαρτάται από τη μορφή της συνάρτησης $G \circ F^{-1}$, θα ήταν παράλογο να αναμένουμε μια ακριβή έκφραση στη γενική περίπτωση. Ωστόσο για την κλάση των εναλλακτικών υποθέσεων τύπου Lehmann, που περιγράφονται στη παρούσα εργασία αξίζει να σημειωθεί ότι όταν η εκτός ελέγχου κατανομή έχει τη μορφή $G(x) = [F(x)]^\gamma$ ($G(x) = 1 - (1 - F(x))^\gamma$), όπου γ είναι θετικός ακέραιος, η εναλλακτική υπόθεση τύπου Lehmann δηλώνει ότι οι Y - μεταβλητές κατανέμονται ως η μεγαλύτερη (μικρότερη) από το πλήθος γX - μεταβλητών. Μια εναλλακτική υπόθεση της μορφής $G(x) = 1 - (1 - F(x))^\gamma$ καλύπτει τις περιπτώσεις που η εντός ελέγχου κατανομή είναι εκθετική με μέσο λ και οι μετατοπίσεις που θέλουμε να ανιχνεύσουμε δημιουργούνται με αλλαγή της παραμέτρου λ . Η χαρακτηριστική συνάρτηση λειτουργίας $p' = p'(m, n, a, b, j, r; \gamma)$ για την εναλλακτική υπόθεση $G(x) = 1 - (1 - F(x))^\gamma$ δίνεται από την ακόλουθη σχέση

$$p'(m, n, a, b, j, r; \gamma) = \sum_{r-1 \leq c+d \leq n-1} \int_0^1 \int_0^t q_{c,d}(1 - (1-s)^\gamma, 1 - (1-t)^\gamma) f(s, t) ds dt$$

$$\text{με } q_{c,d}(v, w) = \binom{n}{j-c-1, c+d+1, n-j-d} v^{j-c-1} (w-v)^{c+d+1} (1-w)^{n-j-d}$$

$$\text{και } f(s, t) = \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b}, 0 < s < t < 1$$

Με ανάλογο τρόπο, η υπόθεση της μορφής $G(x) = [F(x)]^\gamma$ καλύπτει τις περιπτώσεις που η εντός ελέγχου κατανομή είναι ομοιόμορφη στο διάστημα $(0,1)$, ενώ η εκτός ελέγχου διεργασία περιγράφεται από την οικογένεια κατανομών με αθροιστική συνάρτηση κατανομής $G(x) = x^\gamma, 0 < x < 1$

Η χαρακτηριστική συνάρτηση λειτουργίας $p = p(m, n, a, b, j, r; \gamma)$ για την εναλλακτική υπόθεση $G(x) = [F(x)]^\gamma$ δίνεται από την ακόλουθη σχέση

$$p(m, n, a, b, j, r; \gamma) = \alpha(b - a) \binom{m}{a, b - a, m - b} \times$$

$$\sum_{r-1 \leq c+d \leq n-1} \binom{n}{j - c - 1, c + d + 1, n - j - d} \times$$

$$B_\gamma(\gamma(j - c - 1) + a, b - a; c + d + 1) B_\gamma(\gamma(j + d) + b, m - b + 1; n - j - d)$$

$$\text{με } B_\gamma(a, b; l) = \sum_{k=0}^l (-1)^k \binom{l}{k} B(a + k\gamma, b)$$

Οι ποσότητες $1 - p$ και $1 - p'$ εκφράζουν τις πιθανότητες συναγερμού όταν η διεργασία έχει πράγματι μετατοπιστεί σε μια κατανομή τύπου Lehmann με παράμετρο γ .

Ο παρακάτω πίνακας 3.5 παρουσιάζει τις πιθανότητες FAR του νέου διαγράμματος για διάφορους σχεδιασμούς που αντιστοιχούν σε διαφορετικές τιμές των a, b, m, n, j, r .

Πίνακας 3.5: Πιθανότητες λανθασμένου συναγερμού για δεδομένο σχεδιασμό
 $(b = m - a + 1, j = \frac{n+1}{2})$

		Δείγμα αναφοράς μεγέθους m							
n	r	40		60		100		200	
		a	FAR	a	FAR	a	FAR	a	FAR
5	1	5	0.0425	9	0.0616	16	0.0689	30	0.0559
		6	0.0654	10	0.0800	17	0.0805	33	0.0719
		7	0.0943	11	0.1011	18	0.0931	37	0.0971
	2	5	0.0482	8	0.0531	14	0.0572	27	0.0495
		6	0.0759	9	0.0721	15	0.0690	30	0.0665
		7	0.1116	10	0.0950	16	0.0821	34	0.0942
	3	4	0.0647	6	0.0626	10	0.0608	19	0.0520
		5	0.1091	7	0.0911	11	0.0773	21	0.0674
		6	0.1656	8	0.1252	12	0.0959	24	0.0946
11	3	8	0.0459	13	0.0558	23	0.0653	44	0.0481
		9	0.0737	14	0.0770	24	0.0796	48	0.0737
		10	0.1129	15	0.1039	25	0.0961	51	0.0988
	4	7	0.0363	11	0.0384	20	0.0517	40	0.0472
		8	0.0642	12	0.0574	21	0.0657	43	0.0687
		9	0.1061	13	0.0895	22	0.0824	46	0.0968
	5	6	0.0375	10	0.0530	17	0.0532	34	0.0496
		7	0.0715	11	0.0807	18	0.0696	37	0.0750
		8	0.1228	12	0.1172	19	0.0892	39	0.0962
25	7	9	0.0211	15	0.0324	27	0.0475	55	0.0462
		10	0.0431	16	0.0536	28	0.0652	57	0.0651
		11	0.0825	17	0.0857	29	0.0881	59	0.0900
	10	7	0.0144	12	0.0287	22	0.0500	45	0.0525
		8	0.0364	13	0.0526	23	0.0714	47	0.0761
		9	0.0798	14	0.0898	24	0.0991	49	0.1070
	12	6	0.0212	10	0.0327	18	0.0462	37	0.0490
		7	0.0539	11	0.0611	19	0.0675	39	0.0726
		8	0.1148	12	0.1048	20	0.0953	41	0.1036

Από τον παραπάνω πίνακα 3.5 μπορεί να κατασκευαστεί ένα μη παραμετρικό διάγραμμα ελέγχου που στοχεύει σε μια προκαθορισμένη πιθανότητα λανθασμένου συναγερμού f . Με τη χρήση των τεσσάρων παραμέτρων (m, n, a, r) στο συγκεκριμένο διάγραμμα, έχουμε την ευελιξία να σταθεροποιήσουμε ορισμένες από αυτές και στη συνέχεια να αναζητήσουμε τη βέλτιστη επιλογή για τις υπόλοιπες, ή διαφορετικά να βρούμε έναν αποδεκτό συνδυασμό των παραπάνω παραμέτρων που ικανοποιεί συγκεκριμένες απαιτήσεις. Παραδείγματος χάριν, αν έχουμε ένα δείγμα αναφοράς μεγέθους $m = 100$ και τα τυχαία δείγματα που εξετάζουμε από τη διεργασία, επιθυμούμε να είναι μεγέθους $n = 25$, τότε μία πιθανότητα λανθασμένου συναγερμού που είναι σχεδόν ίση με $f = 0.05$ μπορεί να πραγματοποιηθεί χρησιμοποιώντας

- την 27^η μικρότερη και μεγαλύτερη παρατήρηση του δείγματος αναφοράς ($a = 27$) και $r = 7$ ($FAR = 0.0475$) ή
- την 22^η μικρότερη και μεγαλύτερη παρατήρηση του δείγματος αναφοράς ($a = 22$) και $r = 10$ ($FAR = 0.0500$) ή
- την 18^η μικρότερη και μεγαλύτερη παρατήρηση του δείγματος αναφοράς ($a = 18$) και $r = 12$ ($FAR = 0.0462$).

Στις περιπτώσεις των εναλλακτικών υποθέσεων τύπου Lehmann μπορούν να εξαχθούν διάφορα χρήσιμα συμπεράσματα για την απόδοση του διαγράμματος ελέγχου με τη βοήθεια του παρακάτω πίνακα 3.6. Στο συγκεκριμένο πίνακα, υπολογίζονται οι πιθανότητες συναγερμού για διάφορες τιμές των παραμέτρων σχεδιασμού m, n, a, r και της παραμέτρου της εκτός ελέγχου κατανομής $\gamma > 0$. Παρουσιάζονται οι ακόλουθες δυο περιπτώσεις

- Η ομοιόμορφη κατανομή στο διάστημα $(0,1)$ ως εντός ελέγχου κατανομή ($F(x) = x, 0 < x < 1$) και μια εκτός ελέγχου κατανομή της μορφής $G(x) = x^\gamma, \gamma = 0.4, 4$. Τα ίδια αποτελέσματα θα προκύψουν αν η x αντικατασταθεί με οποιαδήποτε συνεχή κατανομή $F(x)$.
- Η εκθετική με μέσο 1 ($F(x) = 1 - \exp(-x), x > 0$) ως εντός ελέγχου κατανομή, και μια εκθετική ως εκτός ελέγχου κατανομή με μέσο $1/\gamma = 1/0.4, 1/4$. Επίσης τα ίδια αποτελέσματα θα προκύψουν αν η $\exp(-x)$ αντικατασταθεί με οποιαδήποτε συνεχή συνάρτηση επιβίωσης $1 - F(x)$.

Παρακάτω παρουσιάζεται ο πίνακας 3.6 με τις πιθανότητες συναγερμού FAR για τον σχεδιασμό ($b = m - a + 1, j = \frac{n+1}{2}$) και για $\gamma=0.4$ και $\gamma=4$. Βασιζόμενοι στα αποτελέσματα των πινάκων 3.5 και 3.6 μπορούμε να γίνει επιλογή ενός σχεδιασμού, ο οποίος όχι μόνο να επιτυγχάνει μια επιθυμητή τιμή πιθανότητας λανθασμένου συναγερμού, αλλά ταυτόχρονα να είναι αποτελεσματικός στην ανίχνευση συγκεκριμένων μετατοπίσεων της εντός ελέγχου κατανομής που αντιστοιχούν σε διάφορες τιμές της παραμέτρου γ των εναλλακτικών υποθέσεων τύπου Lehmann. Παραδείγματος χάριν, στην περίπτωση που $m = 100, n = 25$, εκ των τριών επιλογών σχεδιασμού για την επίτευξη πιθανότητας $FAR \cong 0.05$, ο σχεδιασμός που αντιστοιχεί στις τιμές $a = 27$ και $r = 7$ προσφέρει τη μεγαλύτερη πιθανότητα ανίχνευσης μιας μετατόπισης σε κατανομή τύπου Lehmann με $\gamma = 0.4$ ($AR = 0.8057$) και για $\gamma = 4$ ($AR = 0.9572$) όπου με AR συμβολίζουμε τη πιθανότητα συναγερμού (alarm rate).

Όπως έχει ήδη αναφερθεί, ένα δεύτερο κριτήριο για την αξιολόγηση της απόδοσης ενός διαγράμματος ελέγχου (και κατά συνέπεια για τον προσδιορισμό των κατάλληλων παραμέτρων σχεδιασμού) προκύπτει από την εξέταση της κατανομής του μήκους ροής (run length distribution), δηλαδή την κατανομή της τυχαίας μεταβλητής N που απαριθμεί το πλήθος των δειγμάτων μέχρι να παραχθεί για πρώτη φορά σήμα για εκτός ελέγχου μετατόπιση.

Πίνακας 3.6: Πιθανότητες συναγερωμού (AR) για δεδομένο σχεδιασμό ($b = m - a + 1, j = \frac{n+1}{2}$) και για $\gamma=0.4$ (άνω είσοδος) και $\gamma=4$ (κάτω είσοδος)

		Δείγμα αναφοράς μεγέθους m								
n	r	40		60		100		200		
		a	AR	a	AR	a	AR	a	AR	
5	1	5	0.3637 0.3448	9	0.4298 0.4371	16	0.4580 0.4874	30	0.4388 0.4517	
		6	0.4238 0.4278	10	0.4678 0.5040	17	0.4805 0.5283	33	0.4734 0.5167	
		7	0.4797 0.5226	11	0.5041 0.5674	18	0.5024 0.5677	37	0.5172 0.5975	
		2	5	0.3718 0.3292	8	0.3993 0.3681	14	0.4215 0.4021	27	0.4122 0.3845
	3	6	0.4359 0.4279	9	0.4421 0.4372	15	0.4471 0.4452	30	0.4510 0.4517	
		7	0.4969 0.5227	10	0.4834 0.5040	16	0.4722 0.4874	34	0.5010 0.5376	
		4	0.3660 0.2324	6	0.3728 0.2313	10	0.3785 0.2298	19	0.3646 0.2070	
		5	0.4518 0.3300	7	0.4312 0.2992	11	0.4142 0.2720	21	0.4010 0.2498	
	11	3	6	0.5317 0.4294	8	0.4871 0.3689	12	0.4460 0.3153	24	0.4539 0.3169
			8	0.5389 0.6379	13	0.5943 0.7301	23	0.6382 0.8029	44	0.6193 0.7902
			9	0.6052 0.7319	14	0.6373 0.7878	24	0.6633 0.8337	48	0.6707 0.8547
			10	0.6667 0.8085	15	0.6780 0.8362	25	0.6875 0.8607	51	0.7067 0.8921
4	7	0.4797 0.5287	11	0.5145 0.5876	20	0.5744 0.6870	40	0.5810 0.7078		
	8	0.5561 0.6379	12	0.5660 0.6631	21	0.6047 0.7300	43	0.6266 0.7713		
	9	0.6123 0.6381	13	0.6156 0.7301	22	0.6343 0.7682	46	0.6702 0.8246		
	5	6	0.4349 0.4101	10	0.5049 0.5054	17	0.5252 0.5381	34	0.5314 0.5521	
25	7	7	0.5257 0.5288	11	0.5656 0.5877	18	0.5623 0.5908	37	0.5875 0.6343	
		8	0.6123 0.6381	12	0.6239 0.6632	19	0.5985 0.6406	39	0.6236 0.6844	
		9	0.6365 0.7714	15	0.7338 0.8901	27	0.8057 0.9572	55	0.8311 0.9791	
		10	0.7148 0.8534	16	0.7805 0.9263	28	0.8303 0.9688	57	0.8542 0.9859	
10	11	0.7833 0.9112	17	0.8223 0.9521	29	0.8530 0.9776	59	0.8753 0.9907		
	7	0.4817 0.5345	12	0.6050 0.7001	22	0.7053 0.8314	45	0.7387 0.8850		
	8	0.5898 0.6638	13	0.6759 0.7781	23	0.7451 0.8680	47	0.7775 0.9152		
	9	0.6910 0.7714	14	0.7412 0.8412	24	0.7822 0.8981	49	0.8130 0.9386		
12	6	0.4594 0.3940	10	0.5638 0.5069	18	0.6518 0.6180	37	0.6925 0.6815		
	7	0.5910 0.5347	11	0.6519 0.6087	19	0.7031 0.6810	39	0.7425 0.7456		
	8	0.7100 0.6643	12	0.7313 0.7004	20	0.7503 0.7381	41	0.7876 0.8009		

Παρόλο που το νέο διάγραμμα ελέγχου είναι ένα διάγραμμα τύπου Shewhart, η κατανομή του μήκους ροής (ARL) του δεν είναι γεωμετρική, διότι τα συμβάντα παραγωγής σήματος για εκτός ελέγχου μετατόπιση δεν είναι ανεξάρτητα και συνεπώς το μέσο μήκος ροής (ARL) δεν είναι ίσο με τον αντίστροφο της πιθανότητας παραγωγής σήματος. Αλλά δίνεται από την σχέση

$$ARL_{in} = \int_0^1 \int_0^t \frac{1}{1 - q(s; t; r)} f(s, t) ds dt$$

όπου το $q(s; t; r)$ και $f(s, t)$ έχουν οριστεί παραπάνω.

Στον παρακάτω πίνακα 3.7 παρουσιάζονται οι εντός ελέγχου τιμές ARL_{in} για επιλεγμένες τιμές των m, n . Για την κατασκευή του διαγράμματος επιλέγονται ονομαστικές τιμές FAR ίσες με 0.01, 0.005 και 0.0027 το οποίο και σχετίζεται με το μέσο μήκος ροής $ARL_{in} = 370$ για ένα \bar{X} - διάγραμμα τύπου Shewhart. Η διατεταγμένη παρατήρηση του εξεταζόμενου δείγματος που απεικονίζεται στο διάγραμμα είναι η διάμεσος, ενώ τα a, b επιλέγονται συμμετρικά ($b = m - a + 1$). Η παράμετρος σχεδιασμού r , επιλέγεται σε κάθε περίπτωση κατάλληλα ώστε το διάγραμμα να επιτυγχάνει πιθανότητα λανθασμένου συναγερού, όσο το δυνατόν πιο κοντά στην τιμή στόχο κάθε φορά. Παραδειγματος χάριν, χρησιμοποιώντας τον πίνακα 3.7, αν χρησιμοποιήσουμε ένα δείγμα αναφοράς μεγέθους $m = 500$ και εξετάσουμε το αν η διεργασία παραμένει εντός ελέγχου με τη βοήθεια τυχαίων δειγμάτων μεγέθους $n = 11$ που εξάγονται από αυτή, τότε ο σχεδιασμός $a = 33, b = 468, j = 6, r = 7$ παράγει λανθασμένο συναγερό με πιθανότητα 0.00999, ενώ αναμένεται να δώσει $ARL_{in} = 127.7$

Πίνακας 3.7: Τα ARL_{in} για καθορισμένη τιμή FAR και δεδομένο σχεδιασμό ($b = m - a + 1, j = \frac{n+1}{2}$)

FAR	m	n	(LCL, UCL)	r	Ακρίβες FAR	ARL_{in}	
0.01	50	5	(3,48)	1	0.0072	635.7	
		11	(7,44)	1	0.0093	642.2	
		25	(6,45)	13	0.0097	3957820	
	100	5	(7,94)	1	0.0086	214.9	
		11	(6,95)	7	0.0091	518.8	
		25	(8,93)	16	0.0082	4820	
	500	5	(38,463)	2	0.0091	121	
		11	(33,468)	7	0.00999	127.7	
		25	(120,381)	7	0.0095	221	
	1000	5	(79,922)	2	0.00997	105.2	
		11	(96,905)	6	0.0098	113.2	
		25	(243,758)	7	0.00994	115.6	
	0.005	50	5	(2,49)	1	0.003	5671
			11	(5,46)	2	0.0025	9503.1
			25	(7,44)	11	0.0048	9073150
100		5	(5,96)	1	0.00352	678.4	
		11	(5,96)	7	0.0046	1700.2	
		25	(15,86)	11	0.00445	6223	
500		5	(20,481)	3	0.00477	254.1	
		11	(56,445)	5	0.0049	260.1	
		25	(90,411)	10	0.00499	319.8	
1000		5	(62,939)	2	0.0049	217.7	
		11	(136,865)	4	0.0049	223.9	
		25	(129,872)	13	0.00497	258.7	
0.0027		50	5	(1,50)	1	0.0008	∞
			11	(5,46)	2	0.0025	9503.1
			25	(7,44)	10	0.0021	29347500
	100	5	(4,97)	2	0.00214	1508	
		11	(9,92)	5	0.0023	2540.4	
		25	(12,89)	12	0.0019	45381	
	500	5	(16,485)	3	0.0026	503.8	
		11	(64,437)	2	0.00263	456	
		25	(101,400)	8	0.0026	599.1	
	1000	5	(49,952)	2	0.0025	442.7	
		11	(49,952)	7	0.00267	449.7	
		25	(189,812)	9	0.00269	476.1	

3.3.1 Μια παραλλαγή του νέου διαγράμματος ελέγχου

Ενδιαφέρον παρουσιάζει η περίπτωση κατά την οποία αντί για τη διατύπωση του κανόνα απόφασης, λάβουμε υπόψιν το πλήθος των παρατηρήσεων του εξεταζόμενου δείγματος που βρίσκονται ανάμεσα σε κάθε ένα όριο ελέγχου και τη στατιστική συνάρτηση που χρησιμοποιείται. Συγκεκριμένα, αφού συλλεγεί ένα τυχαίο δείγμα Y_1, Y_2, \dots, Y_n από τη διεργασία, χρησιμοποιείται η $j - \text{οστή}$ διατεταγμένη παρατήρηση σε συνδυασμό με τις επόμενες δύο στατιστικές συναρτήσεις

$$R_1 = R_1(Y_1, Y_2, \dots, Y_n; X_{a:m}, Y_{j:n}) = |\{i \in \{1, 2, \dots, n\} : X_{a:m} \leq Y_i \leq Y_{j:n}\}|$$

$$R_2 = R_2(Y_1, Y_2, \dots, Y_n; Y_{j:n}, X_{b:m}) = |\{i \in \{1, 2, \dots, n\} : Y_{j:n} \leq Y_i \leq X_{b:m}\}|$$

Η διεργασία θα θεωρείται ότι είναι εντός ελέγχου όταν θα ικανοποιούνται οι συνθήκες

$$LCL \leq Y_{j:n} \leq UCL, \quad R_1 \geq r_1 \text{ και } R_2 \geq r_2$$

Οι παράμετροι a, b, j, r_1, r_2 είναι οι παράμετροι σχεδιασμού του προτεινόμενου μη παραμετρικού διαγράμματος ελέγχου. Η πιθανότητα λανθασμένου συναγερού του διαγράμματος δίνεται από την ακόλουθη σχέση

$$FAR = 1 - \sum_{c=r_1}^{n-1} \sum_{d=r_2}^{n-c-1} \frac{\binom{j-c+a-2}{a-1} \binom{m+n-b-d-j}{n-j-d} \binom{b+c+d-a}{c+d+1}}{\binom{m+n}{n}}$$

και το εντός ελέγχου μέσο μήκος ροής δίνεται από τη σχέση

$$ARL_{in} = \int_0^1 \int_0^t \frac{1}{1 - q^*(s, t; r_1, r_2)} f(s, t) ds dt$$

Στον παρακάτω πίνακα 3.8 παρουσιάζονται οι εντός ελέγχου τιμές FAR για επιλεγμένες τιμές των m, n .

Πίνακας 3.8: Πιθανότητες λανθασμένου συναγερωμού για δεδομένο σχεδιασμό
 $(b = m - a + 1, j = (n + 1)/2)$

			Δείγμα αναφοράς μεγέθους m							
n	r_1	r_2	40		60		100		200	
			a	FAR	a	FAR	a	FAR	a	FAR
5	1	0	2	0.0316	2	0.0149	3	0.0111	7	0.0132
			3	0.0616	4	0.0478	7	0.0496	14	0.0476
			4	0.0999	6	0.0969	10	0.0943	20	0.0922
	1	1	1	0.0202	2	0.0279	2	0.0108	5	0.0138
			2	0.0575	3	0.0538	5	0.0505	10	0.0479
			3	0.1090	4	0.0865	7	0.0903	15	0.0988
11	2	1	3	0.0186	4	0.0107	8	0.0144	16	0.0120
			4	0.0392	7	0.0518	12	0.0499	24	0.0452
			6	0.1138	9	0.1053	15	0.0977	30	0.0917
	2	2	2	0.0116	4	0.0183	6	0.0106	13	0.0107
			3	0.0306	6	0.0547	10	0.0479	21	0.0496
			5	0.1095	7	0.0831	13	0.1025	26	0.0955
25	6	5	3	0.0183	5	0.0175	8	0.0097	18	0.0109
			4	0.0455	7	0.0596	12	0.0525	25	0.0505
			5	0.0927	8	0.0960	14	0.0971	29	0.0967
	6	6	2	0.0080	4	0.0118	7	0.0087	16	0.0099
			3	0.0631	6	0.0489	11	0.0539	22	0.0427
			5	0.1246	7	0.0836	13	0.1029	26	0.0879

3.3.2 Αριθμητικές συγκρίσεις του νέου διαγράμματος ελέγχου

Συνοψίζοντας θα μπορούσαμε να πούμε πως ένα βασικό πλεονέκτημα ενός μη παραμετρικού διαγράμματος ελέγχου είναι ότι το εντός ελέγχου μέσο μήκος ροής (ARL_{in}) δεν εξαρτάται από την κατανομή της διεργασίας. Επιπλέον, τα παραμετρικά διαγράμματα ελέγχου επηρεάζονται σημαντικά από τυχόν αλλαγή του τύπου της κατανομής του υπό μελέτη χαρακτηριστικού.

