1
Στοιχεία δεδομένων: τρόποι ελέγχου και αξιολόγησης

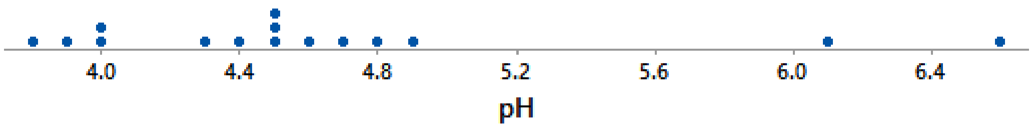
Ας υποθέσουμε ότι σε 15 κούπες γιαούρτης μετρούμε το pH και θέλουμε να συλλέξουμε πληροφορίες σχετικά με το pH του δείγματος, κάτι δηλαδή που να περιγράφει γενικά και ειδικά τη φύση και τη σχέση των αριθμών μεταξύ τους. Οι τιμές του pH που μετρήσαμε είναι:

4.5 4.7 4.4 3.8 4.5 4.0 4.0 6.6 4.3 4.5 4.9 4.8 3.9 4.6 6.1

Μία πρώτη επισκόπηση των αριθμών δεν φανερώνει τίποτα το ιδιαίτερο στις τιμές του pH, τίποτα ασυνήθιστο που να προκαλεί ενδεχομένως πρόβλημα σε μία στατιστική ανάλυση που θα θέλαμε να εφαρμόσουμε στο δείγμα, π.χ. την εκτίμηση του μέσου όρου και της τυπικής απόκλισης. Ο μέσος όρος που εξάγεται ως το άθροισμα των τιμών και διαίρεση με το πλήθος αυτών ισούται με 4.64 και η τυπική απόκλιση που αποτελεί το μέτρο της διασποράς των τιμών από το μέσο όρο ισούται με 0.77. Η ελάχιστη τιμή είναι 3.8 και η μέγιστη 6.6, το πλήθος των τιμών 15 και η διάμεσος του δείγματος που δείχνει την τιμή που αντιστοιχεί στο κέντρο των τιμών ίση με 4.5 (λεπτομερής ανάλυση περιγράφεται στο σχήμα 2.9, κεφάλαιο 2). Θα μπορούσαμε έτσι να καταλήξουμε σε κάποιο συμπέρασμα για το δείγμα μας, περιγράφοντας με τους παραπάνω όρους τι ακριβώς συμβαίνει με τις τιμές του pH, μία προσεκτικότερη ματιά όμως στις τιμές του δείγματος, αφού πρώτα τις κατατάξουμε από μικρότερες σε μεγαλύτερες για την ευκολία μας,

3.8 3.9 4.0 4.0 4.3 4.4 4.5 4.5 4.5 4.6 4.7 4.8 4.9 6.1 6.6

ή ευκρινέστερα τοποθετώντας τες σε ένα σημειόγραμμα,



αποκαλύπτει ότι η μεγάλη μάζα των τιμών εμπεριέχεται στο εύρος 3.8-4.9, περίπου μέσα σε μία μονάδα μεταβολής του pH, ενώ οι δύο τελευταίες τιμές φαίνεται να απέχουν αρκετά από τις άλλες και μάλιστα περίπου μιάμιση φορά πιο πάνω. Τίθεται επομένως το ερώτημα: είναι δυνατόν οι δύο αυτές τιμές να επηρεάζουν τις υπόλοιπες ως προς το μέσο όρο τους και την τυπική απόκλιση ή και οποιαδήποτε άλλη εκτίμηση κάνουμε; Αν αφαιρέσουμε τις δύο αυτές τιμές και εκτιμήσουμε ξανά τις προηγούμενες παραμέτρους προκύπτουν τα εξής: το δείγμα τώρα αποτελείται από 13 παρατηρήσεις αντί 15, το εύρος των τιμών μικραίνει θεαματικά (3.8-4.9), ο μέσος όρος μειώνεται από 4.64 σε 4.38, η τυπική απόκλιση μειώνεται από 0.77 σε 0.35, στο μισό και πλέον της προηγούμενης, ενώ η διάμεσος παραμένει αμετάβλητη σε 4.5 (σχήματα 2.9 και 2.10, κεφάλαιο 2). Η σημαντικότερη μεταβολή επιφέρεται στην τυπική απόκλιση, η οποία τώρα αποκαλύπτει θεαματική μείωση της απομάκρυνσης των τιμών από το μέσο όρο τους, με άλλα λόγια το δείγμα είναι τώρα πιο ομοιογενές, πιο αξιόπιστο στις πληροφορίες του. Μία άλλη επίσης σημαντική πληροφόρηση είναι ότι η διάμεσος του δείγματος παρέμεινε σταθερή, πράγμα που δηλώνει ότι όταν έχουμε κάποιες απομακρυσμένες τιμές από το υπόλοιπο σύνολο των τιμών, είναι προτιμότερο να εκτιμούμε τη διάμεσο ως πλέον αξιόπιστη παράμετρο του δείγματος αντί του μέσου όρου.

Μία αποτελεσματική λύση στο παραπάνω πρόβλημα είναι να μπορέσουμε να φέρουμε τις τιμές εγγύτερα μεταξύ τους, μειώνοντας με κάποιο τέχνασμα τις αποστάσεις τους. Αυτό επιτυγχάνεται λογαριθμώντας τους αριθμούς με το δεκαδικό λογάριθμο (ή το φυσικό), και πραγματικά οι αριθμοί τώρα δείχνουν σαφώς πλησιέστεροι μεταξύ τους:

0.580 0.591 0.602 0.602 0.633 0.643 0.653 0.653 0.653 0.663 0.672 0.681 0.690 0.785 0.819,

με τη μικρότερη και μεγαλύτερη τιμή να απέχουν μόλις 0.239. Επομένως, αν εξάγουμε το μέσο όρο των λογαριθμημένων αριθμών (0.6615) και ακολούθως αντιλογαριθμήσουμε για να επαναφέρουμε την παράμετρο στους κανονικούς αριθμούς προκύπτει μέσος όρος ίσος με 4.59 (ως αντιλογάριθμος του 0.6615), ο οποίος εμφανίζεται συντηρητικότερος από τον πρωταρχικό 4.64 και άρα περισσότερο αντιπροσωπευτικός ως προς το σύνολο των τιμών (Σχ. 2.11). Προκύπτει όμως και ένα άλλο ερώτημα: τελικά ποια από τις δύο περιπτώσεις πρόκειται να ακολουθήσουμε; είναι ικανές οι δύο απόμακρες τιμές να επηρεάζουν τόσο πολύ το μέσο όρο του δείγματος, ώστε να πρέπει να τις αφαιρέσουμε ή ο αντιλογαριθμημένος μέσος όρος (μαζί με άλλες σχετιζόμενες παραμέτρους, όπως τα όρια εμπιστοσύνης) επαρκούν για την αξιόπιστη πληροφόρηση του δείγματος; Το πρόβλημα αυτό λύνεται σχετικά εύκολα είτε υπολογίζοντας το θηκόγραμμα αλλά και κάποια ειδικά κριτήρια (κεφάλαιο 2), τα οποία μας ενημερώνουν για το μέγεθος και τη σοβαρότητα της απομάκρυνσης των ακραίων τιμών είτε υποβάλλοντας τα στοιχεία στη στατιστική ανάλυση της κανονικότητας του δείγματος. Για παράδειγμα, ο έλεγχος των Kolmogorov-Smirnov μας ενημερώνει αν το δείγμα επηρεάζεται σημαντικά από τις ακραίες τιμές, οπότε πράττουμε αναλόγως. Αν πραγματικά εξετάσουμε τα στοιχεία μας σύμφωνα με τα παραπάνω, θα βρούμε ότι οι ακραίες τιμές επηρεάζουν την ομαλή κατανομή των υπόλοιπων και θα είναι προτιμότερο να απομακρυνθούν ως πιθανές ύποπτες τιμές. Αν για κάποιο ιδιαίτερο λόγο αυτές είναι αναγκαίο να παραμείνουν στο δείγμα, τότε θα πρέπει υποχρεωτικά όλα τα στοιχεία του δείγματος να υποστούν κάποιο μετασχηματισμό που να ικανοποιεί τη συνθήκη ομαλοποίησης της διασποράς των αριθμών μεταξύ τους. Περισσότερες λεπτομέρειες για το συγκεκριμένο θέμα αναφέρονται στα κεφάλαια 2 και 5.

1.1 ΣΤΑΤΙΣΤΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΣΤΟΙΧΕΙΩΝ

Η υλοποίηση οποιασδήποτε έρευνας στο πεδίο ή στο εργαστήριο προϋποθέτει τη συλλογή στοιχείων (τιμών) στα οποία θα εφαρμοστεί κάποια συγκεκριμένη στατιστική επεξεργασία. Η στατιστική είναι η γλώσσα επικοινωνίας μεταξύ των αριθμών και των αντικειμένων της έρευνας και επομένως το μέσον έκφρασης των αποτελεσμάτων της. Αυτή διεξάγεται συνοπτικά σε τρία στάδια:

1. Συλλογή, κωδικοποίηση και αποθήκευση των στοιχείων με τη χρήση ενός στατιστικού προγράμματος. Ακολούθως, λεπτομερής εξέταση των στοιχείων με τη βοήθεια γραφικών παραστάσεων ή εφαρμογή της περιγραφικής στατιστικής. Με τον όρο αυτόν νοείται η πρώτη και επισταμένη στατιστική ανάλυση που αφορά πληροφορίες για το μέσο όρο κάθε στήλης αριθμών, τη διάμεσο και όλες τις σχετιζόμενες παραμέτρους όπως, τυπικό σφάλμα, τυπική απόκλιση, συντελεστή διασποράς κτλ. Με τον τρόπο αυτό επισημαίνονται, αν υπάρχουν, ιδιομορφίες στα στοιχεία, όπως ύποπτες τιμές, έλλειψη κανονικότητας των στοιχείων ή χαμένες τιμές (ελλιπή στοιχεία). Έτσι, γίνεται εφικτή μία πρώτη γενική εικόνα των αποτελεσμάτων της μελέτης.
2. Επιλογή του κατάλληλου στατιστικού ελέγχου ή της σειράς με την οποία θα ακολουθήσει η διαδικασία των ελέγχων. Η φύση των στοιχείων και ο τρόπος συλλογής τους απαιτεί και τη χρήση κατάλληλης στατιστικής επεξεργασίας. Για παράδειγμα, μία ποσοτική μεταβλητή όπως είναι το βάρος ενός τροφίμου σε μία κονσέρβα (g) ή μία χημική συγκέντρωση σε ένα υγρό (g/ml γάλακτος ή g/g κρέατος), είναι κατάλληλη μόνο για την εφαρμογή ειδικών ελέγχων (t-test, F-test). Ενώ μία κατηγορική μεταβλητή, η οποία αποτελείται από διατεταγμένους ή μη αριθμούς (π.χ. προτίμηση ενός από 4 προϊόντα κωδικοποιημένα με τους αριθμούς 1-4), είναι κατάλληλη για να εφαρμοστεί ο έλεγχος χ^2 των συχνοτήτων. Στη συνέχεια ακολουθούν διάφορες συγκρίσεις μεταξύ ενός ή περισσότερων μέσων όρων ή συχνοτήτων των στηλών με σκοπό την εξαγωγή αποτελεσμάτων και γενίκευση αυτών για όλο τον πληθυσμό.
3. Εξαγωγή συμπερασμάτων και τελικών αποφάσεων από την επεξεργασία των αποτελεσμάτων. Η διαδικασία αυτή στηρίζεται στην επαρκή συλλογή στοιχείων που αποτελούν το δείγμα που ελήφθη από ένα πληθυσμό και με την προϋπόθεση ότι το δείγμα αυτό είναι αντιπροσωπευτικό του μελετώμενου πληθυσμού. Μόνο τότε μπορούν να εξάγονται συμπεράσματα με βεβαιότητα που αφορούν τις ιδιότητες αυτού του πληθυσμού. Η βεβαιότητα αυτή εκφράζεται με την εκτίμηση μιας πιθανότητας ή ποσοστού αξιοπιστίας των αποτελεσμάτων π.χ. με πιθανότητα σφάλματος 5% ή αντίστροφα με πιθανότητα ακρίβειας 95%. Για παράδειγμα, από ένα δείγμα ανδρών και γυναικών καταναλωτών μιας πόλης προκύπτει μετά την εφαρμογή κατάλληλης στατιστικής ανάλυσης, ότι οι άνδρες καταναλώνουν τρόφιμα με περισσότερες πρωτεΐνες απ' ότι οι γυναίκες, με πιθανότητα σφάλματος 5% στο εκτιμημένο αποτέλεσμα, και το γεγονός αυτό μπορεί να επεκταθεί περαιτέρω και για όλο τον αστικό πληθυσμό. Η δυνατότητα να εξάγονται συμπεράσματα που θα αφορούν ολόκληρο τον υπό μελέτη πληθυσμό από τη στατιστική και μόνο σύγκριση ενός μικρού αλλά αντιπροσωπευτικού τμήματος αυτού (δείγμα), καλείται συμπερασματική στατιστική.

1.2 ΚΛΙΜΑΚΕΣ ΤΩΝ ΣΤΟΙΧΕΙΩΝ

Τα στοιχεία όταν είναι ομοειδούς περιεχομένου γράφονται σε συγκεκριμένες στήλες. Κάθε σειρά μιας στήλης αποτελεί και ένα αριθμό ή επανάληψη, δίνοντας έτσι την εικόνα ενός λογιστικού φύλλου, δηλαδή πολλές στήλες με πλήθος αριθμών κατά μήκος τους. Κάθε στήλη αφορά ένα πλήθος μετρήσεων μιας συγκεκριμένης παραμέτρου π.χ. θερμοκρασίας, υγρασίας, μεγέθους συσκευασίας, βάρους προϊόντων και καλείται μεταβλητή. Αν όλες οι μετρήσεις είναι ίδιες, τότε ομιλούμε για σταθερά. Η φύση των μετρήσεων παρουσιάζει ποικιλομορφία με βάση την οποία οι μεταβλητές προσαρμόζονται σε διάφορες κλίμακες μέτρησης:

1. **Ονομαστική κλίμακα** (nominal scale). Οι αριθμοί αντιπροσωπεύουν μόνο ονόματα χωρίς καμία αριθμητική αξία και η διαφοροποίησή τους είναι απολύτως ποιοτική. Για παράδειγμα, το φύλο παριστάνεται με τους αριθμούς 1 για τα αρσενικά άτομα και 2 για τα θηλυκά. Το χρώμα των ακτινιδίων παριστάνεται επίσης με αριθμούς: 1- πράσινο, 2-κίτρινο ή και συνδυασμούς αυτών (3-κιτρινοπράσινο). Οι δοκιμαστές ενός προϊόντος σε καπνίζοντες (1) και μη καπνίζοντες (2). Οι κονσέρβες κρέατος απαριθμούνται ανάλογα με το μικροβιακό φορτίο σε επικίνδυνες (1) και μη (0) για τη δημόσια υγεία. Η στήλη (μεταβλητή) που περιέχει την

ονομαστική κλίμακα λέγεται και κατηγορική γιατί κατανέμεται σε κατηγορίες αριθμών με ομοειδή χαρακτηριστικά.

2. **Κλίμακα διαστημάτων** (interval scale). Είναι μία σπάνια διάταξη των αριθμών στην οποία ενώ οι αριθμοί έχουν ποσοτική αξία, αυτή προκαθορίζεται μόνο από την οριοθέτηση διαστημάτων στα οποία δεν υπάρχει πραγματικό μηδέν. Συγκεκριμένα, οι κλίμακες θερμοκρασίας των Fahrenheit και Celcius έχουν σταθερό διάστημα η καθεμία, μέσα στο οποίο μία διαφορά 20°C (25-5) έχει την ίδια ποσοτική αξία με μία διαφορά 20°F, όμως η θερμοκρασία των 60°C δεν παρουσιάζει τριπλάσια ζέστη από αυτή των 20°C, αφού το σημείο μηδέν και στις δύο κλίμακες θερμοκρασίας ορίστηκε αυθαίρετα. Το ίδιο ισχύει και για τη χρονολογική κλίμακα, πριν και μετά Χριστό: 2000 π.Χ. έτη δεν είναι διπλάσια από τα 1000 π.Χ. έτη καθότι στο έτος 0 π.Χ. δεν ξεκίνησε η ηλικία του κόσμου.

3. **Ποσοτική κλίμακα** ή **κλίμακα αναλογιών** (quantitative scale). Στην κλίμακα αυτή ανήκουν πλήθος γνωστών μεταβλητών: βάρος, μήκος, πυκνότητα, υγρασία, όγκος, και στις οποίες το μηδέν αποτελεί πραγματικό μηδέν. Έτσι, ένα μήλο διαμέτρου 10cm είναι διπλάσιο σε μέγεθος από ένα άλλο διαμέτρου 5cm και το αποξηραμένο του βάρος είναι υποπενταπλάσιο σε φυσική κατάσταση. Στην αναλογική κλίμακα οι μεταβλητές μπορούν να χωριστούν σε δύο κατηγορίες:

- *Συνεχής* ποσοτική μεταβλητή, στην οποία οι αριθμοί παίρνουν οποιαδήποτε μορφή τιμής: 1, 2, 2½, 2.73, 2.73162. Στην κατηγορία αυτή κατατάσσονται οι μεταβλητές που εύκολα μπορούν να υποδιαιρεθούν ανάλογα με τις δυνατότητες των οργάνων που χρησιμοποιούνται για τη μέτρηση αυτών: βάρος, μήκος, ύψος, όγκος, χημική συγκέντρωση.
- *Ασυνεχής* ποσοτική μεταβλητή, στην οποία οι αριθμοί έχουν συγκεκριμένη μορφή τιμής, π.χ. ο αριθμός των καταναλωτών μιας περιοχής, των δοκιμαστών μιας οργανοληπτικής δοκιμής, ο αριθμός των αυγών σολωμού που διατίθενται σε γυάλινα σκευάσματα πώλησης, ο αριθμός των κλωστηριδίων ανά g κρέατος. Δεν μπορούμε να αναφερόμαστε σε 3.5 κλωστηρίδια ανά g κρέατος ούτε και σε 135.6 αυγά ανά φιαλίδιο.

4. **Κλίμακα διαβάθμισης** ή **τακτική κλίμακα**. Οι αριθμοί της μεταβλητής αντιπροσωπεύουν βαθμίδες αριθμητικής κλίμακας, π.χ. 1,2,3...15,16 κτλ. στις οποίες το 1 είναι μικρότερο του 2, όπως και το 6 από το 7, αλλά σε καμία περίπτωση δεν υπονοούν συγκεκριμένο μέγεθος ποσοτικής μεταβολής. Για παράδειγμα, το μέγεθος των μικροβίων στο γάλα ή στο νερό μετά από μικροσκοπική ανάλυση κατατάσσεται σε 4 κατηγορίες αυξανόμενου μεγέθους: ❍, ❍, ❍, ❍, στις οποίες δίνονται αντίστοιχα οι αριθμοί 1, 2, 3 και 4. Είναι προφανές ότι το 2 είναι μεγαλύτερο του 1, αλλά όχι υποχρεωτικά διπλάσιο σε μέγεθος. Συνεπώς, στην κλίμακα διαβάθμισης μετρείται η σχετική διαφορά δύο αριθμών εκφραζόμενη με ανισότητα και όχι η ποσοτική. Στο σημείο αυτό επιτάσσεται η ανάγκη διασαφήνισης μεταξύ των διαβαθμισμένων και τακτικών μεταβλητών:

- Οι τακτικές μεταβλητές (ordinal variables) περιέχουν κατηγορίες στις οποίες αποδίδεται μία διαφορετική βαθμίδα τη φορά ώστε να δημιουργηθεί μία αυξητική ή μειωτική κλίμακα διαβαθμίσεων της τάξης π.χ. 1-5. Τα διαστήματα μεταξύ δύο βαθμίδων σε καμία περίπτωση δεν έχουν τη μορφή ίσων αποστάσεων, απλώς ισχύει η έννοια της ανισότητας, π.χ. $2<3<4$ κτλ. Ο όρος αυτός είναι δημοφιλής στις κοινωνικές επιστήμες.
- Οι διαβαθμισμένες μεταβλητές (rank variables) αποτελούν μετασχηματισμένες ποσοτικές μεταβλητές σε διάταξη βαθμίδων αυξητικής ή μειωτικής κλίμακας και προορίζονται κυρίως για άμεση χρήση σε μη παραμετρικές στατιστικές αναλύσεις.

Επισημαίνεται ιδιαίτερα, ότι οι επιστήμονες διαχειρίζονται στατιστικά μία τακτική μεταβλητή είτε ως κατηγορική, και επομένως υπό μορφή συχνοτήτων (μη παραμετρική στατιστική), είτε ως ποσοτική και άρα με τη χρήση της παραμετρικής στατιστικής, παρά τη γενική διαπίστωση ότι οι βαθμίδες δεν έχουν πραγματική ποσοτική υπόσταση. Ανεξαρτήτως, πάντως, από τις δύο τάσεις που διαμορφώνονται στην επιστημονική κοινότητα, συνιστάται πρωτίστως η γραφική απεικόνιση των συχνοτήτων των βαθμίδων, δεδομένου ότι ενίοτε ο μέσος όρος των τιμών να μην αντιπροσωπεύει ούτε στο ελάχιστο την κατανομή των βαθμίδων όπως γλαφυρά επεξηγείται στο παράρτημα Α.

1.3 ΚΛΙΜΑΚΕΣ ΣΤΗΝ ΟΡΓΑΝΟΛΗΠΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΤΡΟΦΙΜΩΝ

Στην περίπτωση του οργανοληπτικού ελέγχου ενός προϊόντος (τρόφιμο ή ποτό) ακολουθούνται άλλες πρακτικές που αξίζει να αναφερθούν. Βασική προϋπόθεση για την επιτυχή έκβαση μιας έρευνας και της μετέπειτα στατιστικής ανάλυσης είναι η επιλογή της κατάλληλης μεθόδου βαθμολόγησης (αξιολόγησης) των στοιχείων. Τα στοιχεία αυτά στην επιστήμη τροφίμων προέρχονται από τη μελέτη ορισμένων χαρακτηριστικών των προϊόντων, όπως για παράδειγμα οσμή, γεύση, χρώμα, ευθρυπτότητα, ελαστικότητα, ιξώδες, στα οποία εφαρμόζεται μία ποικιλία κλιμάκων βαθμολόγησης:

▼ Κλίμακα της *ποιοτικής περιγραφικής βαθμολόγησης* (descriptive grading), στην οποία το τρόφιμο ταξινομείται από ειδικούς εκτιμητές ονομαστικά ανάλογα με την παρουσία ή μη κάποιων χαρακτηριστικών. Ένα ψάρι, για παράδειγμα, βαθμολογείται συνολικά ποιοτικά από 0 μέχρι 18 με το *δείκτη ποιότητας* QIM (Quality Index Method) ανάλογα με την εμφάνιση που παρουσιάζει, είτε μεμονωμένα είτε σε συνδυασμό, ως προς τη φύση της βλέννας, οσμή και φωτεινότηα δέρματος, το βαθμό ακαμψίας, το χρώμα των βραγχίων, τη θολότητα του ματιού, ή και άλλα χαρακτηριστικά, καθώς αυτό χάνει τη νωπότητα του με την αύξηση του χρόνου ψύξης σε θρύμματα πάγου. Η αριθμητική αξιολόγηση αυτής της μεθόδου δεν έχει καμία σχέση με την ποσοτική κλίμακα και την κλίμακα διαβάθμισης γι' αυτό και στατιστικά δύσκολα ελέγχεται. Το ενδιαφέρον στην κλίμακα αυτή εστιάζεται απλώς στη συχνότητα εμφάνισης των διαφορετικών κατηγοριών της κλίμακας, η καθεμία των οποίων βαθμολογείται από 0 (νωπό) μέχρι 3 (αλλοιωμένο).

▼ Κλίμακα της *βαθμιδωτής κατάταξης* (ranking scale) της έντασης ενός χαρακτηριστικού στην οποία μία σειρά από τρία ή περισσότερα δείγματα ενός τροφίμου προσφέρονται για οργανοληπτική δοκιμή ταυτόχρονα, με στόχο την αξιολόγησή τους σε μία κλίμακα έντασης χωρίς σαφή ποσοτικό διαχωρισμό του κάθε χαρακτηριστικού. Η κλίμακα αυτή δηλαδή, αδυνατεί να παρέχει πληροφορίες ως προς το μέγεθος των διαφορών μεταξύ των δειγμάτων. Το εύρος της κλίμακας κατάταξης καθορίζεται αποκλειστικά και εξαντλείται από τον αριθμό των δειγμάτων που εξετάζονται κάθε φορά. Όταν δοκιμάζονται για παράδειγμα 5 δείγματα, τότε αυτά κατατάσσονται υποχρεωτικά με τους αριθμούς 1 μέχρι 5 σε μειούμενη κλίμακα έντασης του χαρακτηριστικού (ή και το αντίθετο) και επομένως το εύρος της κλίμακας είναι 1-5. Το δείγμα με τη μεγαλύτερη ένταση, π.χ. της ελαστικότητας, βαθμολογείται με 1 και το δείγμα με τη μικρότερη ένταση βαθμολογείται με 5, τα δε υπόλοιπα τρία δείγματα θα λάβουν τις ενδιάμεσες τιμές. Δεν αποκλείεται το ενδεχόμενο κάποια δείγματα να αξιολογηθούν με τον ίδιο βαθμό της κλίμακας κατάταξης (ισοψηφίες) όταν ο διαχωρισμός της έντασης σε αυτά είναι ασαφής (π.χ. τα παραπάνω 5 δείγματα αξιολογούνται ως 1, 2, 2, 4 και 5, οπότε η τελική τους βαθμολογία τροποποιείται ως, 1, 2.5, 2.5, 4 και 5). Η μέθοδος κατάταξης κρίνεται πολύ κατάλληλη σε δοκιμασίες χρώματος ή οσμής, όχι όμως στις συνήθεις γευστικές δοκιμασίες στις οποίες, αν τελικά χρησιμοποιηθεί, αυξάνεται η ευαισθησία της μόνο όταν διαπιστώνεται μεγάλη γευστική διαφορά μεταξύ των δειγμάτων και ταυτόχρονα επείγεται η ανάγκη ταχείας εκτίμησης πολλών δειγμάτων απευθείας από τον ίδιο δοκιμαστή. Πλεονέκτημα της μεθόδου είναι ότι δεν απαιτούνται ειδικοί δοκιμαστές αλλά ούτε και προηγούμενη εκμάθηση των ανεκπαίδευτων δοκιμαστών στη βαθμιδωτή κλίμακα κατάταξης. Για τη μέθοδο αυτή έχει προσαρμοστεί ειδική στατιστική διαδικασία γνωστή και ως έλεγχος του Kramer (ενότητα 10.5).

▼ Βαθμολογική κατάταξη της έντασης του χαρακτηριστικού σε *δομημένη κλίμακα κατηγοριών* (structural category scale), στην οποία τα μελετώμενα χαρακτηριστικά διαβαθμίζονται σε κατηγορίες κατάταξης (βαθμίδες) αυξανόμενης αισθητικής έντασης που λέγονται χαρακτηρισμοί. Ένα τρόφιμο για παράδειγμα εξετάζεται ως προς το χαρακτηριστικό 'σκληρότητα', το οποίο ακολούθως αξιολογείται ως πολύ σκληρό, μετρίως σκληρό, λίγο μαλακό πολύ μαλακό κοκ., και ταυτόχρονα προσαρμόζεται κάποια κλίμακα μεταβολής βαθμίδων. Ο δοκιμαστής εκτιμά κάθε δείγμα χωριστά ή και ταυτόχρονα με άλλα και αποδίδει ένα βαθμό στο καθένα, με αποτέλεσμα κάποια δείγματα να αξιολογηθούν με την ίδια τιμή ή επίσης κάποιες βαθμίδες της κλίμακας ουδέποτε να καταγραφούν. Για παράδειγμα, αν οι χαρακτηρισμοί της κλίμακας αριθμούνταν από 1 μέχρι 8 βαθμίδες είναι πιθανόν οι βαθμοί αξιολόγησης να κυμαίνονται από το 3 μέχρι το 6. Αντίθετα, στην κλίμακα της βαθμιδωτής κατάταξης οι βαθμοί όλων των προϊόντων είναι διαφορετικοί βαθμοί κατάταξης και όλοι αυτοί αναγκαστικά θα αποτελέσουν και το μέγιστο εύρος της κλίμακας.

Το μεγάλο μειονέκτημα του οργανοληπτικού ελέγχου βρίσκεται στη δυσκολία επινόησης μίας ποσοτικής αριθμητικής κλίμακας για κάθε μελετώμενο χαρακτηριστικό. Η ποσοτική κλίμακα υπονοεί όχι μόνο τη δυνατότητα αφαίρεσης μεταξύ δύο αριθμών (25-5=20) αλλά και την εκτίμηση του υπο/πολλαπλάσιου αυτών (25:5=5), να υπάρχει δηλαδή πραγματικό μηδέν, οπότε θα διέθεται πλήρη στατιστική κάλυψη με δυνατότητα σαφέστερης συμπερασματολογίας. Οι κλίμακες που ευρύτατα ή και αναγκαστικά χρησιμοποιούνται για οργανοληπτικό έλεγχο, είναι κλίμακες διαβάθμισης χαρακτηρισμών και διακρίνονται: σε διπολικές, όταν αντίθετοι χαρακτηρισμοί ενός προϊόντος ορίζουν τα δύο άκρα της κλίμακας, π.χ. σκληρό-μαλακό, και σε πολικές, όταν το ένα άκρο της κλίμακας εκκινεί από το μηδέν π.χ. καθόλου αλμυρό, και περατώνεται στο άλλο άκρο με το χαρακτηρισμό εξαιρετικά πολύ αλμυρό ή κάτι ανάλογο. Κάθε χαρακτηριστικό υποβάλλεται σε αισθητικό έλεγχο από μία ομάδα εκτιμητών (ή κριτών ή δοκιμαστών), οι οποίοι θα το αξιολογήσουν με ένα αριθμό που αντιστοιχεί σε ένα χαρακτηρισμό και ο οποίος αντιπροσωπεύει κάποια βαθμίδα έντασης, π.χ. λίγο αλμυρό, λίγο τραχύ, πολύ αλμυρό, πολύ τραχύ κτλ. Οι πλέον διαδεδομένοι χαρακτηρισμοί στις διαβαθμισμένες κλίμακες είναι οι εξής:

α) Χαρακτηρισμοί σε πολικές κλίμακες με αρχή το 0 ή 1 ως πρώτη βαθμίδα και κατάλληλη αντιστοίχιση τους στις βαθμίδες αναλόγως της φύσης του εξεταζόμενου προϊόντος,

καθόλου	**λίγο**	**μέτρια**	**αρκετά**	**πολύ**
1	2	3	4	5

αλμυρό αλμυρό

καθόλου	**ελαφρά**	**μέτρια**	**έντονα**	**πολύ έντονα**
1	2	3	4	5

ερυθρό ερυθρό

Αν η οργανοληπτική ευαισθησία των κριτών είναι μεγάλη ή με άλλα λόγια είναι έμπειροι, τότε οι κατηγορίες στις κλίμακες μπορούν να υποδιαιρεθούν ως εξής:

καθόλου		λίγο		μέτρια		πολύ		εξαιρετικά πολύ	
καθόλου		ελαφρά		μέτρια		έντονα		πολύ έντονα	
1	2	3	4	5	6	7	8	9	10

όπου σε κάθε χαρακτηρισμό αποδίδονται δύο βαθμίδες. Με την τακτική αυτή διευκολύνονται οι δοκιμαστές να εκφράζουν με μεγαλύτερη ευχέρεια την οργανοληπτική αίσθηση, κυρίως όμως βελτιώνεται σημαντικά η διαχείριση των στοιχείων όταν πρόκειται να εφαρμοστεί η παραμετρική στατιστική. Μεγαλύτερο εύρος κλίμακας επιτρέπει καλύτερη συμπεριφορά της κανονικότητας των στοιχείων και ακριβέστερη εκτίμηση των όρων της περιγραφικής στατιστικής.

