ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ Δ. ΜΑΚΕΔΟΝΙΑΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΑΣ ΓΕΩΠΟΝΙΑΣ
& ΤΕΧΝΟΛΟΓΙΑΣ ΤΡΟΦΙΜΩΝ ΚΑΙ ΔΙΑΤΡΟΦΗΣ

ΑΡΧΕΣ ΚΑΙ ΜΕΘΟΔΟΙ ΒΙΟΜΕΤΡΙΑΣ ΚΑΙ ΓΕΩΡΓΙΚΟΥ ΠΕΙΡΑΜΑΤΙΣΜΟΥ

ΙΩΑΝΝΗΣ Ν. ΞΥΝΙΑΣ, *M. Sc., Ph. D.*
ΚΑΘΗΓΗΤΗΣ

ΦΛΩΡΙΝΑ 2015

ΠΡΟΛΟΓΟΣ

Οι σημειώσεις αυτές γράφηκαν για να καλύψουν τις ανάγκες του μαθήματος της Στατιστικής των φοιτητών της Σχολής Τεχνολόγων Γεωπόνων, του Τεχνολογικού και Εκπαιδευτικού Ιδρύματος Δυτικής Μακεδονίας. Καταβλήθηκε ιδιαίτερη προσπάθεια ώστε οι γνώσεις που περιλήφθηκαν στο κείμενο να είναι όσο το δυνατόν πιο συγκεκριμένες και σαφείς, καλύπτοντας την ύλη που περιλαμβάνεται στο σχετικό περίγραμμα του μαθήματος. Άλλωστε το ανά χείρας πόνημα δε αποσκοπεί να αντικαταστήσει τα υπάρχοντα βιβλία Στατιστικής και Γεωργικού Πειραματισμού. Ο κύριος του στόχος είναι να δώσει στον φοιτητή τις απαραίτητες γνώσεις, ώστε στη συνέχεια να είναι σε θέση να εμβαθύνει περισσότερο, χρησιμοποιώντας πλέον ειδικά βιβλία. Βάση για τη συγγραφή του παρόντος πονήματος ήταν το βιβλίο του καθηγητή Α. Φασούλα "Στοιχεία Πειραματικής Στατιστικής", που θεωρώ ότι είναι το πλέον πρακτικό και κατανοητό βιβλίο που γράφηκε για τις ανάγκες των ασχολουμένων με τη Γεωπονική επιστήμη.

Φλώρινα, Μάρτιος 2015

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1ο

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

1.1. Ορισμός της Βιομετρίας.

Όταν κάποιος θέλει να μελετήσει τα γνωρίσματα ενός φυτικού οργανισμού, θα πρέπει να το κάνει κατά τέτοιο τρόπο, ώστε τα αποτελέσματα και τα συμπεράσματα στα οποία θα καταλήξει να μη επιδέχονται καμία αμφισβήτηση. Θα πρέπει δηλαδή να ανταποκρίνονται στην πραγματικότητα και να δίνουν την πραγματική εικόνα του γνωρίσματος που μελετάται. Η αναγκαιότητα αυτή προέκυψε σταδιακά, καθώς ο άνθρωπος προχωρούσε από την εξημέρωση κάποιων φυτικών ειδών (πρωτόγονος άνθρωπος) έως τη δημιουργία εμπορικών ποικιλιών και υβριδίων (σύγχρονος). Επιλέγοντας λοιπόν κάποιο υλικό κάθε φορά, θα έπρεπε να είναι σίγουρος ότι αυτό που κρατούσε είναι καλύτερο από αυτό που απέρριπτε. Η εργασία αυτή στην αρχή ήταν εύκολη, γιατί οι γενότυποι είχαν μεγάλες διαφορές μεταξύ τους. Με την συνεχή όμως επιλογή, οι διάφοροι γενότυποι άρχισαν να παρουσιάζουν ομοιότητες μεταξύ τους, να στενεύει δηλαδή η γενετική τους βάση, ενώ η εμπορική τους αξία αύξανε συνεχώς. Με την πάροδο λοιπόν των ετών άρχισε να δημιουργείται και η ανάγκη της καλύτερης αξιολόγησης του γενετικού υλικού που δημιουργούσε ο άνθρωπος. Για να μπορέσει να το πετύχει αυτό, άρχισε να παίρνει κάποιες παρατηρήσεις, τις οποίες συγκέντρωνε, ταξινομούσε και άρχισε να αναπτύσσει τρόπους επεξεργασίας των δεδομένων που συγκέντρωνε, ώστε να καταλήγει σε λογικά συμπεράσματα. Για να πάρει όμως τις παρατηρήσεις σωστά θα έπρεπε να βρεθεί και να χρησιμοποιηθεί κάποια κατάλληλη μεθοδολογία, ανάλογα με το είδος της παρατήρησης. Τέλος, για να κατορθώσει να επεξεργασθεί τα δεδομένα του θα έπρεπε να αναπτύξει και να χρησιμοποιήσει κάποιες ειδικές τεχνικές. Έτσι, λοιπόν, άρχισε να αναπτύσσεται σιγά-σιγά μια νέα επιστήμη που ονομάσθηκε Στατιστική. Η Στατιστική είναι μια εφαρμοσμένη επιστήμη που ασχολείται ακριβώς με τη σχεδίαση πειραμάτων, με τη λήψη κάποιων αριθμητικών δεδομένων, με τις μεθόδους με τις οποίους θα πρέπει να αναλυθούν τα δεδομένα αυτά και τέλος με τον τρόπο με τον οποίο θα παρουσιασθούν τα αποτελέσματα της ανάλυσης των δεδομένων. Η Βιομετρία είναι το τμήμα εκείνο της Στατιστικής που ασχολείται με την επίλυση βιολογικών προβλημάτων. Τέλος ο Γεωργικός Πειραματισμός είναι το τμήμα της Βιομετρίας που ασχολείται με το πώς θα πρέπει να γίνει ο πειραματισμός ώστε τα συμπεράσματα να είναι λογικά και αδιαμφισβήτητα.

Στη συνέχεια θα γίνει προσπάθεια να εξηγηθούν κάποιο όροι που είναι απαραίτητοι, ώστε να μπορέσει κάποιος να καταλάβει καλύτερα τις αρχές και τις μεθόδους που χρησιμοποιεί η Βιομετρία και ο Γεωργικός Πειραματισμός. Επειδή οι σημειώσεις αυτές απευθύνονται σε φοιτητές που σπουδάζουν την Γεωπονική Επιστήμη, θα επικεντρωθούν στο τμήμα εκείνο της Στατιστικής που μελετά βιολογικά φαινόμενα (Βιομετρία). Πιο συγκεκριμένα, θα επικεντρωθούν στο πώς θα πρέπει να γίνει ο πειραματισμός και η επεξεργασία των διάφορων δεδομένων που θα συγκεντρωθούν, ώστε το γενετικό υλικό πού έχει δημιουργηθεί να αξιολογηθεί σωστά (Γεωργικός Πειραματισμός).

1.2. Τι είναι παράγων ή μεταβλητή.

Όπως αναφέρθηκε προηγούμενα, ο πρωτόγονος άνθρωπος ήταν αυτός που άρχισε την εξημέρωση και βελτίωση των περισσότερων καλλιεργούμενων ειδών. Αυτό το έκανε με το μάτι, διαλέγοντας και κρατώντας εκείνα τα φυτά που παρουσίαζαν κατά τη γνώμη του, ενδιαφέρον. Γιατί όμως χρειάσθηκε να διαλέγει τα άτομα και πως έκρινε κάποια ότι είναι περισσότερο κατάλληλα από κάποια άλλα; Αυτό το έκανε γιατί τα μέλη ενός συνόλου ατόμων ενός φυτικού είδους είναι δυνατό να διαφέρουν μεταξύ τους σε πολλά γνωρίσματα: ύψος, πρωιμότητα, τίναγμα του σπόρου, ποιότητα του σπόρου κ. ά. Έτσι, παρατηρώντας ότι κάποια φυτά τίναζαν λιγότερο το σπόρο κατά την ωρίμανση στον αγρό, σκέφθηκε ότι αυτά ακριβώς τα φυτά θα έπρεπε να κρατήσει, γιατί με τον τρόπο αυτό θα μάζευε στην αποθήκη του μεγαλύτερη ποσότητα σπόρου.

Η διαφοροποίηση αυτή των ατόμων που απαρτίσουν ένα σύνολο, ως προς τα επιμέρους γνωρίσματά τους εισάγει μια νέα έννοια, αυτή του παράγοντα ή μεταβλητής. Ως παράγων ή μεταβλητή ονομάζεται κάθε γνώρισμα που παραλλάσσει και η παραλλακτικότητα αυτή μπορεί να μετρηθεί. Το ύψος των διαφόρων φυτών διαφέρει μεταξύ τους και είναι δυνατό να μετρηθεί. Για το λόγο αυτό είναι ένας παράγοντας ή μεταβλητή. Για να γίνουν τα πράγματα πιο απλά, στη συνέχεια θα χρησιμοποιείται ο όρος παράγοντας.

Οι παράγοντες διακρίνονται σε τρεις κατηγορίες:
(α) σε αυτούς που μπορούν να μετρηθούν,
(β) σε αυτούς που καθορίζουν τάξη ή βαθμό και
(γ) σε αυτούς που καθορίζουν ποιότητα.

Στους παράγοντες που μπορούν να μετρηθούν, οι διάφορες καταστάσεις μπορούν να εκφρασθούν αριθμητικά. Για το λόγο αυτό οι παράγοντες αυτοί μπορεί να είναι συνεχείς ή ασυνεχείς. Συνεχείς είναι εκείνοι στους οποίους μεταξύ δύο τιμών μπορούν να παρεμβληθούν άπειρες. Για παράδειγμα μεταξύ των τιμών 1,60 και 1,70 μπορεί να υπάρξουν τιμές 1,61, 1,62……1,69, ανάλογα με την ακρίβεια που χρησιμοποιείται. Στη κατηγορία αυτή ανήκουν πολλοί βιολογικοί παράγοντες: ύψος φυτών, μήκος σπάδικα, απόδοση φυτών, πρωιμότητα, θερμοκρασία ανάπτυξης κ. ά. Από την άλλη πλευρά, ασυνεχείς είναι οι παράγοντες εκείνοι που έχουν σταθερές τιμές, χωρίς να μπορεί να υπάρχουν ενδιάμεσες. Για παράδειγμα οι στήμονες ενός φυτού μπορεί να είναι 3, 4, 5, ποτέ όμως 3,5.

Στη δεύτερη κατηγορία παραγόντων ανήκουν αυτοί που ενώ δεν είναι δυνατό να μετρηθούν, μπορούν να καταταγούν, να ταξινομηθούν ή να βαθμολογηθούν κατά μέγεθος. Για παράδειγμα όταν κάποιος θέλει να αξιολογήσει την ανθεκτικότητα μιας ποικιλίας σε μια ασθένεια, βαθμολογεί τα φυτά χωρίς να ενδιαφέρεται πότε έγινε η προσβολή. Η σειρά της ανθεκτικότητας των φυτών είναι μια μεταβλητή που δείχνει βαθμολογία.

Τέλος, οι παράγοντες της τρίτης κατηγορίας δεν μετρώνται αλλά δείχνουν ιδιότητες: π. χ. άσπρο-μαύρο, άνδρας-γυναίκα. Για να γίνει δυνατή η στατιστική επεξεργασία των μεταβλητών αυτών θα πρέπει να συνοδεύονται από συχνότητες.

1.3. Η έννοια του πληθυσμού.

Συνήθως, αυτός που κάνει τη μελέτη και επιλογή ενός φυτικού είδους, έχει μπροστά του ένα μεγάλο αριθμό διαφορετικών ατόμων. Ξεκινά επιλέγοντας κάποια από τα χιλιάδες φυτά που έχει στον αγρό και καταλήγει στη συγκριτική αξιολόγηση κάποιων γενοτύπων, ώστε να βρεθεί ο καλύτερος. Για να το πετύχει αυτό κάνει ατομικές μετρήσεις στα φυτά. Κάθε ατομική μέτρηση που αναφέρεται σε ένα παράγοντα ονομάζεται τιμή του παράγοντα. Το σύνολο όλων των τιμών μαζί ονομάζεται πληθυσμός. Ο πληθυσμός είναι δυνατό να προσδιορισθεί με την εκτίμηση κάποιων συγκεκριμένων ποσοτήτων (μέσος όρος και τυπική απόκλιση), που ονομάζονται παράμετροι.

1.4. Η έννοια του δείγματος.

Όταν τα άτομα ενός συνόλου που μελετά κάποιος είναι λίγα, τότε μπορεί να μετρήσει το ύψος όλων των ατόμων και να βρει στη συνέχεια το ύψος όλου του συνόλου. Τις περισσότερες όμως φορές ο αριθμός των ατόμων που απαρτίζουν ένα σύνολο είναι πολλά και είναι για το λόγο αυτό αδύνατο να τα μετρήσει κάποιος όλα για να βρει το ύψος του συνόλου. Έτσι περιορίζεται στη μέτρηση του ύψους κάποιων ατόμων και προσπαθεί από τη μέτρηση αυτή να βγάλει συμπεράσματα για όλο τον πληθυσμό. Το κομμάτι του πληθυσμού που χρησιμοποιείται στην περίπτωση αυτή ονομάζεται δείγμα. Οι ποσότητες που προσδιορίζουν ένα δείγμα ονομάζονται στατιστικές και με την εκτίμησή τους επιδιώκεται να περιγραφεί ο πληθυσμός. Για να δώσουν όμως οι στατιστικές του δείγματος την καλύτερη δυνατή εκτίμηση των αντίστοιχων παραμέτρων του πληθυσμού, θα πρέπει το δείγμα να είναι, όπως λέγεται, αντιπροσωπευτικό. Τα δείγμα, λοιπόν, είναι αντιπροσωπευτικό όταν παίρνεται κατά τέτοιο τρόπο, ώστε κάθε τιμή του πληθυσμού να μπορεί να περιληφθεί σε αυτό με την ίδια πιθανότητα. Μόνο όταν το δείγμα είναι τυχαίο μπορεί να πληροί την παραπάνω προϋπόθεση. Επειδή στη Στατιστική γενικότερα και στη Βιομετρία ειδικότερα, η περιγραφή του πληθυσμού γίνεται με εκτιμήσεις που βασίζονται σε δείγματα, η αντιπροσωπευτικότητα των τελευταίων αποτελεί τον πιο σπουδαίο συντελεστή για αντικειμενική και αδιαμφισβήτητη έρευνα.

1.5. Δειγματοληψία.

Όπως αναφέρθηκε προηγούμενα, η μελέτη ενός πληθυσμού γίνεται χρησιμοποιώντας ένα μικρό κομμάτι του που ονομάσθηκε δείγμα και το οποίο θα πρέπει να είναι τυχαίο για να είναι αντιπροσωπευτικό. Όταν κάποιος πάρει μετρήσεις στα άτομα που αποτελούν το δείγμα που χρησιμοποιεί, βλέπει ότι οι μετρήσεις αυτές διαφέρουν μεταξύ τους. Τα άτομα βέβαια είναι αυτά που διαφέρουν και η διαφορετικότητα αυτή κάνει φανερή την ύπαρξη της παραλλακτικότητας που υπάρχει. Η γενετική παραλλακτικότητα είναι αυτή που ενδιαφέρει τη βελτίωση των καλλιεργούμενων ειδών. Όμως η συμπεριφορά των ατόμων μπορεί να οφείλεται και σε παραλλακτικότητα που οφείλεται στο περιβάλλον, αλλά και στη αλληλεπίδραση του γενότυπου με το περιβάλλον. Μια άλλη όμως πηγή παραλλακτικότητας πιθανό να είναι όταν συγκρίνονται ανόμοια πράγματα ή η σύγκριση γίνεται σε ανόμοιες συνθήκες, αυτό όμως εύκολα μπορεί κάποιος να το αποφύγει.

Για να μπορέσει να βγάλει κάποιος λογικά και αδιαμφισβήτητα συμπεράσματα χρησιμοποιεί το δείγμα του πληθυσμού για τις μετρήσεις του. Η δειγματοληψία θα πρέπει να γίνει κατά τέτοιο τρόπο ώστε να εξασφαλίζει την αξιοπιστία του πειράματος. Θα πρέπει όλοι οι χειρισμοί να γίνονται με τον ίδιο τρόπο για όλα τα άτομα που μελετώνται. Για παράδειγμα το δείγμα των σπόρων των ποικιλιών που πρόκειται να σπαρθούν για να αξιολογηθούν θα πρέπει να περιλαμβάνει ισομεγέθεις σπόρους για όλες τις ποικιλίες, που έχουν απολυμανθεί κατά τον ίδιο τρόπο. Η δειγματοληψία δηλαδή θα πρέπει να εξασφαλίζει το αξιόπιστο του πειραματισμού. Επίσης, για να είναι το δείγμα πιο λειτουργικό είναι καλό να ταξινομείται (οι μετρήσεις να ταξινομούνται) κατά μέγεθος ή όπως βολεύει καλύτερα, για να διευκολυνθεί η παρουσίαση και η ανάλυση των δεδομένων. Τα δείγματα τέλος μπορούν και να ομαδοποιούνται. Όταν χρησιμοποιείται η τελευταία περίπτωση, ο τρόπος με τον οποίο μια σειρά παρατηρήσεων κατανέμεται κατά ομάδες ονομάζεται κατανομή συχνότητας.

1.5.1. Τρόποι δειγματοληψίας

Α. Μεροληπτική δειγματοληψία

Είναι εύκολο να επιλεγεί ένα δείγμα που να δίνει μια παραμορφωμένη εικόνα ενός πληθυσμού. Π.χ. δείγμα οικογενειών κυρίως μεταξύ φίλων που δέχονται την επίσκεψη. Δείγμα από άτομα που βρίσκονται στα σπίτια τους και είναι πρόθυμα να απαντήσουν από την πρώτη φορά.

Β. Τυχαίο δείγμα

Β1. Με αντικατάσταση

Π. χ. ένας πληθυσμός με τέσσερα μέλη, αριθμημένα 1, 2, 3, 4 και ζητείται δείγμα μεγέθους 2.

Υπάρχουν 16 πιθανά δείγματα: (1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3, 2), (3,3), (3,4), (4,1), (4,2), (4,3), και (4,4).
Κάθε μέλος του πληθυσμού εμφανίζεται μια φορά σε έξι δείγματα και δυο σε ένα. Η δειγματοληψία δεν ευνοεί κάποιο μέλος σε βάρος κάποιου άλλου. Κάθε ένα από τα 16 δείγματα έχει ίση ευκαιρία να είναι το δείγμα που θα ληφθεί

Β2. Χωρίς αντικατάσταση
Υπάρχουν τα παρακάτω πιθανά δείγματα: (1,2), (1,3), (1,4), (2,3), (2,4), (3,4).

Β 3. Στρωματοποιημένη τυχαία δειγματοληψία.
Κάποιοι λογαριασμοί (ή αποδείξεις) αρχικά ταξινομηθήκαν σύμφωνα με το μέγεθος κάθε λογαριασμού. Επιλέχθηκαν όλοι οι λογαριασμοί που ήταν μεγαλύτεροι από $40, γιατί είναι αναγκαίο ένα ακριβές δείγμα των λογαριασμών που είναι μεγαλύτεροι από $40. Στη συνέχει επιλέχθηκε ένα τυχαίο δείγμα του 50% των λογαριασμών που ήταν από $20 έως $40 κ. ο. κ. Μόνο το 1% των λογαριασμών που ήταν μικρότεροι από $5 επιλέχθηκε, γιατί δεν είναι απαραίτητο ένα ακριβές δείγμα όταν το χρηματικό ποσό είναι μικρό.

Γ. Πίνακες τυχαίων αριθμών.
Στην περίπτωση αυτή χρησιμοποιούνται οι πίνακες τυχαίων αριθμών (Πίνακας 1, αλλά και Πίνακες Χ, Χι, ΧΙΙ στο Παράρτημα).

Πίνακας 1. Τυχαίοι αριθμοί

	00-04	05-09	10-14	15-19
00	54463	22662	65905	70639
01	15389	85205	18850	39226
02	85941	40756	82414	02015
03	61149	69440	11286	88218
04	05219	81619	10651	67079
……………………	……………………	……………………	……………………	……………………

Για να βρεθεί λοιπόν η σειρά με την οποία θα σπαρούν 10 ποικιλίες σε έναν πειραματικό αγρό χρησιμοποιούνται οι πρώτοι αριθμοί της πρώτης στήλης, πηγαίνοντας από πάνω προς τα κάτω. Έτσι πρώτη θα σπαρθεί η πέμπτη ποικιλία γιατί πρώτος αριθμός της πεντάδας 54463 είναι το 5. Στη συνέχεια θα σπαρθεί η πρώτη ποικιλία, γιατί πρώτος αριθμός της πεντάδας 15385 είναι το 1, ακολουθεί η όγδοη ποικιλία (γιατί το 8 είναι ο πρώτος αριθμός της πεντάδας 85941, και για τον ίδιο λόγο θα σπαρθούν στη συνέχεια οι ποικιλίες με αριθμό έξι και 10 (το 0 στις περιπτώσεις αυτές αντιστοιχεί με το 10). Με τον τρόπο αυτό βρέθηκε η τυχαία σειρά με την οποία θα σπαρθούν οι πέντε πρώτες ποικιλίες. Επειδή στο παράδειγμα του Πίνακα 1 δεν υπάρχουν άλλοι αριθμοί προχωρώντας προς τα κάτω, στη συνέχεια θα χρησιμοποιηθούν οι δεύτεροι αριθμοί των πεντάδων της πρώτης στήλης, οι τρίτοι αριθμοί κ. ο. κ., έως ότου βρεθεί η σειρά και των δέκα ποικιλιών. Εννοείται ότι αν κάποιος αριθμός βρεθεί για δεύτερη φορά, τότε θα αγνοηθεί και θα χρησιμοποιηθεί ο επόμενος, αν και αυτός με τη σειρά του δεν έχει χρησιμοποιηθεί άλλη φορά. Με τον τρόπο λοιπόν αυτό προκύπτουν και οι αριθμοί τέσσερα, τρία, εννέα, δύο και επτά, που συμπληρώνουν τη δεκάδα.

1.6. Παρατηρήσεις και πειραματικά δεδομένα.

Οι πληροφορίες που συγκεντρώνει αυτός που θέλει να αξιολογήσει κάποιους γενοτύπους αφορούν όσο το δυνατό περισσότερα από τα γνωρίσματα που έχουν τα φυτά. Στην περίπτωση αυτή παίρνει όπως λέγεται παρατηρήσεις για το γνώρισμα που τον ενδιαφέρει. Οι παρατηρήσεις διακρίνονται σε δυο κατηγορίες: στις αντικειμενικές και στις υποκειμενικές. Αντικειμενικές είναι εκείνες που στηρίζονται σε μετρήσεις, όπου η εμπειρία αυτού που παίρνει την παρατήρηση δεν έχει καμιά σημασία. Αντίθετα στις υποκειμενικές, που είναι οι υπόλοιπες, η εμπειρία παίζει τον καθοριστικό ρόλο στη λήψη της παρατήρησης. Το πότε και το πώς θα ληφθεί μια παρατήρηση καθορίζεται από το Ίδρυμα που κάνει την μελέτη.

Το σύνολο των παρατηρήσεων που παίρνει κάποιος καταγράφονται και ονομάζονται πειραματικά δεδομένα. Ο τρόπος με τον οποίο θα καταγραφούν τα πειραματικά δεδομένα είναι τέτοιος ώστε να εξασφαλίζει την ευκολότερη παρουσίαση και επεξεργασία τους. Τα πειραματικά δεδομένα συνήθως παρουσιάζονται σε πίνακες είναι όμως δυνατό μετά από μια πρόχειρη επεξεργασία να παρουσιασθούν και με τη μορφή διαγραμμάτων (αν χρησιμοποιηθούν όλες οι μετρήσεις) ή ιστογραμμάτων (αν γίνει ομαδοποίηση). Τέλος, ως προς την επεξεργασία των δεδομένων μπορούν να χρησιμοποιηθούν διάφορες τεχνικές ανάλογα με το σκοπό της έρευνας.

ΚΕΦΑΛΑΙΟ 2ο

ΠΑΡΟΥΣΙΑΣΗ ΚΑΙ ΣΥΝΟΨΙΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

2.1. Εισαγωγή.

Τα πειραματικά δεδομένα καταγράφονται και παρουσιάζονται με διάφορους τρόπους, ανάλογα με τη χρήση για την οποία προορίζονται. Αν σκοπός είναι η απλή παρουσίασή τους τότε χρησιμοποιούνται πίνακες ή καμπύλες. Αν όμως ο σκοπός είναι η στατιστική του επεξεργασία, τότε χρησιμοποιούνται μόνο πίνακες, στους οποίους τα δεδομένα μπαίνουν κατά τέτοιο τρόπο ώστε να είναι δυνατή η καλύτερη και ευχερέστερη επεξεργασία τους.

2.2. Πίνακες και διαγράμματα κατανομών συχνότητας.

Κάποιος θέλησε να μελετήσει την επίδοση δυο διαφορετικών πληθυσμών ατόμων σε ένα μάθημα. Για να το κάνει αυτό πήρε ένα δείγμα (Α) 100 ατόμων από τον ένα πληθυσμό και ένα δεύτερο δείγμα (Β) 100 ατόμων από τον άλλο. Η επίδοση των δυο ομάδων δίνεται στον Πίνακα 2.

Πίνακας 2. Επίδοση σε ένα μάθημα δυο ομάδων φοιτητών.

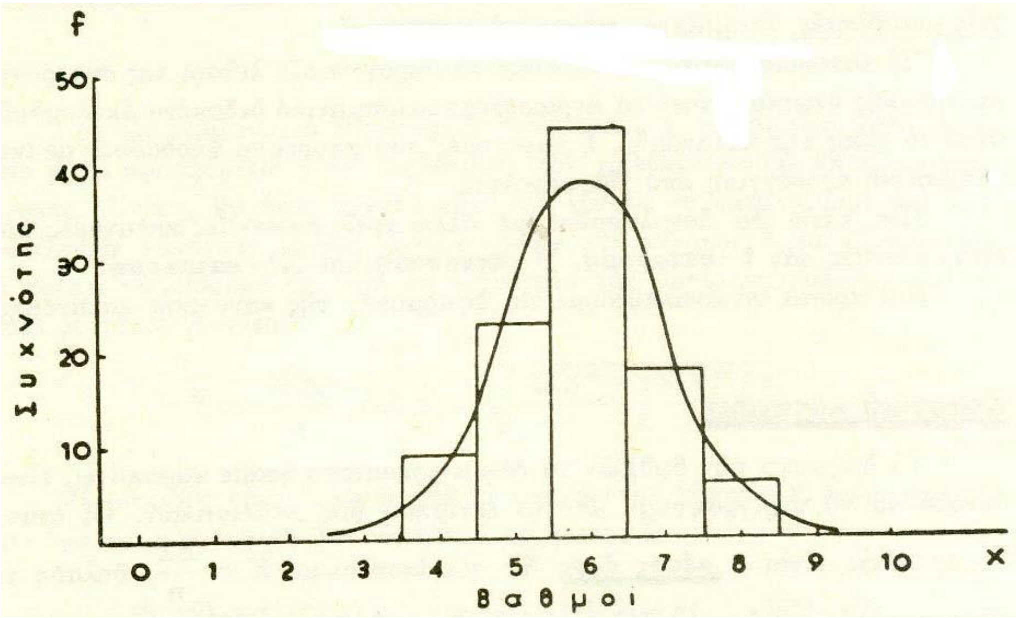
Ομάδα Α										Ομάδα Β									
6	7	5	5	5	6	6	6	7	5	5	6	5	5	5	4	5	4	5	4
8	5	5	6	6	7	8	7	5	5	8	6	2	8	4	3	6	5	4	4
5	6	6	5	6	7	8	6	5	6	5	6	5	6	4	4	7	8	5	7
6	7	6	6	7	6	6	8	4	4	5	8	5	4	5	4	5	3	5	5
6	4	6	6	7	6	6	6	5	7	4	5	6	5	5	6	6	4	3	5
7	6	5	6	6	5	8	6	6	6	4	3	5	4	6	4	5	6	4	6
5	6	4	5	7	7	5	5	6	6	3	6	5	6	6	2	3	7	6	6
4	6	6	4	4	5	6	7	5	6	5	5	5	6	7	5	7	3	5	6
6	8	7	6	4	5	6	6	7	6	6	3	6	5	5	7	7	2	5	5
5	7	6	4	6	7	6	6	7	5	4	6	3	3	5	5	7	4	6	4

Η επίδοση των φοιτητών που παρουσιάζεται στον παραπάνω Πίνακα μπορεί να πάρει οποιαδήποτε τιμή μεταξύ 0 και 10. Για το λόγο αυτό ονομάζεται συνεχής μεταβλητή ή παράγοντας. Στις περιπτώσεις αυτές αντί να αναγραφούν όλες οι δυνατές τιμές, που μπορεί να είναι πολλές, συγκροτούνται επιμέρους ομάδες, που ονομάζονται κλάσεις. Για παράδειγμα, ο αριθμός 6 είναι ο μέσος όρος της κλάσης που περιλαμβάνει τις τιμές από 5,5 έως 6,5. Επίσης, επειδή η κάθε κλάση επαναλαμβάνεται περισσότερες από μια φορές, εκτιμάται πόσες φορές επαναλαμβάνεται, αθροίζοντας όλες τις κλάσεις που έχουν τιμή 1, κατόπι αυτές που έχουν τιμή 2 , αυτές που έχουν τιμή 3 κ. ο. κ., έως ότου ολοκληρωθεί η εκτίμηση για όλες τις κλάσεις. Από το παραπάνω προκύπτει και η έννοια της συχνότητας, με την οποία εννοείται πόσες φορές επαναλαμβάνεται η κάθε κλάση. Με τον τρόπο αυτό γίνεται ο Πίνακας 3, που ονομάζεται πίνακας κατανομής συχνότητας ή πίνακας συχνότητας, γιατί περιλαμβάνει τις κλάσεις του παράγοντα και τις συχνότητες των κλάσεων.

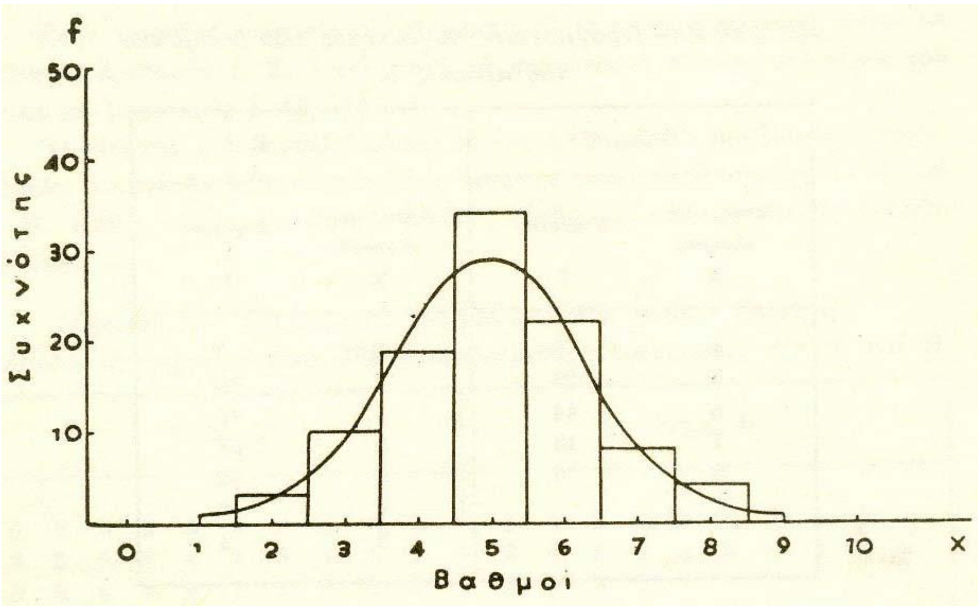
Πίνακας 3. Μέσες τιμές κλάσης και κατανομές συχνότητας των δεδομένων του Πίνακα 2.

Ομάδα Α		Ομάδα Β	
Μέση τιμή κλάσης X	Συχνότητα f	Μέση τιμή κλάσης X	Συχνότητα f
4	9	2	3
5	23	3	10
6	44	4	19
7	18	5	34
8	6	6	22
		7	8
		8	4

Όπως φαίνεται στον Πίνακα 3, οι συχνότητες κατανέμονται με ένα ξεχωριστό τρόπο: οι μεσαίες εμφανίζονται με τη μεγαλύτερη συχνότητα, ενώ προς τα δυο αντίθετα άκρα οι συχνότητες φθίνουν σταδιακά. Για την καλύτερη απεικόνιση των κατανομών της συχνότητας χρησιμοποιούνται ιστογράμματα, όπως αυτό φαίνεται στα Σχήματα 1 και 2. Εκτός όμως από τα ιστογράμματα, χρησιμοποιούνται και οι θεωρητικές καμπύλες, που υπολογίζονται από τα δεδομένα. Οι καμπύλες αυτές έχουν χαρακτηριστικό κωδωνοειδές σχήμα και ονομάζονται κανονικές καμπύλες ή καμπύλες του Gauss. Το είδος αυτό της κατανομής το ακολουθούν οι συχνότητες των συνεχών μεταβλητών και ονομάζεται κανονική κατανομή.



Σχήμα 1. Κατανομή συχνότητας των βαθμών των 100 φοιτητών της Α ομάδας.



Σχήμα 2. Κατανομή συχνότητας των βαθμών των 100 φοιτητών της Β ομάδας.

Από αυτά που αναφέρθηκαν παραπάνω προκύπτει ότι τα δεδομένα των πειραμάτων είναι δυνατό να παρουσιασθούν με διάφορους τρόπους, είτε με τη μορφή αριθμητικών πινάκων, πινάκων κατανομής συχνοτήτων ή με τη μορφή καμπυλών κατανομών συχνότητας.

2.3. <u>Μέτρα θέσης διασποράς</u>.