Η απόδοση ενός διαγράμματος ελέγχου αξιολογείται με τη βοήθεια της κατανομής του μήκους ροής του, αλλά και την πιθανότητα να παράγει (ή όχι) σήμα για εκτός ελέγχου μετατόπιση, όταν η διεργασία έχει πράγματι μετατοπιστεί. Ένας τρόπος σύγκρισης δύο διαφορετικών διαγραμμάτων είναι να επιλεχθεί κατάλληλος σχεδιασμός ώστε να επιτυγχάνουν στην πράξη την ίδια (περίπου) πιθανότητα λανθασμένου σχεδιασμού και στη συνέχεια να υπολογισθούν τα εντός ελέγχου μέσα μήκη ροής τους. Εναλλακτικά, μπορεί να χρησιμοποιηθεί μια κοινή τιμή στόχος για τα ARL_{in} και στη συνέχεια να συγκρίνουμε τις εκτός ελέγχου μέσες τιμές

του μήκους ροής τους (ARL_{out}) για διάφορες μετατοπίσεις της κατανομής της διεργασίας.

Θέλοντας να αξιολογηθεί η απόδοση του διαγράμματος (ΔΙΑΓΡΑΜΜΑ 1) όπως ορίζεται από τις σχέσεις

$$LCL \leq Y_{j;n} \leq UCL \text{ και } R \geq r$$

και του παραλλαγμένου διαγράμματος (ΔΙΑΓΡΑΜΜΑ 2) όπως ορίζεται από τις σχέσεις

$$LCL \leq Y_{j;n} \leq UCL, \quad R_1 \geq r_1 \text{ και } R_2 \geq r_2$$

συγκρίθηκαν τα δυο διαγράμματα των οποίων και τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα. Συγκεκριμένα, καθορίζεται το επίπεδο πιθανότητας λανθασμένου συναγερμού ώστε να προσδιοριστούν κατάλληλα οι παραμέτροι των δύο διαγραμμάτων και να υπολογιστούν τα εντός ελέγχου μέσα μήκη ροής τους. Υπό την επικεφαλίδα <<ΔΙΑΓΡΑΜΜΑ 1>> συνοψίζουμε τα αποτελέσματα που σχετίζονται με το μη παραμετρικό διάγραμμα ελέγχου που ορίζεται στις πρώτες σχέσεις, ενώ για το διάγραμμα που ορίζεται από τις δεύτερες σχέσεις οι αντίστοιχες τιμές βρίσκονται κάτω από τον τίτλο <<ΔΙΑΓΡΑΜΜΑ 2>>.

Πίνακας 3.9: Σύγκριση των ARL_{in} των δυο διαγραμμάτων ελέγχου με την ίδια τιμή FAR

FAR	m	n	ΔΙΑΓΡΑΜΜΑ 1				ΔΙΑΓΡΑΜΜΑ 2			
			(LCL, UCL)	r	Ακριβής FAR	ARL_m	(LCL, UCL)	r_1 r_2	Ακριβής FAR	ARL_{in}
0.05	100	5	(9, 92)	3	0.047	30.40	(7,94)	1 0	0.049	29.97
		11	(21, 80)	3	0.043	34.91	(12,89)	2 1	0.049	33.90
	200	5	(18, 183)	3	0.045	26.21	(14,187)	1 0	0.048	25.58
		11	(26, 175)	6	0.045	30.67	(24,177)	2 1	0.045	28.93
0.10	100	5	(12, 89)	3	0.096	12.98	(10,91)	1 0	0.094	13.40
		11	(25, 76)	3	0.096	13.63	(15,86)	2 1	0.098	14.35
	200	5	(24, 177)	3	0.095	11.76	(20,181)	1 1	0.092	12.19
		11	(46, 155)	4	0.097	12.19	(30,171)	2 1	0.092	12.94

Από τον παραπάνω πίνακα 3.10 μπορούμε να καταλάβουμε πως τα δύο μη παραμετρικά διαγράμματα ελέγχου παρουσιάζουν παρόμοια απόδοση, αφού για την ίδια σχεδόν τιμή FAR , πετυχαίνουν πολύ κοντινές τιμές του μέσου μήκους ροής τους.

Παρακάτω παρουσιάζεται ο πίνακας στον οποίο έχουν καθοριστεί τρία διαφορετικά επίπεδα για την πιθανότητα λανθασμένου συναγερμού και διάφορες τιμές των παραμέτρων m, n . Για την κατασκευή του διαγράμματος επιλέγονται οι παράμετροι σχεδιασμού a, b, r ώστε η ακριβής τιμή της πιθανότητας FAR να είναι όσο το δυνατόν πιο κοντά στο προκαθορισμένο επιθυμητό επίπεδο. Υπολογίζεται το μέσο μήκος ροής του διαγράμματος. Γίνεται σύγκριση μεταξύ του νέου μη παραμετρικού διαγράμματος το οποίο και χαρακτηρίζεται στο πίνακα ως NEO και του διαγράμματος του Chakraborti (2004), το οποίο χαρακτηρίζεται ως $C - L - W$.

Παρατηρώντας τις τιμές του πίνακα 3.10 μπορούμε να πούμε πως το νέο διάγραμμα παρουσιάζει για όλες τις περιπτώσεις που μελετήθηκαν, καλύτερη απόδοση από το διάγραμμα $C - L - W$ ως προς τις τιμές ARL_{in}

Πίνακας 3.10: Σύγκριση των ARL_{in} των διαγραμμάτων ελέγχου με την ίδια τιμή FAR

FAR	m	n	C-L-W			NEO			
			(LCL, UCL)	Ακριβής FAR	ARL _{in}	(LCL, UCL)	r	Ακριβής FAR	ARL _{in}
0.01	50	5	(3, 48)	0.0072	635.7	(3,48)	1	0.0072	635.7
		11	(7, 44)	0.0093	642.2	(7,44)	1	0.0093	642.2
		25	(10, 41)	0.0061	10990	(6,45)	13	0.0097	3957820
	100	5	(7, 94)	0.00864	214.9	(7,94)	1	0.00864	214.9
		11	(15, 86)	0.00856	245	(6,95)	7	0.0091	518.8
		25	(23, 78)	0.00804	510.8	(8,93)	16	0.0082	4820
	500	5	(40, 461)	0.00954	114.5	(38,463)	2	0.0091	121
		11	(83, 418)	0.00988	113.3	(33,468)	7	0.00999	127.7
		25	(127, 374)	0.00946	128.3	(120,381)	7	0.0095	221
	1000	5	(82, 919)	0.00998	104.6	(79,922)	2	0.00997	105.2
		11	(167, 834)	0.00974	108.4	(96,905)	6	0.0098	113.2
		25	(258, 743)	0.00994	109.8	(243,758)	7	0.00994	115.6
0.005	50	5	(2, 49)	0.003	5671	(2,49)	1	0.003	5671
		11	(5, 46)	0.0025	9503	(5,46)	2	0.0025	9503.1
		25	(9, 42)	0.0031	44750	(7,44)	11	0.0048	9073150
	100	5	(5, 96)	0.00352	678.4	(5,96)	1	0.00352	678.4
		11	(13, 88)	0.0045	574.5	(5,96)	7	0.0046	1700.2
		25	(21, 80)	0.004	1488	(15,86)	11	0.00445	6223
	500	5	(31, 470)	0.00466	242.3	(20,481)	3	0.00477	254.1
		11	(72, 429)	0.0048	240.9	(56,445)	5	0.0049	260.1
		25	(118, 383)	0.00486	261	(90,411)	10	0.00499	319.8
	1000	5	(64, 937)	0.00492	215.1	(62,939)	2	0.0049	217.7
		11	(146, 855)	0.00488	219.8	(136,865)	4	0.0049	223.9
		25	(239, 762)	0.0049	227.5	(129,872)	13	0.00497	258.7
0.0027	50	5	(1, 50)	0.0008	∞	(1,50)	1	0.0008	∞
		11	(5, 46)	0.0025	9503	(5,46)	2	0.0025	9503.1
		25	(8, 43)	0.0015	173700	(7,44)	10	0.0021	29347500
	100	5	(4, 97)	0.00204	1550	(4,97)	2	0.00214	1508
		11	(11, 90)	0.00212	1630	(9,92)	5	0.0023	2540.4
		25	(19, 82)	0.00184	5183	(12,89)	12	0.0019	45381
	500	5	(25, 476)	0.00254	460.2	(16,485)	3	0.0026	503.8
		11	(64, 437)	0.00262	456.1	(64,437)	2	0.00263	456
		25	(110, 391)	0.00254	526.2	(101,400)	8	0.0026	599.1
	1000	5	(51, 950)	0.00258	419.5	(49,952)	2	0.0025	442.7
		11	(130, 871)	0.00262	219.8	(49,952)	7	0.00267	449.7
		25	(224, 777)	0.00266	430.2	(189,812)	9	0.00269	476.1

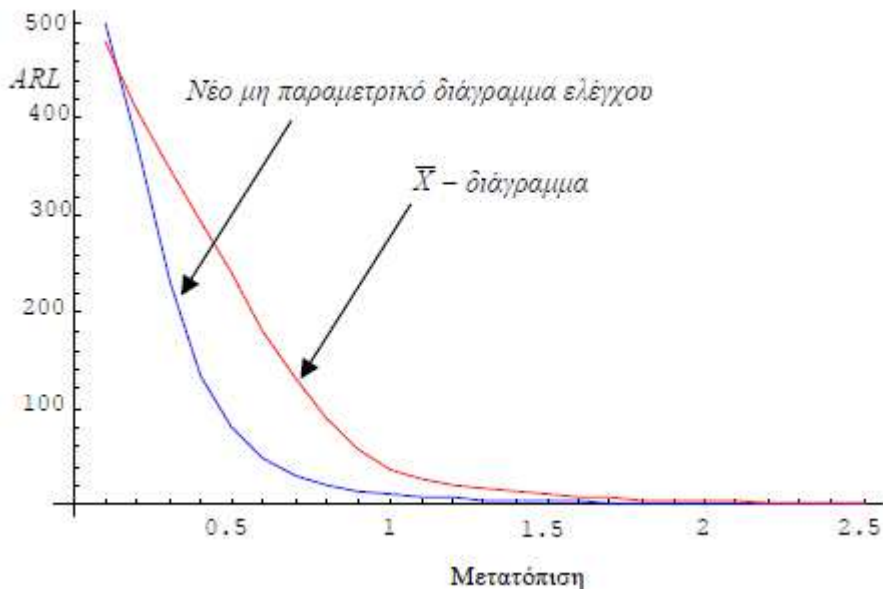
Μια ακόμα σύγκριση έχει πραγματοποιηθεί όσο αναφορά την απόδοση των διαγραμμάτων ελέγχου χρησιμοποιώντας την ίδια τιμή ARL_{in} και υπολογίζοντας τις εκτός ελέγχου τιμές ARL_{out} για συγκεκριμένες μετατοπίσεις. Θεωρούμε μια διεργασία, η οποία αρχικά βρίσκεται σε εντός ελέγχου κατάσταση (ακολουθώντας κανονική κατανομή $N(0,1)$), ενώ στη συνέχεια μετατοπίζεται στην κατανομή $N(0, \theta)$. Έτσι λοιπόν στον Πίνακα 3.10, παρουσιάζονται οι τιμές ARL_{out} για το διάγραμμα $C - L - W$, το προτεινόμενο διάγραμμα, και το \bar{X} - διάγραμμα τύπου Shewhart για $ARL_{in} = 370, 500, m = 200, 500, n = 5, 11, 25 (j = (n + 1) / 2)$ και $\theta = 1.1, 1.2, 1.3$. Οι παράμετροι σχεδιασμού $LCL = X_{\alpha:m}, UCL = X_{m-\alpha+1:m}, r$ έχουν καθορισθεί έτσι ώστε οι τιμές του ARL_{in} να είναι όσο το δυνατόν πιο κοντά στην τιμή στόχο. Στη συνέχεια, η εκτός ελέγχου ARL_{out} τιμή υπολογίζεται αριθμητικά με τη βοήθεια της ακόλουθης σχέσης

$$ARL = \int_0^1 \int_0^t \frac{1}{1 - q(G(\Phi^{-1}(s)), G(\Phi^{-1}(t)); r)} f(s, t) ds dt$$

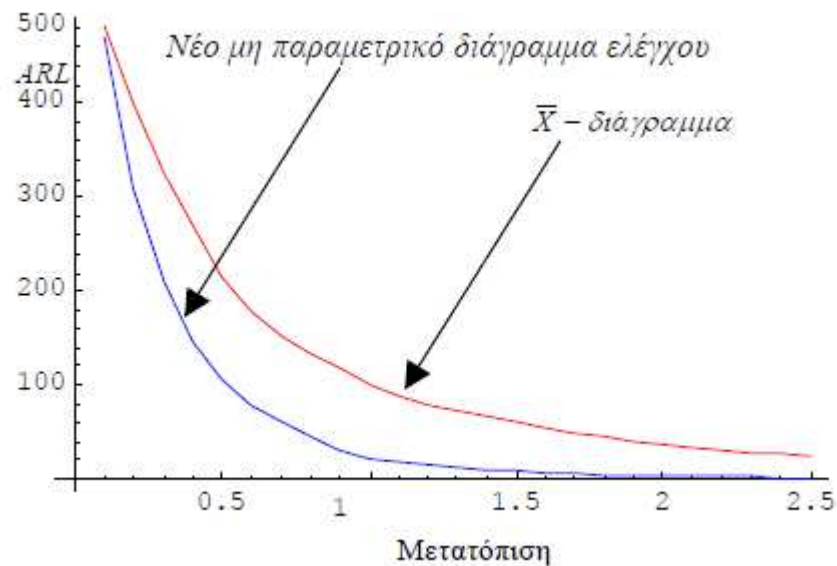
Είναι σημαντικό να αναφερθεί ότι, αν και η μετατόπιση της κατανομής που πραγματοποιείται αφορά μόνο τη διακύμανση της κατανομής, το προτεινόμενο μη παραμετρικό διάγραμμα ελέγχου είναι αποτελεσματικό στην ανίχνευσή της, πράγμα που οφείλεται στην ενσωμάτωση της πρόσθετης συνθήκης να βρίσκεται ανάμεσα στα όρια ελέγχου επαρκές πλήθος των παρατηρήσεων του εξεταζόμενου δείγματος.

Τέλος συγκρίνοντας γραφικά τα εκτός ελέγχου μέσα μήκη ροής του \bar{X} - διαγράμματος τύπου Shewhart και του νέου μη παραμετρικού διαγράμματος, όπως αυτό ορίζεται υπολογίζουμε, υπό δύο διαφορετικές εντός ελέγχου κατανομές (Student t με 4 βαθμούς ελευθερίας και Εκθετική) τα ARL_{out} για διαφορετικές μετατοπίσεις της τυπικής απόκλισης. Οι παράμετροι σχεδιασμού επιλέγονται έτσι ώστε το ARL_{in} να είναι σχεδόν ίσο με 500, ενώ για την εκθετική κατανομή, χρησιμοποιήσαμε ένα μονόπλευρο διάγραμμα, αφού είναι δυνατές μόνο θετικές μετατοπίσεις. Τα συγκεκριμένα αποτελέσματα παρουσιάζονται στα σχήματα, όπου είναι σαφές ότι, σε περιπτώσεις κατανομών με βαριές ουρές (όπως η Student t) ή με έντονη λοξότητα (όπως η Εκθετική), το νέο διάγραμμα είναι πιο αποτελεσματικό από το \bar{X} - διάγραμμα τύπου Shewhart .

Σχήμα 1: Εκτός ελέγχου ARL για την εκθετική κατανομή



Σχήμα 1: Εκτός ελέγχου *ARL* για την κατανομή Student με 4 βαθμούς ελευθερίας



3.4 Μη παραμετρικά διαγράμματα ελέγχου με χρήση ρών και βαθμολογικών συναρτήσεων τύπου Wilcoxon

Σε αυτή την παράγραφο παρουσιάζεται ένα διάγραμμα ελέγχου με χρήση ρών και βαθμολογικών συναρτήσεων τύπου Wilcoxon. Ας υποθέσουμε ότι X_1, X_2, \dots, X_m είναι ένα τυχαίο δείγμα μεγέθους m που προέρχεται από μια συνεχή κατανομή με αθροιστική συνάρτηση $F_x(x)$ και Y_1, Y_2, \dots, Y_n ένα δείγμα μεγέθους n από μια άλλη κατανομή $F_y(x)$. Ένα πρόβλημα που παρουσιάζει ιδιαίτερο ενδιαφέρον είναι ο έλεγχος της μηδενικής υπόθεσης $H_0: F_x(x) = F_y(x)$. Στα πλαίσια της Θεωρίας Αξιοπιστίας και του Ελέγχου Ποιότητας, ο έλεγχος της παραπάνω υπόθεσης πραγματοποιείται συχνά έναντι της εναλλακτικής $H_1: \bar{F}_x(x) > \bar{F}_y(x)$ δηλαδή έναντι της υπόθεσης ότι η X είναι στοχαστικά μεγαλύτερη από την Y (ή της μονόπλευρης εναλλακτικής με αντίστροφη φορά $H_2: \bar{F}_x(x) < \bar{F}_y(x)$). Πρακτικά, η συγκεκριμένη εναλλακτική υπόθεση δηλώνει ότι η X - διεργασία παράγει πιο αξιόπιστα προϊόντα (δηλαδή που έχουν μεγαλύτερο χρόνο ζωής) σε σχέση με την Y - διεργασία. Για έναν τέτοιο έλεγχο μπορεί να χρησιμοποιηθεί η αθροιστική βαθμολογική συνάρτηση τύπου Wilcoxon (Wilcoxon's rank-sum statistic), η οποία είχε αρχικά προταθεί στην εργασία του Wilcoxon (1945). Για την κατασκευή της συνάρτησης αυτής, ενώνουμε τα δύο δείγματα και στη συνέχεια βαθμολογούμε τις παρατηρήσεις στο από κοινού δείγμα από 1 μέχρι $(m + n)$. Τότε, η αθροιστική βαθμολογική στατιστική συνάρτηση τύπου Wilcoxon δίνεται από τον τύπο

$$W_R = \sum_{i=1}^m Rank(X_i)$$

και εκφράζει το άθροισμα των βαθμών των X - παρατηρήσεων στο ενιαίο δείγμα. δεδομένου ότι υπό την εναλλακτική υπόθεση πως η X είναι στοχαστικά μεγαλύτερη από την Y , αναμένουμε ότι το X - δείγμα θα λάβει κυρίως τους μεγάλους βαθμούς, έπεται ότι θα έχουμε ένδειξη για απόρριψη της μηδενικής υπόθεσης H_0 έναντι της

εναλλακτικής H_1 , αν $W_R \geq w$, όπου w είναι μια κατάλληλα επιλεγμένη κρίσιμη τιμή. Προφανώς, αν η εναλλακτική υπόθεση ήταν αμφίπλευρη ($H_1: F_x(x) \neq F_y(x)$) θα είχαμε ένδειξη απόρριψης της H_0 αν ίσχυε $W_R \geq w_2$ ή $W_R \leq w_1$ όπου $w_1 < w_2$ δύο κατάλληλα επιλεγμένες σταθερές.

Οι Ng και Balakrishnan (2004) πρότειναν παραλλαγές της παραπάνω βαθμολογικής συνάρτησης για τον έλεγχο ισονομίας δύο δειγμάτων, κατασκευάζοντας κατάλληλους στατιστικούς ελέγχους προτεραιότητας,

ενώ ο Balakrishnan (2008) μαζί με τους συνεργάτες του τη χρησιμοποίησαν για το σχεδιασμό μη- παραμετρικών διαγραμμάτων ελέγχου. Υπό τη μηδενική υπόθεση H_0 , έχουμε ότι

$$P(\text{Rank}(X_i) = i_1, \dots, \text{Rank}(X_m) = i_m | H_0: F_x = F_y) = \frac{n!}{(n+m)!}$$

για κάθε υποσύνολο $\{i_1, \dots, i_m\}$ του συνόλου $\{1, 2, \dots, m+n\}$. Συνεπώς, προκύπτει ότι

$$P(W_R = w | H_0: F_x = F_y) = \frac{N_w}{\binom{m+n}{m}}$$

όπου N_w είναι το πλήθος των υποσυνόλων $\{i_1, \dots, i_m\}$ με $i_1 + \dots + i_m = w$.

3.5 Νέα μη-παραμετρικά διαγράμματα ελέγχου

Στην παράγραφο αυτή παρουσιάζονται τρία νέα ελεύθερα κατανομής διαγράμματα ελέγχου τύπου Shewhart τα οποία προτάθηκαν από τους Balakrishnan, N., Triantafyllou, I.S. & Koutras, M.V. (2010) και χρησιμοποιούν συναρτήσεις ρών ή βαθμολογικές συναρτήσεις για να οδηγηθούμε σε απόφαση κατά το πόσο μια διεργασία παραμένει εντός ελέγχου ή έχει μετατοπιστεί σε εκτός ελέγχου κατάσταση. Η κατανομή των στατιστικών συναρτήσεων που απεικονίζονται στα νέα μη-παραμετρικά διαγράμματα ελέγχου προσδιορίζεται με τη βοήθεια των ποσοτήτων M_i . Το πρώτο διάγραμμα ρών χρησιμοποιεί το μέγιστο μήκος ροής των παρατηρήσεων του εξεταζόμενου δείγματος στο από κοινού δείγμα, ενώ το δεύτερο προτεινόμενο διάγραμμα λαμβάνει υπόψη το πλήθος των ρών των παρατηρήσεων του εξεταζόμενου δείγματος, των οποίων το μήκος υπερβαίνει μια προκαθορισμένη τιμή. Τέλος, το τρίτο διάγραμμα στηρίζεται στο άθροισμα των βαθμών των παρατηρήσεων του εξεταζόμενου δείγματος, οι οποίες βρίσκονται ανάμεσα στα όρια ελέγχου. Όπως έχει ήδη αναφερθεί, τα όρια ελέγχου των νέων μη παραμετρικών διαγραμμάτων που παρουσιάζονται του παρόντος κειμένου, προσδιορίζονται με τη βοήθεια ενός δείγματος αναφοράς που συλλέγεται από τη διεργασία, όταν βρίσκεται σε εντός ελέγχου κατάσταση. Έστω ότι X_1, X_2, \dots, X_m είναι ένα τυχαίο δείγμα μεγέθους m που προέρχεται από την εντός ελέγχου (αθροιστική) κατανομή $F_x(x) = F(x)$ και ας υποθέσουμε ότι χρησιμοποιούνται ως όρια ελέγχου του διαγράμματος δύο συγκεκριμένες διατεταγμένες παρατηρήσεις $X_{a:m}$ και $X_{b:m}$, δηλαδή

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

Οι παράμετροι a, b είναι παράμετροι σχεδιασμού των προτεινόμενων διαγραμμάτων και ο προσδιορισμός τους επιτυγχάνεται μέσω δύο διαδικασιών. Η πρώτη διαδικασία απαιτεί την επίτευξη μιας συγκεκριμένης πιθανότητας

λανθασμένου συναγερμού, ενώ η δεύτερη στηρίζεται σε μια προκαθορισμένη τιμή του εντός ελέγχου μέσου μήκους ροής ARL_{in} , συνήθως ίση με 370 ή 500. Αξίζει να σημειωθεί ότι το ARL_{in} των τριών διαγραμμάτων ελέγχου παραμένει ίδιο για όλες τις (εντός ελέγχου) συνεχείς κατανομές. Η κατασκευή των νέων διαγραμμάτων ελέγχου βασίζεται στην έννοια της ροής (run) ίδιων συμβόλων (ή αποτελεσμάτων). Πιο συγκεκριμένα, ως ροή αποτελεσμάτων ίδιου τύπου ορίζουμε την εμφάνιση διαδοχικών (χωρίς διακοπή) όμοιων αποτελεσμάτων τα οποία ακολουθούνται και έπονται από διαφορετικά αποτελέσματα. Το πλήθος k των όμοιων στοιχείων μιας ροής ονομάζεται μήκος ροής (k θετικός ακέραιος). Για παράδειγμα, στην ακολουθία δύο (διαφορετικών) συμβόλων X, Y , που φαίνεται παρακάτω

$X X Y Y Y X Y Y X X X X$,

έχουμε διαδοχικά μία ροή των X μήκους 2, ακολουθούμενη από μία ροή των Y μήκους 3, μία ροή των X μήκους 1, μία ροή των Y μήκους 2 και τέλος μία ροή των X μήκους 4.

