

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ



ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Διπλωματική Εργασία

ΔΕΙΚΤΕΣ ΙΚΑΝΟΤΗΤΑΣ
ΣΤΟ ΣΤΑΤΙΣΤΙΚΟ ΕΛΕΓΧΟ ΔΙΕΡΓΑΣΙΩΝ

ΠΕΤΡΟΥΛΙΑ ΜΑΡΙΑ

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

Αθήνα, Ιούλιος 2014

ΠΕΡΙΛΗΨΗ

Στο σύγχρονο ανταγωνιστικό περιβάλλον, προκειμένου να πετύχουμε τη συνεχόμενη βελτίωση της παραγωγικής διαδικασίας, απαιτείται η υψηλή ποιότητα των παραγόμενων προϊόντων της. Για αυτόν το λόγο, έχει επιβάλλει σε κάθε επιχείρηση να υιοθετήσει προγράμματα και να εφαρμόσει ειδικές μεθοδολογίες προκειμένου να παράγει προϊόντα που βρίσκονται εντός των ορίων προδιαγραφών. Επομένως, πολλές βιομηχανικές επιχειρήσεις χρησιμοποιούν ένα από τα πιο χρήσιμα εργαλεία του Στατιστικού Ποιοτικού Ελέγχου, που λέγονται δείκτες ικανότητας μιας διεργασίας.

Οι δείκτες ικανότητας μιας διεργασίας αποτελούν μια άμεση ένδειξη ότι η διεργασία είναι ικανή να παράγει προϊόντα εντός των ορίων προδιαγραφών. Ο λόγος που χρησιμοποιούνται ευρέως στις περισσότερες επιχειρήσεις είναι η ευκολία στη χρήση τους αλλά και στην κατανόησή τους. Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάλυση των σημαντικότερων δεικτών ικανότητας, η παρουσίαση των ιδιοτήτων τους αλλά και η σύγκρισή τους.

Στο πρώτο κεφάλαιο, γίνεται μια συνοπτική παρουσίαση του στατιστικού ελέγχου ποιότητας, και τι ακριβώς εννοούμε με την έννοια ικανότητα μιας διεργασίας.

Στο δεύτερο κεφάλαιο, περιγράφονται οι βασικές τεχνικές με τις οποίες μπορούμε να αναλύσουμε την ικανότητα μιας διεργασίας.

Στο τρίτο κεφάλαιο που είναι και το βασικό στην παρούσα εργασία, γίνεται αναλυτική περιγραφή του κάθε δείκτη ικανότητας ξεχωριστά, σε συνδυασμό με τρόπους εκτίμησης, υπολογισμό διαστημάτων εμπιστοσύνης και παρουσίαση ελέγχων υποθέσεων. Εν συνεχεία, αναφέρονται και τα μειονεκτήματα του κάθε δείκτη ικανότητας.

Τέλος στο τελευταίο κεφάλαιο, με τη βοήθεια των στατιστικών προγραμμάτων Statgraphics και Minitab, παρουσιάζεται μια εφαρμογή των δεικτών ικανότητας.

ABSTRACT

In today's competitive environment, in order to achieve the continuous improvement of the production process, we require high quality of its products. For this reason, it imposes each company to adopt programs and implement specific methodologies so as to produce products that are within the specification limits. Therefore, many industrial companies use one of the most useful tools of Statistical Quality Control, namely, the process capability indices.

Process capability indices are a direct indication that the process is capable of producing products within specification limits. The reason they are widely used in many companies is the ease of their use and understanding. The purpose of this thesis is the analysis of the most important capability indices, the presentation of their properties and their comparison.

The first chapter is an overview of statistical quality control, and what exactly we mean by the ability of a process.

The second chapter describes the basic techniques by which we can analyze the ability of a process.

In the third chapter, which is the key to this work, we provide a detailed description of each capability index separately, combined with estimation techniques, calculation of confidence intervals and hypothesis testing presentation. Thereafter, we report the disadvantages of each capability index.

Finally in the last chapter, with the use of statistical programs Statgraphics and Minitab, we present an application of capacity indices.

ΕΥΧΑΡΙΣΤΙΕΣ

Η εκπόνηση και η ολοκλήρωση της παρούσας διπλωματικής εργασίας πραγματοποιήθηκε υπό την επίβλεψη του Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου, κ. Χρήστο Κουκουβίνο, τον οποίο ευχαριστώ θερμά για την καθοδήγησή του σε όλη τη διάρκεια συγγραφής της εργασίας, καθώς και τη δυνατότητα που μου προσέφερε να ασχοληθώ με ένα ιδιαίτερα ενδιαφέρον θέμα.

Επίσης, θα ήθελα να ευχαριστήσω τον υποψήφιο διδάκτορα Ανδρουλάκη Εμμανουήλ, για την πολύτιμη βοήθεια και το συνεχές ενδιαφέρον του, ο οποίος δίνοντας ιδέες και χρήσιμες παρατηρήσεις συνέβαλε καθοριστικά σε όλη την πορεία της μελέτης μου.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την υποστήριξη, τα εφόδια που μου προσέφερε και την ενθάρρυνσή της στην ολοκλήρωση της προσπάθειάς μου.

ΠΕΡΙΕΧΟΜΕΝΑ

Σελίδα

ΠΕΡΙΛΗΨΗ.....	1
ABSTRACT.....	2
ΕΥΧΑΡΙΣΤΙΕΣ.....	3

ΚΕΦΑΛΑΙΑ

<u>ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ</u>	<u>6</u>
1.1 Γενικά.....	6
1.2 Διαγράμματα ελέγχου.....	9
1.3 Ικανότητα μιας διεργασίας.....	15
1.4 Χρήση ή μη των δεικτών ικανότητας μιας διεργασίας.....	18

ΚΕΦΑΛΑΙΟ 2: ΒΑΣΙΚΕΣ ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ

<u>ΤΗΣ ΙΚΑΝΟΤΗΤΑΣ ΜΙΑΣ ΔΙΕΡΓΑΣΙΑΣ</u>	<u>19</u>
2.1 Εισαγωγή.....	19
2.2 Ανάλυση της ικανότητας μιας διεργασίας με τη βοήθεια του ιστογράμματος.....	20
2.3 Ανάλυση της ικανότητας μιας διεργασίας με τη βοήθεια ενός διαγράμματος πιθανοτήτων.....	23

ΚΕΦΑΛΑΙΟ 3: ΔΕΙΚΤΕΣ ΙΚΑΝΟΤΗΤΑΣ ΔΙΕΡΓΑΣΙΩΝ

3.1 Ο δείκτης ικανότητας C_p	29
3.1.1 Ορισμός και ιδιότητες του δείκτη C_p	30
3.1.2 Εκτίμηση του δείκτη C_p	34
3.1.3 Διαστήματα εμπιστοσύνης του δείκτη C_p	37

3.1.4 Έλεγχοι υποθέσεων του δείκτη C_p	40
3.1.5 Αδυναμίες του δείκτη C_p	46
3.2 Ο δείκτης ικανότητας C_{pk}	46
3.2.1 Ορισμός και ιδιότητες του δείκτη ικανότητας C_{pk}	47
3.2.2 Εκτίμηση του δείκτη ικανότητας C_{pk}	53
3.2.3 Διαστήματα εμπιστοσύνης και έλεγχοι υποθέσεων του δείκτη ικανότητας C_{pk}	54
3.2.4 Αδυναμίες του δείκτη ικανότητας C_{pk}	56
3.3 Ο δείκτης ικανότητας C_{pm}	57
3.3.1 Ορισμός και ιδιότητες του δείκτη ικανότητας C_{pm}	58
3.3.2. Εκτίμηση του δείκτη ικανότητας C_{pm}	60
3.3.3. Διαστήματα εμπιστοσύνης του δείκτη ικανότητας C_{pm}	62
3.3.4 Έλεγχοι υποθέσεων για τον δείκτη ικανότητας C_{pm}	64
3.3.5 Παρόμοιοι δείκτες του δείκτη ικανότητας C_{pm}	65
3.3.6 Αδυναμίες του δείκτη ικανότητας C_{pm}	66
3.4 Ο δείκτης ικανότητας C_{pmk}	67
3.4.1 Ορισμός και χαρακτηριστικά του δείκτη C_{pmk}	67
3.4.2 Βασικές ιδιότητες του δείκτη C_{pmk}	68
3.4.3 Εκτίμηση του δείκτη C_{pmk}	69
3.4.4 Διαστήματα εμπιστοσύνης και έλεγχοι υποθέσεων του δείκτη C_{pmk}	70
3.4.5 Αδυναμίες του δείκτη C_{pmk}	71

ΚΕΦΑΛΑΙΟ 4: ΕΦΑΡΜΟΓΗ ΜΕ ΧΡΗΣΗ ΤΩΝ

ΣΤΑΤΙΣΤΙΚΩΝ ΠΑΚΕΤΩΝ STATGRAPHICS ΚΑΙ MINITAB 72

4.1 Περιγραφή των δεδομένων.....	72
4.2 Χρήση του πακέτου Statgraphics και αποτελέσματα.....	76
4.3 Χρήση του πακέτου Minitab και αποτελέσματα.....	80

ΒΙΒΛΙΟΓΡΑΦΙΑ 84

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Κάθε παραγωγική επιχείρηση προκειμένου να αναπτυχθεί αλλά και να πετύχει τη διαρκή βελτίωση της ποιότητας σε κάθε επίπεδο (Διοίκηση Ολικής Ποιότητας-Total Quality Management ή Total Quality Assurance), έχει ως σημαντικό εργαλείο το Στατιστικό Έλεγχο Ποιότητας (Statistical Quality Control). Ο Στατιστικός Έλεγχος Ποιότητας επηρεάζει τις αποφάσεις της επιχείρησης που έχουν σχέση με την παραγωγική διαδικασία, τις προδιαγραφές και τον έλεγχο των παραγόμενων προϊόντων. Για να μπορέσει η χρήση του Στατιστικού Ελέγχου Ποιότητας να είναι αποτελεσματική, θα πρέπει να ενσωματωθεί σε ένα πλαίσιο λειτουργίας και διοίκησης της επιχείρησης.

Σε κάθε παραγωγική διεργασία, όσο καλά σχεδιασμένη κι αν είναι κι ανεξάρτητα από το πόσο προσεκτικά επιβλέπεται και συντηρείται, πάντα θα υπάρχει μια μορφή μεταβλητότητας. Για να γίνει πιο κατανοητός ο όρος της φυσικής μεταβλητότητας, αρκεί να σκεφτούμε το παρακάτω παράδειγμα. Έστω δύο προϊόντα που παράγονται από την ίδια πρώτη ύλη, με τα ίδια μηχανήματα και με τους ίδιους χειριστές μηχανημάτων. Θα υπάρχει σίγουρα κάποιο μετρήσιμο μέγεθος που θα έχει διαφορετική τιμή για τα δύο αυτά προϊόντα. Δηλαδή, όσο καλά ρυθμισμένα κι αν είναι τα μηχανήματα, όσο ικανοί κι αν είναι οι χειριστές των μηχανημάτων, όσο ικανοποιητική κι αν είναι η πρώτη ύλη, ποτέ δύο παραγόμενα προϊόντα δε θα είναι ίδια. Αυτή η φυσική μεταβλητότητα οφείλεται σε πολλές μικρές αιτίες γνωστές και ως κοινές ή τυχαίες τιμές μεταβλητότητας (common or chance causes of variation). Στις διεργασίες στις οποίες παρατηρούνται μόνο φυσική μεταβλητότητα ονομάζονται

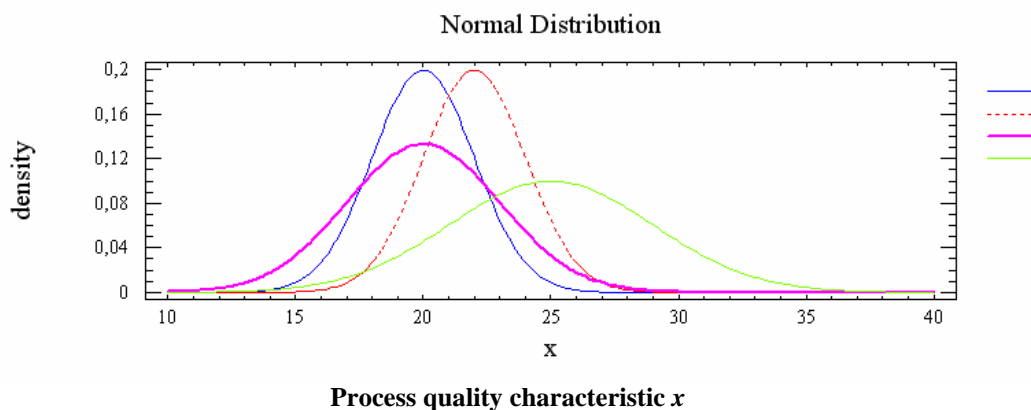
εντός ελέγχου διεργασίες (in control process), ή ότι η διεργασία λειτουργεί σε ευσταθή κατάσταση (stable state).

Να σημειώσουμε όμως, ότι υπάρχουν κάποιες μορφές μεταβλητότητας οι οποίες δεν οφείλονται σε τυχαίες τιμές όπως αναφερθήκαμε παραπάνω. Παρά το γεγονός ότι εμφανίζονται περιστασιακά, καθορίζουν την τελική ποιότητα του προϊόντος σε μεγάλο βαθμό. Αυτές οι μορφές μεταβλητότητας προέρχονται από λανθασμένα ρυθμισμένες μηχανές, κακή χρήση των μηχανημάτων από τους χρήστες, κακή ποιότητα ή ελαττωματική πρώτη ύλη. Η συγκεκριμένη μεταβλητότητα διαδραματίζει σημαντικό ρόλο στο επίπεδο κάποιου ή κάποιων ποιοτικών χαρακτηριστικών (quality characteristics) και για αυτό το λόγο έχουν καθοριστικό ρόλο στην ποιότητα του προϊόντος. Η εν λόγω μεταβλητότητα, αφενός έχει πολύ μεγαλύτερο μέγεθος από την φυσική μεταβλητότητα, αφετέρου όταν παρατηρηθεί σε μια παραγωγική διεργασία οδηγεί σε μη αποδεκτά λειτουργικά επίπεδα. Αυτή η μεταβλητότητα ονομάζεται ειδική μεταβλητότητα και οι αιτίες στις οποίες οφείλεται καλούνται ειδικές ή προσδιορισμένες αιτίες μεταβλητότητας (special or assignable causes of variation). Μια τέτοια παραγωγική διεργασία που παρουσιάζει ειδική μεταβλητότητα λέμε ότι είναι εκτός ελέγχου διεργασία (out of control process) ή αλλιώς ότι λειτουργεί σε ασταθή κατάσταση (unstable state).

Εκτός από τη φυσική και ειδική μεταβλητότητα, πολύ σημαντικό ρόλο διαδραματίζουν και τα όρια προδιαγραφών (specification limits) των ποιοτικών χαρακτηριστικών του προϊόντος, τα οποία καθορίζονται στη φάση σχεδιασμού του και στα οποία αξίζει να αναφερθούμε. Πρόκειται για το άνω και κάτω όριο προδιαγραφών (lower and upper specification limits, LSL and USL). Τα όρια αυτά εκφράζουν το διάστημα μέσα στο οποίο θα πρέπει να βρίσκονται οι τιμές του ποιοτικού χαρακτηριστικού για κάθε παραγόμενο προϊόν προκειμένου να θεωρείται αποδεκτό για την ποιότητά του. Πέρα από τα όρια προδιαγραφών των ποιοτικών χαρακτηριστικών του προϊόντος, στη φάση σχεδιασμού του προϊόντος ορίζεται και μια επιθυμητή τιμή για το ποιοτικό χαρακτηριστικό και είναι συνήθως το μέσον του διαστήματος των ορίων προδιαγραφών δηλαδή του διαστήματος [LSL,USL]. Αυτή η επιθυμητή τιμή αναφέρεται ως τιμή στόχος (target value, T). Αξίζει να σημειωθεί ότι κάτω από συνθήκες φυσικής μεταβλητότητας οι περισσότερες τιμές του ποιοτικού

χαρακτηριστικού στα παραγόμενα προϊόντα βρίσκεται εντός των ορίων προδιαγραφών, κάτι που δε συμβαίνει σε περιπτώσεις ειδικής μεταβλητότητας. Για να γίνουν κατανοητές οι έννοιες της φυσικής και ειδικής μεταβλητότητας, δίνεται το ακόλουθο παράδειγμα. Ας υποθέσουμε ότι η ποιότητα ενός προϊόντος εκφράζεται με την τιμή ενός μετρήσιμου ποιοτικού χαρακτηριστικού του για το οποίο ισχύουν τα εξής: $T=15$, $LSL=10$, $USL=20$. Επιπλέον, ας υποθέσουμε ότι κάτω από συνθήκες φυσικής μεταβλητότητας, δηλαδή όταν έχουμε μια διεργασία εντός ελέγχου οι τιμές του ποιοτικού χαρακτηριστικού στα παραγόμενα προϊόντα ακολουθούν την κανονική κατανομή με μέση τιμή $\mu=20$ και διασπορά $\sigma^2=4$, δηλαδή $N(20,4)$. Τα αποτελέσματα που έχουν οι κοινές και οι ειδικές αιτίες μεταβλητότητας στην παραγωγική διεργασία φαίνονται αναλυτικά στο παρακάτω γράφημα.

Γράφημα 1.1: Ερμηνεία ειδικών αιτίων μεταβλητότητας



Στο παραπάνω διάγραμμα παρατηρούμε ότι όταν η διαδικασία είναι εντός ελέγχου σχεδόν όλα τα παραγόμενα προϊόντα είναι εντός των ορίων προδιαγραφών, ενώ υπό συνθήκες ειδικής μεταβλητότητας σημειώνονται αλλαγές στην κατανομή των τιμών του ποιοτικού χαρακτηριστικού. Αυτές οι αλλαγές οφείλονται είτε στη μετατόπιση του μέσου της κατανομής $N(22, 4)$, είτε στην αλλαγή της διασποράς της κατανομής $N(20, 9)$, είτε στην ταυτόχρονη αλλαγή του μέσου και της διακύμανσης της κατανομής $N(25, 16)$. Σε όλες τις περιπτώσεις όμως, η παρουσία ειδικής μεταβλητότητας επιφέρει αύξηση των παραγόμενων προϊόντων με τιμές ποιοτικού χαρακτηριστικού που βρίσκονται εκτός των ορίων προδιαγραφών.

Σε μια διεργασία η έγκαιρη ανίχνευση της εμφάνισης των ειδικών αιτιών μεταβλητότητας είναι ένα από τα σημαντικότερα αντικείμενα με τα οποία ασχολείται ο Στατιστικός Έλεγχος Ποιότητας. Πιο συγκεκριμένα, όταν παρατηρηθούν τιμές ποιοτικού χαρακτηριστικού εκτός των ορίων προδιαγραφών προχωρούμε σε έρευνα ώστε να προβούμε στις απαραίτητες ενέργειες για να αποφύγουμε τη κατασκευή μεγάλου αριθμού προϊόντων που δεν πληρούν τις προδιαγραφές. Μια από τις πιο διαδεδομένες τεχνικές ανίχνευσης αιτιών μεταβλητότητας σε μια διεργασία, είναι τα διαγράμματα ελέγχου (control charts).

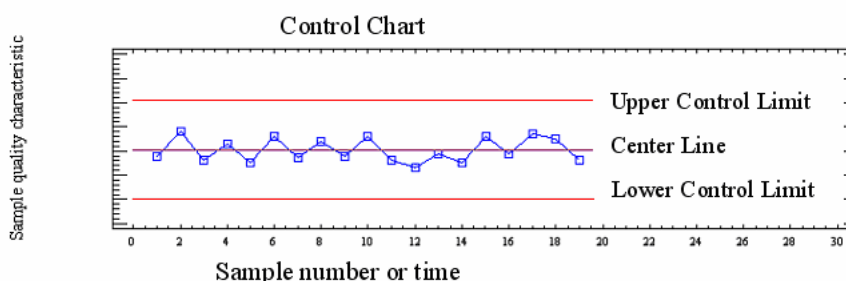
1.2 Διαγράμματα ελέγχου

Στο κεφάλαιο αυτό παρουσιάζονται διαγράμματα που αφορούν μεταβλητές. Ως μεταβλητή ορίζεται ένα ποιοτικό χαρακτηριστικό X που μετρείται σε αριθμητική κλίμακα όπως για παράδειγμα το μήκος το βάρος, όγκος προϊόντων. Η διαδικασία παρακολούθησης της κρίσιμης ποσότητας βασίζεται σε μετρήσεις του χαρακτηριστικού X (τυχαία μεταβλητή). Οι τυχαίες μεταβλητές προκύπτουν από την επιλογή τυχαίων δειγμάτων προϊόντων από την παραγωγή σε διαφορετικές χρονικές στιγμές στα οποία αντιστοιχούν τυχαία δείγματα τιμών του χαρακτηριστικού X (έστω X_1, \dots, X_n), όπου n το μέγεθος του τυχαίου δείγματος. Όταν έχουμε ένα πρόβλημα που αφορά μια μεταβλητή είναι συχνά απαραίτητο να μπορούμε να ελέγξουμε τη μέση τιμή και τη διασπορά του ποιοτικού χαρακτηριστικού. Έτσι λοιπόν χρησιμοποιώντας τα τυχαία δείγματα, υπολογίζουμε την τιμή $W_i = g(X_i)$, μιας κατάλληλης στατιστικής συνάρτησης (τυχαίας μεταβλητής) για να εκτιμήσουμε με τη βοήθεια μιας αμερόληπτης εκτιμήτριας την κρίσιμη ποσότητα που μας ενδιαφέρει (συνήθως η μέση τιμή μ και η τυπική απόκλιση σ της X). Ο έλεγχος για το μέσο κατά τη διάρκεια μιας διαδικασίας γίνεται συνήθως με τα διαγράμματα ελέγχου για το μέσο ή αλλιώς \bar{X} διαγράμματα. Ο έλεγχος της διασποράς μπορεί να γίνει είτε με διαγράμματα ελέγχου για την τυπική απόκλιση, που ονομάζονται S-διαγράμματα, είτε με τα διαγράμματα ελέγχου για το εύρος, που ονομάζονται R διαγράμματα. Τα \bar{X} και

R διαγράμματα αποτελούν τα πιο χρήσιμα εργαλεία του Στατιστικού Ελέγχου Διεργασίας. Ωστόσο, είναι σημαντικό να σημειώσουμε ότι πρέπει να γίνεται συγχρόνως έλεγχος του μέσου αλλά και της διασποράς. Έτσι η παρακολούθηση της συμπεριφοράς της κρίσιμης ποσότητας επιτυγχάνεται με την παρακολούθηση των τιμών που λαμβάνει η στατιστική συνάρτηση W στα διάφορα δείγματα.

Το παρακάτω γράφημα στο οποίο παρουσιάζεται μια γραφική παράσταση είναι ένα τυπικό διάγραμμα ελέγχου Shewhart.