β) Χαρακτηρισμοί σε διπολικές κλίμακες με ή χωρίς την ενσωμάτωση της βαθμίδας του ουδέτερου χαρακτηρισμού στο κέντρο (π.χ. ούτε κρύο – ούτε ζεστό)

εξαιρετικά πολύ	πολύ	μέτρια	ελαφρά	ελαφρά	μέτρια	πολύ	εξαιρετικά πολύ
-4	-3	-2	-1	1	2	3	4

μαλακό
κρύο
εύθρυπτο

σκληρό
ζεστό
συμπαγές

Η επιλογή των βαθμίδων στις κατηγορίες των κλιμάκων είναι αυθαίρετη, γι` αυτό και δεν είναι σπάνια η παρουσία αρνητικών τιμών, οι οποίες όμως καλό είναι να αντικαθίστανται με θετική κλίμακα αριθμών, ειδάλλως προκαλούνται προβλήματα στην εκτέλεση της παραμετρικής στατιστικής. Για παράδειγμα, μετασχηματισμοί του λογαρίθμου και της τετραγωνικής ρίζας δεν μπορούν να εφαρμοστούν παρουσία αρνητικών τιμών.

γ) Χαρακτηρισμοί *γραμμικής μεταβολής ή αδιαβάθμητης κλίμακας* (unstructured scale). Οι κλίμακες αυτές στερούνται ουσιαστικά της δομής των βαθμίδων και χαρακτηρίζονται αποκλειστικά από δύο κατηγορίες χαρακτηρισμών, το ένα άκρο των οποίων είναι είτε πολικό είτε διπολικό:

καθόλου
αλμυρό

εξαιρετικά
πολύ
αλμυρό

εξαιρετικά
πολύ
μαλακό

εξαιρετικά
πολύ
σκληρό

Οι εκτιμητές αξιολογούν κάθε χαρακτηριστικό που δοκιμάζουν σημειώνοντας την ένταση του με μία στίξη (x ή I) πάνω στην ευθεία. Η απόσταση της κλίμακας ανάγεται στη συνέχεια σε αριθμητική συνήθως με τη βοήθεια της κλίμακας μήκους ενός υποδεκάμετρου (χάρακα) και έτσι οι σημειούμενες προτιμήσεις των δοκιμαστών παίρνουν διαστάσεις ποσοτικής μεταβολής. Η κλίμακα μήκους 15cm αποτελεί τη συνηθέστερη επιλογή των ερευνητών. Μειονέκτημα της μεθόδου είναι ότι τα άκρα της ευθείας θα πρέπει να προσδιοριστούν με πολλαπλά δοκιμαστικά πρότυπα (standards), πράγμα που απαιτεί εντατική εκπαίδευση των δοκιμαστών. Η μέθοδος αυτή εφαρμόζεται με ενθαρρυντικά αποτελέσματα και δίχως τη διαδικασία των δοκιμαστικών προτύπων βρίσκοντας ολοένα και μεγαλύτερη απήχηση διεθνώς.

δ) Χαρακτηρισμοί σε κλίμακες αποδοχής ή αρέσκειας ή προτίμησης ή *κλίμακες αισθητικής απόλαυσης* (hedonic scales), οι οποίες σχετίζονται με τις έννοιες αξιολόγησης ενός προϊόντος ως ευχάριστο ή δυσάρεστο. Οι κλίμακες αρέσκειας είναι συνήθως διπολικές και φέρουν στο κέντρο της κλίμακας τους μία ουδέτερη βαθμίδα (ούτε ευχάριστο ούτε δυσάρεστο). Επίσης, χαρακτηρίζονται από 9 κατηγορίες βαθμίδων οι οποίες βαθμολογούνται συνήθως από -4 μέχρι +4 και διαφέρουν από τις άλλες κλίμακες στο ότι σχετίζονται με χαρακτηρισμούς ευχαρίστησης, όπως γευστικής απόλαυσης και όχι με χαρακτηρισμούς έντασης φυσικών χαρακτηριστικών (μαλακό, τραγανό, χυμώδες, εύθρυπτο). Για το λόγο αυτό τα χαρακτηριστικά στα οποία μετρούμε ένταση προτίμησης (π.χ. χρώμα, οσμή, γεύση, γενική εμφάνιση ενός προϊόντος κτλ.) βασίζονται σε υποκειμενική κρίση των δοκιμαστών σε αντιδιαστολή με τα λοιπά χαρακτηριστικά που βασίζονται σε αντικειμενική κρίση (Παράδειγμα 1.1).

εξαιρετικά πολύ	πάρα πολύ	μέτρια	λίγο	ούτε δυσάρεστο ούτε ευχάριστο	λίγο	μέτρια	πάρα πολύ	εξαιρετικά πολύ
-4	-3	-2	-1	0	1	2	3	4
ΔΥΣΑΡΕΣΤΟ								ΕΥΧΑΡΙΣΤΟ

Παράδειγμα 1.1. Δελτίο οργανοληπτικού ελέγχου έξι σκευασμάτων από παριζάκι (μεταχειρίσεις Α-Ζ) με αυξημένα ποσοστά προσθήκης μηχανικά αποστεωμένου κοτόπουλου (ΜΑΚ) και με συμμετοχή n αριθμού δοκιμαστών. Τα εξεταζόμενα χαρακτηριστικά ανήκουν στην ομάδα των χαρακτηριστικών που βαθμολογούνται με υποκειμενική κρίση, πλην της συνεκτικότητας. Οι μεταχειρίσεις (σκευάσματα) με τους κωδικούς αριθμούς 904, 132, 923, 607, 145 και 239 αντιστοιχούν στις μεταχειρίσεις Ε, Β, Α, Ζ, Γ και Δ αντίστοιχα, και με τον τρόπο αυτό διασφαλίζεται ότι αυτές σε καμία περίπτωση δεν γίνονται γνωστές στο δοκιμαστή πριν και κατά την εξέλιξη της οργανοληπτικής δοκιμής, για να μην επηρεάζεται η κρίση του βαθμολογικά.

ΑΥΞΟΝΤΑΣ ΑΡΙΘΜΟΣ......... ΗΜΕΡΟΜΗΝΙΑ.....................

ΟΡΓΑΝΟΛΗΠΤΙΚΟΣ ΕΛΕΓΧΟΣ
ΜΕ ΚΛΙΜΑΚΑ ΒΑΘΜΙΔΩΝ

ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΔΟΚΙΜΑΣΤΗ: ....| |

ΠΡΟΙΟΝ: ΠΑΡΙΖΑΚΙ ΑΠΟ ΜΑΚ 6 ΜΕΤΑΧΕΙΡΙΣΕΩΝ (Α-Ζ)

	ΜΕΤΑΧΕΙΡΙΣΕΙΣ (ΚΩΔΙΚΟΙ)					
ΟΡΓΑΝΟΛΗΠΤΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	904	132	923	607	145	239
ΑΡΩΜΑ						
ΧΡΩΜΑ						
ΣΥΝΕΚΤΙΚΟΤΗΤΑ						

<table>
<tr><th></th><th colspan="10">ΚΛΙΜΑΚΑ ΒΑΘΜΙΔΩΝ</th></tr>
<tr><th></th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th></tr>
<tr><td>ΑΡΩΜΑ</td><td colspan="2">ΔΥΣΑΡΕΣΤΟ</td><td colspan="2">ΕΛΑΦΡΩΣ ΔΥΣΑΡΕΣΤΟ</td><td colspan="2">ΟΥΤΕ ΔΥΣΑΡΕΣΤΟ ΟΥΤΕ ΕΥΧΑΡΙΣΤΟ</td><td colspan="2">ΕΛΑΦΡΩΣ ΕΥΧΑΡΙΣΤΟ</td><td colspan="2">ΕΥΧΑΡΙΣΤΟ</td></tr>
<tr><td>ΧΡΩΜΑ</td><td colspan="2">ΑΣΧΗΜΟ</td><td colspan="2">ΕΛΑΦΡΩΣ ΑΣΧΗΜΟ</td><td colspan="2">ΟΥΔΕΤΕΡΟ</td><td colspan="2">ΕΛΑΦΡΩΣ ΩΡΑΙΟ</td><td colspan="2">ΩΡΑΙΟ</td></tr>
<tr><td>ΣΥΝΕΚΤΙΚΟΤΗΤΑ</td><td colspan="2">ΜΑΛΑΚΟ</td><td colspan="2">ΛΙΓΟ ΜΑΛΑΚΟ</td><td colspan="2">ΟΥΤΕ ΜΑΛΑΚΟ ΟΥΤΕ ΣΦΙΚΤΟ</td><td colspan="2">ΕΛΑΦΡΩΣ ΣΦΙΚΤΟ</td><td colspan="2">ΣΦΙΚΤΟ</td></tr>
<tr><td>ΓΕΝΙΚΗ ΚΡΙΣΗ</td><td colspan="2">ΑΠΑΡΑΔΕΚΤΟ</td><td colspan="2">ΥΠΟΦΕΡΤΟ</td><td colspan="2">ΑΔΙΑΦΟΡΟ</td><td colspan="2">ΜΕΤΡΙΟ</td><td colspan="2">ΚΑΛΟ</td></tr>
</table>

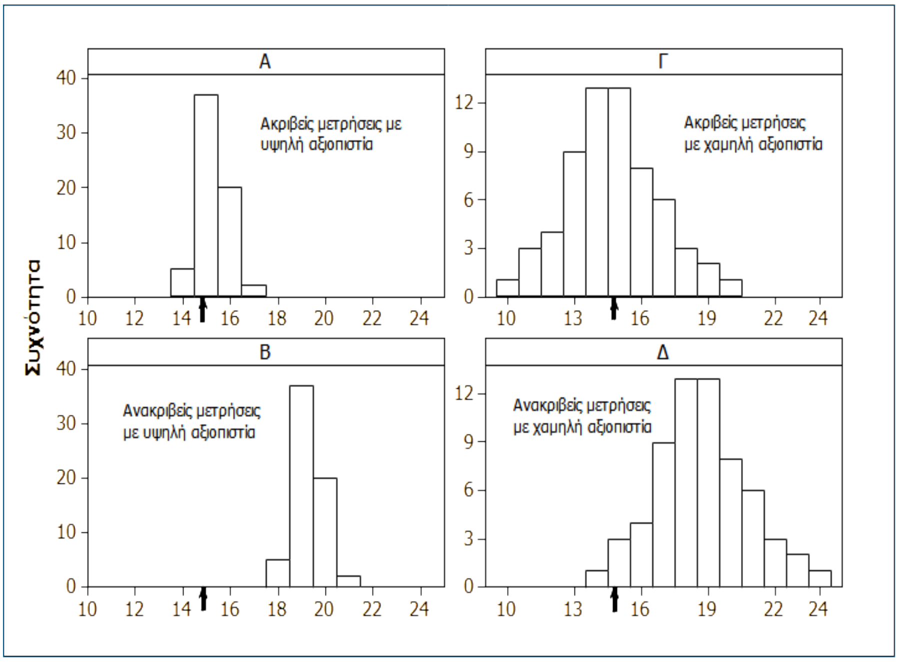
Η αίσθηση της προτίμησης μπορεί άνετα να προσαρμοστεί και σε αδιαβάθμητη κλίμακα με τον προσδιορισμό των δύο άκρων συνήθως ως καθόλου ευχάριστο και πολύ ευχάριστο. Η αισθητική απόλαυση διαφέρει επιπρόσθετα από τις αντικειμενικές μεταβλητές στο ότι οι βαθμοί που αποδίδονται από τους δοκιμαστές δεν αναμένεται να αυξάνουν ή να μειώνονται μονοτονικά με την αύξηση ή μείωση της έντασης του χαρακτηριστικού, αλλά εμφανίζουν μία κορυφή εξαιτίας της συσσώρευσης πολλών επαναληπτικών τιμών

2

Περιγραφική στατιστική

Οι περισσότερες παρατηρήσεις κάθε δείγματος έχουν την τάση να διασπείρονται κοντά και γύρω από το μέσο του εύρους των τιμών τους. Η τάση αυτή των αριθμών λέγεται κεντρική και βοηθάει στη λήψη σημαντικών πληροφοριών για το ποιόν του δείγματος, όπως είναι ο μέσος όρος και η διάμεσος, ενώ ο τρόπος με τον οποίο οι αριθμοί διατάσσονται γύρω από το κέντρο του εύρους τους εκτιμάται με την περιγραφή της διασποράς τους, όπως είναι το στατιστικό εύρος και η τυπική απόκλιση. Οι εκτιμήσεις αυτές που αφορούν την κεντρική τάση και τη διασπορά του δείγματος αποτελούν τις παραμέτρους του πληθυσμού. Προηγουμένως όμως, ένα δείγμα για να θεωρείται ότι αποτελεί επαρκή εκτίμηση του μελετώμενου πληθυσμού, θα πρέπει να πληρεί τις εξής προϋποθέσεις:

1. Να είναι *αμερόληπτο*. Είναι προφανές ότι αν επαναλάβουμε τη λήψη ενός δείγματος αρκετές φορές από τον ίδιο πληθυσμό, οι παράμετροι του πληθυσμού θα δίνουν σίγουρα διαφορετικές εκτιμήσεις. Μερικές θα υπερεκτιμούν τον πληθυσμό και άλλες θα τον υποεκτιμούν. Όμως, τα επαναληπτικά δείγματα τελικά θα δώσουν παραμέτρους που θα τείνουν να πλησιάσουν την αυθεντική (πραγματική) τιμή τους.
2. Να είναι *αξιόπιστο* και *ακριβές* (βλέπε και ενότητα 2.5), δηλαδή οι παράμετροι του δείγματος με μία μόνο λήψη να πλησιάζουν τις αντίστοιχες τιμές του πληθυσμού (Σχ. 2.1). Η ακρίβεια (accuracy) αναφέρεται στην προσέγγιση μιας μέτρησης ως προς την πραγματική τιμή της μεταβλητής ή με άλλα λόγια πόσο κοντά η μέτρηση πλησιάζει στην πραγματική. Η αξιοπιστία (precision) αναφέρεται στην προσέγγιση μεταξύ των αλλεπάλληλων μετρήσεων της ίδιας ποσότητας. Η αξιοπιστία ελαττώνεται όσο αυξάνει η διαφορά μεταξύ των διαδοχικών μετρήσεων. Ως παράδειγμα ακρίβειας αναφέρεται ο αναλυτικός προσδιορισμός της γλυκόζης με τη χρήση φασματοφωτόμετρου. Ένα διάλυμα γλυκόζης γνωστής συγκέντρωσης συγκρίνεται με ένα άλλο ίδιας υποτιθέμενης περιεκτικότητας για να εκτιμηθεί η ακρίβεια της μετρούμενης ποσότητας, πόσο δηλαδή απομακρύνεται η τιμή αυτή από την αληθινή. Κατά την καταγραφή των μετρήσεων μιας συνεχούς μεταβλητής είναι αναγκαίο, επιπρόσθετα, να προσδιοριστεί η ακρίβεια με την οποία ελήφθησαν οι μετρήσεις. Έτσι, μία τιμή βάρους 8g σημαίνει εύρος ακρίβειας 1g, μία τιμή 8.6g υπονοεί εύρος ακρίβειας 0.1g και 8.64g εύρος 0.01g. Όσο αυξάνουν τα δεκαδικά ψηφία τόσο αυξάνει και η ακρίβεια μιας μέτρησης, όμως η αύξηση



Σχ. 2.1. Γραφικός προσδιορισμός (ιστογράμματα συχνοτήτων) των εννοιών της ακρίβειας και αξιοπιστίας ενός δείγματος με πραγματικό μέσο όρο 14.9.

αυτή δεν συνεχίζεται ασταμάτητα, αλλά περιορίζεται στα μέγιστα όρια ακρίβειας που δίνει η συσκευή με την οποία γίνονται οι μετρήσεις. Έτσι, οποιαδήποτε μέτρηση γίνει σε ζυγαριά υψηλής ακρίβειας δυνάμενης να μετρήσει μέχρι 0.0001g, αυτή μπορεί να φτάσει μέχρι και τέσσερα δεκαδικά ψηφία, π.χ. 8.6452g. Τα ψηφία λοιπόν ενός αριθμού που αναφέρονται στην ακρίβεια της μέτρησης λέγονται σημαντικά ψηφία και αξιολογούνται ως εξής: ο αριθμός 8 έχει ένα σημαντικό ψηφίο, ο αριθμός 8.6 δύο, ο αριθμός 8.64 τρία κοκ. Είναι προφανές ότι η έννοια των σημαντικών ψηφίων στις ασυνεχείς μεταβλητές είναι αδόκιμη. Τονίζεται επίσης, ότι μία συνεχής μεταβλητή δεν πρέπει να συγχέεται με μία ασυνεχή εξαιτίας του τρόπου μέτρησης αυτής, όπως π.χ. όταν μετρείται το βάρος ενός αρακά χωρίς δεκαδικά, εξαιτίας της χαμηλής ακρίβειας της ζυγαριάς.

3. Να είναι *συνεπές*, δηλαδή όσο το δειγματοληπτικό μέγεθος αυξάνει τόσο οι παράμετροι του δείγματος να πλησιάζουν περισσότερο στις πραγματικές εκτιμήσεις του πληθυσμού. Με άλλα λόγια, αν το δείγμα αφορούσε όλο το μέγεθος του πληθυσμού, τότε οι παράμετροι θα έδιναν εκτιμήσεις με ακρίβεια 100%.

2.1 ΕΚΤΙΜΗΣΗ ΤΗΣ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ ΤΩΝ ΑΡΙΘΜΩΝ

Οι πιο γνωστές παράμετροι της *κεντρικής τάσης* των αριθμών ενός δείγματος είναι:

(α) Ο *αριθμητικός μέσος όρος* $\bar{X}$ (arithmetic mean) ή απλά μέσος όρος (Παράδειγμα 2.1), ο οποίος είναι το πηλίκο του αθροίσματος όλων των τιμών του δείγματος διαιρούμενο με το πλήθος αυτών,

$$\bar{X} = \frac{\sum X_i}{n}$$

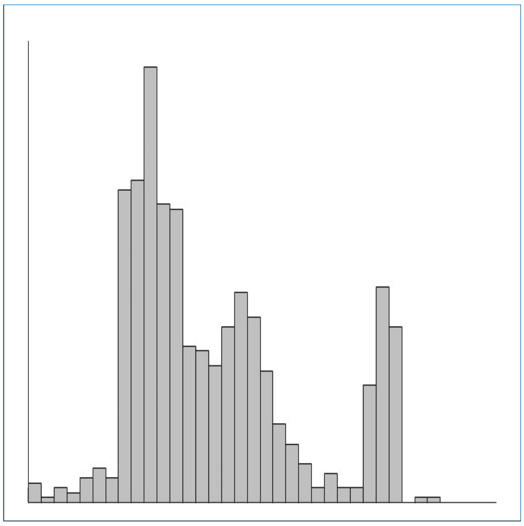
όπου X_i είναι μία τιμή του δείγματος, n το πλήθος των παρατηρήσεων και $\Sigma X_i = X_1 + X_2 + ... + X_n$, είναι το άθροισμα των τιμών του δείγματος. Ειδική περίπτωση αποτελεί ο *σταθμισμένος* ή *εξισορροπημένος* μέσος όρος (weighted mean) ο οποίος αφορά τη μέση τιμή μέσων όρων k δειγμάτων, στους μέσους όρους των οποίων δίνονται ατομικά ειδικές τιμές βαρύτητας w_i γνωστές και ως *συντελεστές στάθμισης* (weights):

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{k} w_i \bar{X}_i}{\sum_{i=1}^{k} w_i} = \frac{w_1 \bar{X}_1 + w_2 \bar{X}_2 + ... + w_k \bar{X}_k}{w_1 + w_2 + ... + w_k}$$

Για παράδειγμα, αν από τρία δείγματα προκύπτουν οι μέσοι όροι 2.34, 3.56 και 2.45, με αντίστοιχους συντελεστές στάθμισης 0.2, 0.7 και 0.1, τότε ο σταθμισμένος μέσος θα ισούται με: (2.34·0.2+3.56·0.7+2.45·0.1)/(0.2+0.7+0.1)=3.205.

(β) Η *διάμεσος m* (median), η οποία προκύπτει ως η τιμή της παρατήρησης που βρίσκεται ακριβώς στο μέσο των παρατηρήσεων, όταν αυτές διαταχθούν με αύξουσα ή φθίνουσα τάξη (Παράδειγμα 2.1). Όταν το πλήθος των παρατηρήσεων είναι μονός αριθμός η διάμεσος είναι η τιμή ακριβώς στο κέντρο της στήλης ($X_{(n+1)/2}$). Σε ζυγό αριθμό παρατηρήσεων η διάμεσος είναι το πηλίκο του αθροίσματος των δύο κεντρικών τιμών διαιρούμενο δια δύο ($[X_{(n/2)}+X_{(n/2)+1}]/2$). Η διάμεσος συμπίπτει με το μέσο όρο μόνον όταν μία κατανομή είναι συμμετρική και ταυτόχρονα κανονική (μεσόκυρτη) και σε εξαιρετική περίπτωση, όταν μία κατανομή είναι συμμετρική και δικόρυφη. Η διάμεσος δίνει ακριβέστερες πληροφορίες απ' ότι ο μέσος όρος στις περιπτώσεις δειγμάτων με μη κανονικές κατανομές (ασύμμετρες), γιατί επηρεάζεται λιγότερο έντονα από τις ακραίες τιμές των δειγμάτων, καθώς και στις περιπτώσεις διαβαθμισμένων μεταβλητών.

(γ) Η *κορυφή* (mode) μιας στήλης αντιστοιχεί στη μέτρηση με τη μεγαλύτερη συχνότητα (επικρατούσα τιμή) και εμφανίζεται στο ιστόγραμμα συχνότητας των μεταβλητών, στο οποίο γίνεται αντιληπτή στο μέγιστο ύψος αυτού (Σχ. 2.2). Η κορυφή εμφανίζεται μόνο μία φορά στην κανονική κατανομή, οπότε και συμπίπτει με το μέσο όρο και τη διάμεσο, και περισσότερες φορές σε μη κανονικές κατανομές, π.χ. δικόρυφη, τρικόρυφη, πολυκόρυφη. Σε στήλες με όχι έντονη ασυμμετρία η κορυφή ισούται εμπειρικά με $3m - 2\bar{X}$ (Παράδειγμα 2.1).



Σχ. 2.2. Μη κανονική καμπύλη με τρεις κορυφές.

Όταν σε ένα γράφημα κατανομής των συχνοτήτων ενός δείγματος η διάμεσος και η κορυφή κείνται αριστερά του μέσου όρου, τότε η κατανομή παρουσιάζει θετική ασυμμετρία (βλέπε ενότητα 2.3 και σχήμα

2.6). Σε κατανομές με αρνητική ασυμμετρία, η διάμεσος και η κορυφή είναι μεγαλύτερες από το μέσο όρο.
(δ) Το *μεσοδιάστημα* (midrange), το οποίο είναι το μισό της απόστασης που ορίζεται από τη μέγιστη και ελάχιστη τιμή (εύρος τιμών) του δείγματος. Δεν θεωρείται σπουδαία μέτρηση της κεντρικής τάσης των αριθμών και η χρήση του περιορίζεται μόνο σε διαχρονικές μεταβολές της θερμοκρασίας, ημερήσιες, μηνιαίες κτλ., στις οποίες καταγράφεται πάντοτε ένα μέγιστο και ένα ελάχιστο αυτής και που λανθασμένα επικρατεί να κατονομάζεται ως μέση θερμοκρασιακή μεταβολή. Σε μεταβλητές με κανονική κατανομή των στοιχείων και χωρίς ακραίες τιμές το μεσοδιάστημα έχει ισχυρή γραμμική σχέση με την τυπική απόκλιση και μπορεί να την υποκαταστήσει σε διάφορες εφαρμογές.
ε) Ο *γεωμετρικός μέσος όρος* GM (geometric mean), είναι η νιοστή ρίζα του παράγωγου

$$GM = \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \ldots X_n} = \sqrt[n]{\prod_{i=1}^{n} X_i}$$

υπολογίζεται επίσης ως ο αντιλογάριθμος του αριθμητικού μέσου των λογαριθμημένων τιμών. Χρησιμοποιείται σπανίως και σε περιπτώσεις μόνο μεταβλητών που έχουν σχέση με ποσοστά και γενικώς σε εκθετικές (λογαριθμικές) μεταβολές.
(ζ) Ο *αρμονικός μέσος όρος* AM (harmonic mean) είναι το αντίστροφο του μέσου των αντίστροφων τιμών,

$$AM = \frac{n}{\sum \left(\frac{1}{X_i} \right)}$$

και απαντάται συνήθως σε μελέτες ρυθμών μεταβολής.

2.2 ΕΚΤΙΜΗΣΗ ΤΟΥ ΜΕΤΡΟΥ ΤΗΣ ΔΙΑΣΠΟΡΑΣ ΤΩΝ ΑΡΙΘΜΩΝ

Οι παράμετροι που υπολογίζουν τη διασπορά των αριθμών γύρω από το κέντρο της κατανομής τους είναι:
(α) Το *εύρος* E (range) των τιμών το οποίο προκύπτει από τη διαφορά της μέγιστης μείον την ελάχιστη τιμή,

$$E = X_{max} - X_{min} \ .$$

Η παράμετρος αυτή είναι περιορισμένης χρήσης, διότι εκτός από τα όρια στα οποία εντάσσονται οι παρατηρήσεις ενός δείγματος, δεν παρέχει πληροφορίες για τον τρόπο που αυτές διασπείρονται. Το εύρος μιας στήλης υποτιμά συνήθως το πραγματικό εύρος του πληθυσμού, καθότι επηρεάζεται έντονα από τις ακραίες τιμές του δείγματος και θεωρείται επομένως μεροληπτική παράμετρος.
(β) Η *μέση απόλυτη απόκλιση* D (mean absolute deviate), η οποία εκφράζει την απόλυτη διαφορά της διασποράς της στήλης,

$$D = \frac{\sum \left| X_i - \bar{X} \right|}{n} \ ,$$

παράμετρος επίσης περιορισμένης χρήσης.
(γ) Η *διακύμανση* s^2 (variance), η οποία είναι το άθροισμα των τετραγώνων της απόκλισης $X_i - \bar{X}$ κάθε αριθμού X_i από το μέσο όρο $\bar{X}$ του δείγματος, διαιρούμενο με το πλήθος των παρατηρήσεων n-1,

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{(n-1)} \ ,$$

ή για λόγους υπολογιστικής ευχέρειας,

$$s^2 = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1} \ ,$$

όπου ΣX_i^2 είναι το άθροισμα των τετραγώνων όλων των τιμών και $(\Sigma X_i)^2$ το άθροισμα των τιμών υψωμένο στο τετράγωνο. Ο αριθμητής της εξίσωσης καλείται *άθροισμα των τετραγωνισμένων* (SS= sum of squared) και έχει μεγάλη υπολογιστική εφαρμογή σε όλους τους στατιστικούς ελέγχους που εντάσσουν τον όρο της διακύμανσης. Ο όρος n-1 γνωστός και ως *βαθμοί ελευθερίας* (degrees of freedom) υποδηλώνει ότι, με την αφαίρεση της μονάδας από το πλήθος των παρατηρήσεων, η διακύμανση προέρχεται από το δείγμα και όχι από τον πληθυσμό και θεωρείται ότι αποτελεί αμερόληπτη εκτίμηση αυτού. Αν όμως, το πλήθος των παρατηρήσεων είναι μεγαλύτερο από 30 τότε θεωρούμε ότι το δείγμα προσεγγίζει την κατανομή του πληθυσμού και στον παρονομαστή της διακύμανσης χρησιμοποιείται μόνο ο όρος n. Η διακύμανση υποκαθιστά τη μέση απόλυτη απόκλιση σε όλες τις περιπτώσεις, διότι έχει μεγάλη υπολογιστική αξία στον προσδιορισμό των στατιστικών ελέγχων. Όπως είναι φυσικό εκφράζεται σε μονάδες μεγέθους υψωμένες στο τετράγωνο, γεγονός που δυσχεραίνει την εφαρμογή της στην περιγραφή του μέσου όρου. Το πρόβλημα λύνεται αν υπολογιστεί η ρίζα της διακύμανσης και ληφθεί υπόψη μόνο το θετικό πρόσημο αυτής:

$$s=+\sqrt{s^2}=+\sqrt{\frac{\sum X_i^2-\frac{\left(\sum X_i\right)^2}{n}}{n-1}}$$

Η νέα παράμετρος *s* καλείται *τυπική απόκλιση* (standard deviation), έχει τις ίδιες μονάδες με τις μετρήσεις της στήλης και αποτελεί το σπουδαιότερο μέτρο διασποράς των τιμών της στήλης (Παράδειγμα 2.1). Η τυπική απόκλιση επηρεάζεται έντονα από το πλήθος των παρατηρήσεων γι' αυτό ως μέτρο της επίδρασης των παρατηρήσεων στην ακρίβεια του μέσου όρου χρησιμοποιείται το *τυπικό σφάλμα SE* (Standard Error) ή $s_{\bar{X}}$,

$$s_{\bar{X}}=\frac{s}{\sqrt{n}} \quad .$$

Ουσιαστικά, η τυπική απόκλιση εκφράζει το μέτρο της μεταβλητότητας η οποία αναπτύσσεται μεταξύ των τιμών και του μέσου όρου σε ένα δείγμα. Το τυπικό σφάλμα εκφράζει το μέτρο της μεταβλητότητας του μέσου όρου του δείγματος, η οποία προσδιορίζεται ακριβέστερα ποσοτικά με την εκτίμηση των α% ορίων εμπιστοσύνης του μέσου όρου (βλέπε παρακάτω). Το τυπικό σφάλμα είναι γνωστό και ως μέτρο της διασποράς των μέσων όρων πολλών δειγμάτων μεγέθους n του ίδιου πληθυσμού. Στην πράξη όμως, το τυπικό σφάλμα δεν υπολογίζεται από ένα πλήθος επαναληπτικών δειγμάτων, αλλά από ένα μόνο δείγμα και ερμηνεύεται ως η αναμενόμενη τυπική απόκλιση ενός μεγάλου πλήθους δειγμάτων μεγέθους n.