Ας υποθέσουμε ότι, μετά τον καθορισμό των ορίων ελέγχου LCL, UCL , συλλέγονται ανεξάρτητα τυχαία δείγματα από τη διεργασία και θέλουμε να εξετάσουμε αν παραμένει εντός ελέγχου ή αν έχει μετατοπιστεί σε εκτός ελέγχου κατάσταση. Αν Y_1, Y_2, \dots, Y_n είναι ένα εξεταζόμενο δείγμα μεγέθους n και $F_Y(x) = G(x)$ η αντίστοιχη αθροιστική συνάρτηση κατανομής του, ο στόχος μας είναι να ανιχνεύσουμε μια πιθανή μετατόπιση της εντός ελέγχου κατανομής από $F(x)$ σε $G(x)$, δηλαδή να ελέγξουμε τη μηδενική υπόθεση $H_0: F(x) = G(x)$ έναντι της δίπλευρης εναλλακτικής $H_1: F(x) \neq G(x)$. Οι στατιστικές συναρτήσεις που απεικονίζονται στα προτεινόμενα διαγράμματα ελέγχου ορίζονται με τη βοήθεια

- a. των ροών των Y -παρατηρήσεων που βρίσκονται ανάμεσα στα όρια ελέγχου LCL, UCL
- b. των βαθμών των Παρατηρήσεως που βρίσκονται ανάμεσα στα όρια ελέγχου LCL, UCL .

Η μεθοδολογία για τα τρία διαγράμματα ελέγχου μπορεί να συνοψιστεί ως εξής: υπό τη μηδενική υπόθεση $H_0: F(x) = G(x)$ (δηλαδή αν, τόσο το δείγμα αναφοράς, όσο και τα εξεταζόμενα τυχαία δείγματα που συλλέγονται από τη διεργασία, προέρχονται από την ίδια κατανομή), το πλήθος των Y -παρατηρήσεων που βρίσκονται μεταξύ δύο διαδοχικών X -παρατηρήσεων δεν θα πρέπει να λαμβάνει ακραίες τιμές, όπου οι ακραίες τιμές ορίζονται βάσει της αναλογίας n/m . Συνεπώς, δύο πιθανές στατιστικές συναρτήσεις που μπορεί να χρησιμοποιηθούν για να αποφασίζουμε αν η διεργασία παραμένει εντός ελέγχου ή όχι, έχουν ως εξής:

- το μέγιστο μήκος ροής των Y -παρατηρήσεων που παρατηρούνται μεταξύ των ορίων ελέγχου,
- το πλήθος των ροών των Y -παρατηρήσεων (μεταξύ των ορίων ελέγχου), των οποίων το μήκος υπερβαίνει μια προκαθορισμένη τιμή k .
- του αθροίσματος των βαθμών (στο από κοινού δείγμα των X και Y παρατηρήσεων) των Y -παρατηρήσεων που βρίσκονται ανάμεσα στα όρια ελέγχου.

Αξίζει να σημειωθεί ότι η τελευταία επιλογή αντιστοιχεί στη γνωστή αθροιστική βαθμολογική συνάρτηση τύπου Wilcoxon.

Όπως αναφέρθηκε και προηγουμένως, οι τρεις στατιστικές συναρτήσεις, είναι δυνατόν να εκφραστούν με τη βοήθεια των στατιστικών συναρτήσεων προτεραιότητας. Πιο συγκεκριμένα, έστω ότι $M_i, i = 1, 2, \dots, m$ εκφράζει το πλήθος των Y -παρατηρήσεων του εξεταζόμενου δείγματος που βρίσκονται μεταξύ της $(i - 1)$ -οστής και i -οστής διατεταγμένης παρατήρησης του X -δείγματος (σύμβαση: $X_{(0)} = -\infty$). Με άλλα λόγια, τα M_i είναι τα μήκη των ρών των Y -παρατηρήσεων μεταξύ διαδοχικών X -παρατηρήσεων. Οι τρεις προαναφερθείσες στατιστικές συναρτήσεις δίνονται από τις σχέσεις

$$R = \max(M_{a+1}, M_{a+2}, \dots, M_b), \quad N_k = |\{M_i: a + 1 \leq i \leq b \text{ και } M_i \geq k\}|, \quad W = \sum_{i=a+1}^b W_i$$

όπου k είναι μια πρόσθετη παράμετρος σχεδιασμού που λαμβάνει ακέραιες τιμές, το W_i εκφράζει το άθροισμα των βαθμών των Y -παρατηρήσεων που βρίσκονται μεταξύ των $X_{(i-1)}$ και X_i . Θέλοντας να εκφράσουμε την αθροιστική βαθμολογική συνάρτηση τύπου Wilcoxon με τη βοήθεια των M_i , μπορούμε να πάρουμε την σχέση

$$W = \frac{1}{2} \left(\sum_{i=a+1}^b M_i \right)^2 + \sum_{i=a+1}^b i M_i + \left(M_0 + a - \frac{3}{2} \right) \sum_{i=a+1}^b M_i$$

Όταν χρησιμοποιείται η στατιστική συνάρτηση R , η διεργασία κρίνεται ότι είναι εντός ελέγχου, αν επαληθεύονται οι δύο ακόλουθες συνθήκες

$$R \leq r \text{ και } M_0 \leq r_0,$$

όπου $M_0 = \sum_{i=1}^a M_i$ εκφράζει το πλήθος των παρατηρήσεων του Y -δείγματος πριν το LCL καθώς και r_0, r είναι παράμετροι σχεδιασμού του διαγράμματος.

Όταν εφαρμόζεται η στατιστική συνάρτηση N_k , η διεργασία κρίνεται ότι είναι εντός ελέγχου, αν επαληθεύονται οι δύο ακόλουθες συνθήκες

$$N_k \leq r_1 \text{ και } M_0 \leq r_0$$

όπου r_1, r_0 αποτελούν παραμέτρους σχεδιασμού.

Τέλος, αν στο διάγραμμα απεικονίζεται η στατιστική συνάρτηση W , τότε δεν θα παράγεται σήμα για εκτός ελέγχου μετατόπιση, εφόσον επαληθεύονται οι ακόλουθες δύο συνθήκες

$$W \leq w \text{ και } M_0 \leq r_0$$

όπου w, r_0 είναι παράμετροι σχεδιασμού του διαγράμματος.

Μπορούμε να αναφέρουμε πως η μέγιστη δυνατή τιμή της W προκύπτει όταν όλες οι Y -παρατηρήσεις βρίσκονται στο διάστημα $(X_{b-1:m}, X_{b:m})$. Στη συγκεκριμένη περίπτωση, οι βαθμοί τους είναι ίσοι με $(b - 1) + j, j = 1, 2, \dots, n$ και η αθροιστική βαθμολογική συνάρτηση θα πάρει την τιμή

$$n(b-1) + \sum_{j=1}^n j = \frac{n(n+2b-1)}{2}$$

Αντίστοιχα, η μικρότερη δυνατή τιμή της W είναι ίση με μηδέν (η οποία επιτυγχάνεται αν καμία Y -παρατήρηση δεν βρίσκεται μεταξύ των LCL και UCL) και συνεπώς το στήριγμα της συνάρτησης W είναι

$$R_W = \left\{ 0, 1, \dots, \frac{n(n+2b-1)}{2} \right\}$$

Στη συνέχεια θα παραθέσουμε την πιθανότητα λανθασμένου συναγερμού των νέων διαγραμμάτων ελέγχου, δηλαδή την πιθανότητα να παραχθεί σήμα για εκτός ελέγχου μετατόπιση της διεργασίας, ενώ στην πραγματικότητα η διεργασία παραμένει εντός ελέγχου. Θα μελετήσουμε την εντός ελέγχου κατανομή των στατιστικών συναρτήσεων που απεικονίζονται στα τρία διαγράμματα ελέγχου. Έστω ότι X_1, X_2, \dots, X_n είναι ένα δείγμα αναφοράς μεγέθους m που έχει συλλεχθεί από την εντός ελέγχου κατανομή $F_X(x) = F(x)$ και Y_1, Y_2, \dots, Y_n ένα τυχαίο δείγμα που είναι διαθέσιμο από την ίδια κατανομή, δηλαδή $F_X(x) = F_Y(x)$ για κάθε x . Οι ακριβείς κατανομές των στατιστικών συναρτήσεων R και N_k μπορούν να μελετηθούν με τη βοήθεια της από κοινού κατανομής των $(M_{a+1}, M_{a+2}, \dots, M_b)$ ενώ η συνάρτηση W μπορεί να αναλυθεί με χρήση της από κοινού κατανομής των $(M_0, M_{a+1}, M_{a+2}, \dots, M_b)$. Οι Balakrishnan και Frattina (2000) και οι Balakrishnan και Ng (2001), πρότειναν, για την εφαρμογή ενός ελέγχου προτεραιότητας τη σύγκριση δύο κατανομών F_X, F_Y , το μέγιστο στατιστικό έλεγχο προτεραιότητας (maximal precedence test). Η στατιστική συνάρτηση M που χρησιμοποιείται στον παραπάνω έλεγχο, ορίζεται ως το μέγιστο πλήθος των Y -παρατηρήσεων που συμβαίνουν πριν την πρώτη, μεταξύ πρώτης και δεύτερης, ..., μεταξύ της $(r-1)$ -οστής και r -οστής διατεταγμένης παρατήρησης του X -δείγματος. Συνεπώς, ισχύει ότι $M = \max(M_1, M_2, \dots, M_r)$ και η κατανομή της, υπό τη μηδενική υπόθεση $H_0: F_X = F_Y$, μπορεί να δοθεί με τη βοήθεια της από κοινού κατανομής των (M_1, M_2, \dots, M_r) . Έτσι λοιπόν καταλήγουμε στις παρακάτω εκφράσεις για τη πιθανότητα λανθασμένου συναγερμού. Όταν χρησιμοποιείται η στατιστική συνάρτηση R τότε η πιθανότητα λανθασμένου συναγερμού (FAR) είναι

$$FAR = 1 - P(R \leq r \text{ και } M_0 \leq r_0)$$

Όταν χρησιμοποιείται η στατιστική συνάρτηση N_k , τότε η πιθανότητα λανθασμένου συναγερμού (FAR) είναι

$$FAR = 1 - P(N_k \leq r_1 \text{ και } M_0 \leq r_0)$$

Όταν χρησιμοποιείται η στατιστική συνάρτηση W τότε η πιθανότητα λανθασμένου συναγερμού (FAR) είναι

$$FAR = 1 - P(W \leq w \text{ και } M_0 \leq r_0)$$

Ενδιαφέρον παρουσιάζει η μελέτη των εκτός ελέγχου κατανομών των τριών στατιστικών συναρτήσεων που ορίστηκαν προηγουμένως. Ας υποθέσουμε λοιπόν, ότι οι παρατηρήσεις Y_1, Y_2, \dots, Y_n ενός τυχαίου δείγματος προέρχονται

από μια συνεχή κατανομή $F_Y(x) = G(x)$. Επίσης εκφράζουμε την πιθανότητα ότι ο κανόνας δεν παράγει σήμα, όταν ισχύει η σχέση

$$p(m, n, a, b; c; F, G) = \sum_{(m_0, m_{a+1}, \dots, m_b) \in A} P(M_0 = m_0 \text{ και } M_j = m_j \text{ για } a + 1 \leq j \leq b)$$

όπου c είναι ένα διάνυσμα σταθερών που πρέπει να προσδιορισθούν ώστε ο κανόνας να ικανοποιεί ορισμένες απαιτήσεις (π.χ. μια προκαθορισμένη πιθανότητα λανθασμένου συναγερμού ή ένα δεδομένο εντός ελέγχου μέσο μήκος ροής) και A είναι το σύνολο τιμών για το διάνυσμα $(M_0, M_{a+1}, M_{a+2}, \dots, M_b)$ ώστε το διάγραμμα ελέγχου να μην παράγει σήμα, προφανώς το A εξαρτάται από τις σταθερές c . Είναι σαφές ότι η πιθανότητα λανθασμένου συναγερμού για την ειδική περίπτωση $G = F$, δίνεται από τη σχέση

$$FAR = 1 - p(m, n, a, b; c; F, F).$$

Αν θεωρήσουμε τώρα τη γενική περίπτωση που οι X – και Y – παρατηρήσεις δεν ακολουθούν την ίδια κατανομή, δηλαδή $G(x) \neq F(x)$. Ας συμβολίσουμε με Y το τυχαίο διάνυσμα Y_1, Y_2, \dots, Y_n και με $X_{\alpha:b}$ το διατεταγμένο X – δείγμα που περιλαμβάνει τις X – παρατηρήσεις από την a – οστή έως την b – οστή διατεταγμένη παρατήρηση, δηλαδή $X_{\alpha:b} = (X_{\alpha:m}, X_{\alpha+1:m}, \dots, X_{b:m})$. Τότε, οι ποσότητες $(M_0, M_{a+1}, M_{a+2}, \dots, M_b)$ είναι συναρτήσεις των Y και $X_{\alpha:b}$ και επομένως οι προσθετοί της παραπάνω σχέσης παίρνουν τη μορφή

$$p_m = P(M_0(Y: X_{\alpha:b}) = m_0 \text{ και } M_j(Y: X_{\alpha:b}) = m_j \text{ για } a + 1 \leq j \leq b)$$

όπου $m = (m_0, m_{a+1}, \dots, m_b)$ το διάνυσμα των δεικτών που αθροίζονται. Επιπρόσθετα, αν η διεργασία μετατοπιστεί σε εκτός ελέγχου κατάσταση, η πιθανότητα το διάγραμμα να μην παράγει σήμα για εκτός ελέγχου μετατόπιση, εξαρτάται τόσο από την εντός ελέγχου, όσο και από την εκτός ελέγχου κατανομή $F(x)$ και $G(x)$ αντίστοιχα. Δεδομένου όμως ότι το τελικό αποτέλεσμα εξαρτάται από τις F και G , μόνο μέσω της συνάρτησης $G \circ F^{-1}$, θα μπορούσαμε εύκολα να λάβουμε ακριβείς εκφράσεις για την εκτός ελέγχου πιθανότητα συναγερμού στην περίπτωση που έχουμε εναλλακτικές τύπου Lehmann. Η από κοινού συνάρτηση πιθανότητας των $(M_0, M_{a+1}, M_{a+2}, \dots, M_b)$ υπό την εναλλακτική υπόθεση τύπου Lehmann $G(x) = (F(x))^Y$ δίνεται στην ακόλουθη σχέση

$$P(M_0 = m_0 \text{ και } M_j = m_j \text{ για } a + 1 \leq j \leq b) = c_1 c_2 \gamma^{-(b-a+1)} \times \\ \prod_{j=a}^{b-1} B\left(\frac{j}{\gamma} + m_0 + \sum_{i=a+1}^j m_i, m_{j+1} + 1\right) \times \\ \sum_{l=0}^{m-b} (-1)^l \binom{m-b}{j} B\left(\frac{b+l}{\gamma} + m_0 + \sum_{i=a+1}^b m_i, n - m_0 - \sum_{i=a+1}^b m_i + 1\right)$$

$$\text{με } C_1 = \frac{m!}{(a-1)!(m-b)!}, \quad C_2 = \frac{n!}{m_0! (\prod_{j=a+1}^b m_j!) (n - m_0 - \sum_{i=a+1}^b m_j)!}$$

Στα νέα μη παραμετρικά διαγράμματα ελέγχου τύπου Shewhart που μελετώνται στο η κατανομή του μήκους ροής τους δεν είναι γεωμετρική, διότι τα συμβάντα παραγωγής σήματος για εκτός ελέγχου μετατόπιση της διεργασίας, δεν είναι ανεξάρτητα. Ωστόσο, κάνοντας χρήση της τεχνικής δέσμευσης του Chakraborti (2000), μπορούμε να καταλήξουμε σε μια ακριβή μορφή της κατανομής του μήκους ροής. Η εντός ελέγχου κατανομή και το αντίστοιχο μέσο μήκος ροής ARL_{in} προκύπτουν θέτοντας $G = F$ και δίνεται από τη σχέση

$$ARL_{in} = \iiint \dots \int_{0 \leq u_a \leq u_{a+1} \leq \dots \leq u_b \leq 1} \frac{1}{1 - Q(u_a, u_{a+1}, \dots, u_b)} f_{a:b}(u_a, u_{a+1}, \dots, u_b) du_a du_{a+1} \dots du_b$$

Το πολλαπλό ολοκλήρωμα μπορεί να υπολογισθεί αριθμητικά για κάθε επιλογή των παραμέτρων σχεδιασμού (δηλαδή των a, b και των παραμέτρων που προσδιορίζουν το σύνολο $A \subseteq \{0, 1, 2, \dots\}^{b-a}$). Πρόσθετα, μπορούμε να κάνουμε το ίδιο σχόλιο για το εκτός ελέγχου μέσο μήκος ροής ARL_{out} , υπό την προϋπόθεση ότι έχουν επιλεγεί συγκεκριμένες εντός και εκτός ελέγχου κατανομές F και G αντίστοιχα. Αξίζει να αναφερθεί ότι, όταν η εκτός ελέγχου κατανομή είναι τύπου Lehmann, δηλαδή είναι της μορφής $G(x) = (F(x))^\gamma$, δεν χρειαζόμαστε τη συναρτησιακή μορφή της $F(x)$, καθώς $GF^{-1}(x) = x^\gamma$ και συνεπώς

$$ARL_{out} = \iiint \dots \int_{0 \leq u_a \leq u_{a+1} \leq \dots \leq u_b \leq 1} \frac{1}{1 - Q(u_a^\gamma, u_{a+1}^\gamma, \dots, u_b^\gamma)} f_{a:b}(u_a, u_{a+1}, \dots, u_b) du_a du_{a+1} \dots du_b$$

3.5.1 Αριθμητικά αποτελέσματα των νέων διαγραμμάτων ελέγχου

Σε αυτή τη παράγραφο θα αναφερθούμε στην απόδοση ενός διαγράμματος ελέγχου που κρίνεται συνήθως με τη βοήθεια της κατανομής του μήκους ροής και την ικανότητα ανίχνευσης μιας πιθανής μετατόπισης της διεργασίας ώστε να αποδεικνύεται η αποτελεσματικότητα και η αντοχή του διαγράμματος, τόσο σε εντός όσο και σε εκτός ελέγχου. Έτσι λοιπόν σε αυτή τη παράγραφο υλοποιούνται αριθμητικοί υπολογισμοί για τα προτεινόμενα διαγράμματα που αποδεικνύουν την αποτελεσματικότητα και την αντοχή τους, τόσο σε εντός όσο και σε εκτός ελέγχου καταστάσεις. Στον παρακάτω πίνακα 3.10, παρουσιάζουμε τις πιθανότητες συναγερμού (alarm rates, AR) των τριών διαγραμμάτων ελέγχου για διαφορετικές επιλογές των m, n . Οι παράμετροι σχεδιασμού a, b, r_0, r, k, w καθορίζονται κατάλληλα ώστε οι κανόνες να επιτυγχάνουν μια συγκεκριμένη πιθανότητα λανθασμένου συναγερμού (false alarm rate) 5% ($FAR = 0.005$) και στη συνέχεια, υπολογίζονται οι τιμές του AR για $\gamma = 1/2, 1/4, 1/8$.