Γράφημα 1.2: Τυπικό Διάγραμμα ελέγχου Shewhart



Στο παραπάνω διάγραμμα μπορούμε να διακρίνουμε τέσσερις γραμμές. Η τεθλασμένη γραμμή ενώνει τις τιμές των παρατηρήσεων W_i , που έχουν απεικονιστεί με τετραγωνάκια. Οι άλλες γραμμές είναι η κεντρική γραμμή (central line), το άνω όριο ελέγχου (upper control limit UCL) και το κάτω όριο ελέγχου (lower control limit LCL). Η κεντρική γραμμή ονομάζεται αλλιώς και μέσο επίπεδο της διεργασίας και παριστάνει ουσιαστικά την μέση τιμή (mean value) της W όπως προκύπτει από μια διεργασία που λειτουργεί μόνο με τη παρουσία φυσικής μεταβλητότητας, δηλαδή μιας εντός στατιστικού ελέγχου διεργασίας. Οι άλλες δύο γραμμές καθορίζουν το διάστημα μέσα στο οποίο θα πρέπει να βρίσκεται η πλειοψηφία των παρατηρούμενων τιμών για να μπορούμε να πούμε ότι πρόκειται για μια εντός ελέγχου διεργασία. Αν όλες οι τιμές της W βρίσκονται μέσα στα όρια ελέγχου και η συμπεριφορά τους είναι τυχαία τότε κι εμείς μπορούμε να υποθέσουμε ότι πρόκειται για μια διεργασία που λειτουργεί υπό συνθήκες φυσικής μεταβλητότητας και μόνο. Εφόσον έχουμε μια εντός ελέγχου διεργασία, δε χρειάζεται να προβούμε σε κάποια διορθωτική ενέργεια. Στη περίπτωση που έστω και ένα σημείο του γραφήματος, δηλαδή μια

παρατηρούμενη τιμή της W βρεθεί έξω από τα όρια ελέγχου, τότε ένα σημείο του γραφήματος θα είναι έξω από τις γραμμές που καθορίζουν τα όρια του διαστήματος. Αυτό σημαίνει ότι έχουμε μια ένδειξη ότι ίσως η διεργασία να είναι εκτός ελέγχου, άρα να λειτουργεί υπό συνθήκες ειδικής μεταβλητότητας και τότε κρίνεται απαραίτητο να γίνει μια έρευνα προκειμένου να ανακαλυφθούν αυτές οι αιτίες που είναι υπεύθυνες για αυτή τη συμπεριφορά. Σε αυτή τη περίπτωση πρέπει να προβούμε σε διορθωτικές ενέργειες έτσι ώστε η διεργασία να είναι εντός ελέγχου.

Στον ακόλουθο πίνακα, παρουσιάζεται ένα γενικό μοντέλο που ονομάζεται μοντέλο ορίων σίγμα (sigma limits model), που μας βοηθάει να κατασκευάσουμε ένα διάγραμμα ελέγχου Shewhart.

Πίνακας 1.1

Μοντέλο ορίων σίγμα
$UCL = \mu_w + L \sigma_w$
$Center\ Line = \mu_w$
$LCL = \mu_w - L\sigma_w$

Οι ποσότητες μ_w και σ_w συμβολίζουν τη μέση τιμή και τη τυπική απόκλιση της στατιστικής συνάρτησης W που απεικονίζεται στο διάγραμμα ελέγχου (συνήθως γίνεται η υπόθεση ότι ακολουθεί την κανονική κατανομή). Η ποσότητα L δηλώνει την απόσταση των ορίων ελέγχου από τη κεντρική γραμμή σε μονάδες τυπικής απόκλισης. Στην περίπτωση που ασχολούμαστε με διαγράμματα Shewhart με 3σ όρια ελέγχου, τότε $L=3$.

Στην πράξη η μέση τιμή μ και η τυπική απόκλιση σ της X είναι άγνωστες οπότε πρέπει πρώτα να εκτιμηθούν για να μπορέσουμε στη συνέχεια να κατασκευάσουμε κατάλληλα διαγράμματα ελέγχου. Παρακάτω διακρίνουμε δύο περιπτώσεις για την εκτίμηση των ποσοτήτων μ και σ ανάλογα με το αν έχουμε δείγματα μεγέθους μεγαλύτερου ή ίσου με τη μονάδα από την τυχαία μεταβλητή X .

A Περίπτωση : m ανεξάρτητα δείγματα μεγέθους $n > 1$

Έστω ότι έχουμε στη διάθεσή μας m ανεξάρτητα τυχαία δείγματα μεγέθους $n > 1$ το καθένα, τα $X_{i1}, X_{i2}, \dots, X_{im}, 1 \leq i \leq m$

A1 Περίπτωση: Εκτίμηση του μ

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

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

Η ποσότητα $\bar{\bar{X}}$ χρησιμοποιείται ως εκτίμηση της ποσότητας μ (αμερόληπτη και συνεπής εκτίμηση του μ), δηλαδή

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

A2 Περίπτωση: Εκτίμηση του σ

i) **Μέθοδος R**: Έστω R_1, R_2, \dots, R_m τα εύρη των m δειγμάτων, δηλαδή

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

και

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

Τότε

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

Όπου d_2 είναι μια σταθερά που εξαρτάται από το μέγεθος του δείγματος n η οποία επιλέγεται με τέτοιο τρόπο έτσι ώστε $E(\sigma) = \sigma$.

ii) **Μέθοδος S**: Έστω S_1, S_2, \dots, S_m οι ποσότητες που ορίζονται από τη σχέση

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

και

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

Τότε

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

Όπου c_4 είναι μια σταθερά που εξαρτάται από το μέγεθος του δείγματος n η οποία επιλέγεται με τέτοιο τρόπο έτσι ώστε $E(\sigma) = \sigma$.

iii) Μέθοδος S^2 : Έστω $S_1^2, S_2^2, \dots, S_m^2$ οι ποσότητες που ορίζονται από τη σχέση

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

και

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

Τότε έχουμε ότι

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

B Περίπτωση: m ανεξάρτητα δείγματα μεγέθους $n=1$

Έστω ότι έχουμε στη διάθεσή μας m ανεξάρτητες παρατηρήσεις

$$X_1, X_2, \dots, X_m \text{ από την } X.$$

B1 Περίπτωση: Εκτίμηση του μ

Σε αυτή τη περίπτωση έχουμε ότι

$$\mu = \bar{X} = \frac{X_1 + X_2 + \dots + X_m}{m} = \frac{\sum_{i=1}^m X_i}{m}.$$

B2 Περίπτωση: Εκτίμηση του σ

i) **Μέθοδος MR:** Θέτουμε

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

και

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

τότε

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

όπου η σταθερά d_2 υπολογίζεται για δείγμα μεγέθους $n = 2$ (προφανώς $E(\sigma) = \sigma$).

ii) **Μέθοδος S :**

Έστω

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

τότε

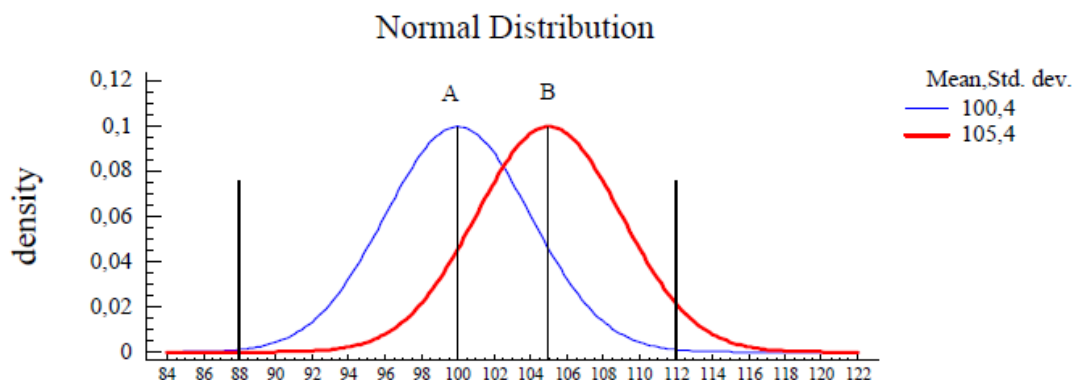
$$\sigma = \frac{S}{c_4}.$$

Όπου c_4 υπολογίζεται για δείγμα μεγέθους $n = m$ (προφανώς $E(\sigma) = \sigma$).

1.3 Ικανότητα μιας διεργασίας

Στο παρακάτω γράφημα παρουσιάζονται δύο διεργασίες A και B για τις οποίες έχουμε υποθέσει ότι βρίσκονται υπό συνθήκες φυσικής μεταβλητότητας, δηλαδή πρόκειται για εντός στατιστικού ελέγχου διεργασίες. Η A διεργασία ακολουθεί την κανονική κατανομή με μέση τιμή $\mu = 100$, και διασπορά 16 ($N(100,16)$), ενώ η B διεργασία ακολουθεί κι αυτή κανονική κατανομή με ίδια διασπορά $\sigma^2 = 16$ και μέση τιμή $\mu = 105$ ($N(105,16)$). Οι δύο αυτές διεργασίες έχουν τα ίδια όρια προδιαγραφών, άνω όριο 112 και κάτω όριο 88, δηλαδή $USL = 112$ και $LSL = 88$.

Γράφημα 1.3: Δύο διεργασίες εντός στατιστικού ελέγχου με ίδια όρια προδιαγραφών.



Παρατηρώντας προσεκτικά το παραπάνω διάγραμμα μπορούμε εύκολα να διαπιστώσουμε ότι η διεργασία A είναι πιο ικανή από την B. Είναι προφανές από το σχήμα ότι η διεργασία B παράγει περισσότερα ελαττωματικά προϊόντα γιατί πολλά από τα σημεία του γραφήματός της βρίσκονται εκτός των ορίων προδιαγραφών.

Από τα παραπάνω, καταλαβαίνουμε την ιδιαίτερη σημασία που έχουν τα διαγράμματα ελέγχου αφού μας δίνουν πληροφορίες για το αν η διεργασία λειτουργεί με ένα σταθερό τρόπο. Από την άλλη όμως διαπιστώνουμε ότι επειδή στην ανάπτυξη των διαγραμμάτων δε λαμβάνονται υπόψη τα όρια προδιαγραφών του προϊόντος, δεν μπορούν να μας δώσουν επιπλέον πληροφορίες που θα ήταν χρήσιμες για το Στατιστικό Έλεγχο Διεργασιών. Μία από αυτές τις πληροφορίες που δε μπορούμε να αντλήσουμε από ένα διάγραμμα ελέγχου, είναι το κατά πόσο η διεργασία είναι ικανή να παράγει προϊόντα που θα γνωρίζουμε εκ των προτέρων ότι θα είναι εντός των ορίων προδιαγραφών. Για αυτόν ακριβώς τον σκοπό υπάρχουν κάποια αριθμητικά μέτρα, τα οποία ονομάζονται δείκτες ικανότητας της διεργασίας (process capability indices).

Η ανάλυση της ικανότητας μιας διεργασίας (process capability analysis) σχετίζει συνήθως το μέσο μ και την τυπική απόκλιση σ της κατανομής ενός ποιοτικού χαρακτηριστικού X με τα όρια προδιαγραφών LSL και USL κι έτσι προκύπτουν οι δείκτες ικανότητας της διεργασίας.

Σε αυτό το σημείο θα πρέπει να σημειώσουμε ότι οι πρώτοι που ασχολήθηκαν με την έννοια της ικανότητας μιας διεργασίας ήταν οι Ιάπωνες στη δεκαετία του 70. Ο λόγος που οδηγήθηκαν οι Ιάπωνες στην αναζήτηση των δεικτών ικανότητας ήταν η επιθυμία τους να έχουν εγγυημένη υψηλή ποιότητα των παραγόμενων προϊόντων. Αρχικά είχαν προσπαθήσει να φέρουν εντός ελέγχου, διεργασίες που ήταν εκτός, κι επειδή αργότερα συνειδητοποίησαν ότι αυτό δεν ήταν αρκετό για να τους εγγυηθεί υψηλή ποιότητα, οδηγήθηκαν στην ανάπτυξη των πρώτων πέντε δημοφιλών Ιαπωνικών δεικτών ικανότητας που είναι οι C_p , C_{pk} , C_{pl} , C_{pu} και k .

Οι περισσότεροι προτεινόμενοι δείκτες ικανότητας μιας διεργασίας συνδέονται με διάφορες ποσότητες. Οι πιο βασικές δίνονται στον ακόλουθο πίνακα:

Πίνακας 1.2

Μέση τιμή του ποιοτικού χαρακτηριστικού	μ
Τυπική απόκλιση του ποιοτικού χαρακτηριστικού	σ
Κάτω όριο προδιαγραφών	LSL
Άνω όριο προδιαγραφών	USL
Τιμή στόχος διεργασίας	T
Μέσο του διαστήματος [LSL, USL]	$M = \frac{LSL + USL}{2}$
Μήκος του διαστήματος [LSL, USL]	$d = \frac{USL - LSL}{2}$

Είναι απαραίτητο να σημειώσουμε ότι τις περισσότερες φορές $\mu = M = T$ αλλά αυτό δε συμβαίνει πάντα, όπως επίσης κι ότι έχουμε μονόπλευρη διεργασία δηλαδή υπάρχει μόνο ένα όριο προδιαγραφών είτε το άνω (USL) είτε το κάτω (LSL).

Από τα παραπάνω συμπεραίνουμε ότι οι δείκτες ικανότητας μιας διεργασίας χρησιμοποιούνται ευρέως και δίνουν με έναν εύχρηστο και κατανοητό τρόπο μια άμεση ένδειξη για το πόσο είναι ικανή μια διεργασία να παράγει προϊόντα εντός των ορίων προδιαγραφών του στατιστικού ελέγχου. Αυτή ακριβώς η ιδιότητά που έχουν τούς κάνει ένα από τα σημαντικότερα στατιστικά εργαλεία των κατασκευαστών. Είναι κάτι που εύκολα διαπιστώνεται, αν σκεφτεί κανείς ότι όλο και αυξάνονται οι κατασκευαστές οι οποίοι στις συμβάσεις των αγορών που κάνουν με τους προμηθευτές, απαιτούν να έχουν τεκμηριωμένη απόδειξη για την ποιότητα των προϊόντων υπολογίζοντας σε κάθε αποστολή τους δείκτες ικανότητας. Αυτός λοιπόν είναι κι ένας από τους λόγους που η χρήση και η ανάπτυξη των δεικτών ικανότητας ως ενός τρόπου έκφρασης της ποιότητας των παραγόμενων προϊόντων αναμένεται να επεκταθεί.

1.4 Χρήση ή μη των δεικτών ικανότητας μιας διεργασίας

Η χρήση των δεικτών ικανότητας μιας διεργασίας έχει απασχολήσει ιδιαίτερα τόσο τους στατιστικούς όσο και τους επαγγελματίες κι αυτό επειδή οι δείκτες ικανότητας, ενώ θεωρούνται από τα πιο χρήσιμα εργαλεία του Στατιστικού Ποιοτικού Ελέγχου, αν δεν χρησιμοποιηθούν σωστά έχουν πολύ σημαντική επίδραση στις βιομηχανίες που τους χρησιμοποιούν, κυρίως αρνητική.

Η σωστή χρήση και προσεκτική εφαρμογή των δεικτών ικανότητας είναι βασική προϋπόθεση για τη χρήση τους γιατί σε διαφορετική περίπτωση μπορούν να γίνουν ένα επικίνδυνο εργαλείο. Αυτό συμβαίνει επειδή αυτοί οι δείκτες προσπαθώντας να αντλήσουν πάρα πολλές πληροφορίες από μία διεργασία και να τις αποτυπώσουν όλες σε έναν και μόνο αριθμό μπορούν να οδηγήσουν σε λανθασμένες ερμηνείες παρά το γεγονός ότι εμφανίζονται απλοί στη χρήση τους.

Πολλοί επαγγελματίες πιστεύουν ότι οι δείκτες ικανότητας είναι πολύ πιο απλοί από κάποιες άλλες στατιστικές μεθόδους που χρησιμοποιούνται από τους κατασκευαστές προϊόντων για την αξιολόγηση της παραγωγικής διεργασίας, για αυτό και τους θεωρούν το πιο χρήσιμο εργαλείο του Σύγχρονου Ποιοτικού Ελέγχου. Προκειμένου λοιπόν να αποφευχθούν ανεπιθύμητα αποτελέσματα από την κακή χρήση τους, κρίνεται αναγκαία αφενός η κατανόηση των υποθέσεων από τους κατασκευαστές που τους χρησιμοποιούν και αφετέρου η χρήση των διαστημάτων εμπιστοσύνης για τους συγκεκριμένους δείκτες.

ΚΕΦΑΛΑΙΟ 2

ΒΑΣΙΚΕΣ ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ ΤΗΣ ΙΚΑΝΟΤΗΤΑΣ ΜΙΑΣ ΔΙΕΡΓΑΣΙΑΣ

2.1 Εισαγωγή

Η ανάλυση ικανότητας μιας διεργασίας είναι ένα μέρος από τη συνολική προσπάθεια βελτίωσης της ποιότητας των παραχθέντων προϊόντων που έχει μεγάλη σημασία.

Τα σημαντικότερα σημεία της ανάλυσης της ικανότητας μιας διεργασίας είναι τα ακόλουθα

1. Να προβλέπει κατά πόσο η διαδικασία θα βρίσκεται εντός ορίων.
2. Να βοηθάει τους κατασκευαστές των προϊόντων να διαλέξουν κατάλληλες πρώτες ύλες ή να τους βοηθάει στην τροποποίηση των διαδικασιών παραγωγής προϊόντων.
3. Να βοηθήσει στην παρακολούθηση του δείγματος, δηλαδή να ελέγχει αν βρίσκεται εντός ενός συγκεκριμένου διαστήματος.
4. Να έχει συγκεκριμένες απαιτήσεις για το είδος και την ποιότητα των πρώτων υλών που θα χρησιμοποιηθούν για την παραγωγή των προϊόντων.
5. Να έχει την ικανότητα να διαλέξει από ποιον θα προμηθευτεί τις πρώτες ύλες και να διαθέτει μια γενική αλυσίδα προμηθειών.
6. Σε περίπτωση που υπάρχουν μεγάλες αποκλίσεις στις τιμές, θα πρέπει να υπάρχει μια σειρά από διαδικασίες, έτσι ώστε να έχουμε μια πιο ολοκληρωμένη εικόνα για την ποιότητα των προϊόντων.

7. Να μειώσει τη μεταβλητότητα σε μια διαδικασία.

Για τους παραπάνω λόγους, η ανάλυση της ικανότητας μιας διεργασίας είναι μια τεχνική που έχει εφαρμογή σε πολλά σημεία του κύκλου ζωής των προϊόντων. Τέτοια σημεία είναι για παράδειγμα, το τελικό προϊόν, η διαδικασία κατασκευής του, η διαδικασία προμηθειών πρώτων υλών και η διαδικασία παραγωγής του.

Στην στατιστική, προκειμένου να αναλύσουμε την ικανότητα μιας διεργασίας, χρησιμοποιούμε κυρίως τέσσερις τεχνικές, τα ιστογράμματα, τα διαγράμματα πιθανοτήτων, τους δείκτες ικανότητας μιας διεργασίας και τα διαγράμματα ελέγχου που αναφέραμε στο προηγούμενο κεφάλαιο. Σε αυτήν την ενότητα θα αναλυθούν τα ιστογράμματα και τα διαγράμματα πιθανοτήτων. Οι δείκτες ικανότητας, που είναι και η βασική τεχνική ανάλυσης της ικανότητας μια διεργασίας, αναλύονται εκτενέστερα στο αμέσως επόμενο κεφάλαιο.

2.2 Ανάλυση της ικανότητας μιας διεργασίας με τη βοήθεια του ιστογράμματος

Το ιστόγραμμα μπορεί να μας βοηθήσει για να εκτιμήσουμε την ικανότητα μιας διαδικασίας. Εναλλακτικά, ένα steam and leaf διάγραμμα, μπορεί να υποκατασταθεί από ένα ιστόγραμμα. Σε αυτό το σημείο, πρέπει να τονίσουμε ότι για να μπορέσει το ιστόγραμμα να δώσει μια αξιόπιστη εκτίμηση για την ικανότητα της διεργασίας πρέπει να υπάρχουν το λιγότερο εκατό παρατηρήσεις. Στην περίπτωση όπου τα μηχανήματα, τα οποία ελέγχουν την ποιότητα των παραγόμενων προϊόντων, έχουν πρόσβαση στην διαδικασία και μπορούν να ελέγξουν πιθανά σφάλματα πρέπει να πραγματοποιηθούν τα παρακάτω.

1. Να διαλέξουν μία μηχανή ή μηχανές που χρησιμοποιούνται. Αν τα αποτελέσματα στηρίζονται σε μία ή γενικά σε λίγες μηχανές τότε πρέπει να διευρύνουν το πλήθος των μηχανών ή εναλλακτικά να προσέξουν ώστε το δείγμα από τις μηχανές που θα συλλέξουν να είναι αντιπροσωπευτικό. Επιπροσθέτως, αν η μηχανή έχει πολλές λειτουργίες ή εφαρμογές, κρίνεται απαραίτητο να συλλέξουμε

ένα πλήθος από τέτοιου είδους μηχανές, έτσι ώστε να περιορίσουμε όσο το δυνατόν περισσότερο την μεταβλητότητα. Αυτό θα γίνει με τη βοήθεια σχεδιασμένων πειραμάτων.

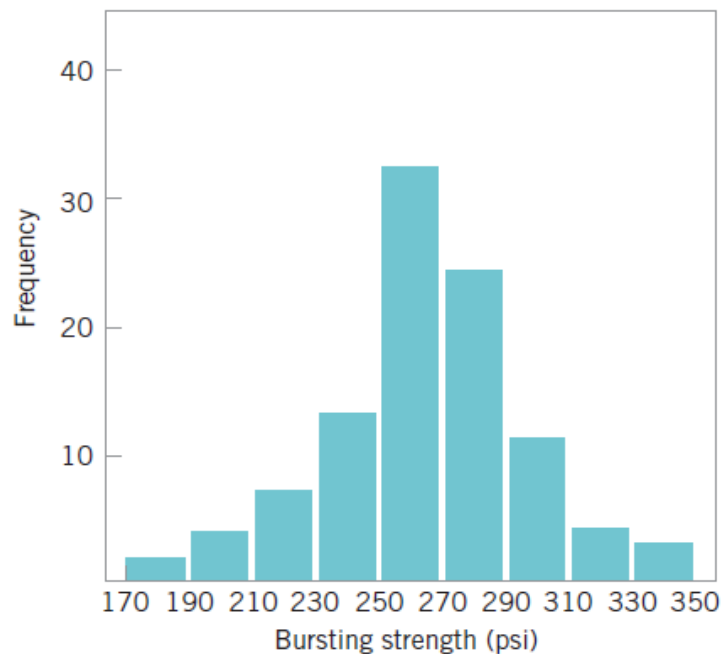
2. Να διαλέξουν τις συνθήκες κάτω από τις οποίες θα γίνει η συγκεκριμένη διεργασία. Αυτό σημαίνει ότι πρέπει να ελέγξουμε ορισμένους παράγοντες, όπως για παράδειγμα η ταχύτητα με την οποία κόβει ένα μηχάνημα, κάτω από συγκεκριμένη θερμοκρασία. Στη συνέχεια είναι απαραίτητο να μελετήσουμε την επίδραση που θα έχουν αυτοί οι παράγοντες στην ικανότητα της διεργασίας.
3. Να διαλέξουν μια αντιπροσωπευτική λειτουργία. Σε πολλές μελέτες είναι σημαντικό να εκτιμήσουμε τη μεταβλητότητα σε μια συγκεκριμένη λειτουργία. Σε αυτήν την περίπτωση η λειτουργία που θα επιλεγεί θα πρέπει να είναι τυχαία από το πλήθος των λειτουργιών.
4. Τέλος πρέπει να καταγράψουμε τα αποτελέσματα της διαδικασίας καθώς και το χρόνο μέσα στον οποίο πραγματοποιήθηκε η διεργασία.