Στο παράδειγμα με τις δυο ομάδες φοιτητών τα δείγματα των βαθμών που χρησιμοποιήθηκαν μπορούν να περιγραφούν με την εκτίμηση δυο στατιστικών: του μέσου όρου ($\overline{X}$) και της τυπικής απόκλισης (S'). Ο μέσος όρος ($\overline{X}$) είναι το πηλίκο της διαίρεσης του αθροίσματος των τιμών του δείγματος (ΣΧ) δια του αριθμού των τιμών (n), δηλαδή

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

Η τυπική απόκλιση (S') μετρά το μέγεθος της διασποράς των τιμών γύρω από τον μέσο όρο και δίνεται από τον τύπο

$$S' = \sqrt{\frac{\sum(X-\overline{X})^2}{n}} \quad ή \quad S' = \sqrt{\frac{\{\sum X^2-(\sum \overline{X})^2\}/n}{n}}$$

Ο μέσος όρος για την ομάδα 1 του Πίνακα 2 είναι

$\overline{X}$=589/100=5,89

και η τυπική απόκλιση, χρησιμοποιώντας το δεύτερο τύπο, που ταιριάζει καλύτερα στις αριθμομηχανές,

$$S' = \sqrt{\frac{\{\sum X^2-(\sum \overline{X})^2\}/n}{n}} = \sqrt{\frac{\{6^2+7^2+\ldots+7^2+5^2-(589)^2\}/100}{100}}$$
$$= \sqrt{0{,}999 =}\ 0{,}999$$

Ο υπολογισμός του μέσου όρου και της τυπικής απόκλισης μπορεί να γίνει λαμβάνοντας υπόψη και τη συχνότητα f, οπότε οι τύποι γίνονται

$$\overline{X} = \frac{\Sigma fX}{n} \quad και \quad S' = \sqrt{\frac{\{\sum fX^2-(\sum f\overline{X})^2\}/n}{n}}$$

Σύμφωνα με τα δεδομένα του Πίνακα 3, με τους δυο νέους τύπους ο μέσος όρος και η τυπική απόκλιση της ομάδας 2 υπολογίζονται σε

$\overline{X}$ = (3Χ2+ 10Χ3+…..8Χ7+4Χ8)/100=5,02 και

$$S' = \sqrt{\frac{\{3 \times 2^2+10 \times 3^2+\ldots+8 \times 7^2+4 \times 8^2-(502)^2\}/100}{100}} = 1{,}7596 = 1{,}326$$

Αν συγκριθούν οι μέσοι όροι των δυο δειγμάτων φαίνεται εύκολα ότι ο μέσος όρος του πρώτου είναι μεγαλύτερος και βρίσκεται δεξιότερα επάνω στον οριζόντιο άξονα (Σχήματα 1 και 2). Ο μέσος όρος λοιπόν αποτελεί ένα <u>μέτρο θέσης</u> επάνω στον οριζόντιο άξονα. Η τυπική απόκλιση μετρά τη διασπορά των τιμών γύρω από τον μέσο όρο και είναι μεγαλύτερη στην ομάδα 2, όπου τα όρια κυμαίνονται από 1 έως 9, ενώ στην ομάδα 1 κυμαίνονται από 3 έως 9.

Από το παραπάνω παράδειγμα γίνεται φανερή μια σπουδαία ικανότητα της Στατιστικής, που είναι η απλούστευση των δεδομένων: οι 100 αριθμοί της ομάδας 1 έγιναν 10 στον Πίνακα της συχνότητας και περιορίσθηκαν σε μόνο δύο, το μέσο όρο και την τυπική απόκλιση.

ΚΕΦΑΛΑΙΟ 3ο

ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ-ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ

3.1. Στοιχεία πιθανοτήτων.

Αν ρίξει κάποιος ένα ζάρι, τότε αυτό μπορεί να πέσει κατά έξι διαφορετικούς τρόπους, δηλαδή με ένα από τους αριθμούς 1 έως 6. Η πιθανότητα να έρθει ο αριθμός 1 είναι μια από τις έξι. Στις περιπτώσεις λοιπόν που υπάρχει το ενδεχόμενο να συμβεί κάτι (Ε) κατά h τρόπους από τους n δυνατούς, τότε η πιθανότητα να συμβεί αυτό ονομάζεται επιτυχία και υπολογίζεται από τον τύπο

$$P = Pr(E) = h\, n^{-1}$$

Η πιθανότητα να μη συμβεί το Ε, ονομάζεται αποτυχία, συμβολίζεται ως –Ε και υπολογίζεται από τον τύπο

$$q = Pr(-E) = (n-h)\, n^{-1} = 1- p = 1-Pr(E)$$

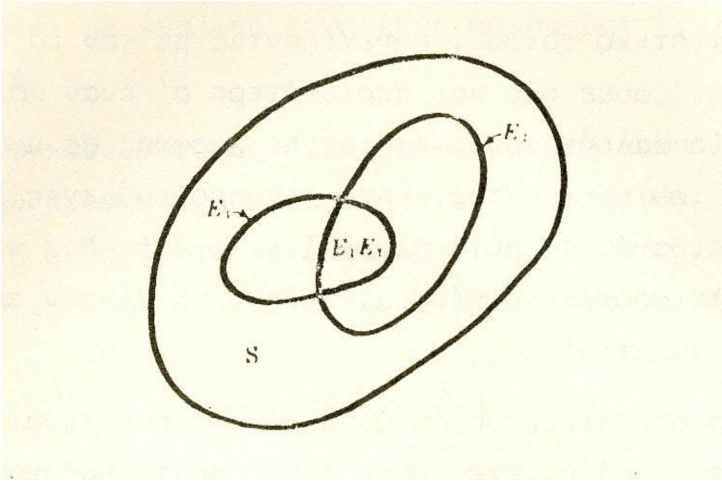
Έτσι στο προηγούμενο παράδειγμα με τα ζάρια η πιθανότητα επιτυχίας να έρθει ο αριθμός 2 είναι $P = Pr(2) = 1\; 6^{-1} = 1/6$, ενώ η πιθανότητα αποτυχίας είναι $q = Pr(-2) = (6-1)\; 6^{-1}= 5/6$.

Από το παραπάνω παράδειγμα γίνεται φανερό ότι η τιμή της πιθανότητας να συμβεί κάτι κυμαίνεται από 0 έως 1. Η τιμή 0 προκύπτει όταν δεν συμβεί αυτό το κάτι, ενώ η τιμή 1 προκύπτει όταν είναι σχεδόν βέβαιο ότι θα συμβεί.

Εκτός από τον παραπάνω ορισμό, υπάρχει και ο στατιστικός όρος για την εμπειρική πιθανότητα. Έτσι, εμπειρική πιθανότητα είναι η σχετική συχνότητα να εμφανισθεί ένα γεγονός, όταν ο αριθμός των παρατηρήσεων είναι πολύ μεγάλος. Στη σύγχρονη θεωρία των πιθανοτήτων, όλα τα δυνατά αποτελέσματα ενός πειράματος θεωρούνται ως σημεία ενός χώρου, που ονομάζεται δειγματικός χώρος S. Αν ο τελευταίος περιέχει ένα αριθμό σημείων, είναι δυνατό με κάθε σημείο να συσχετισθεί ένας θετικός αριθμός, ώστε το άθροισμα των αριθμών να ισούται με τη μονάδα. Στην περίπτωση αυτή ο κάθε αριθμός ισοδυναμεί με την πιθανότητα.

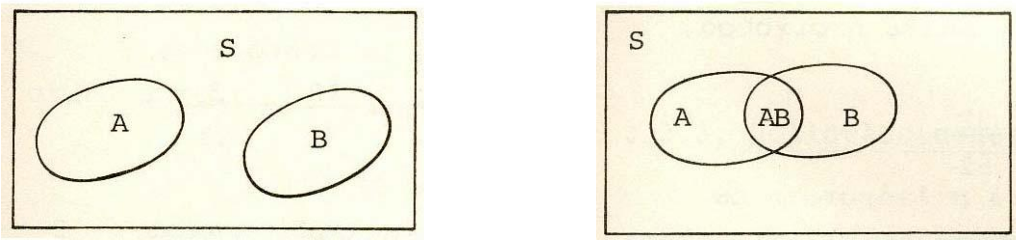
3.1.1. Κανόνες των πιθανοτήτων.

Δειγματικός χώρος (S) καλείται ο χώρος που περιλαμβάνει ως σημεία όλα τα δυνατά αποτελέσματα ενός πειράματος. Κάθε σημείο του χώρου αυτού είναι δυνατό να συσχετισθεί με ένα θετικό αριθμό, κατά τέτοιο τρόπο ώστε το σύνολο των αριθμών αυτών να ισούται με τη μονάδα. Ο καθένας αριθμός χωριστά ισοδυναμεί με την πιθανότητα. Ένα σύνολο σημείων στο δειγματικό χώρο αντιπροσωπεύει ένα ενδεχόμενο (π. χ. τα ενδεχόμενα E_1 και E_2 στο Σχήμα 3).



Σχήμα 3. Δύο ενδεχόμενα (E_1 και E_2) ενός δειγματικού χώρου(S).

Αν τα δύο ενδεχόμενα δεν έχουν κοινά σημεία, τότε ονομάζονται ασυμβίβαστα. Αντίθετα αν έχουν κοινά σημεία ονομάζονται συμβιβαστά (Σχήμα 4).



Σχήμα 4. Ασυμβίβαστα (αριστερά) και συμβιβαστά (δεξιά) ενδεχόμενα.

Στις πιθανότητες διακρίνονται οι εξής πράξεις:

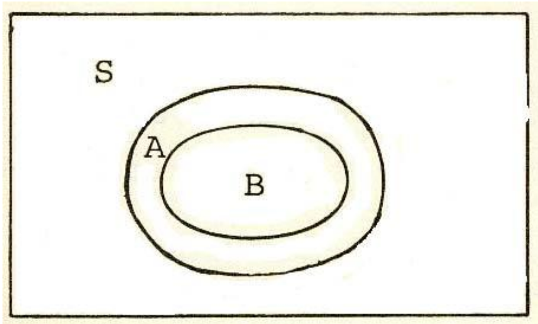
(α) Πρόσθεση πιθανοτήτων - ενδεχόμενα ασυμβίβαστα Σχήμα 4, αριστερά):

P(A+B)= P(A) +P(B)

(β) Πρόσθεση πιθανοτήτων - ενδεχόμενα συμβιβαστά (Σχήμα 4, δεξιά):

P(A+B)= P(A) +P(B) –P(AB)

(γ) Αφαίρεση πιθανοτήτων: P(A-B) = P(A)- P(B) (Σχήμα 5).



Σχήμα 5. Αφαίρεση πιθανοτήτων.

(δ) Πολλαπλασιασμός πιθανοτήτων: $P(E_1 \setminus E_2) = P(E_1)\, P(E_2 \setminus E_1)$
(η παράσταση $P(E_2 \setminus E_1)$ ονομάζεται δεσμευμένη πιθανότητα και είναι η πιθανότητα να συμβεί το E_2, με την προϋπόθεση ότι έχει ήδη συμβεί το E_1)

3.2. Κατανομές συχνότητας.

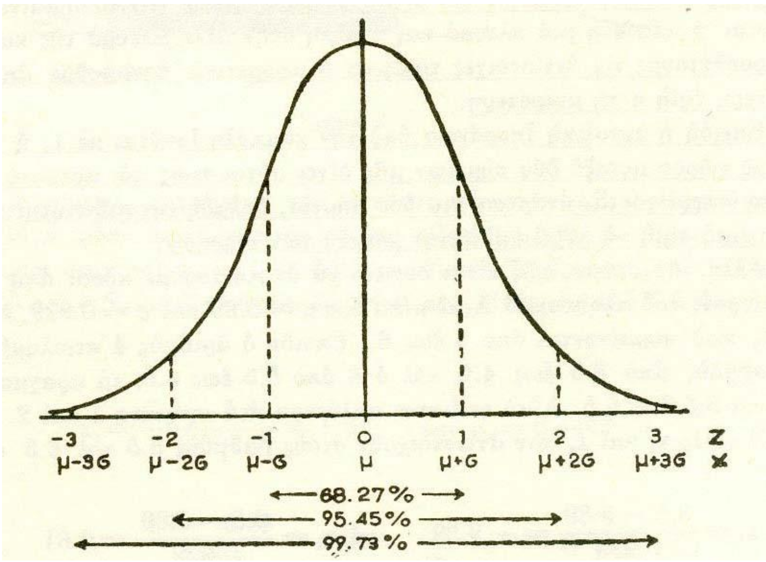
Η κανονική κατανομή στην οποία έγινε αναφορά στο προηγούμενο κεφάλαιο αποτελεί τη βάση της Στατιστικής, γιατί τα περισσότερα πειραματικά δεδομένα ακολουθούν αυτό το είδος της κατανομής, ή κατανομές, που είναι δυνατό να αποδοθούν ικανοποιητικά με την κανονική. Για το λόγο αυτό θα γίνει ειδική αναφορά στην κανονική κατανομή. Υπάρχουν όμως και άλλες κατανομές, τρεις από τις οποίες (t-κατανομή, F-κατανομή και X^2-κατανομή) έχουν σχεδόν ίση σπουδαιότητα με την κανονική και για το λόγο αυτό θα αναλυθούν επίσης με λεπτομέρεια. Μικρότερη αναφορά θα γίνει σε άλλες τρεις (διωνυμική, πολυωνυμική και Poisson) που χρησιμοποιούνται λιγότερο συχνά.

3.2.1. Κανονική κατανομή.

Στο προηγούμενο Κεφάλαιο αναφέρθηκε ότι ο μέσος όρος και η τυπική απόκλιση είναι δυο στατιστικές ικανές να περιγράφουν το δείγμα. Όμως μπορούν να περιγράψουν και ολόκληρο τον πληθυσμό από τον οποίο προέρχεται το δείγμα, μόνο που στην περίπτωση αυτή ονομάζονται παράμετροι και συμβολίζονται με τα γράμματα μ και σ αντίστοιχα. Οι παράμετροι μ και σ υπολογίζονται κατά προσέγγιση από τις αντίστοιχες στατιστικές του δείγματος $\overline{X}$ και S'. Από τις δυο στατιστικές του δείγματος το $\overline{X}$ αποτελεί την καλύτερη εκτίμηση του μ. Δεν συμβαίνει όμως το ίδιο και με το S', που αποτελεί μόνο μια καλή προσέγγιση του σ. Όσο το μέγεθος του δείγματος μεγαλώνει τόσο καλύτερη γίνεται η προσέγγιση αυτή. Έτσι, αν το δείγμα είναι μεγάλο (> 30) τότε χωρίς τον κίνδυνο να είναι μεγάλο το σφάλμα, μπορεί να υποτεθεί ότι σ = S' και η τυπική απόκλιση των μεγάλων δειγμάτων παριστάνεται και αυτή με σ (για να αποφευχθούν τα πολλά σύμβολα). Αν όμως το μέγεθος του δείγματος είναι μικρό (< 30), τότε η καλύτερη προσέγγιση της τυπικής απόκλισης του πληθυσμού δίνεται από τον τύπο:

$$S = \sqrt{\frac{\sum(X - \overline{X})^2}{n - 1}} \text{ αντί του τύπου } S' = \sqrt{\frac{\sum(X - \overline{X})^2}{n}}$$

Οι παράμετροι μ και σ έχουν μεγάλη σπουδαιότητα γιατί προσδιορίζουν με ακρίβεια τον πληθυσμό. Είναι δηλαδή δυνατό να υπολογισθεί το ποσοστό των τιμών % που περιέχονται μεταξύ δυο σημείων X_1 και X_2 του οριζόντιου άξονα. Η μόνη προϋπόθεση γι' αυτό είναι να υπολογισθεί η επιφάνεια της καμπύλης που ορίζεται από τις δυο αυτές τεταγμένες. Αν για παράδειγμα X_1= μ+σ και X_2=μ-σ, τότε μεταξύ των δυο αυτών τιμών περιλαμβάνεται το 68,27 των τιμών (Σχήμα 6) και αυτό ισχύει για οποιαδήποτε κανονική κατανομή. Ομοίως, αν τα σημεία απομακρυνθούν από το μ κατά 2σ και 3σ, τότε περιλαμβάνονται ανάμεσα στα αντίστοιχα όρια το 95,45 και 99,73 των τιμών.



Σχήμα 6. Οι επιφάνειες που περιλαμβάνονται από τα σημεία που ορίζουν διάφορες τεταγμένες.

Από τα παραπάνω προκύπτει ότι αν η απόσταση μιας τιμής από το μέσο όρο εκφρασθεί σε τυπικές αποκλίσεις και η απόσταση αυτή ονομασθεί z, τότε κάθε τιμή του X ισούται με

X = μ ± zσ και συνεπώς

$$z = \pm \frac{X - \mu}{\sigma}$$

Συνεπώς, είναι δυνατό να μετατραπούν όλες οι τιμές X σε μονάδες z και η μορφή καμπύλης που θα αποκτηθεί με τον τρόπο αυτό ονομάζεται τυπική μορφή και χαρακτηρίζεται από το γεγονός ότι η συνολική της επιφάνεια ισούται με τη μονάδα (Σχήμα 6). Οποιαδήποτε κανονική κατανομή είναι δυνατό να μετατραπεί στη τυπική μορφή. Αυτό καθιστά δυνατή τη μετατροπή μιας κατανομής συχνότητας σε κατανομή πιθανότητας. Έτσι, είναι εφικτό να συγκριθούν κατανομές με διαφορετικούς μέσους όρους και τυπικές αποκλίσεις, ακόμα και αν οι μονάδες μέτρησης είναι διαφορετικές. Για το σκοπό αυτό υπάρχουν ειδικοί πίνακες που δίνουν την επιφάνεια α, που βρίσκεται κάτω από την καμπύλη που ορίζουν δυο οποιαδήποτε σημεία (Πίνακας Ι στο Παράρτημα).

Για να γίνουν περισσότερο κατανοητά αυτά που αναφέρθηκαν έως τώρα θα χρησιμοποιηθεί το ακόλουθο παράδειγμα.

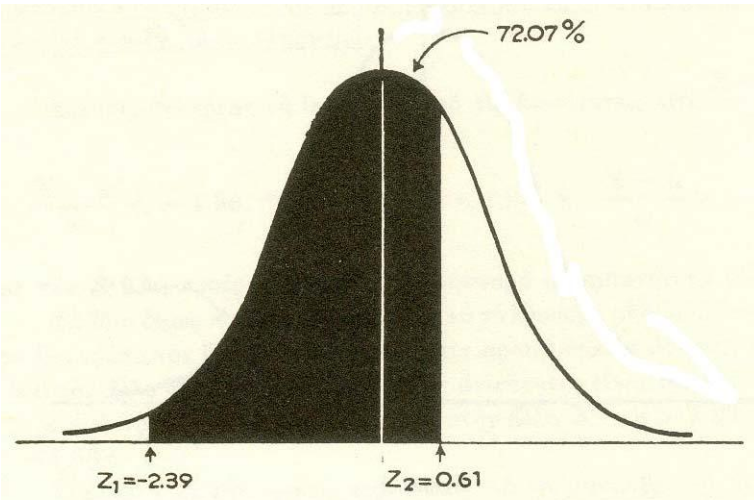
Παράδειγμα: Να υπολογισθεί πόσοι από του 100 φοιτητές ενός πληθυσμού με μέσο όρο = 5,89 και τυπική απόκλιση σ = 0,999, έχουν βαθμούς που κυμαίνονται μεταξύ 4 και 6.

Λύση.

Ο αριθμός 4 περιλαμβάνει τους αριθμούς από 3,5 έως 4,5, ενώ ο αριθμός 6 περιλαμβάνει του αριθμούς από 5,5 έως 6,5. Από αυτά προκύπτει ότι τα πραγματικά όρια είναι μεταξύ 3,5 και 6,5. Οι τιμές z_1 και z_2 που αντιστοιχούν στους βαθμούς 3,5 και 6,5 είναι

$$z_1 = \frac{3{,}5 - 5{,}89}{0{,}999} = -2{,}39 \text{ και } z_2 = \frac{6{,}5 - 5{,}89}{0{,}999} = 0{,}61$$

Από τον Πίνακα Ι στο Παράρτημα προκύπτει ότι για τιμή z_1 = -2,39 αντιστοιχεί επιφάνεια α = 0,4916, ενώ για την τιμή z_2 = 0,61 αντιστοιχεί επιφάνεια α = 0,2291. Από τα προηγούμενα λοιπόν προκύπτει ότι το 49,16 % των φοιτητών έχουν βαθμούς που κυμαίνονται από 3,5 έως 5,89 ενώ το 22,91% έχει βαθμούς που κυμαίνονται από 5,89 έως 6,5. Συνεπώς το 72,07 % των φοιτητών έχει βαθμούς που κυμαίνονται από 4 έως 6 (Σχήμα 7). Αυτό επιβεβαιώνεται αν προστεθούν οι θεωρητικές συχνότητες του Πίνακα 4 που αντιστοιχούν στους βαθμούς 4, 5 και 6 που δίνουν άθροισμα 72,1



Σχήμα 7. Επιφάνεια που περιλαμβάνεται μεταξύ των σημείων z_1 και z_2.

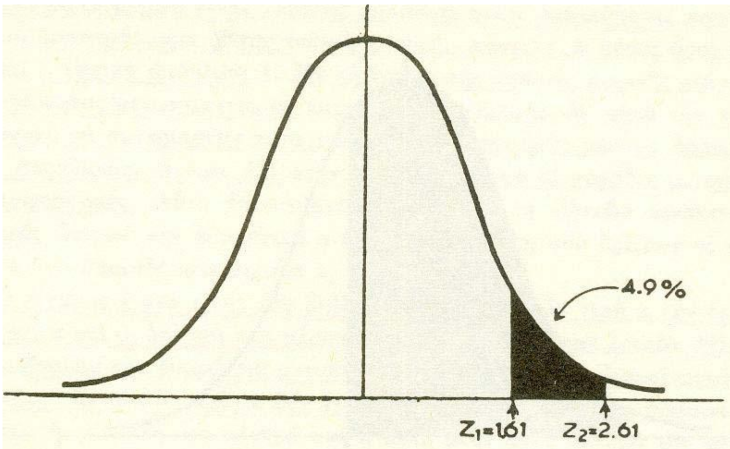
Στο προηγούμενο παράδειγμα οι τιμές z_1 και z_2 βρίσκονταν εκατέρωθεν του άξονα της καμπύλης κατανομής. Στο παράδειγμα που ακολουθεί τα σημεία θα είναι προς τη μια πλευρά του άξονα.
Παράδειγμα: Να υπολογισθεί πόσοι από του 100 φοιτητές ενός πληθυσμού με μέσο όρο = 5,89 και τυπική απόκλιση σ = 0,999, έχουν βαθμό 8.
Λύση.
Ο αριθμός 8 περιλαμβάνει τους αριθμούς από 7,5 έως 8,5. Συνεπώς, τα πραγματικά όρια είναι μεταξύ 7,5 και 8,5. Οι τιμές z_1 και z_2 που αντιστοιχούν στους βαθμούς 7,5 και 8,5 και οι αντίστοιχες α_1 και α_2 είναι

$$z_1 = \frac{7,5-5,89}{0,999} = 1,61 \text{ και } \alpha_1 = 0,4463$$
$$z_2 = \frac{8,5-5,89}{0,999} = 2,61 \text{ και } \alpha_2 = 0,4955$$

Οι δυο τιμές z έχουν το αυτό πρόσημο και αυτό σημαίνει ότι βρίσκονται προς την ίδια πλευρά του άξονα. Για το λόγο αυτό αφαιρούμε από τη μεγαλύτερη τιμή α τη μικρότερη, δηλαδή: $\alpha_2 - \alpha_1$ = 0,4955-0,4463 = 0,0492. Δηλαδή το 4,92 % των φοιτητών έχουν βαθμό 8 (Σχήμα 8). Αυτό επιβεβαιώνεται και από τη θεωρητική συχνότητα του Πίνακα 4 που αντιστοιχεί στο βαθμό 8.

Πίνακας 4. Πραγματικές και θεωρητικές συχνότητες των δεδομένων του Πίνακα 2.

Ομάδα Α			Ομάδα Β		
Μέση τιμή κλάσης X	Συχνότητα πραγματική f	Συχνότητα θεωρητική φ	Μέση τιμή κλάσης X	Συχνότητα πραγματική f	Συχνότητα θεωρητική φ
0	-	-	0	-	-
1	-	-	1	-	0,4
2	-	-	2	3	2,5
3	-	0,8 } 8,2	3	10	9,6
4	9	7,4	4	19	22,3
5	23	26,6	5	34	29,2
6	44	38,1	6	22	22,8
7	18	21,7	7	8	10,1
8	6	4,9 }5,4	8	4	2,6
9	-	0,5	9	-	0,4
10	-	-	10	-	-



Σχήμα 8. Επιφάνεια που περιλαμβάνεται μεταξύ των σημείων z_1 και z_2, όταν αυτά βρίσκονται προς την ίδια πλευρά του άξονα..

Όπως προκύπτει από τον Πίνακα I στο Παράρτημα, η τιμή z ± 1,96 περιλαμβάνει το 95% των τιμών. Έτσι, με τον τύπο

$$z = \pm \frac{X - \mu}{\sigma}$$

είναι δυνατή η μετατροπή οποιασδήποτε τιμής X σε z. Επίσης είναι δυνατό να βρεθεί αν η τιμή αυτή περιλαμβάνεται στα όρια ± 1,96. Αν η τιμή z που θα προκύψει από την παραπάνω μετατροπή περιλαμβάνεται μεταξύ των τιμών –1,96 και +1,96 (δηλαδή $-1{,}96 < z < +1{,}96$), τότε με πιθανότητα 95% η τιμή αυτή προέρχεται από τον πληθυσμό. Αν συμβαίνει το αντίθετο (δηλαδή αν $z < -1{,}96$ και $z > +1{,}96$), τότε η τιμή αυτή με πιθανότητα πάλι 95% δεν περιλαμβάνεται στο πληθυσμό. Η πιθανότητα 95% επιλέχθηκε γιατί στη Στατιστική θεωρείται ικανοποιητικό όταν είναι δυνατό να βεβαιωθεί κάτι με την πιθανότητα αυτή.

Από τη πρώτη σχέση που αναφέρθηκε προηγούμενα ($-1{,}96 < z < +1{,}96$) και τον τύπο

$$z = \pm \frac{X - \mu}{\sigma}$$

Προκύπτουν εύκολα οι ανισότητες $X < \mu + 1{,}96\sigma$ ή $X > \mu - 1{,}96\sigma$ και

$X-\mu < +1{,}96\sigma$ ή $X-\mu > -1{,}96\sigma$.

Σύμφωνα με τις δυο πρώτες, είναι δυνατό να βεβαιωθεί με πιθανότητα 95% ότι μια τιμή που βρίσκεται μέσα στα όρια $\mu-1{,}96\sigma$ και $\mu+1{,}96\sigma$ προέρχεται από τον πληθυσμό. Τα όρια αυτά, για το λόγο αυτό ονομάζονται όρια εμπιστοσύνης και αποτελούν ένα δεύτερο τρόπο για να βεβαιωθεί κάποιος ότι ο εν λόγω αριθμός προέρχεται από το συγκεκριμένο πληθυσμό.

Σύμφωνα με τις δυο τελευταίες ανισότητες η διαφορά Χ-μ πρέπει είναι > +1,96σ και < -1,96σ. Η ποσότητα ±1,96σ, που στη γενική της μορφή συμβολίζεται ως ±zσ, ονομάζεται ελάχιστη σημαντική διαφορά (ΕΣΔ) και όπως αναλύεται με λεπτομέρεια στο 8ο Κεφάλαιο αποτελεί μια αξιόπιστη μέθοδο που επιτρέπει τη σύγκριση δυο μέσων όρων.

Ένα άλλο ερώτημα που μπορεί να δημιουργηθεί είναι το εξής: είναι δυνατό να βρεθεί αν ένας φοιτητής με βαθμό 8 ανήκει στον πληθυσμό 1 ή στον πληθυσμό 2; Η απάντηση στο ερώτημα αυτό μπορεί να δοθεί κατά τους εξής διαφορετικούς τρόπους:

(α) Χρησιμοποίηση του z-κριτηρίου: εκτιμάται η τιμή z που αντιστοιχεί στο βαθμό 8. Αν η τιμή αυτή είναι >1,96 τότε υπάρχει πιθανότητα 95% ο φοιτητής να μη προέρχεται από τον Πληθυσμό 1 ή πιθανότητα 5% να προέρχεται από αυτόν. Η τιμή z ισούται με

$$z = \frac{8 - 5{,}89}{0{,}999} = 2{,}11 > 1{,}96$$

Επομένως, με πιθανότητα 95% ο φοιτητής αυτός δεν ανήκει στον πληθυσμό 1 ή με πιθανότητα. ≤ 5% ο φοιτητής αυτός προέρχεται από τον πληθυσμό 1.

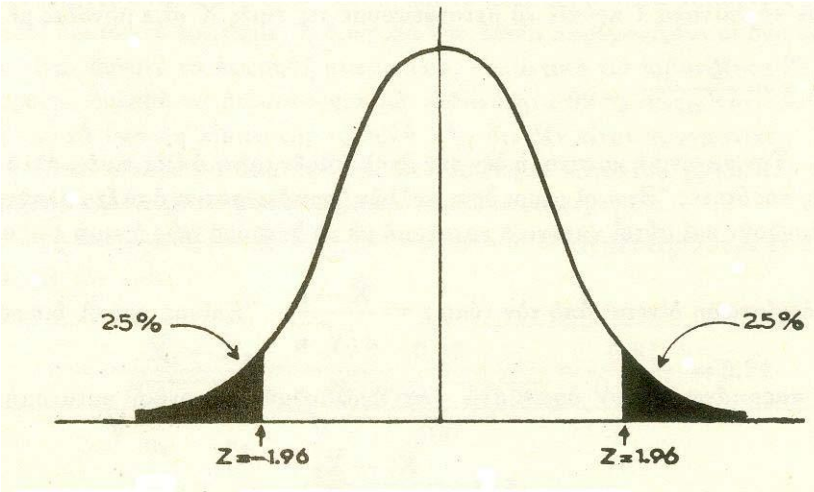
(β) Χρησιμοποίηση των ορίων εμπιστοσύνης: εκτιμώνται τα 95% όρια εμπιστοσύνης του πληθυσμού 1 που είναι από μ-1,96σ έως μ+1,96σ, δηλαδή από 5,89 – 1,96 x 0,999 = 3,93 έως 5,89 + 1,96 x 0,999 = 7,85. Επειδή η τιμή 8 δεν περιλαμβάνεται στα όρια αυτά είναι δυνατό να βγει το συμπέρασμα ότι με πιθανότητα 95% ο φοιτητής αυτός δεν προέρχεται από τον πληθυσμό 1.

(γ) Χρησιμοποίηση της ελάχιστης σημαντικής διαφοράς: υπολογίζεται η διαφορά του αριθμού από το μέσο όρο, δηλαδή Χ –μ =8 – 5,89 = 2,11. Στη συνέχεια η διαφορά αυτή συγκρίνεται με την ελάχιστη σημαντική διαφορά (ΕΣΔ) για πιθανότητα 95% που δίνεται από τον τύπο

$$\text{ΕΣΔ} = \pm z_{05}\,\sigma = \pm 1{,}96 \times 0{,}999 = \pm 1{,}96.$$

Επειδή η διαφορά του αριθμού από το μέσο όρο είναι μεγαλύτερη από την ελάχιστη σημαντική διαφορά, δηλαδή 2,11 > 1,96 βγαίνει εκ νέου το συμπέρασμα ότι ο φοιτητής αυτός με πιθανότητα 95% δεν προέρχεται από τον πληθυσμό 1.

Στην πρώτη περίπτωση, όταν χρησιμοποιήθηκε το z-κριτήριο η πιθανότητα υπολογίσθηκε και προς τις δύο πλευρές της καμπύλης, κατά τρόπο ώστε το 5% να αποτελείται κατά το μισό από τις μεγαλύτερες θετικές τιμές και κατά το υπόλοιπο μισό από τις μεγαλύτερες αρνητικές τιμές (Σχήμα 5). Αυτό το είδος της εφαρμογής του z-κριτηρίου ονομάζεται κριτήριο και προς τις δύο πλευρές της καμπύλης. Είναι όμως δυνατό η εφαρμογή του z-κριτηρίου να γίνει μόνο προς τη μία πλευρά της καμπύλης, οπότε ονομάζεται κριτήριο προς την μία μόνο πλευρά της καμπύλης. Για να γίνει περισσότερο κατανοητή η περίπτωση αυτή θα χρησιμοποιηθεί ένα παράδειγμα.



Σχήμα 9. Εφαρμογή του z-κριτηρίου και προς τις δύο πλευρές της καμπύλης με πιθανότητα 5%.

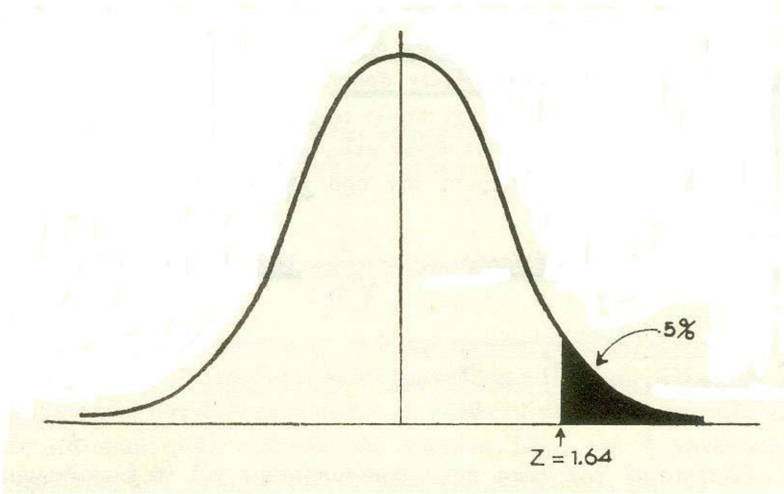
Παράδειγμα εφαρμογής του z-κριτηρίου προς τη μία πλευρά της καμπύλης.

Ζητείται να βρεθεί με πιθανότητα 95% αν ένας φοιτητής που είχε βαθμό 8 ανήκει στο Πληθυσμό 1.

Λύση.

Όπως είναι φανερό, στην περίπτωση αυτή το 5% αποτελείται μόνο από τις μεγαλύτερες θετικές τιμές. Από τον Πίνακα I στο Παράρτημα προκύπτει ότι η τιμή z που αφήνει δεξιά της το 5% των υψηλότερων τιμών είναι 1,64. Κάθε βαθμός που αντιστοιχεί σε τιμή z μεγαλύτερη ή ίση του 1,64 ανήκει στο 5% των υψηλότερων τιμών. Ο βαθμός 8 έχει τιμή z = 2,11 που είναι μεγαλύτερη από την τιμή 1,64. Επομένως ο φοιτητής με βαθμό 8 ανήκει στο 5% των καλύτερων φοιτητών.

Ο πίνακας II στο Παράρτημα δίνει τις τιμές του z για κάθε ένα από τα δύο κριτήρια και για διάφορες πιθανότητες.



Σχήμα 10. Εφαρμογή του z-κριτηρίου και προς τη μία πλευρά της καμπύλης με πιθανότητα 5%.

Θα πρέπει εδώ να σημειωθεί ότι την κανονική κατανομή, εκτός από τις απλές τιμές, την ακολουθούν και άλλες ποσότητες, όπως για παράδειγμα οι μέσοι όροι πολλών ίσων δειγμάτων ενός πληθυσμού. Στην περίπτωση όμως αυτή η τιμή του z δίνεται από τον παρακάτω τύπο

$$z = \overline{X} - \mu / \sigma \sqrt{n}$$

Οι διαφορές των παραπάνω μέσων όρων ανά δύο ακολουθούν επίσης κανονική κατανομή και η τιμή του z δίνεται από τον τύπο

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Ο παρονομαστής του πρώτου τύπου ονομάζεται τυπικό σφάλμα των μέσων όρων και ο παρανομαστής του δεύτερου τυπικό σφάλμα διαφοράς των μέσων όρων.

Ένα τελευταίο ερώτημα που μπορεί να δημιουργηθεί είναι το εξής: είναι δυνατό να βρεθεί αν οι μέσοι όροι των δύο δειγμάτων διαφέρουν μεταξύ τους στατιστικώς σημαντικά για πιθανότητα 99%; (το πρώτο δείγμα ως γνωστό έχει μέσο όρο $\overline{X}_1$= 5,89 και σ_1= 0,999, ενώ το δεύτερο έχει $\overline{X}_2$= 5,02 και σ_2 = 1,326, σελίδα 14).