Πίνακας 3.11: Σύγκριση των AR των τριών διαγραμμάτων ελέγχου με $FAR=0.005$ και ($\gamma = 1/2, 1/4$ και $1/8$)

m	n	R -διάγραμμα				N_2 -διάγραμμα				W -διάγραμμα						
		(LCL,UCL)	r_0	r	Ακριβής FAR	AR	(LCL,UCL)	r_0	r_1	Ακριβής FAR	AR	(LCL,UCL)	w	r_0	Ακριβής FAR	AR
100	5	(7, 10)	2	2	0.0043	0.1201 0.5130 0.8486	(7, 10)	2	1	0.0041	0.1193 0.5123 0.8483	(7, 10)	70	2	0.0041	0.1193 0.5123 0.8483
	11	(10, 13)	4	3	0.0049	0.2461 0.8279 0.9901	(10, 13)	4	1	0.0047	0.2457 0.8277 0.9901	(10, 13)	95	4	0.0048	0.2475 0.8288 0.9902
	25	(9, 12)	7	4	0.0045	0.4668 0.9872 0.9999	(9, 12)	7	1	0.0050	0.4680 0.9872 0.9999	(9, 12)	123	7	0.0049	0.4753 0.9879 0.9999
200	5	(16, 19)	2	2	0.0050	0.1409 0.5526 0.8689	(16, 19)	2	1	0.0050	0.1408 0.5526 0.8689	(15, 18)	65	2	0.0046	0.1326 0.5399 0.8631
	11	(21, 24)	4	3	0.0045	0.2647 0.8488 0.9922	(21, 24)	4	2	0.0044	0.2646 0.8488 0.9922	(21, 24)	135	4	0.0048	0.2669 0.8497 0.9923
	25	(16, 19)	6	5	0.0049	0.5678 0.9952 0.9999	(16, 19)	6	2	0.0049	0.5677 0.9952 0.9999	(16, 19)	190	6	0.0050	0.5688 0.9953 0.9999
500	5	(39, 42)	2	1	0.0047	0.1374 0.5510 0.8690	(39, 42)	2	2	0.0044	0.1367 0.5504 0.8689	(39, 42)	100	2	0.0049	0.1381 0.5515 0.8692
	11	(33, 36)	3	2	0.0048	0.3042 0.8911 0.9965	(33, 36)	3	2	0.0048	0.3041 0.8911 0.9965	(33, 36)	3	2	0.0049	0.3045 0.8912 0.9965
	25	(21, 24)	4	4	0.0045	0.5907 0.9971 0.9999	(21, 24)	4	2	0.0045	0.5907 0.9971 0.9999	(21, 24)	150	4	0.0046	0.5915 0.9971 0.9999

Από τον παραπάνω πίνακα, φαίνεται ότι τα τρία μη παραμετρικά διαγράμματα ελέγχου που μελετώνται, παρουσιάζουν παρόμοια απόδοση, αφού, για την ίδια σχεδόν τιμή FAR , πετυχαίνουν σχεδόν ίσες AR τιμές. Για το λόγο αυτό, οι υπολογισμοί που ακολουθούν, θα γίνουν μόνο σε ένα από αυτά τα διαγράμματα και συγκεκριμένα στο W – διάγραμμα, μιας και έχουμε ισχυρές ενδείξεις ότι τα άλλα δύο διαγράμματα, με κατάλληλες επιλογές των παραμέτρων σχεδιασμού, θα δώσουν ισοδύναμα αποτελέσματα. Στον παρακάτω πίνακα 3.12 παρουσιάζουμε τις πιθανότητες FAR του W – διαγράμματος για διάφορους σχεδιασμούς και επιλογές των παραμέτρων m, n, a, b, w, r_0 . Παρατηρώντας τις τιμές του πίνακα 3.12 μπορούμε να κατασκευάσουμε διαγράμματα ελέγχου ελεύθερα κατανομής που πετυχαίνουν μια προκαθορισμένη πιθανότητα λανθασμένου συναγερμού f . Η χρήση έξι παραμέτρων m, n, a, b, w, r_0 στο διάγραμμα προσφέρει την ευελιξία να σταθεροποιήσουμε ορισμένες από αυτές και στη συνέχεια να αναζητήσουμε τη βέλτιστη επιλογή για τις άλλες, ή εναλλακτικά για έναν αποδεκτό συνδυασμό τους που ικανοποιεί ορισμένες απαιτήσεις. Παραδείγματος χάριν, αν έχουμε συλλέξει δείγμα αναφοράς μεγέθους $m = 200$ και επιθυμούμε να εξετάσουμε τυχαία δείγματα μεγέθους $n = 25$ που εξάγονται από τη διεργασία, τότε η πιθανότητα λανθασμένου συναγερμού σχεδόν ίση με $f = 0.05$ μπορεί να επιτευχθεί χρησιμοποιώντας

- την 19^η και 22^η μικρότερη παρατήρηση στο εξεταζόμενο δείγμα ($a = 19, b = 22$), $w = 81$ και $r_0 = 8$ και με $FAR = 0.0464$ ή
- την 30^η και 33^η μικρότερη παρατήρηση στο εξεταζόμενο δείγμα ($a = 30, b = 33$), $w = 127$ και $r_0 = 14$ και με $FAR = 0.0470$ ή
- την 33^η και 36^η μικρότερη παρατήρηση στο εξεταζόμενο δείγμα ($a = 33, b = 36$), $w = 140$ και $r_0 = 16$ και με $FAR = 0.0468$.

Πίνακας 3.12: Πιθανότητες λανθασμένου συναγερμού για δεδομένο σχεδιασμό

		Δείγμα αναφοράς μεγέθους m												
n	r_0	40			60			100			200			
		(a,b)	w	FAR	(a,b)	w	FAR	(a,b)	w	FAR	(a,b)	w	FAR	
5	2	(7,11)	200	0.0471	(10,13)	45	0.0466	(15,19)	150	0.0292	(20,24)	80	0.0140	
			40	0.0597		30	0.0669		50	0.0464		42	0.0453	
			33	0.0900		23	0.0747		33	0.0696		40	0.0890	
	3	(13,17)	90	0.0499	(20,23)	150	0.0492	(15,19)	50	0.0201	(20,24)	43	0.0143	
			63	0.0601		80	0.0761		33	0.0451		42	0.0369	
			58	0.0923		44	0.0832		32	0.0857		27	0.0953	
	4	(16,20)	100	0.0249	(25,28)	55	0.0439	(25,29)	60	0.0185	(30,34)	63	0.0185	
			72	0.0430		54	0.0581		54	0.0323		62	0.0410	
			65	0.0987		53	0.0943		53	0.0604		27	0.0948	
	11	4	(7,12)	170	0.0443	(11,15)	140	0.0469	(15,19)	78	0.0331	(20,24)	85	0.0164
				100	0.0506		90	0.0556		68	0.0499		44	0.0416
				66	0.0978		55	0.0987		53	0.0960		43	0.0693
6		(10,15)	170	0.0152	(11,15)	100	0.0104	(18,22)	90	0.0143	(25,29)	55	0.0301	
			1022	0.0471		58	0.0486		79	0.0427		54	0.0458	
			79	0.0929		52	0.0913		42	0.0873		53	0.0770	
8		(15,20)	143	0.0472	(17,21)	130	0.0101	(21,25)	100	0.0130	(26,30)	70	0.0227	
			115	0.0749		84	0.0477		91	0.0457		56	0.0471	
			111	0.0961		78	0.0906		48	0.0916		55	0.0789	
25		8	(5,8)	85	0.0496	(10,13)	250	0.0337	(15,18)	155	0.0198	(19,22)	120	0.0104
				70	0.0834		130	0.0481		105	0.0498		81	0.0464
				66	0.0966		85	0.0998		73	0.0906		43	0.0877
	14	(8,12)	150	0.0428	(15,18)	200	0.0115	(28,31)	240	0.0166	(30,33)	130	0.0330	
			140	0.0567		146	0.0486		187	0.0471		127	0.0470	
			120	0.0972		114	0.0963		130	0.0960		67	0.0821	
	16	(12,16)	210	0.0444	(19,22)	210	0.0234	(32,35)	240	0.0155	(33,36)	160	0.0105	
			190	0.0665		185	0.0469		215	0.0431		140	0.0468	
			170	0.0998		143	0.0961		148	0.0946		73	0.0902	

Στη συνέχεια παρουσιάζεται ο πίνακας 3.13 από τον οποίο μπορούμε να εξάγουμε ενδιαφέρουσες παρατηρήσεις σχετικά με την απόδοση του προτεινόμενου διαγράμματος για την περίπτωση των εναλλακτικών τύπου Lehmann . Στο συγκεκριμένο πίνακα, υπολογίζονται οι πιθανότητες συναγερμού για διάφορες τιμές των παραμέτρων σχεδιασμού m, n, a, b, w, r_0 και της παραμέτρου μετατόπισης $\gamma > 0$. Στην πραγματικότητα, οι σχεδιασμοί που παρουσιάζονται στον πίνακα ελέγχονται για την ικανότητα τους να ανιχνεύουν μετατοπίσεις από την εντός ελέγχου κατανομή.

Σύμφωνα με τα αποτελέσματα όπως μπορούμε να δούμε, για την περίπτωση $m = 200, n = 25$ που περιγράφηκε νωρίτερα, ο σχεδιασμός που αντιστοιχεί στις τιμές $a = 19, b = 22, w = 81, r_0 = 8$ (από τους τρεις εναλλακτικούς σχεδιασμούς που πετυχαίνουν πιθανότητα λανθασμένου συναγερμού ίση σχεδόν με 5%), παρουσιάζει την μεγαλύτερη πιθανότητα συναγερμού για ανίχνευση μιας μετατόπισης σε εναλλακτικές τύπου Lehmann με $\gamma = 1/3$ ($AR = 0.8868$) και $\gamma = 1/5$ ($AR = 0.9977$). Παρακάτω παρουσιάζεται ο πίνακας 3.13

Πίνακας 3.13: Πιθανότητες συναγερωμού (AR) για δεδομένο σχεδιασμό. Κάθε κελί περιέχει τις τιμές AR για $\gamma = 1/3$ (άνω είσοδος) και $\gamma = 1/5$ (κάτω είσοδος)

		Δείγμα αναφοράς μεγέθους m											
n	r_n	40			60			100			200		
		(a,b)	w	AR	(a,b)	w	AR	(a,b)	w	AR	(a,b)	w	AR
5	2	(7,11)	200	0.5858	(10,13)	45	0.5910	(15,19)	150	0.5484	(20,24)	80	0.4367
				0.8251			0.8296			0.8067			0.7333
			40	0.5967		30	0.6011		50	0.5661		42	0.4907
		3	(13,17)		0.8300		0.8333		0.8152		0.7643		
				33	0.6340	23	0.6228	33	0.6054	40	0.5259		
					0.8491		0.8452		0.8354		0.7804		
	(20,23)		90	0.4886	150	0.5037	50	0.2474	43	0.2071			
				0.7189		0.7320		0.5012		0.4433			
			63	0.4937	80	0.5168	33	0.3253	42	0.2373			
	4	(16,20)		0.7217		0.7386		0.5641		0.4658			
			58	0.5211	44	0.5352	32	0.3660	27	0.2752			
				0.7373		0.7515		0.5921		0.4824			
(25,28)		100	0.2156	55	0.2529	60	0.1108	63	0.1028				
			0.3912		0.4295		0.2537		0.2079				
		72	0.2319	54	0.2837	54	0.1643	62	0.1262				
11	4	(7,12)		0.4007		0.4552		0.3062		0.2236			
			65	0.2468	53	0.3233	53	0.1981	27	0.1503			
				0.4067		0.4844		0.3323		0.2331			
		(10,15)	170	0.8018	(11,15)	140	0.8309	(15,19)	78	0.7820	(20,24)	85	0.6539
				0.9690		0.9764		0.9670		0.9341		0.9341	
			100	0.8088	90	0.8351	68	0.8059	44	0.7107			
	6	(10,15)		0.9706		0.9771		0.9716		0.9477			
			66	0.8301	55	0.8616	53	0.8107	43	0.7362			
				0.9745		0.9818		0.9721		0.9532			
		(17,21)	170	0.5838	(11,15)	100	0.4374	(18,22)	90	0.4394	(25,29)	55	0.4097
				0.8739		0.7967		0.8021		0.7605		0.7605	
			102	0.6049	58	0.4998	79	0.4864	54	0.4403			
8	(15,20)		0.8803		0.8260		0.8198		0.7698				
		79	0.6477	52	0.5514	42	0.5717	53	0.4589				
			0.8955		0.8416		0.8584		0.7732				
	(21,25)	143	0.3686	(17,21)	130	0.2187	100	0.1411	(26,30)	70	0.0715		
			0.6687		0.5307		0.4167		0.2504		0.2504		
		115	0.3959	84	0.2712	91	0.1762	56	0.2419				
25	8	(5,8)		0.5660		0.5660		0.4291		0.3954			
			111	0.4057	78	0.3109	48	0.3433	55	0.2595			
				0.6855		0.5789		0.5689		0.3985			
		(19,22)	85	0.9181	(10,13)	250	0.9636	(15,18)	155	0.9602	(19,22)	120	0.8681
				0.9978		0.9995		0.9995		0.9972		0.9972	
			70	0.9336	130	0.9672	105	0.9649	81	0.8868			
	14	(8,12)		0.9984		0.9996		0.9996		0.9977			
			66	0.9347	85	0.9756	73	0.9730	43	0.9301			
				0.9984		0.9997		0.9997		0.9987			
		(30,33)	150	0.5543	(15,18)	200	0.6696	(28,31)	240	0.7627	(30,33)	130	0.3777
				0.9212		0.9630		0.9810		0.8703		0.8703	
			140	0.5636	146	0.6848	187	0.7736	127	0.3777			
16	(12,16)		0.9219		0.9650		0.9820		0.8703				
		120	0.6135	114	0.7200	130	0.8066	67	0.5636				
			0.9347		0.9693		0.9848		0.9110				
	(32,35)	210	0.3975	(19,22)	210	0.5866	240	0.6028	(33,36)	160	0.1742		
			0.8187		0.9283		0.9385		0.6775		0.6775		
		190	0.4130	185	0.5890	215	0.6090	140	0.2026				
		0.8210		0.9284		0.9390		0.6803		0.6803			
	170	0.4326	143	0.6240	148	0.6562	73	0.4241					
		0.8311		0.9348		0.9466		0.7692					

Στη συνέχεια συγκρίνεται η απόδοση του W – διαγράμματος και του νέου μη παραμετρικού διαγράμματος που εισήχθηκε στη παράγραφο 3.3. Συγκεκριμένα, χρησιμοποιώντας τρία διαφορετικά επίπεδα FAR και διάφορες τιμές για τις παράμετρος m, n , προσδιορίζουμε τις παραμέτρους σχεδιασμού των διαγραμμάτων έτσι ώστε να επιτυγχάνεται η πιθανότητα FAR ίση με την τιμή στόχο που έχουμε επιλέξει σε κάθε περίπτωση. Στη συνέχεια, υπολογίζουμε αριθμητικά τις AR τιμές για δύο συγκεκριμένες εναλλακτικές τύπου Lehmann

($\gamma = 0.4$ και $\gamma = 0.2$). Τα αποτελέσματα αυτά συνοψίζονται στο τμήμα του παρακάτω πίνακα 3.14 με επικεφαλίδα « W – διάγραμμα».

Πίνακας 3.14: Σύγκριση των AR των διαγραμμάτων ελέγχου με την ίδια FAR ($\gamma = 0.4$ και 0.2)

FAR	m	n	W -διάγραμμα				Διάγραμμα $(j=(n+1)/2)$				
			(LCL,UCL)	n_0	Αριθμός FAR	AR	(LCL,UCL)	r	Αριθμός FAR	AR	
0.01	50	5	(4,7)	31	2	0.0082	0.2538 0.6660	(3,48)	1	0.0072	0.1888 0.5956
		11		41	4	0.0097	0.3848 0.8853	(7,44)	1	0.0093	0.3651 0.8638
		25		90	7	0.0096	0.7018 0.9974	(6,45)	13	0.0097	0.3886 0.9388
	100	5	(10,13)	65	2	0.0100	0.3089 0.9957	(7,94)	1	0.0086	0.2130 0.6455
		11		66	4	0.0095	0.4610 0.9249	(6,95)	7	0.0091	0.3179 0.8539
		25		115	7	0.0096	0.8113 0.9990	(8,97)	16	0.0082	0.5194 0.9548
	500	5	(50,53)	120	2	0.0094	0.5139 0.7317	(38,43)	2	0.0091	0.1472 0.6760
		11		104	4	0.0073	0.4817 0.9398	(35,68)	7	0.0099	0.1643 0.8994
		25		207	7	0.0091	0.8401 0.9995	(120,81)	7	0.0095	0.7411 0.9965
	1000	5	(100,103)	203	2	0.0091	0.3164 0.7354	(79,92)	2	0.0099	0.2569 0.6956
		11		203	4	0.0072	0.4818 0.9369	(96,95)	6	0.0098	0.3188 0.8423
		25		206	7	0.0086	0.8542 0.9996	(243,78)	7	0.0099	0.7545 0.9969
0.005	50	5	(3,6)	45	2	0.0036	0.1885 0.5956	(2,49)	1	0.0030	0.1279 0.5053
		11		39	4	0.0036	0.2878 0.8287	(5,46)	2	0.0025	0.2377 0.7804
		25		80	7	0.0049	0.5068 0.9896	(7,44)	11	0.0048	0.3547 0.9410
	100	5	(7,10)	40	2	0.0046	0.2254 0.6471	(5,96)	1	0.0035	0.1618 0.5730
		11		49	4	0.0048	0.3348 0.8729	(5,96)	7	0.0046	0.2538 0.8141
		25		83	7	0.0049	0.6661 0.9962	(15,86)	11	0.0045	0.4112 0.9659
	500	5	(40,43)	165	2	0.0048	0.2569 0.6961	(20,481)	3	0.0048	0.1197 0.5581
		11		84	4	0.0046	0.3973 0.9100	(56,454)	5	0.0049	0.3044 0.8462
		25		168	7	0.0048	0.7460 0.9988	(90,411)	10	0.0050	0.5391 0.9865
	1000	5	(82,85)	350	2	0.0050	0.2627 0.6922	(62,99)	2	0.0049	0.2043 0.6347
		11		167	4	0.0044	0.4040 0.9131	(136,865)	4	0.0049	0.3715 0.8839
		25		171	7	0.0044	0.7665 0.9990	(129,872)	13	0.0050	0.4760 0.9766
0.0027	50	5	(2,5)	19	2	0.0026	0.1404 0.5207	(1,50)	1	0.0008	0.0634 0.3617
		11	(3,6)	43	4	0.0027	0.2788 0.8264	(5,46)	2	0.0025	0.2377 0.7804
		25		100	7	0.0025	0.5514 0.9873	(7,44)	10	0.0021	0.3509 0.9581
	100	5	(5,8)	26	2	0.0026	0.1790 0.5854	(4,97)	2	0.0021	0.1304 0.5262
		11		39	3	0.0027	0.2437 0.8134	(9,92)	5	0.0023	0.2185 0.7756
		25		73	6	0.0026	0.6461 0.9960	(12,89)	12	0.0019	0.3828 0.9364
	500	5	(32,35)	134	2	0.0026	0.2089 0.6395	(16,485)	3	0.0026	0.1197 0.5116
		11	(38,41)	155	4	0.0025	0.3534 0.8952	(64,437)	2	0.0026	0.3386 0.8692
		25		162	7	0.0026	0.7216 0.9985	(101,400)	8	0.0026	0.6978 0.9915
	1000	5	(66,69)	330	2	0.0027	0.2151 0.6470	(49,952)	2	0.0025	0.1626 0.5861
		11	(64,67)	131	4	0.0027	0.3186 0.8778	(49,952)	7	0.0027	0.2536 0.8355
		25	(66,69)	268	7	0.0026	0.6456 0.9975	(189,812)	9	0.0027	0.5626 0.9891

Από τις τιμές του παραπάνω πίνακα, μπορούμε να πούμε πως για την ίδια τιμή FAR , το νέο διάγραμμα έχει καλύτερη απόδοση από το διάγραμμα που ορίζεται στην παράγραφο 3.3 αφού παρουσιάζει μεγαλύτερη πιθανότητα AR για όλες τις περιπτώσεις που εξετάστηκαν. Για παράδειγμα, χρησιμοποιώντας δείγμα αναφοράς μεγέθους $m = 500$, εξεταζόμενα δείγματα μεγέθους $n = 5$ και προκαθορισμένο επίπεδο $FAR = 0.0027$, το διάγραμμα που ορίζεται στην 3.3 παρουσιάζει για $\gamma = 0.4$ ($\gamma = 0.2$), πιθανότητες συναγερμού ίσες με

0.1197 (0.5116), ενώ οι αντίστοιχες τιμές για το W- διάγραμμα είναι 0.2089 (0.6395).

Συγκρίνουμε τέλος την απόδοση των δύο μη παραμετρικών διαγραμμάτων ελέγχου καθορίζοντας ένα κοινό εντός ελέγχου μέσο μήκος ροής ARL_{in} και υπολογίζοντας τις αντίστοιχες πιθανότητες συναγερμού (AR) για συγκεκριμένες εκτός ελέγχου (τύπου Lehmann) εναλλακτικές. Έτσι λοιπόν στον παρακάτω πίνακα 3.15 παρουσιάζουμε τις τιμές AR των δύο διαγραμμάτων για $ARL_{in} = 370, 500$, $m = 100, 200$, $n = 5, 7, 11$ και $\gamma = 0.2, 0.5$. Οι παράμετροι σχεδιασμού $LCL = X_{a:b}$, $UCL = X_{b:m}$, w, r_0 προσδιορίζονται έτσι ώστε τα ARL_{in} να είναι όσο το δυνατόν πλησιέστερα στην τιμή στόχο που έχουμε επιλέξει σε κάθε περίπτωση. Τα αποτελέσματα του πίνακα 3.15 αποκαλύπτουν την υπεροχή του W – διαγράμματος σύμφωνα και με αυτό τον τρόπο σύγκρισης.

Πίνακας 3.15: Σύγκριση των AR των διαγραμμάτων ελέγχου με την ίδια τιμή ARL_{in} . Κάθε AR κελί περιέχει τις τιμές που επιτυγχάνονται για $\gamma = 0.5$ και 0.2 αντίστοιχα

ARL_m	m	n	W -διάγραμμα					Διάγραμμα $(j=(n+1)/2)$				
			(LCL,UCL)	w	r_0	Ακριβές ARL_m	AR	(LCL,UCL)	r	Ακριβές ARL_m	AR	
370	100	5	(8,11)	51	2	370.30	0.1424 0.6757	(6,95)	1	359.64	0.0990 0.6121	
		7	(6,9)	44	2	379.04	0.2325 0.8578	(9,92)	2	369.35	0.1279 0.7283	
		11	(11,14)	75	4	358.82	0.2937 0.9364	(14,87)	3	360.04	0.1952 0.8770	
	200	5	(14,17)	60	2	367.31	0.1240 0.6575	(11,190)	2	363.60	0.0895 0.6028	
		7	(9,12)	40	2	379.44	0.1823 0.8288	(16,185)	3	372.86	0.1105 0.7103	
		11	(7,10)	50	2	376.58	0.3359 0.9673	(22,179)	5	367.80	0.1410 0.8370	
	500	100	5	(8,11)	88	2	499.46	0.1404 0.6744	(4,97)	3	499.88	0.0752 0.5407
			7	(6,9)	74	2	500.62	0.2280 0.8545	(8,93)	3	490.92	0.1116 0.7019
			11	(11,14)	80	4	494.65	0.2883 0.9360	(13,88)	4	462.98	0.1749 0.8621
200		5	(14,17)	93	2	498.29	0.1195 0.6532	(7,194)	3	500.20	0.0621 0.5229	
		7	(9,12)	41	2	505.36	0.1772 0.8266	(16,185)	2	490.76	0.1062 0.7083	
		11	(18,21)	59	4	504.50	0.2405 0.9192	(21,180)	5	496.80	0.1274 0.8255	

Εφαρμογές

Παραμετρικά και μη-παραμετρικά διαστήματα πρόβλεψης βασισμένα κατά την φάση II των διαγραμμάτων ελέγχου των επαναλαμβανόμενων δεδομένων βιολογικής ανάλυσης

Ο ποιοτικός έλεγχος για τα επαναλαμβανόμενα τρεξίματα βιολογικής ανάλυσης μπορεί να εκτελεσθεί από τη φάση II των διαγραμμάτων ελέγχου, γνωστά από το βιομηχανικό ποιοτικό έλεγχο. Η αξία ενδιαφέροντος είναι η δραστικότητα, της οποίας μια μοναδική τιμή ανά τρέξιμο είναι διαθέσιμη. Τα παραμετρικά και μη-παραμετρικά διαστήματα πρόβλεψης περιγράφονται για να υπολογίσουν τα διαστήματα του ποιοτικού ελέγχου για τα μελλοντικά τρεξίματα. Οι παραβιάσεις ενάντια στην κανονική κατανομή που συχνά εμφανίζονται είναι στην ουσία απόμακρες από τα υπόλοιπα στοιχεία. Τα μη-παραμετρικά διαστήματα πρόβλεψης περιορίζονται σε όχι και πολύ μικρά μεγέθη δειγμάτων καθώς επίσης στις ιστορικές και μελλοντικές φάσεις δειγματοληψίας. Επομένως, τα ισχυρά διαστήματα πρόβλεψης βασίζονται στο μετασχηματισμό ακραίων τιμών.