Το ακόλουθο ιστόγραμμα εξετάζει την αντοχή εκατό γυάλινων δοχείων (Montgomery, 2009). Τα στατιστικά δεδομένα που συλλέξαμε φαίνονται στον παρακάτω πίνακα και το ερώτημα είναι ποια είναι η ικανότητα της διεργασίας σε αυτό το παράδειγμα;

Πίνακας 2.1

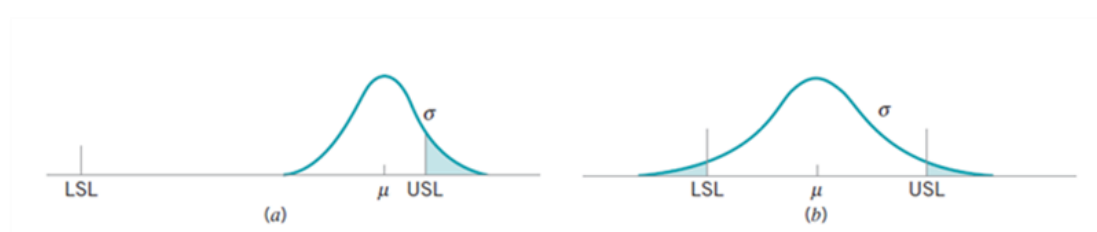
265	197	346	280	265	200	221	265	261	278
205	286	317	242	254	235	176	262	248	250
263	274	242	260	281	246	248	271	260	265
307	243	258	321	294	328	263	245	274	270
220	231	276	228	223	296	231	301	337	298
268	267	300	250	260	276	334	280	250	257
260	281	208	299	308	264	280	274	278	210
234	265	187	258	235	269	265	253	254	280
299	214	264	267	283	235	272	287	274	269
215	318	271	293	277	290	283	258	275	251

Γράφημα 2.1



Έχουμε ότι $\bar{x}=264.06$, $s=32.02$ και $\bar{x} \pm 3s = 264 \pm 96$ psi, που στην ουσία είναι και η εκτίμηση της ικανότητας της διαδικασίας. Όπως επίσης φαίνεται από τη μορφή του ιστογράμματος η κατανομή της αντοχής των γυάλινων δοχείων είναι προσεγγιστικά κανονική. Επομένως μπορούμε να ισχυριστούμε ότι το 99.73% των μπουκαλιών που κατασκευάστηκαν σε αυτή τη διεργασία έχουν αντοχή μεταξύ 168 και 360 psi. Εδώ αξίζει να σημειώσουμε ότι θα μπορούσαμε να εκτιμήσουμε την ικανότητα της συγκεκριμένης διεργασίας ανεξάρτητα από τις προδιαγραφές αντοχής των μπουκαλιών. Ένα σημαντικό πλεονέκτημα του ιστογράμματος σχετικά με την εκτίμηση της ικανότητας μιας διεργασίας είναι ότι μας δίνει μια γρήγορη και ξεκάθαρη εικόνα για την αποτελεσματικότητα της διαδικασίας. Επιπλέον, σε περίπτωση που η διαδικασία δεν είναι ικανή, μας δείχνει και την αιτία που οφείλεται αυτή η αδυναμία της διεργασίας. Για να γίνει περισσότερο κατανοητό αυτό ας παρατηρήσουμε το παρακάτω διάγραμμα.

Γράφημα 2.2



Όπως φαίνεται στο παραπάνω διάγραμμα στην πρώτη περίπτωση έχουμε μια διεργασία που φαίνεται ότι είναι ικανή, αλλά η τιμή-στόχος της διαδικασίας δεν είναι σε καλό σημείο. Στη δεύτερη περίπτωση έχουμε μια διεργασία που δεν είναι ικανή εξαιτίας της μεγάλης μεταβλητότητας.

2.3 Ανάλυση της ικανότητας μιας διεργασίας με τη βοήθεια ενός διαγράμματος πιθανοτήτων

Το διάγραμμα πιθανοτήτων είναι ένας εναλλακτικός τρόπος εκτός από το ιστόγραμμα που μπορεί να χρησιμοποιηθεί προκειμένου να έχουμε πληροφορίες για τη μορφή, το κέντρο και την έκταση της κατανομής των παρατηρήσεων. Ένα από τα πλεονεκτήματα που έχει συγκριτικά με το ιστόγραμμα, είναι το γεγονός ότι δεν είναι απαραίτητο να χωρίσουμε το εύρος των τιμών της μεταβλητής σε μικρότερα διαστήματα, καθώς επίσης μπορεί να δώσει αξιόπιστα αποτελέσματα ακόμα και με δείγματα μικρού μεγέθους, κάτι που δε γίνεται με το ιστόγραμμα. Σε γενικές γραμμές, ένα διάγραμμα πιθανοτήτων είναι ένα γράφημα από τα ταξινομημένα δεδομένα ως προς τη σωρευτική συχνότητα των τιμών του δείγματος, έτσι ώστε η σωρευτική συνάρτηση της υποτιθέμενης κατανομής των δεδομένων να είναι μια ευθεία γραμμή.

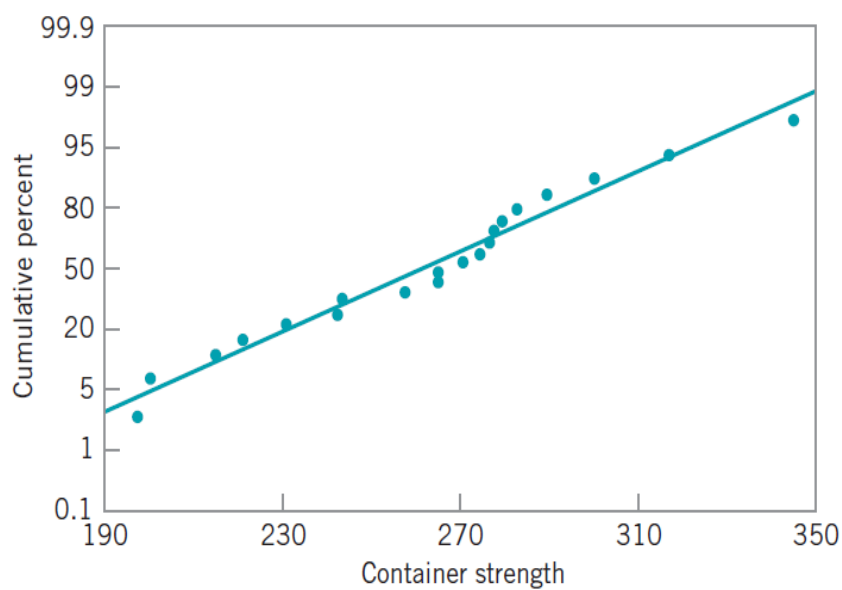
Για να κατανοήσουμε την χρήση του διαγράμματος πιθανοτήτων σε μια μελέτη ικανότητας μιας διεργασίας ας σκεφτούμε τις 20 παρακάτω παρατηρήσεις για τη μελέτη της αντοχής των γυάλινων δοχείων (Montgomery, 2009)..

Πίνακας 2.2

197	271
200	275
215	277
221	278
231	280
242	283
245	290
258	301
265	318
265	346

Οι παραπάνω παρατηρήσεις φαίνονται στο ακόλουθο διάγραμμα πιθανοτήτων σχετικά με την αντοχή των γυάλινων δοχείων.

Γράφημα 2.3



Σε αυτό το σημείο αξίζει να παρατηρήσουμε ότι τα δεδομένα βρίσκονται κοντά στην ευθεία γραμμή, που σημαίνει ότι η κατανομή που ακολουθούν οι παρατηρήσεις είναι προσεγγιστικά κανονική. Ο δειγματικός μέσος της κανονικής κατανομής είναι το πεντηκοστό εκατοστημόριο και μπορούμε να εκτιμήσουμε από το διάγραμμα πιθανοτήτων ότι η τιμή του είναι περίπου 265 psi, ενώ η τυπική απόκλιση της κατανομής είναι η κλίση της ευθείας γραμμής. Είναι γεγονός ότι μας διευκολύνει να υπολογίσουμε την τιμή της τυπικής απόκλισης ως τη διαφορά ανάμεσα στο ογδοεικοστό τέταρτο εκατοστημόριο και στο πεντηκοστό. Χρησιμοποιώντας τα παραπάνω δεδομένα αλλά και το διάγραμμα πιθανοτήτων μπορούμε να βρούμε ότι

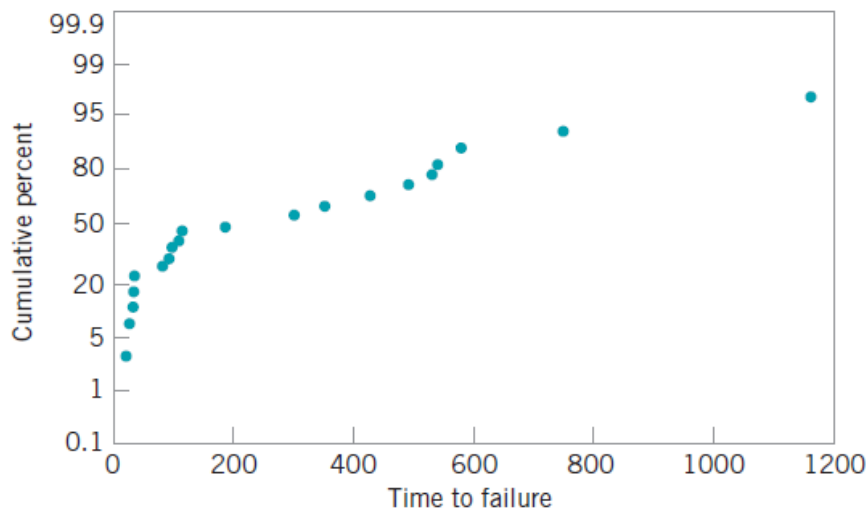
$$\sigma = 840 \text{ εκατοστημόριο} - 500 \text{ εκατοστημόριο} = 298 - 265 \text{ psi} = 33 \text{ psi}.$$

Εδώ πρέπει να τονίσουμε ότι οι εκτιμήσεις του δειγματικού μέσου και της τυπικής απόκλισης $\mu = 265 \text{ psi}$ και $\sigma = 33 \text{ psi}$ δεν απέχουν και πολύ από τις τιμές της μέσης τιμής $\bar{X} = 264.06$ και της τυπικής απόκλισης $S = 32.02$.

Το κανονικό διάγραμμα πιθανοτήτων μπορεί να εκτιμήσει την απόδοση αλλά και τα λάθη της διαδικασίας. Για παράδειγμα ο προσδιορισμός του κατώτερου ορίου προδιαγραφών της αντοχής ενός γυάλινου δοχείου είναι $LSL = 200 \text{ psi}$. Από το παραπάνω διάγραμμα θα μπορούσαμε να εκτιμήσουμε ότι το 5% των γυάλινων δοχείων έχουν κατασκευαστεί με τέτοιο τρόπο ώστε να σπάνε κάτω από αυτό το όριο.

Η χρήση των διαγραμμάτων πιθανότητας θα πρέπει να είναι πολύ σχολαστική, διότι αν τα δεδομένα δεν προέρχονται από την κατανομή που υποθέτουμε τότε τα συμπεράσματα που θα εξάγουμε από το διάγραμμα πιθανοτήτων για την ικανότητα της διεργασίας ενδέχεται να περιέχουν πολλά σοβαρά σφάλματα. Το παρακάτω διάγραμμα κανονικής πιθανότητας παρουσιάζει το χρόνο επιβίωσης μιας βαλβίδας (σε ώρες) σε ένα χημικό εργαστήριο. Καθώς εξετάζουμε αυτό το διάγραμμα παρατηρούμε ότι η κατανομή του χρόνου επιβίωσης της βαλβίδας δεν είναι κανονική.

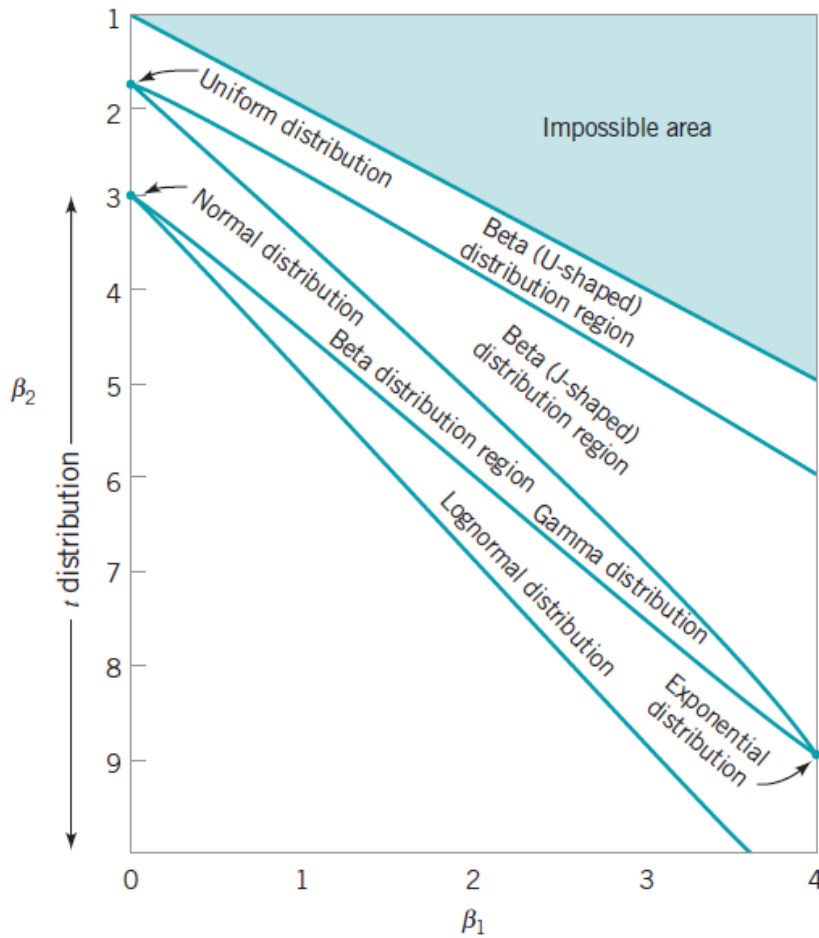
Γράφημα 2.4



Ένα φανερό μειονέκτημα του διαγράμματος πιθανοτήτων είναι ότι δεν είναι μια αντικειμενική διαδικασία. Αυτό σημαίνει ότι υπάρχει περίπτωση δύο αναλυτές χρησιμοποιώντας τα ίδια δεδομένα (παρατηρήσεις), να βγάλουν εντελώς διαφορετικά αποτελέσματα. Για αυτό το λόγο, κρίνεται απαραίτητη η χρήση συμπληρωματικών διαγραμμάτων πιθανοτήτων με πιο επίσημα στατιστικά αποτελέσματα που ονομάζονται τεστ καλής προσαρμογής. Μια πολύ καλή εισαγωγή για αυτά τα τεστ βρίσκεται στο Shapiro (1980). Συνδυάζοντας την ερμηνεία του διαγράμματος κανονικής πιθανότητας με το Shapiro-Wilk τεστ για την κανονικότητα μπορεί να κάνει τη διαδικασία πιο αντικειμενική.

Στο διάγραμμα πιθανοτήτων είναι πολύ σημαντικό να διαλέξουμε μια κατανομή που να ταιριάζει με τα δεδομένα. Πολλές φορές μπορούμε να χρησιμοποιήσουμε τις γνώσεις μας πάνω σε φυσικές καταστάσεις αλλά και την εμπειρία μας ώστε να προτείνουμε μια κατανομή για τα δεδομένα μας. Σε πολλές περιπτώσεις όπως για παράδειγμα στο παρακάτω διάγραμμα το σχήμα του γραφήματος μπορεί να μας βοηθήσει στην επιλογή της κατανομής, η οποία περιγράφει τα δεδομένα.

Γράφημα 2.5



Αυτό το γράφημα μας δείχνει τη περιοχή ανάμεσα στις τιμές β_1, β_2 για αρκετές κατανομές πιθανοτήτων, όπου β_1 και β_2 είναι τα μέτρα της λοξότητας και της κυρτότητας αντίστοιχα. Σύμφωνα με το παραπάνω διάγραμμα, μπορούμε να υπολογίσουμε την εκτίμηση της λοξότητας και της κυρτότητας από το δείγμα ως εξής

$$\sqrt{\beta_1} = \frac{M_3}{(M_2)^{3/2}}$$

και

$$\beta_2 = \frac{M_4}{M_2^2}$$

όπου

$$M_j = \frac{\sum_{i=1}^n (x_i - \bar{x})^j}{n}, j = 1, 2, 3, 4$$

και το σημείο (β_1, β_2) είναι σημείο του γραφήματος. Αν τα σημεία του διαγράμματος βρίσκονται κοντά σε ένα σημείο, γραμμή ή περιοχή που ανταποκρίνεται στο γράφημα, αυτό σημαίνει ότι η κατανομή που διαλέξαμε είναι μία καλή επιλογή για τα δεδομένα μας. Αν το σημείο πέσει σε περιοχές εκτός του β_1, β_2 , όπου καμιά από τις κατανομές δε φαίνεται κατάλληλη, τότε θα ανατρέξουμε σε πιο γενικές κατανομές πιθανοτήτων όπως τις οικογένειες κατανομών του Johnson ή Pearson.

Βασική παρατήρηση σε αυτό το σημείο είναι ότι η στατιστική λοξότητα και η κυρτότητα δεν είναι αξιόπιστες αν το δείγμα δεν είναι πολύ μεγάλο. Παρόμοιες διαδικασίες και γραφήματα στα οποία ψάχνουμε το είδος της κατανομής βρίσκονται στο Hahn and Shapiro (1967).

ΚΕΦΑΛΑΙΟ 3

ΔΕΙΚΤΕΣ ΙΚΑΝΟΤΗΤΑΣ ΔΙΕΡΓΑΣΙΩΝ

3.1 Ο δείκτης ικανότητας C_p

Όπως αναφέρθηκε αναλυτικά στα προηγούμενα κεφάλαια, είναι πολύ σημαντικό για τη στατιστική να μπορέσουμε να καθορίσουμε το πόσο αποτελεσματική είναι μια διεργασία που παράγει προϊόντα με τιμές X εντός των ορίων προδιαγραφών. Για αυτό το λόγο, κρίνεται απαραίτητη η χρήση διαφόρων μέτρων, από τα οποία το πιο γνωστό είναι οι δείκτες ικανότητας διεργασίας. Οι συγκεκριμένοι δείκτες αποτελούν ένα μέτρο της δυνατότητας μιας διαδικασίας να κατασκευάσει προϊόντα που θα πληρούν τις προδιαγραφές. Ο δείκτης αυτός θα πρέπει να υπολογίζεται υπό τις ακόλουθες σημαντικές υποθέσεις:

1. Το ποιοτικό χαρακτηριστικό ακολουθεί την κανονική κατανομή.
2. Η διαδικασία είναι υπό στατιστικό έλεγχο.
3. Στην περίπτωση δίπλευρων προδιαγραφών, ο μέσος της διαδικασίας είναι κεντραρισμένος μεταξύ του άνω και του κάτω ορίου προδιαγραφών.

Οι υποθέσεις αυτές είναι κρίσιμες για την ακρίβεια και την εγκυρότητα των τιμών των δεικτών ικανότητας διεργασίας, κι αν δεν ισχύουν τότε οι τιμές αυτές μπορεί να προκύπτουν εσφαλμένες (βλέπε X. Κουκουβίνος, 2008).

Για να μπορέσουμε να κατανοήσουμε καλύτερα αυτή την ανάγκη αρκεί να δούμε το ακόλουθο παράδειγμα. Έστω ότι σε μια διεργασία το χαρακτηριστικό X που μας ενδιαφέρει να αναλύσουμε ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$ με

μέση τιμή μ και τυπική απόκλιση σ δηλαδή διασπορά σ^2 . Σε αυτήν την περίπτωση μπορούμε να πούμε με βεβαιότητα, ότι το 99.73% των παρατηρούμενων τιμών X θα βρίσκεται μέσα στο διάστημα $(\mu - 3\sigma, \mu + 3\sigma)$ που έχει εύρος 6σ . Το διάστημα αυτό ονομάζεται φυσικό εύρος και μας γνωστοποιεί ότι μέσα σε αυτό το εύρος θα υπάρχει το 99.73% των παραγόμενων προϊόντων.

Το πιο απλό μέτρο για την ικανότητα μιας διεργασίας είναι ο δείκτης C_p . Είναι ο πιο παλιός δείκτης που χρησιμοποιήθηκε και παρά τις αδυναμίες που παρουσιάζει, συνεχίζει να χρησιμοποιείται ευρύτατα μέχρι και σήμερα στις περισσότερες βιομηχανίες.

3.1.1 Ορισμός και ιδιότητες του δείκτη C_p

Ο δείκτης C_p ορίζεται ως εξής:

$$C_p = \frac{USL - LSL}{6\sigma}$$

όπου το USL είναι το άνω όριο προδιαγραφών και το LSL είναι το κάτω όριο προδιαγραφών. Το σ είναι η τυπική απόκλιση της διεργασίας. Από τον τύπο του μπορούμε να διαπιστώσουμε ότι ο δείκτης C_p εκφράζει τη σχέση του εύρους των ορίων προδιαγραφών ενός προϊόντος $(USL - LSL)$ με το φυσικό εύρος τιμών της διεργασίας 6σ . Το φυσικό εύρος προέρχεται από τον τύπο $(\mu - 3\sigma) - (\mu + 3\sigma)$

$$\text{Η ποσότητα } P = \left(\frac{1}{C_p} \right) \cdot 100\% = \left(\frac{6\sigma}{USL - LSL} \right) \cdot 100\%$$

δηλώνει το ποσοστό του εύρους των ορίων προδιαγραφών που χρησιμοποιεί η διεργασία.

Είναι προφανές ότι σε μια διεργασία επιθυμούμε τη μεγιστοποίηση του δείκτη C_p . Βασιζόμενοι λοιπόν, στους τύπους χρειαζόμαστε μικρή τυπική απόκλιση, δηλαδή μικρή φυσική μεταβλητότητα, επομένως μικρή τιμή του δείκτη P (ποσά αντιστρόφως ανάλογα).