Λύση.

Το ερώτημα μπορεί να απαντηθεί με το z-κριτήριο, γιατί τα δείγματα είναι μεγάλα. Αρχικά βρίσκεται η διαφορά $\overline{X}_1 - \overline{X}_2$ = 5,89 –5,02 = 0,87. Η διαφορά αυτή μετατρέπεται σε μονάδες z με τον τύπο

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0{,}87}{\sqrt{\frac{0{,}999+1{,}7596}{100}}} = \frac{0{,}87}{0{,}1661} = 5{,}24$$

Επειδή η τιμή z = 5,24 είναι μεγαλύτερη της τιμής $z_{.01}$ = 2,57 (Πίνακας ΙΙ, πιθανότητα 1%, Παράρτημα) βγαίνει το συμπέρασμα ότι η διαφορά των δυο μέσων όρων είναι στατιστικώς σημαντική για πιθανότητα 99%.

3.2.2. Κατανομή t.

Όταν το μέγεθος του δείγματος είναι μικρό (n < 30) τότε ο τύπος που δίνει την καλύτερη προσέγγιση της τυπικής απόκλισης είναι (βλέπε σελίδα 18)

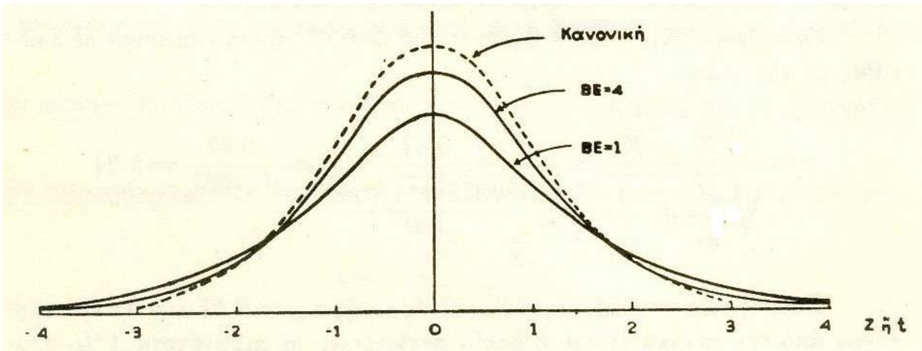
$$S = \sqrt{\frac{\Sigma(X-\bar{X})^2}{n-1}}$$

Στις περιπτώσεις αυτές η ποσότητα που προκύπτει μετά από αντικατάσταση του σ από το s, δεν ακολουθεί την κανονική κατανομή, αλλά μια νέα που ονομάζεται t-κατανομή. Η μορφή της t-κατανομής είναι επίσης κωδωνοειδής, όμως τη φορά αυτή η επιφάνεια της εξαρτάται από το μέγεθος του δείγματος (n), αφού αφαιρεθεί μια μονάδα (n-1). Η τιμή n-1 ονομάζεται βαθμοί ελευθερίας (Β. Ε.). Στην t-κατανομή η απόσταση από το σημείο της μεγαλύτερης συχνότητας σε τυπικές αποκλίσεις συμβολίζεται με το t και η τιμή του τελευταίου εξαρτάται, όπως είναι αναμενόμενο, από τους βαθμούς ελευθερίας. Βαθμοί ελευθερίας (ΒΕ) σημαίνει ελευθερία να ποικίλλουν. Ας υποτεθεί ότι υπάρχουν έξι τιμές και ο μέσος όρος τους είναι 10. Η έκτη τιμή πρέπει να προσαρμοστεί ώστε να διορθωθεί η παραλλακτικότητα που προέκυψε από τις 5 τιμές και τελικά ο μέσος όρος τους να είναι 10. Για παράδειγμα, αν οι 5 τιμές είναι 10, 12, 18, 16, και 4 για να είναι ο μέσος όρος τους 10 η έκτη τιμή πρέπει να είναι 0. Αν οι αριθμοί είναι 2, 8, 4, 6, και 10 τότε η έκτη τιμή πρέπει να είναι 30 ώστε ο μέσος όρος των 6 τιμών να είναι πάλι 10. Και στις δύο αυτές περιπτώσεις οι βαθμοί ελευθερίας είναι 5. Και στις δύο περιπτώσεις οι 5 παρατηρήσεις μπορούν να έχουν οποιαδήποτε τιμή το μέγεθος όμως της έκτης είναι καθορισμένο επειδή ο μέσος όρος των τιμών είναι 10.

Αν ληφθούν n τιμές για τον υπολογισμό του μέσου όρου καταναλώνεται ένας βαθμός ελευθερίας. Μετά τον υπολογισμό του μέσου όρου οι ΒΕ είναι n-1.

Όταν τα δεδομένα είναι σε ζευγάρια ο αριθμός των βαθμών ελευθερίας είναι ίσος με τον αριθμό των ζευγαριών μείον 1.

Στον Πίνακα ΙΙΙ στο παράρτημα δίνονται οι τιμές του t, για διάφορες πιθανότητες και Β. Ε. Από τη μελέτη του Πίνακα ΙΙΙ προκύπτει ότι για πιθανότητα 5% και Β. Ε. 1 και ∞ οι αντίστοιχες τιμές του t είναι 12,706 και 1,96. Δηλαδή, όταν το μέγεθος του δείγματος γίνεται ∞, τότε οι τιμές t και z συμπίπτουν. Αυτό σημαίνει ότι όσο αυξάνονται οι Β, Ε. η καμπύλη της t-κατανομής τείνει να συμπέσει με την αντίστοιχη καμπύλη της κανονικής. Αντίθετα, όσο περισσότερο μικραίνει ο αριθμός των Β. Ε. τόσο τα άκρα της καμπύλης της t-κατανομής απομακρύνονται από τον άξονα (Σχήμα 11).



Σχήμα 11. Η καμπύλη της t-κατανομής σε σχέση με την αντίστοιχη της κανονικής κατανομής, για διάφορους Β. Ε.

Ότι αναφέρθηκε για το z-κριτήριο ισχύει και για το t-κριτήριο, μόνο που το τελευταίο εφαρμόζεται στα μικρά δείγματα και ακολουθεί t-κατανομή. Οι εφαρμογές του t-κριτηρίου είναι περισσότερες από τις αντίστοιχες του z-κριτηρίου γιατί στο γεωργικό πειραματισμό τις περισσότερες φορές χρησιμοποιούνται μικρά δείγματα. Ένα άλλο πλεονέκτημα που έχει το t-κριτήριο είναι ότι μπορεί να χρησιμοποιηθεί τόσο σε μικρά δείγματα, όσο και σε μεγάλα, γιατί στα μεγάλα δείγματα αν διαιρεθεί η τυπική απόκλιση δια του n-1 ή n, η τιμή της επηρεάζεται ελάχιστα. Αντίθετα, στα μικρά δείγματα η τιμή της τυπικής απόκλισης επηρεάζεται πολύ αν διαιρεθεί δια του n-1 ή n και για το λόγο αυτό η z–κατανομή δεν εφαρμόζεται σε αυτά.

Οι μέσοι όροι των μικρών δειγμάτων ακολουθούν t-κατανομή με τυπικό σφάλμα

$$S_{\bar{X}} \doteq \frac{S}{\sqrt{n}}$$

Ο αριθμός των τυπικών σφαλμάτων που χωρίζει το μέσο όρο του δείγματος ($\overline{X}$) από αυτόν του πληθυσμού (μ) είναι

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Οι διαφορές που παρουσιάζουν ανά δυο οι μέσοι όροι ακολουθούν t-κατανομή με μέσο όρο 0 και τυπικό σφάλμα

$$S_d = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} =$$

Τέλος, ο τύπος που δίνει σε αριθμό τυπικών σφαλμάτων την απόσταση από το 0 μιας διαφοράς μέσων όρων είναι

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Οι δυο αυτοί τύποι κάνουν δυνατή τη σύγκριση δυο μέσων όρων και τη διαπίστωση της ύπαρξης ή όχι στατιστικών διαφορών.

3.2.3. Κατανομή F.

Τα δυο είδη κατανομών που μελετήθηκαν έως τώρα επιτρέπουν τη σύγκριση των μέσων όρων ανά δύο. Σε περίπτωση που οι μέσοι όροι είναι πολλοί και επιζητείται η ταυτόχρονη σύγκρισή τους, τα δύο προαναφερθέντα κριτήρια δεν μπορούν να ανταποκριθούν. Για να απαντηθεί το ερώτημα που δημιουργείται θα πρέπει να γίνουν τόσες επιμέρους συγκρίσεις, όσοι είναι οι συνδυασμοί των μέσων όρων, όταν συγκρίνονται ανά δύο. Για να συγκριθούν λοιπόν ταυτόχρονα πολλοί μέσοι όροι, θα πρέπει να βρεθεί κάποιος άλλος τρόπος και εδώ ακριβώς μπαίνει η έννοια του μέσου τετραγώνου ή διακύμανσης. Μέσο τετράγωνο ή διακύμανση είναι το τετράγωνο της τυπικής απόκλισης (s^2). Αν από ένα πληθυσμό ληφθούν π δείγματα και το κάθε δείγμα έχει n τιμές, τότε η τυπική απόκλιση σ του πληθυσμού (και συνεπώς και η διακύμανση σ^2) είναι δυνατό να υπολογισθεί:

(α) υπολογίζοντας τη διακύμανση κάθε δείγματος χωριστά και στη συνέχεια τη μέση διακύμανση, αθροίζοντας τις διακυμάνσεις και διαιρώντας με τον αριθμό των δειγμάτων π. Ο τύπος της μέσης διακύμανσης είναι

$$S_1^2 = \frac{\sum(X_1 - \overline{X}_1)^2 + \sum(X_2 - \overline{X}_2)^2 + \cdots \sum(X_\pi - \overline{X}_\pi)^2}{\pi(n-1)}$$

(β) υπολογίζοντας τη διακύμανση των μέσων όρων των δειγμάτων και από αυτή τη διακύμανση του πληθυσμού. Η διακύμανση των μέσων όρων των δειγμάτων είναι

$$S_{\overline{X}}^2 = \frac{(\overline{X}_1 - \overline{X})^2 + (\overline{X}_2 - \overline{X})^2 + \cdots. + (\overline{X}_\pi - \overline{X})^2}{\pi - 1}$$

όπου $\overline{X}_1$, $\overline{X}_2$, ... $\overline{X}_\pi$ οι μέσοι όροι των δειγμάτων και $\overline{X}$ ο γενικός μέσος όρος.

Από το τύπο $$S_{\bar{X}} = \frac{S}{\sqrt{n}}$$

προκύπτει ότι $$S^2 = n\, S_{\bar{X}}^2$$

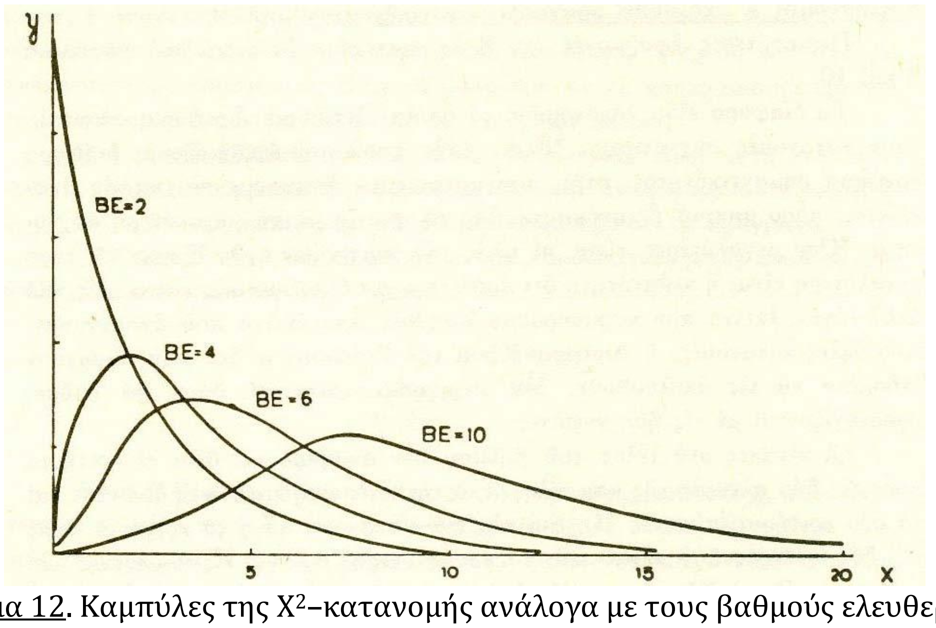
Αν τα δείγματα προέρχονται από τον ίδιο πληθυσμό τότε $S_1^2 = S_2^2$. Αν τώρα από ένα πληθυσμό ληφθούν π δείγματα, που το καθένα τους έχει n τιμές ο λόγος $F= S_2^2/\ S_1^2$ ακολουθεί μια άλλη κατανομή που ονομάζεται F–κατανομή. Όταν τα δείγματα προέρχονται από τον ίδιο πληθυσμό $F= S_2^2/\ S_1^2 = 1$. Όσο μεγαλύτερος γίνεται ο αριθμητής του προηγούμενου κλάσματος τόσο το F παίρνει τιμές > 1 και επομένως αυξάνει και η πιθανότητα οι μέσοι όροι των δειγμάτων να διαφέρουν σημαντικά. Η χρησιμοποίηση του F-κριτηρίου θα αναλυθεί περισσότερο στο 8ο Κεφάλαιο, όπου θα αναλυθούν ορισμένες τεχνικές της βιομετρίας, που είναι γνωστές με το όνομα Ανάλυση παραλλακτικότητας.

3.2.4. Κατανομή x^2.

Το κριτήριο αυτό χρησιμοποιείται στις περιπτώσεις που υπάρχει η ανάγκη να γίνουν συγκρίσεις μεταξύ ποσοτήτων που αναφέρονται στον κατακόρυφο άξονα της κατανομής (π. χ. συχνότητες, αναλογίες). Στις περιπτώσεις αυτές χρησιμοποιείται το x^2-κριτήριο, που ισούται:

$$X^2 = \frac{(f_1 - \varphi_1)^2}{\varphi_1} + \frac{(f_2 - \varphi_2)^2}{\varphi_2} + \cdots . + \frac{(f_\kappa - \varphi_\kappa)^2}{\varphi_\kappa}$$

όπου f_1, f_2 …….f_κ είναι οι πραγματικές συχνότητες και φ_1, φ_2 …..φ_κ είναι οι θεωρητικές. Όσο πιο μεγάλη είναι η τιμή του X^2, τόσο περισσότερες είναι οι πιθανότητες οι συχνότητες να διαφέρουν σημαντικά. Το X^2 ακολουθεί μια κατανομή που ονομάζεται X^2-κατανομή.



Σχήμα 12. Καμπύλες της X^2–κατανομής ανάλογα με τους βαθμούς ελευθερίας.

Η καμπύλη της X^2-κατανομής εξαρτάται από τους βαθμούς ελευθερίας και στο Σχήμα 12 υπάρχει η γραφική παράσταση της κατανομής αυτής για διάφορους βαθμούς ελευθερίας. Τέλος ο Πίνακας VI στο Παράρτημα δίνει τις τιμές του X^2– κριτηρίου για διαφορετικές πιθανότητες και Β. Ε.

Για να γίνει περισσότερο κατανοητή η χρησιμοποίηση του X^2-κριτηρίου θα χρησιμοποιηθούν τα δεδομένα του Πίνακα 4, σελίδα 20 που αναφέρονται στην πρώτη ομάδα φοιτητών. Με τα δεδομένα αυτά θα διερευνηθεί αν οι συχνότητες που παρατηρήθηκαν ακολουθούν ή όχι κανονική κατανομή. Από τα δεδομένα λοιπόν του Πίνακα 3 προκύπτει ότι

$$X^2 = \frac{(9-8{,}2)^2}{8{,}2} + \frac{(23-26{,}6)^2}{26{,}6} + \frac{(44-38{,}1)^2}{38{,}1} + \frac{(18-21{,}7)^2}{21{,}7} + \frac{(6-5{,}4)^2}{5{,}4} =$$
$$= 2{,}17$$

Οι Βαθμοί Ελευθερίας είναι ίσοι με τον αριθμό των κλάσεων μείον 3:Β. Ε. = 5 – 3 = 2. Από τον Πίνακα VI (Παράρτημα) προκύπτει ότι η τιμή του X^2 που αντιστοιχεί σε πιθανότητα 5% και 2 Β. Ε. είναι 5,99, που είναι μεγαλύτερη από την τιμή 2,17 που βρέθηκε προηγούμενα. Αυτό σημαίνει ότι ο Α πληθυσμός των φοιτητών ακολουθεί κανονική κατανομή.

3.2.5. Λοιπές κατανομές.

Οι τέσσερις κατανομές που περιγράφηκαν έως τώρα είναι οι πιο συνηθισμένες. Υπάρχουν όμως και άλλες τρεις κατανομές, που συναντώνται λιγότερο συχνά και για το λόγο αυτό θα περιγραφούν με συντομία.

3.2.5.1. Διωνυμική κατανομή.

Έστω ότι κάποιος έχει δύο φασόλια, ένα άσπρο και ένα μαύρο. Αν τραβήξει στην τύχη ένα φασόλι τότε οι μισές πιθανότητες είναι να τραβήξει το άσπρο και οι υπόλοιπες να τραβήξει το μαύρο. Αν μετά τη λήψη του ενός από τα δύο φασόλια το επανατοποθετήσει και τραβήξει ξανά, τότε η πιθανότητα να τραβήξει εκ νέου άσπρο (ή μαύρο) δεν θα επηρεαστεί από την προηγούμενη λήψη. Στην περίπτωση αυτή, η πιθανότητα της τυχαίας μεταβλητής που δίνει τον αριθμό των άσπρων ή μαύρων φασολιών, ακολουθεί ένα άλλο είδος κατανομής που ονομάζεται <u>διωνυμική</u>. Η δυωνυμική κατανομή είναι αρκετά σπουδαία, γιατί την ακολουθεί η μεταβίβαση των γονιδίων, όταν η αναπαραγωγή των ατόμων είναι εγγενής. Ο γενικός τύπος της διωνυμικής κατανομής είναι

$$\mathbf{f(y)}\left\{\begin{matrix}\mathbf{n}\\ \\ \\ \mathbf{y}\end{matrix}\right\} = \mathbf{p^y q^{n-y}} = \frac{\mathbf{n!}}{\mathbf{y!\,(n-y)!}} = \mathbf{p^y q^{n-y}}$$

όπου: f(y) η πιθανότητα να ληφθούν n άσπρα φασόλια και (n-y) μαύρα.
n ο αριθμός των λήψεων
p η πιθανότητα λήψης άσπρου φασολιού και
q η πιθανότητα λήψης μαύρου φασολιού.
Ο μέσος όρος και η τυπική απόκλιση της διωνυμικής κατανομής δίνονται από τους τύπους

$$\mu = n\,p \quad \text{και} \quad \sigma = \sqrt{npq}$$

Όταν το n τείνει προς το άπειρο, η διωνυμική κατανομή τείνει προς την κανονική. Στην περίπτωση αυτή για να υπολογισθούν οι πιθανότητες της διωνυμικής κατανομής χρησιμοποιείται η κανονική προσέγγιση της τελευταίας. Για να είναι ικανοποιητική η προσέγγιση της κανονικής από τη διωνυμική:
(α) αν p = q = 0,5 θα πρέπει το n ≥ 10 και
(β) αν p≠ q≠ 0,5, θα πρέπει np > 15 και nq > 15.

3.2.5.2. **Πολυωνυμική κατανομή**.

Αποτελεί γενίκευση της διωνυμικής κατανομής Διαφέρει από τη διωνυμική γιατί στην περίπτωση αυτή σε κάθε απλό πείραμα τύχης υπάρχουν περισσότερες από δύο εκβάσεις. Για παράδειγμα, αν γίνει λήψη με επσνατοποθέτηση άσπρων, μαύρων και πιτσιλωτών φασολιών. η συνάρτηση της πιθανότητας της κατανομής αυτής είναι

$$f(y_1, y_2, \dots y_\kappa) = \frac{n!}{y_1!\, y_2! \dots . y_\kappa!}\, p_1^{y_1} p_2^{y_2} \dots . p_\kappa^{y_\kappa}$$

όπου $y_1 + y_2 + \dots + y_k = n$ και $P_1 + P_2 + \dots + P_k = 1$
Το $f(y_1, y_2, \dots y_k)$ είναι η πιθανότητα σε n επαναλήψεις του πειράματος να υπάρχουν y_1 εκβάσεις από την κάθε περίπτωση των k δυνατών εκβάσεων. Ο παρονομαστής του κλάσματος είναι ο αριθμός των επαναληπτικών μεταθέσεων ή πραγμάτων, από τα οποία τα y_1 είναι όμοια μεταξύ τους, τα y_2 είναι επίσης όμοια μεταξύ τους κ. ο. κ. Τα $P_1^{y1}, P_2^{y2}, \dots$ είναι οι πιθανότητες να υπάρχουν αντίστοιχα y_1 εκβάσεις του πρώτου είδους, y_2 εκβάσεις του δεύτερου είδους κ. ο. κ.

3.2.5.3. **Κατανομή Poisson**.

Στην περίπτωση που το n είναι μεγάλο ($n \longrightarrow \infty$) και το p είναι πολύ μικρό ($p \longrightarrow 0$) τότε το np παίρνει τιμή μικρότερη από 15. Αν το np παραμένει σταθερό και ίσο με μ, τότε η διωνυμική κατανομή τείνει προς την κατανομή του Poisson. Η συνάρτηση πιθανότητας της τελευταίας είναι

$$f(y) = \frac{\mu^y}{y!}\, e^{-\mu} \qquad y = 0, 1, 2, \dots$$

όπου f(y) η πιθανότητα εμφάνισης y σπανίων γεγονότων, np = μ και e= 2,71828

3.3. **Εφαρμογές των κριτηρίων**.

Εκτός από τις περιπτώσεις που αναφέρθηκαν όταν εξετάζονταν το κάθε κριτήριο χωριστά, τα τελευταία μπορούν να χρησιμοποιηθούν σε διάφορες επιπλέον περιπτώσεις. Οι πιο κοινές από αυτές θα περιγραφούν στη συνέχεια.

3.3.1. **Σύγκριση του μέσου όρου με ένα γνωστό αριθμό**.

Υπάρχουν πολλές περιπτώσεις όπου ένας ερευνητής επιθυμεί να γνωρίζει αν ο μέσος όρος μιας ποικιλίας διαφέρει από κάποιο συγκεκριμένο αριθμό, γνωστό εκ των προτέρων, που μπορεί να είναι ο μέσος όρος μιας άλλης εμπορικής ποικιλίας. Για να γίνει αυτό κατανοητός ο τρόπος με τον οποίο εργαζόμαστε όταν έχουμε να αντιμετωπίσουμε τέτοια θέματα, θα χρησιμοποιηθούν δύο παραδείγματα, ένα με μικρό μέγεθος δείγματος και ένα στο οποίο το μέγεθος του δείγματος είναι μεγάλο.

Παράδειγμα 1ο: μέγεθος δείγματος μεγάλο.

Ένας ερευνητής μέτρησε το ύψος 48 φυτών μιας ποικιλίας σιταρόβριζας. Στη συνέχεια, υπολόγισε τον μέσο όρο του ύψους των φυτών αυτών, που ήταν 94,63cm, ενώ το τυπικό σφάλμα ήταν $\sigma_{\bar{X}} = 1{,}22$. Το ερώτημα που τίθεται είναι αν διαφέρει η συγκεκριμένη ποικιλία από μια άλλη, της οποίας το ύψος είναι 89,50 cm.

Λύση.

Επειδή το μέγεθος του δείγματος είναι μεγάλο, θα χρησιμοποιηθεί το z-κριτήριο. Αρχικά θα υπολογισθεί από τα δεδομένα μια τιμή z, που θα συγκριθεί με τη θεωρητική, για πιθανότητα 5%, που είναι $z_{.05} = 1.96$ (πίνακας ΙΙ στο Παράρτημα).

Η τιμή του z είναι

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{94{,}63 - 89{,}50}{1{,}22}\ \frac{5{,}13}{1{,}22} = 4{,}2 > 1{,}96$$

Επειδή η τιμή του z = 4,2 είναι μεγαλύτερη από τη θεωρητική τιμή του z, για πιθανότητα σφάλματος 5%, που είναι 1,96, βγαίνει το συμπέρασμα ότι το ύψος της συγκεκριμένης ποικιλίας διαφέρει στατιστικώς σημαντικά από το ύψος της άλλης ποικιλίας.

Το ίδιο ερώτημα μπορεί να απαντηθεί χρησιμοποιώντας τα όρια εμπιστοσύνης που είναι

$$\bar{X} \pm z_{.05}\,\sigma_{\bar{X}} = 94{,}63 \pm 1{,}96 \text{ X } 1{,}22 = 94{,}63 \pm 2{,}39$$

Έτσι, για πιθανότητα 5% τα όρια εμπιστοσύνης του μέσου όρου του ύψους της ποικιλίας είναι από 92,24 έως 97,02 εκατοστά. Επειδή η τιμή 89,50 είναι έξω από τα όρια αυτά, βγαίνει το συμπέρασμα ότι οι δυο μέσοι όροι διαφέρουν μεταξύ τους στατιστικώς σημαντικά.

Παράδειγμα 2ο: δείγματα μικρά.

Κάποιος κατάγραψε τα ύψη 23 φυτών ενός πληθυσμού βρίζας και υπολόγισε το μέσο όρο σε 117,13 cm και το τυπικό σφάλμα σε $S_{\overline{X}} = 2{,}05$. Να βρεθεί αν το ύψος του πληθυσμού αυτού διαφέρει για πιθανότητα 5% από το ύψος ενός άλλου πληθυσμού, που ήταν 115,65 cm.

Λύση.

Εδώ, επειδή το μέγεθος του δείγματος είναι μικρό θα χρησιμοποιηθεί το t-κριτήριο. Αρχικά θα υπολογισθεί από τα δεδομένα μια τιμή t, που θα συγκριθεί με τη θεωρητική, για πιθανότητα 5% και Β. Ε. = n - 1= 23 – 1 = 22, που είναι $t_{.05,\,22} = 2{,}074$ (πίνακας III στο Παράρτημα).

$$t = \frac{\overline{X} - \mu}{S_{\overline{X}}} = \frac{117{,}13 - 115{,}65}{2{,}05} = \frac{1{,}48}{2{,}05} = 0{,}72 < 2{,}064$$

Επειδή η τιμή του t που υπολογίσθηκε από τα δεδομένα (0,72) είναι μικρότερη από τη θεωρητική, βγαίνει το συμπέρασμα ότι οι δυο μέσοι όροι δεν διαφέρουν στατιστικώς σημαντικά.

Όπως και στην προηγούμενη περίπτωση, το ερώτημα αυτό μπορεί να απαντηθεί χρησιμοποιώντας τα όρια εμπιστοσύνης που είναι $\overline{X} \pm t_{.05}\ S_{\overline{X}}$ = 117,13 ± 2,074 x 2,05, δηλαδή 117,13 ± 4,25. Έτσι, για πιθανότητα 5% τα όρια εμπιστοσύνης του μέσου όρου του ύψους του πληθυσμού είναι από 112,88 έως 121,38 εκατοστά. Επειδή η τιμή 115,65 είναι μέσα στα όρια αυτά, βγαίνει το συμπέρασμα ότι οι δυο μέσοι όροι δε διαφέρουν μεταξύ τους στατιστικώς σημαντικά.

3.3.2. Σύγκριση δύο μέσων όρων.

Στην περίπτωση αυτή διακρίνονται τέσσερις περιπτώσεις, που επεξηγούνται με αντίστοιχα παραδείγματα.

3.3.2.1. Μέγεθος δείγματος μεγάλο (n > 30).

Παράδειγμα.

Κάποιος θέλησε να συγκρίνει δύο ποικιλίες σιταριού που είχαν: η πρώτη $\overline{X}_1 = 85,\ 2$, $\sigma_1^2 = 65{,}7$ και $n_1 = 48$ και $\overline{X}_2 = 81{,}4$, $\sigma_2^2 = 84{,}5$ και $n_2 = 65$.

Λύση.

Θα χρησιμοποιηθεί το z-κριτήριο, σύμφωνα με το οποίο θα συγκριθεί η τιμή z που θα υπολογισθεί από τα δεδομένα με τη θεωρητική τιμή για πιθανότητα σφάλματος 5%, που είναι $z_{.05} = 1.96$ (πίνακας II στο Παράρτημα). Η τιμή του z υπολογίζεται από τον τύπο:

$$z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}} = \frac{85{,}2 - 83{,}4}{\sqrt{\frac{65{,}7}{48} + \frac{84{,}5}{65}}} = \frac{1{,}8}{1{,}634} = 1{,}102$$

Επειδή η τιμή z = 1,102 είναι μικρότερη από τη θεωρητική τιμή $z_{.05}$ = 1.96 βγαίνει το συμπέρασμα ότι οι δύο ποικιλίες σιταριού δε διαφέρουν ως προς το ύψος.

Το ίδιο ερώτημα μπορεί να απαντηθεί με άλλους δύο τρόπους:

(α) <u>Με τα όρια εμπιστοσύνης</u>: τα όρια εμπιστοσύνης της διαφοράς δύο μέσων όρων και για πιθανότητα σφάλματος 5% υπολογίζονται από τον τύπο

$$(\overline{X}_1 - \overline{X}_2) \pm z_{.05}\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} = 1{,}8 \pm 1{,}96\ x\ 1{,}634$$

οπότε τα όρια είναι μεταξύ –1,400 και 5,003. Επειδή τα όρια αυτά περιέχουν την τιμή μηδέν οι διαφορές δεν είναι στατιστικά σημαντικές.

(β) Με την ελάχιστη σημαντική διαφορά (ΕΣΔ): η ελάχιστη σημαντική διαφορά υπολογίζεται από τον τύπο

$$ΕΣΔ = \pm z_{.05}\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} = 1{,}96\ x\ 1{,}634 = 3{,}203$$

Επειδή τώρα η διαφορά των δύο μέσων όρων είναι μικρότερη από την ΕΣΔ, δηλαδή $\overline{X}_1 - \overline{X}_2$ = 1,8 < 3,203, οι μέσοι όροι δε διαφέρουν στατιστικώς σημαντικά.

3.3.2.2. <u>Μέγεθος δείγματος μικρό (n < 30) – τα δείγματα ανεξάρτητα με ίσες διακυμάνσεις</u>.

<u>Παράδειγμα</u>.

Μελετήθηκε η απόδοση δυο ομάδων φοιτητών στο μάθημα της Γενετικής. Η πρώτη ομάδα, που την αποτελούσαν 21 φοιτητές, είχε μέσο όρο $\overline{X}_1$ = 6,7 και τιμή $\Sigma(X_1 - \overline{X}_1)^2$ = 52,10. Η δεύτερη ομάδα, που την αποτελούσαν 27 φοιτήτριες, είχε μέσο όρο $\overline{X}_2$ = 6,2 και τιμή $\Sigma(X_2 - \overline{X}_2)^2$ = 41,16. Ζητείται να βρεθεί με πιθανότητα σφάλματος αν οι φοιτητές είχαν καλύτερη επίδοση από τις φοιτήτριες.

<u>Λύση</u>.

Επειδή το μέγεθος του δείγματος είναι μικρό, θα χρησιμοποιηθεί το t–κριτήριο. Η τιμή του t που θα υπολογισθεί από τα δεδομένα θα συγκριθεί με την θεωρητική τιμή t, για πιθανότητα σφάλματος 5% και ΒΕ = $n_1 + n_2 - 2$ = 21 +27 – 2 = 46, που είναι $t_{.05,46}$ = 2,020. Η τιμή του t υπολογίζεται από τον τύπο:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{n_1 + n_2}{n_1 n_2} \, S^2}}$$

Όπου

$$s^2 = \frac{\sum(X_1 - \overline{X}_1)^2 + \sum(X_1 - \overline{X}_1)^2}{n_1 + n_2 - 2}$$

Αν γίνουν οι αντικαταστάσεις και οι πράξεις προκύπτει ότι το $S^2 = 2{,}027$ και $t = 1{,}208$. Επειδή η τιμή αυτή είναι μικρότερη από τη θεωρητική ($t_{.05} = 2{,}020$), βγαίνει το συμπέρασμα ότι οι φοιτητές δεν είναι καλύτεροι από τις φοιτήτριες στο μάθημα της Γενετικής.

Τα όρια εμπιστοσύνης και η ΕΣΔ δίνονται από τους τύπους

$$(\overline{X}_1 - \overline{X}_2) \pm t_{.05}\sqrt{\frac{n_1+n_2}{n_1 n_2} \, S^2}$$

$$\text{και} \quad \text{ΕΣΔ} = t_{.05} \sqrt{\frac{n_1+n_2}{n_1 n_2} \, S^2}$$

3.3.2.3. <u>Μέγεθος δείγματος μικρό (n < 30) – τα δείγματα ανεξάρτητα με διαφορετικές διακυμάνσεις</u>.

Στην περίπτωση αυτή υπολογίζεται μια τιμή t από τον τύπο

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

και η τιμή αυτή συγκρίνεται με τη θεωρητική τιμή που υπολογίζεται από τον τύπο

$$t' = \frac{t_1 {s_1}^2/n_1 + t_2 {s_2}^2/n_2}{{s_1}^2/n_1 + {s_2}^2/n_2}$$

όπου t_1 η τιμή του t για ΒΕ $= n_1 - 1$ και
t_2 η τιμή του t για ΒΕ $= n_2 - 1$

Τα όρια εμπιστοσύνης δίνονται από τον τύπο:

$$\overline{X}_1 - \overline{X}_2 \pm t'_{.05}\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

και η ΕΣΔ από τον τύπο

$$ΕΣΔ = t'_{.05}\sqrt{\frac{{s_1}^2}{n_1}+\frac{{s_2}^2}{n_2}}$$

3.3.2.4. <u>Μέγεθος δείγματος μικρό (n < 30) – οι τιμές είναι κατά ζεύγη</u>.

<u>Παράδειγμα</u>.

Για να μελετηθεί η ικανότητα αδελφώματος δυο ποικιλιών σιταριού χρησιμοποιήθηκαν 18 σπόροι από την πρώτη ποικιλία και άλλοι τόσοι από τη δεύτερη. Το κάθε ένα από τα ζευγάρια που σχηματίσθηκαν σπάρθηκε σε διαφορετικές εδαφοκλιματικές συνθήκες και καταγράφηκε ο αριθμός αδελφιών του κάθε φυτού ως εξής:

1η ποικιλία																	
31	40	28	21	26	38	48	60	46	35	45	14	22	40	39	35	20	26
2η ποικιλία																	
17	52	16	36	26	15	48	34	22	30	12	30	10	38	47	28	14	12
Διαφορά																	
14	-12	12	-15	0	23	0	26	24	5	33	-16	12	2	-8	7	6	14

<u>Λύση</u>.