1. Το πρόβλημα

Σε ένα βιο-φαρμακευτικό εργαστήριο, συνήθως η δραστικότητα των φαρμάκων, τα δείγματα σταθερότητας, κ.λπ. υπολογίζονται με τη σύγκριση τους με τις αντίστοιχες ενώσεις αναφοράς τους σε μια επικυρωμένη βιολογική ανάλυση. Συνήθως οι βιολογικές αναλύσεις εκτελούνται στις ιδιαίτερες τυποποιημένες λεπτές πλάκες βιολογικής αναλύσεως (εδικές λεπτές πλάκες με κυψέλες οι οποίες είναι ευρέως διαδεδομένες για χρήσεις κατά την βιολογική ανάλυση). Μια ενιαία εκτίμηση δραστικότητας ανά τρέξιμο εμφανίζεται, η οποία υπολογίζεται από διάφορες συγκεντρώσεις και αντενστάσεις. Στις περισσότερες βιολογικές αναλύσεις η δραστικότητα ενός δείγματος υπολογίζεται ως σχετική δραστικότητα από μια παράλληλο-γραμμή ή μια αναλογία δοκιμής κλίσεων, δηλαδή η δραστικότητα και το διάστημα εμπιστοσύνης της υπολογίζονται από διάφορες συγκεντρώσεις για κάθε δείγμα και ουσίας ανά τρέξιμο. Συχνά η δραστικότητα της ουσίας όπως το EC50 χρησιμεύει ως ένα ποιοτικό μέτρο της δοκιμής, μερικές φορές αυτό χρησιμοποιείται ακόμη και ως κριτήριο αποδοχής της δοκιμής. Οι εκτιμήσεις της δραστικότητας μιας ένωσης συλλέγονται κατά τη διάρκεια μιας ορισμένης περιόδου, παραδείγματος χάριν εννέα μήνες, για να παρατηρηθεί η σταθερότητα της βιολογικής ανάλυσης. Με βάση έναν περιορισμένο αριθμό νέων τρεξιμάτων μιας ακόλουθης φάσης επανελέγχου, η απόφαση λαμβάνεται εάν η καθιερωμένη ένωση μπορεί να χρησιμοποιηθεί περαιτέρω ή πρέπει να ανανεωθεί. Η τυποποιημένη προσέγγιση για το χαρακτηρισμό της σταθερότητας των εργαστηριακών στοιχείων είναι το διάγραμμα ελέγχου, όπου το κοινό διάγραμμα υποθέτει την ομοιογένεια και την κανονική κατανομή για τις επαναλαμβανόμενες μετρήσεις ανά τρέξιμο. Μέχρι τώρα, τα διαγράμματα ελέγχου για τις εκτιμήσεις μαζί με τα διαστήματα εμπιστοσύνης τους δεν είναι διαθέσιμα. Ένα διάγραμμα ελέγχου για τις εκτιμήσεις της δραστικότητας βασισμένες στις μη γραμμικές καμπύλες συγκέντρωσης με τους ιεραρχικούς τυχαίους παράγοντες, όπως μεταξύ των τρεξιμάτων, μεταξύ των πλακών, και μεταξύ της σειράς και των στηλών είναι μια δυσκολία. Επειδή η μεταβλητότητα μέσα στις κυψέλες των πλακών είναι συνήθως αρκετά μικρή, μια

ενιαία αξία ανά τρέξιμο χρησιμοποιείται στην πράξη. Επομένως, υποτίθεται μια προσέγγιση διαγραμμάτων ελέγχου για μια κάθε φορά συνεχή διαδικασία. Εδώ, εξετάζονται δύο πρόσθετοι όροι:

- i. στερεότυπη χρήση για τις διαφορετικές βιολογικές αναλύσεις, όπου η υπόθεση της κανονικής κατανομής είναι προβληματική, και
- ii. Τα μη-ξαναχρησιμοποιημένα στοιχεία, μια μέτρηση τη φορά (αντίστοιχο τρέξιμο).

Όταν η υπόθεση κανονικότητας δεν ικανοποιείται, οι προϋποθέσεις για να εργαστούμε με τα κοινά διαγράμματα ελέγχου συχνά παραβιάζονται σοβαρά και επομένως οι μη-παραμετρικές προσεγγίσεις είναι συστημένες. Η αξιολόγηση διάφορων ιστορικών στοιχείων βιολογικής ανάλυσης έδειξε στις περισσότερες περιπτώσεις την παραβίαση της κανονικής κατανομής και επομένως μια μη-παραμετρική προσέγγιση διαμορφώθηκε. Εναλλακτικά, προτείνονται τα απλά παραμετρικά διαγράμματα ελέγχου για μιας χρονικής στιγμής στοιχεία. Σε μερικές βιολογικές αναλύσεις τα ακραία σημεία είναι ασύμμετρα κατανεμημένα και επομένως ένας λογαριθμικός μετασχηματισμός φαίνεται να είναι ένας απλός τρόπος να χρησιμοποιηθούν οι παραμετρικές προσεγγίσεις, που μερικές φορές συστήνονται. Εντούτοις, βασισμένοι στα περιορισμένα μεγέθη των δειγμάτων, μια δοκιμή στη λογαριθμική-κανονική κατανομή δεν είναι αρκετά ικανοποιητική. Από έναν μετασχηματισμό στοιχείων όχι μόνο η θέση αλλά και η κλίμακα μετασχηματίζονται, και ιδιαίτερα για τα διαγράμματα ελέγχου μια σωστή εκτίμηση κλίμακας είναι σημαντική. Ιδιαίτερα τα παραμετρικά διαγράμματα ελέγχου δεν είναι αρκετά ισχυρά στις παραβιάσεις των υποθέσεων. Οι μη παραμετρικές προσεγγίσεις που είναι ισχυρές ενάντια σε οποιοδήποτε είδος ασύμμετρης κατανομής, είναι διαθέσιμες και περίπου τόσο αποδοτικές όσο και οι παραμετρικές προσεγγίσεις όταν εκπληρώνονται πραγματικά οι υποθέσεις τους.

Δύο τύποι διαγραμμάτων ελέγχου μπορούν να διακριθούν: τα φάσης I εντός έλεγχος και τα φάσης II διαγράμματα ελέγχου. Πρόσφατα, τα διαστήματα ανοχής χρησιμοποιήθηκαν για την επικύρωση της προμελέτης διάφορων αναλυτικών μεθόδων κατά την εντός ελέγχου διαδικασία. Για το ανωτέρω πρόβλημα επανελέγχου, η φάση II είναι παρόμοια. Χρησιμοποιώντας τα διαγράμματα φάσης II, η ποιότητα των n ταυτόχρονων τρεξιμάτων (στη φάση επανελέγχου) χαρακτηρίζεται από τον υπολογισμό των χαρακτηριστικών διαδικασίας των m ιστορικών τρεξιμάτων. Το πρόβλημα που συζητείται εδώ είναι η ετήσια απόδειξη της σταθερότητας μιας ένωσης μιας βιολογικής ανάλυσης. Μετά από περίπου εννέα μηνών τιμές τα m ιστορικά τρεξίματα συλλέχθηκαν και από τα νέα n τρεξίματα, που εμφανίζονται κατά τη διάρκεια της φάσης επανελέγχου μέσα στους επόμενους δύο έως τρεις μήνες, η απόφαση που απαιτείται είναι, εάν η ένωση είναι ακόμα υπό έλεγχο ή όχι. Έτσι το μήκος ενός αναλυτικού τρεξίματος καθορίζεται. Η ιδέα της φάσης II των διαγραμμάτων ελέγχου είναι πρώτη για να υπολογίσει τη θέση και τη διαφορά από τα m ιστορικά τρεξίματα, όπου m μπορεί να είναι όσο το δυνατόν μεγαλύτερο. Αφετέρου τα όρια ελέγχου υπολογίζονται, όπου ένα ορισμένο ποσό των νέων n τρεξιμάτων, τιμές ή επιλεγμένες στατιστικές, όπως η διάμεσος, πρέπει να είναι μέσα σε αυτά τα όρια για να χαρακτηρίσει τη διαδικασία υπό έλεγχο. Έτσι η φάση II των διαγραμμάτων που υποθέτουν την κανονική κατανομή περιγράφηκε, αλλά σπάνια οι μη-παραμετρικές προσεγγίσεις είναι διαθέσιμες. Η φάση I των διαγραμμάτων ελέγχου Shewhart για τα μοναδικά

ανά χρονική στιγμή στοιχεία είμαι χρήσιμη να χαρακτηρίσουμε τη συμπεριφορά των ιστορικών τρεξιμάτων.

Δύο μη-παραμετρικές προσεγγίσεις χρησιμοποιούνται εδώ:

- i. τα διαγράμματα ελέγχου προτεραιότητας για μια μελλοντική διάμεσο και
- ii. το διάστημα πρόβλεψης για να προσεγγίσουμε τουλάχιστον k από τα n νέα τρεξίματα

Δεδομένου ότι η καθορισμένη στατιστική προσέγγιση πρέπει να εφαρμοστεί από το στατιστικά χωρίς πείρα προσωπικό, η απεικόνιση των τρεξιμάτων από τον ιστορικό εντός έλεγχου και η πρόσφατη φάση ελέγχου σε ένα διάγραμμα ελέγχου είναι χρήσιμη. Επομένως, ένα σχετικό πολυσύνθετο διάγραμμα παρήχθη χρησιμοποιώντας το λογισμικό R.

2. Εξετάζοντας τη συμπεριφορά των ιστορικών τρεξιμάτων

Τρία χαρακτηριστικά των ιστορικών στοιχείων τρεξίματος είναι ενδιαφέρον.

Πρώτον η παραβίαση ενάντια στην κανονική κατανομή. Δεύτερον η πιθανή αυτοσυσχέτιση, και η πιθανότητα ύπαρξης ακραίων τιμών. Η παραβίαση ενάντια στην κανονική κατανομή χαρακτηρίζεται από τρεις μεθόδους:

- i. Quantile-quantile plot (QQ-plot). Ο όρος quantile δηλώνει έναν αριθμό X_p τέτοιος ώστε η αναλογία των p από τον πληθυσμό τιμών να είναι μικρότερη ή ίση από τον αριθμό X_p . Η μέθοδος παριστάνεται από ένα διάγραμμα πιθανότητας, δηλαδή μια γραφική μέθοδο για να συγκρίνουμε δυο κατανομές πιθανότητας η μια ενάντια στην άλλη.
- ii. Lilliefors test μέσω της προσαρμογής του Kolmogorov-Smirnov test. Χρησιμοποιείται για τον έλεγχο της μηδενικής υπόθεσης των δεδομένων που προέρχονται από δείγμα κανονικής κατανομής, όταν η μηδενική υπόθεση δεν διευκρινίζει την κανονική κατανομή, δηλαδή την αναμενόμενη μέση τιμή και διασπορά. Αρχικά εκτιμάται ο μέσος όρος και η διασπορά του πληθυσμού των δεδομένων. Στη συνέχεια βρίσκεται η μέγιστη απόκλιση μεταξύ της εμπειρικής συνάρτησης κατανομής και της εκτιμημένης συνάρτησης κατανομής (CDF) της κανονικής κατανομής με τον εκτιμημένο μέσο όρο και την εκτιμημένη διασπορά. Τελικά αντιμετωπίζεται το θέμα αν η μέγιστη απόκλιση είναι αρκετά μεγάλη ώστε να είναι στατιστικά σημαντική απαιτώντας κατά συνέπεια την απόρριψη της μηδενικής υπόθεσης. Αυτό το κομμάτι κάνει το test να είναι πιο περίπλοκο από το Kolmogorov-Smirnov test. Δεδομένου ότι το υποτιθέμενο CDF έχει κινηθεί πιο κοντά προς τα δεδομένα, η μέγιστη απόκλιση έχει γίνει μικρότερη από ότι θα ήταν αν η μηδενική υπόθεση είχε επιλέξει μόνο μια κανονική κατανομή. Κατά συνέπεια η μηδενική υπόθεση της στατιστικής δοκιμής δηλαδή η κατανομή πιθανότητας υποθέτει την αλήθεια της μηδενικής κατανομής που είναι στοχαστικά μικρότερη από την Kolmogorov-Smirnov κατανομή.
- iii. Shapiro-Wilks test. Η μέθοδος αυτή εξετάζει την μηδενική υπόθεση ενός δείγματος που προέρχεται από κανονική κατανομή. Η μέθοδος QQ-plot παρέχει τις πληροφορίες για τη διαφορά της υποτιθέμενης κανονικής κατανομής, τα αποτελέσματα κλίμακας και τις ακραίες τιμές. Το Durbin-Watson test χρησιμοποιείται για την ανίχνευση της αυτοσυσχέτισης. Εντούτοις, κανένας έλεγχος για ακραίες τιμές δεν χρησιμοποιείται,

επειδή χωρίς τη γνώση της κατανομής, οι επίσημοι έλεγχοι ακραίων τιμών και η αποβολή των ακραίων τιμών μπορούν να προκαταλάβουν σοβαρά τα όρια ελέγχου. Στο παράδειγμα της παραγράφου 4 υπάρχει ένα QQ-σχεδιάγραμμα το οποίο υποδηλώνει την παραβίαση της κανονικής κατανομής καθώς η υπολογισμένη p-value τιμή και από το Shapiro-Wilks test και από το Lilliefors-KS test είναι μικρότερη από 0,10. Το Durbin-Watson test ότι δεν υπάρχει αυτοσυσχέτιση στην διαδικασία.

3. Τα διαγράμματα ελέγχου μιας χρονικής στιγμής δεδομένων

Η στατιστική βιβλιογραφία περιγράφει τουλάχιστον τρεις τύπους διαστημάτων που χρησιμοποιούνται για τα δείγματα των στοιχείων. Το διάστημα εμπιστοσύνης περιέχει ένα άγνωστο χαρακτηριστικό των ιστορικών δειγμάτων των στοιχείων, όπως ο μέσος όρος, με μια προδικρισμένη πιθανότητα εμπιστοσύνης. Το διάστημα ανοχής περιέχει ένα διευκρινισμένο ποσοστό των μελλοντικών δειγμάτων όπου ο αριθμός μελλοντικών δειγμάτων δεν είναι ανάγκη να διευκρινιστεί. Το διάστημα πρόβλεψης περιέχει ένα, όλα, ή τουλάχιστον k από n μελλοντικά δείγματα βασισμένα στα m ιστορικά δείγματα με μια προκαθορισμένη πιθανότητα. Επομένως, τα διαστήματα πρόβλεψης που χρησιμοποιούνται για την ανωτέρω περιγραφημένη βιολογική ανάλυση επανελέγχουν το πρόβλημα μέσω της φάσης II του διαγράμματος ελέγχου. Υπάρχουν μερικές βιολογικές αναλύσεις στις οποίες η τάση για πιθανή αλλαγή είναι εκ των προτέρων γνωστή, σε μερικές ανοδική και σε άλλες καθοδική, επομένως λαμβάνεται υπόψη η μονόπλευρη ή η αμφίπλευρη πρόβλεψη διαστημάτων.

Η λήψη αποφάσεων για τη βιολογική ανάλυση επανελέγχου του προβλήματος είναι βασισμένη στα εκτιμημένα όρια ελέγχου. Επίσης το υλικό αναφοράς μπορεί επιπλέον να χρησιμοποιηθεί, όταν ο αριθμός των μελλοντικών τιμών των δειγμάτων που είναι εκτός των ορίων ελέγχου είναι μικρότερος από ένα προκαθορισμένο επίπεδο τιμών, διαφορετικά ένα νέο υλικό αναφοράς πρέπει να χρησιμοποιηθεί.

3.1. Φάση I παραμετρικά διαγράμματα Shewhart

Η φάση I στα διαγράμματα ελέγχου τύπου Shewhart για τα μιας χρονικής στιγμής στοιχεία χρησιμοποιείται για να χαρακτηρίσει τη συμπεριφορά μόνο των ιστορικών τρεξιμάτων. Λαμβάνοντας υπόψη ένα δείγμα παρατηρήσεων $x_i, i = 1, \dots, m$ και υποθέτοντας κανονική κατανομή, τα όρια ελέγχου υπολογίζονται εύκολα από την σχέση $[LCL_\sigma; UCL_\sigma] = \bar{x} \pm As$, όπου το A είναι μια σταθερά, που επιλέγεται συχνά να είναι 3 ή 2. Αυτός ο παράγοντας πολλαπλασιασμού προσεγγίζει την ποσότητα $t_{m-1, 1-\alpha/2}$ της μονομεταβλητής t-κατανομής η οποία παίρνει τις τιμές 1.96 για ένα $\alpha = 0,05$ και 3 για ένα $\alpha = 0,0027$ για $m = \infty$. Δεδομένου ότι στις μέρες μας δεν είναι ένα μεγάλο υπολογιστικό πρόβλημα να υπολογίσει κάποιος αυτές τις ποσότητες από οποιοδήποτε στατιστικό λογισμικό, τα όρια ελέγχου μπορούν επίσης να υπολογιστούν από την σχέση

$$[LCL_t; UCL_t] = \bar{x} \pm t_{m-1, 1-\alpha/2} s$$

Για να ρυθμιστεί μια πιθανή απόκλιση, που μπορεί να είναι παρούσα στα ιστορικά στοιχεία, η κινητή σειρά (moving range) χρησιμοποιείται στη θέση της τυπικής

χαμηλότερου ορίου ελέγχου (LCL) και του ανώτερου ορίου ελέγχου (UCL), δίνεται από την σχέση

$$[LCL_{pPI}; UCL_{pPI}] = \bar{x} \pm s \cdot t_{m-1, 1-a/2} \sqrt{1 + \frac{1}{m}}$$

και ένα συντηρητικό αμφίπλευρο $(1 - \alpha)$ διάστημα πρόβλεψης για όλα τα n μελλοντικά δείγματα, που χρησιμοποιεί μια απλή Bonferroni προσαρμογή δίνεται από την σχέση

$$[LCL_{pPIall}; UCL_{pPIall}] = \bar{x} \pm s \cdot t_{m-1, 1-a/2n} \sqrt{1 + \frac{1}{m}}$$

Κατά την Bonferroni προσαρμογή αντιμετωπίζεται το πρόβλημα των πολλαπλών συγκρίσεων. Η προσαρμογή βασίζεται στο αν ο ερευνητής ελέγχει n εξαρτημένες ή ανεξάρτητες υποθέσεις από ένα σύνολο στοιχείων. Τότε ένας τρόπος για να διατηρήσουμε το σφάλμα τύπου I είναι να ελέγξουμε κάθε υπόθεση μεμονωμένα με στατιστική σημαντικότητα $\frac{1}{n}$ κάτι που πρακτικά σημαίνει τον έλεγχο μόνο μιας υπόθεσης. Έτσι αν κάποιος θα ήθελε να ελέγξει το επίπεδο σημαντικότητας όλης της οικογενείας να είναι α τότε κατά την Bonferroni προσαρμογή θα έπρεπε κάθε υπόθεση να ελεγχθεί σε επίπεδο σημαντικότητας $\frac{\alpha}{n}$.

Ακριβή διαστήματα βασισμένα σε πολλών μεταβλητών t-κατανομή μπορούν να ληφθούν χρησιμοποιώντας την R-βιβλιοθήκη

$$[LCL_{pPI_all}; UCL_{pPI_all}] = \bar{x} \pm s \cdot t_{n,R,2-sided,m-1,1-a} \sqrt{1 + \frac{1}{m}}$$

με το R να είναι μια μήτρα συσχετισμού $n \times n$ με τα εκτός-διαγωνίου στοιχεία να είναι $\rho = \left(\frac{1}{m+1}\right)$.

Τα σχετικά μονόπλευρα διαστήματα μπορούν να ληφθούν χρησιμοποιώντας $(1 - a)$ αντί για $(1 - \frac{\alpha}{2})$ για την t-κατανομή και τη πολυμεταβλητή t-κατανομή. Εντούτοις, ο όρος ότι όλα τα μελλοντικά τρεξίματα πρέπει να είναι μέσα σε ένα ορισμένο διάστημα είναι αρκετά περιοριστικός. Ρεαλιστικότερο είναι ένα αμφίπλευρο $(1 - \alpha)$ διάστημα πρόβλεψης για k από n μελλοντικά δείγματα.

$$[LCL_{pPI_kofn}; UCL_{pPI_kofn}] = \bar{x} \pm r_{k,n,m,1-a} s \cdot$$

Η ποσότητα $r_{k,n,m,1-a}$ ικανοποιείται όταν $r = \sqrt{\left(\frac{m+1}{m}\right)u^*}$, με το u^* να είναι η αξία του u που ικανοποιεί την σχέση $\sum_{j=k}^m P(F_j(u^*)) = (1 - a)$. Η $P(F_j(u))$ είναι διαθέσιμη μέσω της εκτίμησης του διπλού ολοκληρώματος

$$P(F_j(u)) = \int_0^\infty \left\{ \int_{-\infty}^\infty \binom{n}{j} [\Phi(b) - \Phi(a)]^j \times [\Phi(b) - \Phi(a)]^{n-j} \varphi(y) dy \right\} f_v(s) ds$$

με $\alpha^j = (-us + \sqrt{\rho}y/\sqrt{1-\rho})$, $b = (us + \sqrt{\rho}y/\sqrt{1-\rho})$, $\rho = \frac{1}{m+1}$.

Οι $\Phi(\cdot)$ και $\varphi(\cdot)$ είναι η τυποποιημένη κανονική πυκνότητα και η συνάρτηση κατανομής και η $f_\nu(s)$ είναι η συνάρτηση πυκνότητας του S , όπου το νS^2 ακολουθεί κανονική κατανομή με $m - 1$ βαθμούς ελευθερίας και $\nu = m - 1$. Τα σχετικά μονόπλευρα διαστήματα λαμβάνονται με την παράλειψη του όρου $-\Phi(a)$ κατά τη διάρκεια της αξιολόγησης του ολοκληρώματος, με την υπόθεση ότι $n = k$.