Ο δείκτης C_p έχει οριστεί όταν η διεργασία ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$ με μέση τιμή μ και διασπορά σ^2 κάτι το οποίο δικαιολογεί την επιλογή της τιμής 6 στον παρονομαστή. Έτσι λοιπόν, όταν ο δείκτης C_p λαμβάνει την τιμή 1 και η διεργασία περιγράφεται από μια κανονική κατανομή με μέση τιμή μ , η οποία ισούται με την κεντρική τιμή του εύρους των προδιαγραφών, τότε το 99.73% θα είναι συμμορφούμενα δηλαδή θα βρίσκονται εντός των ορίων προδιαγραφών. Μια άλλη έκφραση που θα μπορούσαμε να χρησιμοποιήσουμε είναι να πούμε ισοδύναμα με βεβαιότητα ότι το 0.27% των παραγόμενων προϊόντων θα είναι ελαττωματικά. Αυτό μπορούμε να το ισχυριστούμε με σιγουριά μόνο όταν η μέση τιμή της διεργασίας μ είναι ίση με την κεντρική τιμή του διαστήματος που ορίζουν τα όρια προδιαγραφών (άνω όριο προδιαγραφών USL και κάτω όριο προδιαγραφών LSL).

Σε αντίθετη περίπτωση, όταν η μέση τιμή μ της διεργασίας δεν είναι ίση με την κεντρική τιμή της περιοχής των προδιαγραφών, όταν δηλαδή $\mu \neq \frac{USL + LSL}{2}$ και επιπλέον ισχύει ότι $C_p = 1$ τότε η εργασία παράγει στη χειρότερη περίπτωση 0.27% ελαττωματικά προϊόντα.

Εν συνεχεία, για να μπορέσουμε να αναλύσουμε τις υπόλοιπες ιδιότητες του δείκτη C_p , θα θεωρούμε ως δεδομένο ότι οι παρατηρούμενες τιμές ακολουθούν την κανονική κατανομή $N(\mu, \sigma^2)$ με μέση τιμή μ και διασπορά σ^2 . Επιπροσθέτως, θεωρούμε ότι στη φάση σχεδιασμού της διεργασίας έχουν οριστεί το άνω όριο προδιαγραφών USL και το κάτω όριο προδιαγραφών LSL .

Στους παρακάτω τύπους ορίζουμε ως την πιθανότητα να παραχθεί ένα προϊόν εντός των ορίων προδιαγραφών και ονομάζεται απόδοση της διεργασίας. Για να την υπολογίσουμε χρησιμοποιούμε τον παρακάτω τύπο:

$$p_c = 1 - p_{nc} = 1 - (P(X > USL) + P(X < LSL)) = 1 - (1 - P(LSL < X < USL)) =$$

$$P(LSL < X < USL) = \Phi\left(\frac{USL - \mu}{\sigma}\right) - \Phi\left(\frac{LSL - \mu}{\sigma}\right) \Rightarrow p_c =$$

$$\Phi\left(\frac{USL - \mu}{\sigma}\right) - \Phi\left(\frac{LSL - \mu}{\sigma}\right)$$

όπου $\Phi(\cdot)$ είναι η αθροιστική συνάρτηση κατανομής της τυποποιημένης κανονικής κατανομής $N(0,1)$. Εύκολα αποδεικνύεται πως αν γνωρίζουμε την τιμή του δείκτη C_p μπορούμε να υπολογίσουμε το όριο το οποίο είναι αδύνατον να υπερβεί η απόδοση της διεργασίας. Πιο συγκεκριμένα, στην περίπτωση που έχουμε

$$\mu = M = \frac{USL + LSL}{2},$$

τότε διαπιστώνουμε ότι

$$LSL = \mu - 3\sigma C_p$$

και

$$USL = \mu + 3\sigma C_p$$

Επομένως έχουμε ότι

$$p_c = \Phi\left(\frac{USL - \mu}{\sigma}\right) - \Phi\left(\frac{LSL - \mu}{\sigma}\right) = \Phi(3C_p) - \Phi(-3C_p) = 2\Phi(3C_p) - 1$$

Στην γενική περίπτωση κι εφόσον έχουμε ότι

$$LSL < \mu < USL$$

τότε

$$\mu - LSL \leq \frac{USL - LSL}{2}$$

και

$$USL - \mu \leq \frac{USL - LSL}{2}$$

οπότε

$$\Phi\left(\frac{USL - \mu}{\sigma}\right) \leq \Phi(3C_p)$$

και

$$\Phi\left(\frac{LSL - \mu}{\sigma}\right) \leq \Phi(3C_p)$$

Επομένως θα ισχύει ότι

$$p_c = 2\Phi(3C_p) - 1$$

Αυτό σημαίνει ότι η υψηλότερη απόδοση που μπορεί να έχει μια διεργασία είναι

$$2\Phi(3C_p) - 1.$$

Στο σημείο αυτό θα γίνει ερμηνεία των τιμών του δείκτη C_p . Πιο συγκεκριμένα,

- $C_p < 1.33$. Αν ο δείκτης πάρει τιμές μικρότερες από 1.33 η διεργασία θεωρείται ότι χρειάζεται παρακολούθηση κι ειδικά στην περίπτωση όπου πάρει τιμές μικρότερες του 1 σημαίνει ότι η διεργασία θεωρείται μη ικανή και χρειάζεται αναθεώρηση ολόκληρου του σχεδιασμού.
- $C_p \geq 1.33$. Στην περίπτωση που ο δείκτης C_p πάρει τιμές μεγαλύτερες ή ίσες από 1.33 θεωρείται ότι η διεργασία είναι ικανή. Εδώ αξίζει να σημειώσουμε ότι όσο πιο μεγάλη τιμή του δείκτη C_p έχουμε, τόσο πιο ικανή είναι η διεργασία που μελετάμε.

Σύμφωνα με το κλασικό σύγγραμμα του Montgomery (2009) οι ελάχιστες τιμές, που πρέπει να υπερβαίνει η τιμή του δείκτη C_p για να θεωρηθεί μια διεργασία ικανή, διαφέρουν ανάλογα με το είδος της διεργασίας. Για να γίνουμε πιο

συγκεκριμένοι, η ελάχιστη τιμή που πρέπει να υπερβαίνει ο δείκτης C_p για να μιλάμε για ικανή διεργασία είναι:

- 1.33 για μια διεργασία που ήδη υπάρχει
- 1.5 για μια νέα διεργασία
- 1.55 για υπάρχουσα διεργασία που συσχετίζεται με την ασφάλεια του κοινού
- 1.67 για μια νέα διεργασία που σχετίζεται με την ασφάλεια του κοινού

Σε αυτό το σημείο πρέπει να επισημάνουμε ότι για να μπορέσουμε να βγάλουμε ακριβή στατιστικά συμπεράσματα υπολογίζοντας την τιμή του δείκτη ικανότητας C_p μιας διεργασίας θα πρέπει να είμαστε σίγουροι ότι :

- Η διεργασία είναι εντός στατιστικού ελέγχου
- Οι παρατηρούμενες τιμές του χαρακτηριστικού X που μελετάμε σε μια διεργασία να ακολουθούν την κανονική κατανομή $N(\mu, \sigma^2)$ με μέση τιμή μ και διασπορά σ^2).
- Οι παρατηρήσεις του χαρακτηριστικού της διεργασίας που μελετάμε να είναι ανεξάρτητες.

3.1.2 Εκτίμηση του δείκτη C_p

Για να μπορέσουμε να εκτιμήσουμε τον δείκτη ικανότητας C_p θα πρέπει να ανατρέξουμε στον τύπο του ορισμού του που είναι ο ακόλουθος:

$$C_p = \frac{USL - LSL}{6\sigma}$$

Παρατηρούμε λοιπόν ότι η μοναδική παράμετρος που περιέχεται στον ορισμό του δείκτη C_p είναι η τυπική απόκλιση σ της διεργασίας. Συνήθως αυτή είναι μια

παράμετρος άγνωστη και θα πρέπει να την εκτιμήσουμε εφόσον θέλουμε να εκτιμήσουμε το δείκτη C_p .

Έστω λοιπόν ότι μια διεργασία παράγει προϊόντα των οποίων οι παρατηρούμενες τιμές του χαρακτηριστικού X της διεργασίας που μελετάμε ακολουθούν την κανονική κατανομή $N(\mu, \sigma^2)$ με μέση τιμή μ και διασπορά σ^2 . Σε αυτήν την περίπτωση λοιπόν, συλλέγουμε ένα τυχαίο δείγμα X_i μεγέθους n έστω X_1, X_2, \dots, X_n από την διεργασία που παρακολουθούμε και θέλουμε να εκτιμήσουμε τη διασπορά σ^2 με την βοήθεια της αμερόληπτης δειγματικής διασποράς

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

όπου

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

Στη συνέχεια εκτιμούμε την τυπική απόκλιση σ μέσω της δειγματικής απόκλισης S και τον δείκτη C_p μέσω της εκτιμήτριας

$$C_p = \frac{USL - LSL}{6S}$$

Άρα λοιπόν είναι προφανές ότι

$$(n-1) \left(\frac{C_p}{C_p} \right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

οπότε

$$\sqrt{n-1} \frac{C_p}{C_p} \sim \chi_{n-1}$$

Αφού η συνάρτηση πυκνότητας πιθανότητας $f(x)$ της κατανομής χ_n δίνεται από τον τύπο

$$f(x) = \frac{1}{2^{(n/2)-1} \Gamma(n/2)} e^{-x^2/2} x^{n-1}, x, n > 0$$

(βλ. Johnson et al.1994) είναι εύκολο να διαπιστωθεί ότι η συνάρτηση πυκνότητας πιθανότητας $g_{C_p}(y)$ του εκτιμητή C_p δίνεται από τον τύπο

$$g_{C_p}(y) = \frac{2(C_p \sqrt{f/2})^f}{\Gamma(f/2)} y^{-f-1} \cdot e^{\frac{fC_p^2}{2y^2}} I_{(0,\infty)}(y)$$

όπου $f = n - 1$.

Με ολοκλήρωση της προκύπτει ότι

$$E(C_p) = c(n)C_p$$

όπου

$$c(n) = \sqrt{\frac{f}{2}} \cdot \frac{\Gamma\left(\frac{f-1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)} = \sqrt{\frac{n-1}{2}} \cdot \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

και

$$E(C_p^2) = \frac{f}{f-2} C_p^2 = \frac{n-1}{n-3} C_p^2.$$

Συνεπώς

$$\text{Var}(C_p^2) = \left(\frac{f}{f-2} - c^2(n) \right) C_p^2 = \left(\frac{n-1}{n-3} - c^2(n) \right) C_p^2$$

Από τα παραπάνω συμπεραίνουμε ότι ο C_p δεν είναι αμερόληπτος εκτιμητής του C_p , αλλά είναι ασυμπτωτικά αμερόληπτος αφού

$$\lim_{n \rightarrow \infty} c(n) = 1$$

Επίσης το μέσο τετραγωνικό σφάλμα δίνεται από τη σχέση

$$MSE(C_p) = E[(C_p - C_p)^2] = 2C_p^2 \left(\frac{n-2}{n-3} - c(n) \right)$$

και συνεπώς ο \hat{C}_p είναι συνεπής εκτιμητής του C_p .

Οι προαναφερθείσες ιδιότητες του εκτιμητή \hat{C}_p μπορούν να βρεθούν στα άρθρα των Chou et.al. (1989, 1990), Chou and Owen (1989), Kane (1986), Li et al. (1990) και Kotz and Johnson(1993).

3.1.3 Διαστήματα εμπιστοσύνης του δείκτη C_p

Τα διαστήματα εμπιστοσύνης που έχουν αναπτυχθεί για όλους τους δείκτες ικανότητας μιας κανονικής διεργασίας υποθέτουν ότι η τυπική απόκλιση σ εκτιμάται με την ποσότητα $\sigma = S$ και ο μέσος μ εκτιμάται με την ποσότητα $\mu = \bar{X}$ ή $\mu = \bar{\bar{X}}$ ανάλογα με το αν έχουμε ένα ή περισσότερα δείγματα, αντίστοιχα. Έχουμε λοιπόν ότι

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

$$(P(\chi_n^2 \geq \chi_{n,\alpha}^2) = \alpha)$$

ή ισοδύναμα

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

ή ισοδύναμα

$$P\left(\frac{USL-LSL}{6S}\sqrt{\frac{\chi_{n-1,1-(\alpha/2)}^2}{n-1}} \leq \frac{USL-LSL}{6\sigma} \leq \frac{USL-LSL}{6S}\sqrt{\frac{\chi_{n-1,\alpha/2}^2}{n-1}}\right) = 1-\alpha.$$

Συνεπώς ένα $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη C_p , όταν η παράμετρος σ εκτιμηθεί με την ποσότητα $\sigma = S$, δίνεται από την ακόλουθη σχέση (Montgomery, 2009)

$$C_p \frac{\chi_{n-1,1-(\alpha/2)}}{\sqrt{n-1}} \leq C_p \leq C_p \frac{\chi_{n-1,\alpha/2}}{\sqrt{n-1}}.$$

Ο Heavlin (1988) πρότεινε ένα διαφορετικό διάστημα εμπιστοσύνης για τον δείκτη C_p . Πιο συγκεκριμένα χρησιμοποίησε τις ακόλουθες προσεγγίσεις:

$$E(S^{-1}) \cong \left(1 + \frac{3}{4(n-1)}\right) \sigma^{-1}$$

και

$$Var(S^{-1}) \cong \frac{1}{2(n-3)\sigma^2}$$

από τις οποίες προκύπτει ότι

$$Var(C_p) \cong \frac{C_p^2}{2(n-3)} \cdot \left(1 + \frac{6}{n-1}\right).$$

Έτσι ένα εναλλακτικό προσεγγιστικό $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη C_p δίνεται από την ακόλουθη σχέση

$$C_p \left[1 - \left(\frac{1}{2(n-3)} \left(1 + \frac{6}{n-1} \right) \right)^{1/2} z_{\alpha/2} \right] \leq C_p \leq C_p \left[1 + \left(\frac{1}{2(n-3)} \left(1 + \frac{6}{n-1} \right) \right)^{1/2} z_{\alpha/2} \right]$$

(αν $Z \sim N(0,1)$ τότε $P(Z \geq z_\alpha) = \alpha$).

Σημειώνουμε εδώ ότι το παραπάνω διάστημα εμπιστοσύνης υποθέτει ότι η παράμετρος σ έχει εκτιμηθεί με την ποσότητα $\sigma = S$.

Σε ένα τυπικό πρόβλημα ποιοτικού ελέγχου η τυπική απόκλιση μιας διεργασίας εκτιμάται μέσω αμερόληπτων εκτιμητών της. Σε προηγούμενα κεφάλαια αναφέραμε ότι μια αμερόληπτη εκτίμηση του σ επιτυγχάνεται μέσω της στατιστικής R/d_2 (όταν έχουμε ένα δείγμα μεγέθους n) ή γενικότερα μέσω στατιστικής \bar{R}_m/d_2 (για m ανεξάρτητα δείγματα μεγέθους n , $\bar{R}_1 = R$). Στο άρθρο των Li et al.(1990) αναφέρεται ότι το στατιστικό

$$\frac{\sqrt{v}}{c} \cdot \frac{\bar{R}_m}{\sigma} \overset{\alpha}{\sim} \chi_v$$

Όπου το σύμβολο $\overset{\alpha}{\sim}$ σημαίνει «κατανέμεται προσεγγιστικά», όπου τα c και v είναι σταθερές που σχετίζονται με τις δύο πρώτες ροπές του εύρους R_1 . Χρησιμοποιώντας το αποτέλεσμα οι Li et al.(1990) έδωσαν μονόπλευρα διαστήματα εμπιστοσύνης για το δείκτη C_p της μορφής $[c, \infty)$ χρησιμοποιώντας ως εκτιμητή του C_p την ποσότητα

$$C_p = \frac{USL - LSL}{6\bar{R}_m / d_2}$$

Το $100\alpha\%$ διάστημα εμπιστοσύνης για το δείκτη C_p με κάτω όριο c_0 ικανοποιεί τη σχέση

$$P(C_p \geq c_0) = \alpha$$

Συνεπώς

$$1 - \alpha = P(C_p \leq c_0) = P\left(\frac{C_p}{C_p} \leq \frac{c_0}{C_p}\right) = P\left(\frac{\bar{R}_m}{\sigma d_2} \leq \frac{c_0}{C_p}\right) = P\left(\frac{\bar{R}_m}{\sigma} \leq \frac{c_0 d_2}{C_p}\right) =$$

$$P\left(\frac{\sqrt{v}}{c} \cdot \frac{\bar{R}_m}{\sigma} \leq \frac{\sqrt{v}}{c} \cdot \frac{c_0 d_2}{C_p}\right) = P\left(\chi_v \leq \frac{\sqrt{v}}{c} \frac{c_0 d_2}{C_p}\right)$$

Οπότε

$$c_0 = \frac{c C_p}{\sqrt{v d_2}} \sqrt{\chi_{v, \alpha}^2} = \frac{c C_p}{\sqrt{v d_2}} \chi_{v, \alpha}$$

3.1.4 Έλεγχοι υποθέσεων του δείκτη C_p

Σύμφωνα με τον Kane (1986), στην ανάλυση της ικανότητας μιας διεργασίας, ο έλεγχος της διεργασίας που μας ενδιαφέρει είναι ο εξής: Η διεργασία δεν είναι ικανή (H_0) – Η διεργασία είναι ικανή (H_1). Η παραπάνω υπόθεση είναι ισοδύναμη με την υπόθεση

$$H_0 : C_p \leq c_0 - H_1 : C_p > c_0$$

Όπου c_0 είναι μια κρίσιμη τιμή για το δείκτη C_p που συνήθως έχει τιμή 1.33 ή 1.67. Η περιοχή απόρριψης της μηδενικής υπόθεσης (κρίσιμη περιοχή) είναι της μορφής $C_p > c$, όπου c είναι μια θετική τιμή (κρίσιμη τιμή). Για σφάλμα τύπου I ίσο με α (η διεργασία να φαίνεται εκτός αλλά να είναι εντός) προκύπτει ότι

$$\alpha = P_{C_p=c_0}(C_p > c) = P_{C_p=c_0} \left[(n-1)(C_p / C_p)^2 \leq (n-1)(C_p / c)^2 \right] = P \left[\chi_{n-1}^2 \leq (n-1)(c_0 / c)^2 \right]$$

οπότε

$$c = c_0 \sqrt{\frac{n-1}{\chi_{n-1,1-\alpha}^2}}$$

Η συνάρτηση ισχύος του ελέγχου δίνεται προφανώς από τον τύπο

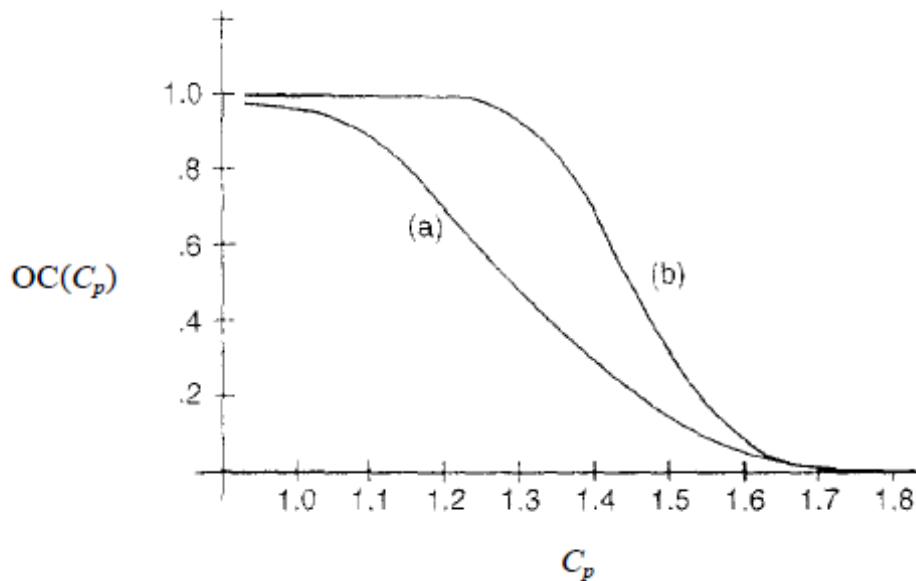
$$P(\text{Απόρριψης } H_0) = \pi(C_p) = P(C_p > c) = P[\chi_{n-1}^2 \leq (n-1)(C_p/c)^2]$$

και η λειτουργική χαρακτηριστική καμπύλη του ελέγχου δίνεται από τη σχέση

$$OC(C_p) = 1 - \pi(C_p)$$

Στο ακόλουθο διάγραμμα δίνονται δύο γραφικές παραστάσεις της λειτουργικής χαρακτηριστικής καμπύλης (Kane 1986).

Γράφημα 3.1: Λειτουργική χαρακτηριστική καμπύλη για
 a) $n = 30$, $c = 1.33$ και b) $n = 70$, $c = 1.46$.



Σύμφωνα με τον Kane (1986) η επιλογή της κατάλληλης κρίσιμης τιμής c πρέπει να βασιστεί σε μια αποδεκτή τιμή $C_p(\text{high})$ του δείκτη C_p και σε μια μη αποδεκτή τιμή $C_p(\text{low})$ του δείκτη C_p , με $C_p(\text{low}) \leq C_p(\text{high})$. Έτσι αν ο δείκτης C_p έχει τιμή μεγαλύτερη της τιμής $C_p(\text{high})$ τότε η διεργασία θα θεωρείται ικανή, ενώ αν ο δείκτης έχει τιμή C_p μικρότερη της τιμής $C_p(\text{low})$ τότε η διεργασία θα θεωρείται μη ικανή. Για τον Kane (1986) η υπόθεση:

H_0 : Η διεργασία δεν είναι ικανή - H_1 : Η διεργασία είναι ικανή
παίρνει τη μορφή

$$H_0 : C_p < C_p(\text{low}) - H_1 : C_p > C_p(\text{high})$$

Από τα προαναφερθέντα είναι σαφής η σχέση της παραπάνω προσέγγισης με την προσέγγιση που ακολουθείται στη δειγματοληψία αποδοχής (acceptance sampling) όπου η έννοια του όρου $C_p(\text{high})$ είναι ισοδύναμη με το αποδεκτό επίπεδο ποιότητας (acceptable quality level, AQL) και η έννοια του όρου $C_p(\text{low})$ είναι ισοδύναμη με το απορριπτέο επίπεδο ποιότητας (rejectable quality level, RQL ή LTPD). Στο αντίστοιχο πρόβλημα της δειγματοληψίας αποδοχής απαιτούμε η λειτουργική χαρακτηριστική καμπύλη να επιβεβαιώνει τα σημεία $(RQL, 1-\alpha)$ και (AQL, β) , όπου τα α και β παριστάνουν το ρίσκο του παραγωγού και του καταναλωτή, αντίστοιχα. Έτσι εδώ απαιτούμε η συνάρτηση ισχύος να επιβεβαιώνει τα σημεία $(C_p(\text{low}), \alpha)$ και $(C_p(\text{high}), 1-\beta)$. Συνεπώς απαιτούμε να ισχύουν οι σχέσεις

$$\alpha = \pi(C_p(\text{low})) = P\left[\chi_{n-1}^2 \leq (n-1)(C_p(\text{high})/c)^2\right]$$

και

$$1-\beta = \pi(C_p(\text{high})) = P\left[\chi_{n-1}^2 \leq (n-1)(C_p(\text{high})/c)^2\right].$$

Από τις δύο παραπάνω σχέσεις προκύπτει ότι

$$\frac{C_p(\text{high})}{C_p(\text{low})} = \sqrt{\frac{\chi_{n-1,\beta}^2}{\chi_{n-1,1-\alpha}^2}}$$

και

$$c = C_p(\text{low}) \sqrt{\frac{n-1}{\chi_{n-1,1-\alpha}^2}}$$

Ο Kane (1986) έδωσε τον παρακάτω πίνακα, σύμφωνα με τον οποίο μπορούμε να καθορίσουμε το μέγεθος του δείγματος n και την κρίσιμη τιμή c για τον έλεγχο της υπόθεσης

$$H_0 : C_p < C_p(\text{low}) - H_1 : C_p > C_p(\text{high})$$

αν γνωρίζουμε τα δεδομένα α , β , $C_p(\text{high})$ και $C_p(\text{low})$.