Πόσα όμως δεκαδικά ψηφία απαιτούνται για την πληρέστερη αναγραφή του μέσου όρου και του τυπικού σφάλματος ή και της τυπικής απόκλισης; Ένας εύκολος τρόπος είναι ο έλεγχος της διαίρεσης του τυπικού σφάλματος με τον αριθμό 3. Αν το πρώτο δεκαδικό ψηφίο του πηλίκου της διαίρεσης δεν είναι μηδέν, τότε ο μέσος όρος εκφράζεται με ένα δεκαδικό ψηφίο και το τυπικό σφάλμα με δύο. Αν το πρώτο δεκαδικό ψηφίο του πηλίκου είναι μηδέν, τότε ο μέσος όρος εκφράζεται με δύο δεκαδικά ψηφία και το τυπικό σφάλμα με τρία κοκ. Για παράδειγμα, το τυπικό σφάλμα 0.764 ενός μέσου όρου ίσου με 8.948 δίνει πηλίκο 0.764:3=0.255, οπότε ο μέσος όρος γράφεται ως 8.9 και το τυπικό σφάλμα ως 0.76. Αντίθετα, αν το τυπικό σφάλμα είναι 0.121 τότε το πηλίκο ισούται με 0.121:3=0.040, οπότε ο μέσος όρος γράφεται ως 8.95 και το τυπικό σφάλμα ως 0.121. Τονίζεται ότι στις κλίμακες βαθμίδων, όπως π.χ. για την αξιολόγηση της συνεκτικότητας ενός προϊόντος χρησιμοποιείται κλίμακα ακεραίων αριθμών από 0-10, ένας υπολογιζόμενος μέσος όρος ίσος με 8.4 πολλές φορές δεν σημαίνει τίποτα το ουσιώδες και γι' αυτό καλύτερα είναι να αποφεύγεται η ερμηνεία των δεκαδικών ψηφίων στο μέσο όρο. Αν δηλαδή, ο βαθμός 8 αξιολογεί τη συνεκτικότητα ως 'πολύ καλή' και ο βαθμός 9 ως 'εξαιρετικά καλή', ο μέσος 8.4 δεν έχει ιδιαίτερη αντιπροσωπευτική αξία.

Το τυπικό σφάλμα, η τυπική απόκλιση και η διακύμανση εκφράζουν μεγέθη, η περιγραφή των οποίων εξαρτάται από τα χαρακτηριστικά των στοιχείων, π.χ. το ανθρώπινο βάρος έχει παραμέτρους διασποράς τελείως διαφορετικές σε μέγεθος απ' ότι το βάρος μιας νιφάδας καλαμποκιού ή μιας χημικής συγκέντρωσης στο γάλα (λίπος). Επίσης, για τα ίδια άτομα ή τρόφιμα το βάρος τους μπορεί να αυξάνει ή ελαττώνεται διαχρονικά (π.χ. λόγω αφυδάτωσης), επομένως και οι παράμετροι διασποράς αλλάζουν. Μολαταύτα, υπάρχει ένας τρόπος μέτρησης του μεγέθους μεταβολής της μελετώμενης μεταβλητής που δεν εξαρτάται από την κλίμακα μεγέθους και καλείται *συντελεστής διασποράς* (coefficient of variation) (Παράδειγμα 2.1),

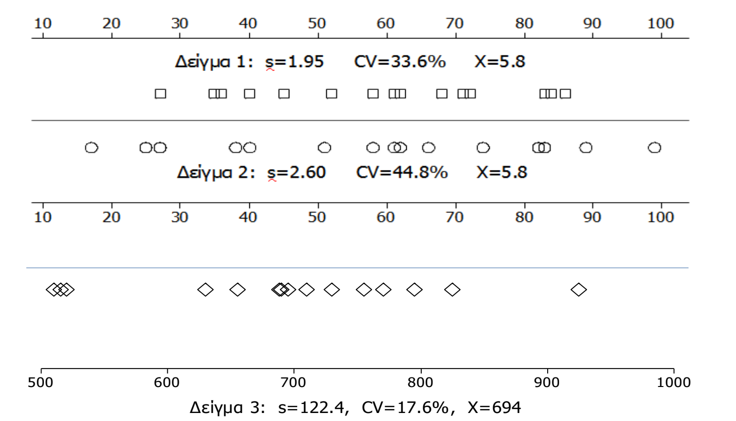
$$CV=\frac{s}{\bar{X}}\cdot 100$$

Ο συντελεστής αυτός μετρά τη μεταβλητότητα του δείγματος ως προς το μέσο όρο και για το λόγο αυτό καλείται και συντελεστής της σχετικής μεταβλητότητας ή σχετικής διασποράς και έχει ισχύ μόνο για ποσοτικές μεταβλητές. Επίσης, δεν επηρεάζεται από τη φύση της κλίμακας μέτρησης των συγκρινόμενων μεταβλητών, η οποία σε κάποιες από αυτές μπορεί να εκφράζεται σε μονάδες όγκου, σε κάποιες άλλες σε μονάδες βάρους, μήκους κτλ. Αναφορικά με το σχήμα 2.3 τα δείγματα 1 και 2 έχουν ίσους μέσους όρους

$(\bar{X}=5.8)$ αλλά άνισες τυπικές αποκλίσεις και κατ` επέκταση διαφορετικούς συντελεστές διασποράς. Μικρότερη τυπική απόκλιση συνεπάγεται μικρότερο συντελεστή διασποράς (δείγμα 1), όταν ο μέσος όρος είναι ίδιος στα συγκρινόμενα δείγματα και για ίσο πάντοτε αριθμό παρατηρήσεων. Το δείγμα 3 παρουσιάζει σαφώς μεγαλύτερο μέσο όρο $(\bar{X}=694)$ και τυπική απόκλιση (s=122.4), όμως παραταύτα δίνει μικρότερο συντελεστή διασποράς απ' ότι τα δείγματα 1 και 2. Αυτό οφείλεται στο στενότερο εύρος της κλίμακας μεγέθους του δείγματος 3 στο οποίο η διασπορά των σημείων κυμαίνεται μεταξύ 500 και 1000, αναλογικά όμως διαιρούμενη δια 100 ισοδυναμεί με την υποπολλαπλάσια κλίμακα 5-10 που επίσης αναλογικά δίνει μέσο όρο 6.94 και τυπική απόκλιση 1.224. Η κλίμακα αυτή αντιστοιχεί στο δεύτερο μισό της κλίμακας 0-10 των δειγμάτων 1 και 2 γι' αυτό και υποχρεωτικά εμφανίζει μικρότερη διασπορά των σημείων.

Πολλές φορές ο μέσος όρος συνοδεύεται με την τυπική απόκλιση για να δηλωθεί έτσι το διάστημα διασποράς των τιμών γύρω από το μέσο όρο. Η αξιοπιστία του διαστήματος αυτού αυξάνει σημαντικά, όταν επιπρόσθετα θέλουμε να δούμε σε τι όρια έκτασης του διαστήματος των τιμών θα βρίσκεται πραγματικά ο μέσος όρος (θα ολισθαίνει δηλαδή ή θα παλινδρομεί) αν δηλώναμε συγκεκριμένη πιθανότητα ακρίβειας, π.χ. 95% ή 99% (ή αντιστρόφως, πιθανότητα σφάλματος 5% ή 1%). Το διάστημα αυτό των τιμών μέσα στο οποίο ολισθαίνει ή εμπεριέχεται ο μέσος όρος λέγεται *a% διάστημα εμπιστοσύνης* (Confidence Interval - C.I.) (συνήθως επιλέγεται το 95%) και με την προϋπόθεση πάντα ότι το δείγμα ακολουθεί την κανονική κατανομή, υπολογίζεται από τη σχέση,

$$a\%\ C.I.:\ \bar{X} \pm t_{a(2),n-1} \cdot s \cdot \sqrt{(1/n)}$$

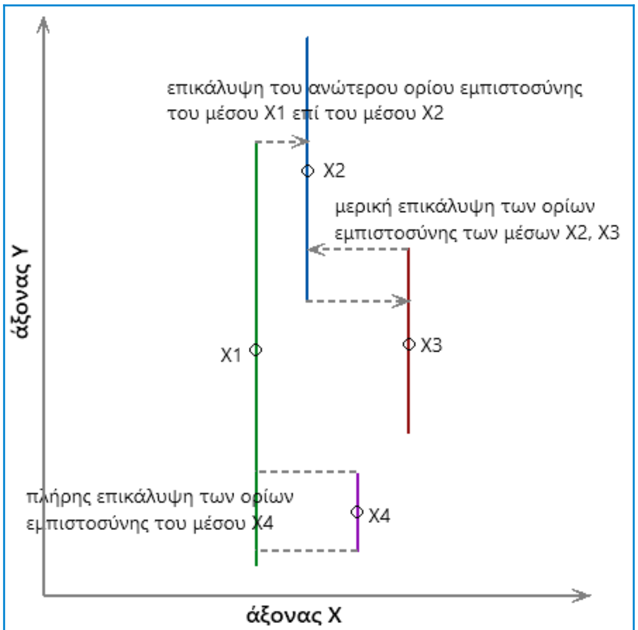


Σχ. 2.3. Σχέση τυπικής απόκλισης και συντελεστή διασποράς σε δείγματα με ίσο αριθμό παρατηρήσεων (n=15) και με διαφορετική κλίμακα μεγέθους στο δείγμα 3.

όπου $t_{α(2),n-1}$ είναι στατιστική τιμή που λαμβάνεται από τον πίνακα Π1 και εξαρτάται από δύο παράγοντες: το ποσοστό ακρίβειας α% που ζητούμε, το οποίο εκφράζεται συνήθως σε πιθανότητα λάθους 0.05 (5%) και αναζητείται στη σειρά α(2) του πίνακα και τους n-1 βαθμούς ελευθερίας του δείγματος. Για παράδειγμα, ένα δείγμα με 10 παρατηρήσεις έχει τιμή $t_{α(2),10-1}$ ίση με 2.262 για 5% πιθανότητα επιτρεπόμενου σφάλματος (3.250 για 1%). Σε περίπτωση που το δείγμα έχει n>30, τότε η τιμή t ισούται με 1.96 για 5% επιτρεπόμενο σφάλμα. Τα άκρα του διαστήματος λέγονται *95% όρια εμπιστοσύνης* (confidence limits) του μέσου όρου μέσα στα οποία ο μέσος όρος θα βρίσκεται οπωσδήποτε (Παράδειγμα 2.1). Τα όρια εμπιστοσύνης διαφοροποιούνται σε *95% όρια πρόβλεψης* (prediction intervals – P.I.) όταν μία νέα ή μελλοντική τιμή λαμβάνεται υπόψη:

$$95\%\ P.I.:\ \bar{X} \pm t_{a(2),n-1} \cdot s \cdot \sqrt{(1+1/n)}\ .$$

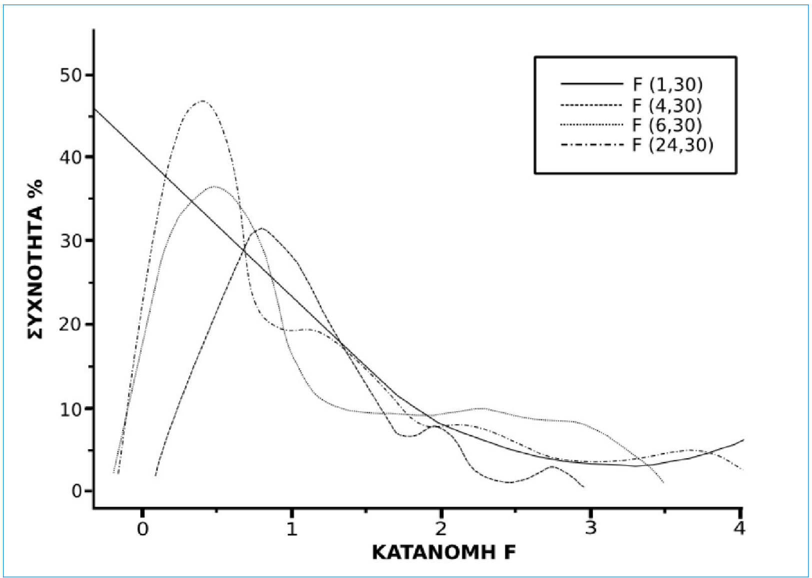
Η σημασία των ορίων εμπιστοσύνης είναι μεγάλη, διότι γρήγορα μπορούμε να συγκρίνουμε ταυτόχρονα δύο ή περισσότερους μέσους όρους της ίδιας μεταβλητής η οποία μεταβάλλεται διαχρονικά ή διαδοχικά (Σχ. 2.4). Αν τα όρια εμπιστοσύνης των μέσων όρων δεν επικαλύπτονται μεταξύ τους, τότε συμπεραίνουμε ότι οι μέσοι όροι διαφέρουν (βλέπε ένθετο σχήμα). Στην περίπτωση που υπάρχει αλληλεπικάλυψη, τότε αν οποιοδήποτε άκρο των ορίων ενός μέσου όρου, ανώτερο ή κατώτερο, υπερκαλύπτει την τιμή ενός δεύτερου μέσου όρου, συμπεραίνουμε ότι οι δύο μέσοι όροι δεν διαφέρουν στατιστικά μεταξύ τους. Η περίπτωση αυτή ισχύει όταν συγκρίνονται οι μέσοι όροι των θηλυκών



7
Έλεγχοι της κατανομής F: ανάλυση διακύμανσης

Η ανάλυση διακύμανσης αποτελεί σημαντικό κεφάλαιο έρευνας πολλών επιστημών σχετιζόμενων με την επιστήμη τροφίμων (βιολογία, μηχανική, χημεία, φυσική), διότι πάνω σε αυτή στηρίζεται πληθώρα πειραματικών σχεδίων. Επιλέγεται κυρίως για τη σύγκριση τριών ή περισσοτέρων μέσων όρων δειγμάτων τα οποία διαφέρουν μεταξύ τους ως προς τη μεταχείριση ενός ή και περισσοτέρων παραγόντων. Για παράδειγμα, μελετούμε την επίδραση της θερμοκρασίας μόνη της ή και σε συνδυασμό με την υγρασία ή και το pH, στο ρυθμό ανάπτυξης των μικροοργανισμών μιας καλλιέργειας ενός τροφίμου (γιαούρτι). Έτσι, οι μετρήσεις της μεταβλητής, που εδώ είναι ο ρυθμός ανάπτυξης, χωρίζονται σε διάφορες ομάδες ανάλογα με τον αριθμό των παραγόντων που υιοθετούνται και τα επίπεδα δράσης αυτών π.χ. υγρασία: 30%, 50% και 70%, θερμοκρασία: 15^o και 30^{o}C, pH: 5.0 και 4.0. Η ανάλυση διακύμανσης ή ο έλεγχος της κατανομής F αναφέρεται διεθνώς ως ANOVA (Analysis Of Variance) και μας πληροφορεί αν οι εξεταζόμενοι μέσοι όροι των δειγμάτων της μελετώμενης μεταβλητής είναι ίσοι μεταξύ τους (έχουν κοινό μέσο όρο και άρα προέρχονται από τον ίδιο πληθυσμό) ή διαφέρουν (προέρχονται από διαφορετικούς πληθυσμούς). Η ANOVA μπορεί να εφαρμοστεί και σε σύγκριση των μέσων όρων δύο ανεξαρτήτων δειγμάτων και δίνει εφάμιλλα αποτελέσματα με το δίπλευρο έλεγχο t των ανεξαρτήτων δειγμάτων.

Η ανάλυση της διακύμανσης F βασίζεται υπολογιστικά στην κατανομή F των μετρήσεων η οποία είναι θετικά ασύμμετρη και η διαμόρφωσή της γίνεται αντιληπτή με τον εξής τρόπο: Αν υποθέσουμε ότι από ένα πληθυσμό παίρνουμε δύο δείγματα με n_1 και n_2 πλήθος στοιχείων το καθένα, τότε αυτά θα έχουν s_1^2 και s_2^2 διακυμάνσεις. Αφού και τα δύο δείγματα είναι εκτιμήσεις του ίδιου πληθυσμού, έπεται ότι οι διακυμάνσεις



Σχ. 7.1. Καμπύλες της κατανομής F σε διάφορους βαθμούς ελευθερίας του αριθμητή της σχέσης $F=s_1^2/s_2^2$.

τους θα πρέπει να είναι ίσες ή αλλιώς,το πηλίκο F της διαίρεσής τους θα ισούται με 1: $F=s_1^2/s_2^2=1$. Οι τιμές των διακυμάνσεων εξαρτώνται από το πλήθος των παρατηρήσεων και επομένως το πηλίκο μπορεί εύκολα να αποκλίνει από τη μονάδα όταν τα δειγματοληπτικά μεγέθη είναι μικρά. Όσο περισσότερο η τιμή F απομακρύνεται από τη μονάδα τόσο περισσότερο αυξάνεται η πιθανότητα οι συγκρινόμενες διακυμάνσεις των δύο πληθυσμών να διαφέρουν μεταξύ τους. Όταν όμως το δεύτερο δειγματοληπτικό μέγεθος n_2 είναι πολύ μεγάλο, τότε η τιμή F προσεγγίζει αυτή του πηλίκου $(n_2-1)/(n_2-3)\approx 1$. Αν ληφθούν πολλά δείγματα από τον ίδιο πληθυσμό, τότε σχηματίζονται διάφορες καμπύλες της κατανομής F, οι μορφές των οποίων καθορίζονται πάντα από ένα ζεύγος τιμών κάθε φορά των βαθμών ελευθερίας n_1-1 και n_2-1. Έτσι, για οποιοδήποτε συνδυασμό των βαθμών ελευθερίας δύο δειγμάτων που κυμαίνονται από 1 μέχρι $+\infty$ αντιστοιχεί και μία ξεχωριστή κατανομή F (Σχ. 7.1). Αν το ένα από τα δύο δειγματοληπτικά μεγέθη είναι μικρό, π.χ. n=1, η καμπύλη παρουσιάζει σχήμα αντίστροφο J, ενώ γίνεται πιο θολωτή όταν είτε το ένα είτε και τα δύο μεγέθη αυξάνουν (Σχ. 7.1).

Παλαιότερα επικρατούσε η αντίληψη ότι αντί της υπόθεσης σύγκρισης πολλών δειγμάτων π.χ. αναφορικά με τρία εξεταζόμενα δείγματα με μηδενική υπόθεση σύγκρισης H_O: $μ_1=μ_2=μ_3$, θα μπορούσε να χρησιμοποιηθεί εξίσου αποτελεσματικά και ο έλεγχος t της υπόθεσης σύγκρισης δύο ανεξαρτήτων δειγμάτων. Στην περίπτωση αυτή θα απαιτούνταν τρεις διαδοχικές συγκρίσεις ανά δύο των μέσων όρων με μηδενικές υποθέσεις, H_O:$μ_1=μ_2$, H_O:$μ_1=μ_3$ και H_O:$μ_2=μ_3$. Τέτοιες υποθέσεις όμως είναι ανεφάρμοστες και πολύ περισσότερο άκυρες διότι οδηγούν σε λανθασμένα συμπεράσματα. Αυτό προκύπτει από το γεγονός ότι, αν εφαρμόσουμε την ανάλυση διακύμανσης και θεωρήσουμε ως επίπεδο σημαντικότητας λάθους το 5% για τη σύγκριση των τριών μέσων όρων, τότε το αντίστοιχο ποσοστό να συμβεί λάθος τύπου I (5% δηλαδή) που αναλογεί αν εφαρμόσουμε επαναληπτικά τον έλεγχο t των ανεξαρτήτων δειγμάτων με δύο μέσους όρους, θα ανέλθει στο 13%. Όσο περισσότεροι μέσοι όροι συμμετέχουν σε μία τέτοια ανορθόδοξη ανάλυση τόσο περισσότερο αυξάνει το ποσοστό λάθους και πράγματι η απόπειρα εφαρμογής του ελέγχου t για σύγκριση 5 δειγμάτων θα αυξήσει το ποσοστό λάθους σε 23% και για σύγκριση 10 δειγμάτων σε 63%.

7.1 ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ ΕΝΟΣ ΠΑΡΑΓΟΝΤΑ Ή ΜΙΑΣ ΚΑΤΕΥΘΥΝΣΗΣ (one-way ANOVA)

Η ορολογία της ANOVA περιλαμβάνει τη σύγκριση δύο ή περισσοτέρων μέσων όρων με μηδενική υπόθεση, H_O: $μ_1=μ_2=μ_3=....=μ_k$, όπου k είναι το πλήθος των συγκρινόμενων μέσων όρων ομάδων και εναλλακτική, H_A: *οι μέσοι όροι δεν είναι όλοι ίσοι*. Το σύμβολο ≠ δεν τίθεται στην εναλλακτική υπόθεση μεταξύ των μέσων όρων, καθότι αρκεί απλούστατα και μόνο ένας μέσος όρος να διαφέρει για να καταρριφθεί η μηδενική υπόθεση.

Ως κλασικά παραδείγματα της ανάλυσης της διακύμανσης ενός παράγοντα αναφέρονται μερικές συχνές περιπτώσεις χρήσης αυτής: η εξέταση της επίδρασης τεσσάρων (ή και παραπάνω) γλυκαντικών υλών στην υφή ενός προϊόντος (μηχανικά χαρακτηριστικά), η εξέταση της δράσης της γεύσης κατά τον ποιοτικό έλεγχο 5 τύπων γιαούρτης ή τυριού ή κρασιών, η εξέταση της εξάπλωσης της μούχλας που ενοφθαλμίστηκε σε τρεις διαφορετικές ποσότητες σε μεγάλο αριθμό γεώμηλων ή της εξάπλωσης των μικροβίων είδους Α που ενοφθαλμίστηκαν σε 3 διαφορετικά θρεπτικά υποστρώματα, η εξέταση της επίδρασης 5 διαφορετικών εκκινητών (starters) μικροοργανισμών στην εξέλιξη του pH γιαούρτης κτλ. Σε όλες αυτές τις περιπτώσεις εξετάζουμε πιθανές μεταβολές των μέσων όρων των υποβαλλόμενων στη μεταχείριση αντικειμένων ή ατόμων ως προς ένα μελετώμενο *παράγοντα*, ή ως προς μία *κατεύθυνση* μεταβολής ή τέλος ως προς ένα *κριτήριο* που αυτό είναι, σύμφωνα με τα αναφερθέντα παραδείγματα, η διαφορετική γλυκαντική ύλη, οι τύποι γιαούρτης, τυριών ή κρασιών, το διαφορετικό θρεπτικό υπόστρωμα και τέλος οι διαφορετικοί εκκινητές. Οι κατηγορίες ή οι τύποι ή τα μερίσματα του κάθε παράγοντα αποτελούν τα *επίπεδα* αυτού, ενώ τα άτομα που συμμετέχουν εκούσια ή ακούσια στον πειραματισμό θα πρέπει να επιλέγονται τυχαία από το συνολικό πληθυσμό τους και να κατανέμονται επίσης τυχαία στα διάφορα επίπεδα του παράγοντα και μάλιστα όσο το δυνατόν σε ισάριθμα μεγέθη. Αυτός ο τρόπος καταμερισμού αποτελεί αναπόσπαστο κομμάτι της ANOVA και είναι γνωστός με το όνομα *πλήρως τυχαιοποιημένο πειραματικό σχέδιο*. Τα επίπεδα ενός μελετώμενου παράγοντα συχνά αναφέρονται και ως *μεταχειρίσεις* ή *δοκιμασίες* στις οποίες υποβάλλουμε τα άτομα ή τα αντικείμενα. Έτσι, ο όρος τρεις μεταχειρίσεις σημαίνει τρία επίπεδα ενός παράγοντα.

Τα επίπεδα ενός παράγοντα μπορεί να έχουν ποσοτική σχέση μεταξύ τους ή καθαρά ποιοτική. Ποσοτική σχέση έχουν τα επίπεδα συγκέντρωσης μιας συγκεκριμένης γλυκαντικής ουσίας η οποία προστίθεται σε αυξημένες ποσότητες στις διάφορες μεταχειρίσεις δειγμάτων και μετρείται η οργανοληπτική δράση αυτών από διαφορετικούς δοκιμαστές. Έτσι, καθίσταται δυνατή η γραφική απεικόνιση της μεταβολής των μέσων όρων της οργανοληπτικής γλυκύτητας (άξονας Y) με τα επίπεδα της αυξητικής συγκέντρωσης της γλυκαντικής ύλης (άξονας X). Ποιοτική σχέση έχουν τα επίπεδα διαφορετικών γλυκαντικών υλών οι οποίες προστίθενται στις διάφορες μεταχειρίσεις δειγμάτων και εξετάζουμε αν παρουσιάζουν διαφορετική οργανοληπτική δράση στους δοκιμαστές.

Επισημαίνεται ότι στον οργανοληπτικό έλεγχο ενός προϊόντος, οι συμμετέχοντες δοκιμαστές λειτουργούν ως βαθμολογητές της έντασης της οργανοληπτικής μεταβλητής και ουσιαστικά επιχειρούν τις μετρήσεις της μεταβλητής. Συνεπώς, για να ισχύει το πλήρως τυχαιοποιημένο σχέδιο στην ANOVA, θα πρέπει οι ίδιοι δοκιμαστές να μην αξιολογούν ταυτόχρονα όλα τα επίπεδα του μελετώμενου παράγοντα, παρά μόνο ένα κάθε φορά. Είτε επίσης, διαφορετικοί δοκιμαστές να αξιολογούν όλα τα επίπεδα του παράγοντα, ένα όμως επίπεδο για τον καθένα. Αντίθετα, αν κάθε δοκιμαστής αξιολογούσε όλα τα επίπεδα του προϊόντος ταυτόχρονα, τότε η ανάλυση της ANOVA αφορά ένα τυχαιοποιημένο σχέδιο κατά ομάδες με επαναληπτικές μετρήσεις, σχέδιο που αποτελεί προέκταση των εξαρτημένων και αναφέρεται εμπεριστατωμένα στην ενότητα 7.9.

Στο παράδειγμα 7.1 εξετάζεται η επίδραση τεσσάρων διαφορετικών χρόνων συντήρησης στη νωπότητα ψαριών που αμέσως μετά την αλίευση τους τοποθετήθηκαν σε θρύμματα πάγου. Η νωπότητα είναι οργανοληπτικό χαρακτηριστικό και η έντασή της ελέγχεται με τη βοήθεια δοκιμαστών που ο καθένας αξιολογεί ένα ψάρι. Κάθε στοιχείο συμβολίζεται με X_{ij}, όπου i σημαίνει τις ομάδες συμμετοχής των ψαριών (ή των αξιολογητών) ή αλλιώς τις στήλες και j τις σειρές μέσα στις ομάδες. Το στοιχείο X_{34} αντιστοιχεί στην ομάδα 3 (6η ημέρα) και σειρά 4, είναι δηλαδή η τιμή 162. Επιπλέον, ο συμβολισμός $\bar{X}_i$ υπονοεί το μέσο όρο της i ομάδας και $\bar{X}$ τον ολικό μέσο όρο όλων των στοιχείων των k ομάδων του πειράματος. Το πλήθος των παρατηρήσεων ανά ομάδα, που λέγονται και *επαναλήψεις,* συμβολίζεται με n_i και το σύνολο όλων των παρατηρήσεων με N. Προσέχοντας τα στοιχεία του παραδείγματος 7.1, εύκολα προκύπτει ότι υπάρχει διαφορετικότητα μεταξύ των στοιχείων που συνοψίζεται:

(α) Στη μεταβλητότητα μεταξύ του κάθε στοιχείου κάθε δείγματος και του μέσου όρου του δείγματος.

(β) Στη μεταβλητότητα μεταξύ του μέσου όρου κάθε δείγματος και του ολικού μέσου όρου του συνόλου των τιμών.

(γ) Στη μεταβλητότητα μεταξύ του κάθε στοιχείου κάθε δείγματος και του ολικού μέσου όρου του συνόλου των τιμών.

Για παράδειγμα η μεταβλητότητα του στοιχείου X_{34}=162 καταμερίζεται με βάση τις τρεις περιπτώσεις ως εξής:

$X_{34} - \bar{X}_3$ =162-169.8=-7.8

$\bar{X}_3 - \bar{X}$ =169.8-167.1=+2.7

$X_{34} - \bar{X}$ =162-167.1=-5.1

Το άθροισμα των δύο πρώτων αποκλίσεων ισούται με την τρίτη απόκλιση και γενικεύοντας για κάθε στοιχείο θα έχουμε: $(X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X}) = (X_{ij} - \bar{X})$. Αν οι δύο ισότητες των αποκλίσεων υψωθούν στο τετράγωνο και ληφθούν υπόψη οι αποκλίσεις όλων των στοιχείων του πειράματος, τότε θα προκύψει με μικρή τροποποίηση:

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2 + n\cdot\sum_{j=1}^{n}(\bar{X}_i - \bar{X})^2 + 2\cdot\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)\cdot(\bar{X}_i - \bar{X}) = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X})^2 \quad .$$

Το γινόμενο της εξίσωσης $2\cdot\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)\cdot(\bar{X}_i - \bar{X})$ διαγράφεται από την εξίσωση, γιατί η ποσότητα αυτού $\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)$ ισούται με το μηδέν. Οι όροι αυτοί αποτελούν τα αθροίσματα SS (sum of squared) των τετραγωνισμένων αριθμών τα οποία και αναφέρονται για λόγους συντομογραφίας ως T(otal)SS= ολικό άθροισμα SS, G(roup)SS= άθροισμα SS *μεταξύ των ομάδων* και E(rror)SS= άθροισμα SS ή σφάλμα SS *μέσα στις ομάδες*.