Η σύγκριση πρώτον με την αμφίπλευρη $1 - \alpha = 0,95$ ποσότητα για $n = k$ χρησιμοποιώντας την προσαρμογή Bonferroni, δεύτερον με ποσοότητες βασισμένες στην πολυμεταβλητή t-κατανομή και τρίτον βασισμένη σε αριθμητικές μεθόδους υποδηλώνει το σωστό υπολογισμό των ποσοτήτων από τη σχετική συνάρτηση της R-βιβλιοθήκης και τη συντηρητικότητα των διαστημάτων τύπου Bonferroni. Επιπλέον, ένα διάστημα πρόβλεψης μπορεί να κατασκευαστεί για να περιέχει το μέσο όρο n μελλοντικών παρατηρήσεων βασισμένων στα αποτελέσματα ενός προηγούμενου δείγματος μεγέθους m .

$$[LCL_{PI_mean}; UCL_{PI_mean}] = \bar{x} \pm s \cdot t_{m-1, 1-\alpha/2} \sqrt{\frac{1}{n} + \frac{1}{m}}$$

Οι ανησυχίες ότι τα στοιχεία βιολογικής ανάλυσης παραβιάζουν την υπόθεση της κανονικής κατανομής προέρχονται στην ουσία από το γεγονός πολλών ακραίων τιμών. Σε αυτήν την κατάσταση τα παραμετρικά διαστήματα μπορούν να σταθεροποιηθούν από ισχυρούς εκτιμητές, όπου ο απλούστερος είναι η μετατροπή των ακραίων τιμών σε αποδεκτές στατιστικά τιμές (winsorization). Ας υποθέσουμε ότι $X_1 \leq \dots \leq X_m$ είναι ένα διαταγμένο, ιστορικό δείγμα και υποθέστε το $\gamma = \frac{w}{m}$ να είναι το ποσοστό των παρατηρήσεων, οι οποίες υποτίθεται ότι μόλυναν την κανονική κατανομή των παρατηρήσεων. Κατόπιν ο μέσος όρος υπολογίζεται από

$$x_w = \left(\frac{1}{m}\right) \left((1 + \lfloor \gamma m \rfloor)(X_{1+\lfloor \gamma m \rfloor} + X_{m-\lfloor \gamma m \rfloor}) + \sum_{i=\lfloor \gamma m \rfloor+2}^{m-\lfloor \gamma m \rfloor-1} X_i \right)$$

με διασπορά

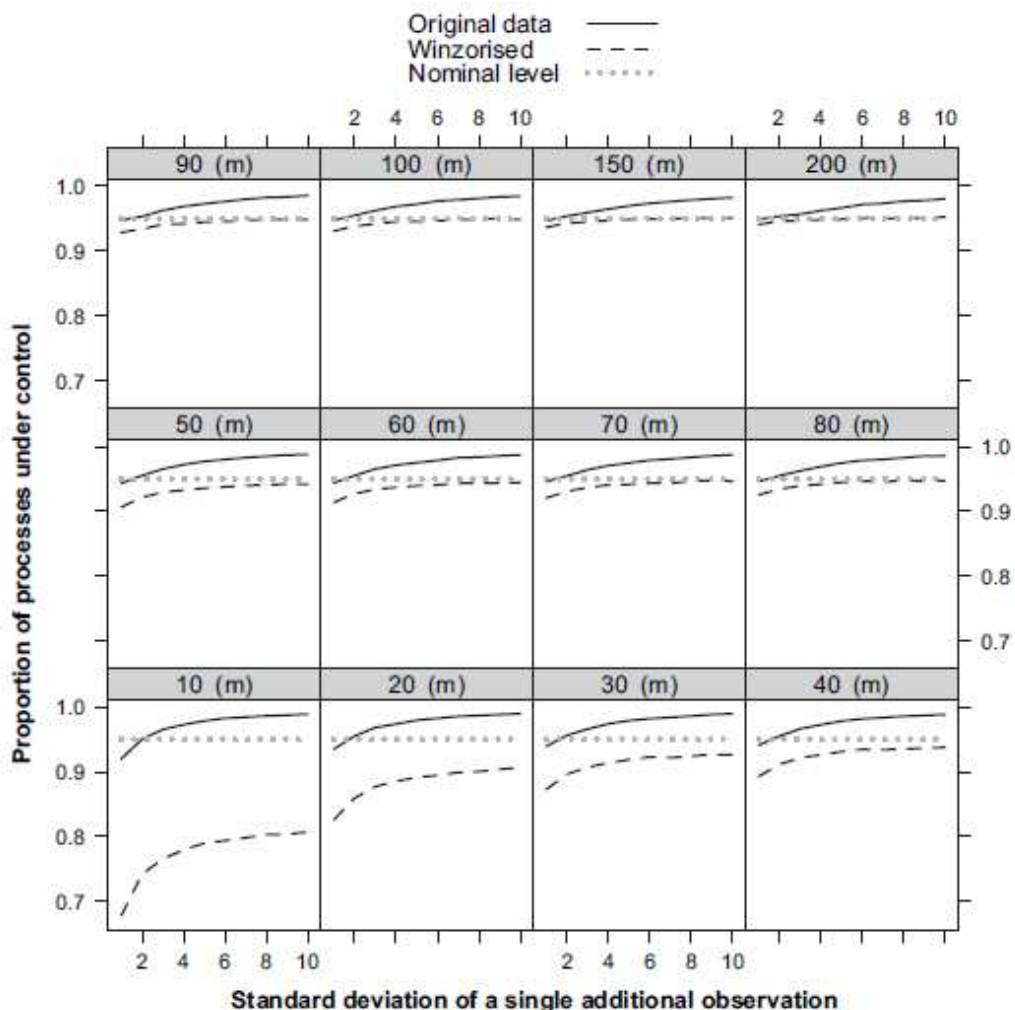
$$s_w = \left(\frac{1}{m}\right) \left((1 + \lfloor \gamma m \rfloor)(X_{1+\lfloor \gamma m \rfloor} - \bar{x}_w)^2 + (X_{m-\lfloor \gamma m \rfloor} - \bar{x}_w)^2 + \sum_{i=\lfloor \gamma m \rfloor+2}^{m-\lfloor \gamma m \rfloor-1} (X_i - \bar{x}_w)^2 \right)$$

Αυτές οι εκτιμήσεις μπορούν να συνδεθούν με τα διαστήματα πρόβλεψης για να παρέχουν μια απόφαση, η οποία είναι λιγότερο επηρεασμένη από μερικές ασυνήθιστες παρατηρήσεις όχι πρώτιστα σχετικές με την ιστορική διαδικασία. Για να ερευνησουμε την απόδοση των παραμετρικών διαγραμμάτων ελέγχου, μια μικρή μελέτη προσομοίωσης πραγματοποιήθηκε. Για 10.000 τρέξιμα προσομοίωσης $m = 10, \dots, 200$ παρατηρήσεις παρήχθησαν με $m - 1$ παρατηρήσεις να είναι κανονικά κατανομημένα με τη τυπική απόκλιση 1, ενώ μια πρόσθετη παρατήρηση λαμβάνεται από μια κανονική κατανομή με τυπική απόκλιση που ποικίλλεται από 1 έως 10. Για κάθε τρέξιμο ένα

παραμετρικό διάστημα πρόβλεψης υπολογίζεται το οποίο βασίζεται στα m ιστορικά δεδομένα. Το ποσοστό των τρεξιμάτων, όπου μια νέα παρατήρηση, που παράγεται από μια κανονική κατανομή με τυπική απόκλιση 1, βρίσκεται μέσα στα όρια διαστήματος πρόβλεψης, μπορεί και υπολογίζεται. Αυτή η αξία αντιπροσωπεύει τα ποσοστά των διαδικασιών υπό έλεγχο, ο οποίος αναμένεται στο ονομαστικό επίπεδο 95% σε ένα προκαθορισμένο λάθος τύπου I 5%. Τα αποτελέσματα προσομοίωσης του ποσοστού των διαδικασιών υπό έλεγχο εξαρτώμενο από την αυξανόμενη εκτροπή της κατανομής, που αντιπροσωπεύεται από τη τυπική απόκλιση μιας ενιαίας παρατήρησης, παρουσιάζονται στο διάγραμμα 2. Ολοκληρώνοντας μπορούμε να πούμε, ότι για τουλάχιστον $m = 40$ ιστορικά τρεξίματα απαιτούνται, ακόμη και για τα κανονικά κατανομημένα στοιχεία, να κρατήσουν το επίπεδο εμπιστοσύνης στο 95%. Εάν η ενιαία απομακρυσμένη αξία είναι πάρα πολύ ακραία (από άποψη τυπικής απόκλισης), η παραμετρική φάση II του διαγράμματος ελέγχου είναι μάλλον συντηρητική και πρέπει να αποφευχθεί.

Για τη χρήση της μεθόδου τροποποίησης των ακραίων τιμών χρειάζονται τουλάχιστον 80 ιστορικά τρεξίματα για να προστατευτεί από μια απρόσμενη συμπεριφορά.

Διάγραμμα 2. Αναλογία υπό έλεγχο για την παραμετρική και τροποποιημένη(winsorized) φάση II των διαγραμμάτων ελέγχου



3.3. Φάση II των μη-παραμετρικών διαγραμμάτων

Δεδομένου ενός διαταγμένου δείγματος των ιστορικών παρατηρήσεων $X_1 \leq \dots \leq X_m$ ένα αμφίπλευρο μη-παραμετρικό $(1 - \alpha)$ διάστημα πρόβλεψης κατασκευάζεται με την επιλογή των ορίων $[X_l, X_u]$ με τέτοιο τρόπο, ώστε να περιλαμβάνουν τουλάχιστον k από n μελλοντικά δείγματα $Y_1 \leq \dots \leq Y_m$. Οι τιμές του l και του u λαμβάνονται με τον υπολογισμό της πιθανότητας ότι τα p από n μελλοντικά δείγματα βασίζονται στο ιστορικό δείγμα X_r , το οποίο μπορεί να γραφτεί ως

$$P(X_r < Y_p \leq \dots \leq Y_n) = \binom{p+m-r}{p} / \binom{n+m}{n}$$

Με την επίλυση της

$$\sum_{k=p}^n P(X_r < Y_k \leq \dots \leq Y_n) \geq (1 - \alpha),$$

η τιμή του r μπορεί να βρεθεί. Για αμφίπλευρα διαστήματα, το r διχοτομείται σε ένα κατώτερο $X_l = X_{r/2}$ και ένα ανώτερο μέρος $X_u = X_{m-\frac{r}{2}+1}$. Τα

μονόπλευρα διαστήματα πρόβλεψης κατασκευάζονται χρησιμοποιώντας το r εξ ολοκλήρου για το κατώτερο όριο $X_l = X_r$ αναλόγως για το ανώτερο όριο $X_u = X_{m-r+1}$. Η ίδια τεχνική χρησιμοποιείται για τα μεσαία διαγράμματα, αποκαλούμενα μη-παραμετρικά διαγράμματα ελέγχου προτεραιότητας. Η πιθανότητα των r ιστορικών παρατηρήσεων που είναι μικρότερα από το Y_p υπολογίζονται από τη σχέση

$$P(X_1 \leq \dots \leq X_r \leq Y_p) = \binom{p+r-1}{r} \binom{n+m-p-r}{m-r} / \binom{n+m}{m}$$

Εάν το p ορίζεται ως ο δείκτης της διαμέσου των μελλοντικών δειγμάτων, τα όρια για το διάστημα προτεραιότητας ώστε να περιέχει το μελλοντικό μέσο $[X_l, X_u]$ υπολογίζεται μέσω της σχέσης

$$\sum_{k=p}^n P(X_r < Y_k \leq \dots \leq Y_n) \geq (1 - \alpha)$$

Τα μονόπλευρα όρια πρόβλεψης βρίσκονται εύκολα με το διπλασιασμό της πιθανότητας ουράς σε $(1 - 2\alpha)$.

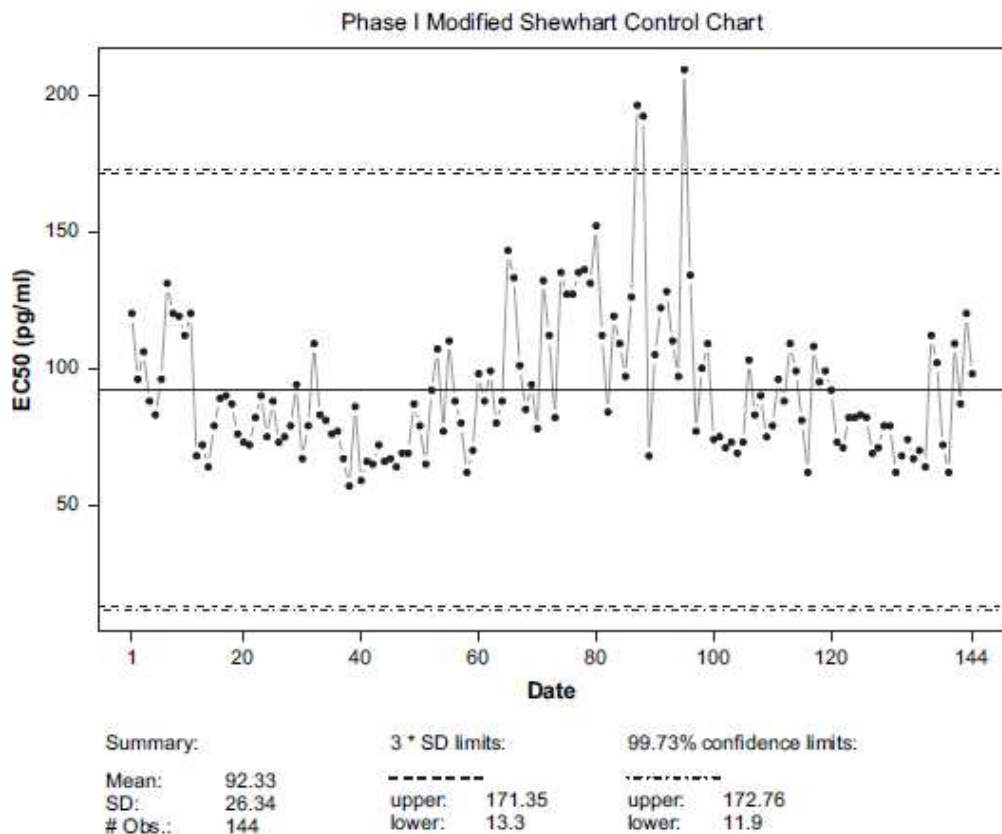
Βασιζόμενοι στην κατανομή των στοιχείων, ο αριθμός των m ιστορικών τρεξιμάτων και ο αριθμός των n επαναλαμβανόμενων τρεξιμάτων, είτε για τα παραμετρικά είτε για τα μη-παραμετρικά διαστήματα πρόβλεψης είναι κατάλληλα.

4. Μια περιπτωσιολογική μελέτη

Τα στοιχεία της δραστηριότητας που παρουσιάζονται εδώ προέρχονται από μια επικυρωμένη βιολογική ανάλυση. Σε αυτή, μια γραμμή κυττάρων επωάζεται με τα τιτλοδοτημένα ποσά μιας βιοφαρμακευτικής αγωγής, και τελικά μια καμπύλη δόσης-αντίδρασης του κυψελοειδούς πολλαπλασιασμού λαμβάνεται. Για κάθε δείγμα οι πρότυπες οκτώ συγκεντρώσεις της βιοφαρμακευτικής αναλύθηκαν εις τριπλούν. Μέσα στους πρώτους εννέα μήνες (ιστορική

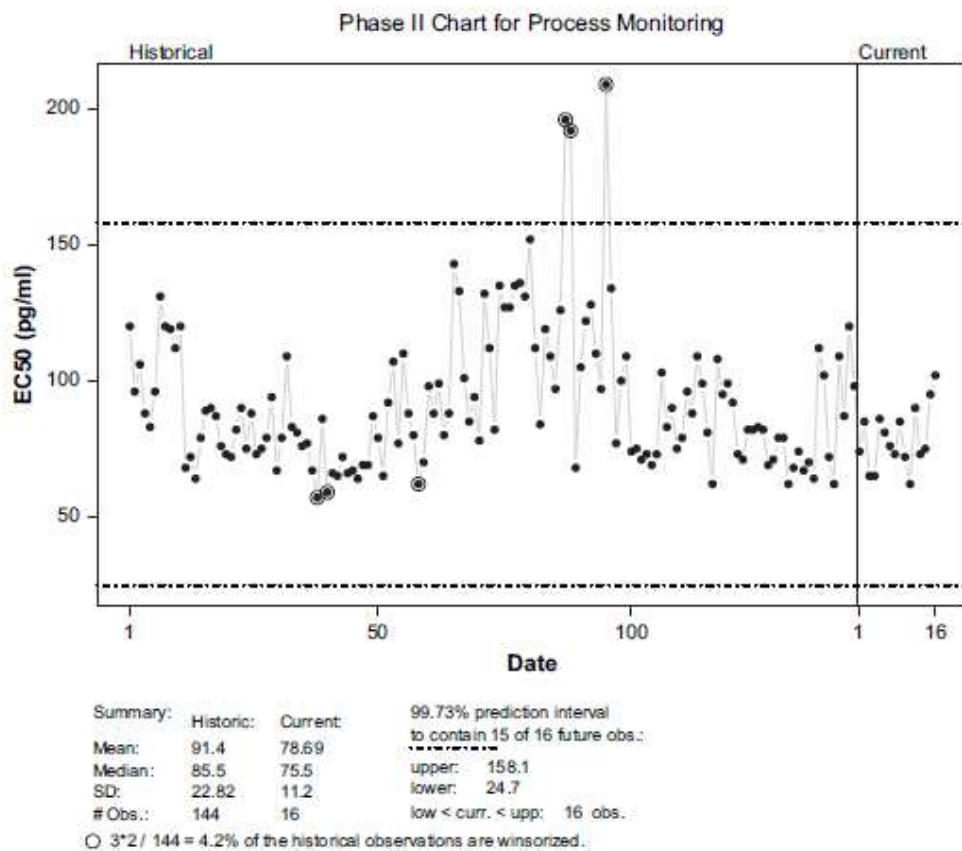
περίοδος), 144 εκτιμήσεις δραστηριότητας των προτύπων των συγκεντρώσεων λήφθηκαν. Στο διάγραμμα 3, η κοινός-χρησιμοποιημένη φάση I του διαγράμματος ελέγχου τύπου Shewhart για 3σ όρια και 99.73 όρια εμπιστοσύνης παρουσιάζεται. Το διάγραμμα αυτό δεν χρησιμοποιείται για την απόφαση, αλλά ακριβώς για το χαρακτηρισμό των ιστορικών δεδομένων.

Διάγραμμα 3: Η φάση I του διαγράμματος ελέγχου των δεδομένων του παραδείγματος



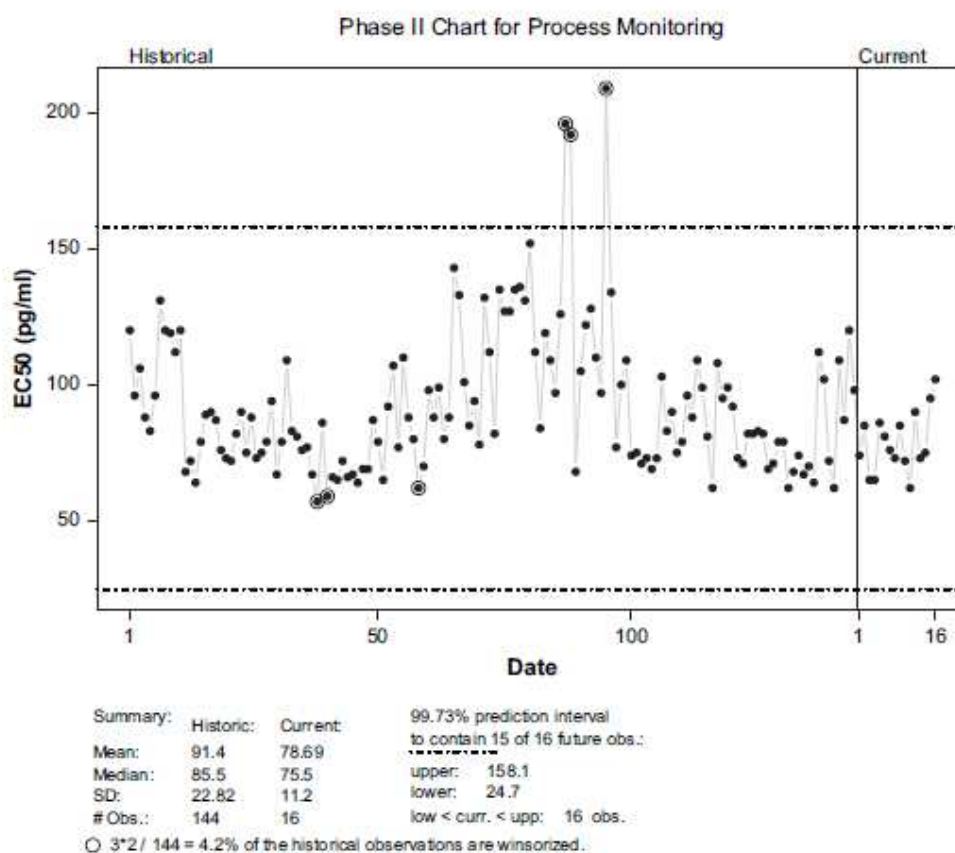
Το QQ-διάγραμμα του διαγράμματος 1 δείχνει την παραβίαση της κανονικής κατανομής, και επομένως η μη παραμετρική προσέγγιση χρησιμοποιήθηκε για τον προσδιορισμό των ορίων εμπιστοσύνης. Στο διάγραμμα 4 το αμφίπλευρο 99.73% μη παραμετρικό διάστημα πρόβλεψης περιέχει τουλάχιστον το $\kappa = 13$ από $n = 16$ μελλοντικά δείγματα.

Διάγραμμα 4: Η φάση II του διαγράμματος ελέγχου των δεδομένων του παραδείγματος



Σαφώς μπορεί να δει κάποιος, ότι δεν είναι απαραίτητη η αλλαγή του δείγματος αναφοράς, επειδή οι $n = 16$ επανελέγχόμενες τιμές τρεξίματος είναι με το κατ' εκτίμηση διάστημα. Πρέπει να σημειωθεί, ότι λόγω του μικρού αριθμού των $n = 16$ μη παραμετρικών διαστημάτων πρόβλεψης, για $k > 13$ μέχρι το $k = n = 16$ δεν μπορούν να υπολογιστούν διαστήματα πρόβλεψης. Επιπλέον, το παραμετρικό διάστημα πρόβλεψης σχεδιάζεται στο διάγραμμα ως εναλλακτική προσέγγιση. Σαφώς η διαφορά στα μη παραμετρικά διαστήματα μπορεί να φανεί, επειδή τα παραμετρικά διάστημα είναι συμμετρικά στον ιστορικό μέσο όρο. Για τα μη κανονικά κατανομημένα στοιχεία, η τροποποίηση των ακραίων σημείων (winsorising) αποκαλύπτει μια καλή εναλλακτική λύση, επειδή οποιοσδήποτε αριθμός τιμών του k μέχρι το $k = n$ είναι δυνατός, ακόμη και για τα μικρά μεγέθη δειγμάτων. Η περίπτωση για $k = 15$ από $n = 16$ παρουσιάζεται στο διάγραμμα 5 μέσω της (3,3) τροποποίησης (winsorization) των ιστορικών στοιχείων τρεξιμάτων.

Διάγραμμα 5: Η φάση II του διαγράμματος ελέγχου των δεδομένων του παραδείγματος – τροποποιημένα διαστήματα



5. Μια εναλλακτική αμφίπλευρη σύγκριση

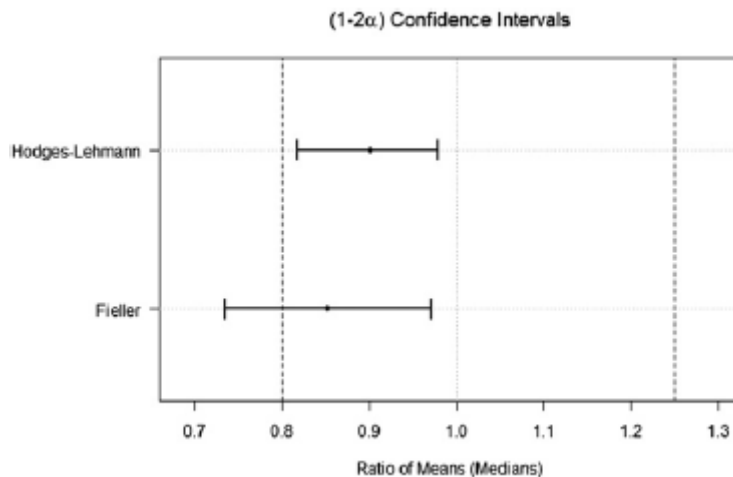
Η χρήση των παραμετρικών και στην ουσία των μη παραμετρικών διαστημάτων πρόβλεψης είναι προβληματική για τα πραγματικά στοιχεία ενός μικρού αριθμού ιστορικών τρεξιμάτων m και/ή των n επανελέγχων. Μια απλή εναλλακτική λύση είναι να καταδειχθεί η ισοδυναμία μεταξύ των δύο δειγμάτων «ιστορικά τρεξιματα» και «επανελέγχος των τρεξιμάτων» ως τυχαίο σχέδιο δύο ομάδων. Συνήθως χρησιμοποιείται και καταλλήλως για την ερμηνεία η αναλογία $\bar{X}_{re-test} / \bar{X}_{historical}$.