Πίνακας 3.1: Τιμές των $C_p(\text{high})/C_p(\text{low})$ και $c/C_p(\text{low})$ για διάφορες τιμές των n, α, β .

n	$\alpha = \beta = 0.10$	$\alpha = \beta = 0.10$	$\alpha = \beta = 0.05$	$\alpha = \beta = 0.05$
n	$C_p(\text{high})/C_p(\text{low})$	$c/C_p(\text{low})$	$C_p(\text{high})/C_p(\text{low})$	$c/C_p(\text{low})$
10	1.88	1.47	2.26	1.65
20	1.53	1.28	1.73	1.37
30	1.41	1.21	1.55	1.28
40	1.34	1.18	1.46	1.23
50	1.30	1.15	1.40	1.20
60	1.27	1.14	1.36	1.18
70	1.25	1.13	1.33	1.16
80	1.23	1.12	1.30	1.15
90	1.21	1.11	1.28	1.14
100	1.20	1.10	1.26	1.13

Για παράδειγμα αν $\alpha = \beta = 0.05$, $C_p(\text{high}) = 1.6$ και $C_p(\text{low}) = 1.2$, τότε $C_p(\text{high}) / C_p(\text{low}) = 1.392$ και από τον πίνακα για $n = 70$ υπάρχει κίνδυνος 5% να θεωρηθεί η διεργασία μη ικανή, που σημαίνει ότι αποδεχόμαστε την υπόθεση H_0 , με τιμή του δείκτη C_p πάνω από 1.6 και κίνδυνος 5% να κριθεί ότι η διεργασία είναι ικανή (απόρριψη H_0) με τιμή του δείκτη C_p κάτω από 1.2.

Σε αυτό το σημείο, αξίζει να σημειώσουμε ότι σύμφωνα με την Hoffman (1993) ακολουθεί μια γενική προσέγγιση για τον έλεγχο της υπόθεσης

$$H_0 : C_p < C_p(\text{low})$$

ο οποίος βασίζεται στην κατανομή της ποσότητας σ / S .

Σε αντίθεση με τον Kane 1986, που ανέπτυξε μια μελέτη για ένα τυχαίο δείγμα μεγέθους n οι Kirmani et al (1991), ασχολήθηκαν με την περίπτωση όπου έχουμε δεδομένα από m ανεξάρτητα τυχαία δείγματα μεγέθους n το καθένα.. Ο εκτιμητής C_p που χρησιμοποιήθηκε είναι ο ακόλουθος.

$$C_p = \frac{(USL - LSL) d_p}{6}$$

όπου

$$d_p = \sqrt{\frac{m(n-1)-1}{m(n-1)}} \cdot \frac{\varepsilon_{m(n-1)-1}}{S_p}$$

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

$$\varepsilon_k = \sqrt{\frac{2}{k}} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}$$

Οι Kirmani et al. (1991) χρησιμοποιώντας τη σχέση

$$\frac{m(n-1)S_p^2}{\sigma^2} \sim \chi_{m(n-1)}^2$$

κατέληξαν στην κρίσιμη τιμή

$$c^2 = \frac{c_0^2(m(n-1)-1)\varepsilon_{m(n-1)-1}}{\chi_{m(n-1),1-\alpha}^2}$$

για τον έλεγχο της υπόθεσης

$$H_0 : C_p \leq c_0 - H_1 : C_p > c_0$$

σε επίπεδο σημαντικότητας α .

Μια άλλη προσέγγιση στο πρόβλημα που μελέτησαν οι Kirmani et al (1991) είναι αυτή που πρότεινε ο Kocherlakota (1992), ο οποίος χρησιμοποιεί τον ακόλουθο εκτιμητή.

$$C_p = \frac{(USL - LSL)}{6\sigma^2}$$

όπου

$$\sigma = \frac{\bar{S}_m}{\varepsilon_{n-1}} = \frac{(S_1 + S_2 + \dots + S_m) / m}{\varepsilon_{n-1}}$$

Ως εκτιμητή του δείκτη C_p . Ο Kocherlakota (1992) παρατηρώντας ότι η ασυμπτωτική κατανομή του $\frac{\bar{S}_m}{\varepsilon_{n-1}}$ είναι η

$$\frac{\bar{S}_m}{\varepsilon_{n-1}} \stackrel{a}{\sim} N\left(\sigma, \frac{\sigma}{\sqrt{m}} \sqrt{\frac{1}{\varepsilon_{n-1}^2} - 1}\right)$$

κατέληξε στην προσεγγιστική κρίσιμη τιμή

$$c = \frac{c_0}{1 + z_{1-\alpha} \sqrt{\frac{1 - \varepsilon_{n-1}^2}{m\varepsilon_{n-1}^2}}}$$

Αξίζει να σημειώσουμε ότι αυτή η μέθοδος θα μπορούσε να εφαρμοστεί όταν χρησιμοποιούμε \bar{X} και S διάγραμμα ελέγχου για την παρακολούθηση της διεργασίας.

3.1.5 Αδυναμίες του δείκτη C_p

Το βασικό μειονέκτημα του δείκτη C_p είναι ότι δε λαμβάνει υπόψη το μέσο της διεργασίας. Αυτό έχει ως συνέπεια να μην επηρεάζεται από την διαφορά του μέσου από την κεντρική τιμή της περιοχής προδιαγραφών (τιμή στόχο T) της διεργασίας. Επιπλέον, για να υπολογιστεί ο δείκτης C_p θα πρέπει να υπάρχουν και τα δύο όρια προδιαγραφών του προϊόντος, τα οποία όμως δεν καθορίζονται πάντα και τα δύο. Η αδυναμία του δείκτη, να παρέχει ουσιαστικές ενδείξεις για την πραγματική απόδοση της διεργασίας, ήταν η αιτία για την εισαγωγή του δείκτη C_{pk} που θα αναλύσουμε στην επόμενη ενότητα.

3.2 Ο δείκτης ικανότητας C_{pk}

Σε αυτήν την ενότητα, θα μελετήσουμε τον δείκτη ικανότητας C_{pk} , ο οποίος είναι ο δεύτερος δημοφιλέστερος δείκτης ικανότητας μια διεργασίας, μετά τον δείκτη C_p . Όπως αναφέραμε παραπάνω αιτία για την εισαγωγή αυτού του δείκτη ήταν οι αδυναμίες που παρουσίασε ο δείκτης C_p . Μια από τις βασικές αδυναμίες του ήταν το γεγονός ότι κατά τον υπολογισμό του δε λαμβάνει υπόψη του το μέσο της διεργασίας.

Αντίθετα λοιπόν, ο δείκτης επίδοσης C_{pk} κατά τον υπολογισμό του, εκτός από την τυπική απόκλιση σ , λαμβάνει υπόψη του και το μέσον μ της διεργασίας.

3.2.1 Ορισμός και ιδιότητες του δείκτη ικανότητας C_{pk}

Όπως αναφέραμε στο προηγούμενο κεφάλαιο, ένα μειονέκτημα του δείκτη C_p είναι το γεγονός ότι για να οριστεί χρειάζεται και τα δύο όρια προδιαγραφών (άνω όριο προδιαγραφών USL και κάτω όριο προδιαγραφών LSL), επομένως στην περίπτωση που μας δίνεται μόνο το ένα από τα δύο όρια είναι αδύνατον να τον υπολογίσουμε. Για αυτό το λόγο έχει προταθεί ο δείκτης C_{pk} , τον οποίο θα αναλύσουμε σε αυτό το κεφάλαιο. Αξίζει να σημειώσουμε σε αυτό το σημείο ότι ο δείκτης C_{pk} είναι ένας συνδυασμός από τους δείκτες C_{pl} και C_{pu} , αφού ο ορισμός του δείκτη C_{pk} βασίζεται στους τύπους ορισμού αυτών των δύο δεικτών. Οι ορισμοί των C_{pl} και C_{pu} δίνονται από τους ακόλουθους τύπους.

$$C_{pl} = \frac{\mu - LSL}{3\sigma}$$

$$C_{pu} = \frac{USL - \mu}{3\sigma}.$$

Εδώ παρατηρούμε ότι ο δείκτης C_{pl} μπορεί να χρησιμοποιηθεί όταν γνωρίζουμε το κάτω όριο προδιαγραφών LSL , ενώ ο δείκτης C_{pu} έχει προταθεί για τις περιπτώσεις όπου έχει καθοριστεί μόνο το ανώτερο όριο προδιαγραφών USL .

Συγκεκριμένα τώρα ο δείκτης ικανότητας C_{pk} μιας διεργασίας ορίζεται ως εξής:

$$C_{pk} = \min(C_{pl}, C_{pu})$$

Εύκολα αποδεικνύεται ότι ο δείκτης C_{pk} ισούται αλγεβρικά με τον αριθμητικό μέσο των δεικτών C_{pl} και C_{pu} , με τη διαφορά όμως ότι οι δύο τελευταίοι λαμβάνουν υπόψη και τις δύο παραμέτρους μιας διεργασίας, δηλαδή το μέσον και τη διασπορά, συγκριτικά με τον δείκτη C_p .

Είναι φανερό ότι ο δείκτης C_{pk} μπορεί να γραφεί και στη μορφή

$$C_{pk} = \min(C_{pl}, C_{pu}) = \min\left(\frac{\mu - LSL}{3\sigma}, \frac{USL - \mu}{3\sigma}\right) = \frac{d - |\mu - M|}{3\sigma}$$

όπου

$$d = \frac{USL - LSL}{2}$$

και

$$M = \frac{USL + LSL}{2}.$$

Δηλαδή το M ορίζεται ως η κεντρική τιμή του διαστήματος που ορίζουν τα όρια προδιαγραφών σύμφωνα με τον Kotz and Lovelace (1998).

Με τη βοήθεια του πρώτου ορισμού, προκύπτει κι ένας δεύτερος ορισμός ως εξής:

$$C_{pk} = \frac{d - |\mu - M|}{3\sigma} = \frac{\frac{USL - LSL}{2} - \left|\mu - \frac{USL + LSL}{2}\right|}{3\sigma} = C_p \left[1 - \frac{\left|\mu - \frac{USL + LSL}{2}\right|}{d} \right] = C_p(1 - k)$$

όπου (βλέπε Αντζουλάκος, 2006).

$$k = \frac{\left|\mu - \frac{USL + LSL}{2}\right|}{d}.$$

Επομένως ένας δεύτερος ισοδύναμος ορισμός για τον δείκτη ικανότητας C_{pk} είναι ο εξής:

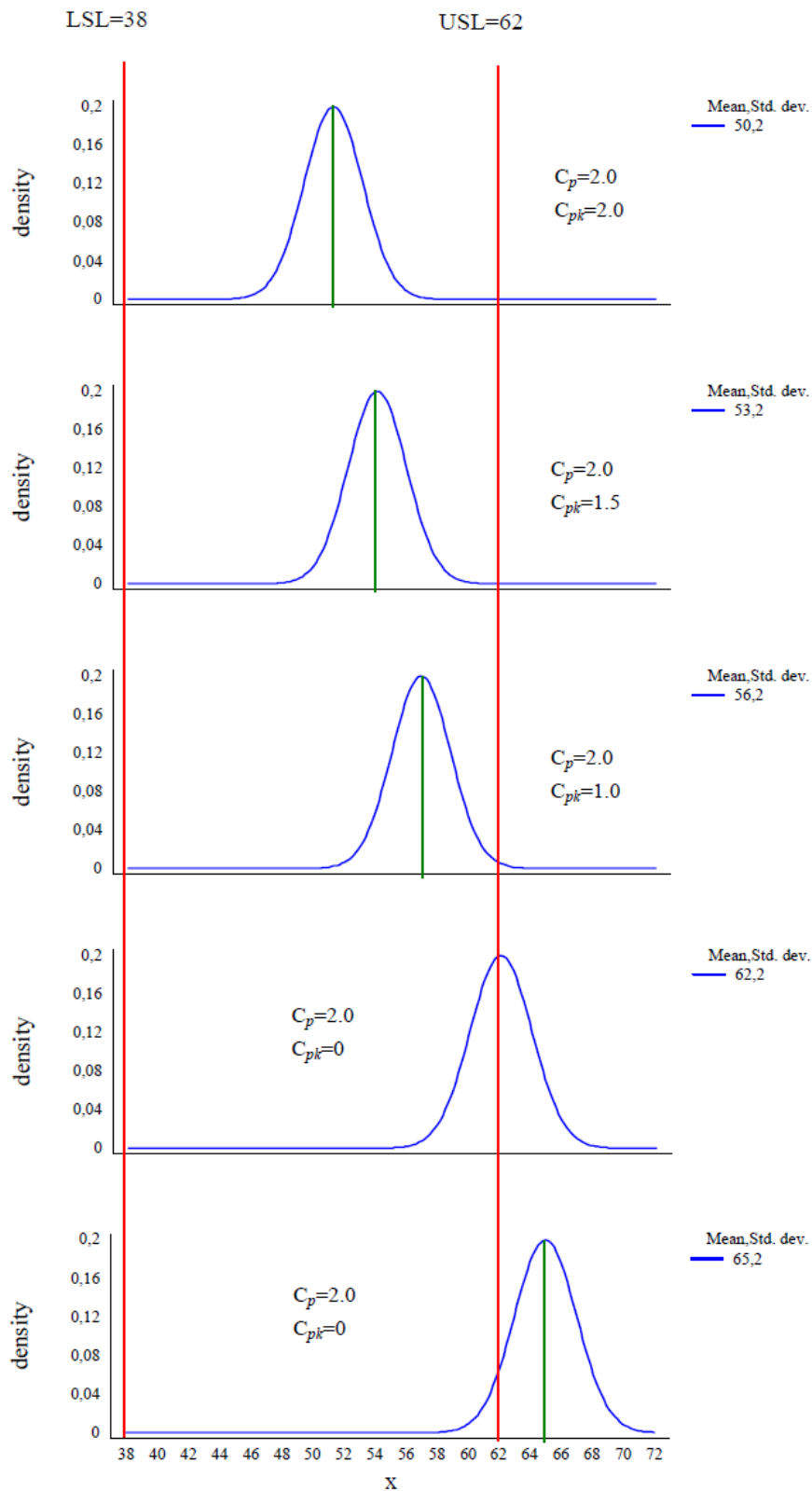
$$C_{pk} = C_p(1-k)$$

ο οποίος περιγράφει την ικανότητα της διεργασίας λαμβάνοντας υπόψη στον υπολογισμό του, όχι τη μεταβλητότητα της διεργασίας αλλά τη θέση της ενώ επιπλέον μετράει την απόσταση του μέσου της διεργασίας από το μέσο του διαστήματος που ορίζουν το άνω και το κάτω όριο προδιαγραφών, σχετικά με αυτά τα όρια.

Αν υποθέσουμε πως ο μέσος βρίσκεται εντός της περιοχής που ορίζουν τα όρια προδιαγραφών, τότε σύμφωνα με τον ορισμό μπορούμε να συμπεράνουμε ότι η τιμή του k βρίσκεται πάντα στο διάστημα $[0,1]$. Στην αντίθετη περίπτωση, όταν δηλαδή ο μέσος είναι εκτός του διαστήματος προδιαγραφών, η τιμή του C_{pk} είναι ίση με το μηδέν. Άρα ο δείκτης C_{pk} δεν μπορεί να υπερβεί την τιμή του C_p , ενώ όταν η τιμή του C_{pk} πάρει την ίση τιμή με αυτή του C_p , τότε σημαίνει ότι ο μέσος της διεργασίας ισούται με το μέσο του διαστήματος που ορίζουν τα όρια προδιαγραφών, δηλαδή $k = 0$.

Στο παρακάτω γράφημα δίνονται μερικές διεργασίες ($LSL = 38$, $USL = 62$) με κοινή τυπική απόκλιση $\sigma = 2$ αλλά με διαφορετικούς μέσους μ ($\mu = 50, 53, 56, 62, 65$). Αν και όλες οι διεργασίες έχουν κοινό δείκτη C_p , δε συμβαίνει το ίδιο και με τον δείκτη C_{pk} . Όσο το μ απομακρύνεται από το M τόσο η τιμή του δείκτη ικανότητας C_{pk} μειώνεται και τόσο λιγότερο ικανή είναι η διεργασία. Προφανώς η ικανότερη διεργασία είναι αυτή που έχει τον υψηλότερο δείκτη C_{pk} . Συνεπώς ο δείκτης C_p δεν μας δίνει επαρκείς πληροφορίες σε αυτή την περίπτωση για το ποία από τις διεργασίες είναι ικανότερη, σε αντίθεση με τον δείκτη C_{pk} . Άλλωστε η πιο βασική αδυναμία του δείκτη C_p είναι ότι δεν λαμβάνει υπόψη του το μέσον της διεργασίας και αυτή την αδυναμία έρχεται να καλύψει ο δείκτης C_{pk} .

Γράφημα 3.2: Διεργασίες με κοινό δείκτη C_p , αλλά διαφορετικό δείκτη C_{pk}



Αν η τιμή στόχος της διεργασίας δε συμπίπτει με το μέσον του διαστήματος που ορίζουν τα όρια προδιαγραφών, χρησιμοποιείται και ο δείκτης

$$k' = \frac{|T - \mu|}{\min(T - LSL, USL - T)}$$

Σε αυτό το σημείο μπορούμε να διακρίνουμε τις τρεις παρακάτω περιπτώσεις

1. Αν $k' = 0$, τότε ο μέσος της διεργασίας ισούται με την τιμή στόχο
2. Αν $0 < k' < 1$, τότε ο μέσος της διεργασίας βρίσκεται ανάμεσα στη τιμή στόχο και σε ένα από τα δύο όρια προδιαγραφών. Σε αυτό το σημείο αξίζει να σημειώσουμε ότι η απόσταση του μέσου από την τιμή στόχο μετριέται από το δείκτη k και τέλος.
3. Αν $k' = 1$, τότε η διεργασία είναι κεντραρισμένη, σε ένα από τα δύο όρια προδιαγραφών.

Με μια προσεκτική ματιά, βλέπουμε ότι ο δείκτης C_{pk} αποτελείται από γινόμενο δύο παραγόντων, από τον δείκτη C_p που μετράει το μέγεθος της διασποράς και τον παράγοντα $(1 - k)$, που λαμβάνει υπόψη του τη θέση της διεργασίας, δηλαδή μετράει την απόσταση του μέσου από την κεντρική τιμή της περιοχής των προδιαγραφών.

Εδώ πρέπει να τονίσουμε ότι για να υπολογίσουμε την ικανότητα και την απόδοση μιας διεργασίας μόνο σε συνδυασμό με τον δείκτη C_p . Για αυτό το λόγο δεν μπορούμε να ισχυριστούμε ότι ο δείκτης C_{pk} είναι επαρκής για τον ακριβή υπολογισμό της απόδοσης μιας διεργασίας.

Στην περίπτωση που μια διεργασία αποδίδει προϊόντα των οποίων οι τιμές του ποιοτικού χαρακτηριστικού που μελετάμε κατανέμονται κανονικά με μέση τιμή μ και τυπική απόκλιση σ , η πιθανότητα να παραχθεί ελαττωματικό προϊόν, δηλαδή εκτός των ορίων προδιαγραφών θα ικανοποιεί την παρακάτω ανίσωση.

$$2\Phi(3C_{pk}) \leq p_c \leq \Phi(3C_{pk})$$

όπου $\Phi(\cdot)$ είναι η αθροιστική συνάρτηση κατανομής της τυποποιημένης κανονικής κατανομής $N(0,1)$ (βλέπε Kotz and Lovelace 1998)

Θα δείξουμε στη συνέχεια ότι η απόδοση μιας διεργασίας ισούται με

$$p_c = \Phi[3(2C_p - C_{pk})] - \Phi(-3C_{pk})$$

Ας υποθέσουμε αρχικά ότι $LSL \leq \mu \leq M$. Στην περίπτωση αυτή θα έχουμε

$$C_p = \frac{USL - LSL}{6\sigma}$$

και

$$C_{pk} = \frac{\mu - LSL}{3\sigma}.$$

Επομένως η απόδοση της διεργασίας είναι ίση με

$$\Phi\left(\frac{USL - \mu}{\sigma}\right) - \Phi\left(\frac{LSL - \mu}{\sigma}\right)$$

Επίσης έχουμε ότι

$$2C_p - C_{pk} = \frac{USL - LSL}{3\sigma} - \frac{\mu - LSL}{3\sigma} = \frac{USL - \mu}{3\sigma}.$$

Οπότε

$$\Phi\left(\frac{USL - \mu}{\sigma}\right) - \Phi\left(\frac{LSL - \mu}{\sigma}\right) = \Phi\left(3\frac{USL - \mu}{3\sigma}\right) - \Phi\left(-3\frac{\mu - LSL}{3\sigma}\right) = \Phi[3(2C_p - C_{pk})] - \Phi(-3C_{pk})$$

Παρόμοια γίνεται η απόδειξη για τη περίπτωση που $M \leq \mu \leq USL$. Επομένως η απόδοση μιας διαδικασίας ισούται με

$$p_c = \Phi[3(2C_p - C_{pk})] - \Phi(-3C_{pk})$$

Ενδεικτικά αναφέρουμε ότι συνήθως η ελάχιστη τυπική τιμή που πρέπει να έχει ο δείκτης C_{pk} μιας διεργασίας είναι 1.33.