Μία σημαντική ιδιότητα του αθροίσματος των τετραγωνισμένων αριθμών SS είναι η προσθετική, δηλαδή το άθροισμα ESS+GSS δίνει TSS, όπως αντίστοιχα και οι βαθμοί ελευθερίας. Τα αθροίσματα SS για λόγους υπολογιστικής ευχέρειας επαναπροσδιορίζονται ως εξής:

$$TSS = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij}^2 - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}$$, με βαθμούς ελευθερίας N-1

$$GSS = n\cdot\sum_{j=1}^{n}(\bar{X}_i - \bar{X})^2 = \sum_{i=1}^{k}\frac{(\sum_{j=1}^{n}X_{ij})^2}{n_i} - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}$$, με βαθμούς ελευθερίας k-1

$$ESS = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2 = TSS - GSS$$, με βαθμούς ελευθερίας N-k.

Η ποσότητα $\frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}$ που υπάρχει και στις δύο εξισώσεις καλείται *συντελεστής διόρθωσης **C.*** Τα παραπάνω αθροίσματα διαιρούμενα με τους αντίστοιχους βαθμούς ελευθερίας, αποτελούν τις επιμέρους διακυμάνσεις των στοιχείων σχετικά με το σύνολό τους, μεταξύ των ομάδων και μέσα στις ομάδες. Οι

διακυμάνσεις αυτές είναι γνωστότερες με την ονομασία μέσα αθροίσματα των τετραγωνισμένων (MS):
GMS=GSS/(k-1), μέσο άθροισμα ή μέσο σφάλμα των τετραγωνισμένων μεταξύ των ομάδων και
EMS=ESS/(N-k), μέσο άθροισμα ή μέσο σφάλμα των τετραγωνισμένων μέσα στις ομάδες.
Στην ANOVA απαραίτητη προϋπόθεση είναι οι διακυμάνσεις όλων των δειγμάτων να είναι ίσες $\sigma_1^2=\sigma_2^2=\sigma_3^2=....\sigma_k^2$. Έτσι, υπολογίζεται μία κοινή διακύμανση για όλες τις ομάδες (pooled variance) που καλείται μέσο σφάλμα του αθροίσματος των τετραγωνισμένων και αντιστοιχεί στον όρο EMS που αναφέρθηκε. Οι υπολογισμοί συνοψίζονται στον πίνακα 7.1.

Πίνακας 7.1. Υπολογισμός των παραμέτρων της ανάλυσης διακύμανσης ενός παράγοντα.

Τύποι διακύμανσης	Βαθμοί ελευθερίας	Αθροίσματα των τετραγωνισμένων SS	Μέσα αθροίσματα των τετραγωνισμένων MS	F τιμή
Ολική	N-1	$TSS=\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij}^2-C$		
Μεταξύ των ομάδων	k-1	$GSS=\sum_{i=1}^{k}\frac{(\sum_{j=1}^{n}X_{ij})^2}{n_i}-C$	$GMS=\frac{GSS}{k-1}$	
Μέσα στις ομάδες	N-k	$ESS=TSS-GSS$	$EMS=\frac{ESS}{N-k}$	$\frac{GMS}{EMS}$

Για να επαληθεύεται η μηδενική υπόθεση αντί της εναλλακτικής, θα πρέπει μετά την εφαρμογή του ελέγχου της διακύμανσης, το πηλίκο F της διαίρεσης της διακύμανσης (μεταβλητότητας) μεταξύ των ομάδων διά της διακύμανσης μέσα στις ομάδες να ισούται με 1 (F=GMS/EMS), δηλαδή να αποτελούν και οι δύο διακυμάνσεις κοινή μέτρηση της διακύμανσης των συγκρινόμενων k ομάδων. Αντίθετα, αν το σφάλμα μεταξύ των ομάδων είναι μεγαλύτερο του σφάλματος μέσα στις ομάδες, τότε η τιμή F θα είναι μεγαλύτερη της μονάδας και τότε μόνο οι μέσοι όροι των ομάδων θα είναι άνισοι μεταξύ τους. Δηλαδή στην ανάλυση διακύμανσης ισχύει μόνο ο μονόπλευρος έλεγχος αυτής, με βαθμούς ελευθερίας το ζεύγος τιμών που αντιστοιχεί σε k-1 βαθμούς στον αριθμητή της διαίρεσης και σε N-k βαθμούς στον παρονομαστή. Η τιμή F του ελέγχου συγκρίνεται με την οριακή τιμή που είναι $F_{\alpha(1)(k-1)(N-k)}$. Οι βαθμοί ελευθερίας του αριθμητή k-1 διατάσσονται σε οριζόντια σειρά στον πίνακα Π3 και οι βαθμοί ελευθερίας του παρονομαστή N-k στην πρώτη κάθετη αριστερή στήλη του πίνακα. Αν $F \geq F_{ορ}$ ισχύει η εναλλακτική υπόθεση και αν $F<F_{ορ}$ η μηδενική υπόθεση. Θα πρέπει να σημειωθεί ότι, όταν η τιμή F υποδεικνύει την επαλήθευση της εναλλακτικής υπόθεσης, αυτό σημαίνει ότι υπάρχει διαφορά μεγέθους μεταξύ των μέσων όρων που εξετάζονται και τίποτα παραπάνω. Η ANOVA μας πληροφορεί απλά ότι κάποιοι μέσοι όροι διαφέρουν, **ΔΕΝ** μας πληροφορεί όμως πού εντοπίζονται οι διαφορές αυτές, δηλαδή ποιοι μέσοι όροι διαφέρουν από ποιους και με ποια σειρά. Οι υποθέσεις αυτές τεκμηριώνονται με τη χρήση ειδικών ελέγχων που θα περιγραφούν στην ενότητα 7.2.

Παράδειγμα 7.1. Ανάλυση της διακύμανσης ενός παράγοντα.
Για τη μελέτη της επίδρασης του χρόνου συντήρησης στη νωπότητα των αλιευμάτων χρησιμοποιήθηκαν λιθρίνια που μετά την αλίευση τους ψύχθηκαν με θρύμματα πάγου. Κάθε 3 μέρες ψύξης επιλέγονται 6 λιθρίνια και υποβάλλονται σε οργανοληπτικό έλεγχο της νωπότητας τους από 6 δοκιμαστές με βάση την οσμή που αναδύουν, το βαθμό νεκρικής ακαμψίας και τον αποχρωματισμό των βραγχιακών επικαλυμμάτων. Η ένταση της νωπότητας βαθμολογήθηκε με τη βοήθεια αδιαβάθμητης κλίμακας αυξανόμενης παλαιότητας των ψαριών από 0-250mm. Η νωπότητα εξετάστηκε και σε ψάρια τη στιγμή της αλίευσης τους (χρόνος συντήρησης 0), έτσι χρησιμοποιήθηκαν συνολικά 4 επίπεδα του παράγοντα χρόνου: 0, 3, 6 και 9 ημέρες. Το πρώτο και βασικό ερώτημα που τίθεται είναι, αν τα διαφορετικά επίπεδα του χρόνου συντήρησης προκάλεσαν μεταβολή στη νωπότητα του αλιεύματος.
H_O:μ1=μ2=μ3=μ4
H_A: οι μέσοι όροι δεν είναι όλοι ίσοι
α=0.05

Αξιολόγηση της νωπότητας αλιευμάτων

	0 μέρα	3η μέρα	6η μέρα	9η μέρα	
1	138	149	160	212	
2	156	172	182	196	
3	135	164	174	184	
4	140	160	162	179	
5	152	156	176	210	
6	153	150	165	186	
ni	6	6	6	6	$N=\Sigma n_i=24$
$\bar{X}_i$	145.7	158.5	169.8	194.5	
ΣX_i^2	127718	151117	173445	227953	$\bar{X}$ =167.1
ΣX_i	874	951	1019	1167	

Οι υπολογισμοί των αθροισμάτων τετραγωνισμένων έχουν ως ακολούθως:

$$C=\frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}=\frac{(138+156+135+........+210+186)^2}{24}=670338$$

$$TSS=\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij}^2-\frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}=138^2+156^2+135^2+.....+210^2+186^2-670338=9895$$ N-1=24-1=23

$$GSS=\sum_{i=1}^{k}\frac{(\sum_{j=1}^{n}X_{ij})^2}{n_i}-\frac{(\sum_{i=1}^{k}\sum_{j=1}^{n}X_{ij})^2}{N}=\frac{874^2}{6}+\frac{951^2}{6}+\frac{1019^2}{6}+\frac{1167^2}{6}-670338=7749$$ k-1=4-1=3

ESS=TSS-GSS=2146 N-k=24-4=20

Τα μέσα αθροίσματα των τετραγωνισμένων υπολογίζονται ως εξής:

GMS=GSS/k-1=7749/3=2583

EMS=ESS/N-k=2146/20=107 και $F=\frac{GMS}{EMS}=\frac{2583}{107}=24.1$

Όλες οι πράξεις συνοψίζονται στον παρακάτω πίνακα:

Ανάλυση διακύμανσης στη νωπότητα του προϊόντος

Πηγές μεταβλητότητας	DF	SS	MS	F	p
Ολική	23	9895			
Μεταξύ των ομάδων	3	7749	2583	24.1	0.000
Μέσα στις ομάδες	20	2145	107		

Μέσοι όροι και 95% όρια εμπιστοσύνης υπολογισμένα με βάση την κοινή διακύμανση EMS του ελέγχου

```
Επίπεδα    N    Μέσοι      SD     --+---------+---------+---------+----
   1       6    145.7     9.00          (---*----)
   2       6    158.5     8.76                (---*---)
   3       6    169.8     8.77    (----*---)
   4       6    194.5    13.94                            (---*----)
                                  --+---------+---------+---------+----
                                   140       160       180       200
```

Συμπέρασμα: Η υπολογισθείσα τιμή F της ANOVA είναι σαφώς μεγαλύτερη της οριακής $F_{0.05(1)3,20}$=3.10 και επομένως ισχύει η εναλλακτική υπόθεση, δηλαδή οι μέσες τιμές της αξιολόγησης της νωπότητας δεν είναι όλες ίσες μεταξύ των διαφορετικών χρόνων συντήρησης. Το συμπέρασμα μπορεί επίσης να διατυπωθεί και ότι υπάρχει ισχυρή επίδραση των διαφορετικών ημερών ψύξης στη νωπότητα των ψαριών. Πώς επιτελείται όμως αυτή η επίδραση και με ποια ιεραρχία μεταξύ των χρόνων συντήρησης, η ANOVA δεν μπορεί να εξηγήσει και θα πρέπει να καταφύγουμε σε κάποιους από τους ελέγχους πολλαπλών συγκρίσεων των μέσων όρων.

Η τιμή F ουσιαστικά μας πληροφορεί για την ύπαρξη σημαντικότητας ($F \geq F_{op}$) ή μη ($F < F_{op}$) της δράσης της μελετώμενης μεταβλητής η οποία είναι η βαθμολόγηση της νωπότητας των δοκιμαστών, όπως αυτή αποδίδεται μεταξύ των τεσσάρων μεταχειρίσεων. Πέρα όμως από την ύπαρξη σημαντικότητας της νωπότητας μεταξύ των τεσσάρων ομάδων δεν αναφέρει τίποτα για την ένταση αυτής της σχέσης μεταξύ των ομάδων. Αυτή προκύπτει από την τιμή του στατιστικού κριτηρίου η^2 που ισούται με:

$$\eta^2 = \frac{GSS}{TSS} \quad \text{ή} \quad \eta^2 = 1 - \frac{ESS}{TSS} \ ,$$

και λαμβάνει τιμές από 0 μέχρι 1. Τιμές κοντά στο 1 (π.χ. 0.892) σημαίνουν ισχυρή γραμμική σχέση μεταξύ των επιπέδων του εξεταζόμενου παράγοντα ως προς τη μελετώμενη μεταβλητή. Το υπόλοιπο $1\text{-}\eta^2$ (0.108) που δεν εξηγείται, οφείλεται σε άλλους παράγοντες άγνωστης προέλευσης. Το κριτήριο η^2 παρουσιάζει ερμηνεία ανάλογη με αυτήν του συντελεστή προσδιορισμού R^2 της γραμμικής παλινδρόμησης (βλέπε ενότητα 8.1) και ισχύει αποκλειστικά και μόνο όταν τα επίπεδα του παράγοντα έχουν ποσοτική σχέση.

Η ANOVA μπορεί να εφαρμοστεί και σε περιπτώσεις σύγκρισης δύο μόνο μέσων όρων αντί του ελέγχου t της υπόθεσης των ανεξαρτήτων δειγμάτων και μόνο για δίπλευρο έλεγχο, με μηδενική υπόθεση δηλαδή, H_O:μ_1=μ_2 και H_A:$\mu_1 \neq \mu_2$. Στην περίπτωση αυτή το σφάλμα μέσα στις ομάδες είναι ίδιο με την κοινή διακύμανση s_p^2 του ελέγχου t και με βαθμούς ελευθερίας k-1=1 στον αριθμητή και N-k=N-2 στον παρονομαστή. Η οριακή τιμή F σχετίζεται με την οριακή t σύμφωνα με την αντιστοιχία, $F_{(\alpha(1),(N-2)}=(t_{\alpha(2),(N-2)})^2$.

Η ανάλυση της διακύμανσης ενός παράγοντα έχει δύο μοντέλα (πρότυπα) που δεν διαφέρουν καθόλου στους υπολογισμούς, αλλάζει όμως η διατύπωση της μηδενικής υπόθεσης:

✰ Το *πρότυπο I*, γνωστό και ως πρότυπο *καθορισμένης δράσης* (fixed effects), χαρακτηρίζεται από την ειδική επιλογή των επιπέδων ενός παράγοντα. Στο παράδειγμα 7.1 τα 3 διαφορετικά χρονικά επίπεδα (ημέρες) επιλέχτηκαν όχι τυχαία, αλλά απλώς όπως εμείς επιδιώξαμε. Εξαίρεση αποτελεί η περίπτωση της περιόδου 0 ημερών η οποία μπορεί να εκληφθεί ως περίοδος *μάρτυρας*. Με τον όρο μάρτυρα νοείται ένα οποιοδήποτε επίπεδο του εξεταζόμενου παράγοντα, στο οποίο τα άτομα (οργανισμοί) ή αντικείμενα (κονσέρβα με υγρό ή στερεό περιεχόμενο) που το απαρτίζουν δεν υποβάλλονται στην παραμικρή μεταχείριση. Στο πρότυπο I η εναλλακτική υπόθεση διατυπώνεται όπως και στο παράδειγμα 7.1. Παραδεχόμαστε δηλαδή ότι ενδεχόμενες διαφορές μεταξύ των μέσων όρων των επιπέδων της διαχρονικής συντήρησης των ψαριών οφείλονται στην προκαθορισμένη δράση των μεταχειρίσεων (διαφορετικών περιόδων συντήρησης) που επιλέχτηκαν σκόπιμα από τον ερευνητή. Όταν, μετά την εφαρμογή της ANOVA βρεθεί ότι ισχύει η μηδενική υπόθεση, ότι δηλαδή δεν υπάρχει στατιστικά σημαντική διαφορά στην επίδραση των διαφορετικών ημερών συντήρησης πάνω στον πληθυσμό των ψαριών, αυτό δεν σημαίνει απαραίτητα ότι ο εξεταζόμενος παράγοντας, που είναι ο χρόνος συντήρησης, δεν προκαλεί κανένα αποτέλεσμα στην επιβράδυνση της αλλοίωσης των ψαριών. Απλώς, τα συγκεκριμένα χρονικά επίπεδα συντήρησης που επιλέχτηκαν μπορεί να μην επηρεάζουν διαφοροποιημένα την νωπότητα των ψαριών. Η σταδιακή μείωση της νωπότητας όμως των ψαριών, είναι σαφώς μη παλίνδρομη βιολογική δράση και οφείλεται στην αυτόλυση των κυττάρων με ταυτόχρονη ανάπτυξη της μικροβιακής χλωρίδας. Η χρήση του πάγου παρατείνει μόνο τη νωπότητα των ψαριών.

✰ Αν η εκλογή των επιπέδων του παράγοντα ήταν τυχαία, τότε η ANOVA χαρακτηρίζεται ως *πρότυπο II* ή *πρότυπο τυχαίας δράσης* (random effects). Για παράδειγμα, έστω ότι επιλέχτηκαν τυχαία 6 φοιτητές για να πάρουν 5 μετρήσεις pH ο καθένας σε μία ποσότητα γάλακτος. Στην περίπτωση αυτή, μας ενδιαφέρει αν υπάρχει διαφορά στον προσδιορισμό των μετρήσεων που έκανε ο κάθε φοιτητής και η διατύπωση της μηδενικής υπόθεσης είναι H_O: οι τιμές του pH δεν διαφέρουν μεταξύ των φοιτητών. Ουσιαστικά δηλαδή, ελέγχουμε αν η μεταβλητότητα των τιμών μεταξύ των ομάδων των μετρήσεων που προέρχεται από διαφορετικό φοιτητή είναι μεγαλύτερη από τη μεταβλητότητα των τιμών του pH του γάλακτος ανά ομάδα. Έτσι, το ενδιαφέρον επικεντρώνεται στο γενικό πρόβλημα του πειραματικού σχεδίου στο οποίο ερευνώνται τυχόν διαφορές που αναπτύσσονται στη μεταβλητότητα του τρόπου χειρισμού του οργάνου μέτρησης από τους πειραματιστές και επίσης στη μεταβλητότητα που μπορεί να εμφανίσει η εξεταζόμενη ποσότητα του γάλακτος ως προς το pH. Η μόνη ουσιαστική διαφοροποίηση από το πρότυπο I της ANOVA είναι το γεγονός ότι, σε περίπτωση ισχύος της εναλλακτικής υπόθεσης (ότι δηλαδή υπάρχει διαφορά εκτίμησης του pH οφειλόμενη στο χειρισμό από διαφορετικούς φοιτητές), αρκούμαστε μόνο στη διατύπωση αυτής και δεν προχωρούμε στην εύρεση συγκεκριμένων διαφορών μεταξύ των μέσων όρων.

Οι περιορισμοί (προϋποθέσεις) που τίθενται πριν την εφαρμογή της ANOVA είναι δύο:

(α) Καθεμία ομάδα (δείγμα) θα πρέπει να προέρχεται από κανονικό πληθυσμό (κανονική κατανομή). Κατανομές που παρουσιάζουν αρκετές τιμές και στα δύο άκρα εντείνουν το πρόβλημα της έλλειψης κανονικότητας πολύ περισσότερο απ' ότι άλλες που είναι απλώς ασύμμετρες θετικά ή αρνητικά. Στις ασύμμετρες κατανομές η τιμή F της ANOVA επηρεάζεται λιγότερο έντονα με παρόντες επιλέξιμους παράγοντες και περισσότερο όταν οι παράγοντες είναι τυχαίοι. Η εγκυρότητα όμως της ANOVA επηρεάζεται ελάχιστα από μη κανονικούς πληθυσμούς, ακόμα και αν η μη κανονικότητα είναι ισχυρή, καθόσον έχει διαπιστωθεί στην πράξη ότι η έλλειψη κανονικότητας προκαλεί μικρή μείωση του πραγματικού επιπέδου σημαντικότητας και της ισχύος του ελέγχου της ANOVA. Έτσι, στην ANOVA η ύπαρξη μη κανονικότητας στα στοιχεία δεν αποτελεί ισχυρό περιορισμό, όπως συμβαίνει σε άλλους ελέγχους συγκρίσεων.

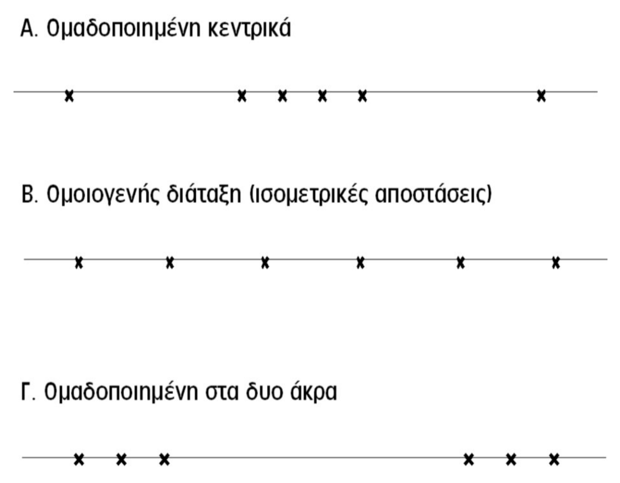
Η κανονικότητα για κάθε ομάδα εξετάζεται με τους ελέγχους του D' Agostino, των Kolmogorov-Smirnov, των τυποποιημένων διαβαθμίσεων, των Anderson-Darling κ.ά. Η κανονικότητα των ομάδων μπορεί να ελεχθεί χάριν συντομίας και με τη βοήθεια των υπολειμμάτων (βλέπε ενότητα 7.13). Στην περίπτωση αυτή, ο έλεγχος της κανονικότητας με μία από τις παραπάνω μεθόδους χρησιμοποιείται μοναδικά στη στήλη υπολειμμάτων των ομάδων και όχι επαναληπτικά ανά ομάδα, όπως για παράδειγμα θα απαιτείτο τέσσερις φορές αν ήταν τόσος και αριθμός των εξεταζόμενων ομάδων.

(β) Οι διακυμάνσεις όλων των ομάδων θα πρέπει να είναι ίσες (ομοιογενείς) μεταξύ τους: $\sigma_1^2=\sigma_2^2=....\sigma_\kappa^2$. Αυτές εξετάζονται με διάφορους ελέγχους π.χ. έλεγχος του Bartlett, του F_{max}, και του Levene. Ο έλεγχος του Bartlett επηρεάζεται έντονα από μη κανονικούς πληθυσμούς και επομένως θα πρέπει να αποφεύγεται η χρήση του όταν συναντώνται τέτοιες περιπτώσεις. Εναλλακτικά, όταν η ετερογένεια εξακολουθεί να ισχύει παρά την ενδεδειγμένη χρήση μετασχηματισμών, εφαρμόζεται η ανάλυση διακύμανσης F_w κατά Welch, αρκεί να ισχύει η κανονικότητα στα υπολείμματα της ανάλυσης και το δειγματοληπτικό μέγεθος στις ομάδες να είναι σχετικά μικρό (βλέπε σχετικό παράδειγμα στο παράρτημα Β). Σημειώνεται επίσης ότι η ανάλυση διακύμανσης εξακολουθεί μερικές φορές να είναι έγκυρη ακόμη και σε περιπτώσεις σημαντικής ετερογένειας των διακυμάνσεων, αρκεί οι επαναλήψεις σε κάθε ομάδα να είναι ίσες ή περίπου ίσες. Η ομοιογένεια των διακυμάνσεων μπορεί να ελεγχθεί επίσης και γραφικά με την εξέταση του ομοσκεδασμού ή μη των υπολειμμάτων (ενότητα 7.13). Στην περίπτωση αυτή δημιουργείται ένα γράφημα διασποράς με τα υπολείμματα στον άξονα Y και τους μέσους όρους στον άξονα X. Αν το διάγραμμα σχηματίζει αξιοσημείωτη διασπορά των σημείων σε όλη του την έκτασή (ομοσκεδασμός), τότε υποδεικνύει ομοιογένεια στις διακυμάνσεις, αν παρατηρείται κάποια τάση μεταβολής των σημείων, τότε αποτελεί ένδειξη συγκεκριμένης ετερογένειας (ετεροσκεδασμός).

7.2 ΠΟΛΛΑΠΛΕΣ ΣΥΓΚΡΙΣΕΙΣ ΤΩΝ ΜΕΣΩΝ ΟΡΩΝ

Όπως αναφέρθηκε προηγούμενα, η ανάλυση της διακύμανσης δείχνει αν υπάρχουν διαφορές ή όχι μεταξύ των ομάδων ενός πειράματος. Αν πράγματι υπάρχουν διαφορές, που σημαίνει ότι ισχύει η εναλλακτική υπόθεση, τότε χρησιμοποιούνται διάφοροι έλεγχοι που γι' αυτό το λόγο λέγονται *a posteriori* ή *post hoc*, δηλαδή μεταγενέστεροι αφού προηγείται υποχρεωτικά ο έλεγχος της ANOVA. Οι έλεγχοι αυτοί σκοπό έχουν να ανιχνεύσουν ποιοι μέσοι όροι των ομάδων διαφέρουν από τους υπόλοιπους, υποβάλλοντας αυτούς σε διαδοχικές διαφορετικές συγκρίσεις ανά δύο τη φορά. Οι πολλαπλές συγκρίσεις εφαρμόζονται μόνο στο πρότυπο Ι της ANOVA (ή και στο πρότυπο ΙΙΙ όπως θα αναλύσουμε αργότερα). Στο πρότυπο ΙΙ το ενδιαφέρον απλώς εστιάζεται στο αν υπάρχει διαφορά στις διακυμάνσεις μεταξύ των ομάδων και μέσα στις ομάδες.

Η ποσοτική διαφορά μεταξύ δύο συγκρινόμενων μέσων όρων δεν είναι απαραίτητο να είναι σχετικά ισομεγέθης. Οι συνηθέστερες διατάξεις που λαμβάνουν οι μέσοι όροι μεταξύ τους, συγκλίνουν σε μία από τις τρεις περιπτώσεις κατανομής τους, όπως παρίστανται στο σχήμα 7.2.



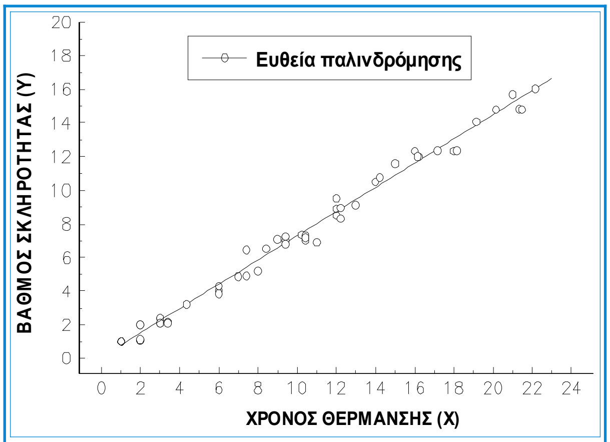
Σχ. 7. 2. Συνηθέστεροι τύποι διάταξης των μέσων όρων.

Υπάρχουν περιπτώσεις η ANOVA, με βάση την υπολογιζόμενη τιμή F, να δείξει στατιστικά σημαντικές διαφορές μεταξύ των μέσων όρων ενός παράγοντα, ενώ παράλληλα αυτές να μην ανιχνεύονται με την εφαρμογή κάποιου από τους ελέγχους των πολλαπλών συγκρίσεων π.χ. του Tukey. Έτσι, από τον έλεγχο αυτό και παρά τις επιδιώξεις μας, μπορεί να προκύψει ότι μετά τις κατά ζεύγη συγκρίσεις όλοι οι μέσοι όροι είναι ίσοι. Οι περιπτώσεις αυτές είναι σπάνιες και εκδηλώνονται όταν η ακριβής πιθανότητα p του ελέγχου της ANOVA προσεγγίζει αυτήν που ορίστηκε ως πιθανότητα αναφοράς (π.χ. α=0.05) ή αλλιώς όταν

8
Γραμμική παλινδρόμηση και συσχέτιση

Στην παραγοντική ανάλυση, παρά την πολυπλοκότητα των σχέσεων μεταξύ των επιπέδων των διαφόρων παραγόντων, γινόταν αναφορά πάντα για μία μόνο μεταβλητή. Στο κεφάλαιο αυτό θα μελετηθεί η ανάλυση δύο ή και περισσοτέρων μεταβλητών, αναφορικά με τις ειδικές σχέσεις που αναπτύσσονται μεταξύ τους. Οι σχέσεις αυτές εκφράζονται με δύο τεχνικές υπολογισμού που είναι η παλινδρόμηση και η συσχέτιση. Στην παλινδρόμηση εκτιμούμε τη σχέση μιας μεταβλητής (Χ) με μία άλλη (Υ), εκφράζοντας αυτή τη σχέση ως γραμμικό αποτέλεσμα της πρώτης (Χ) επί της δεύτερης (Υ). Στη συσχέτιση, εκτιμούμε το βαθμό της έντασης με τον οποίο δύο μεταβλητές μεταβάλλονται ταυτόχρονα. Στην παλινδρόμηση και στη συσχέτιση οι μεταβλητές πρέπει να είναι συνεχείς. Τακτικές μεταβλητές είναι επίσης αποδεκτές και τις μεταχειριζόμαστε ως ποσοτικές αν και υπάρχει αμφισβήτηση ως προς την εγκυρότητα των αποτελεσμάτων, ιδιαίτερα όταν οι διαβαθμίσεις είναι ολιγάριθμες. Αντίθετα, οι ονομαστικές μεταβλητές δεν χρησιμοποιούνται καθόλου, τουλάχιστον με την κατηγορική μορφή που αυτές συνίστανται, αλλά ύστερα από μετασχηματισμό τους σε εικονικές μεταβλητές. Η παλινδρόμηση εκφράζει την εξάρτηση της μεταβλητής Υ από την ανεξάρτητη Χ με μορφή ευθείας γραμμής με τη χρήση της μαθηματικής εξίσωσης, $Y= a + bX$ και η έννοιά της προήλθε από την αρχική απόδοση του όρου σε θέματα βιομετρίας ότι 'επιστρέφει' ή 'επανέρχεται' ή 'γυρίζει πίσω'.