Στην περίπτωση αυτή μεταχειριζόμαστε τις διαφορές ως απλές τιμές και το t υπολογίζεται από τον τύπο

$$t = \frac{\overline{D}}{\sqrt{\frac{\sum D^2 - (\sum D)^2/n}{n(n-1)}}}$$

ΣD = 14 + (-12) + 12 + (-15) + 0 + ………………………… + 2 + (-8) + 7 + 6 + 14 = 127

$\overline{D}$ = Σ D/ n = 127 / 18 = 7,06

$ΣD^2 = 14^2 + (-12)^2 + 12^2 +$ ……………………………………………… $+ 7^2 + 6^2 + 14^2 = 4353$

$(ΣD)^2/n = (127)^2 / 18 = 896{,}06$

και t = 1,8907

Η τιμή αυτή συγκρίνεται με τη θεωρητική για ΒΕ = n – 1 = 17 και πιθανότητα σφάλματος 5%, που είναι $t_{.05,17}$ = 2,11. Επειδή η θεωρητική τιμή είναι μεγαλύτερη από αυτή που υπολογίσθηκε, οι διαφορές δεν είναι σημαντικές και επομένως οι δυο ποικιλίες σιταριού δε διαφέρουν ως προς το αδέλφωμα.

Τα όρια εμπιστοσύνης στην περίπτωση αυτή είναι

$$\overline{D} \pm t_{.05}\sqrt{\frac{\sum D^2 - (\sum D)^2/n}{n(n-1)}}$$

και η ελάχιστη σημαντική διαφορά θα είναι

$$ΕΣΔ = t_{.05}\sqrt{\frac{\sum D^2 - (\sum D)^2/n}{n(n-1)}}$$

3.3.3. Σύγκριση ενός ποσοστού με ένα γνωστό αριθμό.

Παράδειγμα.

Ένας παραγωγός σιταριού παρέδωσε τους σπόρους του σε ένα κέντρο σποροπαραγωγής. Έγινε δοκιμή φυτρωτικής ικανότητας και από τους 90 σπόρους που μετρήθηκαν φύτρωσαν οι 50. Αν το όριο φυτρωτικής ικανότητας που χρησιμοποιεί το Κέντρο Σποροπαραγωγής είναι 65%, ζητείται να βρεθεί αν το ποσοστό φυτρώματος των σπόρων του παραγωγού πληροί το παραπάνω όριο με πιθανότητα σφάλματος 5%.

Λύση.

Τα δεδομένα στο παράδειγμα αυτό ακολουθούν τη διωνυμική κατανομή. Η τελευταία όμως μπορεί να προσεγγισθεί με αρκετή ακρίβεια με την κανονική κατανομή, όταν το μέγεθος του δείγματος είναι μεγάλο. Έτσι, για να λυθεί το παραπάνω πρόβλημα θα χρησιμοποιηθεί το z-κριτήριο. Η τιμή z δίνεται από τον τύπο

$$z = \frac{\frac{X}{n - p}}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\frac{50}{90 - 0{,}65}}{\sqrt{\frac{0{,}65(1-0{,}65)}{n}}} = \frac{-0{,}094}{0{,}050} = 1{,}88$$

Επειδή η τιμή αυτή είναι μεγαλύτερη από τη θεωρητική $z_{.05} = 1{,}645$ (Πίνακας ΙΙ στο Παράρτημα) βγαίνει το συμπέρασμα ότι το δείγμα του παραγωγού δεν εκπληρώνει τους όρους του Κέντρου Σποροπαραγωγής.

Τα όρια εμπιστοσύνης του p θα είναι

$$p \pm z_{.05} \sqrt{\frac{p(1-p)}{n}}$$

3.3.4. Σύγκριση δύο ποσοστών.

Παράδειγμα.

Ένας ερευνητής θέλησε να δοκιμάσει την αποτελεσματικότητα δύο υποστρωμάτων που χρησιμοποιούνται στην καλλιέργεια ανθήρων. Για το σκοπό αυτό καλλιέργησε 85 ανθήρες στο 1ο υπόστρωμα, από τους οποίους αντέδρασαν οι 57 και 70 ανθήρες στο 2ο υπόστρωμα, από τους οποίους αντέδρασαν οι 46. Ζητείται να βρεθεί, για πιθανότητα σφάλματος 5%, αν τα δυο αυτά υποστρώματα είναι εξίσου αποτελεσματικά.

Λύση.

Στην περίπτωση αυτή χρησιμοποιείται το z-κριτήριο και προς τις δυο πλευρές. Ο τύπος που δίνει την τιμή του z- κριτηρίου είναι

$$z = \frac{\frac{X_1}{n_1} - \frac{X_2}{n_2}}{\sqrt{\frac{\frac{X_1}{n_1}(1 - \frac{X_1}{n_1})}{n_1}} + \sqrt{\frac{\frac{X_2}{n_2}(1 - \frac{X_2}{n_2})}{n_2}}} = \frac{0{,}67 - 0{,}66}{0{,}076} = 0{,}1316$$

Η τιμή λοιπόν το z που υπολογίσθηκε από τα δεδομένα είναι μικρότερη από την θεωρητική που είναι $z_{.05}$ = 1,96. Από τη σύγκριση αυτή βγαίνει το συμπέρασμα ότι τα δύο υποστρώματα δε διαφέρουν μεταξύ τους.

Τέλος, τα όρια εμπιστοσύνης είναι

$$\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) \pm z_{.05} \sqrt{\frac{\frac{X_1}{n_1}\left(1-\frac{X_1}{n_1}\right)}{n_1}} + \sqrt{\frac{\frac{X_2}{n_2}\left(1-\frac{X_2}{n_2}\right)}{n_2}}$$

3.3.5. <u>Σύγκριση μιας διακύμανσης με ένα γνωστό αριθμό</u>.

Στην περίπτωση αυτή όταν το μέγεθος του δείγματος είναι μικρό χρησιμοποιείται το X^2-κριτήριο, ενώ όταν τα δείγματα είναι μεγάλα χρησιμοποιείται το z-κριτήριο.

3.3.5.1. <u>Σύγκριση μιας διακύμανσης με ένα γνωστό αριθμό: δείγματα μικρά</u>.

<u>Παράδειγμα</u>.

Ένας παραγωγός κριθαριού θέλησε να παραδώσει σε μια ζυθοποιεία μια ποσότητα κριθαριού. Επειδή οι Βιομηχανίες ζυθοποιείας ενδιαφέρονται για μεγάλους και ομοιόμορφους σπόρους, έχουν βάλει ως όριο η διακύμανση του μεγέθους των σπόρων να μη υπερβαίνει την τιμή σ_0^2 = 23,69. Αν η διακύμανση του μεγέθους ενός δείγματος 29 σπόρων από την ποικιλία του παραγωγού ήταν s^2 = 38,50, να βρεθεί με πιθανότητα σφάλματος 5%, αν η συνολική ποσότητα κριθαριού του παραγωγού καλύπτει το κριτήριο της Βιομηχανίας ζυθοποιείας.

<u>Λύση</u>.

Για να απαντηθεί το ερώτημα αυτό θα χρησιμοποιηθεί το X^2-κριτήριο και θα υπολογισθεί μια τιμή που θα συγκριθεί με τη θεωρητική τιμή X^2, για πιθανότητα σφάλματος 5% και Β. Ε. = n – 1 = 28.

$$X^2 = \frac{\sum(X-\overline{X})^2}{\sigma_0^2} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{28 \text{ x } 38{,}50}{23{,}69} = 45{,}50$$

Επειδή η τιμή του X^2 που υπολογίσθηκε από τα δεδομένα είναι μεγαλύτερη από την θεωρητική για πιθανότητα σφάλματος 5% και Β. Ε. 28 που είναι 41,34 (Πίνακας VI στο Παράρτημα), βγαίνει το συμπέρασμα ότι η διακύμανση του δείγματος είναι μεγαλύτερη από την επιτρεπτή και η ποικιλία αυτή είναι ακατάλληλη για ζυθοποίηση.

3.3.5.2. Σύγκριση μιας διακύμανσης με ένα γνωστό αριθμό: δείγματα μεγάλα.

Στην περίπτωση αυτή χρησιμοποιείται το z-κριτήριο, που υπολογίζεται από τον τύπο

$$z = \frac{\sigma - \sigma_0}{\frac{\sigma}{\sqrt{2n}}}$$

όπου σ η τυπική απόκλιση του δείγματος και σ_0 η τυπική απόκλιση που χρησιμοποιείται ως βάση σύγκρισης.

3.3.6. Σύγκριση δύο διακυμάνσεων.

Παράδειγμα.

Να βρεθεί αν οι διακυμάνσεις s_1^2 = 2,85, με Β. Ε. = 20 και s_2^2 = 1,98, με Β. Ε. = 27 διαφέρουν για πιθανότητα σφάλματος 5%.

Λύση.

Στην περίπτωση αυτή χρησιμοποιείται το F- κριτήριο, που υπολογίζεται από τον τύπο F= s_1^2/ s_2^2. Η τιμή αυτή συγκρίνεται με τη θεωρητική για πιθανότητα σφάλματος 5% και Β. Ε. 20 και 27. Η τιμή αυτή είναι 1,98 (Πίνακας IV, στο Παράρτημα). Η τιμή του F που είναι 2,85/1,98 = 1,43, είναι μικρότερη από τη θεωρητική. Αν τώρα θέλουμε να δούμε αν οι δυο διακυμάνσεις είναι διαφορετικές τότε θα πρέπει να έχουμε υπόψη ότι οι πίνακες του F έχουν υπολογισθεί με βάση τη μια πλευρά της καμπύλης. Έτσι, αν χρησιμοποιηθεί η πιθανότητα σφάλματος 5% τότε στην ουσία το ερώτημα απαντάται για πιθανότητα 10%. Αν χρησιμοποιηθεί πιθανότητα 1% τότε στην πραγματικότητα το ερώτημα απαντάται με πιθανότητα 2%. Για το λόγο αυτό θα χρησιμοποιηθεί η πιθανότητα σφάλματος 1%. Η θεωρητική τιμή του F για πιθανότητα 1% και Β. Ε. 20 και 27 είναι 2,61 (Πίνακας V, στο Παράρτημα). Επειδή η τιμή του F, που υπολογίσθηκε από τα δεδομένα είναι μικρότερη από τη θεωρητική για πιθανότητα σφάλματος 1%, οι δυο διακυμάνσεις μπορούν να θεωρηθούν ίσες και το σφάλμα θα είναι το πολύ 2%.

3.3.7. Σύγκριση συχνοτήτων.

Τα άτομα των διαφόρων πληθυσμών είναι δυνατό να καταταγούν σε διάφορες κατηγορίες, ανάλογα με το κριτήριο που θα χρησιμοποιηθεί. Για παράδειγμα οι γενότυποι που υπάρχουν σε ένα πληθυσμό με βάση τα γνωρίσματά τους μπορούν να καταταγούν σε νάνους, σε μετρίου ύψους και σε υψηλούς. Κάθε μια από τις κατηγορίες αυτές περιλαμβάνει ένα αριθμό φυτών. Το ίδιο μπορεί να γίνει και για ένα δεύτερο γνώρισμα, όπως για παράδειγμα το σχήμα των φύλλων. Τα δεδομένα αυτά μπορούν να καταγραφούν σε ένα πίνακα διπλής κατάταξης, όπου οι οριζόντιες γραμμές να αντιπροσωπεύουν τους διαφορετικούς γενοτύπους και οι στήλες τις συχνότητες των διαφόρων γνωρισμάτων. Η ανάλυση αυτών των πινάκων που είναι διπλής κατάταξης συνίσταται στην εκτίμηση αν οι συχνότητες των σειρών παρουσιάζουν μεταξύ τους ετερογένεια ή ομοιογένεια. Έτσι είναι αναγκαίο να ελεγχθεί η συμφωνία μεταξύ των συχνοτήτων, ή πρέπει να διευκρινισθεί αν οι συχνότητες συμφωνούν με μια γνωστή αναλογία. Για να γίνει πιο κατανοητές οι περιπτώσεις αυτές αυτή θα χρησιμοποιηθούν αντίστοιχα παραδείγματα.

3.3.7.1. Εκτίμηση της συμφωνίας μεταξύ συχνοτήτων.

Παράδειγμα.

Για να μελετηθεί ο τρόπος κληρονόμησης του ύψους και του σχήματος των φύλλων χρησιμοποιήθηκαν τρία διαφορετικά γενετικά υλικά. Τα φυτά που διέφεραν στα παραπάνω γνωρίσματα απομονώθηκαν και διασταυρώθηκαν μεταξύ τους. Τα φυτά της F_2 γενεάς, με τις τιμές στα αντίστοιχα γνωρίσματα δίνονται στο παρακάτω πινακίδιο (οι τιμές που είναι έξω από την παρένθεση).

Γνώρισμα φυτών F_2 / Γενετικό υλικό	Φυτά ψηλά, φύλλα πλατιά	Φυτά ψηλά, φύλλα στενά	Φυτά κοντά, φύλλα πλατιά	Φυτά κοντά, φύλλα στενά	Σύνολο
Α	93 (90,86)	28 (26,93)	26 (31,80)	12 (9,41)	159
Β	107 (92,57)	15 (27,44)	30 (32,40)	10 (9,597)	162
Γ	80 (96,57)	40 (28,63)	42 (33,80)	7 (10,00)	169
Σύνολο κατηγοριών	280	83	98	29	490

Ζητείται να βρεθεί αν ο τρόπος διάσπασης είναι κοινός ή κάποιο υλικό παρουσιάζει διαφορετική διάσπαση.

Λύση.

Για να απαντηθεί το παραπάνω ερώτημα θα χρησιμοποιηθεί το X^2– κριτήριο με το οποίο θα υπολογισθεί μια τιμή που θα συγκριθεί με τη θεωρητική για πιθανότητα σφάλματος 5% και Β. Ε. = (Β. Ε. στηλών −1) (Β. Ε. γραμμών – 1) = (4-1) (3 – 1) = 6. Η τιμή του X^2 υπολογίζεται από τον τύπο

$$X^2 = \frac{(f_1 - \varphi_1)^2}{\varphi_1} + \frac{(f_1 - \varphi_2)^2}{\varphi_2} + \cdots \frac{(f_1 - \varphi_\kappa)^2}{\varphi_\kappa}$$

όπου f οι συχνότητες που παρατηρήθηκαν και φ οι αναμενόμενες, αν κάθε μια από τις τρεις γραμμές παρουσίαζε την ίδια αναλογία διάσπασης με τις συχνότητες του συνόλου των κατηγοριών (δηλαδή 280, 83, 98, 29). Το κάθε ένα φ ισούται με ένα κλάσμα με αριθμητή το γινόμενο των δυο ακραίων συνόλων και παρονομαστή το γενικό σύνολο. Έτσι, οι τιμές των φ που δίνονται στο πινακίδιο μέσα στις παρενθέσεις, υπολογίσθηκαν ως εξής:

φ_1= 280(159/490)= 90,86, φ_2 = 83(159/490)= 26,93,.....φ_{12} = 29(169/490) = 10,0

Στη συνέχεια υπολογίζεται η τιμή του X^2:
X^2 = (93 – 90,86)^2/90,86 + (28 – 26,93)^2/26,93+…………..+ (7 – 10,0)^2/10,0 = 20,20

Από τον Πίνακα VI, στο Παράρτημα προκύπτει η θεωρητική τιμή του $X^2_{.05,\ 6}$ = 12,592, που είναι μικρότερη από αυτήν που υπολογίσθηκε από τα δεδομένα. Από αυτό βγαίνει το συμπέρασμα ότι ο τρόπος διάσπασης δεν είναι κοινός.

3.3.7.2. Εκτίμηση συμφωνίας μεταξύ συχνοτήτων με μια γνωστή αναλογία.

Παράδειγμα.

Να βρεθεί αν τα δεδομένα που παρουσιάσθηκαν στο πινακίδιο του προηγούμενου παραδείγματος συμφωνούν με την αναλογία που προϋποθέτει η ανεξάρτητη κληρονόμηση και είναι 9:3:3:1.

Λύση.

Εδώ είναι δυνατό να συμβούν τα εξής
(α) οι συχνότητες των υλικών που χρησιμοποιήθηκαν να ακολουθούν την αναλογία 9:3:3:1,
(β) οι συχνότητες να ακολουθούν μια άλλη αναλογία και
(γ) να υπάρχουν περισσότερες από μια αναλογίες.

Για να απαντηθεί ποια από τις παραπάνω περιπτώσεις ισχύει ακολουθούνται τα παρακάτω βήματα. Πρώτα ελέγχεται αν οι συχνότητες του 1^{ου} γενετικού υλικού ακολουθούν την αναλογία 9:3:3:1. Για να γίνει αυτό πρέπει να υπολογισθούν οι αναμενόμενες συχνότητες φ, κατά τρόπο ανάλογο με αυτόν που χρησιμοποιήθηκε στο προηγούμενο παράδειγμα:

93 (89,44)	28 (29,81)	26 (29,81)	12 (9.94)	159
9	3	3	1	16

Οι αναμενόμενες συχνότητες θα είναι: φ_1 = 9(159/16) = 89,44, φ_2 = 3(159/16) = 29,81, φ_3 = 3(159/16) = 29,81και φ_4 = 1(159/16) = 9,94, οπότε
X_A^2= (93–89,44)^2/89,44+(28-29,81)^2/29,81+(26-29,81)^2/29,81+(12-,94)^2/9,84=
= 0,1417 + 0,1099 + 0,4870 + 0,4741 = 1,2127

Κατά το ίδιο τρόπο υπολογίζονται οι αναμενόμενες συχνότητες για το 2^o γενετικό υλικό:

107(91,25)	15(30,38)	30(30,38)	10(10,13)	162
9	3	3	1	16

Οι αναμενόμενες συχνότητες θα είναι: φ_5 = 9(162/16) = 91,25, φ_6 = 3(162/16) = 30,38, φ_7 = 3(162/16) = 30,38 και φ_8 = 1(162/16) = 10,13, οπότε
X_B^2=(107-91,25)^2/91,25+(15-30,38)^2/30,38+(30-30,38)^2/30,38+(10-10,13)^2/10,13=
2,72 + 7,79 + 0,005 +0,002 = 10,52

Για το 3ο γενετικό υλικό:

80(95,04)	40(31,68)	42(31,68)	7(10,56)	169
9	3	3	1	16

Οι αναμενόμενες συχνότητες θα είναι: $φ_9$ = 9(169/16)= 95,04, $φ_{10}$ = 3(169/16)= 31,68, $φ_{11}$ = 3(169/16)= 31,68 και $φ_{12}$ =1(169/16)=10,56, οπότε

$X_Γ^2$=(80-95,04)^2/95,04+(40-31,68)^2/31,68+(42-31,68)^2/31,68+(7-10,56)^2/10,56=

= 2,38 + 2,19 +3,36 + 1,20 = 9,13

ΚΕΦΑΛΑΙΟ 4^{ο}

ΕΚΤΙΜΗΣΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΟΥ ΠΛΗΘΥΣΜΟΥ

4.1. Εισαγωγή.

Ένας πληθυσμός αποτελείται από πολλά άτομα και αυτό κάνει δύσκολο να προσδιορισθούν οι στατιστικές εκείνες που είναι απαραίτητες για την περιγραφή του. Για το λόγο αυτό χρησιμοποιείται από τη Βιομετρία το δείγμα, κατά τέτοιο τρόπο ώστε οι μετρήσεις που γίνονται σε αυτό να έχουν εφαρμογή σε όλο την πληθυσμό. Ο αριθμός των τιμών από τις οποίες απαρτίζεται ένα δείγμα συμβολίζεται με το Λατινικό γράμμα n. Όταν το $n > 30$, τότε το δείγμα θεωρείται μεγάλο. Αντίθετα, όταν το $n < 30$, τότε το δείγμα θεωρείται μικρό. Όπως αναφέρθηκε στο 1^{ο} Κεφάλαιο, οι δυο στατιστικές που περιγράφουν ένα πληθυσμό είναι ο μέσος όρος και η τυπική απόκλιση. Οι δυο αυτές στατιστικές ονομάζονται παράμετροι. Ο μέσος όρος και η τυπική απόκλιση του πληθυσμού συμβολίζονται με το ελληνικό γράμματα μ και σ (ενώ του δείγματος με τα γράμματα $\overline{X}$ και S αντίστοιχα). Ενώ όμως η μέσος όρος του δείγματος μπορεί να θεωρηθεί ότι αποτελεί την καλύτερη δυνατή εκτίμηση του μέσου όρου του πληθυσμού, δεν ισχύει το ίδιο και για την τυπική απόκλιση του δείγματος, όταν το τελευταίο είναι μικρό. Στη συνέχεια, θα περιγραφούν οι τρόποι υπολογισμού των παραμέτρων του πληθυσμού και των ορίων εμπιστοσύνης του μέσου όρου του πληθυσμού.

4.2. Υπολογισμός του μέσου όρου και της διακύμανσης.

Για να γίνουν κατανοητοί οι υπολογισμοί αυτοί θα χρησιμοποιηθεί το παρακάτω παράδειγμα. Κάποιος θέλησε να μελετήσει τη γεννητικότητα των κατοίκων ενός χωριού. Για το σκοπό αυτό πήρε ένα δείγμα 15 οικογενειών, στις οποίες μέτρησε τον αριθμό των παιδιών / οικογένεια και είχε τα εξής αποτελέσματα: 2, 4, 1, 3, 2, 5, 2, 3, 2, 6, 3, 4, 2, 2, 3. Ζητείται να υπολογισθεί ο μέσος όρος του δείγματος.

Λύση.

Θα χρησιμοποιηθεί ο τύπος

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

$$= \frac{(2+4+1+3+2+5+2+3+2+6+3+4+2+2+3)}{15} = \frac{44}{15} = 2{,}9$$

Ο μέσος όρος λοιπόν του δείγματος είναι 2,9.

Η διασπορά των τιμών γύρω από το μέσο όρο εκφράζεται με διάφορους τρόπους. Ένας από αυτούς είναι το άθροισμα των τετραγώνων των αποκλίσεων των τιμών γύρω από το μέσο όρο, που ονομάζεται άθροισμα τετραγώνων (ΑΤ).Το ΑΤ υπολογίζεται από τον τύπο ΑΤ= ΣX^2-$(\Sigma X)^2/n$. Αν το άθροισμα των τετραγώνων διαιρεθεί δια του n στα μεγάλα δείγματα και δια του n –1 στα μικρά, προκύπτει ένα άλλο μέτρο εκτίμησης της διασποράς, που ονομάζεται διακύμανση ή μέσο τετράγωνο (ΜΤ). Διακρίνονται λοιπόν δυο περιπτώσεις

(α) μεγάλα δείγματα:

$$\sigma^2 = \frac{\sum X^2 - \frac{\left(\sum X^2\right)}{n}}{n}$$

(β) μικρά δείγματα:

$$S^2 = \frac{\sum X^2 - \frac{\left(\sum X^2\right)}{n}}{n-1}$$

Όταν τα δείγματα είναι μεγάλα η διακύμανση συμβολίζεται με το σ^2, όπως δηλαδή και η διακύμανση του πληθυσμού. Χρησιμοποιώντας τα δεδομένα του προηγούμενου παραδείγματος στον τύπο της διακύμανσης (για μικρά δείγματα) προκύπτει:

$$S^2 = \frac{\sum X^2 - \frac{\left(\sum X^2\right)}{n}}{n-1} = \frac{(22 + 42 + 12 + \cdots .. + 22 + 22 + 32) - \frac{(44^2)}{15}}{15-1}$$

$$= \frac{154 - \frac{1936}{15}}{14} = 1{,}78$$

Η τετραγωνική ρίζα της διακύμανσης ονομάζεται τυπική απόκλιση και αποτελεί τον τρίτο τρόπο έκφρασης του μεγέθους της διασποράς. Επίσης, εκτός από την τυπική απόκλιση των τιμών ενός δείγματος, ενδιαφέρον παρουσιάζει η τυπική απόκλιση των μέσων όρων, των διαφορών των μέσων όρων κ. ά., οπότε η τυπική απόκλιση ονομάζεται τυπικό σφάλμα. Αν η τυπική απόκλιση εκφραστεί επί τοις εκατό, προκύπτει ένας πολύ χρήσιμος συντελεστής, που ονομάζεται συντελεστής παραλλακτικότητας (CV). Ο συντελεστή αυτός ισούται με

$$CV = \frac{S}{\overline{X}}$$

Ο συντελεστής παραλλακτικότητας εκφράζεται πάντα σε εκατοστά και είναι ανεξάρτητος από τις χρησιμοποιούμενες μονάδες μέτρησης. Χρησιμοποιείται για να συγκριθεί το μέγεθος της παραλλακτικότητας που υπάρχει σε πληθυσμούς με διαφορετικούς μέσους όρους και όταν έχουν χρησιμοποιηθεί διαφορετικές μονάδες μέτρησης. Όσο πιο μικρή τιμή παίρνει ο συντελεστής αυτός τόσο πιο αξιόπιστο είναι το πείραμα. Τέλος, κατά τον Φασούλα, ο συντελεστής αυτός δείχνει το φορτίο των εκφυλιστικών γονιδίων που υπάρχουν σε ένα πληθυσμό.

4.3. Υπολογισμός των ορίων εμπιστοσύνης του μέσου όρου ενός πληθυσμού.

Όταν κάποιος είναι σε θέση να βεβαιώσει με μια συγκεκριμένη πιθανότητα ότι ο μέσος όρος ενός πληθυσμού περικλείεται από δυο τιμές, τότε αυτές καθορίζουν ένα διάστημα που δίνει τα όρια εμπιστοσύνης. Μέσα στα όρια αυτά βρίσκεται ο μέσος όρος του πληθυσμού και για τη συγκεκριμένη πιθανότητα που ορίσθηκε. Η πιθανότητα αυτή εκφράζεται είτε ως ποσοστό επιτυχίας (π. χ. 95%) είτε ως ποσοστό αποτυχίας (π. χ. 5%). Συνήθως χρησιμοποιείται το ποσοστό αποτυχίας. Διακρίνονται δυο περιπτώσεις, μια όταν τα δείγματα είναι μεγάλα και μια όταν τα δείγματα είναι μικρά.

4.3.1. Δείγματα μεγάλα.

Κάποιος κατάγραψε τα ύψη 48 φυτών σιταρόβριζας, που αποτελούν ένα τυχαίο δείγμα από τα φυτά μιας ποικιλίας. Τα ύψη αυτά ήταν τα ακόλουθα:

75, 90, 80, 88, 105, 88, 95, 100, 110, 97, 108, 93, 90, 92, 80, 85, 96, 110, 100, 100, 90, 102, 98, 90, 84, 85, 85, 100, 110, 90, 104, 90, 85, 100, 100, 100, 95, 95, 100, 80, 104, 100, 105, 98, 92, 90, 90, 98.

Ζητείται να βρεθούν τα όρια μέσα στα οποία βρίσκεται ο μέσος όρος της ποικιλίας αυτής για πιθανότητα 5%.

Λύση.

Επειδή το μέγεθος του δείγματος είναι μεγάλο τα όρια εμπιστοσύνης βασίζονται στην κανονική κατανομή και δίνονται από τον τύπο $\overline{X} \pm z_{.05}\,\sigma_{\overline{X}}$.

Ο μέσος όρος ισούται με

$$\overline{X} = \Sigma X / n =$$
$$= 4542 / 48 = 94{,}63$$

Το τυπικό σφάλμα των μέσων όρων είναι

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

$$\sigma = \sqrt{\frac{\sum X^2 - \dfrac{(\sum X^2)}{n}}{n}} = \sqrt{\frac{433208 - 429786{,}8}{48}} = 8{,}44$$

Οπότε $$\sigma_{\overline{X}} = \frac{8{,}44}{\sqrt{48}} = 1{,}22$$

Η τιμή του $z_{.05}$ δίνεται στον Πίνακα III στο Παράρτημα και για πιθανότητα 5% είναι 1,96. Στη συνέχεια συμπληρώνονται τα όρια εμπιστοσύνης που είναι

$\overline{X} \pm z_{.05}\,\sigma_{\overline{X}} = 94{,}63 \pm 1{,}96 \times 1{,}22 = 94{,}63 \pm 2{,}39$, Έτσι, για πιθανότητα 5% τα όρια εμπιστοσύνης του μέσου όρου του ύψους της ποικιλίας είναι από 92,24 έως 97,02 εκατοστά.

4.3.1. Δείγματα μικρά.

Κάποιος κατάγραψε τα ύψη 23 φυτών βρίζας, που αποτελούν ένα τυχαίο δείγμα από τα φυτά ενός πληθυσμού. Τα ύψη αυτά ήταν τα ακόλουθα:

120, 108, 128, 120, 108, 108, 122, 128, 124, 132, 112, 112, 96, 112, 108, 108, 104, 120, 116, 128, 132, 120, 128.

Ζητείται να βρεθούν τα όρια μέσα στα οποία βρίσκεται ο μέσος όρος του πληθυσμού αυτού για πιθανότητα 5%.

Λύση.

Επειδή το μέγεθος του δείγματος είναι μικρό (μικρότερο του 30) τα όρια εμπιστοσύνης βασίζονται στην t-κατανομή και δίνονται από τον τύπο

$$\overline{X} \pm t_{.05}\,\sigma_{\overline{X}}$$

Ο μέσος όρος είναι

$$\overline{X} = \Sigma X / n = 2694/23 = 117{,}13$$

$$S_{\overline{X}} = \frac{S}{\sqrt{n}}$$

$$S = \sqrt{\frac{\Sigma x^2 - \frac{(\Sigma x^2)}{n}}{n-1}} = \frac{317684 - 315549{,}39}{22} = 9{,}85$$

Οπότε

$$S_{\overline{X}} = \frac{S}{\sqrt{n}} = \frac{9{,}85}{\sqrt{23}} = 2{,}05$$

Από τον Πίνακα ΙΙΙ στο Παράρτημα βρίσκεται η τιμή του t για πιθανότητα 5% και Β. Ε. 23-1 =22 είναι $t_{.05}$ = 2,074, οπότε τα όρια είναι X ± $t_{.05}$ σ_X = 117,13 ± 2,074 x 2,05, δηλαδή 117,13 ± 4,25. Έτσι, για πιθανότητα 5% τα όρια εμπιστοσύνης του μέσου όρου του ύψους του πληθυσμού είναι από 112,88 και 121,38.

ΚΕΦΑΛΑΙΟ 5^{0}

ΕΛΕΓΧΟΣ ΣΤΑΤΙΣΤΙΚΩΝ ΥΠΟΘΕΣΕΩΝ

5.1. Έλεγχος στατιστικών υποθέσεων.

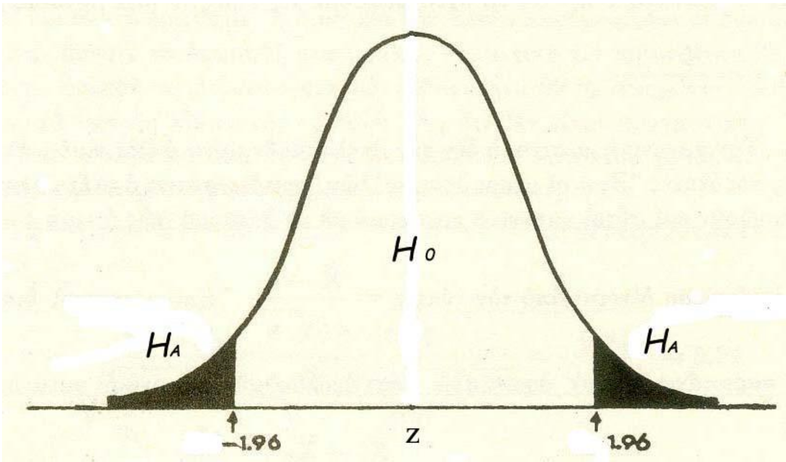
Ο μέσος όρος ενός πληθυσμού μπορεί να εκτιμηθεί υπολογίζοντας τα όρια εμπιστοσύνης. Αντ' αυτού όμως είναι δυνατό να εκτιμηθεί ελέγχοντας την υπόθεση ότι ο μέσος όρος έχει τιμή μηδέν. Η υπόθεση αυτή ονομάζεται μηδενική ή υπόθεση μηδέν και συμβολίζεται με H_0. Αν τώρα υπάρχουν πολλοί μέσοι όροι για να συγκριθούν, τότε η υπόθεση μηδέν είναι ότι αυτοί δεν διαφέρουν μεταξύ τους.

5.1.1. Έλεγχος H_0 $\mu=\mu_0$.

Η υπόθεση μηδέν εδώ είναι ότι ο μ έχει τη συγκεκριμένη τιμή μ_0. Έστω τώρα ένας πληθυσμός N, όπου το σ^2 είναι γνωστό (αν δεν είναι υπολογίζεται από το S^2 του δείγματος). Αφού ληφθεί από τον πληθυσμό ένα δείγμα n, υπολογίζεται το $\overline{X}$ και γίνεται η H_0: $\mu= \mu_0$. Εκτός όμως από την υπόθεση αυτή γίνεται και η εναλλακτική υπόθεση H_A ότι H_A: $\mu=\mu_A$. Αν ο μέσος όρος του δείγματος ($\overline{X}$) είναι κοντά στο μ_0 τότε γίνεται αποδεκτή η H_0. Αν όμως ο $\overline{X}$ είναι κοντά στο μ_A, τότε γίνεται αποδεκτή η εναλλακτική υπόθεση. Το ερώτημα που δημιουργείται είναι πόσο κοντά στο μ_0 πρέπει να είναι ο $\overline{X}$ για να μπορεί κάποιος να πει ότι ισχύει η H_0; Για να απαντηθεί το ερώτημα αυτό χρησιμοποιείται το z-κριτήριο

$$z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Αν η απόλυτη τιμή του z είναι μικρότερη του 1,96, δηλαδή αν $|z| < 1,96$, αν δηλαδή η τιμή του είναι αριστερά του 1,96 ή δεξιά του –1,96, τότε γίνεται αποδεκτή η H_0. Αν η απόλυτη τιμή του z είναι μεγαλύτερη από 1,96, αν δηλαδή $|z| > 1,96$, τότε ισχύει η εναλλακτική υπόθεση H_A (Σχήμα 13). Η σύγκριση έγινε με το 1,96, γιατί το πιο συνηθισμένο επίπεδο σημαντικότητας στο οποίο εργάζεται κάποιος είναι το 95%.



Σχήμα 13. Όταν η τιμή του z είναι μεταξύ -1,96 και +1,96 ισχύει η H_0, αλλιώς ισχύει η H_A.

Το σύνολο των τιμών που περιέχονται μεταξύ –1,96 και +1,96 καθορίζει μια περιοχή που ονομάζεται περιοχή αποδοχής. Αντίθετα, οι τιμές που βρίσκονται αριστερά του –1,96 και δεξιά του +1,96 καθορίζουν μια άλλη περιοχή, που ονομάζεται περιοχή απόρριψης ή κριτική περιοχή. Το κριτήριο μπορεί να είναι αμφίπλευρο, αν η κριτική περιοχή περιλαμβάνει τις επιφάνειες και στα δύο άκρα της καμπύλης. Ως εμβαδόν της κριτικής περιοχής επιλέγεται οποιοδήποτε, συνήθως όμως το 5% ή το 1%. Όταν το επίπεδο σημαντικότητας είναι 5% τότε z=1,96, ενώ όταν είναι 1% τότε z= 2,576 ≈ 2,58. Το εμβαδόν της κριτικής περιοχής είναι το επίπεδο σημαντικότητας. Όταν η απόλυτη τιμή του z είναι μεγαλύτερη του 1,96, δηλαδή όταν βρίσκεται στην κριτική περιοχή, οι μέσοι όροι οι οποίοι συγκρίνονται παρουσιάζουν μεταξύ τους διαφορά που είναι στατιστικώς σημαντική. Αυτό συμβολίζεται με (*). Όταν η απόλυτη τιμή του z είναι μεγαλύτερη του 2,576, τότε η διαφορά είναι στατιστικώς πολύ σημαντική και συμβολίζεται με (**). Το σύμβολο (***) σημαίνει ότι η σύγκριση έγινε στο επίπεδο 1‰.