3.2.2 Εκτίμηση του δείκτη ικανότητας C_{pk}

Για να μπορέσουμε να εκτιμήσουμε τον δείκτη C_{pk} , πρέπει να κάνουμε αντικατάσταση τη μέση τιμή με τον δειγματικό μέσο και την τυπική απόκλιση με την δειγματική τυπική απόκλιση. Στην προηγούμενη παράγραφο, αναφέραμε τον ορισμό του δείκτη C_{pk} , τον οποίο θα χρησιμοποιήσουμε για να καταλήξουμε στην ακόλουθη εκτιμήτρια.

$$C_{pk} = \min(C_{pl}, C_{pu}) = \min\left(\frac{\bar{X} - LSL}{3S}, \frac{USL - \bar{X}}{3S}\right) = \frac{d - |\bar{X} - M|}{3S}.$$

Ακολουθώντας όμως τον δεύτερο ορισμό του δείκτη C_{pk} και δουλεύοντας ανάλογα, δηλαδή αντικαθιστώντας την τυπική απόκλιση με τη δειγματική τυπική απόκλιση και το μέσο με το δειγματικό μέσο, καταλήγουμε στην εκτιμήτρια

$$C_{pk} = C_p (1 - k)$$

όπου

$$k = \frac{2|M - \bar{X}|}{USL - LSL}.$$

Οι δύο παραπάνω εκτιμήτριες είναι ισοδύναμες και για το λόγο αυτό χρησιμοποιήσαμε τον ίδιο συμβολισμό.

Οι δείκτες C_{pl} και C_{pu} μπορούν να εκτιμηθούν μέσω των τύπων

$$C_{pl} = \frac{\bar{X} - LSL}{3S}$$

και

$$C_{pu} = \frac{USL - \bar{X}}{3S}$$

αντίστοιχα. Εδώ αξίζει να σχολιάσουμε ότι οι δύο αυτές εκτιμήτριες δεν είναι αμερόληπτες εκτιμήτριες των αντίστοιχων δεικτών αφού

$$E\left(\frac{1}{S}\right) \neq E(S).$$

3.2.3 Διαστήματα εμπιστοσύνης και έλεγχοι υποθέσεων του δείκτη ικανότητας C_{pk}

Η κατασκευή διαστημάτων εμπιστοσύνης για τον δείκτη C_{pk} απασχόλησε πολλούς ερευνητές, οι οποίοι έχουν δώσει και διάφορες προσεγγίσεις. Αυτό συνέβη διότι η κατασκευή τους δεν είναι εύκολη υπόθεση αφού η κατανομή του εκτιμητή C_{pk} περιλαμβάνει την από κοινού κατανομή δύο μη κεντρικών t κατανομών. Ο Heavlin (1988) πρότεινε το ακόλουθο $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη ικανότητας C_{pk}

$$C_{pk} - z_{\alpha/2} \sqrt{\frac{n-1}{9n(n-3)} + C_{pk}^2 \frac{1}{2(n-3)} \left(1 + \frac{6}{n-1}\right)}$$

και

$$C_{pk} + z_{\alpha/2} \sqrt{\frac{n-1}{9n(n-3)} + C_{pk}^2 \frac{1}{2(n-3)} \left(1 + \frac{6}{n-1}\right)}$$

Οι Franklin and Wasserman (1992) πρότειναν το ακόλουθο κάτω $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη ικανότητας C_{pk}

$$\left(C_{pk} - z_{\alpha} \sqrt{\frac{1}{9n} + C_{pk}^2 \frac{1}{2(n-1)}}, +\infty \right)$$

το οποίο δίνει αρκετά ακριβή αποτελέσματα $n \geq 30$. Ο Bissel (1990) υποθέτοντας ότι η κατανομή του εκτιμητή C_{pk} είναι κανονική πρότεινε το ακόλουθο διάστημα εμπιστοσύνης 95% για τον δείκτη C_{pk}

$$C_{pk} \left(1 \pm 1.96 \sqrt{\frac{1}{9nC_{pk}^2} + \frac{1}{2(n-1)}} \right)$$

που παρέχει ικανοποιητική ακρίβεια. Οι Nagata and Nagahata (1992) ανέπτυξαν το ακόλουθο $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για το δείκτη C_{pk}

$$\left(C_{pk} - z_{1-\alpha/2} \sqrt{\frac{C_{pk}^2}{2(n-1)} + \frac{1}{9n}}, C_{pk} + z_{1-\alpha/2} \sqrt{\frac{C_{pk}^2}{2(n-1)} + \frac{1}{9n}} \right).$$

Αργότερα οι Nagata and Nagahata (1994) πρότειναν το ακόλουθο βελτιωμένο $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη C_{pk}

$$\left(\sqrt{1 - \frac{2}{5(n-1)}} C_{pk} - z_{1-\alpha/2} \sqrt{\frac{C_{pk}^2}{2(n-1)} + \frac{1}{9n}}, \sqrt{1 - \frac{2}{5(n-1)}} C_{pk} + z_{1-\alpha/2} \sqrt{\frac{C_{pk}^2}{2(n-1)} + \frac{1}{9n}} \right)$$

Οι Kusler and Hurley (1992) πρότειναν κάτω $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη ικανότητας C_{pk}

$$\left(C_{pk} \left[\frac{1 - z_\alpha}{(2n-2)^{1/2}} \right], +\infty \right).$$

Σε αυτό το σημείο αξίζει να αναφέρουμε ότι τα διαστήματα εμπιστοσύνης που αναπτύχθηκαν για τον δείκτη ικανότητας C_{pk} μιας κανονικής διεργασίας υποθέτουν ότι η τυπική απόκλιση σ εκτιμάται με την ποσότητα $\sigma = S$. Οι Li et al. (1990) έδωσαν πίνακες με κατώτερα όρια εμπιστοσύνης για το δείκτη C_{pk} στην περίπτωση που η παράμετρος σ εκτιμάται από την ποσότητα $\sigma = \bar{R} / d_2$.

Στην ανάλυση της ικανότητας μιας διεργασίας, ο έλεγχος υπόθεσης που μας ενδιαφέρει είναι ο ακόλουθος

H_0 : Η διεργασία δεν είναι ικανή - H_1 : Η διεργασία είναι ικανή.

Η παραπάνω υπόθεση είναι ισοδύναμη με την υπόθεση

$$H_0 : C_{pk} \leq c_0 - H_1 : C_{pk} > c_0$$

Όπου c_0 είναι μια κρίσιμη τιμή για τον δείκτη C_{pk} που συνήθως έχει τιμή 1.33 ή 1.67. Η περιοχή απόρριψης της μηδενικής υπόθεσης (κρίσιμη περιοχή) είναι της μορφής $C_{pk} > c$, όπου c είναι μια θετική τιμή. Για τον παραπάνω έλεγχο δεν υπάρχει σχετική βιβλιογραφία.

3.2.4. Αδυναμίες του δείκτη ικανότητας C_{pk}

Όπως έχουμε ήδη αναφέρει, οι πληροφορίες που μας δίνει ο δείκτης C_{pk} , αναφέρονται κυρίως στην ικανότητα της διεργασίας προς τη μεριά εκείνου του ορίου προδιαγραφών που παράγει περισσότερα ελαττωματικά προϊόντα. Ο δείκτης C_{pk} , σε αντίθεση με τον δείκτη ικανότητας C_p , παρόλο που λαμβάνει υπόψη του και τις δύο παραμέτρους της διεργασίας μ και σ , δε μας πληροφορεί για το πόσο κοντά βρίσκεται ο μέσος στην κεντρική τιμή της περιοχής προδιαγραφών. Επιπλέον, ακόμα μια αδυναμία του είναι το γεγονός ότι μπορεί να αυξηθεί ανεξέλεγκτα, ανεξάρτητα από την τιμή του μέσου. Αυτό συμβαίνει επειδή η τυπική απόκλιση εκτός του ότι αγνοεί την τιμή στόχο, τείνει να γίνει πολύ μικρή. Για αυτό ακριβώς το λόγο, υπάρχει ο δείκτης C_{pm} , ο οποίος θα μελετηθεί στο επόμενο κεφάλαιο. Τέλος, να σημειώσουμε ότι ο δείκτης αυτός έχει νόημα να χρησιμοποιηθεί, μόνο όταν έχουμε κανονικά, ανεξάρτητα δεδομένα, υπό στατιστικό έλεγχο. Χρειάζεται επίσης μεγάλο μέγεθος δείγματος για να μπορέσουμε να τον εκτιμήσουμε με ακρίβεια (Chou et al, 1990).

3.3 Ο δείκτης ικανότητας C_{pm}

Όπως αναφέραμε στο προηγούμενο κεφάλαιο, οι δείκτες C_p και C_{pk} παρά το γεγονός ότι είναι ευρέως διαδεδομένοι αλλά και οι πιο δημοφιλείς από τους δείκτες ικανότητας, παρουσιάζουν ορισμένες αδυναμίες. Μία από αυτές είναι ότι δε λαμβάνουν υπόψη τους την προκαθορισμένη τιμή στόχο T της διεργασίας. Οι δείκτες C_p και C_{pk} παρόλο που μάς ενημερώνουν για την ικανότητα της διεργασίας να παράγει προϊόντα μέσα στα όρια προδιαγραφών, δε μας δίνουν καμία πληροφορία σχετικά με το πόσο κοντά στην τιμή στόχο βρίσκονται οι παρατηρήσεις μας. Αυτός ακριβώς είναι και ο λόγος που εμφανίστηκε ο δείκτης C_{pm} .

Οι Hsiang and Taguchi (1985) ήταν οι πρώτοι που ασχολήθηκαν με το δείκτη C_{pm} , ενώ αργότερα οι Chan et al. (1988) ήταν αυτοί που ουσιαστικά καθιέρωσαν τον δείκτη ικανότητας C_{pm} . Ο συγκεκριμένος δείκτης διαφέρει από τους άλλους δύο που αναφέραμε, επειδή μετράει την απόδοση της διεργασίας με διαφορετικό τρόπο.

Ο δείκτης C_{pm} , όχι μόνο δηλώνει την ικανότητα μιας διεργασίας, αλλά και την ποιότητα των παραγόμενων προϊόντων. Σύμφωνα με το δείκτη ικανότητας C_{pm} , για να παράγονται προϊόντα υψηλής ποιότητας θα πρέπει ο μέσος της κατανομής να βρίσκεται όσο το δυνατόν πιο κοντά στην τιμή στόχο T . Αν δεν συμβαίνει αυτό, στην περίπτωση δηλαδή που οι περισσότερες παρατηρήσεις της διεργασίας είναι απομακρυσμένες από την τιμή στόχο, δηλαδή βρίσκονται πιο κοντά στα όρια προδιαγραφών, σημαίνει ότι η διεργασία παράγει προϊόντα χαμηλής ποιότητας. Επομένως, για να είναι μια διαδικασία ικανή δεν αρκεί μόνο οι περισσότερες παρατηρήσεις να βρίσκονται εντός των ορίων προδιαγραφών (άνω όριο προδιαγραφών USL κάτω όριο προδιαγραφών LSL), αλλά και η μέση τιμή μ να βρίσκεται όσο το δυνατόν πιο κοντά στην τιμή στόχο T της διεργασίας.

3.3.1 Ορισμός και ιδιότητες του δείκτη ικανότητας C_{pm}

Ο ορισμός του δείκτη C_{pm} δίνεται από τον παρακάτω τύπο

$$C_{pm} = \frac{USL - LSL}{6\tau}$$

όπου

$$\tau^2 = E(X - T)^2 = E[(X - \mu) + (\mu - T)]^2 = E[(X - \mu)^2] + (\mu - T)^2 = \sigma^2 + (\mu - T)^2.$$

Εδώ αξίζει να σημειώσουμε ότι ο ορισμός του συγκεκριμένου δείκτη ικανότητας βασίζεται κατά ένα μεγάλο ποσοστό στον δείκτη ικανότητας C_p , αφού με μια προσεκτική ματιά η μόνη διαφορά των δύο δεικτών είναι στον παρονομαστή. Συγκεκριμένα, ο δείκτης C_p έχει στον παρονομαστή την τυπική απόκλιση των παρατηρήσεων της διεργασίας ενώ ο δείκτης C_{pm} περιέχει την τετραγωνική ρίζα της μέσης τετραγωνικής απόκλισης από την προκαθορισμένη τιμή στόχο της διεργασίας. Σε αυτό το σημείο παρατηρούμε ότι οι δύο αυτοί δείκτες C_{pm} και C_p είναι ακριβώς ίδιοι στην περίπτωση που η τιμή στόχος T της διεργασίας είναι ίση με την μέση τιμή μ της διεργασίας.

Ένας άλλος τύπος για τον δείκτη C_{pm} είναι ο ακόλουθος

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}}$$

Από τον παραπάνω τύπο είναι ξεκάθαρο ότι η τιμή του δείκτη C_{pm} είναι αντιστρόφως ανάλογη με την τυπική απόκλιση σ των παρατηρήσεων μιας διεργασίας, δηλαδή όσο αυξάνεται η τυπική απόκλιση σ , πράγμα που σημαίνει ότι ο δειγματικός μέσος της διεργασίας μ απομακρύνεται από την κεντρική τιμή στόχο T , μειώνεται ο δείκτης C_{pm} και αντίστροφα, όταν μειώνεται η τυπική απόκλιση σ , που σημαίνει ότι η μέση τιμή της διεργασίας πλησιάζει την κεντρική τιμή στόχο της

διεργασίας, αυξάνεται ο δείκτης C_{pm} , πράγμα που ήταν αναμενόμενο αφού η τυπική απόκλιση σ βρίσκεται στον παρονομαστή.

Στην περίπτωση όπου $M = T$ οι τρεις δείκτες μπορούν να συνδεθούν με μία και μόνο σχέση την εξής

$$C_{pm} = \frac{C_p}{\sqrt{1+9(C_p - C_{pk})^2}}$$

Επιπλέον, σύμφωνα με τους Kotz and Lovelace (1998) ισχύει και ο παρακάτω τύπος ο οποίος μας δείχνει τη σχέση των C_{pm} , C_p και C_{pk}

$$C_{pm} = \frac{C_p}{\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}} = \frac{C_{pk}}{\left(1-\frac{|\mu-M|}{d}\right)\sqrt{1+\left(\frac{\mu-T}{\sigma}\right)^2}}$$

Στην ειδική περίπτωση που έχουμε $\mu = T = M$ τότε $C_p = C_{pm} = C_{pk}$. Ειδικά αν, $|\mu-T| \neq 0$ τότε $C_{pm} \leq C_p$, ενώ αν $|\mu-T| \rightarrow \infty$, τότε $C_{pm} \rightarrow 0$. Επίσης, επειδή ισχύει ότι

$$\sqrt{\sigma^2 + (\mu-T)^2} \geq |\mu-T|$$

έχουμε ότι

$$C_{pm} \leq \frac{USL - LSL}{6|\mu-T|}.$$

Από την παραπάνω σχέση μπορούμε να καταλήξουμε στα εξής συμπεράσματα: Οι δείκτες C_{pk} και C_p μπορούν να πάρουν οποιαδήποτε τιμή, αν αλλάξει η τιμή της τυπικής απόκλισης σ , δηλαδή αν αυξηθεί ή αν μειωθεί ακόμα κι αν ο δειγματικός μέσος είναι πολύ μακριά από την τιμή στόχο της διεργασίας. Ο δείκτης C_{pm} όμως δεν μπορεί να πάρει οποιαδήποτε τιμή, διότι στην περίπτωση που η ποσότητα $|\mu-T|$ είναι σταθερή τότε ο δείκτης C_{pm} είναι άνω φραγμένος, κάτι που σημαίνει ότι αν ο C_{pm} είναι πολύ μεγάλος δε σημαίνει ότι και ο δειγματικός μέσος μ απέχει πολύ από την κεντρική τιμή T .

Αν υποθέσουμε ότι $M = T$ τότε το ανώτερο δυνατό όριο για οποιαδήποτε τιμή του δείκτη C_{pm} δίνεται από τον ακόλουθο τύπο

$$|\mu - T| < \frac{USL - LSL}{6C_{pm}}$$

Εδώ αξίζει να σημειώσουμε ότι αν $C_{pm} \geq 1$ τότε η ελάχιστη απόσταση που μπορεί να έχει ο δειγματικός μέσος μ της διεργασίας, είναι μικρότερη από το 1/6 του εύρους του διαστήματος των ορίων προδιαγραφών. Αυτό σημαίνει ότι ο μέσος μ βρίσκεται εντός της περιοχής των προδιαγραφών και πιο συγκεκριμένα στο μεσαίο τρίτο της περιοχής αυτής.

Ο πίνακας που παρουσιάζεται παρακάτω σύμφωνα με τον Αντζουλάκο (2006) μας δείχνει τις πιο βασικές ιδιότητες του δείκτη C_{pm}

Πίνακας 3.2

$\mu = T \Leftrightarrow C_{pm} = C_p$
$\mu = T = M \Leftrightarrow C_{pm} = C_p = C_{pk}$
$\mu \neq T \Leftrightarrow C_{pm} < C_p$
$ \mu - T \rightarrow \infty \Rightarrow C_{pm} \rightarrow 0$
$\sigma \downarrow \Rightarrow C_{pm} \uparrow$
$ \mu - T \downarrow \Rightarrow C_{pm} \uparrow$

3.3.2 Εκτίμηση του δείκτη ικανότητας C_{pm}

Για να μπορέσουμε να εκτιμήσουμε τον δείκτη C_{pm} θα συλλέξουμε ένα τυχαίο δείγμα $X_1 \dots X_n$ μεγέθους n του οποίου οι παρατηρήσεις του χαρακτηριστικού

X της διεργασίας που μελετάμε ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$ με μέση τιμή μ και τυπική απόκλιση σ . Επιπλέον τα όρια προδιαγραφών είναι τα γνωστά, κάτω όριο προδιαγραφών LSL και άνω όριο προδιαγραφών USL στη φάση σχεδιασμού της διεργασίας. Στηριζόμενοι στην εκτιμήτρια που πρότειναν οι Chan et al.(1988), θα εκτιμήσουμε την ποσότητα τ , με τη βοήθεια της στατιστικής συνάρτησης

$$\tilde{\tau} = \sqrt{\frac{\sum_{i=1}^n (X_i - T)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} + \frac{n(\bar{X} - T)^2}{n-1}} = \sqrt{S^2 + \frac{n}{n-1}(\bar{X} - T)^2}$$

Μπορούμε να εκτιμήσουμε τον δείκτη C_{pm} σύμφωνα με τους Chan et al.(1988) με τον παρακάτω ορισμό

$$C_{pm} = \frac{USL - LSL}{6\tilde{\tau}}.$$

Εκτός από τους Chan et al.(1988), ο Boyles(1991) πρότεινε για την εκτίμηση του δείκτη C_{pm} τον ακόλουθο τύπο

$$C_{pm} = \frac{USL - LSL}{6\hat{\tau}}.$$

Η εκτιμήτρια της ποσότητας του τ , η $\hat{\tau}$ σύμφωνα με τους Kotz and Johnson(1993) δίνεται από τον τύπο

$$\hat{\tau} = \sqrt{\frac{\sum_{i=1}^n (X_i - T)^2}{n}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} + (\bar{X} - T)^2} = \sqrt{\frac{n-1}{n}S^2 + (\bar{X} - T)^2} = \sqrt{\frac{n-1}{n}}\tilde{\tau}.$$

Σε αυτό το σημείο πρέπει να τονίσουμε ότι οι δύο εκτιμήτριες που αναφέρθηκαν για τον δείκτη C_{pm} συνδέονται με τον ακόλουθο τύπο από τον οποίο μπορούμε να εξάγουμε το συμπέρασμα ότι έχουνε παρόμοιες στατιστικές ιδιότητες.

$$C_{pm} = \sqrt{\frac{n}{n-1}} C_{pm}$$

Εδώ αξίζει να σημειώσουμε ότι αν $n \rightarrow \infty$ τότε οι εκτιμήτριες συμπίπτουν ασυμπτωτικά.

Όπως τονίσαμε παραπάνω, αυτές οι δύο εκτιμήτριες, παρά το γεγονός ότι έχουν παρόμοιες στατιστικές ιδιότητες έχουν μια βασική διαφορά. Σύμφωνα με τους Subbaiah and Taam (1993), οι οποίοι έκαναν προσομοιώσεις, συμπέραναν ότι όταν χρειαζόμαστε να υπολογίσουμε ένα διάστημα εμπιστοσύνης είναι πιο σωστό να χρησιμοποιούμε την εκτιμήτρια C_{pm} παρά την C_{pm} , την οποία την προτιμάμε σε περιπτώσεις όπου μας έχει ζητηθεί απλή σημειακή εκτίμηση.

3.3.3 Διαστήματα εμπιστοσύνης του δείκτη ικανότητας C_{pm}

Ο Boyles (1991) ήταν ο πρώτος που έδωσε διαστήματα εμπιστοσύνης για το δείκτη C_{pm} που μελετάμε, αποδεικνύοντας ότι η ποσότητα

$$n \frac{\tau_\beta}{\sigma^2} \sim \chi_n^2(\lambda)$$

όπου

$$\lambda = \frac{n(\mu - T)^2}{\sigma^2} = n\zeta^2$$

Έπειτα, απέδειξε ότι

$$\frac{\tau_\beta}{\tau} \overset{\alpha}{\sim} \frac{\chi_\nu^2}{\nu}$$

όπου

$$\nu = \frac{n(1+\zeta^2)^2}{1+2\zeta^2}$$

Σε αυτό το σημείο πρέπει να σημειώσουμε ότι για πολύ μεγάλες τιμές του ν έχουμε ότι

$$\frac{\tau_\beta}{\tau} \stackrel{\alpha}{\sim} N\left(1, \frac{1}{2\nu}\right)$$

Σύμφωνα με τα παραπάνω αποτελέσματα που έδωσε ο Boyles (1991) μπορούμε εύκολα να κατασκευάσουμε προσεγγιστικά διαστήματα εμπιστοσύνης για τον δείκτη επίδοσης C_{pm} , αν παρατηρήσουμε πιο προσεκτικά την ακόλουθη σχέση

$$\frac{C_{pm}}{C_{pm}(\beta)} = \frac{\tau_\beta}{\tau}$$

Επίσης η εκτίμηση μέγιστης πιθανοφάνειας του ν δίνεται από τον παρακάτω τύπο

$$\nu = \frac{n(1+\zeta^2)^2}{1+2\zeta^2}$$

όπου

$$\zeta = \frac{\bar{X} - T}{\sqrt{(n-1)S^2/n}}$$

Επομένως ένα προσεγγιστικό κάτω $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη ικανότητας C_{pm} είναι το εξής

$$\left(C_{pm}(\beta) \sqrt{\frac{\chi_{\nu,1-\alpha}^2}{\hat{\nu}}}, \infty \right)$$

Ο παραπάνω τύπος για πολύ μεγάλα $\hat{\nu}$ διαμορφώνεται ως εξής

$$\left(C_{pm}(\beta) \left(1 - z_\alpha \sqrt{\frac{1}{2\hat{\nu}}} \right), \infty \right)$$

Εδώ πρέπει να αναφέρουμε ότι εκτός από τον Boyles, διαστήματα εμπιστοσύνης για την εκτίμηση του δείκτη $C_{pm}(\beta)$ έδωσαν οι Kotz and Johnson (1993) αλλά και οι Pearn et al (1992).