Συχνά, ο ερευνητής θέλει να μάθει αν υπάρχει κάποια σχέση μεταξύ δύο μεταβλητών που εξετάζονται ταυτόχρονα π.χ. σχέση της οξύτητας με το χρόνο της γαλακτικής ζύμωσης στα ξυνολάχανα, σχέση της σκληρότητας της υφής ενός προϊόντος με το χρόνο θέρμανσης, σχέση της νωπότητας ενός ψαριού με το χρόνο συντήρησης κτλ. Η σχέση αυτή μεταξύ των δύο μεταβλητών είναι ουσιαστικά εξάρτηση της πρώτης από τη δεύτερη, δηλαδή το μέγεθος της μεταβολής της πρώτης προσδιορίζεται ως το τυχαίο αποτέλεσμα της δράσης της δεύτερης, χωρίς να ισχύει το αντίθετο. Τονίζεται επίσης ότι τα όρια μέσα στα οποία μεταβάλλεται η δεύτερη μεταβλητή καθορίζονται συνήθως επιλεκτικά από τον οργανωτή του πειράματος και σπανιότερα με τυχαίο τρόπο. Η πρώτη μεταβλητή καλείται *εξαρτημένη* ή μεταβλητή *απόκρισης* ή *προβλέψιμη* και συμβολίζεται με Υ και η δεύτερη καλείται *ανεξάρτητη* ή *προβλεπτική* ή *επεξηγηματική* και ορίζεται με Χ. Για παράδειγμα, στη σχέση μεταξύ της σκληρότητας της υφής και του χρόνου θέρμανσης σε ένα τρόφιμο, η σκληρότητα της υφής (μηχανική ιδιότητα) εκλαμβάνεται ως εξαρτημένη μεταβλητή και ο χρόνος υποβολής του προϊόντος σε θέρμανση εκλαμβάνεται ως ανεξάρτητη. Ευνόητο είναι ότι, δεν είναι λογικό να θεωρήσουμε ότι ο χρόνος θέρμανσης ενός τροφίμου εξαρτάται από τη σκληρότητά του, δηλαδή δεν ισχύει η αντίστροφη σχέση. Η μεταβλητή Χ καλείται ανεξάρτητη, διότι ελέγχεται με μετρήσεις που διεξάγουμε (οι τιμές ελέγχονται από τον ερευνητή), το αποτέλεσμα των οποίων αναμένεται να φανεί στη μεταβλητή Υ και της οποίας οι τιμές εξαρτώνται άμεσα από τις τιμές της Χ. Τέτοια εξαρτημένη σχέση καλείται παλινδρόμηση και πιο συγκεκριμένα απλή παλινδρόμηση, όταν εμπλέκονται δύο μόνο μεταβλητές (Σχ. 8.1). Για περισσότερες της μιας μεταβλητές Χ, π.χ. X_1, X_2, X_3 κτλ., ως προς μία ίδια μεταβλητή Υ, η παλινδρόμηση καλείται *πολλαπλή* και αποτελεί σημαντικό κεφάλαιο εφαρμογών στη βιομηχανία τροφίμων.



Σχ. 8.1. Γραμμική σχέση μεταξύ δύο μεταβλητών Χ και Υ: επίδραση του χρόνου θέρμανσης στην υφή (σκληρότητα) ενός προϊόντος.

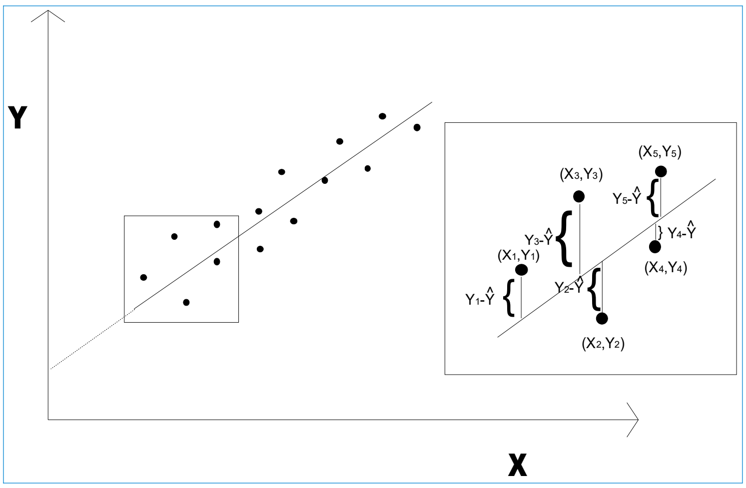
Η μεταβλητή Χ, στην απλή παλινδρόμηση, μπορεί άνετα να συγκριθεί αλλά και να αντικατασταθεί με τα επίπεδα του παράγοντα στην ANOVA μιας κατεύθυνσης. Απαραίτητη προϋπόθεση για να ισχύει η αντικατάσταση θα πρέπει τα επίπεδα του παράγοντα να έχουν ποσοτική μεταβολή, π.χ. αυξημένα ποσοστά συγκέντρωσης μιας ουσίας σε ένα τρόφιμο. Στην περίπτωση αυτή, η μεταβλητή Υ καταγράφεται από τις επαναληπτικές τιμές που λαμβάνονται σε κάθε επίπεδο του παράγοντα ο οποίος παίζει το ρόλο της μεταβλητής Χ. Αν πρόκειται για παραγοντική ανάλυση της ANOVA με k παράγοντες, τα επίπεδα των οποίων έχουν οπωσδήποτε ποσοτική σχέση, τότε οι τιμές των παρατηρήσεων που λαμβάνονται αποτελούν τη μεταβλητή Υ και οι k παράγοντες τις μεταβλητές X_1, $X_2 ... X_k$. Η ανάλυση αυτή είναι ιδιαίτερα γνωστή ως ανάλυση της επιφάνειας απόκρισης (κεφάλαιο 13).

Στην πράξη, οι παρατηρήσεις μιας γραμμικής σχέσης Χ-Υ ποτέ δεν πρόκειται όλες να αποτελούν σημεία της ευθείας γραμμής, αλλά θα κείνται λίγο πάνω και λίγο κάτω αυτής. Αυτό οφείλεται κυρίως στη φυσική μεταβλητότητα των εξεταζόμενων υλικών, εξαιτίας περιβαλλοντικών ή και γενετικών παραγόντων και δευτερογενώς σε πειραματικό σφάλμα. Συνεπώς, στην παλινδρόμηση η γραμμική εξίσωση δεν σημαίνει ότι η τιμή της Υ θα είναι a + bX για συγκεκριμένη τιμή της Χ, αλλά η μέση τιμή όλων των σημείων της Υ που κείνται πάνω και κάτω της ευθείας γραμμής για την ίδια τιμή της Χ.

Υπάρχουν δύο πρότυπα στην παλινδρόμηση τα οποία διαφέρουν κυρίως ως προς τον τρόπο επιλογής της μεταβλητής Χ. Στο πρότυπο Ι η ανεξάρτητη μεταβλητή είναι ελεγχόμενη από τον ερευνητή και άρα επιλέξιμη. Στο πρότυπο ΙΙ είναι τυχαία και επομένως οι δύο μεταβλητές Χ και Υ έχουν τυχαία μεταβλητότητα. Το δεύτερο πρότυπο κυριαρχεί στις περιπτώσεις που λαμβάνουμε ταυτόχρονα δύο διαφορετικές μετρήσεις Χ-Υ από το ίδιο προϊόν χωρίς να έχουμε προελέγξει τον τρόπο μεταβολής της Χ. Αντίθετα, στο πρότυπο Ι λαμβάνουμε πρώτα μία ή περισσότερες μετρήσεις Χ και μετά από παρέλευση κάποιου χρόνου ικανού για να δράσει η μεταβλητή Χ, εκτιμούμε την ή τις αντίστοιχες μετρήσεις της Υ. Δύο ταυτόχρονες μετρήσεις διενεργούνται όταν μετρούμε π.χ. το βάρος και τη διάμετρο δημητριακών καρπών ή όταν μετρούμε μία μεταβλητή με δύο διαφορετικά όργανα μέτρησης που στηρίζονται σε διαφορετικές αρχές λειτουργίας. Επειδή στο πρότυπο ΙΙ και οι δύο μεταβλητές συμμεταβάλλονται αφού μετρούνται ταυτόχρονα στο ίδιο υλικό, θα πρέπει να προκαθορισθεί ποια μεταβλητή θα αποτελέσει την Υ και ποια την Χ. Συνήθως καθορίζεται ως Χ εκείνη της οποίας ο τρόπος μέτρησής της είναι συγκριτικά ευκολότερος και ταχύτερος ή οικονομικότερος. Έτσι, για παράδειγμα, η μέτρηση του παχύρρευστου τοματοπολτών με τη χρήση αμφικλινούς επίστρωτου εργαλείου σε μικρο-χιλιομετρικό χαρτί (blotter), επειδή είναι σαφώς οικονομικότερη και ταχύτερη της μεθόδου Bostwick, θα πρέπει να εκληφθεί ως η προβλεπτική (ανεξάρτητη Χ).

8.1 ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ (simple linear regression)

Στην απλή παλινδρόμηση η σχέση των δύο μεταβλητών απεικονίζεται με ευθεία γραμμή (Σχ. 8.1) με την μεταβλητή Υ στον άξονα των τεταγμένων και την Χ στον άξονα των τετμημένων. Στο σχήμα 8.2 εμφανίζεται επίσης μία διασπορά σημείων, καθένα των οποίων αντιπροσωπεύει ένα ζεύγος από Χ και Υ τιμές. Το πρώτο ζεύγος ορίζεται από τις τιμές X_1,Y_1, το δεύτερο από τις X_2,Y_2, το τρίτο από τις X_3,Y_3 κτλ. Η γενική εξίσωση για μία ευθεία γραμμή που προκύπτει ως αποτέλεσμα της σχέσης της μιας μεταβλητής με την άλλη, αναφορικά ως προς ένα εξεταζόμενο πληθυσμό τυριών, λαδιών, κονσερβών κτλ., καλείται *απλή γραμμική* παλινδρόμηση: $\hat{Y} = a + \beta X$ και αναφορικά με ένα δείγμα του πληθυσμού $\hat{Y} = a + bX$,
όπου η *παράμετρος α* (ή a), γνωστή και ως *ύψος της ευθείας,* είναι η απόσταση του σημείου τομής της προέκτασης της ευθείας στον άξονα Υ από το σημείο τομής των αξόνων Χ και Υ και β ή b ο *συντελεστής παλινδρόμησης* ή η *κλίση της ευθείας*.



Σχ. 8.2. Τμήμα της γραμμικής σχέσης σε μεγέθυνση στο οποίο περιγράφεται ο προσδιορισμός των αποστάσεων των σημείων από την ευθεία προσαρμογής (υπολειμμάτων).

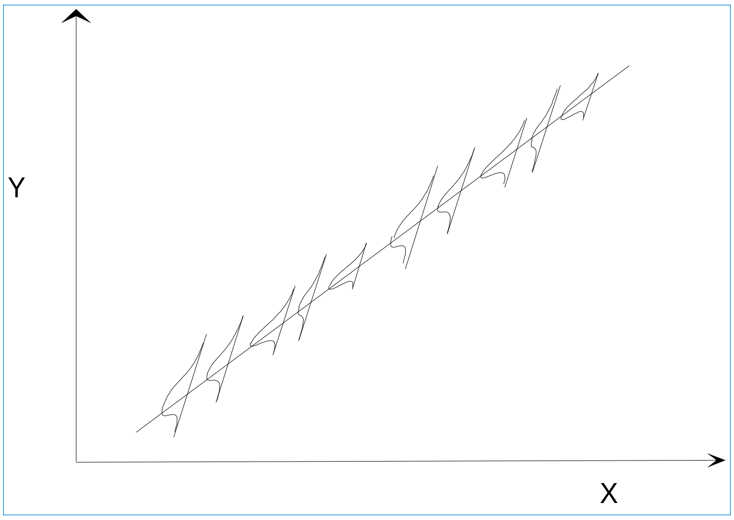
Όπως φαίνεται στο σχήμα 8.2 (και γενικά σε όλες τις γραμμικές σχέσεις δύο μεταβλητών), υπάρχει μία διασπορά των στοιχείων (σημείων) γύρω από μία υποθετική ευθεία γραμμή που θα θέλαμε να σχηματίσουμε και η οποία θα διερχόταν όσο πιο κοντά και τέλεια γινόταν μεταξύ των διεσπαρμένων σημείων. Η ευθεία αυτή καλείται ευθεία της *άριστης προσαρμογής* ή απλά ευθεία προσαρμογής και προσδιορίζεται με τη μέθοδο των *ελαχίστων τετραγώνων.* Η έννοια της μεθόδου γίνεται αντιληπτή ως εξής:
Κάθε τιμή της Χ έχει μία αντίστοιχη τιμή της Υ που βρίσκεται επί της ευθείας που θα θέλαμε να σχηματίσουμε (ευθεία προσαρμογής). Η τιμή αυτή της Υ ορίζεται με $\hat{Y}$ (καπελάκι) για να διαφέρει από την τιμή Υ που πραγματικά μετρήθηκε στο δείγμα. Έτσι, στο σχήμα 8.2 ένα μετρημένο ζεύγος στοιχείων ορίζεται ως (X_1,Y_1) και ένα σημείο στην ευθεία προσαρμογής ως $(X_1,\hat{Y})$. Κριτήριο της μεθόδου των ελαχίστων τετραγώνων είναι η κατακόρυφη απόκλιση κάθε σημείου από την ευθεία προσαρμογής Y_i-$\hat{Y}_i$, γνωστή και ως υπόλειμμα και καθορίζει ως άριστη ευθεία εκείνη που θα προκύψει από την ελάχιστη τιμή του αθροίσματος των

τετραγώνων αυτών των αποκλίσεων απ' όλες τις τιμές Y_i και $\hat{Y}_i$. Το άθροισμα των τετραγώνων αυτών των αποκλίσεων καλείται *υπόλειμμα* ή *σφάλμα του αθροίσματος των τετραγωνισμένων* (residual SS): $ESS=\Sigma(Y_i-\hat{Y}_i)^2 = (Y_1-\hat{Y}_1)^2 + (Y_2-\hat{Y}_2)^2 +...+(Y_n-\hat{Y}_n)^2$, όπου n ο αριθμός των σημείων (παρατηρήσεων) και όσο μικρότερο ποσοτικά είναι, τόσο πλησιέστερα προς την ευθεία προσαρμογής διατάσσονται τα σημεία.

Περιορισμοί της ανάλυσης της παλινδρόμησης

Για να εκφραστεί στατιστικά η μεταβλητή Y ως συνάρτηση (εξάρτηση) της μεταβλητής X με τη βοήθεια της παλινδρόμησης, θα πρέπει να ληφθούν υπόψη ορισμένες προϋποθέσεις:

1. Η ανεξάρτητη μεταβλητή X μετρείται χωρίς σφάλμα, είναι δηλαδή επιλέξιμη (fixed variable) γιατί βρίσκεται κάτω από τον έλεγχο του οργανωτή του πειράματος και έχει ανάλογη σημασία με τον επιλέξιμο παράγοντα στην ANOVA. Αντίθετα, η εξαρτημένη μεταβλητή μεταβάλλεται τυχαία (random variable). Πρακτικά, είναι αδύνατον η ανεξάρτητη μεταβλητή να μετρηθεί χωρίς σφάλμα, γι' αυτό απλά θεωρούμε ότι το σφάλμα στη λήψη των στοιχείων (μετρήσεων) της X είναι αμελητέο ή ελάχιστο συγκριτικά με εκείνο που παρατηρείται στα στοιχεία της Y.
2. Η σχέση μεταξύ των X και Y είναι γραμμική και περιγράφεται από την εξίσωση $\hat{Y} = a + bX$.
3. Για κάθε δεδομένη τιμή της X αντιστοιχεί μία κανονική κατανομή επαναληπτικών μετρήσεων της Y, ανεξαρτήτων μεταξύ τους οι οποίες αποτελούν επίσης ένα τυχαίο δείγμα της κατανομής αυτής (Σχ. 8.3).
4. Οι διακυμάνσεις των κατανομών των επαναληπτικών μετρήσεων της τυχαίας μεταβλητής Y, για κάθε συγκεκριμένη τιμή της X, πρέπει να είναι όλες ίσες μεταξύ τους.



Σχ. 8.3. Παράδειγμα παράθεσης των επαναληπτικών μετρήσεων της μεταβλητής Y σε κανονική κατανομή για συγκεκριμένες τιμές της X.

5. Οι τιμές της μεταβλητής Y είναι όλες ανεξάρτητες μεταξύ τους.
6. Το σφάλμα στην μεταβλητή Y πρέπει να είναι προσθετικό και όχι πολλαπλασιαστικό. Με άλλα λόγια, η διακύμανση των σημείων γύρω από την ευθεία προσαρμογής είναι σταθερή και γι' αυτό ανεξάρτητη από το μέγεθος μεταβολής των X και Y.

Οι προϋποθέσεις που διατυπώθηκαν συνήθως ισχύουν αυτόματα, χωρίς ιδιαίτερη δυσκολία κατά την εφαρμογή της στατιστικής ανάλυσης στην παλινδρόμηση και ως εκ τούτου παρεμβαίνουμε ενεργητικά μόνο όταν παρατηρείται ισχυρή παρέκκλιση σε κάποια από αυτές, προβαίνοντας σε διορθωτικές κινήσεις όπως είναι οι μετασχηματισμοί.

Έλεγχος της σημαντικότητας της παλινδρόμησης (Παράδειγμα 8.1)

Πρακτικά, όσο τα σημεία μιας παλινδρόμησης βρίσκονται πλησιέστερα μεταξύ τους τόσο πιο ισχυρή είναι η σχέση των δύο μεταβλητών, πλην όμως, μόνο στατιστικά μπορεί να βρεθεί η σημαντικότητα και η ένταση της σχέσης αυτής. Ο έλεγχος της σημαντικότητας της παλινδρόμησης βρίσκεται από την ανάλυση της διακύμανσης της γραμμικής παλινδρόμησης με μηδενική υπόθεση H_0: δεν υπάρχει γραμμική σχέση μεταξύ X και Y και με εναλλακτική H_A: υπάρχει γραμμική σχέση μεταξύ X και Y. Οι υπολογισμοί για την εκτίμηση των παραμέτρων α και β και της σημαντικότητας της παλινδρόμησης είναι οι ακόλουθοι:

- Ολικό άθροισμα των τετραγώνων των αποκλίσεων των τιμών Y_i από το μέσο όρο της μεταβλητής Y:

$$TSS = \sum y^2 = \sum(Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$, με βαθμούς ελευθερίας n-1

- Άθροισμα των τετραγώνων των αποκλίσεων των τιμών X_i από το μέσο όρο της μεταβλητής X:

$$\sum x^2 = \sum (X_i - \bar{X})^2 = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

❀ Άθροισμα του γινομένου των τιμών X_i,Y_i, για κάθε αντιστοιχία σειρών,

$$\sum xy = \sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}$$

❀ Άθροισμα των τετραγωνισμένων της γραμμικής παλινδρόμησης RSS (regression SS),

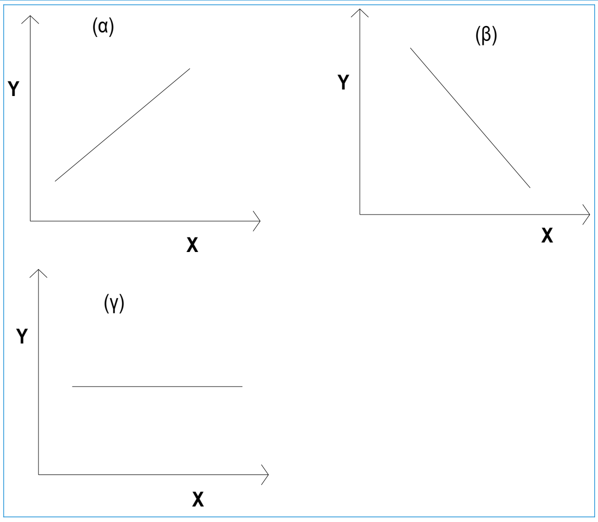
$$RSS = \frac{(\sum xy)^2}{\sum x^2} \quad \text{ή} \quad RSS = b \sum xy \quad ,$$

με βαθμούς ελευθερίας k-1=1, όπου k στην προκείμενη περίπτωση είναι οι δύο στήλες X και Y της παλινδρόμησης (k=2).

❀ Υπολογισμός του συντελεστή της παλινδρόμησης,

$$b = \frac{\sum xy}{\sum x^2} \quad .$$

Ο παρονομαστής στην παραπάνω σχέση είναι πάντα θετικός, ο αριθμητής όμως μπορεί να είναι θετικός ή αρνητικός ή 0 και η τιμή του b μπορεί να κυμαίνεται από -∞ μέχρι +∞ (Σχ. 8.4). Ο συντελεστής b εκφράζει πόση μεταβολή στη μεταβλητή Y σχετίζεται με μία μονάδα μεταβολής στη μεταβλητή X.



Σχ. 8.4. Τρόποι προσανατολισμού της ευθείας προσαρμογής: α) ανοδική κλίση εκφραζόμενη με θετικό συντελεστή b, β) καθοδική κλίση εκφραζόμενη με αρνητικό συντελεστή b, γ) κλίση σταθερή με b=0.

❀ Υπολογισμός της παραμέτρου a,

$a = \bar{Y} - b\bar{X}$

❀ Άθροισμα του υπολείμματος των τετραγωνισμένων, ή σφάλμα ESS,

$ESS = \Sigma (Y_i - \hat{Y}_i)^2 = TSS - RSS$ με βαθμούς ελευθερίας n-2

❀ Υπολογισμός των μέσων αθροισμάτων των τετραγωνισμένων της γραμμικής παλινδρόμησης RMS και του υπολείμματος EMS,

RMS = RSS/(k-1)
EMS = ESS/(n-2)

Η ποσότητα EMS συμβολίζεται και ως s^2_{xy} και εκφράζει τη διακύμανση της Y ως αποτέλεσμα της εξάρτησης της Y από τη X. Η ρίζα αυτής της ποσότητας s_{xy} καλείται *τυπικό σφάλμα* της παλινδρόμησης και δείχνει την ακρίβεια με την οποία η παλινδρόμηση επεξηγεί την εξάρτηση της Y από την X η οποία γράφεται ως εξής: $\hat{Y} = a + bX \pm s_{xy}$.

❀ Υπολογισμός του κριτηρίου της ανάλυσης της διακύμανσης F,

$$F = \frac{RMS}{EMS} \quad .$$

Η τιμή F συγκρίνεται με την οριακή τιμή $F_{0.05(1)1,n-2}$ που προκύπτει από τον πίνακα Π3. Αν $F \geq F_{op}$, τότε απορρίπτεται η μηδενική υπόθεση, δηλαδή υπάρχει γραμμική σχέση μεταξύ των δύο μεταβλητών. Η ανάλυση της διακύμανσης για τη σημαντικότητα της γραμμικής παλινδρόμησης συνοψίζεται στον πίνακα 8.1.

Πίνακας 8.1. Ανάλυση διακύμανσης για τη σημαντικότητα της παλινδρόμησης.

Πηγές μεταβλητότητας	Βαθμοί ελευθερίας (df)	Αθροίσματα τετραγωνισμένων (SS)	Μέσα αθροίσματα τετραγωνισμένων (MS)	F τιμή
Ολική	n-1	$TSS = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$		
Παλινδρόμηση	k-1=1	$RSS = \frac{(\Sigma xy)^2}{\Sigma x^2}$	$RMS = \frac{RSS}{1}$	
Υπόλειμμα	n-2	$ESS = TSS - RSS$	$EMS = \frac{ESS}{n-2}$	$F = \frac{RMS}{EMS}$

❀ Υπολογισμός του *συντελεστή προσδιορισμού της προσαρμογής* R^2 *ή* r^2,

$$R^2 = \frac{RSS}{TSS} \quad .$$

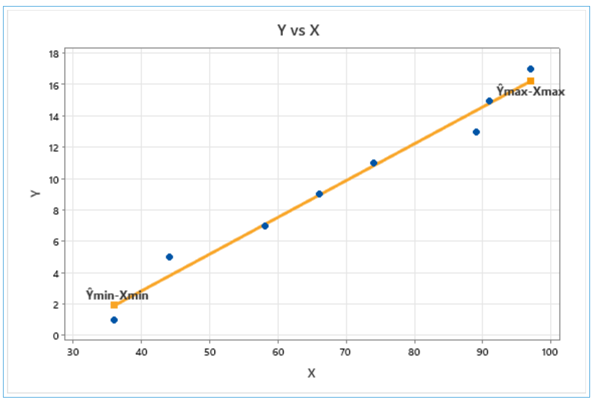
Ο συντελεστής R^2 δηλώνει την ποιότητα προσαρμογής της ευθείας της γραμμικής παλινδρόμησης ή με άλλα λόγια την αναλογία (ή ποσοστό) της συνολικής διακύμανσης των παρατηρούμενων τιμών της Y και ουσιαστικά μετρά την ισχύ της γραμμικής σχέσης των δύο μεταβλητών. Ο συντελεστής παίρνει τιμές από 0 μέχρι 1 (ή σε ποσοστά 0-100%) και όσο πιο κοντά πλησιάζει στο 1 ή 100%, τόσο μεγαλύτερο ποσοστό της διακύμανσης της Y εξηγεί, δηλαδή τόσο πιο καλή (άριστη) είναι η εξίσωση της παλινδρόμησης. Όταν όμως το δείγμα είναι μικρό (n<10), ο προσδιοριστικός συντελεστής δεν είναι ιδιαίτερα ακριβής και γι' αυτό χρησιμοποιείται ο πιο αμερόληπτος *διορθωμένος συντελεστής προσδιορισμού της προσαρμογής* R^2_δ,

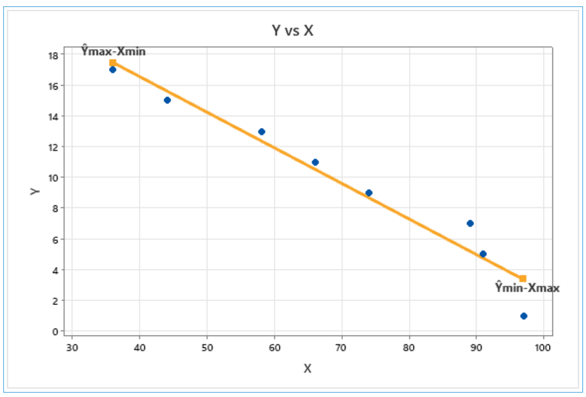
$$R^2_\delta = 1 - \frac{ESS\ (n-1)}{TSS\ (n-2)} \quad ,$$

ο οποίος είναι πάντα μικρότερος του προηγούμενου.

Γραφική απεικόνιση της ευθείας προσαρμογής

Ο σχηματισμός της ευθείας προσαρμογής σε γραφική παράσταση πραγματοποιείται εύκολα με τη χρήση γραφικών και σχεδιαστικών προγραμμάτων, σε έλλειψη όμως αυτών τηρείται η ακόλουθη διαδικασία:
(α) Εύρεση των παραμέτρων a και b.
(β) Σχηματισμός των αξόνων X και Y και τοποθέτηση των σημείων που αντιστοιχούν σε κάθε ζεύγος παρατηρήσεων.
(γ) Από την εξίσωση παλινδρόμησης $\hat{Y} = a \pm bX$ βρίσκουμε την ελάχιστη και μέγιστη τιμή της $\hat{Y}$ που προκύπτουν εισάγοντας την ελάχιστη και μέγιστη τιμή της X από τις παρατηρήσεις του δείγματος. Σημειώνουμε τα σημεία των δύο ζευγών (ελάχιστο και μέγιστο ζεύγος) στη γραφική παράσταση και φέρουμε την ευθεία που διέρχεται από αυτά τα ζεύγη, και δεν είναι άλλη από την ευθεία προσαρμογής.
Σε ανοδικές κλίσεις τα δύο ζεύγη τιμών σημειώνονται ως $\hat{Y}_{min}$-X_{min} και $\hat{Y}_{max}$-X_{max}, και σε καθοδικές κλίσεις ως $\hat{Y}_{max}$-X_{min} και $\hat{Y}_{min}$-X_{max}, όπως παριστάνονται στα παρακάτω σχήματα παλινδρόμησης.





Έλεγχος της σημαντικότητας των συντελεστών παλινδρόμησης α και β

Εναλλακτικά, αντί της ανάλυσης διακύμανσης για τη σημαντικότητα της παλινδρόμησης, μπορούμε να χρησιμοποιήσουμε απευθείας τον έλεγχο της σημαντικότητας της κλίσης b που είναι ουσιαστικά δίπλευρος έλεγχος της υπόθεσης ενός δείγματος με μηδενική υπόθεση H_O: β=0 και εναλλακτική H_A: β≠0. Πρώτα υπολογίζεται το *τυπικό σφάλμα* του συντελεστή παλινδρόμησης και μετά το στατιστικό κριτήριο t,

$$s_b = \sqrt{\frac{EMS}{\sum x^2}} \quad , \qquad t = \frac{b}{s_b} \quad .$$

Η τιμή t συγκρίνεται με την οριακή $t_{0.05(2)n-2}$ που προκύπτει από τον πίνακα Π1 και η μηδενική υπόθεση απορρίπτεται όταν $|t| \geq t_{op}$.

Στην περίπτωση που θέλουμε να συγκρίνουμε το συντελεστή b με κάποιον άλλο συντελεστή b_o, τότε ο έλεγχος είναι είτε μονόπλευρος με μηδενική υπόθεση H_0: $\beta \geq \beta_0$ (ή $\beta \leq \beta_0$) και εναλλακτική υπόθεση H_A: $\beta < \beta_0$ (ή $\beta > \beta_0$), είτε δίπλευρος με H_0: $\beta = \beta_0$ και H_A: $\beta \neq \beta_0$. Το κριτήριο t υπολογίζεται από τη σχέση,

$$t = \frac{b - b_o}{s_b} \quad ,$$

και η οριακή τιμή στον πίνακα είναι $t_{0.05(1)n-2}$ για μονόπλευρο έλεγχο και $t_{0.05(2)n-2}$ για δίπλευρο.

Με παρόμοιο τρόπο ελέγχεται και η σημαντικότητα της παραμέτρου a, υιοθετώντας το δίπλευρο έλεγχο της υπόθεσης ενός δείγματος, H_0: α=0 και H_A: α≠0 και υπολογίζοντας πρώτα το τυπικό σφάλμα της παραμέτρου a και μετά το στατιστικό κριτήριο t,

$$s_a = \sqrt{EMS \cdot \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum x^2} \right)} \quad , \qquad t = \frac{a}{s_a} \quad .$$

Η τιμή t συγκρίνεται με την οριακή $t_{0.05(2)n-2}$ που προκύπτει από τον πίνακα Π1 και η μηδενική υπόθεση απορρίπτεται όταν $|t| \geq t_{op}$.

Παράδειγμα 8.1. Υπολογισμός της εξίσωσης της απλής γραμμικής παλινδρόμησης με τη μέθοδο των ελαχίστων τετραγώνων.

Η επίδραση της ψύξης με πάγο στη νωπότητα του μπακαλιάρου, μελετήθηκε χρησιμοποιώντας διαβαθμισμένη κλίμακα οσμής των προϊόντων από 0 μέχρι 20 και για χρονικό διάστημα 13 ημερών. Τα ερωτήματα που τίθενται είναι η περιγραφή του τρόπου μεταβολής της οσμής με την πάροδο του χρόνου ψύξης και ο ρυθμός της μεταβολής αυτής.