Έστω τώρα μια αμφίπλευρη κριτική περιοχή. Το ερώτημα που δημιουργείται είναι γιατί να απορρίπτεται η H_0 αριστερά του –1,96 και δεξιά του +1,96; Ας υποτεθεί ότι ενώ κάποιος έχει δεχθεί ως σωστή την H_0 αυτή είναι λάθος. Αυτό σημαίνει ότι το μ δεν είναι μ_0 αλλά έτυχε με τη δοκιμή που χρησιμοποιήθηκε να βρεθεί μια τιμή του μέσου όρου κοντά στο μ_0, ενώ η H_0 είναι εσφαλμένη. Είναι όμως δυνατό να συμβεί και το αντίθετο: να απορριφθεί η H_0 ενώ είναι σωστή. Στις περιπτώσεις αυτές διακρίνουμε δυο ειδών σφάλματα

(α) σφάλμα τύπου Ι ή σφάλμα 1ου είδους: το σφάλμα αυτό διαπράττεται αν απορριφθεί η H_0 ενώ είναι ορθή.

(β) σφάλμα τύπου ΙΙ ή σφάλμα 2ου είδους: αυτό διαπράττεται όταν κάποιος δεχθεί την H_0 ενώ είναι λανθασμένη.

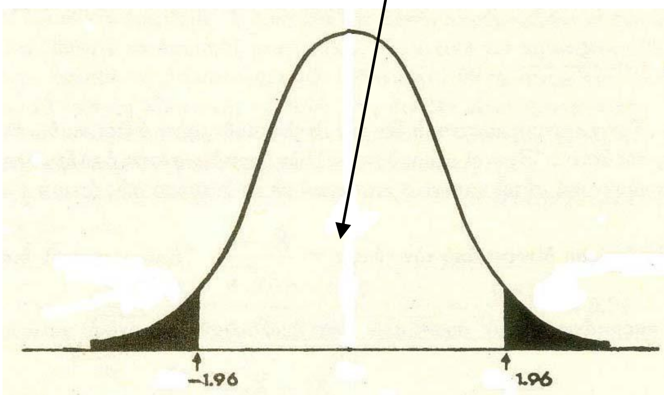
Παράδειγμα:

Κάποιος μελέτησε την αύξηση βάρους χοιριδίων, έχοντας τα εξής δεδομένα:σ=10, n=5110 και $\overline{X}$=29,823. Να ελεγχθεί η H_0 ότι μ=30 με εναλλακτική υπόθεση ότι H_A ότι μ ≠ 30.

Λύση.

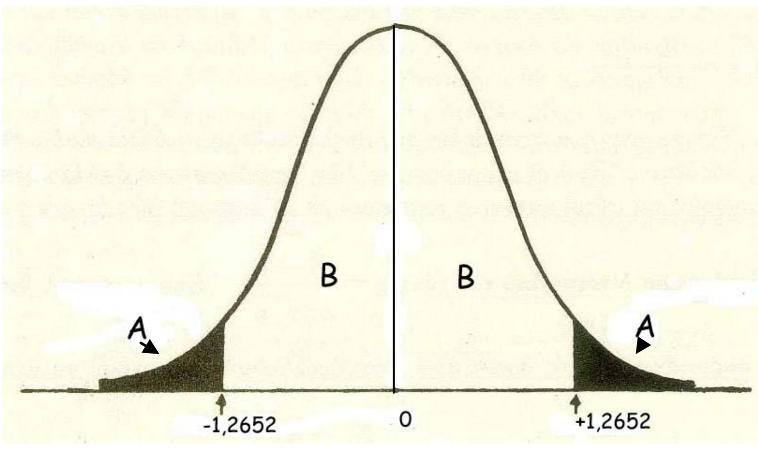
Χρησιμοποιείται το z-κριτήριο και υπολογίζεται μια τιμή

$$z = \frac{29{,}823 - 30}{\frac{10}{\sqrt{5110}}} = -1{,}2652 > -1{,}96$$



Η τιμή όμως αυτή βρίσκεται μέσα στην περιοχή αποδοχής. Αυτό σημαίνει ότι πρέπει να γίνει αποδεκτό ότι η H_0 είναι ορθή, ότι δηλαδή μ = 30. Μερικές φορές όμως αντί κάποιος να πει ότι η απόλυτη τιμή $|1{,}2652| < 1{,}96$ μπορεί να πει ότι P = 0,206. Αυτή είναι η πιθανότητα της μεγαλύτερης τιμής.

Υπολογισμός της πιθανότητας μεγαλύτερης τιμής (P).



Από τον Πίνακα Ι, στο Παράρτημα προκύπτει ότι στην τιμή z = -1,2652 αντιστοιχεί επιφάνεια 0,3971 (από 0 έως z). Επομένως η επιφάνεια A είναι

0,5000 – 0,3971 = 0,1029. Αν αθροιστούν οι δυο επιφάνειες Α προκύπτει: 0,1029 + 0,1029 = 0,206 > 0,05. Η τιμή Ρ = 20 σημαίνει ότι είμαστε μέσα στην κριτική περιοχή.

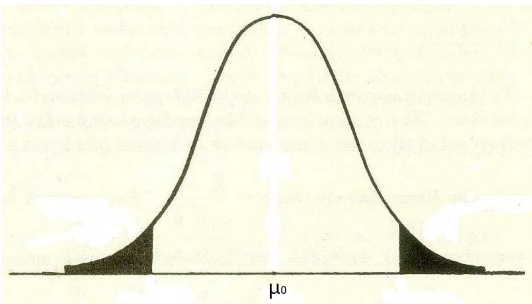
Όταν γίνεται αποδεκτή η H_0, αυτό σημαίνει ότι ισχύει το κριτήριο:

$$-1{,}96 \leq \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \leq 1{,}96$$

Η σχέση όμως αυτή χρησιμοποιείται για τον υπολογισμό των ορίων εμπιστοσύνης και μπορεί να γραφεί ως

$$\bar{X} - 1{,}96\ (\sigma/\sqrt{n}) \leq \mu \leq \bar{X} + 1{,}96\ (\sigma/\sqrt{n})$$

Αυτό σημαίνει ότι αν η H_0 είναι σωστή, τότε ο μέσος όρος είναι μέσα στα όρια εμπιστοσύνης.



Όταν απορρίπτεται η H_0: τότε είμαστε κατηγορηματικοί ότι $\mu \neq \mu_0$. Αν όμως γίνει αποδεκτή η H_0 αυτό δεν σημαίνει ότι έγινε αποδεκτό και ότι $\mu = \mu_0$.

Όταν το $\bar{X}$ είναι μέσα στα σκιασμένα τμήματα της καμπύλης τότε $\mu \neq \mu_0$.

Όταν η H_0 γίνεται αποδεκτή: αυτό δεν σημαίνει ότι το μ ισούται ακριβώς με το μ_0. Απλά, γίνεται δεκτό ότι το μ είναι κοντά στο μ_0. Το σωστό είναι να γραφεί ότι για την H_0 το $\mu \approx \mu_o$, αλλά επειδή απαιτείται μια συγκεκριμένη τιμή για να υπολογισθεί το z, για το λόγο αυτό γράφουμε ότι $\mu = \mu_0$.

Η πρακτική συνέπεια αποδοχής της H_0 εξαρτάται:

(α) από τις συνέπειες της αποδοχής της H_0 και

(β) από το μέγεθος του δείγματος: όσο μεγαλύτερο είναι το δείγμα, τόσο μεγαλύτερη είναι η ακρίβεια εκτίμησης του πληθυσμού. Αν το δείγμα είναι μικρό το κριτήριο δεν είναι αξιόπιστο και γίνεται αποδεκτό με επιφύλαξη ότι $\mu = \mu_0$. Στην περίπτωση αυτή βοηθούν τα όρια εμπιστοσύνης.

Στο προηγούμενο παράδειγμα

$$29{,}823 \pm 1{,}96\ (10/\sqrt{5110}) = 29{,}549 \text{ και } 30{,}097$$

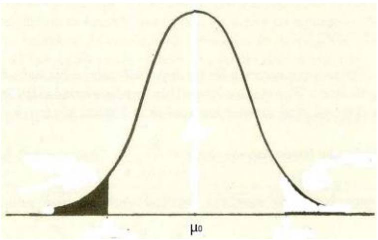
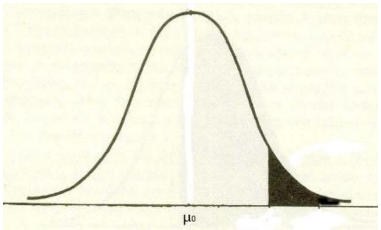
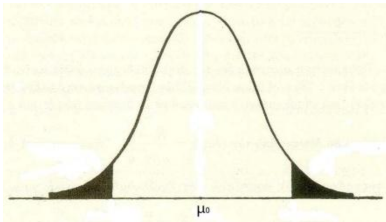
Αυτές είναι δυο τιμές που βρίσκονται κοντά η μια στην άλλη. Αν το μέγεθος του δείγματος είναι μικρό (n = 10) τότε

$$z = \frac{29{,}823-30}{\frac{10}{\sqrt{10}}} = -0{,}05$$

Επειδή και η τιμή αυτή είναι μέσα στην περιοχή αποδοχής η H_0 θεωρείται αποδεκτή. Αν όμως υπολογισθούν τα όρια εμπιστοσύνης 29,823 ± 1,96 $(10/\sqrt{10})$ = 23,62 και 36,02. Από το παραπάνω προκύπτει ότι ναι μεν η H_0 είναι αποδεκτή, τα όρια εμπιστοσύνης όμως απέχουν πολύ. Ένα μέγεθος δείγματος ≥ 30 θεωρείται ικανοποιητικό.

Όταν το σ είναι γνωστό, όπως στο προηγούμενο παράδειγμα, τότε χρησιμοποιείται το z-κριτήριο. Όταν όμως το σ δεν είναι γνωστό τότε χρησιμοποιείται το t-κριτήριο.

Προηγουμένως χρησιμοποιήθηκε η κριτική περιοχή και στα δυο άκρα της καμπύλης και το κριτήριο ήταν αμφίπλευρο. Το κριτήριο όμως μπορεί να είναι αμφίπλευρο ή <u>μονόπλευρο</u>.

<u>Πότε χρησιμοποιείται το μονόπλευρο κριτήριο</u>.

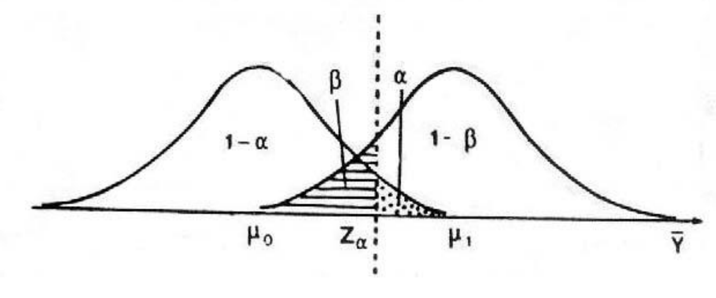
(α) <u>όταν είναι γνωστό από πείρα ότι $\mu \geq \mu_0$</u> (αλλά όχι $\mu \leq \mu_0$). Για παράδειγμα αν ψεκάσει κάποιος έναν αγρό καλαμποκιού με ένα εντομοκτόνο, επειδή είναι γνωστό ότι δεν βλάπτει τα φυτά, είναι δυνατό να συμβούν δύο πράγματα: (1) θα ωφελήσει και θα αυξηθεί η απόδοση και (2) δεν θα ωφελήσει και η απόδοση θα είναι ίση με το μέσο όρο χωρίς ψεκασμό.

(β) <u>όταν ενδιαφέρει αν $\mu > \mu_o$</u> (το $\mu < \mu_o$ δεν ενδιαφέρει). Για παράδειγμα ενδιαφέρει αν το νέο εντομοκτόνο είναι καλύτερο από τους μάρτυρες, ενώ δεν ενδιαφέρει αν είναι χειρότερο.

Στις περιπτώσεις αυτές κάνουμε την H_0: $\mu=\mu_0$ ή H_0: $\mu \leq \mu_0$ και H_A: $\mu > \mu_o$. Χρησιμοποιείται μονόπλευρο κριτήριο, προς το ένα ή το άλλο άκρο. Θα πρέπει να σημειωθεί ότι όταν κάποιος εργάζεται στο 5% στο αμφίπλευρο κριτήριο, τότε έχει 2,5% στο ένα άκρο και 2,5% στο άλλο. Στο μονόπλευρο όμως κριτήριο έχει 5% στο μοναδικό άκρο. Όταν το α = 0,05 το z = 1,645. Όταν το α= 0,01 το z = 2,326.

5.1.2.<u>Ισχύς του κριτηρίου</u>.

<u>Ισχύς του κριτηρίου</u> είναι η πιθανότητα απόρριψης της H_0 όταν αυτή είναι εσφαλμένη. Για παράδειγμα H_0: $\mu = \mu_0$ και H_A: $\mu = \mu_A$.



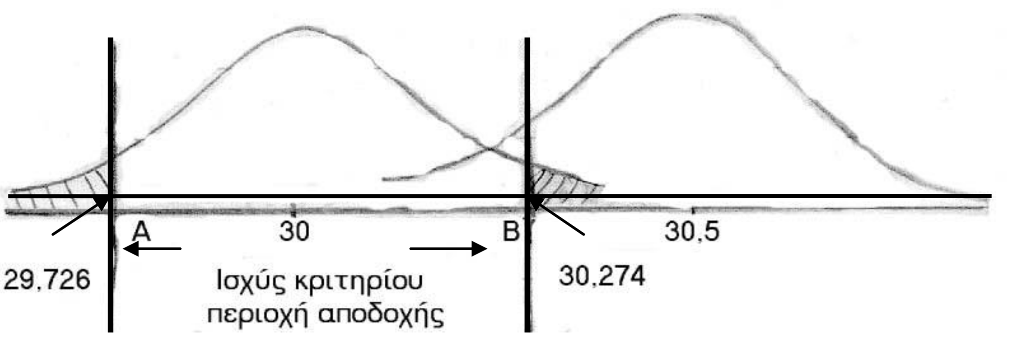
Αν η H_0 είναι ορθή ισχύει η αριστερή καμπύλη. Αν είναι η H_A ορθή, τότε ισχύει η δεξιά καμπύλη. Η διαχωριστική γραμμή καθορίζεται από το z (= 1,96).

Το <u>εμβαδόν α</u> είναι η κριτική περιοχή και είναι η πιθανότητα σφάλματος Ι είδους, ενώ το <u>εμβαδόν Β</u> είναι η πιθανότητα σφάλματος ΙΙ είδους. Το 1-α είναι ο <u>συντελεστής εμπιστοσύνης</u> και το 1-β η <u>ισχύς του κριτηρίου</u>.

Για να γίνουν όλα αυτά περισσότερο κατανοητά θα χρησιμοποιηθούν τα δεδομένα από το προηγούμενο παράδειγμα: n = 5110, σ = 10, μ_0 = 30. Έστω μ_A = 30,5 και ότι το μ_0 είναι το σωστό.

<u>Λύση</u>.

Το κριτήριο είναι αμφίπλευρο.



Οι δυο καμπύλες είναι ίδιες, εκτός από τη θέση τους. Έστω ότι α = 0,05. Ποια είναι τα σημεία Α και Β; Για να βρεθούν χρησιμοποιείται η σχέση:

$$30 \pm 1{,}96(10/\sqrt{5110}) =$$

$$= 30 \pm 0{,}274 \begin{cases} 29{,}726 \\ 30{,}274 \end{cases}$$

Η σκιασμένη περιοχή είναι η κριτική περιοχή. Θα γίνει αποδεκτό ότι μα αν πάρουμε τιμές κάτω από 29,726 και πάνω από 30,274.

Για να βρεθεί η ισχύς του κριτηρίου θα πρέπει να υπολογισθεί η επιφάνεια της καμπύλης.

(α) <u>αριστερά του 29,723</u>:

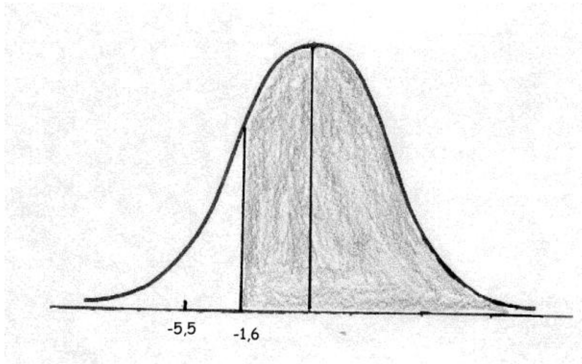
$$z = \frac{29{,}723 - 30{,}5}{\frac{10}{\sqrt{5110}}} = \text{-}5{,}529$$

Όπως προκύπτει από τον Πίνακα Ι στο Παράρτημα, οι τιμές σε αυτόν φθάνουν έως το 4 όπου αντιστοιχεί επιφάνεια 0,5000. Αριστερά λοιπόν του –5,529 η πιθανότητα είναι μηδέν. Αυτό μπορεί να γραφεί ως P (z < -5,529) ≈ 0.

(β) <u>δεξιά του 30,274</u>:

$$z = \frac{30{,}274 - 30{,}5}{1{,}40} = -1{,}614$$

και P (z > -1,614) = 0,4468 +0,5000 = 0,9468. Επομένως, η ισχύς του κριτηρίου είναι 0,9468 (0 + 0,947), δηλαδή η επιφάνεια της καμπύλης H_A είναι μέσα στην κριτική περιοχή.



και P (z > -1,614) = 0,4468 +0,5000 = 0,9468. Επομένως, η ισχύς του κριτηρίου είναι 0,9468 (0 + 0,947), δηλαδή η επιφάνεια της καμπύλης H_A είναι μέσα στην κριτική περιοχή.

5.1.2.<u>Από τι εξαρτάται η ισχύς του κριτηρίου</u>.

Η ισχύς του κριτηρίου εξαρτάται από:

(α) τη διαφορά $\mu_0 - \mu_A$,

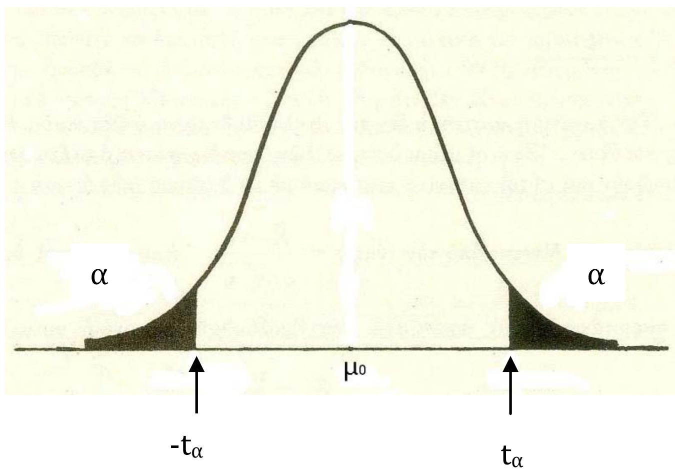
(β) το n και το σ, δηλαδή από το κλάσμα $\sigma/\sqrt{n}$. Όσο πιο μικρό είναι το σ και όσο πιο μεγάλο το n, τόσο πιο μικρό θα είναι το τυπικό σφάλμα και οι καμπύλες θα μαζεύονται γύρω από το μ και θα επικαλύπτονται ελάχιστα.

(γ) τον τύπο του κριτηρίου, αν είναι δηλαδή αμφίπλευρο ή μονόπλευρο.

Όταν το σ δεν είναι γνωστό και εκτιμάται από το δείγμα, όπως και στην περίπτωση των ορίων εμπιστοσύνης χρησιμοποιείται ο τύπος

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

για Β. Ε. = n – 1, όσοι είναι δηλαδή οι ΒΕ του σφάλματος. Το κριτικό βρίσκεται από τον Πίνακα ΙΙΙ στο Παράρτημα.



Όταν το t είναι μέσα στην περιοχή αποδοχής, το $\bar{X}$ είναι μέσα στα όρια $\mu_0 \pm t(S / \sqrt{n})$

<u>Παράδειγμα</u>: σε ένα πείραμα ψεκάσθηκαν 14 γραμμές φυτών καλαμποκιού, ενώ άλλες 14 δεν ψεκάσθηκαν. Στη συνέχεια πήραν τις διαφορές των αποδόσεων

-5,7, 3,7, 6,4, 1,5, 4,3, 4,8, 3,3, 3,6, 0,5, 5,0, 24,0, 8,8, 4,5, και 1,1

$$\overline{D} = 4{,}7,\ S_D = 6{,}48$$

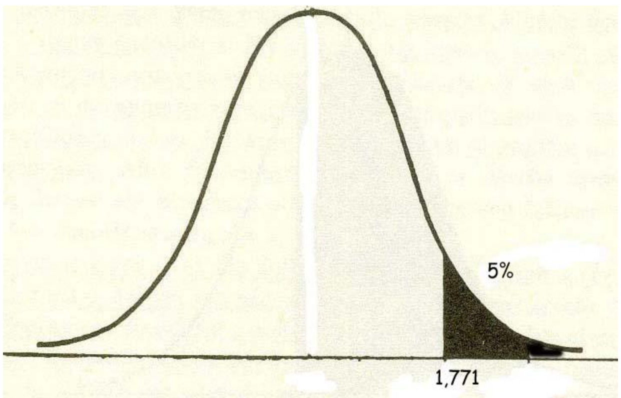
(από το δείγμα υπολογίζεται ο μέσος όρος, που ονομάζεται $\overline{D}$ γιατί αφορά διαφορές μέσων όρων). Αν το κόστος ψεκασμού είναι 3 € / acre και η τιμή του προϊόντος 1,5 € / bushel, ζητείται να βρεθεί αν μια διαφορά απόδοσης 2 bu / acre μπορεί να καλύψει το κόστος.

<u>Λύση</u>.

Το κριτήριο στο συγκεκριμένο πρόβλημα είναι μονόπλευρο γιατί ενδιαφέρει μόνο η περίπτωση που ο ψεκασμός αυξάνει την απόδοση. Για να απαντηθεί το πρόβλημα γίνεται η H_0: $\mu \leq 2$, με H_A: $\mu > 2$.

Το τυπικό σφάλμα υπολογίζεται σε

$$S_{\overline{D}} = \overline{D}/\sqrt{n} = \frac{6{,}48}{\sqrt{14}} = 1{,}73 \text{ bu/acre.}$$



Για ΒΕ = 13 από τον Πίνακα ΙΙΙ στο Παράρτημα, στήλη 0,10 αντιστοιχεί μια τιμή t= 1,771. Επειδή το κριτήριο είναι μονόπλευρο, ενδιαφέρει μόνο η σκιασμένη πλευρά.

Υπολογίζεται η τιμή του t = (4,7- 2) / 1,73 = 1,561. Η τιμή όμως αυτή είναι μικρότερη της θεωρητικής 1,771. Επομένως είμαστε στην περιοχή αποδοχής και η H_0 είναι δεκτή, ότι δηλαδή $\mu_0 \leq 2$. Αυτό σημαίνει ότι μια διαφορά της απόδοσης 2 bu/acre δεν καλύπτει το κόστος.

Το κατώτερο όριο εμπιστοσύνης είναι μονόπλευρο.

$\overline{X} - t\,S_{\overline{D}}$ = 4,70 –(1,771) (1,73) = 1,64 bu/acre

Επομένως η διαφορά μπορεί να κατεβεί έως 1,64 bu και κατά συνέπεια δεν συμφέρει ο ψεκασμός. Η ζημιά είναι 2 - 1,64 = 0,36 bu/acre. Θα πρέπει να τονισθεί εδώ ότι η τιμή 24 είναι ύποπτη, γιατί είναι πολύ μεγαλύτερη από τις υπόλοιπες.

Από το παραπάνω παράδειγμα γίνεται φανερό ότι εκείνο που έχει σημασία είναι το πώς καθορίζεται η H_0. Στο παράδειγμα έγινε με βάση το κόστος ψεκασμού ενός acre. Επίσης σπουδαίο ρόλο παίζει και το πώς καθορίσθηκε το κριτήριο, αν θα είναι αμφίπλευρο ή μονόπλευρο. Στο παράδειγμα, επειδή είναι επιθυμητό να γίνει γνωστό αν με τον ψεκασμό αυξάνει η απόδοση ή όχι, το κριτήριο ήταν μονόπλευρο. Το συμπέρασμα που βγαίνει από το παράδειγμα είναι ότι η αύξηση της απόδοσης δεν είναι αρκετά μεγάλη για να καλύψει το κόστος του ψεκασμού.

ΚΕΦΑΛΑΙΟ 6^{o}

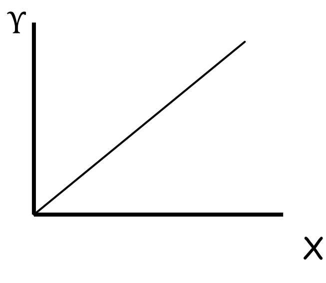
ΣΥΜΜΕΤΑΒΟΛΗ

6.1. Συμμεταβολή.

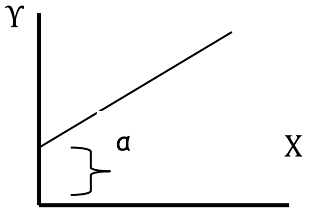
Έως τώρα ασχοληθήκαμε με μια μόνο μεταβλητή: πως διαφέρουν οι γενότυποι του μαλακού σιταριού ως προς την πρωιμότητα. Πολλές φορές όμως υπάρχουν δυο μεταβλητές (Χ και Υ) και το ζητούμενο είναι να διερευνηθεί αν υπάρχει κάποια σχέση μεταξύ των μεταβλητών αυτών και ποια είναι ακριβώς η σχέση αυτή. Για παράδειγμα υπάρχει σχέση μεταξύ της ποσότητας του λιπάσματος που χορηγείται και της απόδοσης των φυτών; Ή υπάρχει σχέση μεταξύ της ποσότητας ενός εντομοκτόνου και του αριθμού των εντόμων που θανατώνονται και ποια είναι αυτή. Στην περίπτωση αυτή υπάρχουν δυο γνωρίσματα που μεταβάλλονται ταυτόχρονα, γιατί έχουν κάποια σχέση. Η σχέση αυτή παρουσιάζεται υπό μορφή συνάρτησης Υ= f(x) και ονομάζεται συμμετεβολή ή παλινδρόμηση. Η μεταβλητή Χ ονομάζεται ανεξάρτητη μεταβλητή και η μεταβλητή Υ ονομάζεται εξαρτημένη, γιατί θεωρείται ότι το Χ είναι το αίτιο της τιμής Υ. Για παράδειγμα Χ μπορεί να είναι η ποσότητα του λιπάσματος και Υ η απόδοση των φυτών. Είναι όμως δυνατό να υπάρχει σχέση μεταξύ Χ και Υ, στην οποία όμως δεν μπορεί κάποιος να πει ότι η μια είναι το αίτιο και η άλλη το αποτέλεσμα. Παράδειγμα της σχέσης αυτής είναι το ύψος δυο αδελφιών, από τα οποία το ένα είναι αγόρι και το άλλο κορίτσι. Στην περίπτωση αυτή, η σχέση που υπάρχει μεταξύ των Χ και Υ ονομάζεται συσχέτιση (δες στο επόμενο Κεφάλαιο).

Το είδος της σχέσης που υπάρχει μεταξύ των δυο μεταβλητών μπορεί να παρασταθεί είτε με μια ευθεία γραμμή είτε με μια καμπύλη. Στην πρώτη περίπτωση η συμμεταβολή ονομάζεται γραμμική ενώ στη δεύτερη καμπυλόγραμμη. Όταν επιδιώκεται να εκτιμηθεί μια εξαρτημένη μεταβλητή από μια μόνο ανεξάρτητη, τότε η συμμεταβολή είναι απλή. Στην περίπτωση που οι ανεξάρτητες μεταβλητές είναι περισσότερες από μια, τότε η συμμεταβολή ονομάζεται πολλαπλή. Οι εξισώσεις που περιγράφουν τις πιο συνηθισμένες καμπύλες είναι δυο τύπων, οι πολυωνυμικές και οι εκθετικές (ή λογαριθμικές).

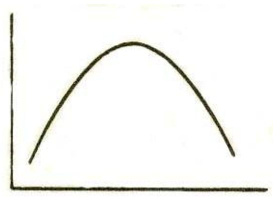
Η σχέση λοιπόν μεταξύ δυο μεταβλητών μπορεί να είναι της μορφής:

1. 

Υ= βΧ: στην περίπτωση αυτή η γραμμή συμμεταβολής είναι ευθεία που περνά από την αρχή των αξόνων. Για x=0 → Υ=0

2. 

Υ= α + βΧ: η γραμμή συμμεταβολής είναι ευθεία και για x= 0 ⟶ Υ=α

3. Υ 

Υ= α +βx + γx^2: η συμμεταβολή είναι καμπυλόγραμμη.

Στο Κεφάλαιο αυτό θα γίνει αναφορά στις δυο πρώτες περιπτώσεις.

6.2. <u>Προσδιορισμός ευθείας συμμεταβολής</u>.

Έστω μια σειρά δεδομένων Χ και Υ και επιζητείται να προσδιορισθεί η ευθεία που δίνει τη σχέση των δυο μεταβλητών.

<u>Παράδειγμα</u>.

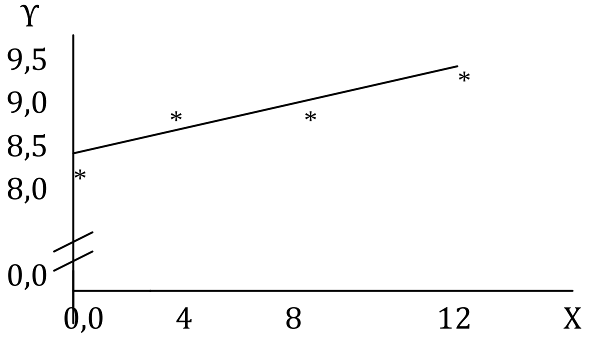
Σε ένα πείραμα θέλησαν να ερευνήσουν την επίδραση της ποσότητας του λιπάσματος στην απόδοση φυτών πατάτας. Χ είναι η ποσότητα του λιπάσματος και Υ η απόδοση των φυτών.

	Χ: ποσότητα λιπάσματος	Υ: απόδοση φυτών
	0	8,34
	4	8,89
	8	9,16
	12	9,50
Σύνολο	24	35,89

Το ερώτημα που τίθεται είναι αν υπάρχει σχέση μεταξύ των Χ και Υ και ποια είναι η μορφή της.

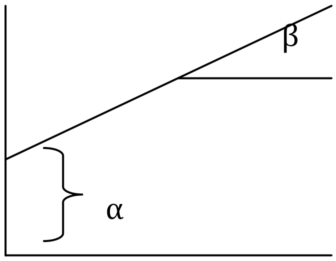
<u>Λύση</u>.

Το πρώτο πράγμα που πρέπει να γίνει για να απαντηθεί το ερώτημα είναι ένα σχεδιάγραμμα



Για x=0 ⟶ Υ=8,34

Στην περίπτωση αυτή η γραφική παράσταση λέγει ότι η σχέση των δυο μεταβλητών είναι ευθεία.



Όπως δε είναι γνωστό η γενική μορφή της εξίσωσης της ευθείας είναι Υ = α + βx, όπου α: η τεταγμένη από την αρχή και β: η κλίση της ευθείας.

Επειδή τα σημεία δεν πέφτουν στην ευθεία, λέμε ότι το πρότυπο της εξίσωσης της ευθείας είναι Υ = α +βx + ε, όπου ε: είναι το σφάλμα, που άλλοτε ανεβάζει και άλλοτε κατεβάζει τις τιμές από την ευθεία. Επειδή οι παράμετροι εκτιμώνται, η εξίσωση παίρνει τη μορφή Υ= a + bx + e, όπου a η εκτίμηση του α, b η εκτίμηση του β και e η εκτίμηση του ε.

Για να βρεθεί η εξίσωση του παραδείγματος θα πρέπει να υπολογιστούν οι συντελεστές a και b. Για το σκοπό αυτό χρησιμοποιούνται οι τύποι:

$$a = \overline{Y} - b\overline{X} \quad \text{και} \quad b = \frac{\Sigma XY - (\Sigma X \Sigma Y/n)}{\sum X^2 - \frac{(\Sigma X)^2}{n}}$$

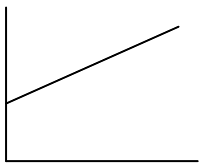
$$\text{Το } b = \frac{222{,}84 - (24)\,(35{,}89)\,/\,4}{224 - (24)2\,/4} = 0{,}09375$$

Οπότε το a = 8,97 – (0,09375) (6) = 8,4075.

Από τους υπολογισμούς αυτούς προκύπτει ότι η ευθεία της εξίσωσης κόβει τον άξονα των Υ στο 8,4075 και έχει κλίση 0,09375.

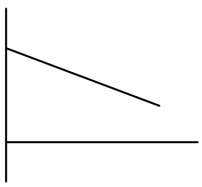
Ο συντελεστής b ονομάζεται <u>συντελεστής συμμεταβολής</u> και δίνει τη μεταβολή του Υ όταν το Χ μεταβάλλεται κατά μια μονάδα. Εδώ διακρίνονται οι εξής περιπτώσεις:

(α) b > 0



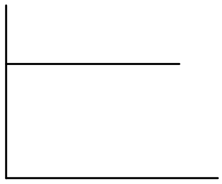
Όταν συμβαίνει αυτό, τότε αυξανομένου του Χ αυξάνεται και το Υ.

(β) b < 0



Εδώ αυξανομένου του Χ το Υ μειώνεται.

(γ) b = 0



Εδώ δεν υπάρχει σχέση μεταξύ Χ και Υ.

6.3. <u>Συμμεταβολή με δεδομένα παρατηρήσεων</u>.

Σε αυτά που αναφέρθηκαν έως τώρα το Χ ήταν δεδομένο: π. χ. μονάδες λιπάσματος 0, 4, 8, 12. Αν το πείραμα επαναλαμβάνονταν, το Χ θα ήταν πάλι το ίδιο, δηλαδή 0. 4. 8. 12. Πολλές φορές όμως χρησιμοποιείται η συμμεταβολή χωρίς να μπορεί να καθορισθεί το Χ. Για παράδειγμα αν Χ είναι ο συνολικός αριθμός μήλων και Υ ο αριθμός των μήλων που έχουν προσβληθεί από έντομα. Αν το πείραμα επαναληφθεί σε ένα άλλο δένδρο ή σε μια άλλη χρονική στιγμή, τότε το Χ θα αλλάξει. Για παράδειγμα, αν το 2002 μετρήθηκε ο συνολικός αριθμός μήλων σε 10 δένδρα και το 2003 καταμετρηθεί πάλι ο συνολικός αριθμός μήλων στα 10 αυτά δένδρα, τότε το Χ δεν θα είναι το ίδιο.