3.3.4 Έλεγχοι υποθέσεων για τον δείκτη ικανότητας C_{pm}

Έστω ο έλεγχος της υπόθεσης σχετικά με την ικανότητα μια διεργασίας είναι ο εξής

$$H_0 : \text{Μη ικανή διεργασία} - H_1 : \text{Ικανή διεργασία}$$

Για τον έλεγχο της υπόθεσης που αναφέραμε, με βάση το δείκτη C_{pm} ο Kane (1986) θεώρησε ότι $\mu = T$. Σε αυτήν ακριβώς την παραδοχή βασίστηκαν οι Chan et al (1988) και διαμόρφωσαν τον έλεγχο όπως ορίζεται παρακάτω.

$$H_0 : C_{pm} < C_{pm}(\text{low}) - H_1 : C_{pm} > C_{pm}(\text{high})$$

Εδώ πρέπει να υπενθυμίσουμε ότι αν βρισκόμαστε στην περίπτωση όπου $\mu = T$, τότε ισχύουν τα παρακάτω

$$W = \frac{(n-1)\tau^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi_n^2$$

$$W = \frac{C_{pm}^2(n-1)}{C_{pm}^2}$$

Συμπεραίνουμε από τα παραπάνω ότι $P(C_{pm} \geq c) = P(\chi_n^2 \leq C_{pm}^2(n-1)/c^2)$. Έστω ότι η συνάρτηση επαληθεύει τα σημεία $(C_p(\text{low}), \alpha)$ και $(C_p(\text{high}), 1-\beta)$ κι επίσης η περιοχή της μορφής $C_{pm} \geq c$ θεωρείται ως κρίσιμη περιοχή, δηλαδή περιοχή της απόρριψης της μηδενικής υπόθεσης. Σύμφωνα με τα προηγούμενα δεδομένα καταλήγουμε στην ακόλουθη σχέση

$$\alpha = \pi(C_p(\text{low})) = P[\chi_n^2 \leq (n-1)(C_{pm}(\text{low})/c)^2]$$

και

$$1 - \beta = \pi(C_p(\text{high})) = P[\chi_n^2 \leq (n-1)(C_{pm}(\text{high})/c)^2]$$

Άμεση συνέπεια των δύο προηγούμενων τύπων είναι οι ακόλουθες σχέσεις

$$\frac{C_{pm}(\text{high})}{C_{pm}(\text{low})} = \sqrt{\frac{\chi_{n,\beta}^2}{\chi_{n,1-\alpha}^2}}$$

και

$$c = C_{pm}(\text{low}) \sqrt{\frac{n-1}{\chi_{n,1-\alpha}^2}}$$

3.3.5 Παρόμοιοι δείκτες του δείκτη ικανότητας C_{pm}

Οι Chan et al (1988) πρότειναν έναν δείκτη που είναι παρόμοιος με τον C_{pm} , ο οποίος καλύπτει ουσιαστικά κάποιες αδυναμίες του τελευταίου, οι οποίες θα αναλυθούν στην αμέσως επόμενη ενότητα. Ο δείκτης που πρότειναν ονομάζεται C_{pm}^* και ορίζεται ως εξής

$$C_{pm}^* = \frac{\min(USL - T, T - LSL)}{3\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d - |T - M|}{3\sqrt{\sigma^2 + (\mu - T)^2}}.$$

Θα μπορούσαμε να ισχυριστούμε ότι ο συγκεκριμένος δείκτης λειτουργεί ως συμπληρωματικός του C_{pm} , αφού μας είναι χρήσιμος στην περίπτωση που $\mu = T \pm \delta, \delta > 0$, διότι δεν αλλάζει τιμή επομένως μας δίνει ελλιπείς πληροφορίες.

Η εκτιμήτρια του δείκτη C_{pm}^* είναι η παρακάτω

$$C_{pm}^* = \frac{(d - |T - M|)}{3} \left[\frac{1}{n} \sum_{i=1}^n (X_i - T)^2 \right].$$

Ένας ακόμα δείκτης που έχει παρόμοιες ιδιότητες με τον C_{pm} είναι ο C_{pm}^+ Boyles(1995), τον οποίο τον χρησιμοποιούμε στην περίπτωση όπου $USL - T \neq T - LSL$ και ορίζεται σύμφωνα με τον Boyles(1995), ως εξής

$$C_{pm}^+ = \frac{1}{3} \left[(T - LSL)^{-2} E_{X < T} [(X - T)^2] + (USL - T)^{-2} E_{X > T} [(X - T)^2] \right]^{-1/2}.$$

3.3.6 Αδυναμίες του δείκτη ικανότητας C_{pm}

Όπως αναφέραμε παραπάνω, ο δείκτης C_{pm} έχει πολλά πλεονεκτήματα, αφού με τη βοήθειά του μπορούμε, όχι μόνο να δούμε κατά πόσο μια διεργασία μπορεί να παράγει μη ελαττωματικά προϊόντα, δηλαδή εντός ορίων προδιαγραφών, αλλά και να ελέγξουμε την ποιότητα των παραχθέντων προϊόντων, αφού υπολογίζει το μέγεθος της τυπικής απόκλισης. Αυτό σημαίνει ότι γνωρίζουμε κατά πόσο απέχουν οι παρατηρήσεις του χαρακτηριστικού μιας διεργασίας από την κεντρική τιμή στόχο. Παρόλα αυτά όμως ο δείκτης C_{pm} έχει και αδυναμίες. Η πιο βασική αδυναμία του είναι το γεγονός ότι, ενώ λαμβάνει υπόψη πόσο απέχει ο δειγματικός μέσος από την κεντρική τιμή στόχο T , δε λαμβάνει υπόψη τη θέση του μ σε σχέση με τα όρια προδιαγραφών πράγμα που έχει αντίκτυπο στην απόδοση μιας διεργασίας. Αυτό γίνεται εύκολα κατανοητό, αν σκεφτούμε ότι με αυτά τα δεδομένα τα όρια προδιαγραφών δεν είναι συμμετρικά, επομένως ο δείκτης C_{pm} αποδίδει την ίδια ικανότητα όταν έχει ίσες τυπικές αποκλίσεις και οι δειγματικοί μέσοι ισαπέχουν από την κεντρική τιμή στόχο, κάτι που μπορεί να διαφέρει κατά πολύ από την πραγματικότητα.

3.4 Ο δείκτης ικανότητας C_{pmk}

3.4.1 Ορισμός και χαρακτηριστικά του δείκτη C_{pmk}

Ο τελευταίος δείκτης που θα εξεταστεί είναι ο δείκτης C_{pmk} , ο οποίος από το όνομά του φαίνεται ότι προέρχεται από τους δύο προηγούμενους δείκτες που περιγράψαμε αναλυτικά τον C_{pm} και τον C_{pk} . Είναι κάτι που φαίνεται και από τον ορισμό του δείκτη που θα αναλύσουμε παρακάτω.

Ο λόγος που χρησιμοποιείται ο δείκτης C_{pmk} είναι για καλύψει τις αδυναμίες των δύο προηγούμενων δεικτών C_{pm} και C_{pk} αλλά και του δείκτη C_p , αφού από τον ορισμό του που θα δούμε παρακάτω είναι ο μόνος δείκτης, ο οποίος λαμβάνει υπόψη του την τυπική απόκλιση σ , πόσο απέχει ο δειγματικός μέσος μ από την κεντρική τιμή στόχο T αλλά και από την κεντρική τιμή του διαστήματος των ορίων προδιαγραφών χωρίς να είναι απαραίτητο το άνω όριο προδιαγραφών USL να είναι συμμετρικό με το κάτω όριο προδιαγραφών LSL .

Οι πρώτοι που πρότειναν τον δείκτη C_{pmk} ήταν οι Choi and Owen (1990) και αργότερα οι Pearn et al. (1992). Ο ορισμός του C_{pmk} δίνεται από τον παρακάτω τύπο

$$C_{pmk} = \min \left(\frac{\mu - LSL, USL - \mu}{3\tau} \right) = \min \left(\frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right)$$

Όπως αναφέραμε παραπάνω αφού ο δείκτης C_{pmk} είναι συνάρτηση των δύο δεικτών C_{pk} και C_{pm} μπορεί να γραφεί και ως εξής

$$C_{pmk} = \frac{C_{pk}}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}} = \left(1 - \frac{|\mu - M|}{d}\right) C_{pm} = \frac{d - |\mu - M|}{3\sqrt{\sigma^2 + (\mu - T)^2}}$$

Ειδικά, αν $\mu = T$, τότε σημαίνει ότι $|\mu - T| = 0$, άρα και $C_{pmk} = C_{pk}$. Στην αντίθετη περίπτωση όπου $\mu \neq T$, τότε $|\mu - T| \neq 0$ άρα $C_{pmk} < C_{pk}$. Επιπλέον, αφού ισχύει ότι $C_{pk} \leq C_p$, σημαίνει ότι $C_{pmk} \leq C_{pm}$ και στην περίπτωση όπου $\mu \neq M$, τότε $C_{pmk} < C_{pm}$. Αν όμως $\mu = T = M$ τότε $C_{pm} = C_p$.

3.4.2 Βασικές ιδιότητες του δείκτη C_{pmk}

Οι πιο βασικές ιδιότητες του δείκτη C_{pmk} δίνονται επιγραμματικά στον παρακάτω πίνακα από τον Αντζουλάκο (2006) και είναι ουσιαστικά μια συνοπτική παρουσίαση των παραπάνω σχέσεων.

$\mu = M \Rightarrow C_{pmk} = C_{pm} (C_{pk} = C_p)$ $\mu = M \Rightarrow C_{pmk} = C_{pk} (C_{pm} = C_p)$ $\mu = T = M \Rightarrow C_{pmk} = C_{pk} = C_{pm} = C_p$ $C_{pmk} \leq C_{pm}$

Σε αυτό το σημείο αξίζει να σημειώσουμε ότι αν $\sigma \rightarrow 0$, $\mu \rightarrow T$ τότε $C_{pmk} \rightarrow \infty$ άρα ο τύπος του διαμορφώνεται ως εξής

$$C_{pmk} = \frac{USL - LSL}{6|\mu - T|} - \frac{1}{3}.$$

Επομένως για κάθε τιμή του δείκτη C_{pmk} υπάρχει το εξής όριο

$$|\mu - T| < \frac{USL - LSL}{2(1 + 3C_{pmk})}.$$

Τέλος αν

$$C_{pmk} = 1 \Rightarrow |\mu - T| < \frac{USL - LSL}{8}.$$

Αυτό σημαίνει ότι ο δειγματικός μέσος βρίσκεται στο μεσαίο τέταρτο της περιοχής του άνω ορίου προδιαγραφών και κάτω ορίου προδιαγραφών.

3.4.3 Εκτίμηση του δείκτη C_{pmk}

Οι Pearn, Kotz and Johnson (1992), προκειμένου να εκτιμήσουν το δείκτη C_{pmk} , πρότειναν την εκτιμήτρια εκείνη, η οποία προκύπτει από τον ορισμό του C_{pmk} . Η εκτιμήτρια του δείκτη C_{pmk} δίνεται από τον ακόλουθο τύπο

$$C_{pmk} = \min \frac{(\bar{X} - LSL, USL - \bar{X})}{3\sqrt{\frac{n}{n-1}S^2 + (\bar{X} - T)^2}} = \min \left(\frac{\bar{X} - LSL}{3\sqrt{\frac{n}{n-1}S^2 + (\bar{X} - T)^2}}, \frac{USL - \bar{X}}{3\sqrt{\frac{n}{n-1}S^2 + (\bar{X} - T)^2}} \right)$$

$$= \frac{d - |\bar{X} - M|}{3\sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} + (\bar{X} - T)^2}}.$$

Σύμφωνα με τον τύπο της εκτιμήτριας μπορούμε να συμπεράνουμε ότι η C_{pmk} δεν είναι αμερόληπτη εκτιμήτρια για τον δείκτη C_{pmk}

3.4.4 Διαστήματα εμπιστοσύνης και έλεγχοι υποθέσεων του δείκτη

C_{pmk}

Όπως αναφέραμε στην προηγούμενη ενότητα αυτού του κεφαλαίου, ο Boyles (1991) πρότεινε διάστημα εμπιστοσύνης για τον δείκτη ικανότητας C_{pm} στην περίπτωση που ισχύει $\mu = M \Rightarrow C_{pmk} = C_{pm}$.

Οι Chen and Hsu (1995), με τη σειρά τους έδωσαν ασυμπτωτικά κάτω $100(1-\alpha)\%$ διάστημα εμπιστοσύνης για τον δείκτη επίδοσης C_{pmk} για δύο περιπτώσεις, αν $\mu = M$ κι αν $\mu \neq M$.

Στην πρώτη περίπτωση όπου $\mu = M$ πρότειναν το ακόλουθο διάστημα εμπιστοσύνης για τον δείκτη C_{pmk}

$$\left(C_{pmk} \sqrt{\frac{\chi_{\nu, 1-\alpha}^2}{\nu}}, \infty \right)$$

όπου

$$\nu = \frac{n(1+\zeta^2)^2}{1+2\zeta^2}$$

και

$$\zeta = \frac{\bar{X} - T}{\sqrt{(n-1)S^2/n}}$$

ενώ για την περίπτωση όπου $\mu \neq M$ έχουμε

$$C_{pmk} \pm z_{\alpha/2} \frac{\sigma_{pmk}}{\sqrt{n}}$$

όπου

$$\sigma_{pmk}^2 = \left(\frac{1}{9(1+\lambda^2)} + \frac{2\lambda}{3(1+\lambda^2)^{3/2}} \right) C_{pmk} + \frac{144\lambda^2 + (USL - LSL) \left(\frac{m_4}{s_n^4} - 1 \right)}{144(1+\lambda^2)^2} C_{pmk}^2$$

και

$$\lambda = \frac{\bar{X} - T}{S}$$

και

$$m_4 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4$$

Σε αυτό το σημείο πρέπει να αναφέρουμε ότι οι μόνοι οι οποίοι ανέπτυξαν μια μεθοδολογία για έλεγχο υποθέσεων σχετικά με την τιμή του δείκτη C_{pmk} ήταν οι Pearn et al. (2001), που όμως δεν έχουν αναλυθεί επαρκώς.

3.4.5 Αδυναμίες του δείκτη C_{pmk}

Ο δείκτης C_{pmk} όπως αναφέραμε παραπάνω προτάθηκε για να καλύψει τις αδυναμίες όλων των προηγούμενων δεικτών που αναφέραμε σε αυτό το κεφάλαιο κι αυτό διότι είναι ο μόνος δείκτης από όσους ο οποίος συνδυάζει όλα τα απαραίτητα στατιστικά δεδομένα μ, σ, T, USL, LSL . Αυτό όμως έχει ως αποτέλεσμα να έχει περίπλοκο τύπο και με αυτόν τον τρόπο είναι δύσκολη η εύρεση εκτιμητριών πράγμα που μας κάνει δύσκολο το έργο εύρεσης διαστημάτων εμπιστοσύνης, όπως διαπιστώσαμε παραπάνω.

ΚΕΦΑΛΑΙΟ 4

ΕΦΑΡΜΟΓΕΣ ΜΕ ΧΡΗΣΗ ΤΩΝ ΣΤΑΤΙΣΤΙΚΩΝ ΠΑΚΕΤΩΝ STATGRAPHICS ΚΑΙ MINITAB

4.1 Περιγραφή των δεδομένων

Ας υποθέσουμε ότι εργαζόμαστε σε ένα εργοστάσιο κατασκευής αυτοκινήτων, σε ένα τμήμα που συναρμολογεί τις μηχανές. Ένα από τα μέρη, η εκκεντροφόρος, πρέπει να είναι $600 \text{ mm} \pm 2 \text{ mm}$ σε μήκος, για να πληροί τις προδιαγραφές. Παρατηρήθηκε ένα χρόνιο πρόβλημα με τα μήκη κάποιων εκκεντροφόρων να είναι εκτός προδιαγραφών – ένα πρόβλημα που προκάλεσε κακή συναρμολόγηση των μηχανών, οδηγώντας σε υψηλά έξοδα επισκευής και εν τέλει σε δυσμενείς συνέπειες στην παραγωγή (δεδομένα camshaft από το minitab).

Μετά την εξέταση των ιστορικών αρχείων του εργοστασίου, διαπιστώσαμε ότι υπήρχαν δύο προμηθευτές για τις εκκεντροφόρους (supp1 και supp2). Ο προϊστάμενος, ζήτησε να εξετασθεί η καταλληλότητα των εξαρτημάτων κάθε προμηθευτή και να βρεθεί μία λύση για την βελτίωση της παραγωγής. Οπότε, για ένα μήνα, συλλέξαμε συνολικά 100 παρατηρήσεις (20 δείγματα από 5 εκκεντροφόρους το καθένα), από όλους τους εκκεντροφόρους που χρησιμοποιούνται στο εργοστάσιο,

καθώς επίσης συλλέξαμε και 100 παρατηρήσεις από κάθε προμηθευτή, αναφορικά με τα μήκη των εκκεντροφόρων. Τα δεδομένα περιγράφονται στον ακόλουθο πίνακα

Πίνακας 4.1: Περιγραφή δεδομένων

Column	Name	Count	Description
C1	Length	100	Μήκη εκκεντροφόρων σε mm
C2	Supp 1	100	Μήκη εκκεντροφόρων του προμηθευτή 1
C3	Supp 2	100	Μήκη εκκεντροφόρων του προμηθευτή 2

Αναλυτικά, οι παρατηρήσεις είναι οι εξής:

Πίνακας 4.2

Length	Supp1	Supp2
601,4	598,0	601,6
601,6	599,8	600,4
598,0	600,0	598,4
601,4	599,8	600,0
599,4	600,0	596,8
600,0	600,0	602,8
600,2	598,8	600,8
601,2	598,2	603,6
598,4	599,4	604,2
599,0	599,6	602,4
601,2	599,4	598,4
601,0	599,4	599,6
600,8	600,0	603,4
597,6	598,8	600,6
601,6	599,2	598,4
599,4	599,4	598,2
601,2	599,6	602,0
598,4	599,0	599,4

599,2	599,2	599,4
598,8	600,6	600,8
601,4	598,8	600,8
599,0	598,8	598,6
601,0	599,8	600,0
601,6	599,2	600,4
601,4	599,4	600,8
601,4	600,0	600,8
598,8	600,2	597,2
601,4	600,2	600,4
598,4	599,6	599,8
601,6	599,0	596,4
598,8	599,0	600,4
601,2	599,8	598,2
599,6	600,8	598,6
601,2	598,8	599,6
598,2	598,2	599,0
598,8	600,0	598,2
597,8	599,2	599,4
598,2	599,8	599,4
598,2	601,2	600,2
598,2	600,4	599,0
601,2	600,2	599,4
600,0	599,6	598,0
598,8	599,6	597,6
599,4	599,6	598,0
597,2	600,2	597,6
600,8	599,2	601,2
600,6	599,0	599,0
599,6	599,6	600,4
599,4	600,4	600,6
598,0	600,0	599,0
600,8	599,0	602,2
597,8	599,6	599,8
599,2	599,4	599,8
599,2	599,2	601,0
600,6	597,8	601,6
598,0	600,4	601,6
598,0	599,6	600,2

598,8	600,0	601,8
601,0	600,8	601,2
600,8	600,4	597,6
598,8	599,4	599,8
599,4	599,0	602,8
601,0	598,4	600,0
598,8	599,0	599,6
599,6	599,6	602,2
599,0	598,8	603,8
600,4	599,2	603,6
598,4	599,6	601,8
602,2	598,6	602,0
601,0	599,8	603,6
601,4	599,6	600,8
601,0	599,2	600,2
601,2	599,6	600,4
601,4	600,2	600,2
601,8	599,8	602,2
601,6	599,6	598,0
601,0	600,0	598,4
600,2	599,6	600,8
599,0	599,2	602,8
601,2	598,6	597,6
601,2	599,6	601,6
601,2	601,2	603,4
601,2	599,6	597,0
601,0	600,2	599,8
601,0	600,0	597,8
601,4	600,0	602,4
601,4	599,4	602,2
598,8	599,8	600,6
598,8	599,2	596,2
598,8	599,6	602,4
598,2	599,4	601,4
601,8	600,0	599,2
601,0	600,0	601,6
601,4	599,2	600,4
601,4	599,4	598,0
599,0	599,6	601,2

601,4	599,8	604,2
601,8	599,0	600,2
601,6	599,6	600,0
601,2	599,4	596,8

Με τη βοήθεια ενός \bar{X} γραφήματος (με τιμή στόχο τη μέση τιμή των δεδομένων, βάσει των οδηγιών του προϊσταμένου), παρατηρήσαμε ότι οι εκκεντροφόροι που προέρχονταν από τον δεύτερο προμηθευτή (supp2) ήταν εκτός ελέγχου, συνεπώς αποφασίσαμε να σταματήσουμε να δεχόμαστε προμήθεια από εκεί. Έπειτα από τη λήξη της συνεργασίας με τον δεύτερο προμηθευτή, ο αριθμός των ελαττωματικών εκκεντροφόρων μειώθηκε σημαντικά, χωρίς όμως να εξαλειφθεί τελείως το πρόβλημα. Συνεπώς, αποφασίσαμε να ελέγξουμε την ικανότητα της διαδικασίας, ώστε να επιβεβαιώσουμε ή όχι αν ο προμηθευτής 1 είναι ικανός να ανταποκριθεί στις προδιαγραφές του εργοστασίου. Η ανάλυση των δεδομένων θα γίνει με τη βοήθεια δύο στατιστικών πακέτων, του Minitab και του Statgraphics. Στη συνέχεια θα γίνει μια περιγραφή για το πώς μπορούμε να χρησιμοποιήσουμε αυτά τα στατιστικά πακέτα.

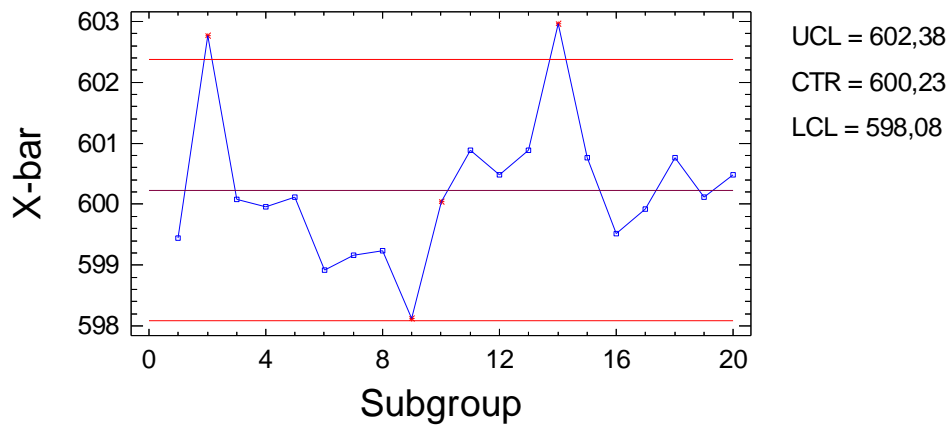
4.2 Χρήση του πακέτου Statgraphics και αποτελέσματα

Ακολουθούμε τα παρακάτω βήματα για την ανάλυση των δεδομένων:

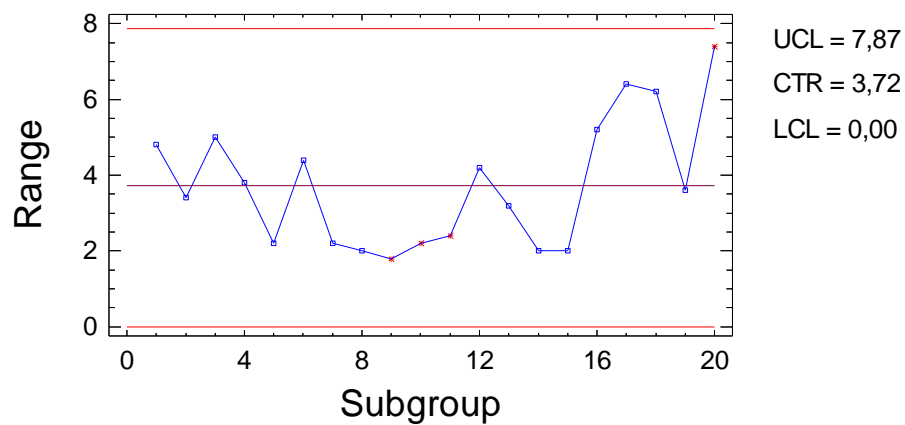
- a) \bar{X} -bar διάγραμμα και αποτελέσματα για τον supp2
 - Επιλέγουμε Special και συνεχίζουμε ως εξής:
 - Special → Quality Control → Variable Control Charts → \bar{X} -bar and R
 - Ανοίγει ένα παράθυρο με τα ακόλουθα πλαίσια, όπου εμείς συμπληρώνουμε τα εξής:
 - Observation: Supp2
 - Subgroup Numbers or Size: 5

- Πατάμε OK.