Βαθμοί οσμής (μεταβλητή Y):	19	18	16	17	14	13	12	11	8	5	6	4	2
Ημέρες στον πάγο (μεταβλητή X):	1	2	3	4	5	6	7	8	9	10	11	12	13

1. Έλεγχος της σημαντικότητας της παλινδρόμησης.

H_O: δεν υπάρχει γραμμική σχέση μεταξύ των μεταβλητών X και Y
H_A: υπάρχει γραμμική σχέση μεταξύ των μεταβλητών X και Y

n=13

$\bar{X}$ =7 ΣX=91 ΣX^2 =819 $\qquad \sum x^2 = 819 - \frac{(91)^2}{13} = 182$

$\bar{Y}$ =11.1 ΣY=145 ΣY^2 =2005 $\qquad \sum y^2 = TSS = 2005 - \frac{(145)^2}{13} = 387.69$

ΣXY=753 Σxy=753-(91·145)/13=-262

$$RSS = \frac{\left(\sum xy\right)^2}{\sum x^2} = \frac{(-262)^2}{182} = 377.16$$

ESS=TSS-RSS=387.69-377.16=10.53

$EMS = s^2 = ESS/(n-2) = 10.53/11 = 0.96$ $\qquad s_{xy} = 0.980$

$R^2 = RSS/TSS = 377.16/387.69 = 0.973$ $\qquad R^2_\delta = 1 - \frac{ESS(n-1)}{TSS(n-2)} = 1 - \frac{10.53 \cdot 12}{387.69 \cdot 11} = 0.970$

$F = RMS/EMS = 377.16/0.96 = 392.87$

Ανάλυση διακύμανσης

Πηγές μεταβλητότητας	df	SS	MS	F	p
Ολική	12	387.69			
Παλινδρόμηση	1	377.16	377.16	392.87	0.000
Υπόλειμμα	11	10.53	0.96		

Η τιμή F του ελέγχου της ανάλυσης της διακύμανσης της παλινδρόμησης είναι στατιστικά σημαντική ($p<0.05$) και επομένως υπάρχει γραμμική σχέση μεταξύ της νωπότητας και του χρόνου ψύξης.

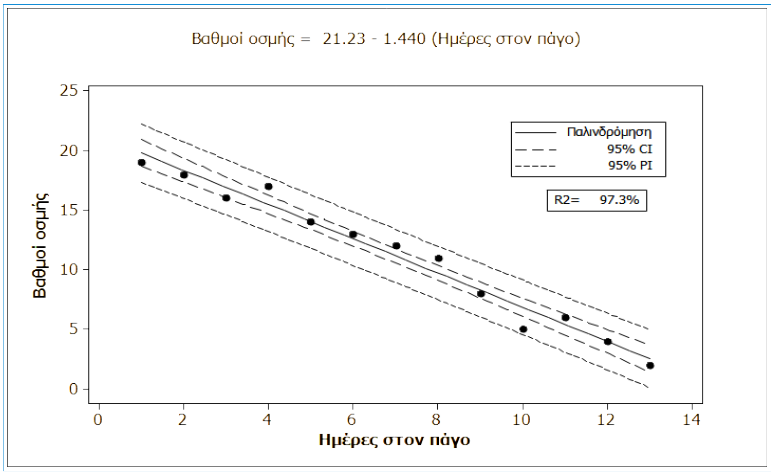
2. Έλεγχος της σημαντικότητας των παραμέτρων a και b της παλινδρόμησης.

H_O: $\beta=0$ H_O: $a=0$

H_A: $\beta\neq0$ H_A: $a\neq0$

$$b=\frac{\sum xy}{\sum x^2}=\frac{-262}{182}=-1.44 \qquad s_b=\sqrt{\frac{EMS}{\sum x^2}}=\sqrt{\frac{0.96}{182}}=0.073 \quad , \qquad a=\bar{Y}-b\cdot\bar{X}=11.1-(-1.44)\cdot 7=21.18$$

$$s_a=\sqrt{EMS\left(\frac{1}{n}+\frac{\bar{X}^2}{\sum x^2}\right)}=\sqrt{0.96(\frac{1}{13}+\frac{7^2}{182})}=0.576$$

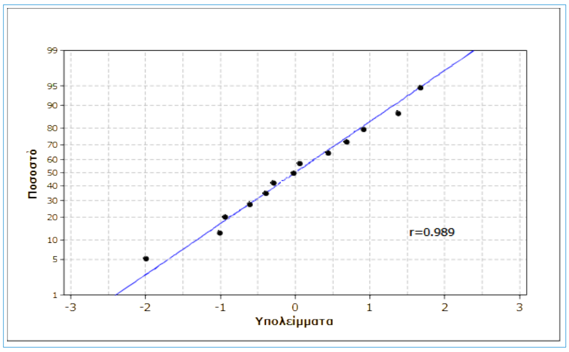


Παράμετροι	Τιμή	SE	Τιμή t	p
Παράμετρος a	21.18	0.576	36.89	0.000
Συντελεστής b	-1.44	0.073	-19.85	0.000

Ο έλεγχος t έδειξε την παράμετρο a και το συντελεστή b στατιστικά σημαντικούς ($p<0.05$). Έτσι, η εξίσωση μεταξύ της νωπότητας Y και του χρόνου ψύξης X περιγράφεται ως: $\hat{Y}$ = 21.2 - 1.44 X, δηλαδή η αύξηση του χρόνου ψύξης προκαλεί μείωση της νωπότητας του μπακαλιάρου με ρυθμό μείωσης 1.44 βαθμούς οσμής την ημέρα.

3. Διαγνωστικά κριτήρια της παλινδρόμησης (περιγράφονται σε ειδική ενότητα παρακάτω).

α) Γραφικός έλεγχος της κανονικότητας των υπολειμμάτων με τη μέθοδο των πιθανοτήτων της κανονικής κατανομής (βλέπε κεφάλαιο 5).



Τα σημεία ακολουθούν γραμμική πορεία και η σημαντικότητα της γραμμικότητας τους είναι αποδεκτή στατιστικά σε 0.05 επίπεδο λάθους (r=0.989>0.938), επομένως οι αποκλίσεις της Y ακολουθούν την κανονική κατανομή.

β) Γραφικός έλεγχος της ομοιογένειας των υπολειμμάτων σε σχέση με την ανεξάρτητη μεταβλητή X και τη στήλη των προσαρμοσμένων τιμών.

Τα υπολείμματα διασπείρονται ομοιόμορφα στα δύο γραφήματα και άρα η διακύμανση είναι ανεξάρτητη από τη δράση των X

11
Στατιστικός έλεγχος της παραγωγικής διαδικασίας

Σε κάθε στάδιο της παραγωγικής διαδικασίας ενός προϊόντος οφείλουμε να γνωρίζουμε αν το προϊόν επεξεργάζεται σύμφωνα με κάποιες προδιαγραφές ποιότητας που αποτελούν απαραίτητα στοιχεία ευνοϊκής αποδοχής του προϊόντος από τους καταναλωτές. Κατά τη διάρκεια της παραγωγικής διαδικασίας δύο είναι τα μείζονα προβλήματα που απαντώνται στην ποιότητα του τελικού προϊόντος:

(α) αποκλίσεις από τις προδιαγραφές που θέτουμε ως στόχο

(β) έντονη μεταβλητότητα γύρω από την επίτευξη των προδιαγραφών.

Για τη βελτίωση ή και εξάλειψη αυτών των προβλημάτων έχουν αναπτυχθεί ορισμένοι τομείς της στατιστικής μεθοδολογίας ποιοτικού ελέγχου, οργανωμένοι έτσι ώστε να μπορούμε να επιβλέπουμε κάθε στιγμή την ποιότητα των παραγόμενων προϊόντων στα διάφορα στάδια επεξεργασίας τους και να επεμβαίνουμε διορθωτικά στο στάδιο που εμφανίζει το πρόβλημα. Η φιλοσοφία του ελέγχου της διαδικασίας παραγωγής σε κάθε στάδιο της, γνωστό και ως διεργασία, είναι αρκετά απλή:

Συλλέγουμε κάποια δείγματα ορισμένου μεγέθους από τη διαδικασία παραγωγής, μετρούμε μία ή περισσότερες καθορισμένες μεταβλητές (π.χ. βάρος ενός περιέκτη) και δημιουργούμε διαγράμματα ελέγχου της μεταβλητότητας του μετρούμενου χαρακτηριστικού, καθώς και της προσέγγισης των δειγμάτων στις τελικές προδιαγραφές. Αν παρουσιάζεται κάποια αισθητή τάση μεταβολής στα διαγράμματα ελέγχου ή επίσης αν τα δείγματα αποκλίνουν από τα όρια των προδιαγραφών, τότε δηλώνουμε ότι η διεργασία της παραγωγής είναι εκτός ελέγχου και δραστηριοποιούμαστε για την εξεύρεση της αιτίας που προκαλεί τη γένεση του προβλήματος. Τα διαγράμματα ελέγχου είναι γνωστά ως *διαγράμματα Shewhart*.

Το κέρδος από τη χρήση του στατιστικού ελέγχου παραγωγής συνοψίζεται στα εξής σημεία:

1) Βελτιωμένη τυποποίηση του προϊόντος με αποτέλεσμα:

- μείωση του κόστους παραγωγής αφού λιγότερο υλικό θα παράγεται εκτός προδιαγραφών, λιγότερες παρεμβάσεις θα απαιτούνται, όπως και λιγότερες εξετάσεις του προϊόντος
- αύξηση της φήμης του προϊόντος με συνακόλουθη αύξηση των πωλήσεων και αίσθημα ικανοποίησης στον καταναλωτή
- καθορισμός αυστηρότερων ορίων προδιαγραφών και άρα αυστηρότερα ελεγμένο προϊόν

2) Αύξηση του ρυθμού παραγωγής. Η μεγιστοποίηση της παραγωγής με το ελάχιστο δυνατόν κόστος έχει ως αποτέλεσμα την αλματώδη αύξηση των κερδών της εταιρίας, ακόμα και αν αυτό συντελείται με μικρή σχετικά αύξηση δαπανών για την αναβάθμιση της παραγωγικής διαδικασίας.

Ο στατιστικός έλεγχος της παραγωγής διενεργείται σε δύο φάσεις:

(α) Στον προληπτικό έλεγχο, επιθεωρούμε τη διαδικασία παραγωγής και εφαρμόζουμε ειδικούς ελέγχους στις διεργασίες για να αποφύγουμε την παραγωγή ελαττωματικών προϊόντων. Στον προληπτικό έλεγχο εντάσσονται:

- η δειγματοληπτική εξέταση του εισαγόμενου υλικού
- τα διαγράμματα ελέγχου των ποσοτικών μεταβλητών
- τα διαγράμματα των αθροιστικών (συσσωρευτικών) αποκλίσεων
- η συνεχής εξέταση του παραγόμενου προϊόντος

(β) Στον έλεγχο με διαλογή, επιθεωρούμε το παραγόμενο υλικό και αν η ποιότητά του δεν είναι ικανοποιητική, απομονώνουμε τις ελαττωματικές μονάδες είτε προς αποκατάσταση των ατελειών τους, αν είναι εφικτό, είτε προς διάθεση με μειωμένη τιμή. Αυτό γίνεται με δειγματοληπτική εξέταση των προϊόντων και ο έλεγχος με διαλογή πολλές φορές αποφεύγεται επειδή είναι εξαιρετικά δαπανηρός.

11. 1 ΔΙΑΓΡΑΜΜΑΤΑ ΕΛΕΓΧΟΥ ΤΟΥ SHEWHART

Η βασική ιδέα στο στατιστικό έλεγχο της παραγωγικής διαδικασίας είναι να επινοήσουμε ένα απλό μηχανισμό ελέγχου που να απεικονίζει και να παρακολουθεί το μέσο επίπεδο μεταβλητότητας και την έκταση της διασποράς μιας μεταβλητής μιας διεργασίας. Η επιλεγμένη μεταβλητή θεωρείται η πλέον κρίσιμη και αποτελεί βασικό κριτήριο για τον έλεγχο της διεργασίας. Επομένως, δύο τουλάχιστον διαγράμματα απαιτούνται, ένα να ελέγχει τη μέση τιμή της μεταβλητής της διεργασίας και ένα να ελέγχει το εύρος της διασποράς της μεταβλητής.

Η μεταβλητότητα της ποιότητας του προϊόντος διαμερίζεται σε δύο κατηγορίες, στην τυχαία μεταβλητότητα και σε αυτήν που οφείλεται σε συγκεκριμένα ή αναγνωρίσιμα αίτια. Η τυχαία μεταβλητότητα προέρχεται από πολλούς συνδυασμούς αιτίων. Το αποτέλεσμα όμως του καθενός αιτίου είναι μικρό σε έκταση και τίποτα σχεδόν δεν μπορούμε να πράξουμε για να διορθωθούν ή να εξαλειφθούν τα αίτια αυτά. Τα συγκεκριμένα αίτια μεταβλητότητας, όπως η μεταβλητότητα (διακύμανση) στην ποιότητα της πρώτης ύλης ή το ανειδίκευτο προσωπικό, είναι συνήθως διακριτά και μπορούν να διορθωθούν μέχρις ενός βαθμού.

Αν τα στοιχεία της μεταβλητής μιας διεργασίας έχουν συγκεκριμένη κατανομή (κανονική, διωνυμική, Poisson) με συγκεκριμένες παραμέτρους ανάπτυξής της (όπως μέσος όρος, μέση αναλογία, μέση απαρίθμηση) και κυμαίνονται σε καθορισμένα όρια μεταβολής μέσα στο διάγραμμα, τότε αποδεχόμαστε ότι η *διεργασία είναι υπό έλεγχο*. Από την άλλη πλευρά, έντονη μεταβλητότητα των παραμέτρων που υπερβαίνει μερικές φορές τα όρια μεταβολής τους, βγάζει τη διεργασία εκτός ελέγχου.

Χαρακτηριστικά των συλλεγόμενων στοιχείων

Ανάλογα με τον τρόπο που συλλέγονται τα στοιχεία για τη διενέργεια του στατιστικού ελέγχου της διαδικασίας παραγωγής, διακρίνουμε τρεις κατηγορίες:

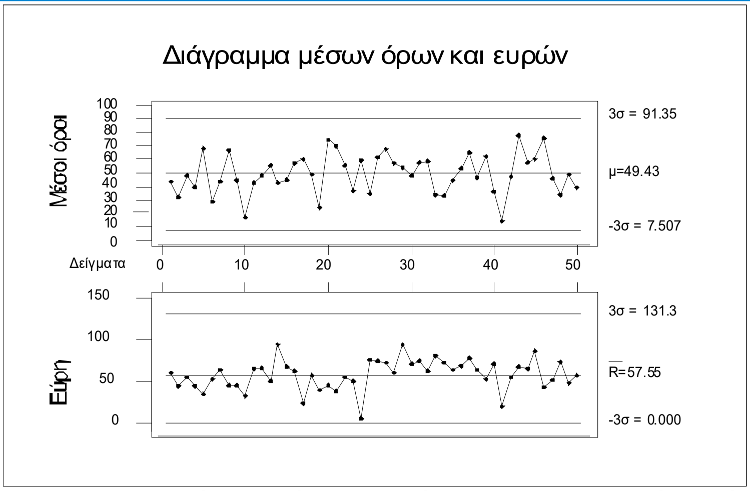
* Στοιχεία στα οποία αποδίδουμε δύο κατηγορίες χαρακτηρισμών των μονάδων που επιθεωρήσαμε, προϊόν ελαττωματικό και κανονικό. Οι μονάδες κατά την εξέταση τους εύκολα κατατάσσονται σε ελαττωματικές ή μη, πράγμα που κάνει τη συλλογή των στοιχείων τους απλή υπόθεση. Η απόδοση των χαρακτηρισμών στις μονάδες δεν είναι πάντα εύκολη υπόθεση, όπως συμβαίνει στην περίπτωση που πρέπει το υλικό να υποστεί κάποια χημική επεξεργασία και το αποτέλεσμα της οποίας θα κρίνει την έκβαση του χαρακτηρισμού στο υλικό ως συμβατικό ή μη. Τα στοιχεία της κατηγορίας αυτής, όταν απαριθμούνται ατομικά αποτελούν τη γνώριμη κατηγορία των ποιοτικών ή κατηγορικών μεταβλητών, συνήθως όμως συναντώνται στον ποιοτικό έλεγχο υπό μορφή συχνοτήτων (αθροίσματα των ελαττωματικών προϊόντων πάσης μορφής) οι οποίες είναι γνωστές με το όνομα χαρακτηρισμοί ή *ιδιότητες* (attributes).
* Στοιχεία που συλλέγονται με απλή καταμέτρηση χωριστά για κάθε συγκεκριμένο ασύμβατο χαρακτηριστικό του προϊόντος, το οποίο ονομάζεται *ελάττωμα* (defect) τύπου Α, Β κτλ., και το οποίο μετράται στη μονάδα του χρόνου ή στη μονάδα μάζας του υλικού. Επιπρόσθετα δηλαδή, της προηγούμενης κατηγορίας των ιδιοτήτων, εδώ μπορούμε να καταγράψουμε όλων των ειδών τα ελαττώματα που απαντώνται στα ελαττωματικά προϊόντα, πράγμα που δεν γίνεται με τους χαρακτηρισμούς, γιατί αυτοί είναι δυαδικής φύσης.

Οι δύο κατηγορίες (ιδιότητες ή ελαττώματα) αποτελούν ως συχνότητες ασυνεχείς ποσοτικές μεταβλητές και στην πράξη τις μεταχειριζόμαστε με τον ίδιο τρόπο αφού και οι δύο αναφέρονται ουσιαστικά σε καταμετρήσεις ελαττωματικών μονάδων. Ο μόνος βασικός τρόπος διαχωρισμού και κατάταξης των στοιχείων σε μία από τις δύο αυτές κατηγορίες είναι η συχνότητα με την οποία εμφανίζονται τα ελαττωματικά προϊόντα. Η συχνότητα είναι πολύ μικρή στις απαριθμήσεις (π.χ. 0.5%) και σχετικά υψηλή (π.χ. 5%) στους χαρακτηρισμούς. Αν τα ποσοστά τους συγκλίνουν οριακά, τότε μόνο ειδικοί έλεγχοι της διασποράς της κατανομής τους στα δείγματα μπορούν να καταδείξουν τη φύση της προέλευσής τους.

* Στοιχεία που συλλέγονται από τη μάζα του προϊόντος με τη χρήση διαβαθμισμένης ή συνεχούς κλίμακας, αφορούν δηλαδή ποσοτικές συνεχείς μεταβλητές. Τέτοιες μεταβλητές μπορεί να είναι η σκληρότητα ενός μετάλλου, το βάρος ενός περιέκτη ή του τροφίμου μέσα σε αυτόν, το ποσοστό του χυμού μιας πορτοκαλάδας, η υγρασία κτλ.

Χαρακτηριστικά των διαγραμμάτων ελέγχου

Τα χαρακτηριστικά και η ερμηνεία των διαγραμμάτων ελέγχου της εξεταζόμενης μεταβλητής εμφανίζουν κοινό σκεπτικό διαμόρφωσης των συμπερασμάτων. Στα περισσότερα από αυτά χρησιμοποιούμε το διάγραμμα του μέσου όρου (διάγραμμα $\bar{X}$) και το διάγραμμα του εύρους (διάγραμμα R), όπως δείχνονται στο σχήμα 11.1. Στα διαγράμματα αυτά ο οριζόντιος άξονας X αντιπροσωπεύει τον αύξοντα αριθμό των διαφορετικών δειγμάτων που ελήφθησαν διαδοχικά και ο κάθετος άξονας Y παραθέτει τους μέσους όρους ή τα εύρη μεταβολής ή τις διακυμάνσεις της εξεταζόμενης μεταβλητής ή ακόμα και τις απαριθμήσεις των ελαττωματικών μονάδων. Για παράδειγμα, ας υποθέσουμε ότι θέλουμε να ελέγξουμε τη διάμετρο των πωμάτων που χρησιμοποιούμε για την εμφιάλωση αναψυκτικών. Η κεντρική ευθεία στο διάγραμμα $\bar{X}$ αντιστοιχεί στη συνολική μέση τιμή μ (διάμετρο) των πωμάτων στα δείγματα που εξετάστηκαν, ενώ η κεντρική ευθεία στο διάγραμμα R αντιστοιχεί στο συνολικό μέσο εύρος $\bar{R}$ των διαμέτρων μέσα στα δείγματα (οι μέθοδοι εκτίμησης των μέσων όρων και μέσων ευρών περιγράφονται στην ενότητα 11.2). Το δεύτερο αυτό διάγραμμα ουσιαστικά εκφράζει το μέτρο της μεταβλητότητας της μέσης τιμής της μεταβλητής της διεργασίας, γιατί όσο αυξάνει το εύρος τόσο αυξάνει και η διάμετρος του πώματος. Η κεντρική ευθεία ουσιαστικά αποτελεί το μέσο όρο της μεταβλητής της διεργασίας ή γενικότερα τη μέση τιμή της διεργασίας υπολογιζόμενη πάντα από συγκεκριμένο πλήθος παρατηρήσεων (δειγμάτων). Η μέση τιμή της μεταβλητής οφείλει να προσεγγίζει σε μεγάλο βαθμό μία μέση τιμή που έχει οριστεί ως δείκτης ποιότητας για τη συγκεκριμένη διεργασία και που συνήθως αποτελεί το *στόχο* (target) της διεργασίας. Δεν είναι σπάνιο το γεγονός, ένα διάγραμμα να εμφανίζεται ότι είναι υπό έλεγχο ενώ παράλληλα η μέση τιμή της μεταβλητής να αποκλίνει έντονα από το στόχο. Πλην της κεντρικής ευθείας, το τυπικό διάγραμμα περιλαμβάνει και δύο οριζόντιες ευθείες που αντιπροσωπεύουν τα όρια ελέγχου, ανώτερο και κατώτερο. Τα σημεία στο διάγραμμα αποτελούν τους μέσους όρους ή εύρη ανά δείγμα ή απλώς ατομικές παρατηρήσεις (αν κάθε δείγμα περιέχει μία μόνο τιμή) και συνδέονται μεταξύ τους με ευθεία γραμμή. Η γραμμή αυτή στο διάγραμμα περνώντας από τα διάφορα σημεία διαγράφει μία τεθλασμένη πορεία και είναι γνωστή ως *γραμμή ελέγχου*. Αν η γραμμή αυτή κινείται έξω από το κατώτερο ή ανώτερο όριο ελέγχου, τότε η ποιότητα του προϊόντος παρουσιάζει κάποιο πρόβλημα.



Σχ. 11.1. Τυπικά διαγράμματα ελέγχου της μεταβολής ενός μετρούμενου χαρακτηριστικού κατά τη διεργασία παραγωγής ενός προϊόντος.

Προσδιορισμός των ορίων ελέγχου

Βασίζεται στην αρχή ότι ο μέσος όρος και η διακύμανση του χαρακτηριστικού που διερευνάται δεν μεταβάλλεται έντονα κατά τη διάρκεια μιας διαδικασίας παραγωγής του προϊόντος και επομένως οι μέσοι όροι των δειγμάτων που λαμβάνονται διαδοχικά για τη μέτρηση του χαρακτηριστικού θα διαφοροποιούνται ομαλά (κατανέμονται κανονικά) γύρω από τον πραγματικό μέσο όρο. Ο μέσος όρος αυτός συνήθως εξάγεται από την ενσωμάτωση όλων των μέσων όρων των δειγμάτων και εκφράζει τη μέση τιμή της μεταβλητής. Η συμπεριφορά αυτή των μέσων όρων των δειγμάτων είναι καθοριστική για τη διαμόρφωση και εξέλιξη κάθε διαγράμματος ελέγχου, εφόσον ισχύει η κανονική κατανομή. Οπότε, η μεταβλητή θα έχει τυπική απόκλιση σ των μέσων μετρήσεων του χαρακτηριστικού και τυπικό σφάλμα $\sigma/\sqrt{n}$, όπου n είναι το πλήθος των παρατηρήσεων ανά δείγμα. Σύμφωνα με τους κανόνες της παραμετρικής στατιστικής, το 95% των μέσων όρων των δειγμάτων θα εμπεριέχεται στα όρια εμπιστοσύνης $\mu \pm 2\sigma/\sqrt{n}$ (ή 1.96 ακριβέστερα) και το 99.7% αυτών θα εμπεριέχεται στα όρια εμπιστοσύνης $\mu \pm 3\sigma/\sqrt{n}$ (ή 3.09 ακριβέστερα). Έχει καθιερωθεί στην πράξη, τα όρια ελέγχου να καταγράφονται ως *±2σ όρια* και *±3σ όρια*. Τα 99.7% όρια ελέγχου είναι αυτά που συναντώνται συχνότερα.

Γενικεύοντας, τα όρια ελέγχου στα διαγράμματα προκαθορίζονται από το μέσο όρο της μεταβλητής που παρίσταται στην κεντρική ευθεία, προσθέτοντας (ανώτερο όριο) και αφαιρώντας (κατώτερο όριο) το γινόμενο 3σ (διαιρούμενο πάντα με την τετραγωνική ρίζα του n). Τα όρια 3σ καλούνται και όρια *δράσης* ή όρια *επέμβασης* ή όρια *κινδύνου* και όταν η γραμμή ελέγχου τέμνει τα όρια αυτά, τότε η συγκεκριμένη διεργασία κρίνεται εκτός ελέγχου και διακόπτεται η παραγωγή μέχρι να αναζητηθεί η αιτία της βλάβης. Τα όρια 2σ καλούνται όρια *προειδοποίησης* ή όρια *επιφυλακής* ή όρια *προσοχής* και αν η γραμμή ελέγχου τα τμήσει, τότε απαιτείται να ληφθούν μέτρα για τη διόρθωσή της, για την επαναφορά δηλαδή της πορείας της γραμμής ελέγχου μέσα στα όρια 2σ. Όταν η γραμμή ελέγχου παρουσιάζει ομαλή μεταβλητότητα σε όλο το μήκος του διαγράμματος ελέγχου, τότε η διαδικασία είναι υπό έλεγχο.

Η διαίρεση της τυπικής απόκλισης με τον όρο $\sqrt{n}$ δεν εφαρμόζεται πάντοτε για την απεικόνιση των ορίων ελέγχου, αντ' αυτού χρησιμοποιούνται ειδικοί πίνακες όπως αναφέρεται παρακάτω. Έτσι, η χρήση του όρου σ της τυπικής απόκλισης θέλει ιδιαίτερη προσοχή, γιατί σε μερικές περιπτώσεις υπονοείται άλλοτε το τυπικό σφάλμα ($\sigma\sqrt{n}$) και άλλοτε η τυπική απόκλιση σ.

Πλεονεκτήματα και μειονεκτήματα των δύο ομάδων διαγραμμάτων ελέγχου

Στον ποιοτικό έλεγχο μας δίνεται συχνά η ευκαιρία να επιλέξουμε προς έλεγχο στοιχεία που αποτελούν μετρήσεις ποσοτικών μεταβλητών, ασυνεχών ή συνεχών και καλό θα ήταν να γνωρίζουμε τα υπέρ και κατά των δύο τύπων διαγραμμάτων που δίνουν:

- Τα *διαγράμματα ελέγχου χαρακτηρισμών* (διαγράμματα p) επιτρέπουν την ταχεία συλλογή αποτελεσμάτων από διάφορα σημεία του ποιοτικού ελέγχου, δηλαδή ο ελεγκτής μπορεί να ταξινομήσει τις μονάδες ως αποδεκτές ή απορριπτέες (και κατά συνέπεια και όλη την παρτίδα) με βάση διάφορα κριτήρια ποιότητας. Τα *διαγράμματα ελέγχου απαριθμήσεων ελαττωμάτων* (διαγράμματα c), στοχεύουν στο διαχωρισμό όλων των ασύμβατων υλικών σε περισσότερες κατηγορίες, διότι προσβλέπουν στην απλή καταγραφή των ελαττωματικών μονάδων συγκεκριμένου τύπου ελαττώματος και περαιτέρω απεικόνισή τους. Τα παραπάνω επιτυγχάνονται χωρίς την ανάγκη προμήθειας ακριβών συσκευών ελέγχου και χωρίς την ανάγκη συλλογής ποσοτικών μετρήσεων με χρονοβόρες διαδικασίες οι οποίες αποτελούν γνωρίσματα των διαγραμμάτων συνεχών μεταβλητών. Η κυριότερη όμως διάσταση είναι ότι τα διαγράμματα ασυνεχών μεταβλητών γίνονται εύκολα κατανοητά από το διευθυντή παραγωγής και άλλους ιθύνοντες, γι' αυτό και αυτά αποτελούν πειστικότερες μαρτυρίες των προβλημάτων ποιότητας.

- Τα *διαγράμματα ελέγχου συνεχών μεταβλητών* είναι σαφώς πιο ευαίσθητα από τα διαγράμματα των ασυνεχών μεταβλητών. Τα διαγράμματα αυτά μπορούν να προειδοποιήσουν για την ύπαρξη προβλημάτων ποιότητας στο στάδιο γένεσης τους, προτού δηλαδή τα προϊόντα κριθούν απορριπτέα, όπως αναπόφευκτα συνήθως συμβαίνει με τα διαγράμματα ελέγχου χαρακτηρισμών. Έτσι, προλαβαίνουμε να πάρουμε τα κατάλληλα μέτρα διόρθωσης του προβλήματος. Για το λόγο αυτό, τα διαγράμματα συνεχών μεταβλητών θεωρούνται δικαίως ως δείκτες εντοπισμού προβλημάτων, οι οποίοι θα σημάνουν έγκαιρα συναγερμό προτού ο αριθμός των ελαττωματικών μονάδων αυξηθεί υπερβολικά κατά τη διεργασία της παραγωγικής διαδικασίας.