Στις περιπτώσεις λοιπόν που δεν είναι καθορισμένο το Χ είναι δυνατό να χρησιμοποιηθεί η συμεταβολή, αρκεί να ισχύουν οι παρακάτω τρεις προϋποθέσεις, προκειμένου να υπολογισθεί το σφάλμα. Οι προϋποθέσεις αυτές είναι:

(α) για κάθε τιμή Χ να αντιστοιχεί ένας πληθυσμός Υ,

(β) ο πληθυσμός των Υ να έχει μέσο όρο $\mu_{YX} = \alpha + \beta x$ και

(γ) όλοι οι πληθυσμοί των Υ που αντιστοιχούν σε κάθε τιμή Χ έχουν κοινή παραλλακτικότητα σ^2_{YX}.

Τέλος, θα πρέπει να γνωρίζει καλά το υπό μελέτη φαινόμενο αυτός που το εξετάζει.

6.4. <u>Πρόβλεψη του Υ από το Χ</u>.

Από τη εξίσωση της συμμεταβολής είναι δυνατό να υπολογισθούν ορισμένες τιμές του Υ από τις τιμές του Χ. Διακρίνονται δυο περιπτώσεις:

(α) είναι δυνατό να εκτιμηθεί μια ποσότητα $\mu_{YX} = \alpha + \beta x = \mu + \beta x$, δηλαδή ένα σημείο της ευθείας συμμεταβολής και

(β) είναι δυνατό να εκτιμηθεί ένα μεμονωμένο σημείο $\hat{Y}_X$, εκτός της ευθείας.

Και στις δυο αυτές περιπτώσεις η προβλεπόμενη τιμή, είτε είναι η μ_{YX} είτε είναι η $\hat{Y}_X$, είναι η ίδια, δηλαδή

$$(\mu_{YX} \text{ ή } \hat{Y}_X) = \overline{Y} + b\,(X - \overline{X}).$$

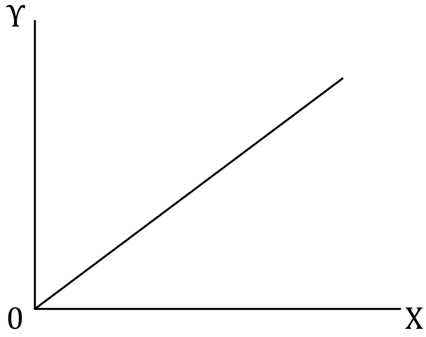
Το σφάλμα όμως διαφέρει σε κάθε περίπτωση. Έτσι είναι:

(α) για την πρώτη περίπτωση $S(\mu_{YX}) = S_{YX}\sqrt{\frac{1}{n} + \frac{x^2}{\sum x^2}}$ και τα όρια εμπιστοσύνης είναι $\mu_{YX} \pm t_{.05,BE}\, S\, \mu_{YX}$.

β) για τη δεύτερη περίπτωση $S(\hat{Y}_X) = S_{YX}\sqrt{1 + \frac{1}{n} + \frac{x^2}{\sum x^2}}$ και τα όρια εμπιστοσύνης είναι $\hat{Y}_X \pm t_{.05,BE}\, S(\hat{Y}_X)$.

και στις δυο περιπτώσεις $x = X - \overline{X}$.

6.5. <u>Ευθεία συμμεταβολής από την αρχή των αξόνων</u>.



Η ευθεία της συμμεταβολής σε ορισμένες περιπτώσεις περνά από την αρχή των αξόνων. Η εξίσωση εδώ είναι Υ = βΧ. Το <u>μαθηματικό πρότυπο</u> είναι Υ = βΧ + ε και η εκτίμηση θα είναι Υ = bX + e.

Ο συντελεστής συμμεταβολής υπολογίζεται από τον τύπο

$$b = \frac{\Sigma X\Upsilon}{\sum X^2}$$

6.6. <u>Πολλαπλή ευθύγραμμη συμμεταβολή</u>.

Πολλές φορές είναι δυνατό να υπάρχει ένας εξαρτημένος παράγοντας και δυο ανεξάρτητοι. Για παράδειγμα η μεταβολή του βάρους του σπόρου του σιταριού (εξαρτημένος) σε σχέση με το μήκος και το πλάτος (ανεξάρτητοι παράγοντες). Στην περίπτωση αυτή η συμμεταβολή ονομάζεται <u>πολλαπλή ευθύγραμμη συμμεταβολή</u>. Η εξίσωση της πολλαπλής ευθύγραμμης συμμεταβολής είναι η ακόλουθη

$$\Upsilon = \alpha + bX + cZ$$

Όπου:

$$b = \frac{[\Sigma Z^2 - (\Sigma Z)^2/n]\,[\Sigma X\Upsilon - \Sigma\Upsilon\Sigma X/n] - [\Sigma\Upsilon Z - \Sigma\Upsilon\Sigma Z/n]\,[\Sigma XZ - \Sigma X\Sigma Z/n]}{[\Sigma X^2 - (\Sigma X)^2/n]\,[\Sigma Z^2 - (\Sigma Z)^2/n] - [\Sigma XZ - \Sigma X\Sigma Z)/n]^2}$$

$$c = \frac{[\Sigma X^2 - (\Sigma X)^2/n]\,[\Sigma\Upsilon Z - \Sigma\Upsilon\Sigma Z/n] - [\Sigma X\Upsilon - \Sigma\Upsilon\Sigma X/n]\,[\Sigma XZ - \Sigma X\Sigma Z/n]}{[\Sigma X^2 - (\Sigma X)^2/n]\,[\Sigma Z^2 - (\Sigma Z)^2/n] - [\Sigma XZ - \Sigma X\Sigma Z)/n]^2}$$ και

$$a = \frac{\Sigma\Upsilon - b\Sigma X - c\Sigma Z}{N}$$

ΚΕΦΑΛΑΙΟ 7ο

ΣΥΣΧΕΤΙΣΗ

7.1. Συσχέτιση.

Το προηγούμενο Κεφάλαιο ασχολήθηκε με την περίπτωση εκείνη όπου υπάρχουν δυο μεταβλητές (Χ και Υ) και η μια από αυτές αποτελεί το αίτιο και η άλλη το αποτέλεσμα. Είναι όμως δυνατό να υπάρχει σχέση μεταξύ Χ και Υ, στην οποία όμως δεν μπορεί κάποιος να πει ότι η μια είναι το αίτιο και η άλλη το αποτέλεσμα. Παράδειγμα της σχέσης αυτής είναι το ύψος δυο αδελφιών, από τα οποία το ένα είναι αγόρι και το άλλο κορίτσι. Στην περίπτωση αυτή, η σχέση που υπάρχει μεταξύ των Χ και Υ ονομάζεται συσχέτιση. Επειδή οι μεταβλητές έχουν κάποια σχέση μεταξύ τους για το λόγο αυτό χρησιμοποιούνται τα σύμβολα X_1 και X_2. Η ποσότητα που μετρά τη σχέση μεταξύ των δυο μεταβλητών ονομάζεται συντελεστής συσχέτισης.

7.2. Ορισμός και ιδιότητες του συντελεστή συσχέτισης.

7.2.1. Ορισμός.

Ο συντελεστής συσχέτισης (r) ορίζεται

$$r = \frac{\sum x_1 \sum x_2}{\sqrt{\sum x_1^2 \sum x_2^2}} = \frac{\sum x_1 x_2 - \sum x_1 \sum x_2 / n}{\sqrt{\left[\sum x_1^2 - \frac{(\sum x_1)^2}{n}\right]\left[\sum x_2^2 - \frac{(\sum x_2)^2}{n}\right]}}$$

Όπου $x_1 = X_1 - \bar{X}_1$ και $x_2 = X_2 - \bar{X}_2$

Ο συντελεστής συσχέτισης είναι καθαρός αριθμός και παίρνει τιμές από –1 έως +1.

7.2.2. Ιδιότητες του συντελεστή συσχέτισης.

Ο συντελεστής συσχέτισης έχει τις παρακάτω ιδιότητες:

(α) είναι καθαρός αριθμός. Αν για παράδειγμα το X_1: cm και X_2: g, τότε ο συντελεστής συσχέτισης είναι

$$r = \frac{(\mathrm{cm})\,(\mathrm{g})}{\sqrt{(\mathrm{cm}^2)(\mathrm{g}^2)}} = \text{καθαρός αριθμός}$$

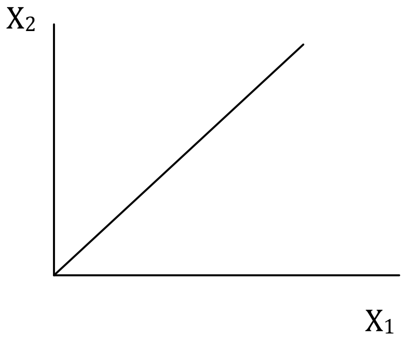
(β) ο συντελεστής συσχέτισης παίρνει τιμές από –1 έως +1.

Όταν $r > 0$: τα X_1 και X_2 μεταβάλλονται προς την ίδια κατεύθυνση.

<u>r < 0</u>: τα X_1 και X_2 μεταβάλλονται προς αντίθετες κατευθύνσεις: αύξηση του X_1 συνεπάγεται μείωση του X_2.

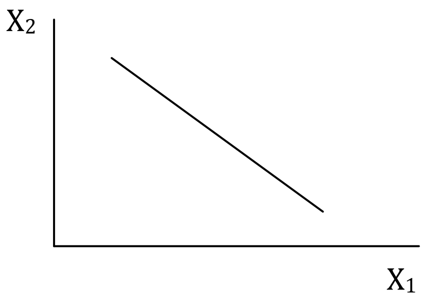
<u>r = 0</u>: οι δυο μεταβλητές είναι ανεξάρτητες.

<u>Διερεύνηση των περιπτώσεων r = 1 και r =-1</u>.



<u>Παράδειγμα</u>: έστω ότι $X_1 = KX_2$.. Αυτό σημαίνει ότι μεταξύ των δυο μεταβλητών υπάρχει η σχέση που φαίνεται στο διπλανό σχήμα. Επάνω στην ευθεία πέφτουν τα σημεία που συνδέουν το X_1 με το X_2. Αν στον τύπο του συντελεστή συσχέτισης αντικατασταθεί το X_1 με το KX_2 και γίνουν οι πράξεις τότε r = 1.

Από το παραπάνω συνεπάγεται ότι όταν $X_1 = KX_2$ τότε δεν υπάρχει καθόλου σφάλμα.



Αν τώρα $X_1 = -KX_2$ τότε r = -1. Και στην περίπτωση αυτή τα σημεία που συνδέουν το X_1 με το X_2 πέφτουν επάνω στην ευθεία, μόνο που αυτή έχει κλίση προς τα κάτω.

7.3. <u>Σχέση μεταξύ συμμεταβολής και συσχέτισης</u>.

Αν ο συντελεστής συσχέτισης υψωθεί στο τετράγωνο:

$$r^2 = \frac{(\sum xy)^2}{\sum x^2 \sum y^2} = \frac{\Sigma xy \quad \Sigma xy}{\sum x^2 \sum y^2} = \mathrm{byx\ bxy}$$

Δηλαδή το τετράγωνο του συντελεστή συσχέτισης ισούται με το γινόμενο του συντελεστή συμμεταβολής, όταν το X είναι η ανεξάρτητη μεταβλητή επί το συντελεστή συμμεταβολής, όταν το Y είναι η ανεξάρτητη μεταβλητή. Αν τώρα ο αριθμητής και ο παρονομαστής διαιρεθούν με το Σx^2, τότε

$$r^2 = \frac{(\sum xy)^2}{\sum x^2 \sum y^2} = \frac{(\sum xy)^2/\sum x^2}{\sum y^2} = \frac{\text{Αθροισμα τετραγώνων συμμεταβολής}}{\text{Συνολικό άθροισμα τετραγώνων}}$$

Το τετράγωνο λοιπόν του συντελεστή συσχέτισης εκφράζει το ποσοστό του συνολικού αθροίσματος τετραγώνων (ΣΑΤ) που οφείλεται στη συμμεταβολή και ονομάζεται <u>συντελεστής προσδιορισμού</u>. Το 1 – r^2 δίνει το ποσοστό του σφάλματος που οφείλεται στις αποκλίσεις από τη συμμεταβολή.

7.4. <u>Παραδείγματα</u>.

7.4.1. <u>Απλή ευθύγραμμη συσχέτιση</u>.

Από την ποικιλία μαλακού σιταριού Βεργίνα εκτιμήθηκαν τρία γνωρίσματα του σπόρου: το βάρος (Υ), το μήκος (Χ) και το πλάτος (Ζ).Δίνονται τα: ΣΧ = 684, ΣΥ = 430, ΣΖ = 314, $ΣΥ^2$ = 19034, $ΣΧ^2$ = 46900, $ΣΖ^2$ = 9898, ΣΧΥ = 29650, ΣΥΖ = 13623, ΣΧΖ = 21523 και n = 10. Ζητείται να βρεθούν οι συντελεστές συσχέτισης μεταξύ βάρους σπόρου και μήκους, βάρους σπόρου και πλάτους και μήκους και πλάτους.

<u>Λύση</u>.

Ο συντελεστής συσχέτισης μεταξύ βάρους και μήκους σπόρου είναι

$$r_{YX} = \frac{ΣΧΥ - ΣΧΣΥ/n}{\sqrt{[ΣΧ^2 - (ΣΧ)^2/n]\,[ΣΥ^2 - (ΣΥ)^2/n]}}$$

Από τον τύπο αυτό αν γίνουν οι πράξεις προκύπτει r_{YX} = 0,95*.

Καθ' όμοιο τρόπο ο συντελεστής συσχέτισης μεταξύ βάρους και πλάτους σπόρου είναι

$$r_{YZ} = \frac{ΣΥΖ - ΣΥΣΖ/n}{\sqrt{[ΣΥ^2 - (ΣΥ)^2/n]\,[ΣΖ^2 - (ΣΖ)^2/n}} = 0{,}84^*$$

Τέλος, ο συντελεστής συσχέτισης μεταξύ μήκους και πλάτους σπόρου είναι

$$r_{XZ} = \frac{ΣΧΖ - ΣΧΣΖ/n}{\sqrt{[ΣΧ^2 - (ΣΧ)^2/n]\,[ΣΖ^2 - (ΣΖ)^2/n}} = 0{,}68^*$$

Στη συνέχεια θα πρέπει να ελεγχθεί αν οι συντελεστές που υπολογίσθηκαν δείχνουν την ύπαρξη ή όχι συσχέτισης. Για να γίνει αυτό χρησιμοποιείται το t-κριτήριο με το οποίο συγκρίνεται στην ουσία αν ο συντελεστής συσχέτισης διαφέρει από το 0. Αν το t > 0, τότε ο συντελεστής συσχέτισης διαφέρει από το μηδέν και επομένως υπάρχει ευθύγραμμη συσχέτιση. Αν το t < 0, τότε ο συντελεστής συσχέτισης δεν διαφέρει από το μηδέν και επομένως δεν υπάρχει ευθύγραμμη συσχέτιση.

Το t υπολογίζεται από τον παρακάτω τύπο

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Άρα το $t = \frac{0{,}95\sqrt{10-2}}{\sqrt{1-0{,}95^2}}$ =8,61. Η τιμή αυτή συγκρίνεται με τη θεωρητική τιμή, που παίρνεται από τον Πίνακα ΙΙΙ στο Παράρτημα για ΒΕ = n – 2 και για όποια πιθανότητα είναι επιθυμητή (συνήθως 5%). Με τον τρόπο αυτό υπολογίζεται η τιμή $t_{.05,\,8}$ = 2,306. Επειδή το t = 8,61 που υπολογίσθηκε από τα δεδομένα είναι μεγαλύτερη από τη θεωρητική $t_{.05,\,8}$ = 2,306 συμπεραίνεται ότι ο συντελεστής συσχέτισης είναι στατιστικά σημαντικός και ότι υπάρχει ευθύγραμμη συσχέτιση μεταξύ βάρους και μήκους σπόρου και για να δηλωθεί αυτό σημειώνεται ένα αστεράκι δίπλα στην τιμή 0.95. Το ίδιο επαναλαμβάνεται και με τους υπόλοιπους συντελεστές συσχέτισης και η ύπαρξη αστεριού δίπλα σε καθένα από αυτούς δηλώνει την ύπαρξη ευθύγραμμης συσχέτισης και στις άλλες δυο περιπτώσεις.

Τέλος, είναι δυνατό να ελεγχθεί αν δυο συντελεστές συσχέτισης διαφέρουν μεταξύ τους στατιστικώς σημαντικά. Αυτό γίνεται με τη βοήθεια του z- κριτηρίου, που ισούται εδώ με $z = (Z_1 - Z_2)/\sigma_{z1\text{-}z2}$, όπου Z_1 και Z_2 οι τιμές των συντελεστών συσχέτισης r_1 και r_2 αφού προηγουμένως μετατραπούν σε τιμές Z με τη βοήθεια του Πίνακα VI στο Παράρτημα. Έτσι, οι τιμές $r_{YX} = 0{,}95$ και $r_{YZ} = 0{,}84$ αντιστοιχούν στις τιμές $Z_1 = 1{,}832$ και $Z_2 = 1{,}221$. Το $\sigma_{z1\text{-}z2}$ υπολογίζεται από τον τύπο

$$\sigma_{z1-z2} = \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}} = \sqrt{\frac{1}{7} + \frac{1}{7}} = 0{,}535$$

Συμπληρώνοντας τώρα στον παραπάνω τύπο του Z προκύπτει

$$z = (1{,}832 - 1{,}221)/0{,}535 = 1{,}14$$

Επειδή το $z = 1{,}14$ είναι μικρότερο από το $z_{.05} = 1{,}96$ συμπεραίνεται ότι οι δυο συντελεστές δεν διαφέρουν μεταξύ τους σημαντικά.

7.4.2. Πολλαπλή ευθύγραμμη συσχέτιση.

Είναι επίσης δυνατό να βρεθεί ο συντελεστής συσχέτισης μεταξύ βάρους σπόρου, και μήκους με πλάτος, Στην περίπτωση αυτή ο συντελεστής ονομάζεται συντελεστής πολλαπλής συσχέτισης. Για να γίνει αυτό κατανοητό θα χρησιμοποιηθούν τα δεδομένα του προηγούμενο παραδείγματος.

Λύση.

Ο συντελεστής πολλαπλής συσχέτισης υπολογίζεται από τον τύπο:

$$R_{y.xz} = \sqrt{\frac{r_{yx}^2 + r_{yz}^2 - 2r_{yx}r_{yz}r_{xz} +}{1 - r_{xz}^2}}$$

Αν τώρα αντικατασταθούν οι συντελεστές της τετραγωνικής ρίζας με τις τιμές που υπολογίσθηκαν προηγούμενα προκύπτει $R_{y.xz} = 0{,}99$.

Στο προηγούμενο παράδειγμα βρέθηκε συντελεστής συσχέτισης μεταξύ βάρους και μήκους σπόρου $r_{yx} = 0{,}95$, ενώ ο συντελεστής μεταξύ βάρους και μήκους με πλάτος βρέθηκε $R_{y.xz} = 0{,}99$. Οι αντίστοιχοι συντελεστές προσδιορισμού είναι $r_{yx}^2 = 0{,}90$ και $R_{y.xz}^2 = 0{,}98$. Τα αποτελέσματα αυτά φανερώνουν ότι η εξίσωση πολλαπλής ευθύγραμμης συσχέτισης δίνει καλύτερη προσέγγιση της εκτίμησης του βάρους του σπόρου σε σχέση με το μήκος και πλάτος.

ΚΕΦΑΛΑΙΟ 8^{0}

ΑΡΧΕΣ ΓΕΩΡΓΙΚΟΥ ΠΕΙΡΑΜΑΤΙΣΜΟΥ-ΜΕΘΟΔΟΙ ΕΥΡΕΣΗΣ ΔΙΑΦΟΡΩΝ ΜΕΤΑΞΥ ΤΩΝ ΕΞΕΤΑΖΟΜΕΝΩΝ ΓΕΝΟΤΥΠΩΝ

8. 1. Αρχές του Γεωργικού Πειραματισμού.

Όπως είναι λογικό, όταν κάποιος μελετά την συμπεριφορά ορισμένων γενοτύπων, θα πρέπει να έχει εξασφαλίσει όλες τις προϋποθέσεις ώστε τα συμπεράσματά που θα βγάλει να είναι αξιόπιστα και αδιαμφισβήτητα. Για να γίνει αυτό χρησιμοποιούνται ορισμένοι κανόνες του Γεωργικού Πειραματισμού που εξασφαλίζουν ακριβώς αυτή την αξιοπιστία. Οι κανόνες αυτοί είναι οι ακόλουθοι.

8.1.1. Το τυχαίο του δείγματος

Όταν κάποιος πειραματίζεται στον αγρό, συνήθως αντιμετωπίζει το εξής πρόβλημα. Θα χρησιμοποιήσει όλα τα φυτά που έχει στον πειραματικό του και πόσο εφικτό είναι αυτό; Η απάντηση στο ερώτημα αυτό είναι απλή και κατηγορηματική: όχι, δεν μπορεί να χρησιμοποιήσει όλα τα φυτά καθενός από τους γενότυπους που μελετά. Αυτό είναι ανέφικτο γιατί είναι αδύνατο να χειριστεί τα χιλιάδες φυτά του κάθε εξεταζόμενου γενότυπου. Το πρόβλημα γίνεται ακόμη πιο δύσκολο, γιατί δεν έχει να εξετάσει ένα, αλλά πολλούς γενότυπους. Το ερώτημα που δημιουργείται στη συνέχεια είναι τι θα πρέπει να κάνει για να μπορέσει να μελετήσει τα φυτά και να βγάλει σωστά συμπεράσματα; Εδώ ακριβώς έρχεται ο γεωργικός πειραματισμός, που προτείνει αντί του συνόλου των ατόμων ενός πληθυσμού να χρησιμοποιηθεί ένα πολύ μικρό μέρος του, που ονομάζεται δείγμα. Αυτόματα όμως δημιουργείται η εξής απορία: πόσο σωστό είναι αυτό και πόσο αξιόπιστα θα ήταν τα συμπεράσματα που στηρίζονται σε ένα μικρό μέρος του πληθυσμού και αφορούν όλο τον πληθυσμό. Η λύση που προτείνει ο Γεωργικός Πειραματισμός για να αντιμετωπισθεί το πρόβλημα αυτό είναι να χρησιμοποιηθεί τυχαίο δείγμα του πληθυσμού. Αν συμβεί αυτό, το δείγμα θεωρείται αντιπροσωπευτικό όλου του πληθυσμού. Από τη στιγμή που ένα δείγμα είναι τυχαίο, κάθε μέλος του πληθυσμού των γενοτύπων που μελετώνται, θα έχει την ίδια πιθανότητα να συμπεριληφθεί σε αυτό.

8.1.2. Χρησιμοποίηση επαναλήψεων

Ο ρόλος τους είναι αποφασιστικός γιατί εξασφαλίζουν την αξιοπιστία του πειράματος, δίνοντας ίσες ευκαιρίες σε κάθε γενότυπο να αποδώσει το μέγιστο των δυνατοτήτων του. Ένας γενότυπος αν δοκιμάζονταν σε ένα μόνο τμήμα ενός αγρού, θα μπορούσε να αποδώσει ικανοποιητικά ή όχι. Το ερώτημα που τίθεται είναι κατά πόσο η συμπεριφορά του αυτή ανταποκρίνεται στο δυναμικό του. Με άλλα λόγια ένας γενότυπος που θα αποδώσει ικανοποιητικά είναι πραγματικά αποδοτικός ή έτυχε να βρεθεί σε κάποιο γόνιμο σημείο του αγρού. Από την άλλη πλευρά ένας άλλος γενότυπος που δεν θα αποδώσει ικανοποιητικά είναι πραγματικά χειρότερος ή είχε την ατυχία να βρεθεί σε κάποιο άγονο σημείο του αγρού. Χωρίζοντας λοιπόν τον αγρό σε μικρότερα τεμάχια-μικροπεριβάλλοντα, που ονομάζονται επαναλήψεις, και συμπεριλαμβάνοντας τους ίδιους γενοτύπους σε κάθε επανάληψη, αντιμετωπίζεται το πρόβλημα της ετερογένειας του εδάφους. Αυτό γίνεται γιατί ενώ ένας γενότυπος είναι δυνατό σε μια επανάληψη να βρεθεί σε ευνοϊκό έδαφος, αυτό δεν είναι δυνατό να συμβεί και στις υπόλοιπες. Όσο αυξάνεται ο αριθμός των επαναλήψεων, το πειραματικό σφάλμα ελαττώνεται, αυξάνεται όμως αρκετά το κόστος του πειράματος. Γι’ αυτό αναζητήθηκε να βρεθεί ο ελάχιστος αριθμός επαναλήψεων που να εξασφαλίζει αξιοπιστία. Επίσης και με δεδομένο ότι οι γενότυποι μελετώνται τόσο διατοπικά (γιατί μας ενδιαφέρει να γνωρίζουμε τη συμπεριφορά τους σε διάφορα περιβάλλοντα) όσο και διαχρονικά (γιατί μας ενδιαφέρει να γνωρίζουμε τη συμπεριφορά των γενοτύπων στις διαφορετικές χρονιές), αναζητήθηκαν και οι αντίστοιχοι αριθμοί των απαραίτητων τοποθεσιών και ετών. Έτσι, βρέθηκε ότι για να εξασφαλισθεί η αξιοπιστία του πειραματισμού οι ελάχιστοι αριθμοί που απαιτούνται είναι τρεις επαναλήψεις, τρεις τοποθεσίες και τρεις συνεχόμενες χρονιές. Επίσης, θα πρέπει να σημειωθεί ότι όπως έδειξαν οι Χριστίδης και Fischer, για να εξασφαλισθεί μεγαλύτερη αξιοπιστία του πειραματισμού θα πρέπει το σχήμα των επαναλήψεων να είναι τετράγωνο και των πειραματικών τεμαχίων (πειραματικό τεμάχιο είναι ο ιδιαίτερος χώρος που καταλαμβάνει ένας γενότυπος σε κάθε επανάληψη) να είναι παραλληλόγραμμο.

Σε κάθε πειραματικό τεμάχιο θα πρέπει να σπέρνονται 5-7 γραμμές από τον κάθ’ ένα γενότυπο και αργότερα να συγκομίζονται οι 3-5 μεσαίες γραμμές. Αυτό γίνεται για να αποφευχθεί το φαινόμενο του περιθωρίου, του ανταγωνισμού δηλαδή μεταξύ διαφορετικών γενοτύπων.

8.2. Απλά πειραματικά σχέδια – ανάλυση της παραλλακτικότητας.

Ανάλυση παραλλακτικότητας (Analysis of Variance, ANOVA), είναι μια μέθοδος που επιτρέπει να αποφανθεί ένας ερευνητής αν οι μέσοι όροι του γνωρίσματος που μελετά, παρουσιάζουν μεταξύ τους στατιστικά σημαντικές διαφορές ή όχι. Για να καταλήξει κάποιος στο συμπέρασμα ότι οι μέσοι όροι που συγκρίνει διαφέρουν, υπολογίζει μια τιμή F. Όσο η τιμή F που υπολογίζεται είναι μεγαλύτερη, τόσο αυξάνει η πιθανότητα ότι οι διαφορές που συγκρίνονται είναι σημαντικές.

8.2.1. Αιτίες που επηρεάζουν το F.

(α) όσο μεγαλύτερες είναι οι διαφορές που παρουσιάζουν μεταξύ τους οι παράγοντες, τόσο αυξάνει η τιμή του F.

(β) όσο μεγαλύτερος είναι ο αριθμός των τιμών του δείγματος (n) του κάθε παράγοντα τόσο αυξάνει το F.

(γ) όσο μεγαλύτερη είναι η εσωτερική ανομοιογένεια των τιμών σε κάθε παράγοντα, τόσο μειώνεται η τιμή του F.

8.2 2. Από που εξαρτάται το πειραματικό σφάλμα.

(α) Από την ανομοιογένεια των τιμών μέσα σ΄ ένα δείγμα, λόγω της φύσης του παράγοντα.

(β) Από το αν η δειγματοληψία έγινε σωστά ή όχι και

(γ) Από το περιβάλλον που επηρεάζει διαφορετικά τους διάφορους παράγοντες. Αυτό αντιμετωπίζεται με δυο τρόπους: (α) εξασφαλίζοντας ομοιογενές περιβάλλον και (β) εκτιμώντας με το κατάλληλο πειραματικό σχέδιο την ανομοιογένεια του περιβάλλοντος και αφαιρώντας την επίδρασή της από το πείραμα.

8.3. Πειραματικό Σχέδιο.

Με το πειραματικό σχέδιο ελέγχεται η επίδραση του περιβάλλοντος κατά την εξής γενική αρχή: κατανέμονται οι παράγοντες σε ομάδες, έτσι ώστε κάθε παράγοντας μέσα στην ομάδα να αντιπροσωπεύεται από μια ή περισσότερες τιμές (ομάδες πλήρεις). Κατόπιν, ανάλογα με τις διαφορές που παρουσιάζει το περιβάλλον, χωρίζεται αυτό σε μικροπεριβάλλοντα και σ’ αυτά τοποθετούνται οι ομάδες. Με τον τρόπο αυτό επιτυγχάνεται κάθε ομάδα να πέσει σε περισσότερο ομοιογενές περιβάλλον, και η παραλλακτικότητα που παρουσιάζουν οι ομάδες μεταξύ τους, να είναι μέτρο της επίδρασης του περιβάλλοντος και να αφαιρείται.

Όταν πρέπει να μελετηθούν πολλοί παράγοντες μαζί, τότε το μέγεθος του δείγματος αυξάνει τόσο, ώστε να μην είναι δυνατός ο έλεγχος του περιβάλλοντος. Για να αντιμετωπισθεί η κατάσταση αυτή, συγκροτούνται μικρότερες ομάδες, που δεν περιλαμβάνουν όλους τους παράγοντες (ατελείς ομάδες).

Τα πειραματικά σχέδια διακρίνονται σε: (α) σχέδια χωρίς ομάδες, (β) σχέδια με πλήρεις ομάδες και (γ) σχέδια με ατελείς ομάδες.

8.3.1. Το σχέδιο χωρίς ομάδες, δείγματα ίσα.

Το σχέδιο αυτό, όπως λέει και το όνομά του δεν χρησιμοποιεί επαναλήψεις και από κάθε γενότυπο που μελετάται υπάρχει ο ίδιος αριθμός φυτών.

<u>Παράδειγμα</u>.

Κάποιος ερευνητής θέλησε να δοκιμάσει 10 ποικιλίες μαλακού σιταριού, για να δει κατά πόσο διαφέρουν μεταξύ τους ως προς την ημερομηνία ξεσταχυάσματος. Για το λόγο αυτό σπάρθηκαν στο χωράφι 10 φυτά από κάθε ποικιλία. Η πρωιμότητα του ξεσταχυάσματος υπολογίσθηκε σε ημέρες μετά από μια σταθερή ημερομηνία και τα αποτελέσματα δίδονται στον Πίνακα 5. Διαφέρουν μεταξύ τους οι ποικιλίες;

<u>Λύση</u>.

Η ανάλυση της παραλλακτικότητας, όταν το σχέδιο είναι χωρίς ομάδες με ίσα δείγματα, δίδεται στον Πίνακα 6.

<u>Πίνακας 5</u>. Πρωιμότητα ξεσταχυάσματος 10 ποικιλιών μαλακού σιταριού.

Ποικιλία											Σύνολα ποικιλιών
Α	6	7	6	6	3	4	3	2	6	5	48
Β	3	3	4	4	4	3	3	2	3	3	32
Γ	3	7	7	4	2	2	4	4	2	4	39
Δ	3	3	9	4	2	3	4	4	5	4	41
Ε	9	6	9	5	4	4	5	4	7	6	59
Ζ	2	8	9	5	4	5	4	6	3	2	48
Η	8	9	2	6	3	3	2	4	2	2	41
Θ	8	4	4	3	2	2	4	3	4	2	36
Ι	5	9	9	5	4	3	2	4	4	4	49
Κ	4	4	7	3	4	4	4	4	3	4	41
ΣΥΝΟΛΟ											434

<u>Πίνακας 6</u>. Ανάλυση της παραλλακτικότητας (διακύμανσης) σε σχέδιο χωρίς ομάδες με ίσα δείγματα.

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες	π-1	ΑΤΠ	ΜΤΠ=ΑΤΠ/π-1	ΜΤΠ/ΜΤΣ
Σφάλμα	π(n-1)	ΑΤΣ	ΜΤΣ=ΑΤΣ/π (n-1)	
Σύνολο	πn-1	ΣΑΤ		

όπου π =ο αριθμός των παραγόντων και n = ο αριθμός των τιμών κάθε παράγοντα.

Για να βρεθεί αν οι ποικιλίες παρουσιάζουν μεταξύ τους διαφορές, πρέπει να αναλυθεί η διακύμανση, σύμφωνα με τον Πίνακα 6. Ο υπολογισμός του αθροίσματος των τετραγώνων (ΑΤ) γίνεται με ένα από τους τύπους

$$\mathrm{AT}=\Sigma(\mathrm{X}-\mathrm{X})^2 \text{ ή}$$

$$\mathrm{AT}=\sum \mathrm{X}^2-\frac{(\sum \mathrm{X})^2}{\mathrm{n}}$$

Από τους τύπους αυτούς ο δεύτερος προσφέρεται καλύτερα για τις επιστημονικές αριθμομηχανές.

Η παράσταση $\frac{(\Sigma X)^2}{n}$ ονομάζεται <u>διορθωτικός όρος</u> (ΔΟ), γιατί επαναλαμβάνεται αυτούσια σ' όλα τα ΑΤ.

Η σειρά με την οποία γίνονται οι υπολογισμοί, είναι η ακόλουθη:

(α) ΔΟ= $\frac{(434)^2}{100}$ =1883,56.

(β) ΑΤΠ= $\frac{48^2+32^2+....+49^2+41^2}{10}$ -ΔΟ=1937,4-1883,56=53,84.

(γ) ΣΑΤ=$6^2+7^2+6^2+.........+4^2+3^2+4^2$-ΔΟ=2264-1883,56=380,44

(δ) ΑΤΣ=ΣΑΤ-ΑΤΠ=380,44-53,84=326,6.

Κατόπιν συμπληρώνεται ο Πίνακας της Ανάλυσης της Παραλλακτικότητας.

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες Σφάλμα	9 90	53,84 326,6	5,98 3,63	1,65
Σύνολο	99	380,44		

Από τους Πίνακες που δίδουν τις τιμές του F, για επίπεδο 5% και βαθμούς ελευθερίας 9 (αριθμητή) και 90 (παρονομαστή), ευρίσκεται η τιμή F_{05}=2,0. Η τιμή όμως F που υπολογίσθηκε από τα δεδομένα του Πίνακα 5 είναι μικρότερη από την τιμή F_{05}. Επομένως οι ποικιλίες δεν παρουσιάζουν μεταξύ τους στατιστικά σημαντικές (ΣΣ) διαφορές.