Η ισοδυναμία για την παράμετρο θέσης μπορεί να απαιτηθεί εάν το αμφίπλευρο $(1 - 2\alpha)$ διάστημα εμπιστοσύνης για την αναλογία είναι μέσα στα προκαθορισμένα πολλαπλασιαστικά κατώτατων ορίων $[\eta_l, \eta_u]$. Τέτοια κατώτατα όρια δεν υπάρχουν για τις εκτιμήσεις βιολογικής ανάλυσης. Είτε το γνωστό $[0.80, 1.25]$ διάστημα που προτείνεται από το FDA για τον ορό AUC των γενικών φαρμάκων ή των επεκταμένων κατώτατων ορίων $[0: 70, 1:43]$ για τις κινητικές παραμέτρους όπως τα όρια C_{max} . Εντούτοις, αυτά τα περιθώρια επιλέγονται αυθαίρετα. Για τα πρακτικά στοιχεία βιολογικής ανάλυσης αυτά τα περιθώρια θα επιλεχτούν βάση της ικανότητας για τη χρήση δηλαδή βασισμένα σε μια διαφορά που ασκεί επίδραση στη χρήση της βιολογικής ανάλυσης στη καθημερινότητα. Τα σχετικά διαστήματα εμπιστοσύνης μπορούν να υπολογιστούν υποθέτοντας την κανονική κατανομή σύμφωνα με τον Fieller ή

για οποιασδήποτε κατανομής μέσω μιας μη παραμετρικής προσέγγισης Hodges-Lehman.

Και οι δύο προσεγγίσεις είναι αριθμητικά διαθέσιμες στην R-βιβλιοθήκη. Το διάγραμμα 6 παρουσιάζει και τις δύο προσεγγίσεις για τα στοιχεία του παραδείγματος.

Διάγραμμα 6: Η ισοδυναμία των σχετικών διαστημάτων εμπιστοσύνης



6. Συμπεράσματα

Τα παραμετρικά και μη παραμετρικά διαστήματα πρόβλεψη κατά την φάση II βασίζονται στα συμπεράσματα του διαγράμματος ελέγχου για τα επαναλαμβανόμενα στοιχεία βιολογικής ανάλυσης. Η απόφαση, αν τα παραμετρικά ή τα μη παραμετρικά διαστήματα πρέπει να χρησιμοποιηθούν, βασίζεται στο QQ-διάγραμμα και τις δοκιμές της κανονικότητας. Όταν η υπόθεση κανονικότητας παραβιάζεται, τα μη παραμετρικά διαστήματα συστήνονται. Όσα περισσότερα m ιστορικά τρεξίματα είναι διαθέσιμα, τόσο πιο διασφαλισμένη είναι η επιλογή χρήσης ή μη της κανονική κατανομής. Εκτιμήσεις μεγέθους δειγμάτων, δηλαδή ο ελάχιστος αριθμός των m ιστορικών τρεξιμάτων και των n μελλοντικών επανελέγχων των τρεξιμάτων, είναι αρκετά δύσκολοι. Για τη μη παραμετρική προσέγγιση, το m πρέπει να είναι σαφώς μεγαλύτερο από 50 (βλέπε διάγραμμα 2). Ομοίως ο αριθμός των n μελλοντικών επανελέγχων των τρεξιμάτων πρέπει να είναι αρκετά μεγάλος, δηλαδή τουλάχιστον πάνω από 20 για να χαρακτηρίσει μια ακριβή διαδικασία. Η τροποποίηση (winsorization) του παραμετρικού διαστήματος μπορεί να είναι μια εναλλακτική λύση στην περίπτωση μικρών δειγμάτων. Τέλος οι συναρτήσεις για τον υπολογισμό των παραμετρικών και μη παραμετρικών διαστημάτων πρόβλεψης εφαρμόζονται στην R-βιβλιοθήκη.

Η χρήση της μεθόδου του διαγράμματος ποιότητας ελέγχου για την ανάλυση της μεταβλητότητας του PEF του επαγγελματικού άσθματος

1. Εισαγωγή

Το άσθμα είναι η πιο κοινή επαγγελματική αναπνευστική αναταραχή στο βιομηχανοποιημένο κόσμο, [Cartier (1994), Chan- Yeung και Malo (1995), Venables και Chan-Yeung (1997)]. Υπολογίζεται ότι οι επαγγελματικοί παράγοντες επηρεάζουν περίπου 1 στις 10 περιπτώσεις ενήλικου άσθματος, συμπεριλαμβανομένης της εκ του μη όντως εκδήλωσης άσθματος αλλά και της εκδήλωσης προϋπάρχοντος άσθματος (Blanc και Toren, 1999). Το κόστος του επαγγελματικού άσθματος έχει αυξηθεί σταθερά κατά τη διάρκεια των τελευταίων 20 ετών. Η μελέτη του ποσοστού δαπανών και της θνησιμότητας του επαγγελματικού άσθματος για το έτος 1996 υπολόγισε 805 θανάτους λόγω του επαγγελματικού άσθματος κάθε έτος στις ΗΠΑ ανερχόμενο το συνολικό κόστος λόγω του επαγγελματικού άσθματος σε \$1.6 δισεκατομμύρια (Leigh 2002). Το επαγγελματικό άσθμα φανερώνεται από την προοδευτική επιδείνωση της λειτουργίας πνευμόνων κατά τη διάρκεια της εβδομάδας εργασίας επομένως, ο σκοπός της διάγνωσης είναι να καθιερωθεί η σχέση μεταξύ της εργασίας και του άσθματος. Είναι κρίσιμο να καθιερωθεί μια αντικειμενική σχέση μεταξύ των συμπτωμάτων του άσθματος και της έκθεσης του προσώπου στον εργασιακό του χώρο. Διάφοροι ερευνητές έχουν προτείνει την τμηματική μέτρηση της μέγιστης εκπνευστικής ροής (PEF) στον τόπο εργασίας και μακριά από τον τόπο εργασίας ως αντικειμενικό μέτρο και κατάλληλο πρώτο βήμα στην καθιέρωση μιας αιτιώδους σχέσης ή την επιβεβαίωση μιας σχέσης μεταξύ άσθματος και έκθεσης εργασίας [Chan-Yeung (1990), Malo (1993), Cartier (1994) και Burge (2000)]. Το PEF είναι η μέγιστη ροή αέρα που μπορεί να επιτευχθεί κατά τη διάρκεια μιας ξαφνικής αναγκασμένης εκπνοής από μια κατάσταση πλήρους εισπνοής. Οι μετρήσεις πρέπει να πραγματοποιούνται για έναν παρατεταμένο χρόνο, και όσο το δυνατόν συχνότερα κατά τη διάρκεια της ημέρας ώστε να ανιχνευτεί οποιαδήποτε ελλοχεύουσα μεταβλητότητα. Ο Walter A. Shewhart ανέπτυξε τα διαγράμματα ελέγχου στη δεκαετία του '20 στα εργαστήρια της Bell και δημοσίευσε την εργασία του στο βιβλίο *‘Οικονομικός έλεγχος της ποιότητας των κατασκευασμένων προϊόντων’* το 1931.

Μια περιεκτική βιβλιογραφική έρευνα οδήγησε σε διάφορες δημοσιεύσεις σχετικά με την εφαρμογή των διαγραμμάτων ποιοτικού ελέγχου στις κλινικές και διαδικασίες υγειονομικής περίθαλψης επιγραμματικά αναφέρουμε

- διαγράμματα ποιοτικού ελέγχου στις εργαστηριακές διαδικασίες [Keorke (1981), Quam (1985) και Linford (1997)]
- διοικητικές διαδικασίες υγειονομικής περίθαλψης [Schwab (1999), Coontz (2000) και Benneyan (2003)]
- ουρική ακράτεια (Fogerty και Schnelle, 1989)
- στεφανιαίες προεγχειριτικές παρακάμψεις νοσηρότητα και θνησιμότητα (Shahian, 1996)
- εκβάσεις ψυχοθεραπείας [Green (1999), Pfadt και Wheeler (1995)]
- παράκαμψη στεφανιαίων αρτηριών που μπολιάζει τη θνησιμότητα (Levett και Carey, 1999)
- παραλλαγή στο αντιβιοτικό που ορίζει μεταξύ δύο ιατρών παθολόγων (Marshall και Mohammed, 2003).

- Στην αντιμετώπιση του άσθματος, ο Boggs (1998, 1999) έχει διευθύνει την περιεκτικότερη εφαρμογή των διαγραμμάτων ελέγχου στην ανάλυση των κλινικών στοιχείων και τον έλεγχο των ασθενών με το άσθμα
- Ο Gibson (1995) χρησιμοποίησε PEF διάγραμμα ελέγχου για να καθοδηγήσει τη θεραπεία αντιμετώπισης του άσθματος
- Ο Hayati (2006) χρησιμοποίησε καθημερινές ελάχιστες τιμές PEF για να ερευνήσει την αποτελεσματικότητα των διαγραμμάτων ελέγχου στην ανίχνευση και τον έλεγχο του επαγγελματικού άσθματος.

2. Εργαλεία και μέθοδοι

Μια ομάδα ερευνητών πραγματοποίησε μια μελέτη στον Καναδά για να βελτιώσει τις μεθόδους για τη διάγνωση του επαγγελματικού άσθματος. Οι εργαζόμενοι που περιλήφθηκαν σε αυτήν την μελέτη αναφέρθηκαν σε τέσσερα καναδικά κέντρα: Νοσοκομείο, Sacré-Cœur, Μόντρεαλ, Κεμπέκ, Νοσοκομείο Laval, Sainte-Foy, Κεμπέκ, Νοσοκομείο St. Joseph's, Χάμιλτον, Οντάριο και Νοσοκομείο TWH/GOEHU, Τορόντο. Μεταξύ των άλλων στόχων, ένας σκοπός της έρευνάς τους ήταν να βελτιωθεί η ευαισθησία και η ιδιομορφία της διάγνωσης του επαγγελματικού άσθματος χρησιμοποιώντας το δείκτη PEF. Επιπροσθέτως των PEF μετρήσεων, η πληθυσμογραφία σώματος (ευαίσθητη μέτρηση των πνευμόνων για την ανίχνευση παθολογίας των πνευμόνων), η επαγωγή πτυέλου και οι δοκιμές εισπνοής (χρυσά πρότυπα) πραγματοποιήθηκαν. Οι παθολόγοι που ειδικεύονται στην επαγγελματική αναπνευστική υγεία στις κλινικές εξωτερικών ιατρείων των νοσοκομείων εξέτασαν αρχικά τα θέματα. Εντούτοις, κατά την έναρξη της μελέτης, η συγκεκριμένη δοκιμή εισπνοής πραγματοποιήθηκε για να επιβεβαιωθεί η συγγένεια εργασίας και άσθματος. Αυτό το τεστ περιλαμβάνει τη συμμετοχή ενός υποψηφίου επαγγελματικού πράκτορα κάτω από ασφαλείς και ελεγχόμενες συνθήκες. Οι συγκεκριμένες προκλήσεις εισπνοής (SIC) εκτελέστηκαν όπως περιγράφονται από τον Cartier (1994), χρησιμοποιώντας μια κλειστού κυκλώματος συσκευή, με έναν ρεαλιστικό τρόπο, που αναπαράγει όσο το δυνατόν περισσότερο τους όρους που αντιμετωπίστηκαν στον εργασιακό χώρο. Η διάρκεια της έκθεσης στον προσβάλλοντα πράκτορα ήταν μεταβλητή και διάρκεσε είτε μέχρι το περιστατικό μιας ασθματικής επιπλοκής είτε μέχρι και 4 ώρες. Όταν το SIC ήταν αρνητικό στο εργαστήριο, τότε δεν μπόρεσε να προσδιοριστεί ο προσβάλλοντα πράκτορας ή εάν οι όροι στην εργασία δεν θα μπορούσαν να αναπαραχθούν στο εργαστήριο, το SIC εκτελέστηκε στον εργασιακό χώρο. Το SIC θεωρήθηκε θετικό όταν τεκμηριώθηκε μια συνεχής πτώση στο FEV1 που έφθασε σε τουλάχιστον 20% από την τιμή των αντικειμενικών γραμμών.

Τα συμπεράσματα της καναδικής μελέτης αναφέρθηκαν από το Girard (2004). Οι μετρήσεις PEF των θεμάτων και τα αποτελέσματα της δοκιμής SIC για όλα τα θέματα παρέχονται με σκοπό τη μελέτη αυτή. Η μελέτη εγκρίθηκε από την επιτροπή ερευνητικής ηθικής και το θεσμικό ερευνητικό καταστατικό κάθε συμμετέχοντος κέντρου καθώς όλοι έδωσαν τη γραπτή συγκατάθεσή τους.

Οι τιμές PEF των 45 θεμάτων αναθεωρήθηκαν και αξιολογήθηκαν με σκοπό αυτήν την μελέτη. Οι τιμές PEF συλλέχθηκαν χρησιμοποιώντας έναν μικρό μετρητή ροής. Τα θέματα ήταν εξοικειωμένα με χρήση του μετρητή καθώς κλήθηκαν να καταγράψουν το PEF τέσσερις φορές καθημερινά, τρεις καταγραφές σε κάθε περίπτωση καθώς και η καταγραφή της καλύτερης εκ των τριών τιμών, για μια

περίοδο 2 εβδομάδων στην εργασία και 2 εβδομάδων μακριά από την εργασία. Το επαγγελματικό άσθμα φανερώνεται από την προοδευτική επιδείνωση της λειτουργίας των πνευμόνων κατά τη διάρκεια της εργασίας (Burge, 1993). Επομένως, για να ανιχνεύσουν αντικειμενικά οποιοσδήποτε λειτουργικές αλλαγές των πνευμόνων, οι μετρήσεις PEF πρέπει να γίνουν και όταν είναι ο υπάλληλος στην εργασία και όταν είναι μακριά από την εργασία. Οι Bright και Burge (1996) συνέστησαν την ανάγνωση για μια περίοδο 2 εβδομάδων στην εργασία και 2 εβδομάδων μακριά από την εργασία. Η ομάδα εργασίας της ευρωπαϊκής αναπνευστικής κοινότητας (Quanjer 1997) πρότεινε ότι στην περίπτωση του επαγγελματικού άσθματος, μια ελάχιστη θετική προσαύξηση πρέπει να συμπεριληφθεί στις μετρήσεις PEF που λαμβάνονται για μία περίοδο 2 εβδομάδων στην εργασία και 2 Σαββατοκύριακα μακριά από την εργασία. Εννέα θέματα αποκλείστηκαν λόγω ποικίλων λόγων όπως οι ασυνέπειες στις καταγραφές PEF, που καταγράφουν μόνο μερικές ημέρες, καταγράφοντας μόνο δύο καθημερινές μετρήσεις. Μετά από την αναθεώρηση των καθημερινών καταγραφών PEF όλων των θεμάτων και στην εργασία και μακριά από την εργασία, χρονικές σειρές αναπτύχθηκαν με τέτοιο τρόπο ώστε να γίνει αντιληπτή η μέγιστη καθημερινή διακύμανση χρησιμοποιώντας τα συμπεράσματα του Boorsboom (1999). Όλα τα θέματα υπογράφηκαν και η μελέτη εγκρίθηκε από το θεσμικό συμβούλιο επιθεώρησης, Auburn University.

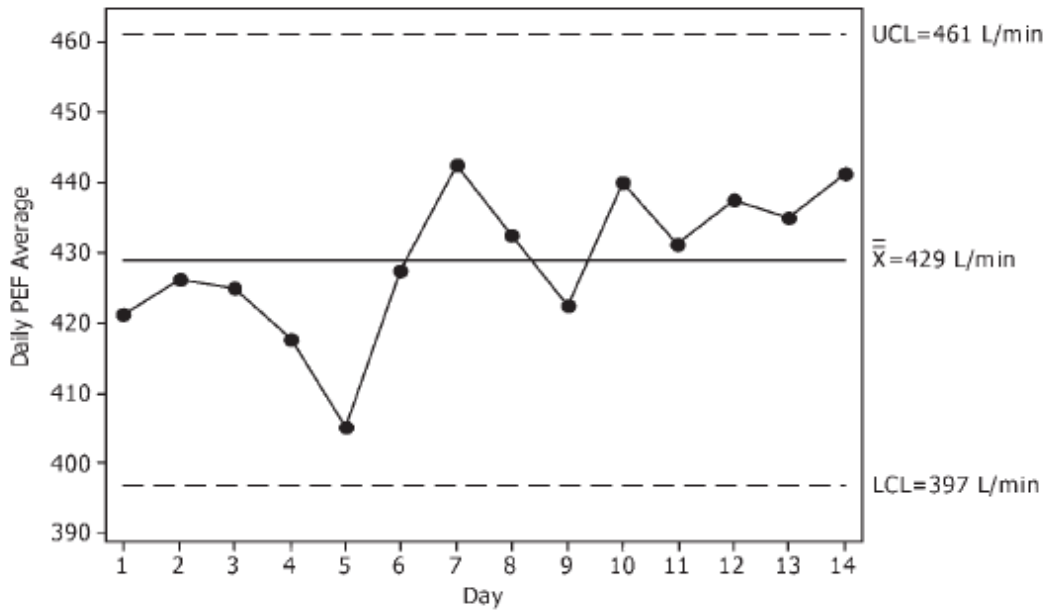
Ο πίνακας I παρουσιάζει καθημερινές μετρήσεις PEF του ενός υπαγόμενου και των καθημερινών μέσων όρων. Το καθημερινό εύρος, η διαφορά μεταξύ των υψηλότερων και χαμηλότερων τιμών, χρησιμοποιείται ως μέτρο της ημερήσιας παραλλαγής στην έρευνα άσθματος. Για τους υπόλοιπους 36 εργαζομένους, διαγράμματα ελέγχου τύπου Shewhart για τις μέσες τιμές κατασκευάστηκαν, ένα για την περίοδο εργασίας και ένα για την περίοδο μακριά από την εργασία.

Πίνακας 1: Μετρήσεις PEF, μέσοι όροι και το εύρος προερχόμενες κατά τη διάρκεια μακριά από την εργασία

Time	Day													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
7:00–8:00	400	420	380	380	400	430	400	450	420	420	430	450	420	420
11:00–13:00	430	430	450	410	400	440	490	450	420	470	410	440	440	470
16:00–17:00	420	425	440	420	410	430	440	410	440	440	455	440	460	450
21:00–22:00	435	430	430	460	410	410	440	420	410	430	430	420	420	425
Daily average	421.3	426.3	425	417.5	405	427.5	442.5	432.5	422.5	440	431.3	437.5	435	441.3
Daily range	35	10	70	80	10	30	90	40	30	50	45	30	40	50

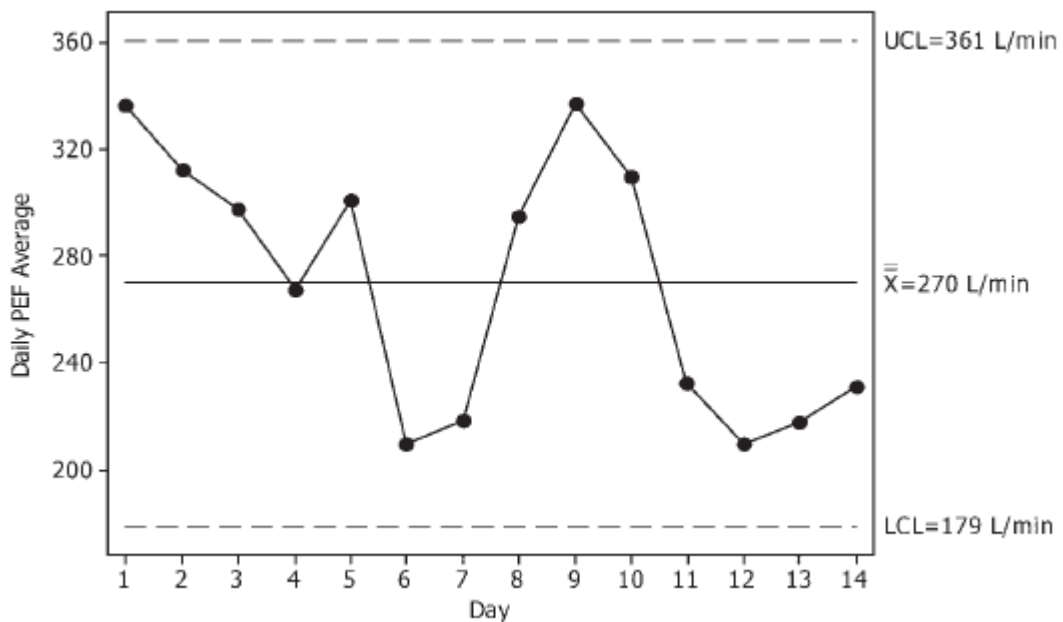
Το διάγραμμα 1 παρουσιάζει ένα PEF διάγραμμα ελέγχου ενός εργαζομένου κατά τη διάρκεια της περιόδου μακριά από την εργασία. Αυτή η περίοδος θεωρείται ως η αντικειμενική γραμμή του θέματος. Το διάγραμμα 1 αναφέρεται ως μέσο διάγραμμα ελέγχου και παρακολουθεί τη συμπεριφορά των καθημερινών μέσων όρων του PEF.

Διάγραμμα 1: Διάγραμμα ελέγχου ποιότητας για τις μετρήσεις PEF κατά την περίοδο μακριά από την δουλειά



Το διάγραμμα 2 παρουσιάζει το διάγραμμα ελέγχου PEF του ίδιου εργαζομένου κατά την περίοδο εργασίας.

Διάγραμμα 2: Διάγραμμα ελέγχου ποιότητας για τις μετρήσεις PEF κατά την περίοδο της εργασίας



Το διάγραμμα 1 παρουσιάζει το διάγραμμα ελέγχου του καθημερινού μέσου όρου PEF μακριά από τον εργασιακό χώρο (βασική γραμμή) και το διάγραμμα 2 παρουσιάζει το διάγραμμα ελέγχου του καθημερινού μέσου όρου PEF κατά τη διάρκεια της περιόδου εργασίας. Η περιοχή που βρίσκεται μεταξύ των ανώτερων

και χαμηλότερων ορίων ελέγχου αντιπροσωπεύει το μέγεθος της παρούσας μεταβλητότητας στη διαδικασία λειτουργίας των πνευμόνων. Δεδομένου ότι τα όρια τοποθετούνται σε απόσταση τρία-σίγμα επάνω και κάτω από το μέσο όρο η μεταβλητότητα μπορεί να εκφραστεί ως έξι-σίγμα. Για το τεστ αναλογίας το μέγεθος αυτής της μεταβλητότητας μπορεί να υπολογιστεί μέσω της σχέσης

$$\begin{aligned} \text{Ratio test} &= \frac{\text{Μέγεθος μεταβλητότητας στην εργασία}}{\text{Μέγεθος μεταβλητότητας στην κεντρική γραμμή}} \\ &= \frac{6\sigma_{\bar{x}} \text{ (εργασία)}}{6\sigma_{\bar{x}} \text{ (Κεντρική γραμμή)}} \end{aligned}$$

Δεδομένου ότι η διασπορά της στατιστικής σειράς χρησιμοποιείται για να υπολογιστεί το σίγμα($\sigma_{\bar{x}}$), η σχέση μπορεί να απλοποιηθεί

$$\text{Ratio test} = \frac{6\sigma_{\bar{x}} \text{ (εργασία)}}{6\sigma_{\bar{x}} \text{ (Κεντρική γραμμή)}} = \frac{\bar{R}_{\text{εργασία}}/d_2\sqrt{n}}{\bar{R}_{\text{κεντρική γραμμή}}/d_2\sqrt{n}} = \frac{\bar{R}_{\text{εργασία}}}{\bar{R}_{\text{κεντρική γραμμή}}}$$

όπου το d_2 είναι μια σταθερά και το n είναι ο αριθμός των καθημερινών μετρήσεων PEF ή το μέγεθος του δείγματος. Δεδομένου ότι το \bar{R} αντιπροσωπεύει τη μέση ημερήσια μεταβλητότητα, η αναλογία λαμβάνεται απλά με τη διαίρεση του μέσου όρου των ημερήσιων μεταβολών στην εργασία προς το μέσο όρο των ημερήσιων μεταβολών μακριά από τον εργασιακό χώρο (βασική γραμμή).