Ανισομεγέθη δείγματα ελέγχου

Τα όρια ελέγχου κατά μήκος του διαγράμματος ελέγχου παραμένουν σταθερά, μόνον όταν όλα τα δείγματα έχουν ίσο αριθμό παρατηρήσεων. Σε αντίθετη περίπτωση, η γραμμικότητα των ορίων ελέγχου παύει να υφίσταται και για τη ρύθμιση του προβλήματος αυτού αναφέρονται τρεις τρόποι:

1. Θεώρηση ενός μέσου δειγματοληπτικού μεγέθους. Αν επιμένουμε στη διατήρηση της οριζόντιας γραμμικότητας των ορίων ελέγχου και ο λόγος που το επιδιώκουμε είναι η πληρέστερη κατανόηση του μελετώμενου διαγράμματος, τότε υπολογίζουμε το μέσο δειγματοληπτικό όρο $\bar{n}$ που προκύπτει αθροίζοντας όλα τα δειγματοληπτικά μεγέθη των δειγμάτων και διαιρώντας με το πλήθος αυτών. Η ρύθμιση αυτή δεν θεωρείται επακριβής, αυξάνει όμως σημαντικά την αξιοπιστία της, όταν τα δειγματοληπτικά μεγέθη αυξομειώνονται ελάχιστα μεταξύ τους.
2. Μεταβαλλόμενο όριο ελέγχου. Στην περίπτωση αυτή υπολογίζουμε ατομικά, για κάθε δείγμα, τα αντίστοιχα όρια ελέγχου βασιζόμενοι στο εκάστοτε μέγεθος δείγματος n. Το αποτέλεσμα αυτών των ενεργειών διαφαίνεται στο διάγραμμα ελέγχου από τα μεταβλητά όρια ελέγχου που παίρνουν τώρα το σχήμα τεθλασμένης γραμμής (με μορφή σκαλοπατιών, βλέπε σχήμα στο παράδειγμα 11.5). Το αντίτιμο για την ακρίβεια της μεταβολής των ορίων ελέγχου που επιτυγχάνουμε με τη ρύθμιση αυτή είναι η σχετικώς αυξημένη δυσκολία στην εξαγωγή συμπερασμάτων από την πορεία της γραμμής ελέγχου μέσα στα μεταβαλλόμενα όρια ελέγχου.
3. Τυποποίηση των ορίων ελέγχου. Αυτή πραγματοποιείται αφαιρώντας κάθε τιμή του διαγράμματος από τη μέση τιμή της μεταβλητής (που μπορεί να αφορά μέσο όρο, μέσο εύρος, μέση τυπική απόκλιση) και διαιρώντας τη διαφορά με την τυπική απόκλιση κάθε δείγματος. Τα όρια ελέγχου, με τη ρύθμιση αυτή, κείνται επί ευθείας γραμμής, αλλά η θέση των σημείων των δειγμάτων στο διάγραμμα εξαρτάται τώρα εκτός από τη μελετώμενη παράμετρο (μέσος όρος, εύρος κτλ.) και από το δειγματοληπτικό μέγεθος κάθε δείγματος. Το μειονέκτημα της ρύθμισης αυτής εντοπίζεται στον κάθετο άξονα του γραφήματος, οι μονάδες του οποίου είναι μονάδες τυπικής απόκλισης της εξεταζόμενης παραμέτρου π.χ. σε διάγραμμα εύρους της μεταβλητής και τιμή ενός σημείου ίση με 1.2 σημαίνει ότι το σημείο αυτό απέχει 1.2σ φορές από τα όρια ελέγχου.

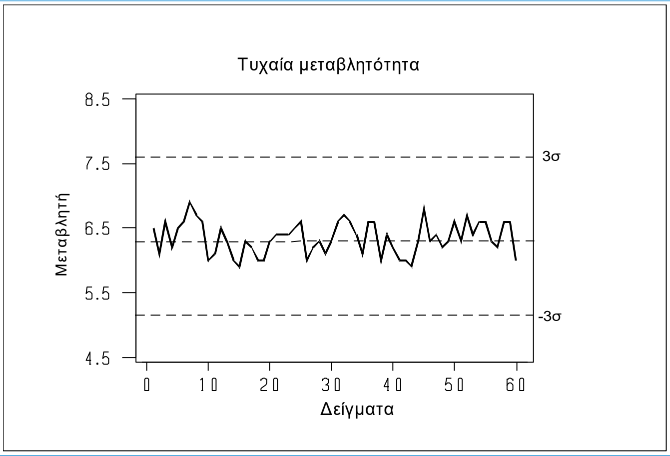
Πηγές μεταβλητότητας κατά τη διαδικασία παραγωγής των μονάδων

Ένας μεγάλος αριθμός αιτίων γένεσης προβλημάτων που απαντώνται κατά τη διαδικασία παραγωγής, μπορούν να προκαλέσουν ισχυρές αποκλίσεις του επεξεργαζόμενου προϊόντος από τις βασικές προδιαγραφές. Τα αίτια αυτά εμφανίζονται κυρίως στις ποσοτικές μετρήσεις των μονάδων και η γένεση τους οφείλεται σε μεταβλητότητα ποικίλης προέλευσης:

1. Ενδογενής μεταβλητότητα. Παρατηρείται στις μονάδες που παρασκευάζονται κάτω από τις ίδιες συνθήκες και προδιαγραφές. Αυτή τη μεταβλητότητα δεν μπορούμε να διορθώσουμε παρεμβατικά.
2. Εξωτερικά αίτια, όπως θερμοκρασία και υγρασία περιβάλλοντος.
3. Αίτια που πηγάζουν μέσα από τη διαδικασία παραγωγής, όπως συσσώρευση άχρηστων υποπροϊόντων, φθορά του καταλύτη, ασταθής απόδοση της θερμαντικής πηγής κ.ά.
4. Καθορισμένα ή ευδιάκριτα αίτια. Αυτά κυρίως οφείλονται στη μεταβαλλόμενη ποιότητα της παρτίδας, εξαιτίας ασταθούς ποιότητας της εισαγόμενης πρώτης ύλης, όπως επίσης στην εσφαλμένη ή ατελή προετοιμασία του εργαστηριακού εξοπλισμού.

Οι περιπτώσεις 3 και 4 συνοψίζονται ως *ειδικά αίτια* (special causes).

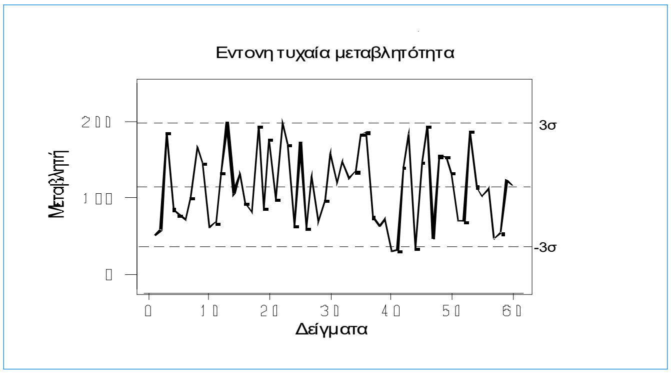
Με την εφαρμογή του στατιστικού ελέγχου της διαδικασίας παραγωγής προσπαθούμε να κάνουμε διακριτό το διαχωρισμό της μεταβλητότητας που οφείλεται σε αναγνωρίσιμα αίτια από τα υπόλοιπα (πλην της ενδογενούς μεταβλητότητας που δεν ανιχνεύεται). Έτσι, τα διαγράμματα ελέγχου $\bar{X}$, κατά κύριο λόγο, ανάλογα με τη φύση της μεταβολής της γραμμής ελέγχου, μπορούν να μας πληροφορήσουν για την προέλευση των αιτίων γένεσης της μεταβλητότητας:



Σχ. 11.2. Διάγραμμα ελέγχου με απλή τυχαία μεταβλητότητα. Η μεταβλητότητα αυτή θεωρείται φυσιολογική και αποτελεί μέτρο αναφοράς μιας ιδανικής διεργασίας.

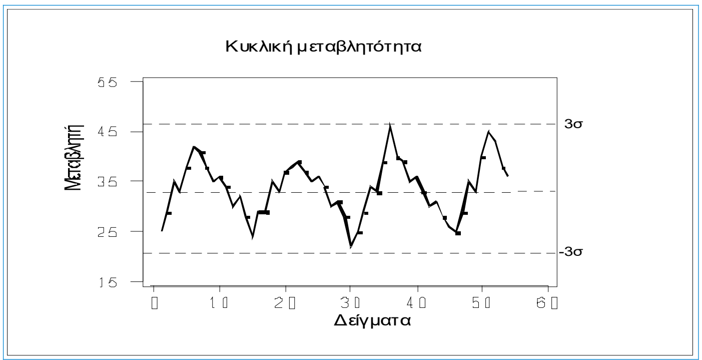
1. Απλή τυχαία μεταβλητότητα (Σχ. 11.2) που οφείλεται στη διασπορά των μέσων όρων των δειγμάτων μέσα στο διάγραμμα ελέγχου (ενδογενής μεταβλητότητα). Αυτή η διασπορά μπορεί να θεωρηθεί ότι προέρχεται από ένα συνολικό δείγμα μέσων όρων που έχει κανονική κατανομή με μέσο όρο μ και τυπική απόκλιση σ. Αναμένεται επίσης μία φυσιολογική διασπορά των μέσων όρων στο διάγραμμα ελέγχου, τέτοια που να κυμαίνεται το πολύ μεταξύ των ορίων που διαμορφώνονται από το μέγεθος τυπικής απόκλισης ίσης με ±3σ (όρια επέμβασης). Σημεία πέρα των ορίων αυτών δηλώνουν την ύπαρξη αναγνωρίσιμης μεταβλητότητας, καθότι τα διαγράμματα ελέγχου προϋποθέτουν την ύπαρξη αποκλειστικά και μόνο της απλής τυχαίας μεταβλητότητας μέσα στα όρια μεταβολής ±3σ.

2. Τυχαία αυξημένης έντασης μεταβλητότητα των σημείων (Σχ. 11.3). Αυτή παρατηρείται εξαιτίας της αυξημένης μεταβλητότητας της μέσης τιμής της μεταβλητής που οφείλεται στην έντονη διακύμανση μέσα στα δείγματα, οι τιμές δηλαδή των δειγμάτων εμφανίζουν μεγάλη διασπορά. Το πρόβλημα αυτό περιορίζεται αυξάνοντας το δειγματοληπτικό μέγεθος στα δείγματα και μειώνοντας έτσι τη μεταβλητότητα που αναπτύσσεται μέσα στα δείγματα. Η αύξηση του δειγματοληπτικού μεγέθους δεν επηρεάζει την απλή τυχαία μεταβλητότητα που αναπτύσσεται μεταξύ των μέσων όρων των δειγμάτων.



Σχ. 11.3. Διάγραμμα ελέγχου με έντονη μεταβλητότητα της γραμμής ελέγχου που οφείλεται στην αυξημένη διασπορά των επαναληπτικών τιμών μέσα σε κάθε δείγμα.

Έντονη μεταβλητότητα παρατηρείται συνήθως στην υπερβολική ρύθμιση ή στην προσπάθεια προσαρμογής των μετρήσεων μιας διεργασίας. Όταν ο χειριστής ρυθμίζει την τιμή του pH ενός διαλύματος προς τα κάτω, επειδή απλώς μερικές μετρήσεις βρέθηκαν να δίνουν τιμές πάνω από τη μέση επιδιωκόμενη τιμή (στόχος), τότε πολύ πιθανόν, αναμένεται να πέσει σύντομα η τιμή του pH κάτω του κατώτερου ορίου τιμών που έχει καθοριστεί. Η συγχώνευση επίσης δύο ή περισσοτέρων διαδικασιών παραγωγής ίδιας φάσης στο ίδιο διάγραμμα ελέγχου R μπορεί να οδηγήσει το διάγραμμα εκτός ορίων. Αυτό συμβαίνει όταν μία διεργασία παράγει υλικό σε σακούλες με μέγιστο μήκος του προϊόντος και μία δεύτερη παράγει υλικό σε σακούλες με ελάχιστο μήκος. Στα διαγράμματα R και p η δυσλειτουργία των συσκευών και η διαλείπουσα χρήση μεταβαλλόμενης πρώτης ύλης μπορεί να προκαλέσει έντονη μεταβλητότητα της διασποράς των μέσων όρων. Παράλληλα, η ανορθόδοξη εξέταση των μονάδων παραγωγής μπορεί να προκαλέσει επίσης σημαντικές

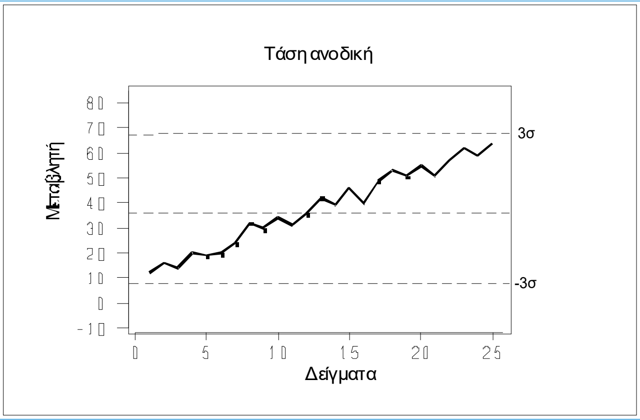


Σχ. 11.4. Διάγραμμα ελέγχου με συνεχή περιοδικότητα μεταβολής της γραμμής ελέγχου.

διακυμάνσεις στο διάγραμμα ελέγχου.

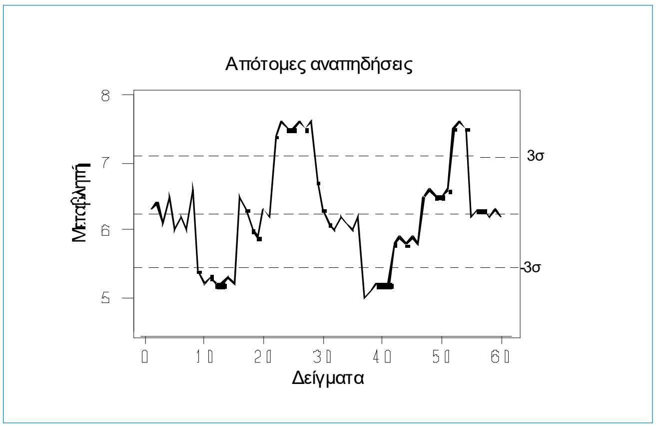
3. Περιοδική μεταβλητότητα (Σχ. 11.4). Η περιοδική επαναληπτικότητα της γραμμής ελέγχου εκφράζεται με κυκλική συνήθως μορφή και μπορεί να οφείλεται στη λανθασμένη περιοδικότητα της περιστροφής των συσκευών, σε μεταβολές του περιβάλλοντος, όπως η θερμοκρασία και η υγρασία ή σε κόπωση του προσωπικού. Επίσης, στην εκ νέου επεξεργασία ήδη κατεργασμένου υλικού στη γραμμή παραγωγής.

Στα διαγράμματα R η περιοδικότητα μπορεί να οφείλεται στην ανεπαρκή συντήρηση των συσκευών με αποτέλεσμα τη διαλείπουσα λειτουργία τους. Στα διαγράμματα p η περιοδικότητα οφείλεται στη συγχώνευση δύο διαδικασιών. Κυκλικές μεταβολές προκαλεί και η ανορθόδοξα σχεδιασμένη δειγματοληψία. Σε φυσιολογικές συνθήκες, η περιοδικότητα σχετίζεται με δραστηριότητες που εμφανίζονται επαναληπτικά ανά συγκεκριμένες μέρες ή ώρες. Μη ομοιόμορφη περιοδικότητα παρατηρείται περιστασιακά, όταν διακόπτεται η λειτουργία των συσκευών ελέγχου, με σκοπό τη ρύθμιση ή επισκευή τους. Τα διαγράμματα $\bar{X}$ και p είναι περισσότερο ευαίσθητα στην εμφάνιση κυκλικών μεταβολών απ' ότι τα διαγράμματα R.



Σχ. 11.5. Διάγραμμα ελέγχου με βαθμιαία αυξητική τάση μεταβολής της γραμμής ελέγχου.

4. Εμφανείς τάσεις βαθμιαίας μεταβολής αυξητικές ή φθίνουσες (Σχ. 11.5). Η μεταβλητότητα αυτή οφείλεται συνήθως στη σταδιακή επιδείνωση λειτουργίας των συσκευών, στη φθορά του καταλύτη μιας διεργασίας, στη συσσώρευση άχρηστων υλικών κτλ. Επίσης, στις συνεχείς αλλαγές του περιβάλλοντος, στην εξάντληση των αποθεμάτων της πρώτης ύλης και αντικατάστασή της με νέα ή στην αλλαγή του ρυθμού παραγωγής. Στα διαγράμματα R η τάση οφείλεται στη σταδιακή βελτίωση ή μείωση της απόδοσης του προσωπικού και στην αλλαγή του ρυθμού παραγωγής. Τα διαγράμματα p επηρεάζονται με τον ίδιο τρόπο.



Σχ. 11.6. Διάγραμμα ελέγχου με ξαφνικές εκτινάξεις της γραμμής ελέγχου και εκτός των ορίων επέμβασης.

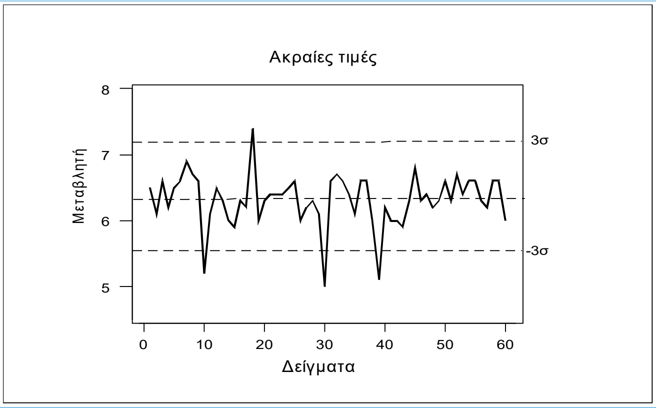
5. Απότομες εκτινάξεις ή αναπηδήσεις (Σχ. 11.6). Η μεταβλητότητα αυτή οφείλεται σε παρτίδα με νέο υλικό, σε αλλαγές του προσωπικού ή συσκευών ή επίσης σε τροποποιήσεις στο πρόγραμμα επεξεργασίας.

6. Ανορθόδοξη ομοιομορφία μεταβολής (Σχ. 11.7). Αυτή αναγνωρίζεται εύκολα από την περιορισμένη διασπορά των παρατηρήσεων στο διάγραμμα ελέγχου στο οποίο η κεντρική ευθεία παριστάνεται από τον στόχο T της διεργασίας και όχι από τη μέση τιμή μ της διεργασίας. Παρατηρείται σε περιπτώσεις που η μονάδα στελεχώνεται από καλά εκπαιδευμένο προσωπικό, όταν η διεργασία αλλάζει εσκεμμένα ή μη, όταν κάποια φάση της διεργασίας παραλείπεται, όταν η πρώτη ύλη είναι απαλλαγμένη από ελαττωματικές μονάδες και τέλος όταν χρησιμοποιούνται στο διάγραμμα λανθασμένα όρια ελέγχου. Μη κανονική ομοιομορφία διασποράς παρατηρείται επίσης και όταν μη τυχαία δείγματα συλλέγονται από προσωπικό που δείχνει είτε υπερβάλλοντα ζήλο είτε έλλειψη εκπαίδευσης. Δείγματα που ελέγχονται σε δυσλειτουργούντα όργανα (π.χ. σκόνη στο φακό του χρωματόμετρου, αρρύθμιστος ζυγός) ή δείγματα που υποβάλλονται σε χημική/βιολογική κατεργασία με αντιδραστήρια παρωχημένης ημε-

Σχ. 11.7. Διάγραμμα ελέγχου με ανορθόδοξη ομοιομορφία γιατί η γραμμή ελέγχου μεταβάλλεται μόνο πάνω από την κεντρική ευθεία που εδώ εκλαμβάνεται ως στόχος της διεργασίας T.

ρομηνίας λήξης, μπορούν να εκδηλώσουν ανορθόδοξη ομοιομορφία διασποράς. Τα διαγράμματα R θεωρούνται τα πλέον ευαίσθητα στην αναγνώριση της ομοιομορφίας μεταβολής.

7. Παρουσία εξαιρετικά απομακρυσμένων τιμών στο διάγραμμα από τη μάζα των υπολοίπων (Σχ. 11.8). Η μεταβλητότητα αυτή εκδηλώνεται περιστασιακά, όταν η γραμμή ελέγχου εξέρχεται των ορίων ελέγχου σε μεγάλο βαθμό, χωρίς εμφανή αίτια, ως απόρροια συνήθως λανθασμένης διενέργειας ποιοτικού ελέγχου. Τέτοια ποιοτικά σφάλματα οφείλονται σε κακούς μαθηματικούς υπολογισμούς, στο ελλιπές μέγεθος του δείγματος, στην εσφαλμένη διαδικασία δειγματοληψίας, στη μόλυνση ευπαθών δειγμάτων, σε προβλήματα της γραμμής παραγωγής που οφείλονται στη διακοπή καυσίμου ή ηλεκτρικού ρεύματος ή και στη διακοπή μιας φάσης της διεργασίας.



Σχ. 11.8. Διάγραμμα ελέγχου με 4 εξαιρετικά ακραία σημεία, κείμενα εκτός των ορίων επέμβασης.

Μέθοδοι εκτίμησης της μεταβλητότητας στα διαγράμματα ελέγχου

Τα διαγράμματα ελέγχου, στην ουσία, εξετάζουν την τυχαία μεταβλητότητα που αναπτύσσεται μέσα στις ομάδες των παρατηρήσεων (δείγματα), δηλαδή μεταξύ των τιμών σε κάθε δείγμα και επίσης μεταξύ των ομάδων παρατηρήσεων, δηλαδή μεταξύ των μέσων όρων. Οι δύο αυτές ποσότητες μεταβλητότητας είναι γνωστές από την ανάλυση της διακύμανσης (κεφάλαιο 7). Επομένως, για τον προσδιορισμό της μεταβλητότητας αυτών απαιτείται η στατιστική εκτίμηση του μέσου όρου της μεταβλητής της διεργασίας (ή απλώς της μέσης διεργασίας) και των τυπικών αποκλίσεων μέσα στις ομάδες και μεταξύ των ομάδων. Αυτά βέβαια ισχύουν στις περιπτώσεις που τα συλλεγόμενα στοιχεία προέρχονται από δείγματα και όχι από ατομικές παρατηρήσεις, μία τη φορά για κάθε δείγμα.

1. Σε ομαδοποιημένα στοιχεία, δύο μέθοδοι έχουν αναπτυχθεί για το σκοπό αυτό:

● Εκτίμηση των παραμέτρων της μεταβλητότητας της μέσης τιμής της μεταβλητής με τη *μέθοδο της τυπικής απόκλισης σ*:

Συλλέγουμε k δείγματα (τουλάχιστον 20) αποτελούμενα από n παρατηρήσεις το καθένα και υπολογίζουμε το μέσο όρο X_i σε κάθε δείγμα (ομάδα):

$$\bar{X}_i = \frac{\sum_{j=1}^{n} X_{ij}}{n}$$, όπου X_{ij} είναι η τιμή της παρατήρησης j στη i ομάδα.

Υπολογίζουμε τη διακύμανση s^2_i της μεταβλητότητας που αναπτύσσεται μέσα σε κάθε ομάδα,

$$s_i^2 = \frac{\sum_{j=1}^{n} X_{ij}^2 - \frac{(\sum_{j=1}^{n} X_{ij})^2}{n}}{n-1} \qquad \text{είτε} \qquad s_i^2 = \frac{\sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2}{n-1} .$$

Υπολογίζουμε την ολική τυπική απόκλιση s_w που αφορά την ολική διασπορά μέσα στις ομάδες,

$$s_w = \sqrt{\frac{\sum_{i=1}^{k} s_i^2}{k}} = \sqrt{\frac{s_1^2 + s_2^2 + \ldots + s_k^2}{k}} .$$

Αν τα δείγματα είναι ανισομεγέθη, τότε χρησιμοποιούμε τον τύπο,

15
Ανάλυση της σταθερότητας τροφίμων

Με τον όρο σταθερότητα (stability) νοείται η δυνατότητα ενός προϊόντος να συμμορφώνεται με συγκεκριμένες προδιαγραφές οι οποίες προσδιορίζουν την ταυτότητα, ποιότητα και γνησιότητά του. Έχει ως αντικειμενικό στόχο την περιγραφή και προσδιορισμό της πορείας αποδόμησης του προϊόντος διαχρονικά και την εκτίμηση της ημερομηνίας λήξης του. Έπεται, ότι ο χρόνος ζωής γενικώς ενός υλικού ορίζεται από το διάστημα μέσα στο οποίο το υλικό αναμένεται να παραμείνει εντός των ορίων συγκεκριμένων προδιαγραφών μετά την παρασκευή του. Στην πράξη, στα περισσότερα τρόφιμα μεσολαβεί μεγάλο χρονικό διάστημα μέχρι να εμφανίσουν φαινόμενα αποδόμησης, όπως απώλεια της συνοχής, αποσύνθεση και γενικώς οτιδήποτε υποβαθμίζει την ποιότητά τους, και για το λόγο αυτό σχεδιάζονται ειδικές πειραματικές δοκιμές οι οποίες επιταχύνουν ή επιβραδύνουν τη διεργασία αποδόμησης, γνωστές με τον όρο **επιταχυνόμενες δοκιμές** (accelerated tests) ή **δοκιμές παλαίωσης**. Οι δοκιμές αυτές επισπεύδουν τους ρυθμούς εξέλιξης των χημικών ή φυσικών διεργασιών υποβάλλοντας τα τρόφιμα σε ακραίες συνθήκες συντήρησης με σκοπό τη δημιουργία ενός προσωρινού χρόνου ζωής, έτσι ώστε να επακολουθήσει μία μακράς διάρκειας μελέτη σταθερότητας του προϊόντος κάτω από φυσιολογικές συνθήκες.

Διεθνώς, ορίζονται διάφορα κριτήρια των αποδεκτών ορίων σταθερότητας, αναλόγως της φύσης των προϊόντων που σχετίζονται με χημικά, φυσικά, μικροβιολογικά ή και τοξικολογικά χαρακτηριστικά π.χ. δοκιμές σταθερότητας ως προς την εμφάνιση, ευθρυπτότητα, σκληρότητα, οσμή, υγρότητα και συνεκτικότητα μεταβάλλοντας κατάλληλα τα επίπεδα θερμοκρασίας, πίεσης, δόνησης, δόσης χημικής ένωσης, μικροβιολογικής δράσης (ολική μεσόφιλη χλωρίδα) κτλ.

Η σταθερότητα ή καλύτερα η διάρκεια στη συντήρηση ενός προϊόντος παρέχει πληροφόρηση στην ετικέτα του με δύο διαφορετικές σημάνσεις. Στα ευπαθή προϊόντα κυρίως από μικροβιολογικής σκοπιάς, αναγράφεται η υπόδειξη "κατανάλωση μέχρι" (use by), η οποία συνδέεται άμεσα με την ασφάλεια του τροφίμου. Σε προϊόντα λιγότερο ευπαθή, όπως εκείνα που αφυδατώθηκαν ή ψύχθηκαν, αναγράφεται η ένδειξη "κατανάλωση πριν από" μία ημερομηνία λήξης (best before), και προσαρμόζεται ανάλογα με το είδος της διεργασίας του τροφίμου: μερικές μέρες για ψημένα προϊόντα, πάνω από ένα χρόνο για κονσερβοποιημένα, αποξηραμένα ή κατεψυγμένα.

Οι κυριότεροι παράγοντες που επηρεάζουν δραστικά τη διάρκεια ζωής των τροφίμων θεωρούνται οι ακόλουθοι:

1) Η πρώτη ύλη. Υλικά που προστίθενται σε έτοιμα προϊόντα (π.χ. μουστάρδα σε ψωμάκι σάντουιτς), πρώτη ύλη που υφίσταται κάποια διεργασία όπως βρασμό ή αλλαγές στις συνθήκες συντήρησης ή κλείνεται σε συσκευασία με τροποποιημένη ατμόσφαιρα που μπορεί να επηρεάσει το χρόνο ζωής είτε επιβραδύνοντας είτε επιμηκύνοντάς τον.
2) Σύσταση του προϊόντος. Διάφορα πρόσθετα που περιγράφονται στη συνταγή του τροφίμου μπορούν να αλληλεπιδράσουν αρνητικά και να μειώσουν το προσδόκιμο του χρόνου ζωής ή να προκαλέσουν οξείδωση και αποχρωματισμό επηρεάζοντας την αποδοχή του από τους καταναλωτές.
3) Τύπος συσκευασίας. Η χρήση τροποποιημένης ατμόσφαιρας ή η συσκευασία κενού αέρος αυξάνει τη διατηρησιμότητα λαμβάνοντας παράλληλα σοβαρά υπόψη και τις εφοδιαστικές αλυσίδες τροφοδοσίας αλλά και τους ενδιάμεσους χειρισμούς μέχρι την τελική αποθήκευση του τροφίμου σε χώρους συντήρησής του.
4) Η θερμοκρασία έκθεσης του τροφίμου η οποία επηρεάζει δραματικά το χρόνο ζωής ως προς την ασφάλεια και την ποιότητά του.
5) Τήρηση των κανόνων υγιεινής, όπως π.χ. στους χώρους χρήσης του προϊόντος ή την αποθήκευσή του και τήρηση των κανόνων ασφαλείας προς αποφυγή μικροβιακής ανάπτυξης ή δημιουργίας εστιών μόλυνσης.
6) Προσδόκιμο ζωής μετά την αποσυσκευασία του προϊόντος. Ειδικές οδηγίες στην ετικέτα προσδιορίζουν τον ασφαλή χρόνο περαιτέρω συντήρησης του τροφίμου αν δεν καταναλωθεί άμεσα.

Πέραν της στατιστικής ανάλυσης και της προβλεπτικής εφαρμογής των εξισώσεων παλινδρόμησης για τη διατηρησιμότητα του προϊόντος, τίθενται σε υψηλή προτεραιότητα και τα κριτήρια διασφάλισης της ποιότητάς του. Επαναληπτικοί οργανοληπτικοί έλεγχοι και μικροβιολογικές αναλύσεις θα επιβεβαιώσουν ότι το προϊόν είναι ασφαλές και έτοιμο προς κατανάλωση. Αναλόγως της ευπάθειας του τροφίμου εφαρμόζονται και ειδικοί μικροβιολογικοί έλεγχοι για την ανίχνευση επικίνδυνων για τη δημόσια υγεία βακτηριδίων (π.χ. *Listeria* spp., *Clostridium* spp.).

Ένα γενικώς αποδεκτό πλάνο εργασίας προτείνεται για τα στάδια που συνδέουν και καθορίζουν τη διάρκεια ζωής του προϊόντος:

1) Καθορισμός της συνταγής του τροφίμου και των συνθηκών αποθήκευσης/συντήρησης.
2) Ταυτοποίηση όλων των αιτίων που καθιστούν το τρόφιμο μη αποδεκτό ή μη ασφαλές.
3) Πρόβλεψη του χρόνου ζωής με τη βοήθεια των στατιστικών μοντέλων.
4) Έλεγχος της ασφάλειας του τροφίμου και της οργανοληπτικής ποιότητας στο τέλος του προβλεπόμενου χρόνου ζωής.
5) Οριστική καταγραφή του χρόνου ζωής με όλους τους παράγοντες υπό έλεγχο.
6) Επανεξέταση του πλάνου για τον εντοπισμό εάν και πότε ένα συστατικό ή μία διεργασία διαφοροποιείται.