8.3.2. Το σχέδιο χωρίς ομάδες, δείγματα άνισα.

Η ανάλυση της παραλλακτικότητας για το σχέδιο αυτό δίδεται στον Πίνακα 7.

<u>Πίνακας 7</u>. Ανάλυση της παραλλακτικότητας (διακύμανσης) σε σχέδιο χωρίς ομάδες με άνισα δείγματα.

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες Σφάλμα	π-1 Ν-π	ΑΤΠ ΑΤΣ	ΜΤΠ=ΑΤΠ/π-1 ΜΤΣ=ΑΤΣ/Ν-π	ΜΤΠ/ΜΤΣ
Σύνολο	Ν-1	ΣΑΤ		

όπου π =ο αριθμός των παραγόντων και Ν = ο αριθμός τιμών όλων των δειγμάτων.

Το ΑΤΠ=($\Pi_1^2/n_1+\Pi_2^2/n_2+.........+\Pi_\pi^2/n_\pi$)-ΔΟ,

όπου Π_1, Π_2,......τα σύνολα των παραγόντων και

n_1, n_2,.......ο αριθμός τιμών κάθε παράγοντα.

8.3.3. <u>Το σχέδιο των πλήρων ομάδων σε ελεύθερη διάταξη</u>.

Για τα πειράματα που διεξάγονται στον αγρό, η εξασφάλιση ομοιόμορφου περιβάλλοντος είναι σχεδόν αδύνατη. Έτσι, προκύπτει η ανάγκη συγκρότησης επαναλήψεων (ομάδων) που θα πρέπει να διαταχθούν στο χώρο κατά τέτοιο τρόπο, ώστε να υπάρξει αποτελεσματικότερος έλεγχος της επίδρασης του περιβάλλοντος. Αν για παράδειγμα υπάρχουν κηλίδες εδάφους διαφορετικής σύστασης, τότε οι επαναλήψεις θα πρέπει να τοποθετηθούν, αν αυτό είναι δυνατό, μια σε κάθε διαφορετική κηλίδα. Το πειραματικό σχέδιο που χρησιμοποιείται στην περίπτωση αυτή ονομάζεται σχέδιο <u>πλήρων ομάδων σε ελεύθερη διάταξη</u> (ΠΟΕΔ).

<u>Παράδειγμα</u>.

Θα χρησιμοποιηθούν τα δεδομένα του παραδείγματος της ανάλυσης σε σχέδιο χωρίς ομάδες, με ίσο αριθμό δείγματος (Πίνακας 8), για να φανούν και οι διαφορές μεταξύ των δυο πειραματικών σχεδίων. Η επί πλέον διαφορά που υπάρχει στο σχέδιο των πλήρων ομάδων σε ελεύθερη διάταξη είναι ότι οι παράγοντες συγκροτούν ομάδες προς τη μια κατεύθυνση του αγρού, για την οποία υπάρχει η υπόνοια ότι το περιβάλλον είναι ανομοιογενές. Συνολικά υπάρχουν 10 ομάδες, O_1 έως O_{10}. Με βάση την πληροφορία αυτή να βρεθεί αν οι γενότυποι διαφέρουν μεταξύ τους.

<u>Λύση</u>.

Επειδή κάθε ομάδα περιλαμβάνει τους ιδίους γενότυπους, θα έπρεπε και τα σύνολά τους να είναι ίδια. Αυτό όμως δεν συμβαίνει γιατί υπάρχει η αλληλεπίδραση του γενότυπου με το περιβάλλον. Η παραλλακτικότητα που παρουσιάζουν μεταξύ τους οι ομάδες, οφείλεται αποκλειστικά στην προηγούμενη αλληλεπίδραση. Έτσι, για να μειωθεί το σφάλμα θα πρέπει η παραλλακτικότητα αυτή να αφαιρεθεί από την συνολική. Εδώ ακριβώς ευρίσκεται η διαφορά του σχεδίου αυτού από το σχέδιο χωρίς ομάδες. Η ανάλυση της παραλλακτικότητας του σχεδίου των πλήρων ομάδων σε ελεύθερη διάταξη δίδεται στον Πίνακα 9.

<u>Πίνακας 8</u>. Πρωιμότητα ξεσταχυάσματος 10 ποικιλιών μαλακού σιταριού.

Ποικιλία	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	Σύνολα ποικιλιών
Α	6	7	6	6	3	4	3	2	6	5	48
Β	3	3	4	4	4	3	3	2	3	3	32
Γ	3	7	7	4	2	2	4	4	2	4	39
Δ	3	3	9	4	2	3	4	4	5	4	41
Ε	9	6	9	5	4	4	5	4	7	6	59
Ζ	2	8	9	5	4	5	4	6	3	2	48
Η	8	9	2	6	3	3	2	4	2	2	41
Θ	8	4	4	3	2	2	4	3	4	2	36
Ι	5	9	9	5	4	3	2	4	4	4	49
Κ	4	4	7	3	4	4	4	4	3	4	41
ΣΥΝΟΛΟ	51	60	66	45	32	33	35	37	39	36	434

Πίνακας 9. Ανάλυση της παραλλακτικότητας (διακύμανσης) σε σχέδιο πλήρων ομάδων σε ελεύθερη διάταξη (ΠΟΕΔ).

Πηγή παραλ-κότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες Ομάδες Σφάλμα	π-1 ο-1 (π-1)(ο-1)	ΑΤΠ ΑΤΟ ΑΤΣ	ΜΤΠ=ΑΤΠ/π-1 ΜΤΟ=ΑΤΟ/ο-1 ΜΤΣ=ΑΤΣ/(π -1)(n-1)	ΜΤΠ/ΜΤΣ ΜΤΟ/ ΜΤΣ
Σύνολο	πο-1	ΣΑΤ		

όπου π =ο αριθμός των παραγόντων και ο = ο αριθμός των ομάδων.

Η σειρά των υπολογισμών είναι η ακόλουθη:

(α) ΔΟ= $\frac{(434)^2}{100}$ =1883,56.

(β) ΑΤΠ= $\frac{48^2+32^2+......49^2+41^2}{10}$ -ΔΟ=1937,4-1883,56=53,84.

(γ) ΑΤΟ= $\frac{51^2+60^2+......+39^2+36^2}{10}$ -ΔΟ=127,04.

(δ) ΣΑΤ=$6^2+7^2+6^2+.........+4^2+3^2+4^2$-ΔΟ=2264-1883,56=380,44.

(ε) ΑΤΣ=ΣΑΤ-ΑΤΠ-ΑΤΟ=380,44-53,84-127,04=199,56.

Με βάση τα δεδομένα αυτά συμπληρώνεται ο Πίνακας της ανάλυσης της παραλλακτικότητας.

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες	9	53,84	5,98	2,43
Ομάδες	9	127,04	14,12	5,74
Σφάλμα	81	199,56	2,46	
Σύνολο	99	380,44		

Για τους παράγοντες: από τους πίνακες του F, για πιθανότητα σφάλματος 5% και για ΒΕ αριθμητή=9 και παρονομαστή=81 αντιστοιχεί μια τιμή F=2,012 (κατ' όμοιο τρόπο ευρίσκεται και η αντίστοιχη τιμή F των ομάδων). Συγκρίνοντας την τιμή F=2,43 με την τιμή του F_{05}=2,012 που υπολογίσθηκε από τους πίνακες, προκύπτει ότι $F_{δεδ}>F_{05}$. Συνεπώς οι 10 ποικιλίες μαλακού σιταριού διαφέρουν μεταξύ τους.

Χρησιμοποιώντας λοιπόν τα ίδια δεδομένα με το σχέδιο χωρίς ομάδες και εκεί που το τελευταίο δεν μπορούσε να εντοπίσει διαφορές, το πειραματικό σχέδιο των ΠΟΕΔ βρήκε ότι οι ποικιλίες που δοκιμάσθηκαν διαφέρουν μεταξύ τους. Αυτό επιτεύχθηκε γιατί με τις επαναλήψεις που χρησιμοποιήθηκαν στο σχέδιο ΠΟΕΔ αφαιρέθηκε το σφάλμα που οφείλεται στην ετερογένεια του εδάφους (το σφάλμα μειώθηκε σε 2,46, από 3,63 που ήταν με το σχέδιο χωρίς ομάδες). Η μείωση αυτή του σφάλματος είχε ως συνέπεια την αύξηση της τιμής του F, με αποτέλεσμα να βρεθούν διαφορές μεταξύ των ποικιλιών που δοκιμάσθηκαν.

Για να απαντηθεί το ερώτημα πόσο μεγάλες είναι αυτές οι διαφορές το $F_{δεδ}$ θα πρέπει να συγκριθεί με την αντίστοιχη τιμή των πινάκων, για πιθανότητα σφάλματος 1%. Έτσι από τον αντίστοιχο πίνακα και ΒΕ 9 και 81 αντιστοιχεί μια τιμή F_{01}=2,66> $F_{δεδ}$ =2,43. Επομένως οι διαφορές <u>δεν είναι πολύ σημαντικές</u>.

Γενικά, στον Γεωργικό Πειραματισμό και για να αποφευχθούν τα πολλά λόγια, χρησιμοποιείται ο παρακάτω συμβολισμός:

(α) $F_{δεδ}< F_{05}$, τότε δεν υπάρχουν διαφορές και δεν σημειώνεται τίποτε διπλά στην τιμή $F_{δεδ}$.

(β) $F_{05} < F_{δεδ} < F_{01}$, τότε υπάρχουν σημαντικές διαφορές, αλλά αυτές είναι μικρές και δίπλα στην τιμή του $F_{δεδ}$ μπαίνει ένας αστερίσκος, δηλ. $F_{δεδ}$ *, και

(γ) $F_{δεδ} > F_{01}$, τότε οι διαφορές που υπάρχουν είναι πολύ σημαντικές και δίπλα στην τιμή του $F_{δεδ}$ μπαίνουν δύο αστερίσκοι, δηλ. $F_{δεδ}$ **.

8.4. <u>Λατινικό τετράγωνο</u>.

Εκτός από την περίπτωση που αναλύθηκε στο παράδειγμα των πλήρων ομάδων σε ελεύθερη διάταξη, υπάρχουν περιπτώσεις όπου η ετερογένεια του περιβάλλοντος προχωρά προς δυο κατευθύνσεις που είναι κάθετες μεταξύ τους. Όταν συμβαίνει αυτό, είναι δυνατό να διαταχθούν οι επαναλήψεις κατά τέτοιο τρόπο, ώστε να είναι δυνατός ο έλεγχος της ετερογένειας και προς τις δυο κατευθύνσεις. Η διάταξη αυτή ονομάζεται <u>Λατινικό τετράγωνο</u> (Πίνακας 10). Ο Πίνακας του Λατινικού τετραγώνου έχει 10 γραμμές και 10 στήλες. Κάθε μια από τις γραμμές αποτελεί μια πλήρη επανάληψη, που περιλαμβάνει και τους 10 γενοτύπους που δοκιμάζονται, πλήρως τυχαιοποιημένους.

<u>Πίνακας 10</u>. Πρωιμότητα ξεσταχυάσματος 10 ποικιλιών μαλακού σιταριού.

Γραμμές	Στήλες										Σύνολα γραμμών
	1	2	3	4	5	6	7	8	9	10	
1	Η 8	Ζ 2	Ε 9	Γ 3	Α 6	Κ 4	Θ 8	Δ 3	Β 3	Ι 5	51
2	Γ 7	Α 7	Κ 4	Η 9	Θ 4	Β 3	Ζ 8	Ι 9	Δ 3	Ε 6	60
3	Β 4	Η 2	Θ 4	Α 6	Ζ 9	Ε 9	Ι 9	Γ 7	Κ 7	Δ 9	66
4	Ε 5	Θ 3	Ζ 5	Κ 3	Ι 5	Δ 4	Α 6	Η 6	Γ 4	Β 4	45
5	Θ 2	Ι 4	Α 3	Β 4	Δ 2	Η 3	Κ 4	Ε 4	Ζ 4	Γ 2	32
6	Κ 4	Δ 3	Ι 3	Θ 2	Γ 2	Α 4	Ε 4	Β 3	Η 3	Ζ 5	33
7	Δ 4	Ε 5	Η 2	Ι 2	Κ 4	Γ 4	Β 3	Ζ 4	Α 3	Θ 4	35
8	Α 2	Β 2	Γ 4	Δ 4	Ε 4	Ζ 6	Η 4	Θ 3	Ι 4	Κ 4	37
9	Ζ 3	Γ 2	Β 3	Ε 7	Η 2	Ι 4	Δ 5	Κ 3	Θ 4	Α 6	39
10	Ι 4	Κ 4	Δ 4	Ζ 2	Β 3	Θ 2	Γ 4	Α 5	Ε 6	Η 2	36
Σύνολα στηλών	43	34	41	42	41	43	55	47	41	47	434

Τα σύνολα τόσο των γραμμών όσο και των στηλών περιέχουν τους ίδιους γενότυπους. Παρ' όλα αυτά δεν είναι ίδια αλλά διαφέρουν. Επίσης, τα σύνολα των γραμμών παρουσιάζουν μεγαλύτερη διακύμανση. Αυτό σημαίνει ότι η ετερογένεια του εδάφους ήταν μεγαλύτερη προς αυτήν την κατεύθυνση. Το λατινικό τετράγωνο επιτρέπει να αφαιρεθεί από τη συνολική διακύμανση, τόσο η διακύμανση των γραμμών, όσο και η διακύμανση των στηλών. Για το λόγο αυτό το πειραματικό σφάλμα μειώνεται ικανοποιητικά και αυξάνεται η ακρίβεια του πειραματισμού. Ο Πίνακας XII στο Παράρτημα δείχνει τα διάφορα σχέδια των λατινικών τετραγώνων. Για να γίνει κατανοητό αυτό το πειραματικό σχέδιο του λατινικού τετραγώνου, θα χρησιμοποιηθεί το παρακάτω παράδειγμα.

Παράδειγμα.

Με βάση τα δεδομένα του Πίνακα 10, που προέρχονται από λατινικό τετράγωνο, να βρεθεί αν οι 10 ποικιλίες σιταριού διαφέρουν ως προς την πρωϊμότητα.

Λύση.

Ο Πίνακας 11 δίνει την ανάλυση της παραλλακτικότητας για το σχέδιο του λατινικού τετραγώνου.

Πίνακας 11. Ανάλυση της παραλλακτικότητας σε σχέδιο λατινικού τετραγώνου.

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F
Παράγοντες	π - 1	ΑΤΠ	ΜΤΠ = ΑΤΠ/π-1	ΜΤΠ/ΜΤΣ
Γραμμές	π - 1	ΑΤγ	ΜΤγ = ΑΤγ/π-1	ΜΤγ/ΜΤΣ
Στήλες	π - 1	ΑΤσ	ΜΤσ = ΑΤσ/π-1	ΜΤσ/ΜΤΣ
Σφάλμα	(π – 1)(π – 2)	ΑΤΣ	ΜΤΣ= ΑΤΣ/(π-1)(π-2)	
Σύνολο	$π^2$ - 1	ΣΑΤ		

Αθροίζοντας όλες τις τιμές κάθε μιας ποικιλίας του Πίνακα 10 προκύπτουν τα σύνολα από την κάθε μια τους:

Ποικιλία	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ
Σύνολα	48	32	39	41	59	48	41	36	49	41
μ. ό.	4,8	3,2	3,9	4,1	5,9	4,8	4,1	3,6	4,9	4,1

Η σειρά των υπολογισμών είναι η ακόλουθη:

(α) $ΔΟ= \dfrac{(434)^2}{100}=1883{,}56.$

(β) $ΑΤΠ= \dfrac{48^2+32^2+......49^2+41^2}{10}-ΔΟ=1937{,}4-1883{,}56=53{,}84.$

(γ) $ΑΤγ= \dfrac{51^2+60^2+......+39^2+36^2}{10}-ΔΟ=127{,}04.$

(δ) $ΑΤσ = \dfrac{43^2+34^2+......+41^2+47^2}{10}- ΔΟ= 1910{,}4-1883{,}56=26{,}84$

(ε) $ΣΑΤ=8^2+2^2+9^2+.........+5^2+6^2+2^2-ΔΟ=2264-1883{,}56=380{,}44.$

(ζ) ΑΤΣ=ΣΑΤ – ΑΤΠ - ΑΤγ – ΑΤσ =380,44-53,84-127,04- 26,84=172,72.

Με βάση τα δεδομένα αυτά συμπληρώνεται ο Πίνακας ανάλυσης της παραλλακτικότητας

Πηγή παραλλακτικότητας	ΒΕ	ΑΤ	ΜΤ	F	F_{05}
Παράγοντες	9	53,84	5,98	2,49	2,01
Γραμμές	9	127,04	14,12	5,88	2,01
Στήλες	9	26,84	2,98	1,24	2,01
Σφάλμα	72	172,72	2,4		
Σύνολο	99	380,44			

Χρησιμοποιώντας λοιπόν τα ίδια δεδομένα με το σχέδιο πλήρεις ομάδες σε ελεύθερη διάταξη το ΜΤΣ μειώθηκε σε 2,4 από 2,46 (που ήταν με το σχέδιο πλήρεις ομάδες σε ελεύθερη διάταξη). Αυτό σημαίνει ότι το σχέδιο του λατινικού τετραγώνου μείωσε το σφάλμα περισσότερο από το σχέδιο των πλήρων ομάδων σε ελεύθερη διάταξη και αποδείχθηκε έτσι πιο αποτελεσματικό. Η αποτελεσματικότητα ενός πειραματικού σχεδίου εξαρτάται από το μέγεθος της επίδρασης του περιβάλλοντος.

8.5. Συντελεστής παραλλακτικότητας.

Η ακρίβεια ενός πειράματος μπορεί να αξιολογηθεί με τη βοήθεια ενός συντελεστή που ονομάζεται συντελεστής παραλλακτικότητας (CV: coefficient of variation). Ο συντελεστής παραλλακτικότητας είναι το τυπικό σφάλμα εκφρασμένο ως επί τοις % του μέσου όρου, είναι ανεξάρτητος από τις μονάδες μέτρησης που χρησιμοποιούνται, ενώ εκφράζεται πάντα σε εκατοστά. Αποτελεί ένα μέτρο για να εκτιμηθεί η σχετική παραλλακτικότητα όταν οι μονάδες μέτρησης και οι μέσοι όροι είναι διαφορετικοί. Ο συντελεστής παραλλακτικότητας υπολογίζεται από τον τύπο

$$\mathrm{CV} + \frac{\sqrt{\mathrm{ΜΤΣ}}}{\overline{\mathrm{X}}}$$

Η τετραγωνική ρίζα του ΜΤΣ είναι το τυπικό σφάλμα. Το $\overline{\mathrm{X}}$ είναι ο γενικός μέσος όρος του πειράματος.

Ο κύριος σκοπός της ανάλυσης της παραλλακτικότητας είναι να μειωθεί το πειραματικό σφάλμα (και όχι μόνο να υπολογισθεί ένα σημαντικό F). Όσο μικραίνει το σφάλμα τόσο μικραίνει η ελάχιστη σημαντική διαφορά (για την ελάχιστη σημαντική διαφορά δες παρακάτω). Μια μικρή τιμή ελάχιστης σημαντικής διαφοράς σημαίνει ότι περισσότεροι μέσοι όροι θα διαφέρουν μεταξύ τους, ενώ ακόμα και μια μικρή διαφορά μεταξύ δυο συγκρινόμενων μέσων όρων μπορεί να είναι σπουδαία. Μια τιμή του συντελεστή παραλλακτικότητας γύρω στο 10% θεωρείται ιδιαίτερα ικανοποιητική στον γεωργικό πειραματισμό. Κατά το Φασούλα, ο συντελεστής αυτός μετρά το φορτίο των εκφυλιστικών γονιδίων που υπάρχουν σε ένα φυτικό είδος.

8.6. <u>Σύγκριση των μέσων όρων των ποικιλιών που μελετήθηκαν</u>.

Στο παράδειγμα της σελίδας 68 βρέθηκε ότι το $F_{\delta\varepsilon\delta}$*. Αυτό σημαίνει ότι ανάμεσα στους γενότυπους υπάρχουν κάποιοι που διαφέρουν από κάποιους άλλους σημαντικά. Οι πιο κοινές μέθοδοι για να αποφανθεί κάποιος ποίοι είναι οι γενότοιποι που διαφέρουν είναι οι εξής: (α) <u>Ελάχιστη Σημαντική διαφορά</u> (ΕΣΔ), (β) <u>δοκιμή του</u> <u>Duncan</u> και (γ) <u>μέθοδος του Scheffe</u>. Από αυτές συνήθως χρησιμοποιείται η πρώτη, που είναι εύκολη και αξιόπιστη, αν ο αριθμός των μέσων όρων δεν είναι μεγάλος.

8.6.1. <u>Η μέθοδος της ελάχιστης σημαντικής διαφοράς</u>.

8.6.1.1. <u>ΕΣΔ, σχέδιο χωρίς ομάδες, δείγματα ίσα</u>.

Χρησιμοποιείται ο τύπος

$$ΕΣΔ=t_{05}\sqrt{\frac{2(ΜΤΣ)}{n}}$$

όπου n: ο αριθμός των τιμών του δείγματος.

Οι ΒΕ είναι π(n-1).

8.6.1.2. <u>ΕΣΔ, σχέδιο χωρίς ομάδες, δείγματα άνισα</u>.

Χρησιμοποιείται ο τύπος

$$ΕΣΔ=t_{05}\sqrt{ΜΤΣ\frac{n_1+n_2}{n_1 n_2}}$$

όπου n_1 και n_2: οι τιμές που αντιστοιχούν στους υπό σύγκριση μέσους όρους.

Οι ΒΕ είναι Ν-π.

8.6.1.3. <u>ΕΣΔ, σχέδιο πλήρεις ομάδες σε ελεύθερη διάταξη</u>.

Χρησιμοποιείται ο τύπος

$$ΕΣΔ=t_{05}\sqrt{\frac{2(ΜΤΣ)}{ο}}$$

όπου ο: ο αριθμός των ομάδων.

Οι ΒΕ είναι (π-1)(ο-1)

Για το παράδειγμα των πλήρων ομάδων σε ελεύθερη διάταξη (Πίνακας 8)

$$ΕΣΔ=1,99\sqrt{\frac{2(2,46)}{10}}=1,40.$$

Τα βήματα που ακολουθούνται για να βρεθεί ο καλύτερος γενότυπος είναι τα ακόλουθα:

(α) υπολογίζεται η ΕΣΔ=1,40.

(β) οι μέσοι όροι του γνωρίσματος που μελετάται, κατατάσσονται κατά σειρά μεγέθους, αρχίζοντας από τον μεγαλύτερο και προχωρώντας προς τον μικρότερο. Για το προηγούμενο παράδειγμα:

$\overline{X}_E$ $\overline{X}_I$ $\overline{X}_{A,Z}$ $\overline{X}_{\Delta,H,K}$ $\overline{X}_\Gamma$ $\overline{X}_\Theta$ $\overline{X}_B$
5,9 4,9 4,8 4,1 3,9 3,6 3,2

(γ) όλοι οι μέσοι όροι συγκρίνονται ανά δύο. Όταν η διαφορά δύο συγκρινόμενων μέσων όρων είναι μικρότερη από την ΕΣΔ, τότε οι μέσοι όροι δεν διαφέρουν και ενώνονται με μία γραμμή. Όταν η διαφορά δύο μέσων όρων είναι μεγαλύτερη από την ΕΣΔ, τότε οι μέσοι όροι διαφέρουν και η γραμμή σταματά. Έτσι για το προηγούμενο παράδειγμα, αφού γίνει η σύγκριση των διαφορών των μέσων όρων με την ΕΣΔ και αυτοί που διαφέρουν ενωθούν με μία γραμμή, η εικόνα διαφοροποιείται όπως φαίνεται παρακάτω:

$\overline{X}_E$ $\overline{X}_I$ $\overline{X}_{A,Z}$ $\overline{X}_{\Delta,H,K}$ $\overline{X}_\Gamma$ $\overline{X}_\Theta$ $\overline{X}_B$
5,9 4,9 4,8 4,1 3,9 3,6 3,2

8.6.1.3. <u>ΕΣΔ, σχέδιο λατινικό τετράγωνο</u>.

Χρησιμοποιείται ο τύπος

$$\text{ΕΣΔ}=t_{05}\sqrt{\frac{2(\text{M}(\text{MT}}{\pi}}$$

όπου π: ο αριθμός των παραγόντων ή γραμμών ή στηλών και το υπολογίζεται για τους ΒΕ του σφάλματος που είναι (π-1)(π-2)

Για το παράδειγμα του Λατινικού Τετραγώνου (Πίνακας 10) η ΕΣΔ είναι

$$\text{ΕΣΔ}=1{,}98\sqrt{\frac{2(2{,}4)}{10}} = 1{,}372$$

Όσοι μέσοι όροι παρουσιάζουν διαφορές μεγαλύτερες από την ΕΣΔ διαφέρουν μεταξύ τους σημαντικά.

8.6.2. <u>Η μέθοδος του ελάχιστου σημαντικού εύρους (ή δοκιμή του Duncan)</u>.

Η μέθοδος της ΕΣΔ έχει ένα μειονέκτημα: το t-κριτήριο χρησιμοποιείται όχι μόνο για να συγκριθούν δύο μέσοι όροι αλλά και για να συγκριθούν όλοι οι μέσοι όροι. Στην περίπτωση αυτή μερικές φορές, συγκρίνονται δυο μέσοι όροι με μια ΕΣΔ που είναι μικρότερη από εκείνη που θα προέκυπτε αν αγνοούνταν όλοι οι υπόλοιποι μέσοι όροι. Αυτό ακριβώς το σφάλμα προσπαθεί να αντιμετωπίσει η <u>δοκιμή του Duncan</u> ή <u>μέθοδος του ελάχιστου σημαντικού εύρους</u>. Σύμφωνα με τη μέθοδο αυτή, αν τοποθετηθούν οι μέσοι όροι κατά σειρά μεγέθους, τότε δυο γειτονικοί μέσοι όροι ορίζουν ένα εύρος που περικλείει 2 μέσους όρους. Ο τρίτος με τον έκτο ορίζουν ένα άλλο εύρος, που περικλείει 4 μέσους όρους κ. ο. κ. Ο συμβολισμός που χρησιμοποιεί η δοκιμή του Duncan για το εύρος που περικλείουν δύο μέσοι όροι είναι E_2, E_3, E_4..., όταν οι δυο συγκρινόμενοι μέσοι όροι περικλείουν ένα εύρος με 2, 3, 4,...., μέσους όρους αντίστοιχα. Οι τιμές του Ε υπολογίζονται από τον τύπο

$$ΕΣΕ = ε\sqrt{\frac{ΜΤΣ}{n}}$$

Οι τιμές του ε λαμβάνονται από τον Πίνακα VIII στο Παράρτημα και εξαρτώνται από τους Β. Ε. και από τον αριθμό των μέσων όρων που περικλείονται στο εύρος. Η παράσταση $\sqrt{\frac{ΜΤΣ}{n}}$ είναι το τυπικό σφάλμα.. Ο παρονομαστής του κλάσματος είναι:

(α) για την περίπτωση του σχεδίου χωρίς επαναλήψεις, ο αριθμός των τιμών του δείγματος (n),

(β) για την περίπτωση του σχεδίου των πλήρων ομάδων σε ελεύθερη διάταξη, ο αριθμός των ομάδων (ο) και

(γ) για την περίπτωση του Λατινικού τετράγωνου, ο αριθμός των σειρών ή γραμμών (π).

Για να γίνει περισσότερο κατανοητή η εφαρμογή της δοκιμής του Dunkan θα χρησιμοποιηθούν τα δεδομένα του παραδείγματος της σελίδας 57. Όπως και στην περίπτωση της ΕΣΔ, έτσι και εδώ ακολουθούνται κάποια βήματα που είναι:

(α) κατατάσσονται οι μέσοι όροι κατά σειρά μεγέθους

$\overline{X}_E$	$\overline{X}_I$	$\overline{X}_{A,Z}$	$\overline{X}_{Δ,Η,Κ}$	$\overline{X}_Γ$	$\overline{X}_Θ$	$\overline{X}_B$
5,9	4,9	4,8	4,1	3,9	3,6	3,2

(β) Υπολογίζεται το τυπικό σφάλμα

$$S_{\overline{X}} = \sqrt{\frac{2,46}{10}} = 0,5$$

(γ) από τον Πίνακα VIII του Παραρτήματος, και για Β. Ε. = 60 (κανονικά οι Β. Ε. είναι 81, χρησιμοποιούνται όμως 60 γιατί οι 81 δεν περιλαμβάνονται στον Πίνακα VIII) υπολογίζονται οι τιμές του ε, που είναι

$$ε_2 = 2,83,\ ε_3 = 2,98 \ldots ε_{10} = 3,33.$$

(δ) οι τιμές των ε πολλαπλασιάζονται με το τυπικό σφάλμα, για να βρεθεί το ελάχιστο σημαντικό εύρος, που αντιστοιχεί σε κάθε περίπτωση E_2 = 2,83 Χ 0,5 = 1,415, E_3 = 2,98 Χ 0,5 =1,49.......κ. ο. κ. Έτσι προκύπτουν οι παρακάτω τιμές

E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
1,42	1,49	1,54	1,57	1,6	1,62	1,64	1,66	1,67

(ε) στη συνέχεια συγκρίνονται οι γενότυποι κατά τη σειρά: ο Ε με το Β, Θ, Γ, κ. λ. π., ο Ι με το Β, Θ, Γ, κ. λ. π., Ο Α με το Β, Θ, Γ, κ. λ. π., έως ότου γίνει η τελική σύγκριση μεταξύ Θ και Β.

Έτσι για το παραπάνω παράδειγμα, συγκρίνεται πρώτα ο γενότυπος Ε με τον γενότυπο Β. Από την τιμή του Ε αφαιρείται η τιμή E_{10}, (γιατί οι δύο αυτοί γενοτύποι καθορίζουν ένα εύρος που περικλείει 10 γενοτύπους), δηλαδή 5,9 – 1,67 = 4,23. Η τιμή αυτή είναι μεγαλύτερη από την τιμή των γενοτύπων Β, Θ, Γ, Κ, Η και Δ. Επομένως ο γενότυπος Ε διαφέρει από όλους αυτούς που αναφέρθηκαν προηγούμενα. Στη συνέχεια ο γενότυπος Ε συγκρίνεται με το γενότυπο Ζ, αφαιρώντας από την τιμή του Ε την τιμή E_4, γιατί μεταξύ Ε και Ζ περικλείονται τέσσερις μέσοι όροι, δηλαδή 5,9 – 1,54 = 4, 36. Επειδή το 4,36 είναι μικρότερο από τον μέσο όρο του Ζ γενοτύπου, αυτό σημαίνει ότι οι γενότυποι Ε και Ζ δεν διαφέρουν μεταξύ τους σημαντικά και όπως και στην περίπτωση της ΕΣΔ, ενώνονται οι δυο γενότυποι με μια γραμμή. Η ίδια εργασία συνεχίζεται και με τους υπόλοιπους γενότυπους (ο Ι συγκρίνεται με τον Β, στη συνέχεια με το Θ κ. ο. κ.) Όταν το ελάχιστο σημαντικό εύρος βρεθεί να είναι μικρότερο από το γενότυπο με τον οποίο συγκρίνεται ο συγκεκριμένος, τότε η σύγκριση σταματά.

8.6.3. Μέθοδος του Scheffe.

Όταν οι συγκρίσεις που θέλει να κάνει κάποιος δεν έχουν προγραμματισθεί εκ των προτέρων, αλλά αντίθετα αυτές υπαγορεύονται εκ των υστέρων από τα δεδομένα, τότε δεν μπορεί να χρησιμοποιηθεί το t-κριτήριο. Στις περιπτώσεις αυτές για να συγκριθούν όλοι οι μέσοι όροι ανά δύο χρησιμοποιείται η μέθοδος του Scheffe. Σύμφωνα με τη μέθοδο αυτή οι διαφορά δύο μέσων όρων συγκρίνεται με το πηλίκο $|L| / S_L$ και μια διαφορά θεωρείται ότι είναι στατιστικώς σημαντική όταν είναι μεγαλύτερη από το παραπάνω πηλίκο. Το $L = \overline{X}_i - \overline{X}_j$ και το $S_L = \frac{\sqrt{2S^2}}{n}$.

ΠΑΡΑΡΤΗΜΑ

Πίνακας Ι. Επιφάνειες κάτω από την κανονική καμπύλη από το 0 έως το z.

z	0	1	2	3	4	5	6	7	8	9
0,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
0,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0754
0,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
0,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
0,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
0,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
0,6	,2258	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2518	,2549
0,7	,2580	,2612	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
0,8	,2881	,2910	,2939	,2967	,2996	,3023	,3051	,3078	,3106	,3133
0,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389
1,0	,3413	,3438	,3461	,3485	,3508	,3531	,3554	,3577	,3599	,3621
1,1	,3643	,3665	,3686	,3708	,3729	,3749	,3770	,3790	,3810	,3830
1,2	,3849	,3869	,3888	,3907	,3925	,3944	,3962	,3980	,3997	,4015
1,3	,4032	,4049	,4066	,4082	,4099	,4115	,4131	,4147	,4162	,4177
1,4	,4192	,4207	,4222	,4236	,4251	,4265	,4279	,4292	,4306	,4319
1,5	,4332	,4345	,4357	,4370	,4382	,4394	,4406	,4418	,4429	,4441
1,6	,4452	,4463	,4474	,4484	,4495	,4505	,4515	,4525	,4535	,4545
1,7	,4554	,4564	,4573	,4582	,4591	,4599	,4608	,4616	,4625	,4633
1,8	,4641	,4649	,4656	,4664	,4671	,4678	,4686	,4693	,4699	,4706
1,9	,4713	,4719	,4726	,4732	,4738	,4744	,4750	,4756	,4761	,4767
2,0	,4772	,4778	,4783	,4788	,4793	4798	,4803	,4808	,4812	,4817
2,1	,4821	,4826	,4830	,4834	,4838	,4842	,4846	,4850	,4854	,4857
2,2	,4861	,4864	,4868	,4871	,4875	,4878	,4881	,4884	,4887	,4890
2,3	,4893	,4896	,4898	,4901	,4904	,4906	,4909	,4911	,4913	,4916
2,4	,4918	,4920	,4922	,4925	,4927	,4929	,4931	,4932	,4934	,4936
2,5	,4938	,4940	,4941	,4943	,4945	,4946	,4948	,4949	,4951	,4952
2,6	,4953	,4955	,4956	,4957	,4959	,4960	,4961	,4962	,4963	,4964
2,7	,4965	,4966	,4967	,4968	,4969	,4970	,4971	,4972	,4973	,4974
2,8	,4974	,4975	,4976	,4977	,4977	,4978	,4979	,4979	,4980	,4981
2,9	,4981	,4982	,4982	,4983	,4984	,4984	,4985	,4985	,4986	,4986
3,0	,4987	,4987	,4987	,4988	,4988	,4989	,4989	,4989	,4990	,4990
3,1	,4990	,4991	,4991	,4991	,4992	,4992	,4992	,4992	,4993	,4993
3,2	,4993	,4993	,4994	,4994	,4994	,4994	,4994	,4995	,4995	,4995
3,3	,4995	,4995	,4995	,4996	,4996	,4996	,4996	,4996	,4996	,4997
3,4	,4997	,4997	,4997	,4997	,4997	,4997	,4997	,4997	,4997	,4998
3,5	,4998	,4998	,4998	,4998	,4998	,4998	,4998	,4998	,4998	,4998
3,6	,4998	,4998	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999
3,7	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999
3,8	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999	,4999
3,9	,5000	,5000	,5000	,5000	,5000	,5000	,5000	,5000	,5000	,5000

<u>Πίνακας II</u>. Τιμές z για διάφορες πιθανότητες και κριτήρια.