Συμπερασματικά τρεις γραμμές απαιτούνται για να κατασκευαστεί ένα διάγραμμα ελέγχου: μια μέση γραμμή (\bar{X}), χαμηλότερο όριο ελέγχου (LCL) και ανώτερο όριο ελέγχου (UCL). Τα όρια τοποθετούνται σε μια απόσταση τρία σίγμα $3\sigma_{\bar{x}}$ επάνω και κάτω από τη μέση γραμμή, αντίστοιχα. Το σίγμα είναι ένα μέτρο της διασποράς. Ο Shewhart προτείνει τη χρήση των μέσων στατιστικών διασποράς για να υπολογίσει τα όρια ελέγχου. Εδώ οι μέσες στατιστικές διασποράς που χρησιμοποιούνται είναι το μέσο εύρος ή R και το $\sigma_{\bar{x}} = \bar{R}/d_2\sqrt{n}$, όπου το n είναι ο αριθμός των καθημερινών καταγραφών PEF ή το μέγεθος των δειγμάτων και d_2 είναι μια σταθερά. Κατά συνέπεια οι $3\sigma_{\bar{x}}$ τιμές προστίθενται και αφαιρούνται από τους μέσους όρους για να ληφθούν τα τρία-σίγμα όρια ελέγχου. Στο διάγραμμα 1, ο καθημερινός μέσος όρος PEF για τις 14 ημέρες μακριά από την εργασία είναι 429 L/min με τα όρια της μεταβλητότητας μεταξύ του 397 L/min και 461 L/min. Για κάθε θέμα, η σχετική αλλαγή στο καθημερινό PEF υπολογίστηκε ως η αναλογία του μέσου εύρους στην εργασία (\bar{R}_w) προς το μέσο εύρος της αντικειμενικής γραμμής (\bar{R}_B) εκφρασμένο ως ποσοστό, δηλαδή $(\bar{R}_w/\bar{R}_B) \times 100$.

3. Αποτελέσματα

Για κάθε θέμα η αναλογία της μέσης καθημερινής διασποράς στην εργασία με βάση την αντικειμενική γραμμή υπολογίστηκαν και εκφράστηκαν επί τοις εκατό. Αυτές οι αναλογίες εξετάστηκαν περισσότερο από 15% και 20% και συγκρίθηκαν με τα αποτελέσματα των χρυσών προτύπων για να καθιερώσουν την σχέση μεταξύ εργασίας και άσθματος. Η ευαισθησία, η ιδιομορφία, η θετική αξία πρόβλεψης (PPV), και η αρνητική αξία πρόβλεψης (NPV) υπολογίστηκαν για κάθε κατηγορία. Η ευαισθησία παρέχει τη δυνατότητα μέσω της αναλογίας να προσδιοριστούν σωστά

οι εργαζόμενοι που έχουν το επαγγελματικό άσθμα, ενώ η ιδιομορφία παρέχει τη δυνατότητα μέσω της αναλογίας να προσδιοριστεί σωστά η σχέση μη-εργασίας και του άσθματος. Το PPV είναι το ποσοστό εκείνων που ήταν θετικοί στα τεστ ότι έχουν πραγματικά το επαγγελματικό άσθμα ενώ NPV είναι το ποσοστό εκείνων που εξεταζόμενοι ήταν αρνητικοί και δεν έχουν επαγγελματικό άσθμα. Η τιμή της αναλογίας της ευαισθησίας και της ιδιομορφίας παρουσιάζονται στον πίνακα II.

Πίνακας 2: Τα αποτελέσματα του τεστ αναλογίας χρησιμοποιώντας το διάγραμμα ποιότητας ελέγχου

Ratio test	Gold standard				
	Positive	Negative	Total	Sensitivity	Specificity
$\frac{R_w}{R_B}$					
≥ 15%	17	7	24	94.44%	61.11%
< 15%	1	11	12	(17/18)	(11/18)
Total	18	18	36		
$\frac{R_w}{R_B}$					
≥ 20%	16	7	23	88.89%	61.11%
< 20%	2	11	13	(16/18)	(11/18)
Total	18	18	36		

Μια αναλογία μεγαλύτερη από 15% παράγει ευαισθησία 94.44% και ιδιομορφία 61.11% και μια αναλογία μεγαλύτερη από 20% παράγει ευαισθησία 89.89% χωρίς καμία αλλαγή στην ιδιομορφία. Σε αυτόν τον πληθυσμό μελέτης, η επικράτηση του επαγγελματικού άσθματος ήταν 50% (18/36). Το PPV σε αναλογία 15% ήταν 70.83% και το NPV ήταν 91.67%. Εντούτοις, εάν η επικράτηση του επαγγελματικού άσθματος στον καλυμμένο πληθυσμό αυξήθηκε σε 90% μέσω του προσδιορισμού των ιδιαίτερα υψηλού κινδύνου ατόμων πριν από τη δοκιμή, κατόπιν PPV δείχνει να αυξάνονται σε 93.3% ενώ NPV να μειώνονται σε 56.2%. Η ευαισθησία και η ιδιομορφία θα ήταν αμετάβλητες ανεξάρτητα από την επικράτηση του επαγγελματικού άσθματος στο δοκιμασμένο πληθυσμό.

4. Παρατηρήσεις

Έχει δειχθεί ότι το καθημερινό εύρος του PEF (που μετριέται ως υψηλότερη καθημερινή τιμή μείον τη χαμηλότερη καθημερινή τιμή) μεγαλύτερο από 20% πρέπει να χρησιμοποιηθεί για τη διάγνωση του άσθματος (Hetzl και Clark, 1980). Επίσης, η μεταβλητότητα του 15% PEF ενδιάμεσα της μέρας, έχει προταθεί να προσδιορίσει το άσθμα (Venables et al., 1984). Όταν το επαγγελματικό άσθμα υποψιάζεται, μια ουσιαστική αλλαγή στο PEF σε σχέση με την έκθεση στο χώρο εργασίας πρέπει να καταδειχθεί. Διάφοροι ερευνητές έχουν χρησιμοποιήσει PEF και

την ημερήσια μεταβλητότητα PEF για να ανιχνεύσουν ή να επιβεβαιώσουν το επαγγελματικό άσθμα. Η διάγνωση του επαγγελματικού άσθματος από τα καταγεγραμμένα PEF έχει αποδειχθεί για να έχει μια ευαισθησία 77% και μια ιδιομορφία 100% στους ηλεκτρονικούς εργαζομένους και την ευαισθησία 100% και ιδιομορφία 83% στους εργαζομένους ισοκυανικών αλάτων [Burge (1979, 1982)]. Η οπτική ανάλυση της σειράς PEF έχει παρουσιάσει μια ευαισθησία 81% και μια ιδιομορφία 74% όταν συγκρίνεται με τα αποτελέσματα των χρυσών προτύπων (Perrin, 1992). Μια ανάλυση της καθημερινής καταγραφής PEF κάθε 2 ώρες οδήγησε σε μια ευαισθησία 73% και μια ιδιομορφία 78% συγκρινόμενο με το ειδικό τεστ εισπνοής (Malo, 1993). Σε αυτό το άρθρο χρησιμοποιώντας τη μέθοδο του διαγράμματος ελέγχου και την ημερήσια μεταβλητότητα φαίνεται ότι η αναλογία της μέσης καθημερινής διακύμανσης στην εργασία στη μέση καθημερινή διακύμανση μακριά από την εργασία είναι μεγαλύτερη από 15% παράγει μια ευαισθησία 94.4% και μια ιδιομορφία 61.11%.

Οι επιπτώσεις των λαθών τύπου I και II, δηλαδή τα λανθασμένα θετικά και τα λανθασμένα αρνητικά αποτελέσματα, είναι ένας λόγος ανησυχίας. Η ευαισθησία μπορεί να εκφραστεί ως $1 - \beta$ (Type II) και ιδιομορφία ως $1 - \alpha$ (Type I). Για να αυξήσουμε την ιδιομορφία, το α πρέπει να μειωθεί κάτι που αυξάνει το β . Η ευαισθησία και η ιδιομορφία του τεστ αναλογίας βρίσκουν μια μέση λύση μεταξύ αυτών των δύο τύπων λαθών. Η ερώτηση προκύπτει όσο αναφορά το κόστος, διαπράττοντας αυτού του τύπου των λαθών. Σε μια αναλογία 15%, τα ψευδώς θετικά θα εμφανιστούν κατά 29.17% από εκείνους που ένδειξη όταν η επικράτηση του επαγγελματικού άσθματος μεταξύ εκείνων που δοκιμάστηκαν είναι 50%. Εντούτοις, εάν η επικράτηση ανέλθει σε 90%, τα ψευδώς θετικά θα μειωθούν σε μόνο 6.7% από αυτούς που εξετάστηκαν και είχαν θετική ένδειξη. Το χειρότερο σενάριο ψευδώς θετική τιμή είναι ότι το άσθμα του εργαζομένου να έχει προσδιοριστεί ως σχετικό με την εργασία του ενώ αυτό δεν είναι, ώστε ο εργαζόμενος να παροτρυνθεί να αλλάξει εργασία ενώ αυτό δεν είναι απαραίτητο. Εξετάζοντας τη χαμηλή ιδιομορφία των αποτελεσμάτων, επιπρόσθετη έρευνα πρέπει να διεξαχθεί για να επιβεβαιώσει τη διάγνωση του επαγγελματικού άσθματος. Τα ψευδώς αρνητικά αποτελέσματα των τεστ θα εμφανιστούν σε μόνο 8.33% από εκείνους που είχαν αρνητική ένδειξη όταν η επικράτηση του επαγγελματικού άσθματος είναι κατά 50%. Όταν το άσθμα ενός εργαζομένου δεν προσδιορίζεται ως σχετικό με την εργασία του, τότε ο εργαζόμενος μπορεί να έχει την άδεια να επιστρέψει στην εργασία. Εντούτοις, θα πρέπει να τους δίνεται η ευκαιρία συνέχισης των εξετάσεων και περεταίρω δοκιμών ώστε να διαπιστωθεί εάν τα συμπτώματα επιδεινώνονται.

Η οπτική ανάλυση των τμηματικών μετρήσεων PEF προτείνεται καθώς είναι πιο ευαίσθητη και πιο συγκεκριμένη από τη στατιστική ανάλυση, και είναι η συνιστώμενη μέθοδος ανάλυσης [Burge (1993), Perrin (1992) και Anees (2003)]. Κάνοντας μια αναζήτηση στη βιβλιογραφία βρέθηκαν δύο μέθοδοι τμηματικής ανάλυσης PEF που χρησιμοποιούνται συνήθως από τους ερευνητές.

Η μια είναι αυτή της οπτικής ανάλυσης και η άλλη η μέθοδος Oasys. Στη πρώτη προσέγγιση είναι μια χρονική σειρά όλων των μεμονωμένων μετρήσεων PEF και οι νοσοκομειακοί γιατροί καλούνται να κρίνουν την κατάσταση του άσθματος με μια υποκειμενική μέθοδο όπως επαγγελματική ή μη-επαγγελματική. Η δεύτερη προσέγγιση είναι επίσης ένα διάγραμμα μιας χρονικής σειράς για κάθε ημέρα του μέγιστου, του ελάχιστου και των μέσων τιμών PEF κάθε ημέρας. Αυτή η προσέγγιση

προσδιορίζει την καθημερινή ημερήσια μεταβλητότητα και προτείνει ότι ο καθημερινός μέσος όρος παρέχει ένα ισχυρό μέτρο της μέγιστης ροής για εκείνη την ημέρα (Burge, 1993).

Εξετάζοντας το τεράστιο οικονομικό κόστος του επαγγελματικού άσθματος και του αντίκτυπού της στην επιχείρηση και τη βιομηχανία (Leigh, 2002), των νομικών επιπτώσεων και του δημόσιου ενδιαφέροντος για τα ζητήματα των εργασιακών χώρων, οι μέθοδοι των διαγραμμάτων ελέγχου που παρουσιάζονται σε αυτό το άρθρο προσφέρουν μια απλή και αποτελεσματική μέθοδο για την πρόωρη επέμβαση στους εργαζομένους πιθανού επαγγελματικού άσθματος. Η διεύθυνση πρέπει να μελετήσει τις μεθόδους αυτές κάνοντας μια πρόσκληση ενδιαφέροντος προς όλη την επιχείρηση με σκοπό την βελτίωση της υγείας και της ασφάλειας των εργαζομένων. Στις περισσότερες επιχειρήσεις αυτές οι μέθοδοι μπορούν να εφαρμοστούν και να ρυθμιστούν από τους επαγγελματίες υγείας και ασφάλειας με ένα πολύ λογικό κόστος. Οι ηλεκτρονικές συσκευές ελέγχου PEF μπορούν να αγοραστούν με ελάχιστο κόστος και η εισαγωγή δεδομένων και η σχεδίαση μπορεί να αυτοματοποιηθεί.

Ένας από τους περιορισμούς αυτής της μελέτης είναι το μικρό μέγεθος δειγμάτων ($n < 50$). Εντούτοις, το ενδιαφέρον χαρακτηριστικό γνώρισμα αυτής της μελέτης είναι η γενικότητα του επιλεγέντος πληθυσμού. Τα θέματα σε αυτήν την μελέτη αντιπροσωπεύουν διαφορετικές βιομηχανίες, γεωγραφικές θέσεις και περιβάλλοντα εργασίας. Επαρχίες του Καναδά όπως το Κεμπέκ και το Οντάριο. Γραφεία διοίκησης, πάτωμα εργοστασίων, ξύλο, βιομηχανίες χημικών ουσιών και τροφίμων, και όργανα υγειονομικής περίθαλψης. Οι μέθοδοι που παρουσιάζονται εδώ απαιτούν τις απλές αριθμητικές διαδικασίες και επομένως είναι πολύ απλές να χρησιμοποιηθούν. Αν και δεν είναι απαραίτητο, υπάρχουν διάφορα πακέτα λογισμικού διαθέσιμα για να παραγάγουν τα διαγράμματα ελέγχου με ελάχιστη προσπάθεια. Το διάγραμμα ποιότητας ελέγχου και το τεστ της αναλογίας που παρουσιάζεται σε αυτό το άρθρο είναι ενδεχομένως μια πολύ αποτελεσματική μέθοδος για να ελεχθεί η μεταβλητότητα του δείκτη PEF στο επαγγελματικό άσθμα. Τέλος, η έρευνα μέσω της εφαρμογής των διαγραμμάτων ποιοτικού ελέγχου στο επαγγελματικό άσθμα είναι περιορισμένη. Τα διαγράμματα ελέγχου μπορούν να εφαρμοστούν σε οποιουδήποτε άλλους δείκτες όπως της παρεμπόδισης εναέριων διαδρόμων ή της ανάφλεξης. Η ελπίδα είναι ότι αυτή η εφαρμογή του διαγράμματος ποιοτικού ελέγχου στην ανάλυση των μετρήσεων λειτουργίας πνευμόνων στο επαγγελματικό άσθμα θα προσκαλέσει το μεγαλύτερο ενδιαφέρον για αυτόν τον τομέα της έρευνας.

5. Συμπεράσματα

Τα διαγράμματα ελέγχου χρησιμοποιήθηκαν για να ερευνηθεί και να κατανοηθεί η μεταβλητότητα του δείκτη PEF στο επαγγελματικό άσθμα ώστε να καθιερωθεί η σχέση της εργασίας με το άσθμα. Οι μέθοδοι είναι απλές και ενδεχομένως πολύ αποτελεσματικές. Το διάγραμμα ελέγχου μπορεί να είναι μια αποτελεσματική μέθοδος στην πρόωρη επέμβαση των εργαζομένων που υποψιάζονται για το επαγγελματικό άσθμα.

Βιβλιογραφία

Ξένη

- Aki, S., Hirano, K., (1989). Estimation of parameters in discrete distribution of order k , *Annals of the Institute of Statistical Mathematics*, **47**, 557-568.
- Albin, S.L., Friedman, D.J., (1992). Off-line quality-control in electronics assembly-identifying the critical problem, *IIE Transactions*, **24**, 58-65.
- Amin, R.W., Reynolds, M.R., Jr, Bakir, S.T. (1995). Nonparametric Quality Control Charts based on the Sign Statistic, *Communication in Statistics-Theory and Methods*, **24**, 1597-1623.
- Amin, R.W., Searcy A.J., (1991). A nonparametric Exponentially Weighted Moving Average Control Scheme, *Communication in Statistics-Simulation and Computation*, **20**, 1049-1072.
- Antzoulakos, D.L., Koutras, M.V. and Rakitzis, A.C., (2008). Start-up demonstration tests based on run and scan statistics, *Journal of Quality Technology*, (to appear).
- Arnold, B.C., (1977). Recurrence relations between expectations of functions of order statistics, *Scand. Actuar. J.*, **2**, 169-174.
- Arnold, B.C. , (1985). The sign test in current control, *Statist. Hefte.*, **26**, 253-262.
- Arnold, B.C. and Balakrishnan, N., (1989). Relations, Bounds and Approximations for Order Statistics, *Lecture Notes in Statistics*, **53**, Springer-Verlag, New York.
- Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N., (1992). *A First Course in Order Statistics*, John Wiley & Sons, New York.
- Bakir, S.T., (2004). A distribution-free Shewhart quality control chart based on signed-ranks, *Quality Engineering*, **16**, 613-623.
- Bakir, S.T., (2006). Distribution-free quality control charts based on signed-rank-like statistics, *Communication in Statistics-Theory and Methods*, **35**, 743-757.
- Bakir, S.T. and Reynolds, M.R., Jr., (1979). A nonparametric procedure for process control based on within-group ranking, *Technometrics*, **21**, 175-183.
- Balakrishnan, N. and Cohen, A.C., (1991). *Order Statistics and Inference: Estimation Methods*, Academic Press, Inc, New York.
- Balakrishnan, N. and Kocherlakota, S., (1986). On the moments of order statistics from doubly truncated logistic distribution, *Journal of Statistical Planning and Inference*, **13**, 117-129.
- Balakrishnan, N. and Koutras, M.V., (2001). *Runs and Scan with Applications*, John Wiley & Sons, New York.
- Balakrishnan, N., Triantafyllou, I.S. & Koutras, M.V. (2009). Nonparametric control charts based on runs and Wilcoxon-type rank-sum statistic, *Journal of Statistical Planning and Inference*, **139**, 3177-3192
- Balakrishnan, N., Triantafyllou, I.S. & Koutras, M.V. (2008c). A new control chart based on order statistics, In Proceedings of International Workshop on Applied Probability 2008, Paris.

- Balakrishnan, N., Triantafyllou, I.S. & Koutras, M.V. (2010) A distribution-free control chart based on order statistics, *Communications in Statistics - Theory and Methods*, **39: 20**, 3652 — 3677
- Chakraborti, S., Van Der Laan P. and Bakir S.T., (2001). Nonparametric control charts: An overview and some Results, *Journal of Quality Technology*, **33**, 304-315.
- Fazel H., Maghsoodloo, S., Devivo, M.J., Thomas, R.E., Lemiere C., (2008). *Quality control chart method for analyzing PEF variability in occupational asthma*.
- Guilbaud, O. (1983). *Nonparametric prediction intervals for sample medians in the general case*, *Journal of the American Statistical Association*, **78**, 937-941.
- Rogerson, P.A., (2006). Formulas for the design of CUSUM quality control charts, *Communication in Statistics-Theory and Methods*, **35**, 373-383.
- Shewhart, W.A., (1926). Quality control charts, *Bell Syst. Technical J.*, 593-603.
- Taghuchi, G., (1986). *Introduction to Quality Engineering*, Asian Productivity Organization, White Plains, New York.
- Hothorn, L.A., Gerhard, D. and Hofmann M., (2009). Parametric and non-parametric prediction intervals based phase II control charts for repeated bioassay data, *Biological*, **37** (5), 323-330.
- Triantafyllou, I.S., (2005). *Preservation of the attributes IFR/DFR*, MSc Thesis, Department of Statistics and Insurance Science, University of Piraeus, Greece.
- Yaschin, E., (1992). Analysis of Cusum and other Markov-type control schemes by using empirical distributions, *Technometrics*, **34**, 54-63.
- Zantek, P.F., (2006). Design of cumulative sum schemes for start-up processes and short runs, *Journal of Quality Technology*, **38**, 365-375.

Ελληνική

- Αντζουλάκος, Δ. (2006). *Στατιστικός Έλεγχος Ποιότητας, Πανεπιστημιακές Σημειώσεις, ΠΜΣ «Εφαρμοσμένη Στατιστική».*
- Δαμιανού, Χ.Χ. (1996). *Στατιστικός Έλεγχος Ποιότητας και Αξιοπιστία, Πανεπιστημιακές Σημειώσεις, Τμήμα Μαθηματικών, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.*
- Κουκουβίνος Χ., (2008). *Στατιστικός Έλεγχος ποιότητας.*
- Κούτρας, Μ. Β. (2001). *Εισαγωγή στη Συνδυαστική, Εκδόσεις Α.Σταμούλης .*
- Κούτρας, Μ. Β. (2002). *Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές: Μέρος Ι, Εκδόσεις Α. Σταμούλης .*
- Κούτρας, Μ. Β. (2004). *Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές: Μέρος ΙΙ, Εκδόσεις Α. Σταμούλης .*
- Κούτρας, Μάρκος Β. (2007). *Στατιστική Θεωρία αξιοπιστίας και έλεγχοι χρόνων ζωής, Πανεπιστημιακές Σημειώσεις, ΠΜΣ «Εφαρμοσμένη Στατιστική».*

- Κούτρας Μ. Β., Τριανταφύλλου Ι. Σ., Balakrishnan N., (2007). Ένα μη παραμετρικό διάγραμμα ελέγχου με χρήση διατεταγμένων δειγμάτων.
- Κυριακούσης, Α.Γ. (2008). Στατιστικές Μέθοδοι, Εκδόσεις Συμμετρία.
- Μπερσίμης, Σ. (2005). Θεωρία ροών επιτυχιών και εφαρμογές, Διδακτορική Διατριβή, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.
- Μπούτσικας, Μ. (2000). Συμβολή στη Θεωρία συναφών τυχαίων μεταβλητών, Διδακτορική Διατριβή, Τμήμα Μαθηματικών, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.
- Ξεναλάκη Ε. (2001). Μη Παραμετρική Στατιστική.
- Ξεκαλάκη Ε., (1994). Ειδικά θέματα Μη παραμετρικής Στατιστικής.
- Ρακιτζής, Α.Χ., (2008). Στατιστικός Έλεγχος Ποιότητας και Θεωρία Ροών και Σχηματισμών, Διδακτορική Διατριβή, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.
- Χαραλαμπίδης, Χ. Α., (1996). Συνδυαστική, Τεύχος 1, Εκδόσεις Συμμετρία, Αθήνα.