Μία γενικώς αποδεκτή διάρκεια ζωής θεωρείται εκείνη που επιτρέπει τη διατήρηση των επιθυμητών οργανοληπτικών, χημικών, μηχανικών, φυσικών και μικροβιολογικών χαρακτηριστικών, τα οποία ορίζονται ως παράμετροι της λήξης του χρόνου ζωής (EOSL-End of Shelf Life parameters). Κάθε τρόφιμο διακρίνεται από τις δικές του ιδιότητες και επομένως από συγκεκριμένα όρια διακύμανσής τους. Η συνηθισμένη διαδικασία προσδιορισμού του χρόνου ζωής περιλαμβάνει τη μελέτη επιλεγμένων συνθηκών για μία περίοδο επιμηκέστερη του αναμενόμενου χρόνου ζωής και έλεγχό τους σε τακτά χρονικά διαστήματα για να εντοπισθεί ή έναρξη των αλλοιώσεων στο τρόφιμο. Ως επιλεγμένες συνθήκες λαμβάνονται συνήθως κάποιες ευαίσθητες φάσεις της διαδικασίας παρασκευής, ο τύπος της συσκευασίας και ο τρόπος αποθήκευσης /συντήρησης.

Αναφορικά με τις μελέτες εμπορικής σταθερότητας, αυτές θα πρέπει να διασφαλίζουν ότι το τρόφιμο στην αγορά συμμορφώνεται με τις δηλωμένες προδιαγραφές μέχρι τη λήξη της διάρκειας ζωής του. Είναι δηλαδή σημαντικό να διαπιστωθεί ή και να πιστοποιηθεί ότι ο πραγματικός χρόνος ζωής είναι μεγαλύτερος από τον αναγραφόμενο στην ετικέτα. Επομένως, τρία μείζονα πεδία ενδιαφέροντος θα πρέπει να διερευνηθούν:

1) Αν ο εκτιμώμενος χρόνος ζωής από τη μελέτη εμπορικής σταθερότητας στα πλαίσια της 95% στατιστικής σημαντικότητας δεν είναι μεγαλύτερος (δεν διαφέρει δηλαδή) από τον αναγραφόμενο σε ένα τρόφιμο, το προϊόν θεωρείται ότι δεν τηρεί τις προδιαγραφές λήξης της διάρκειας ζωής του και θα πρέπει να ανακληθεί.
2) Αν ο εκτιμώμενος χρόνος είναι μικρότερος από τον αναγραφόμενο, αποτελεί σοβαρή ένδειξη κάποιας μεταβολής στις διεργασίες παρασκευής του με κίνδυνο να θεωρηθεί ακόμα και ως διεργασία εκτός ελέγχου.
3) Αν ο εκτιμώμενος χρόνος ζωής είναι μακρύτερος του αναγραφόμενου, τότε αναμένεται, συνήθως, μία αλλαγή (νέα προσαρμογή) στο χρόνο λήξης που δηλώνεται στην ετικέτα προς όφελος του παραγωγού ή μεταποιητή.

Μοντέλα περιγραφής της μελέτης σταθερότητας

Σε γενικές γραμμές η αποδόμηση ενός τροφίμου διενεργείται με τη συλλογή και μελέτη πλήθους μεταβλητών η καθεμία των οποίων εκφράζεται σε συνάρτηση με την πάροδο του χρόνου συντήρησης και με το γραμμικό μοντέλο, $\hat{Y}_i = \alpha + \beta X_i$, όπου $\hat{Y}_i$ η αναμενόμενη απόκριση της μεταβλητής σε χρόνο X_i, α η δράση της παρτίδας του παραγόμενου τροφίμου ή αλλιώς της μεταχείρισης και β ο ρυθμός επιδείνωσης της ποιότητας της μεταβλητής. Η μεταβλητή μετρείται ως % μεταβολή διαχρονικά, φθίνουσα ή αύξουσα, από την αρχική τιμή της και συνεπώς η απώλεια σταθερότητας με την παρέλευση του χρόνου ορίζεται ως βX.

Τρία γραμμικά μοντέλα αποδόμησης (υποβάθμισης της ποιότητας) επιλέγονται συνήθως για την εκτίμηση του χρόνου λήξης του τροφίμου:

- μοντέλο I, αποδέχεται διαφορετικές κλίσεις της εξεταζόμενης μεταβλητής και διαφορετικά ύψη των κλίσεων στις παρτίδες της μελέτης,
- μοντέλο II, αποδέχεται κοινή κλίση και διαφορετικά ύψη κλίσεων στις παρτίδες,
- μοντέλο III, αποδέχεται κοινή κλίση και κοινή παράμετρο α (ύψος κλίσης) ως αντιπροσωπευτικές όλων των παρτίδων.

Το προτεινόμενο τελικά μοντέλο προσδιορίζεται σύμφωνα με την παρακάτω ακολουθία ενεργειών:

- Εκκίνηση με προσαρμογή του μοντέλου I, με τη δράση δηλαδή αρχικά του χρόνου, ακολούθως με τη δράση της παρτίδας και στο τέλος με τη δράση της αλληλεπίδρασής τους. Είναι εφαρμογή ουσιαστικά της ανάλυσης διακύμανσης και έλεγχος της σημαντικότητας των κλίσεων.
- Αν η ακριβής πιθανότητα σφάλματος της αλληλεπίδρασης βρεθεί $p<0.25$, οι κλίσεις θεωρείται ότι διαφέρουν στατιστικά σημαντικά μεταξύ των παρτίδων. Η διαδικασία περατώνεται και το μοντέλο I χρησιμοποιείται για την εκτίμηση του χρόνου λήξης.
- Αν η πιθανότητα της αλληλεπίδρασης υπολογιστεί ότι είναι $p\geq 0.25$, οι κλίσεις θεωρείται ότι είναι κοινές (ίσες) και προβαίνουμε στην εφαρμογή του μοντέλου II εξετάζοντας πλέον την ισότητα των υψών των κλίσεων.
- Αν η πιθανότητα σφάλματος της παρτίδας ισούται με $p<0.25$, οι παράμετροι α παραδεχόμαστε ότι

διαφέρουν μεταξύ των παρτίδων και το μοντέλο II χρησιμοποιείται για την εκτίμηση της ημερομηνίας λήξης.

- Αν η πιθανότητα της παρτίδας είναι $p \geq 0.25$, τα ύψη των κλίσεων συμπεραίνουμε ότι είναι κοινά (ίσα) μεταξύ των παρτίδων και το μοντέλο III χρησιμοποιείται για την εκτίμηση του χρόνου λήξης.

Όταν το μοντέλο I (διαφορετικές κλίσεις και ύψη των κλίσεων) χρησιμοποιείται για την εκτίμηση του χρόνου λήξης, η ποσότητα EMS δεν μοιράζεται από κοινού στις παρτίδες. Τα όρια πρόβλεψης υπολογίζονται χωριστά για κάθε παρτίδα με βάση τις ατομικές ποσότητες EMS και το διάστημα που τέμνει (διαπερνά) το όριο αποδοχής (προκαθορισμένη προδιαγραφή του προϊόντος) επιλέγεται πρωτίστως για την εκτίμηση της ημερομηνίας λήξης. Τρεις τουλάχιστον παρτίδες απαιτούνται για τη μελέτη σταθερότητας των προϊόντων.

Γενικεύοντας, ο χρόνος ζωής ορίζεται ως το χρονικό σημείο στο οποίο το 95% όριο εμπιστοσύνης, ανώτερο ή κατώτερο, της μέσης ευθείας αποδόμησης τέμνει ένα προκαθορισμένο όριο προδιαγραφών (συνήθως 90%) και το οποίο επισημαίνεται στην ετικέτα του προϊόντος.
Ο χρόνος ζωής εκτιμάται είτε γραφικά (Σχ. 15.1) είτε από την εξίσωση:

$x = \frac{\eta - a}{b}$ όπου x=t (χρόνος λήξης) και η το προκαθορισμένο όριο αποδοχής.

Η εξίσωση αυτή δεν θεωρείται αποδεκτή από τη συνολική διεθνή επιστημονική κοινότητα ή και κάποιους διεθνείς οργανισμούς. Ένας δεύτερος τρόπος βασίζεται στην τεχνική της αντίστροφης παλινδρόμησης (βλέπε ενότητα 8.1) και ο χρόνος ζωής υπολογίζεται ως:

$$x = c + d \cdot y \quad όπου \quad c = \bar{X} - d \cdot \bar{Y} \quad και \quad d = \frac{\Sigma xy}{\Sigma y^2} \quad .$$

Το 95% όριο εμπιστοσύνης που υπολογίζεται κατ' αυτόν τον τρόπο και ισχύει για την τιμή x χρησιμοποιείται για την εκτίμηση του χρόνου ζωής.

Παρουσίαση και ερμηνεία της μελέτης σταθερότητας

Τα βασικά σημεία για την ερμηνεία της μελέτης σταθερότητας περιλαμβάνουν τη στατιστική σημαντικότητα της ανάλυσης διακύμανσης (p-τιμή), την εξίσωση και τους συντελεστές της παλινδρόμησης, τη συνοπτική περιγραφή των στατιστικών ποσοτήτων, τα διαγράμματα των τυποποιημένων υπολειμμάτων και τις εκτιμήσεις του χρόνου ζωής του προϊόντος. Τα σημεία αυτά συνοψίζονται σε πέντε βήματα ενεργειών:

1. Εκτίμηση της στατιστικής σημαντικότητας της σχέσης που διέπει τη μεταβλητή απόκρισης (εξαρτημένη) με τους στατιστικούς όρους της παλινδρόμησης. Εκκινούμε με το πλήρες μοντέλο το οποίο περιλαμβάνει το χρόνο, την παρτίδα και την αλληλεπίδραση των δύο όρων (παραγόντων) συγκρίνοντας καταρχήν την ακριβή πιθανότητα σφάλματος p της αλληλεπίδρασης με την οριακή τιμή αναφοράς η οποία ειδικά στη μελέτη σταθερότητας ισούται με α=0.25. Αν $p<0.25$ τότε το μοντέλο δεν μπορεί να αναχθεί και διατηρεί στους υπολογισμούς και τους τρεις όρους. Αν $p \geq 0.25$ τότε η αλληλεπίδραση απομακρύνεται από την ανάλυση διακύμανσης, η οποία εξετάζει πλέον το ανηγμένο μοντέλο με τους δύο όρους (χρόνος και παρτίδα). Αν η πιθανότητα της παρτίδας στο νέο μοντέλο είναι $p<0.25$, το μοντέλο δεν μπορεί να αναχθεί εκ νέου και διατηρεί τους δύο όρους. Αν όμως $p \geq 0.25$ τότε ο όρος της παρτίδας αφαιρείται από την ανάλυση και το τελικό μοντέλο περιλαμβάνει μόνο το χρόνο.

Ως παράδειγμα, αναφέρεται η μελέτη σταθερότητας μίας δραστικής ουσίας φαρμακευτικού σκευάσματος επί 48μηνο με τη χρήση 5 παρτίδων παρασκευής του. Το ζητούμενο αφορά στην εκτίμηση του χρόνου ζωής της ουσίας όταν η αποδόμηση του προϊόντος προσεγγίσει το 90% της αρχικής του ισχύος. Ο ακόλουθος ένθετος πίνακας της μελέτης σταθερότητας περιγράφει τα αποτελέσματα της ανάλυσης διακύμανσης (ANOVA) των 5 παρτίδων με τη 48μηνη παρακολούθηση της δραστικότητας του σκευάσματος.

Επιλογή μοντέλου με βάση την πιθανότητα αναφοράς α=0.25

Πηγή διακύμανσης	DF	ESS	EMS	F-τιμή	p-τιμή
Μήνες	1	184.97	184.97	427.52	0.000
Παρτίδα	4	7.00	1.75	4.04	0.008
Μήνες*Παρτίδα	4	2.72	0.68	1.57	0.203
Σφάλμα	35	15.14	0.43	0.00	
Ολική	44	209.83	0.00	0.00	

Όπως καταφαίνεται, η αλληλεπίδραση είναι στατιστικά σημαντική, p=0.203<0.25, οπότε παραμένει στο μοντέλο. Επίσης και οι κλίσεις των παρτίδων είναι στατιστικά σημαντικές, p=0.008<0.25 και έπεται ότι διαφέρουν στατιστικά σημαντικά μεταξύ των παρτίδων.
Οι υπολογισμοί των στατιστικών παραμέτρων που συνοδεύουν μία πλήρη στατιστική ανάλυση της μελέτης

σταθερότητας περιγράφονται λεπτομερώς στον πίνακα 15.2 στο τέλος του κεφαλαίου.

2. Εκτίμηση της διάρκειας ζωής η οποία συνοδεύεται από το προκαθορισμένο όριο 90% αποδοχής της αποδόμησης (ή αλλιώς το 90% της αρχικής τιμής) και από το 95% όριο εμπιστοσύνης (ανώτερο ή κατώτερο αναλόγως της κατευθυντήριας τάσης της απόκρισης) το οποίο ανταποκρίνεται στο χρόνο ζωής. Όταν το πείραμα εκτελείται με τη χρήση μίας μόνο παρτίδας, τότε εκτιμάται μία τιμή για τη διάρκεια ζωής του προϊόντος (Παράδειγμα 15.1 και σχήμα 15.1). Ειδάλλως, υπολογίζεται ένας χρόνος ζωής ανά παρτίδα και ο ολικός χρόνος ζωής για όλες τις παρτίδες θα ισούται υποχρεωτικά με την μικρότερη ατομική τιμή που καταγράφηκε μεταξύ των παρτίδων (όταν βέβαια ισχύει για τις παρτίδες $p<0.25$). Στο παράδειγμα του φαρμάκου, στο τελικό μοντέλο με παρούσες στατιστικά σημαντικές τις παρτίδες, εκτιμώνται 5 τιμές και η ολική τιμή ισούται με $t=56.5$ ημέρες και αντιστοιχεί στη μικρότερη τιμή του πίνακα εκτίμησης της διάρκειας ζωής των 5 παρτίδων, δηλαδή στην παρτίδα 2. Στον ακόλουθο ένθετο πίνακα παρατίθενται το φύλλο με τα στοιχεία της μελέτης σταθερότητας.

Εκτίμηση του χρόνου ζωής

Παρτίδα	Χρόνος ζωής
1	71.15
2	56.50
3	58.73
4	69.12
5	68.02
Ολικός	56.50

Φύλλο εργασίας

Μήνες	Παρτίδα	Απόκριση%	Μήνες	Παρτίδα	Απόκριση%	Μήνες	Παρτίδα	Απόκριση%
0	4	99.27	9	2	97.00	24	3	96.70
0	3	100.10	9	5	99.28	24	1	97.60
0	5	100.40	9	3	98.49	24	5	97.20
0	2	100.20	9	1	99.00	24	2	96.01
0	1	99.80	9	4	98.53	24	4	96.92
3	2	99.48	12	2	98.00	36	2	94.30
3	4	99.76	12	5	97.68	36	3	95.00
3	1	99.60	12	3	99.05	36	5	94.40
3	5	98.00	12	1	98.13	36	4	94.61
3	3	99.62	12	4	97.70	36	1	96.50
6	2	98.00	18	3	98.34	48	4	94.30
6	5	98.60	18	4	97.75	48	1	93.30
6	1	99.05	18	5	97.40	48	3	92.24
6	3	99.81	18	2	95.00	48	5	94.07
6	4	99.39	18	1	97.90	48	2	92.60

3. Εξέταση του τρόπου που σχετίζονται οι όροι της εξίσωσης με την απόκριση. Αν το ανηγμένο μοντέλο περιλαμβάνει μόνο το χρόνο, τότε οι παρτίδες μοιράζονται ισότιμα ένα κοινό ύψος και μία κοινή κλίση και το μοντέλο περιγράφεται από μία απλή εξίσωση παλινδρόμησης. Αν πάλι ο επιλέξιμος παράγοντας της παρτίδας εισάγεται στο μοντέλο χωρίς όμως την αλληλεπίδραση με το χρόνο, τότε όλες οι παρτίδες περιγράφονται με διαφορετικό ύψος κλίσης αλλά με τον ίδιο ρυθμό αποδόμησης (κοινή κλίση). Αν τέλος το μοντέλο περιλαμβάνει και τους τρεις όρους, τότε οι παρτίδες έχουν διαφορετικές κλίσεις και ύψη των κλίσεων. Στο παράδειγμα του φαρμάκου οι εξισώσεις των παρτίδων διαμορφώνονται ως εξής:

Παρτίδα	Εξισώσεις παλινδρόμησης		
1	Απόκριση%	=	99.971 - 0.120 Μήνες
2	Απόκριση%	=	99.257 - 0.146 Μήνες
3	Απόκριση%	=	100.480 - 0.160 Μήνες
4	Απόκριση%	=	99.678 - 0.121 Μήνες
5	Απόκριση%	=	99.553 - 0.121 Μήνες

Σύμφωνα με τον πίνακα της ANOVA, οι παράγοντες παρτίδα και χρόνος*παρτίδα είναι στατιστικά σημαντικές, επομένως οι εξισώσεις παλινδρόμησης έχουν διαφορετικές κλίσεις και ύψη των κλίσεων. Οι κλίσεις εκφράζονται μετρικά με τους συντελεστές παλινδρόμησης $β_i$ και τα ύψη των κλίσεων με τις παραμέτρους $α_i$, τις αποστάσεις δηλαδή των κλίσεων από το σημείο 0 των αξόνων X και Y, αφού προηγουμένως προεκταθούν μέχρι να τμήσουν τον άξονα Y. Ο συντελεστής παλινδρόμησης εκφράζει το μέγεθος και την κατεύθυνση της σχέσης που διέπει μια προβλεπτική μεταβλητή με την απόκριση Y και εκτιμά τη μεταβολή της μέσης απόκρισης όταν ο χρόνος αυξάνεται κατά μία μονάδα. Αν είναι αρνητικός, η μέση απόκριση μειώνεται με το χρόνο, αν είναι θετικός αυξάνεται.

Η παρτίδα 3 εμφανίζει την ταχύτερη πτώση (-0.160), η οποία εξηγεί ότι κάθε μήνα η συγκέντρωση της χημικής ένωσης (απόκριση%) για την παρτίδα 3 μειώνεται κατά 0.160 ποσοστιαίες μονάδες. Η παρτίδα 2 έχει το μικρότερο ύψος της κλίσης (99.257) γεγονός που σημαίνει ότι είχε τη μικρότερη συγκέντρωση σε χρόνο 0 (κατά την έναρξη). Παρατηρώντας τις κλίσεις στον πίνακα αλλά και στο σχήμα 15.2 εύκολα διαπιστώνεται ότι οι παρτίδες 1, 4 και 5 έχουν ισότιμες κλίσεις, δηλαδή παρουσιάζουν ομοιόμορφο ρυθμό αποδόμησης, όπως επίσης διαφορετικό αλλά σχετικά παρόμοιο ρυθμό μεταβολής εμφανίζουν οι παρτίδες 2 και 3. Η συμπεριφορά των ρυθμών αποδόμησης είναι εμφανής στο σχήμα 15.2 και εφόσον υπάρχει ενδιαφέρον, ο ερευνητής μπορεί να ανατρέξει στη διαδικασία σύγκρισης των κλίσεων που παρέχεται λεπτομερώς στην ενότητα 8.2.

4. Εκτίμηση του ποσοστού (βαθμού) προσαρμογής των στοιχείων στην εξίσωση της παλινδρόμησης.

Αυτή περιγράφεται από τρεις διαφορετικούς συντελεστές της ευθείας προσαρμογής με τιμές που κυμαίνονται μεταξύ 0% (έλλειψη προσαρμογής) και 100% (πλήρης προσαρμογή).
Ο προσδιοριστικός συντελεστής R^2 εκφράζει το ποσοστό της διακύμανσης που περιγράφεται από το μοντέλο, όσο υψηλότερος τόσο επαρκέστερο το μοντέλο. Ο συντελεστής αυξάνεται συνεχώς όταν προστίθενται νέοι όροι στο μοντέλο ακόμα και όταν δεν παρατηρείται ουσιαστική βελτίωση στην προσαρμογή του μοντέλου και είναι χρήσιμος μόνο σε συγκρίσεις μοντέλων με ίσο αριθμό όρων.
Ο διορθωτικός συντελεστή $R^2_δ$ είναι χρήσιμος για συγκρίσεις μοντέλων με διαφορετικό αριθμό όρων σε αυτά, καθότι δεν αυξάνεται σημαντικά προστιθέμενων των όρων στο μοντέλο.
Στο παράδειγμα του φαρμάκου, το μοντέλο εμφανίζει υψηλή προσαρμοστικότητα όπως φαίνεται από τον προσδιοριστικό συντελεστή (R^2=92.78%) και πολύ καλή προβλεπτική προσαρμογή νεοεισερχόμενων τιμών η οποία ανέρχεται σε ποσοστό R^2_p=86.34%.

s	**R^2**	**$R^2_δ$**	**R^2_p**
0.658	92.78%	90.93%	86.34%

Παράδειγμα 15.1. Τυπική μελέτη σταθερότητας μίας δραστικής χημικής ουσίας, η ισχύς της οποίας εξασθενεί με την πάροδο του χρόνου, με δύο επαναληπτικές τιμές ανά περίοδο και με τη χρήση μίας παρτίδας παραγωγής.
Τα στοιχεία της ισχύος (%) προκύπτουν αφαιρετικά από την έναρξη του πειράματος. Ως αρχική τιμή λαμβάνεται είτε η αναγραφόμενη στην ετικέτα (και άρα 100%) είτε προκύπτει ως μέσος όρος από δύο ή περισσότερες επαναληπτικές τιμές στο χρόνο 0. Για παράδειγμα, αν οι δύο πρώτες τιμές της ουσίας βρέθηκαν 37 και 34 μονάδες τότε η αρχική τιμή ισούται με 35.5 [(37+34)/2] και οι ποσοστιαίες τιμές διαμορφώνονται αντιστοίχως σε 104.2% (37/35.5) και 95.7%(34/35.5).

Η ανάλυση διακύμανσης της παλινδρόμησης έδειξε υψηλή στατιστική σημαντικότητα της σχέσης που διέπει τη δραστικής ουσία με το χρόνο ($p<0.001$).

Ημέρες	Ισχύς%
0	99.3
0	100.0
3	100.6
3	99.6
6	99.8
6	99.4
9	99.7
9	99.3
12	98.1
12	97.3
18	98.3
18	97.9
24	96.0
24	96.9
36	94.1
36	94.6
48	92.2
48	94.1

Ανάλυση Διακύμανσης

Πηγή διακύμανσης	**DF**	**ESS**	**EMS**	**F-τιμή**	**p-τιμή**
Ημέρες	1	95.97	95.97	202.34	0.000
Σφάλμα	16	7.59	0.47		
Ολική	17	103.56			

Το μοντέλο εμφανίζει υψηλή προσαρμοστικότητα όπως δείχνουν ο προσδιοριστικός συντελεστής (92.67%) και ο προβλεπτικός (89.91%).

Προσαρμογή του μοντέλου

s	**R^2**	**$R^2_δ$**	**R^2_p**
0.689	92.67%	92.21%	89.91%

Οι συντελεστές του ύψους της κλίσης της ευθείας (παράμετρος α) και του χρόνου είναι στατιστικά σημαντικοί ($p<0.001$).

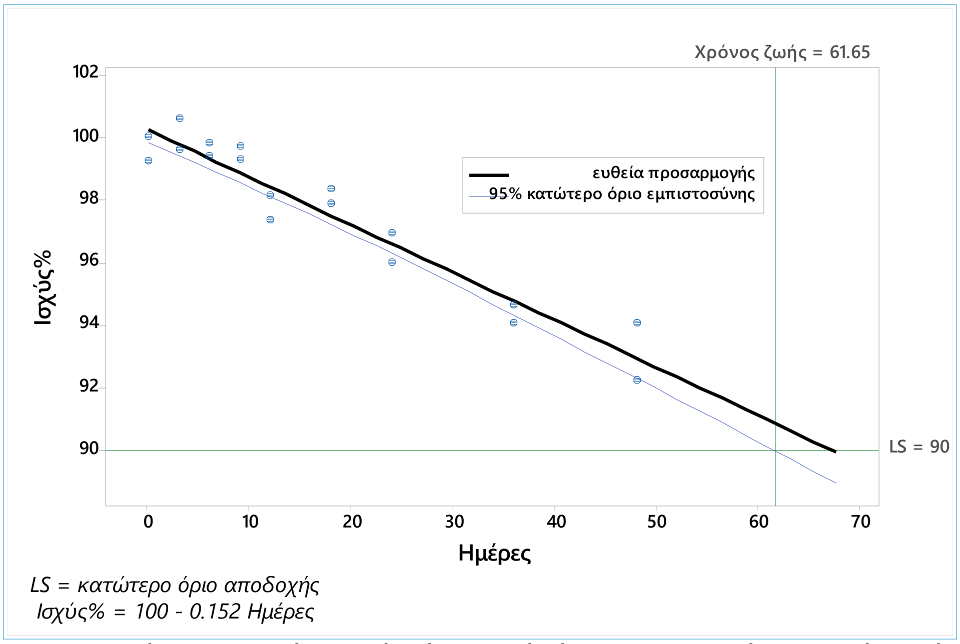
Συντελεστές του μοντέλου

Όρος	**β**	**SE(β)**	**t-τιμή**	**p-τιμή**
Παράμετρος α	100.272	0.247	406.53	0.000
Ημέρες	-0.152	0.011	-14.22	0.000

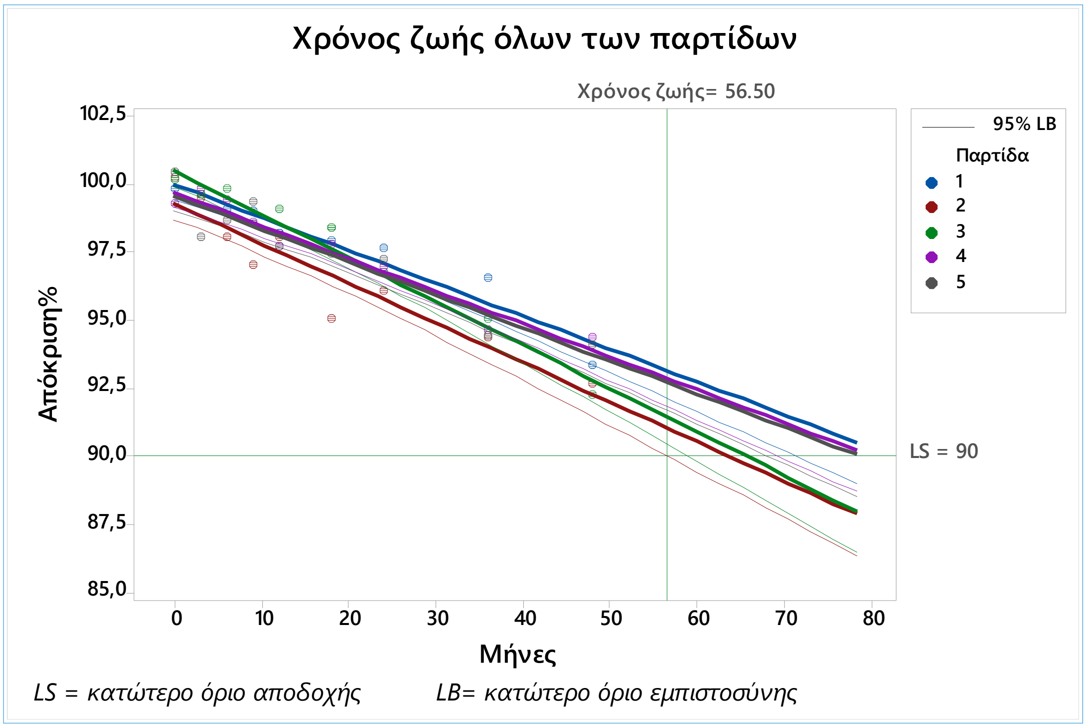
και συνθέτουν την εξίσωση της παλινδρόμησης

Ισχύς% = 100.272 - 0.152 Ημέρες

η οποία πληροφορεί ότι η αύξηση του χρόνου κατά μία ημέρα προκαλεί μείωση της δραστικότητας της ουσίας κατά 0.152%.
Ο χρόνος ζωής της δραστικής ουσίας είναι η περίοδος που με 95% πιθανότητα εμπιστοσύνης αποφαινόμαστε ότι η μέση απόκριση είναι άνω του χαμηλότερου ορίου αποδοχής.
Εκτίμηση χρόνου ζωής
κατώτερο όριο αποδοχής = 90%
Χρόνος ζωής = 61.65 (ημέρες)
Όπως επεξηγείται στο σχήμα 15.1 ο χρόνος ζωής προκύπτει από το σημείο τομής του 95% κατώτερου ορίου της παλινδρόμησης και του ορίου αποδοχής εξασθένησης της δραστικότητας στο 90%.



Σχ. 15.1. Εκτίμηση του χρόνου ζωής μίας χημικής ένωσης διαχρονικά. Το σημείο τομής της γαλάζιας ευθείας η οποία αντιστοιχεί στο 95% κατώτερο όριο της μέσης απόκρισης με το 90% όριο αποδοχής της αποδόμησης δίνει προεκβαλόμενο χρόνο ζωής 61.65 μήνες.



Σχ. 15.2. Εκτίμηση του χρόνου ζωής μίας δραστικής ουσίας φαρμάκου διαχρονικά με τη χρήση 5 παρτίδων παρασκευής επί 48μηνο. Το σημείο τομής της κόκκινης λεπτής ευθείας η οποία αντιστοιχεί στην παρτίδα 2 και στο 95% κατώτερο όριο της μέσης απόκρισης με το 90% όριο αποδοχής της αποδόμησης δίνει προεκβαλόμενο χρόνο ζωής 56.5 μήνες. Η τεχνική προϋποθέτει την εφαρμογή του μοντέλου I της μελέτης σταθερότητας, το οποίο αποδέχεται διαφορετικές κλίσεις και ύψη των κλίσεων.

5. Εξέταση της συμμόρφωσης του μοντέλου με τις προϋποθέσεις της κανονικότητας και ομοιογένειας των διακυμάνσεων των υπολειμμάτων της εξίσωσης στα αντίστοιχα διαγράμματα. Περιπτώσεις που οδηγούν