X-bar Chart for Supp2



Range Chart for Supp2



X-bar and Range - Initial Stud

Number of subgroups = 20
Subgroup size = 5,0
0 subgroups excluded

X-bar Chart

UCL: +3,0 sigma = 602,376
Centerline = 600,23
LCL: -3,0 sigma = 598,084

2 beyond limits

Range Chart

UCL: +3,0 sigma = 7,86542
Centerline = 3,72
LCL: -3,0 sigma = 0,0

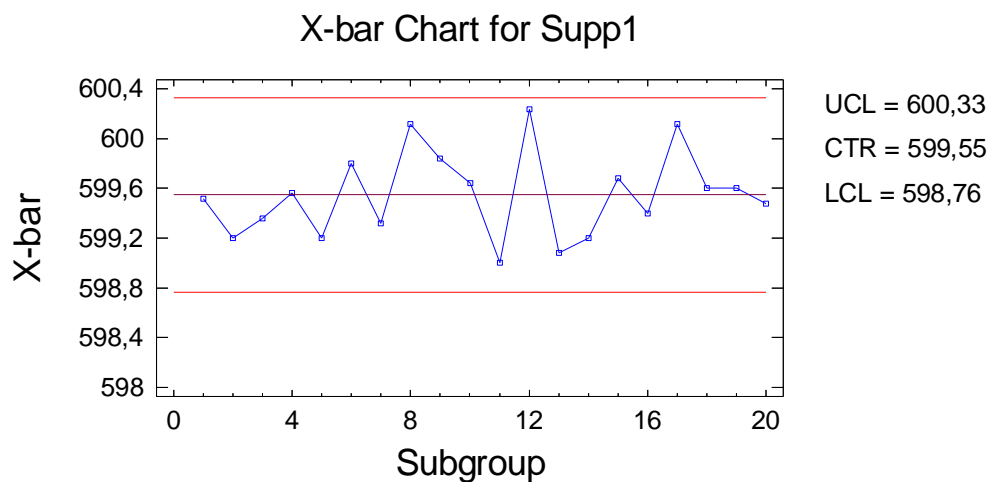
0 beyond limits

Estimates

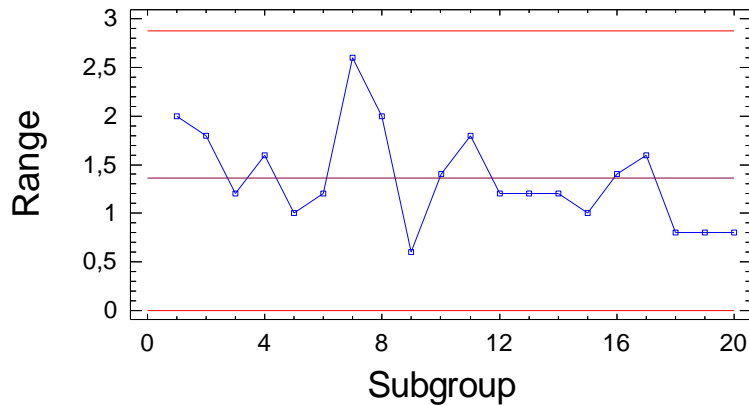
Process mean = 600,23
Process sigma = 1,59931
Mean range = 3,72

b) Συνεχίζουμε με τα αναλυτικά αποτελέσματα για τον supp1, για τον οποίο και θα εξετάσουμε τους δείκτες ικανότητας:

Αρχικά, ακολουθούμε ακριβώς την ίδια διαδικασία και για τον Supp1 για να προκύψει το παρακάτω X-bar διάγραμμα (μαζί και με το Range chart).



Range Chart for Supp1



UCL = 2,88

CTR = 1,36

LCL = 0,00

X - bar and Range - Initial Stud

Number of subgroups = 20

Subgroup size = 5,0

0 subgroups excluded

X - bar Chart

UCL: +3,0 sigma = 600,332

Centerline = 599,548

LCL: -3,0 sigma = 598,764

0 beyond limits

Range Chart

UCL: +3,0 sigma = 2,87553

Centerline = 1,36

LCL: -3,0 sigma = 0,0

0 beyond limits

Estimates

Process mean = 599,548

Process sigma = 0,584695

Mean range = 1,36

- Στο σημείο αυτό, συνεχίζουμε ως εξής:
- Επιλέγουμε Tabular Options → Capability Indices → OK
- Μετά πατάμε δεξί κλικ και στη συνέχεια επιλέγουμε Pane Options για να δώσουμε τα όρια. Στο συγκεκριμένο παράδειγμα συμπληρώνουμε ως εξής:
- LSL: 598
- USL: 602
- Norminal: 600
- Και τέλος επιλέγουμε OK.

Capability Indices fc

Specifications

USL = 602,0

Nominal = 600,0

LSL = 598,0

Cp = 1,1402

Cpk = 0,882512

Cpk (upper) = 1,39788

Cpk (lower) = 0,88251

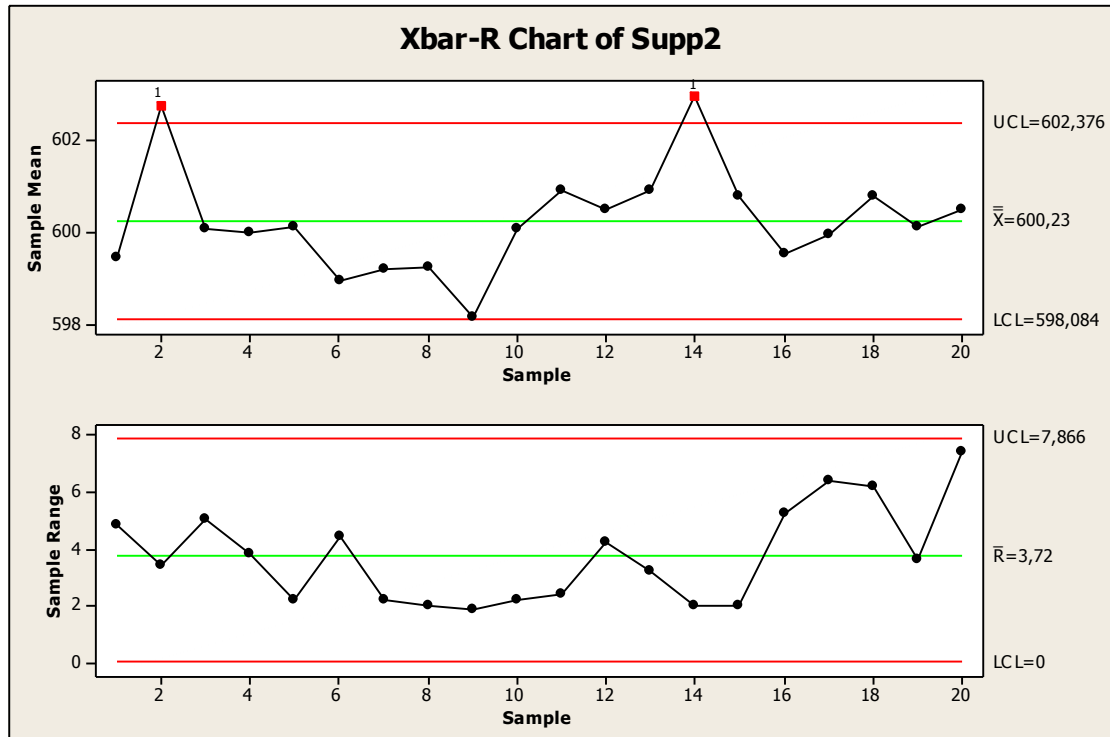
Cr = 0,877042

Cpm = 0,868

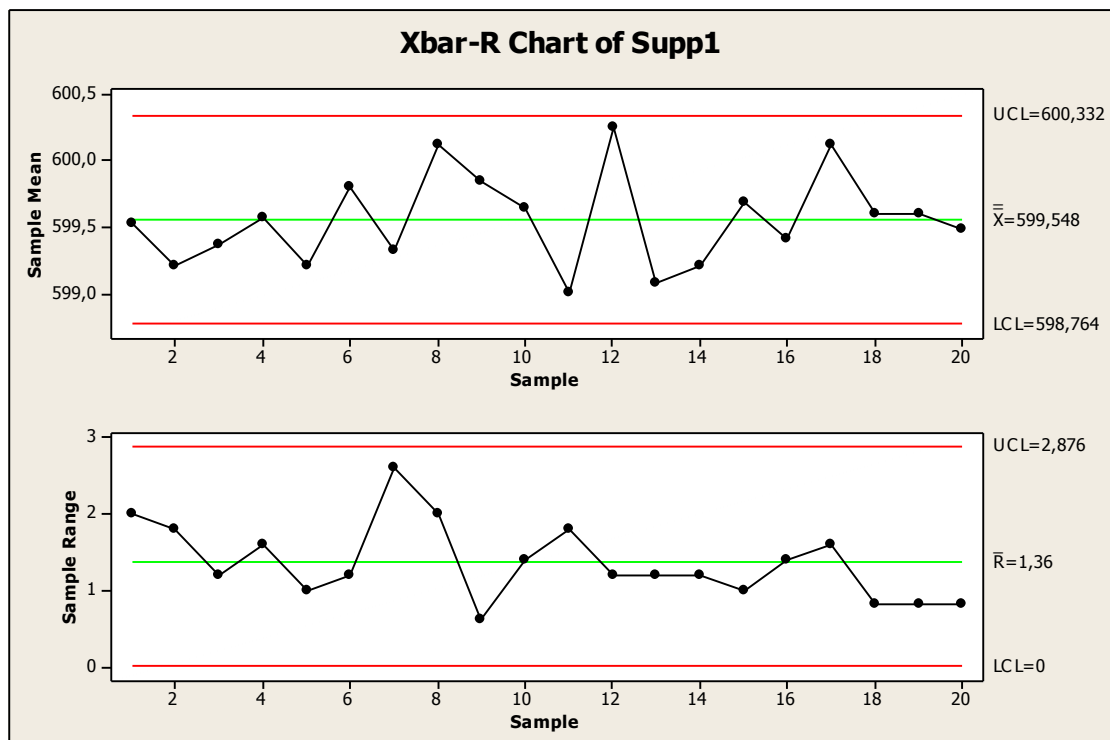
Based on 6.0 sigma li

4.3 Χρήση του πακέτου Minitab και αποτελέσματα

- Επιλέγουμε Stat → Control Charts → Variable Charts for Subgroups → X-bar-R.
- Ανοίγει ένα παράθυρο με πλαίσια, όπου εμείς επιλέγουμε τα εξής:
- All observations for a chart are in one column
- Ακριβώς στο κάτω πλαίσιο επιλέγουμε τη στήλη με τον Supp2
- Βάζουμε Subgroups sizes 5.
- Επιλέγουμε X-bar R Options και στο tab Estimate επιλέγουμε Rbar
- Πατάμε OK
- Πατάμε πάλι OK

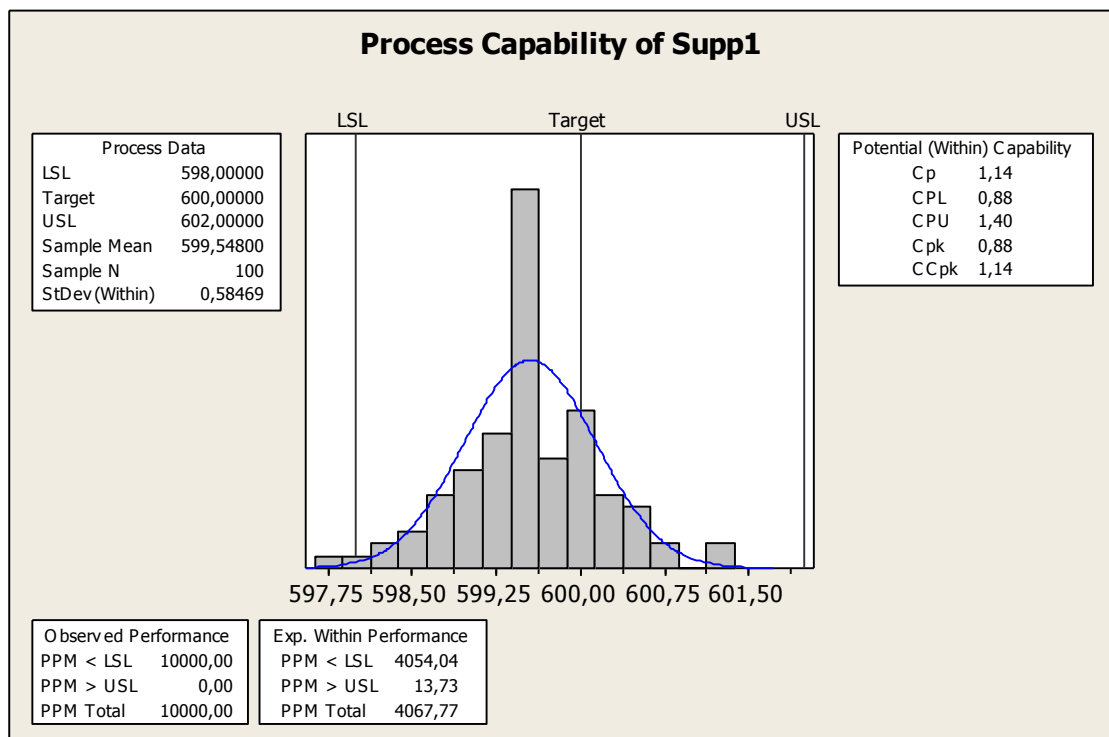


Τα ίδια ακριβώς βήματα ακολουθούμε και για τον Supp1 για να προκύψει το παρακάτω X-bar διάγραμμα.



Για να συνεχίσουμε τώρα με την ανάλυση της ικανότητας της διαδικασίας σε αυτόν τον προμηθευτή, εκτελούμε τα εξής βήματα:

- Stat → Quality Tools → Capability Analysis → Normal
- Σε αυτό το σημείο ανοίγει ένα παράθυρο εμφανίζοντας τα παρακάτω πλαίσια, όπου συμπληρώνουμε τα εξής:
- Single Column: Supp1
- Subgroup size: 5
- Lower Specification: 598
- Upper specification: 602
- Στη συνέχεια επιλέγουμε options και ανοίγει πάλι ένα παράθυρο με ένα πλαίσιο, όπου συμπληρώνουμε τα παρακάτω:
- Target: 600
- Πατάμε OK.
- Επιλέγουμε Estimate.
- Methods of estimating standard deviation επιλέγουμε Rbar.
- Πατάμε OK.
- Πατάμε OK.



Ανάλυση των αποτελεσμάτων:

Παρατηρούμε ότι αναφορικά με τον Supp2, η διαδικασία είναι ξεκάθαρα εκτός ελέγχου, σε αντίθεση με τον Supp1 που είναι εντός. Συνεπώς, καλώς διακόπηκε η συνεργασία με τον Supp2. Όσον αφορά τώρα τους δείκτες ικανότητας για τον Supp1, αρχικά ελέγξαμε την κανονικότητα των δεδομένων, μια απαίτηση που βάσει του ιστογράμματος, ικανοποιείται. Παρ' όλα αυτά, ο μέσος της διαδικασίας (599.548) διαφέρει λίγο από το στόχο των προδιαγραφών (600). Επιπλέον η αριστερή ουρά της κατανομής ξεφεύγει από το κάτω όριο προδιαγραφών (LSL). Αυτό σημαίνει ότι είναι πιθανόν να προκύψουν εκκεντροφόροι που δε πληρούν το κάτω όριο των 598 mm.

Σχετικά τώρα με τον C_p δείκτη, έχει τιμή >1 . Συνεπώς, αν λάβουμε υπόψη μόνο αυτόν τον δείκτη, η διαδικασία είναι ικανή. Αν επίσης υπολογίσουμε το λόγο

$$P = \left(\frac{1}{C_p} \right) 100 = \left(\frac{1}{1.14} \right) 100 = 87.72,$$

έχουμε ότι η διαδικασία χρησιμοποιεί το 87.72% της ζώνης των προδιαγραφών. Να μην ξεχνάμε όμως ότι ο δείκτης αυτός δε λαμβάνει υπόψη που βρίσκεται ο μέσος της διαδικασίας σε σχέση με τις προδιαγραφές. Απλά μετράει την έκταση των προδιαγραφών σχετικά με την έκταση 6σ στη διαδικασία. Χρησιμοποιώντας λοιπόν τον εναλλακτικό δείκτη $C_{pk} = 0.88$, παρατηρούμε ότι $C_p > C_{pk}$, άρα η διαδικασία είναι εκτός κέντρου. Επίσης, λόγω της μικρής του τιμής, όπως επίσης λόγω και της μικρής τιμής του δείκτη C_{pm} , οδηγούμαστε στο συμπέρασμα, ότι πρέπει μεν να συνεχίσουμε τη συνεργασία με τον πρώτο προμηθευτή, αλλά συγχρόνως απαιτείται η βελτίωση της διαδικασίας μειώνοντας τη μεταβλητότητα, κεντράροντας τη διαδικασία όσο γίνεται στο στόχο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Αντζουλάκος, Δ. Λ. (2006). *Στατιστικός Έλεγχος Ποιότητας, Σημειώσεις παραδόσεων του μαθήματος του Στατιστικού Ελέγχου Ποιότητας στα πλαίσια του Μ.Π.Σ. στην Εφαρμοσμένη Στατιστική, Πανεπιστήμιο Πειραιώς.*
- Bissell, A.F. (1990). How Reliable is Your Capability Index?. *Applied Statistics*, 39, 331-340.
- Boyles, R.A. (1991). The Taguchi Capability Index. *Journal of Quality Technology*, 23, 17-26.
- Boyles, R.A. (1995). Cpm for Asymmetric Tolerances. *Manuscript*, Precision Castparts Corp., Portland, OR.
- Chan, L.K., Cheng, S.W. and Spiring, F.A. (1988). A New Measure of Process Capability Cpm. *Journal of Quality Technology*, 20, 162-175.
- Chen, S.M. and Hsu, N.F. (1995). The Asymptotic Distribution of the Process Capability Index Cpmk. *Communications in Statistics - Theory and Methods*, 24, 1279-1291.
- Chou, Y. M. and Owen, D. B. (1989). On the Distributions of the Estimated Process Capability Indices. *Communications in Statistics-Theory and Methods*, 18, 4549-4560.
- Chou, Y. M., Owen, D. B. and Borrego, A. S. A. (1990). Lower Confidence Limits on Process Capability Indices. *Journal of Quality Technology*, 22, 223-229.
- Franklin, L. A. and Wasserman, G. S. (1992). A Note on the Conservative Nature of the Tables of Lower Confidence Limits for Cpk with a Suggested Correction. *Communications in Statistics - Simulation and Computation*, 21, 1165-1169.

- Hahn, G. J., and S. S. Shapiro (1967). *Statistical Models in Engineering*, Wiley, New York.
- Heavlin, W.D. (1988). Statistical Properties of Capability Indices. *Technical Report no 320*, Tech. Library, Advanced Micro Devices, Inc., Sunnyvale, CA.
- Hsiang, T. C. and Taguchi, G. (1985). A Tutorial on Quality Control and Assurance: The Taguchi Methods. *Proceedings of the ASA Annual Meeting*, Las Vegas, Nevada.
- Johnson, N.L. and Kotz, S. (1995). Comment on: Percentages of units within specification limits associated with given values of Cpk and Cpm (by Chan L.K. and Tang J.H.). *Communications in Statistics-Simulation and Computation*, 24.
- Kane, V.E. (1986). Process Capability Indices. *Journal of Quality Technology*, 18, 41-52.
- Κουκουβίνος, Χ. (2008). Στατιστικός Έλεγχος Ποιότητας. Ε.Μ.Π.
- Kotz, S. and Johnson, N.L. (1993). *Process Capability Indices*. Chapman and Hall, London.
- Kotz, S., and Lovelace, C. R. (1998). *Process Capability Indices in Theory and Practice*, Arnold, London.
- Kocherlakota, S. (1992). Process Capability Index: Recent Developments, *Sankhya: The Indian Journal of Statistics*, 54, 352-369.
- Kirmani, S.N.U.A., Kocherlakota, K. and Kocherlakota, S. (1991). Estimation of σ and the Process Capability Index Based on Subsamples. *Communications in Statistics - Theory and Methods*, 20, 275-291.

- Kushler, R.H. and Hurley, P. (1992) Confidence Bounds for Capability Indices. *Journal of Quality Technology*, 24, 188-195.
- Li, H., Owen, D.B. and Borrego, S.A. (1990). Lower Confidence Limits on Process Capability Indices Based on the Range. *Communications in Statistics –Simulation and Computation*, 19, 1-24.
- Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*, 6th edition, John Wiley & Sons, Hoboken, NJ.
- Nagata, Y. and Nagahata, H. (1992). Approximation Formulas for the Confidence Intervals of Process Capability Indices. *Reports of Statistical Approx. Research*, JUSE 39, 15-29.
- Nagata, Y. and Nagahata, H. (1994). Approximation formulas for the lower confidence limits of process capability indices, *Okayama Economic Review*, 25, 301- 314.
- Pearn, W.L., Kotz, S. and Johnson, N.L. (1992). Distributional and Inferential Properties of Process Capability Indices. *Journal of Quality Technology*, 24, 216-231.
- Pearn, W.L., Yang, S.L., Chen, K.S. and Lin, P.C. (2001). Testing Process Capability Using the Index Cpmk with an Application. *International Journal of Reliability, Quality and Safety Engineering*, 8, 15-34.
- Shapiro, S. S. (1980). *How to Test Normality and Other Distributional Assumptions*, Vol 3, The ASQC Basic References in Quality Control: Statistical Techniques, ASQC, Milwaukee, WI.
- Subbaiah, P. and Taam, W. (1993). Inference on the Capability Index, *Communications in Statistics-Theory and Methods*, 22(2), 537-560.