Κριτήριο	$Z_{,10}$	$Z_{,05}$	$Z_{,01}$	$Z_{,003}$	$Z_{,005}$
Προς τη μια πλευρά	1,28	1,645	2,33	2,58	2,88
Προς τις δυο πλευρές	1,645	1,96	2,58	2,81	3,08

Πίνακας III. Τιμές t για διάφορες πιθανότητες και βαθμούς ελευθερίας.

Β. Ε.	$t_{,10}$	$t_{,05}$	$t_{,02}$	$t_{,01}$	$t_{,001}$
1	6,314	12,706	31,821	63,657	636,619
2	2,920	4,303	6,965	9,925	31,598
3	2,353	3,182	4,541	5,841	12,941
4	2,132	2,776	3,747	4,604	8,610
5	2,015	2,571	3,365	4,032	6,859
6	1,943	2,447	3,143	3,707	5,959
7	1,895	2,365	2,998	3,499	5,405
8	1,860	2,306	2,896	3,355	5,041
9	1,833	2,262	2,821	3,250	4,781
10	1,812	2,228	2,764	3,169	4,587
11	1,796	2,201	2,718	3,106	4,437
12	1,782	2,179	2,681	3,055	4,318
13	1,771	2,160	2,650	3,012	4,221
14	1,761	2,145	2,624	2,977	4,140
15	1,753	2,131	2,602	2,947	4,073
16	1,746	2,120	2,583	2,921	4,015
17	1,740	2,110	2,567	2,898	3,965
18	1,734	2,101	2,552	2,878	3,922
19	1,729	2,093	2,539	2,861	3,883
20	1,725	2,086	2,528	2,845	3,850
21	1,721	2,080	2,518	2,831	3,819
22	1,717	2,074	2,508	2,819	3,792
23	1,714	2,069	2,500	2,807	3,767
24	1,711	2,064	2,492	2,797	3,745
25	1,708	2,060	2,485	2,787	3,725
26	1,706	2,056	2,479	2,779	3,707
27	1,703	2,052	2,473	2,771	3,690
28	1,701	2,048	2,467	2,763	3,674
29	1,699	2,045	2,462	2,756	3,659
30	1,697	2,042	2,457	2,750	3,646
40	1,684	2,021	2,423	2,704	3,551
60	1,671	2,000	2,390	2,660	3,460
120	1,658	1,980	2,358	2,617	3,373
∞	1,645	1,960	2,326	2,576	3,291

<u>Πίνακας IV</u>. Τιμές F για πιθανότητα 5% και διάφορους βαθμούς ελευθερίας.

ΒΕ	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Οι Β. Ε. του αριθμητή είναι στον οριζόντιο άξονα και του παρονομαστή στον κάθετο.

<u>Πίνακας V</u>. Τιμές F για πιθανότητα 1% και διάφορους βαθμούς ελευθερίας.

ΒΕ	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14 0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.17	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Οι Β. Ε. του αριθμητή είναι στον οριζόντιο άξονα και του παρονομαστή στον κάθετο.

Πίνακας VI. Τιμές X^2 για διάφορες πιθανότητες και βαθμούς ελευθερίας.

ΒΕ	$X^2_{.99}$	$X^2_{.98}$	$X^2_{.95}$	$X^2_{.90}$	$X^2_{.80}$	$X^2_{.70}$	$X^2_{.50}$	$X^2_{.30}$	$X^2_{.20}$	$X^2_{.10}$	$X^2_{.05}$	$X^2_{.02}$	$X^2_{.01}$	$X^2_{.001}$
1	.00016	.00063	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	.12	.18	.35	.58	1.00	1.42	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	.30	.43	.71	1.06	1.65	2.20	3.36	4 88	5.99	7.78	9.49	11.67	13.28	18.46
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	.87	1.13	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	3.57	4.18	5.23	6.30	7.81	9.03	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	5.23	5.98	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.46	23.54	26.30	29.63	32.00	39.25
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.79
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.60	22.76	25.99	28.87	32.35	34.80	42.31
19	7.63	8.57	10.12	11.65	13.72	15.35	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	8.90	9.92	11.59	13.24	15.44	17.18	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	9.54	10.60	12.34	14.04	16.31	18.10	21.34	24.94	27.30	30.81	33.92	37.66	40.29	48.27
23	10.20	11.29	13.09	14.85	17.19	19.02	22.34	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	10.86	11.99	13.85	15.66	18.06	19.94	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	11.52	12.70	14.61	16.47	18.94	20.87	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	12.20	13.41	15.38	17.29	19.82	21.79	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	12.88	14.12	16.15	18.11	20.70	22.72	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	13.56	14.85	16.93	18.94	21.59	23.65	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	14.26	15.57	17.71	19.77	22.48	24.58	28.34	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	14.95	16.31	18.49	20.60	23.36	25.51	29.34	33.53	36.25	40.26	43.77	47.96	50.89	59.70

Πίνακας VII. Μετατροπή του r σε z.

r	Z	r	Z	r	Z
.00	.000				
.01	.010	.36	.377	.71	.887
.02	.020	.37	.388	.72	.908
.03	.030	.38	.400	.73	.929
.04	.040	.39	.412	.74	.950
.05	.050	.40	.424	.75	.973
.06	.060	.41	.436	.76	.996
.07	.070	.42	.448	.77	1.020
.08	.080	.43	.460	.78	1.045
.09	.090	.44	.472	.79	1.071
.10	.100	.45	.485	.80	1.099
.11	.110	.46	.497	.81	1.127
.12	.121	.47	.510	.82	1.157
.13	.131	.48	.523	.83	1.188
.14	.141	.49	.536	.84	1.221
.15	.151	.50	.549	.85	1.256
.16	.161	.51	.563	.86	1.293
.17	.172	.52	.576	.87	1.333
.18	.182	.53	.590	.88	1.376
.19	.192	.54	.604	.89	1.422
.20	.203	.55	.618	.90	1.472
.21	.213	.56	.633	.91	1.528
.22	.224	.57	.648	.92	1.589
.23	.234	.58	.662	.93	1.658
.24	.245	.59	.678	.94	1.738
.25	.255	.60	.693	.95	1.832
.26	.266	.61	.709	.96	1.946
.27	.277	.62	.725	.97	2.092
.28	.288	.63	.741	.98	2.298
.29	.299	.64	.758	.99	2.647
.30	.310	.65	.775		
.31	.321	.66	.793		
.32	.332	.67	.811		
.33	.343	.68	.829		
.34	.354	.69	.848		
.35	.365	.70	.867		

Πίνακας VIII. Τιμές ε για πιθανότητα 5%, διαφορετικό αριθμό μέσων όρων και διαφερετικούς Β. Ε.

ΒΕ	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8	ε_9	ε_{10}	ε_{12}	ε_{14}	ε_{16}
1	17,97											
2	6,09	6,09										
3	4,50	4,52	4,52									
4	3,93	4,01	4,03	4,03								
5	3,64	3,75	3,80	3,81	3,81							
6	3,46	3,59	3,65	3,68	3,69	3,70						
7	3,34	3,48	3,55	3,59	3,61	3,62	3,63					
8	3,26	3,40	3,48	3,52	3,55	3,57	3,57	3,58				
9	3,20	3,34	3,42	3,47	3,50	3,52	3,54	3,54	3,55			
10	3,15	3,29	3,38	3,43	3,47	3,49	3,51	3,52	3,52			
11	3,1 1	3,26	3,34	3,40	3,44	3,46	3,48	3,49	3,50	3,51		
12	3,08	3,23	3,31	3,37	3,41	3,44	3,46	3,47	3,48	3,50		
13	3,06	3,20	3,29	3,35	3,39	3,42	3,44	3,46	3,47	3,48	3,49	
14	3,03	3,18	3,27	3,33	3,37	3,40	3,43	3,44	3,46	3,47	3,48	
15	3,01	3,16	3,25	3,31	3,36	3,39	3,41	3,43	3,45	3,47	3,48	3,48
16	3,00	3,14	3,23	3,30	3,34	3,38	3,40	3,42	3,44	3,46	3,47	3,48
17	2,98	3,13	3,22	3,28	3,33	3,37	3,39	3,41	3,43	3,45	3,46	3,47
18	2,97	3,12	3,21	3,27	3,32	3,36	3,38	3,40	3,42	3,44	3,46	3,47
19	2,96	3,11	3,20	3,26	3,31	3,35	3,38	3,40	3,41	3,44	3,46	3,47
20	2,95	3,10	3,19	3,25	3,30	3,34	3,37	3,39	3,41	3,43	3,45	3,46
24	2,92	3,07	3,16	3,23	3,28	3,31	3,35	3,37	3,39	3,42	3,44	3,45
30	2,89	3,03	3,13	3,20	3,25	3,29	3,32	3,35	3,37	3,40	3,43	3,45
40	2,86	3,01	3,10	3,17	3,22	3,27	3,30	3,33	3,35	3,39	3,42	3,44
60	2,83	2,98	3,07	3,14	3,20	3,24	3,28	3,31	3,33	3,37	3,40	3,43
120	2,80	2,95	3,04	3,12	3,17	3,22	3,25	3,29	3,31	3,36	3,39	3,42
∞	2,77	2,92	3,02	3,09	3,15	3,19	3,23	3,27	3,29	3,34	3,38	3,41

Πίνακας IX. Μετατροπή εκατοστών σε μοίρες.

%	0	1	2	3	4	5	6	7	8	9
0	0	5,7	8,1	10,0	11,5	12,9	14,2	15,3	16,4	17,5
10	18,4	19,4	20,3	21,1	22,0	22,8	23,6	24,4	25,1	25,8
20	26,6	27,3	28,0	28,7	29,3	30,0	30,7	31,3	31,9	32,6
30	33,2	33,8	34,4	35,1	35,7	36,3	36,9	37,5	38,1	38,6
40	39,2	39,8	40,4	41,0	41,6	42,1	42,7	43,3	43,9	44,4
50	45,0	45,6	46,1	46,7	47,3	47,9	48,4	49,0	49,6	50,2
60	50,8	51,4	51,9	52,5	53,1	53,7	54,3	54,9	55,6	56,2
70	56,8	57,4	58,1	58,7	59,3	60,0	60,7	61,3	62,0	62,7
80	63,4	64,2	64,9	65,6	66,4	67,2	68,0	68,9	69,7	70,6
90	71,6	72,5	73,6	74,7	75,8	77,1	78,5	80,0	81,9	84,3

Πίνακας Χ. Τυχαιοποιημένοι μονοψήφιοι αριθμοί.

3	3	2	6	1	6	8	0	4	5	6	0	8	0	0	7	3	5	4	5	5	4
2	7	0	7	3	6	0	7	5	1	2	4	6	1	9	9	9	7	3	8	1	8
1	3	5	5	3	8	5	8	5	9	8	8	6	7	0	0	3	7	2	6	7	1
5	7	1	2	1	0	1	4	2	1	8	8	6	5	2	2	4	2	8	6	7	3
0	6	1	8	4	4	3	2	5	3	2	3	0	9	2	7	8	5	0	0	9	4
8	7	3	5	2	0	9	6	4	3	8	4	6	7	4	2	9	6	3	7	2	7
2	1	7	6	3	3	5	0	2	5	8	3	3	0	8	7	8	6	1	0	4	8
1	2	8	6	7	3	5	8	0	7	4	4	1	7	5	5	7	3	9	5	2	5
1	5	5	1	0	0	1	3	4	2	9	9	3	2	0	0	8	2	8	0	0	4
9	0	5	2	8	4	7	7	2	7	0	8	1	3	1	3	7	5	6	9	2	9
0	6	7	6	5	0	0	3	1	0	5	5	2	6	7	2	5	1	9	2	7	5
2	0	1	4	8	5	8	8	4	5	1	0	2	3	8	1	7	0	5	0	0	3
3	2	9	8	9	4	0	7	7	2	9	3	2	8	3	4	5	0	3	4	1	4
8	0	2	2	0	2	5	3	5	3	8	6	4	0	6	6	8	9	2	0	8	7
5	4	4	2	0	6	8	7	9	8	3	5	1	7	3	1	8	8	9	2	2	9
1	7	7	6	3	7	1	3	0	4	0	7	0	0	8	8	3	8	7	9	3	1
7	0	3	3	2	4	0	3	5	4	9	7	2	4	3	3	9	3	4	8	6	6
0	4	4	3	1	8	6	6	7	9	9	4	9	7	7	2	7	8	0	5	7	3
1	2	7	2	0	7	3	4	4	5	9	9	6	1	9	9	1	3	9	5	0	9
5	2	8	5	6	6	6	0	4	4	3	8	5	7	0	0	4	5	0	5	0	9
0	4	3	3	4	6	0	9	5	2	6	8	5	0	4	2	4	2	4	6	7	7
1	3	5	8	1	8	2	4	7	6	1	5	5	8	2	7	7	4	4	7	6	2
9	6	4	6	9	2	4	2	4	5	9	7	3	2	8	5	3	8	2	7	5	7
1	0	4	5	6	5	0	4	2	6	1	1	1	2	1	8	7	7	3	4	9	6
3	4	2	5	2	0	5	7	2	7	4	0	5	1	7	2	7	5	9	5	4	1
6	0	4	7	2	1	2	9	6	8	0	2	5	4	6	7	0	2	6	5	5	4
7	6	7	0	9	0	3	0	8	6	3	8	3	9	4	7	4	8	2	5	8	7
1	6	9	2	5	3	5	6	1	6	0	2	7	6	3	9	9	2	2	5	0	4
4	0	0	1	7	4	9	1	6	2	4	8	8	6	7	4	8	3	3	1	2	7
0	0	5	2	4	3	4	8	8	5	2	7	8	4	6	4	5	9	7	9	4	9

Πίνακας XI. Τυχαιοποιημένοι διψήφιοι αριθμοί.

03 47 43 73 86	36 96 47 36 61	46 98 63 71 62	33 26 16 80 45	60 11 14 10 95
97 74 24 67 62	42 81 14 57 20	42 53 32 37 32	27 07 36 07 51	24 51 79 89 73
16 76 62 27 66	56 50 26 71 07	32 90 79 78 53	13 55 38 58 59	88 97 54 14 10
12 56 85 99 26	96 96 68 27 31	05 03 72 93 15	57 12 10 14 21	88 26 49 81 76
55 59 56 35 64	38 54 82 46 22	31 62 43 09 90	06 18 44 32 53	23 83 01 30 30
16 22 77 94 39	49 54 43 54 82	17 37 93 23 78	87 35 20 96 43	84 26 34 91 64
84 42 17 53 31	57 24 55 06 88	77 04 74 47 67	21 76 33 50 25	83 92 12 06 76
63 01 63 78 59	16 95 55 67 19	98 10 50 71 75	12 86 73 58 07	44 39 52 38 79
33 21 12 34 29	78 64 56 07 82	52 42 07 44 38	15 51 00 13 42	99 66 02 79 54
57 60 86 32 44	09 47 27 96 54	49 17 46 09 62	90 52 84 77 27	08 02 73 43 28
18 18 07 92 46	44 17 16 58 09	79 83 86 19 62	06 76 50 03 10	55 23 64 05 05
26 62 38 97 75	84 16 07 44 99	83 11 46 32 24	20 14 85 88 45	10 93 72 88 71
23 42 40 64 74	82 97 77 77 81	07 45 32 14 08	32 98 94 07 72	93 85 79 10 75
52 36 28 19 95	50 92 26 11 97	00 56 76 31 38	80 22 02 53 53	86 60 42 04 53
37 85 94 35 12	83 39 50 08 30	42 34 07 96 88	54 42 06 87 98	35 85 29 48 39
70 29 17 12 13	40 33 20 38 26	13 89 51 03 74	17 76 37 13 04	07 74 21 19 30
56 62 18 37 35	96 83 50 87 75	97 12 25 93 47	70 33 24 03 54	97 77 46 44 80
99 49 57 22 77	88 42 95 45 72	16 64 36 16 00	04 43 18 66 79	94 77 24 21 90
16 08 15 04 72	33 27 14 34 09	45 59 34 68 49	12 72 07 34 45	99 27 72 95 14
31 16 93 32 43	50 27 89 87 19	20 15 37 00 49	52 85 66 60 44	38 68 88 11 80
68 34 30 13 70	55 74 30 77 40	44 22 78 84 26	04 33 46 09 52	68 07 97 06 57
74 57 25 65 76	59 29 97 68 60	71 91 38 67 54	13 58 18 24 76	15 54 55 95 52
27 42 37 86 53	48 55 90 65 72	96 57 69 36 10	96 46 92 42 45	97 60 49 04 91
00 39 68 29 61	66 37 32 20 30	77 84 57 03 29	10 45 65 04 26	11 04 96 67 24
29 94 98 94 24	68 49 69 10 82	53 75 91 93 30	34 25 20 57 27	40 48 73 51 92
16 90 82 66 59	83 62 64 11 12	67 19 00 71 74	60 47 21 29 68	02 02 37 03 31
11 27 94 75 06	06 09 19 74 66	02 94 37 34 02	76 70 90 30 86	38 45 94 30 38
35 24 10 16 20	33 32 51 26 38	79 78 45 04 91	16 92 53 56 16	02 75 50 95 98
38 23 16 86 38	42 38 97 01 50	87 75 66 81 41	40 01 74 91 62	48 51 84 08 32
31 96 25 91 47	96 44 33 49 13	34 86 82 53 91	00 52 43 48 85	27 55 26 89 62
66 67 40 67 14	64 05 71 95 86	11 05 65 09 68	76 83 20 37 90	57 16 00 11 66
14 90 84 45 11	75 73 88 05 90	52 27 41 14 86	22 98 12 22 08	07 52 74 95 80
68 05 51 18 00	33 96 02 75 19	07 60 62 93 55	59 33 82 43 90	49 37 38 44 59
20 46 78 73 90	97 51 40 14 02	04 02 33 31 08	39 54 16 49 36	47 95 93 13 30
64 19 58 97 79	15 06 15 93 20	01 90 10 75 06	40 78 78 89 62	02 67 74 17 33
05 26 93 70 60	22 35 85 15 13	92 03 51 59 77	59 56 78 06 83	52 91 05 70 74
07 97 10 88 23	09 98 42 99 64	61 71 62 99 15	06 51 29 16 93	58 05 77 09 51
68 71 86 85 85	54 87 66 47 54	73 32 08 11 12	44 95 92 63 16	29 56 24 29 48
26 99 61 65 53	58 37 78 80 70	42 10 50 67 42	32 17 55 85 74	94 44 67 16 94
14 65 52 68 75	87 59 36 22 41	26 78 63 06 55	13 08 27 01 50	15 29 39 39 43
17 53 77 58 71	71 41 61 50 72	12 41 94 96 26	44 95 27 36 99	02 96 74 30 83
90 26 59 21 19	23 52 23 33 12	96 93 02 18 39	07 02 18 36 07	25 99 32 70 23
41 23 52 55 99	31 04 49 69 96	10 47 48 45 88	13 41 43 89 20	97 17 14 49 17
60 20 50 81 69	31 99 73 68 68	35 81 33 03 76	24 30 12 48 60	18 99 10 72 34
91 25 38 05 90	94 58 28 41 36	45 37 59 03 09	90 35 57 29 12	82 62 54 65 60
34 50 57 74 37	98 80 33 00 91	09 77 93 19 82	74 94 80 04 04	45 07 31 66 49
85 22 04 39 43	73 81 53 94 79	33 62 46 86 28	08 31 54 46 31	53 94 13 38 47
09 79 13 77 48	73 82 97 22 21	05 03 27 24 83	72 89 44 05 60	35 80 39 94 88
88 75 80 18 14	22 95 75 42 49	39 32 82 22 49	02 48 07 70 37	16 04 61 67 87
90 96 23 70 00	39 00 03 06 90	55 85 78 38 36	94 37 30 69 32	90 89 00 76 33

Πίνακας XII. Τυχαιοποιημένοι πενταψήφιοι αριθμοί.

58649	85086	16502	97541	76611	94229	34987	86718	87208	05426
97306	52449	55596	66739	36525	97563	29469	31235	79276	10831
09942	79344	78160	11015	55777	22047	57615	15717	86239	36578
83842	28631	74893	47911	92170	38181	30416	54860	44120	73031
73778	30395	20163	76111	13712	33449	99224	18206	51418	70006
88381	56550	47467	59663	61117	39716	32927	06168	06217	45477
31044	21404	15968	21357	30772	81482	38807	67231	84283	63552
00909	63837	91328	81106	11740	50193	86806	21931	18054	49601
69882	37028	41732	37425	80832	03320	20690	32653	90145	03029
26059	78324	22501	73825	16927	31545	15695	74216	98372	28547
38573	98078	38982	33078	93524	45606	53463	20391	81637	37269
70624	00063	81455	16924	12848	23801	55481	78978	26795	10553
49806	23976	05640	29804	38988	25024	76951	02341	63219	75864
05461	67523	48316	14613	08541	35231	38312	14969	67279	50502
76582	62153	53801	51219	30424	32599	49099	83959	68408	20147
16660	80470	75062	75588	24384	27874	20018	11428	32265	07692
60166	42424	97470	88451	81270	80070	72959	26220	59939	31127
28953	03272	31460	41691	57736	72052	22762	96323	27616	53123
47536	86439	95210	96386	38704	15484	07426	70675	06888	81203
73457	26657	36983	72410	30244	97711	25652	09373	66218	64077
11190	66193	66287	09116	48140	37669	02932	50799	17255	06181
57062	78964	44455	14036	36098	40773	11688	33150	07459	36127
99624	67254	67302	18991	97687	54099	94884	42283	63258	50651
97521	83669	85968	16135	30133	51312	17831	75016	80278	68953
40273	04838	13661	64757	17461	78085	60094	27010	80945	66439
57260	06176	49963	29760	69546	61336	39429	41985	18752	98128
03451	47098	63495	71227	79304	29753	99131	18419	71791	81515
62331	20492	15393	84270	24396	32962	21632	92965	38670	44923
32290	51079	06512	38806	93327	80086	19088	59887	98416	24918
28014	80428	92853	31333	32648	16734	43418	90124	15086	48444
18950	16091	29543	65817	07002	73115	94115	20271	50250	25061
17403	69503	01866	13049	07263	13039	83844	80143	39048	62654
27999	50489	66613	21843	71746	65868	16208	46781	93402	12323
87076	53174	12165	84495	47947	60706	64034	31635	65169	93070
89044	45974	14524	46906	26052	51851	84197	61694	57429	63395
98048	64400	24705	75711	36232	57624	41424	77366	52790	84705
09345	12956	49770	80311	32319	48238	16952	92088	51222	82865
07086	77628	76195	47584	62411	40397	71857	54823	26536	56792
93128	25657	46872	11206	06831	87944	97914	64670	45760	34353
85137	70964	29947	27795	25547	37682	96105	26848	09389	64326
32798	39024	13814	98546	46585	84108	74603	94812	73968	68766
62496	26371	89880	52078	47781	95260	83464	65942	91761	53727
62707	81825	40987	97656	89714	52177	23778	07482	91678	40128
05500	28982	86124	19554	80818	94935	61924	31828	79369	23507
79476	31445	59498	85132	24582	26024	24002	63718	79164	43556

Πίνακας XII. Σχέδια λατινικών τετραγώνων.

3 × 3

Α	Β	Γ
Β	Γ	Α
Γ	Α	Β

4 × 4

Α	Β	Γ	Δ
Β	Λ	Δ	Γ
Γ	Δ	Β	Α
Δ	Γ	Α	Β

Α	Β	Γ	Δ
Β	Γ	Δ	Α
Γ	Δ	Α	Β
Δ	Α	Β	Γ

Α	Β	Γ	Δ
Β	Δ	Α	Γ
Γ	Α	Δ	Β
Δ	Γ	Β	Α

Α	Β	Γ	Δ
Β	Α	Δ	Γ
Γ	Δ	Α	Β
Δ	Γ	Β	Α

5 × 5

Α	Β	Γ	Δ	Ε
Β	Α	Ε	Γ	Δ
Γ	Δ	Α	Ε	Β
Δ	Ε	Β	Α	Γ
Ε	Γ	Δ	Β	Α

6 × 6

Α	Β	Γ	Δ	Ε	Ζ
Β	Ζ	Δ	Γ	Α	Ε
Γ	Δ	Ε	Ζ	Β	Α
Δ	Α	Ζ	Ε	Γ	Β
Ε	Γ	Α	Β	Ζ	Δ
Ζ	Ε	Β	Α	Δ	Γ

7 × 7

Α	Β	Γ	Δ	Ε	Ζ	Η
Β	Γ	Δ	Ε	Ζ	Η	Α
Γ	Δ	Ε	Ζ	Η	Α	Β
Δ	Ε	Ζ	Η	Α	Β	Γ
Ε	Ζ	Η	Α	Β	Γ	Δ
Ζ	Η	Α	Β	Γ	Δ	Ε
Η	Α	Β	Γ	Δ	Ε	Ζ

8 × 8

Α	Β	Γ	Δ	Ε	Ζ	Η	Θ
Β	Γ	Δ	Ε	Ζ	Η	Θ	Α
Γ	Δ	Ε	Ζ	Η	Θ	Α	Β
Δ	Ε	Ζ	Η	Θ	Α	Β	Γ
Ε	Ζ	Η	Θ	Α	Β	Γ	Δ
Ζ	Η	Θ	Α	Β	Γ	Δ	Ε
Η	Θ	Α	Β	Γ	Δ	Ε	Ζ
Θ	Α	Β	Γ	Δ	Ε	Ζ	Η

9 × 9

Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι
Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Α
Γ	Δ	Ε	Ζ	Η	Θ	Ι	Α	Β
Δ	Ε	Ζ	Η	Θ	Ι	Α	Β	Γ
Ε	Ζ	Η	Θ	Ι	Α	Β	Γ	Δ
Ζ	Η	Θ	Ι	Α	Β	Γ	Δ	Ε
Η	Θ	Ι	Α	Β	Γ	Δ	Ε	Ζ
Θ	Ι	Α	Β	Γ	Δ	Ε	Ζ	Η
Ι	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ

10 × 10

Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ
Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Α
Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Α	Β
Δ	Ε	Ζ	Η	Θ	Ι	Κ	Α	Β	Γ
Ε	Ζ	Η	Θ	Ι	Κ	Α	Β	Γ	Δ
Ζ	Η	Θ	Ι	Κ	Α	Β	Γ	Δ	Ε
Η	Θ	Ι	Κ	Α	Β	Γ	Δ	Ε	Ζ
Θ	Ι	Κ	Α	Β	Γ	Δ	Ε	Ζ	Η
Ι	Κ	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ
Κ	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι

11 × 11

Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ
Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Α
Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Α	Β
Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Α	Β	Γ
Ε	Ζ	Η	Θ	Ι	Κ	Λ	Α	Β	Γ	Δ
Ζ	Η	Θ	Ι	Κ	Λ	Α	Β	Γ	Δ	Ε
Η	Θ	Ι	Κ	Λ	Α	Β	Γ	Δ	Ε	Ζ
Θ	Ι	Κ	Λ	Α	Β	Γ	Δ	Ε	Ζ	Η
Ι	Κ	Λ	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ
Κ	Λ	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι
Λ	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ

<u>Πίνακας XIII</u>. Λογάριθμοι ακεραίων αριθμών, από το 1 έως το 10000.

											'Ανάλογα μέρη								
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374	4	8	12	17	21	25	29	33	37
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755	4	8	11	15	19	23	26	30	34
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106	3	7	10	14	17	21	24	28	31
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430	3	6	10	13	16	19	23	26	29
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732	3	6	9	12	15	18	21	24	27
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014	3	6	8	11	14	17	20	22	25
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279	3	5	8	11	13	16	18	21	24
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529	2	5	7	10	12	15	17	20	22
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765	2	5	7	9	12	14	16	19	21
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989	2	4	7	9	11	13	16	18	20
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201	2	4	6	8	11	13	15	17	19
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2	4	6	8	10	12	14	16	18
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2	4	6	8	10	12	14	15	17
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2	4	6	7	9	11	13	15	17
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2	4	5	7	9	11	12	14	16
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2	3	5	7	9	10	12	14	15
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2	3	5	7	8	10	11	13	15
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2	3	5	6	8	9	11	13	14
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2	3	5	6	8	9	11	12	14
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1	3	4	6	7	9	10	12	13
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1	3	4	6	7	9	10	11	13
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1	3	4	6	7	8	10	11	12
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1	3	4	5	7	8	9	11	12
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1	3	4	5	6	8	9	10	12
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1	3	4	5	6	8	9	10	11
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1	2	4	5	6	7	9	10	11
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1	2	4	5	6	7	8	10	11
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1	2	3	5	6	7	8	9	10
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1	2	3	5	6	7	8	9	10
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1	2	3	4	5	7	8	9	10
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1	2	3	4	5	6	8	9	10
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1	2	3	4	5	6	7	8	9
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1	2	3	4	5	6	7	8	9
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1	2	3	4	5	6	7	8	9
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1	2	3	4	5	6	7	8	9
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1	2	3	4	5	6	7	8	9
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1	2	3	4	5	6	7	7	8
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1	2	3	4	5	5	6	7	8
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1	2	3	4	4	5	6	7	8
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1	2	3	4	4	5	6	7	8
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1	2	3	3	4	5	6	7	8
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1	2	3	3	4	5	6	7	8
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1	2	2	3	4	5	6	7	7
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1	2	2	3	4	5	6	6	7
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1	2	2	3	4	5	6	6	7
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

Πίνακας XIII (συνέχεια). Λογάριθμοι ακεραίων αριθμών, από το 1 έως το 10000.

N	0	1	2	3	4	5	6	7	8	9	Ἀνάλογα μέρη 1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1	2	2	3	4	5	5	6	7
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1	2	2	3	4	5	5	6	7
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1	2	2	3	4	5	5	6	7
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1	1	2	3	4	4	5	6	7
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1	1	2	3	4	4	5	6	7
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1	1	2	3	4	4	5	6	6
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1	1	2	3	4	4	5	6	6
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1	1	2	3	3	4	5	6	6
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1	1	2	3	3	4	5	5	6
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1	1	2	3	3	4	5	5	6
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1	1	2	3	3	4	5	5	6
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1	1	2	3	3	4	5	5	6
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1	1	2	3	3	4	5	5	6
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1	1	2	3	3	4	4	5	6
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1	1	2	2	3	4	4	5	6
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1	1	2	2	3	4	4	5	6
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1	1	2	2	3	4	4	5	5
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1	1	2	2	3	4	4	5	5
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1	1	2	2	3	4	4	5	5
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1	1	2	2	3	4	4	5	5
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1	1	2	2	3	3	4	5	5
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859	1	1	2	2	3	3	4	5	5
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1	1	2	2	3	3	4	4	5
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1	1	2	2	3	3	4	4	5
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1	1	2	2	3	3	4	4	5
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1	1	2	2	3	3	4	4	5
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1	1	2	2	3	3	4	4	5
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1	1	2	2	3	3	4	4	5
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1	1	2	2	3	3	4	4	5
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1	1	2	2	3	3	4	4	5
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1	1	2	2	3	3	4	4	5
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1	1	2	2	3	3	4	4	5
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	0	1	1	2	2	3	3	4	4
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	0	1	1	2	2	3	3	4	4
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	0	1	1	2	2	3	3	4	4
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0	1	1	2	2	3	3	4	4
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0	1	1	2	2	3	3	4	4
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0	1	1	2	2	3	3	4	4
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0	1	1	2	2	3	3	4	4
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0	1	1	2	2	3	3	4	4
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0	1	1	2	2	3	3	4	4
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0	1	1	2	2	3	3	4	4
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0	1	1	2	2	3	3	4	4
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0	1	1	2	2	3	3	4	4
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0	1	1	2	2	3	3	3	4
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

ΒΙΒΛΙΟΓΡΑΦΙΑ

Α. Ελληνική.

ΒΑΦΙΑΣ, Β. Ν. και ΜΗΤΡΟΓΕΩΡΓΙΟΥ, Π. Δ. 1999. Σημειώσεις Βιομετρίας. Τ. Ε. Ι Λάρισας, Σ. Τ. Ε. Γ. *Λάρισα* 239 σελ.

ΔΑΛΛΙΑΝΗΣ, Κ. Δ. 1972. Σχεδίαση και Ανάλυση Πειραμάτων. *Αθήνα*. 586 σελ.

ΖΩΤΗΣ, Σ. 1992. Σημειώσεις Βιομετρίας. ΤΕΙ Δ. Μακεδονίας, ΣΤΕΓ, Τμήμα Φυτικής Παραγωγής. 206 σελ.

ΚΑΛΤΣΙΚΗΣ, Π. Ι. 1981. Γεωργικός Πειραματισμός. Τόμοι I-IV. *Αθήνα*. Τόμος I 510 σελ., Τόμος II 310 σελ. Τόμος III 164 σελ. Τόμος IV. 258 σελ.

ΛΑΖΑΡΙΔΟΥ, Θ. 2013. Βιομετρία, Σημειώσεις. ΤΕΙ Δ. Μακεδονίας, ΣΤΕΓ, Τμήμα Φυτικής Παραγωγής.

ΤΖΑΛΗΣ, Κ. 2000. Σημειώσεις Βιομετρίας. ΤΕΙ Δ. Μακεδονίας, ΣΤΕΓ, Τμήμα Ζωικής Παραγωγής. 141 σελ.

ΥΦΟΥΛΗΣ, Α. Χ. και ΓΕΛΕΚΗΣ, Σ. Β. Βιομετρία. Ο. Ε .Δ. Β. *Αθήνα.* 290 σελ.

ΦΑΣΟΥΛΑΣ, Α. Κ. 1979. Στοιχεία Πειραματικής Στατιστικής. *Θεσσαλονίκη.* 255 σελ.

ΦΩΤΙΑΔΗΣ, Ν. Α. 1995. Εισαγωγή στη Στατιστική για Βιολογικές Επιστήμες. University Studio Press. *Θεσσαλονίκη.* 227 σελ.

Β. Αγγλική.

COCHRAN, W. G and Cox, G. M. 1957. Experimental Designs. J. Wiley and Sons. *New York*. 617 p.

FISHER, R. A. and YATES, F. 1963. Statistical Tables for Biological, Agricultural and Medical Research. Oliver and Boyd. *Edinburgh*. 146 p.

KEMPTHORNE, O. 1952. The Design and Analysis of Experiments. J. Wiley and Sons. *New York*. 631 p.

SNEDECOR, G. and COCHRAN, W. 1980. Statistical Methods. Iowa State University Press. *Ames.* 507 p.

SOKAL, R. R. and ROHLF, F. J. 1981. Biometry. The principles and practice of Statistics in Biological Research. W. H. Freeman and Co. *San Francisco.* 857 p.

STEEL, R., G., and TORRIE, J., H. 1980. Principles and procedures of statistics. Mc Graw –Hill. *New York.* 633 p